Measuring Theoretic Probability

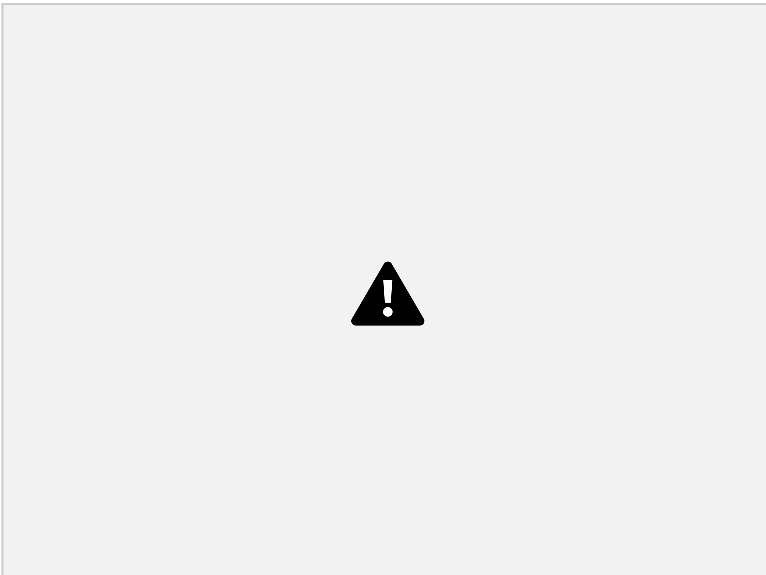
Professor: Suprio Bhar
Department of Mathematics & Statistics,
Indian Institute of Technology, Kanpur

Lecture 32

Sets of measure zero and Measure Theoretic Integration

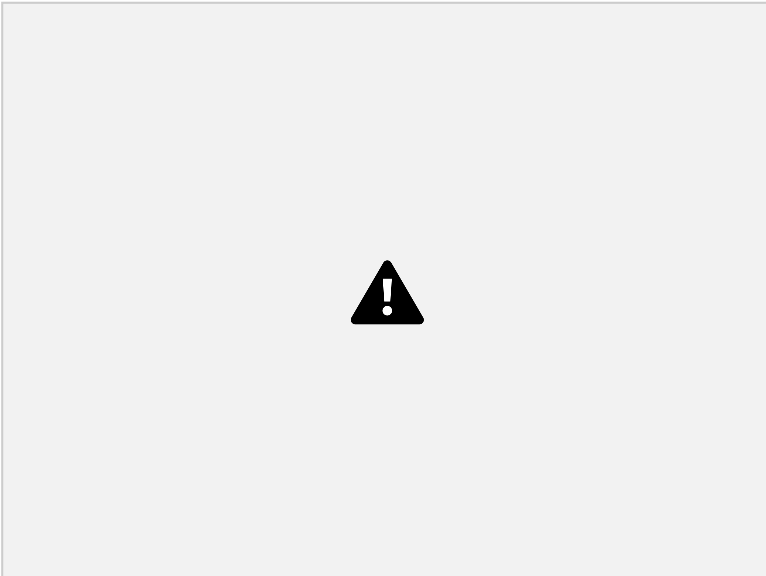
Welcome to this lecture. In this week, we are focusing on the limiting behavior of measure theoretic integration, and we are also going to see applications of this in obtaining other major properties. So, in the first lecture of this week, we have already seen proof of the monotone class theorem, and as an application, we have seen a brief sketch of the proof of the additivity of the integration. So, we are continuing that discussion, and we are going to look at the structures that follow when we focus on the limiting behaviors. So, let us move ahead and start the discussion with the notes.

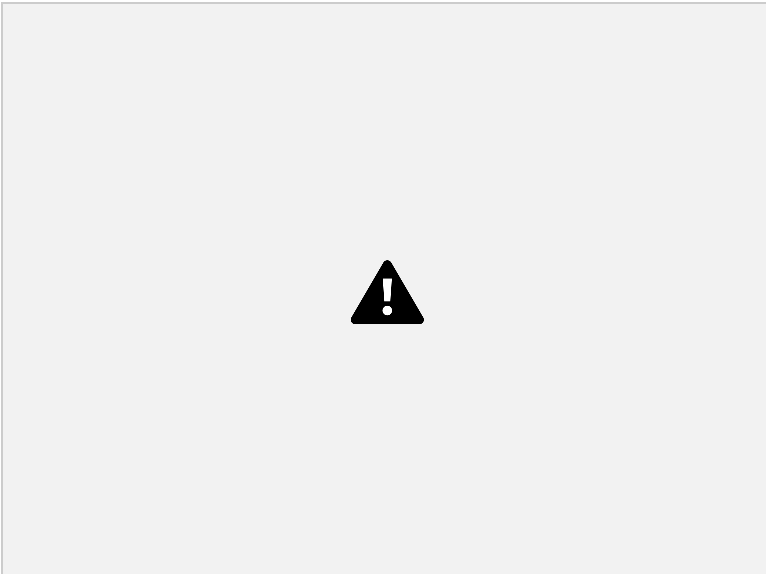
(Refer Slide Time: 00:41)



So, in this lecture, we are going to discuss the effect of null sets in integration. So, what do I mean by null sets? Null sets mean, the sets of measures 0 in the given σ-field. So, for the purpose of the discussion we are going to fix this measure space $(\Omega, \mathcal{F}, \mu)$.

(Refer Slide Time: 01:24)

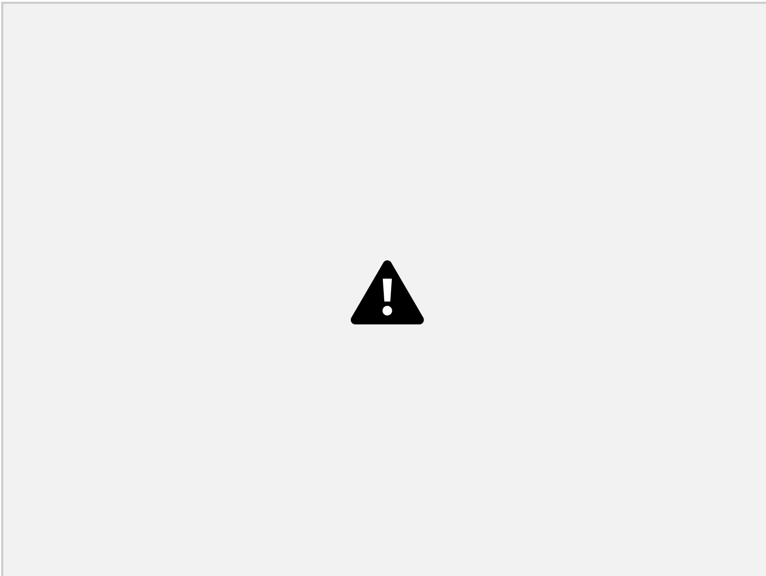




First definition is about certain properties holding μ almost everywhere. What does it mean? Property P, $\wp$ is said to hold μ almost everywhere if there exists a null set in your σ-field. So, that means, that it has measured 0, such that, outside the null set on the complement of the null set the property $\wp$ holds at that points ω.

So, therefore, you are going to obtain some null set outside which you are going to get the required property or you can verify the property on those points. So, shorthand notation for this kind of a situation will be μ a . e . μ almost everywhere. So, that will be short hand notation that we are going to use.

(Refer Slide Time: 02:26)

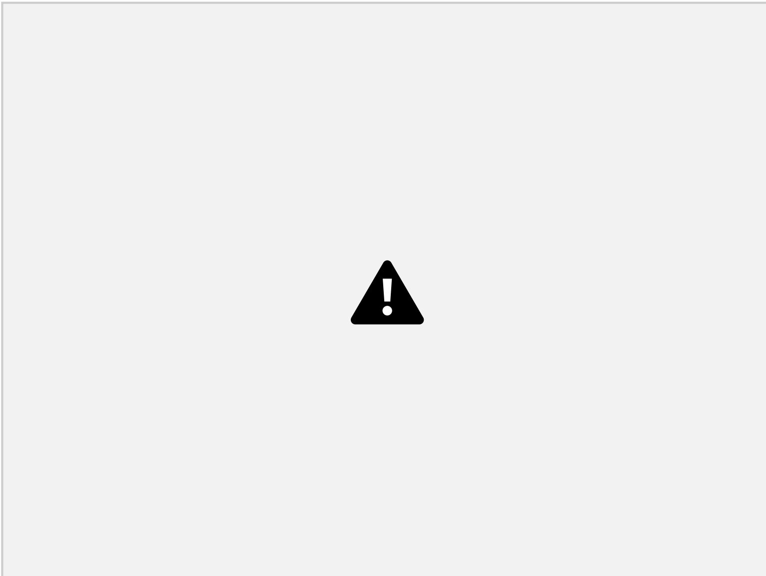


So, let us first try to understand what do these almost everywhere statements look like and what do they mean. So, consider some function h, it need not be measurable, but suppose, you know that there is a function h, such that, on the points where h takes negative values, that set has measured 0. Suppose you are given this fact, then you are going to say that $h \geq 0$, h is non-negative, μ a. e.

So again, on the complement of this set, you get the other condition which is $h \geq 0$. So, that is what you are looking for, you are looking for such null sets outside, which your property or the statement hold. A similar situation might involve two functions or more functions. So, let us look at this situation.

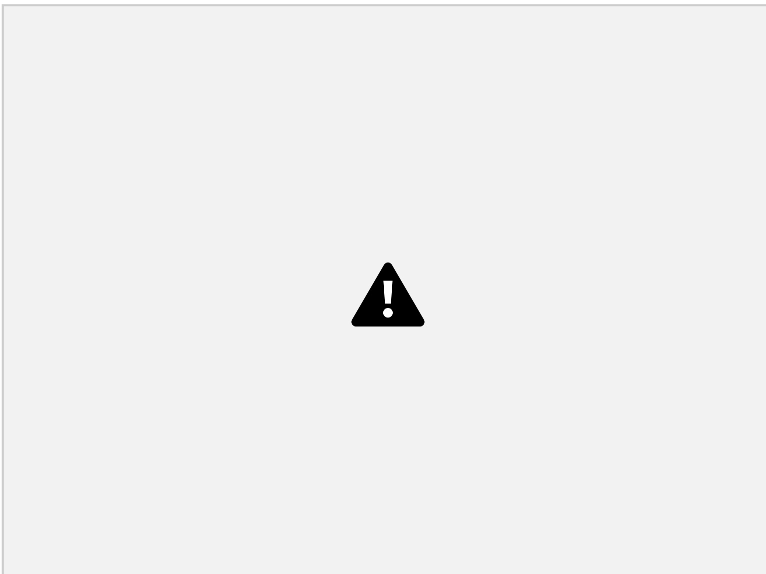
So, again, let us consider functions f and g, again, this need not be measurable. As long as you know that the set of points where $f(\omega) > g(\omega)$, if that measure is 0, then, what you say is that $f \leq g$, μ a. e. So, therefore, as long as the measure of the complimentary event gets 0 mass or if the complimentary event is contained in a set of measures 0, then you are going to say that this happens almost everywhere.

(Refer Slide Time: 04:02)



However, you have to introduce a separate terminology for the special case when you are dealing with a probability measure. So, suppose, you are dealing with a probability measure and you are going to talk about this almost everywhere statements. So instead of using this term, almost everywhere here, what do we prefer to use is the term almost surely or $\mathbb{P}$ a.s. If the probability measure $\mathbb{P}$ is clear from the context.

So, once you have fixed it beforehand and then started the discussion, then you might just go with almost surely or a.s. You do not have to mention the probability measure explicitly. But we are going to focus on such situations, and we are going to look at what is the effect of such sets in integration.



(Refer Slide Time: 04:51)

So here is the first result. So, consider a measurable function now, call it h. It is real-valued function. Such that $h = 0$ a. e. So, then, we are going to claim that $\int h \, d\mu$ is 0. So, to prove such a statement we are going to follow the usual procedure the standard procedure that you are going to work with indicators of simple functions then you go to non-negative measurable functions and finally, to general measurable functions.

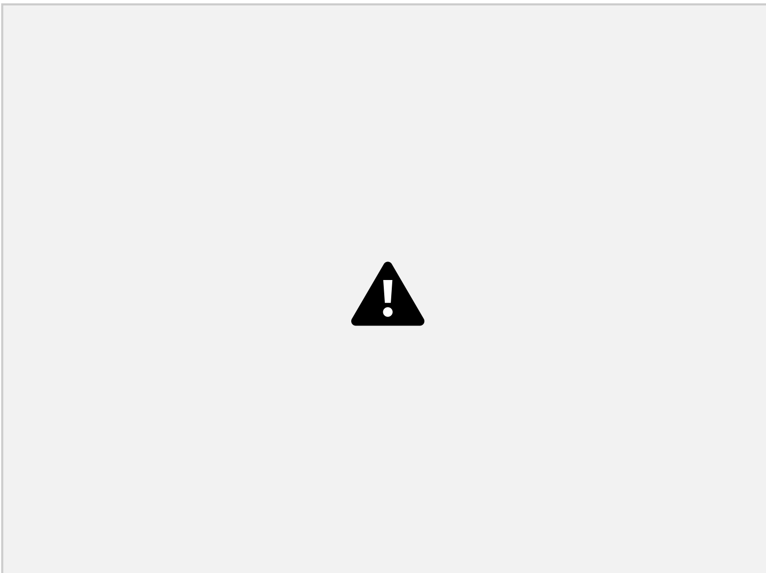
So, let us start with the case of a simple function. So, let us see what happens here. So, you have a linear combination where you have $h = \sum_{i=1}^{n} x_i 1_{A_i}$, that is a standard representation for a simple function. Now, if it so happens that for some value x_i, you have that is nonzero. So, suppose you have that then you observe that the corresponding set A_i will be contained in the set of points where h takes nonzero values. Why?

First of all, A_is are taken to be pairwise disjoint. So therefore, on the points inside A_i, these indicator will only contribute other terms will not contribute. So therefore, on those points, the value of h will be exactly x_i. And if x_i is nonzero, you get the value of h is nonzero there. So therefore, if for some i, x_i is nonzero, then the set A_i is contained in the set of points where $h(\omega)$ takes nonzero quantities. Then, use the fact that, μ measure associates 0 mass to such a set, you use the fact to claim that $\mu(A_i)$ is 0 whenever the x_i is nonzero, so you have that conclusion.

Now let us look at the $\int h\, d\mu$. So as per definition, we have to make sense of this summation first, that you are going to multiply the scalars x_i by the size or the measure of the sets A_i. But here something interesting happens. Either, the value of x_i is 0, in which case the product is 0. Here we are using the fact that $0. \infty$ is taken as 0. Otherwise, if x_i's are nonzero they are some real numbers, but you are multiplying sets of measures 0 here, so you are considering sets of measure 0 here, $\mu(A_i)$ is 0. So therefore, $x_i\mu(A_i) = 0$.

So, therefore, all the terms here are 0. If you sum them up, you still end up with 0. So, therefore, what you have obtained is that $\int h\, d\mu$ exists and s 0. So, here we first made sense of this summation as per the definition.

(Refer Slide Time: 07:58)

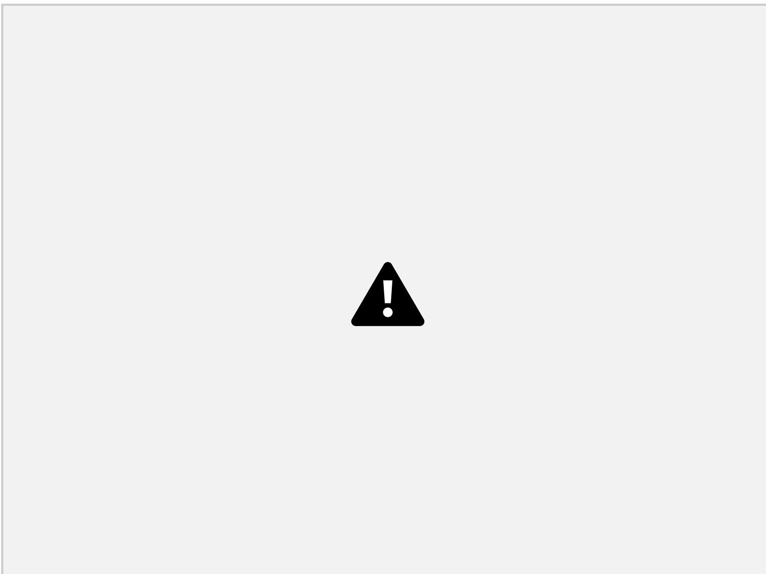


Let us move ahead, and consider the case when h is non-negative and measurable. In this case, what do you do? Follow the definition, take any simple function s, which falls between 0 and the function h. In this case, observe that the points where the simple function takes nonzero values on those set of points, each will also take nonzero values. Why? Because you are dealing with functions h and s to be non-negative.

And if you know that s takes strictly positive values, so these are the points exactly where s takes nonzero values. So, on those points h also must take nonzero values, as it is taking at least that much value on those points. So, it is above the function s. So, therefore, you have this inclusion. But now, since the set of points where h takes nonzero values, that set of points has 0 mass, you will immediately claim that the set of points where s takes nonzero values that also has 0 mass. And, hence, you now apply the result that you have obtained in the previous paragraph for simple functions.

You immediately claim, that the integration of such a simple function must be 0. But you have chosen the simple functions to be below h, such that, it falls between 0 and h. And for all such simple functions, you have no claim that the integration is 0 and $\int h \, d\mu$ by definition is the supremum of all these quantities, therefore, it turns out to be 0. So therefore, you have shown that $\int h \, d\mu$ here it is defined, and it takes the value 0.

(Refer Slide Time: 09:49)



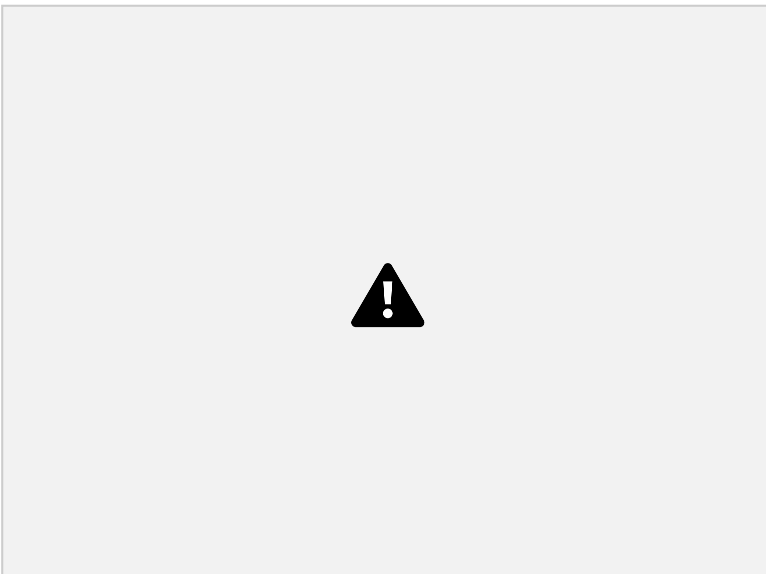
So, let us move ahead to the case of general measurable functions, which takes now real values it can take positive values or negative values. So, you follow the definition and split the function h into its positive part and negative part. So, again, observe that the set of points where h takes positive values on those points you know that h^- the negative part will be taking 0 value. This identification has been discussed earlier.

Therefore, on the points, where h^+ is strictly positive on those points $h(\omega)$ must be nonzero. Similarly, you can make the same statement for h^- that the set of points for h^- takes nonzero values it is contained in the set of points where h takes non-zero follows, but you are given that the set of points for h takes nonzero values that has measures 0 and hence, he will claim that both for h^+ and h^- the set of points where these functions takes nonzero values that has measures 0.

But since h^+ and h^- separately they are non-negative and measurable functions. So, therefore, using the previous steps argument you will claim that the $\int h^+ \, d\mu$ and $\int h^- \, d\mu$ both must be 0. And hence, as a difference of these two quantities h will have a integral, so $\int \, h d\mu$ exists and takes the value 0. So, to complete the proof for all the type of functions that you consider in measure theory.

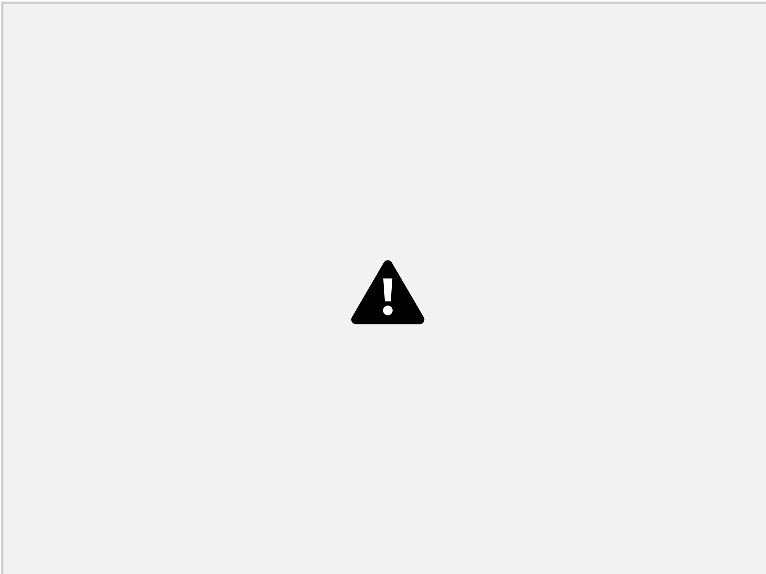
So, one important step that we would like to highlight is that in the above proof in each step we have established the existence of $\int \, h \, d\mu$ and then we are showing that $\int \, h \, d\mu$ is actually. So, let us just go back to the statement just to highlight this fact.

(Refer Slide Time: 11:58)



So, in the statement, we are not saying that h has integral or $\int hd\mu$ exists, we are directly claiming that we are starting off with a measurable function so we do not know whether $\int h\, d\mu$ exists, but it is a part of the statement that $\int h\, d\mu$ exists and takes the value 0. So, this clarification is important.

(Refer Slide Time: 12:24)



So, let us see what happens if you try to use this result, and obtain further nice properties. So, here you start off with two measurable functions g and h now. And suppose it happens that $g = h$ μ a. e. So, that means, the set of points where inequality holds will be contained in a set of measures 0. So that is good.

So, therefore, you can now try to say about their connection involving the $\int g\, d\mu$ or $\int h\, d\mu$. So, here the statements says, that if one of the integrals exists, then so does the other and they must be equal. So, how do you show this?

(Refer Slide Time: 13:22)

So, suppose you start by assuming without loss of generality, that $\int g \, d\mu$ exists then you consider the function f, which is $g - h$. So, g and h separately they are given to the measurable functions, so therefore, they are difference which is f is also measurable. But then you are given the fact that $g = h$ almost everywhere and therefore, you will claim that f must be 0 almost everywhere.

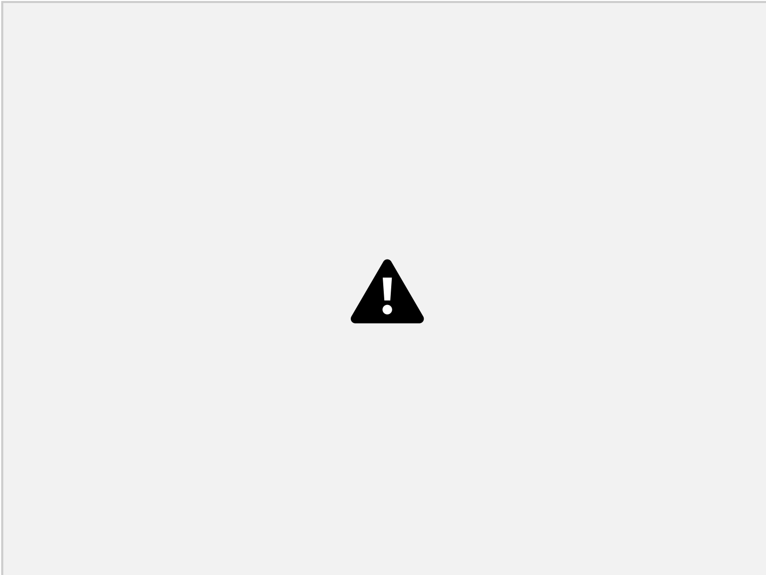
So therefore, by the previous proposition, you will immediately claim that the $\int f \, d\mu$ exists and it is 0. Now, look at the fact that h now can be expressed as $g - f$. And as such, you are going to use the additivity property to claim that $\int h\, d\mu$ exists. Why? You have assumed $\int g\, d\mu$ exists and you have shown that $\int f\, d\mu$ exists, so therefore, their difference makes sense, and will imply the existence of $\int h\, d\mu$.

But then, what is the value? The value is exactly the difference of the two integrations, but the second integration is nothing but 0, so therefore, you get back the value of $\int g\, d\mu$ itself. So, therefore, if one of the integrations exists, and the functions match almost everywhere, you get the same value for the other functions integral. So, this is a very important property.

Now you can try to extend this kind of a notion when you are dealing with external real-valued functions and their integrations. So, suppose, you consider g and h to be extended real value, and Borel measurable then consider μ integrable functions. So again, here you are assuming that $\int |h|\, d\mu$ is finite. Then the statements says that the set of points where $|h|$ takes $+\ \infty$ as its value that has measure 0.

So, and restatement of this, in terms of the function h is this where the function h takes $\pm\ \infty$ as its value that set as measure 0. Now, there is a way to state this in terms of almost everywhere statements, and you are going to say, that h is finite μ a. e. By that we mean that h takes these infinite values $\pm\ \infty$ on sets of measures 0 or such possibilities are included in our set of measures 0.

(Refer Slide Time: 16:07)



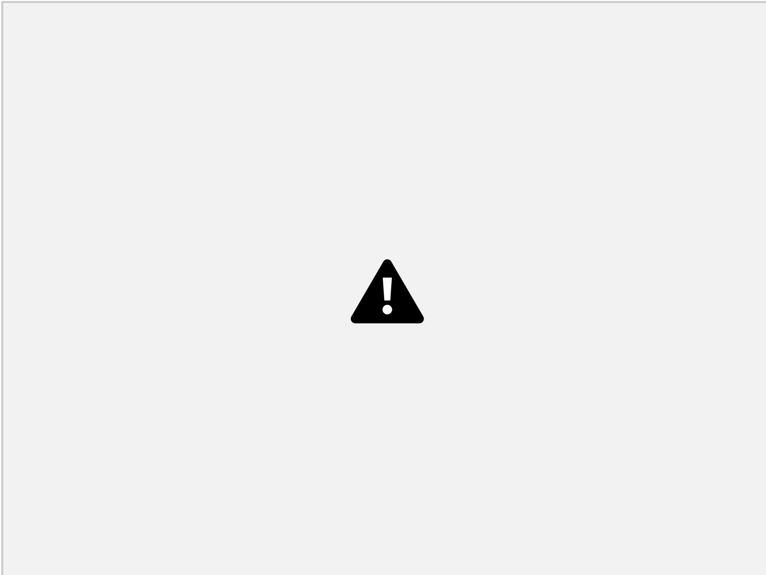
We prove this by contradiction. So, assume, that the set of points consisting of the situations where $h(\omega)$ takes ∞ values, so that set of points, if it has positive mass under the measure μ then consider the $\int |h|\, d\mu$. So, here, $|h|$ is a non-negative function, and therefore, its integrals exists could be $+\ \infty$, but let us look at this lower bound.

So, we consider the multiplication by this specific indicator. So, this is exactly the set where $|h|$ is taking value $+\ \infty$. So, we are just multiplying by $1_{(|h|=\pm\infty)}$, and therefore, $|h|\ \geq |h|\ 1_{(|h|=\pm\infty)}$. So, we are using that inequality for the integrations. But on this set, $|h|$takes the value $+\ \infty$, so it is just a simple function here.

So, you get this value while you compute the integration as per the definition. So, this is the value of the function multiplied by the size of the set. And that turns out to be ∞ since, the measure of this set is assumed to be positive. Therefore, we ended up with the fact that $\int |h|\, d\mu$ is infinity, but this will contradict the fact that h is integrable.

Remember, h is integrable means, that $\int |h|\, d\mu$ is finite. So, you have a contradiction, which arise in from this condition, which we assume that this set of points has positive mass. So, therefore, this set of points must have 0 mass, and that will tell you or give you the result.

(Refer Slide Time: 17:54)



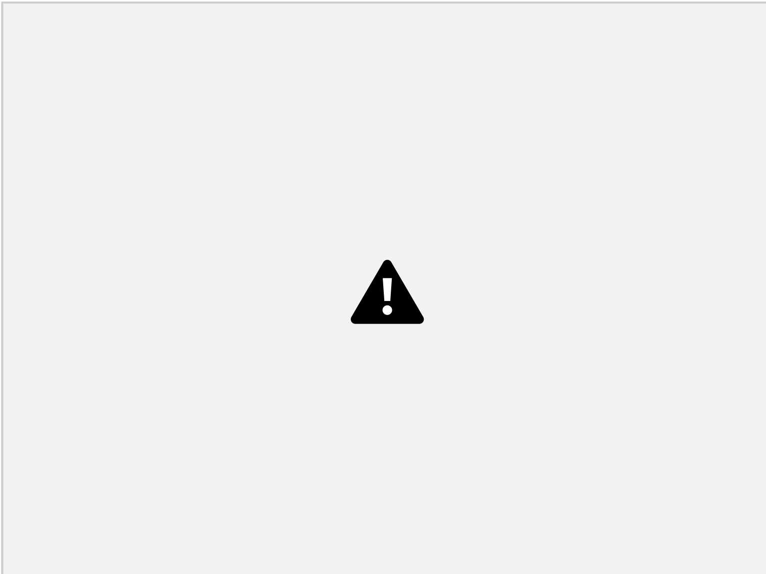
So, let us move forward and use this observation in the next result. Consider a non-negative function h now, so then, you can always consider the $\int h\, d\mu$. But if it so happens that the $\int h\, d\mu$ is 0, then we can claim that $h = 0$ μ almost everywhere. So, again, how do you approach this. So, it is given, that h is non-negative, so you do not have to worry about h taking negative values. So, forget about that situation.

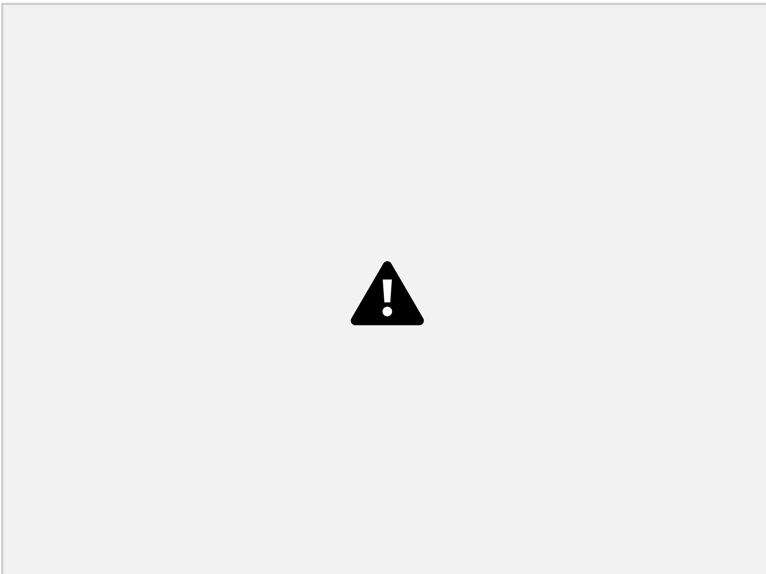
So now we only have to consider two situations when h takes the value 0, and h takes positive values. So, now, you want to show that $h = 0$ μ a. e. so it is therefore enough to show that h takes positive values on a set of measures 0, so that is the target. Now this set, instead of directly working with it, considered a slightly different set, so for integers n 12,3,4, and so on, consider this sequence of sets where $h > \frac{1}{n}$.

So, you just considered this set of values. So, we first claim that these set has 0 mass under the measure μ. If the claim is true, then observe that the set h taking positive values that set has this approximation from below by the sets $(h > \frac{1}{n})$. So, that is all you have to observe and use the continuity from below of the measure μ.

Now if it so happens, that μ assigns mass 0 to $h > \frac{1}{n}$ type sets, then the limit value is also 0. And that will tell you that hh takes positive values on a set of measure 0. So that will end the argument as soon as you prove the claim. So, you want to prove that for all natural numbers n, $h > \frac{1}{n}$, that set of points will get 0 mass.

(Refer Slide Time: 20:02)



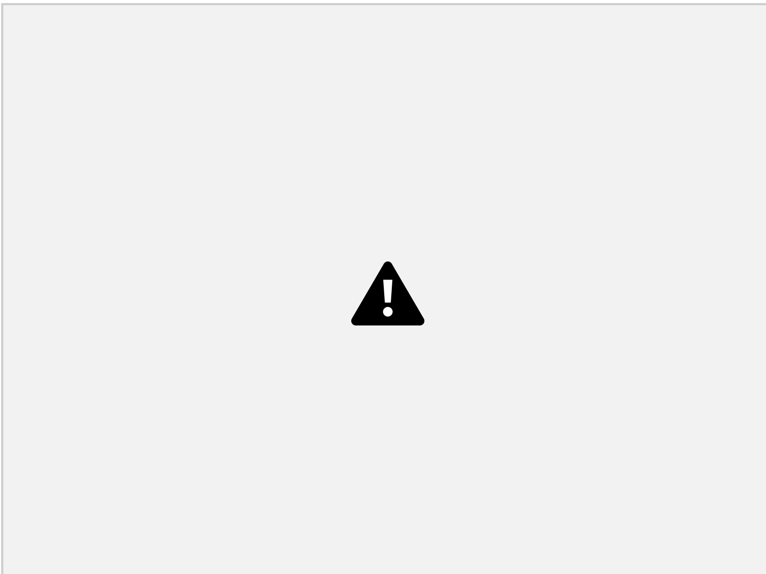


Again, use a contradiction argument. So, this is a proof by contradiction. So, suppose for some natural number n naught, it so happens that the measure of this set is positive then try to compare the functions h and h times the indicator of this set. So, here we have assumed that the measure of this set is positive. But observed that h is non-negative so, therefore, inequalities pointwise holds and therefore, the same inequality holds for the integrations.

But on this set $h > \frac{1}{n_0}$, so you get this lower bound for the integration. Remember, it is always greater equals to here. So, now, you put that lower bound here $\frac{1}{n_0}$ and then consider the indicator function. So, that is all we are writing down, but then this is no a simple function, write down the integration so, that is positive.

So, now, what does it tell you? It says that $\int h \, d\mu$ is strictly positive. So, that will contradict the hypothesis. And, therefore, this contradiction is arising from the assumption that for some natural number and not measured up this set is positive. So, therefore, we must have that for all possible natural numbers measure of $h > \frac{1}{n}$ is 0, so we must have that.

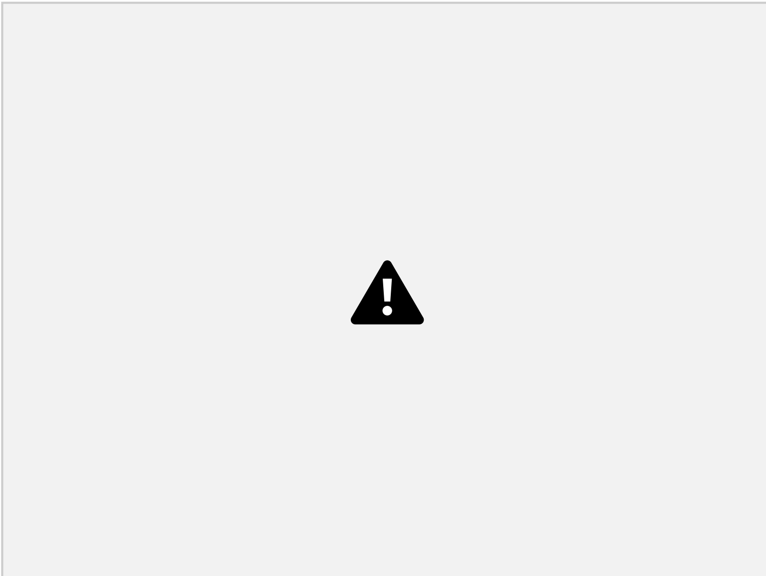
(Refer Slide Time: 21:28)

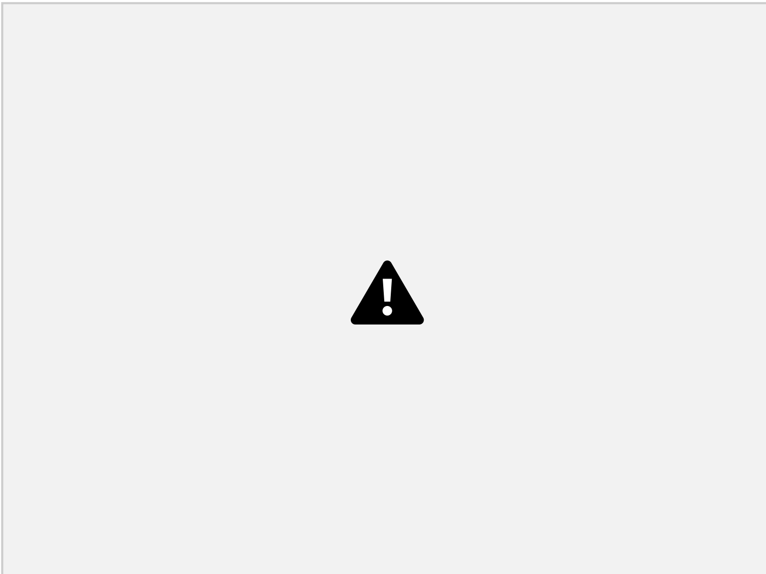


Now, you use this while comparing two functions. So, if it so happens that you take two measurable functions h and g and you observe that $|h - g|$ that is a non-negative function integration of that is 0 then you will claim that $h = g$ μ almost everywhere. So, how do you prove this? So, here, you take the function f to be $|h - g|$. So, this is now a non-negative function, which integration is 0.

So, therefore, by the previous part you will immediately claim that f must be 0 μ a. e. But that will immediately tell you that on this set of points where the function f is taking value 0, there you must have $h = g$ that is pointwise happens. So, therefore, outside the appropriate null set, you must have $h = g$, so that completes the proof.

(Refer Slide Time: 22:25)



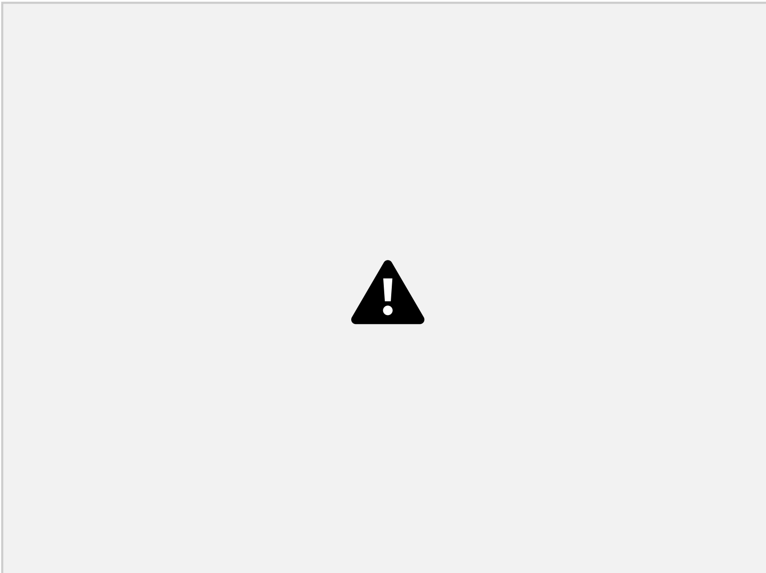


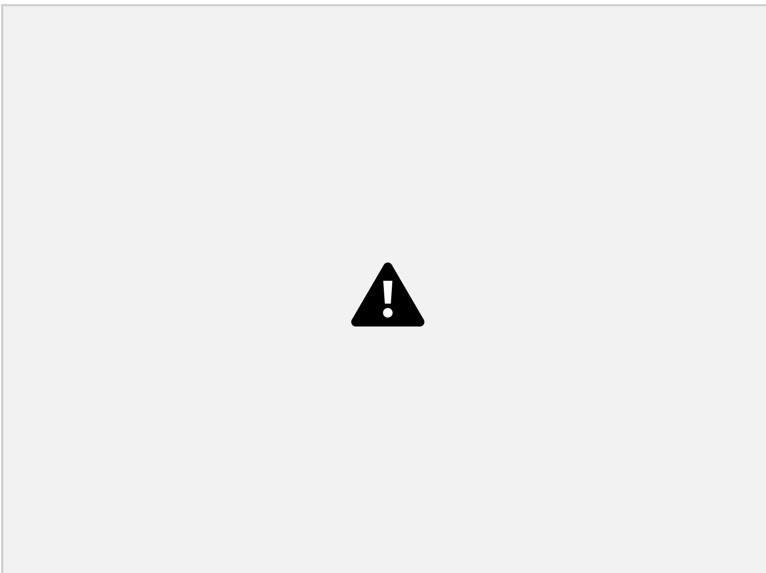
But now here is an interesting question. So, suppose it so happens that instead of assuming that h is non-negative throughout you assume that h is non-negative μ almost everywhere, and take it together with the fact that $\int h\, d\mu$ is 0, then, can you show that $h = 0$ μ a. e. So, this is the only difference in the hypothesis.

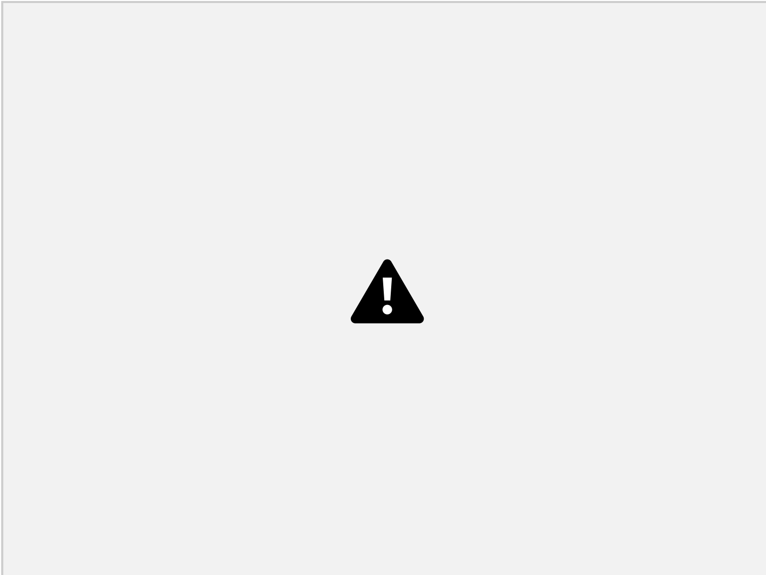
Now, that we had earlier assumed that h is non-negative throughout, but now we are saying h is non-negative μ almost everywhere, but we are taking the $\int h\, d\mu$ to be 0 then can you now claim that h is 0 μ one most everywhere, take it as an exercise. But an important fact that we have understood from all this discussion is that the hellos have a measurable function on a set of measure 0 will not affect the value of the integration.

So, you can forget about the values of the function, whatever they are on a set of measure 0. And more generally, in the setup of this proposition where we had done all these compositions of functions and so on, if we are only interested in the integration part, then the equality $g = h$ μ almost everywhere may be treated as an equality of functions. So, let us just go back to this proposition 2 once more.

(Refer Slide Time: 23:53)



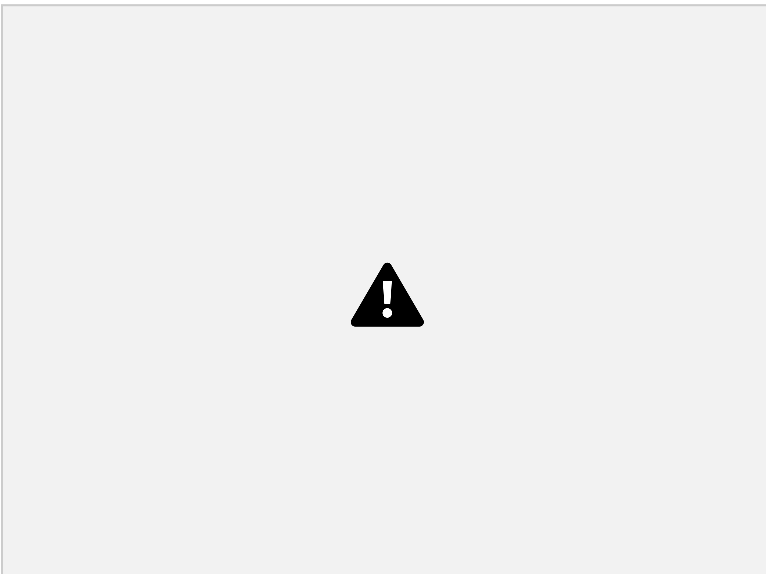




So, in proposition two, we stated that if two functions are measurable, and they agree almost everywhere, if one of the integrals exists, then so does the other, you also get the equality of the integration. So, if you are only interested in the equality of the integration and existence of the integration, then, you can as well treat the functions g and h to be the same in the setup.

So, if two functions agree, and you are only interested in the integration value, then you treat the two functions to be the same. So, in the setup of this proposition 2, you are going to consider that two functions agreeing almost everywhere is almost equivalent to equality of functions.

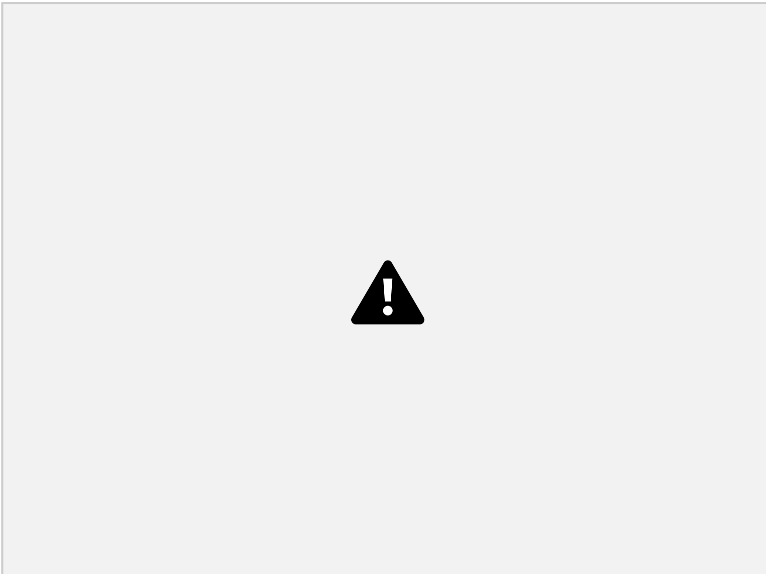
(Refer Slide Time: 24:39)



But we have also mentioned in proposition 3, that you can talk about this external real valued functions, but they are integrable, then this is essentially real-value. So, we said that this is finite valued, finite valued means it is real value. So therefore, it can only take the values $\pm\ \infty$ on a set of measure 0. It cannot take $\pm\ \infty$ on a set of positive measure.

So, therefore, here also you can treat it as a real-valued function, as long as you are dealing with integrations. So, as long as the function is integrable then the function cannot take $\pm\ \infty$, as its value on a set of positive measure.

(Refer Slide Time: 25:23)

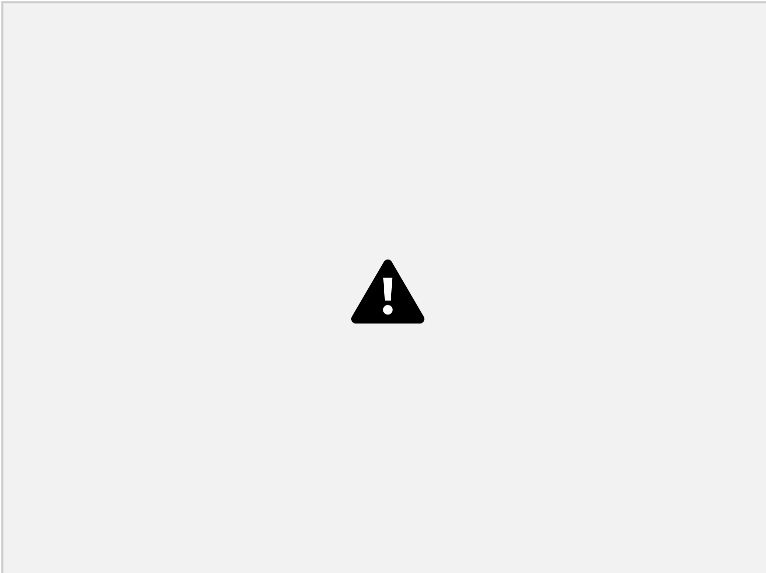


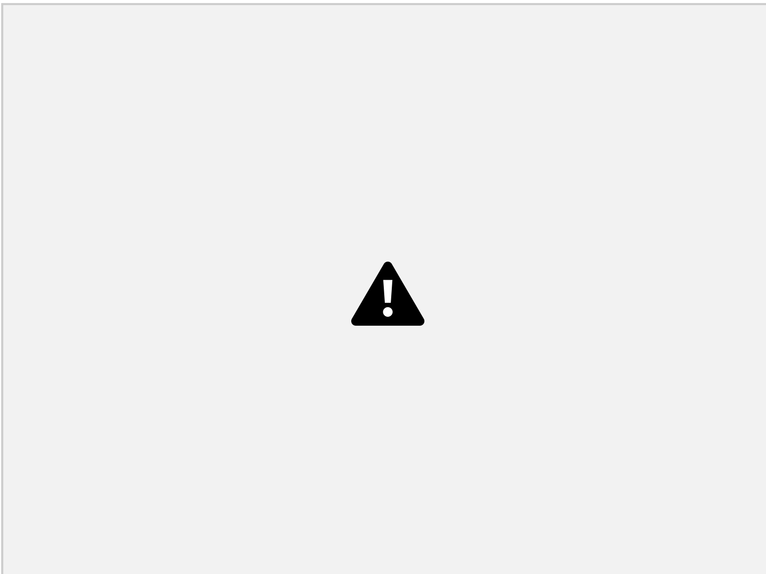
So, with that understanding let us now consider the case of random variables. So, we have so far discussed the situation for general measurable functions, but suppose you want to discuss things when the measure, the given measure is a probability measure. So, highlight that problem dimension here and you consider the random variable X.

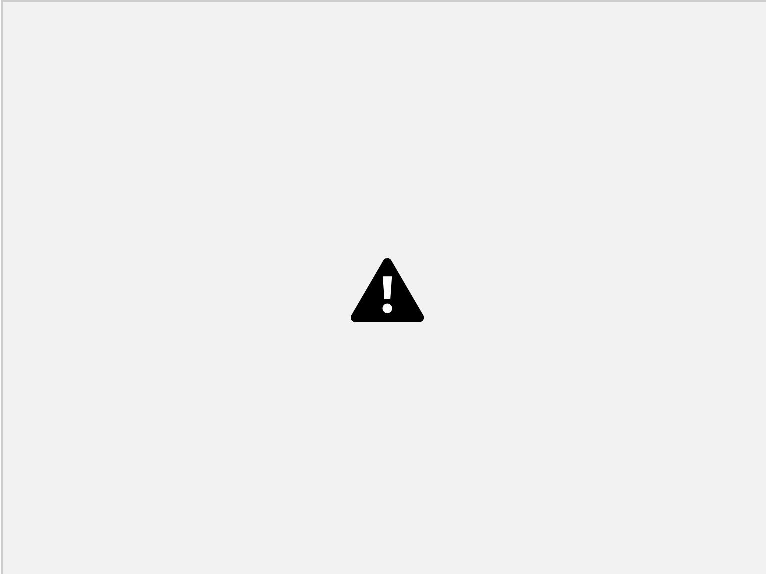
If it so happens that you get this bound point wise that there exists a constant $C > 0$ such that $|X(\omega)| \leq C, \forall\omega$ is dominated from above by this constant for all points, then you get this bound for the integration and therefore the expectation. So, $\mathbf{E}\,|X|$ is $\int |X|\, d\mathbb{P}$, but if the function $|X|$ is dominated from above by this constant function, C, you observe this inequalities and find that $\mathbf{E}\,|X|$ is finite. So therefore, X is integrable in this situation.

So, if X is bounded by such constants, X is integrable. However, in the discussion above, we have considered these kind of inequalities in the almost sure framework. So, whenever you are talking about proclamations you talk about instead of saying almost everywhere you say almost surely, so a.s.

(Refer Slide Time: 26:42)





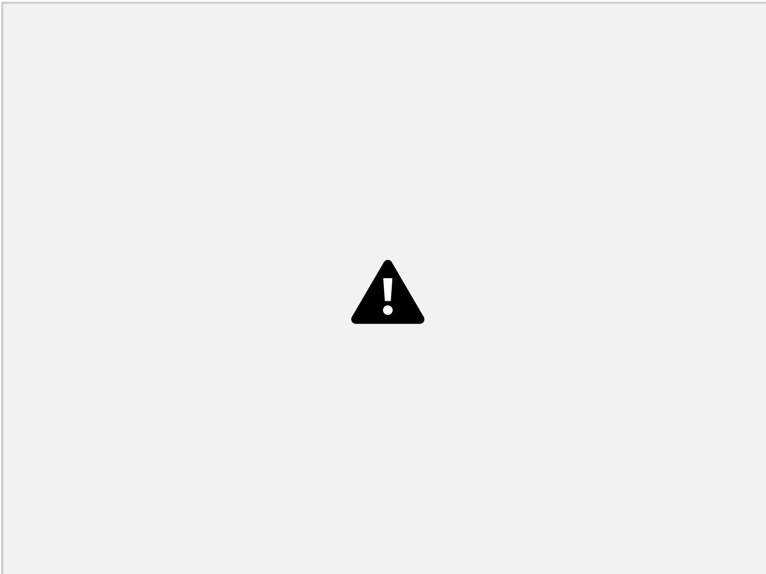


So, suppose this happens then for again an appropriate $\mathbb{P}$ null set you get these appropriate statements that these domination holds, but what happens to this inequality now. So, here you already have a bound on n complement. So, if N is the null set, then on the complement of the null set you get the bound. So, that is what almost surely statements imply.

But you can consider this inequality that $|X(\omega)| \leq C.1_{N^c} + \infty.1_N$. Observe that this inequality holds now point for all points in the domain. Now, try to look at the bound function here. So, here you are using the constant value C on N^c and then value ∞ on the set of measured 0.

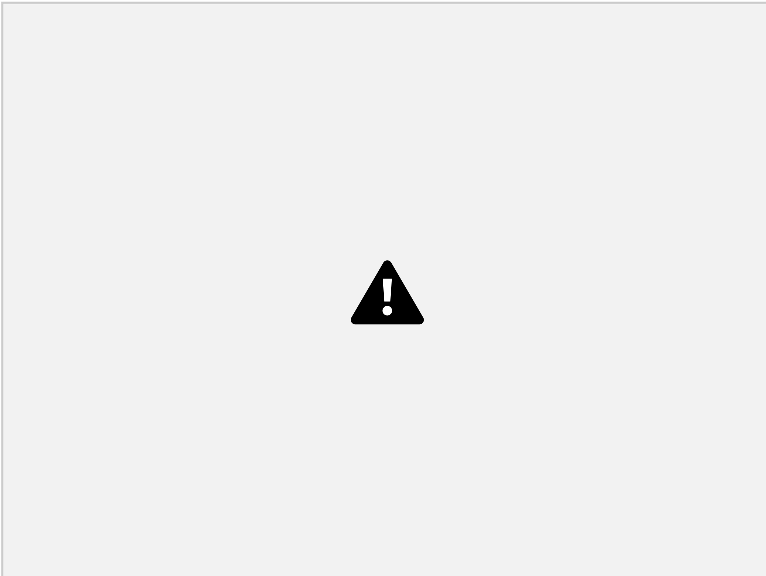
So, now try to show that this upper bound that you have constructed this function is integrable and that will tell you that X will also be integrable. So, verify that X is integrable if X is bounded from above by a constant almost surely. So, you can reduce the, this restriction that it has to be bounded point-wise to bounded almost surely. So, try to work this out this is left as an exercise.

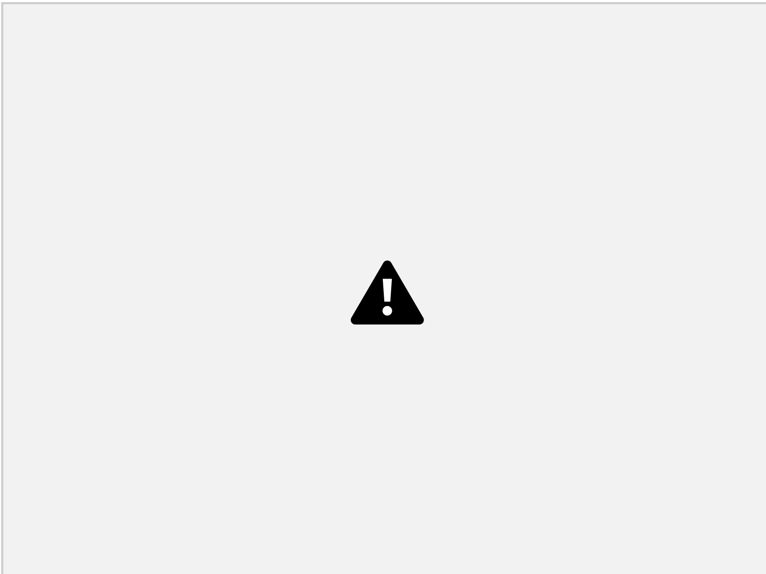
(Refer Slide Time: 27:58)



So, now here is an interesting comment. So, look at this very specific situation that Ω is a set of three elements $\{a,\ b,\ c\}$ and consider the Dirac measures portrayed at the point c here, δ_c. Now, spatially you consider this σ-field $\mathcal{F}=\{\Phi,\ \{a,\ b\},\ \{c\},\ \Omega\}$ these four sets. So, you can immediately verify that this is a σ-field. Now consider that Dirac measure here, and consider these two special functions. So, g and h, so $g = 1_{\{a,\ b\}}$, but take the function h like this, that you define the function value $h(a) = 1$, $h(b) = 2$, $h(c) = 0$. So, we will see the reason for these choices.

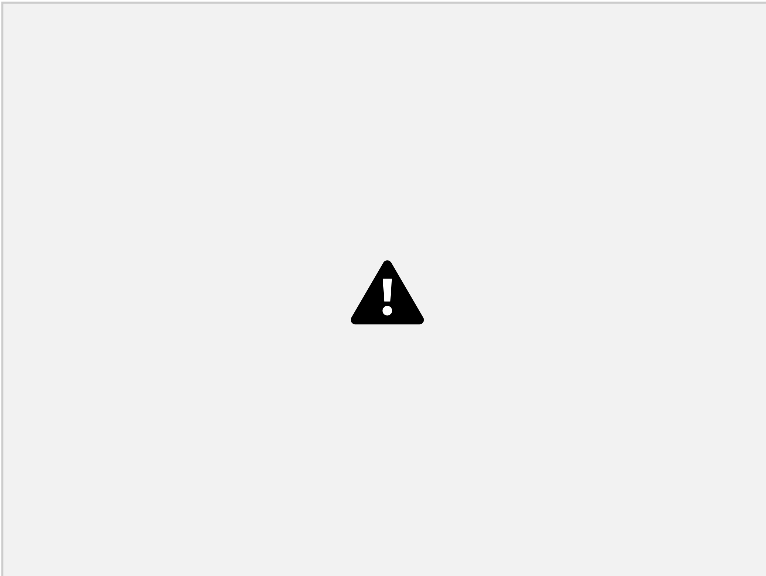
(Refer Slide Time: 28:51)

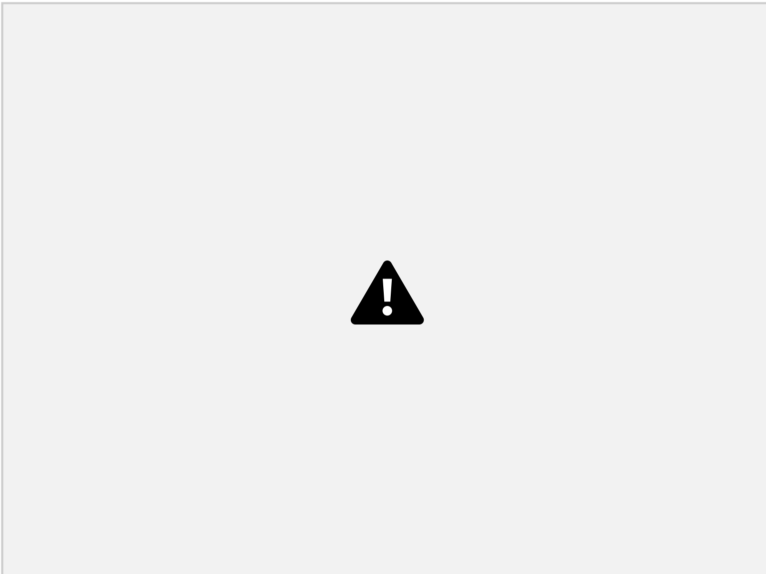




Now, observe here that the δ_c will not put any mass on $\{a, b\}$ so that will be 0. Here, one important comment, is that, we have not included the subsets, $\{a\}$ and $\{b\}$ in the σ-field. So, remember, our σ- field is made up of the empty set, the doubleton set, and the singleton set and the whole set Ω. So, the singleton sets $\{a\}$ or $\{b\}$ are not there in the σ-field.

(Refer Slide Time: 29:18)





So, here, I have this set of measure 0, which is this doubleton set, but its subsets $\{a\}$ and $\{b\}$ singleton sets are not included in the σ-field. So, what we have observed is that subsets of null sets, here the doubleton set is a null set, the subsets of it need not be in the σ-field. So, this gives you an interesting observation. But now consider this situation for the set of points where g and h does not agree.

(Refer Slide Time: 29:50)

So again, let us just go back to the values of g and h. So, g takes the value 1 on the points a and b, takes the value 0 on the point c. So, therefore g and h defers only on the point b, all the other points a and c, the functions g and h match. So, only at the point b the functions do not agree. So, if you consider the set of points where g and h do not agree so that will be the singleton set $\{b\}$, but this is the subset of this null set that we have just observed. But this specific null set is not having this property that all its subsets are also in the σ-field. So, specifically, this subset of singleton set $\{b\}$ is not included in the σ-field.

(Refer Slide Time: 30:34)

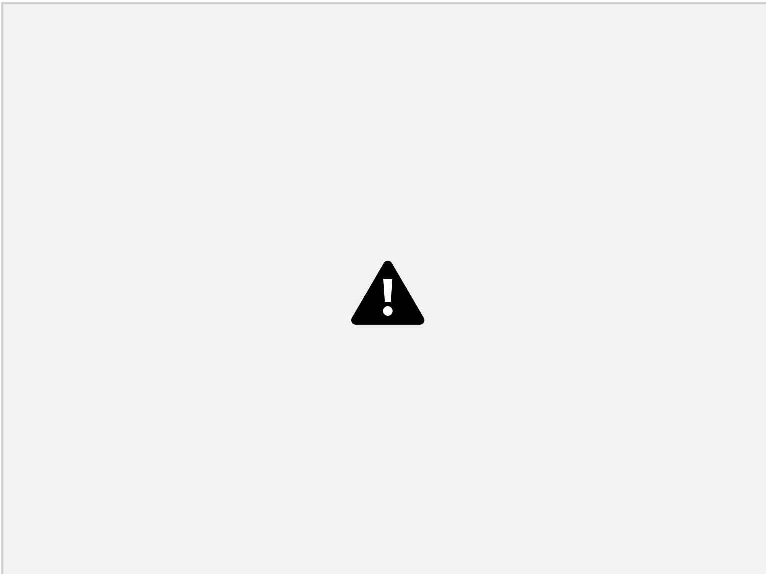
Now, by definition $g = h$ almost surely, Dirac measure is a probability measure, so you use the term almost surely. So, it does not agree with these function h only on a set of measures. So, you can, as well consider this set $\{a, b\}$ that will be a superset of the points where g and h do not agree, and the super set has 0 mass.

But here you note that g, which is indicator of the doubleton set $\{a, b\}$ is measurable because the doubleton set is in the σ-field, but h is not a measurable function. So, observe this that g is measurable, g agrees with h almost surely, but h is not a measurable function. So, this is an important observation. So, please check this.

(Refer Slide Time: 31:23)

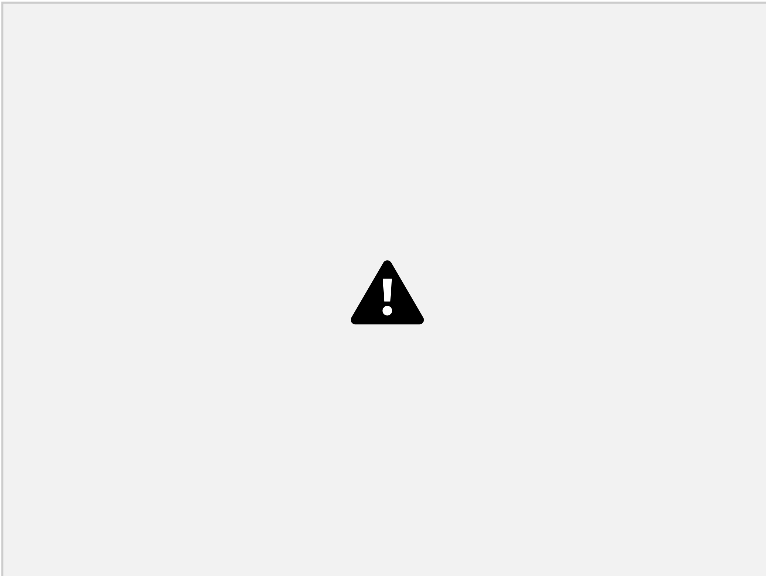
In continuation of this observation given a measure, we usually are going to enlarge the σ-field by including all subsets of null sets. So, typically it may not be true that all subsets of a null sets are already included. Now, if you include all these possible subsets of null sets, this procedure of enlarging the σ-field is known as the completion of the σ-field. So, by including all the null sets and by an appropriate construction you can get a σ-field once more, which will be containing a larger number of sets. Now, this is known as completion of the σ-field with respect to the given measure.

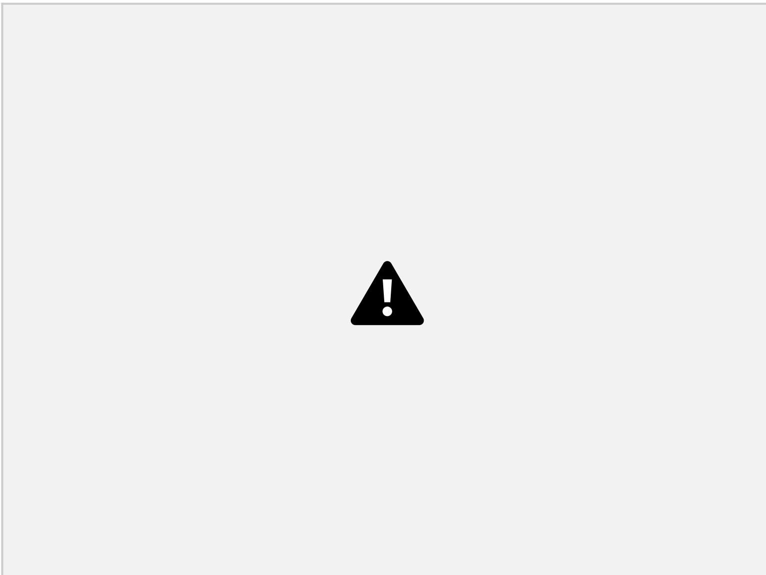
(Refer Slide Time: 32:05)

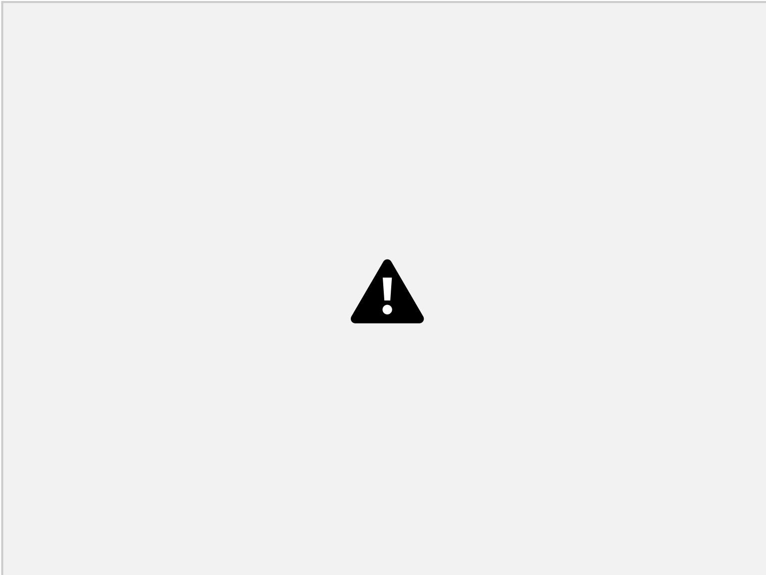


So, here what we are considering is the Dirac measure here, and we had observed that the singleton set $\{b\}$ was not there. If you include this by considering that these are subsets of measure 0 sets under the measure δ_c that will be the competition, if you include all those appropriate subsets.

(Refer Slide Time: 32:24)







So, here, we are not going to discuss this topic in further detail. We shall, however, briefly connect this topic in a later lecture. There we are going to talk about the Lebesgue measure on the Borel σ-field. So, take it as a fact that the Borel σ-field is not complete with respect to the Lebesgue measure, by that I mean, not all subsets of null sets are included. And what we are going to do is, to include all the subsets of measure 0 sets included in the Borel σ- field.

And if you include all those subsets, which are subsets of measures 0 sets, you are going to enlarge the Borel σ-field, and you are going to get a σ-field by an appropriate construction that is referred to as the Lebesgue σ-field. This latter σ-field now contains all subsets of the null sets with respect to the Lebesgue measure.

We are going to use this σ-field in as very specific construction, but otherwise, we are always going to work with the Borel σ-field. So, we are assuming the fact that Borel σ-field is not complete with respect to the Lebesgue measure by that, I mean, not all subsets are measure 0 sets are included. However, we are not going to discuss the completion structure, the procedure of completion in further detail. So, we stop here and we will continue the discussion in the next lecture.