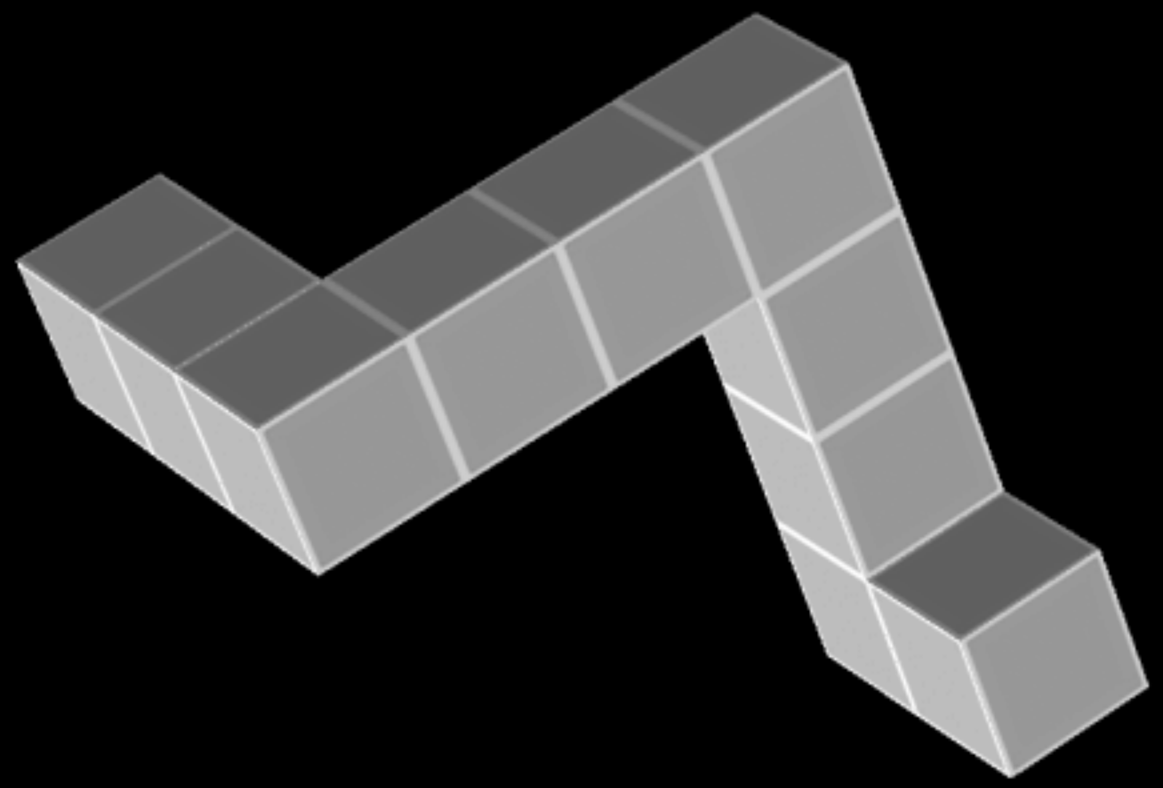
Computing Perceptual Invariants

Sophia Sanborn
VS265 Guest Lecture
Fall 2020

Outline

1. The problem of invariance
2. The mathematics of invariance
3. Invariance in visual cortex
4. Invariance in convolutional neural networks

The Problem of Invariance



The Problem of Invariance

The Problem of Invariance

The Problem of Invariance

VOL. IX. OCTOBER, 1898. No. 1.

THE MONIST

Henri Poincaré

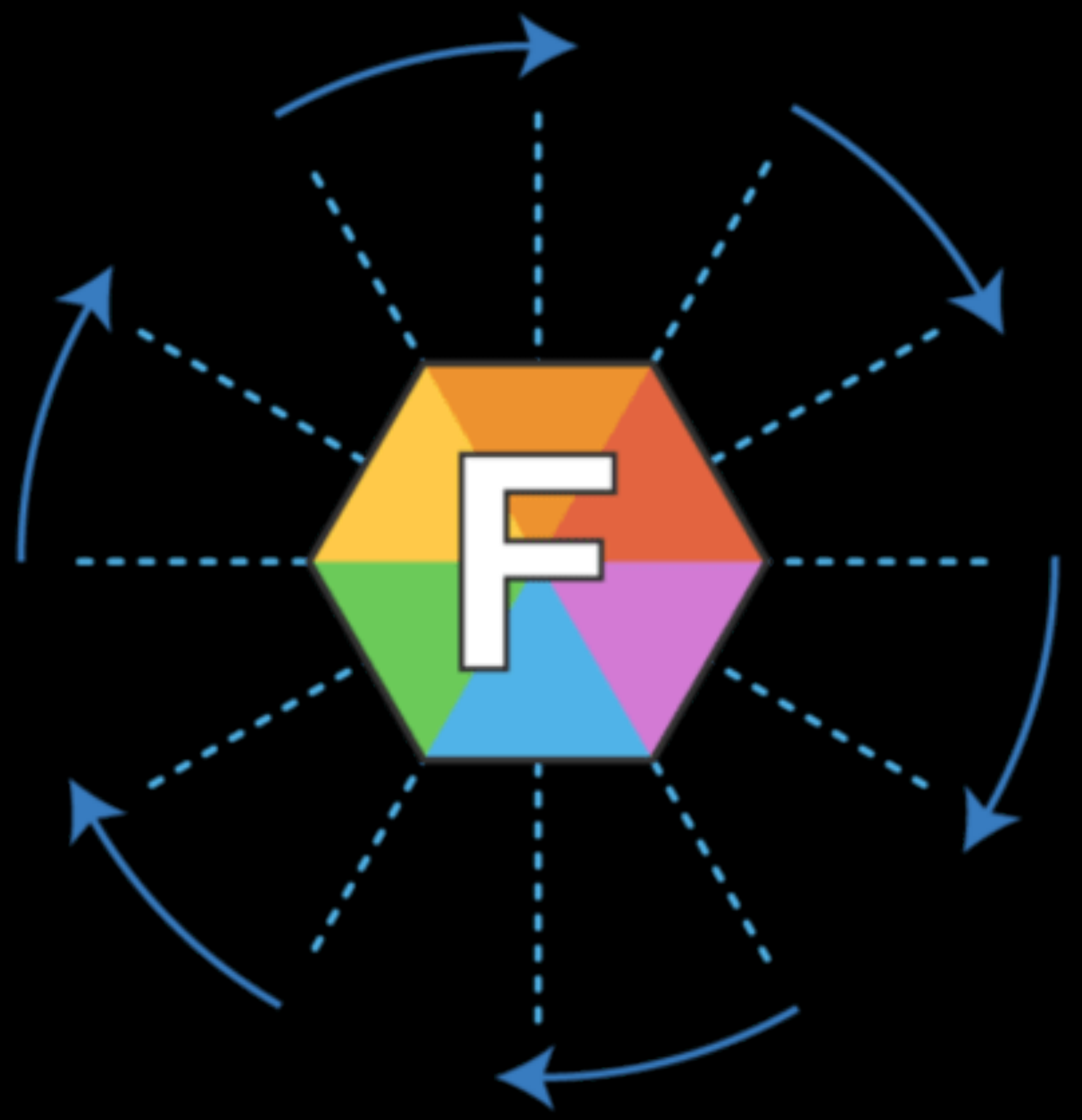
ON THE FOUNDATIONS OF GEOMETRY.[1]

ALTHOUGH I have already had occasion to set forth my views on the foundations of geometry,[1] it will not, perhaps, be unprofitable to revert to the question with new and ampler developments, and seek to clear up certain points which the reader may have found obscure. It is with reference to the definition of the point and the determination of the number of dimensions that new light appears to me most needed; but I deem it opportune, nevertheless, to take up the question from the beginning.

SENSIBLE SPACE.

Our sensations cannot give us the notion of space. That notion is built up by the mind from elements which pre-exist in it, and external experience is simply the occasion for its exercising this power, or at most a means of determining the best mode of exercising it.

The Mathematics of Invariance



The Mathematics of Invariance

The Mathematics of Invariance

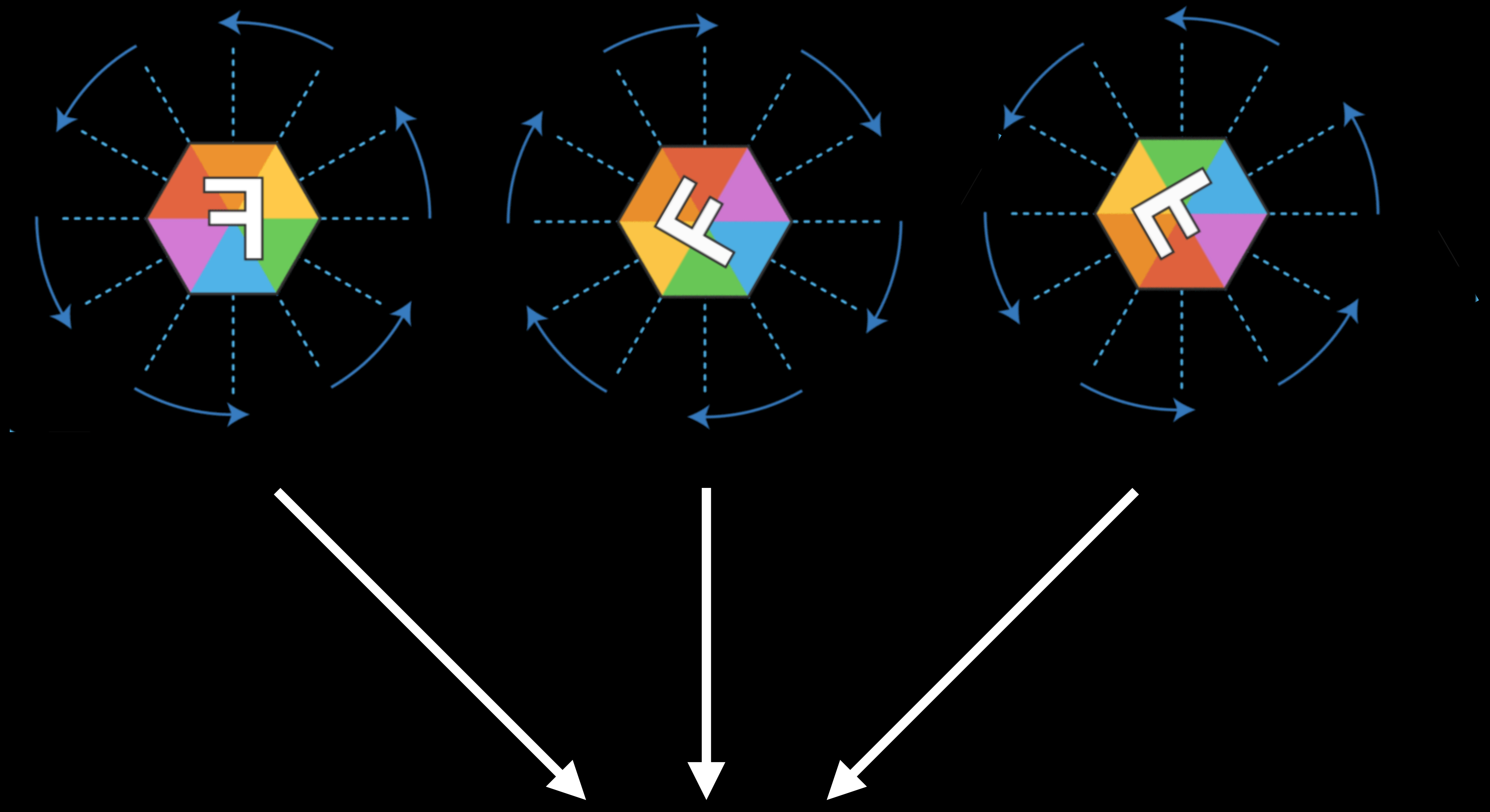
Invariance

$$f(\phi(x)) = f(x)$$

The Mathematics of Invariance

Invariance

$$f(\phi(x)) = f(x)$$



F

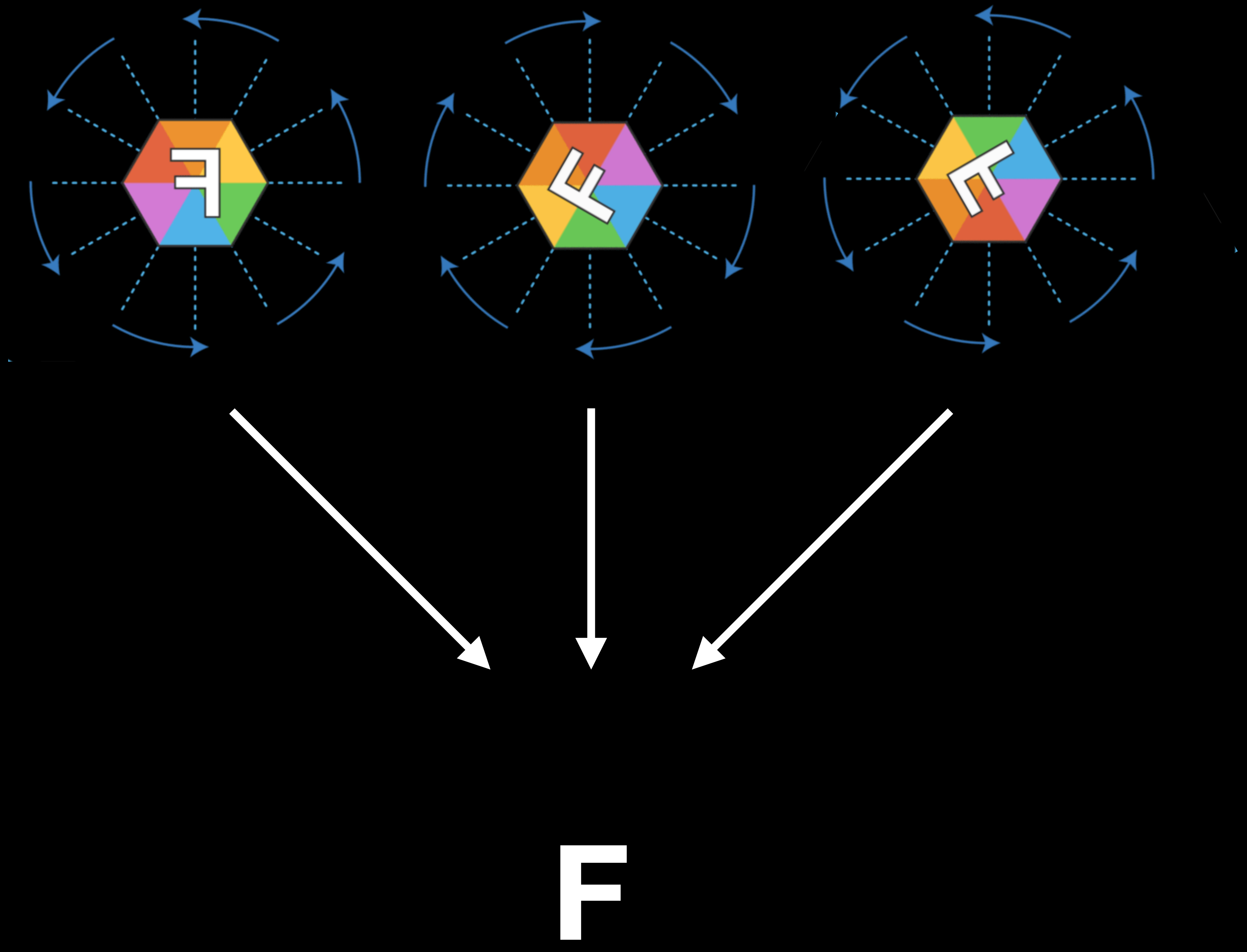
The Mathematics of Invariance

Invariance

$$f(\phi(x)) = f(x)$$

Equivariance

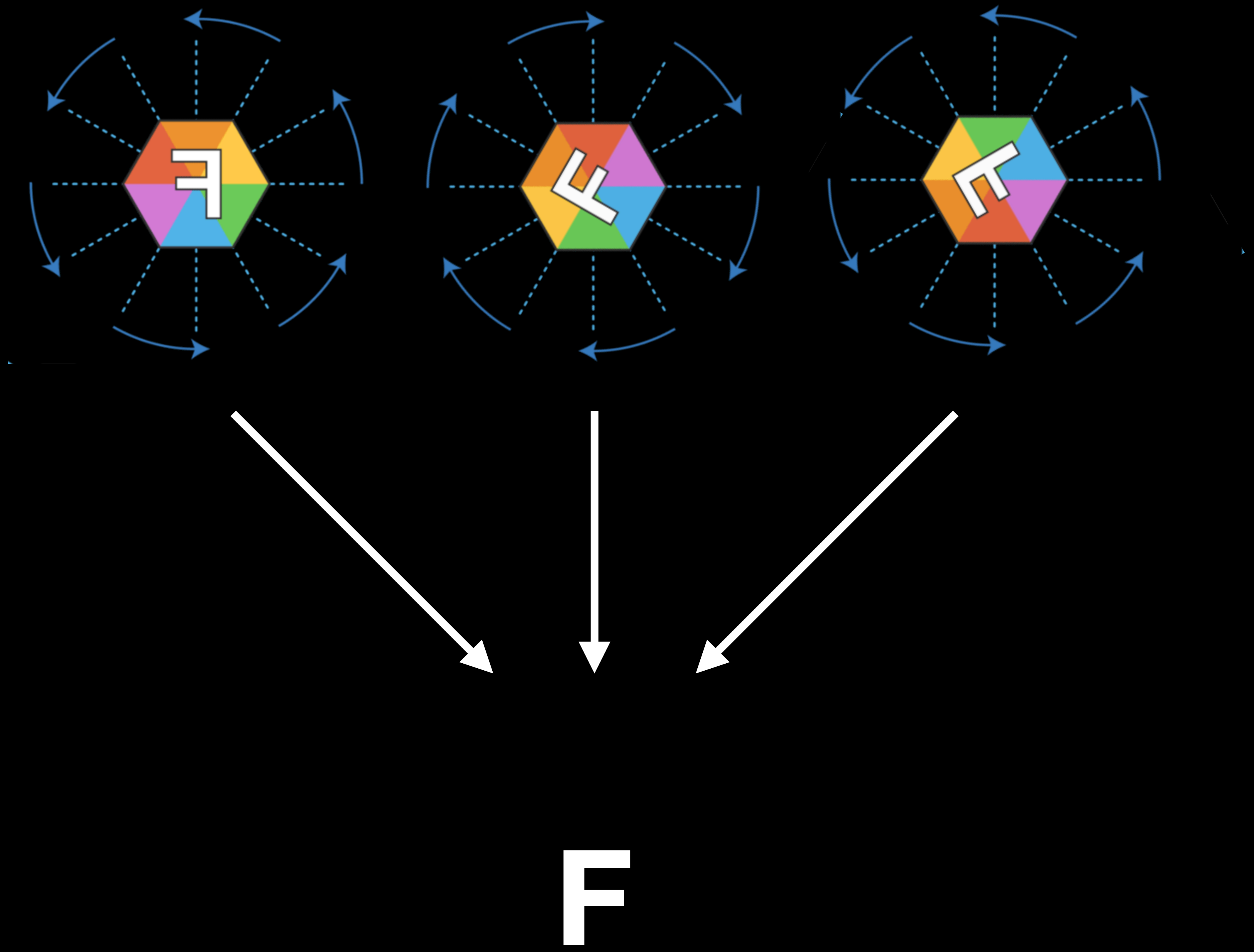
$$f(\phi(x)) = \phi'(f(x))$$



The Mathematics of Invariance

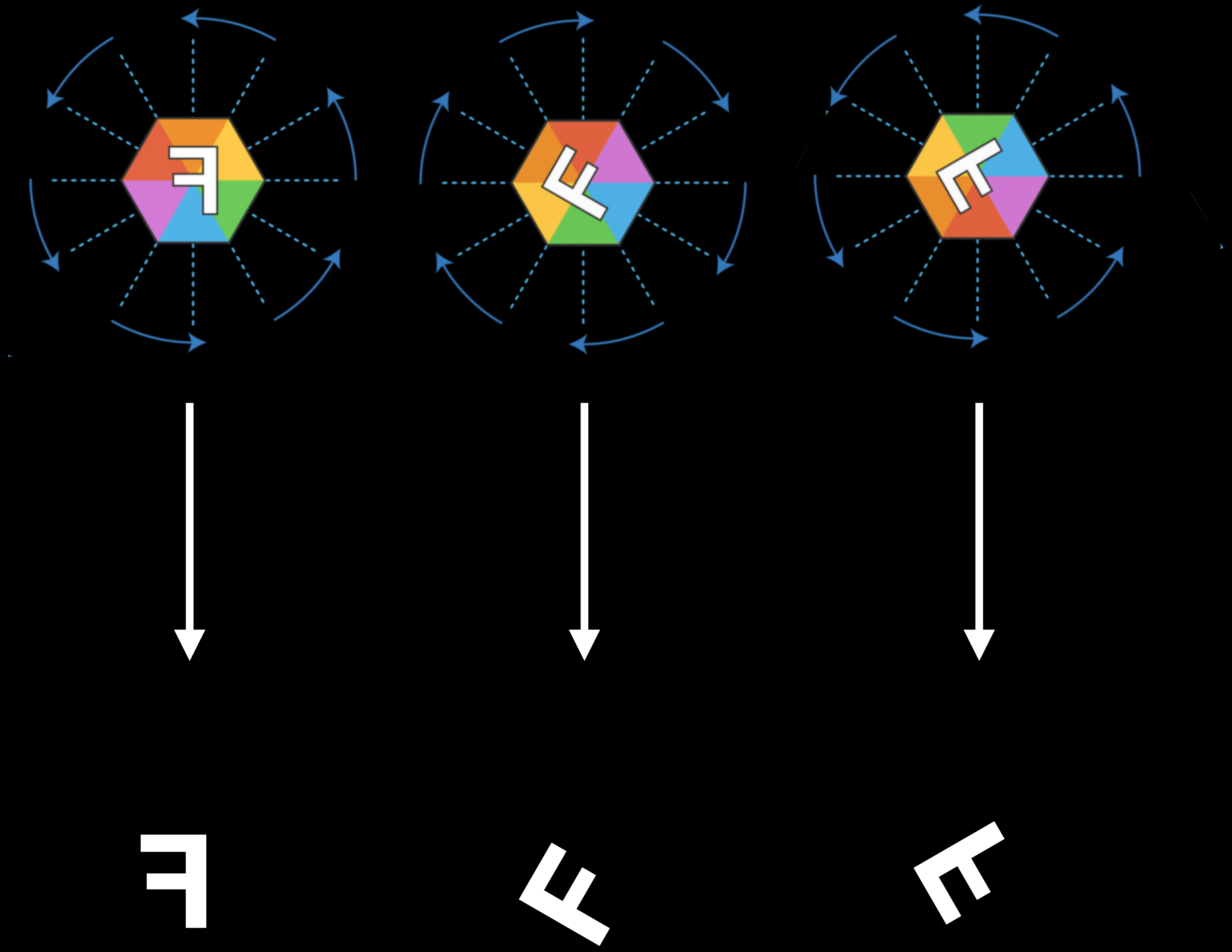
Invariance

$$f(\phi(x)) = f(x)$$



Equivariance

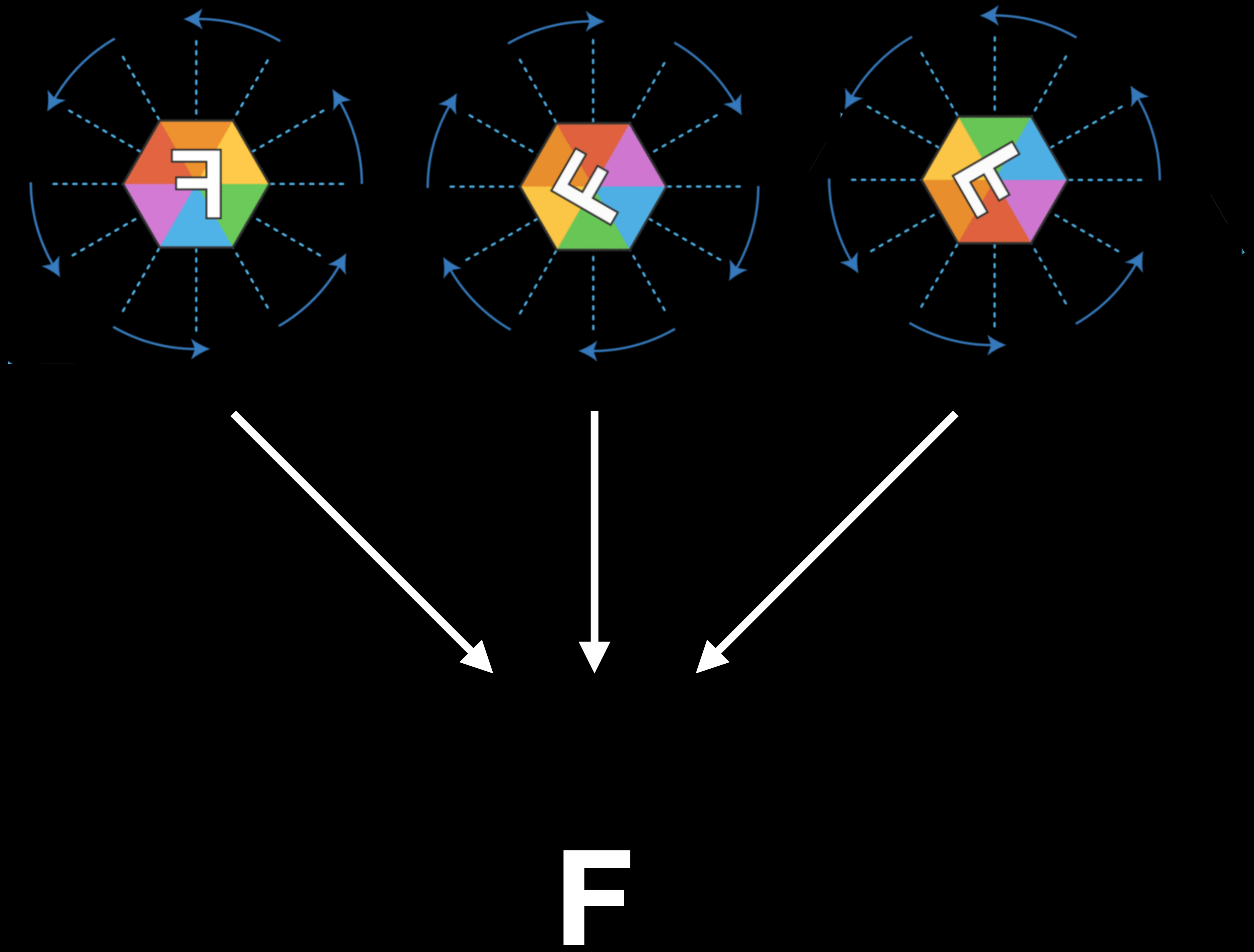
$$f(\phi(x)) = \phi'(f(x))$$



The Mathematics of Invariance

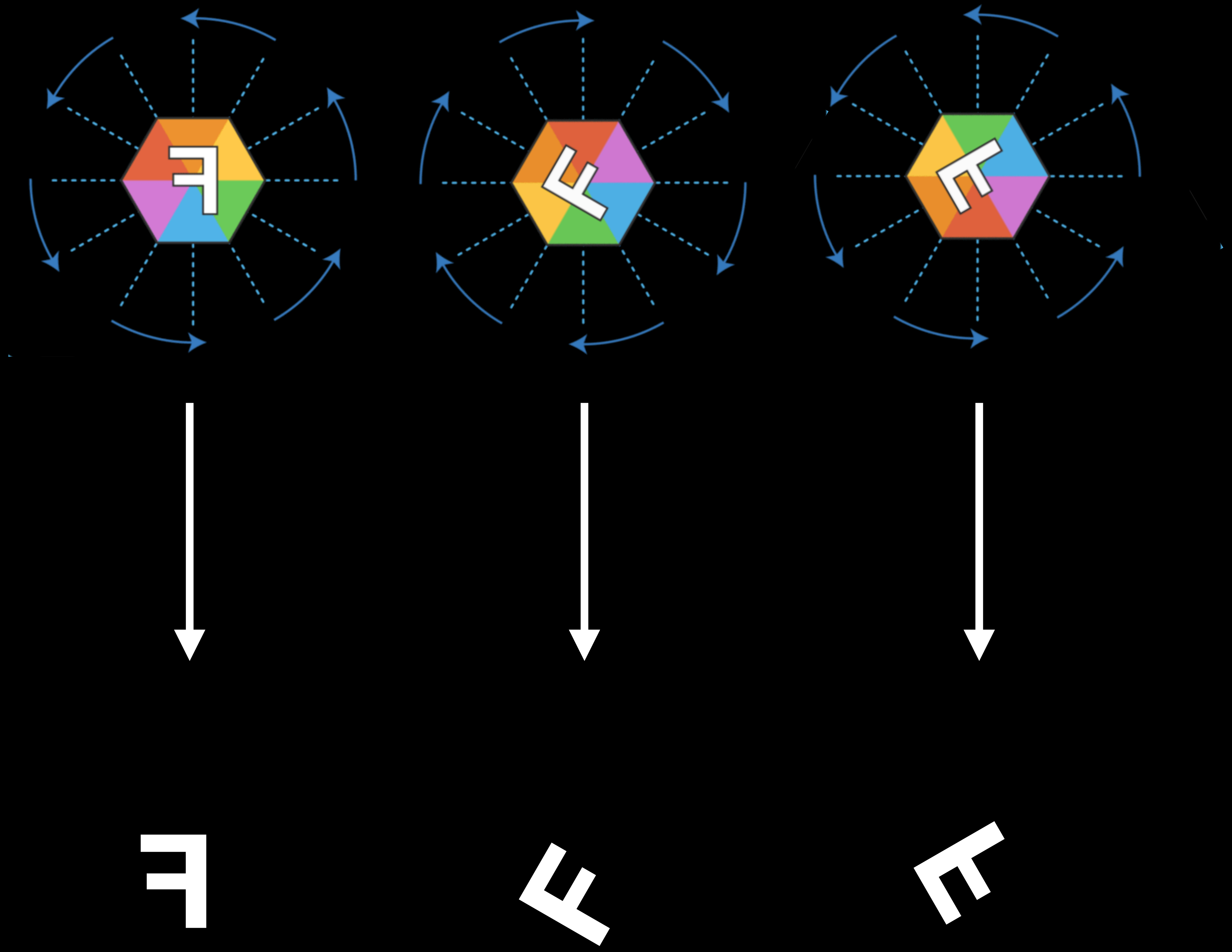
Invariance

$$f(\phi(x)) = f(x)$$



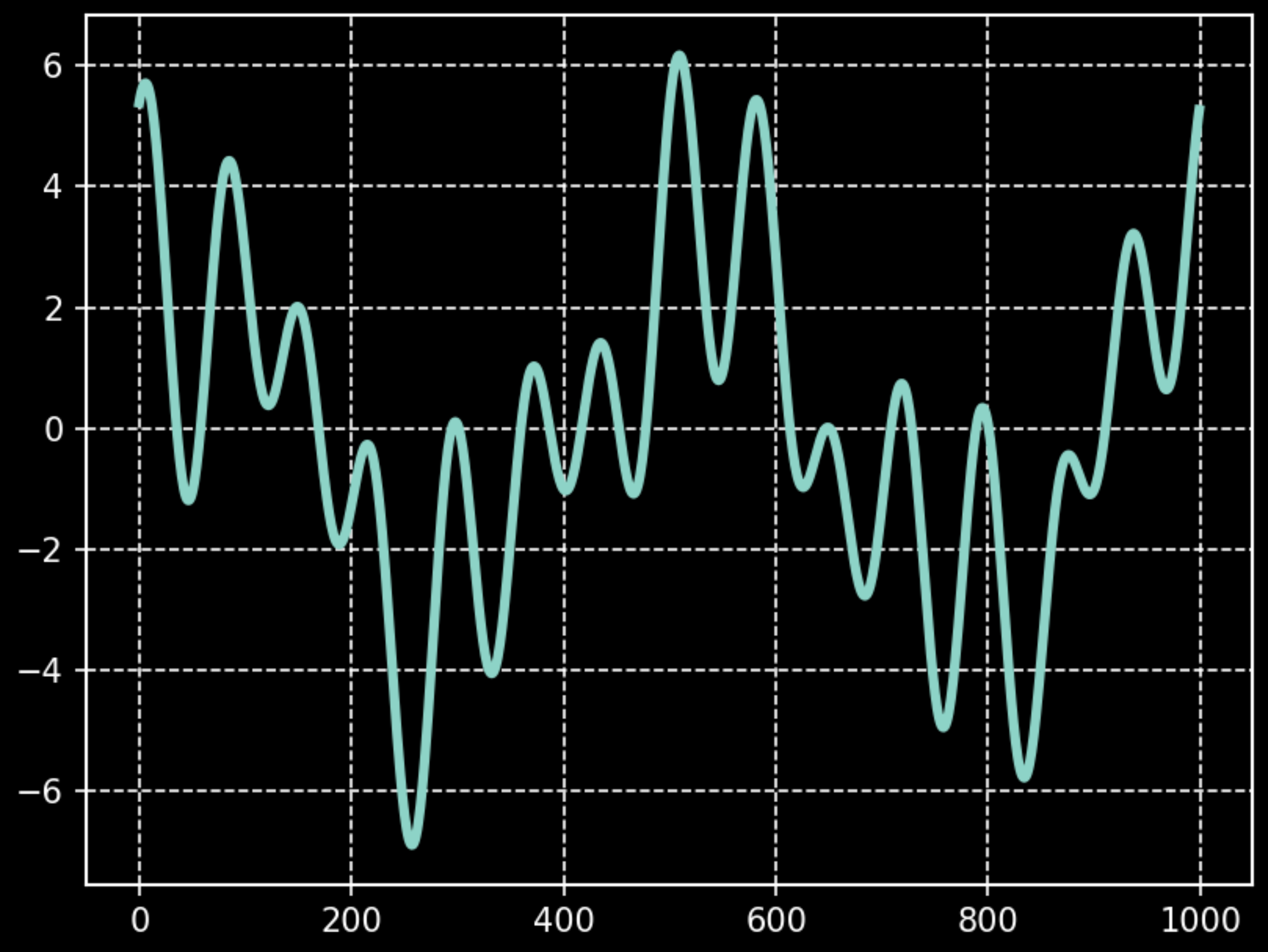
Equivariance

$$f(\phi(x)) = \phi'(f(x))$$



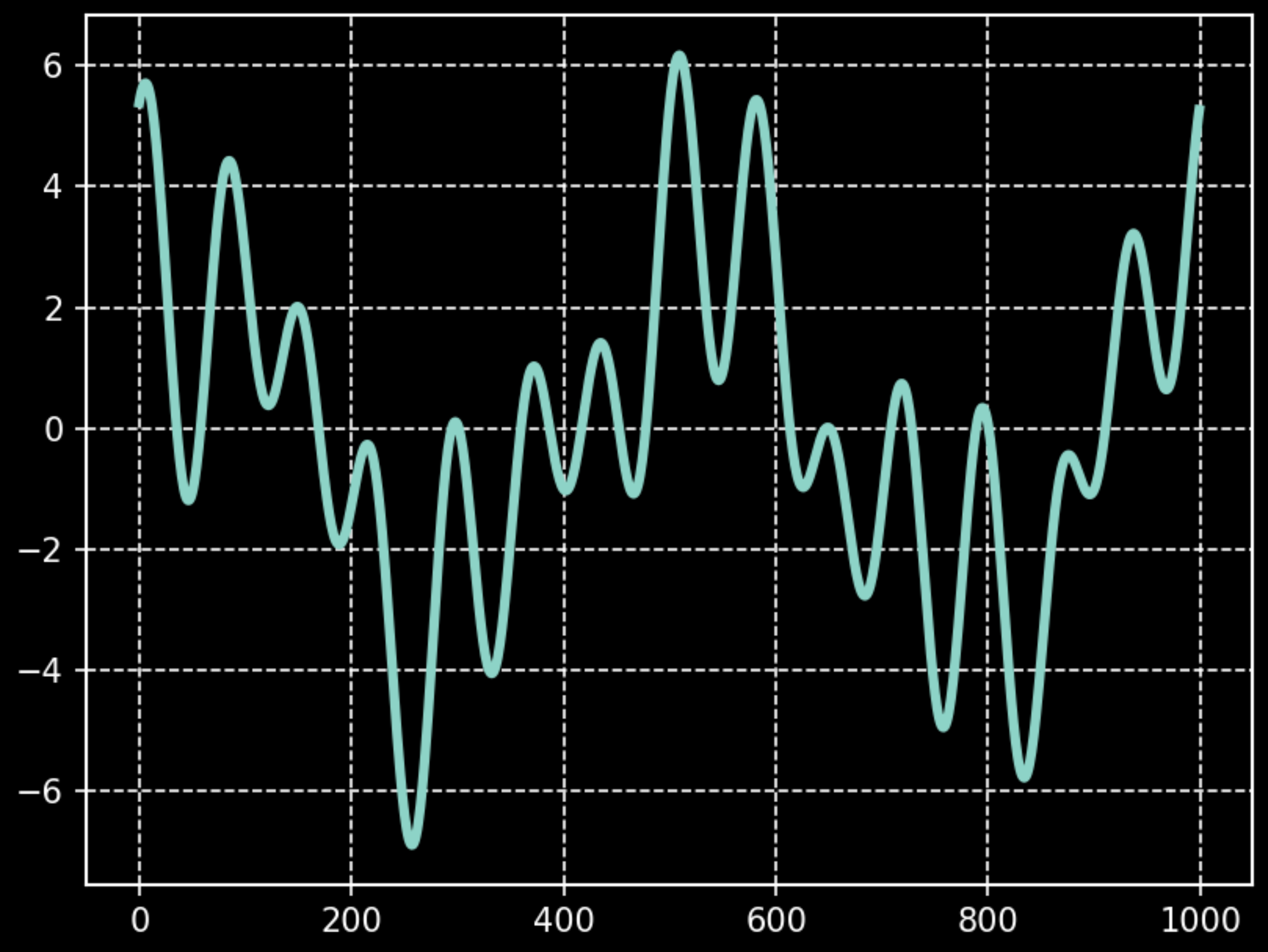
The Mathematics of Invariance

$$f_n$$



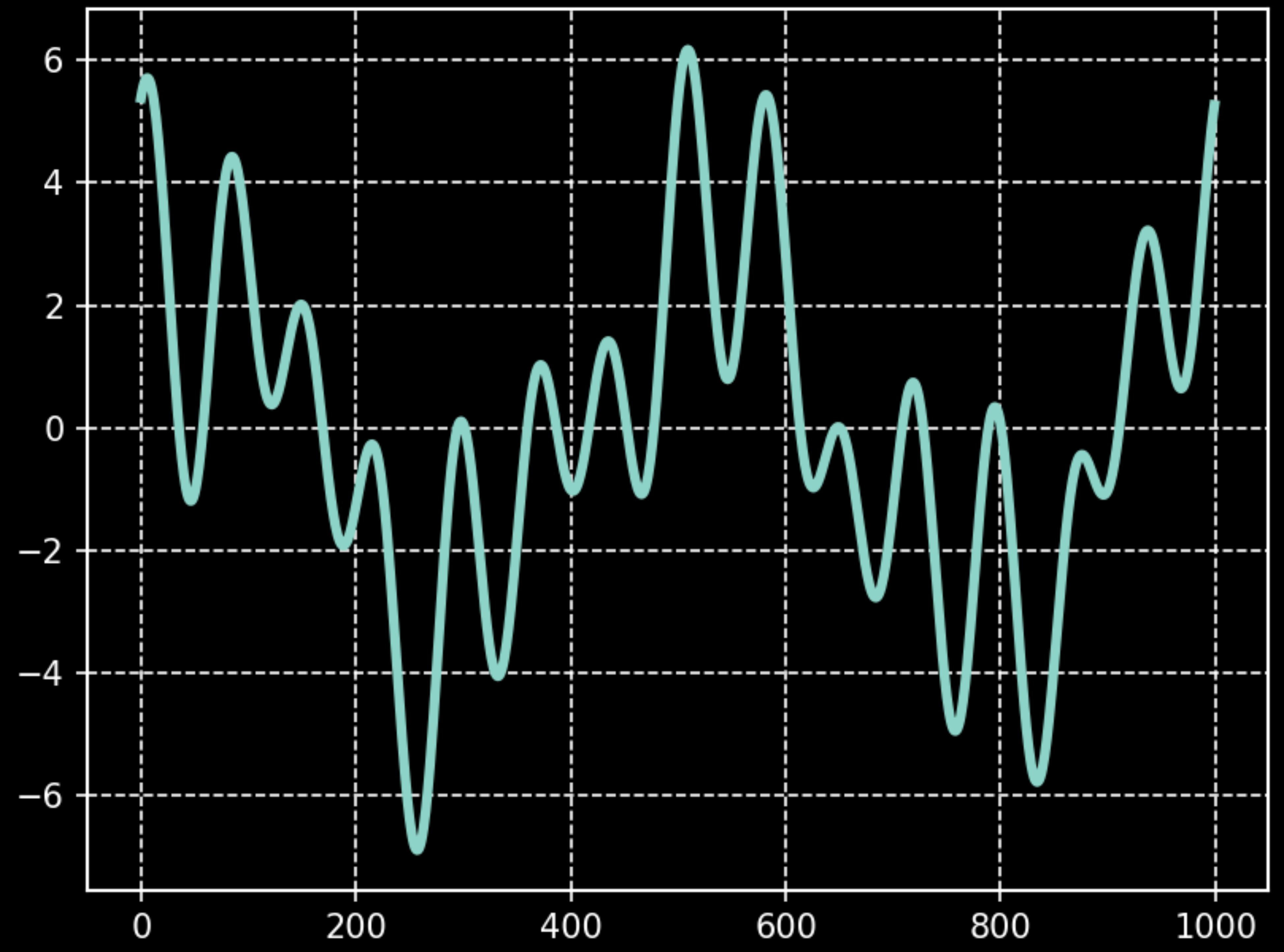
The Mathematics of Invariance

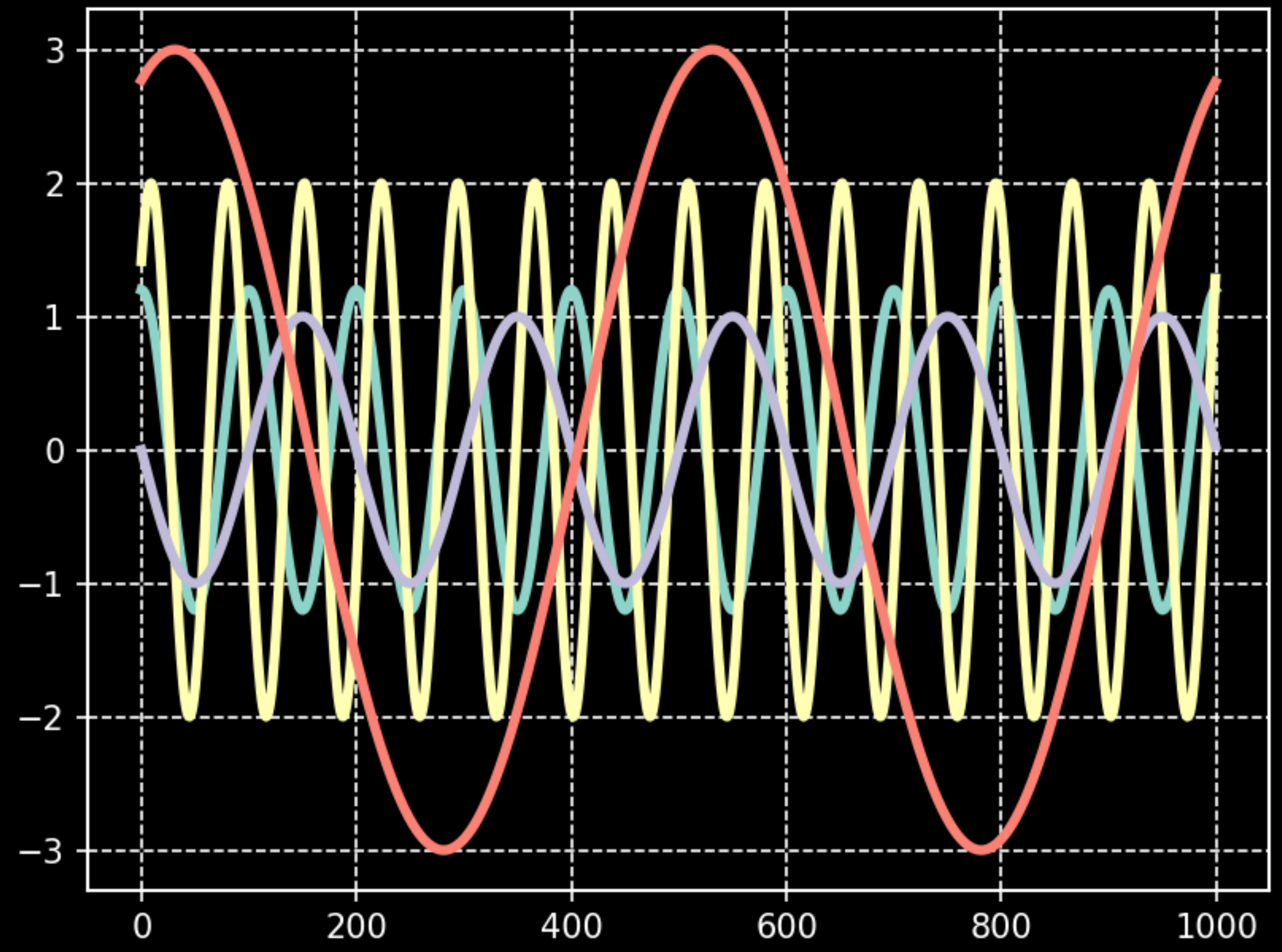
$$f_n =$$



The Mathematics of Invariance

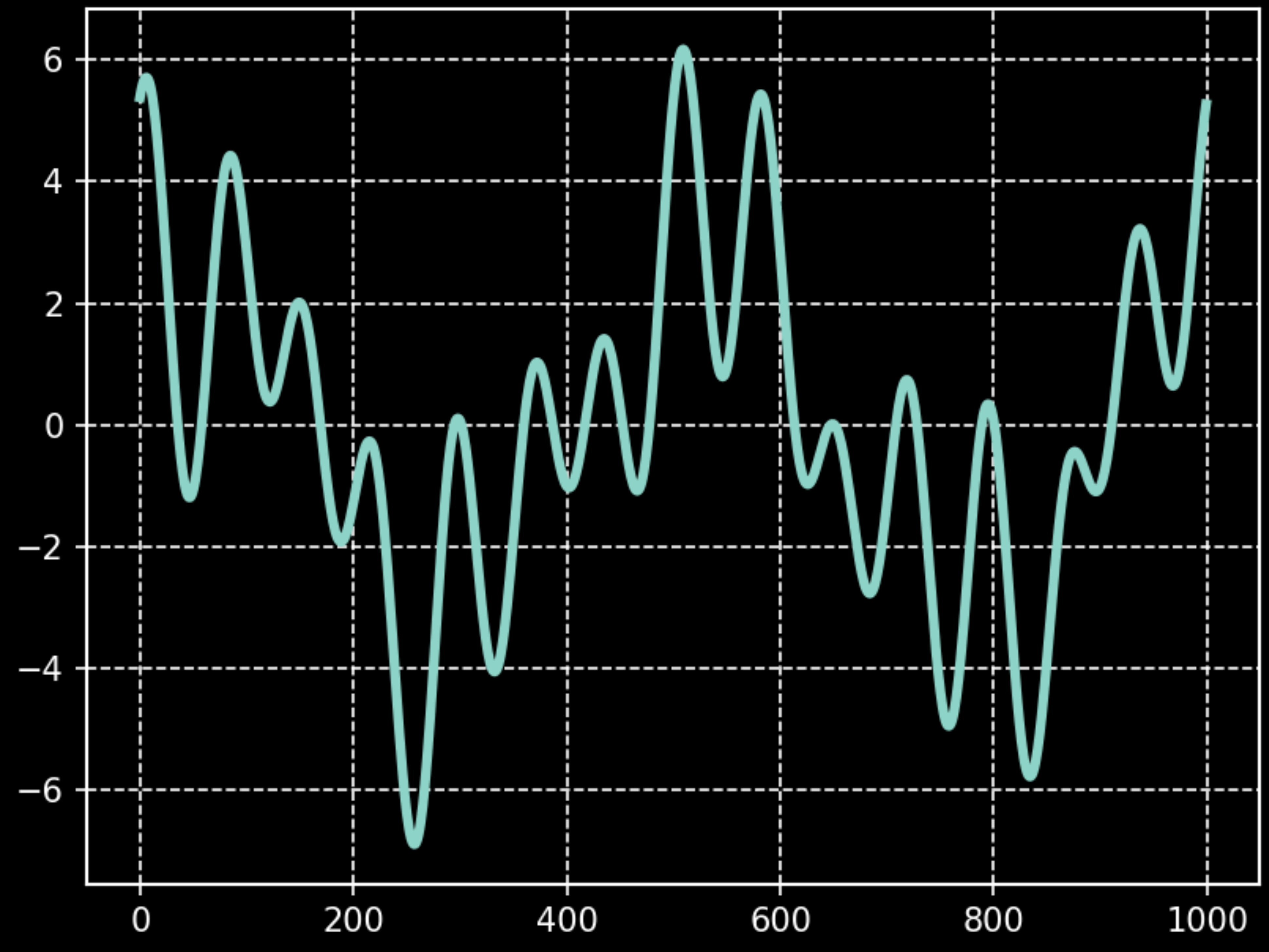
$$f_n = \sum_{n=0}^{N-1} \hat{f}_\omega \, e^{\frac{i2\pi\omega n}{N}}$$



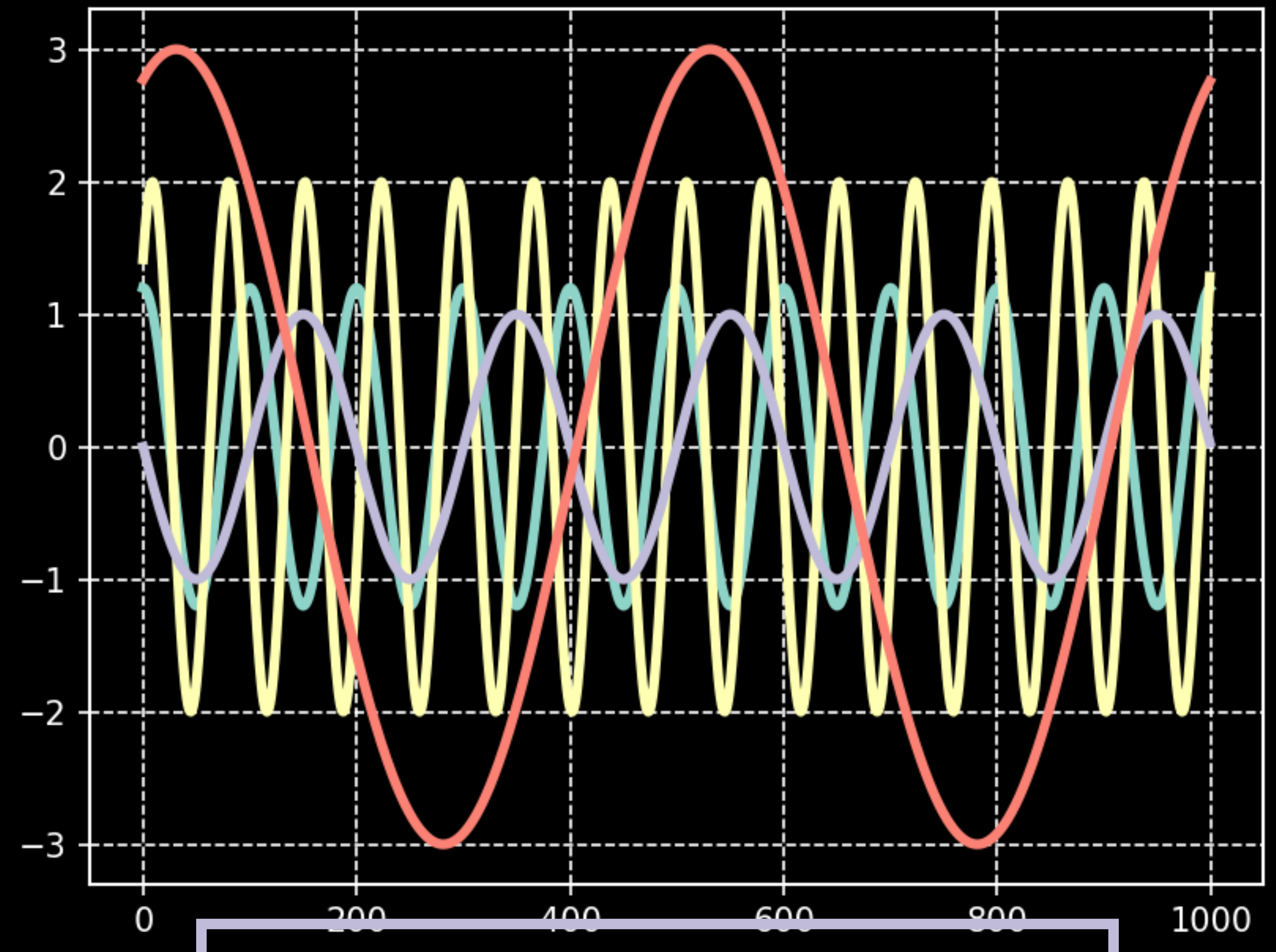


The Mathematics of Invariance

$$f_n = \sum_{n=0}^{N-1} \hat{f}_\omega \boxed{e^{\frac{i2\pi\omega n}{N}}}$$



→

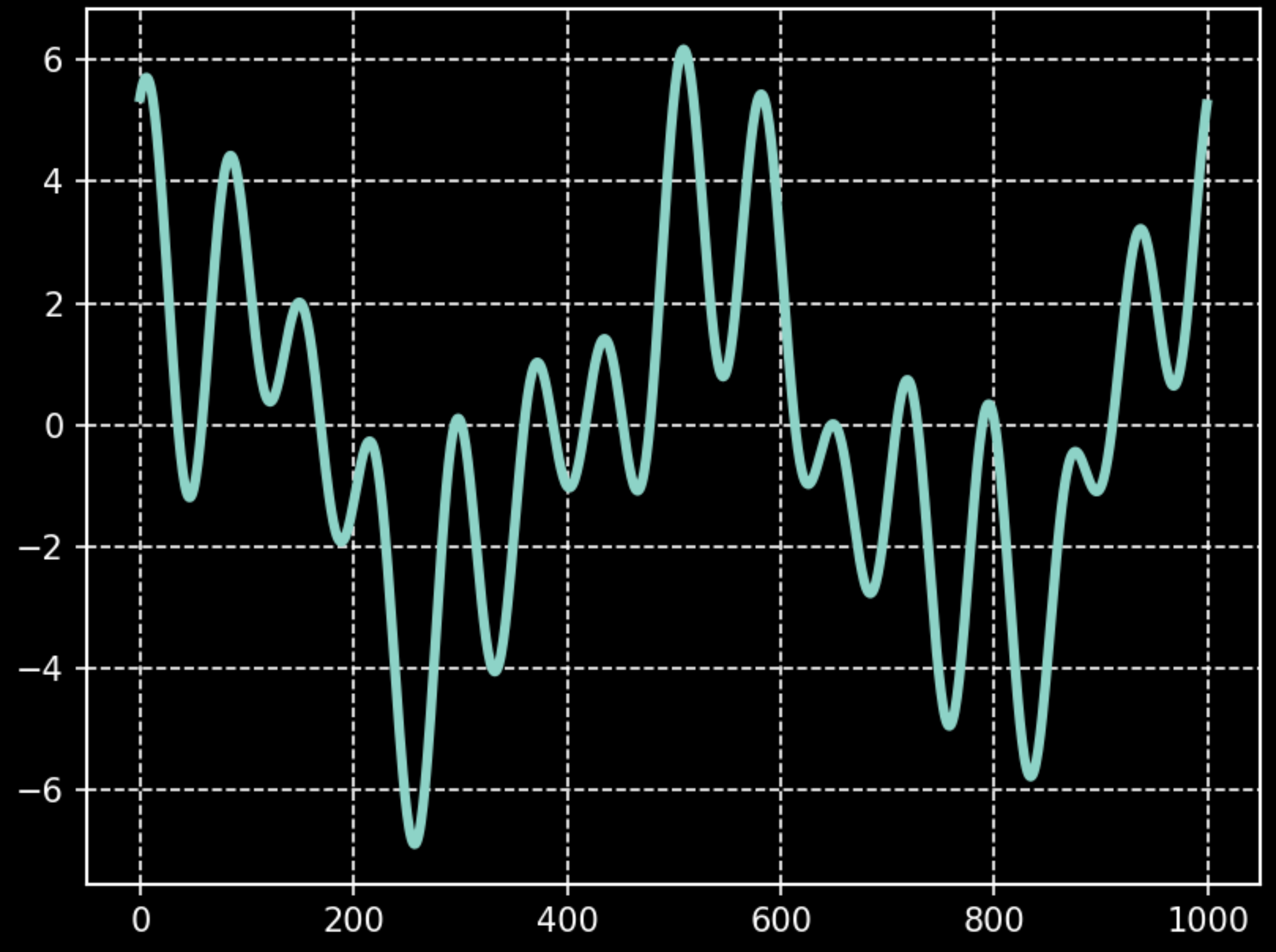


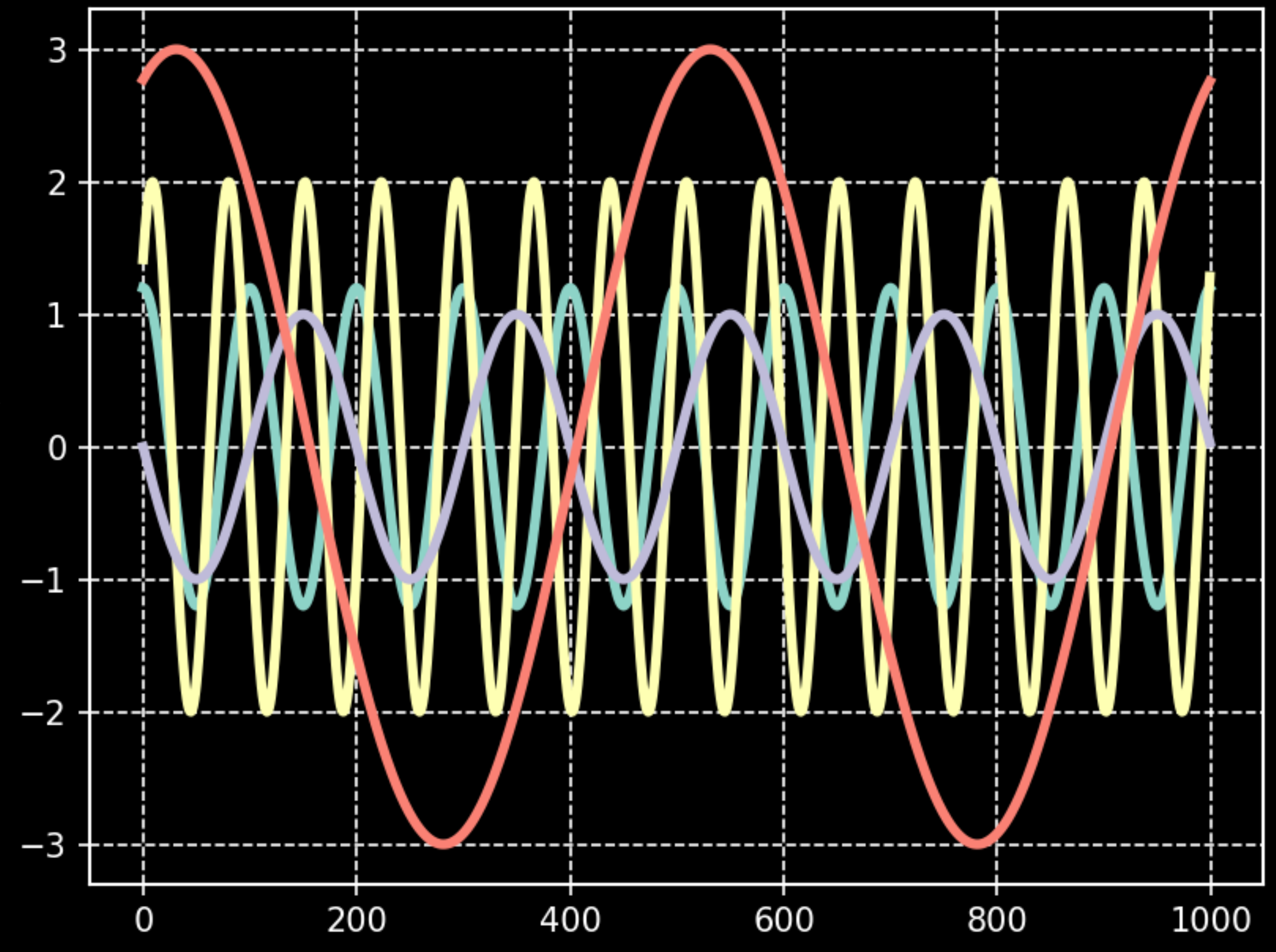
$$e^{i\theta} = \cos\theta + i\sin\theta$$

Euler's Formula

The Mathematics of Invariance

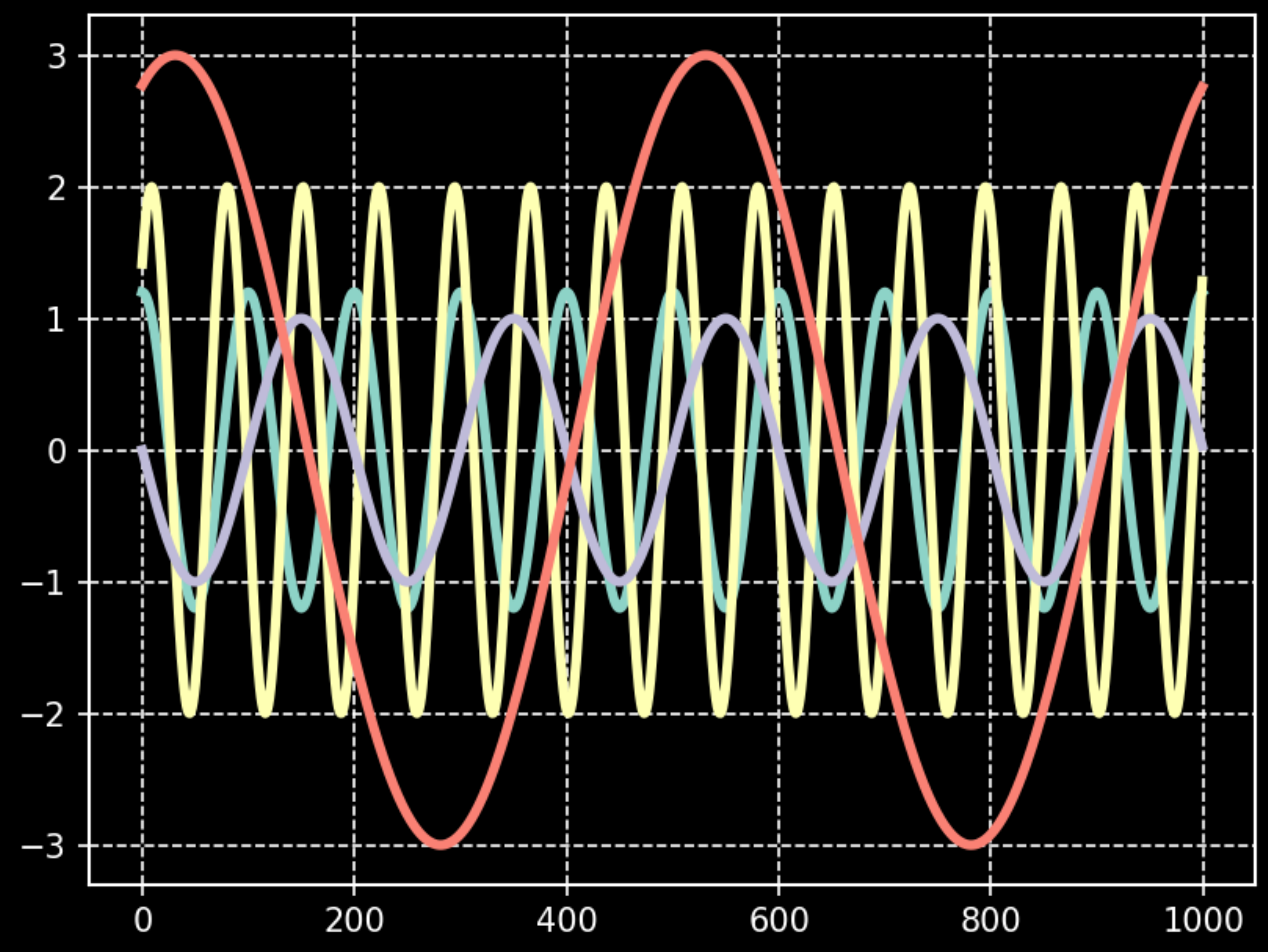
$$f_n = \sum_{n=0}^{N-1} \boxed{\hat{f}_\omega} e^{\frac{i2\pi\omega n}{N}}$$





The Mathematics of Invariance

$$f_n = \sum_{n=0}^{N-1} \boxed{\hat{f}_\omega} e^{\frac{i2\pi\omega n}{N}}$$

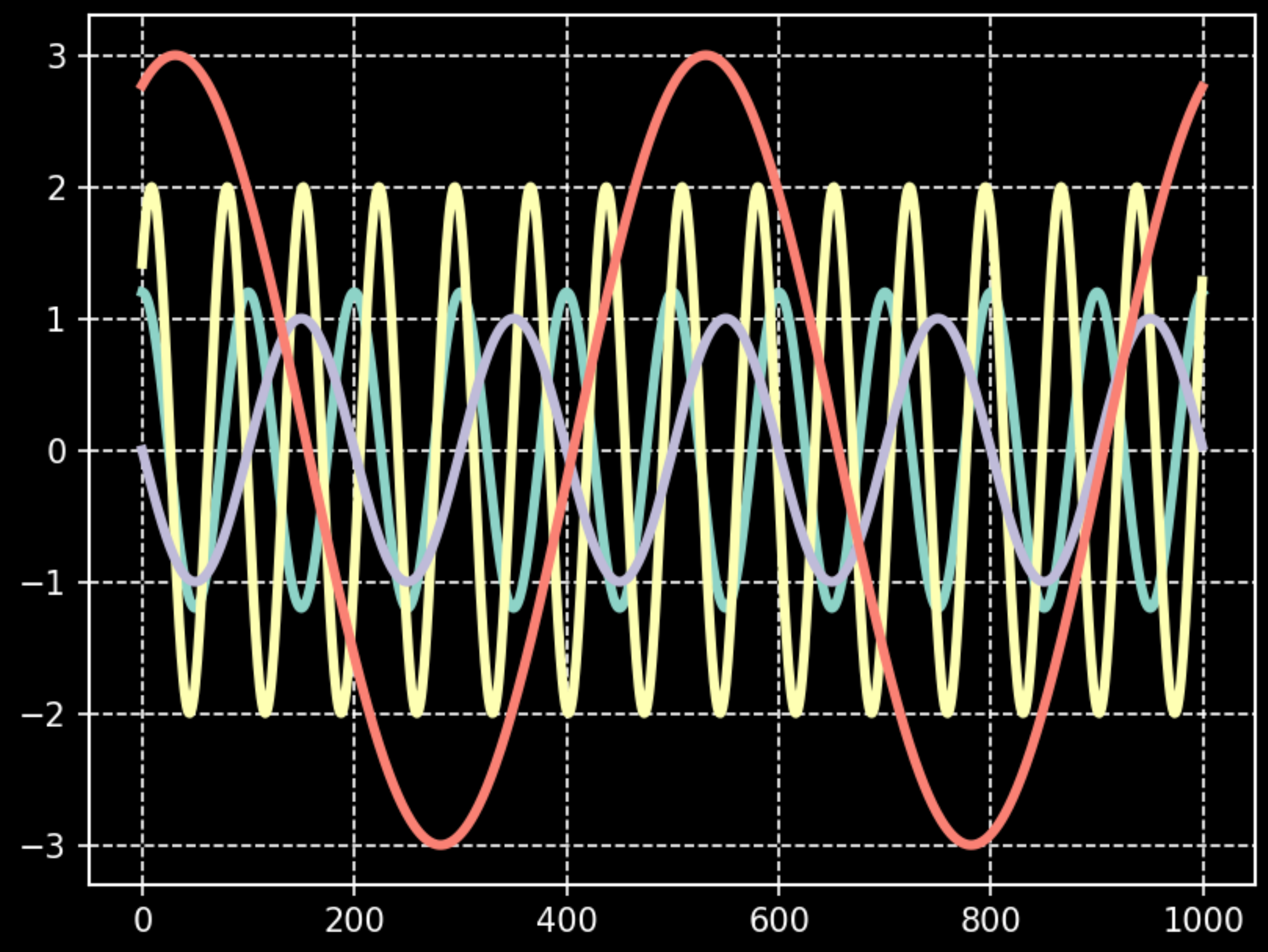


The Mathematics of Invariance

$$f_n = \sum_{n=0}^{N-1} \boxed{\hat{f}_\omega} e^{\frac{i2\pi\omega n}{N}}$$

$$\boxed{\hat{f}_\omega} = \sum_{n=0}^{N-1} f_n \, e^{\frac{-i2\pi\omega n}{N}}$$

The Discrete Fourier Transform

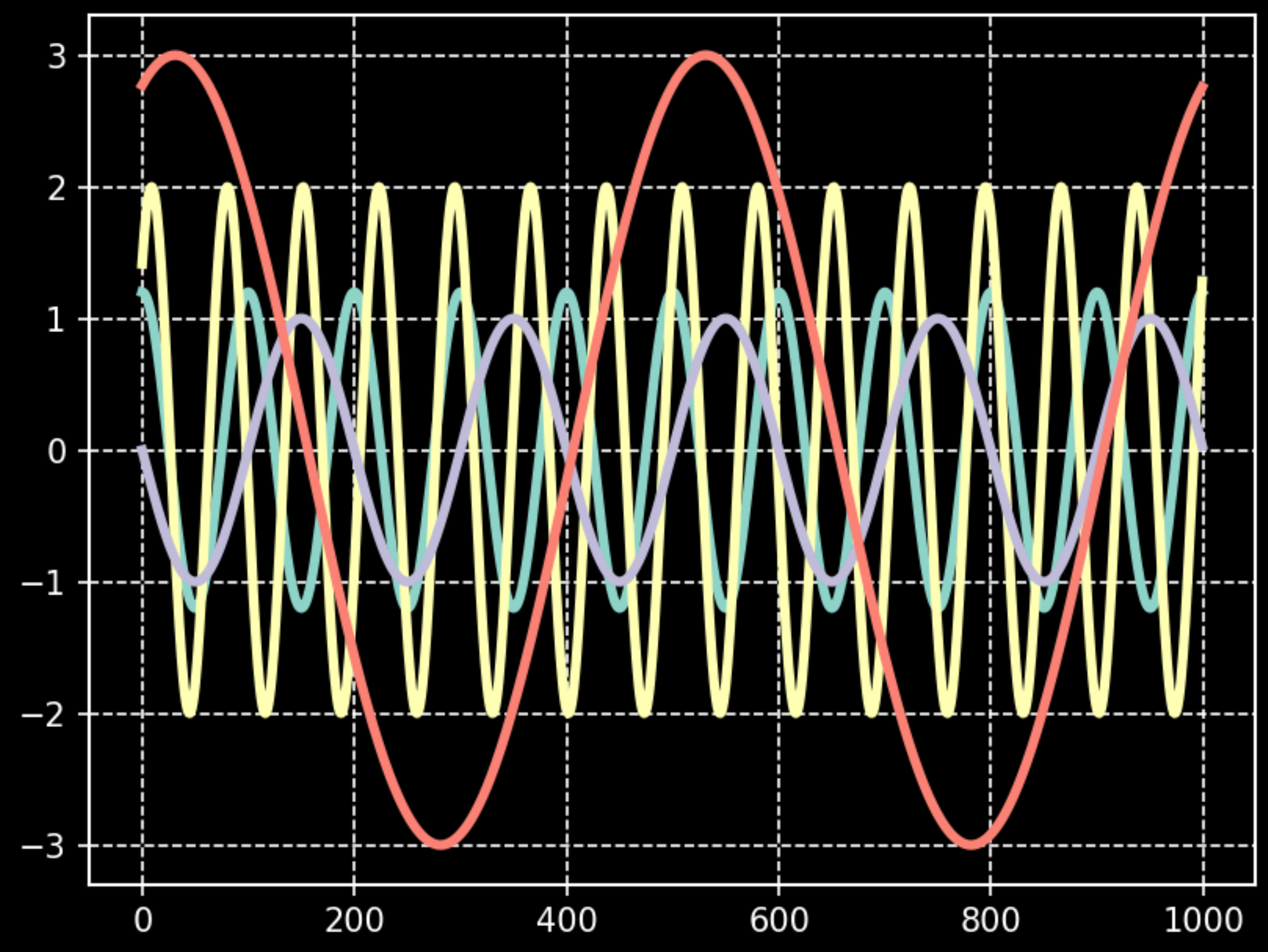


The Mathematics of Invariance

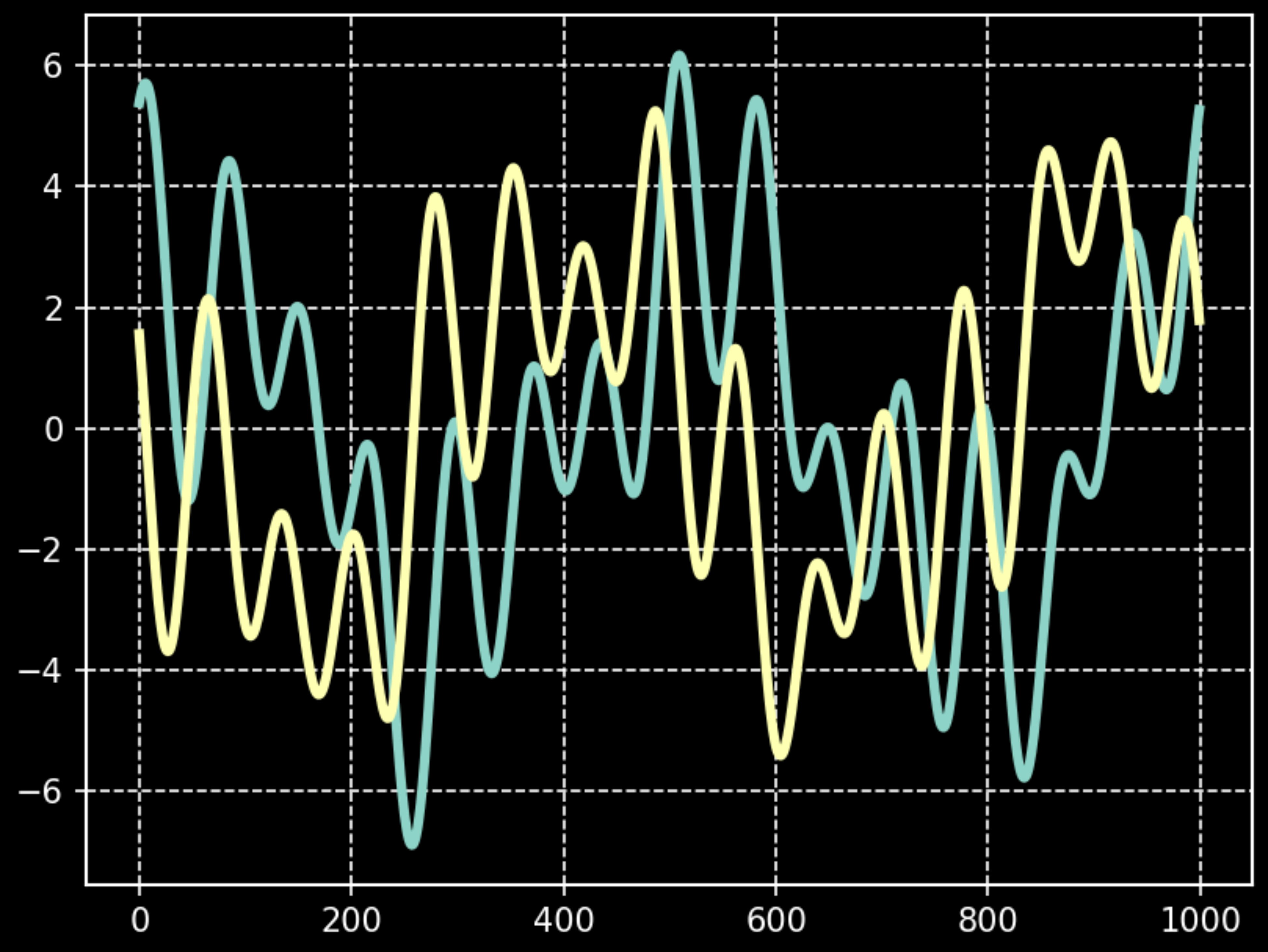
$$f_n = \sum_{n=0}^{N-1} \boxed{\hat{f}_\omega} e^{\frac{i2\pi\omega n}{N}}$$

$$\boxed{\hat{f}_\omega} = \sum_{n=0}^{N-1} f_n \, e^{\frac{-i2\pi\omega n}{N}}$$

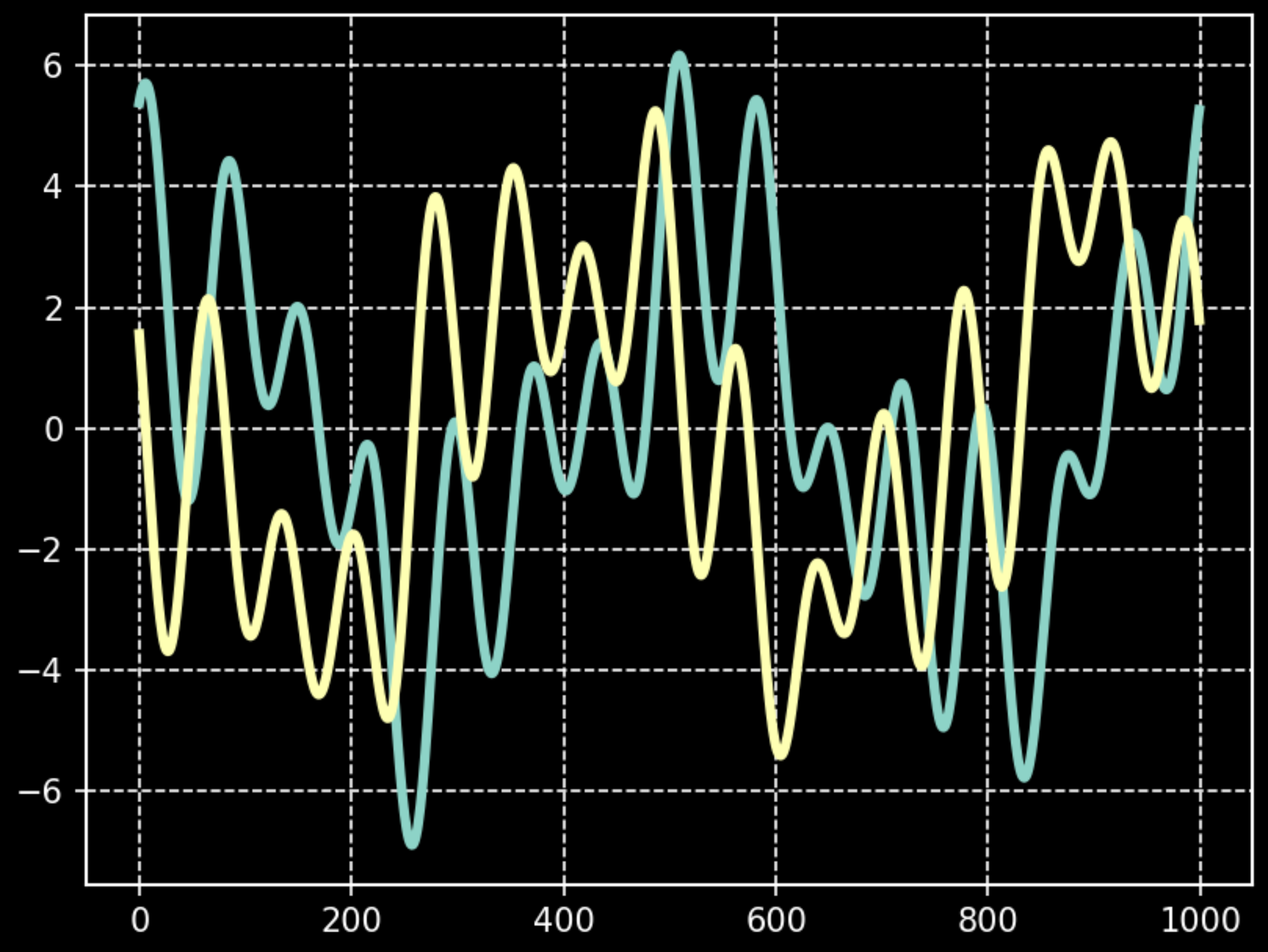
The Discrete Fourier Transform

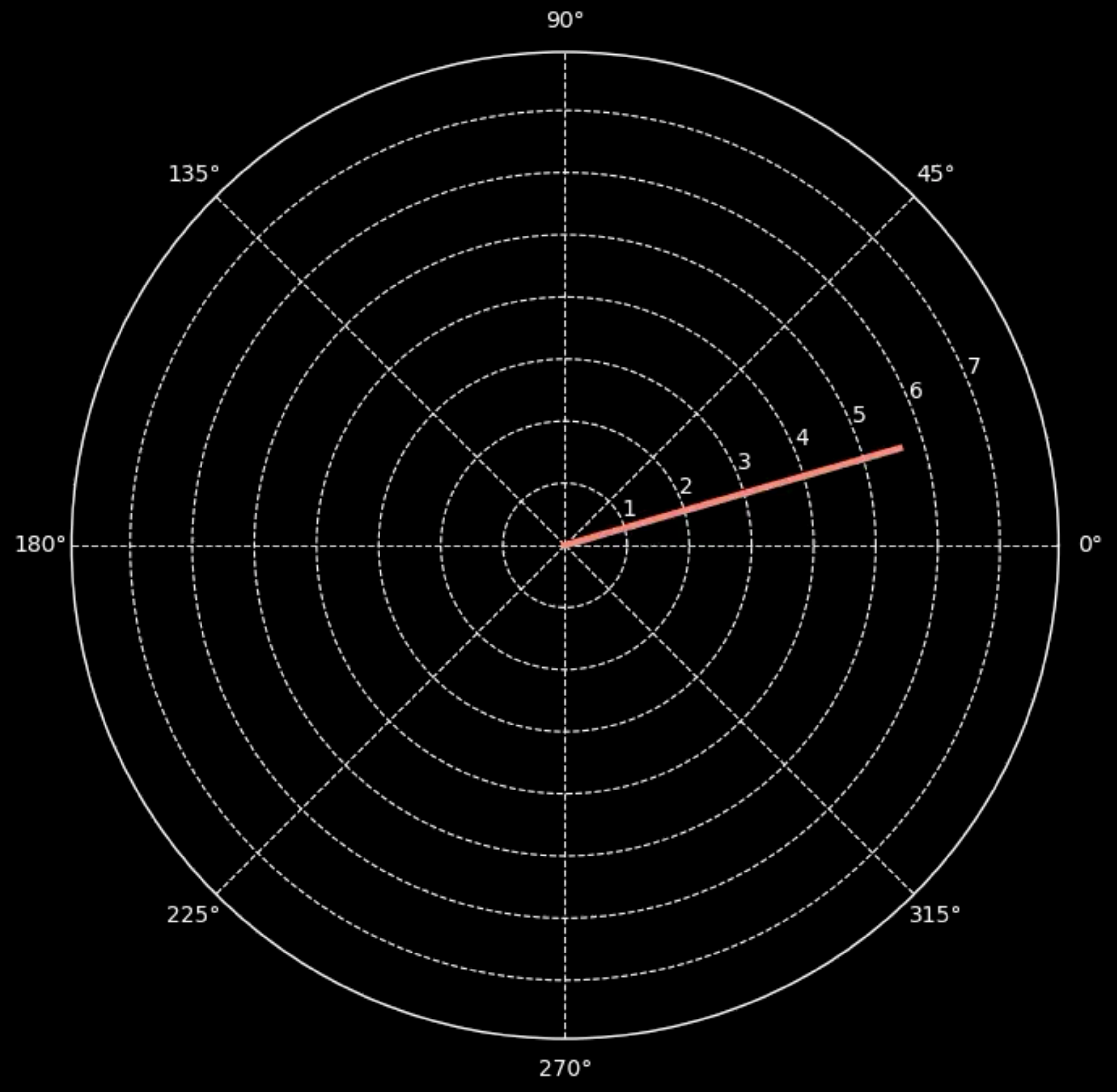


The Mathematics of Invariance

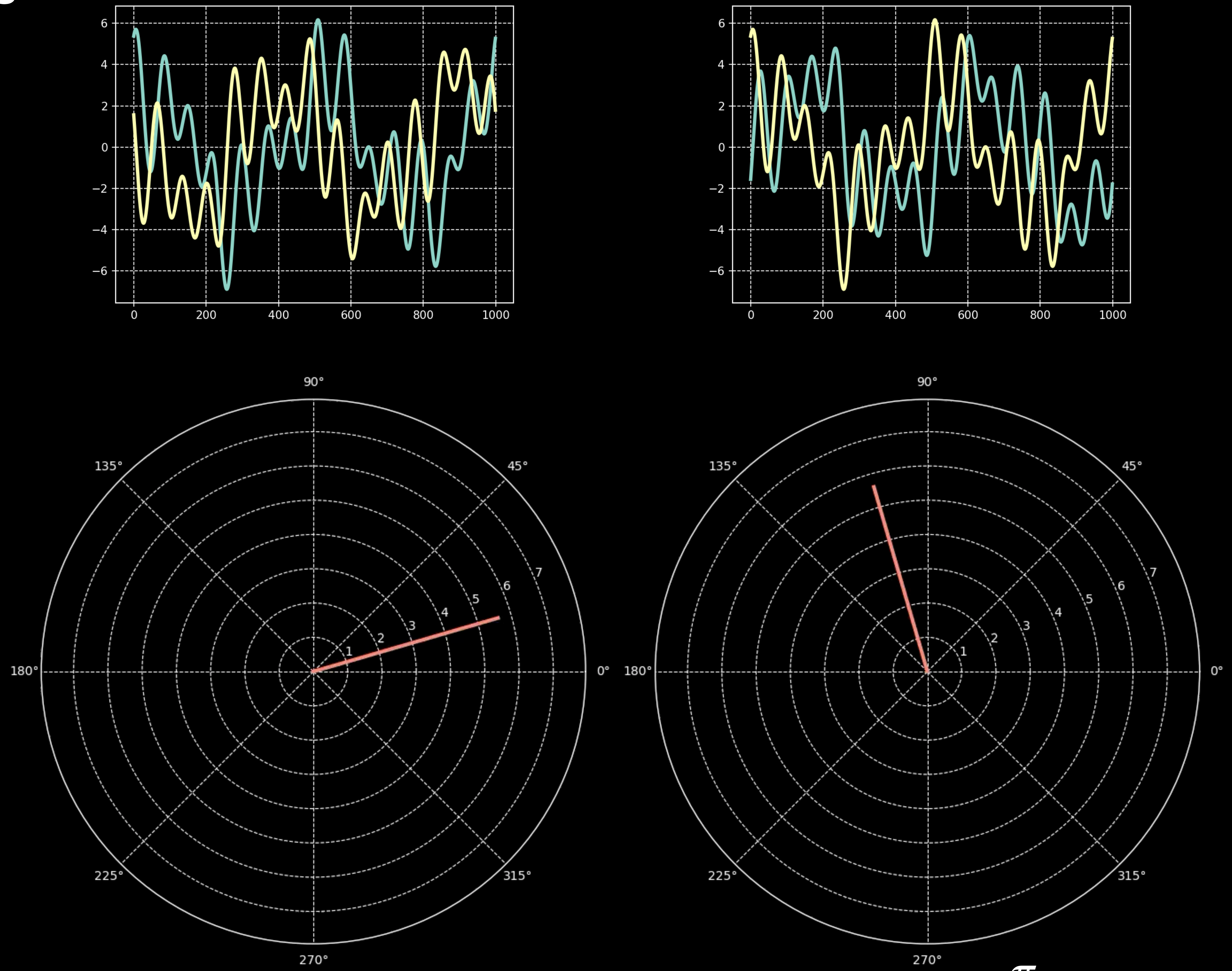


The Mathematics of Invariance





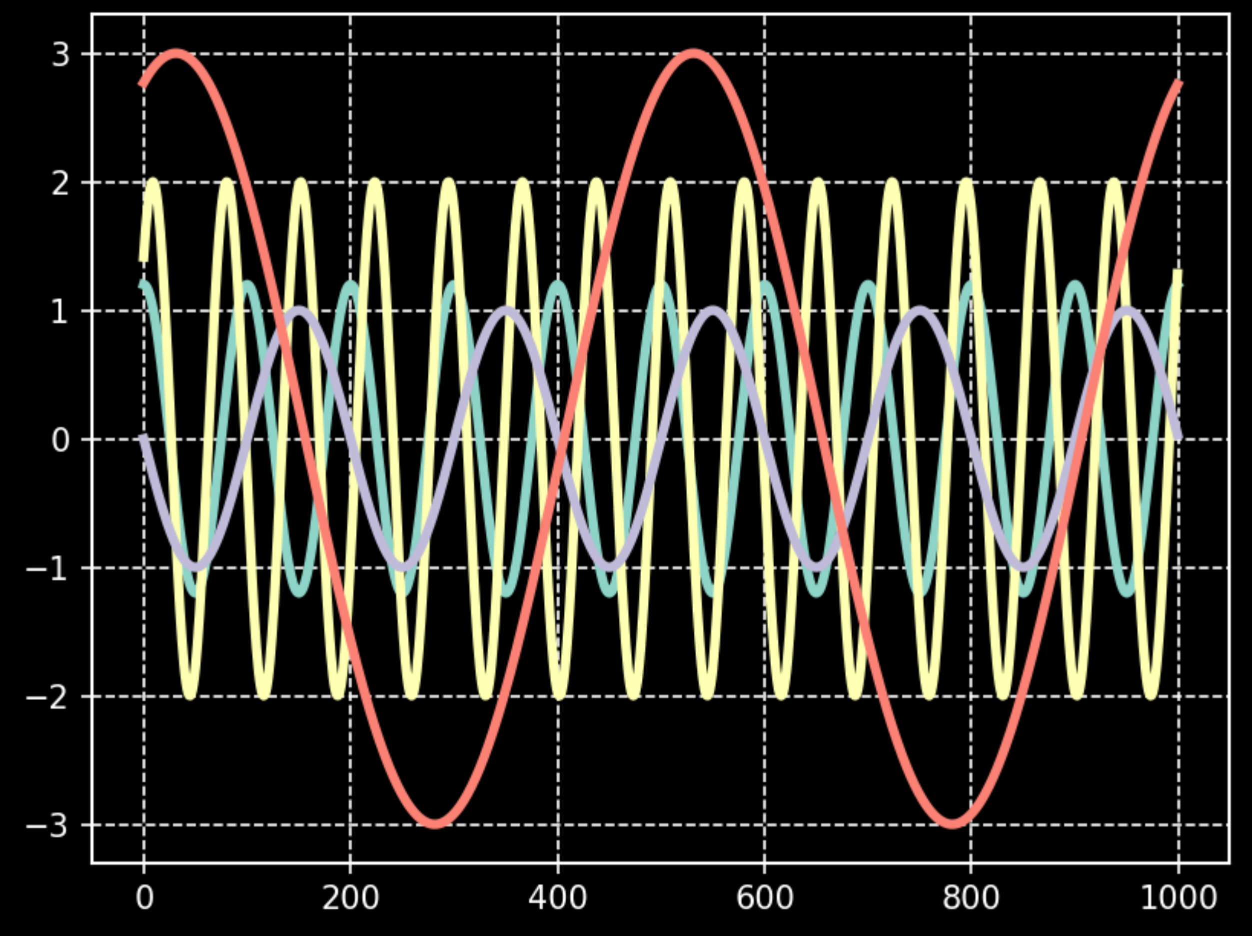
The Mathematics of Invariance



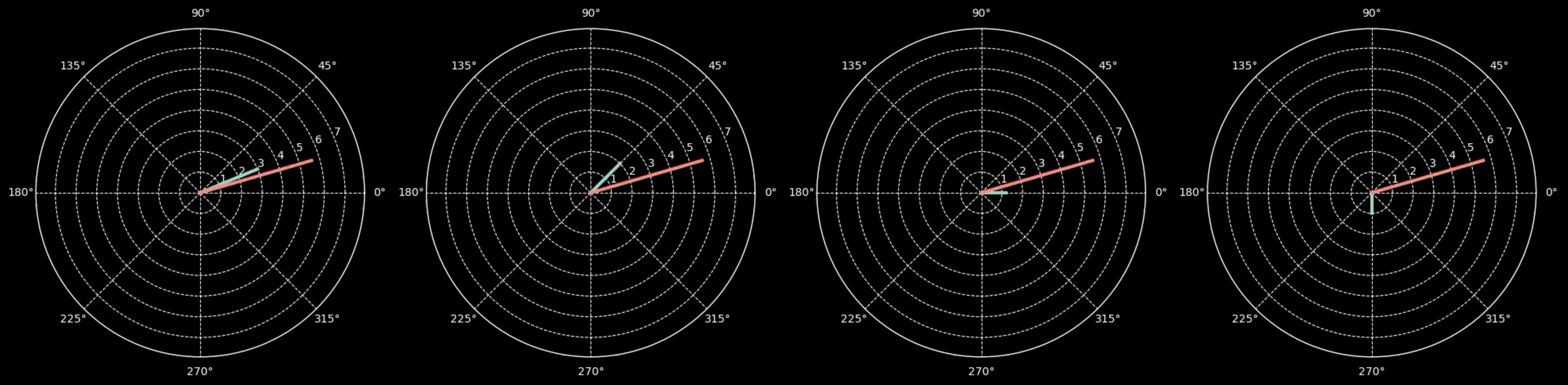
Original Signal

Shifted by $\frac{\pi}{2}$

The Mathematics of Invariance

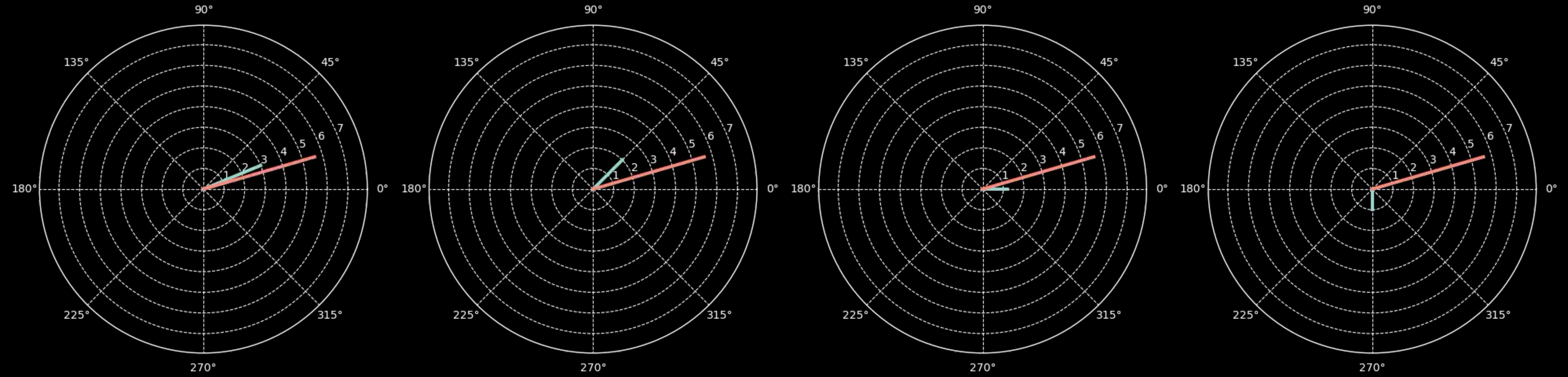


Top Fourier Coefficients

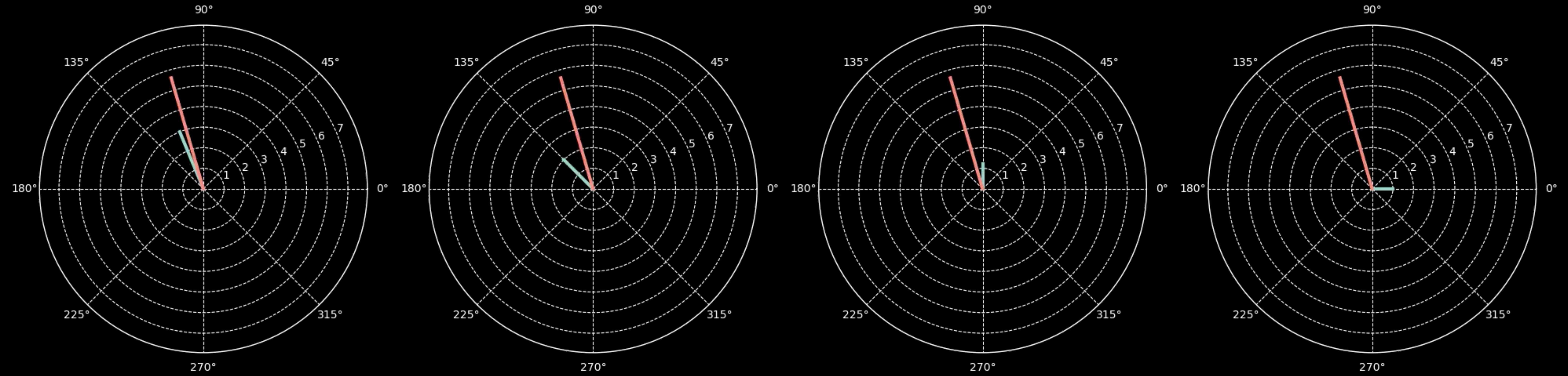


The Mathematics of Invariance

Original Fourier Coefficients



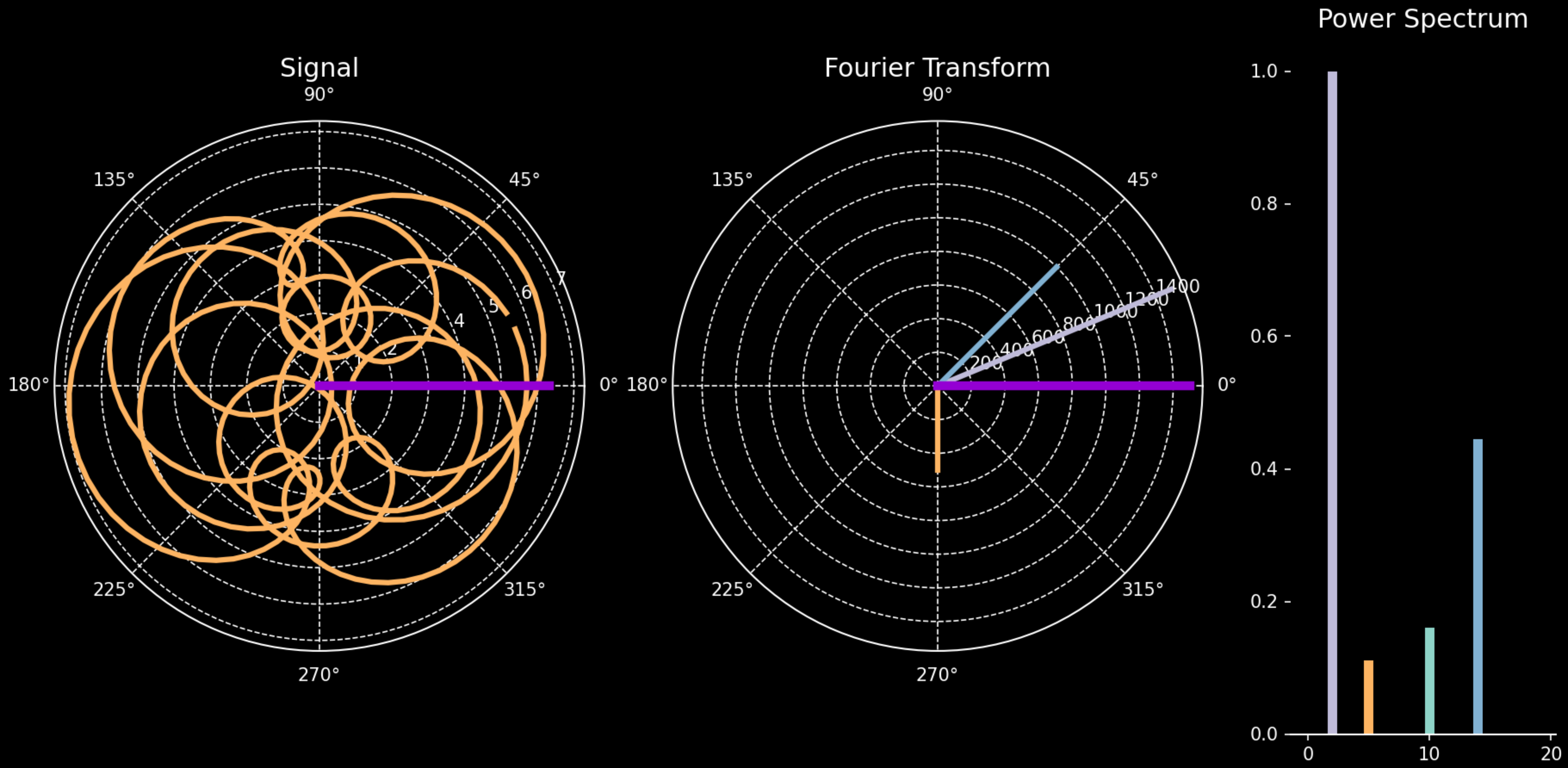
Fourier Coefficients of Shifted Signal



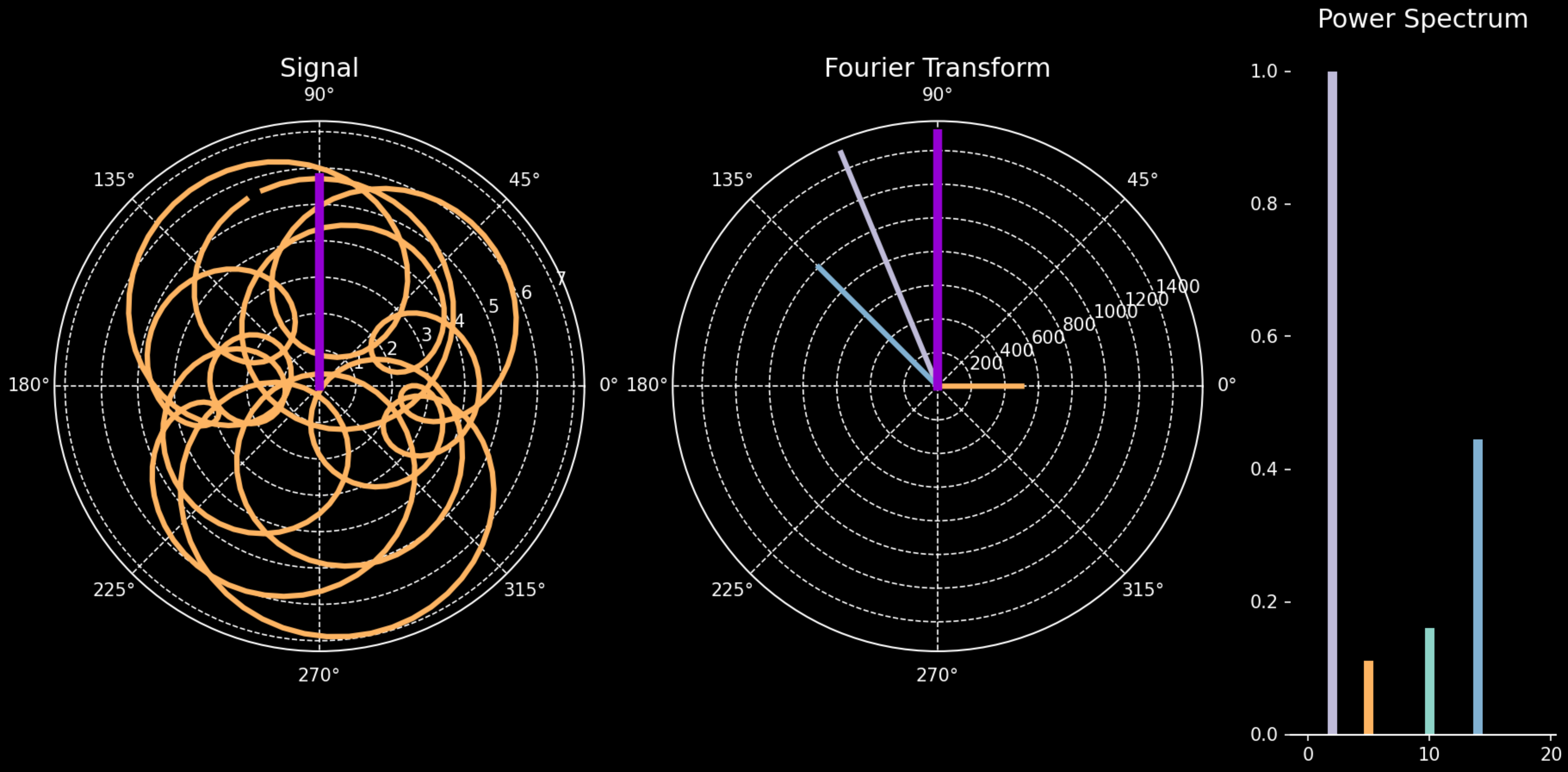
Power Spectrum

$$P_\omega = |\hat{f}_\omega|^2$$

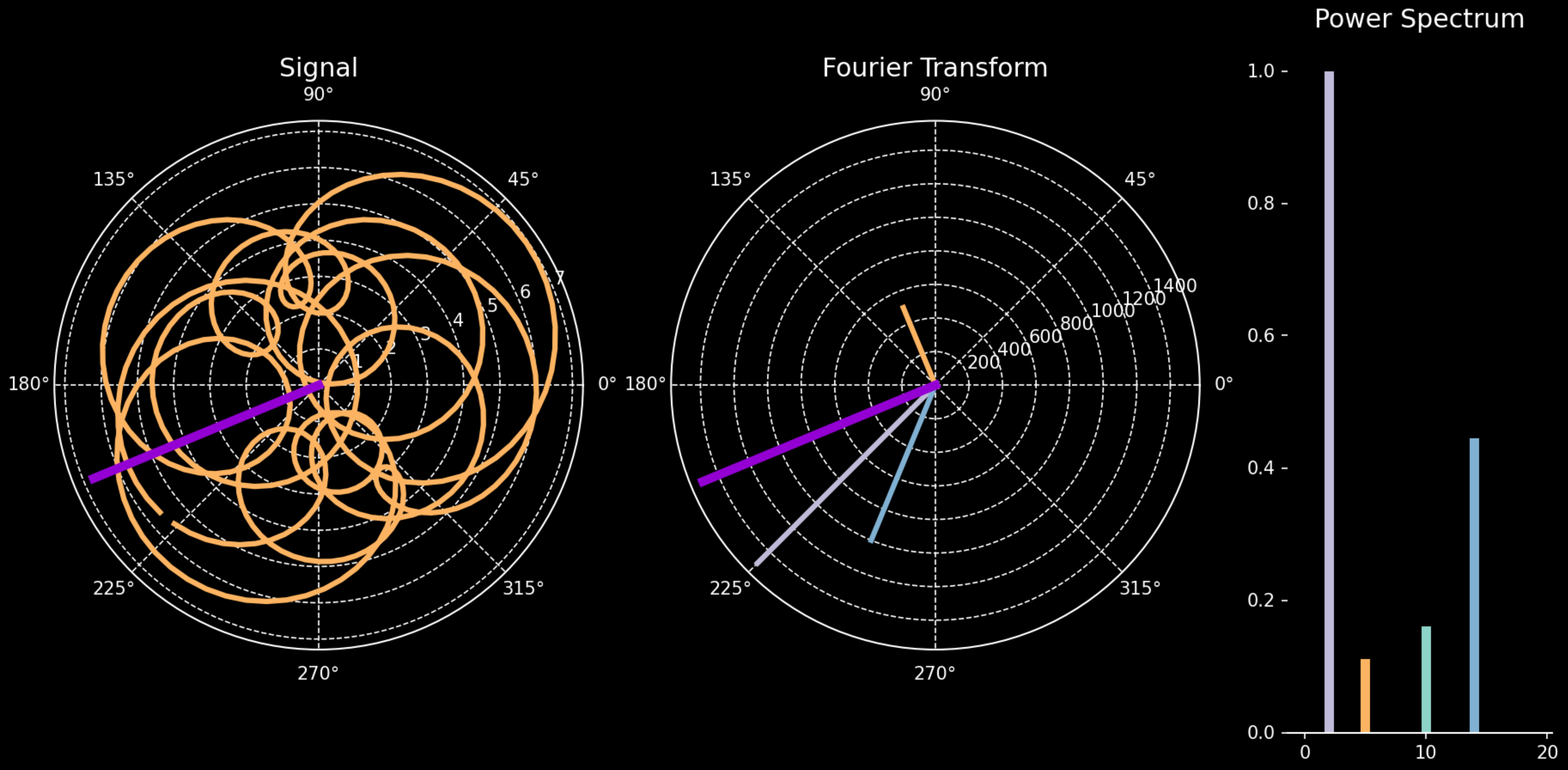
The Mathematics of Invariance



The Mathematics of Invariance

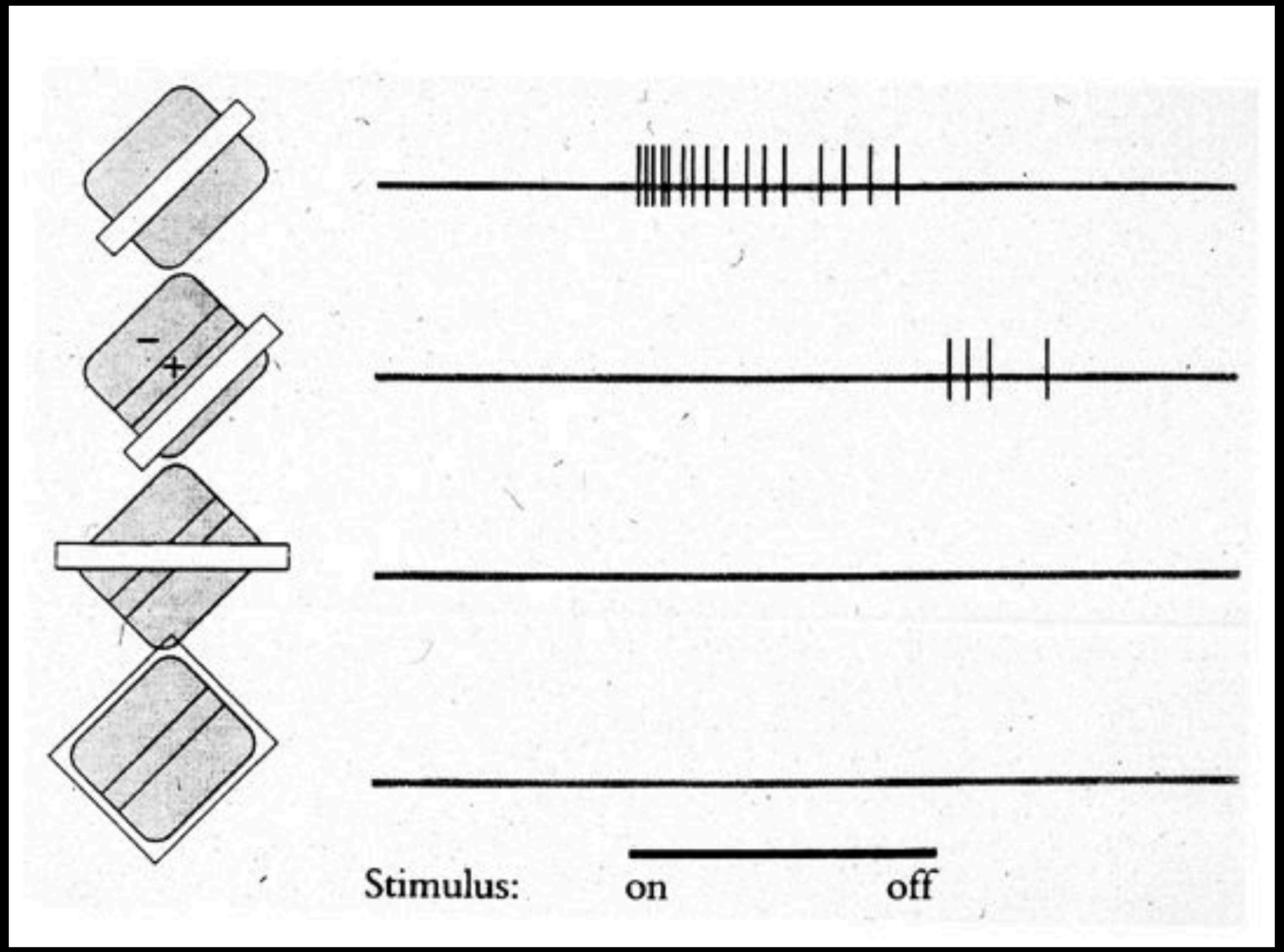


The Mathematics of Invariance

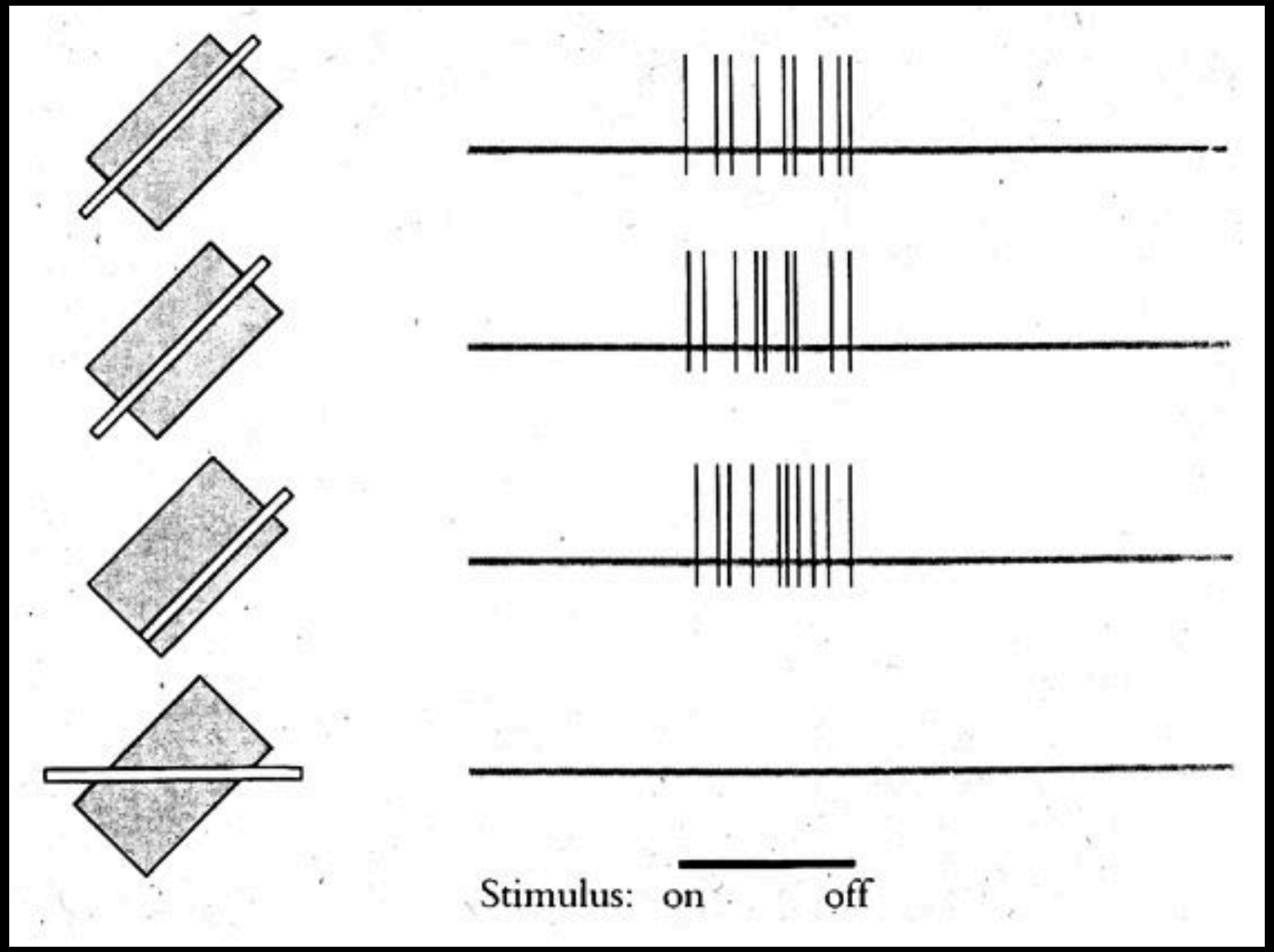


Invariance in Visual Cortex

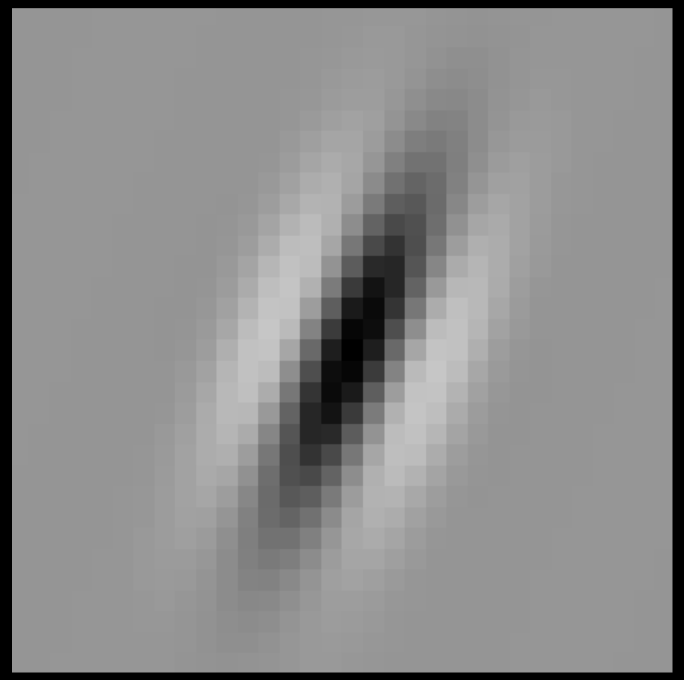
Simple Cells

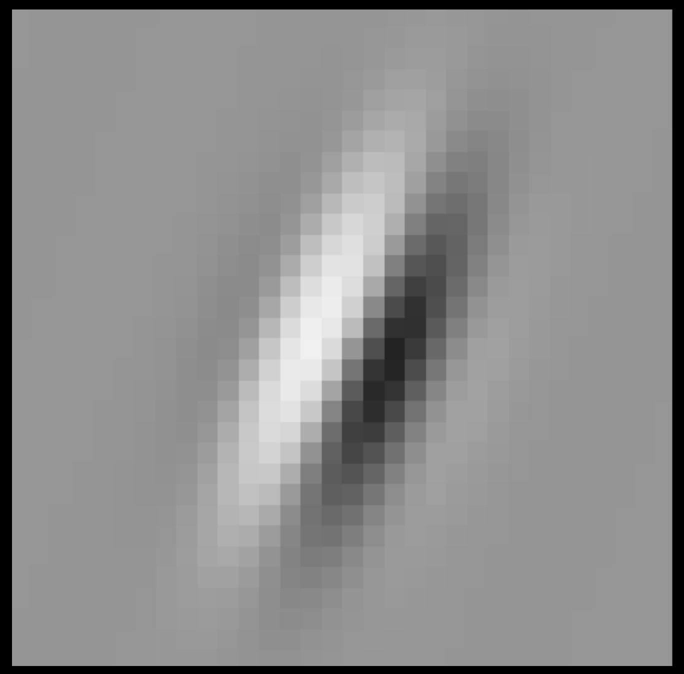


Complex Cells



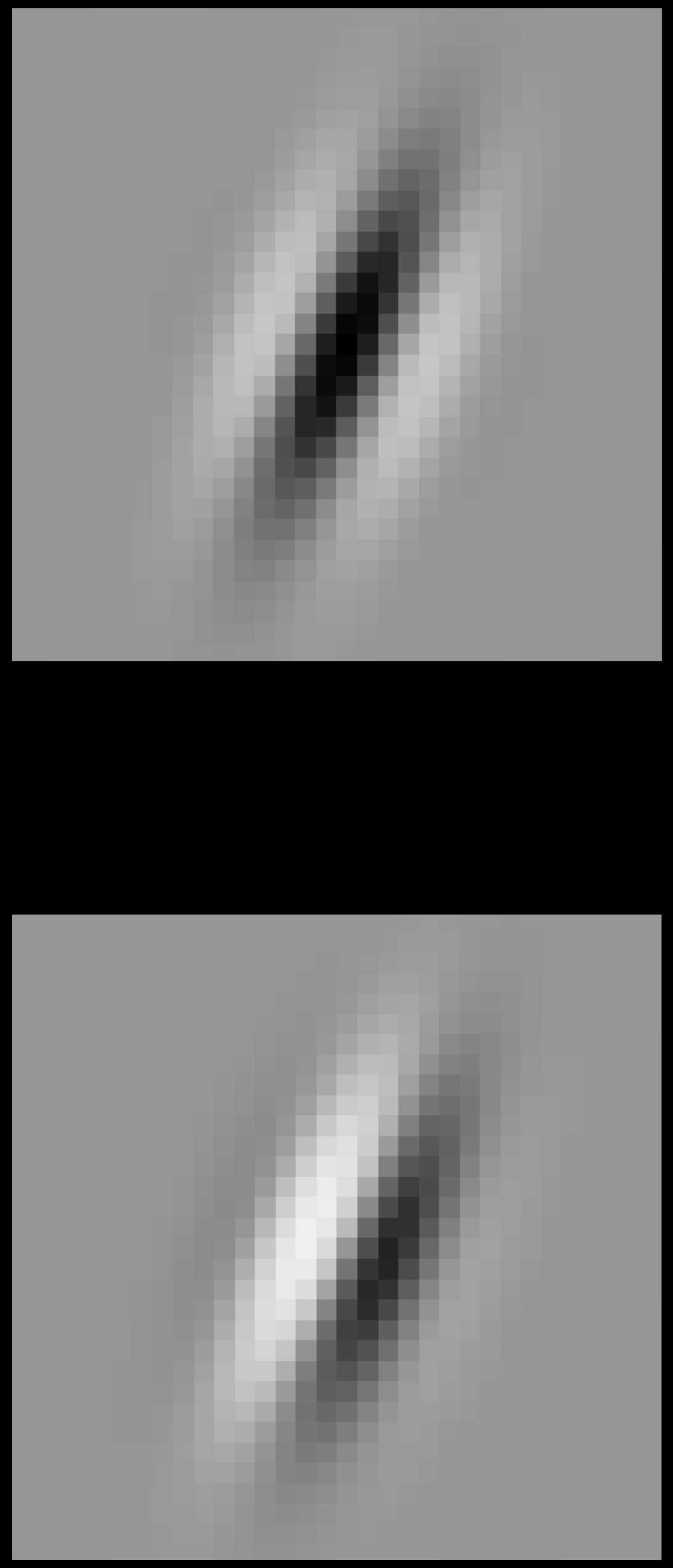
Invariance in Visual Cortex



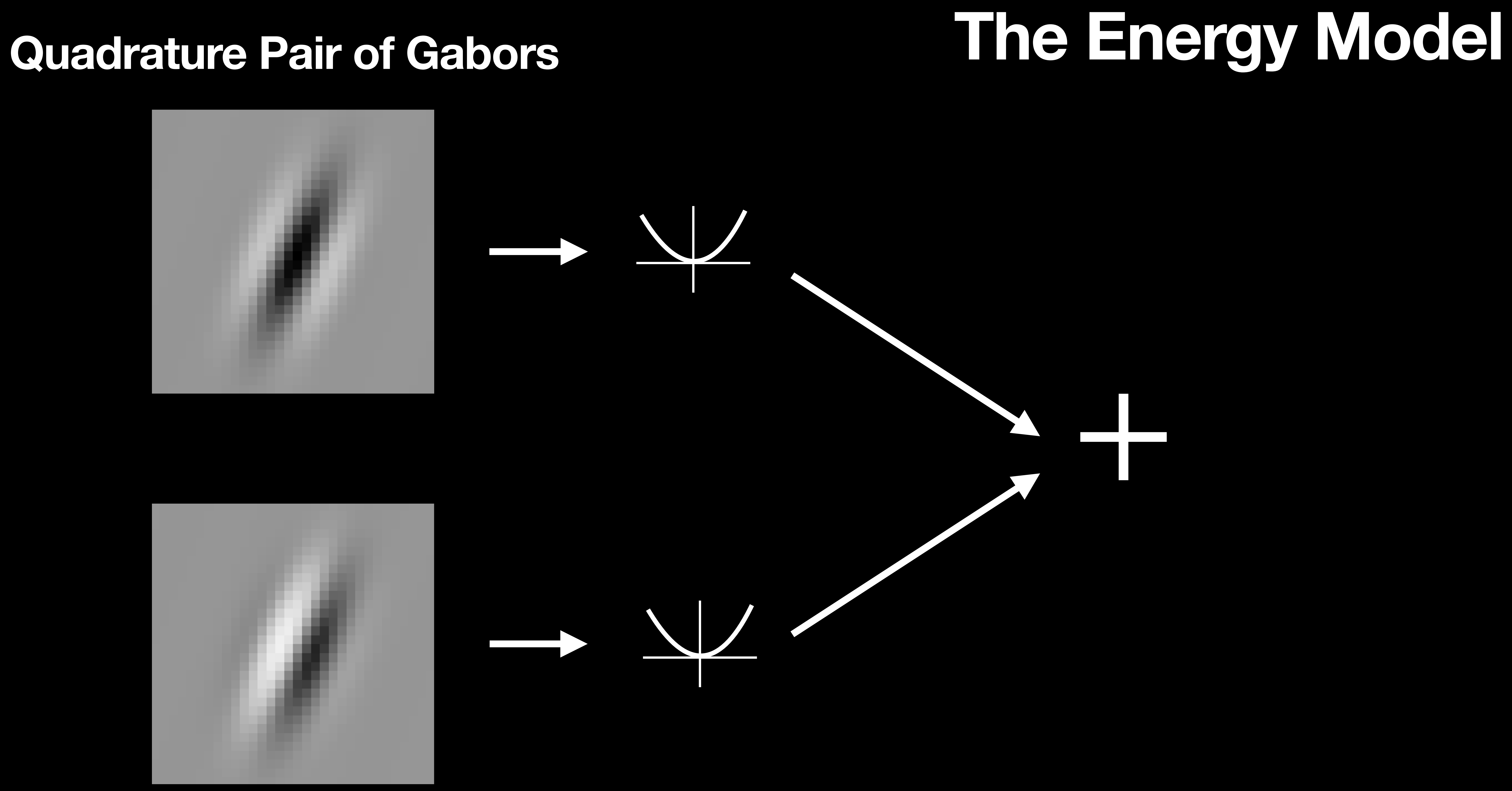


Invariance in Visual Cortex

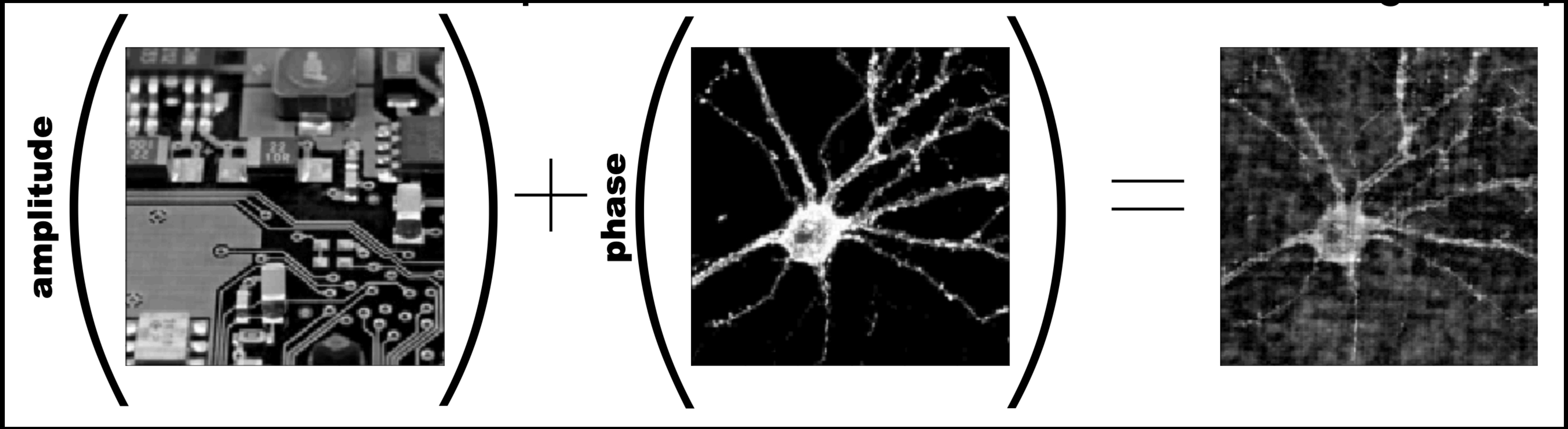
Quadrature Pair of Gabors



Invariance in Visual Cortex



The Mathematics of Invariance



Bispectrum

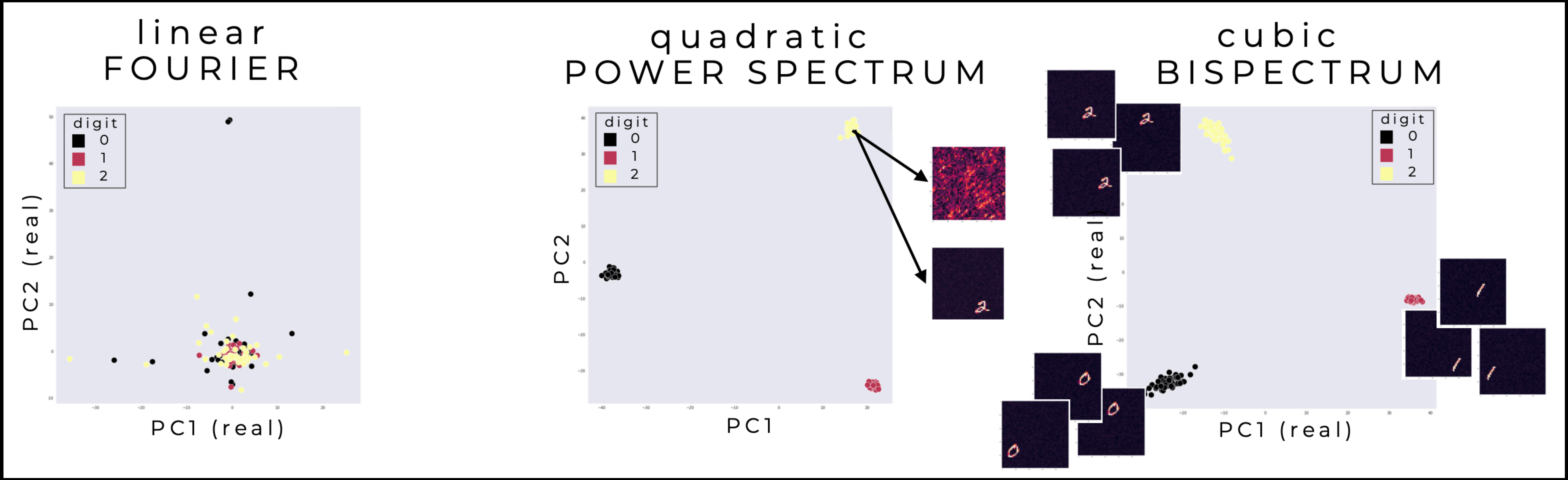
$$\beta(f)_{j,k} = \hat{f}_j^* \hat{f}_k^* \hat{f}_{j+k}$$

Bispectrum

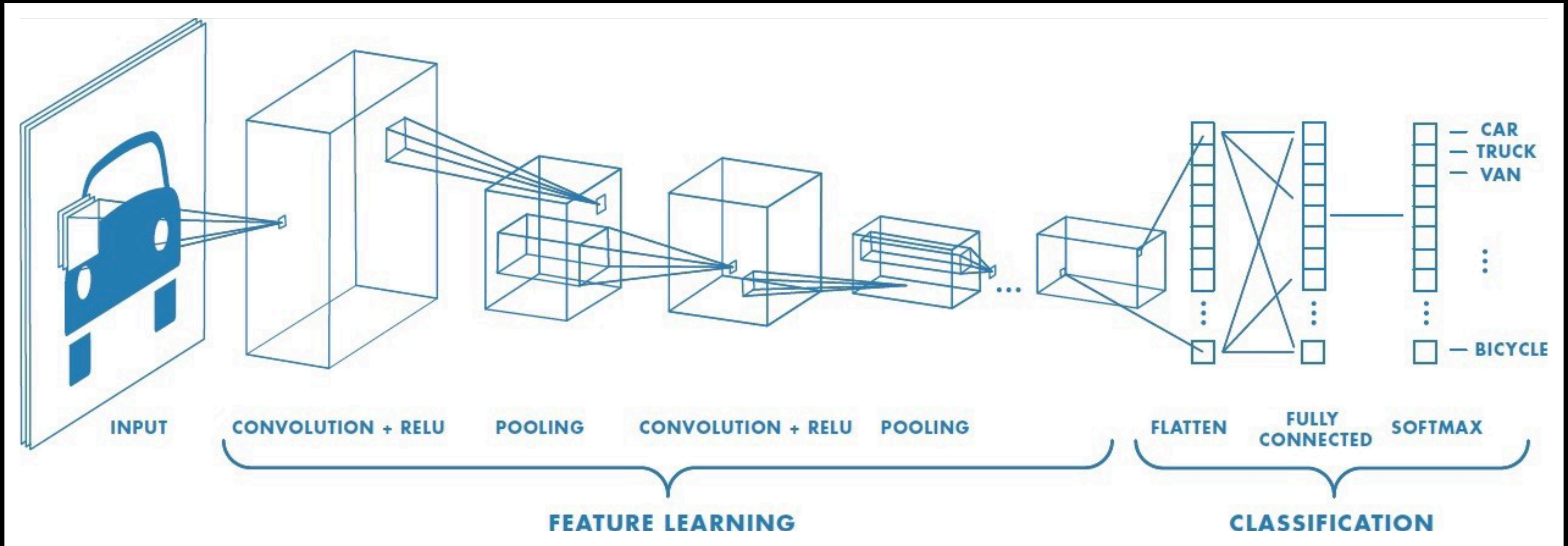
$$\beta(f)_{j,k} = \hat{f}_j^* \hat{f}_k^* \hat{f}_{j+k}$$

$$\beta(f)_{j,k} = \left(\sum_{t=0}^{\tau} e^{\mathrm{i}2\pi tj/\tau} f_t \right) \left(\sum_{t=0}^{\tau} e^{\mathrm{i}2\pi xk/\tau} f_t \right) \left(\sum_{t=0}^{\tau} e^{-\mathrm{i}2\pi t(j+k)/\tau} f_t \right)$$

The Mathematics of Invariance

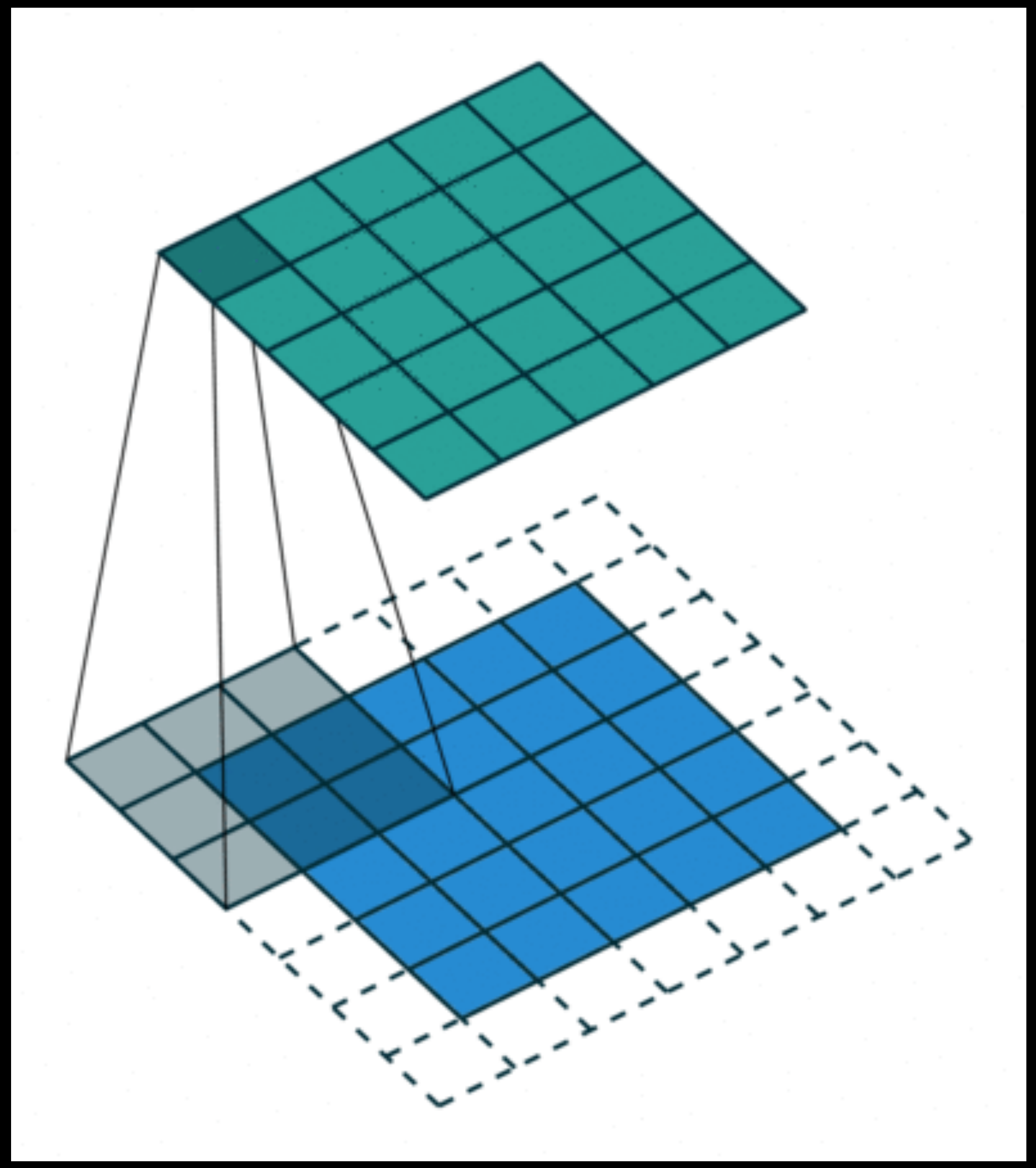


Invariance in Computer Vision

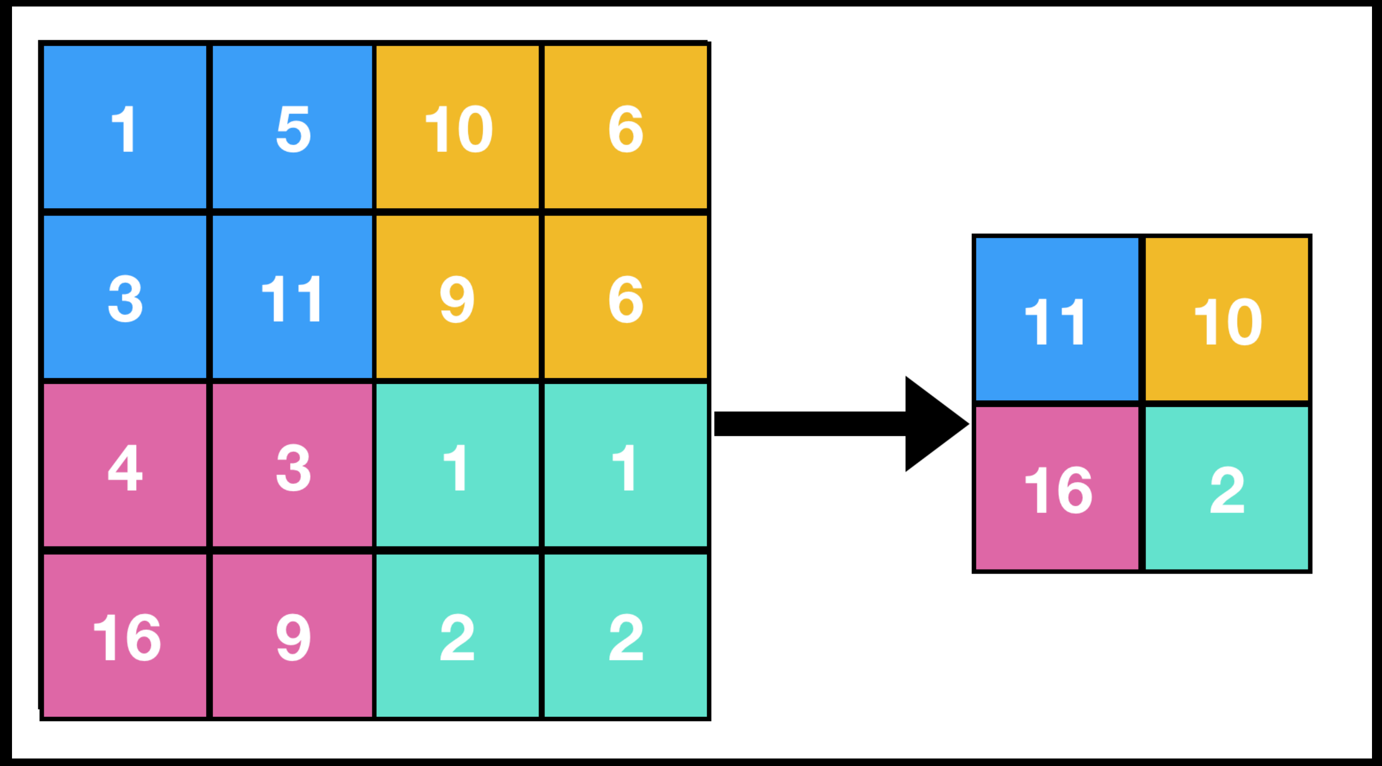


Invariance in Computer Vision

Convolution

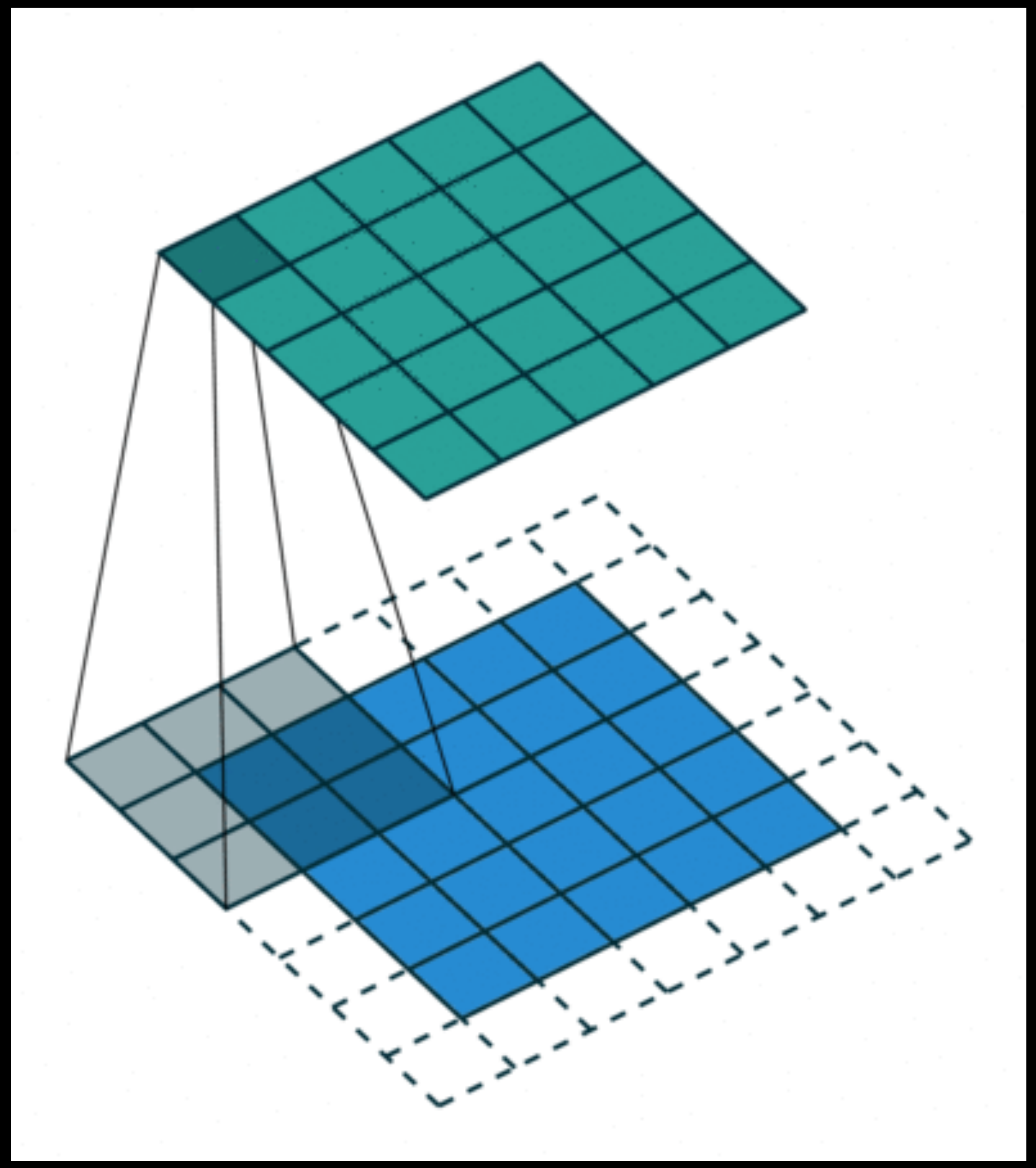


Pooling



Invariance in Computer Vision

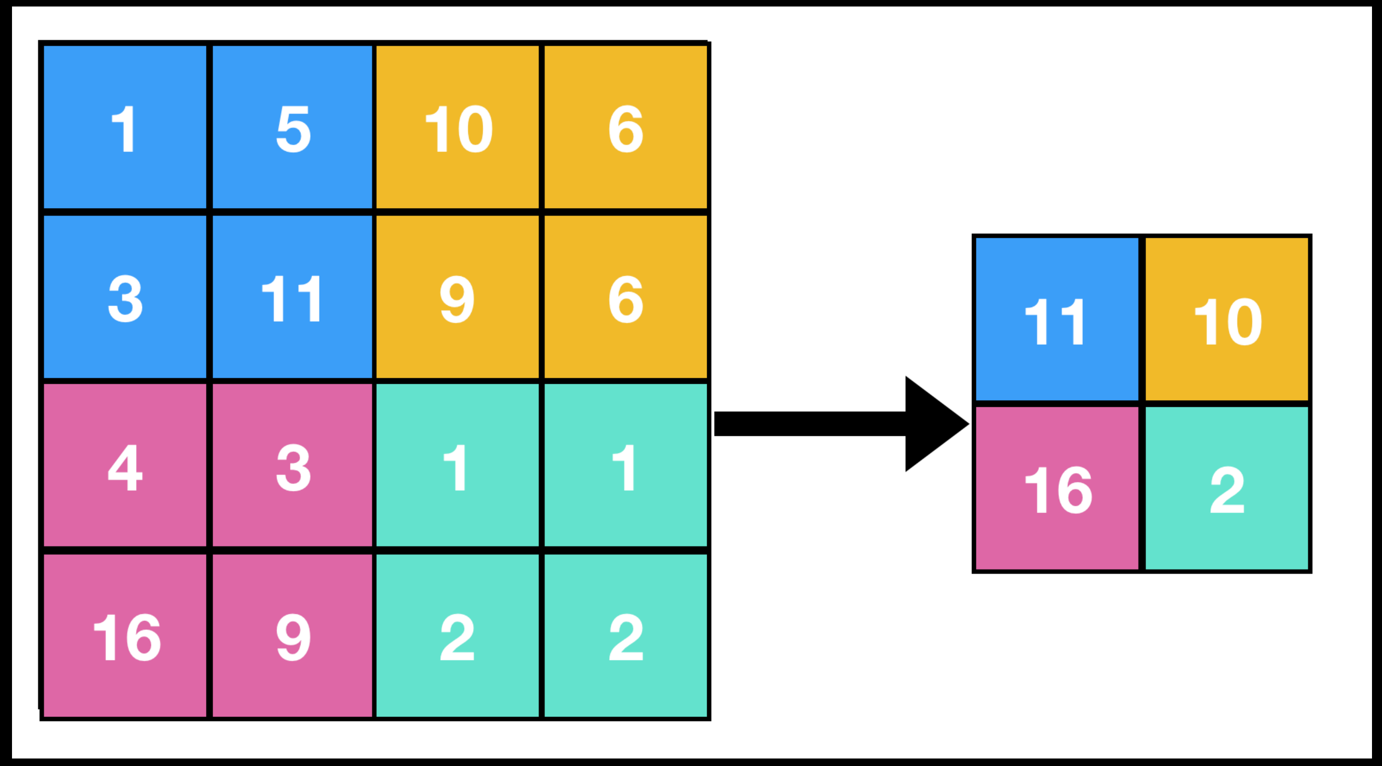
Convolution



Equivariance

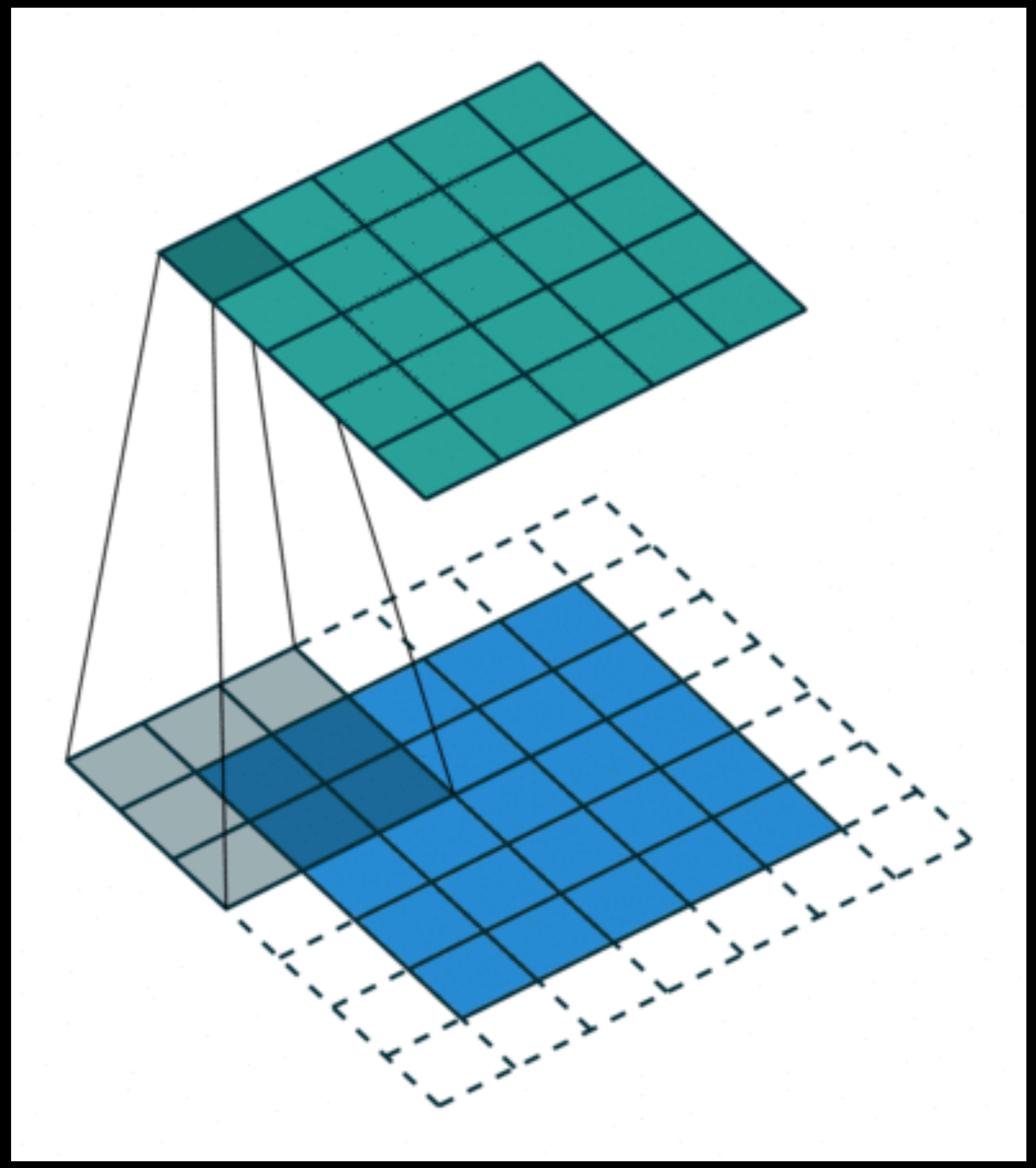
to translation within the global image plane

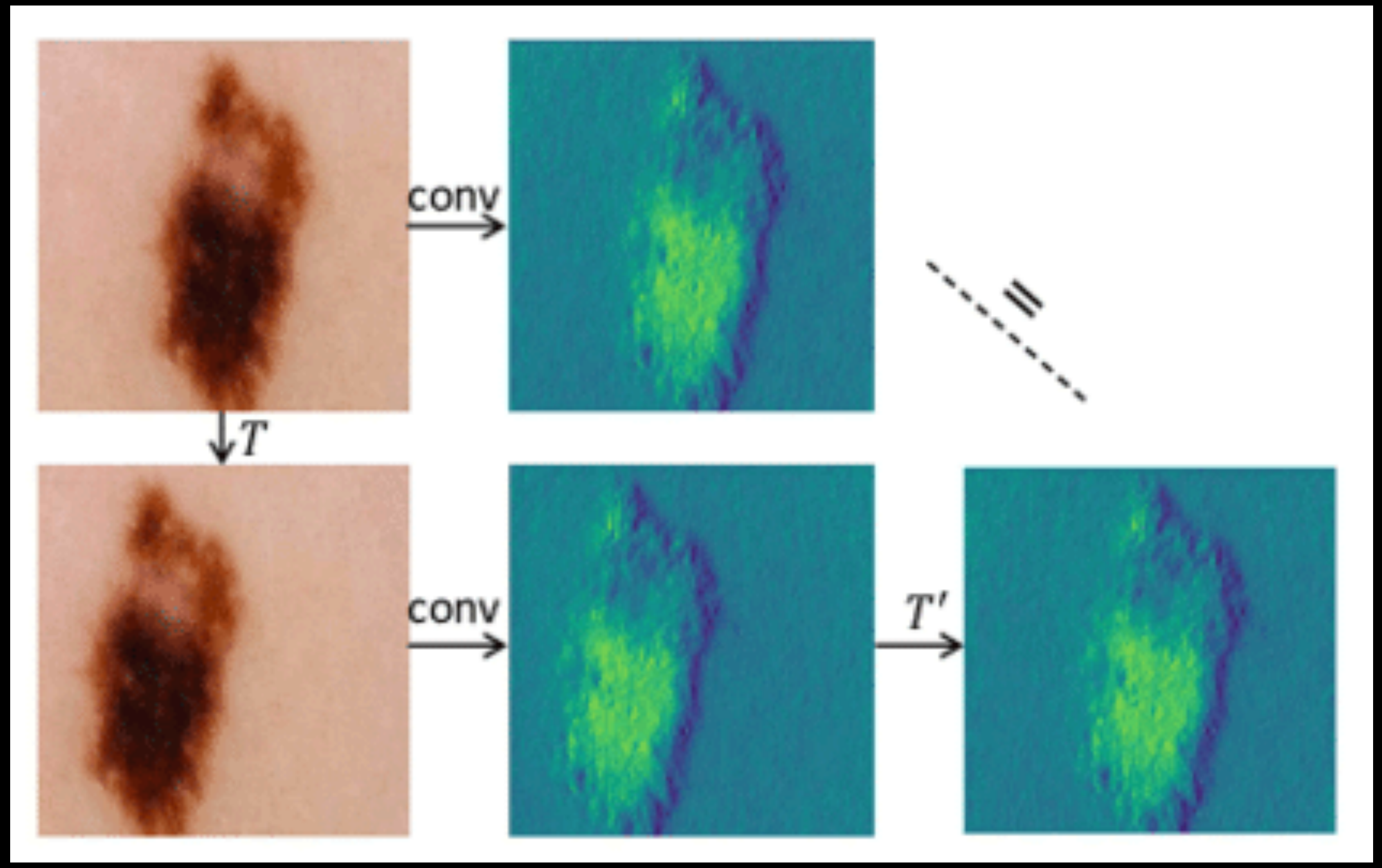
Pooling



Invariance in Computer Vision

Convolution



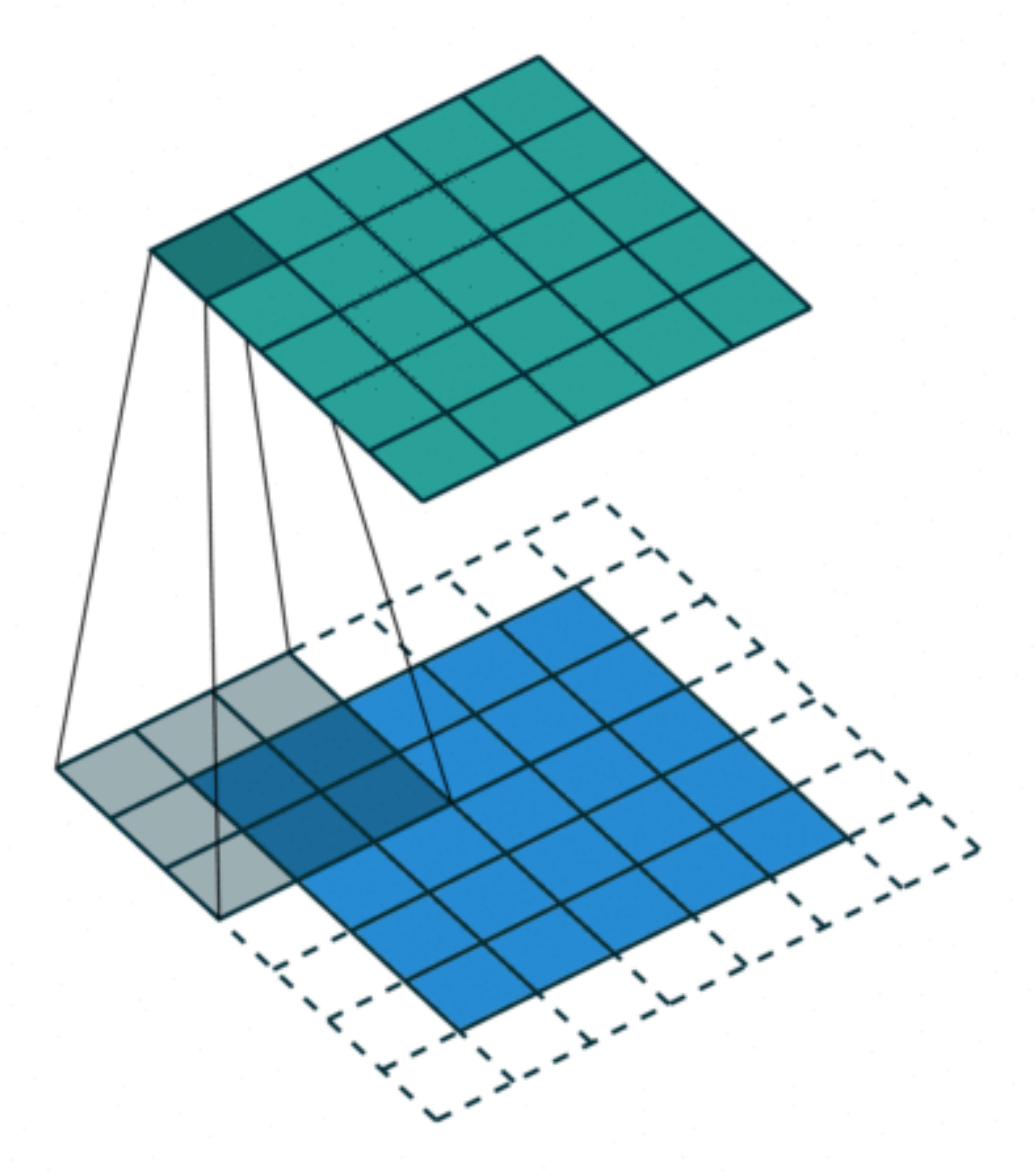


Equivariance

to translation within the global image plane

Invariance in Computer Vision

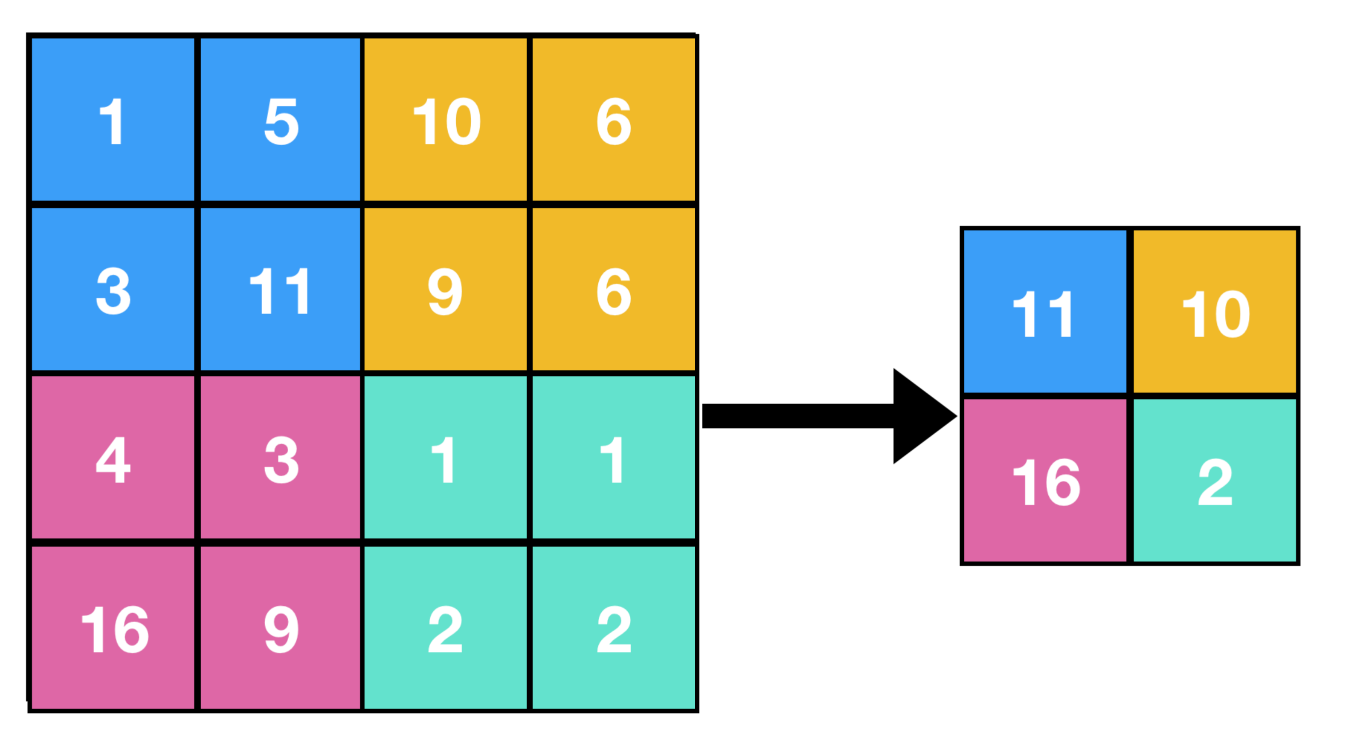
Convolution



Equivariance

to translation within the global image plane

Pooling

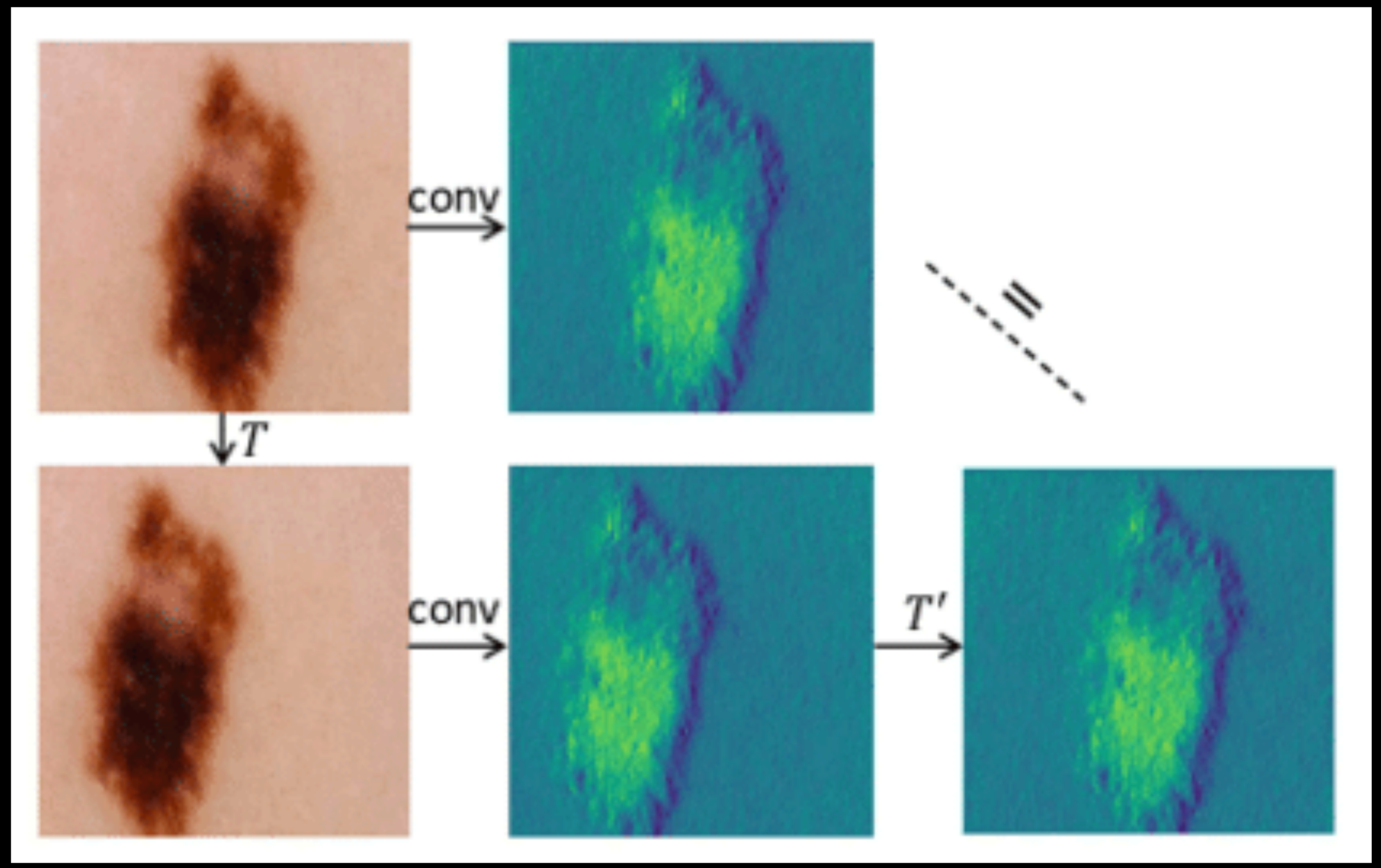


Invariance

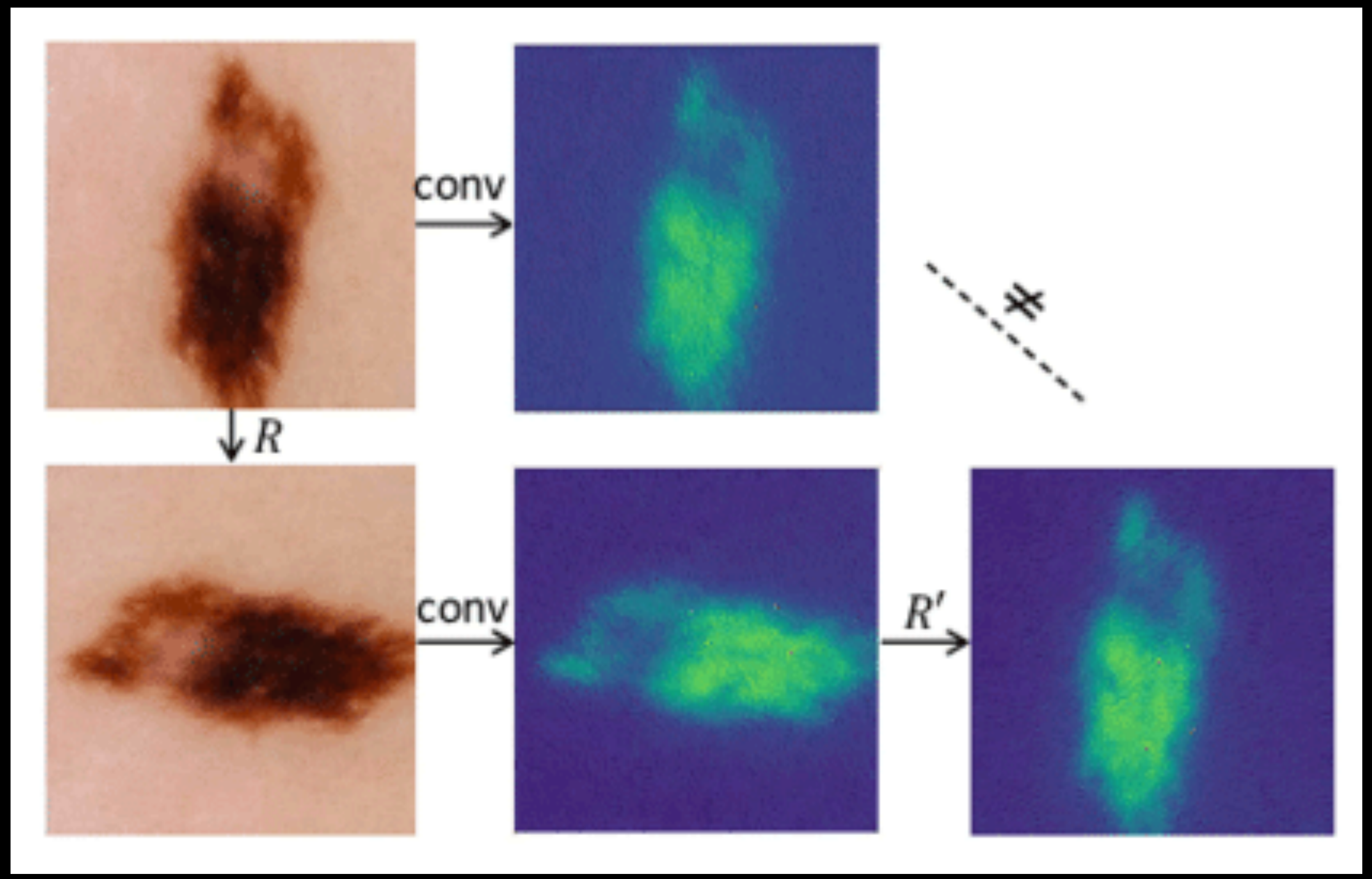
to slight perturbations within a local region

Invariance in Computer Vision

Translation

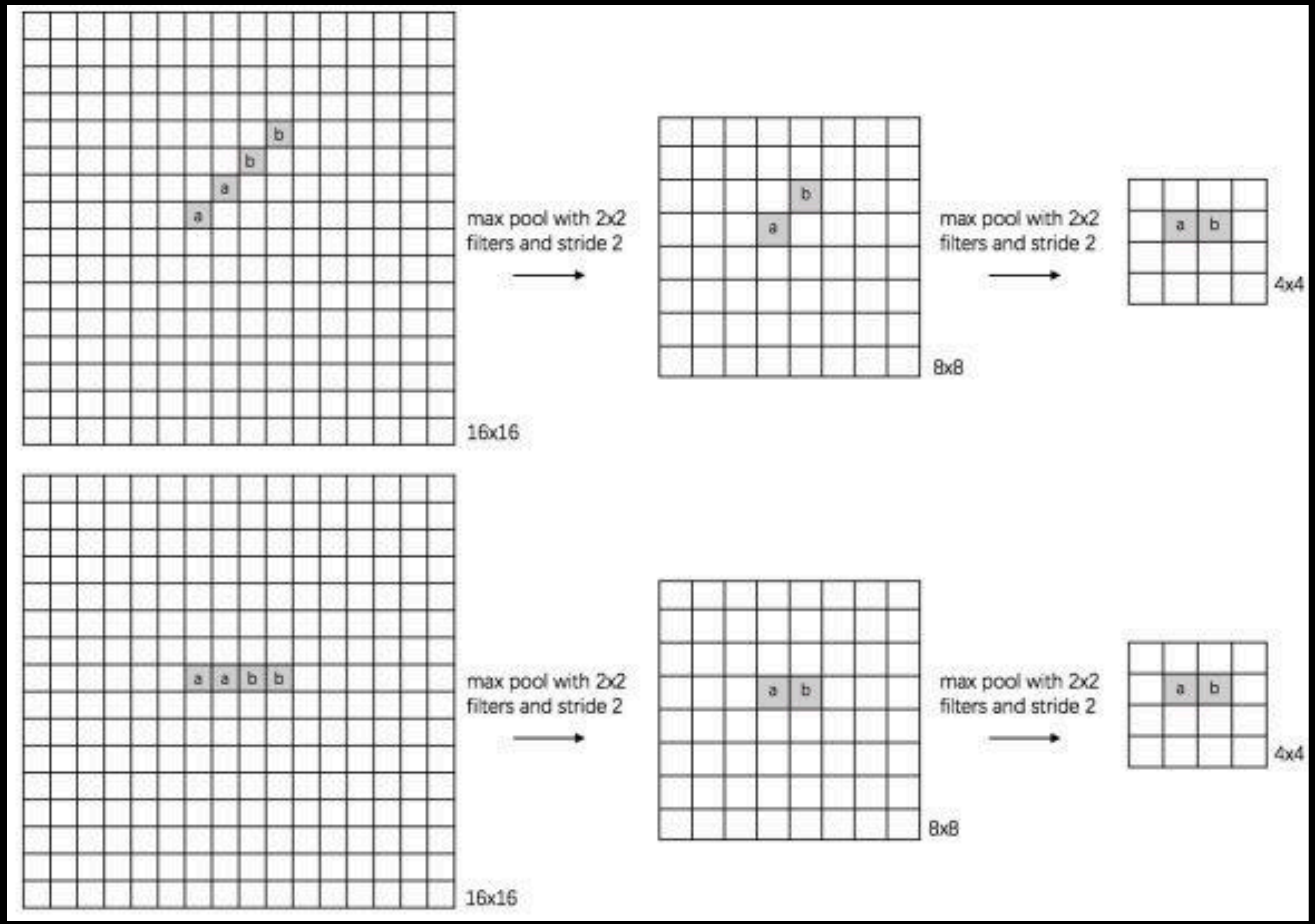


Rotation



Invariance in Computer Vision

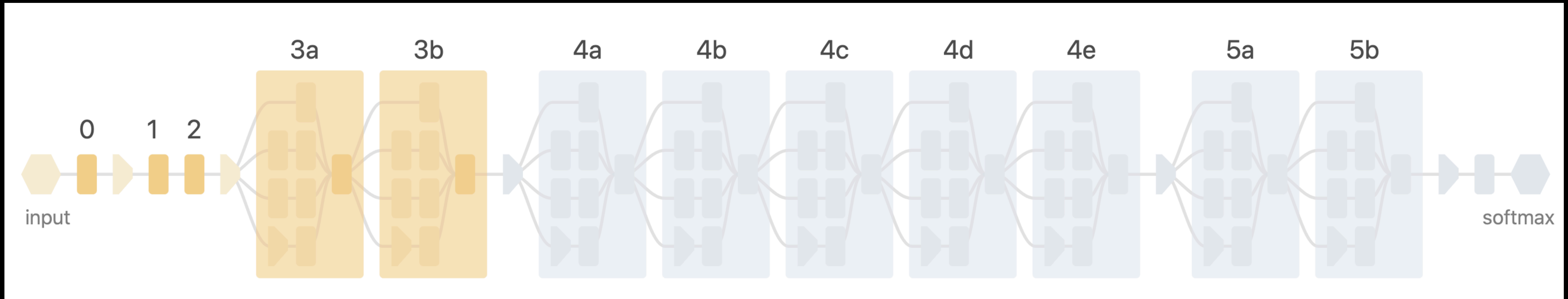
Max Pooling



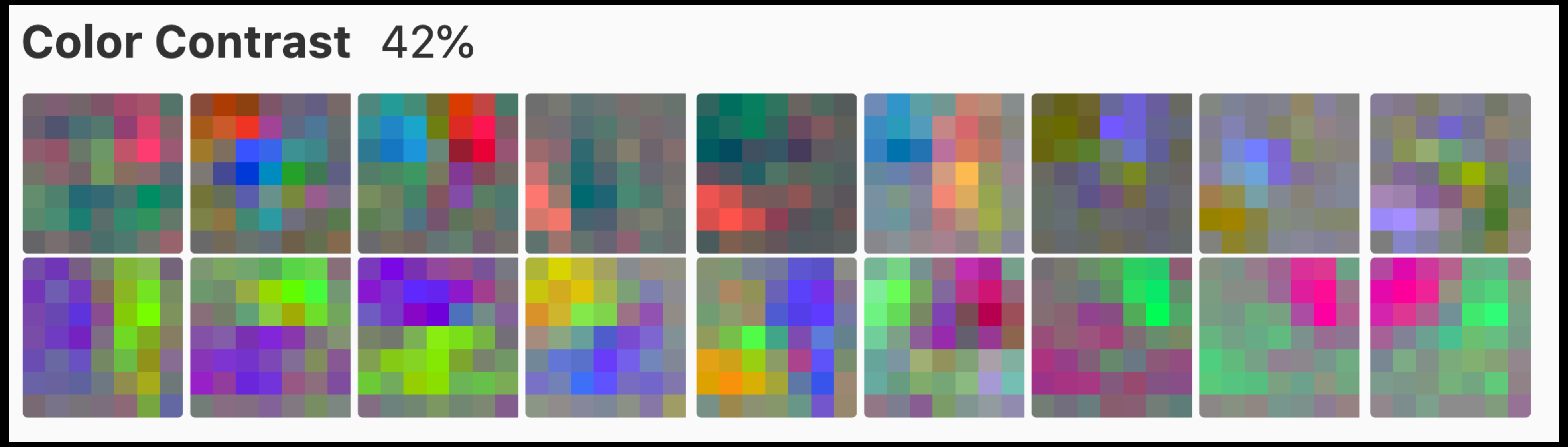
Invariance in Computer Vision

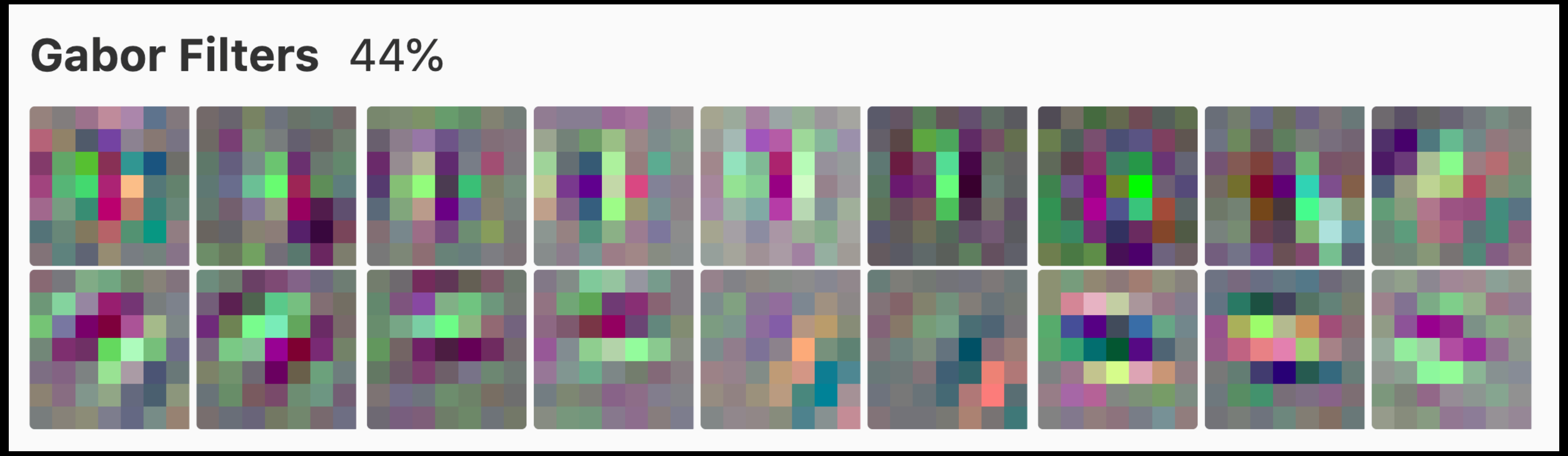
Data Augmentation

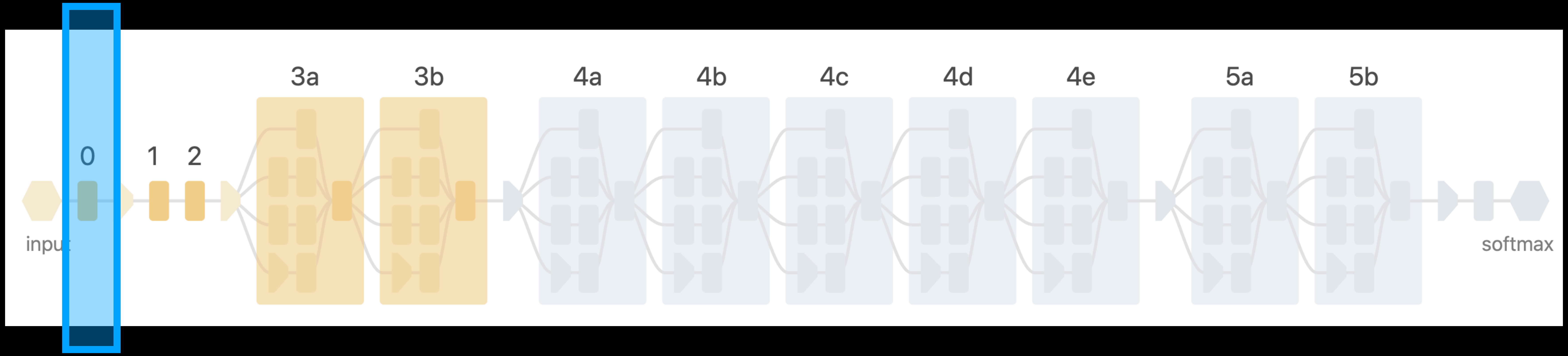
Invariance in Computer Vision



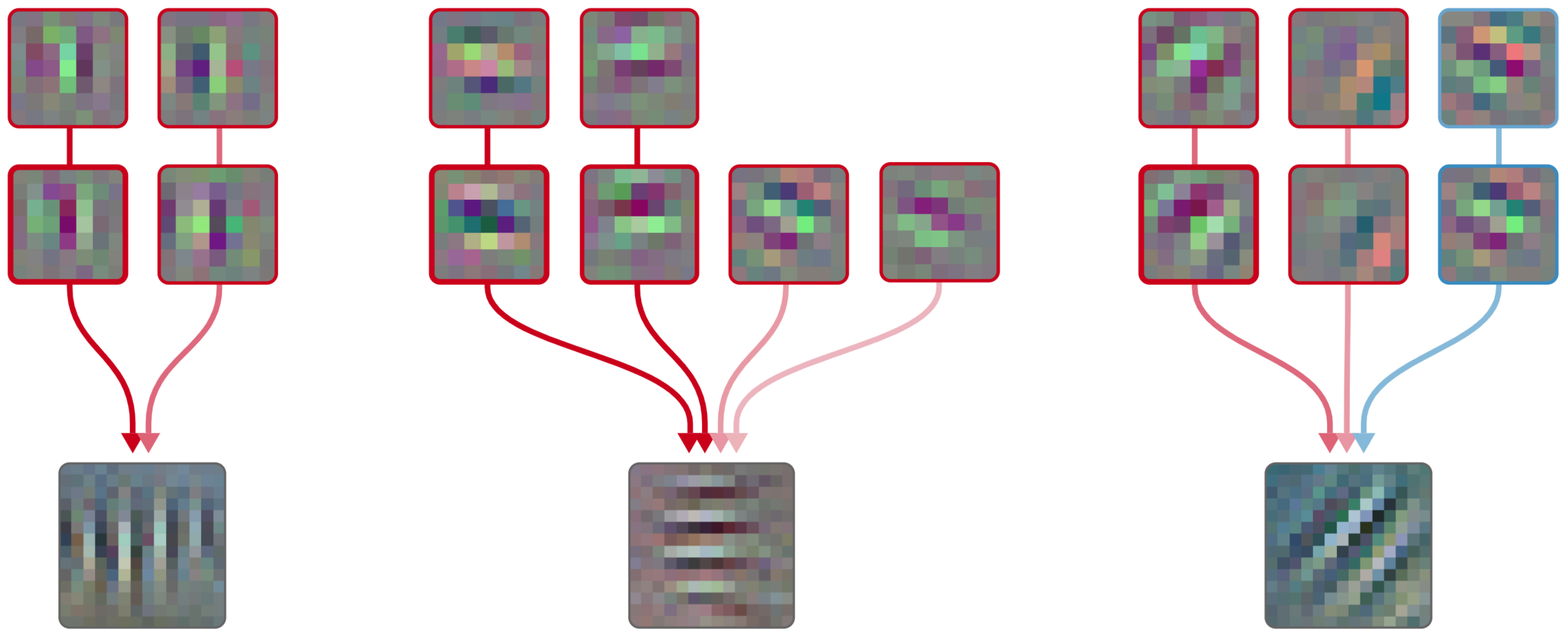
Invariance in Computer Vision



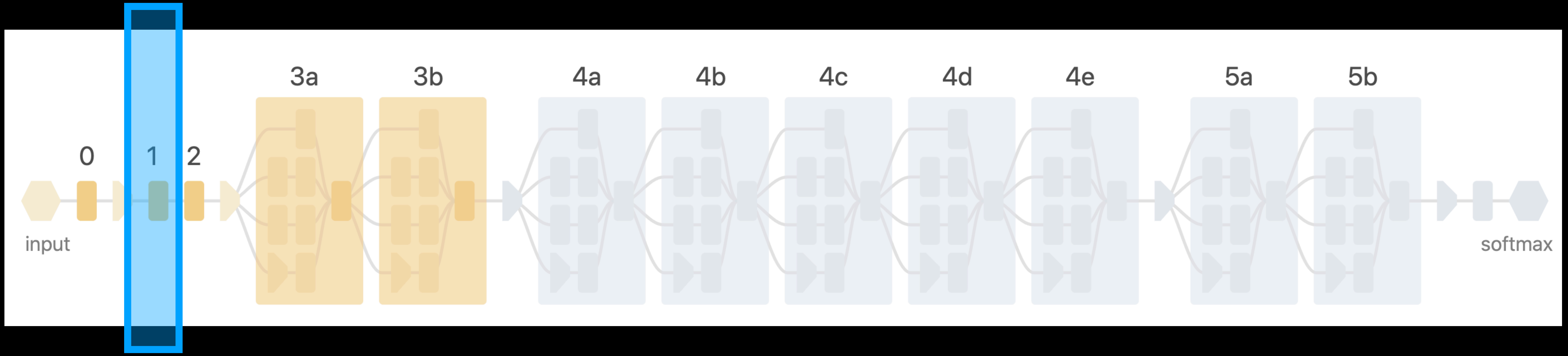




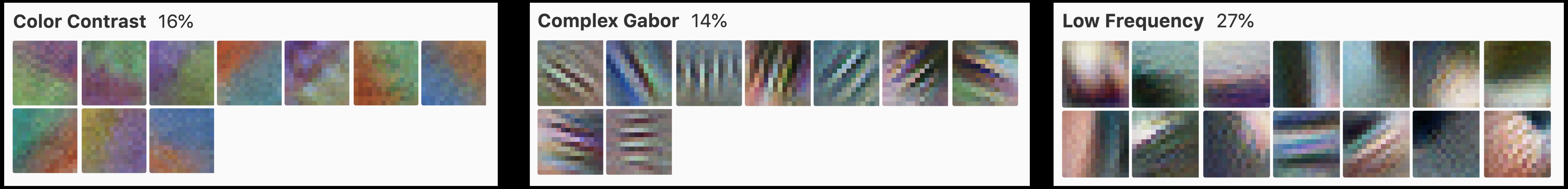
Hierarchical Vision Models

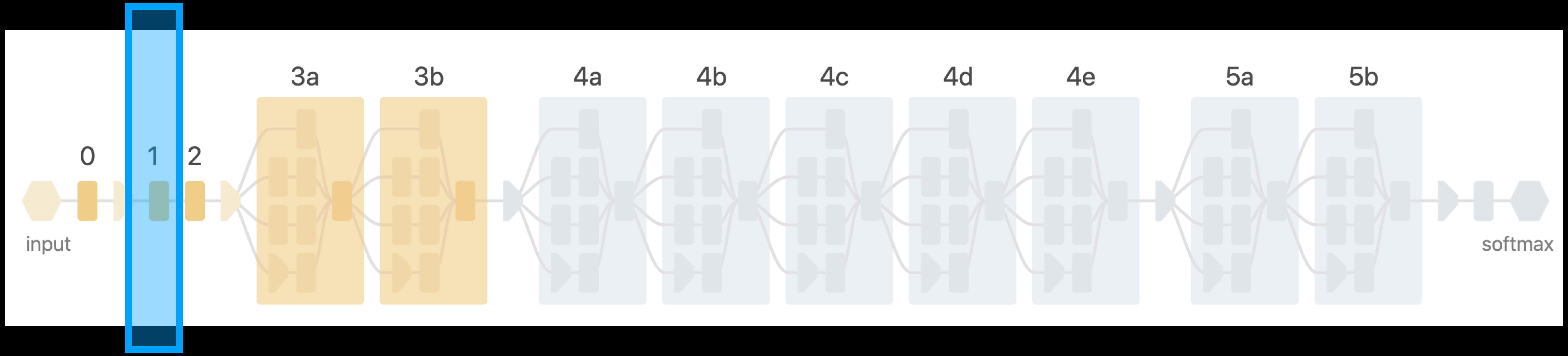


All neurons in the previous layer with at least 30% of the max weight magnitude are shown, both **positive (excitation)** and **negative (inhibition)**. Click on a neuron to see its forwards and backwards weights.

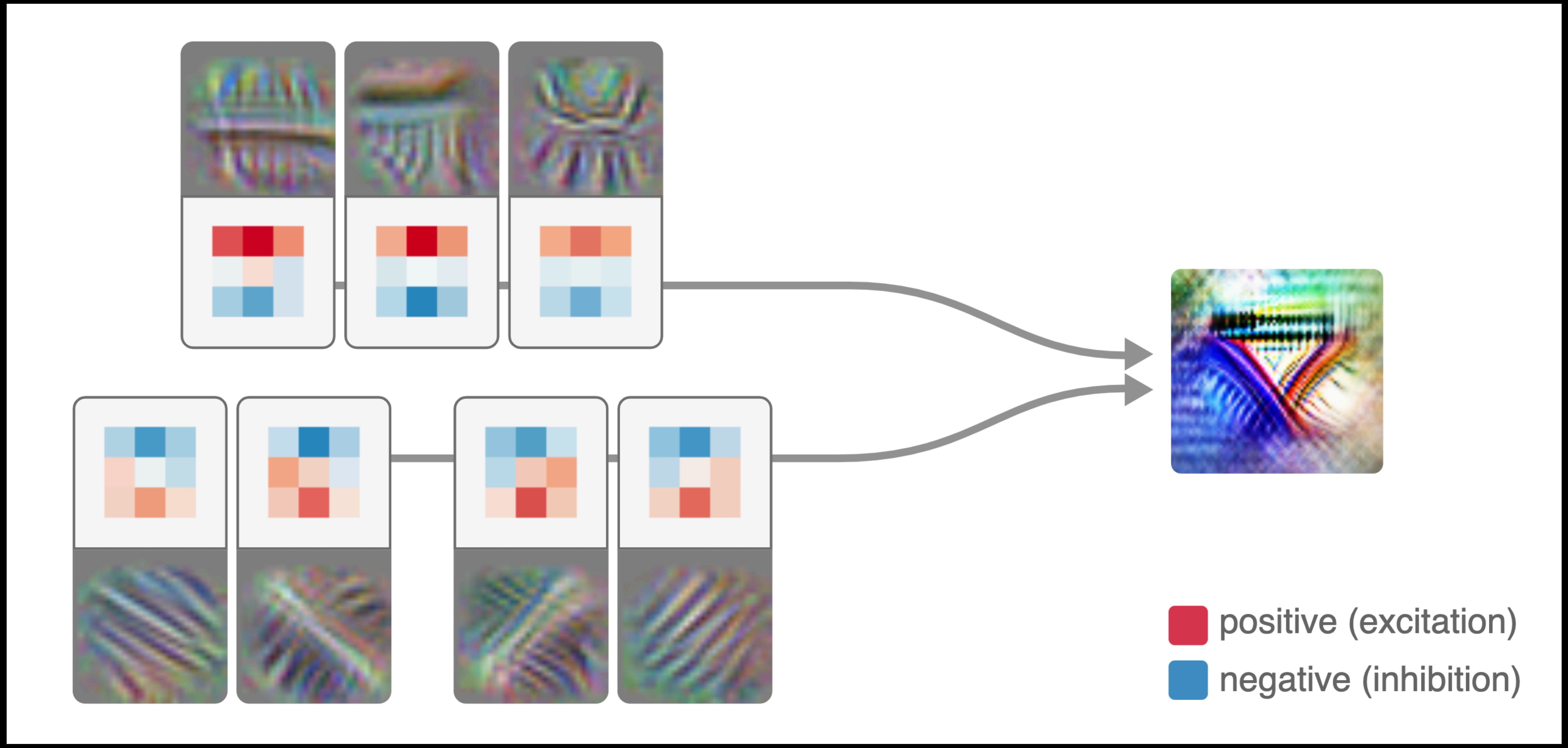


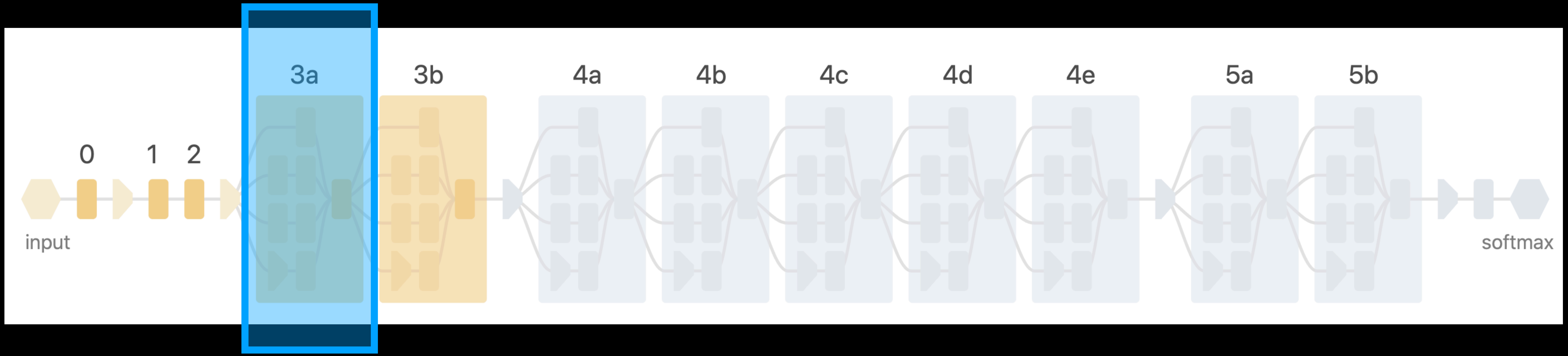
Hierarchical Vision Models



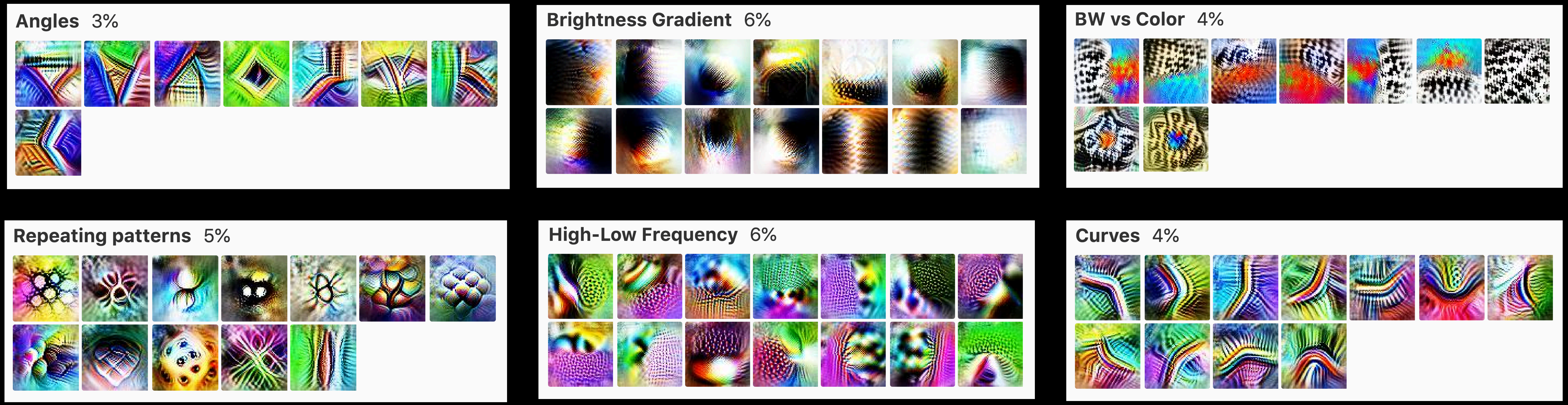


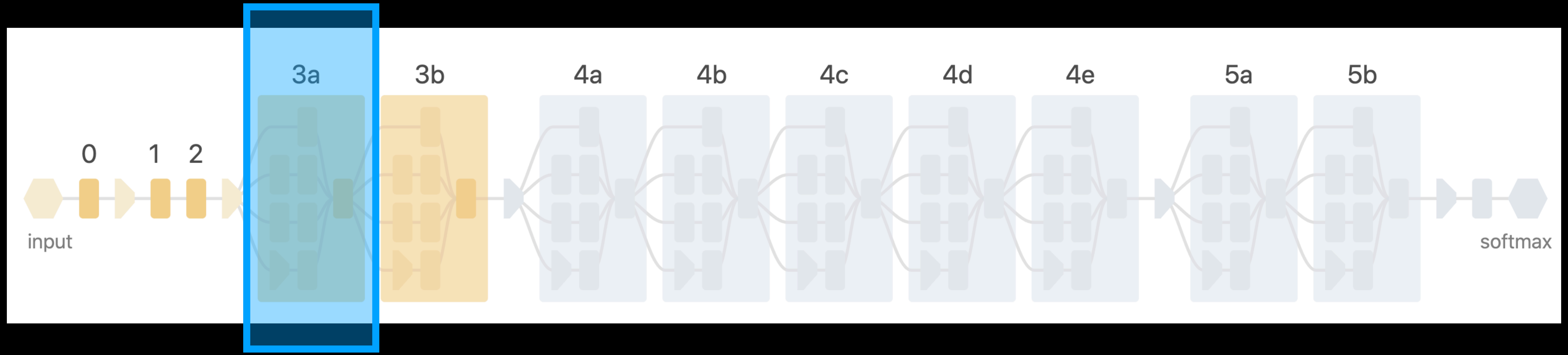
Hierarchical Vision Models



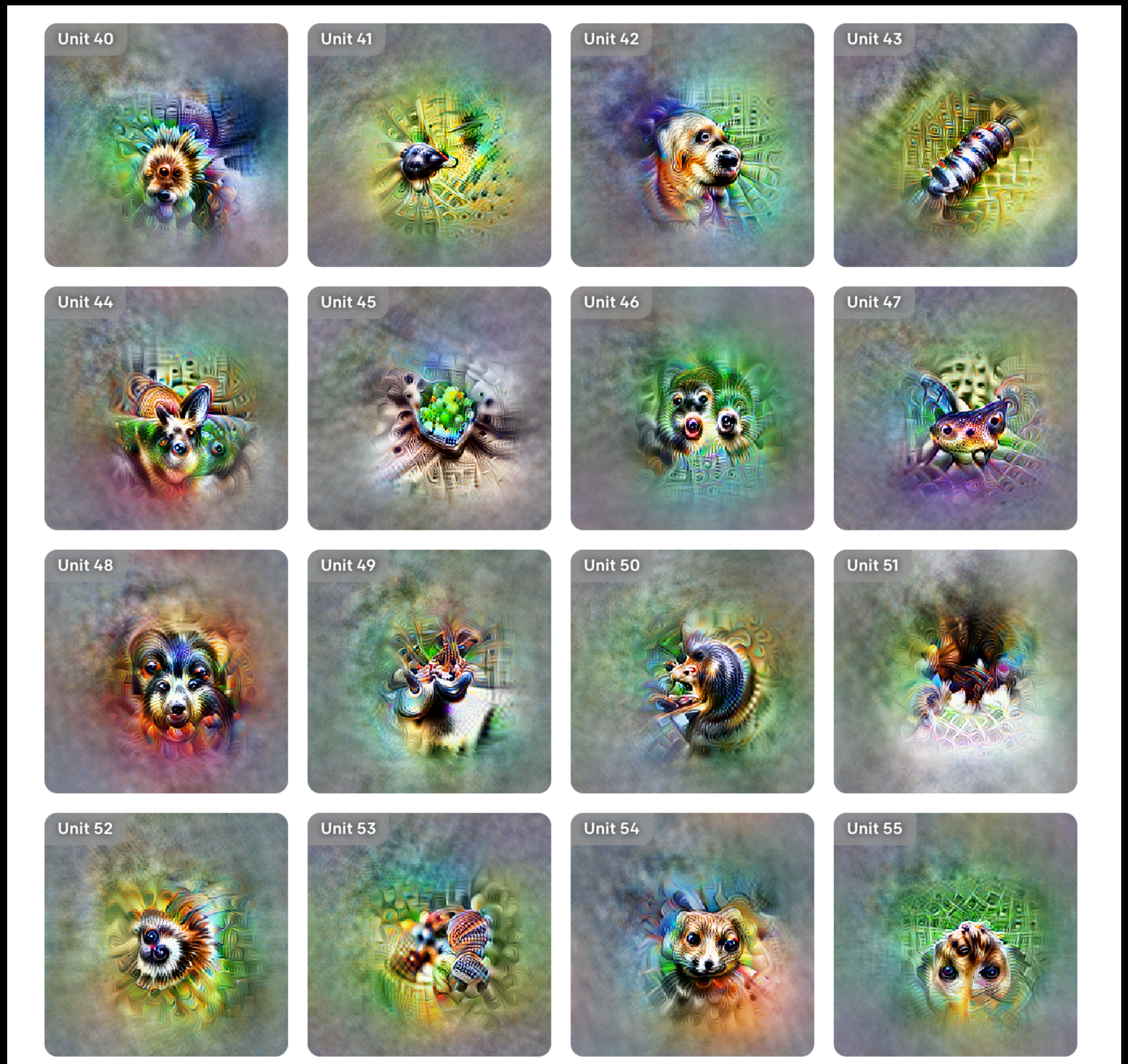


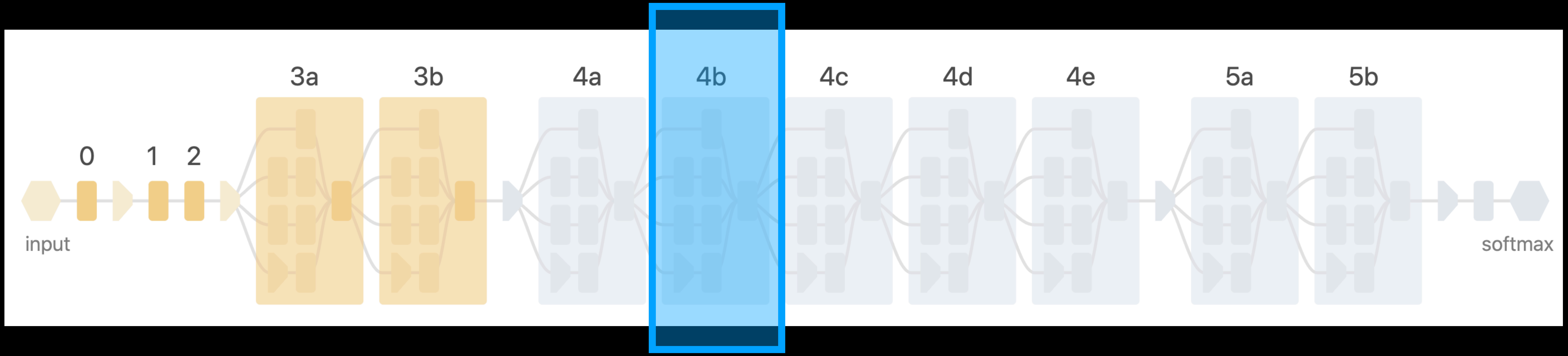
Hierarchical Vision Models





Hierarchical Vision Models

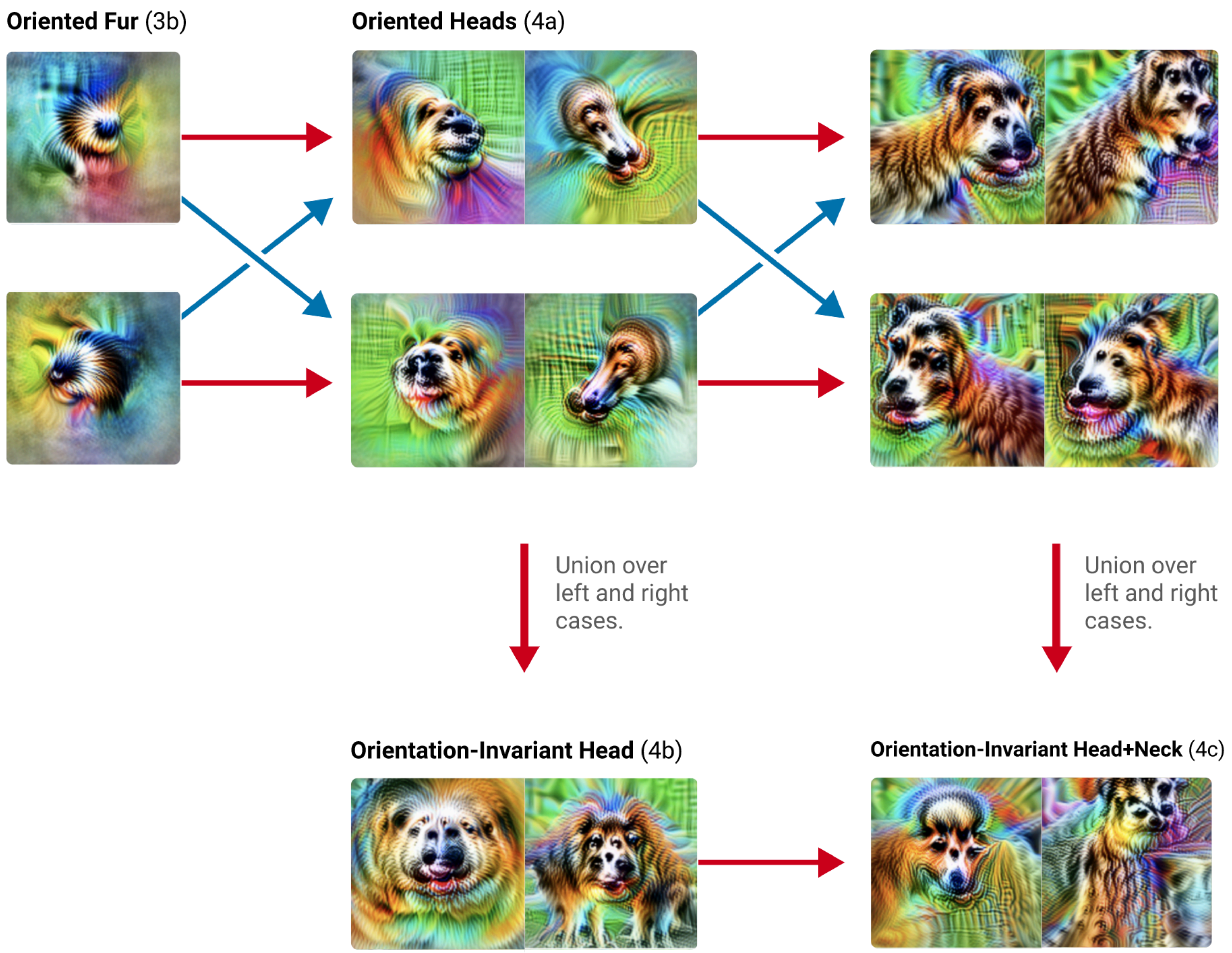




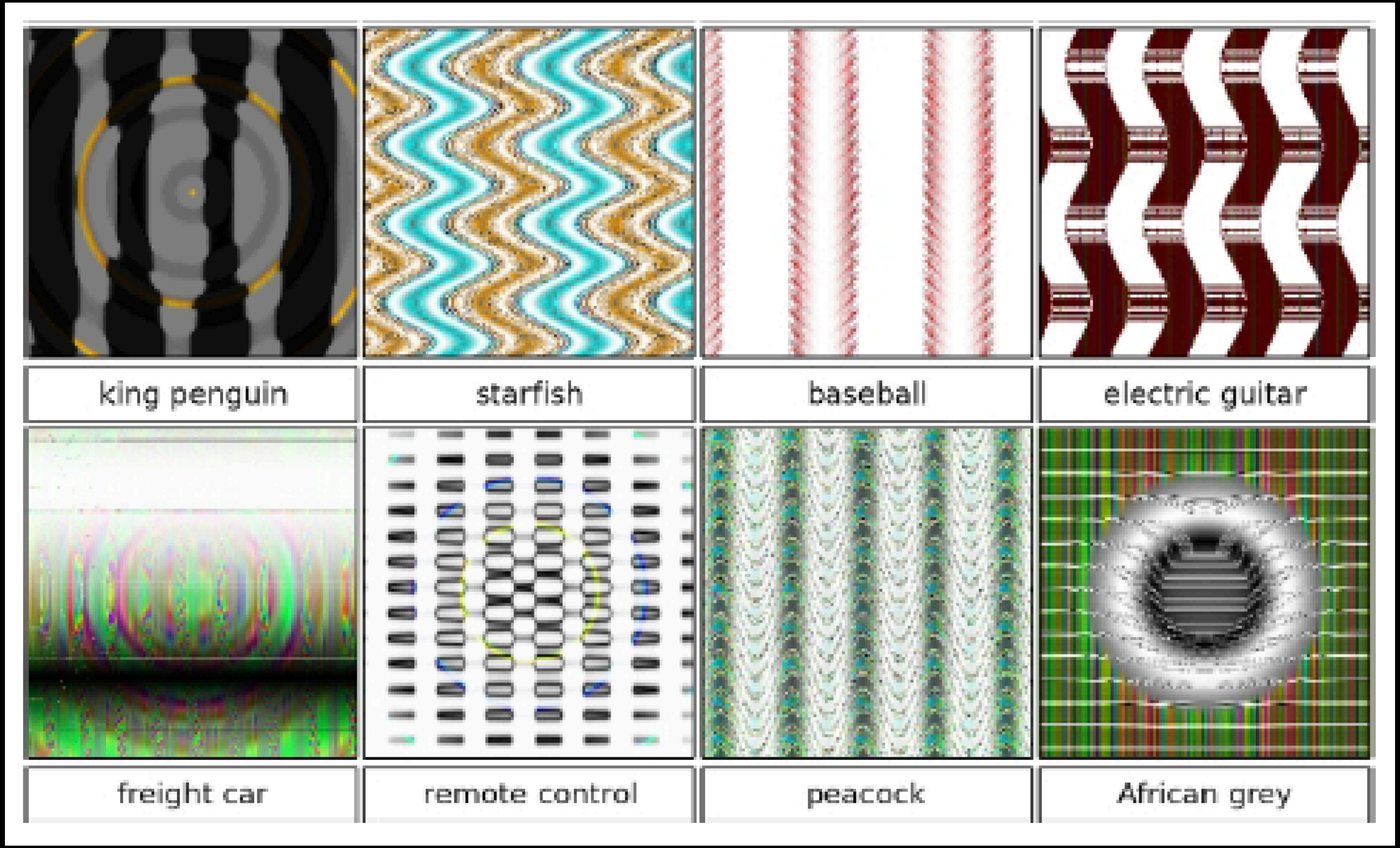
Hierarchical Vision Models

InceptionV1 has a **left-oriented** pathway detecting dogs facing left...

... and a symmetric **right-oriented** pathway detecting dogs facing right. At each step, the two pathways **inhibit** each other and **excite** the next stage.



Too Invariant?



Not Invariant Enough?