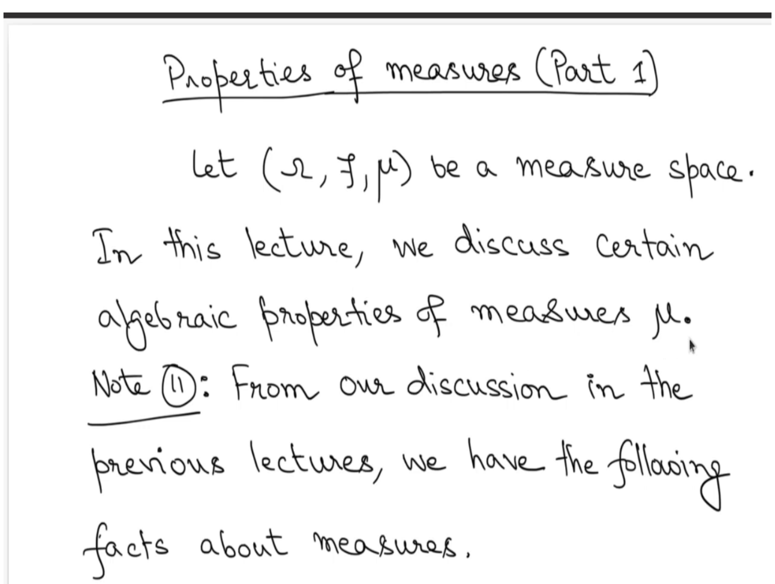
Measure Theoretic Probability 1
Prof. Suprio Bhar
Department of Mathematics & Statistics
Indian Institute of Technology, Kanpur
Lecture No. 08
Properties of Measures (Part 1)

Welcome to this lecture. So, as before, let us quickly recall what we have done so far. So, in this week, we are looking at measures which allow us to look at how likely certain events are in some random experiment or how large or important a set is in a general measurable space. We had looked at the probability measures, which associates to the whole set the total mass 1 and therefore, all the other subsets will have mass less than 1.

So, you are looking at such measures in the last lecture and then we had obtained certain nice properties of probability measures. And in particular, we looked at certain explicit ways of constructing probability measures like using convex combinations of known measures, known probability measures.

Now, with those examples at hand, this is the right time to look at the properties of measures or the set functions. Some of the results that we are going to discuss will be specifically for probability measures, but otherwise this will be for general measures. Let us start and go through the slides.

(Refer Slide Time: 1:37)



So, as before we are starting off with a measure space $(\Omega, \mathcal{F}, \mu)$. So, μ is a measure on the measurable space $(\Omega, \mathcal{F})$. And we are going to concentrate on algebraic properties of the measures μ. So, what do we mean by algebraic properties, that means, we are going to look at certain kinds of inequalities, which involve addition, subtractions and maybe multiplications, and so on. So, let us look at these inequalities.

(Refer Slide Time: 2:11)

algebraic properties of measures μ.

Note (11): From our discussion in the previous lectures, we have the following facts about measures.

(i) $\mu(\phi) = 0$.

(ii) μ is finitely additive and Countably additive.

Countably additive.

(iii) $0 \leq \mu(A) \leq \mu(\Omega), \forall A \in \mathcal{F}$.

(iv) $\mu(A) + \mu(A^c) = \mu(\Omega) \;\forall A \in \mathcal{F}$.

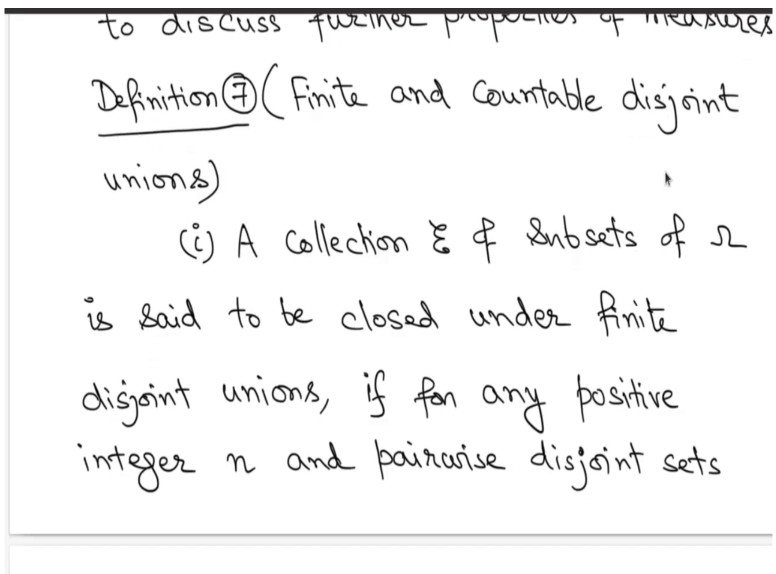
(v) measures μ with $\mu(\Omega) = 1$ are said to be probability measures.

We need the following terminology to discuss further properties of measures

So, just to recall, we have, we already have certain nice properties about measures. So, what are these? So, let us just quickly go about these facts. Let us first quickly recall these facts about measures. The measure of the empty set is 0. A measure is a non-negative set function and it is finitely additive and countably additive. The measures of each arbitrary subset in your σ-field is dominated from above by the measure of the whole set. So, the measure of the whole set is the maximum value achieved by the measure μ.

Then, if you add up the measures of A and A^c, you get back the measure of the whole set. So, in fact, this was proved earlier and by means of this we prove that measures of arbitrary subsets is dominated from above by measures of the whole set. And finally, we defined probability measures as measures with the property that $\mu(\Omega) = 1$. So, with these basic facts about measures at hand, we are now going towards looking at certain algebraic properties of measures. But before that, before we start the discussion, we introduced this terminology.

(Refer Slide Time: 3:30)



$A_1, A_2, \ldots, A_n \in \mathcal{E}$, we have $\bigcup_{i=1}^{n} A_i \in \mathcal{E}$.

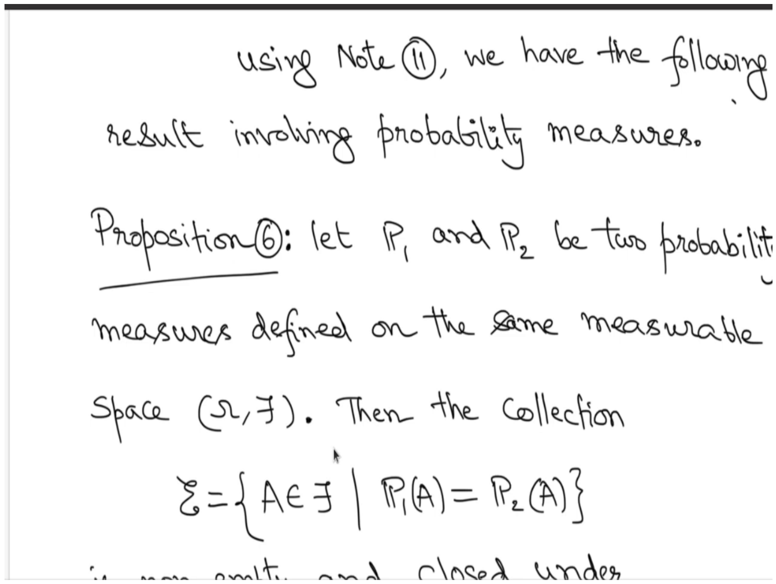
(ii) A collection $\mathcal{E}$ of subsets of Ω is said to be closed under countable disjoint unions, if for any sequence $\{A_n\}_n$ of pairwise disjoint sets in $\mathcal{E}$, we have $\bigcup_{n=1}^{\infty} A_n \in \mathcal{E}$.

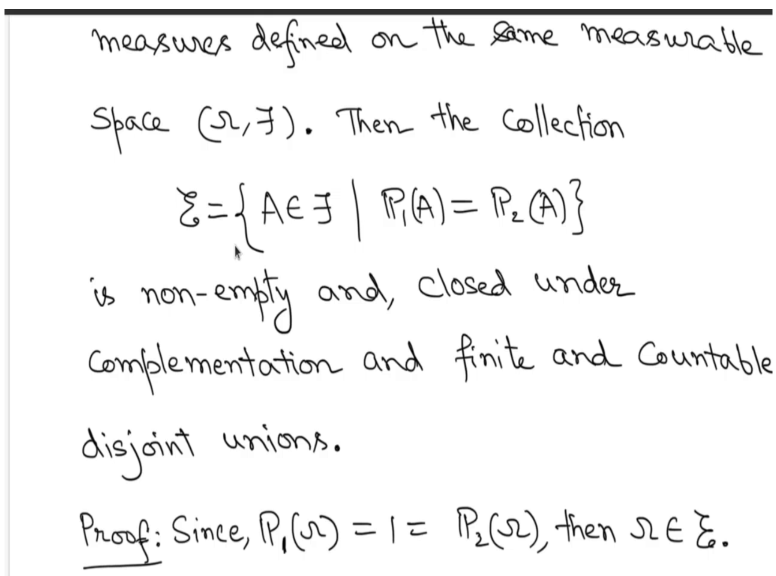
First, we are going to talk about something called finite disjoint union subsets. So, suppose you will look at certain collection of subsets, which we call a $\mathcal{E}$. So, this is just a collection, arbitrary collection. We say this collection $\mathcal{E}$ is closed under finite disjoint unions if for arbitrary positive integers n, you look at pairwise disjoint sets $A_1, A_2, \ldots, A_n$ coming from the collection $\mathcal{E}$, then their union is again in $\mathcal{E}$.

So, again we are not allowing arbitrary subsets $A_1, A_2, \ldots, A_n$, we want pairwise disjoint subsets for any arbitrary number n, but then I want their union, this is a finite union, I want their union to be in epsilon. So, if this happens for all arbitrary such n and $A_1, A_2, \ldots, A_n$, then I will call $\mathcal{E}$ that collection is closed under finite disjoint unions. Similarly, we now define a corresponding concept for countable disjoint unions.

So, again we look at a collection $\mathcal{E}$ and then we say it is closed under countable disjoint unions, if for a sequence of sets now, the set should be first of all pairwise disjoint, then for those kinds of sequences I want their countable union to be back in the collection $\mathcal{E}$. If it so happens for arbitrary such sequences with pairwise disjoint sets, then I will call that $\mathcal{E}$ is closed under countable disjoint unions. So, with these two operations, we are now ready to discuss certain interesting properties of measures.

(Refer Slide Time: 5:22)





So, earlier we have already recalled some of the basic facts about measures that we have already seen. And now, using those properties, we now derive certain more interesting properties. So, first properties involving probe two probability measures, later on we will see some more properties involving general measures. So, let us start with $\mathbb{P}_1$ and $\mathbb{P}_2$ be two probability measures on some appropriate measurable space $(\Omega, \mathcal{F})$. So, they are defined on the same measurable space.

Now, we look at this collection. So, this is defined in a very special form. We want those subsets in your σ-field such that the probability of the sets under $\mathbb{P}_1$ and $\mathbb{P}_2$ match. So, if the probability of the set A has the same probability under these two different way of measuring $\mathbb{P}_1$ and $\mathbb{P}_2$, then I will put it in the collection $\mathcal{A}$. So, that is the collection. Now, what is the claim of the proposition? So, we claim that this is a first of all a non-empty collection, and then it is also closed under complementation and it is also closed under finite disjoint unions and countable disjoint unions.

So, this is a very special collection of subsets of the original set Ω and we are choosing these subsets from the σ-field. And we are saying that this collection is non empty and it is closed under these important operations, complementation, finite disjoint unions and countable addition to unions. So, how do you show this? Let us prove them on by one.

First observe that the measure of the whole set is 1, because we are looking at property measures. So, therefore, Ω which is, of course, content in your σ-field must be such a set that we have described. So, therefore, Ω is in your $\mathcal{E}$. So, therefore, this is non-empty. So, let us try to see why this is closed under complementation.

(Refer Slide Time: 7:25)

Proof: Since, $\mathbb{P}_1(\Omega) = 1 = \mathbb{P}_2(\Omega)$, then $\Omega \in \mathcal{E}$.

Thus $\mathcal{E}$ is non-empty.

If $A \in \mathcal{E}$, then $\mathbb{P}_1(A) = \mathbb{P}_2(A)$. In this case,

$$\mathbb{P}_1(A^c) = \mathbb{P}_1(\Omega) - \mathbb{P}_1(A) = \mathbb{P}_2(\Omega) - \mathbb{P}_2(A) = \mathbb{P}_2(A^c).$$

Hence $A^c \in \mathcal{E}$ and $\mathcal{E}$ is closed under complementation.

So, you look at arbitrary set here. I only want to show that its complement is also in the same collection. So, to check that, first observe that as per the definition the probability of the set A according to $\mathbb{P}_1$ and according to P2 must match, but then you will now try to compute the probability of the complement.

But according to the formula that we have probability of A^c according to $\mathbb{P}_1$ is nothing but $1 - \mathbb{P}_1(A)$. $\mathbb{P}_1(\Omega)$ is nothing but 1. That is the probability measure itself. But then $\mathbb{P}_1(A) = \mathbb{P}_2(A)$ i so just write that and immediately you get that this equality holds that A^c has the same mass whenever you are looking at under $\mathbb{P}_1$ and whenever you are looking at under P2.

So, no matter which probability measure you use $\mathbb{P}_1$ or $\mathbb{P}_2$ you get the same mass for A^c provided A already has that same property. And hence, A^c, which is in your σ-field, must be that in the special collections $\mathcal{E}$. And therefore, $\mathcal{E}$ is closed under complementation.

(Refer Slide Time: 8:37)

let $A_1, A_2, \ldots, A_n$ be pairwise disjoint sets in $\mathcal{E}$. Then $\mathbb{P}_1(A_i) = \mathbb{P}_2(A_i)$ for $i = 1, 2, \ldots, n$. By finite additivity of $\mathbb{P}_1$ and $\mathbb{P}_2$, we have,

$$\mathbb{P}_1\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}_1(A_i) = \sum_{i=1}^{n} \mathbb{P}_2(A_i) = \mathbb{P}_2\left(\bigcup_{i=1}^{n} A_i\right)$$

and hence $\bigcup_{i=1}^{n} A_i \in \mathcal{E}$, i.e. $\mathcal{E}$ is closed and hence $\bigcup_{i=1}^{n} A_i \in \mathcal{E}$, i.e. $\mathcal{E}$ is Closed under finite disjoint unions.

The proof involving Countable disjoint unions is similar and is left as an exercise.

We now discuss further properties of measures.

So, let us now try to show finite disjoint unions is also such a nice property for such a collection $\mathcal{E}$. So, choose some arbitrary number of sets $A_1, A_2, \ldots, A_n$, but finitely many. Suppose they are pairwise disjoint, I want to claim that their union is also in $\mathcal{E}$. So, how do you show this? First of all, since Ais are in your collection, then their probability according to $\mathbb{P}_1$ and according to $\mathbb{P}_2$ must match. So, for all $1 \leq i \leq n$, this is the equality that you already have.

But then, you use finite additivity of the individual probability measures $\mathbb{P}_1$ and $\mathbb{P}_2$. If you use that, then look at the probability under $\mathbb{P}_1$ of the union of the sets that can be split as the finite summation of the probabilities of Ai according to $\mathbb{P}_1$, but then $\mathbb{P}_1(A_i) = \mathbb{P}_2(A_i)$, write that and again use the finite additivity of $\mathbb{P}_2$ to get this last expression on your right hand side. And immediately it says that this union A_i is also in your collection $\mathcal{E}$.

Therefore, $\mathcal{E}$ is closed under finite disjoint unions and here we have used the finite additivity of the probability measures $\mathbb{P}_1$ and $\mathbb{P}_2$, the remaining property is about countable disjoint unions. But then, the proof goes again in this exactly same way as done for finite disjoint unions, you look at a sequence of sets which are pairwise disjoint, use the fact that $\mathbb{P}_1$ and $\mathbb{P}_2$ are probability measures and they are countably additive using that you can prove this fact. This part is left as an exercise for you. Please work this out.

So, therefore, just to consolidate, we have looked at these algebraic operations involving set operations and we are looking at these spatial collections for any two arbitrary property measures defined on the same measurable space and then this collection of these spatial subsets turns out to have this nice properties. We will see more properties of this collection later on.

(Refer Slide Time: 10:45)

We now discuss further properties of measures.

Proposition 7: For any $A, B \in \mathcal{F}$, we have

$$\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B).$$

Proof: Note that, A and B can be written as a pairwise disjoint union of sets in

Proof: Note that, A and B can be written as a pairwise disjoint union of sets in

the following form:

$$A = (A \cap B) \sqcup (A \cap B^c)$$

and $B = (A \cap B) \cup (A^c \cap B)$.

Then by finite additivity of μ,

$$\mu(A) + \mu(B) = \mu(A \cap B) + [\mu(A \cap B^c) + \mu(A \cap B) + \mu(A^c \cap B)].$$

Again, by finite additivity of μ,

$$\mu(A \cup B) = \mu(A \cap B^c) + \mu(A \cap B) + \mu(A^c \cap B).$$

Hence, the equality follows.

<u>Corollary to Proposition 7</u>

For any $A, B \in \mathcal{F}$ with $B \subseteq A$,

So, now we concentrate our attention to general measures. So, suppose you take any two arbitrary subsets in your σ-field and suppose μ is a measure as taken earlier, now I claim that this equality holds. What is this? This says a $\mu(A \cup B)$ and $\mu(A \cap B)$. If you add them up, this agrees with $\mu(A) + \mu(B)$. So, how do you show this? So, start with the right hand side. The trick here is to start with the right hand side.

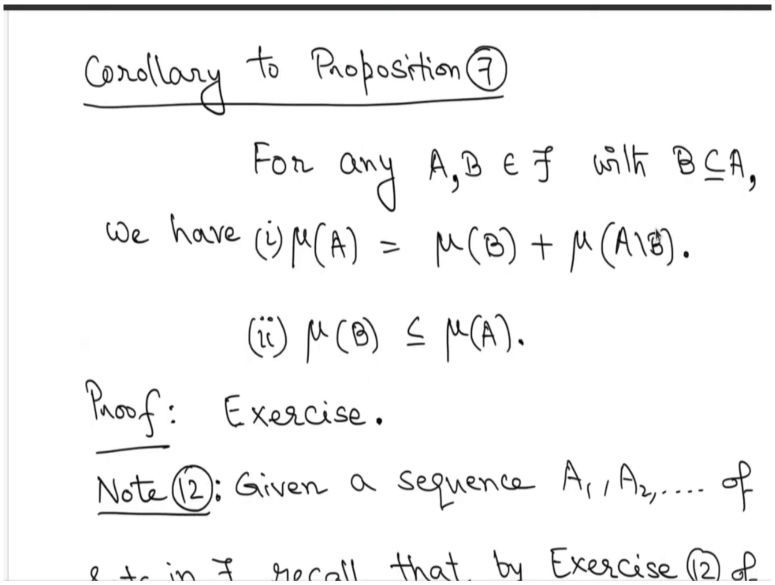
So, start with the set A and set B and observe that this can be written as pairwise disjoint union of certain other sets. What are these? So, I write A as $A \cap B$ and $A \cap B^c$. So, I write $A \cap B$ as one set and $A \cap B^c$ as another set. But then by the construction $A \cap B$ and $A \cap B^c$ are pairwise disjoint.

Similarly, I split the set B as this pairwise disjoint union of $A \cap B$ and $A^c \cap B$. Use the final relativity of mu, you immediately get this relation that $\mu(A) + \mu(B) = \mu(A \cap B)$. So, that was one of the terms that is here. But the other three terms I put them together under this bracket.

So, there is a reason for doing that. But you see that you have $\mu(A \cap B)$ appearing twice. So, we keep one outside and put the other one with the remaining terms within this bracket. So, now, you use finite additivity to claim that whatever term is within the bracket is exactly the $\mu(A \cup B)$. Why, because $A \cup B$ is the finite disjoint union of the sets $A \cap B^c$, $A \cap B$ and $A^c \cap B$.

So, this is a very standard set theoretic property that $A \cup B$ can be split as pairwise, as the disjoint union of these three sets. Therefore, you apply the finite additivity and combine the sizes of these three sets together to write a $\mu(A \cup B)$, and hence, the required equality follows. So, let us go back. So, for any arbitrary subsets A and B, you get that $\mu(A \cup B)$ added with $\mu(A \cap B)$, you get back $\mu(A) + \mu(B)$.

(Refer Slide Time: 13:22)



Now, as a consequence, you get this corollary that if you take two subsets now, A and B, but with this additional property that B is a further subset of A actually. Then what happens, then that earlier relation reduces to this relation. How, A you can write it as the pairwise disjoint union of B and $A \setminus B$. So, that is one way to go about this. Another way is to simply apply previous proposition and simplify it to this one. So, there are two ways of proving this equality.

So, one here, since A is a bigger set than B, you first write A as a disjoint union of B and $A \setminus B$. So, write using finite additivity. Another option is to use the previous proposition. But then from this equality, there is a very interesting relation that follows. Since $\mu(A \setminus B)$ that set is non-negative, since μ associates non-negative values, $\mu(A \setminus B)$ that is non-negative, you get that $\mu(B) \leq \mu(A)$.

So, this is a non-negative quantity and hence you get this inequality. So, here as long as $B \subseteq A$, $\mu(B)$ cannot be more than $\mu(A)$. So, that is an interesting property, some kind of a modern city relation, algebraic relation for the measurement. So, try to work them out, write them out, but I have already discussed the main ideas for the proofs.

(Refer Slide Time: 14:53)

Note (12): Given a sequence $A_1, A_2, \ldots$ of sets in $\mathcal{F}$, recall that by Exercise (12) of Week (1), we can construct sets $B_1, B_2, \ldots$ also in $\mathcal{F}$ such that

(i) $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n$

(ii) $B_n \subseteq A_n \; \forall n$

(i) $\bigcup_{n=1} A_n = \bigcup_{n=1} B_n$

(ii) $B_n \subseteq A_n \; \forall n$

and (iii) B_n's are pairwise disjoint

In fact, Exercise (12) was for any finit number of sets $A_1, A_2, \ldots, A_n$. Here, we are repeating the construction of B_n's for each n and thereby we obtain number of sets $A_1, A_2, \ldots, A_n$. Here, we are repeating the construction of B_n's for each n and thereby we obtain

a sequence $\{B_n\}_n$ as above, for any given sequence $\{A_n\}_n$.

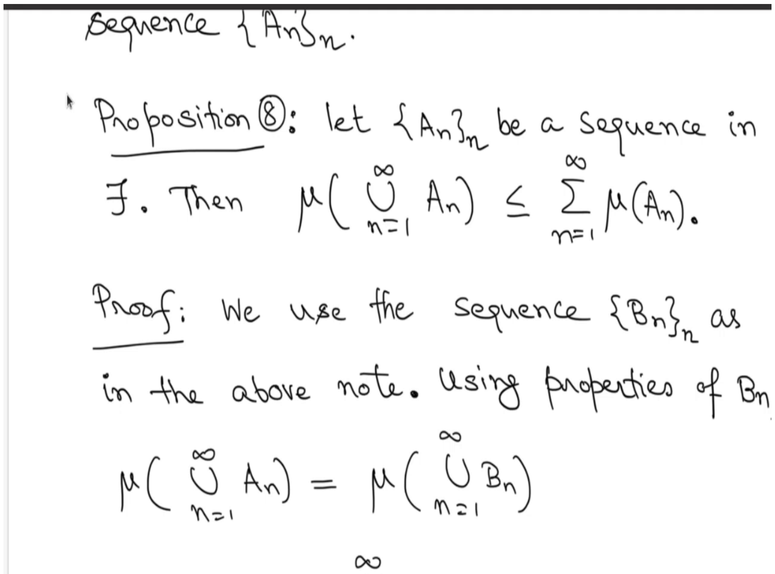
Proposition ⑧: let $\{A_n\}_n$ be a sequence in

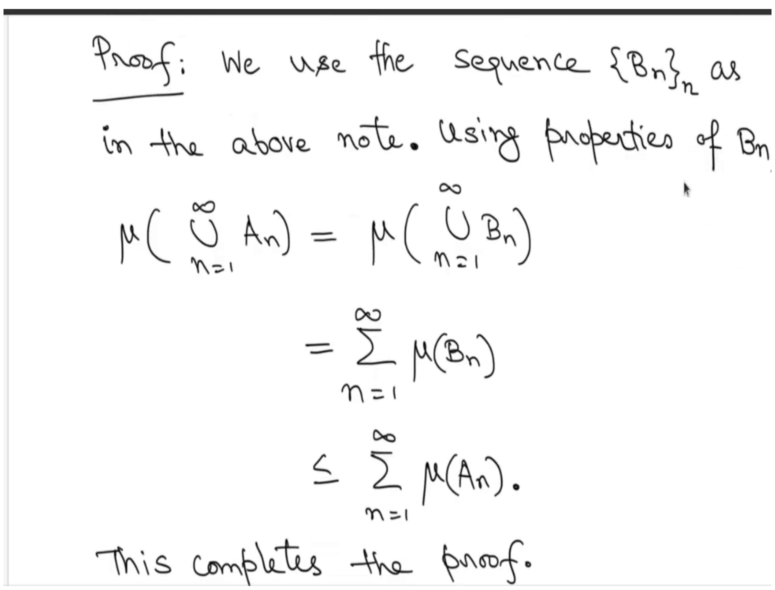
But given a sequence $\{A_1, A_2, \ldots, A_n\}$, recall that earlier we had discussed the interesting construction in exercise 12 this was in week one. So, let us recall what we had done. So, given the sequence of sets or maybe a finite number of sets $A_1, A_2, \ldots, A_n$, you can construct certain sets $B_1, B_2, \ldots, B_n$ such that these properties holds that the $\cup A_n$ is matching with $\cup B_n$. B_n's are subsets of A_n, but further the B_n should be pairwise disjoint.

So, you can construct sets $B_1, B_2, \ldots, B_n$, which are pairwise disjoint which are subsets of original sets A_n, but their union matches with the unions of A_n. And notice that the union of B_ns is nothing but a countable disjoint union now, because Bns are pairwise disjoint. So, this is something we are going to use from now on.

Just to recall about this exercise 12 that we are talking about, this was actually for any finite number of sets $A_1, A_2, \ldots, A_n$ but if you fix the first n elements, first n sets $A_1, A_2, \ldots, A_n$ and construct these sets $B_1, B_2, \ldots, B_n$, then you exactly get this property. So, here you vary for any n, then you can construct the corresponding Bn by the same construction method. So, that is the idea that it was earlier explained for finitely number of sets $A_1, A_2, \ldots, A_n$ but then you can repeat for any n and obtain the B_n and prove these three properties. So, this is easy to observe, but then we are going to use these specific type of $B_n's$ in our later results.

(Refer Slide Time: 16:35)





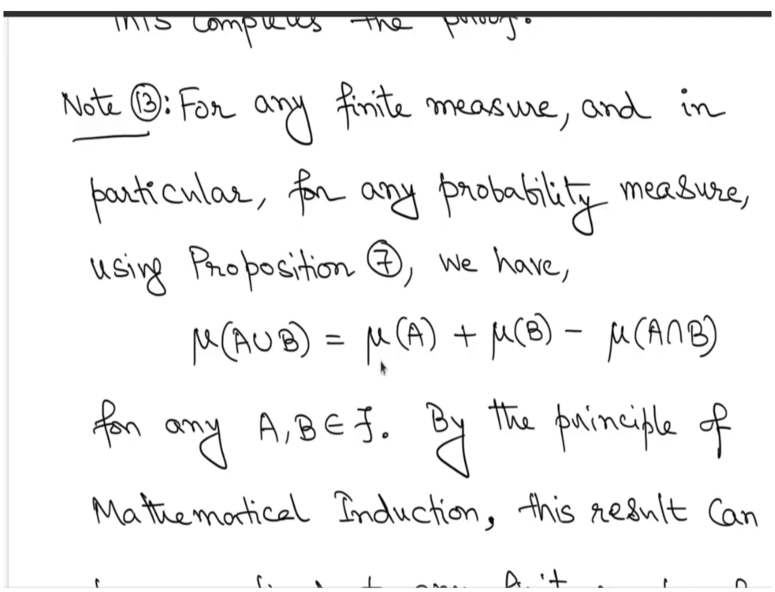
So, this is immediately going to be used in proposition 8. So, remember, for sequence of sets $\{A_n\}$, which was pairwise disjoint, we have countable additivity of the measure μ, which said that the measure of the countable union is obtained as the sum of the individual sizes. So, that was what we called as countable additivity of a set function of a measure μ.

But then, what we actually have is that for arbitrary sequences, where there may be intersection between the sets, you actually can claim an inequality. You can no more claim equality, but you can claim an inequality. What does it say? It says that the measure of the, this countable union is at most the series value which is obtained by adding up the individual sizes of the sets.

To prove this we use the sequence Bn as constructed above. What do we do? So, take those Bns. Remember, B_n's are pairwise disjoint, $\cup B_n$ is same as $\cup A_n$ and $B_n \subset A_n$. So, we use all these properties. First, observe that this countable union of A_ns is nothing but countable union of B_n, so that you write first. Since B_n's are pairwise disjoint, you apply the countable additivity of the set function μ, of the measure μ and you get this summation that you have to add up the individual sizes and you get this series. So, that is the countable additivity of the measure mu.

But then $B_n \subset A_n$, so therefore $\mu(B_n)$ cannot be more than $\mu(B_n)$ and hence you get this inequality this upper bound that is given by adding up the individual sizes of the sets A_n and therefore, you get this inequality that was stated in the proposition. So, therefore, when you are working with pairwise disjoint sequence, you get an equality. But otherwise, for general sequences, you can at most claim an inequality and this is the upper bound.

(Refer Slide Time: 18:39)



particular, for any probability measure, using Proposition 7, we have,

$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$$

for any $A, B \in \mathcal{F}$. By the principle of Mathematical Induction, this result can be generalized to any finite number of sets. This is stated in the next result.

Now, this is an interesting comment. So, this is about the discussions in proposition 7. Remember, we proved that for any measure, $\mu(A \cup B)$ together with $\mu(A \cap B)$ if you add them up, that will be exactly $\mu(A) + \mu(B)$. But then you would like to transfer the $\mu(A \cap B)$ to the other side. And thereby, you would like to write $\mu(A \cup B)$ in terms of these three things, $\mu(A) + \mu(B) - \mu(A \cap B)$.

But you have to be careful here, if it so happens that you are dealing with an infinite measure, then $\mu(A \cap B)$ may be infinity. If you are dealing with an infinite measure, then $\mu(A \cap B)$ may be ∞. In that case, you have to be careful when you are subtracting. So, remember, infinity minus infinity is not defined. So, if you want to subtract this infinite term $\mu(A \cap B)$, if it is infinite, then you are not allowed to do so. You cannot define $\infty - \infty$. So, therefore, you cannot cancel off the $\mu(A \cap B)$ that was earlier on the left hand side.

So, you cannot subtract it and put it on the right hand side. So, you can do it for finite measures or if you are given the information that $\mu(A \cap B)$ is finite, only then you are allowed to write this formula that $\mu(A \cup B)$ is $\mu(A) + \mu(B) - \mu(A \cap B)$. But this is true for any finite measure, so you do not have to worry about that.

And in particular, this is of course, true for probability measures. And that gives you back your familiar formula that $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. So, that is your familiar formula that is now generalized to arbitrary finite measures for example. Now, by the principle of mathematical induction, you can now extend it to n many sets. So, you can extend it to finitely many sets and you get the result, which is called the inclusion exclusion formula.

(Refer Slide Time: 20:55)

Proposition 9: (Inclusion- Exclusion formula)

let μ be a finite measure on $(\Omega, \mathcal{F})$

and let $A_1, A_2, \ldots, A_n \in \mathcal{F}$. Then,

$$\mu\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu(A_i) - \sum \mu(A_i \cap A_j)$$

$$\mu\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu(A_i)$$

$$- \sum_{\substack{i,j=1 \\ i<j}}^{n} \mu(A_i \cap A_j)$$

$$+ \cdots$$

$$+ (-1)^{n-1} \mu(A_1 \cap A_2 \cap \cdots \cap A_n).$$

Proof: Exercise.

So, here you are considering n many sets now. So, here I am taking finite measure once more, just for the sake of well defineness. We do not want to subtract any term which is infinite. So, just for simplicity, let us work with finite measures. Then we are saying that if you choose n many sets, $A_1, A_2, \ldots, A_n$, arbitrary n many sets, then the measure of their union, this is a finite union, can be written in terms of an algebraic expression like this. So, it is an addition and subtraction of certain terms. What you observe is this.

So, first, you add up the size of the individual sets according to the measure μ. But look at pairwise intersections and subtract out those sizes. Here, you should choose i and j, so that they do not get repeated. So, that is why I have put $i < j$. So, as i and j varies from 1 to n, you get all possible pairs A_i and A_j. So, you look at all possible such intersections, pairwise intersections, and look at the size of that you subtract them out.

Then next step, you add the term, which is intersections of three terms like this, three sets like this. So, you take A_i, A_j, A_k, for example, with i, j, k distinct i, j run, i, j, k runs from 1 to n. But here you add them. So, that is the sign gets flipped. So, start with plus sign, then for pairwise intersections, you add minus sign and so on. And at the end, you will end up with this terms, which is the intersection of all the sets and look at the size of that. You associate $(-1)^{n-1}$. So, that is the general formula.

So, here, what you are doing, you are first taking the sets A, then you are subtracting out the common parts, then adding up the triple intersections and so on. This is again a generalization of the results that you know for probability measures. And this is left as exercises for you. This can be proved in many different ways. One of the ways is just by extending it by mathematical induction, by the principle of mathematical induction, using the case for two sets.

From two sets, you can go to n sets by principle of mathematical induction. So, that is one way of doing it. Other way of doing it is by counting the appearance of the, appearance of an arbitrary element in these sets. So, basically, what you are doing is that if you go via a Venn diagram of these sets $A_1, A_2, \ldots, A_n$, you had just looking at how many times you are counting these sets. If you have counted the sets A_i's once at a time, then you have already counted the intersections twice. So, that is why you are subtracting out the pairwise intersections.

But if you subtract out the pairwise intersections, what you have already removed are the triple intersections. So, that is what you have to add. So, that is why this formula comes with plus and minus signs in the alternative sense. So, alternatively, you get plus or minus signs and you get this expression which is known as the inclusion exclusion principle. So, try to prove this, that again, the hint is you can use the principle of mathematical induction and apply and prove this. Use the principle of mathematical induction and prove this. So, this is again left as an exercise.

So, we have discussed some algebraic properties of measures and we are going to see certain nice properties involving continuity in the next lecture. So, we will stop here.