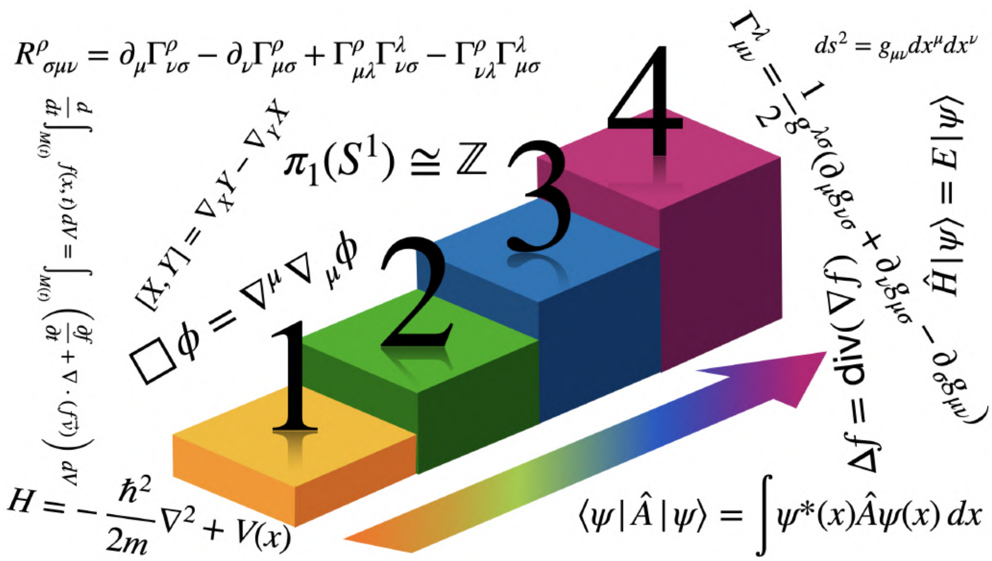
The 4-Step Process to Learn Anything in Mathematics

by DiBeos

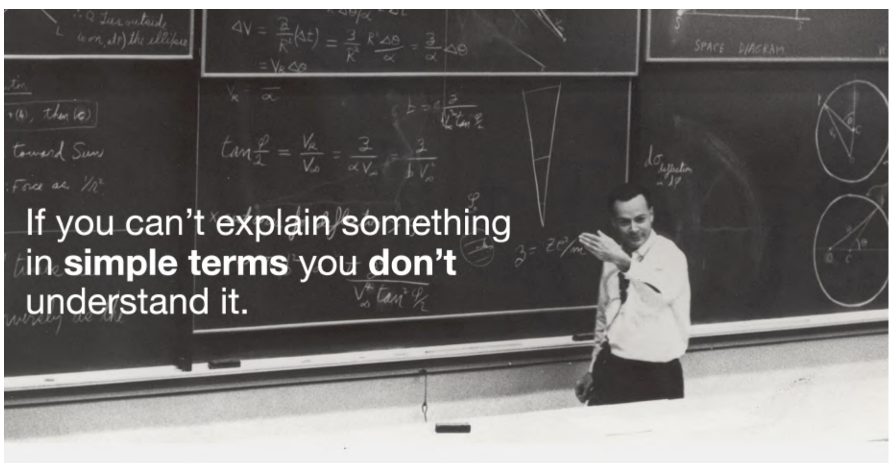


If you follow the 4-step process that we will share here, you can literally learn anything in mathematics in a much easier and faster way than traditional methods used by most people – like reading from a super abstract book with bad explanations, for example. If you don't believe me, just try it out right after reading this file and you will see it by yourself. Let's get to it!

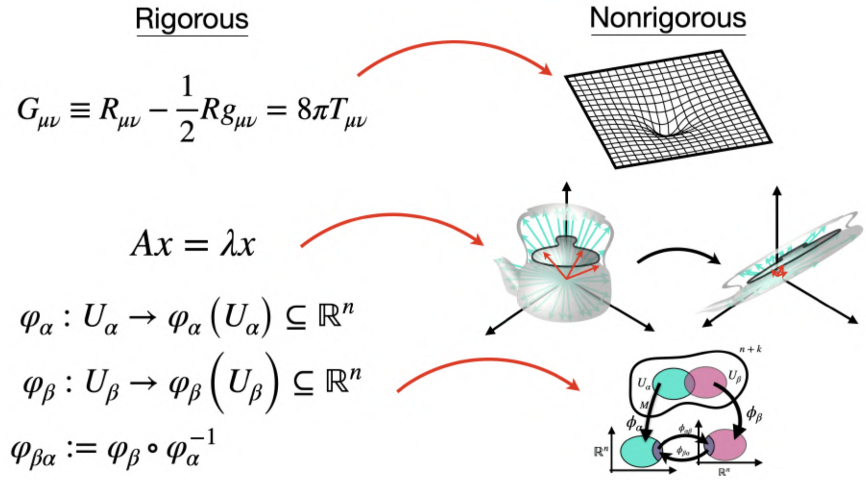
Intuition

Richard Feynman believed that "If you can't explain something in simple terms, you don't understand it".

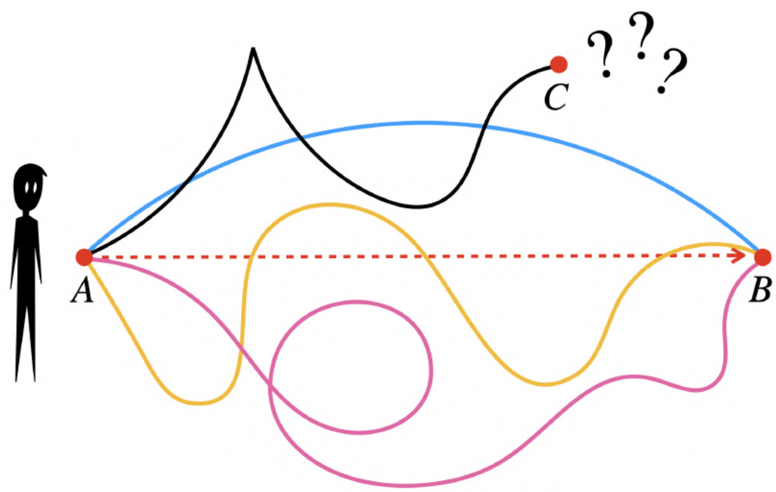
Intuition



Before jumping into theorems, corollaries, definitions, lemmas, you should start with a very intuitive (and even imprecise!) picture of what the concept, or subject you are trying to learn, is all about. This easy mental framework will be very important later on, especially when you study abstract definitions because it is a quick way to remind yourself why (so, the motivation) this particular concept matters to you and to keep track of the final goal, i.e. what you want to achieve with it.



By the way, this is a very recurring pattern throughout all mathematics: you want to go from a point A to a point B, and there might be many different paths, sometimes completely non-intuitive and abstract, so what happens is that when you are trying to prove something, or calculate a result, you end up forgetting in the middle of the road what your initial goal was (i.e., to get to point B).



A quick mental picture that summarizes your goal, like an imperfect analogy, does help a lot along the way! Beyond that, this initial contact with the concept or subject will dictate your subconscious motivation throughout the whole process. If you get it right, you will be motivated, because it will be more and more interesting as you study it, since you will experience the feeling of getting closer and closer to your desired destination.

Let's see an illustration of it: say you want to understand what "*Self-Adjoint Operators on Hilbert Spaces*" are.

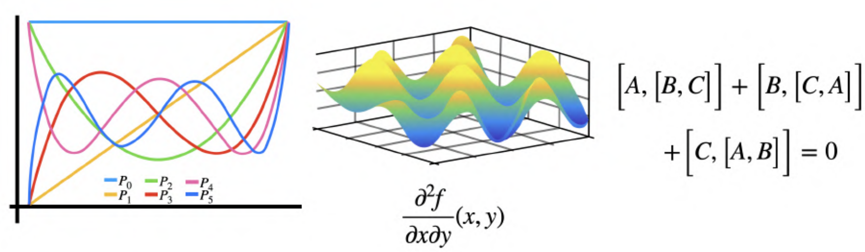
Self-Adjoint Operators on Hilbert Spaces

$$(\mathscr{H}, \|\cdot\|)$$

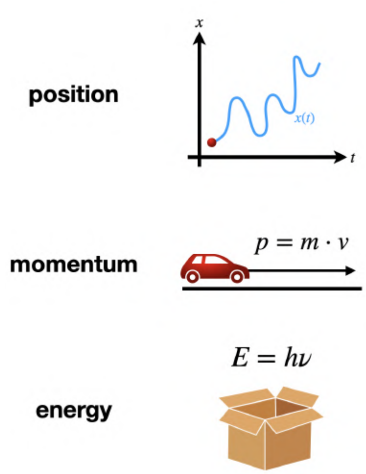
$$\langle T\psi, \phi \rangle = \langle \psi, T\phi \rangle$$

$$\forall \psi, \phi \in \text{Dom}(T)$$

I mean, it makes sense that you want to study them, they are the backbone of functional analysis, PDEs, and operator algebras. They are super important in understanding the structure of Hilbert spaces. In applied mathematics they are even more important!



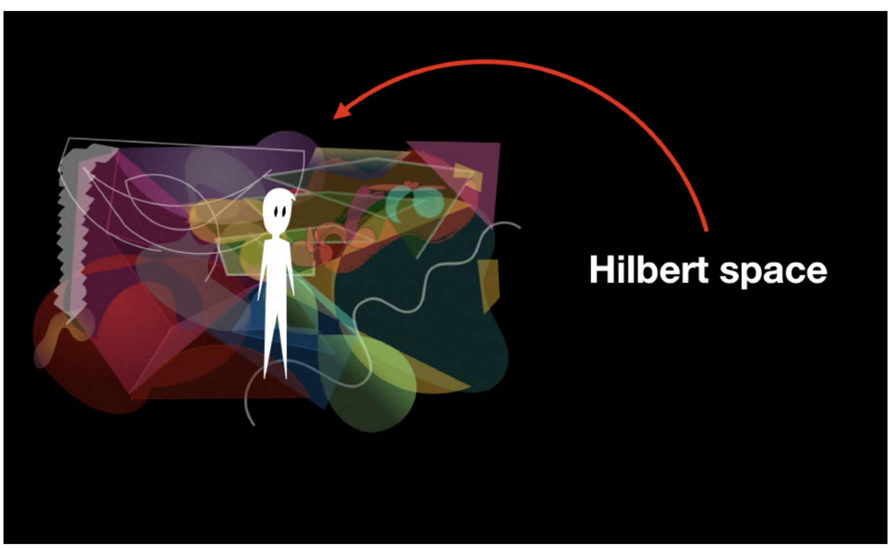
In QM, for example, every observable – so, position, momentum, energy – is modeled by a self-adjoint operator because only these give us real eigenvalues.



$$T = T^\dagger \implies \lambda \in \mathbb{R}$$

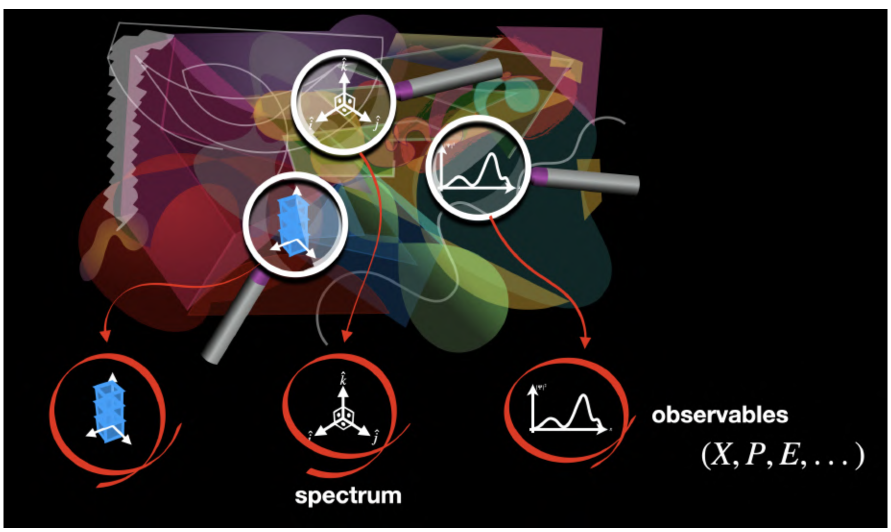
Did you see what was just done?! I gave you motivation to study it. Now you probably have a strong conviction that this is an important concept to learn, and worth your effort to keep paying attention to this file. But what about the intuitive picture?

Imagine you're floating in a vast infinite-dimensional space. There are so many things going on that your chimp brain can't really recognize any patterns and therefore everything just looks like a huge mess to you. It seems that there is no well-defined structure. Absolutely, nothing interesting. This represents a Hilbert Space.

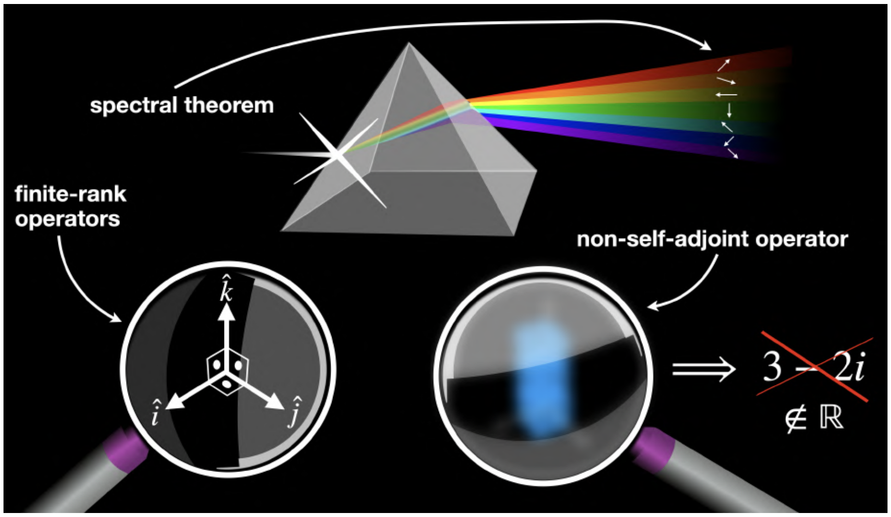


Now, wouldn't it be nice if you had a special lens that filtered out the noise, and depending on the perspective with which you observed this bizarre space, you could see different well-defined structures?

This lens is the self-adjoint operator, and what you see is “the spectrum”. Going back to quantum mechanics (just to help us to grasp the intuition, and then go back to pure math), we call this structure “observables”, because these are the things we can measure in our limited physical world. In other words: A self-adjoint operator is a mathematical object that reveals what’s observable about a Hilbert space.



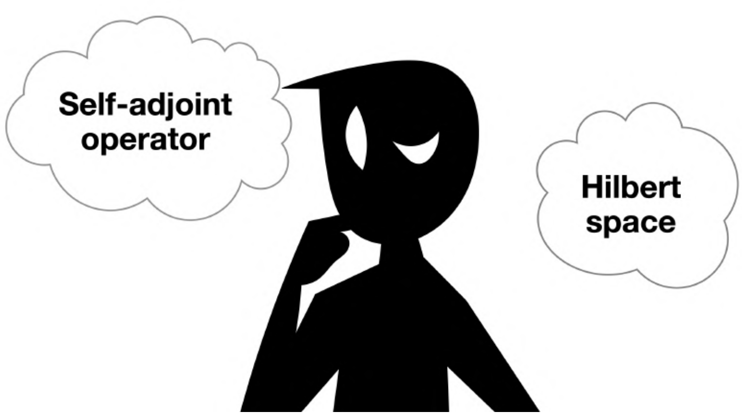
Some lenses are very simple: they only show a few axes. Others act like prisms: they break the whole space into a continuous rainbow of components you can analyze separately. Some lenses are blurry: they distort or return complex measurements, which is not acceptable if you’re trying to extract something “real”.



Great! I'd say that after this strong intuitive picture you are ready to move to the next step:

Concrete Examples

That's not the time to perform complex calculations on your own, yet. This is the moment that you connect your beautiful and intuitive (but imprecise) mental picture, with the mathematics we are trying to model. This is best done by concrete examples. Basically, it's time to get back to Earth! Let's continue with our study of self-adjoint operators so that you will better understand what we mean:



1. The first concrete example is a self-adjoint operator called A on the Hilbert space $\mathbb{R}^2$ with the standard inner product. Notice that we still didn't even define rigorously what a Hilbert space or self-adjoint operators are – it doesn't matter yet, since all we want at this point is to go back from abstract/intuitive land to a more concrete one.

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

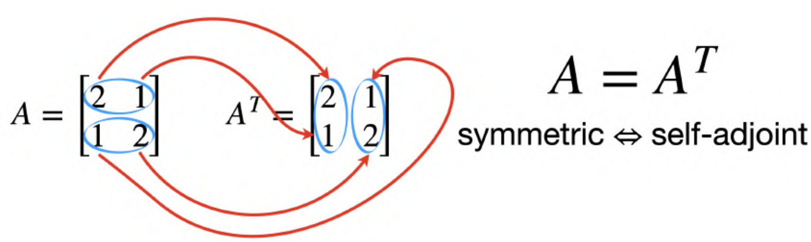
This Hilbert Space is finite-dimensional. Nice, this example won't be that hard.

Let's calculate the transpose matrix (so, rows become columns and vice-versa). As you can see, it is still equal to the original matrix.

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = A^T$$

$$A = A^T$$

So, this is a symmetric matrix, and in finite-dimensional real Hilbert spaces, this is the precise definition of self-adjoint.



But what kind of structure can we see using this lens? Since elements of this particular Hilbert space are vectors of the form $\begin{bmatrix} x \\ y \end{bmatrix}$, then we can calculate its eigenvalues and eigenvectors and find the results:

$$\lambda_1 = 3 \implies v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\lambda_2 = 1 \implies v_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Let's see how we calculated these results. But first, if you want to learn more about eigenvectors and eigenvalues check out this video and PDF link:

The Core of Eigenvalues & Eigenvectors
PDF link: Eigenvalues & Eigenvectors

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\boxed{\boxed{\det(A - \lambda I) = 0}} \implies \begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = 0 \implies (2-\lambda)^2 - 1 = 0 \implies$$

$$\implies \boxed{\lambda^2 - 4\lambda + 3 = 0}$$

$$\lambda_{1,2} = \frac{4 \pm \sqrt{16-12}}{2} = \frac{4 \pm 2}{2} = 3 \vee 1$$

$\underline{\lambda_1 = 3}$:

$$\boxed{A\,\vec{v}_1 = \lambda_1\,\vec{v}_1} \implies \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} 3\,v_x \\ 3\,v_y \end{bmatrix} \implies \begin{cases} 2\,v_x + v_y = 3\,v_x \\ v_x + 2\,v_y = 3\,v_y \end{cases} \implies v_y = v_x \implies$$

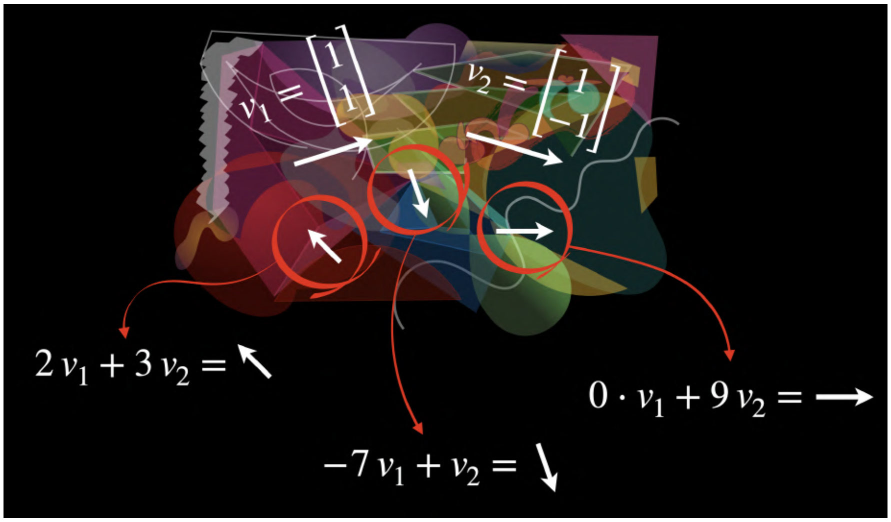
$$\implies \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$\underline{\lambda_2 = 1}$:

$$\boxed{A\,\vec{v}_2 = \lambda_2\,\vec{v}_2} \implies \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} v_x \\ v_y \end{bmatrix} \implies \begin{cases} 2\,v_x + v_y = v_x \\ v_x + 2\,v_y = v_y \end{cases} \implies v_y = -v_x \implies$$

$$\implies \vec{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Notice how the fact that the operator is self-adjoint results in real eigenvalues and the eigenvectors are orthogonal – which is awesome, because it means they qualify for a basis of this Hilbert space. You can basically decompose any vector in this Hilbert space as a linear combination of these two eigenvectors, which is a super powerful property!



Beyond that, since the operator A is self-adjoint, the Spectral Theorem says that:
"Self-adjointness $\implies$ diagonalizability.". In other words, in the basis $\{\vec{v}_1\,,\,\vec{v}_2\}$, the matrix becomes:

$$Q^T A\,Q = \begin{bmatrix}3 & 0\\0 & 1\end{bmatrix}$$

Let's see where this diagonal matrix came from:

Fist, we *normalize* the eigenvectors $\vec{v}_1$ and $\vec{v}_2$ that we found. This means that we can compute other vectors, called $\hat{v}_1$ and $\hat{v}_2$, that point in the same direction to $\vec{v}_1$ and $\vec{v}_2$, but are *unit vectors* – they magnitude is 1.

$$\hat{v}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}$$

$$\hat{v}_2 = \frac{\vec{v}_2}{\|\vec{v}_2\|} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}$$

This gives us a *orthonormal basis* $\{\hat{v}_1\,,\,\hat{v}_2\}$ for the Hilbert space in question.

Now, we build the orthogonal matrix Q:

$$Q = [\hat{v}_1 \quad \hat{v}_2] = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}$$

Since the columns are orthonormal, $Q^T = Q^{-1}$.

And finally we diagonalize the matrix:

$$Q^T A\,Q = \left(\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\right)\begin{bmatrix}3 & 0\\ 0 & 1\end{bmatrix}\left(\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\right) = \frac{1}{\sqrt{2}}\begin{bmatrix}3 & 3\\ 1 & -1\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} \implies$$

$$\implies Q^T A\,Q = \begin{bmatrix}3 & 0\\ 0 & 1\end{bmatrix}$$

Indeed, 3 and 1 are the eigenvalues we found before!

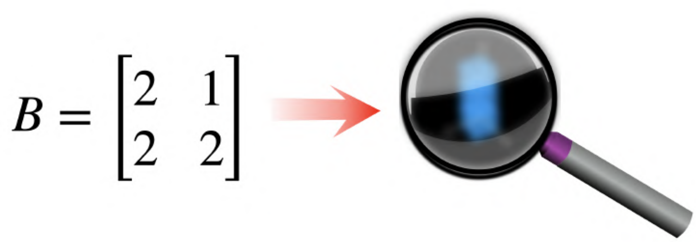
$$Q^T A\,Q = \begin{bmatrix}3 & 0\\ 0 & 1\end{bmatrix}$$

Let's see another concrete example now:

2. The non-self-adjoint operator B, shown below, on the same Hilbert space $\mathbb{R}^2$ with the standard inner product.

$$B = \begin{bmatrix}2 & 1\\ 2 & 2\end{bmatrix}$$

This is like the blurry lens in our analogy. Not very useful...



Notice a few things:

(a) $B^T \neq B \implies$ B is not symmetric $\implies$ B is not self-adjoint;

(b) Its eigenvalues are real (which is a good thing), but its eigenvectors are not orthogonal!

$$\boxed{\boxed{\det(B - \lambda I) = 0}} \implies \begin{vmatrix} 2-\lambda & 1 \\ 2 & 2-\lambda \end{vmatrix} = 0 \implies (2-\lambda)^2 - 2 = 0 \implies$$

$$\implies \boxed{\lambda^2 - 4\lambda + 2 = 0}$$

$$\lambda_{1,2} = 2 \pm \sqrt{2} \quad (\in \mathbb{R})$$

$\underline{\lambda_1 = 2 + \sqrt{2}}$:

$$\boxed{B\,\vec{v}_1 = \lambda_1\,\vec{v}_1} \implies \begin{bmatrix} 2 & 1 \\ 2 & 2 \end{bmatrix} \cdot \begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} \left(2+\sqrt{2}\right)\, v_x \\ \left(2+\sqrt{2}\right)\, v_y \end{bmatrix} \implies \begin{cases} 2\,v_x + v_y = \left(2+\sqrt{2}\right)\, v_x \\ 2\,v_x + 2\,v_y = \left(2+\sqrt{2}\right)\, v_y \end{cases} \implies$$

$$\implies \vec{v}_1 = \begin{bmatrix} 1 \\ \sqrt{2} \end{bmatrix}$$

$\underline{\lambda_2 = 2 - \sqrt{2}}$:

$$\boxed{B\,\vec{v}_2 = \lambda_2\,\vec{v}_2} \implies \begin{bmatrix} 2 & 1 \\ 2 & 2 \end{bmatrix} \cdot \begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} \left(2-\sqrt{2}\right)\, v_x \\ \left(2-\sqrt{2}\right)\, v_y \end{bmatrix} \implies \begin{cases} 2\,v_x + v_y = \left(2-\sqrt{2}\right)\, v_x \\ 2\,v_x + 2\,v_y = \left(2-\sqrt{2}\right)\, v_y \end{cases} \implies$$

$$\implies \vec{v}_2 = \begin{bmatrix} 1 \\ -\sqrt{2} \end{bmatrix}$$

Are they orthogonal?

$$\langle \vec{v_1}\,,\,\vec{v_2} \rangle = \begin{bmatrix} 1 \\ \sqrt{2} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -\sqrt{2} \end{bmatrix} = 1 \cdot 1 + \sqrt{2} \cdot (-\sqrt{2}) = 1 - 2 = -1 \neq 0$$

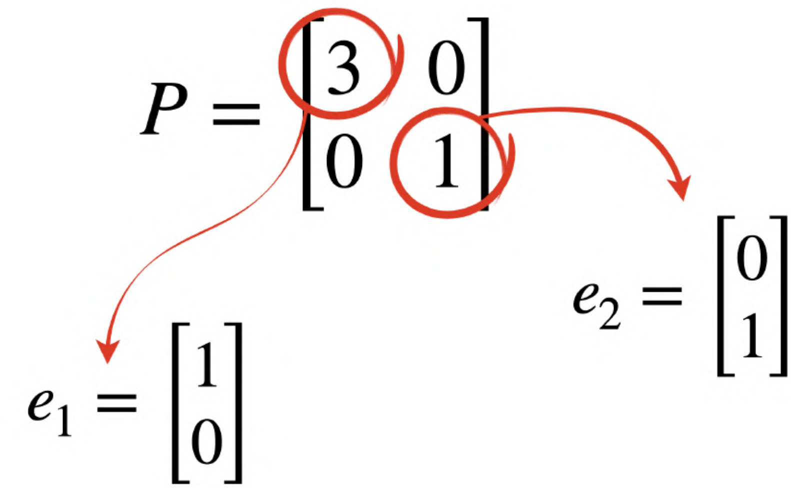
After calculating their dot product we get -1, which is not zero... That's not good! They do not form a basis of the Hilbert space in question.

3. The third example is the "prism" in our analogy.

$$P = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$$

This is just the diagonalized version of the operator A in the first example.

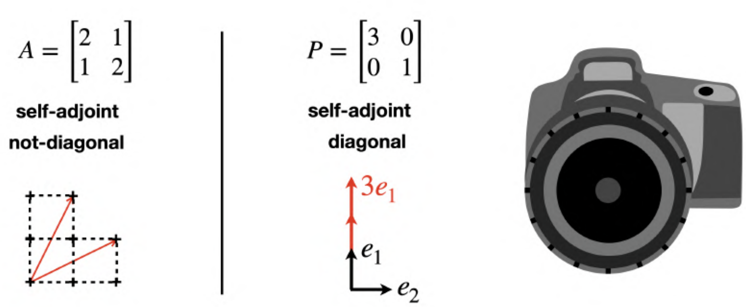
Its eigenvalues are clearly 3 and 1, and its eigenvectors are the elements $\vec{e}_1$ and $\vec{e}_2$ of the standard basis of this Hilbert space: $\left\{\vec{e}_1 = \begin{bmatrix}1\\0\end{bmatrix}, \vec{e}_2 = \begin{bmatrix}0\\1\end{bmatrix}\right\}$.



Every vector $x \in \mathbb{R}^2$ can be written as a liner combination of the basis elements:

$$x = x_1\,\vec{e}_1 + x_2\,\vec{e}_2 \implies Px = 3\,x_1\,\vec{e}_1 + 1\,x_2\,\vec{e}_2$$

So, we say that the action of P is *pure spectral*, i.e. it just stretches in the direction of $\vec{e}_1$ by 3, and in the direction of $\vec{e}_2$ by 1 – no change in the latter.



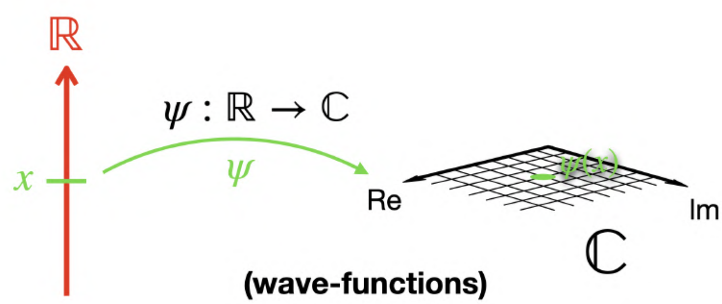
Comparing it to the original operator, we find out that even though that matrix A is self-adjoint, it is not not diagonal in the standard basis. And as a consequence, we had to rotate the eigenbasis in order to “see” the decomposition. It was a lens that needed adjustment to get the full spectral view. However, with P, the decomposition is already aligned with the space. So, no rotation is needed.

4. Now, for the last concrete example, we’ll look at a self-adjoint operator in a Hilbert space where the elements aren’t arrows in space, but rather a special kind of function. These are the wave functions used in quantum mechanics.

This operator is called the “position operator”. And the Hilbert space we are studying here is the space $L^2(\mathbb{R})$.

$$L^2\left(\mathbb{R}^2\right) = \left\{\psi : \mathbb{R} \to \mathbb{C} \ : \ \int_{-\infty}^{\infty} |\psi(x)|^2 \, dx < -\infty\right\}$$

The elements in this space are functions, denoted as ψ, that depend on the position $x \in \mathbb{R}$ related to where the measurement is being performed. These are usually called wave-functions in Quantum Mechanics, and they have the property that if you compute their modulus squared and then integrate along their whole domain (i.e., all real numbers) the result is a finite number.



$$\int_{\mathbb{R}} \left|\psi(x)\right|^2 dx < \infty$$

This requirement is important because in Quantum Mechanics the result of this integration is interpreted as the probability of finding the particle, or event that you are measuring, in each specific region along the x-axis. If this integral diverges to infinity, then it is impossible to normalize this function, i.e., to multiply it by a constant such that the probability of all possible outcomes add up to 100%.

$$\Psi = A \cdot \psi$$

The position operator X will act on a wave-function $\psi(x)$. That’s how we represent this action:

$$(X\psi)(x)$$

The result, so what the operator X effectively does to the function ψ, is to multiply it by the position x. That's all. Pretty simple.

$$(X\psi)(x) = x\psi(x)$$

Now, why is it considered a self-adjoint operator? And what is this specific lens revealing about the Hilbert space? Well, let's see.

The operator is still said to be symmetric, just like when dealing with matrices, but in this context it is represented in the following way:

$$\int \left(X\,\psi_1\right)(x) \cdot \psi_2(x)\,dx = \int \psi_1(x) \cdot \left(X\,\psi_2\right)(x)\,dx$$

, for all ψ_1 and ψ_2 in the domain of X.

Notice the different positioning of the X operator in the LHS (left-hand side) with respect to the RHS (right-hand side). That's why we say it is symmetric.

But being symmetric is not enough for being self-adjoint here. We also need that the domain of the operator X is the same as the domain of the operator X^T:

$$\mathrm{Dom}\,(X) = \mathrm{Dom}\left(X^T\right)$$

Actually, this equation is not correct when written this way... The superscript T (of transpose) is meaningful only when dealing with matrices. But when we are working with operators on Hilbert spaces in general, possibly infinite-dimensional, we use the this notation to indicate symmetry:

$$X = X^\dagger$$

This symbol here is called a *dagger*. In linear algebra, $X^\dagger$ is referred to as the *conjugate transpose* of a matrix X. In our case though (and in Quantum Mechanics), $X^\dagger$ is called the *Hermitian adjoint* of the operator X.

So, the two requirements for being self-adjoint, in this context, are:

$$X = X^\dagger \quad \text{and} \quad \text{Dom}(X) = \text{Dom}(X^\dagger)$$

We don't need to prove (not yet) that the position operator X is self-adjoint. But we can confirm that it does satisfy the first condition by picking two concrete examples of function ψ_1 and ψ_2 in this Hilbert space.

Let's check the first condition: $\vdash X = X^\dagger$

Our *test functions* are:

$$\psi_1(x) = e^{-x^2}$$
$$\psi_2(x) = x\, e^{-x^2}$$

Let's compute the LHS:

$$\int_{-\infty}^{\infty} (X\,\psi_1)\,(x) \cdot \psi_2(x)\, dx = \int_{-\infty}^{\infty} x\, e^{-x^2} \cdot x\, e^{-x^2}\, dx = \int_{-\infty}^{\infty} x^2\, e^{-2x^2}\, dx$$

Now, the RHS:

$$\int_{-\infty}^{\infty} \psi_1(x) \cdot (X\,\psi_2)\,(x)\, dx = \int_{-\infty}^{\infty} e^{-x^2} \cdot x^2\, e^{-x^2}\, dx = \int_{-\infty}^{\infty} x^2\, e^{-2x^2}\, dx$$

As you can see, both sides are identical. Let's calculate this integral just to complete the example.

Consider, first, the more general case:

$$\boxed{I = \int_{-\infty}^{\infty} e^{-ax^2}\, dx} \quad , \; a > 0$$

Now, square both sides:

$$I^2 = \left(\int_{-\infty}^{\infty} e^{-ax^2}\, dx\right)^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-a\,(x^2+y^2)}\, dx\, dy$$

At this point we can switch to a different coordinate system (i.e., polar coordinates):

$$\begin{cases} x = r\,\cos\theta \\ y = r\,\sin\theta \end{cases} \qquad \Longrightarrow \qquad dx\,dy = r\,dr d\theta$$

Then:

$$I^2 = \int_0^{2\pi}\int_0^{\infty} e^{-a\,r^2} r\,dr d\theta = \text{(Fubini's theorem)} = \left(\int_0^{2\pi} d\theta\right)\left(\int_0^{\infty} r\,e^{-a\,r^2}\,dr\right) =$$
$$= 2\pi\cdot\left(\int_0^{\infty} r\,e^{-a\,r^2}\,dr\right)$$

We need another substitution of variables here. This time though: $t := a\,r^2 \implies dt = 2a\,r\,dr \implies r\,dr = \frac{dt}{2a}$.

$$I^2 = 2\pi\cdot\int_0^{\infty} e^{-t}\,\frac{dt}{2a} = \frac{\cancel{2}\pi}{\cancel{2}a}\int_0^{\infty} e^{-t}\,dt = \frac{\pi}{a} \implies$$
$$\implies \boxed{I = \sqrt{\frac{\pi}{a}}}$$

In our case, $a \equiv 2$, and the integral is slightly different:

$$\boxed{\int_{-\infty}^{\infty} x^2\,e^{-2x^2}\,dx}$$

Let's integrate it by parts. this is the general formula of integration by parts:

$$\int u\,dv = uv - \int v\,du$$

In our specific case, we have:

$$\begin{cases} u = x \implies du = dx & (1) \\ dv = x\,e^{-2x^2}\,dx \implies v = \int x\,e^{-2x^2}\,dx & (2) \end{cases}$$

Before moving on, let's calculate a simplified version of equation (2) by performing the following substitution of variables:

$$\boxed{-2x^2 =: t} \implies dt = -4x\,dx \implies \boxed{\frac{-dt}{4} = x\,dx}$$

$$v = \int x\,e^{-2x^2}\,dx = \int \frac{-dt}{4} \cdot e^t = -\frac{1}{4}\int e^t\,dt = -\frac{1}{4}\,e^t = -\frac{1}{4}\,e^{-2x^2}$$

Now we can proceed with applying the integration by parts formula:

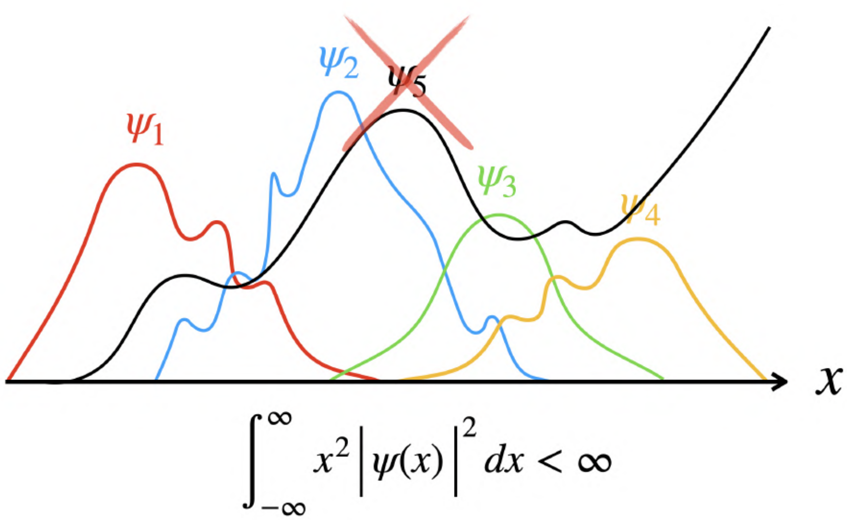
$$\int_{-\infty}^{\infty} x^2\,e^{-2x^2}\,dx = \left[x \cdot \left(-\frac{1}{4}\,e^{-2x^2}\right)\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty}\left(-\frac{1}{4}\,e^{-2x^2}\right)dx =$$
$$= -\frac{1}{4}\left(\frac{x}{e^{2x^2}}\right)_{-\infty}^{\infty} 0 + \frac{1}{4}\int_{-\infty}^{\infty} e^{-2x^2}\,dx =$$

Notice how (in the first term) the function e^{-2x^2} in the denominator is the *dominant term* – it grows way faster than the function x in the denominator – ans that's why the whole thing tends to zero in both extremes of the real line.

$$= \frac{1}{4}\int_{-\infty}^{\infty} e^{-2x^2}\,dx = \frac{1}{4}\sqrt{\frac{\pi}{2}} = \frac{\sqrt{\pi}}{4\sqrt{2}}$$

Great! And what is this lens revealing about the space $L^2(\mathbb{R})$?

The position operator shows that you can organize your wavefunctions according to position. It shows that this space has a coordinate-wise interpretation. So, you can ask how much of the wavefunction lives near each position.



Beyond that this operator filters out wavefunctions that grow too fast to infinity, and doing so allows us to focus our attention only on functions that are "physically well-behaved".

$$\int_{-\infty}^{\infty} x^2 \, |\psi(x)| \, dx < \infty$$

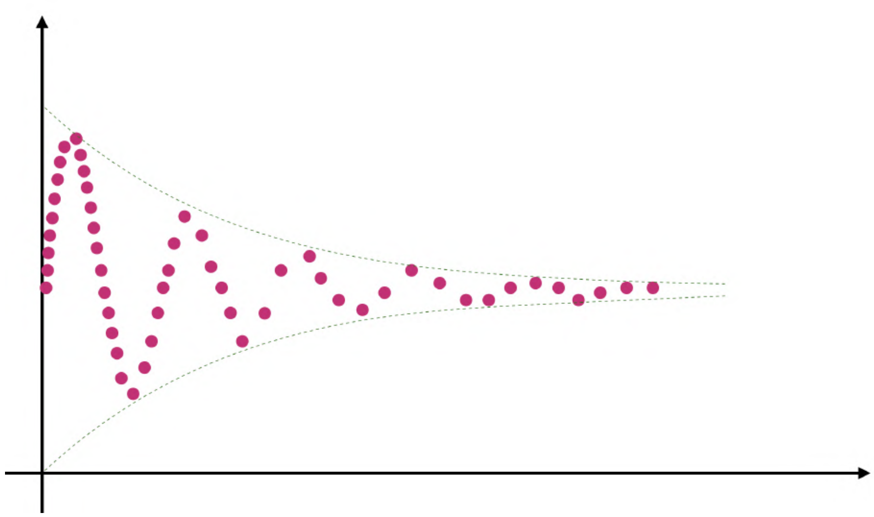
Awesome! I'd say that at this point you are more than ready to move on to the third step in this learning method:

Rigor

Definition 1: A *Hilbert space* is a complete inner product space. That is:

∗ A vector space $\mathcal{H}$ over $\mathbb{R}$ or $\mathbb{C}$.
∗ Equipped with an inner product $\langle \cdot \, , \cdot \rangle$.
∗ Such that the norm $\|x\| = \sqrt{\langle x \, , x \rangle}$.
∗ And it's *complete* with respect to this norm, i.e. every *Cauchy sequence* converges in $\mathcal{H}$.

Cauchy sequences can be thought of (intuition) as list of ordered numbers, functions, or any other mathematical object such that each term get arbitrarily close to each other, even if you don't yet know what they're approaching.



The crucial point about Cauchy sequences, as we mentioned, is that we don't always know where they converge, or even if they converge within the original space. They may converge to a set or space outside the original one. And in our case (Hilbert spaces) we want every Cauchy sequence of terms in $\mathcal{H}$ to converge to another element of $\mathcal{H}$ itself. This is what makes the space to be *complete*.

Let's see an illustration of it (concrete example):

$$\left(\mathbb{C}^{2\times 2}\ ;\ \langle \cdot\,,\cdot\rangle_{\text{Frobenius}}\right)$$

This is the Hilbert space of matrices $\mathbb{C}^{2\times 2}$ with the *Frobenius inner product*. Ok, but what does it mean?

The elements of this space are 2×2 matrices with complex entries – some of them would be:

$$\begin{bmatrix} i & 3 \\ 2-4i & 0 \end{bmatrix} \quad , \quad \begin{bmatrix} 8i & 4i \\ 4-7i & 2i \end{bmatrix} \quad , \quad \begin{bmatrix} 1 & 0 \\ 8 & 2 \end{bmatrix}$$

The Frobenius inner product is defined by the following rule (assuming A and B to be matrices in $\mathbb{C}^{2\times 2}$):

$$\boxed{\langle A\,, B\rangle_{\text{Frobenius}} := \text{Tr}\left(B^\dagger A\right)}$$

$\text{Tr}(\cdot)$ means the trace of the function acted upon – i.e., the sum of its diagonal terms.

Now, imagine we have the following Cauchy sequence:

$$A_k = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{k} \end{bmatrix} \quad , \quad k \in \mathbb{N}$$

This is a Cauchy sequence, under the Frobenius norm, because it converges:

$$\begin{aligned} \|A_k - A_m\|_{\text{Frobenius}} &= \sqrt{\langle (A_k - A_m)\ ,\ (A_k - A_m)\rangle_{\text{Frobenius}}} = \sqrt{\text{Tr}\left((A_k - A_m)^\dagger (A_k - A_m)\right)} = \\ &= \sqrt{\text{Tr}\left(\begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{k} - \frac{1}{m} \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{k} - \frac{1}{m} \end{bmatrix}\right)} = \sqrt{\text{Tr}\left(\begin{bmatrix} 0 & 0 \\ 0 & (\frac{1}{k} - \frac{1}{m})^2 \end{bmatrix}\right)} \\ &= \sqrt{\left(\frac{1}{k} - \frac{1}{m}\right)^2} = \left|\frac{1}{k} - \frac{1}{m}\right| \xrightarrow{k,m\to\infty} 0 \end{aligned}$$

In fact, it converges to:

$$A_k = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{k} \end{bmatrix} \xrightarrow{k\to\infty} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} =: A \in \mathbb{C}^{2\times 2}$$

This is not only true for this Cauchy sequence (we won't prove it here), but for all Cauchy sequences in this space. And that's what characterizes it as a *complete* – and also *Hilbert*, in this case – space.

Definition 2: Let $T : \mathrm{Dom}(T) \subset \mathcal{H} \to \mathcal{H}$ be a linear operator. Then, T is:

∗ *Symmetric* if:

$$\langle T\psi\,, \phi\rangle = \langle \psi\,, T\phi\rangle \;\; , \forall \psi, \phi \in \mathrm{Dom}(T)$$

∗ *Self-adjoint* if:

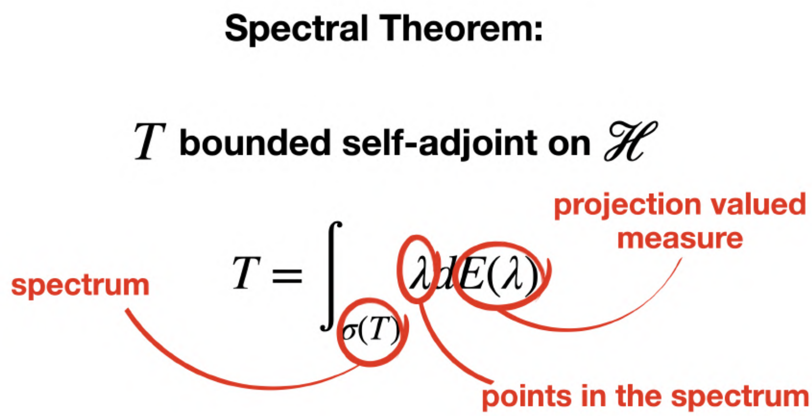
$$T = T^{\dagger} \;\text{ and }\; \mathrm{Dom}(T) = \mathrm{Dom}\left(T^{\dagger}\right)$$

Theorem 1: (Spectral Theorem) Let T be a bounded self-adjoint operator on a Hilbert space $\mathcal{H}$, then:

$$T = \int_{\sigma(\lambda)} \lambda \, dE(\lambda)$$

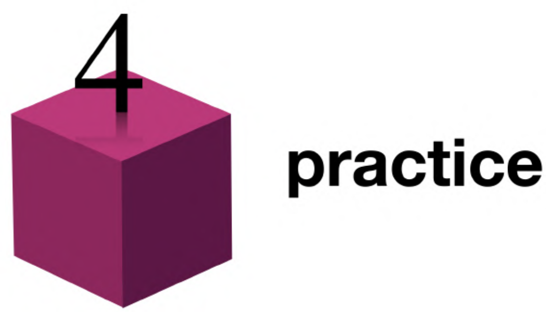
$\sigma(\lambda)$ is the *spectrum* of T. Think of it as the set of all complex numbers λ (points in the spectrum) for which you can't fully undo the operation $T - \lambda I$. Some λ are eigenvalues, others are not – but they all represent places where the operator becomes "non-reversible" in some way.

$E(\lambda)$ is a *projection valued measure*. It is the projection onto the part of the Hilbert space where the operator "acts like" the number λ. We won't get into details here, but it basically "filters out" the component of a vector that corresponds to the spectral value λ.



The Spectral Theorem treats each bounded self-adjoint operator T as a tool to "break apart" the Hilbert space into orthogonal spectral components.

Practice



The last point is to practice as much as possible with a bunch of exercises.

Exercise 1. Let

$$A = \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix} \text{ on } \mathbb{R}^2$$

Tasks:

a) Show that A is self-adjoint, i.e. symmetric.

b) Compute its eigenvalues and eigenvectors.

c) Verify that the eigenvectors are orthogonal.

d) Normalize the eigenvectors and write the orthonormal basis.

e) Express A as $A = PD^{-1}P$, where D is diagonal and P is built from your eigenvectors.

Solutions:

$$A = \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix}$$

a)

$$A^T = \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix} = A$$

b)

$$\boxed{\boxed{\det(A - \lambda I) = 0}} \implies \begin{vmatrix} 4-\lambda & -2 \\ -2 & 1-\lambda \end{vmatrix} = 0 \implies (4-\lambda)(1-\lambda) - 4 = 0 \implies$$

$$\implies \boxed{\lambda^2 - 5\lambda = 0}$$

$$\lambda_{1,2} = 0 \vee 5$$

$\underline{\lambda_1 = 0}$:

$$\boxed{A\,\vec{v}_1 = \lambda_1\,\vec{v}_1} \implies \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix} \cdot \begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies \begin{cases} 4\,v_x - 2v_y = 0 \\ -2\,v_x + v_y = 0 \end{cases} \implies v_y = 2\,v_x \implies$$

$$\implies \vec{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$\underline{\lambda_2 = 5}$:

$$\boxed{A\,\vec{v}_2 = \lambda_2\,\vec{v}_2} \implies \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix} \cdot \begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} 5\,v_x \\ 5\,v_y \end{bmatrix} \implies \begin{cases} 4\,v_x - 2\,v_y = 5\,v_x \\ -2\,v_x + v_y = 5\,v_y \end{cases} \implies v_x = -2\,v_y \implies$$

$$\implies \vec{v}_1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

c)

$$\langle \vec{v}_1\,, \vec{v}_2 \rangle = \langle \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \end{bmatrix} \rangle =$$

$$= 1 \cdot (-2) + 2 \cdot 1 = 0 \implies \vec{v}_1 \perp \vec{v}_2$$

d)

$$\hat{v}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|} = \frac{1}{\sqrt{\langle \vec{v}_1 \, , \vec{v}_1 \rangle}} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\hat{v}_2 = \frac{\vec{v}_2}{\|\vec{v}_2\|} = \frac{1}{\sqrt{\langle \vec{v}_2 \, , \vec{v}_2 \rangle}} \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

$$\therefore \text{Orthonormal basis:} \quad \{\hat{v}_1 \, , \hat{v}_2\}$$

e)

$$P = [\hat{v}_1 \;\; \hat{v}_2] = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}$$

$$P^{-1} = P^T = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \quad \text{(since } P \text{ is orthogonal)}$$

$$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix}$$

$$\therefore A = PDP^{-1} = PDP^T = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$$

Exercise 2. Let $\mathcal{H} = L^2([0,1])$, the space of square-integrable functions on the interval $[0,1]$, with the inner product:

$$\langle f, g \rangle = \int_0^1 f(x)g(x)\,dx$$

Define the operator T by:

$$(Tf)(x) = xf(x) \quad \text{(multiplication operator)}$$

Tasks:

a) Prove that T is symmetric.

b) Show that T has real spectrum.

c) Find a function $f \in L^2([0,1])$ that is *not* an eigenfunction of T, but for which $\langle Tf, f \rangle$ still gives a finite expected value.

Solutions:

a)

$\vdash T = T^\dagger \implies \vdash \langle Tf \, , g \rangle = \langle f \, , Tg \rangle \quad (*) \; , \forall f, g \in \text{Dom}(T).$

$$\langle Tf \, , g \rangle = \int_0^1 (Tf)(x) \cdot g(x)\,dx = \int_0^1 xf(x)g(x)\,dx$$

$$\langle f \, , Tg \rangle = \int_0^1 f(x) \cdot (Tg)(x)\,dx = \int_0^1 f(x)xg(x)\,dx$$

$$\therefore (*)\checkmark \implies \text{T is symmetric.}$$

b) This is a general property of self-adjoint operators: *Every self-adjoint operator on a Hilbert space has real spectrum.* But specifically here, T is a multiplication operator:

$$(Tf)(x) = xf(x)$$

The "spectrum" is the set of all values $x \in [0, 1]$ where the multiplication acts.

So the spectrum of T is the closure of the set of values being multiplied:

$$\sigma(T) = [0, 1]$$

This is a subset of $\mathbb{R}$, and thus real.

c) Try:

$$f(x) = \sqrt{3} \cdot x \quad \text{on } [0, 1]$$

Clearly $f \in L^2([0, 1])$, since:

$$\int_0^1 |f(x)|^2 \, dx = \int_0^1 3x^2 \, dx = 3 \cdot \frac{1}{3} = 1$$

However, it is **not** an eigenfunction of T, because:

$$Tf(x) = xf(x) = \sqrt{3}x^2 \not\propto f(x)$$

If you found this document useful let us know. If you found typos and things to improve, let us know as well. Your feedback is very important to us. We're working hard to deliver the best material possible. Contact us at: dibeos.contact@gmail.com