Let's talk about Vector Spaces...

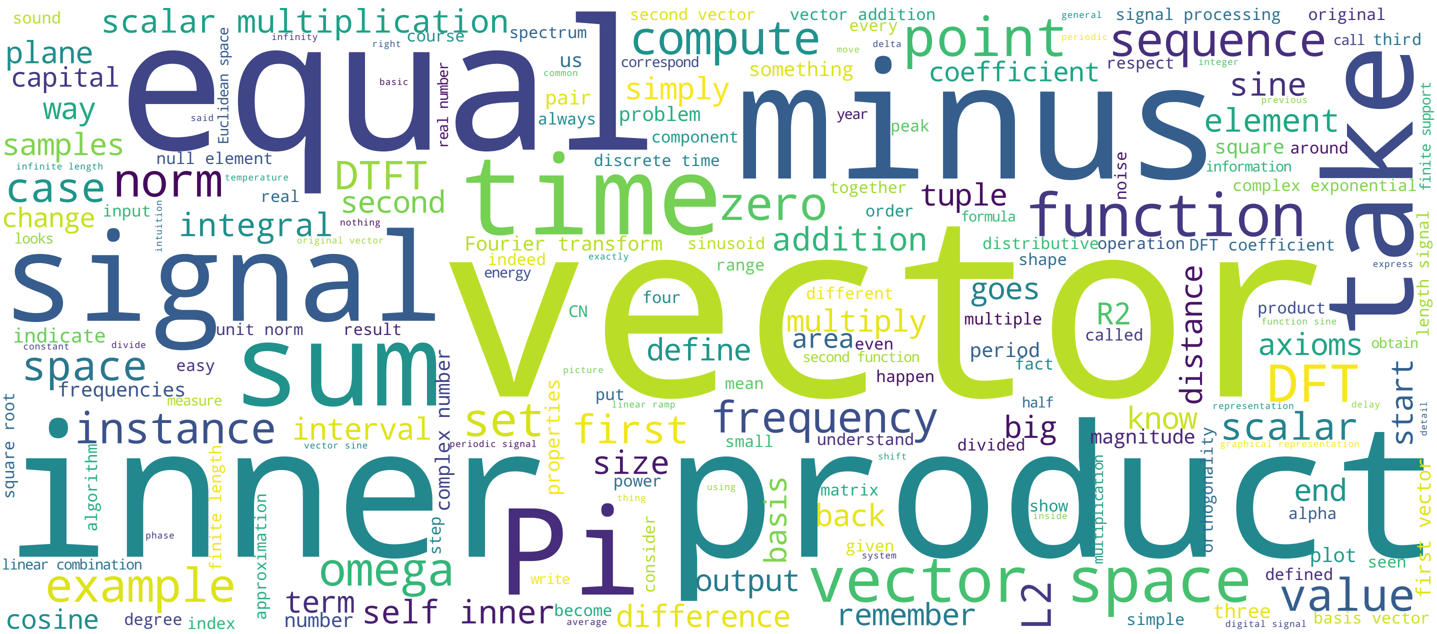
Some spaces should be very familiar:

- $\mathbb{R}^2, \mathbb{R}^3$: Euclidean space, geometry
- $\mathbb{R}^N, \mathbb{C}^N$: linear algebra

Others perhaps not so much...

- $\ell_2(\mathbb{Z})$: space of square-summable infinite sequences
- $L_2([a, b])$: space of square-integrable *functions* over an interval

yes, vectors can be functions!



Search MOOC

Video

Some spaces can be represented graphically

$$\mathbb{R}^2 : \qquad \mathbf{x} = [x_0 \quad x_1]^T$$

2

Instead of introducing vector space inductively, namely by listing the axioms that regulate the behaviour of vectors in vector space, let's start with some examples that you should be familiar with. R2 and R3 are also known as the Euclidean plane and Euclidean space. These are very familiar entities. The extension of Euclidean space to an arbitrary number of dimension is called RN, that is the subject of linear algebra. Another extension uses complex numbers instead of real numbers, so we have CN, vectors that are made of tuples of n complex numbers, again the subject of linear algebra. Other vector spaces are perhaps less known. We have L2 of Z. This is the space of square-summable infinite sequences. We will see that this is useful for modelling infinite-length signals. L2 of a, b is the space of square integrable functions over an interval. Here the key point is that vectors can be arbitrarily complex entities, such as functions for instance, and this would be very useful in unifying our approach to signal processing. Some spaces, some vector spaces can be represented graphically. This is the case for instance of planar geometry.

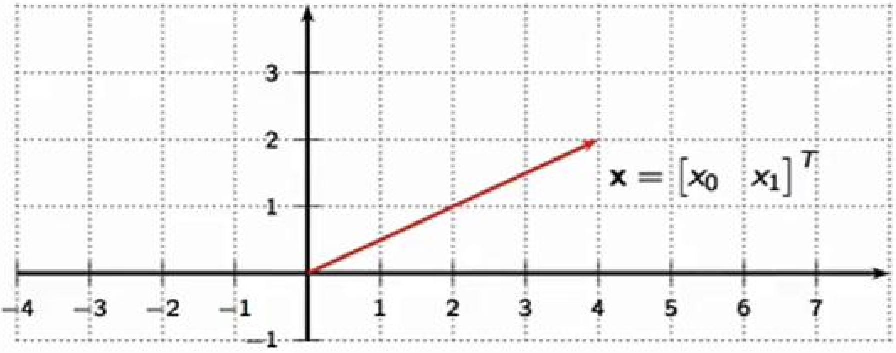
Notes

Summary

0m 00s

Some spaces can be represented graphically

$$\mathbb{R}^2: \qquad \mathbf{x} = \begin{bmatrix} x_0 & x_1 \end{bmatrix}^T$$



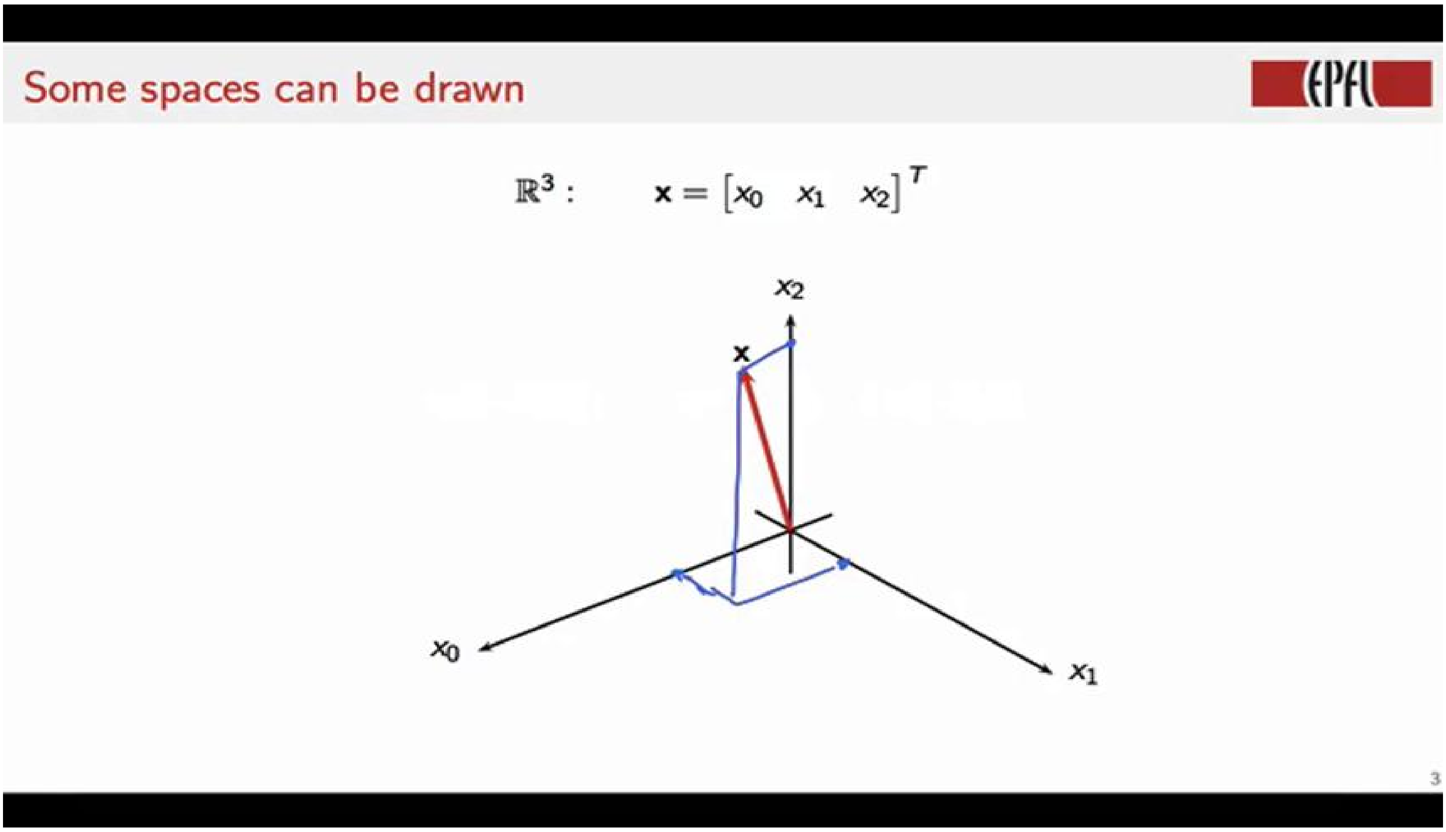
2

The intuition is that if you take the plane and set up a Cartesian reference system like in this figure here, then you can identify any point on the plane with a pair of real-valued coordinates like so. This correspondence between points on the plane and vectors in R2 allow us to interpret the properties of vectors in R2 in terms of geometric properties on the Euclidean plane. This analogy is extremely useful, especially when extended also to three dimensions because it will allow us to use our geometric intuition to understand the properties of arbitrarily complex vector spaces that don't allow for an intuitive representation or a graphical representation.

Notes

Summary

1m 19s



In three dimensions, we use triples of coordinates to locate uniquely a point in space and therefore we have a complete correspondence between Euclidean space and R3. Here we have our point in space and its 3D coordinates will be like so. Again, geometrical intuition is pretty straightforward.

Notes

Summary

2m 05s

Some spaces can be drawn

$$L_2([-1,1]): \qquad \mathbf{x} = x(t), \quad t \in [-1,1]$$

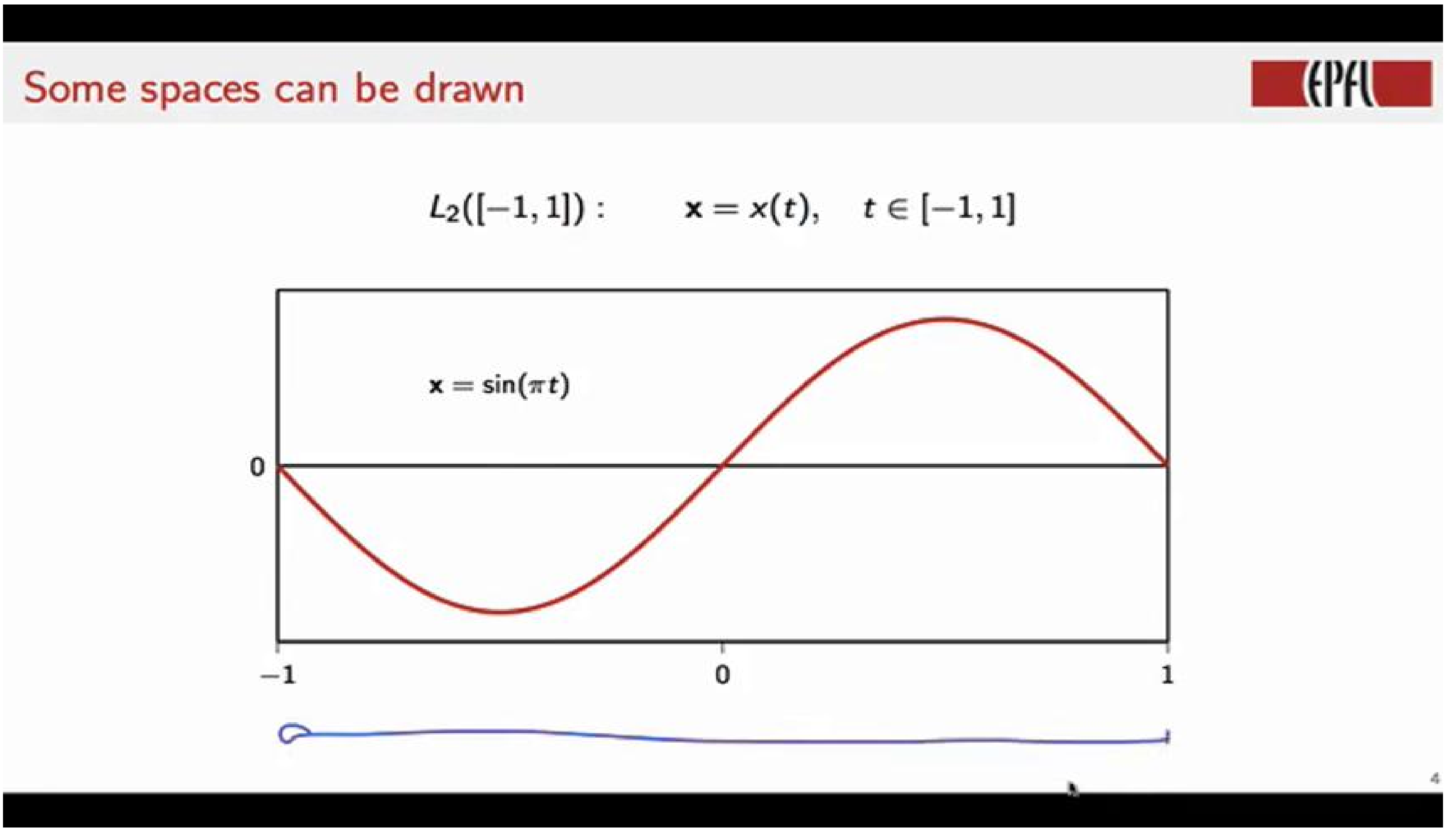
4

Let's look now at a function vector space. This is L2 of minus 1, 1 so the space of square integrable function on the interval minus 1, 1. Any element of this vector space is a function, so this is the vector notation and then explicitly, it will be a function of a variable t that goes from minus 1 to 1.

Notes

Summary

2m 25s



An element, for instance, is the sine function. We can take x equal to sine Pi t and the vector can be drawn by plotting the function over the interval.

Notes

Summary

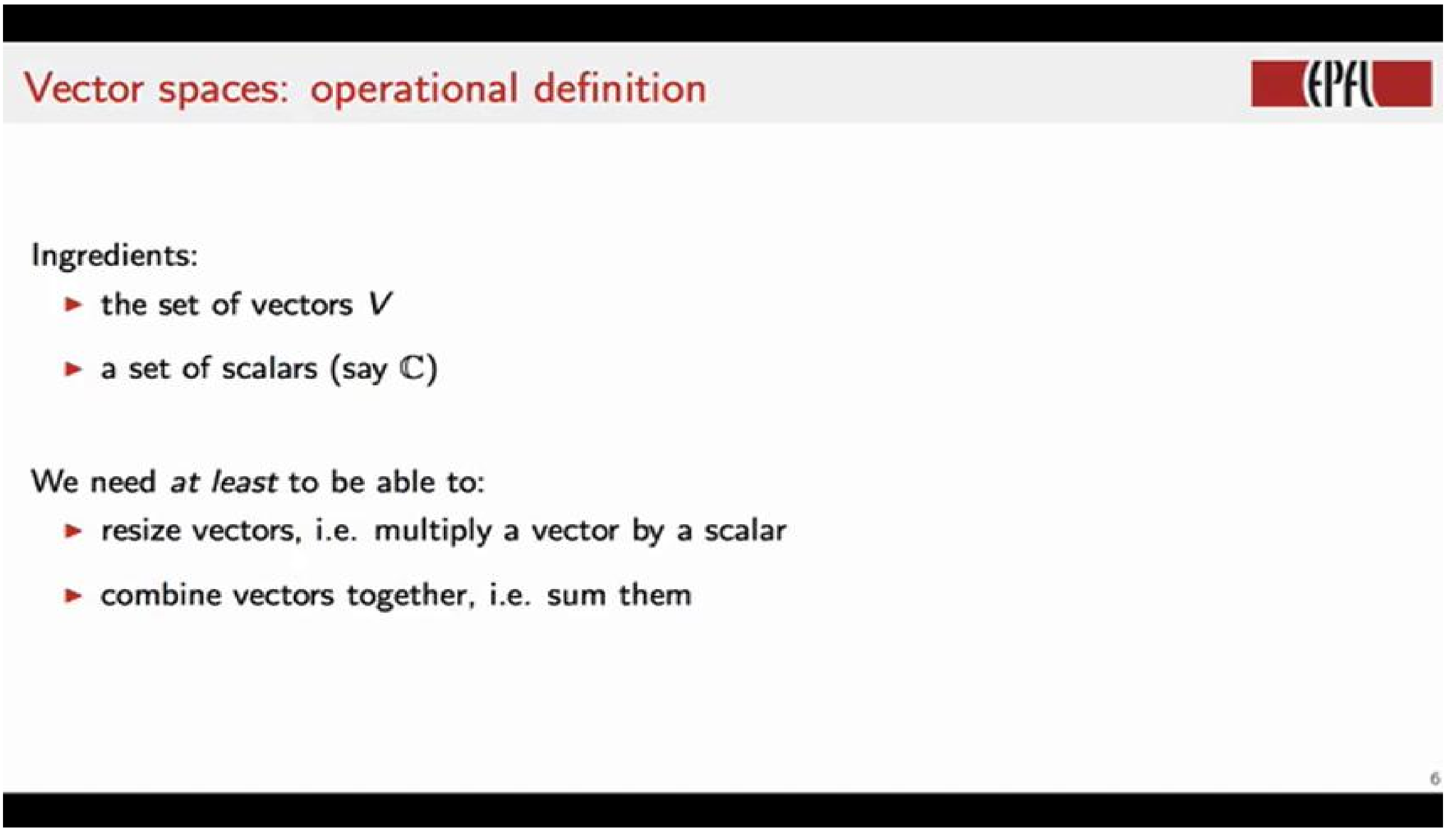
2m 46s

Other vector spaces do not admit graphical representation. For instance, RN for N greater than 3, well, there is no way that we can plot that in a way that we can comprehend. CN, so the tuples of complex numbers cannot be drawn for anything but N equal to 1, in which case, we have the classic complex plane.

Notes

Summary

2m 57s



From these examples, we can see that vector spaces can be a very diverse bunch and what all these vector spaces have in common is that they obey to a set of axioms that define the properties of the vectors and what we can do with these vectors. The ingredients for a basic vector space are a set of vectors, V and we say nothing about what is inside these vectors and a set of scalars that in our examples, will always be the set of complex numbers. We need to be able to at least resize vectors, so multiply them by a scalar to change their size and combine vectors together, sum them together.

Notes

Summary

3m 17s

Formal properties of a vector space:

For $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ and $\alpha, \beta \in \mathbb{C}$:

- $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
- $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$
- $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{y} + \alpha\mathbf{x}$
- $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$
- $\alpha(\beta\mathbf{x}) = (\alpha\beta)\mathbf{x}$
- $\exists 0 \in V \quad | \quad \mathbf{x} + 0 = 0 + \mathbf{x} = \mathbf{x}$
- $\forall \mathbf{x} \in V \ \exists(-\mathbf{x}) \quad | \quad \mathbf{x} + (-\mathbf{x}) = 0$

7

The way we do this is by defining some axiomatic properties for the sum and for the scalar multiplication. This is pretty dry but we will go through these axioms and then we will understand that all the vector spaces that we have seen in the examples actually do fulfil these axioms. We want the addition to be commutative, we want the addition to be distributive, we want the scalar multiplication to be distributive with respect to vector addition, and we want scalar multiplication to be distributive with respect to the addition in the field of scalars. Scalar multiplication is associative as well. We want to have a null element in the vector space so that the sum of any vector plus the null element and because of commutativity the sum of the null element plus the vector will give back the original vector. Then we want that for every element in the vector space, there exists an inverse element for the addition so that x plus minus x as we indicate the inverse for the addition will be equal to the null vector.

Notes

Summary

3m 57s

Vector space example: $\mathbb{R}^N$

$$\mathbf{x} = \begin{bmatrix} x_0 & x_1 & \dots & x_{N-1} \end{bmatrix}^T$$

$$\mathbf{y} = \begin{bmatrix} y_0 & y_1 & \dots & y_{N-1} \end{bmatrix}^T$$

$$\alpha\mathbf{x} = \begin{bmatrix} \alpha x_0 & \alpha x_1 & \dots & \alpha x_{N-1} \end{bmatrix}^T$$

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_0 + y_0 & x_1 + y_1 & \dots & x_{N-1} + y_{N-1} \end{bmatrix}^T$$

8

All right so let's look at an example. If we take RN, the space of tuples of capital N real numbers, and we define the basic operations like so. We have scalar multiplication. We take a tuple and we multiply each element in the tuple by a scalar alpha. Then the sum is defined as a tuple where each element is the sum of the corresponding elements of the two tuples. If we define the operations like so, then we can easily verify that these operations fulfil the axioms and therefore RN is a valid vector space.

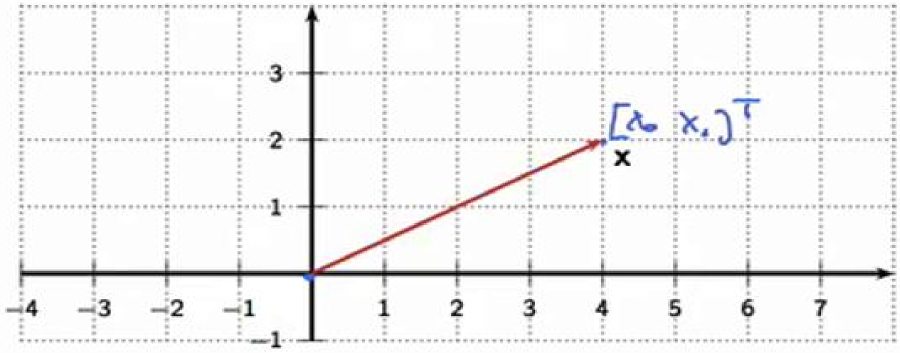
Notes

Summary

5m 04s

Scalar multiplication in $\mathbb{R}^2$

$$\alpha \mathbf{x} = [\alpha x_0 \quad \alpha x_1]^T$$



9

We can, as we said, use the geometric intuition in low dimensions to understand what these operations do. Consider the scalar multiplication of a vector in R2 while the vector is identified by a pair of coordinates x0 and x1. If we multiply each component in the pair by alpha, we are in fact lengthening the vector by a factor-alpha without changing its direction.

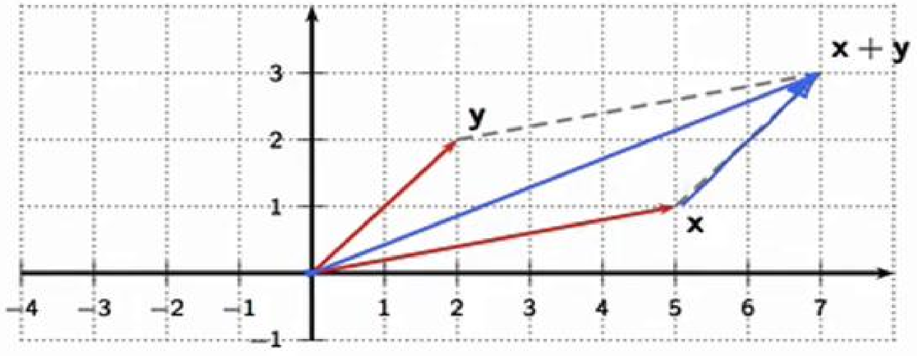
Notes

Summary

5m 39s

Addition in $\mathbb{R}^2$

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_0 + y_0 & x_1 + y_1 \end{bmatrix}^T$$



10

Vector addition in two dimensions corresponds to summing the first element of each tuple together and this becomes the new first element of the sum and then we sum the second components together. Graphically we take our first vector, we take the second vector and we place it at the end of the first vector without changing its direction and the new vector will be the arrow that goes from the origin to the end of this combination of vectors. This is the famous parallelogram rule.

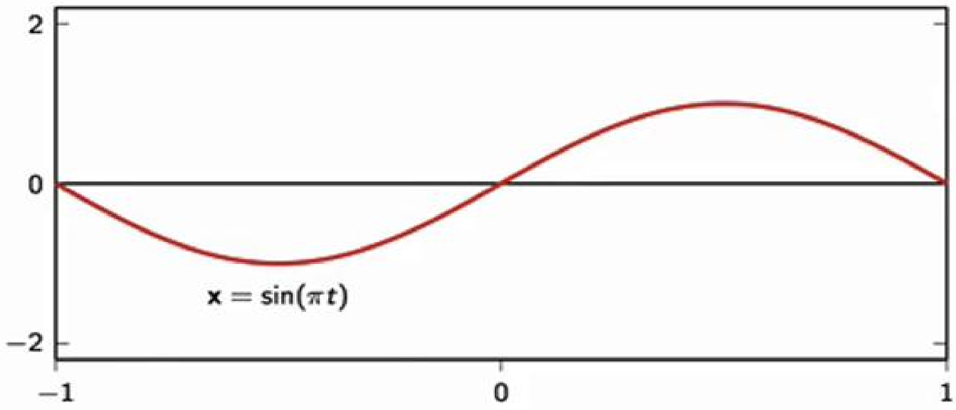
Notes

Summary

6m 07s

Scalar multiplication in $L_2[-1, 1]$

$$\alpha\mathbf{x} = \alpha\, x(t)$$



11

In L2 of minus 1, 1, scalar multiplication is intuitively defined as multiplying the function by a scalar alpha. Here we have the vector sine of Pi t, we multiply this vector by a scalar 1.5 and what we get is a function with a wider amplitude.

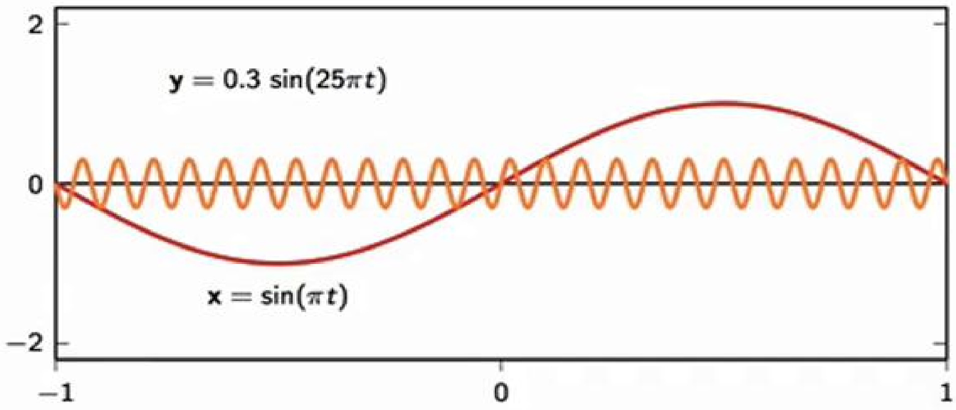
Notes

Summary

6m 37s

Addition in $L_2[-1, 1]$

$$\mathbf{x} + \mathbf{y} = x(t) + y(t)$$



12

Sum is defined intuitively as well in L2 of minus 1, 1 as simply the sum of the two functions over the interval. We take a first vector sine of Pi t, a second vector 0.3 times sine of 25 Pi t.

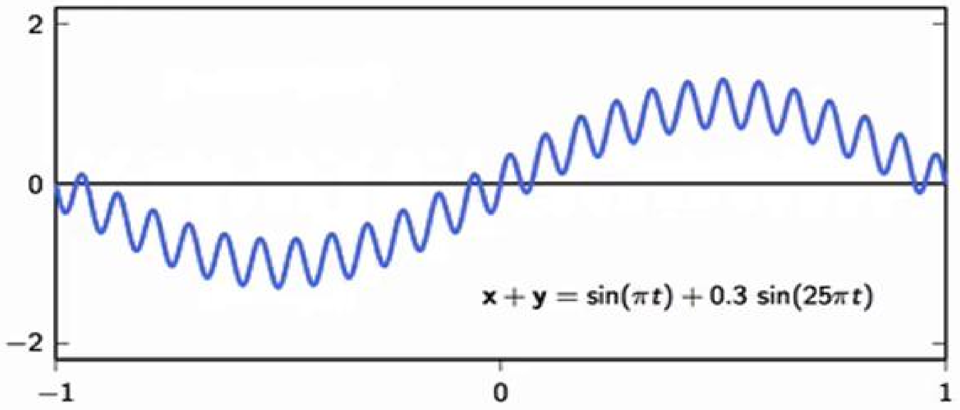
Notes

Summary

6m 57s

Addition in $L_2[-1, 1]$

$$\mathbf{x} + \mathbf{y} = x(t) + y(t)$$



12

We sum them together and we get this other element of the vector space.

Notes

Summary

7m 12s

Vector spaces: we need something more

- the set of vectors V
- a set of scalars (say $\mathbb{C}$)
- scalar multiplication
- addition

We need something to measure and compare:
inner product (aka dot product)

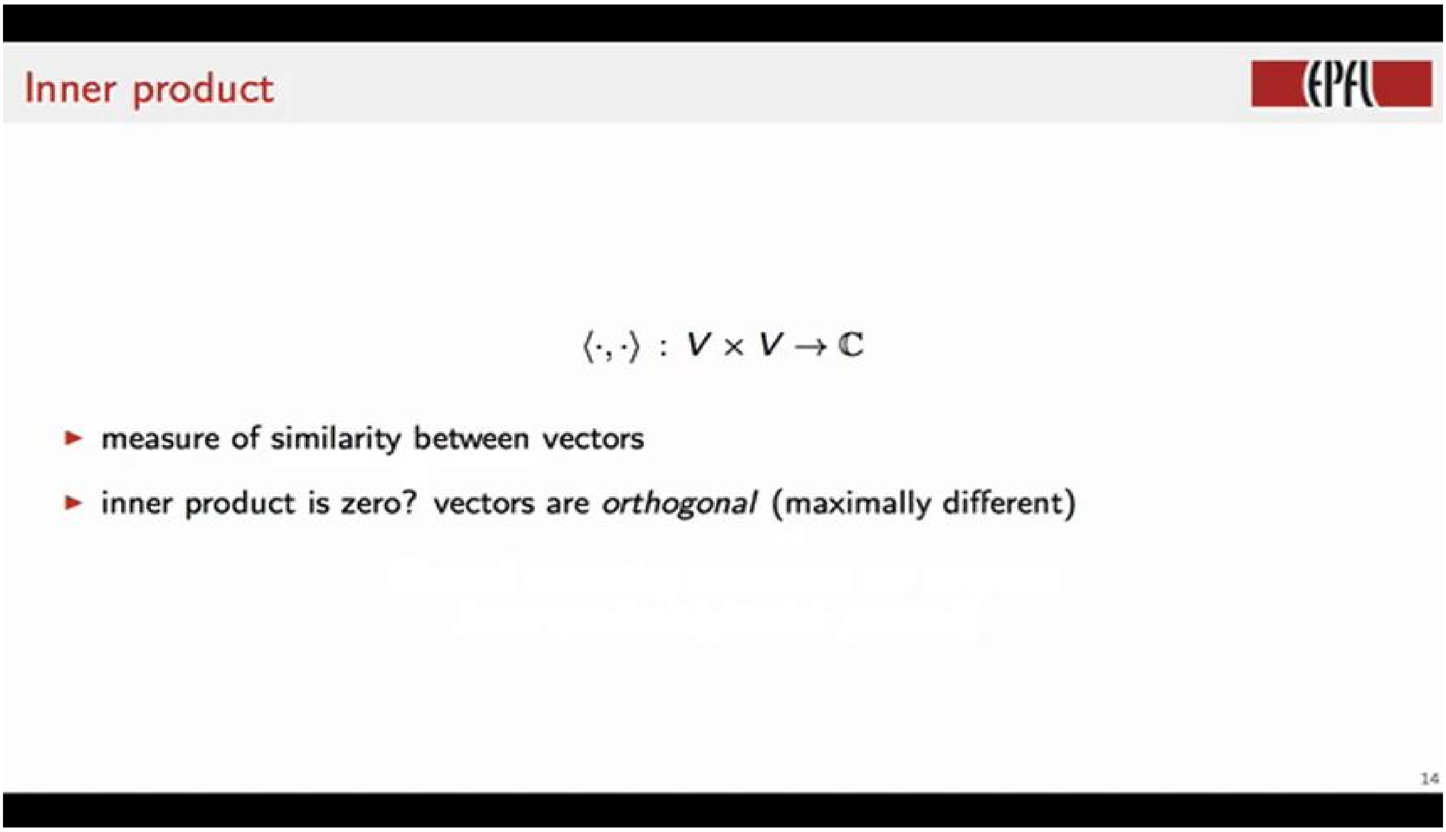
13

Now, we have the set of vectors, set of scalars, we know how to multiply vectors by a scalar and we know how to put vectors together by addition. But we need something more. We need something to measure and compare vectors. This is provided by the inner product which is an additional new operator that we defined for the vector space.

Notes

Summary

7m 18s



The inner product takes a couple of vectors and returns a scalar which is a measure of similarity between the two vectors. If the inner product is 0, this is a very important concept that we will examine in detail, we say that the vectors are maximally different or in other words orthogonal.

Notes

Summary

7m 36s

Formal properties of the inner product

For $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ and $\alpha \in \mathbb{C}$:

- $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$
- $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle^*$
- $\langle \alpha\mathbf{x}, \mathbf{y} \rangle = \alpha^* \langle \mathbf{x}, \mathbf{y} \rangle$
  $\langle \mathbf{x}, \alpha\mathbf{y} \rangle = \alpha \langle \mathbf{x}, \mathbf{y} \rangle$
- $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$
- $\langle \mathbf{x}, \mathbf{x} \rangle = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$
- if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ and $\mathbf{x}, \mathbf{y} \neq \mathbf{0}$ then $\mathbf{x}$ and $\mathbf{y}$ are called orthogonal

15

Like for scalar multiplication and addition, the inner product is defined axiomatically and these are the properties that it will have to fulfil. We want the inner product to be distributive with respect to vector addition. We want the inner product to be commutative with conjugation. This of course applies when our set of scalars is complex-valued. We want the inner product to be distributive with respect to scalar multiplication and we conjugate the scalar if it affects the first upper hand in the inner product. We want the self inner product to be greater than or equal to 0, which means that the self inner product is necessarily a real number. Finally, the self-inner product can be 0 only if the element is the null element for the vector addition.

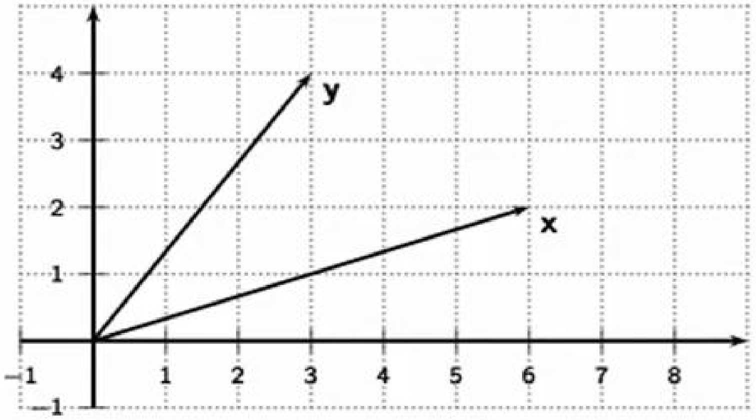
Notes

Summary

7m 55s

Inner product in $\mathbb{R}^2$

$$\langle \mathbf{x}, \mathbf{y} \rangle = \underbrace{x_0 y_0} + \underbrace{x_1 y_1}$$



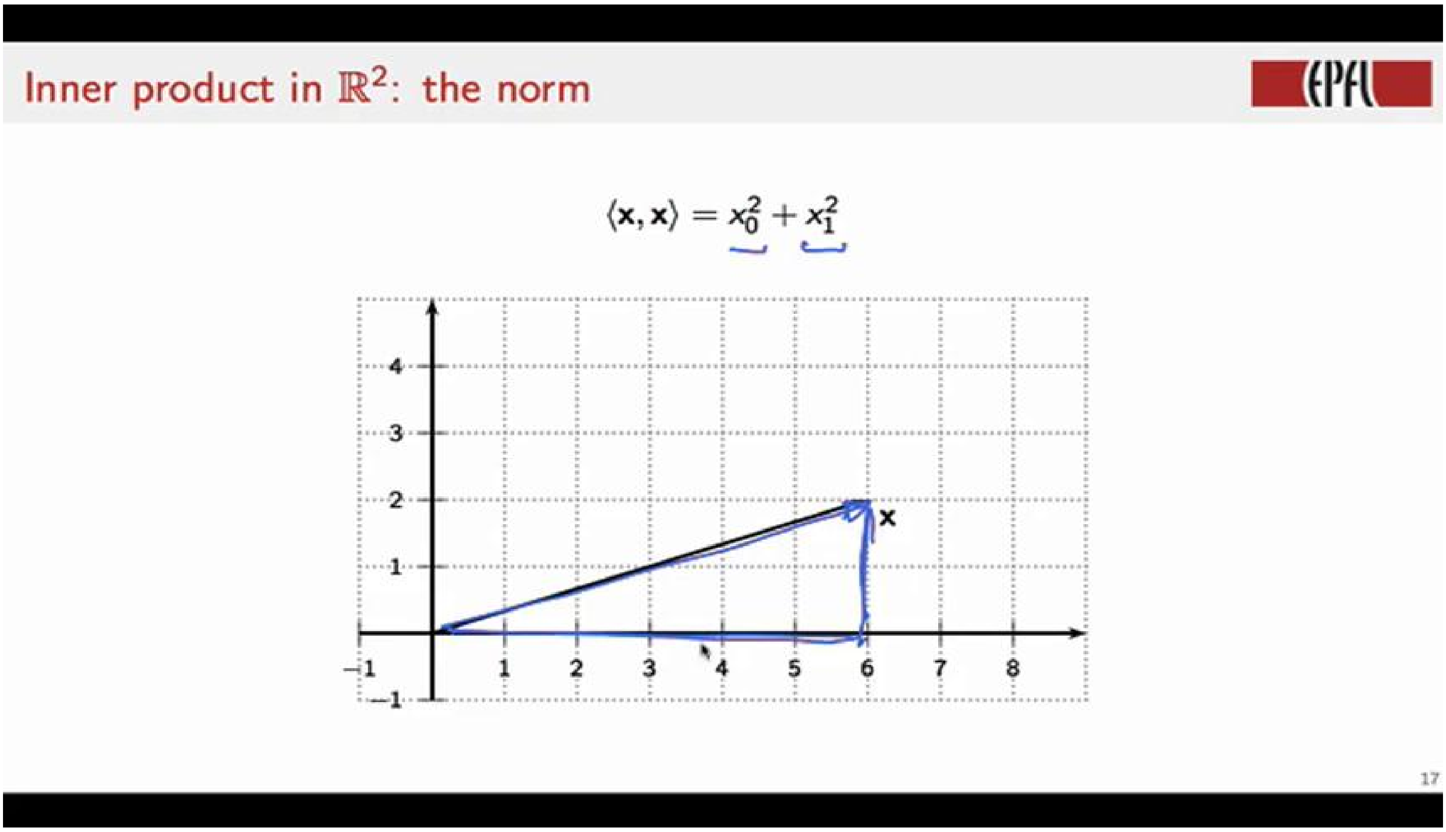
16

Let's look at some examples of inner products to foster our intuition. In R2, the inner product between two vectors is defined as the product of the first component of each pair plus the product of the second components of each pair. Now, this is not immediately intuitive so let's take it apart a bit.

Notes

Summary

8m 45s



First of all, if we look at the self-inner product of a vector, we can see that it's equal to x0 square plus x1 square, so the sum of the coordinates square. Now, if we look at the graphical representation of the vector, we can see that this Pythagoras theorem is equal to the length of the vector squared.

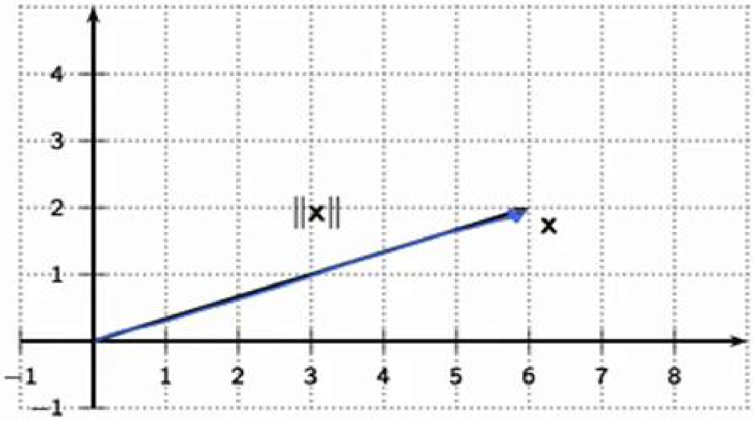
Notes

Summary

9m 04s

Inner product in $\mathbb{R}^2$: the norm

$$\langle \mathbf{x}, \mathbf{x} \rangle = x_0^2 + x_1^2 = \|\mathbf{x}\|^2$$



17

We actually have a name for this, we call this the norm of a vector. The norm of a vector is the square root of the self-inner product and it indicates the length of the vector. The self-inner product is a very good measure of the size of a vector in R2 and by extension in all vector spaces.

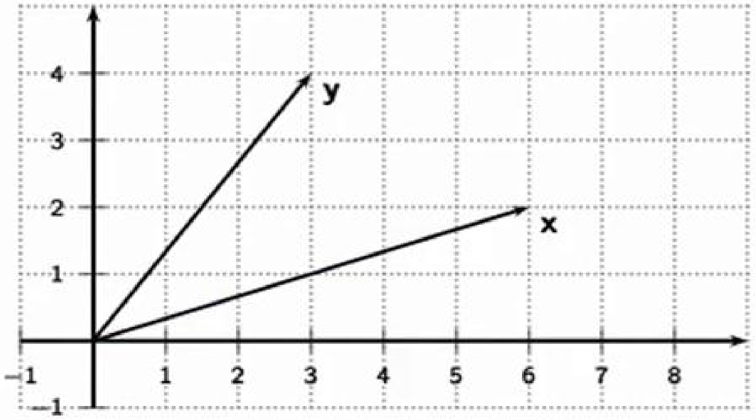
Notes

Summary

9m 26s

Inner product in $\mathbb{R}^2$

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_0 y_0 + x_1 y_1$$



18

Let's go back to the general definition for the inner product. We could use some simple trigonometry to show that this formulation is actually equivalent to this.

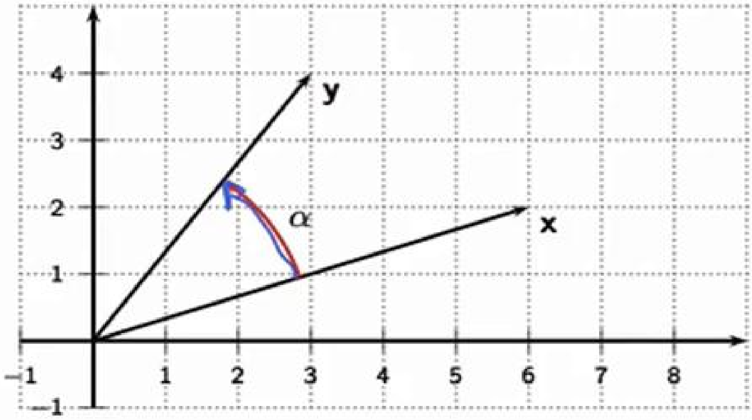
Notes

Summary

9m 46s

Inner product in $\mathbb{R}^2$

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_0 y_0 + x_1 y_1 = \|\mathbf{x}\| \, \|\mathbf{y}\| \, \cos\alpha$$



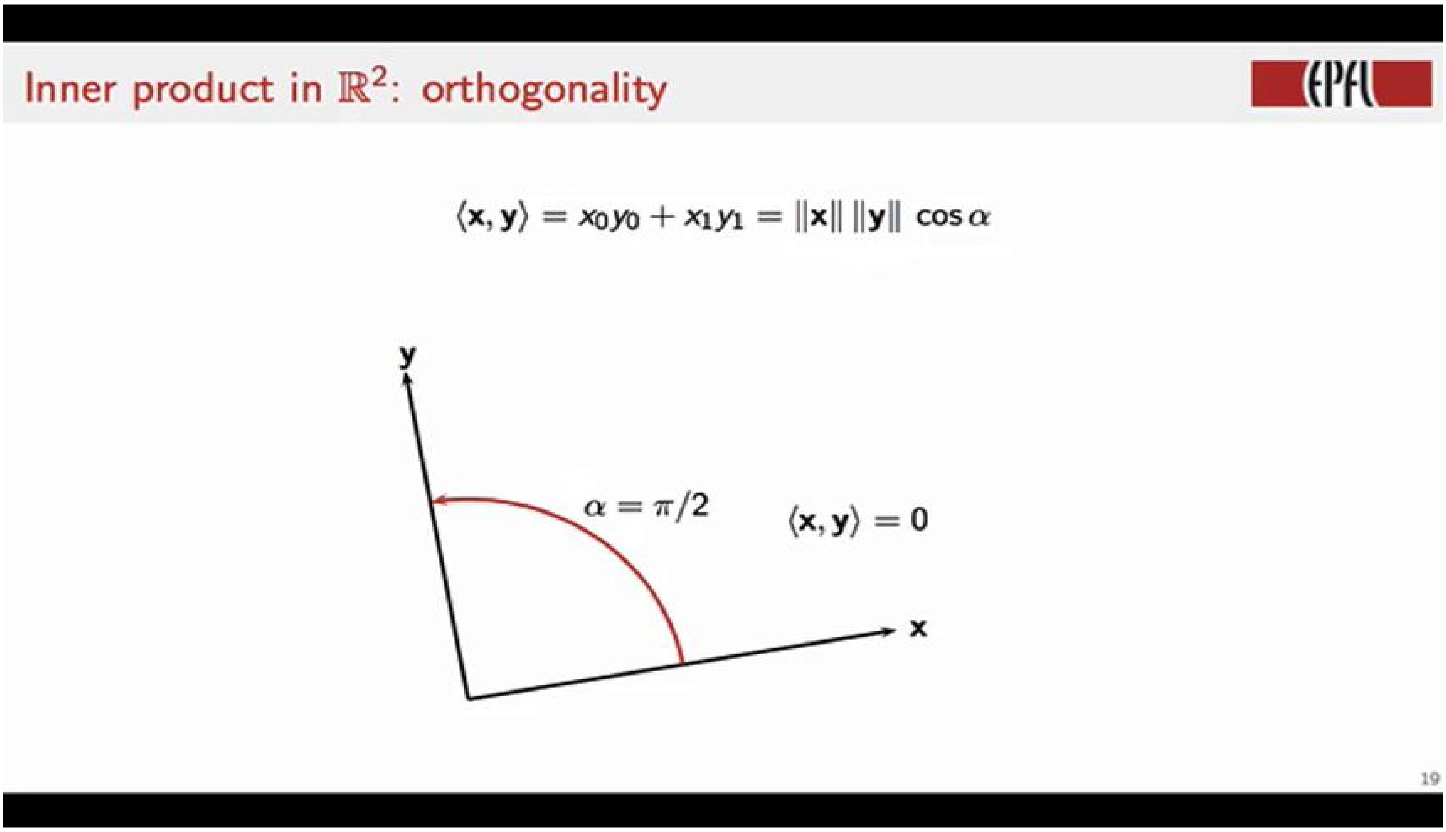
18

The inner product between two vectors is equal to the norm of the first vector times the norm of the second vector times the cosine of the angle formed by the two vectors. Now if the two vectors have equal norm, the inner product will be a good measure of similarity between these two vectors because it will go from the norm square when the vector coincide and therefore the angle alpha is zero to zero when the vectors are at 90 degrees and therefore are pointing in orthogonal directions.

Notes

Summary

9m 55s



Orthogonality is indeed the maximal difference between two vectors on the plane.

Notes

Summary

10m 27s

Inner product in $L_2[-1, 1]$

$$\langle \mathbf{x}, \mathbf{y} \rangle = \int_{-1}^{1} x(t)y(t)\, dt$$

20

As another example, let's now look at the inner product in L2 of minus 1, 1. This is usually defined as the integral of the product of the two functions that corresponds to the two vectors. It's very easy to verify that this inner product fulfils the axioms.

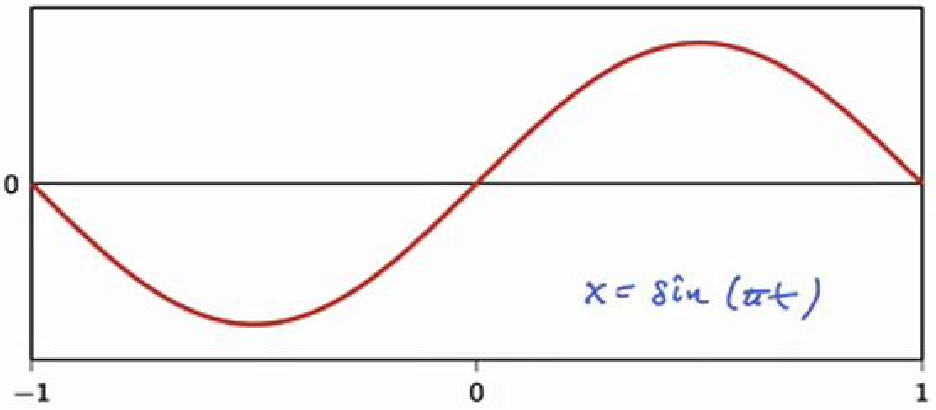
Notes

Summary

10m 32s

Inner product in $L_2[-1,1]$: the norm

$$\langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|^2 = \int_{-1}^{1} \sin^2(\pi t)dt = 1$$



21

Let's now use this inner product to compute the norm of the vector x equal to sine of Pi t. This will be equal to the self-inner product and so the integral from minus 1 to 1 of sine square of Pi t in the t.

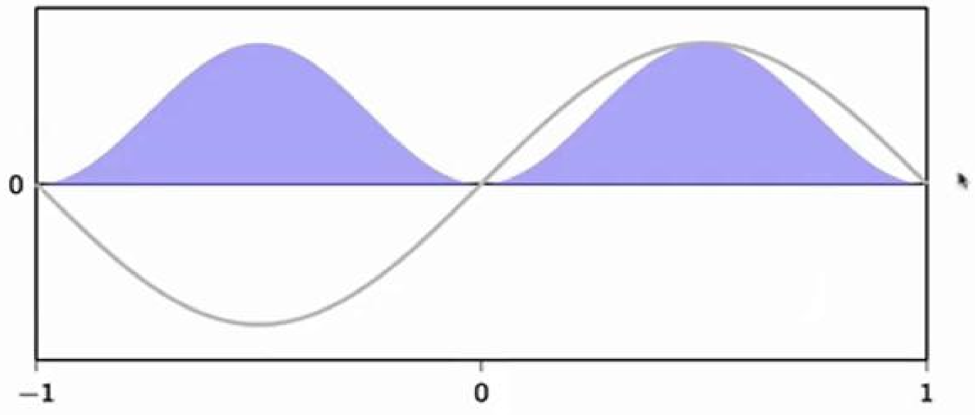
Notes

Summary

10m 48s

Inner product in $L_2[-1,1]$: the norm

$$\langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|^2 = \int_{-1}^{1} \sin^2(\pi t)dt = 1$$



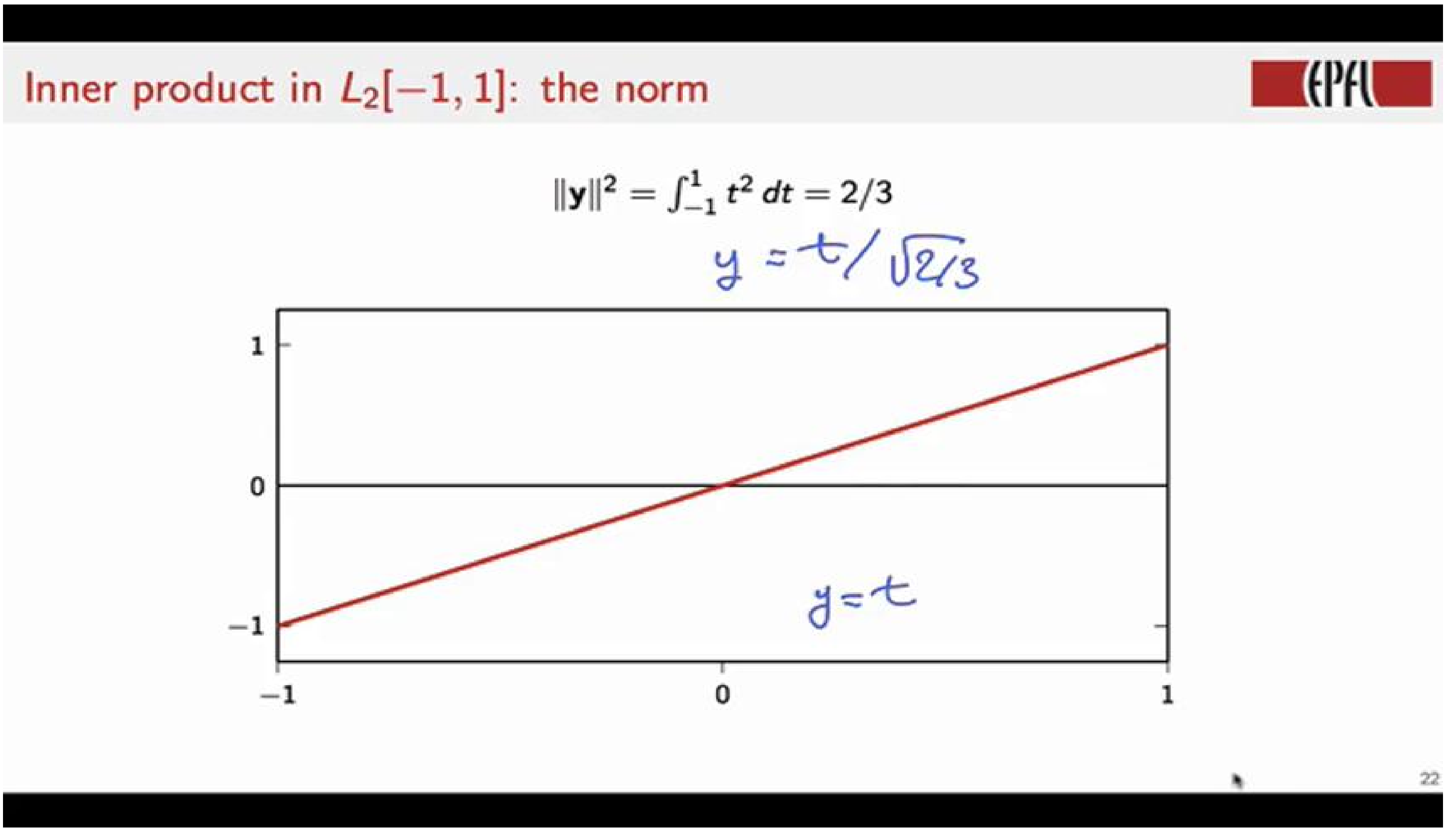
21

This is the area that is shaded in the picture and it's equal to 1.

Notes

Summary

11m 03s

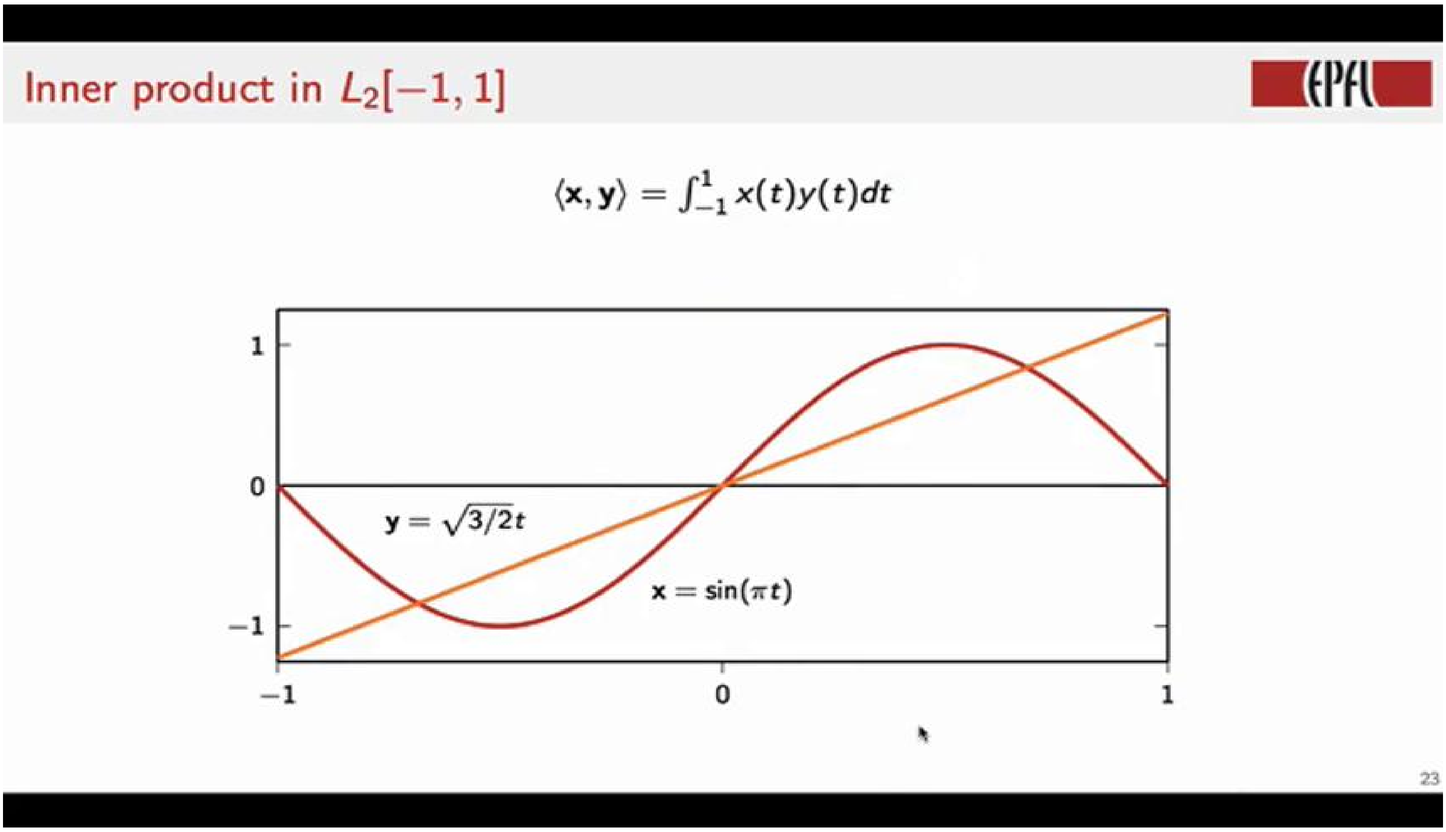


Let's take another element of L2 minus 1, 1. The vector y equal to the function t, so the linear ramp between minus 1 and 1. If we compute the norm as the self inner product of this vector, we have the integral from minus 1 to 1 of t square in the t which is equal to two-thirds. This vector doesn't have unit norm. If we want to normalise this vector, we can define an alternate version where you say y is equal to t divided by square root of two-thirds. We divide the vector by the norm and this vector will now have unit norm. The reason why we do that is because now, we can use the inner product to compare in L2 of minus 1, 1 the function sine of Pi t with the function t divided by square root of two-thirds.

Notes

Summary

11m 09s

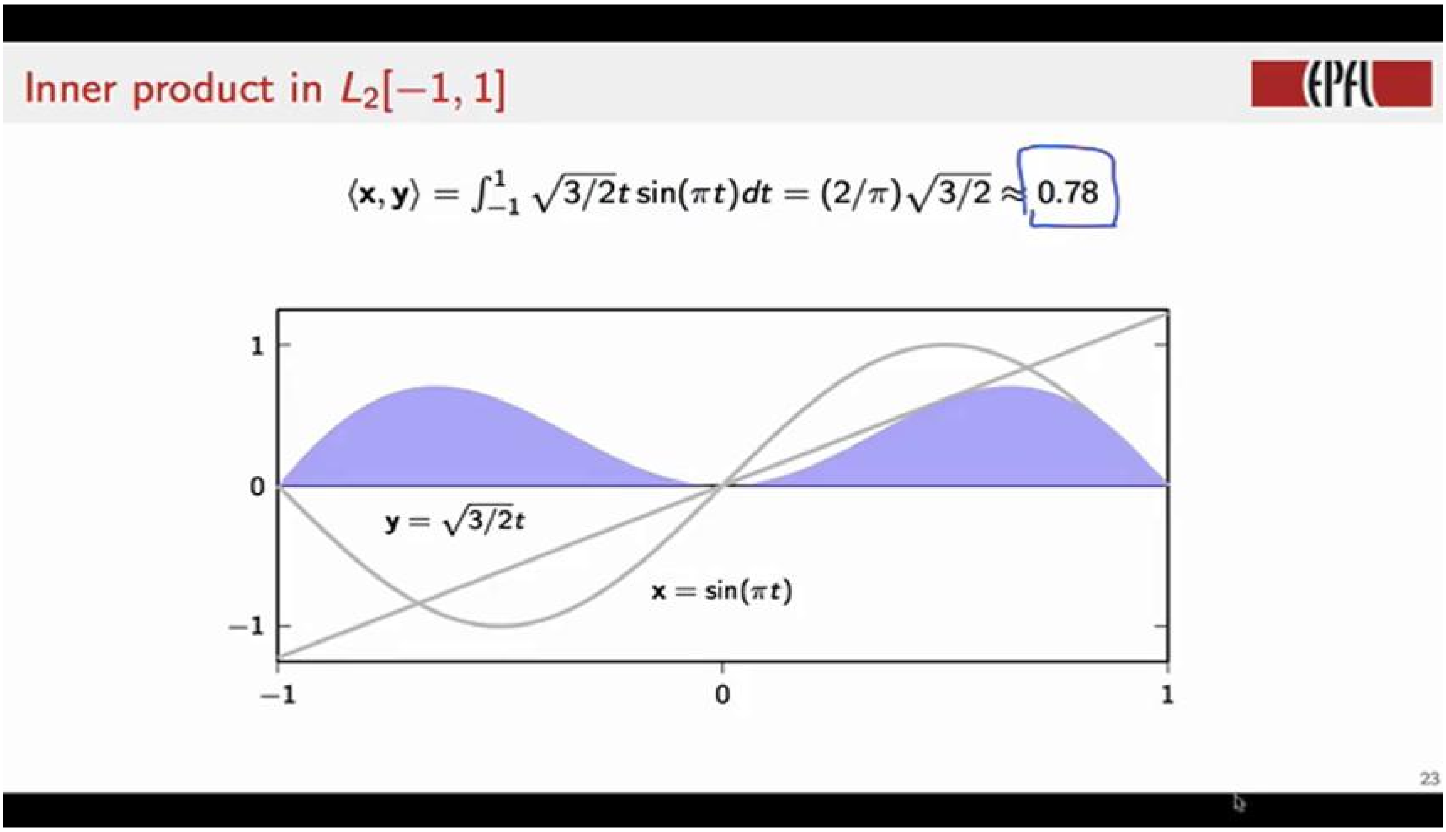


Here we have the first function, here we have the second function. We compute the inner product between the two.

Notes

Summary

11m 57s



We have to compute this area here, the result of the inner product is 0.78. Now remember that we are comparing unit norm vectors so the inner product can range between 1 in the case of maximal similarity and 0 in the case of orthogonality. Here we have a value of 0.78 which indicates that the linear ramp and the sine functions are actually pretty close.

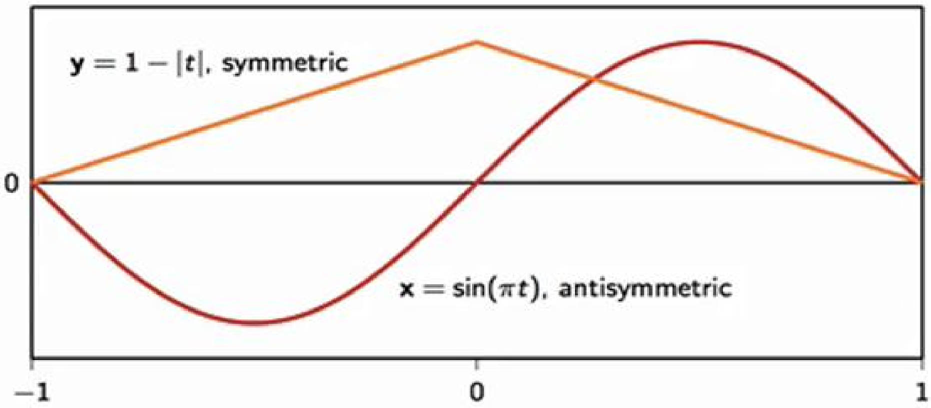
Notes

Summary

12m 03s

Inner product in $L_2[-1,1]$

x, y from orthogonal subspaces:



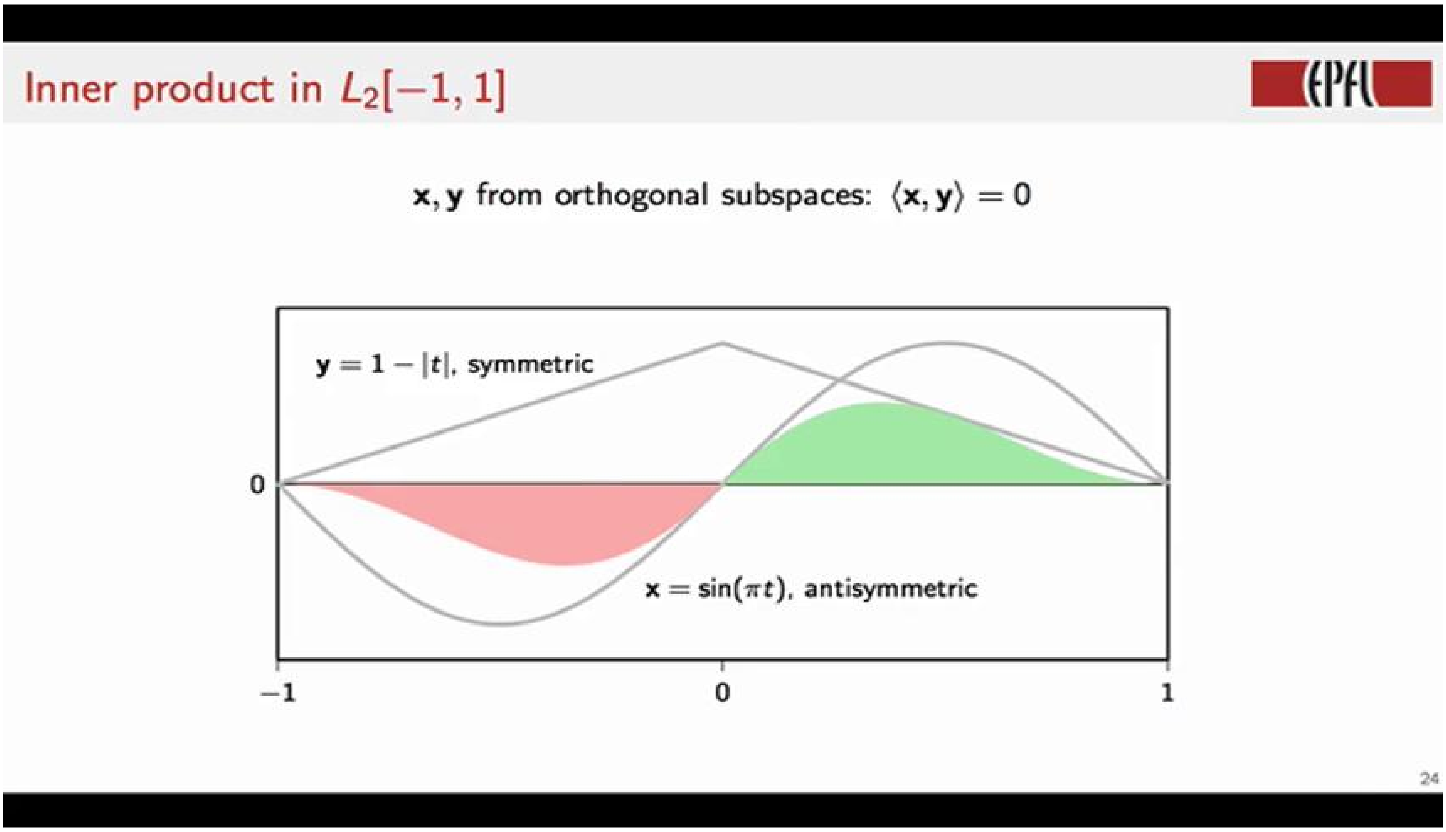
24

We can take the inner product with maximally dissimilar functions if we take our usual x equal to sine of Pi t which is an antisymmetric function. We take the inner product of this function with a symmetric function like for instance a triangle function, 1 minus absolute value of t.

Notes

Summary

12m 28s



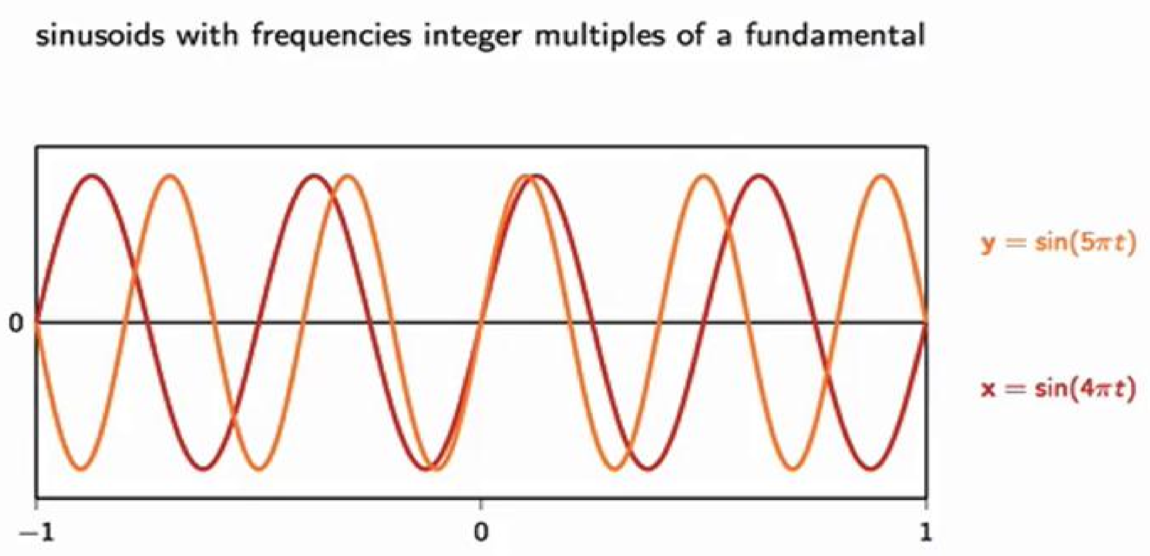
If we take the integral of the product of these two functions, we have to sum the green area with the red area. These two areas have opposite signs and so the integral and therefore the inner product will be equal to 0. The inner product has successfully captured the fact that we will never be able to express any symmetric function in terms of an antisymmetric shape. They are maximally different, they have nothing in common, their inner product is 0.

Notes

Summary

12m 45s

Inner product in $L_2[-1, 1]$



Another famous case of orthogonality between functions is given by sinusoids whose frequencies are harmonically related. We're still in L2 of minus 1, 1. Let's pick as the first function sine of 4 Pi t and as the second function sine of 5 Pi t. These two frequencies are multiples of the fundamental frequency for the interval which is Pi.

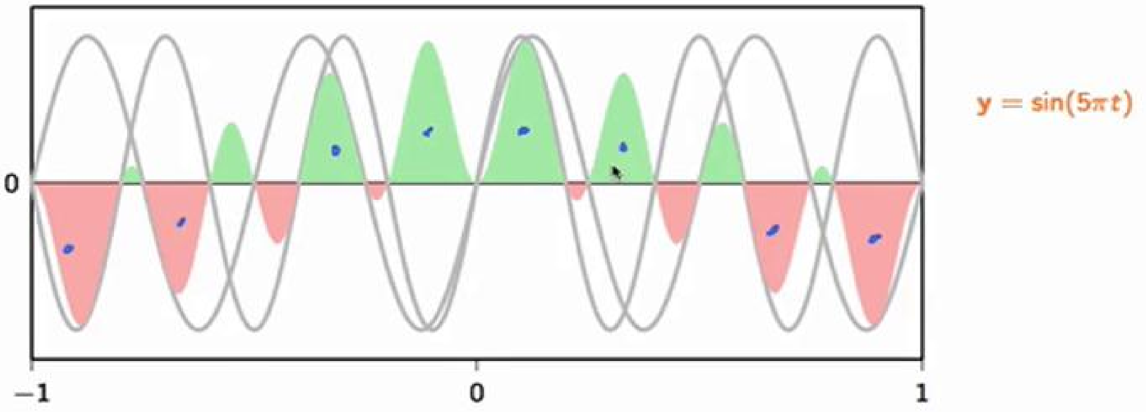
Notes

Summary

13m 13s

Inner product in $L_2[-1, 1]$

sinusoids with frequencies integer multiples of a fundamental



25

If we compute the inner product between the two, even graphically we can see that we have to compute these areas here and sum them together. Now for each red coloured area, you have a corresponding green coloured area that has opposite sign and so as you sum them together, you end up having a total inner product of 0.

Notes

Summary

13m 38s

Norm vs Distance

- inner product defines a norm: $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$
- norm defines a distance: $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$

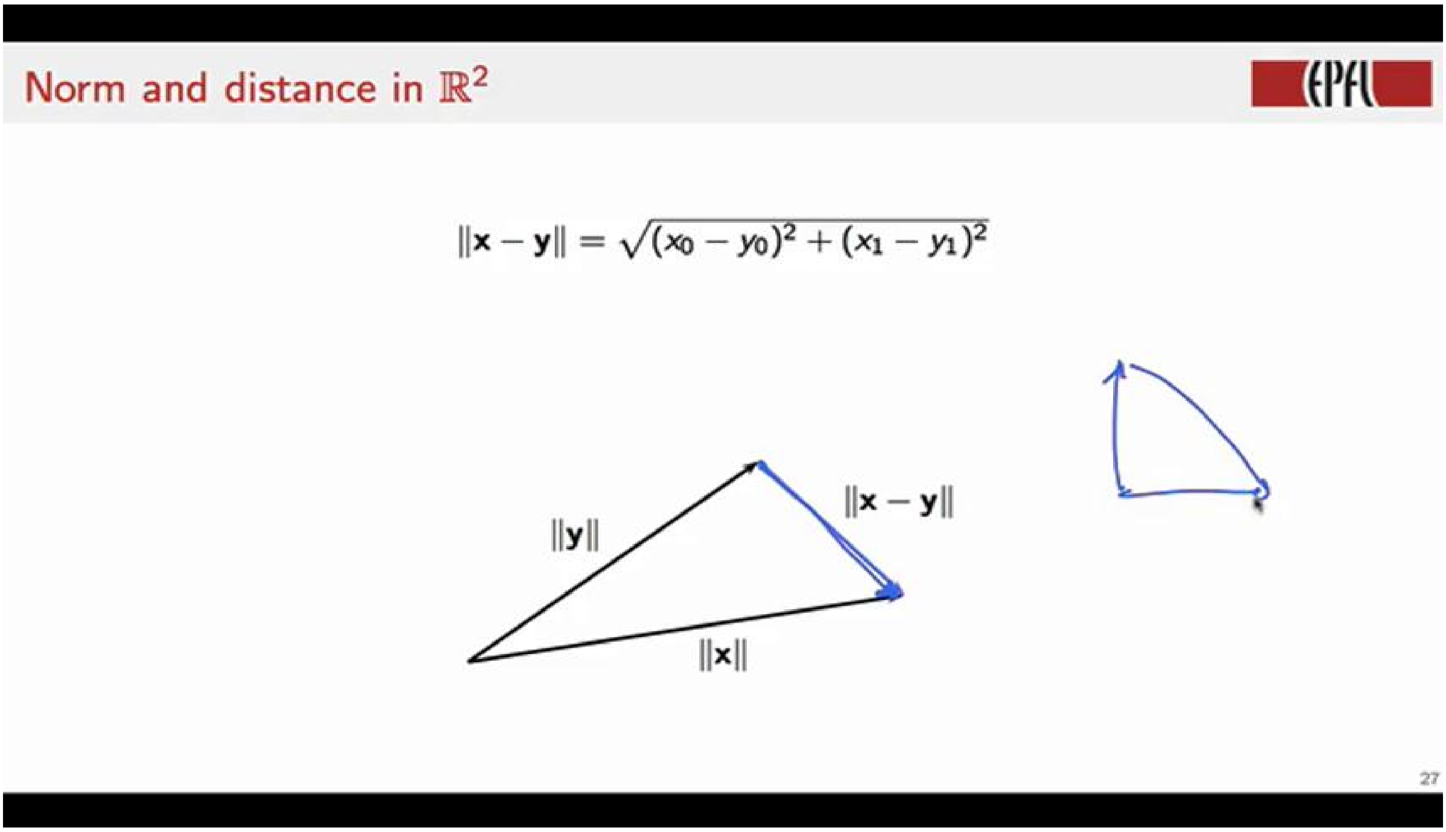
26

In general, if we have a vector space equipped with an inner product, we have a natural norm defined on the space which is the square root of the self inner product. We can use this notion of norm to define a distance between vectors as the norm of the difference between vectors.

Notes

Summary

13m 59s

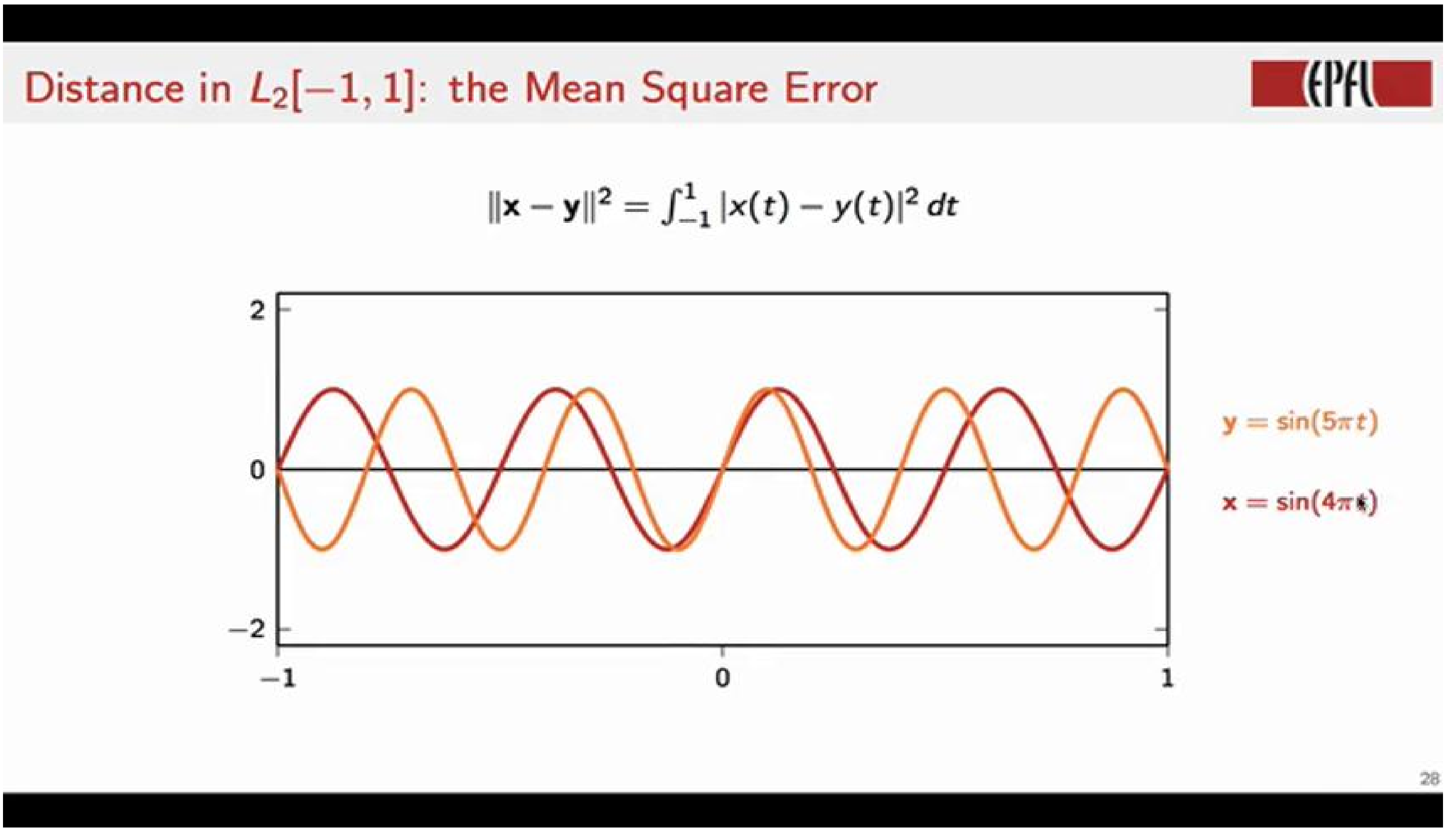


Again, in R2, this is very easy to visualise, take one vector, take the other vector. Now the difference between two vectors is the vector that connects their endpoints and the norm of this connecting vector is a quantitative measure of distance between the original vectors. Note that the distance is not like orthogonality, as a matter of fact orthogonal vectors can have a very large distance, the distance is certainly not 0.

Notes

Summary

14m 16s

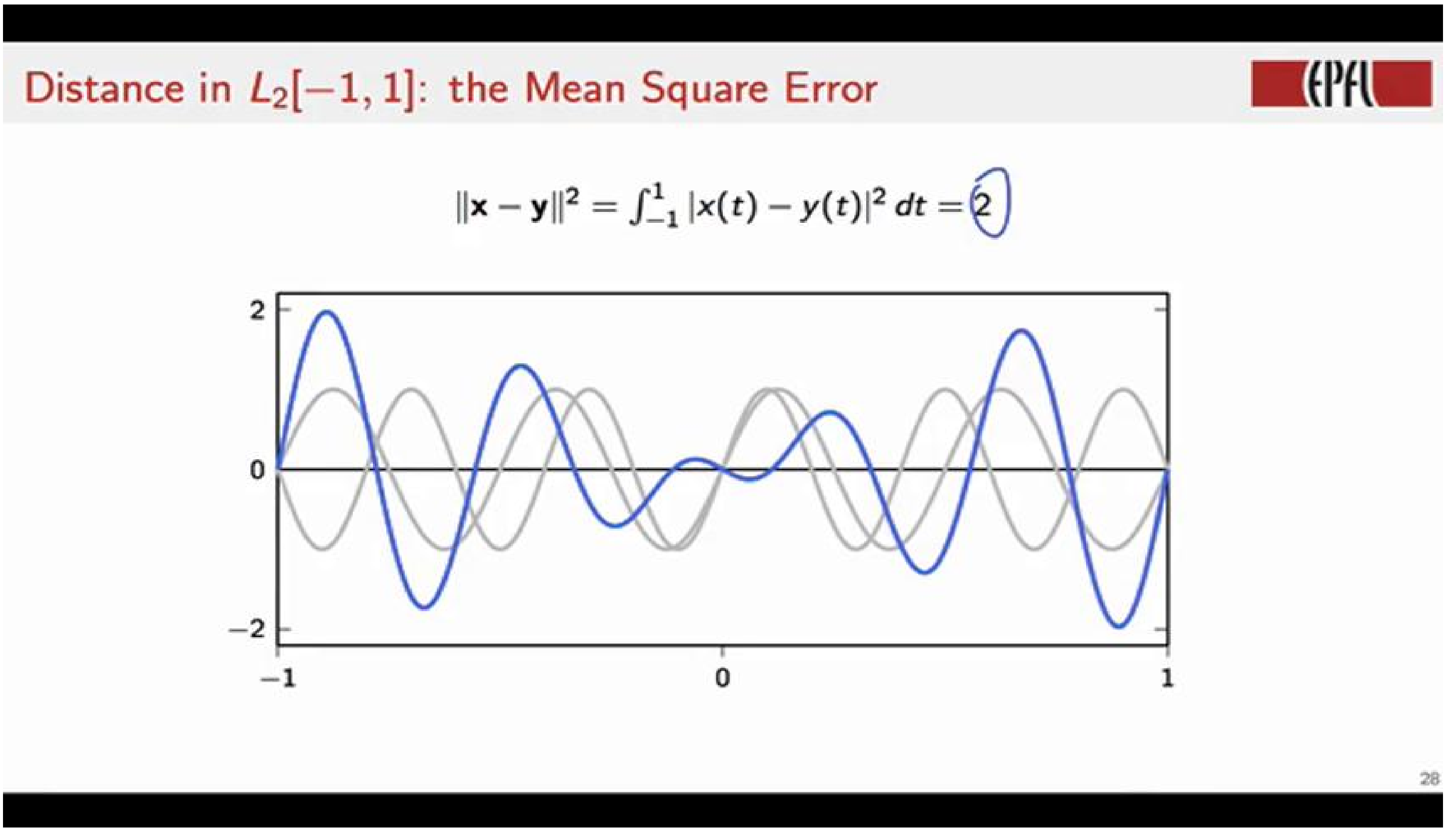


In function vector spaces, the norm of the difference between vectors is usually known as the mean square error. It is indeed the integral, for instance in the case of L2 of minus 1, 1 of the difference squared between two functions.

Notes

Summary

14m 40s



Here we have as an example a first vector sine of Pi 4 t, the second function is sine of 5 Pi t, their difference will be this weird function here and if we compute the area of this function squared, we have a value of 2 which is the distance between these two harmonically related functions.

Notes

Summary

14m 56s