

[SQUEAKING]

[RUSTLING]

[CLICKING]

CASEY
RODRIGUEZ:

All right, so let's continue with our discussion of measurable functions. So last time we introduced the notion of measurable functions. So if I have a measurable set, E from-- or f from E to the extended real numbers is measurable, or is a measurable function, if for all α in \mathbb{R} , the inverse image of the half infinite open interval α to infinity is measurable.

And then we proved that if we have a function which is measurable, then not only does is the pre-image of these intervals measurable, but the pre-image of any Borel set, any member of the Borel sigma algebra, which includes open sets, close sets, and so on, the pre-image of those sets is also measurable.

And we proved that being measurable is closed undertaking sups, imps, lim sups, lim imps, limits in particular, and changing the function a little bit on-- or changing a function on the set of measure 0 also preserves measurability. And also being measurable is closed under the algebraic operations of taking linear combinations and products.

Now, this definition and properties we've worked out was for extended real-valued functions. Quite often we'll be dealing with-- or in general, functions that take values in the complex numbers. So let me define what it means for a complex value function to be measurable. It's not too crazy. Let E be a subset of \mathbb{R} be measurable.

We say a function f from E now to the complex numbers is measurable if the two functions given by the real part of f , which now goes from E to \mathbb{R} and the imaginary part of f , which now is a function from E to \mathbb{R} , are measurable.

So function into the complex numbers you can always write as f equals the real part of f plus the imaginary part of f times i . So we just-- to say that f is measurable, just require that it's real and imaginary parts are measurable, OK.

And then you can verify the following simple theorem. Maybe I'll put it on assignment or not, haven't decided yet-- or just parts of it. If f, g were measurable-- and I didn't say so, but if we're talking about measurable functions, the domain E , the domain is always assumed to be a measurable subset of \mathbb{R} .

So if these are measurable functions and α is a complex number, then the function's α times f , f plus g , f times g -- and then we can do a few other things to complex numbers that change it. The complex conjugate of f and modulus of f are measurable functions.

And then we have the following theorem. If f_n from E to \mathbb{C} is measurable for all n . And these functions converge pointwise to a function f , then the limiting function is measurable, OK. And this follows, again-- so both of these theorems follow immediately from what we know about measurable extended real-value functions.

So, for example, this one here follows from the fact that $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ if and only if $\lim_{n \rightarrow \infty} \operatorname{Re}(f_n(x)) = \operatorname{Re}(f(x))$ and $\lim_{n \rightarrow \infty} \operatorname{Im}(f_n(x)) = \operatorname{Im}(f(x))$.

And if we're assuming f_n is measurable for all n , then the real part and imaginary parts of f_n are measurable for all n . And therefore there are pointwise limits, which is the real part of f and imaginary part of f are measurable by the theorem we proved about extended real-value measurable functions.

And so we conclude that the real part of f and the imaginary part are measurable, and therefore f is measurable. So you don't have to work very hard. You can just use of what we proved from the previous lecture about extended real-valued functions which are measurable.

All right, so now as far as measurable functions go, we've shown that if I have a continuous function, then that is - so this is from the previous time, from previous lecture, continuous functions are measurable. And if I have a measurable subset of E , then the indicator function of that set is measurable.

And we know that linear combinations of measurable functions are measurable. So if I take linear combinations of indicator functions with, say, complex coefficients, then that will remain measurable as well. And those functions are the simplest type in that they only take finitely many values.

They're so simple that we glorify them by giving them that name. And we'll show that every measurable function is, in a sense, approximately a simple function. So we have the following definition, if E is measurable, a measurable function ϕ from E to C is simple, or we call it a simple function.

So we'll say ϕ is a simple function if the range of ϕ is equal to finitely many values. So a measurable function is a simple function if its range is finite. So let me make a general remark about simple functions.

And also when I write that ϕ is equal to a_1 up to a_n , this is a set. And when I write a_1, a_2, \dots, a_n , I am implicitly writing here. I'm not saying the set-- I'm saying each of these elements are different from each other. So a simple function that just takes the value 1, I would not write its range as $1, 1, 1, 1$. So although that's a kind of simple and silly remark to make, I just want to make it now.

So if ϕ is a simple function, then we can write it in a canonical way with-- so if ϕ is a simple function with ϕ of E equals a_1 up to a_n , then for all i , the set A_i , which is equal to the inverse image of-- so here ϕ is going from E to C .

The inverse image of the single set, singleton, which is a closed-- well, this is going to be-- so the inverse image of this guy here is measurable because it's equal to the intersection of two measurable sets.

It's equal to the intersection of when the real part of ϕ equals the real part of a_i intersect the set of all x 's where the imaginary part of ϕ equals the imaginary part of a_i . So that's why the set is measurable, is measurable.

And we have a few properties of this guy. If I take two different elements in the range-- again, this is why I made that comment that when I write the range this way, I'm listing the distinct elements of the range-- or the image of E -- For all i equals not equal to j , these two sets are disjoint.

And if I take the union of i equals 1 to n of the A_i 's, this is equal to the total set E because this is just the inverse image of this set, which is just E , since E of E equals that, OK. And finally for all x in E , I can write ϕ of x as sum from i equals 1 to n of $a_i \chi_{A_i}$ of x .

So for a simple function, I can write it in a canonical way where it's just a linear combination of indicator functions where the sets those indicator functions are non-zero on or disjoint from each other. And their union gives me E , the domain, OK.

So these three are the simple but important properties of how to represent a simple function. So it's not difficult to verify, again, just from the definition of a simple function, that scalar multiples, linear combinations, and products of simple functions are, again, simple functions.

So I've said that these functions are so simple that that's what we call them, and that they are somehow universal that, in a sense, any measurable function is almost a simple function. So in what sense do I mean that?

So that's the content of the following theorem, which is that if f from E to \mathbb{R} and let's first-- we're going to do this for extended real-valued functions. And then the proof will essentially carry over to the complex value.

So I'm just going to do it for the extended real-value non-negative functions. And, again, I'll indicate what the difference is once we go to complex valued measurable functions. So if E is from \mathbb{R} to \mathbb{R} so now this is an extended real value, but it's non-negative.

So if this is a measurable function, then there exists a sequence of simple functions ϕ_n such that three things hold. We can do this-- let's write down-- let's go over here. three things hold. For all x in E -- or let me hold off on stating that.

f dominates these simple functions. And these simple functions are pointwise increasing. For all x in E , 0 is less than or equal to ϕ_0 of x is less than or equal to ϕ_1 of x is less than or equal to ϕ_2 of x and so on. And they all sit below f of x .

For all x in E , these ϕ_n 's are converging to f of x . So A and B says that there exists a sequence, a measure of simple functions that increase to f of x . The last part is that if this function, if this measurable function is bounded, or wherever it's bounded, this convergence is not just pointwise but uniform. For all B bigger than or equal to 0 , the sequence ϕ_n converges to f uniformly on the set where f is bounded.

So, again, the take home is that for every non-negative extended real-value measurable function, we can find a sequence of simple functions that well approximates f . And in what sense does this will approximate f , they increase to f .

And if f is bounded, then that convergence is, in fact, uniform. Or more precisely, wherever f is bounded, convergence is uniform. So this is in what sense every measurable function is almost a simple function. So let me-- let's get started with the proof.

And first I'm going to draw a few pictures so that you can get the idea of how we're going to do this, or how we're going to construct this sequence of simple functions, and then turn that into math which might be a little bit jarring if I went that route first and then drew pictures.

And this picture I'm going to draw kind of looks like what I was talking about where we split the range up rather than the domain when we were motivating why we would even introduce the concept of a measurable function. So let's say we have our function f .

Now how I'm going to build these phi's are that-- what I'm going to do is-- so here I'm going to draw phi 0 now. Where's the yellow chalk. What I do is this will indicate the power of 2, how high up I'm going, and also the resolution, how much of my dividing how high up I'm going into smaller parts.

So phi 0 you should think that my height will be-- well, my height is going to be 1. And so what I do is, how I define this simple function phi 0 is I look at where f is above the final value 1. And my simple function will be that final value 1 there.

And move this over a little bit. And that just leaves where f is less than 1. And there I set the value of my simple function to be 0, and 0 simply because that's the lower bound, or that's the smallest value of this interval 0, 1.

And that's how I define phi 0. It goes up to height 1. And I only split each part into 1. So there's only one part here. Now maybe some of that didn't make any sense to you. That's OK. We're going to phi 1 and then I'm going to stop there because this is going to grow exponentially, which means I'll probably draw an exponentially worse picture each time.

So now we're going to draw phi 2. And what I do is-- there's one-- or phi 1, sorry. One, again, should indicate the power of 2 that I'm both going up and resolving the axis. So now I go up to-- let's make this a-- so, again, this will not be quite the scale, but hopefully that's OK. So let's make this a little closer to scale.

OK, so one, now I have two parts, 2 to the 1. So this is parameterized how-- tallest I'm going and cutting up the range of my function f. And then I'm going to now resolve each part in $1/2$, 2 to the minus 1.

So now I take-- this is now $3/2$, and this is now $1/2$, 2 to the minus 1. So why am I writing to 2 to the minus 1? Again because this one here corresponds to-- and that one here corresponds to-- I'm cutting the whole increments into halves.

If I go on to phi 2, I'm going to be going up to 2 to the 2, which is now 4; and from 4 to 3; 3 to 2; 2 to 1; I'm going to be cutting those into fourths. And, again, what I do is I look at-- I cut-- and now I look at where the function is in these widths. So if it's above my highest bound, then my simple function will be 2 on that set of x's where it's at its highest, or where f goes past the largest number which I'm resolving the axis in.

And then now here, for example-- so you should try and draw your own picture on not just go off mine because, again, mine is looking kind of rough already. But now I look at the function-- the set of X's where the function is between 2 and $3/2$.

So that's going to be this piece and this piece. And then I set my simple function to be equal to the value, the lower bound on this interval that I've cut the range into. And then I do that from $3/2$ to 1. And, for example, that's happening here on this x.

And I set it equal to-- I set my simple function equal to, again, the lower bound on this interval that I've cut the range up into, which is 1, and so, and so on. And now the function between 1 and $1/2$, so this is kind of only the last piece. Everywhere else is filled in. And there's no f in between here, so I don't assign a value there.

All right, now that was me talking my way through how this function how the sequence of functions look. Now I'll just write down how these functions are defined. So for n equals 0, 1, 2, and so on for k between 2 to the 2^n minus 1, I define sets E_{kn} .

This is equal to the set of all X 's in E such that f of x is between k times 2 to the minus n is less than or equal to k plus 1 2 to the minus n , which is just me explicitly writing out that this is the inverse image of k times 2 to the minus n k plus 1 2 to the minus n closed.

And this is an interval. And since f is measurable, assumed to be measurable, this is a measurable-- the inverse image of that interval is a measurable set, OK.

And then I define f_n to be the inverse image of when f exceeds my top value of how I'm cutting up the range, 2 to the n , again, which is measurable. And, finally, I will take my simple function ϕ_n to be sum from k equals 0 to 2 to the $2n$ minus 1 , k times 2 to the minus n , which is-- again, this is the lower-- this would correspond to the lower part of my E_{k_n} that I'm looking at.

So for this example, if this E_{k_n} is the two $3/2$, and $3/2$ is the lower part times $\chi_{E_{k_n}}$ plus 2 to the n times where f exceeds 2 to the n . So I encourage you to maybe write out ϕ_1 , what that actually is.

I mean, there's-- in fact here, I'll do that. I drew the picture that goes over here. Let me write out what ϕ_1 actually looks like. ϕ_1 is equal to 0 times the indicator function where f is between 0 and $1/2$ plus a $1/2$ times the indicator function where f is between $1/2$ and 1 plus $3/2$ times the indicator function of when f is between-- no, that's not $3/2$. That should be $1/2$.

No, no, no, what am I doing, 1 -- times indicator function where f is between 1 and $3/2$ plus $3/2$ times $\chi_{f \text{ inverse of where } f \text{ is between now } 3/2 \text{ and } 2}$. So 2 is how high up I break up the range. And then plus this last part, this f_n part, which is 2 times $\chi_{\text{where } f \text{ is bigger than } 2 \text{ to the } n}$, So bigger than 2 .

So all of these sets are disjoint. The E_{k_n} 's and f_n 's for k different from k prime and fixed n are disjoint. And they are disjoint from this set. So this is a simple function, the finitely many values it takes is 2 to the n along with k times 2 to the minus n . Or at least the finitely many values it can take is a subset of that. So this is a simple function for each n .

And by design it's always sitting below f . So let me in fact bring it up here. Well, let me make the statement and then I'll say. So by definition for all x in E , E_n is non-negative and always sits below f of x .

So how do we see that? You can see it from the general formula, but I'll just indicate why just for safety. One, let's say-- so x has to be in one of these sets. Let's say it's here where f is between $1/2$ and 1 . Then ϕ_1 of x is equal to $1/2$ which is less than f of x because f takes on the value between $1/2$ and 1 .

In fact, here, I'll give a the brief argument here. If x is in E_{k_n} then by definition this means that k times 2 to the minus n is less than f of x is less than or equal to k plus 1 2 to the minus n , which implies that ϕ_n of x , which is by definition k times 2 to the minus n -- just, again, by how we've defined the ϕ_n 's-- where did we define the ϕ_n 's-- there is less than f of x .

So that's for x and E_{k_n} . And if x is in $f \text{ sub } n$, then that means f of x is greater than 2 to the n , which is always bigger than or equal to the ϕ_n of x -- well, I mean this is actually equal to ϕ_n of x . So we always have-- the ϕ_n 's are non-negative. And they always sit below f of x . So now let's prove that they are, in fact, increasing, so part A.

So now we're proving part A, the ϕ_n 's increase to f . And, again, for fixed n , E_k^n and F_n , the E_k^n 's, and F_n form a disjoint union of E . I just need to check that what I want holds on each of the E_k^n 's and F_n 's.

So suppose x is in E_k^n , then f of x is less than or equal to $k + 1 \cdot 2^{-n}$. n is bigger than $k + 2^{-n}$, and which by a silly trick of just multiplying and dividing by 2 tells me that f is between $2^k \cdot 2^{-n} - 1$ is less than f of x is less than or equal to $2^k + 2 \cdot 2^{-n} - 1$, which implies that x is in the union of $E_{2^k}^{n+1}$, $2^k + 1 \cup 2^k + 2^{-n+1}$.

If x is in $E_{2^k}^{n+1}$, then I get that ϕ_n of x is equal to, by definition, $k \cdot 2^{-n}$, which equals $2^k \cdot 2^{-n} - 1$, which because x is in $E_{2^k}^{n+1}$, this is equal to ϕ_{n+1} of x .

And if x is in $E_{2^k+2^{-n+1}}$ -- no, I shouldn't-- this shouldn't be a 2. This should be a 1, I'm sorry-- because this goes from 2^k up to $2^k + 1$, and then from $2^k + 1$ up to $2^k + 2^{-n+1}$. So it's either in E_{2^k} or $2^k + 1$.

And if x is in $2^k + 1$, then the ϕ_n of x is still-- I mean, x is still in E_k^n , so the ϕ_n of x is equal to still $k \cdot 2^{-n}$, which is equal to $2^k \cdot 2^{-n} - 1$, which is less than $2^k + 1 \cdot 2^{-n} - 1$, which is by definition, since x is in $E_{2^k}^{n+1}$, ϕ_{n+1} of x .

And, similarly, if x is in F_n , then the ϕ_n of x is less than or equal to ϕ_{n+1} of x . So we verified for all x since E is equal to this union over k equals 0 to $2^n - 1$ $E_k^n \cup F_n$. This implies that for all x in E , ϕ_n of x is less than or equal to ϕ_{n+1} of x . And this proves A.

Now, how to prove B and C-- these things will follow from part A in a simple estimate that we're going to prove. So B and C will follow immediately from the following claim in part a, which the claim is that for all x in set Y in E that f of y is less than or equal to 2^{-n} .

This part we already know that f of x minus ϕ_n of x is non-negative. But, in fact, this is bounded by 2^{-n} . So then B and C from A and this claim. Why does B follow from this claim?

Well, wherever x is-- I don't want to have to also explain what happens if f is equal to infinity. That also follows basically from the definitions, not necessarily from this estimate. But the more important part is when f is finite. So let's assume f is just finite for every x .

So then every x in E is eventually in one of these sets. So there exist let x be fixed. Then for n sufficiently large, x is in one of these sets since f of x is finite. And therefore for all n -- for all capital N sufficiently large, this minus ϕ_n of x is less than or equal to 2^{-n} .

But then f of x minus ϕ_m is also less than or equal to 2^{-n} for every m bigger than or equal to n because the ϕ_n 's are increasing. If ϕ_n is this close to f , then ϕ_m is also that close to f if m is bigger than or equal to n , again, because they're increasing.

So that proves-- that's why pointwise convergence B follows from this estimate. As far as part C, that also follows from this estimate since if I have a fixed B , then I can choose a natural number just depending on what B is so that that's set depending on B and the original statement I'm pointing at it, but I don't think you can see it from the camera.

That set which depends on B is contained in one of these sets, and therefore for all x in the set where f is bounded by B , this holds uniformly in x . And that's where the uniform convergence comes from. But the whole point is that this is the estimate that gives us B and C once we've proved A .

So let's prove this claim. It's not hard. It's basically because we are cutting up the range not only to height 2^{-n} , but with resolution 2^{-n} at each stage, at each n . So to prove the claim we have that the set of y 's in E such that $f(x)$ is less than or equal to 2^{-n} , this is equal to the union $\bigcup_{k=0}^{2^n-1} E_{kn}$ of the E_{kn} 's.

So if I want to check that bound, I just have to check it for each, if x is in one of these. So suppose x is in E to the kn , then-- I mean, it really is just following from the fact that we're cutting up the range with resolution 2^{-n} . So I'm just going to draw a small little picture here.

We have the axis here. And here's $k + 1$ times 2^{-n} . Here's k times 2^{-n} . If x is in there, that means there's a-- we're looking at the portion of f that sits between $k 2^{-n}$ and $(k + 1) 2^{-n}$, then $k 2^{-n} < f(x) \leq (k + 1) 2^{-n}$.

And remember the simple function on this piece evaluated in here-- so here's the x -- gives the value at the lower bound. And therefore we get that $f_n(x) - \phi_n(x)$, this is equal to-- not f_n , sorry-- $f(x) - k 2^{-n}$.

Now x , again, is an E_{kn} . So $f(x)$ is between these two numbers. So this is less than or equal to $(k + 1) 2^{-n} - k 2^{-n} = 2^{-n}$. And this equals 2^{-n} . So whenever we have an x in one of these sets, it has to be within 2^{-n} .

The simple function evaluated at that x has to be within 2^{-n} of f . This is just by construction, by how we've cut up the y -axis, if you like, the range. We're cutting it up with resolution 2^{-n} . All right, so that proves the claim which, as I said along with part A , proves B and C .

So that proves that every measurable function is almost at least non-negative extended real valued measurable function is a limit of a sequence of simple functions. Now, this theorem carries over without difficulty to complex value functions after I just introduce a breakup of a function in general.

So if E is a function from minus infinity to infinity, so an extended real-value function now, we define its positive and negative part f^+ of x to equal the max of $f(x)$, 0 . So this is the positive part of f . f^- of x is equal to the min of-- the max, I'm sorry-- minus $f(x)$, 0 . This is the negative part of f .

So what about these positive and negative parts? Then f is equal to $f^+ - f^-$. You can just check that. I mean, take any x if f is positive or non-negative on that, then I get $f(x)$ out. If it is negative, then I will get minus $f(x)$, which is the absolute value times minus gives me back $f(x)$. And the absolute value of f is equal to $f^+ + f^-$.

So let me make a comment that if-- this is just a definition for an arbitrary function from E to the extended real numbers-- if f is measurable, then each of these functions is measurable because this is-- if you like the supremum of the sequence of functions given by f and then 0 afterwards.

And this is given by the supremum of functions given by minus f and 0 afterwards. So if f is measurable, it's positive. And negative parts are also measurable functions. And they're also non-negative. So maybe that wasn't clear, or at least it should be clear. This is the max always involving 0, so it's always non-negative.

So now the construction we did a minute ago essentially carries over to the case of complex value measurable functions. So let E be measurable and f from E to \mathbb{C} be measurable. Then there exists sequence of simple functions ϕ_n such that-- analogs of those three properties hold, A is for all x in E . These are increasing in modulus, in absolute value.

0 is bigger than or equal to-- of course ϕ_0 of x is less than or equal to the absolute value of ϕ_1 of x less than or equal to the, let's say absolute value and modulus, the same thing-- is less than or equal to the absolute value of f of x .

These ϕ_n 's are converging to f pointwise. And, finally, convergence is uniform on sets where f is bounded. So for all B bigger than or equal to 0-- What did I say a minute ago. Oh, ϕ_n convergence to f uniformly on the set x in E such that the absolute value of f of x , or the modulus of f of x , is less than or equal to B .

So this theorem follows immediately from the previous theorem because now what I do is I just take f . I split it into its real and imaginary parts. And then I split the imaginary-- the real and imaginary parts into both positive and negative parts.

So I will leave it to you to actually fill in the details, but you apply the previous theorem to the real part of f plus or minus, which are now non-negative measurable functions. And the positive and negative parts of the imaginary part of f , again, which are non-negative measurable functions, and then you just take linear combinations of these simple functions that add up to f .

You will take the sequence of simple functions corresponding to the real part of f and subtract the sequence of simple functions that you got for the minus for the negative part of f , real part of f . And you will take that and add i times the positive part or the sequence of simple functions converging to the positive part of the imaginary part of f minus the sequence of functions, simple functions converging to the negative part of the imaginary part of f . So that's what I mean by apply previous theorem to the positive and negative parts of the real and imaginary parts of f .

So what's the significance of this theorem? Not only-- let me just say if-- also showing that somehow measurable functions are well approximated or almost simple functions, this also gives us a way of maybe defining the integral at least of non-negative functions that way we don't have to deal with possible deals with subtracting infinity from infinity, but by simply defining its integral to be the limit of these integrals of these simple functions.

And for a simple function, we would presumably know how to define an integral. It would just be the numbers time the measure of the sets that appear in the indicator functions for these simple functions. And like I said last lecture, if you wanted to define the Lebesgue integral that way, you would run up against, well, does this number depend on the sequence of simple functions you chose to approximate f ?

But we're not going to define the Lebesgue integral in that way. We're going to define it a little bit differently, which is what we're going to move on to now, which is the Lebesgue integral of a non-negative function. And then we will define the Lebesgue integral, or Lebesgue integrable functions.

These will be complex value functions now for which we can define an integral for. And that is the full theory of-- and that's the end of the game for as far as defining Lebesgue integral. And then we'll prove some convergence theorems along the way, which make Lebesgue integral stronger than Riemann integral.

So now we're moving on to the Lebesgue integral of a non-negative function. And why start with a non-negative function? Again, because I just pulled this trick on you a minute ago, that if somehow we know how to do stuff for non-negative measurable functions, then by playing this game where we take the real and imaginary parts and split it up into positive and negative parts, hopefully we can do something for general functions.

So that's why we start out with introducing or defining the Lebesgue integral of a non-negative function. So definition, if E is a measurable subset of \mathbb{R} , we define L^+ of E . This is a set of all extended real-value functions, non-negative real-value functions that are measurable .

And now what the goal is is to define the integral of a function like that's in this class. And it may be an infinite number. It may not be. And in order to do that we're first going to define how to integrate the simplest type of functions-- well, simple functions.

So let ϕ be a simple function. And let's write ϕ in this canonical way. So ϕ is equal to a sum of $a_j \chi_{A_j}$ where for all i not equal to j -- well, first off, for all j , A_j is a subset of E . For all i not equal to j , A_i intersect A_j is empty. So these are disjoint. And their union gives me the set E .

The Lebesgue integral of the simple function ϕ is the number, which I said is the simplest way, or what you would expect to define-- how to define the integral of a simple function.

So we know how to measure sets. And remember we initially built up measure so that the integral, which would be a theory of area underneath the curve should be-- the integral of the indicator function should be the area underneath the curve of 1, area being the measure of the set where that indicator function is 1.

And therefore by linearity, this will be how we define the Lebesgue integral of a simple function. So the Lebesgue of ϕ is the following number, $\int_E \phi$ -- this is how it's defined-- sum from j equals 1 to n a_j measure of A_j . And this number could be infinite.

All right, and instead of writing just the integral over E of ϕ , I might add a dx in there. I'm just warning you ahead of time. All right, so this is the Lebesgue integral of a simple function. Again, we split it up into this canonical way where it's just the indicator function of disjoint sets which whose union gives the set of the domain and where these numbers out in front, these coefficients, give you the numbers that go-- of the elements of the range.

So, for example again, I mean, this is-- let's say A is from-- or let's say E is the interval a, b . You can check that-- well, I mean, it's just from the definition that if my simple function is, in fact, these sets are just intervals, so if this is what my simple function looks like, it takes finitely many values, this one, this one, this one, this one, this one.

And on these sets whose disjoint union forms a, b , then as I've defined and since-- as I've defined the integral and since the measure of an interval equals the length of the interval, again, you should count this as some more comments. This integral equals the sum of-- so this would be $\sum a_j$.

Then how I've defined the Lebesgue integral, it spits out the area underneath this simple function that is taking these values on these intervals. I hope that's clear. It's towards the end of the day so maybe my explanations are getting a bit wonky, but I hope it's clear.

And then we're going to use this definition of simple-- of how to integrate simple functions. We're going to extend it to general elements of L^1 plus L^2 , the non-negative measurable functions. But first let's prove a few properties of how we've defined the integral for simple functions.

Let's take two simple functions. First is that if c is bigger than or equal to 0, then the integral of c times ϕ over E , this is equal to c times the integral of ϕ over E . Two, the Lebesgue integral over E of ϕ plus ψ is equal to the Lebesgue integral of ϕ plus the Lebesgue integral of ψ .

And third property is if ϕ is less than or equal to ψ on E , so I'm writing that shorthand by that. So what I'm saying in that statement is for all x in E , $\phi(x)$ is less than or equal to $\psi(x)$, then what you expect is the integral of ϕ is less than or equal to the integral of ψ .

And let me include one more very simple property. If f is-- let's add a little punctuation. If f is a measurable subset of E , then ϕ which is a simple function on E is also a simple function on f .

It takes only finitely many values on f as well, and therefore it has an integral over f . And my claim is that is equal to the integral over E of ϕ times the indicator function of f . And that's less than or equal to the integral over E .

Now I will leave problem 4 to you. I might even put on the assignment. But it will follow once you've seen how we prove 1, 2, and 3. You'll be like, OK, so I know how to do this. So number 1 is pretty easy simply because multiplying by a non-negative constant just carries through and just changes the constant, but not the sets.

So c times ϕ is equal to c times a sum $\sum_j \phi_j$. And therefore the integral of c times ϕ over E , this is by definition equal to c times the sum from $j=1$ to n of c times $\int \phi_j$ measure of a sub J_j . And this is equal to c times sum from $j=1$ to n of $\int \phi_j$ measure of a sub J_j . And this equals c times the integral of ϕ over E .

All right, so to prove 2 we write ϕ in this canonical form, sum over a sub J_j times the indicator function of a sub J_j 's where, again, these sets are disjoint. And their union gives me E . And then also I do the same thing for ψ , now k equals 1 to m , maybe take on different number of values, ψ_k , where again the B_k 's are disjoint and their union gives me E .

So since the union of the A_j 's gives me E and the union of the B_k 's give me E , this implies that if I want to look at one of these A_j 's it's equal to a union of a certain type, is equal to the union $k=1$ to m of $A_j \cap B_k$ because this is just going to be equal to $A_j \cap$ the union of the B_k 's, but the union of the B_k 's gives me E , and similarly for the B_k 's because the union of A_j 's gives me E .

And, again, and these unions are disjoint. The A_j 's are disjoint from each other. The B_k 's are disjoint from each other. So for all j and k , $A_j \cap B_k$ is going to be disjoint from another $A_{j'} \cap B_{k'}$ when $j \neq j'$ or $k \neq k'$.

So since these are disjoint unions, we have from the additivity property of Lebesgue measure, we get that the integral over ϕ on E plus the integral over E of ψ , this is equal to, by definition, $\sum_{j=1}^n \mu(A_j)$ plus $\sum_{k=1}^m \mu(B_k)$.

Now, A_j is written as this disjoint union. And therefore the measure of A_j is equal to the sum of the measure of these, of $A_j \cap B_k$ plus... and then the same thing here. The measure of... so B_k is equal to this disjoint union. So its measure is the sum of the measure. So this sum now includes j . So let me just rewrite this as sum over j, k , $\mu(A_j \cap B_k)$.

But the point is that the sum of these two simple functions you can check, this is equal to the sum over j, k of $\mu(A_j \cap B_k)$ plus B_k times the indicator function of $A_j \cap B_k$. So at those x 's when $\phi = A_j$, and at those x 's where $\phi = B_k$, then I'm in this set and the two sides agree.

This implies that the integral over E of $\phi + \psi$ is equal to measure of $A_j \cap B_k$, which as we just saw is equal to the sum of the integrals. So that was 2. 3 is not too difficult either.

So, again, let's assume that ϕ and ψ are written in that way, in this canonical way. Then for all x in E , $\phi(x)$ is less than or equal to $\psi(x)$. This is equivalent to $A_j \cap B_k$ is nonempty whenever $A_j \cap B_k$ is nonempty.

Thus, again, we're going to use the additivity of the Lebesgue measure and the fact that the unions of these B_k 's give me E , then if I look at the interval of E of ϕ , this is $\sum_{j=1}^n \mu(A_j)$. This is equal to the sum $\sum_{j,k} \mu(A_j \cap B_k)$ because the A_j is a union over $A_j \cap B_k$. And that's a disjoint union.

And now whenever this is non-zero, that means that $A_j \cap B_k$ is nonempty. And therefore that $\mu(A_j \cap B_k)$ appearing there is going to be less than or equal to $\mu(B_k)$. And whenever this is 0, well whatever's there is almost always less than or equal to $\mu(B_k)$ times the measure of $A_j \cap B_k$.

So this is less than or equal to $\sum_{j,k} \mu(B_k) \mu(A_j \cap B_k)$. So whenever this is... so maybe that previous explanation was not good. Whenever this is nonempty I will always have $\mu(A_j \cap B_k)$ less than or equal to $\mu(B_k)$.

So I should have just said that I don't know what I was going on about measure zero stuff, but ignore that. Whenever this is nonempty, like I just said a minute ago, $\mu(A_j \cap B_k)$ is less than or equal to $\mu(B_k)$, all right. And now we just reverse course. And this is equal to $\sum_{k=1}^m \mu(B_k)$ because union over j 's of these sets give me B_k . And this is equal to the integral of ψ on E . And, again, I will leave 4 to you as a very simple exercise.

OK, I'm about out of time. So what we've done is we've defined the Lebesgue integral of a simple function. And as that picture shows, I hope, or at least convinces you that if the simple function takes on that of a step function, meaning the A_j 's are just intervals, then the Lebesgue integral of that step function will, in fact, be the area underneath ϕ .

So it's, again, this is-- you can think of that as two ways, as the Lebesgue integral giving a theory of the area underneath the curve. You may also think of that as being the first indication that when I have a Riemann integral function, it will be also Lebesgue integral because if I have a step function like ϕ is in the picture, then that is a Riemann integral function. And the Riemann integral is the area underneath the curve, which also agrees with the definition of the Lebesgue integral.

So that, like I said, should maybe indicate that the Lebesgue integral reduces to the Riemann integral whenever we're integrating a Riemann integrable function. And so next time we will define the integral of a non-negative measurable function using how we've defined the integral of simple functions, and prove some basic properties, including two of the main convergence theorems that go along with this theory of integration. So we'll stop there.