Tutorial 1: Linear Algebra Revisited

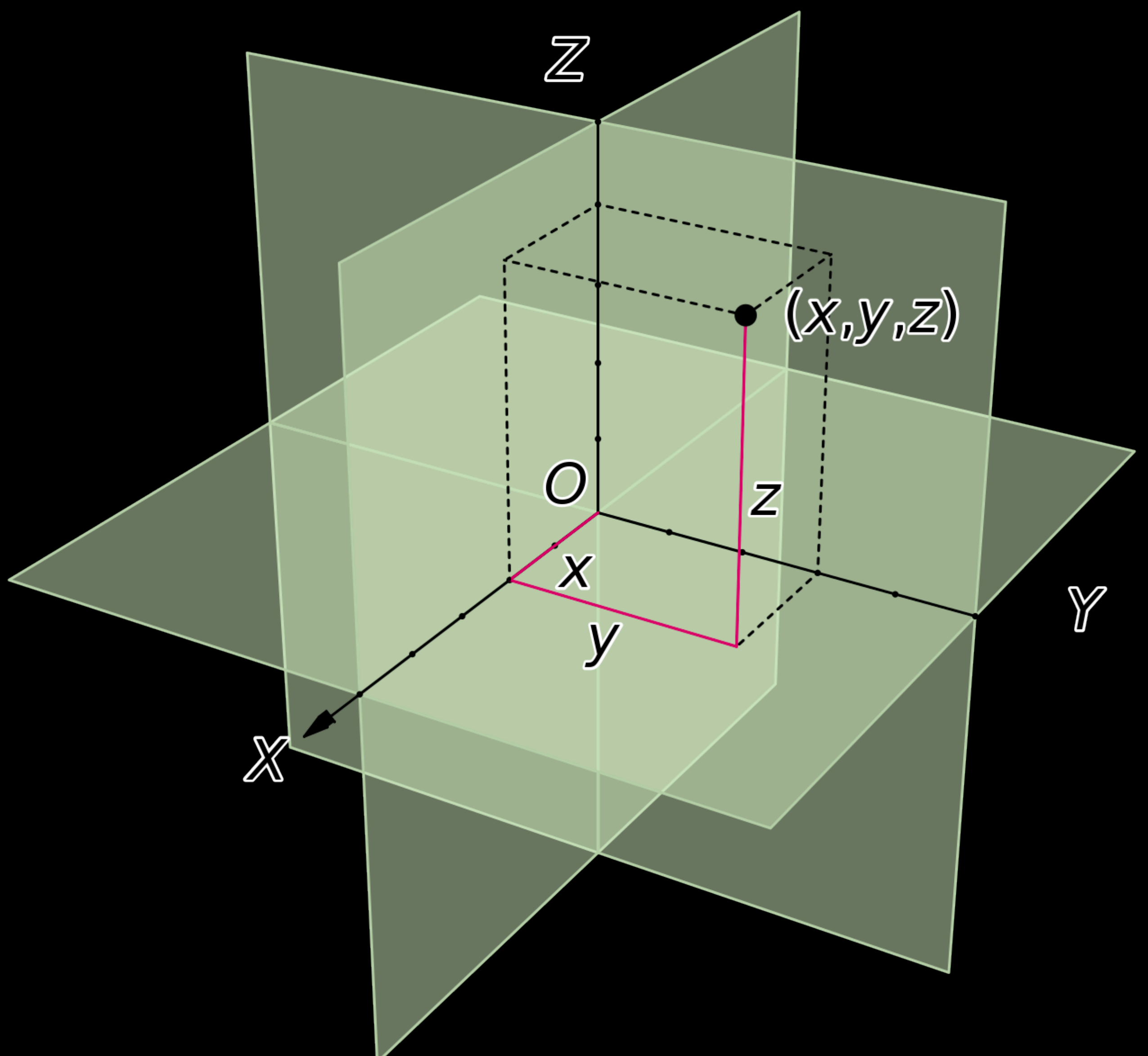
Digital Image Processing (236860)

Vector spaces

- A **vector space** over a field $\mathscr{F}$ is a set $\mathscr{V}$ that has the following properties:
  - A sum $\mathscr{V} \times \mathscr{V} \to \mathscr{V}$, denoted as $(x, y) \to x + y$, is define, satisfying:
    - Commutativity: $x + y = y + x$
    - Associativity: $x + (y + z) = (x + y) + z$
    - Null element: $x + 0 = x$
    - Opposite element: $u + (-u) = 0$
  - A product $\mathscr{F} \times \mathscr{V} \to \mathscr{V}$, denoted as $(\alpha, x) \to \alpha x$, is defined, satisfying:
    - Distributivity in $\mathscr{V}$: $\alpha(x + y) = \alpha x + \alpha y$.
    - Distributivity in $\mathscr{F}$: $\alpha(x + y) = (\alpha + \beta)x = \alpha x + \beta x$.
    - Homogeneity in $\mathscr{F}$: $\alpha(x + y) = \alpha(\beta x) = (\alpha\beta)x$.
    - Scalar unit element: $1 \cdot x = x$.

Vector spaces

- The elements of $\mathscr{V}$ are called **vectors**.
- The elements of $\mathscr{F}$ are called **scalars**.
- Throughout the course we will assume $\mathscr{F} \in \{\mathbb{R}, \mathbb{C}\}$.



Example 1: $\mathbb{R}^d$

Example 2: function spaces

- $\mathbb{F}\left(\mathbb{R}^d, \mathbb{R}\right) = \left\{f : \mathbb{R}^d \to \mathbb{R}\right\}$

- For $d = 2, f\left(x, y\right)$ is interpreted as an image.

- For $d = 3, f\left(x, y, t\right)$ is interpreted as a video.

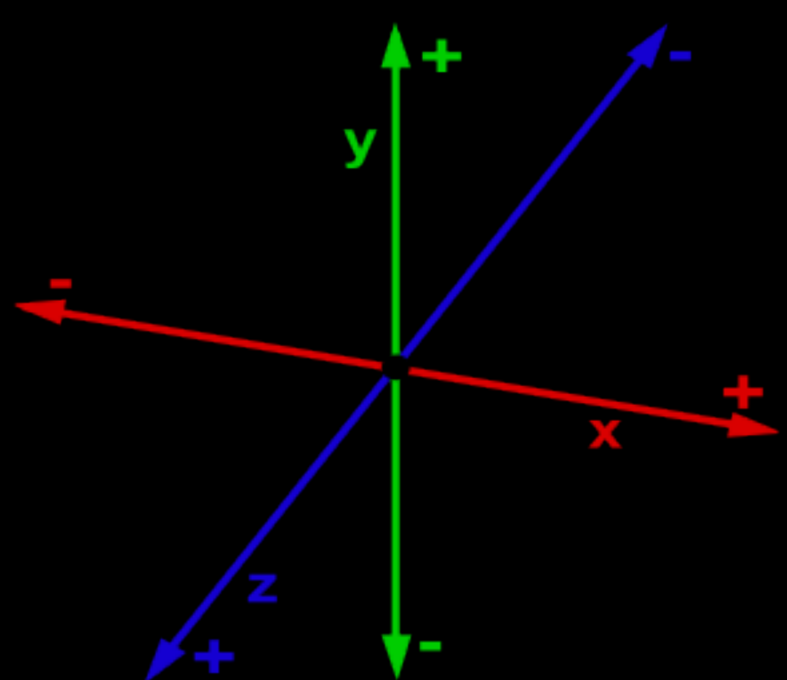
Bases

- Let $\mathscr{V}$ be a vector space. A finite set of vectors $\{x_1, \ldots, x_n\} \subseteq \mathscr{V}$ is called **linearly independent**, if the identity $\displaystyle\sum_{k=1}^{n} \alpha_k x_k = 0$ implies that $\alpha_k = 0$ for all k. Otherwise, the set is said to be **linearly dependent**.
  - $\{x_1, \ldots, x_n\} \subseteq \mathscr{V}$ are linearly dependent iff $\exists i$ s.t. $x_i \in \text{Span}\left\{x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n\right\}$.
- If a vector space $\mathscr{V}$ contains n linearly independent vectors and every $n+1$ vectors are linearly dependent, then we say that $\mathscr{V}$ has **dimension** n: $\dim \mathscr{V} = n$. If $\dim \mathscr{V} \neq n$ for every $n \in \mathbb{N}$, then we say that $\mathscr{V}$ has **infinite dimension**.

Bases

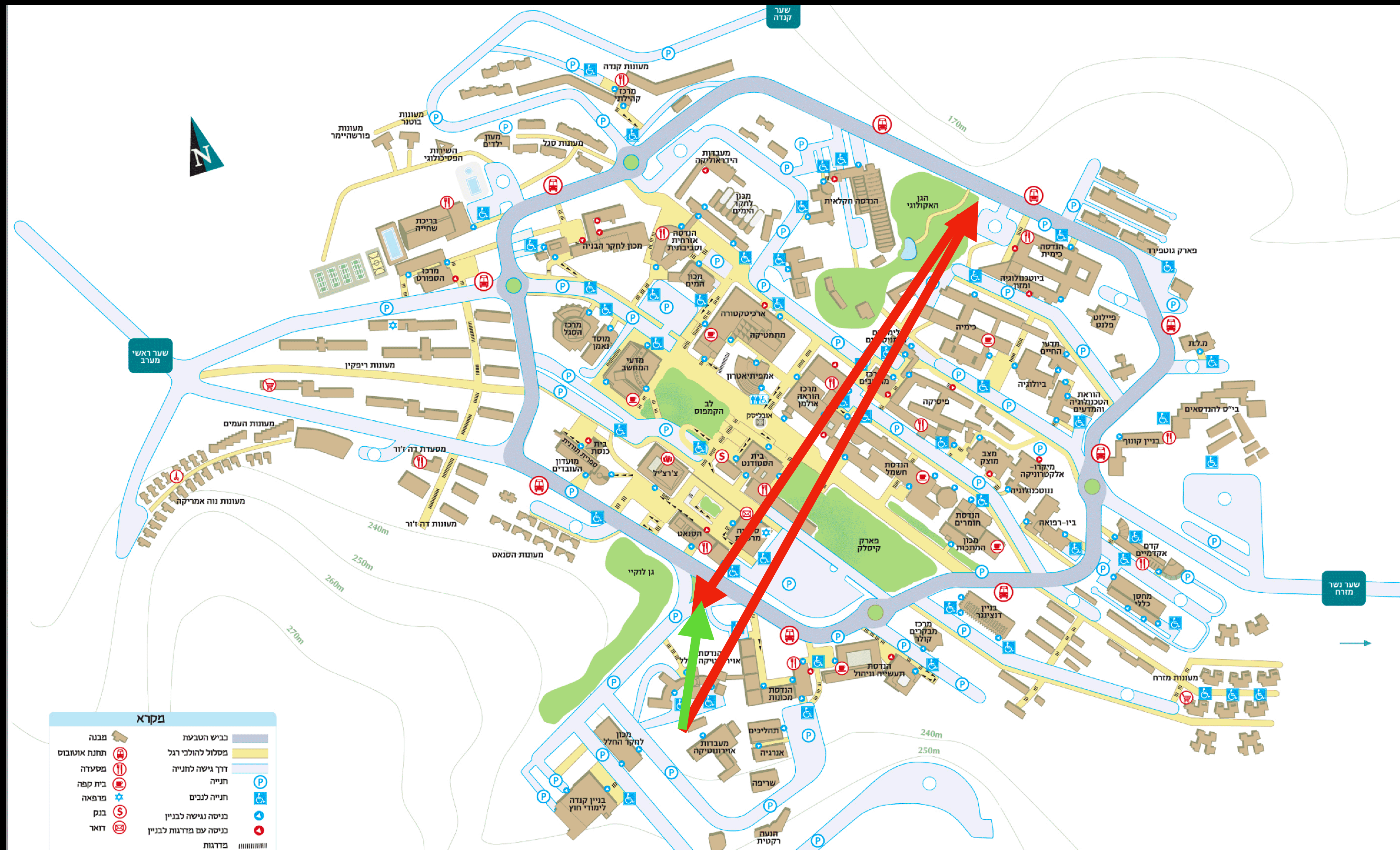
- Let $\mathscr{V}$ be a vector space. Suppose that $\dim \mathscr{V} = n$ and let $\{x_1, \ldots, x_n\} \subseteq \mathscr{V}$ be linearly independent (a **basis**). Then, every $y \in \mathscr{V}$ has a unique representation
$$y = \sum_{k=1}^{n} \alpha_k x_k.$$

- Bases give us a formal way to generalise “axis systems”.



Bases

- Examples:
  - every two non-parallel vectors in $\mathbb{R}^2 \ \Rightarrow \dim \mathbb{R}^2 = 2$.
  - $\{e_i\}_{i=1}^n$ in $\mathbb{R}^n \Rightarrow \dim \mathbb{R}^n = n$.
  - The monomials in the first sense $\{x^i\}_{i=1}^N$ are linearly independent in $\mathbb{F}(\mathbb{R}, \mathbb{R})$ for every $N \in \mathbb{N}$ $\Rightarrow \dim \mathbb{F}(\mathbb{R}, \mathbb{R}) = \infty$.



$$\alpha \nearrow + \beta \swarrow$$

Go α kms that way and β miles the other way

$$\alpha \uparrow + \beta \rightarrow$$

Go α kms straight and then β kms to the right

Normed spaces

- A **norm** over a vector space $\mathscr{V}$ is a mapping $\|\cdot\| : \mathscr{V} \to \mathbb{R}$ such that:
  - Positivity: $\|x\| \geq 0$ with equality iff $x = 0$.
  - Homogeneity: $\|\alpha x\| = |x|\|x\|$
  - Triangle inequality: $\|x + y\| \leq \|x\| + \|y\|$
- A **normed space** is a pair $\left(\mathscr{V}, \|\cdot\|\right)$, where $\mathscr{V}$ is a vector space and $\|\cdot\|$ is a norm over $\mathscr{V}$.

Normed spaces

- Norms allows us to compute the size of our vectors.
- We say $x \in \mathscr{V}$ is a **unit vector**, if $\|x\| = 1$.
- Examples:
  - $\ell_p^n = \left(\mathbb{R}^n, \|\cdot\|_p\right)$ where $\|x\|_p = \left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}}$.
  - For $p = 2$ we get the Euclidean norm $\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$.
  - $\left(\mathbb{F}\left(\mathbb{R}^n, \mathbb{R}\right), \|\cdot\|_p\right)$ where $\|f\|_p = \left(\int f(x)^p dx\right)^{\frac{1}{p}}$.

Normed spaces

- Norms allows us to compute the size of our vectors.
- We say $x \in \mathscr{V}$ is a **unit vector**, if $\|x\| = 1$.
- Examples:
  - $\ell_p^n = \left(\mathbb{R}^n, \|\cdot\|_p\right)$ where $\|x\|_p = \left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}}$.
  - For $p = 2$ we get the Euclidean norm $\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$.
  - $\left(\mathscr{L}_p^n, \|\cdot\|_p\right)$ where $\|f\|_p = \left(\int f(x)^p dx\right)^{\frac{1}{p}}$ and $\mathscr{L}_p^n = \left\{f : \mathbb{R}^n \to \mathbb{R} \mid \|f\|_p < \infty\right\}$.

Inner-product spaces

- A vector field $\mathscr{V}$ is called an **inner-product space** if there exists a product $\langle \cdot , \cdot \rangle : \mathscr{V} \times \mathscr{V} \to \mathbb{C}$, satisfying:
  - Symmetry: $\langle x, y \rangle = \overline{\langle y, x \rangle}$
  - Bilinearity: $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
  - Homogeneity: $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$
  - Positivity: $\langle x, x \rangle \geq 0$ with equality iff $x = 0$.

Inner-product spaces

- Properties:
  - $\langle x, y+z \rangle = \langle x, y \rangle + \langle x, z \rangle$
  - $\langle x, \alpha y \rangle = \overline{\alpha} \langle x, y \rangle$
  - $\| \cdot \| = \sqrt{(\cdot , \cdot)}$ is a norm.
  - **Cauchy-Schwarz inequality**: $|\langle x, y \rangle| \leq \|x\| \|y\|$.

Inner-product spaces

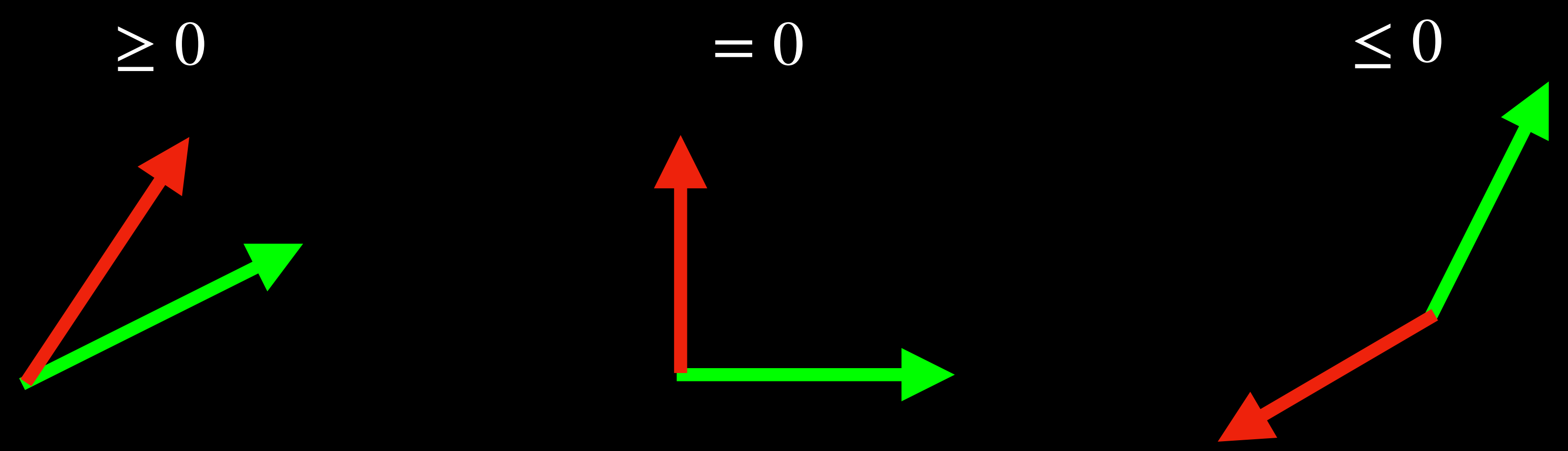
- Inner-products allows us to compute the similarity between our vectors.

- Examples:

  - The scalar product on $\mathbb{R}^n$

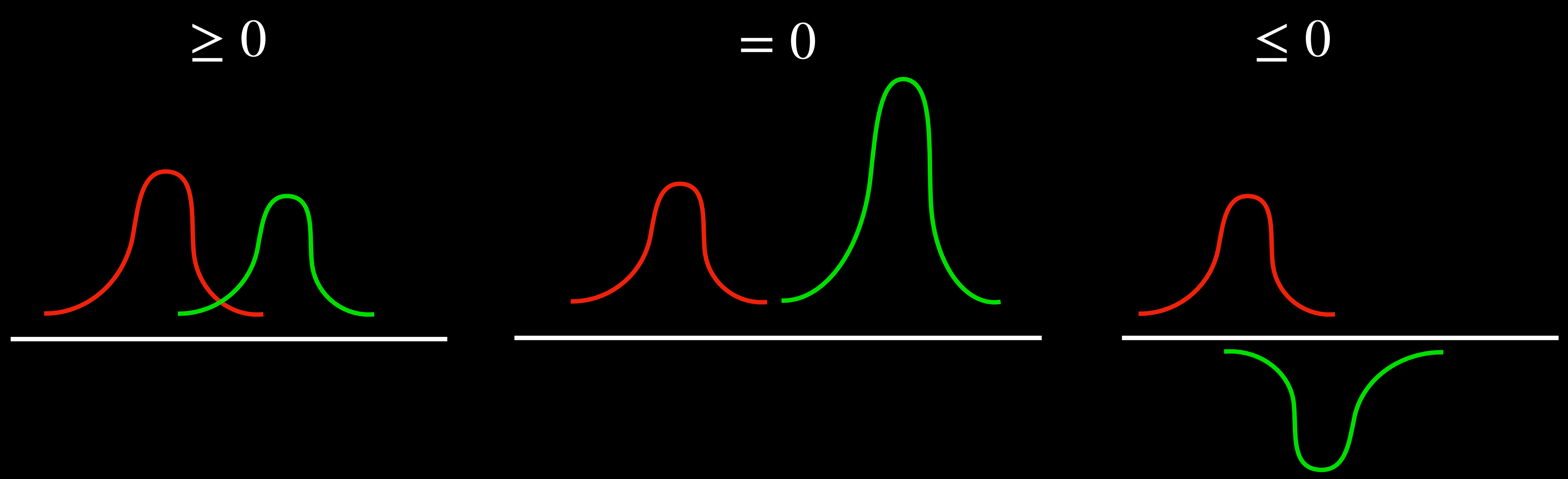
$$\langle x, y \rangle = x^T \overline{y} = \sum_{i=1}^{n} x_i \overline{y_i} = |x||y| \cos\angle(x, y)$$

induces the Euclidean norm ℓ_2^n.



Inner-product spaces

- Examples:
  - The product $\langle f, g \rangle = \int_a^b f(x)\overline{g(x)}dx$ induces the $\mathscr{L}_2^n$ norm.

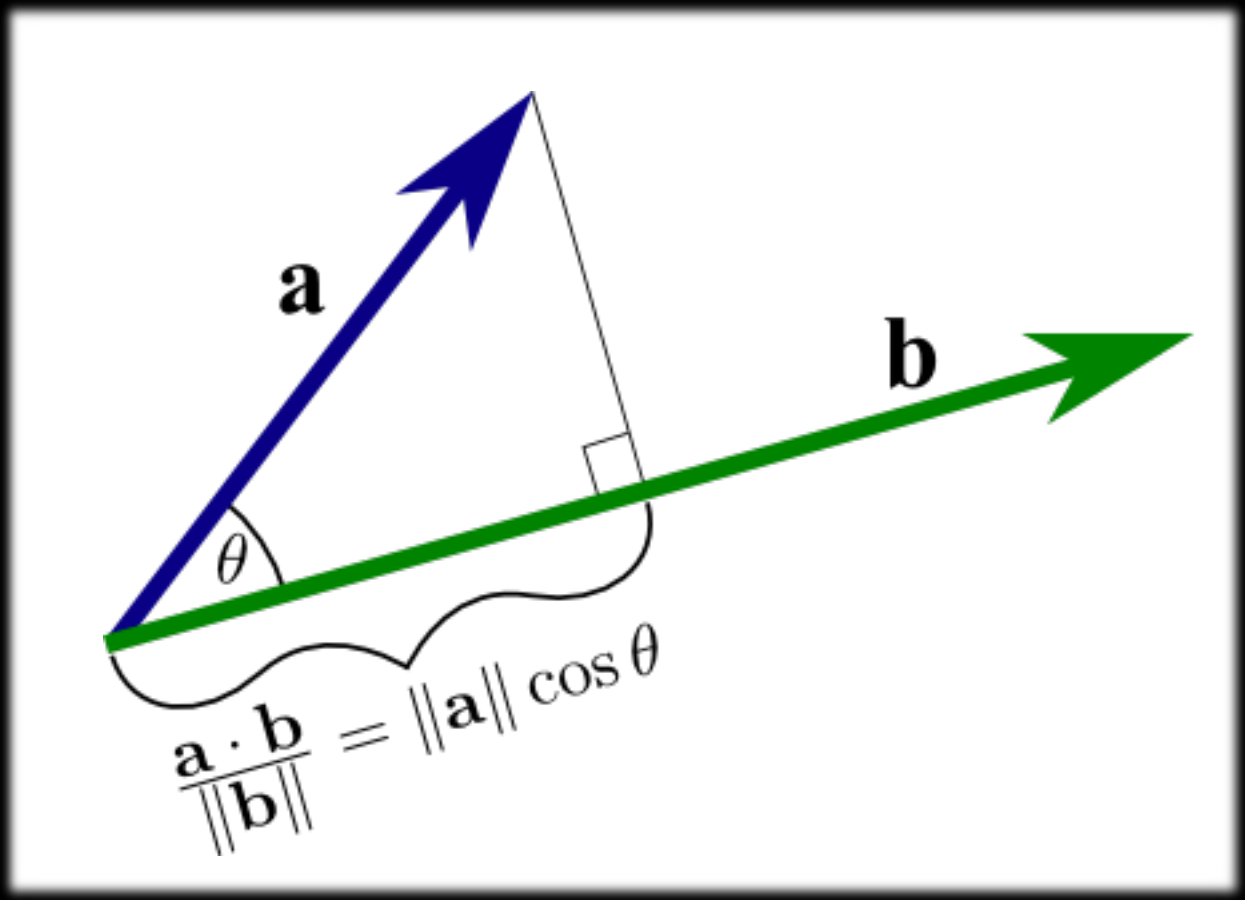


Orthonormal bases

- Let $(\mathscr{V}, \langle \cdot , \cdot \rangle)$ be an inner-product space.
- We say $x, y \in \mathscr{V}$ are orthogonal, or $x \perp y$, if $\langle x, y \rangle = 0$.
- Let A be some index set, and let $\{x_\alpha\}_{\alpha \in A} \subseteq \mathscr{V}$ be a set of vectors. The set $\{x_\alpha\}_{\alpha \in A}$ is called an **orthonormal system** if $\forall \alpha, \beta \in A : \left\langle x_\alpha, x_\beta \right\rangle = \delta_{\alpha\beta}$, where $\delta_{\alpha\beta} = 1$ if $\alpha = \beta$, and $\delta_{\alpha\beta} = 0$ otherwise.
  - $\Leftrightarrow \{x_\alpha\}_{\alpha \in A}$ are orthogonal unit vectors.
- **Gram-Schmidt orthonormalization:** Let (x_n) be either a finite or a countable sequence of linearly independent vectors in an inner-product space $\mathscr{V}$. Then it is possible to construct an orthonormal sequence (y_n) that has the same cardinality as the sequence (x_n), such that: $\forall n \in \mathbb{N}$: Span $\left\{y_k \,|\, 1 \leq k \leq n\right\}$ = Span $\left\{x_k \,|\, 1 \leq k \leq n\right\}$.

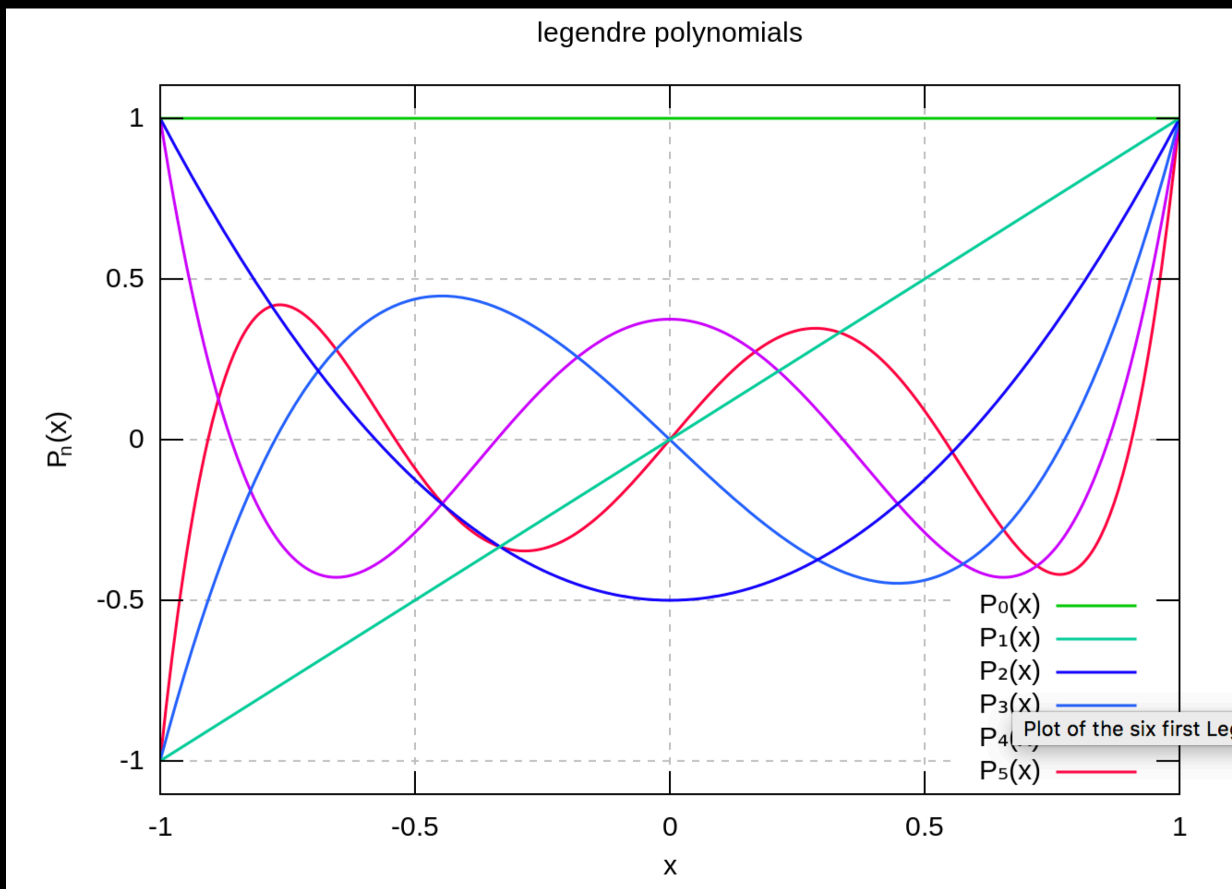
Orthonormal bases

- If $\{x_i\}_{i=1}^{n}$ is an orthonormal basis of $\mathscr{V}$, then $\forall y \in \mathscr{V}$:
  $$y = \sum_{i=1}^{n} \langle y, x_i \rangle \, x_i.$$



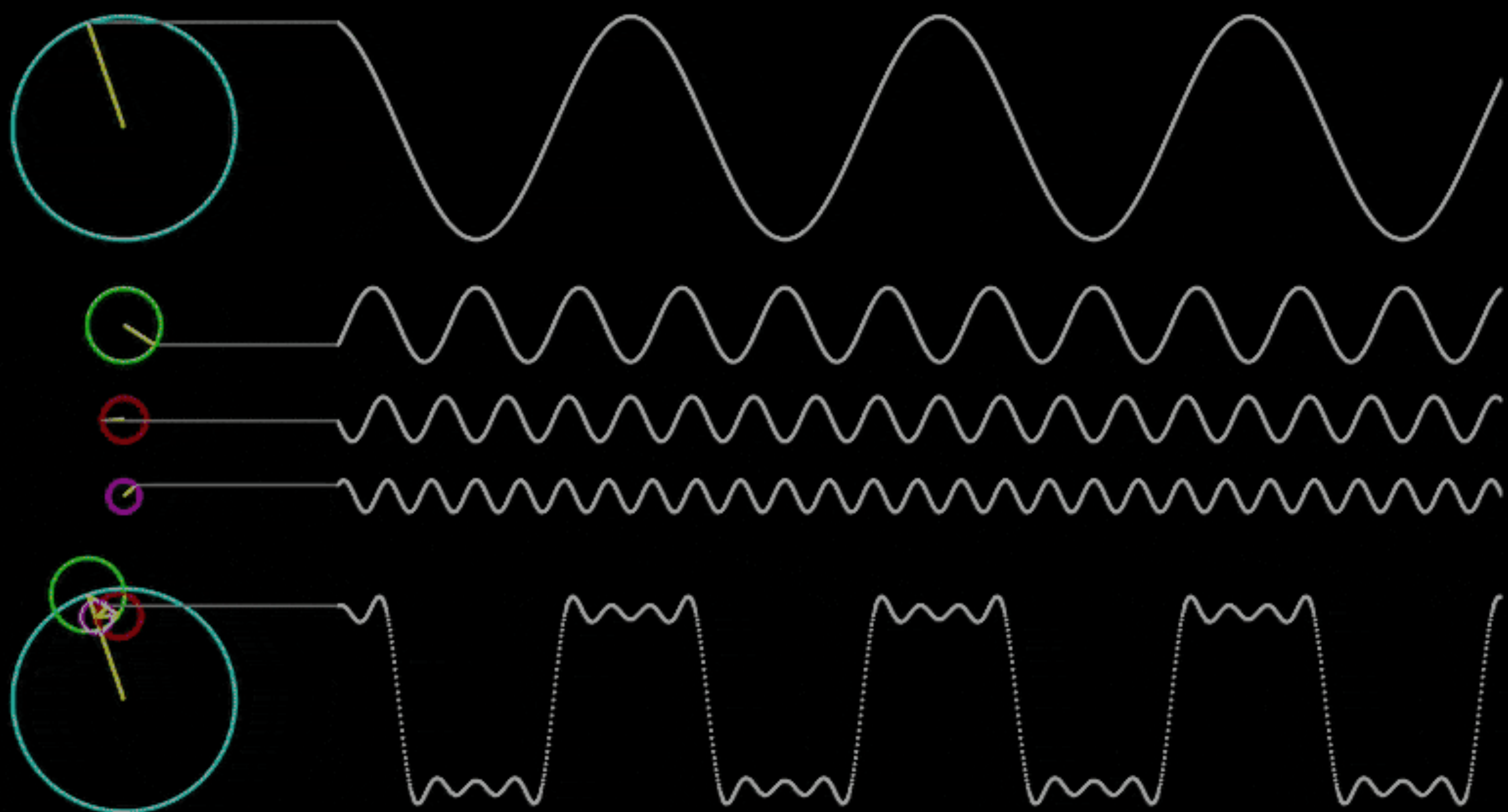
Basis for L_2

- We have seen that if $\left\{x_i\right\}_{i=1}^{n}$ is an orthonormal basis of $\mathscr{V}$, then $\forall y \in \mathscr{V}$: $y = \sum_{i=1}^{n} \left\langle y, x_i \right\rangle x_i$.
- It can be proven that there is an orthonormal "basis" $\{g_i\}_{i=1}^{\infty}$ for $\mathscr{L}_2$ such that $\forall f \in \mathscr{L}_2$: $f \approx \sum_{i=1}^{\infty} \left\langle f, g_i \right\rangle g_i$.



n	$P_n(x)$
0	1
1	x
2	$\frac{1}{2}\left(3x^2-1\right)$
3	$\frac{1}{2}\left(5x^3-3x\right)$
4	$\frac{1}{8}\left(35x^4-30x^2+3\right)$
5	$\frac{1}{8}\left(63x^5-70x^3+15x\right)$
6	$\frac{1}{16}\left(231x^6-315x^4+105x^2-5\right)$
7	$\frac{1}{16}\left(429x^7-693x^5+315x^3-35x\right)$
8	$\frac{1}{128}\left(6435x^8-12012x^6+6930x^4-1260x^2+35\right)$
9	$\frac{1}{128}\left(12155x^9-25740x^7+18018x^5-4620x^3+315x\right)$
10	$\frac{1}{256}\left(46189x^{10}-109395x^8+90090x^6-30030x^4+3465x^2-63\right)$

Example: Legendre polynomials



Example: Fourier

Linear operators

- Let $\mathscr{V}$ and $\mathscr{Y}$ be vector spaces over the same field $\mathscr{F}$. A mapping $T : \mathscr{V} \to \mathscr{Y}$ is said to be a **linear transformation** if for all $x_1, x_2 \in \mathscr{V}$ and $\alpha_1, \alpha_2 \in \mathscr{F}$:
$T\left(\alpha_1 x_1 + \alpha_2 x_2\right) = \alpha_1 T\left(x_1\right) + \alpha_2 T\left(x_2\right)$.

- A linear transformation $T : \mathscr{V} \to \mathscr{V}$ is called a **linear operator**.

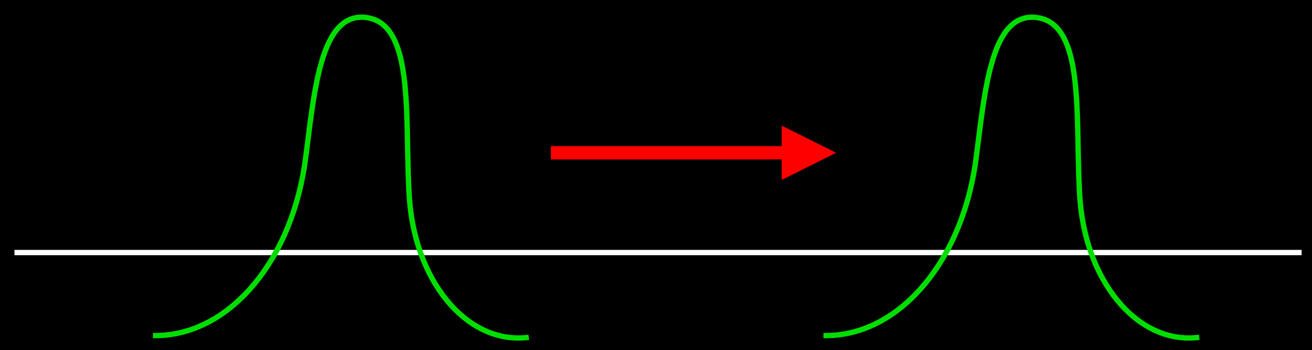
- All linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ can be represented by a $\mathbb{R}^{m\times n}$ matrix.

Linear operators

- $\mathscr{D} : \mathscr{L}_2 \to \mathscr{L}_2$ where $\mathscr{D}f = f'$.
  - Proof: $\mathscr{D}\left(\alpha f + \beta g\right) = \alpha f' + \beta g' = \alpha \mathscr{D}f + \beta \mathscr{D}g$.

Linear operators

- $\tau_p : \mathscr{L}_2 \to \mathscr{L}_2$ where $\tau_p f(x) = f(x - p)$.
   - Proof: $\Big(\tau_p\left(\alpha f + \beta g\right)\Big)(x) = \left(\alpha f + \beta g\right)\left(x - p\right)$
   $= \alpha f\left(x - p\right) + \beta g\left(x - p\right) = \alpha \tau_p\left(f\right) + \beta \tau_p\left(g\right)$.



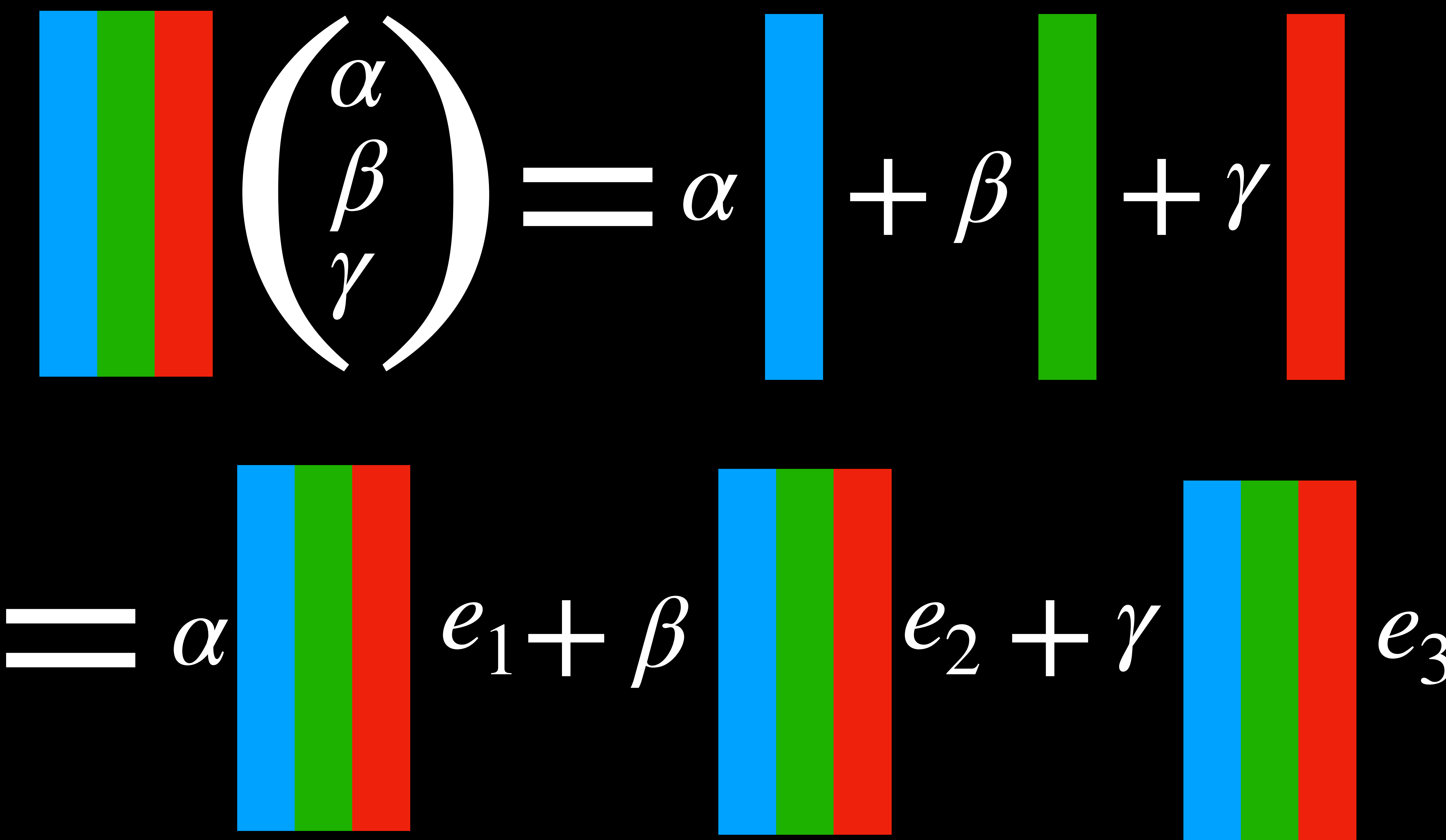
Linear functionals

- A linear transformation $T : \mathscr{V} \to \mathscr{F}$ is called a **linear functional**.
- Examples:
  - Given $y \in \mathscr{V}$: $T_y(x) = \langle y, x \rangle$.
    - If $\dim \mathscr{V} < \infty$, these are the only functionals.
    - In this case there is an isomorphism between functionals and vectors.

Linear functionals

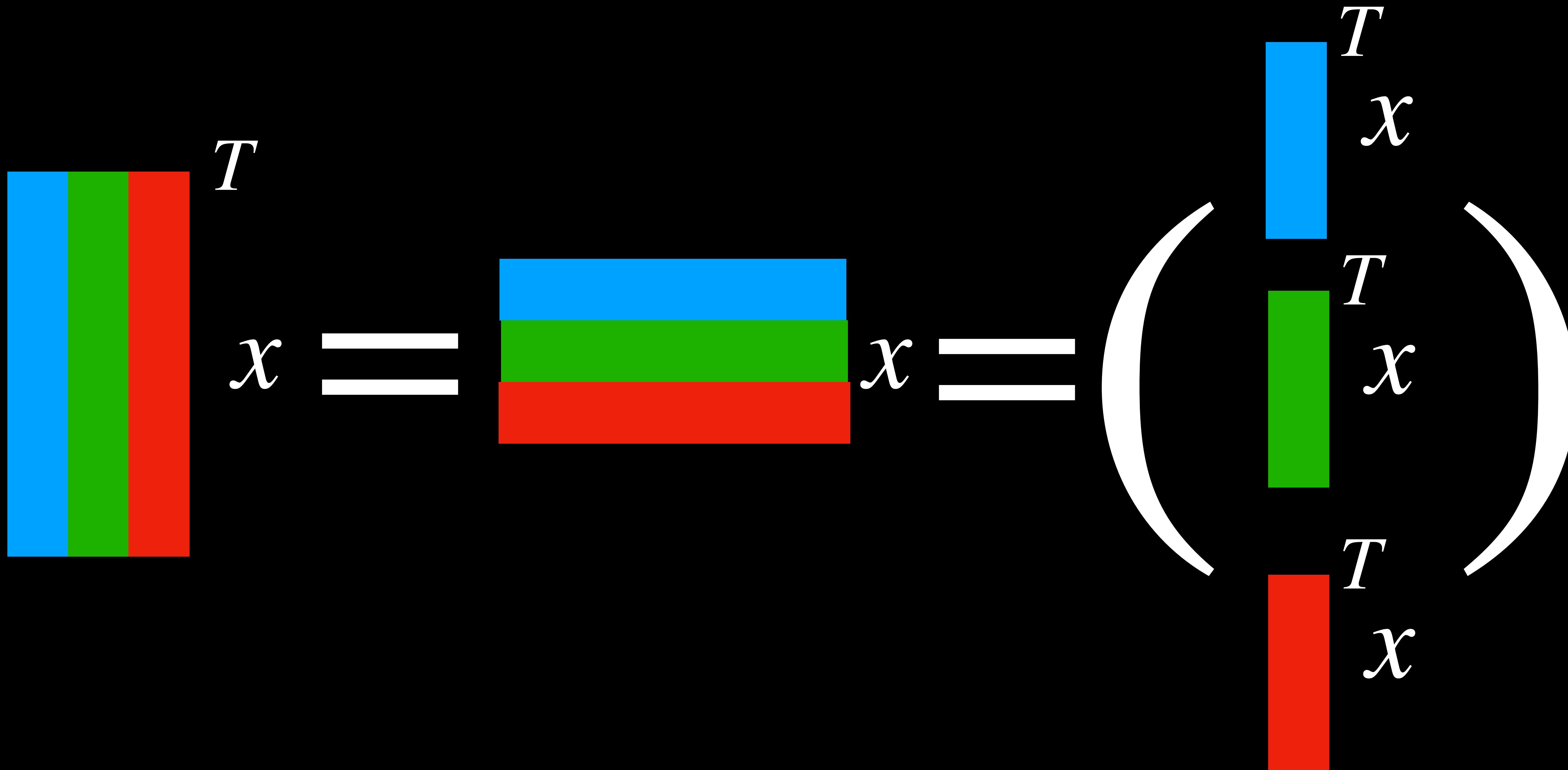
- Examples:
  - $\delta : \mathscr{L}_2 \to \mathbb{R}$ where $\delta(f) = f(0)$.
    - There is no $g \in \mathscr{L}_2$ s.t. $\delta = \langle g, \cdot \rangle$.
    - We will sometimes write $\delta(f) = \langle \delta, f \rangle = \displaystyle\int \delta(x) f(x) dx$.
  - $\delta\tau_p : \mathscr{L}_2 \to \mathbb{R}$ where
    $\delta\tau_p \left(f(x)\right) = \delta \left(f\left(x - p\right)\right) = f\left(0 - p\right) = f\left(-p\right)$.

Kernels



$\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \alpha + \beta + \gamma$
$= \alpha \quad e_1 + \beta \quad e_2 + \gamma \quad e_3$

Kernels



Kernels

- Let $\mathscr{H} : \mathscr{L}_2 \to \mathscr{L}_2$ be a linear operator, and let $\{g_i\}_{i=1}^{\infty}$ be a basis for $\mathscr{L}_2$.
- For each $f \in \mathscr{L}_2$, we can write $f \approx \sum_{i=1}^{\infty} \langle f, g_i \rangle g_i$.
- Then $\mathscr{H} f \approx \mathscr{H} \left(\sum_{i=1}^{\infty} \langle f, g_i \rangle g_i \right) = \sum_{i=1}^{\infty} \langle f, g_i \rangle \mathscr{H} g_i$.
- This is analogous to multiplying a matrix by a vector: $Hu = \sum_{j=1}^{n} u_j H e_j$.

Kernels

- Hence

$$\mathscr{H} f(x) = \sum_{i=1}^{\infty} \langle f, g_i \rangle \mathscr{H} g_i(x) = \sum_{i=1}^{\infty} \left(\int f(y)\, g_i(y)\, dy \right) \mathscr{H} g_i(x)$$
$$= \int \left(\sum_{i=1}^{\infty} \mathscr{H} g_i(x)\, g_i(y) \right) f(y)\, dy.$$

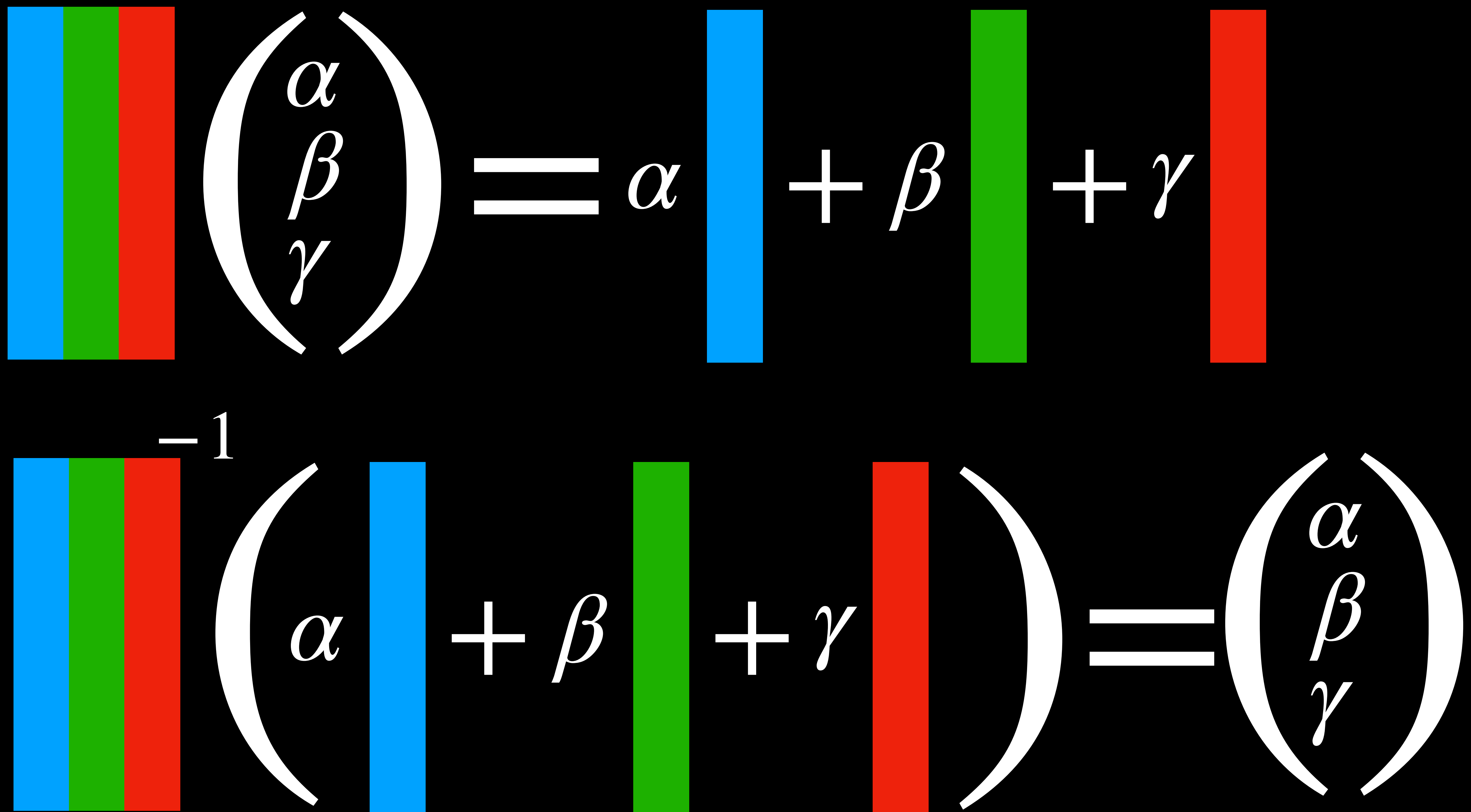
- Denote $h(x, y) = \sum_{i=1}^{\infty} \mathscr{H} g_i(x)\, g_i(y)$.

- Then $\mathscr{H} f(x) = \int h(x, y)\, f(y)\, dy$.

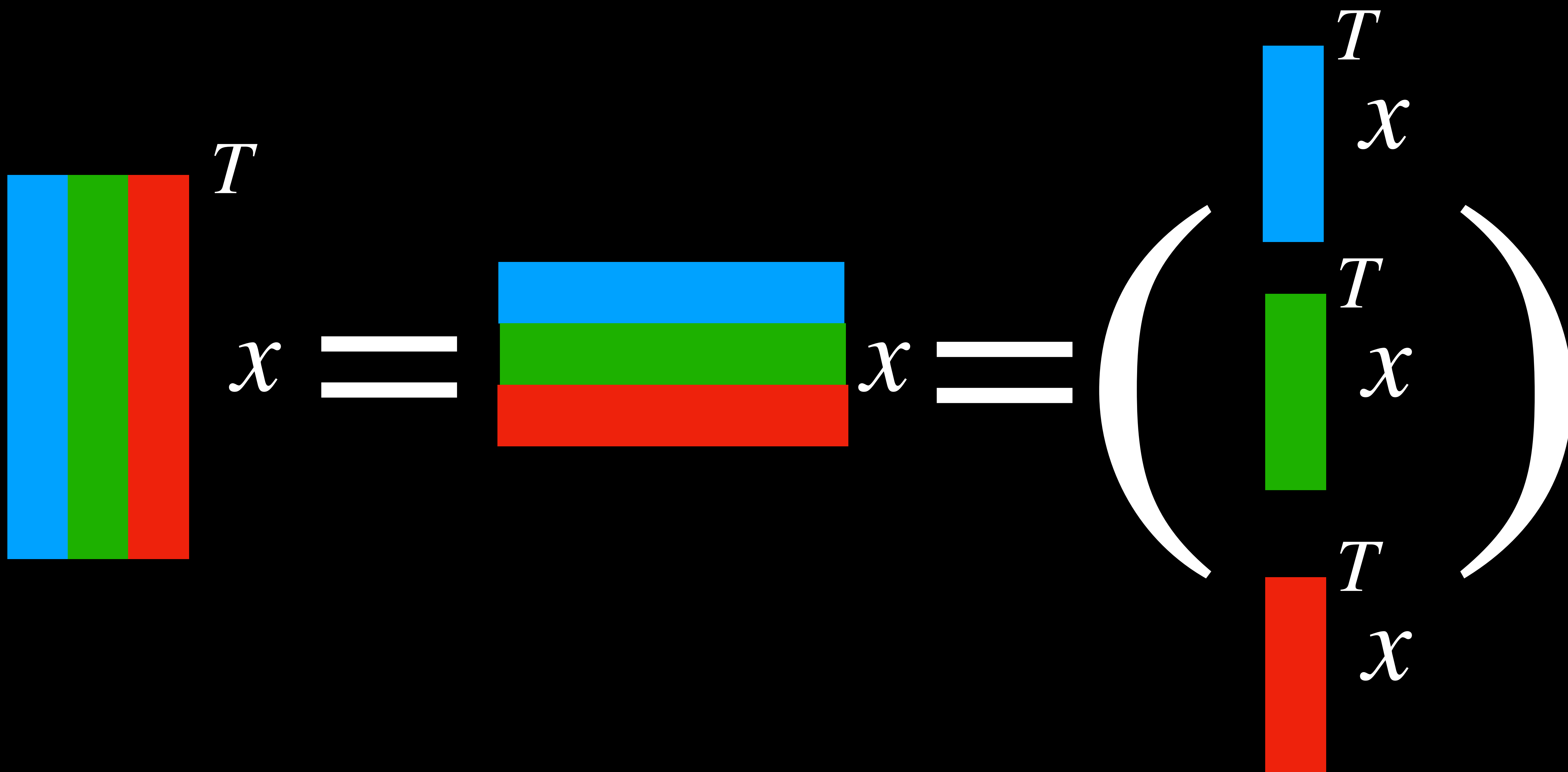
Kernels

- $\mathscr{H}f(x) = \int h\left(x, y\right) f\left(y\right)\, dy = \langle h(x, \cdot\,), f \rangle.$
- h is called the kernel, or the impulse response of $\mathscr{H}$.
- This is analogous to multiplying a matrix by a vector: $(Hu)_i = \sum_{j=1}^{n} h_{ij} u_j = \left\langle h_i, u \right\rangle.$

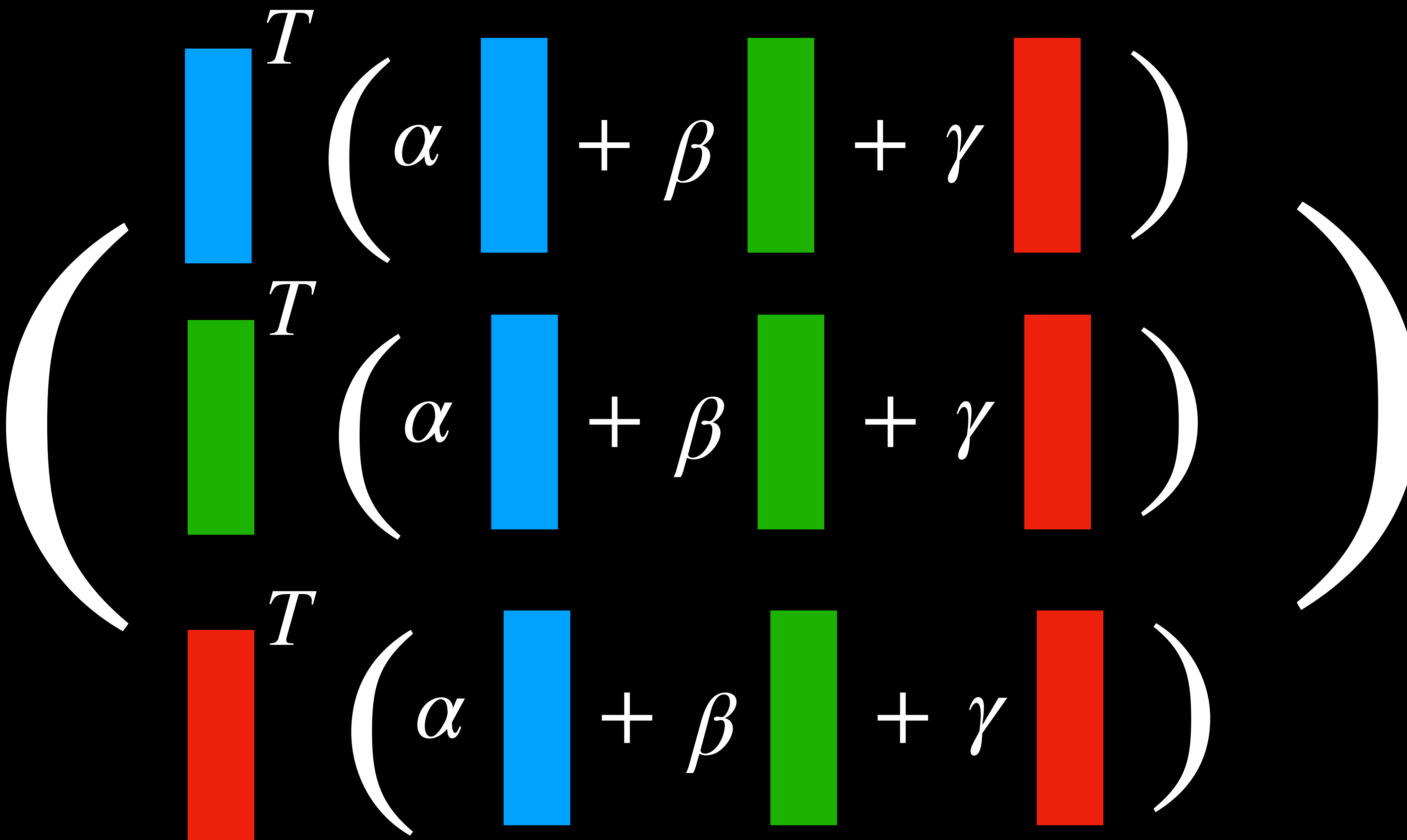
Coordinate systems



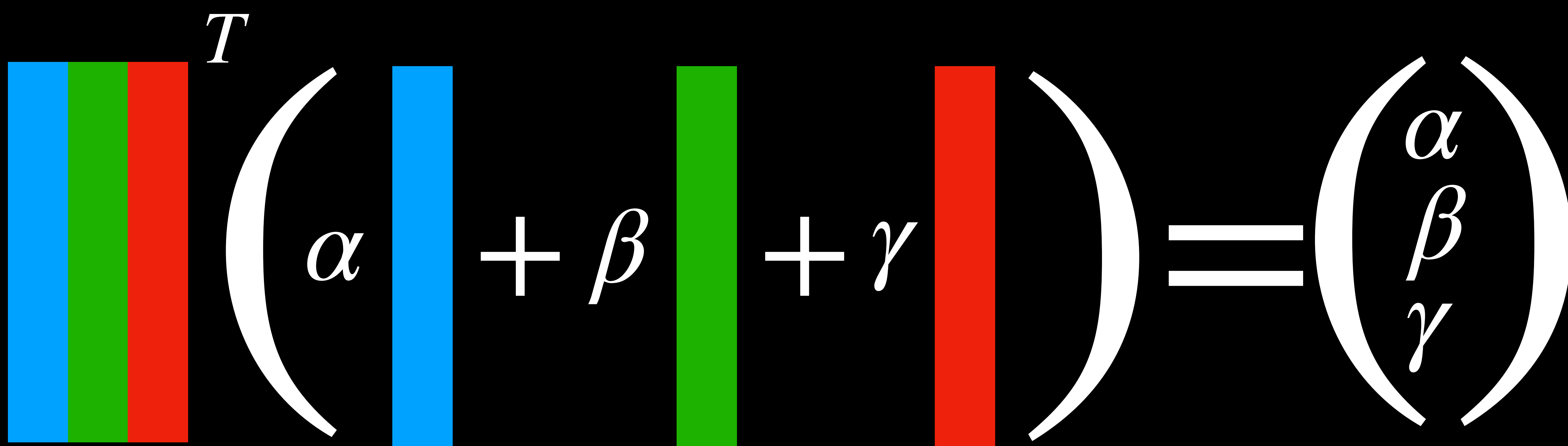
Orthonormal coordinates



Orthonormal coordinates



Orthonormal coordinates



Orthonormal matrices

- A matrix $U \in \mathbb{R}^{n \times n}$ is called **orthonormal**, if its columns are made of orthonormal vectors.

- Properties:

  - The set of columns and the set of rows of U are orthonormal bases.

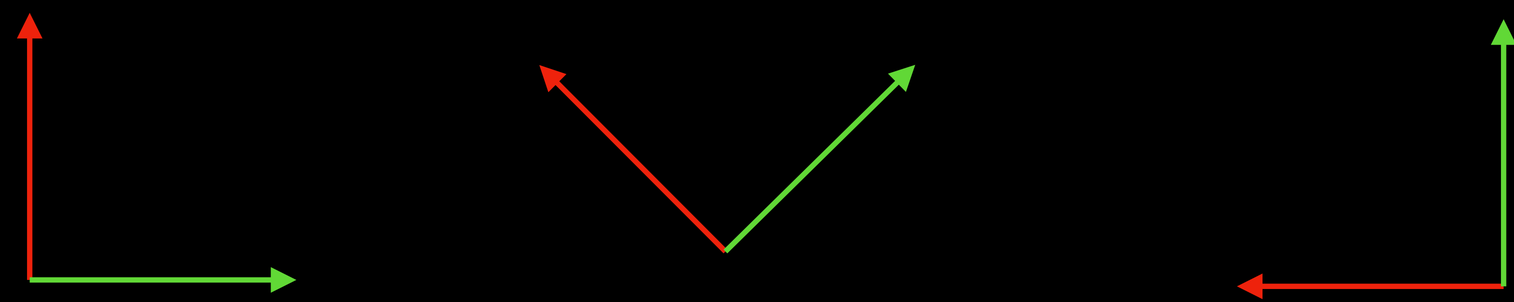
  - $U^{-1} = U^T$.

  - The vector Ux is the vector x converted from the U-coordinate system to the natural coordinate system.

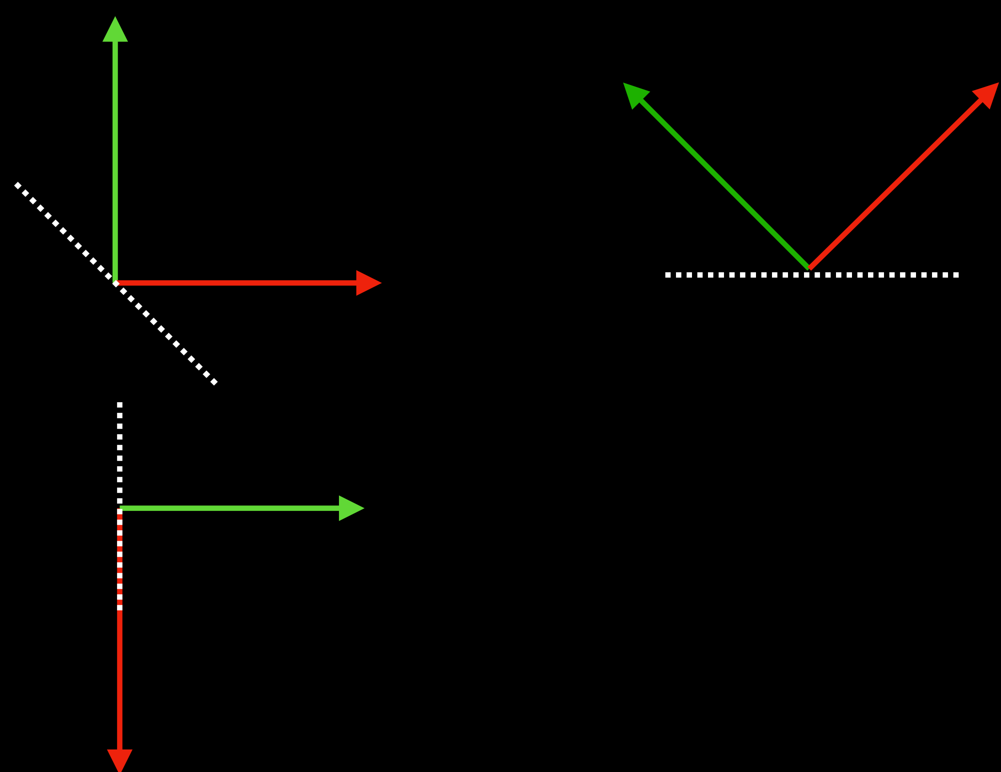
  - The vector $U^T x$ is the vector x converted from the natural coordinate system to the U-coordinate system.

  - $\langle Ux, Uy \rangle = \langle x, y \rangle$, and thus $\|Ux\| = \|x\|$.

Orthonormal matrices



Rotations



Reflections

Unitary matrices

- A matrix $U \in \mathbb{C}^{n\times n}$ is called **unitary**, if its columns are made of orthonormal vectors.
- Properties:
    - The set of columns and the set of rows of U are orthonormal bases.
    - $U^{-1} = U^*$.
    - The vector Ux is the vector x converted from the U-coordinate system to the natural coordinate system.
    - The vector U^*x is the vector x converted from the natural coordinate system to the U-coordinate system.
    - $\langle Ux, Uy\rangle = \langle x, y\rangle$, and thus $\|Ux\| = \|x\|$.

Unitary transformations

- Generally, a linear transformation $U : \mathscr{V} \to \mathscr{Y}$ is called **unitary** if $U^{-1} = U*$.
- How do we define $U*$ in $\mathscr{L}_2$? We will found out in the lecture.
- In the meantime think - is τ_p unitary?

Spectral decomposition

- An **eigenvector** of A is a non-zero vector $u \neq 0$ satisfying $Au = \lambda u$, with the scalar λ called an **eigenvalue**.

- The collection of eigenvalues is called the **spectrum** of a matrix.

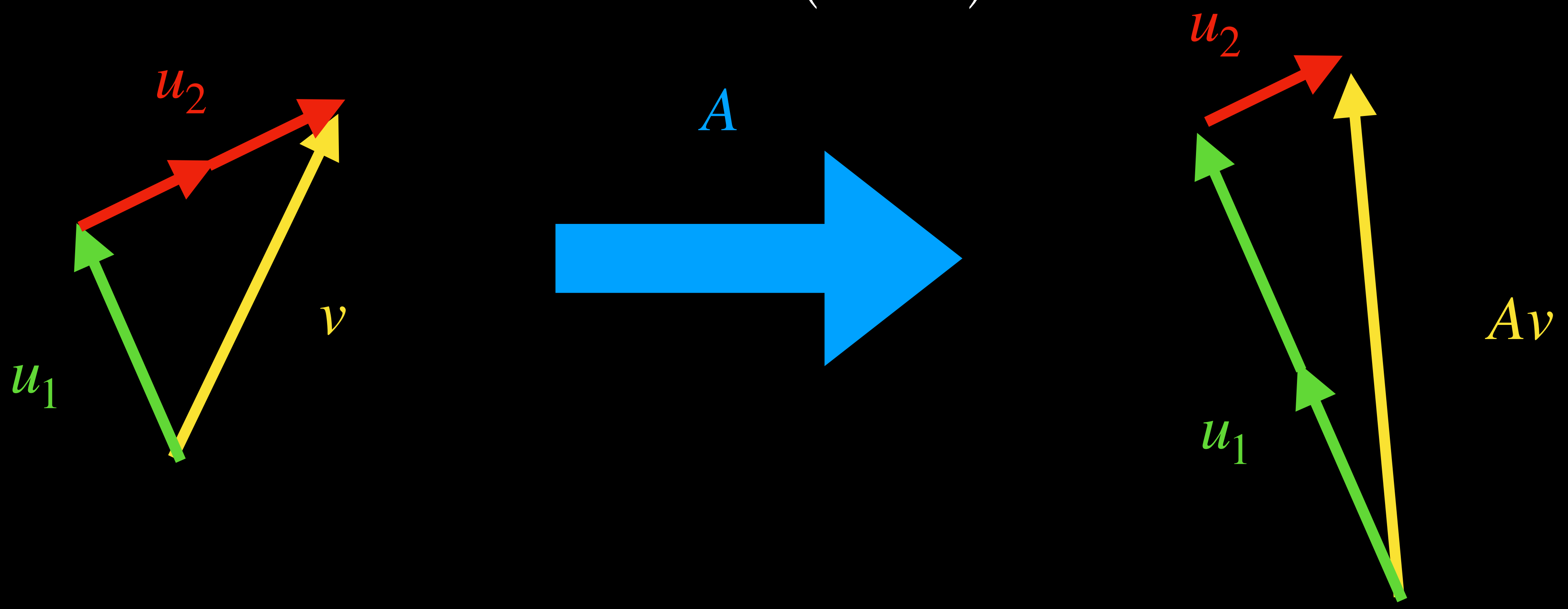
- For an $n \times n$ matrix A with n linearly independent eigenvectors, we can write $Au_i = \lambda_i u_i$ for each $i = 1,\ldots,n$, or $AU = U\Lambda$ in matrix form. This leads to the decomposition $A = U\Lambda U^{-1}$.

Spectral decomposition

- Let us examine how A operates on a vector $v \in \mathbb{R}^n$.
- Since $\{u\}_{i=1}^{n}$ are a set of n linearly independent vectors, they form a basis, meaning $v = \sum_{i=1}^{n} \alpha_i u_i$ for some $\left\{\alpha_i\right\}_{i=1}^{n} \subseteq \mathbb{R}$.
- Then $Av = U\Lambda U^{-1} \sum_{i=1}^{n} \alpha_i u_i = U\Lambda\alpha = U(\lambda \odot \alpha)$, where λ, α are the vectors containing the appropriate scalar entries, and $\odot$ is the point-wise product.

Spectral decomposition

$$\Lambda = \textbf{diag}\left(2, 0.5\right)$$



Spectral decomposition

- If the matrix A is symmetric, it can be shown that there is an orthonormal basis of eigenvectors. Thus, the spectral decomposition becomes $A = U\Lambda U^T$.

$$\Lambda = \textbf{diag}\left(3,0.5\right)$$