📘

Proof of the Cauchy-Schwarz Inequality

Source: Proof of the Cauchy-Schwarz inequality (video) | Khan Academy

Introduction to the Cauchy–Schwarz Inequality 🌟

Imagine you've got two vectors, $\vec{x}$ and $\vec{y}$, quietly sitting in $\mathbb{R}^n$ 🛋️. They're nonzero—no lazy zero vector allowed in this party. The Cauchy–Schwarz inequality tells us something magical about them:

$$\left|\vec{x} \cdot \vec{y}\right| \;\leq\; \|\vec{x}\|\,\|\vec{y}\|$$

Translation: The magnitude of their dot product can never exceed the product of their lengths.

It's like saying: the "overlap" between two arrows in space can't be more than what you'd get if they were perfectly aligned ⚡.

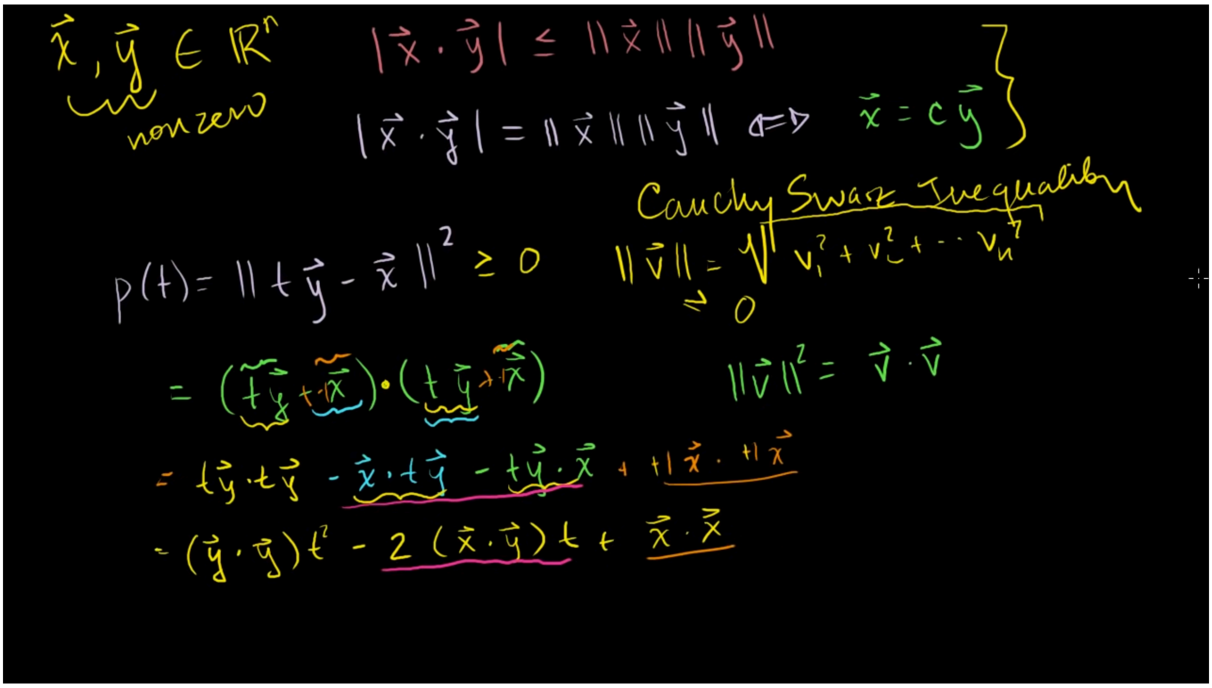
Equality Condition 🎯

The **only** time this turns into an equality (a perfect tie) is when one vector is a scalar multiple of the other (collinear):

$$\vec{x} = \lambda \vec{y} \quad (\lambda \in \mathbb{R})$$

In plain terms: they point in exactly the same or exactly opposite direction—like two train tracks running in the same direction 🚄🚄.

Constructing the Proof 🛠️



Proving this gem is not about memorizing—it's about setting up the right "mathematical trap" so the inequality falls out naturally. Here's the clever strategy:

Setting Up the Problem 🎨

We define a new function:

$$p(t) = \|t\vec{y} - \vec{x}\|^2$$

where $t \in \mathbb{R}$ is just a scalar variable we can adjust freely—like turning a knob on a sound mixer 🎚️.

Why this definition? Because:

- **Fact 1:** The length (norm) of any real vector is **always** non-negative:

  $$\|\vec{v}\| \geq 0$$

This comes from the fact that it's **the square root of a sum of squares**—just like the distance formula can't give you a negative result 🚫➖.

- **Fact 2:** The square of the length can be rewritten using the dot product:

  $||\vec{v}||^2 = \vec{v} \cdot \vec{v}$

  This is a key identity—it lets us trade norms for dot products, which are easier to expand.

Applying Fact 2 to $p(t)$:

$p(t) = (t\vec{y} - \vec{x}) \cdot (t\vec{y} - \vec{x})$

Expanding the Dot Product 💥

The dot product behaves almost like regular multiplication when it comes to distributive, commutative, and associative laws (just remember it's *not* exactly multiplication—it's a projection-based operation). Expanding step-by-step:

1. $(t\vec{y}) \cdot (t\vec{y}) = t^2(\vec{y} \cdot \vec{y})$
2. $(t\vec{y}) \cdot (-\vec{x}) = -t(\vec{y} \cdot \vec{x})$
3. $(-\vec{x}) \cdot (t\vec{y}) = -t(\vec{x} \cdot \vec{y})$ (same as above by commutativity)
4. $(-\vec{x}) \cdot (-\vec{x}) = (\vec{x} \cdot \vec{x})$

Combining:

$p(t) = (\vec{y} \cdot \vec{y})t^2 - 2(\vec{x} \cdot \vec{y})t + (\vec{x} \cdot \vec{x})$

And since $p(t) \geq 0$ for **all** real t, we've got a quadratic in t that never dips below the horizontal axis 📈.

Introducing Constants for Simplicity ✂️

Let's clean this up by introducing friendlier symbols:

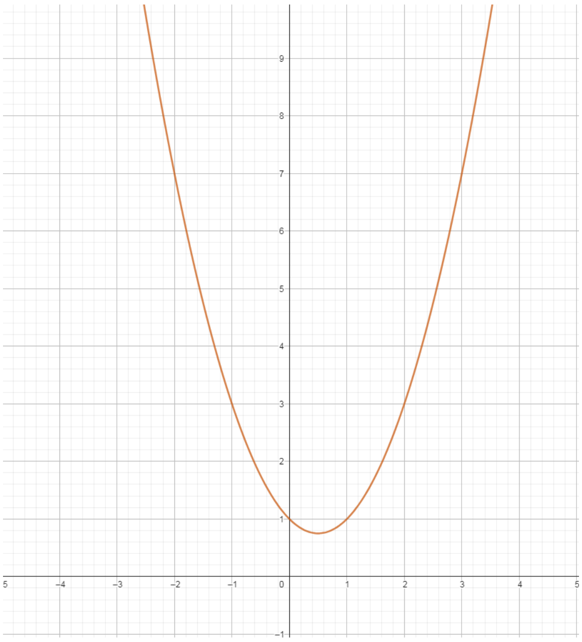
- $a = \vec{y} \cdot \vec{y} = ||\vec{y}||^2$ 📏
- $b = 2(\vec{x} \cdot \vec{y})$ 🔗
- $c = \vec{x} \cdot \vec{x} = ||\vec{x}||^2$ 📏

Now:

$$p(t) = at^2 - bt + c \quad \text{and} \quad p(t) \geq 0 \; \forall t \in \mathbb{R}$$

This re-labeling makes the algebra sleeker—like swapping bulky hiking boots for sleek running shoes 👟.

Using Quadratic Properties 📐



p(t) = at² - bt + c

We've got:

$$p(t) = at^2 - bt + c \quad \text{with} \quad a > 0$$

and we know $p(t) \geq 0$ for **every** real t.

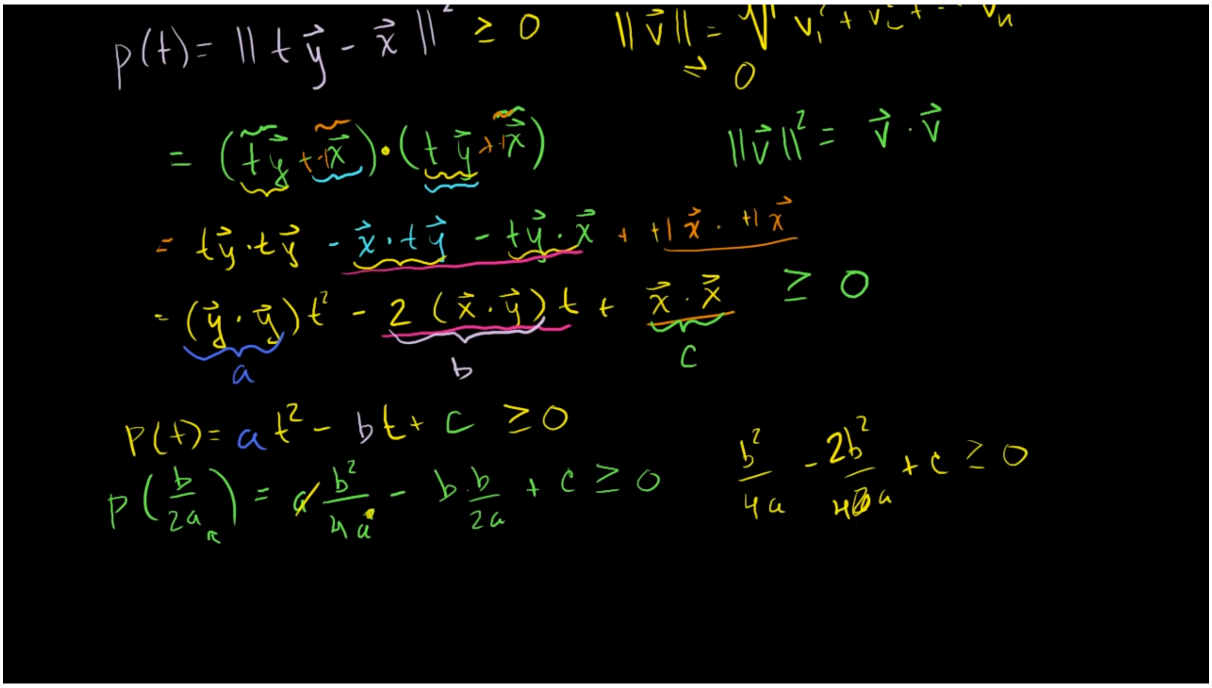
In math-speak, this means our quadratic curve never dips below the t-axis—it's like a hammock that's always hanging above the ground 🪢.

The **minimum** of a quadratic $at^2 - bt + c$ (when $a > 0$) happens at:

$$t_{\min} = \frac{b}{2a}$$

So, we plug $t = \dfrac{b}{2a}$ into $p(t) \geq 0$.

Substituting the Minimum Value 🧮



At $t = \dfrac{b}{2a}$:

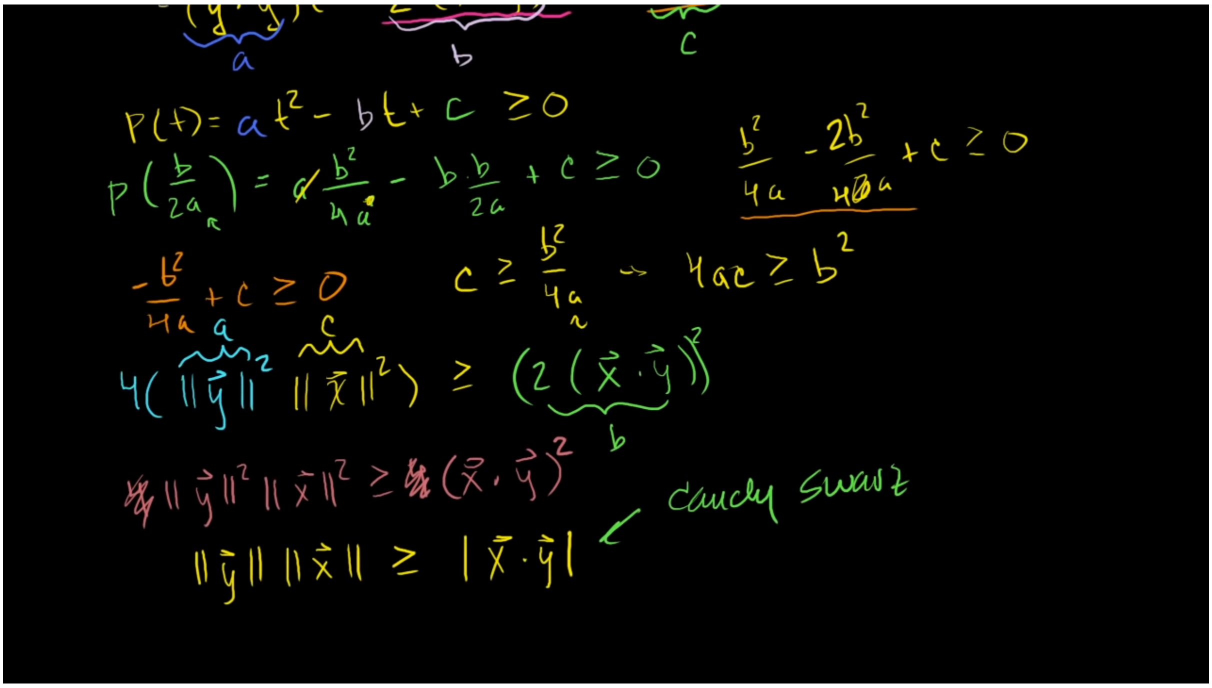
- $p\left(\dfrac{b}{2a}\right) = a\left(\dfrac{b}{2a}\right)^2 - b\left(\dfrac{b}{2a}\right) + c \;\geq\; 0$
- $a \cdot \dfrac{b^2}{4a^2} - b \cdot \dfrac{b}{2a} + c \;\geq\; 0$
- $\dfrac{b^2}{4a} - \dfrac{b^2}{2a} + c \;\geq\; 0$
- Combine them:

  $$-\frac{b^2}{4a} + c \geq 0$$

- Rearranges to:

  $$c \geq \frac{b^2}{4a}$$

Back-Substitution and Final Inequality 🎯



Now we replace a, b, c with their original meanings:

- $a = ||\vec{y}||^2$
- $b = 2(\vec{x} \cdot \vec{y})$
- $c = ||\vec{x}||^2$

Plugging in:

$$||\vec{x}||^2 \ \geq \ \frac{\left[2(\vec{x} \cdot \vec{y})\right]^2}{4 \, ||\vec{y}||^2}$$

Simplify:

$$||\vec{x}||^2 \ \geq \ \frac{4(\vec{x} \cdot \vec{y})^2}{4 \, ||\vec{y}||^2}$$

Cancel the 4:

$$||\vec{x}||^2 \ \geq \ \frac{(\vec{x} \cdot \vec{y})^2}{||\vec{y}||^2}$$

Multiply through by $||\vec{y}||^2$ (positive, so inequality direction stays the same):

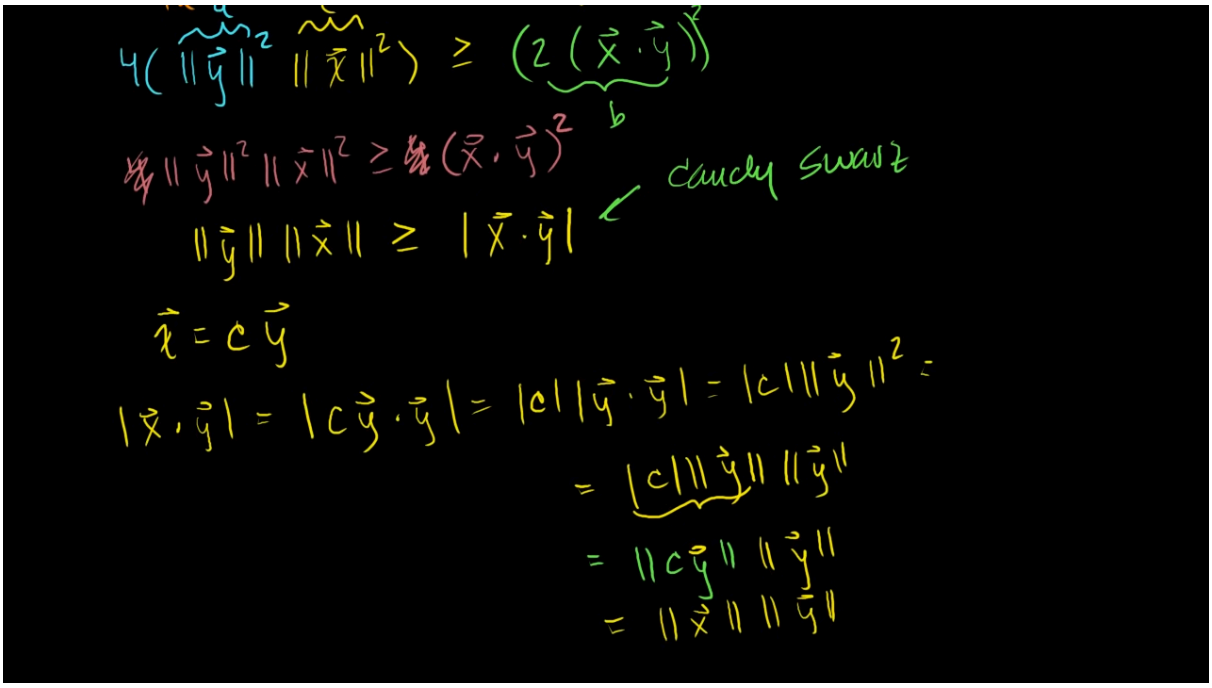
$$\|\vec{x}\|^2\,\|\vec{y}\|^2 \;\geq\; (\vec{x}\cdot\vec{y})^2$$

Taking square roots (principal root to keep things positive):

$$\|\vec{x}\|\,\|\vec{y}\| \;\geq\; \left|\vec{x}\cdot\vec{y}\right|$$

Boom 💥—we've just re-discovered the Cauchy–Schwarz inequality in all its glory.

Equality Condition ⚖️



When does equality happen?

It's only when our quadratic $p(t)$ actually *touches* the t-axis—i.e., has a single root. That's the case if:

$$\vec{x} = \lambda\vec{y}$$

for some scalar $\lambda \in \mathbb{R}$.

In that case:

$\vec{x} \cdot \vec{y} = (\lambda \vec{y}) \cdot \vec{y} = \lambda(\vec{y} \cdot \vec{y})$

$\vec{x} \cdot \vec{y} = \lambda \|\vec{y}\|^2$

Taking absolute values:

$|\vec{x} \cdot \vec{y}| = |\lambda| \, \|\vec{y}\|^2$

$|\vec{x} \cdot \vec{y}| = |\lambda| \, \|\vec{y}\| \, \|\vec{y}\|$

Then, taking magnitudes ($\|\vec{x}\| = |\lambda| \, \|\vec{y}\|$), you find:

$|\vec{x} \cdot \vec{y}| = \|\vec{x}\| \, \|\vec{y}\|$

✅ Meaning they're perfectly aligned (or perfectly opposite).

Applications & Notes 📚

The Cauchy–Schwarz inequality is **everywhere** in linear algebra and analysis:

- **Vector projections**: It guarantees projection lengths behave nicely.
- **Angles between vectors**: Leads to the cosine formula $\cos\theta = \dfrac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \, \|\vec{y}\|}$.
- **Norm properties**: It's a key ingredient in proving the triangle inequality.

Think of it as the "gravity" of vector math—it's always there, quietly pulling results into place 🌌.

Key Takeaways 📝

- **Statement**:

  $$\left|\vec{x} \cdot \vec{y}\right| \leq \|\vec{x}\| \, \|\vec{y}\|$$

  for all nonzero $\vec{x}, \vec{y} \in \mathbb{R}^n$.

- **Equality**: Happens **only** if vectors are scalar multiples (collinear).
- **Proof strategy**:
  1. Build a non-negative quadratic $p(t) = \|t\vec{y} - \vec{x}\|^2$.
  2. Expand via dot products.
  3. Use quadratic minimization to reveal the inequality.
  4. Take square roots for the final form.
- **Why it matters**: It links **dot products** and **norms**, and appears in almost every advanced proof in linear algebra, functional analysis, and even probability theory.