Ultrafast Processes in Chemistry
Prof. Anindya Dutta
Department of Chemistry
Indian Institute of Technology-Bombay

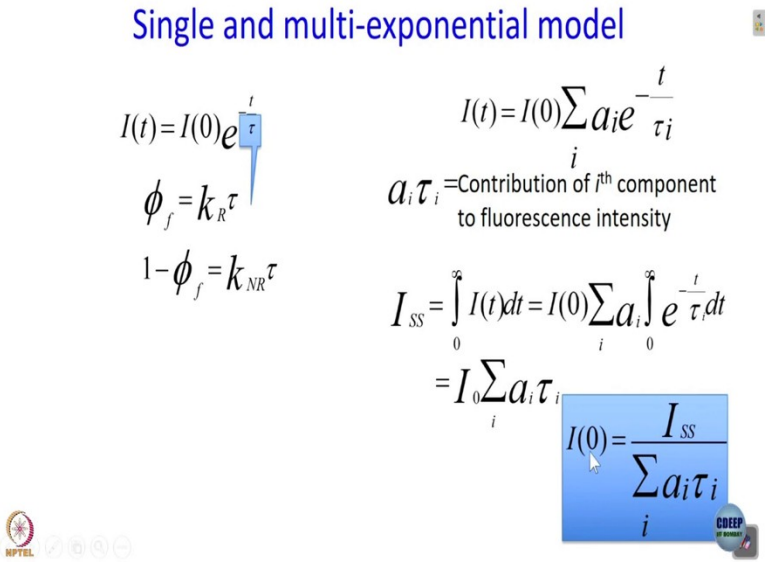
Lecture # 09
Data Fitting 1

Last time we met, we talked about iterative deconvolution, and how to fit data. So today we will try to learn 2 things. First is what do we fit the data to and second, how do we know that the fitting is good or not? I mean, if we look at the data, looking at it, we might be able to say that is a good fit or not a good fit. But the question is, computer does not have eyes, how does the computer know whether the fit is good or otherwise, but before we go there, let us discuss the data fitting models.

(Refer Slide Time: 00:51)

thereafter all, if you remember we have said that in the simplest case scenario, we have a single exponential decay and we said that in more complicated scenario, the most popular way of fitting data, not necessarily always the correct way of fitting data is by a multi-exponential function. And I think we are more or less familiar with this kind of functions.

(Refer Slide Time: 01:17)



The first one here, I of t = I of 0 multiplied by e to the power – t by tau, this is the simplest way you can fit the data I of at t is the fluorescence intensity at time t. I at 0 is a person is the fluorescence intensity at time of excitation or 0 time and Tau is a lifetime. So, this is essentially the integrated rate law for a first order process. It cannot get any simpler than this. But then we said that it is not necessary that life will be so simple and we can have more complicated data.

So, as of the first kind of complication we can think of, is in the form of a multi-exponential data, let us say you have several independent decay pathways, then what do you get I of time t after excitation is the same I of time 0. But now this is multiplied by not one exponential term, but rather a linear combination or weighted average or weighted sum of several exponential terms, I can have several values 2 3 4 5 6 10 100 200 in principle can have any number.

And then you have as many exponential terms and as many amplitudes. square of amplitudes as you know gives the contribution. Now, the thing is if you increase the number of exponential terms, generally, you'd get a better decay a better fit, because there is something called over parameterization. So, first question to ask is my decay single exponential or is it not? How do we know whether a single exponential or not.

The best way to do that is to make a similar plot where y axis now is in logarithmic scale, x axis time is in linear scale and have a look at it. What will the shape of this curve B, if it is a single

exponential decay is a straight line. And if it is not a single exponential decay, then it will not be a straight line. So, that is the first question to ask. So, if it is a straight line, then actually if you fit it to 2 or 3 exponential terms, then it will still fit very well, but it would not make any sense.

So, first thing that one needs to do is a visual inspection. And by visual inspection, the first thing to ask is whether it is at all exponential, single exponential, or whether it is more complicated. Of course, looking at the decay, you will never be able to tell whether it is bi exponential or tri exponential or what you can only tell whether it is single exponential or not. Now, what are the implications of these terms? Let us start with the discussion of the single exponential decay tau.

We have discussed tau already. so what is the meaning of tau? Can you tell me? Tau is called the lifetime that is right, why is it called the lifetime because it is average time spent by the molecule in its excited state. And that is something that was given to you as a homework, you are supposed to work it out. I is worked out in standard textbooks, like that of lakowicz, principles of person spectroscopy by lakowicz.

And then, this lifetime tau is also related to some other quantity that we have discussed very early on in this course. And what is that quantity? So, let us put it this way. If lifetime is longer, do we expect the fluorescence to be more intense or less intense? Then you expect it to be more intense. Why? Because of this simple relationship that phi f , the fluorescence quantum yield = k r multiplied by tau. what is k r here? it is a radiative rate constant. A word of caution here, very often in literature, you will see people say it is radiative rate.

But let us not forget that it is a radiative rate constant. So, please be careful and remember it is a rate constant and not rate right. Now, this radiative rate constant is related to some fundamental quantities that we might have studied in spectroscopy courses during our MSc or something. Can you tell me what the radiative rate constant is related with? Einstein coefficient which one actually if it is related to ‘a’ It will be related to ‘b’ also. Einstein ‘a’ coefficient is the coefficient for spontaneous emission.

But then that is also linearly related to be at what is ‘b’ related to? ‘b’ when you say b, einsteins ‘b’ coefficient it is for stimulated transition between 2 states right. Upward transition and downward transition so, what about upward transition is there a ‘b’ associated with it that actually equal right b12 equal to b2 1 in terms of experiment, what is the experimental quantity that is associated with b12, where one is a lower level 2 is a higher level for absorption the Einstein ‘b’ coefficient for absorption which experimental quantity should it be related to yes, little louder please. So, what do we what is it called of course, it is related to transition moment integral, but there is a that is something that you get from quantum mechanics. What is it that I can get using some instrument experimentally without knowing any quantum mechanics perhaps that will be related to radiative rate constant. Molar extinction coefficient or molar absorption coefficient?

That would be related to the radiative rate constant. And there is a relationship between the 2, which was a you can study, we are not going to the detail right now. Now, so the good thing about knowing fluorescence quantum yield from steady state measurement and lifetime from a time resolved measurement is that you can work out the radiative rate constant. More importantly, 1 – phi f = k NR multiplied by tau.

So you can work out the non-radiative rate constant. So as you go further in our discussions, we will see that will more and more want to know what is the rate constant associated with some non-radiative process that takes place in the excited state of a molecule? And this is how we will get the answer. Now, the problem is we get the answer very nicely if it is a single exponential decay. The moment it is multi exponential situation become complicated.

So, when it is multi exponential, what is the implication of a_i? What is the implication of tau i let us ask that question and the answer is a_i multiplied by tau i gives you the contribution of the i th component to fluorescence intensity. Now, this point needs to be understood very clearly, in order to go further ahead in the discussion of time reserved fluorescence spectroscopy. a_i into tau i. i is some component right.

So, let us think like this, that we use this example once again a little later. Let us think that there are 2 components tau 1 and tau 2, tau 1 is because of a fluorophore that is free and tau 2 is due to the same fluorophore that is bound to say cyclodextrin or protein or something like that and tau 2 is longer than tau 1. what will be the intensity will the intensity be more will the intensity be less? That depends not only on tau 1 and tau 2, but also on how much of it is bound and how much of it is free.

Let us only 20% of the fluorophore is bound to cyclodextrin and let us say for the free form of the fluorophore lifetime is one nanosecond for the bound form lifetime is 10 nanosecond. what will be the intensity is 20% is bound and what will be the intensity if 80% is bound. naturally intensity will be much more but 80% is where does that come from that comes from here that contribution of the ith component to fluorescence intensity is actually a_i multiplied by tau i. tau i you remember is an intrinsic quantity lifetime characteristic quantity.

but a i is the contribution and this can have actually severe implication. suppose think of a Nano particle that we have made which is almost completely non fluorescent. The only photoluminescence it has is due to some trapped states right. So, let us say that the time for the recombination of electron and hole in the nano particle is something like 1 picosecond 1 picosecond is a small time so, fluorescence intensity should be low.

but let us say there is some trap state and concentration of traps which is really very low, but lifetime of the traps it is hundred nanosecond what will the photo luminescence of this Nano particle be due to mainly the trap state which is very few in number or the intrinsic Band-Aids recombination of electron and whole which is taking place all the time in photo luminescence, you will actually see a much greater contribution of the trap state because this lifetime is hundred nanosecond.

But this is only an example, there are cases in which a small a i can be overcome by a large tau i what we just discussed, there are cases in which a small tau i can be overcome by a large a_i. Think of an extreme case think of say warfarin is a very common fluorophore that is used in

fluorescence and study of protein. Lifetime of free warfarin is something like hundred picosecond lifetime of bound warfarin is about 2 nanoseconds.

Now, let us say a very little protein almost all the warfarin is free will intensity be high or low? It will be low because then this tau i tau 1 a 100 picosecond that component will have almost hundred percent contribution a i will be large for it, but when it is bound to protein, even if say 10% of it is bound to protein then what will happen? Contribution of fluorescence intensity will be much more because lifetime has now increased 100 picosecond to 2000 picosecond, two nanosecond, 20 fold increase.

So, what the composite intensity will be is governed by a i be the relative values of amplitude as well as lifetime. So, a i tau i remember is the contribution of the I th component to fluorescence intensity. So, what is steady state that what is a steady state intensity see a very talked about a single exponential decay we could easily correlate the quantum yield which is a measure of steady state intensity with lifetime. can we do some such correlation in case of multi exponential decay.

Let us see, in case of multi exponential decay will you agree with or for any decay Actually, I hope you agree with me when I say that the steady state intensity is integral of intensity at some time t. From time 0 to infinity of after excitation. Of course, when I say 0 to infinity, right infinity only to make it a general statement, it is not really infinity. For all practical purposes, what is infinity? Infinity is a point where the decay has become almost 0.

If it is an exponential of multi exponential decay, it becomes 0 asymptotically. But for all practical purposes, suppose I at times 0 is 5000 counts. And then you go to 10 nanosecond. And there you see that the intensity has become 5 counts. So 5 is much, much less than 5000. So you say to yourself, so what we are saying is, intensity of steady state is really integral of i of t dt. For limits 0 to infinity of time or rather we can say that it is the area under the decay.

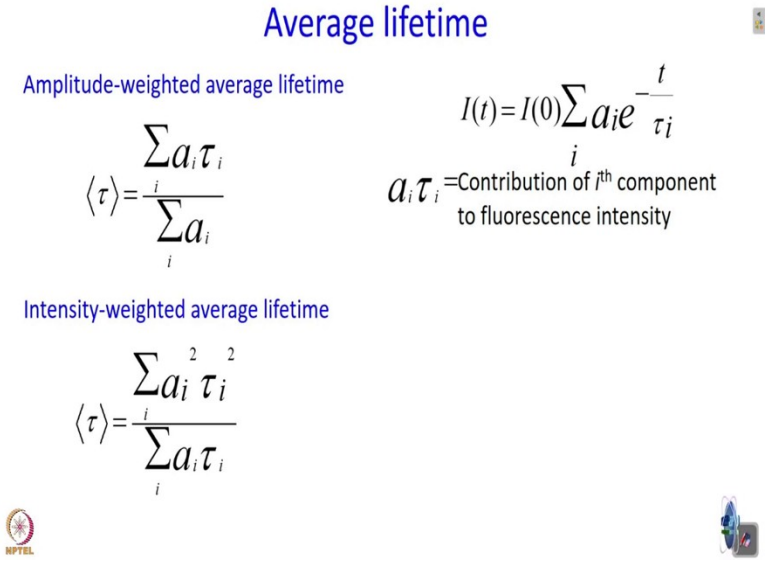
Of course, we are talking about a particular wavelength is this understood? that steady state intensity at any particular emission wavelength is the area under the decay or it is the integral

from 0 to infinity of i of t. Then let us substitute the expression. Since I at 0 is a constant it comes out and I can take the summation outside the integral. So, I get i at times 0 sum over I, a i integral e to the power - t by tau i dt.

And an advantage of setting the limits from 0 to infinity is that this becomes a standard integral solution of which is known. And when you put the solution, we get something like this. I steady state is I at times 0. I missed that. 0 in brackets here about that this has become small i at time 0 sum over i a i tau i or you can write i at times 0 = i steady state divided by some over i a i tau i. So, here there is a correlation between steady state intensity and the lifetimes.

The take home message is that it is not enough to look at only lifetimes, you have to look at their amplitude as well contributions as well. But, actually, it is better to stop here and not get over enthusiastic and take it a little further like say 80% of people do in fluorescence spectroscopy. So, what you see is that almost all the decays are fitted to multi exponential function and everywhere people happily work with what they call.

(Refer Slide Time: 15:56)



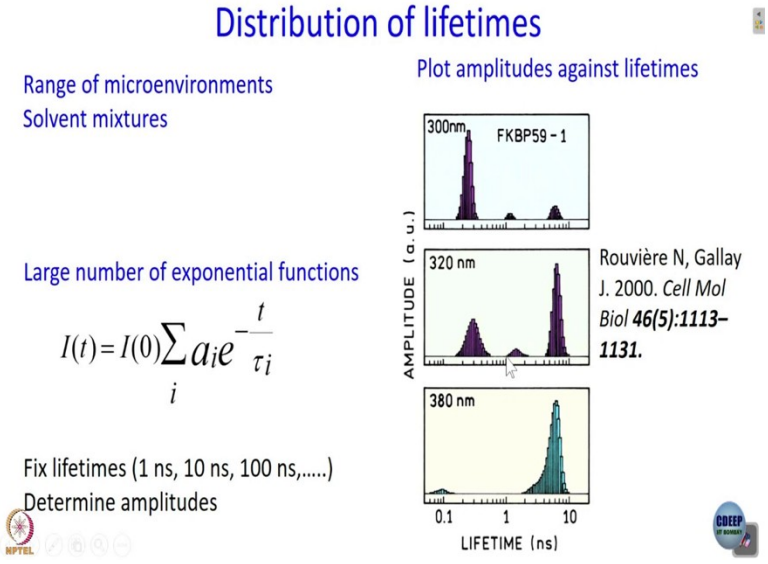
Average lifetime this thing that you see average lifetime is some over i a i tau i divided by some over i a i dimensionally is this fine right because this will have the magnitude of time. But this amplitude weighted average lifetime to be honest, has no meaning other than steady state intensity.

So basically gives you a measure of steady state intensity. And if we are going to talk about steady state intensity only, then what is the point of doing a time resolved measurement in the first place. So, as far as possible, it is better to avoid using average lifetimes and also this is not really average lifetime.

What is really average lifetime is this intensity weighted average lifetime sum over i a i square tau i square divided by sum over i a i tau i. So, you see here the denominator is actually the total intensity so this average lifetime may have some meaning. So it is related to the area under the curve. But then from here trying to work out radiative rate constant nonradiative rate constant is not a very sensible thing to do. Because after all you are saying that different lifetimes are associated with different processes. Which would have different non radiative constants or radiative constants or whatever.

So, if you take an average lifetime, all that individual, information and amplification everything is lost. So, if we have to work multi exponential decay, if we have to use average lifetime for some reason, then let us not try to take it too far and work out the rate constants in the first place. They are not completely useless. This amplitude related lifetimes actually are used when you talk about say, foster resonance energy transfer. That is where this amplitude weighted lifetime have some application. But generally, it is not really correct to talk to call this the average lifetime. This is average time lifetime if at all.

(Refer Slide Time: 18:12)



And it is not very useful. That being said, Let us move over to something that is more complicated and therefore, closer to reality many times. So, next model we want to discuss is distribution of lifetimes. And this distribution of lifetimes is a much better model, then some of exponent but the problem is this, when you do some of exponential, then what you imply is that you have that many lifetimes, discreetly, but sometimes that may not be the case. Suppose you have a range of micro environments you do not have a 01 situation.

You have some kind of a micro heterogeneous medium, very you have gradient polarity or gradient viscosity say there is a polymer right and maybe at the core the polymer is very dense and on the outside it is not dense at all. And let us say your fluorophore is distributed from core to the end many places. Now, a multi-exponential model is not valid, if it is simply bound versus free, then it is valid, but even when you go back to this bound versus free model that we used, think of some fluorophore that is bound to a protein.

It is not always the case that it is bound specifically to one site and experiences one kind of environment, more often than not, you can have nonspecific binding, and if it is nonspecific binding, then even bound fluorophore actually experienced different kinds of environments. Or in other words, they experience a distribution of environment and this environment, it might be convenient if we talk in terms of polarity, we are all familiar with the dielectric constant even though dielectric constant is not a good parameter of polarity in micro heterogeneous media.

But still for the sake of simplicity, let us see dielectric constant, let us say our fluorophore is bound to a protein not specifically, and it experiences a range of dielectric constants. The modal dielectric constant, let us say, is 20 and there is a distribution so 20 + - 5, that is a distribution and a distribution is going to have some kind of a shape, it can be a gaussian distribution, it can be a Lorentzian distribution, it can be 2 sided exponential, it can be whatever.

But some distribution function may be there for such a case, a better fitting function then the mundane multi exponential model is distribution of lifetimes. And here you need to look at the function a little carefully, because it might actually look like multi exponential function to the

untrained eye. i as time t = integral 0 to infinity alpha tau = e to the power - 1 by tau t by tau t tau. Please note it is not dt.

Of course, an integration is a summation, but here alpha tau means distribution function of lifetime and we are integrating over lifetime I am not treating the distribution function explicitly because you might have to use different distribution function depending on what kind of system it is. But this is more often than not much better fitting model, than multi exponential. See multi exponential function might fit your decay.

I am not saying it would not fit. Because I might have said in this course, I have actually seen an elephant shape of an elephant drawn by a clever combination of 30 exponential functions using a sufficient number of exponential functions, but as you can draw a self-portrait or Also, if you play with the amplitude correctly, and if you play with the shifts correctly, but that would not mean anything is an elephant made up of 30 exponential functions.

And it does not even make sense. It is funny, right? It is laughable. Similarly, just because your decay might fit to all the exponential function does not mean that it is the correct model to use. And if we are going to do a quantitative study, if we are going to extract as much juice as you can from your lifetime data, then it is important to go beyond the convenient multi exponential model and think what your system is like and think what kind of fitting model would be appropriate for your system.

Unfortunately, this distribution of lifetime and all they actually come with commercial data fitting packages now. In our lab, we have 2 programs. One is from picoquanttropical one, the other is from IBH, which is now horiba jobin yvon. Both the programs, I believe, have this option of fitting to a distribution of lifetime. it is more difficult. It takes more time. It requires more playing around, but it is doable. Of course, if you use a better algorithm, then it is easier to do it.

But maybe we will postpone the discussion until we talk about it actual data fitting and goodness of it. Let us move on this distribution of lifetime is often a better model to use, depending on

what kind of system you are looking at. But as we discussed, it is also a more complicated model. Multiple exponential is easier. In fact, even fitting is easier. So often, what we do is and these programs usually have provision of letting you do it.

Often what you do is you try to get away with the trouble of using explicitly a distribution function like gaussian lorenzian etc. By instead using a large number of exponential function was at this point, it might be a little confusing, because 10 minutes ago, I was saying not very kind things about multiple exponential functions. And here I am saying that you can fit the data to a large number of exponential functions, but bear with me for a while, it will start making sense.

So, what you do is you fit to a large number of exponential functions, but what you do is that you tell the system what the lifetimes are.so if you fit two, not two exponential or three exponential function. Fit to 100 exponentials. If your computer and if your program are good enough free to 1000 exponentials and use a wide range something like this, you 6 lifetimes. So, the way you fit now is that you say that the lifetimes which I have are 0.1 nanosecond, one nanosecond 10 nanosecond 100 nanoseconds so on and so forth.

Usually they are arranged logarithmically not thereafter one, but logarithamically so that you can look at small lifetimes as well as large lifetimes. And you fit your data to this function, where all these tau i values are forcibly preset. So, what is the only play you have? What is the only parameter that is going to change the amplitudes right a i. So what you will get is you will get if you are using hundred lifetimes, you will get 100 amplitude.

Now, what you do is you brought the amplitude against lifetime and then you get thoughts like this is actual data taken from this 2000 cell molecular biology paper. So, here you see at the looked at different emission wavelengths, 300 nanometer 320 nanometers 380 nanometers, note the y axis amplitude, note the x axis lifetime. And here of course, they are not there are 100 nanosecond rather they have gone from less than .1 nanosecond.

Do not ask me how they did it using time correlated single photon counting up to say, 10 nanosecond. And if you look carefully at the x axis and you see that it is logarithamic because

you want to look at .1 picosecond .1 nanosecond kind of lifetime as well as 5 6 nanosecond kind of life. If it is not logarithmic, you are going to miss this. So, you see, let us not worry about what you what FKBP59 - 1 some kind of a protein.

But what you see is at 380 nanometer, you have 2 kinds of lifetimes something that is very small .1 nanosecond or So, something that is quite large, say, what is this 2 3 4 5 6 nanosecond and there is a distribution about 6 nanosecond, there is a distribution of .1 nanosecond as well. That means that, first of all, there are 2 broad kinds of environments. Moreover, within each of the kind of environment, there are sub domain.

That is where you get this distribution. And one reason why if you can if you have the capability of fitting you are data to 100 exponentials. One way, this approach is better than using an explicit distribution is that how do you know what the distribution is? How do you know that it is gaussian? Look at what we see here is this gaussian is actually lognormal kind of distribution. But there is no way in which I can know beforehand.

Whether it is going to be gaussian lorenzian lognormal what? Right? So good thing about fitting your data too many exponential model, where lifetime is fixed amplitude is varied. And you make a product amplitude versus lifetime, is that you do not care about what the kind of distribution is, but it comes out automatically in your result. Right here you see that the short lifetime but of course, this may not be lognormal. Also, do not forget, the x axis is not linear.

It is actually logarithmic. So I do not know what it is, but the point is, I am not working with any particular kind of distribution, whatever is the distribution is expected to show up in the process. Now, when you go from 380 nanometer to 320 nanometers. Now, what do you see? Now, this .1 nanosecond kind of component is completely gone. Rather, you have a broad effect if you work out the area under this one, and this one, I do not know which one will be more, I do not know even more, because the scale is logarithmic.

But here you have quite a good distribution around .3 nanosecond. So, the .1 is that nanosecond component is gone, you get a 0.3 nanosecond component and you have this distribution there.

We have something new between one and 2 nanoseconds and that also has a broad distribution and whatever you heard earlier is there. But now what it appears is that this thick edge you had has given way to a completely new distribution that is there.

I do not know what the system is and at the moment I do not care, just try to show you some data and try to discuss what this data would mean. Now, when you go to 300 nanometer, what do you see? You see that this .3 nanosecond component that had come, that is now the major component. It has some distribution, but it is not so much. But of course, you can see that here fully testament maximum is several nanosecond here fully width half maximum is hundreds of picosecond again of course, you have to work a percentage.

So, now, this 200 300 picosecond component is a major one, this long component has become very small, and this one has also gone down compared to what it is. So here, I hope we have been able to convey that by doing this kind of data fitting, we actually get a wealth of information that we do not get if we mindlessly fit our data to double exponential triple exponential model and resort to your average lifetime that means nothing this actually tells you what your system is like.

So, what we have learned so far is that more often than not you might have to work with a system where you have a distribution of lifetimes. The one way of handling distribution of lifetimes is to use a specific distribution. Danger of that is that may not be the case. Other way of doing it is go back to good old multi exponential function, but this time, plot, amplitude versus lifetime and do not stop at 2 and 3.

Since you are doing multi exponential go all the away and fit 200 exponential, but for that you have to a good computer. You have pressed out algorithm. We will talk a little bit about algorithm towards the end. But here we take a break and we come back in the next module and continue with more data fitting models. And then we also learn about goodness of it.