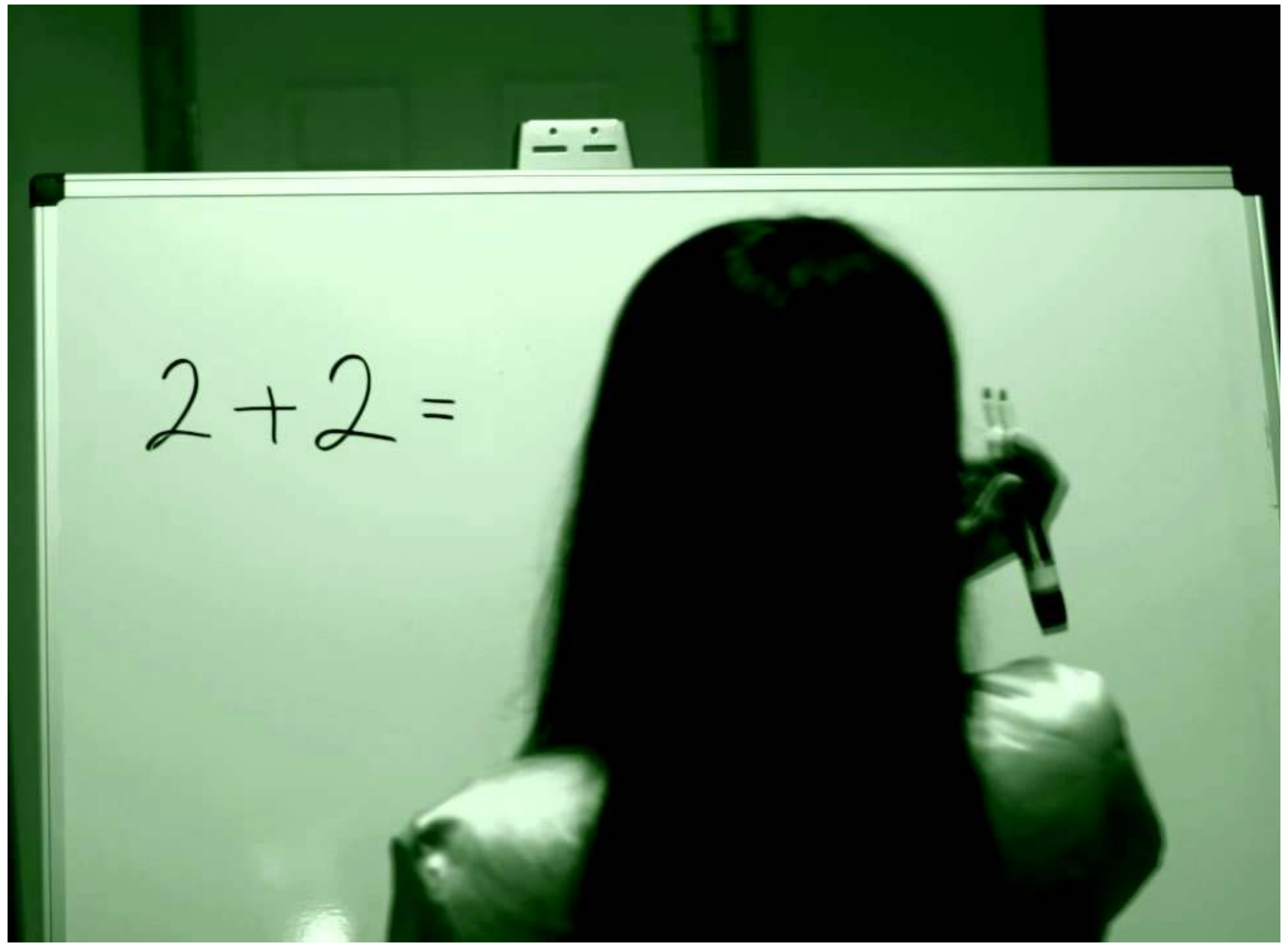
Math Fundamentals

Providing universal access to AI education and practice

Maths and **Data Science**

Machine Learning theory is a field that intersects **statistical**, **probabilistic**, **computer science** and **algorithmic aspects** arising from learning iteratively from data and finding hidden insights which can be used to build intelligent applications. Despite the immense possibilities of Machine and Deep Learning, a thorough **mathematical understanding** of many of these techniques is necessary for a good grasp of the **inner workings** of the **algorithms** and getting good results.



Maths in **Life**

It is not just data science. **Math is** essentially **everywhere**, even in nature. Nowadays, we use lots of technology without understanding their inner details.

How many amps of current can go through your body? **0.1 ampere** for a mere **2 seconds** can be fatal.

Hopefully, an engineer has done the **calculations** to ensure that doesn't happen. That has a value.

Common **excuses** for not learning Math

- My cell phone has a calculator.
- It requires too much thinking
- I was never taught math well
- I am not “Math” smart
- The dog ate my homework
- I will never do math in my life

Common **counterarguments** for learning Math

- There is no substitute for exercising your mind.

- Thinking is generally good for survival. From an evolution perspective, the ones that don't think, die.

- The internet has unlimited resources. You can start today!

- Funny think is, no one is. Some people work hard and some people do not.

My **personal** story

- **Primary School:** I could do fast calculations, was lazy with everything else.

- **High School:** I could do fast calculations, was lazy with everything else.

- **Montreal:** Found out Koreans could calculate way faster than me, was jealous.

- **USA:** Got an extremely inspiring math teacher. Essentially, he tricked me into doing Calculus -which I would have probably thought it was too hard-. Ended up signing up for Honors Math, Statistics, Calculus I and Calculus II.

  The principal did not think I could do that much math, specially because I was from Spain and that is a province of Mexico. I actually dropped out of Calculus and ended up signing for study sessions, that very coincidentally took place in the room and time where Calculus was held. A korean and me were the only one's that got AP Calc certified.

My **personal** story

- **1st year of University:** I thought I had everything solved because I already did Calculus. Turns out, University Calculus was way harder. Had to relearn everything from scratch. It was incredibly frustrating to see my friends that had not done Calculus before get things quicker than me.

- **2nd year of University:** We had a really cocky teacher, smart but horrible at teaching and insulted us when we could not figure out the value of a 3D integral or a Fourier transform. Learned from Youtube. PatrickJMT and indian guys were way better teachers.

PRINCIPI DE CAVALIERI

Donat $W = \{(x,y,z) : x \in [a,b]\} \subset \mathbb{R}^3$, sigui $\Omega(x_0) \subset \mathbb{R}^2$ la secció que s'obté al tallar W amb el pla $x = x_0$ (vegi's la figura 2.1 de [1]). Diem $A(x_0)$ a l' àrea de $\Omega(x_0)$, llavors el volum de W és

$$V(W) = \int_a^b A(x)dx.$$

Exemple:(Volum d'un con)
$W = \{z^2 = x^2 + y^2, z \in I = [0,H]\}$. Llavors

$$V(W) = \int_0^H A(z)dz = \int_0^H \pi z^2 dz = \pi H^3/3.$$

Exemple:(Volum d'un el.lipsoïde)

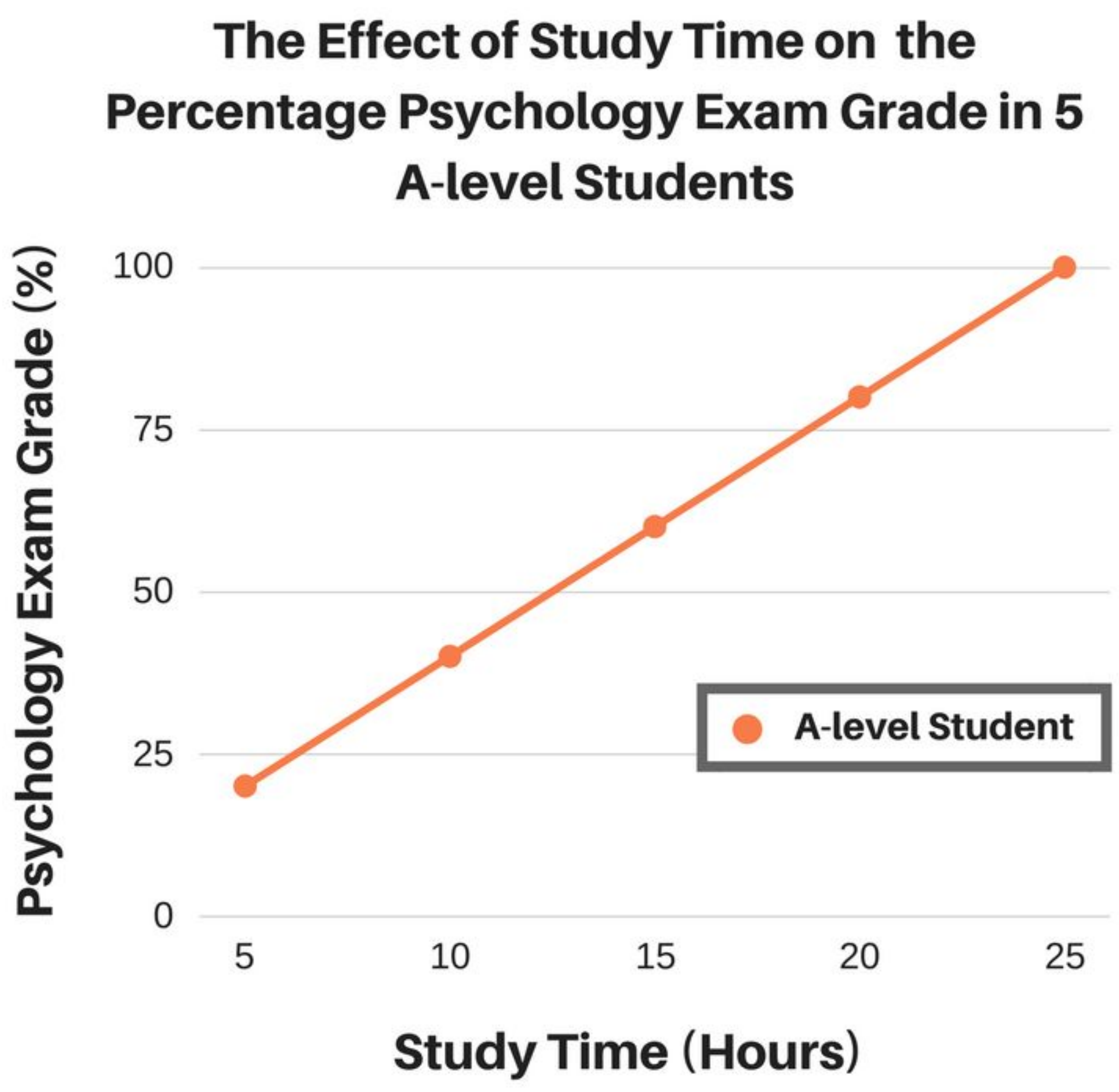
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1.$$

Al tallar amb plans z constant s'obtenen el.lipses de la forma $\frac{x^2}{a^2\left(1-\frac{z^2}{c^2}\right)} + \frac{y^2}{b^2\left(1-\frac{z^2}{c^2}\right)} \leq 1$.
Com calculem l'àrea d'una el.lipse? Doncs, aplicant Cavalieri en dimensió 2, i.e., tallant amb rectes perpendiculars als eixos calculant la longitut del segment intersecció i despres integrant respecte de la variable que parametritza aquestes rectes.
En aquest cas la longitut és $L(x) = 2b\sqrt{1 - \frac{x^2}{a^2} - \frac{z^2}{c^2}}$ i per tant $A(z) = \int_{-a}^{a} L(x)dx = \pi ab\left(1 - \frac{z^2}{c^2}\right)$,
així doncs, $V = \int_{-c}^{c} A(z)dz = \frac{4}{3}\pi abc$.

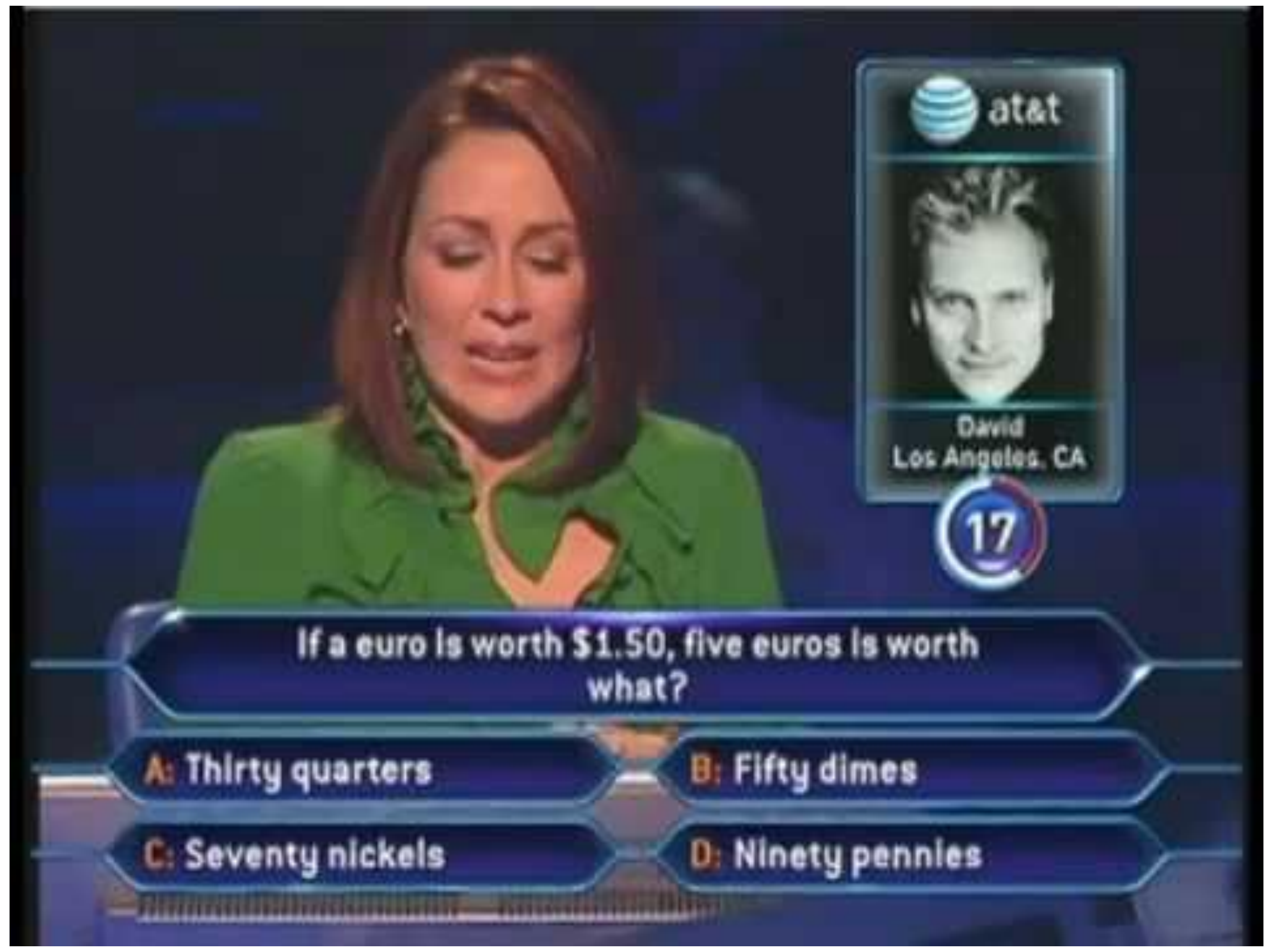
My **personal** story

- **3rd year of University:** I recall that Differential equations was extremely hard. I was genuinely scared so we studied a lot. Due to the amount of hard work we put in, we actually knew it by hard. However, I almost failed the easiest subject (Quality techniques for Statistics) because I did not study at all.

- **4rd year of University:** No more hard math! Turns out, real world math is easier than having to calculate the volume of a turtle. In a way, I am glad to have done that since it gave me the resilience to persist through hard problems.

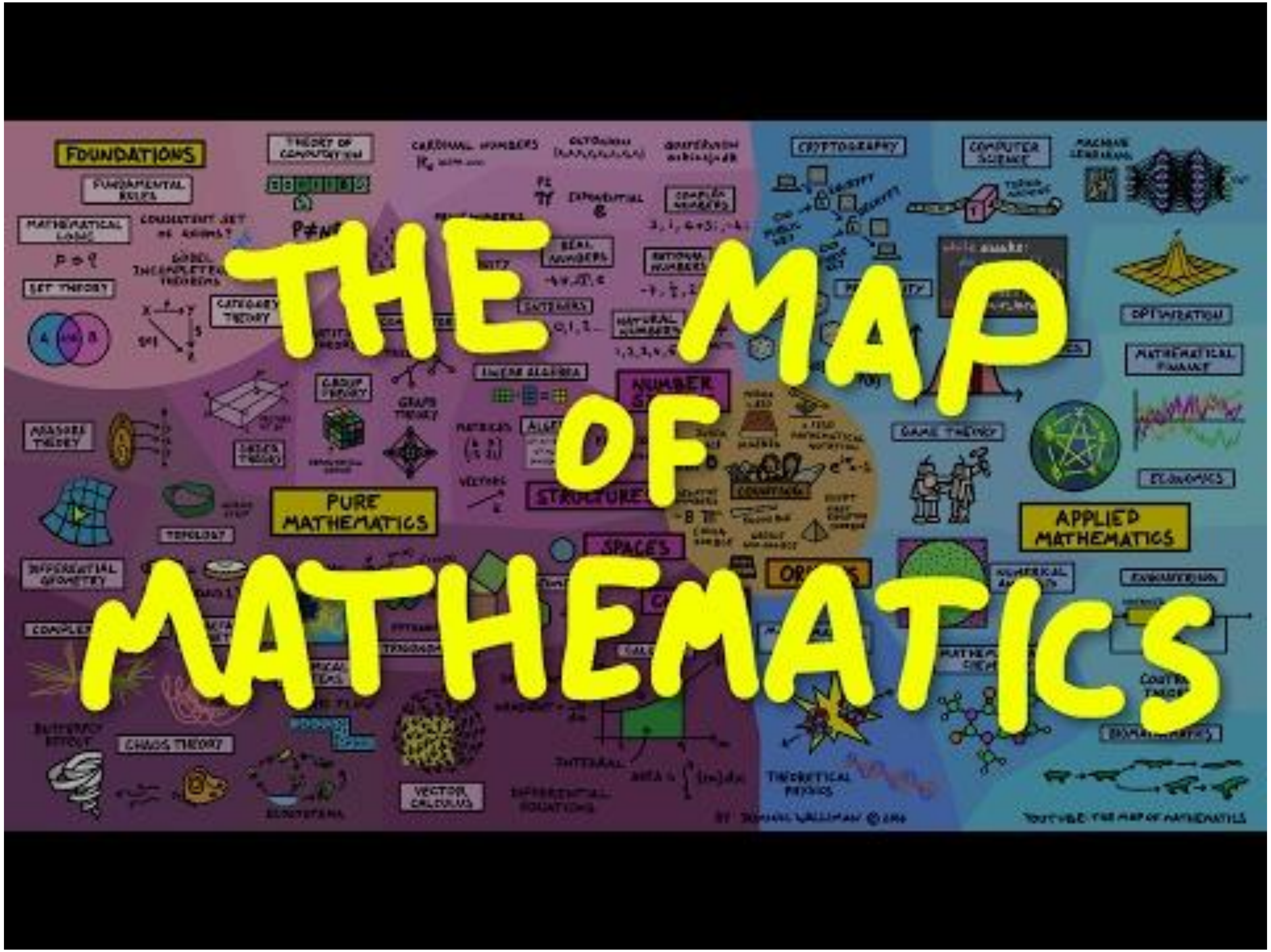


Let's get **inspired**



[Wikipedia article]

Let’s begin with the **overall map**



[MindMap]

[Khan Academy]

[PatrickJMT]

[Numberphile]

[3Blue1Brown]

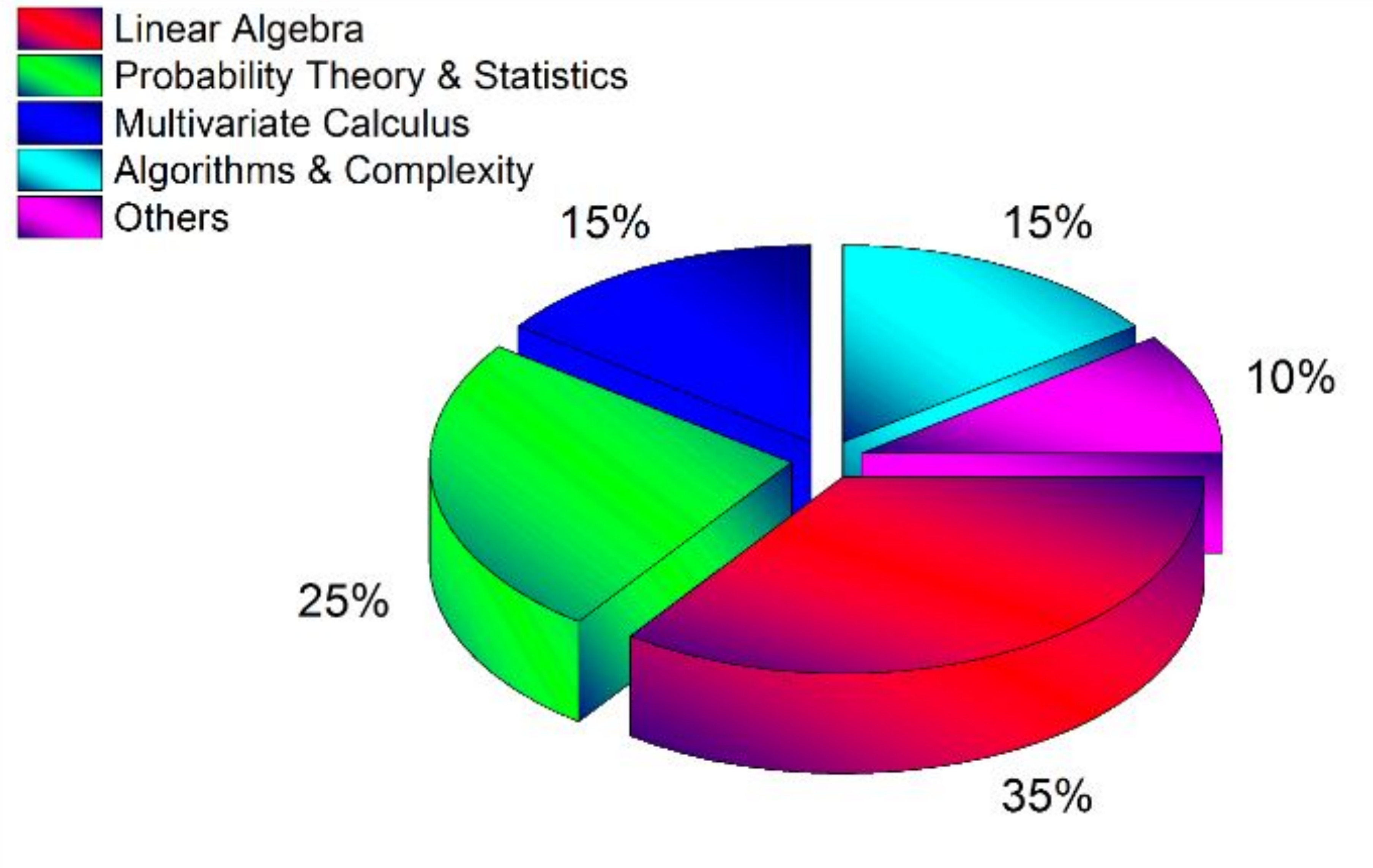
Why Worry About The Maths in ML?

1. Selecting the right algorithm which includes giving considerations to **accuracy**, **training time**, **model complexity**, number of parameters and number of features.
2. Choosing **parameter settings** and validation strategies.
3. Identifying **underfitting** and **overfitting** by understanding the **Bias-Variance tradeoff**.
4. Estimating the **right confidence interval** and **uncertainty**.

We cannot cover all of this in one day and **not everything is required to get into the field**. Just seeking to be brutally honest with the long term plan.

Math in Machine Learning

1. Use of Linear Algebra and ML: **Principal Component Analysis (PCA)**, **Singular Value Decomposition (SVD)**, Eigendecomposition of a matrix, Symmetric **Matrices,** Orthogonalization & Orthonormalization, **Matrix Operations**, **Projections, Eigenvalues & Eigenvectors**, Vector Spaces and Norms are very helpful towards for understanding the optimization methods used for machine learning.
2. **Probability Theory and Statistics**: **Bayes' Theorem**, **Variance** and Expectation, Conditional and Joint Distributions, Standard Distributions (Bernoulli, Binomial, Multinomial, Uniform and **Gaussian**), Moment Generating Functions, Maximum Likelihood Estimation (MLE), and **Sampling Methods**.
3. **Multivariate Calculus**: Some of the necessary topics include Differential and Integral Calculus, Partial Derivatives, Vector-Values Functions, Directional Gradient, Hessian, Jacobian, Laplacian and Lagragian Distribution.
4. **Algorithms and Complex Optimizations:** Knowledge of data structures (Binary Trees, Hashing, Heap, Stack etc), Dynamic Programming, Randomized & Sublinear Algorithm, Graphs, Gradient/Stochastic Descents



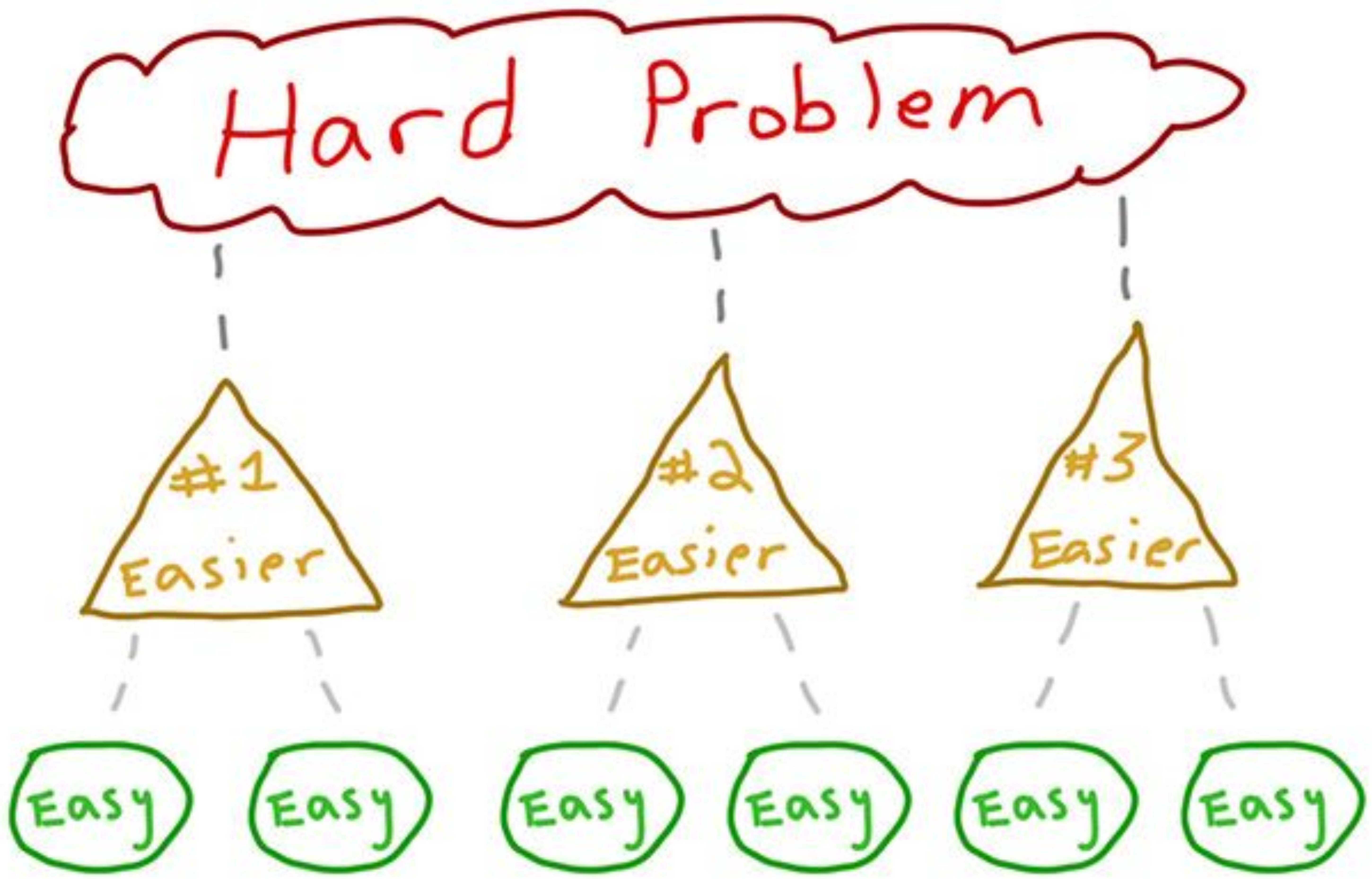
Importance of Maths Topics Needed For Machine Learning

More resources

- Khan Academy's Linear Algebra, Probability & Statistics, Multivariable Calculus and Optimization.
- Coding the Matrix: Linear Algebra through Computer Science Applications by Philip Klein, Brown University.
- Linear Algebra — Foundations to Frontiers by Robert van de Geijn, University of Texas.
- Applications of Linear Algebra, Part 1 and Part 2. A newer course by Tim Chartier, Davidson College.
- Joseph Blitzstein — Harvard Stat 110 lectures.
- Larry Wasserman's book — All of statistics: A Concise Course in Statistical Inference.
- Boyd and Vandenberghe's course on Convex optimization from Stanford.
- Linear Algebra — Foundations to Frontiers on edX.
- Udacity's Introduction to Statistics.

Before we begin...

Small victories, big steps



Overview of the Plan

- Introduction to Linear Algebra
- Vectors
- Linear Combination
- Linear Transformation and Matrices
- Practice with Vectors and Linear Combinations
- Autonomous Learning
- **Phase 2 (Calculus)**

Our goal

In this "mini" Linear Algebra lesson we will cover some **important tools in Linear Algebra**. This will help set a mathematical foundation for understanding and creating a Neural Network.

This lesson is NOT a complete stand-alone Linear Algebra class.

Our goal is to give you the **fundamental tools** you will need to train neural networks while showing you the **beautiful visual world of Vectors** and **Matrices**.

Mathematical Definition of Vectors

What is a Vector? In a plain explanation, a vector is an ordered list of numbers.

Each **element** in the vector, also called **component** or **coordinate**, is a number, denoted here by ***vn***.

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_{n-1} \\ v_n \end{bmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_{n-1} \\ v_n \end{pmatrix}$$

This specific vector (in the picture above) has ***n*** elements and can be in the field of Real Numbers **R**.

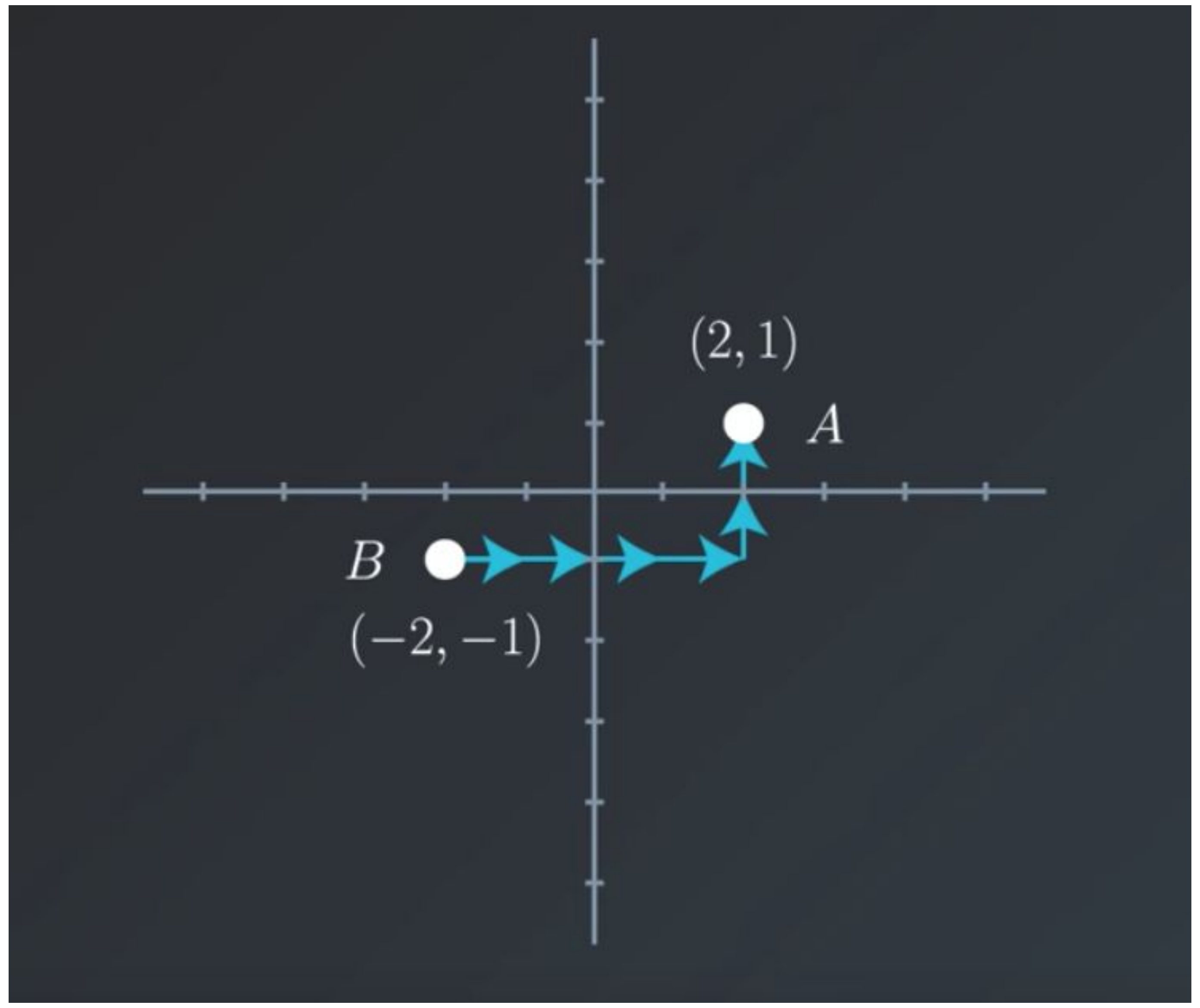
A vector of ***n*** real elements defines an ***n*** dimensional vector and belongs to **R**.

In the example above:

- $\vec{x} \in \mathbb{R}^n$
- $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$

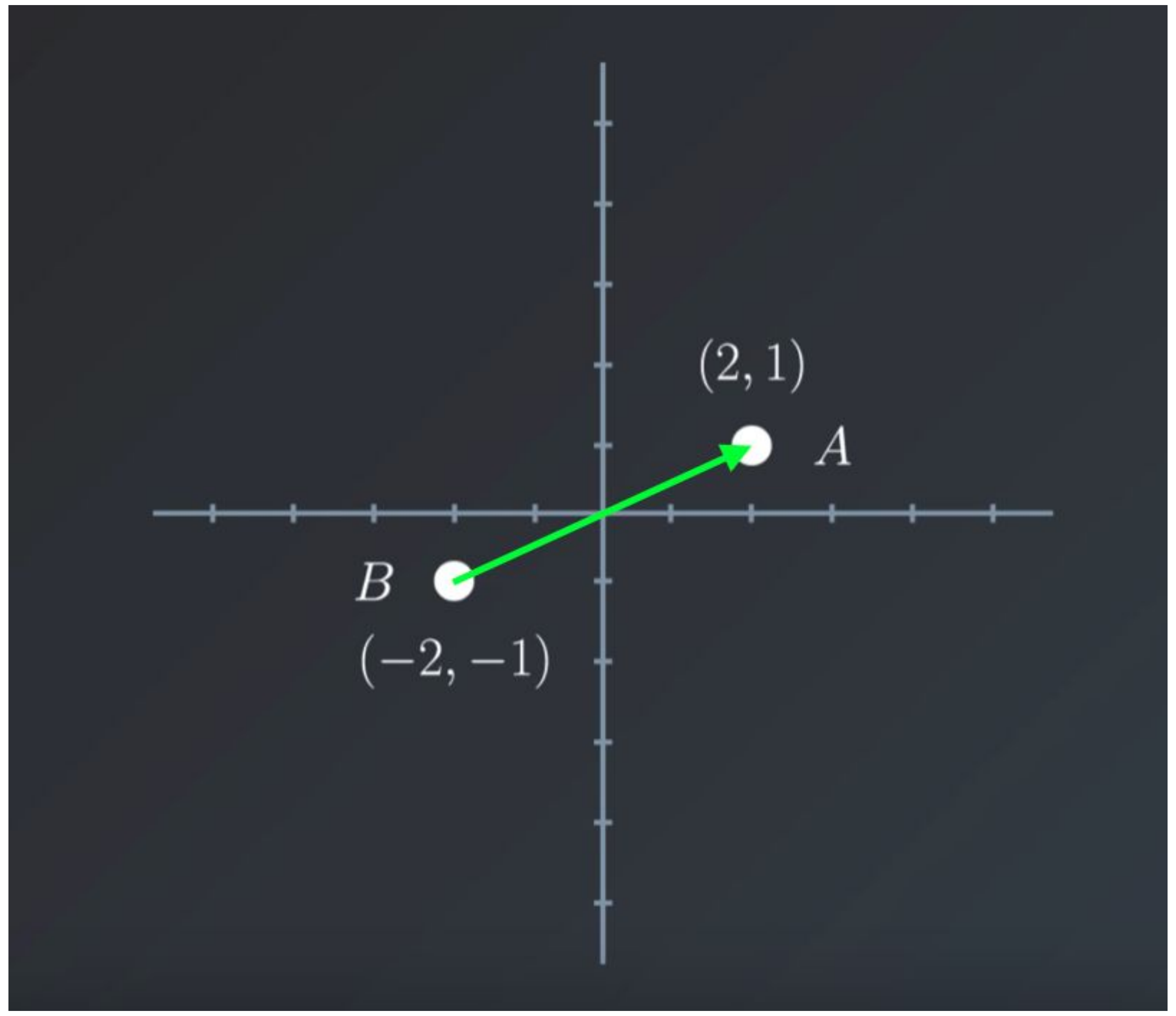
Mathematical Definition of Vectors

As in the video let's put this into a more visual context and focus on a 2D vector of the field of real numbers. In other words, we will focus on a vector in $\mathbf{R}^2$, which defines all points on the plane.



To go from point B to point A we will need to take 4 steps to the right and 2 steps up. This defines the vector $\vec{\mathbf{x}} = [4 , 2]$.

The green arrow below defines the **vector** $\vec{\mathbf{x}}$.



Mathematical Definition of Vectors

To calculate the **direction of the movement** we will use an angle. We can use Degrees or Radians. In this example we will focus on Degree.

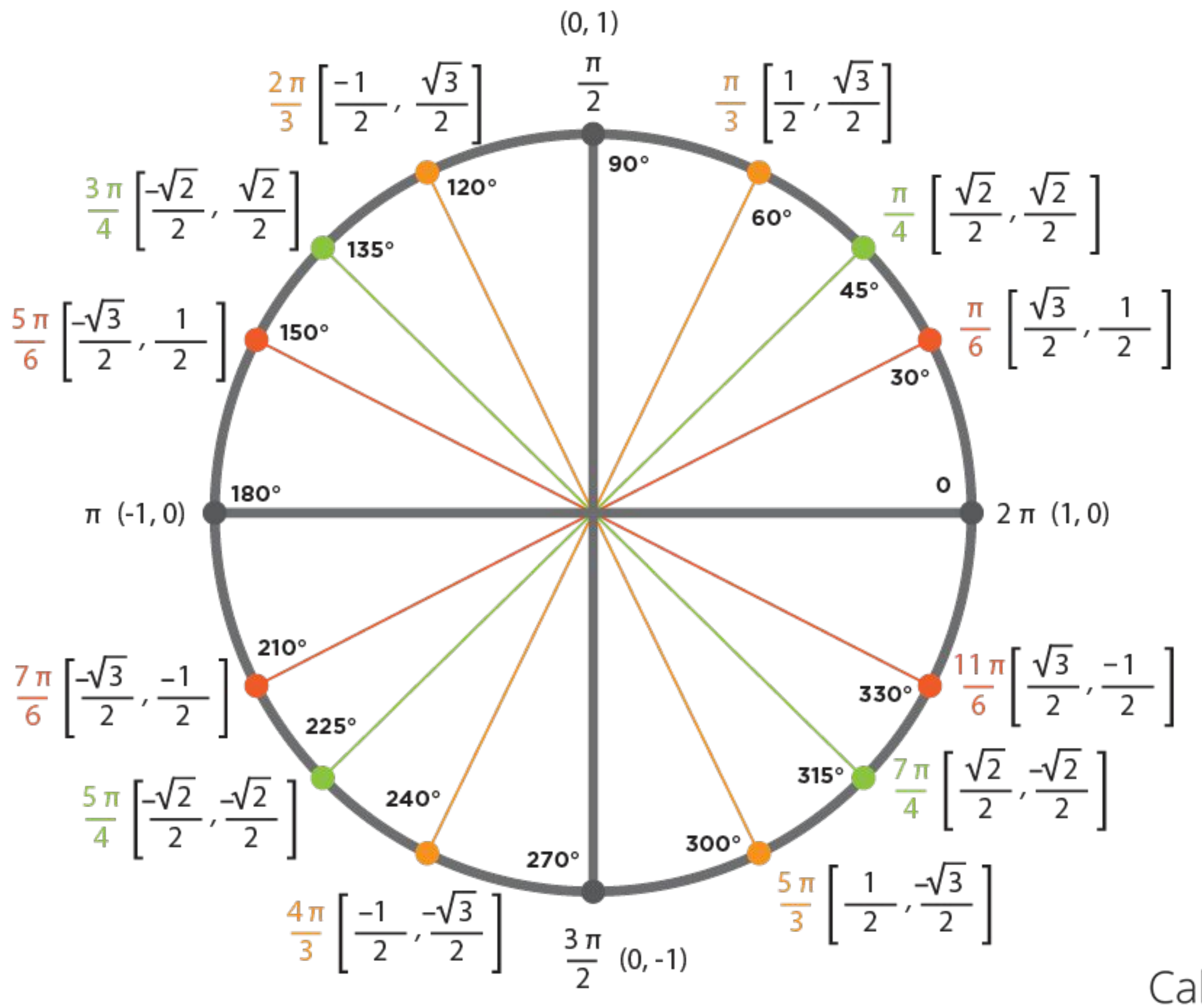
It is always possible to [move Degrees to Radians and vice versa](#).

$$\textbf{degrees} = \textbf{r} \cdot \frac{180^\circ}{\pi}$$

$$\textbf{radians} = \textbf{degrees} \cdot \frac{\pi}{180^\circ}$$

Pedantic legend:

degree symbol (°) is a typographical symbol used to represent degrees of an arc.



Calc

Transpose

It's very important to note that in this lesson we emphasize the **column vector**.

Vector $\vec{x} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix}$ is a column vector.

But vectors can also be represented at **row vectors**.

Vector $\vec{y}$=[a1 a2 a3 … an] is a row vector.

If you look closely at both vectors, $\vec{x}$ and $\vec{y}$, you will notice that they have the same elements, only one is a column vector and the other is a row.

It's as if one vector was actually tilted by 90°.

Transpose

In the world of Linear Algebra we call this change a **transpose**. The mathematical symbol of a transpose is **T** and it's used the following way:

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix}^T = \begin{bmatrix} a_1 & a_2 & a_3 & \dots & a_n \end{bmatrix}$$

Or:

$$\begin{bmatrix} a_1 & a_2 & a_3 & \dots & a_n \end{bmatrix}^T = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix}$$

$$\vec{x}^T = \vec{y}$$

or

$$\vec{y}^T = \vec{x}$$

Magnitude and Direction

Each vector holds the magnitude as well as the direction of the movement.

Lets calculate both for vector $\vec{x}$=[4 2].

The symbol we use for the magnitude is *// //* .

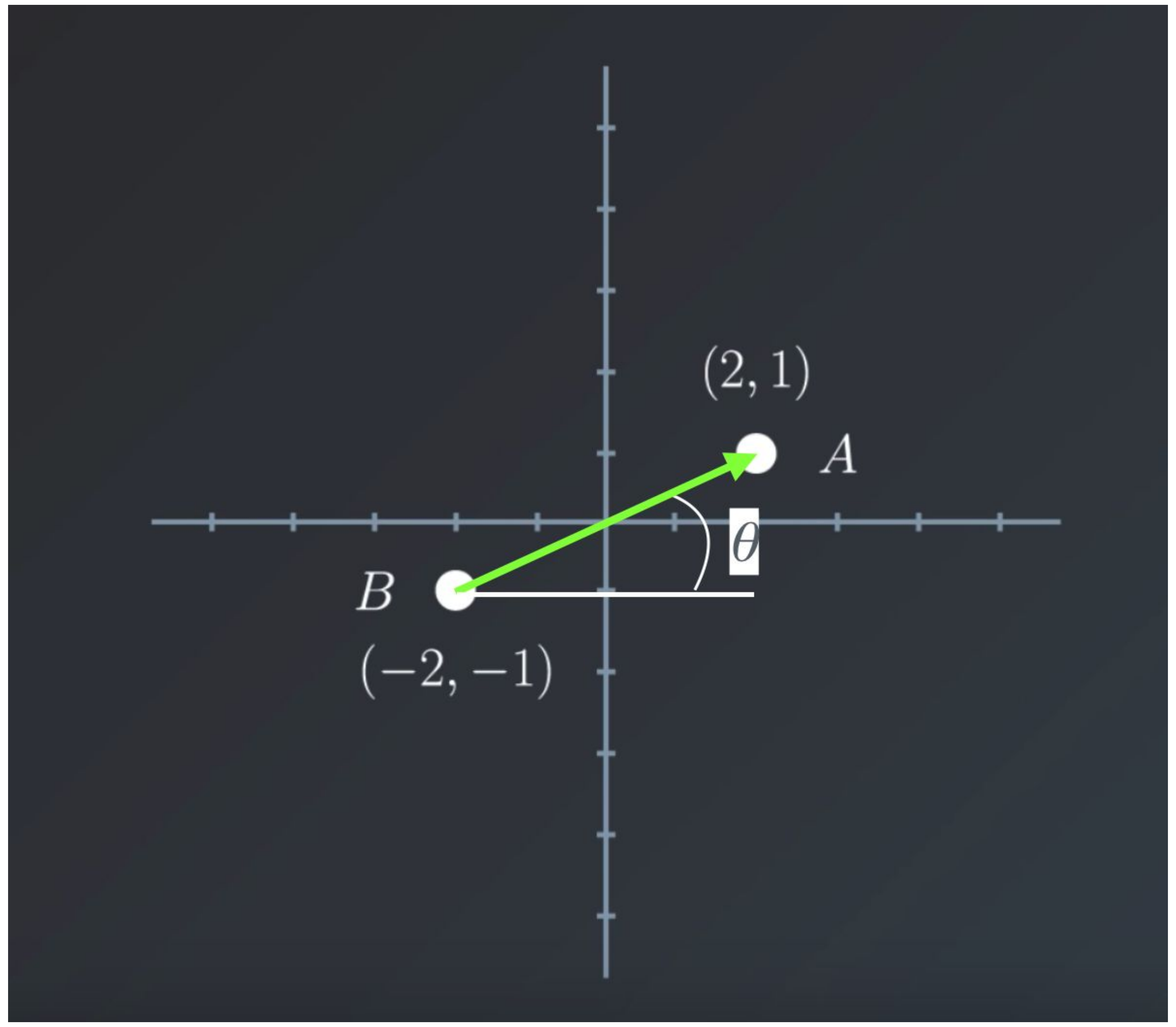
To calculate the magnitude of a ***2D*** **vecto**r we will use the Pythagorean Theorem.

In our example the magnitude will be calculated the following way:

$$||\vec{x}|| = \sqrt{4^2 + 2^2}$$

Where 4 is the horizontal component of the vector and 2 is the vertical.

Magnitude and Direction



To calculate the direction of the movement we will use an angle. As we stated above, let's use Degrees for simplicity.

Let's look at the vector $\vec{x}$ again. It has an angle θ with respect to the horizontal axis.

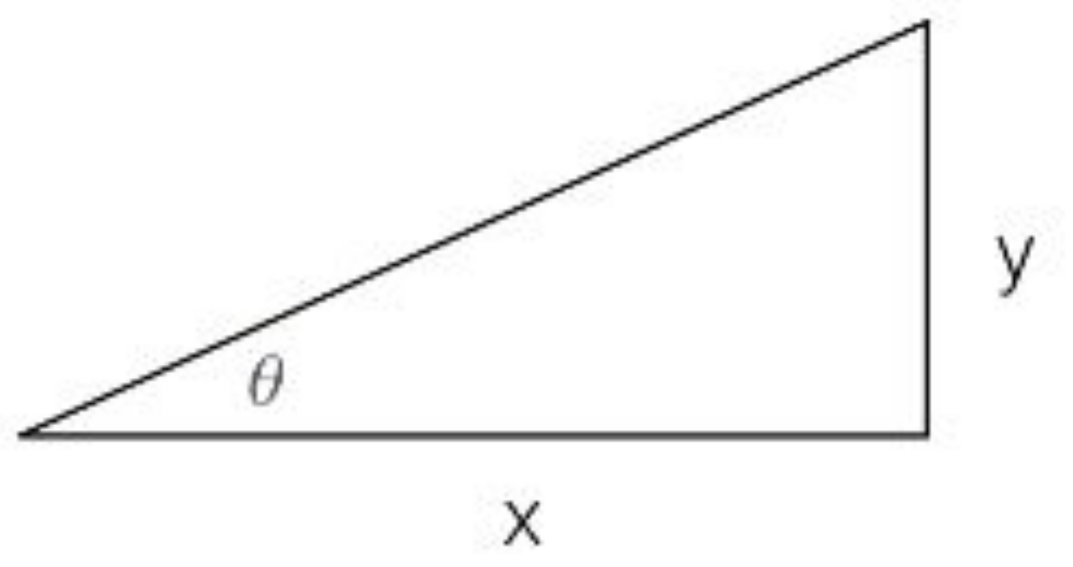
To calculate θ we will use what we remember from Trigonometry!

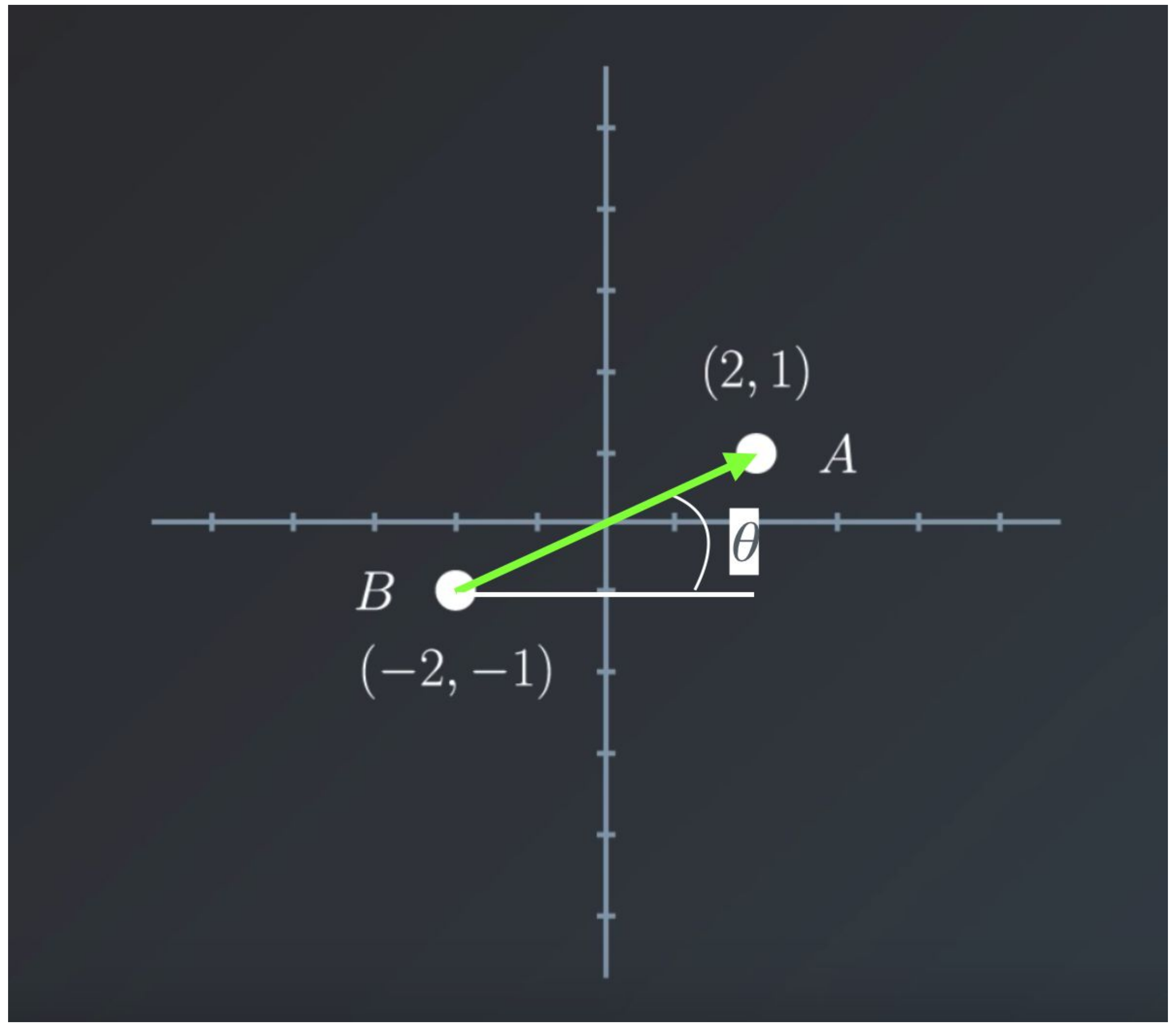
Magnitude and Direction

In the specific angle illustrated below θ is calculated the following way:

$$\theta = tan^{-1}\frac{y}{x}$$

Equation 3

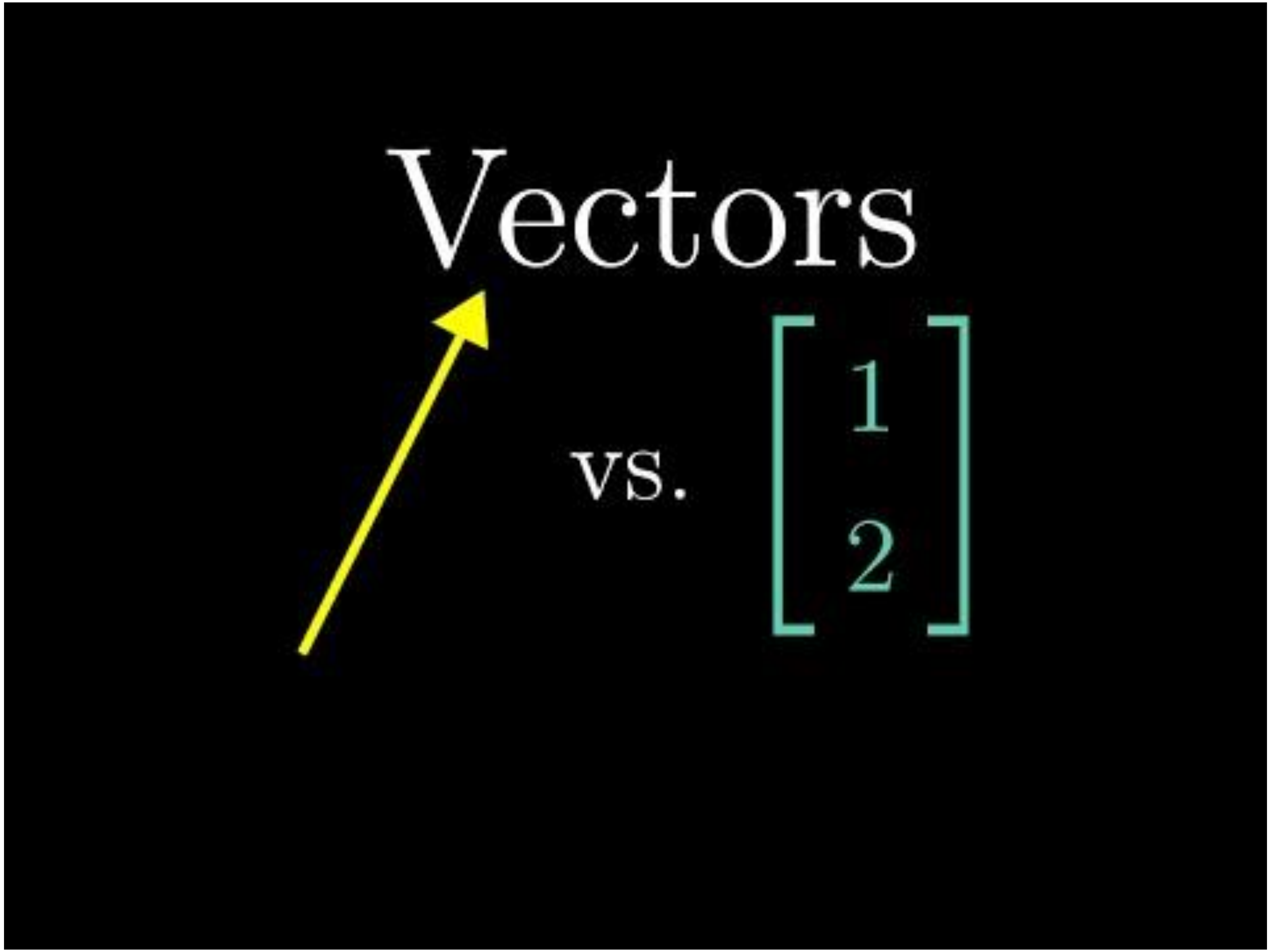




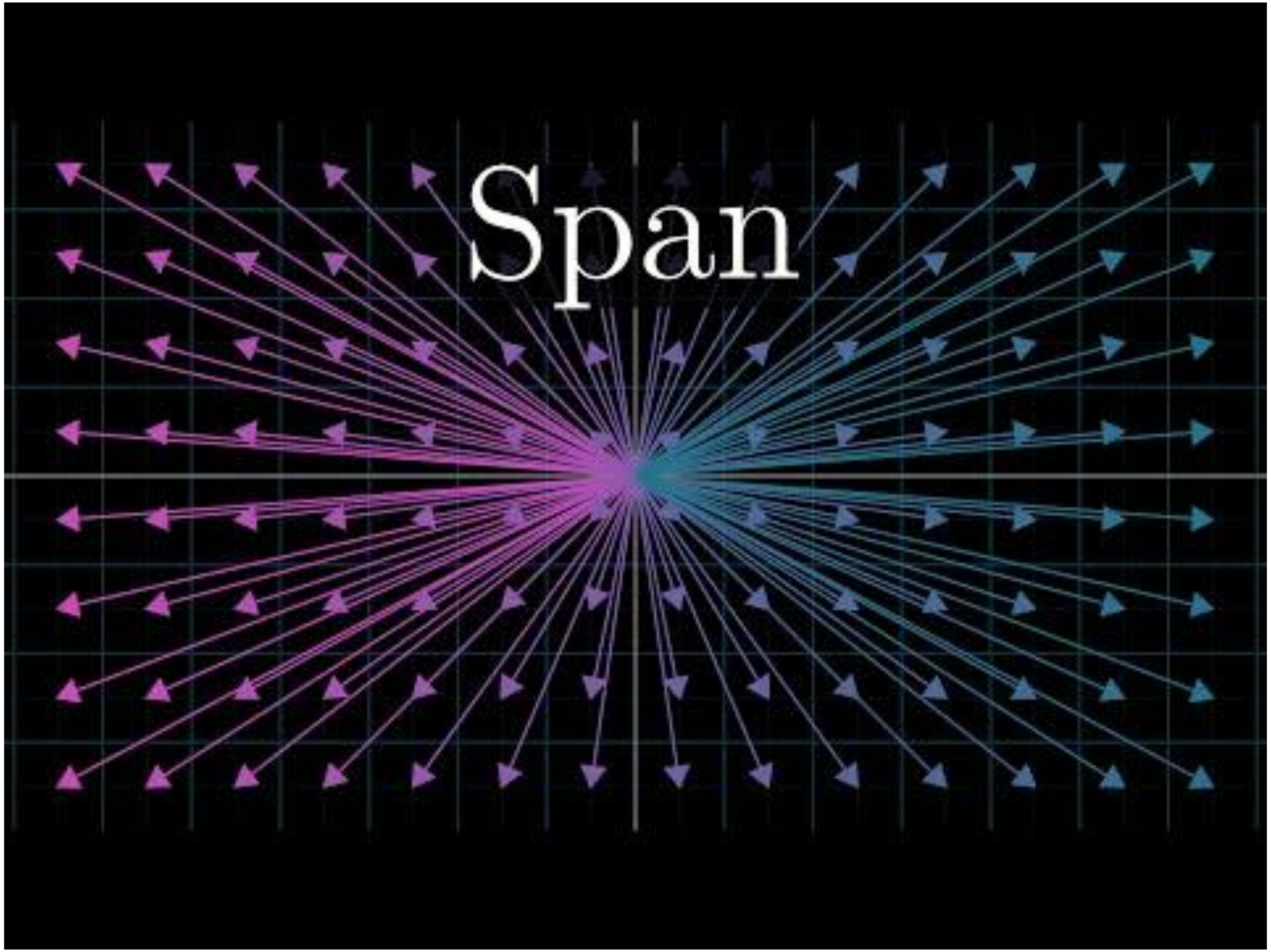
Therefore, in our case

$$\theta = tan^{-1}\frac{2}{4} = 26.56505^{\circ}$$

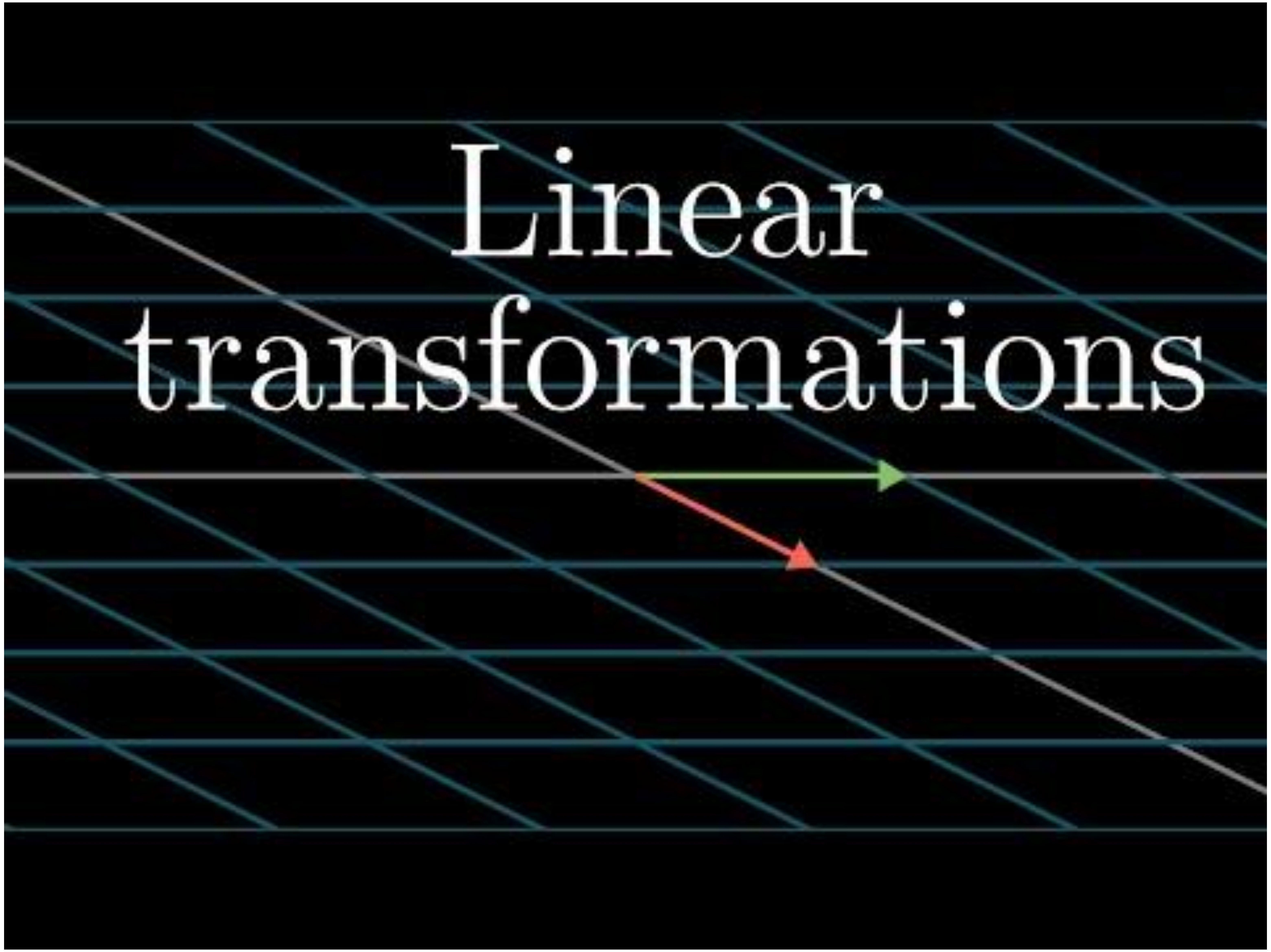
The Essence of Linear Algebra



Linear Combinations



Linear Transformations



Autonomous part… How far can you get?

- 12 videos left
- 130 minutes to complete all of them
- **2 hours, 20 minutes**
- If you use pomodoro, you can get it done in **2hours , 50 minutes**
- Feel free to pick your own pace. This is not a race.
- Ask questions in **Slack!!**

Autonomous part… How far can you get?