Course Name: Turbulence Modelling

Professor Name: Dr. Vagesh D. Narasimhamurthy

Department Name: Department of Applied Mechanics

Institute Name: Indian Institute of Technology, Madras

Week - 1

Lecture - Lec04

4. Statistical Analysis and Cartesian tensors - I

Let us begin and welcome to the class again. So, in the last class just to recollect we saw that the arithmetic mean that we are going to use in reality when we work with data is also a random process right. So, we saw that the arithmetic mean is a random process and we are and therefore, we cannot work with arithmetic mean. The objective is to separate turbulence from the statistical mean and therefore to study them for that objective we define what is called a true mean or an ensemble mean which is only theoretically possible we cannot have access to this and that is because there are two conditions that is required that is samples must be statistically independent and you need infinite number of samples these two are not possible to get statistically independent events that is one event not influencing the other one right. So, this kind of statistical independence as well as infinite numbers number of samples both are not possible and therefore, it is only a theoretical definition, but we proceed with that believing that this ensemble mean or a true mean exists and therefore, we can separate the all the random components from the data ok.

So, we defined what is an ensemble mean? And then we also saw that the time mean or a time averaging is you know one interesting concept people want to use it, but you have to be careful this time mean becomes ensemble mean only when flows are stationary. If you have a flow like this where the mean is not a function of time then this ϕ' bar, over bar represents ensemble averaging then this is fine in a stationary flow. But if we have a sinusoidal or an oscillatory flow then the mean itself is changing with time. So in these cases, time averaging doesn't work.

So in general, any type of flow, whether you take stationary flows, periodic flows, or transient flows, ensemble mean works. And therefore, we proceed with using ensemble mean. And we also saw that two signals can look very different, but can have same mean. So, the ϕ' bar represents ensemble averaging, ϕ' is any random process it could be your velocity temperature pressure. So, these two are having same mean, but they look different.

So, the randomness has to be studied for that we define an higher order term called variance. So, the variance can help us to know when two signals are having same mean here right and we also saw that the when you ensemble average a fluctuation x' is the fluctuation component when you ensemble average it actually becomes 0 because x' can be split into or written it as x which is the instantaneous signal itself minus its mean true mean. When you ensemble average one more time, then you get basically ensemble averaged value minus ensemble average value. So, performing an ensemble average on an already a true mean does not change anything. It has already achieved its true mean state.

So, it has no effect and therefore, this will be 0. And when it comes to reality because in lab or in simulations you do not have access to true mean, you are going to use arithmetic mean. There you must know that averaging the fluctuation will not be 0. Some residue will be there. So, smaller the residue, more closer it is going to 0 is good for you.

when you work in reality in the lab. So, we defined also variance which is can also be defined in terms of an RMS or a standard deviation. This has a physical meaning I will come to that what it means and yeah. So, I also discussed that we can get two signals with same mean ϕ' bar as well as same variance ϕ'ϕ' over bar. like, these two signals.

Let's say, but then one signal is biased towards the one side of the it is having a greater excursion towards one side of the mean than the other compared to here where it is more or less becoming flat So, for that when two signals or two different data sets, two turbulence data having same mean same variance then we search for higher order terms that is what I mentioned here, but the only thing is here. these two terms the higher order terms x' cube over bar or x' raise to the power of 4 over bar. We need such higher order terms and I will explain what they physically what is the physical meaning of this when I go to governing equations. And I did mention that this is related to skewness and flatness only to remember here is that this skewness and flatness are non-dimensional, they are non-dimensional. They are non-dimensional equivalents of these two.

So, we need to non-dimensionalize the third order and the fourth order term to get what is called a skewness and a flatness. So, yesterday we look into all these parameters. One thing to notice is that when we look at fluctuations about the mean, fluctuations about the mean. We have already seen that ensemble averaging the fluctuation becomes 0. So, I would like to graphically see how it looks like that is x' over bar is 0.

So, suppose if I have a signal let us say I will take a bit complicated signal. Let us say I have a signal like this, this is the instantaneous component, let us call x or ϕ' or anything.

Then, this is the number of samples, of course, and then upon averaging, I would achieve its mean. So, this will be my true mean x bar. So, now what is ϕ' minus? So, if I want to represent ϕ' which is ϕ' minus ϕ' over bar.

If I want this, how does that look like here? If I subtract the mean from these fluctuations, it essentially goes to this. Correct? I am subtracting the mean which is that red signal ϕ' bar from the black signal which is the instantaneous raw data. When I do that, your fluctuations go, it will have excursions about zero, right? So, mean of that true mean of the signal is 0, right and therefore, ϕ' ' is 0, right? We already saw the mathematical proof of it, but this is how it looks like. And therefore, here in this particular signal, you should not do time average. If you do time average, then this ϕ' ' over bar will not be 0.

You will have a sinusoidal signal also overlapping into it. So, that is why ensemble average is the best in any class of flows, ok. So, this is one thing that I wanted to highlight. Another thing is we also saw that the arithmetic mean is a random process. So, if I am going to plot the arithmetic mean itself.

So, let us say x_n as n changes. So, we had this throwing the dice game yesterday where x_n was changing 3.5, 4.1 and so on as n increases. Even this is not a true mean that we have already seen.

So, the value will go like this. So, as n changes so this is also a random process. So, in practice what should you do is the question. So, the question you have to ask is what to do in practice not theory right. So, in a theoretically we have the ensemble mean that is fine, but when you work with data that is coming from your simulations or experiments, what do you do in practice? We cannot get statistically independent data and we cannot get infinite number of samples.

So, first thing to check here is convergence. So, as n increases this x_n should converge to x, x is the true mean or x bar. So, increase n that is the samples to achieve convergence. Convergence of arithmetic mean something like you have seen in a signal it is no longer So, when it is coming to this part here it is no longer varying so much minor variations is exhibiting that is what some of you said I in the laboratory I took 7, 8 say samples and averaged and it was not varying in some other data you may take 7 million depends on what is your error tolerance here, but you should note that the ensemble mean of that fluctuation will still be not 0 some small value make sure it is small value. So, that is the first thing you have to increase n take as many number of samples.

So, that you achieve convergence. So, then the other thing is you cannot get statistically independent samples, but what you should do is you should check for samples that are not correlated samples ok, or you can say uncorrelated, you have to collect uncorrelated samples ok, collect uncorrelated samples. So, what does this mean uncorrelated samples means? So, physically let us say I have, let us say I have a pipe flow here. okay, pipe flow, I'm only looking into a certain cross-section, okay, so here you would obviously see some turbulent eddies, some vertices such vertices are coming, so what you have to make sure here is that, let's say if you are taking data at a certain point. Let's say I'm taking data at this position If I am taking a data at this certain XYZ location and let us say the eddy is now sitting here.

This particular eddy would have let us say there is an eddy sitting around it. You must wait for certain amount of time so that this eddy pass over it and then take the next sample. So, this eddy should move out. and be outside its region of influence then you collect the next sample at XYZ so don't collect samples since you are going to do time averaging here right. I'm talking about in reality not the theory in reality you are going to take continuous samples let's say in a wind tunnel or a pipe flow calculation you're looking into let's say velocity at this position continuously taking data don't do that wait for some amount of time so that this data is not correlated with each other if you take data in the which is let's say this eddy let's say it takes 10 seconds to move out of the zone of influence then if you take samples every millisecond all this data is influencing each other you will not learn anything new out of that one you wait for a certain amount of time I'll tell you how to compute this there is a way to compute what we call an autocorrelation coefficient the time required to wait in sampling This experimentalists also do.

This is all data collection and processing, right? So, uncorrelated samples and then check for convergence,. Ok, So now we look into what is called single point statistics, ok single point statistics. We kind of already saw this in this pipe flow example. So, I am looking into data at a particular XYZ location and seeing recording the turbulence data. So, in this single point statistics let us say you are collecting data at a particular XYZ, at a particular XYZ location.

When you do this we have already seen the fluctuations about the mean variance and so on. So, this special parameter appears here you let us call it u‘ u‘ over bar v‘ v‘ over bar w‘ w‘ over bar. right, the three variances for u, v, w velocity component. This is called covariances, they vary together.

So, this is called covariances. And then you get the other t terms in the cross terms or u‘v‘ bar, u‘w‘ over bar and v’w‘ over bar. These are cross covariances or simply cross-correlation terms. Most popular is to use cross-correlation ok, or simply sometimes used as correlation terms. Because you are looking into correlation of one velocity fluctuation, one random component with the other component u‘ with the v‘ at a particular point to see whether they are varying together or not. So, this I already discussed if this covariances are zero, you are essentially looking into a laminar flow.

These are the random component that has to be there. In a turbulent flow, they are positive. It is squared, u‘ square, right? So, it is always positive. ok, So, these are always positive, the covariances, but the other ones this can be positive, zero or negative. Zero of course that means it is not correlated.

So, these terms it can be positive ok, positive implies that you have, let us say both the components are varying together one increases the other one is also increases. Let us say u‘ example: if u‘ increases, v‘ also increases at the same point at another xyz location, they may not be. It is different from position to position. So, positive implying u‘ increases, v‘ also increases.

So, zero implies no correlation. Both are not correlated at all. Negative implies you have let us say u‘ increases, v‘ decreases. an opposite trend it could be v‘w‘ also or u‘w‘, I have just given an example here. And this will have a profound meaning when we go to turbulence this correlation terms as well as the covariances we use a different terminology when we go there, I will explain what it means. So, zero does not mean that, like u‘ increases, v‘ has to be constant.

Like, it can be random. So, like in one instance it was increasing, in another instance it can also decrease. Like for a single point it can increase or decrease or remain constant, but it does not have to mean that if u‘ is increasing it has to stay constant. They are basically exhibiting no correlation with each other. Exactly same as let us say zero in the covariance terms. So, u‘ v‘ zero is also not a good condition in, I mean the zone where correlation is not there that is also indicates that there is no turbulence there.

Yes, you can also do that this is a single point statistics we are looking into this u‘ is a function of x, y, z time. And the v‘ here is also the same location x, y, z and at the same time t. That is what we are looking into. And it is ensemble averaged.

So, this is if I write it differently. So, this is basically $u'(x, y, z, t), v'(x, y, z, t)$. And of course, his one question is that whether we can get correlation in time. This is useful in this pipe flow example I said like how much time you have to wait. So, you need to wait till the correlation has to go to zero. So, before that let us define what is called a correlation coefficient.

So, let us define what is called a correlation. correlation coefficient. So, let me call this as rho uv I am looking into the correlation of this or let us say u‘v‘. So, then this is defined as your correlation term and it is normalized using its own rms which is u‘ square average square root v‘ square average square root or you can write this as u‘v‘ bar over u rms v rms. So, this value of course, will take only value from 1 to minus 1 is a correlation coefficient we are looking into.

So, this value will be minus correlation coefficient will go from positive correlation or anti-correlation. Now, this is still at a single point statistics. If we do this in time that is the question one student has asked can we get correlation in time? Yes, this is useful to know how long I have to wait to get a new sample.