Image Processing II

Computer Vision
Fall 2018
Columbia University

Convolution Review

Cross Correlation

$$F[x, y]$$

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	0	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	0	0	0	0	0	0	0
0	0	90	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

$$\frac{1}{9}$$

0	0	0
-1	0	1
0	0	0

$$G[x, y]$$

Cross Correlation

$F[x,y]$

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	0	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	0	0	0	0	0	0	0
0	0	90	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

$\frac{1}{9}$

0	0	0
-1	0	1
0	0	0

$G[x,y]$

	0								

Cross Correlation

$$F[x, y]$$

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	0	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	0	0	0	0	0	0	0
0	0	90	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

$$\frac{1}{9}\begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad G[x, y]$$

	0	0							

Cross Correlation

$F[x, y]$

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	0	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	0	0	0	0	0	0	0
0	0	90	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

$$\frac{1}{9}\begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad G[x, y]$$

	0	0	0						

Cross Correlation

$F[x, y]$

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	0	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	0	0	0	0	0	0	0
0	0	90	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

$$\frac{1}{9} \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad G[x, y]$$

	0	0	0						

Cross Correlation

$F[x, y]$

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	0	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	0	0	0	0	0	0	0
0	0	90	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

$$\frac{1}{9}\begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} G[x, y]$$

	0	0	0						
			90						

Cross Correlation

$F[x, y]$

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	0	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	0	0	0	0	0	0	0
0	0	90	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

$$\frac{1}{9}\begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} G[x, y]$$

	0	0	0						
			90						

Cross Correlation

$F[x, y]$

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	0	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	0	0	0	0	0	0	0
0	0	90	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

$$\frac{1}{9}\begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad G[x, y]$$

	0	0	0						
			90	0					

Convolution

$F[x, y]$

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	90	0	90	90	90	0	0
0	0	0	90	90	90	90	90	0	0
0	0	0	0	0	0	0	0	0	0
0	0	90	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

$\frac{1}{9}$

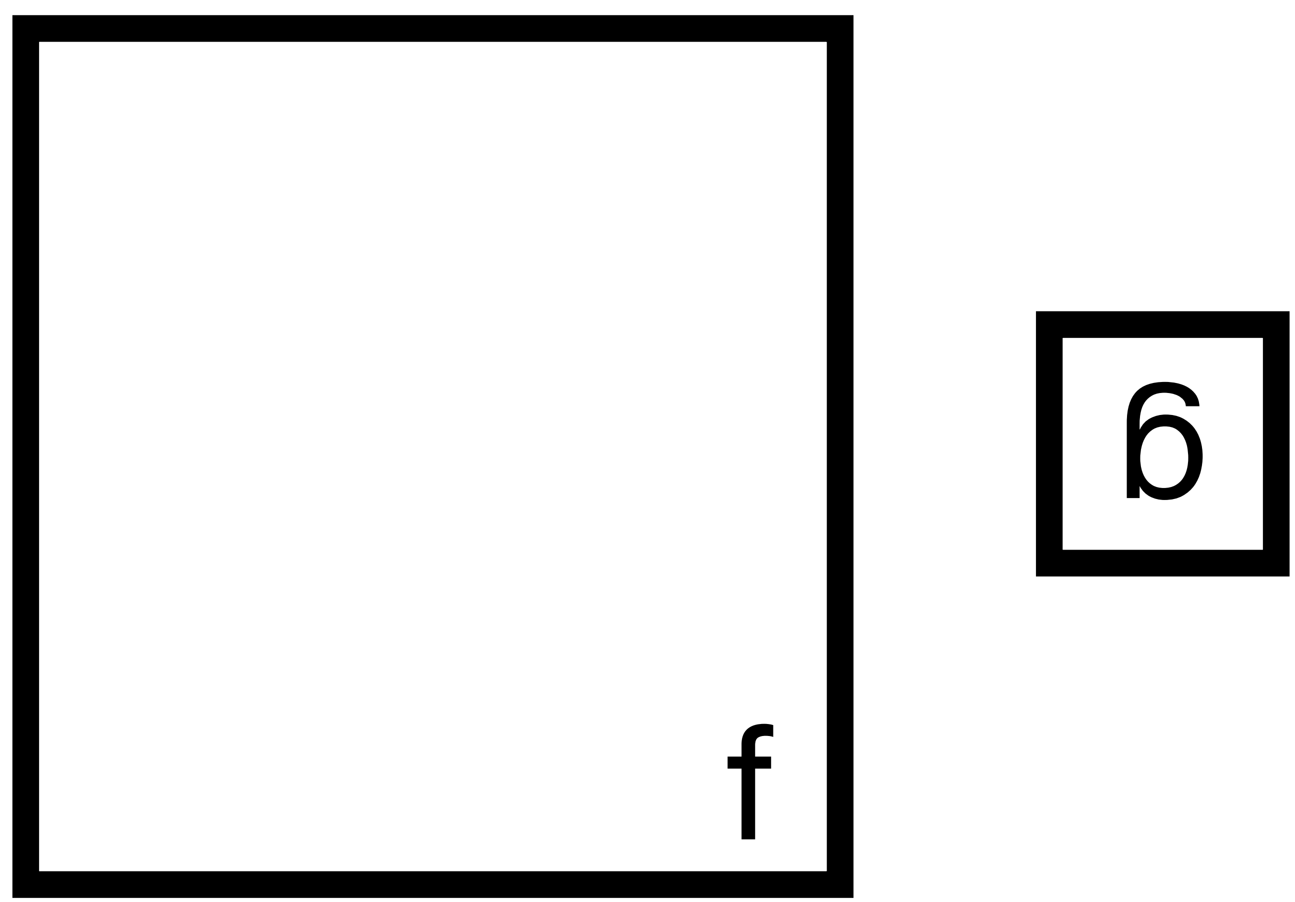
0	0	0
1	0	-1
0	0	0

$G[x, y]$

	0	0	0						
			90	0					

Convolution

$$(f * g)[x, y] = \sum_{i,j} f[x - i, y - j]g[i, j]$$



Flip LR, UD

We flip for the nice properties

Commutative:

$$F * H = H * F$$

Associative:

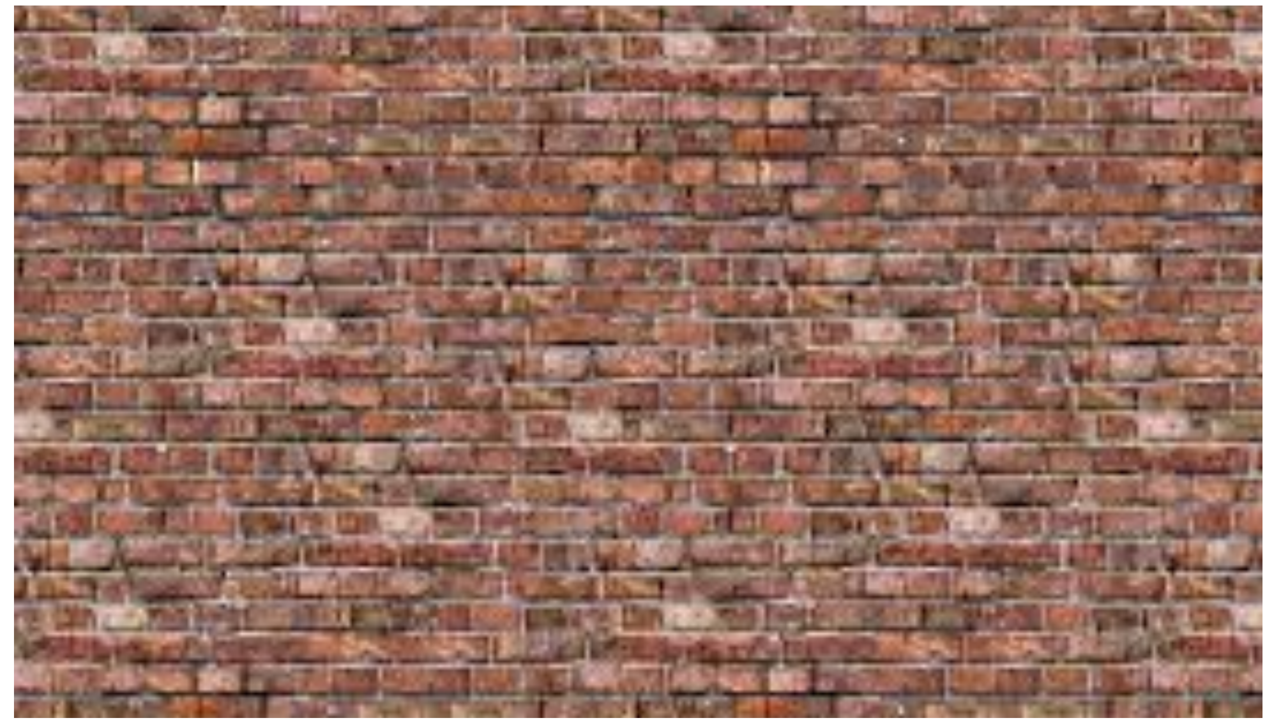
$$(F * H) * G = F * (H * G)$$

Distributive:

$$(F * G) + (H * G) = (F + H) * G$$

Fourier Transforms

Vision is Repetitive

Joseph Fourier

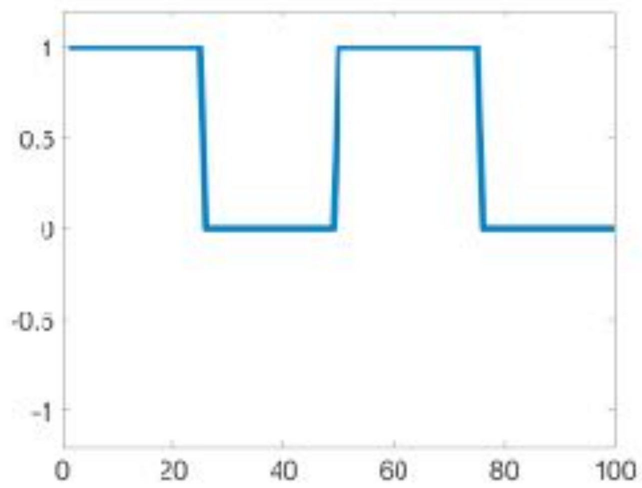
A bold idea (1807):
Any univariate function can be rewritten as a weighted sum of sines and cosines of different frequencies.

Don't believe it?
Neither did Lagrange, Laplace, Poisson and other bigwigs

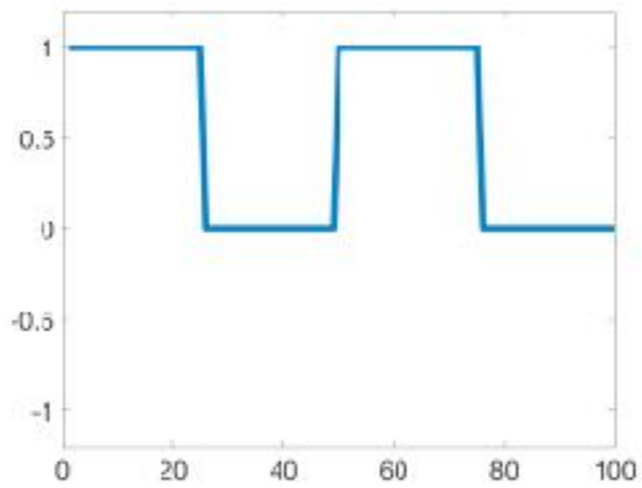
Not translated into English until 1878!

Wikipedia

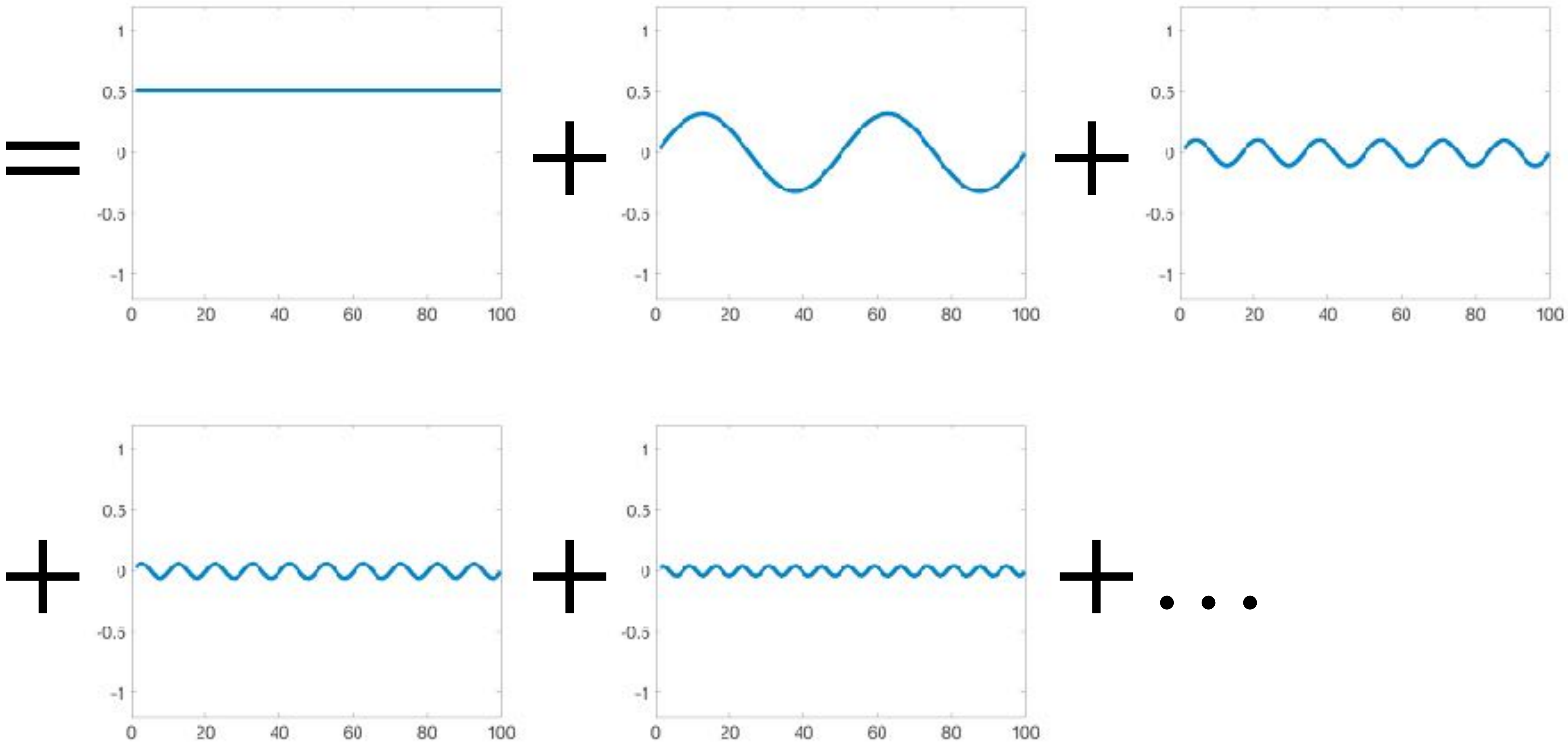
Source: James Hays

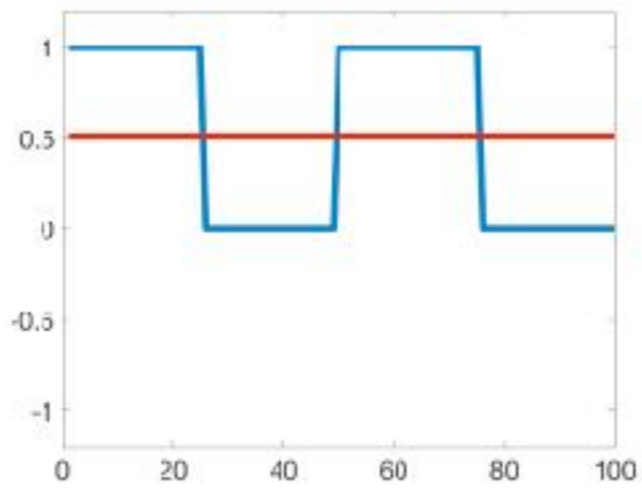


How to build this 1D signal using sin waves?

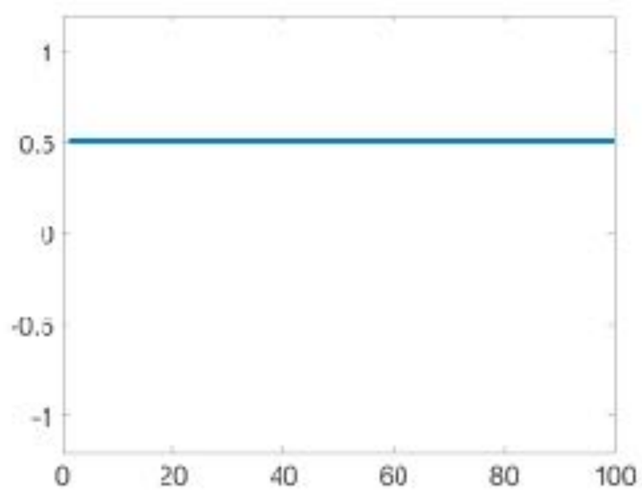


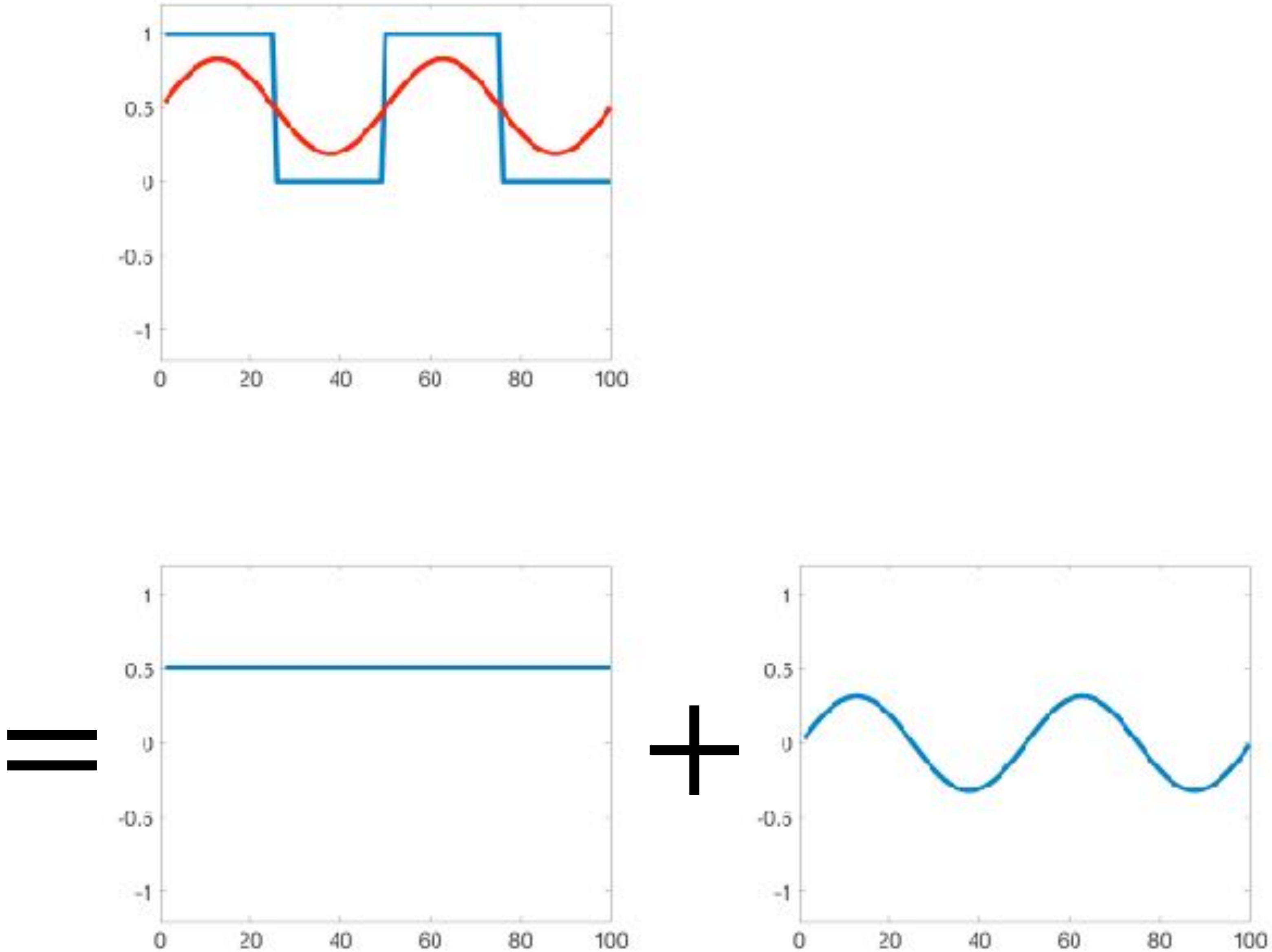
How to build this 1D signal
using sin waves?

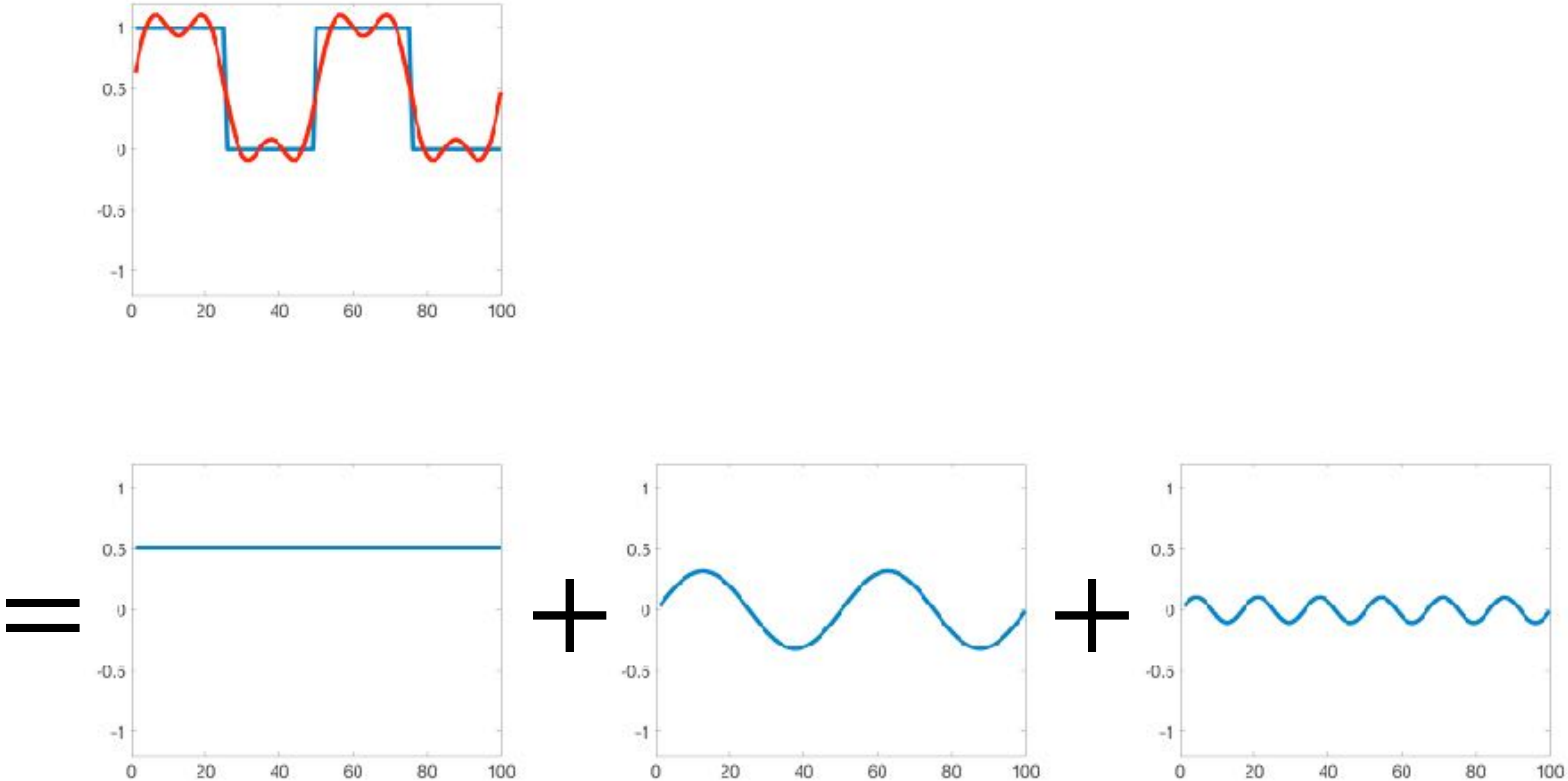


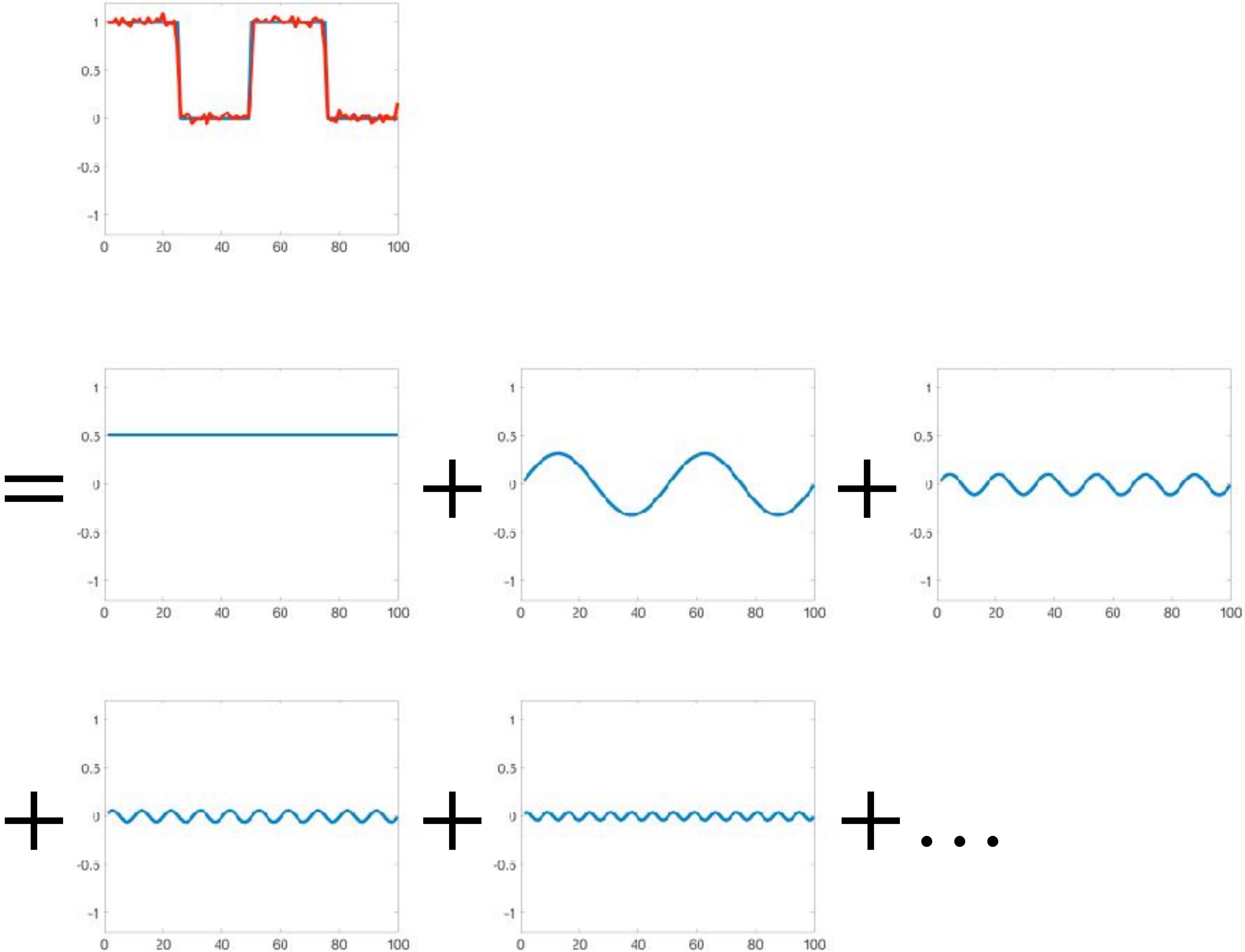

1
0.5
0
-0.5
-1
0 20 40 60 80 100

=


1
0.5
0
-0.5
-1
0 20 40 60 80 100

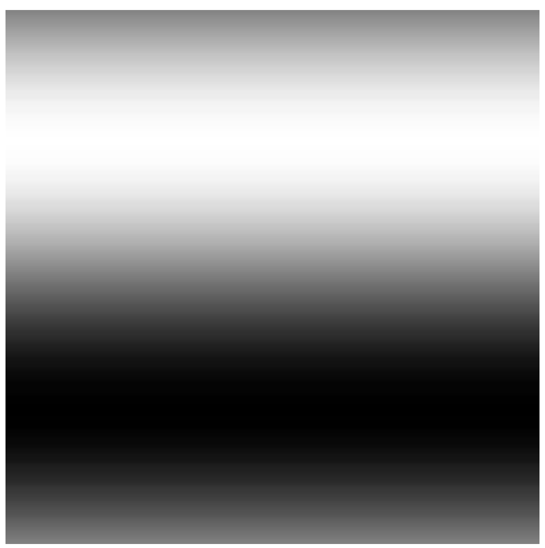

1
0.5
0
-0.5
-1
0
20
40
60
80
100
=
+


1
0.5
0
-0.5
-1
0
20
40
60
80
100
=
+
+

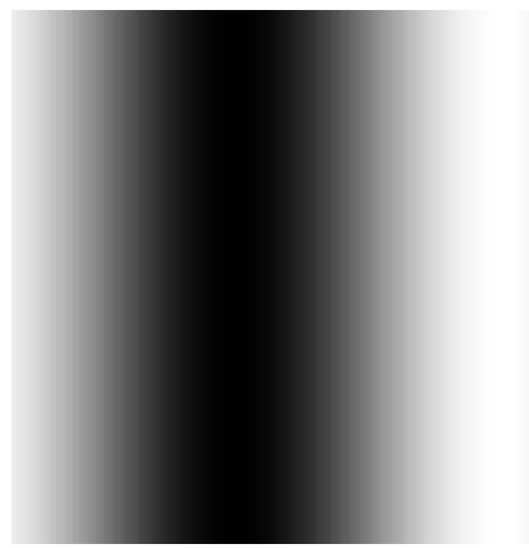

1
0.5
0
-0.5
-1
0 20 40 60 80 100
=
+
+
+
+
+ . . .

Where we are going…

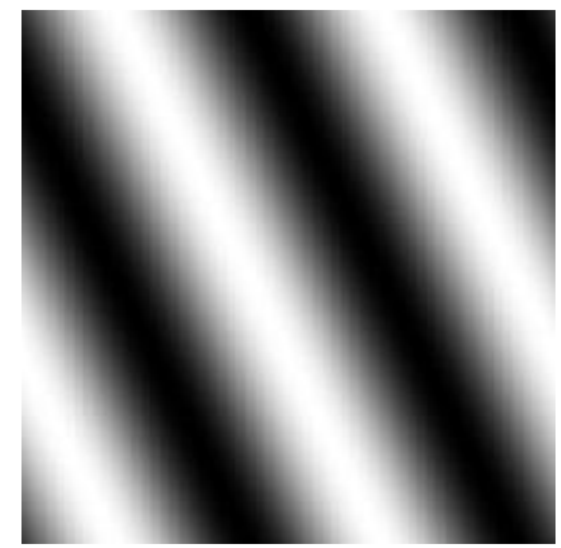
=



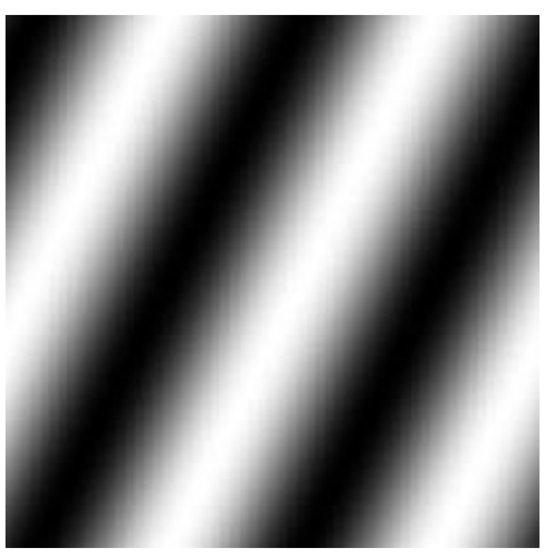
+



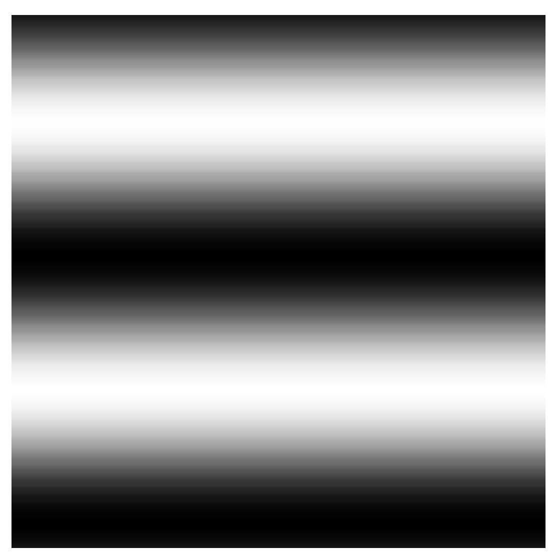
+



+



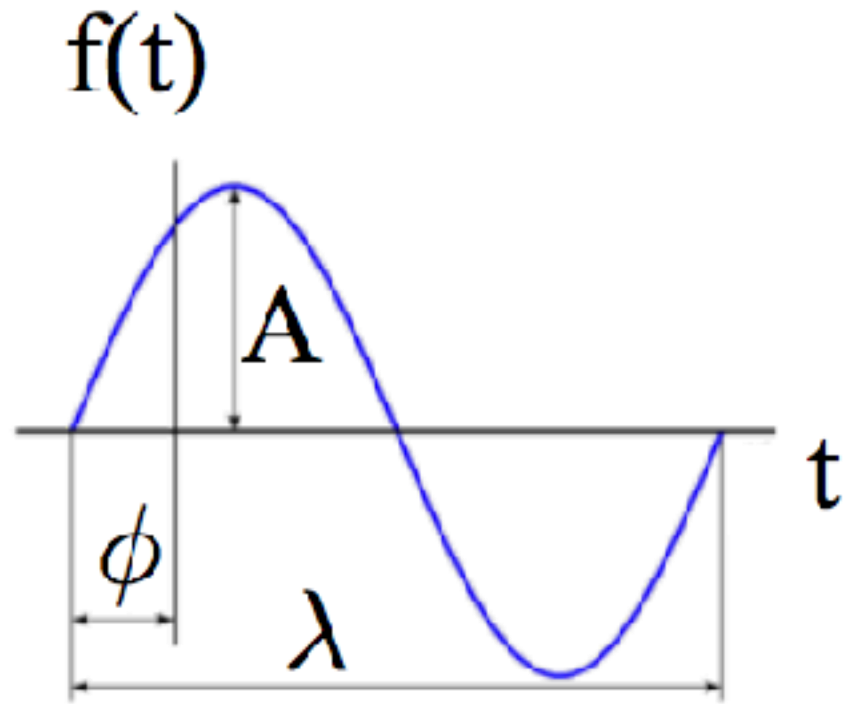
+



+…

Background: Sinusoids

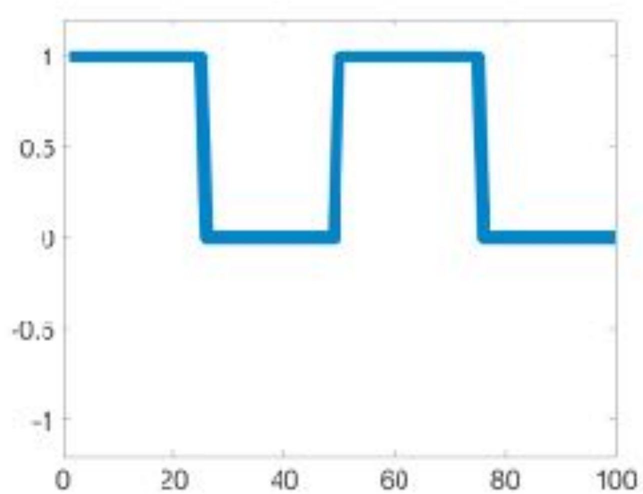
A = amplitude
ϕ = phase
f = frequency



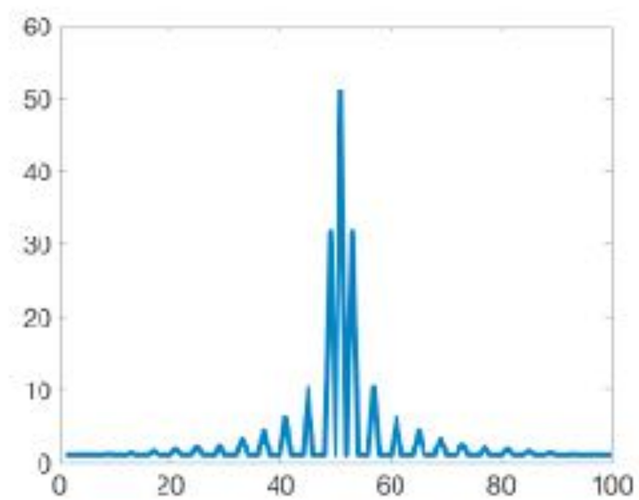
$$f(t) = A\sin(2\pi ft + \phi) = A\sin(\omega x + \phi)$$

Source: Deva Ramanan

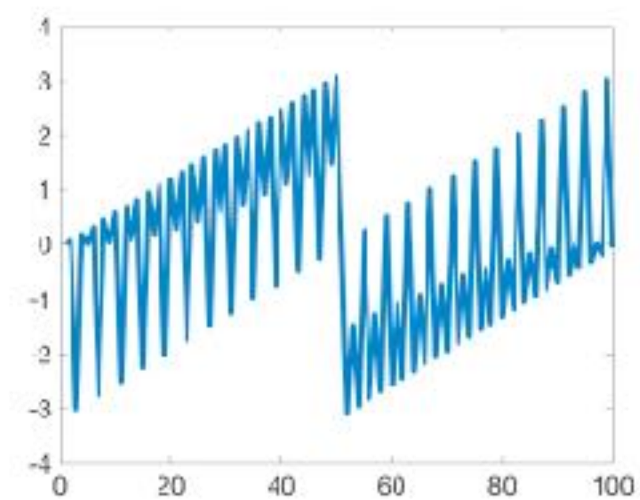
Fourier Transform



Signal



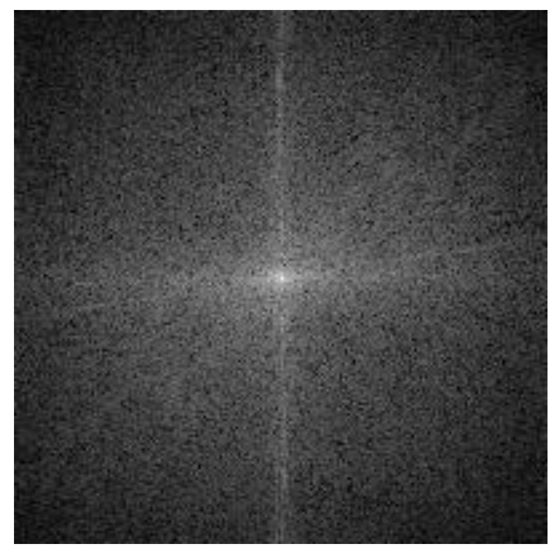
Amplitude



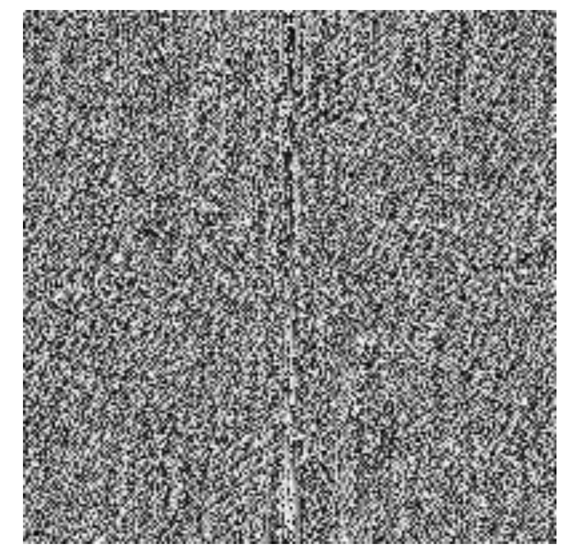
Phase

2D Fourier Transform

Signal

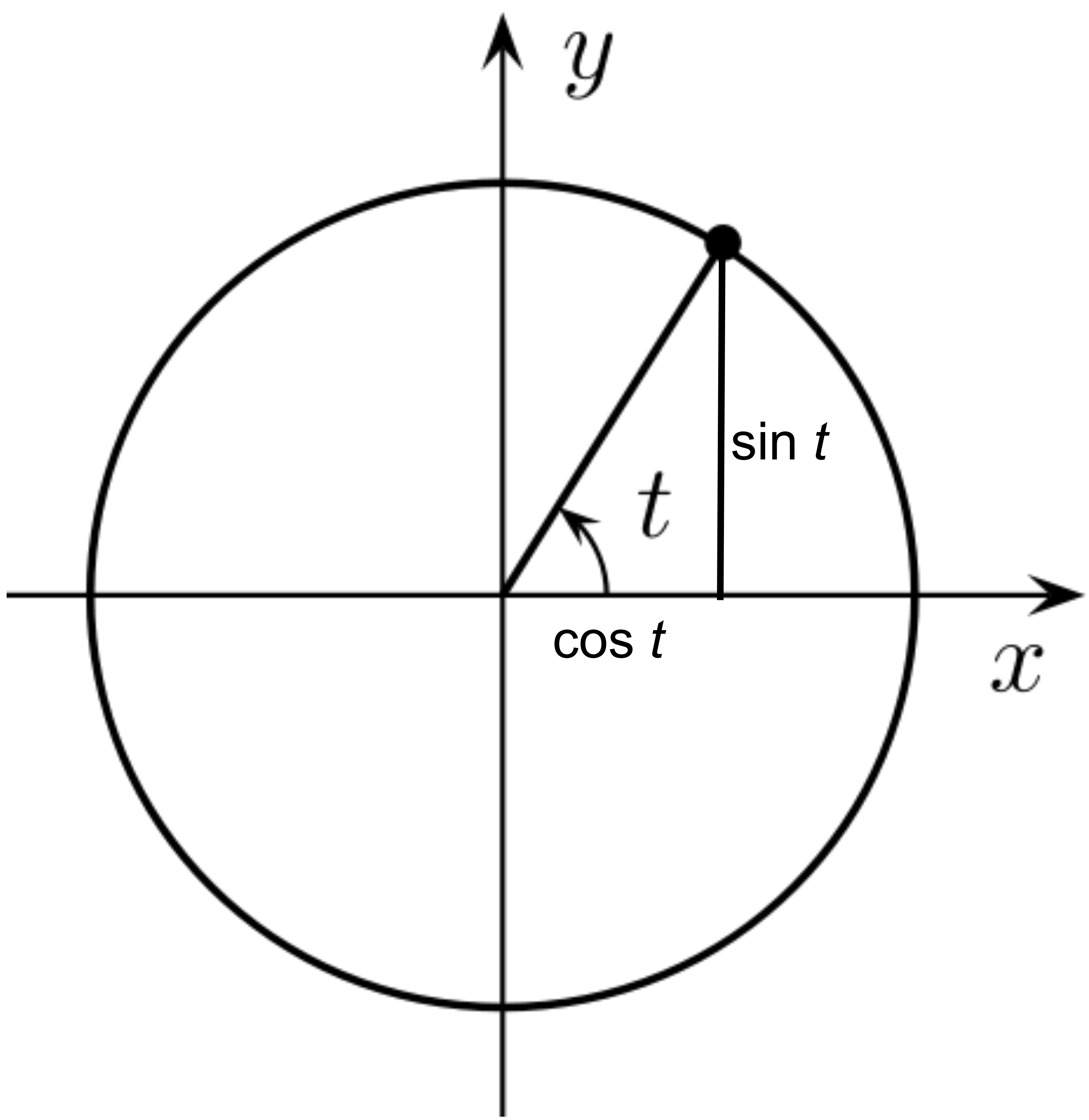


Amplitude



Phase

Sine/cosine and circle



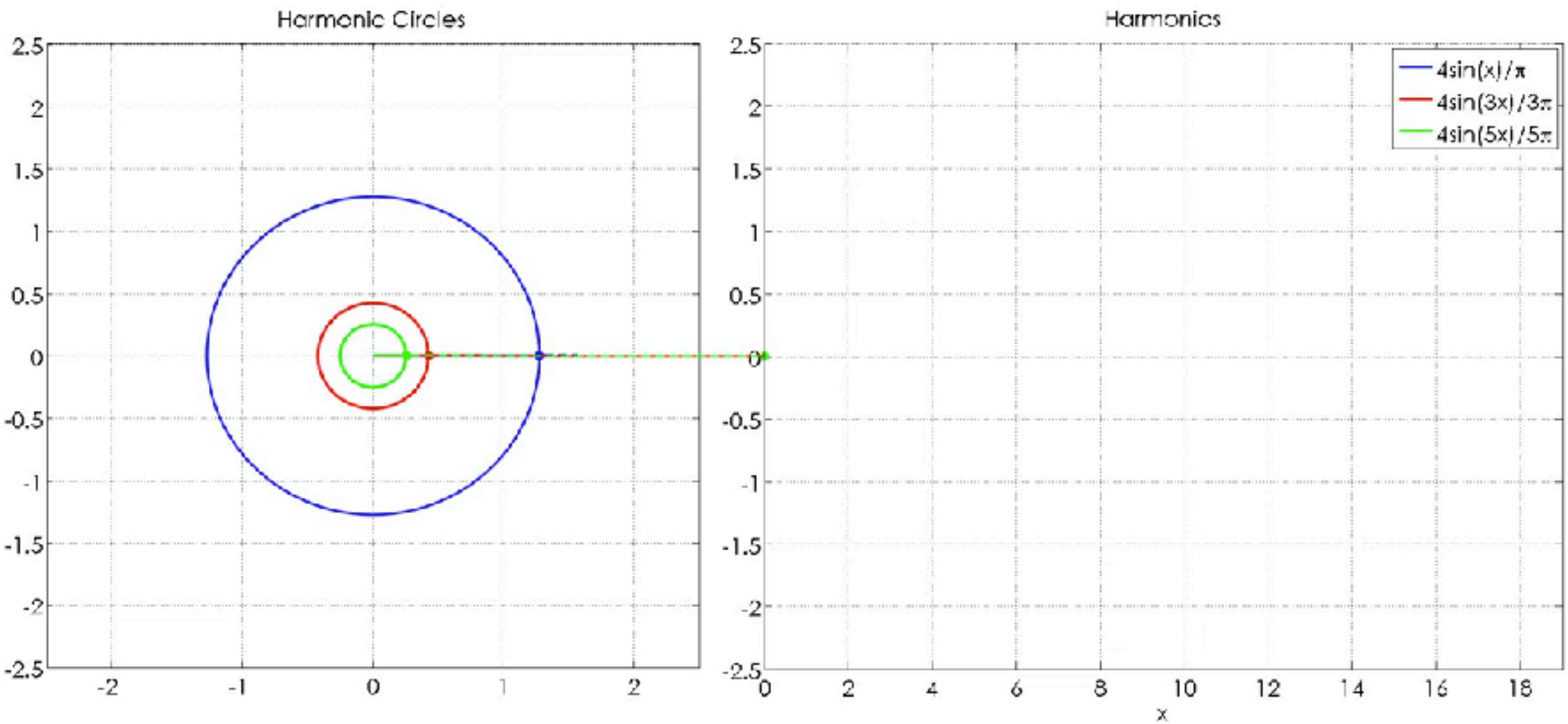
$$e^{ift} = \cos ft + i \sin ft$$

Amplitude: Radius of circle

Frequency: How fast you change t

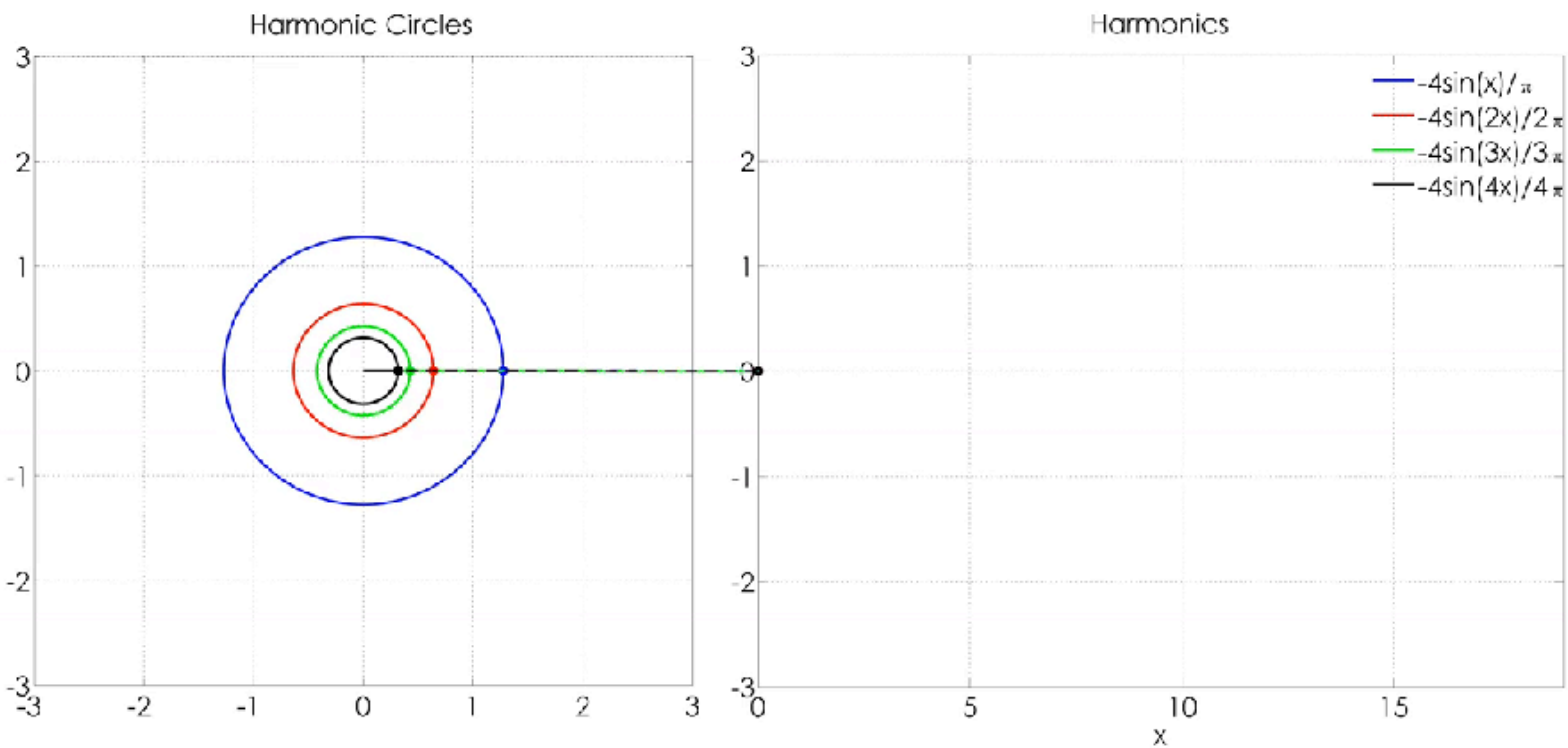
Wikipedia – Unit Circle

Square wave (approx.)

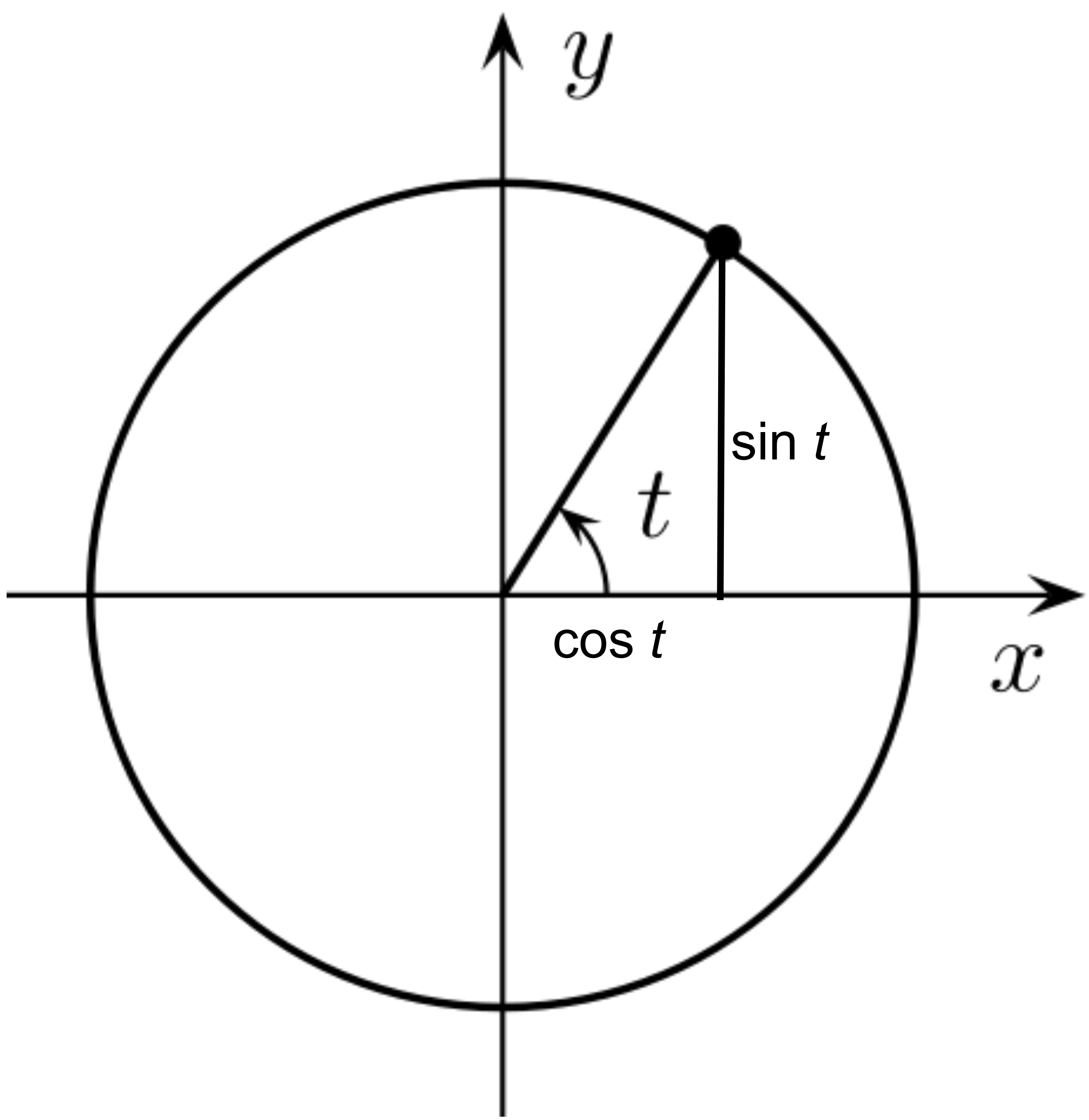


Mehmet E. Yavuz

Sawtooth wave (approx.)



Sine/cosine and circle



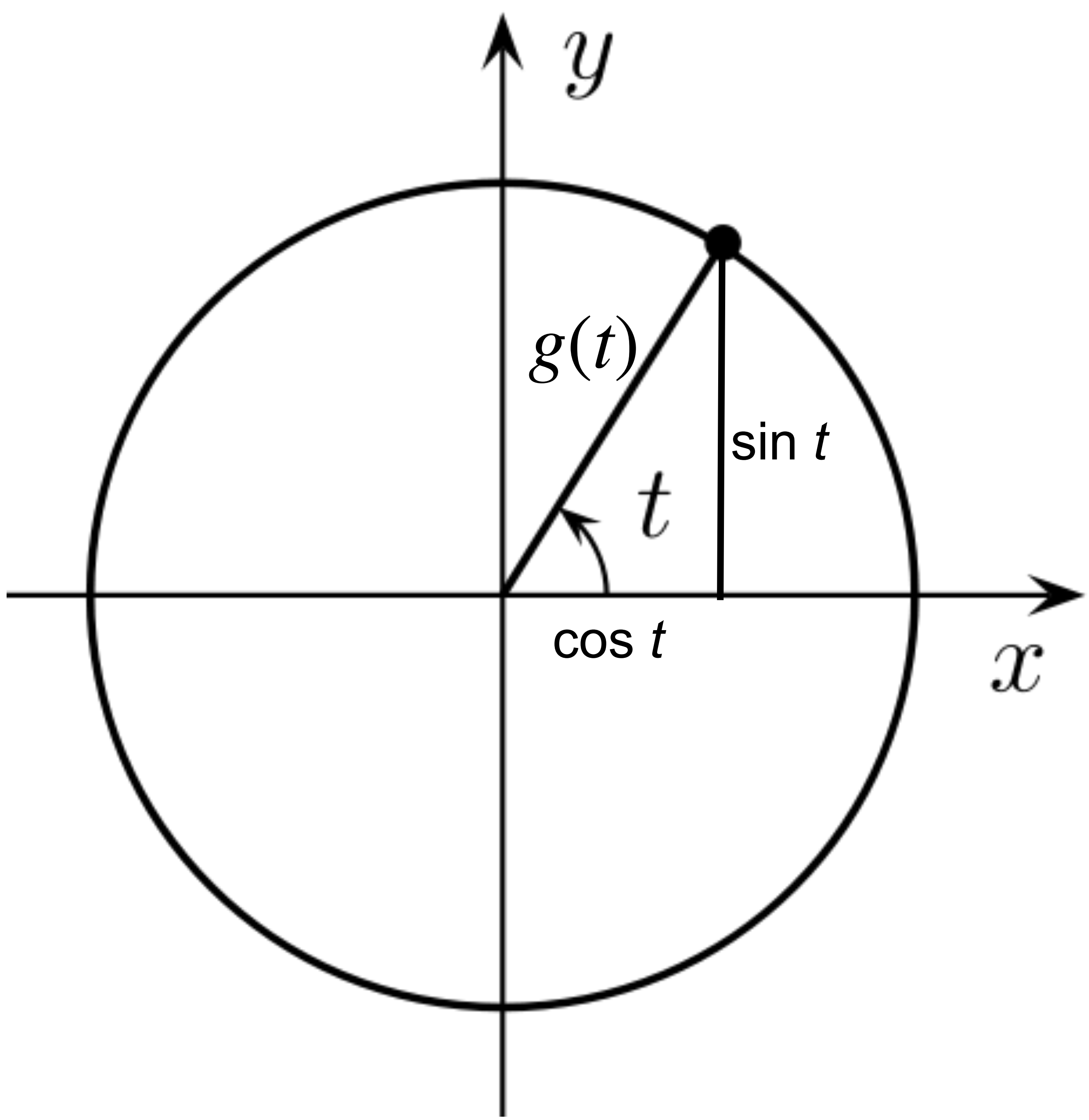
$$e^{ift} = \cos ft + i \sin ft$$

Amplitude: Radius of circle

Frequency: How fast you change t

Wikipedia – Unit Circle

Towards Fourier Transform

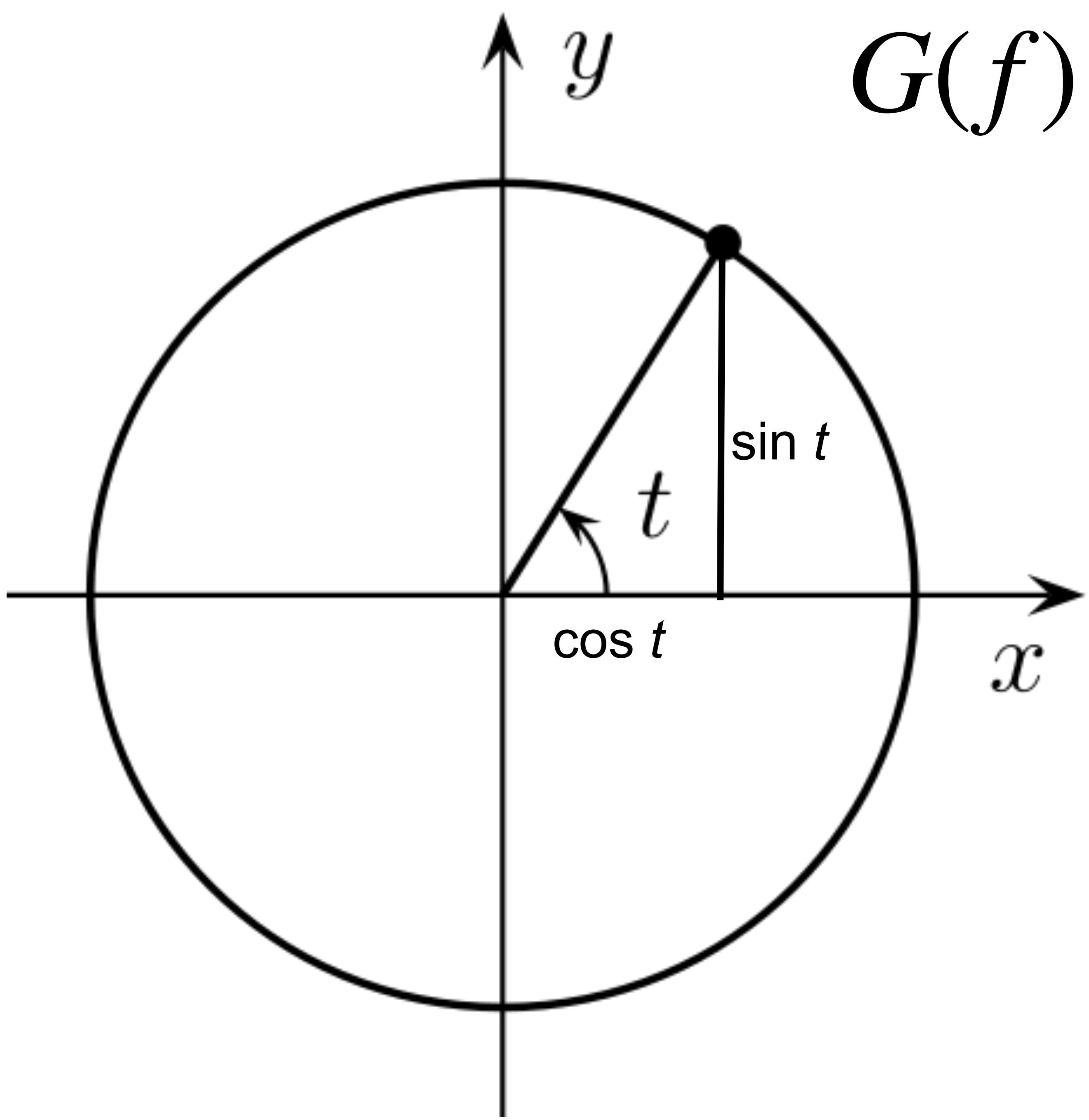


$$g(t)e^{-2\pi i f t}$$

Maps g(t) on to the unit circle with frequency f

Wikipedia – Unit Circle

The Fourier Transform

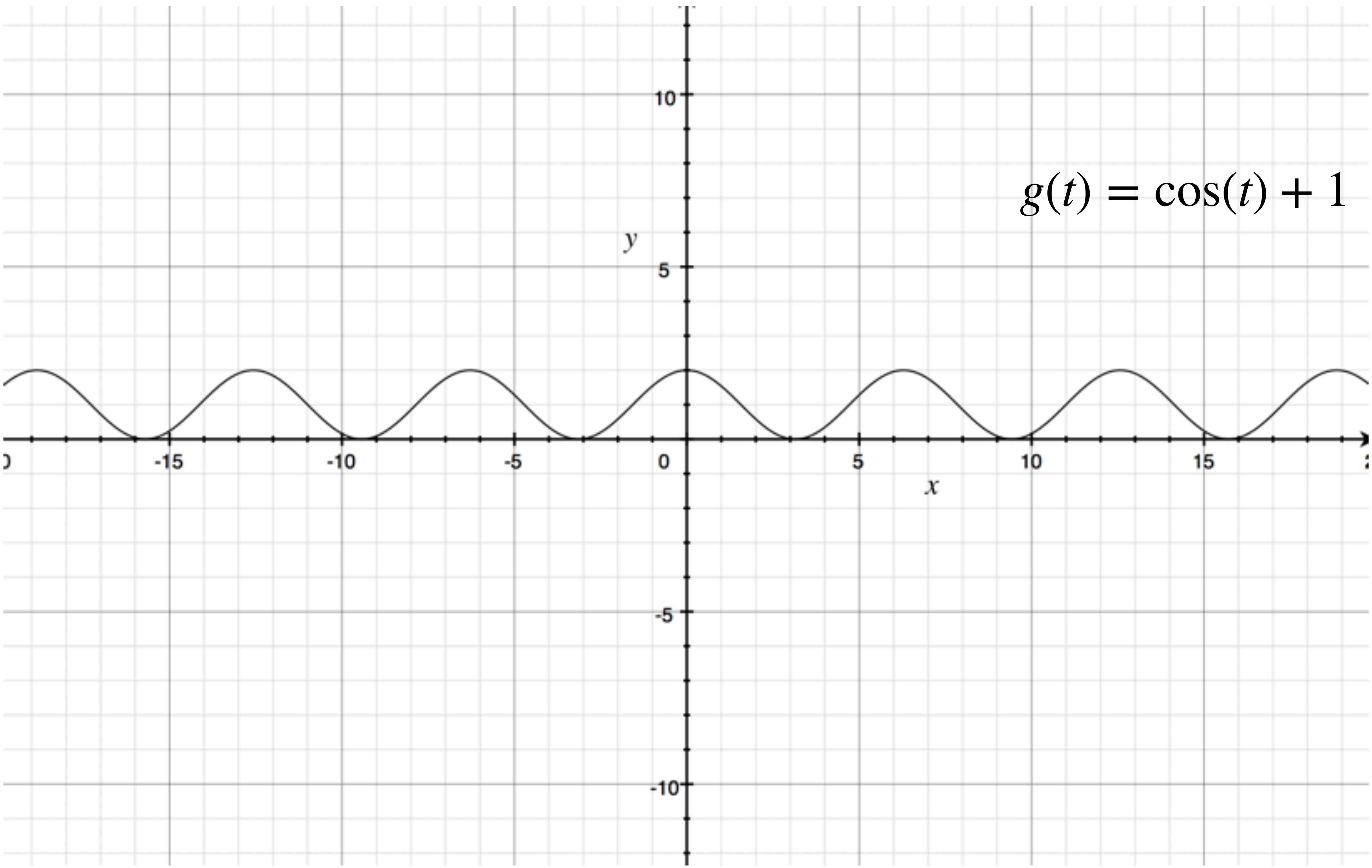


$$G(f) = \int_{-\infty}^{\infty} g(t)e^{-2\pi i f t}dt$$

How I think of it:
You wrap g(t) around the circle with frequency f, then calculate average position of g(t)

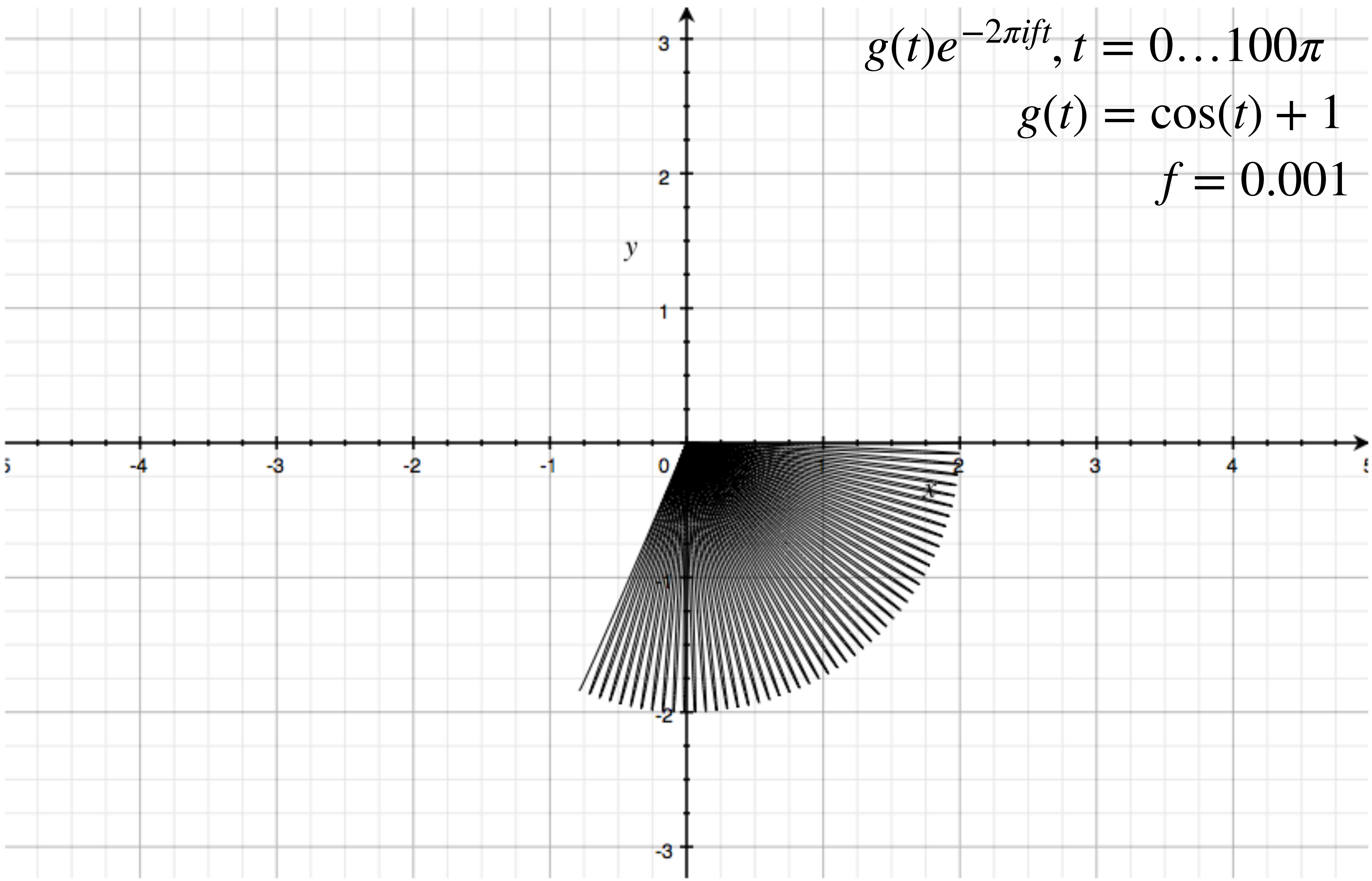
Wikipedia – Unit Circle

Signal that we want to compute FT on



The Fourier Transform

$$G(f) = \int_{-\infty}^{\infty} g(t)e^{-2\pi ift}dt$$



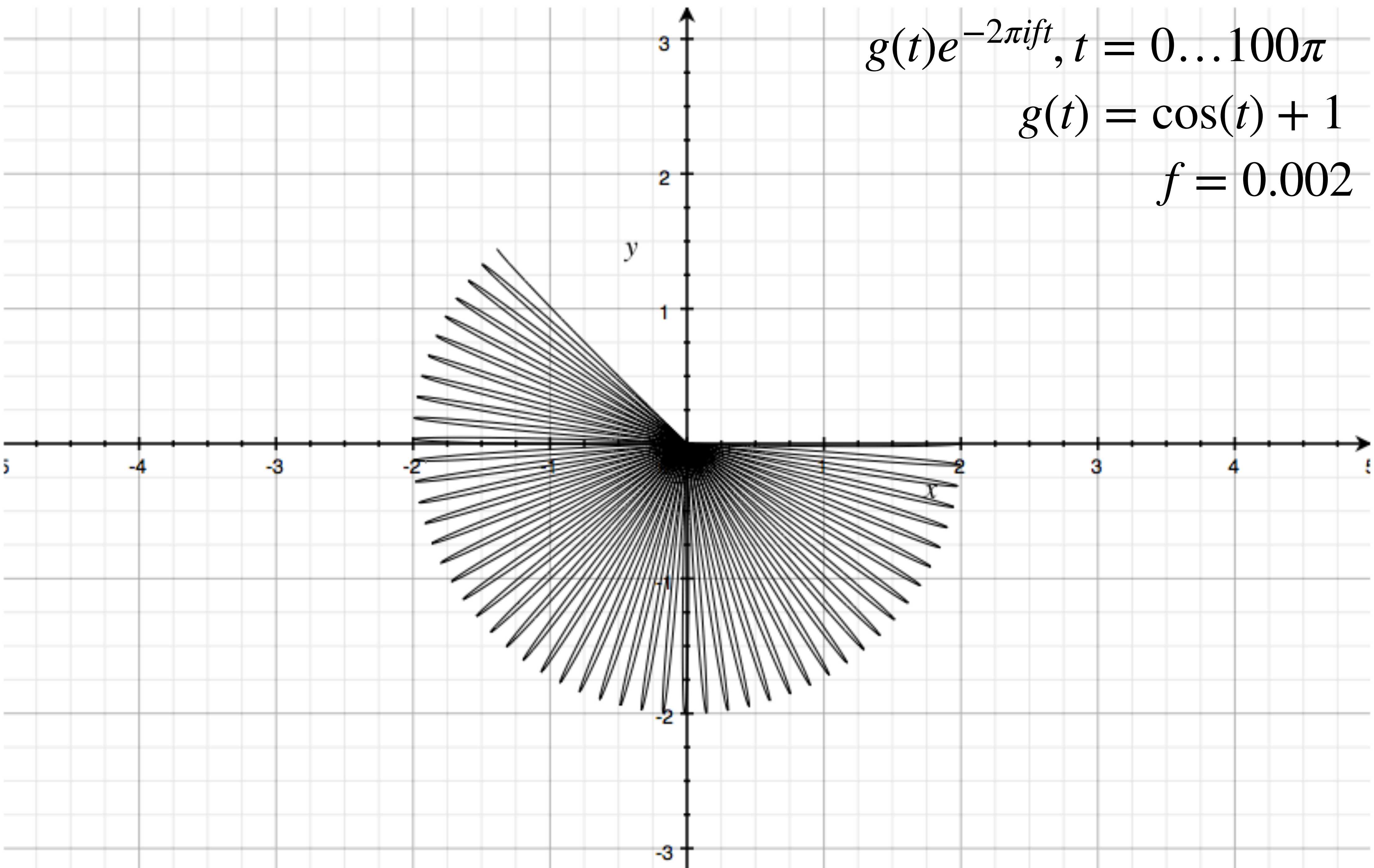
The Fourier Transform

$$G(f) = \int_{-\infty}^{\infty} g(t)e^{-2\pi ift}dt$$

$$g(t)e^{-2\pi ift}, t = 0\ldots100\pi$$

$$g(t) = \cos(t) + 1$$

$$f = 0.002$$



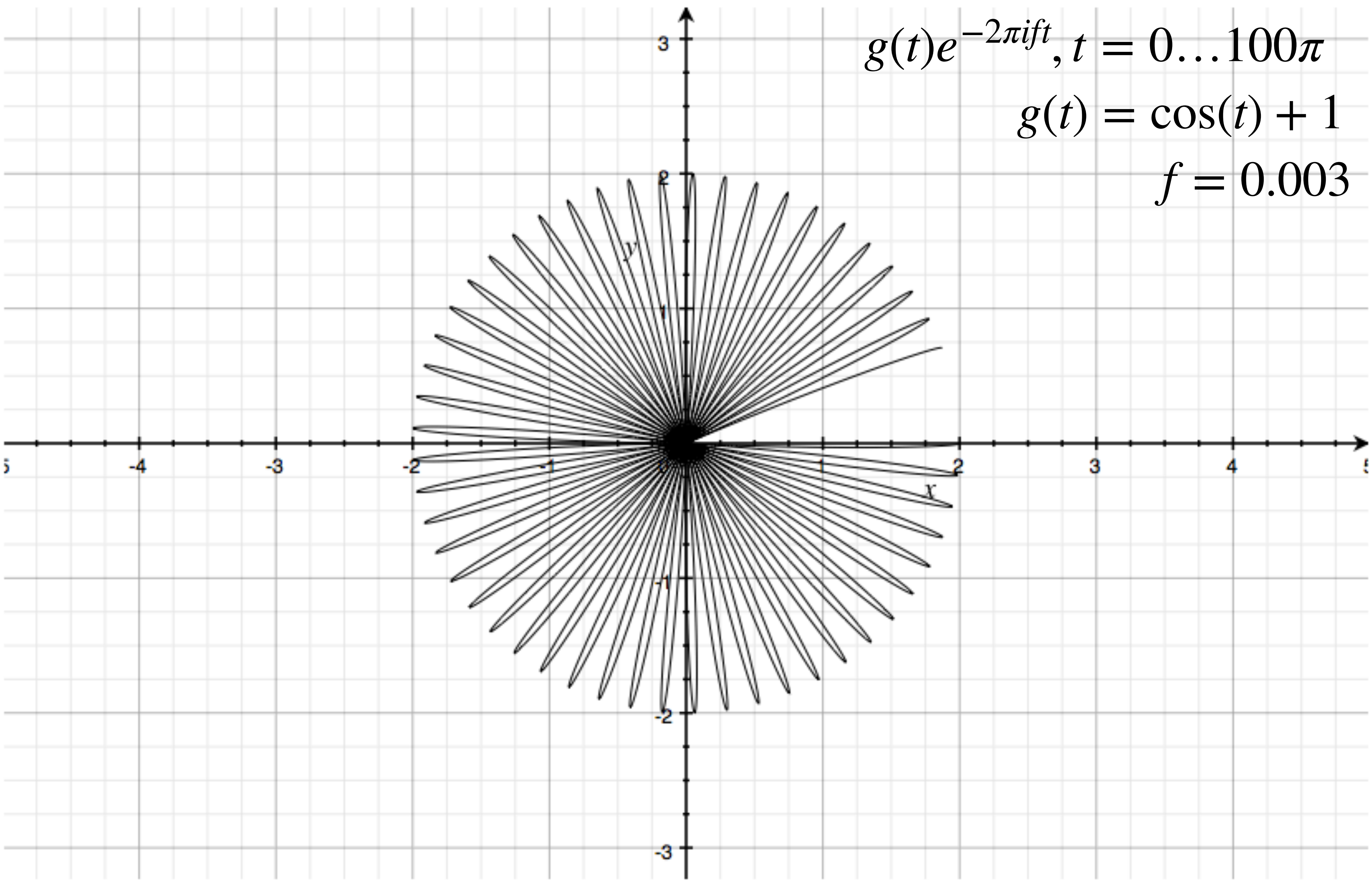
The Fourier Transform

$$G(f) = \int_{-\infty}^{\infty} g(t)e^{-2\pi i f t}dt$$

$$g(t)e^{-2\pi i f t}, t = 0\ldots100\pi$$

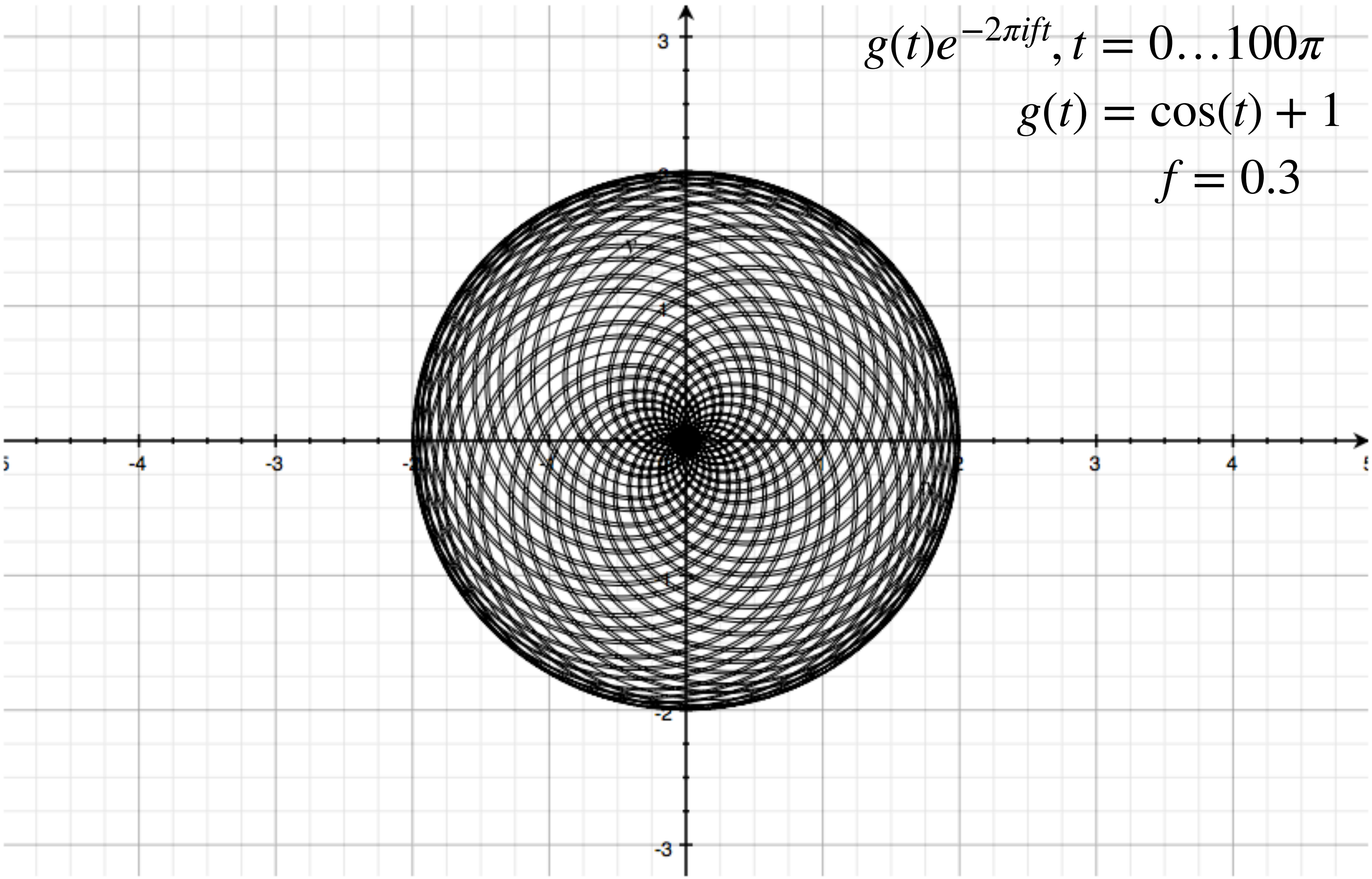
$$g(t) = \cos(t) + 1$$

$$f = 0.003$$



The Fourier Transform

$$G(f) = \int_{-\infty}^{\infty} g(t)e^{-2\pi ift}dt$$



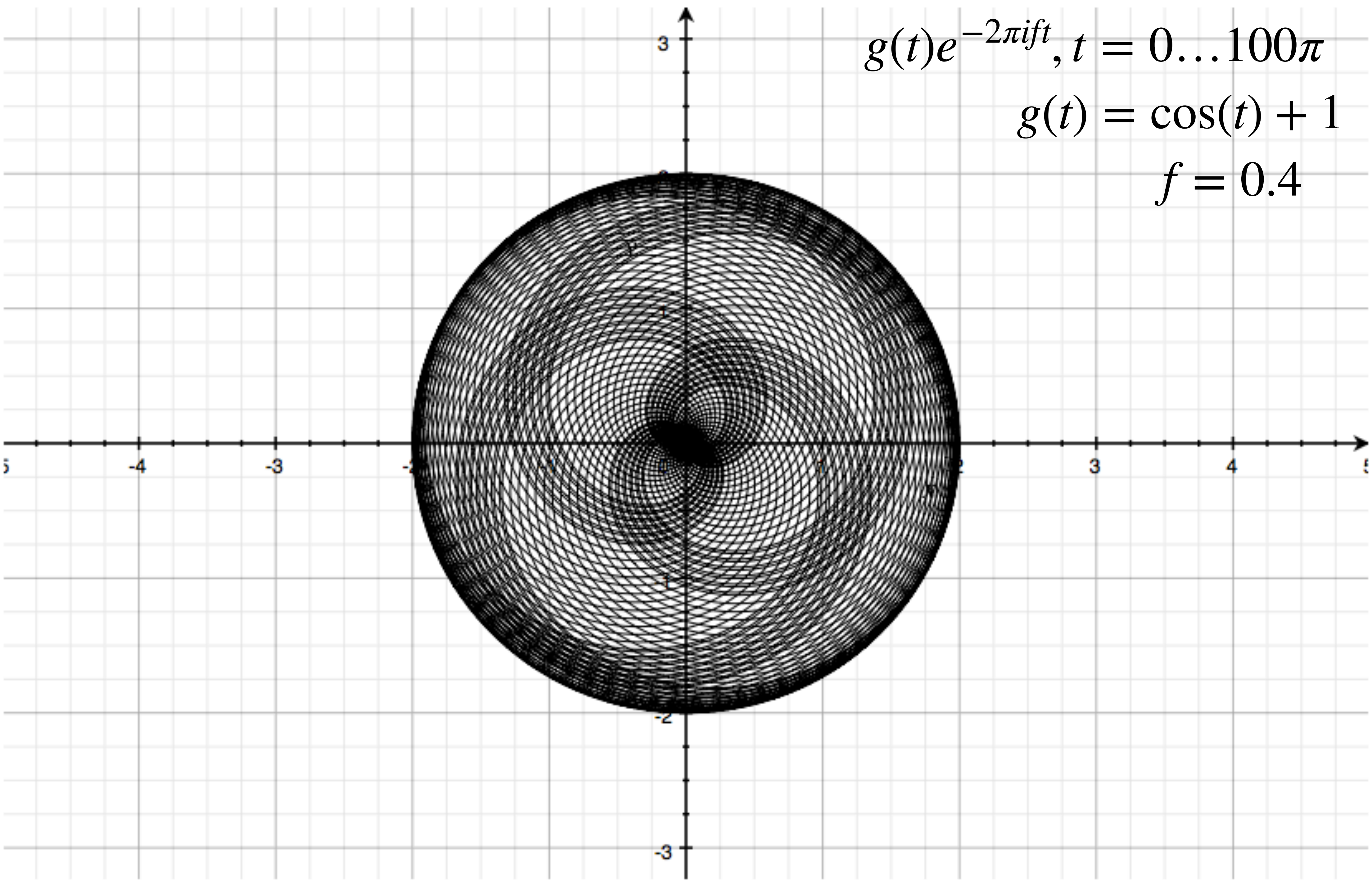
$$g(t)e^{-2\pi ift}, t = 0\ldots100\pi$$

$$g(t) = \cos(t) + 1$$

$$f = 0.3$$

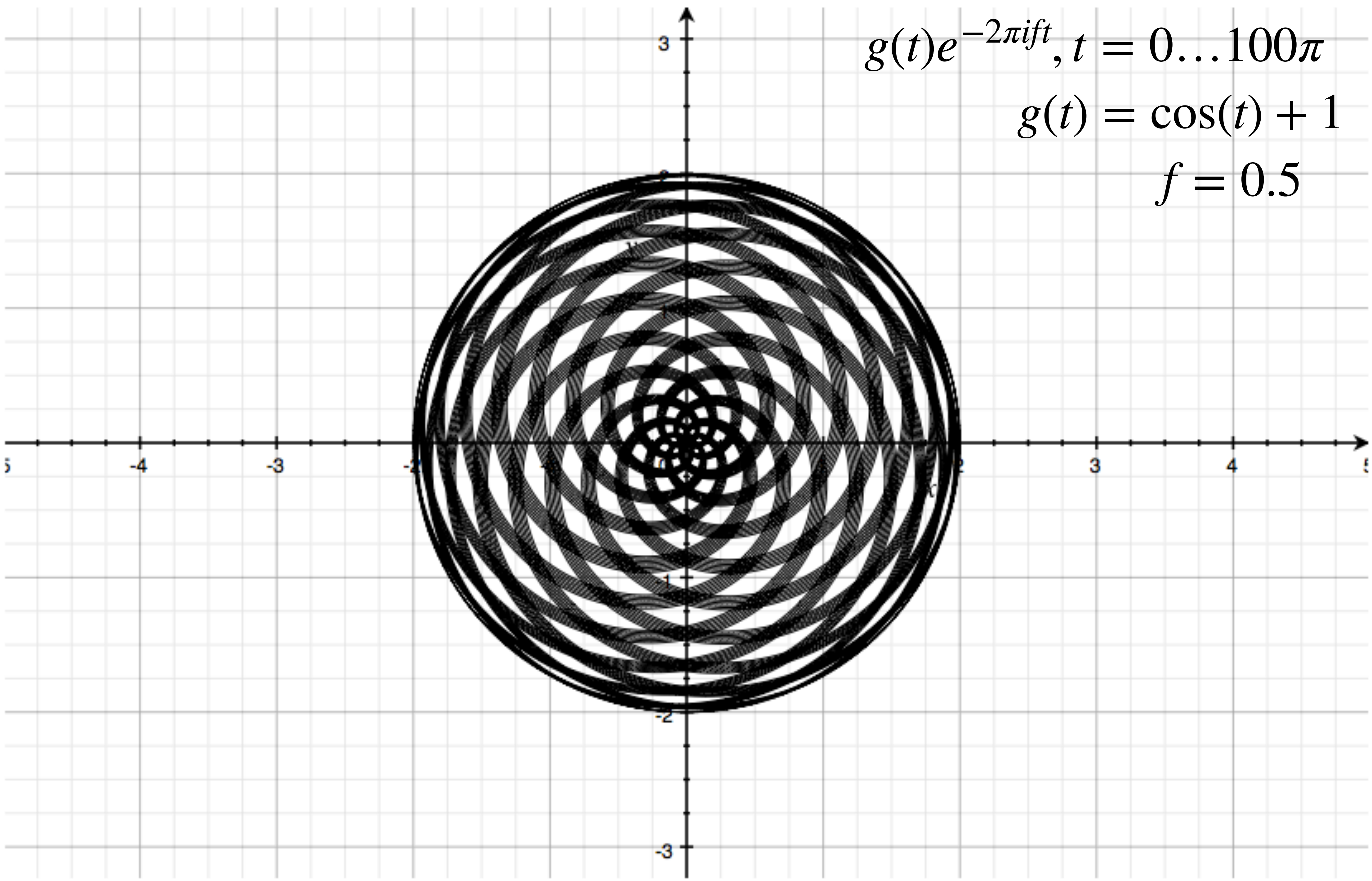
The Fourier Transform

$$G(f) = \int_{-\infty}^{\infty} g(t)e^{-2\pi ift}dt$$



The Fourier Transform

$$G(f) = \int_{-\infty}^{\infty} g(t)e^{-2\pi ift}dt$$



$$g(t)e^{-2\pi ift}, t = 0\ldots100\pi$$

$$g(t) = \cos(t) + 1$$

$$f = 0.5$$

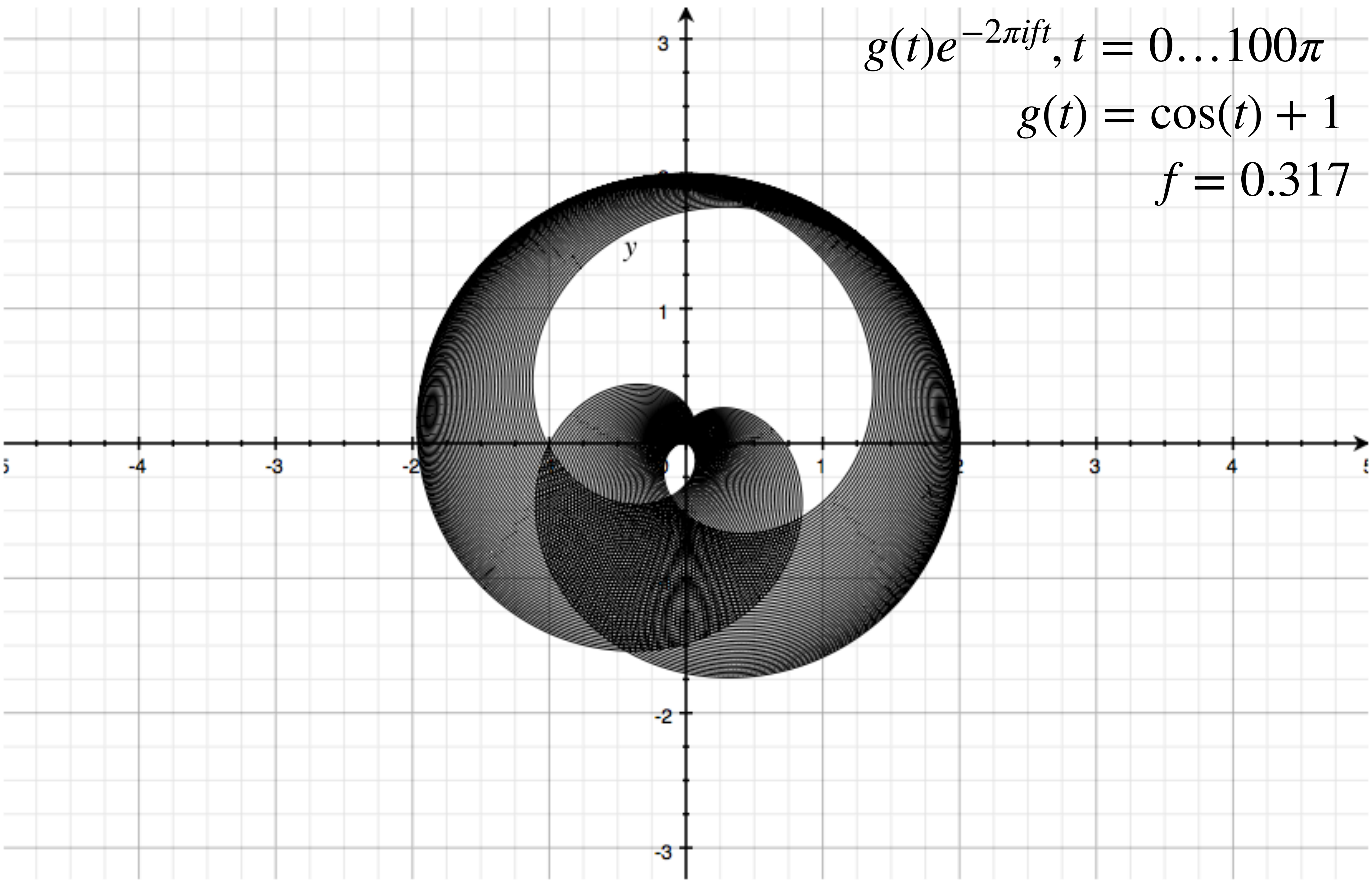
The Fourier Transform

$$G(f) = \int_{-\infty}^{\infty} g(t)e^{-2\pi i f t}dt$$

$$g(t)e^{-2\pi i f t}, t = 0\ldots100\pi$$

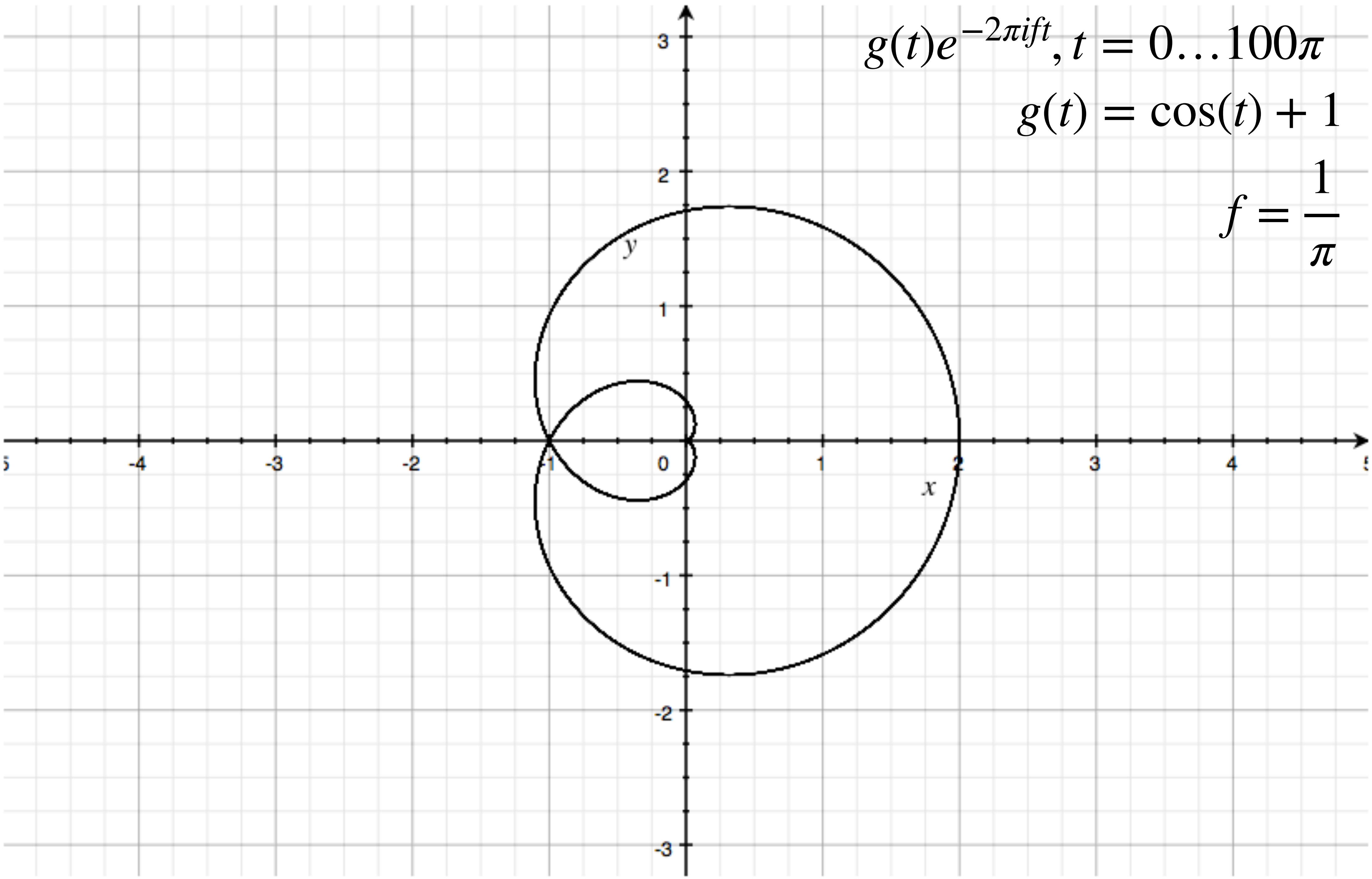
$$g(t) = \cos(t) + 1$$

$$f = 0.317$$



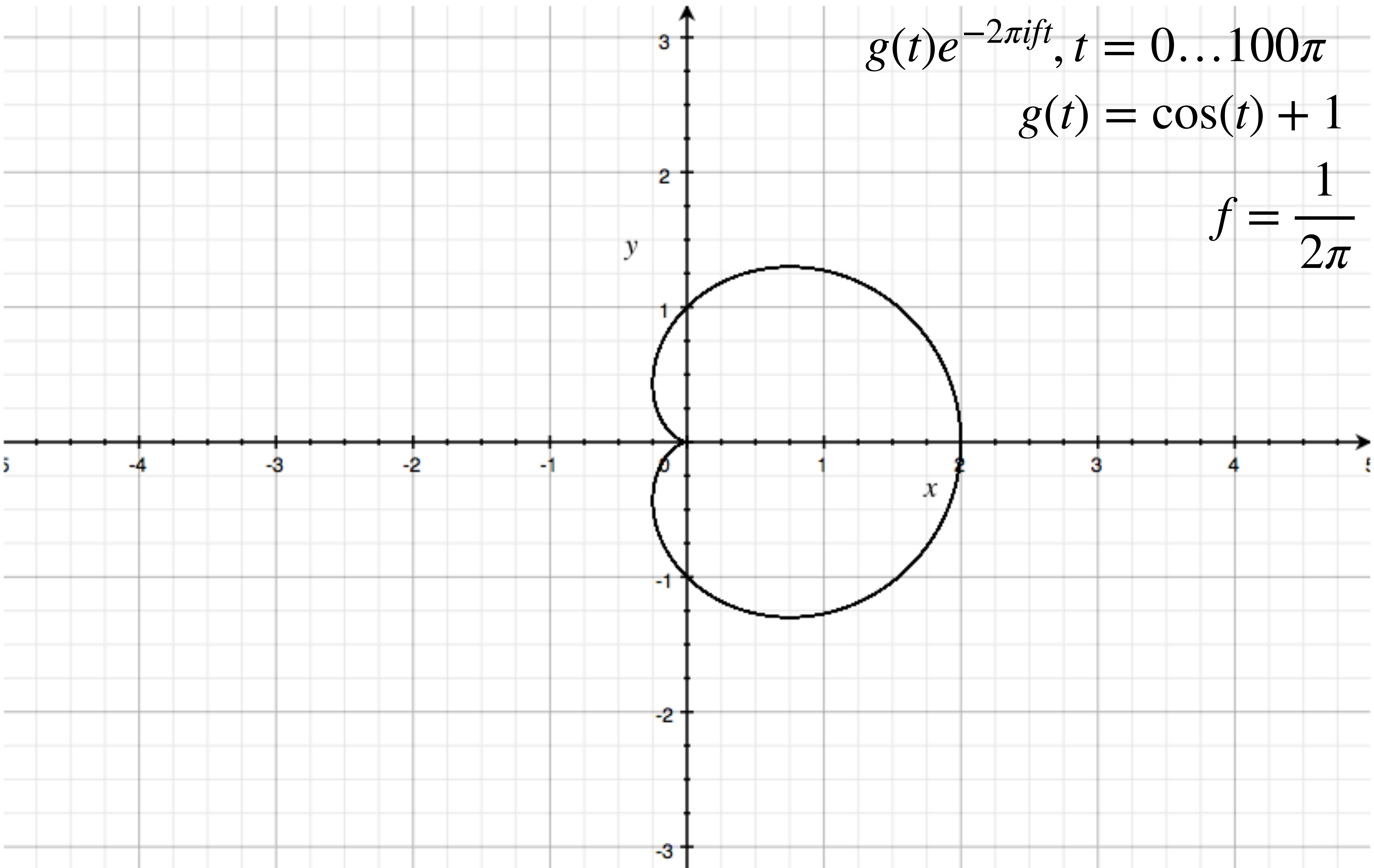
The Fourier Transform

$$G(f) = \int_{-\infty}^{\infty} g(t)e^{-2\pi i f t}dt$$



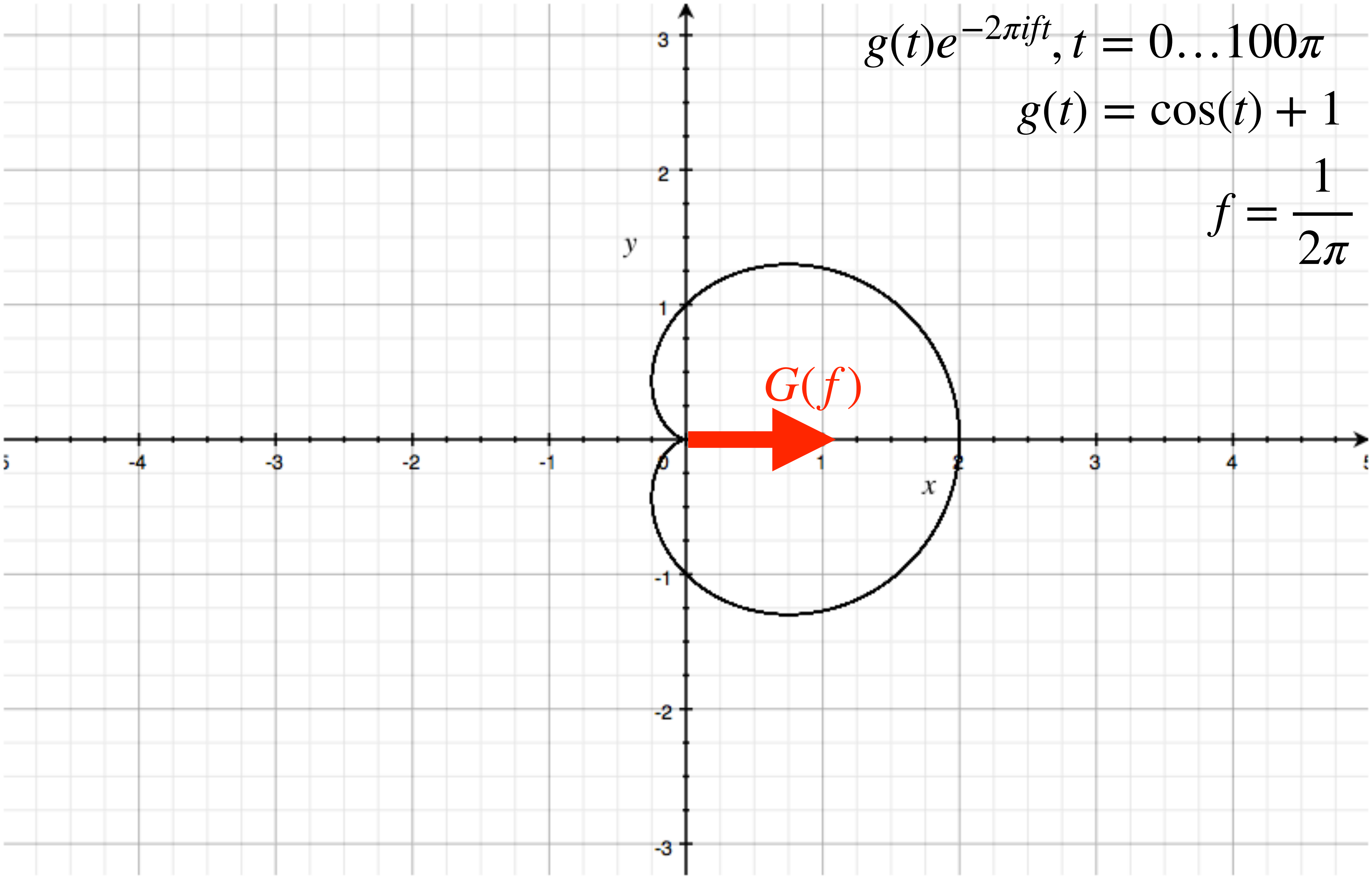
The Fourier Transform

$$G(f) = \int_{-\infty}^{\infty} g(t) e^{-2\pi i f t} dt$$



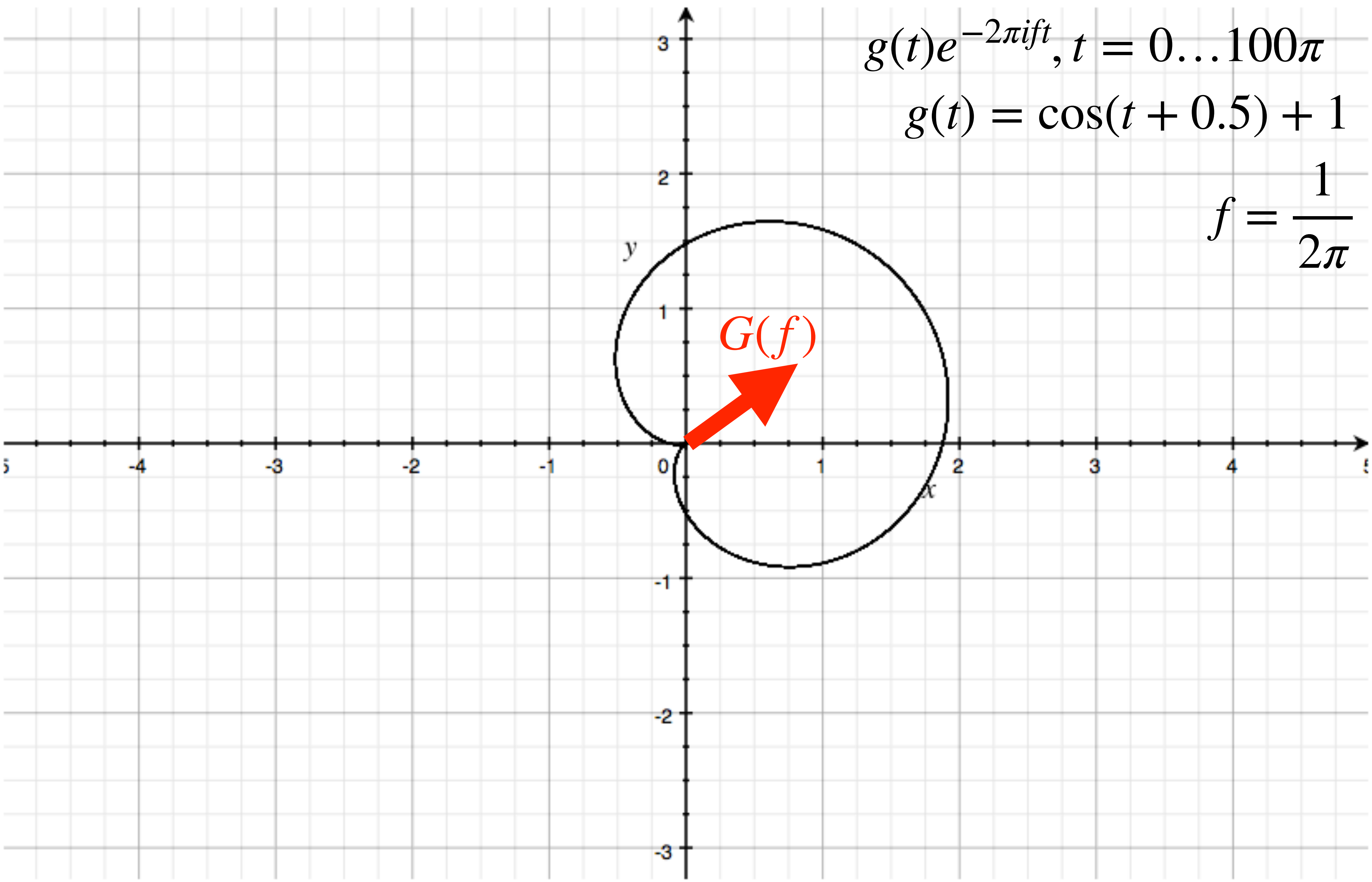
The Fourier Transform

$$G(f) = \int_{-\infty}^{\infty} g(t)e^{-2\pi i f t}dt$$



The Fourier Transform

$$G(f) = \int_{-\infty}^{\infty} g(t)e^{-2\pi ift}dt$$



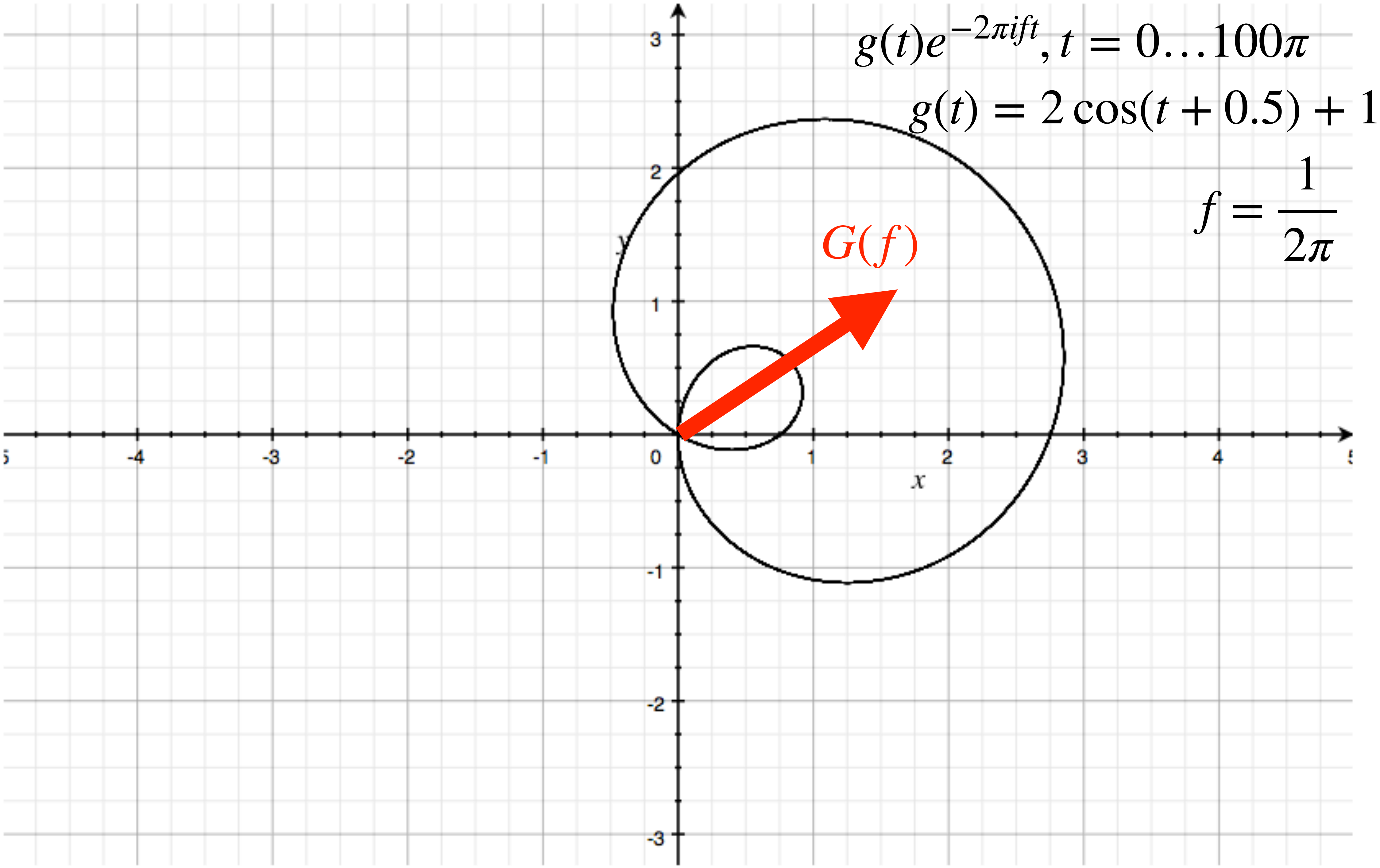
$$g(t)e^{-2\pi ift}, t = 0\ldots100\pi$$

$$g(t) = \cos(t + 0.5) + 1$$

$$f = \frac{1}{2\pi}$$

The Fourier Transform

$$G(f) = \int_{-\infty}^{\infty} g(t)e^{-2\pi ift}dt$$



The Fourier Transform

$$G(f) = \int_{-\infty}^{\infty} g(t) e^{-2\pi i f t} dt$$

Amplitude:

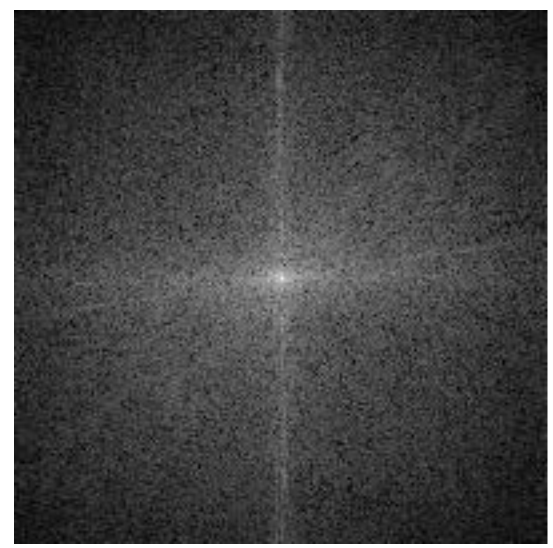
$$\sqrt{\Re[G(f)]^2 + \Im[G(f)]^2}$$

Phase:

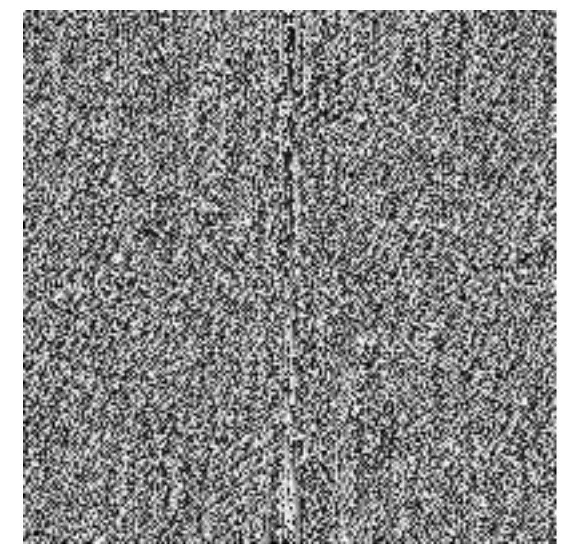
$$\tan^{-1} \frac{\Im[G(f)]}{\Re[G(f)]}$$

Inverse Fourier Transform

Signal



Amplitude



Phase

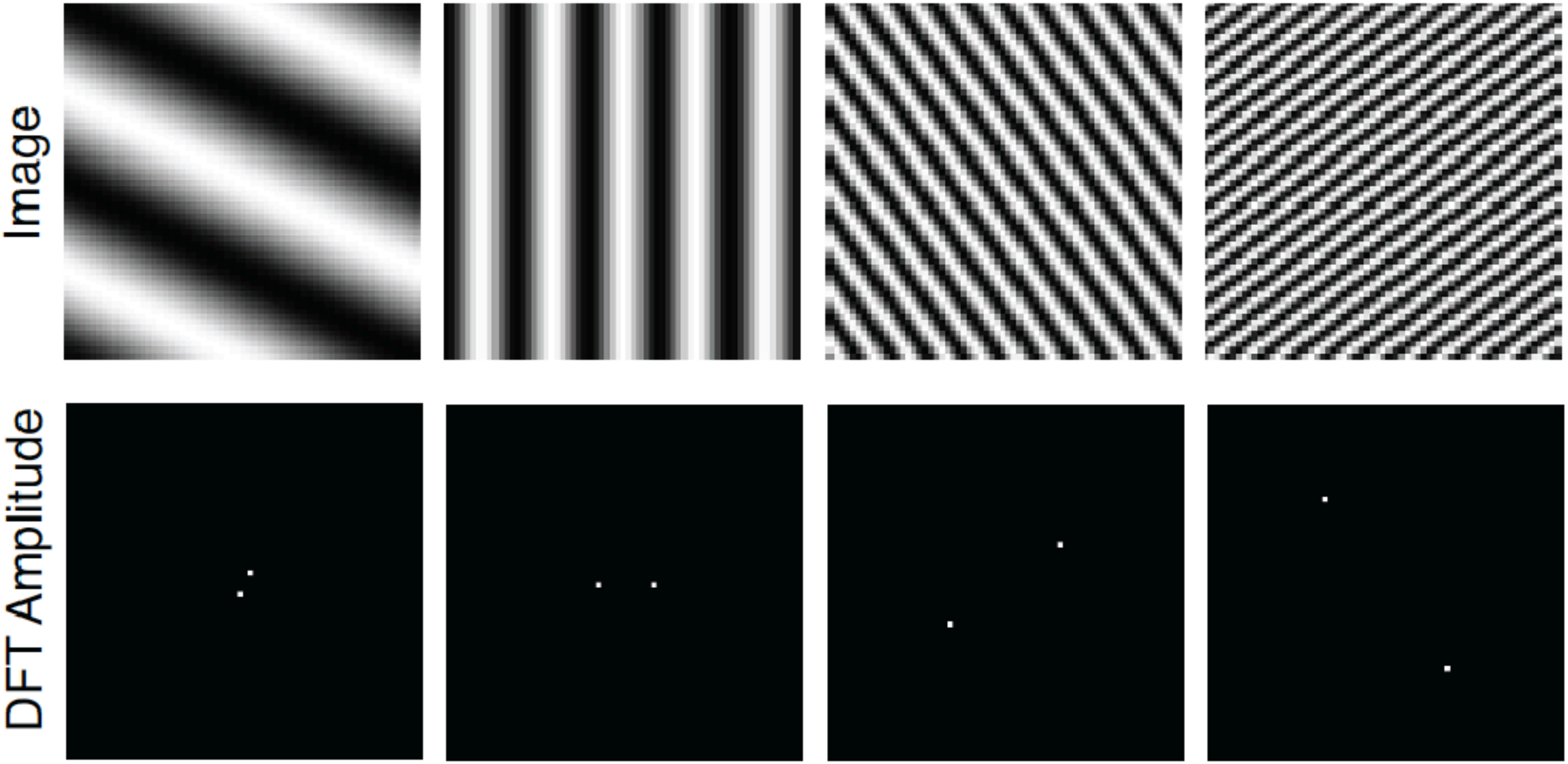
Fourier Transform:

$$G(f) = \int_{-\infty}^{\infty} g(t) e^{-2\pi i f t} dt$$

Inverse Fourier Transform:

$$g(t) = \int_{-\infty}^{\infty} G(f) e^{2\pi i f t} dt$$

Frequencies

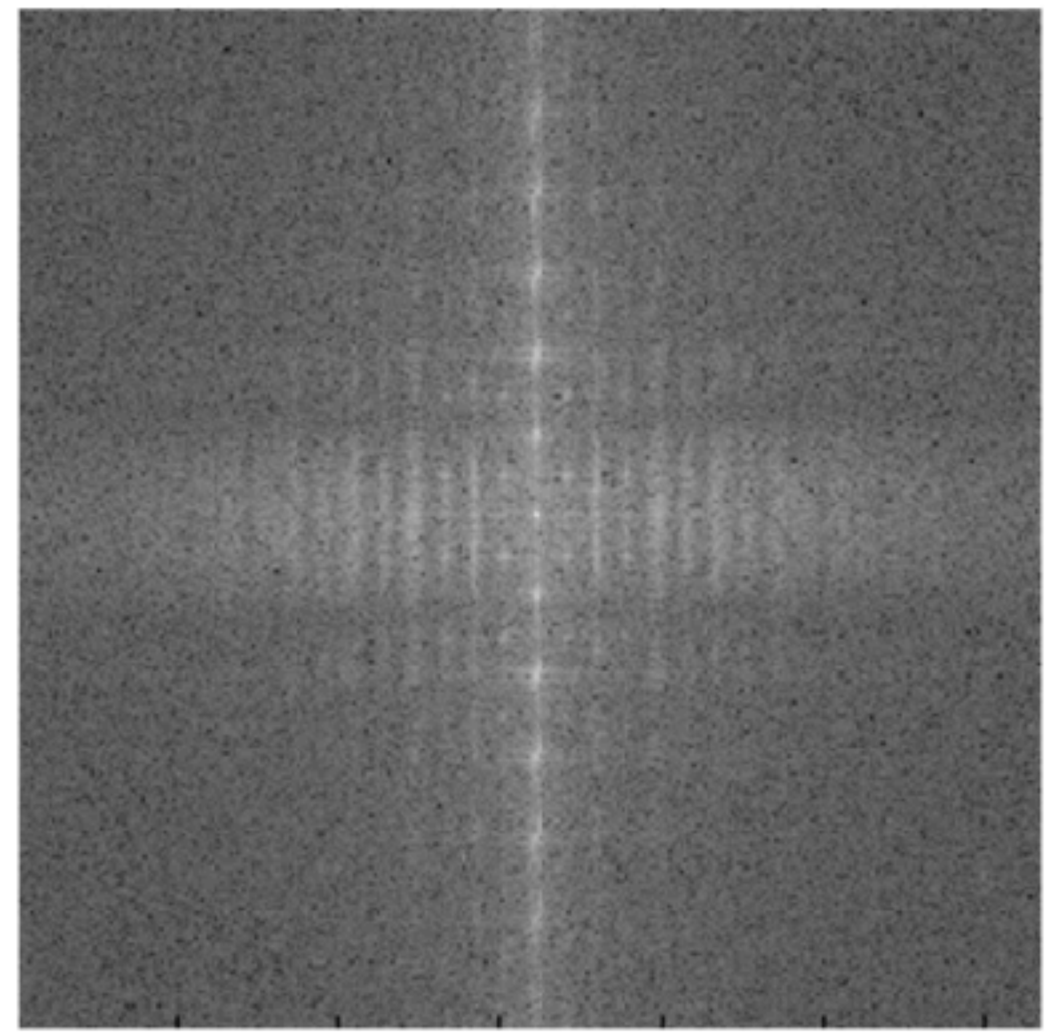


Images are 64x64 pixels. The wave is a cosine (if phase is zero).

Source: Bill Freeman

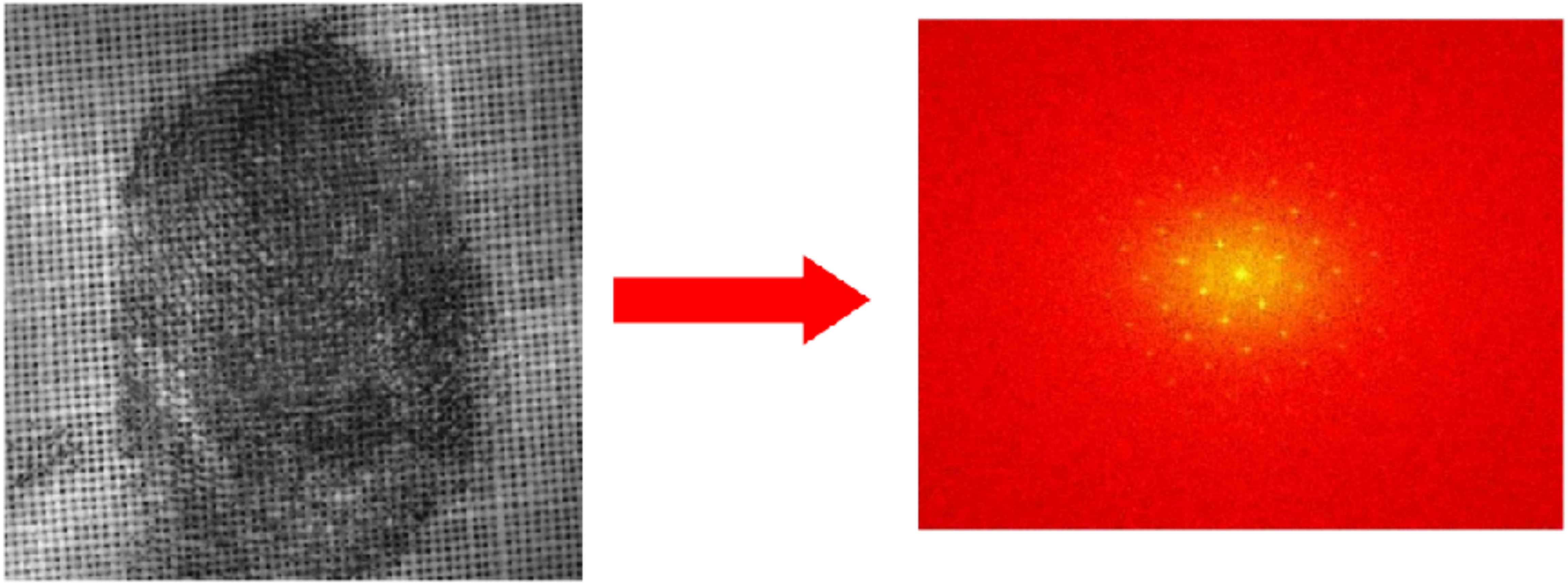
FT has peaks at spatial frequencies of repeated structure

Image

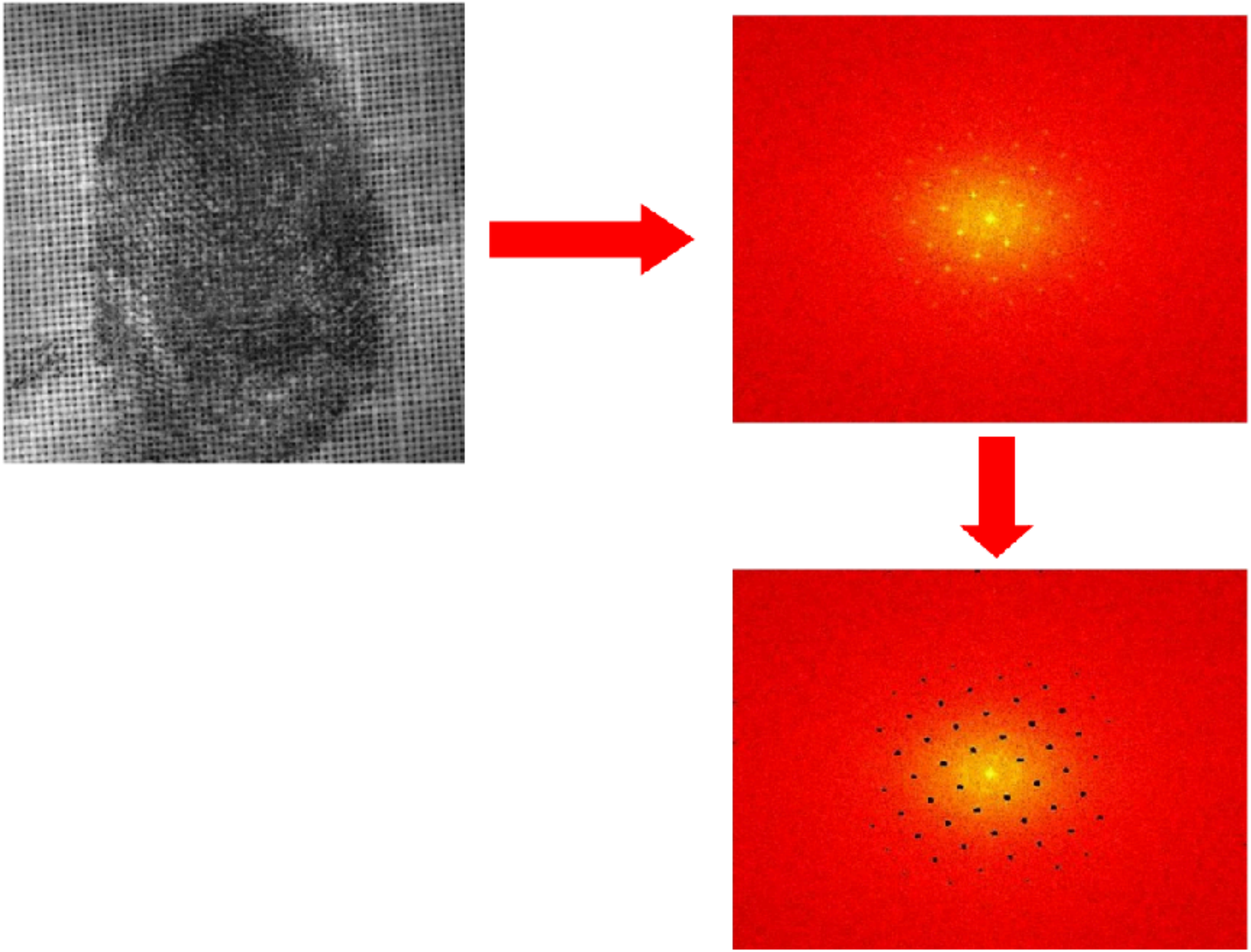


Amplitude

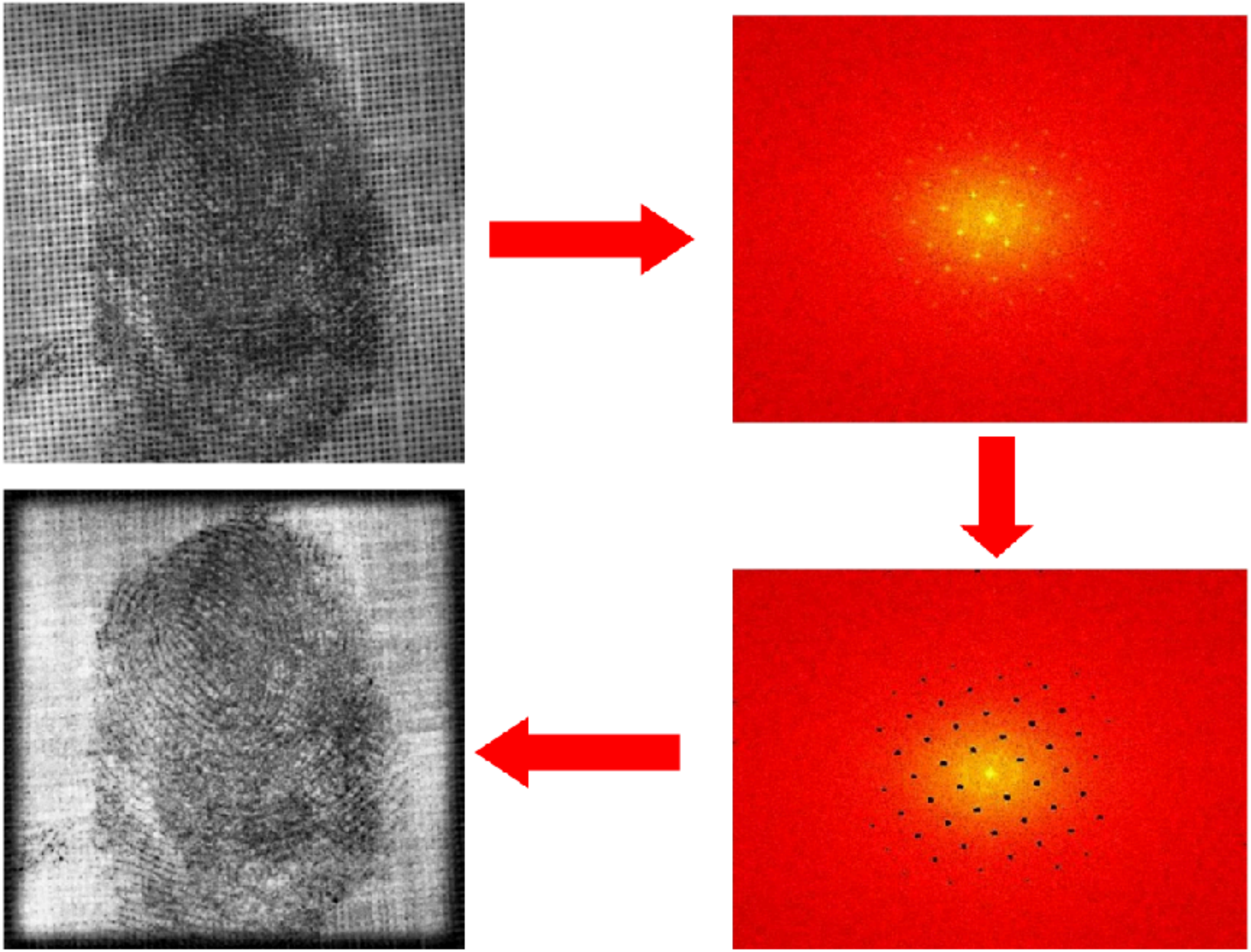
Source: Deva Ramanan



Source: Deva Ramanan

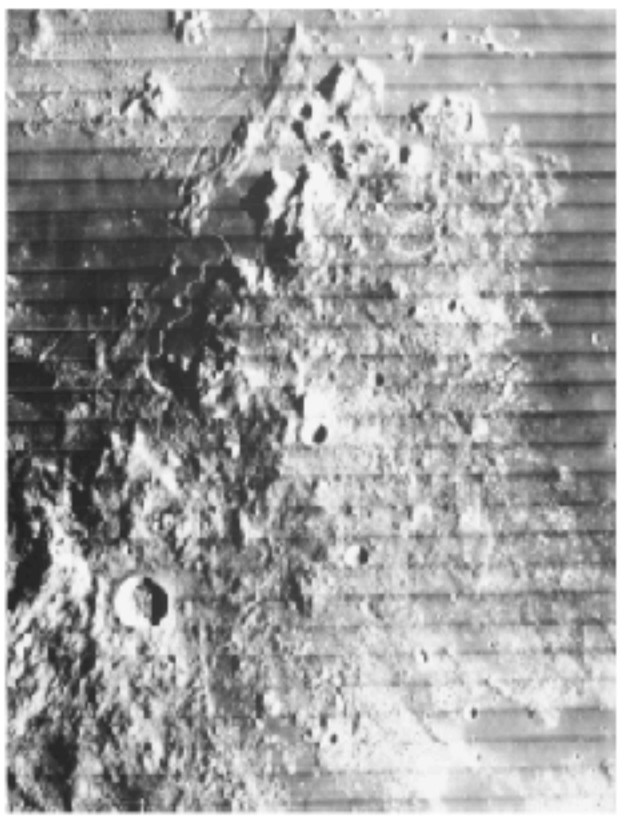


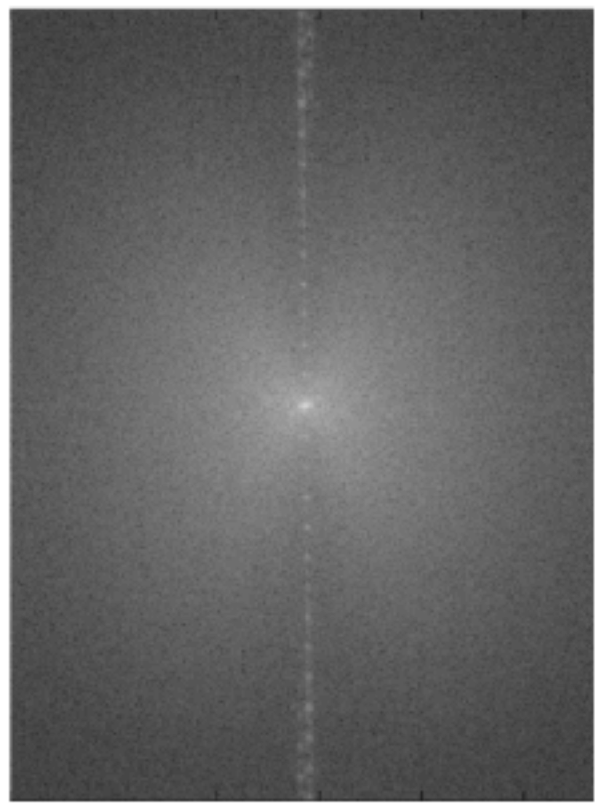
Source: Deva Ramanan



Source: Deva Ramanan

Lunar Orbital Image (1966)

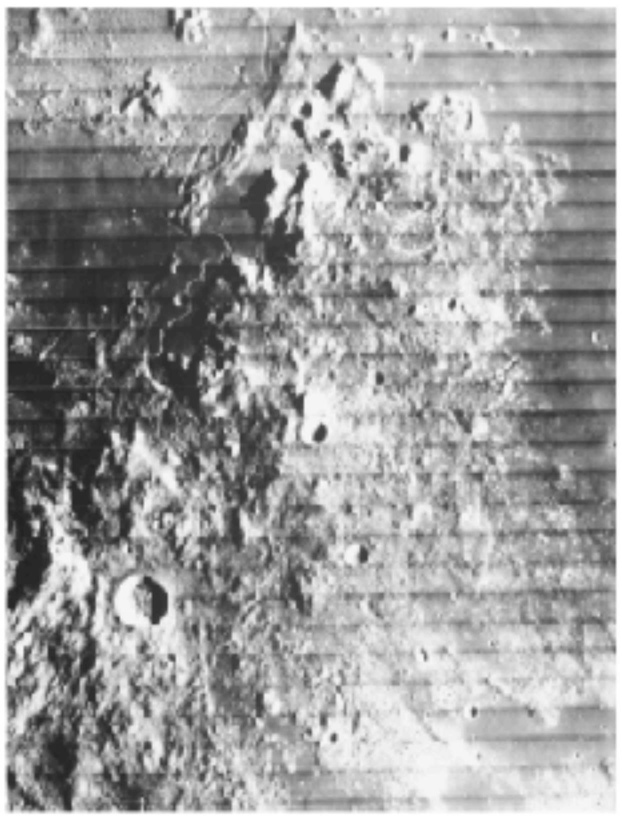


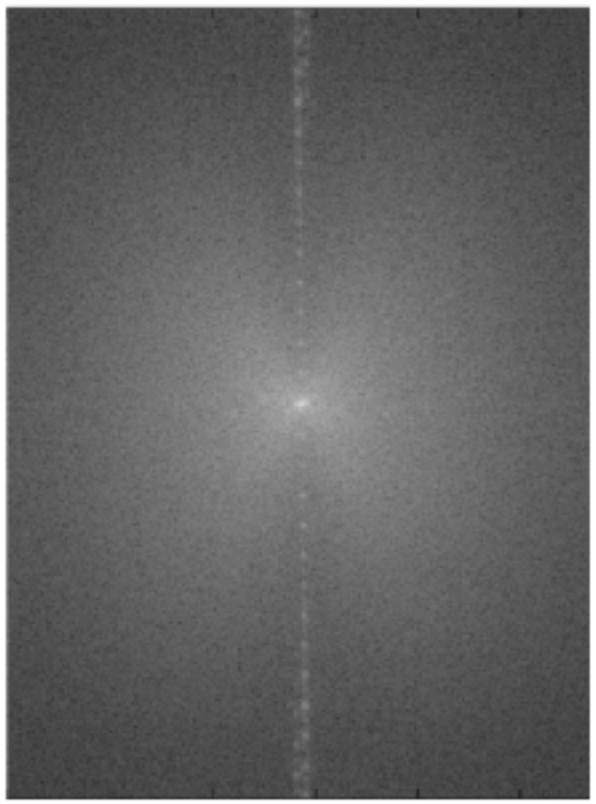


Amplitude

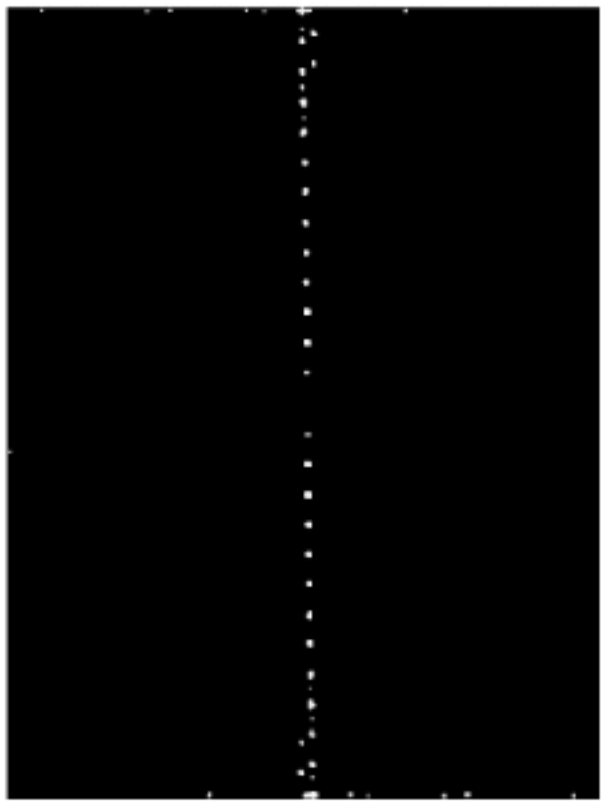
Source: Deva Ramanan

Lunar Orbital Image (1966)





Amplitude



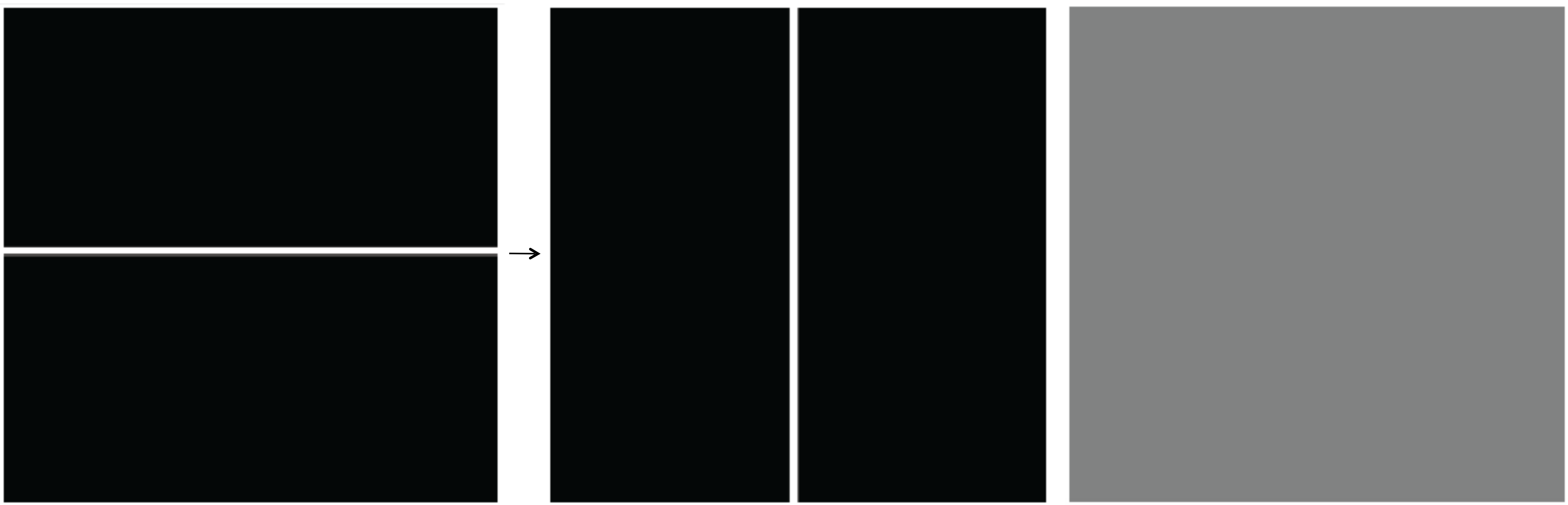
Remove Peaks

Source: Deva Ramanan

Let’s Practice

Some important Fourier transforms

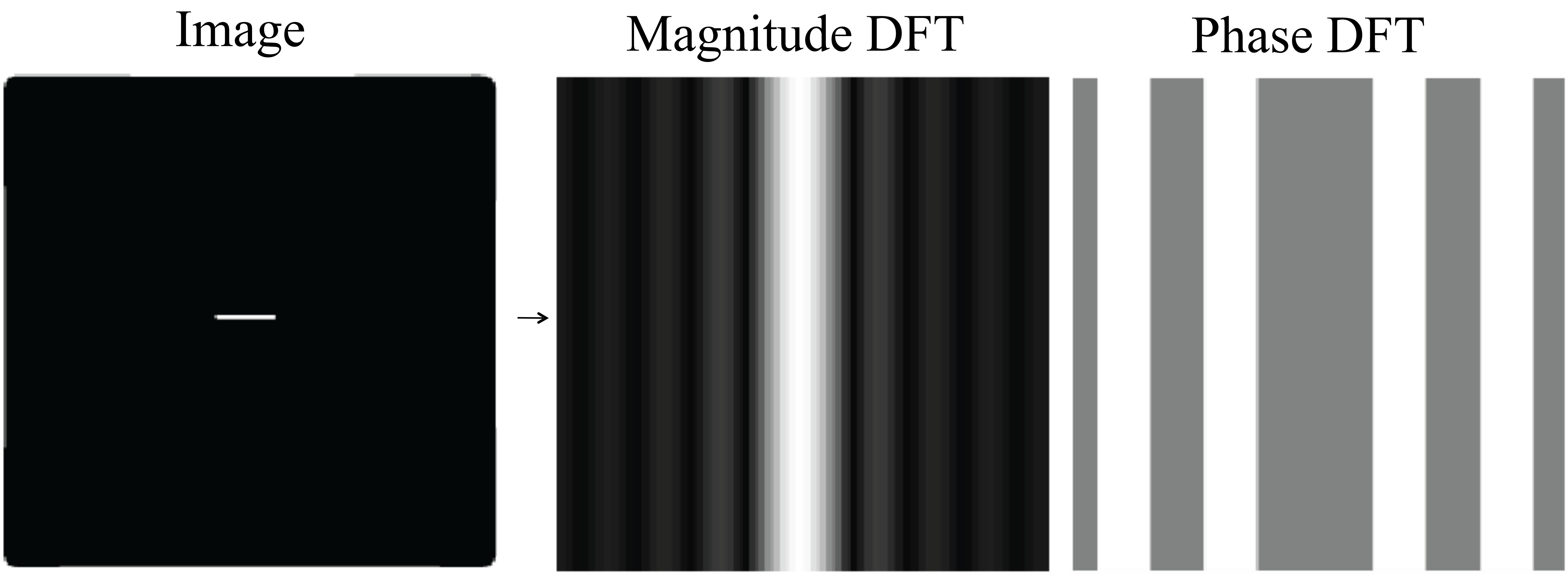
Image | Magnitude DFT | Phase DFT



Images are 64x64 pixels.

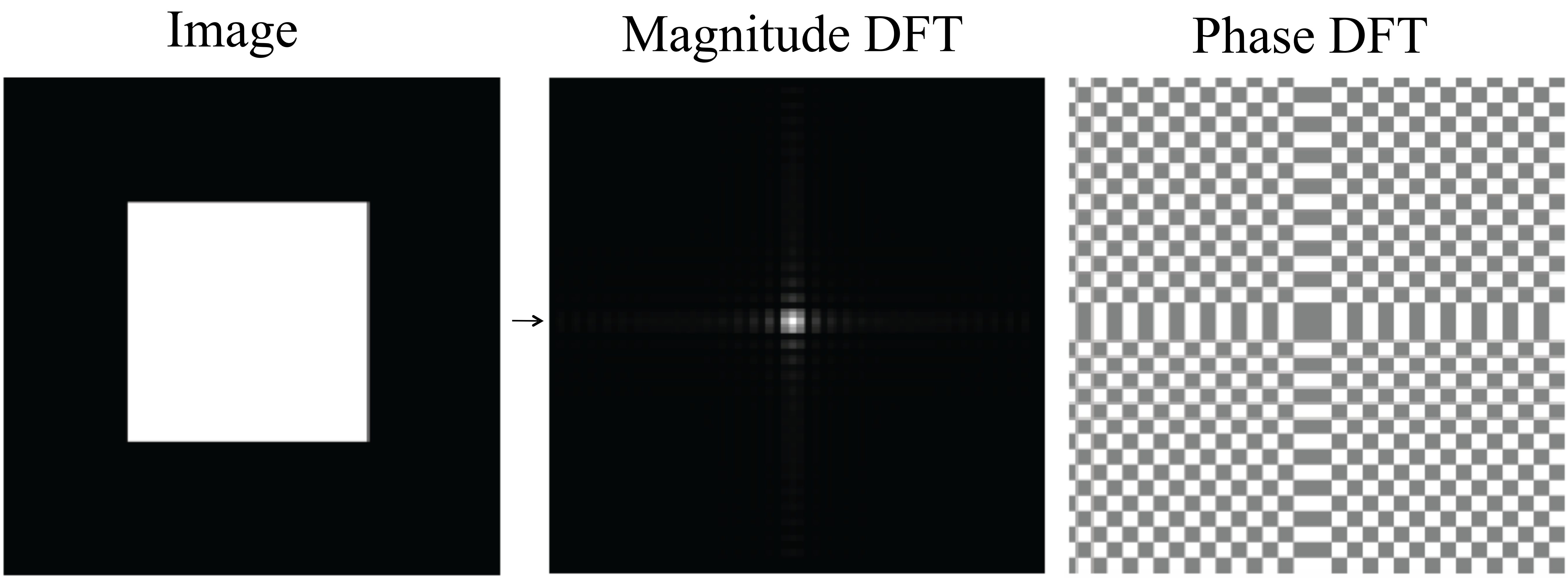
Source: Bill Freeman

Some important Fourier transforms



Source: Bill Freeman

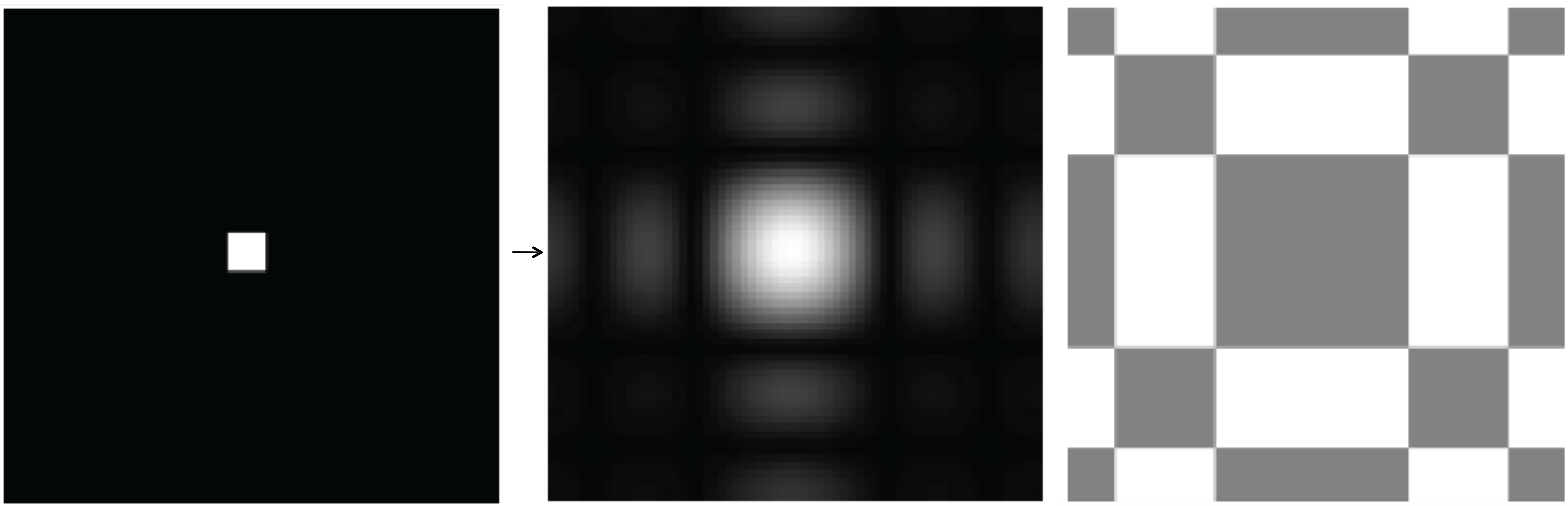
Some important Fourier transforms



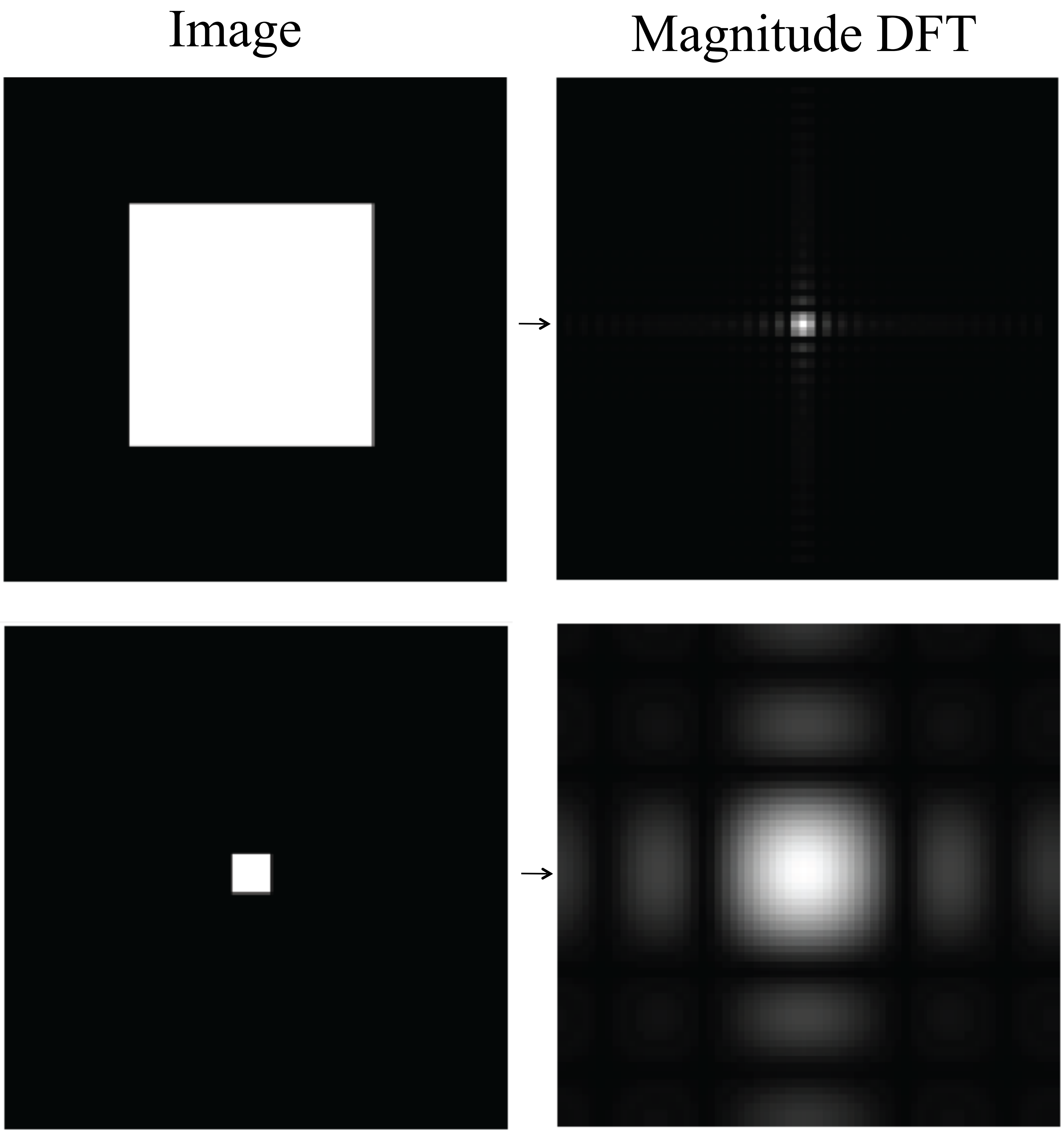
Source: Bill Freeman

Some important Fourier transforms

Image | Magnitude DFT | Phase DFT



Source: Bill Freeman



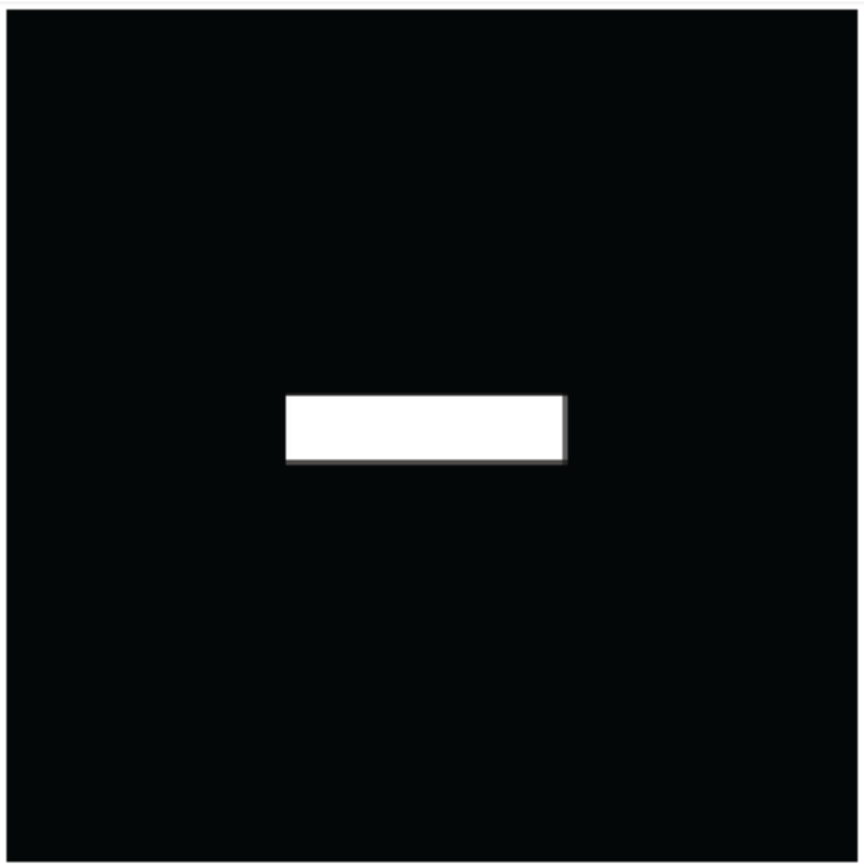
Scale

Small image details produce content in high spatial frequencies

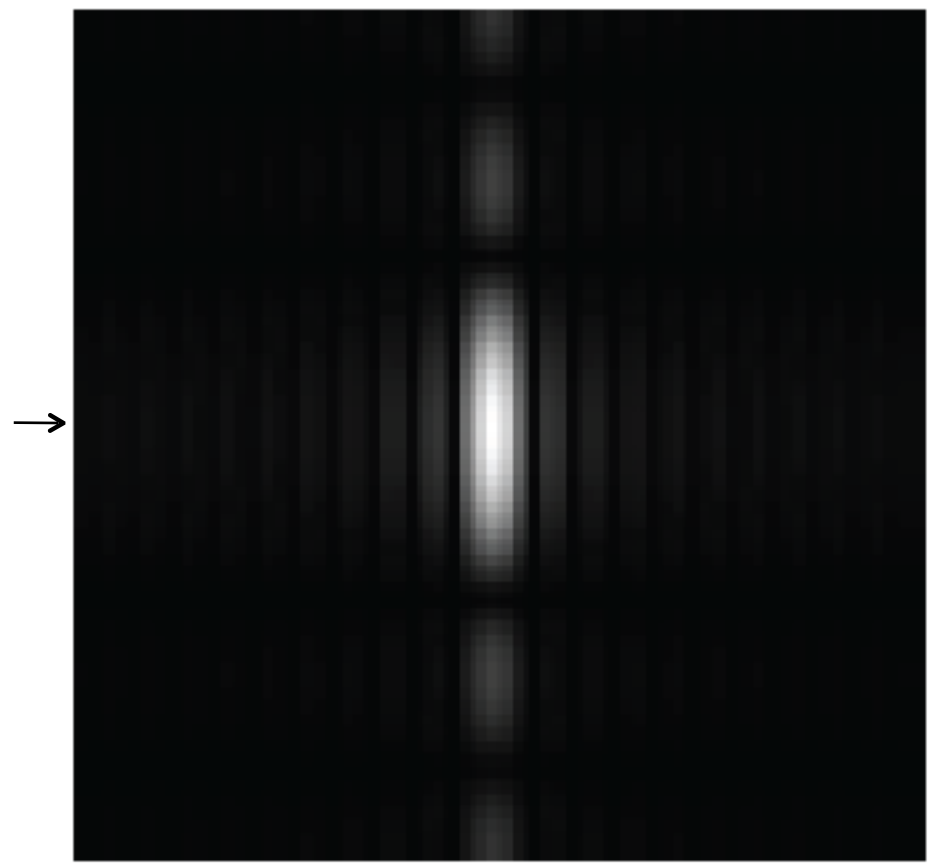
Source: Bill Freeman

Some important Fourier transforms

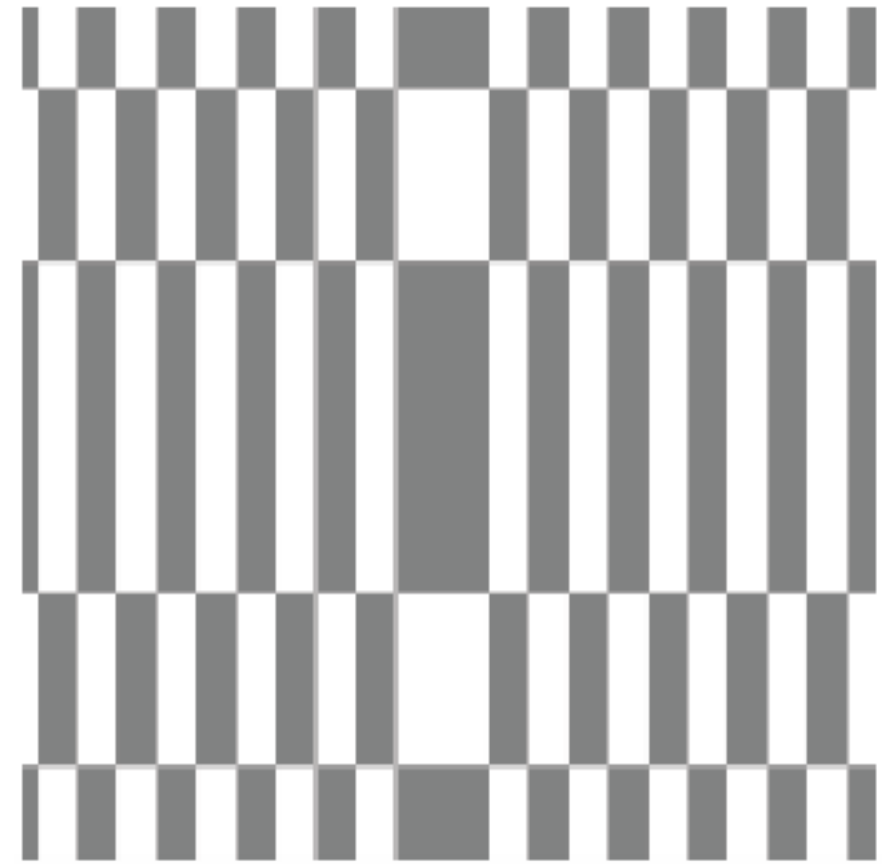
Image



Magnitude DFT

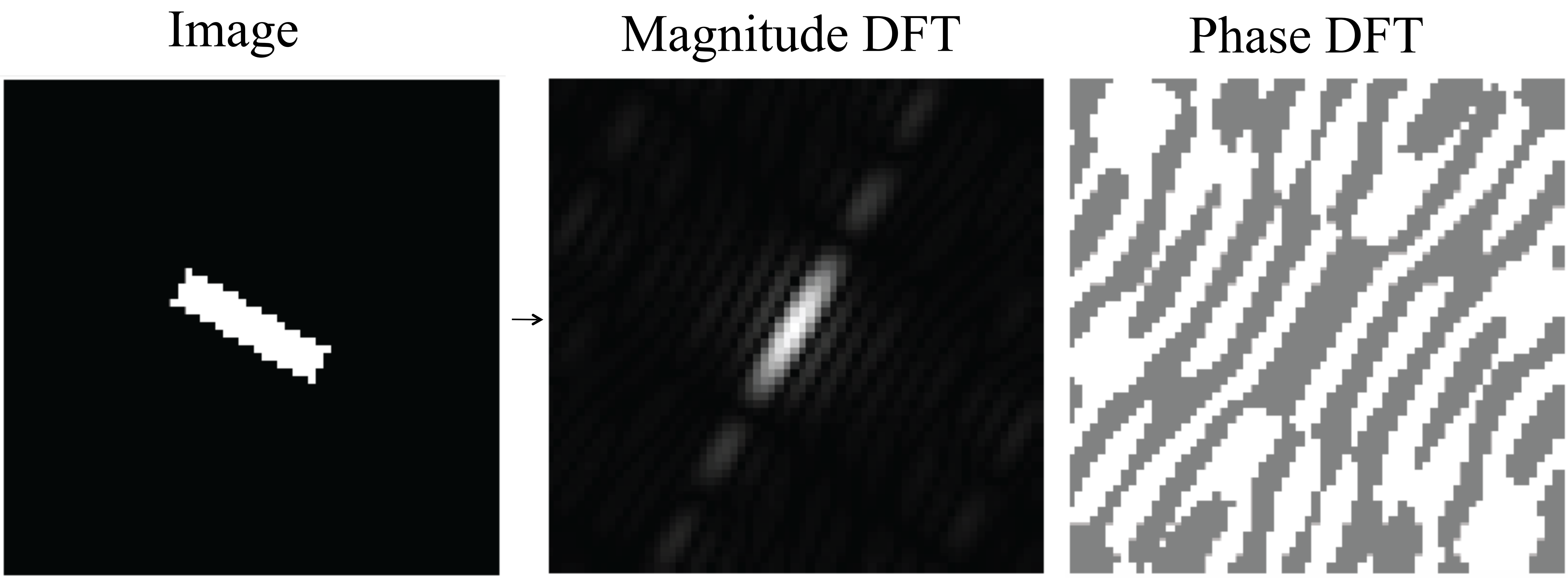


Phase DFT

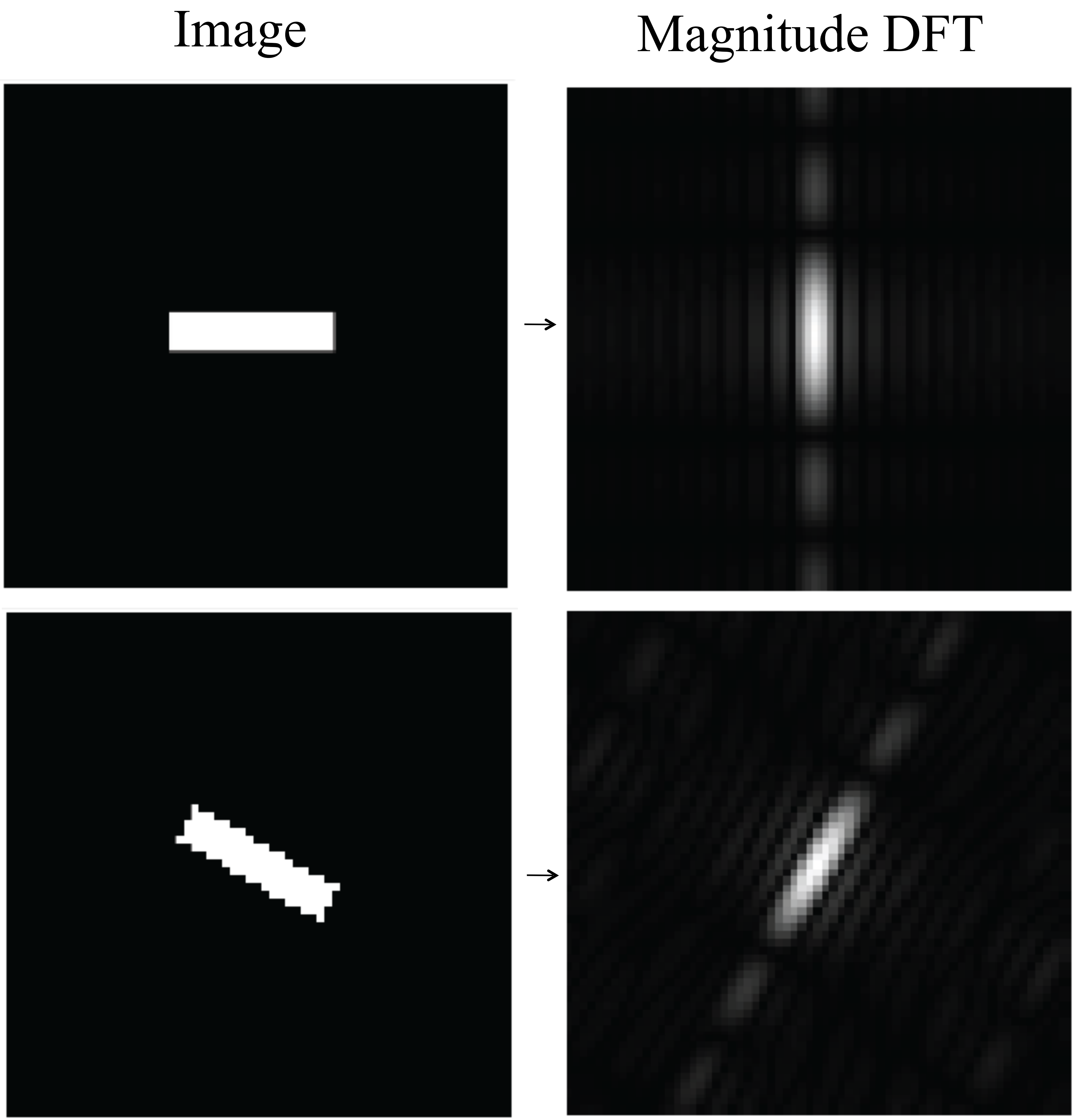


Source: Bill Freeman

Some important Fourier transforms



Source: Bill Freeman



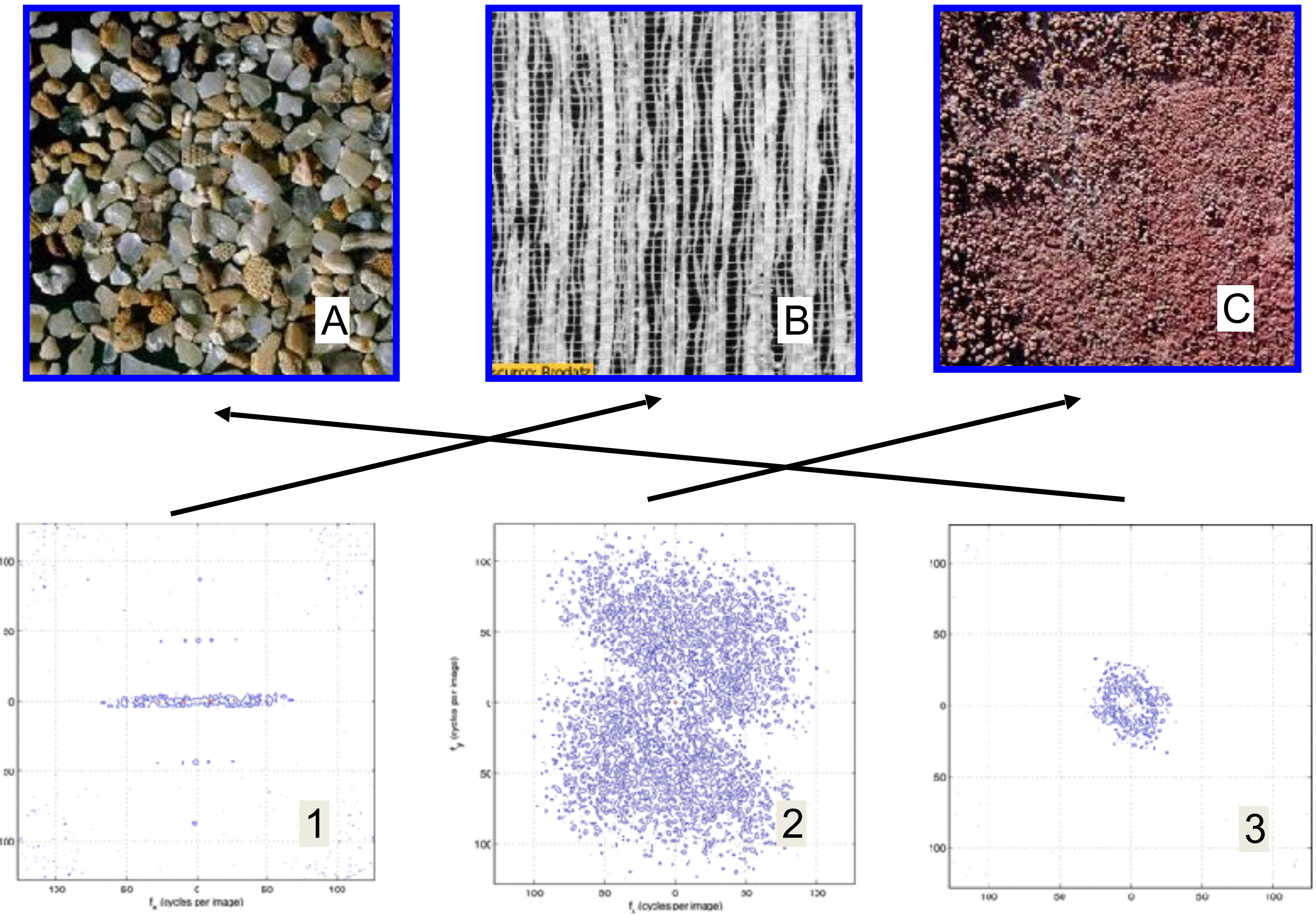
Orientation

A line transforms to a line oriented perpendicularly to the first.

Source: Bill Freeman

Game: find the right pairs

Images

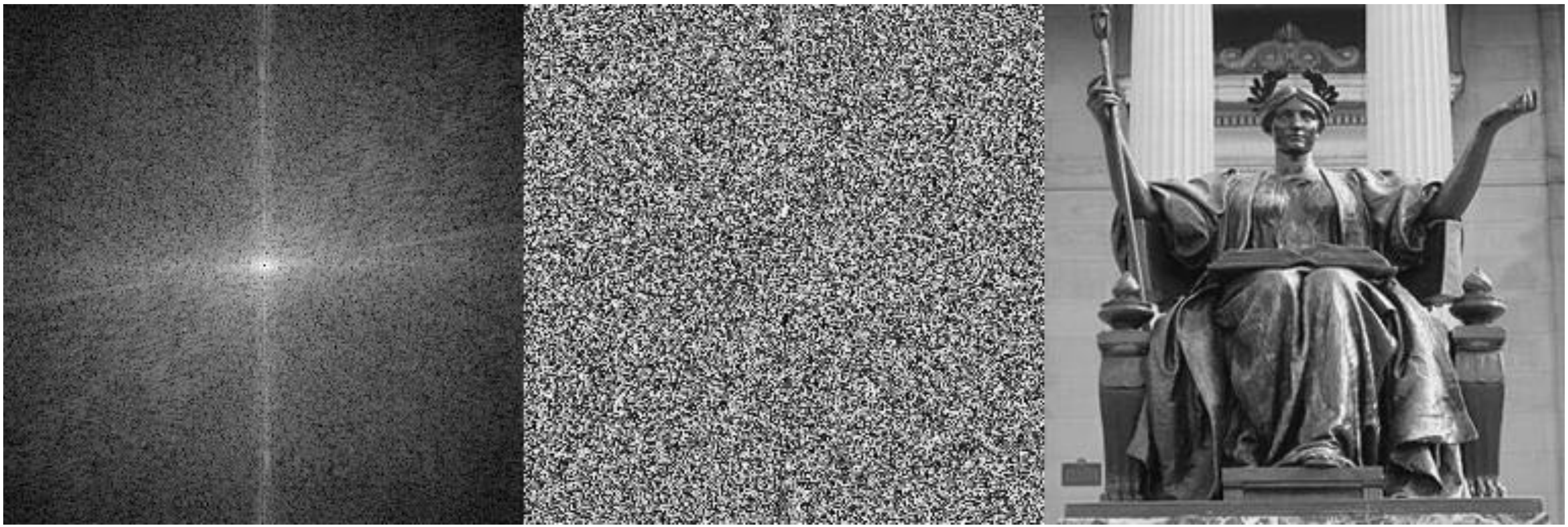


DFT magnitude

fx(cycles/image pixel size)   fx(cycles/image pixel size)   fx(cycles/image pixel size)

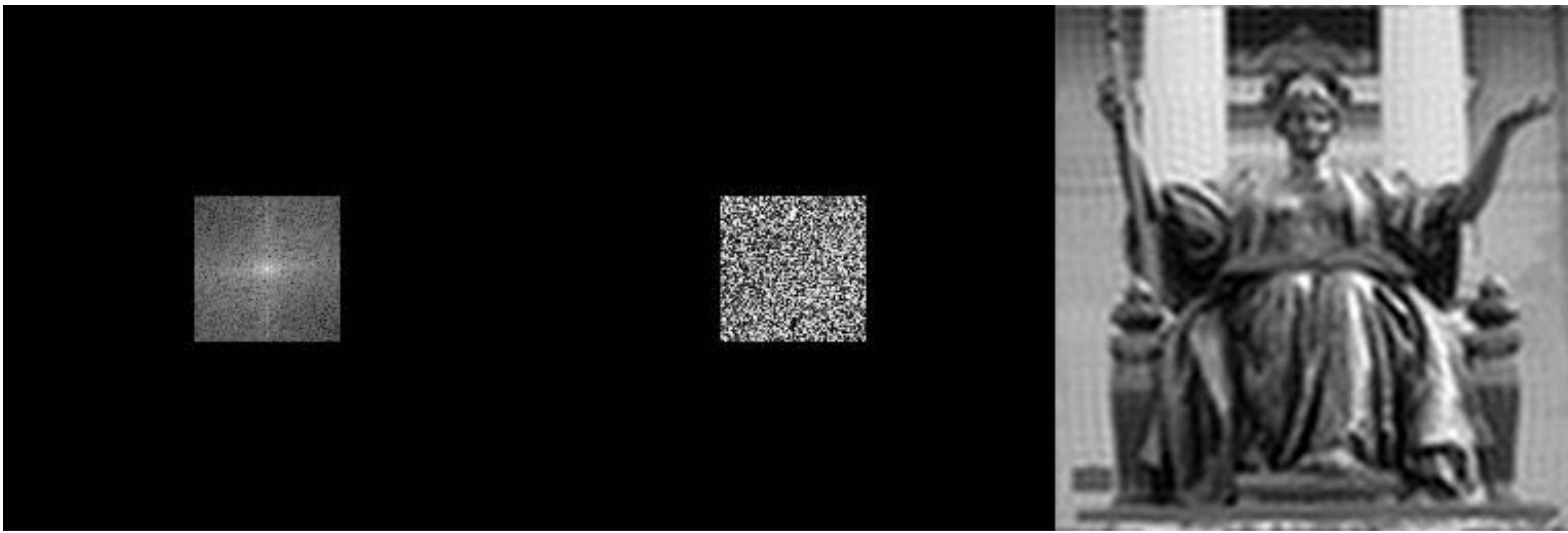
Source: Bill Freeman

Magnitude Phase Reconstruction



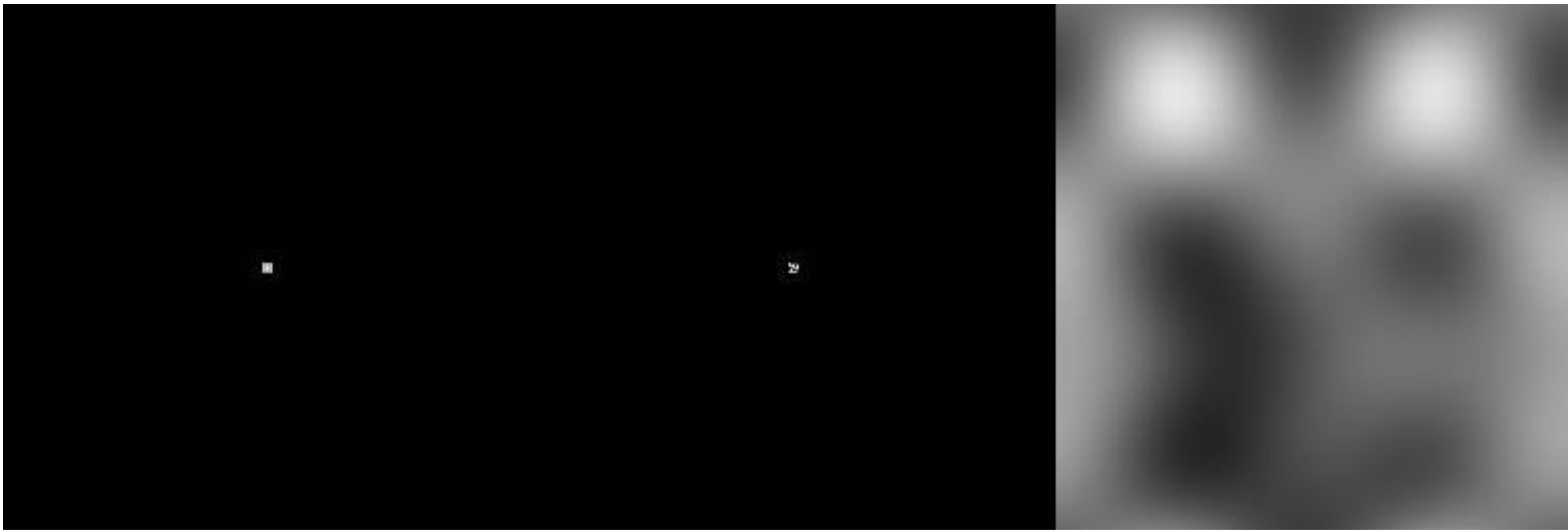
Removing High Frequencies

Magnitude Phase Reconstruction

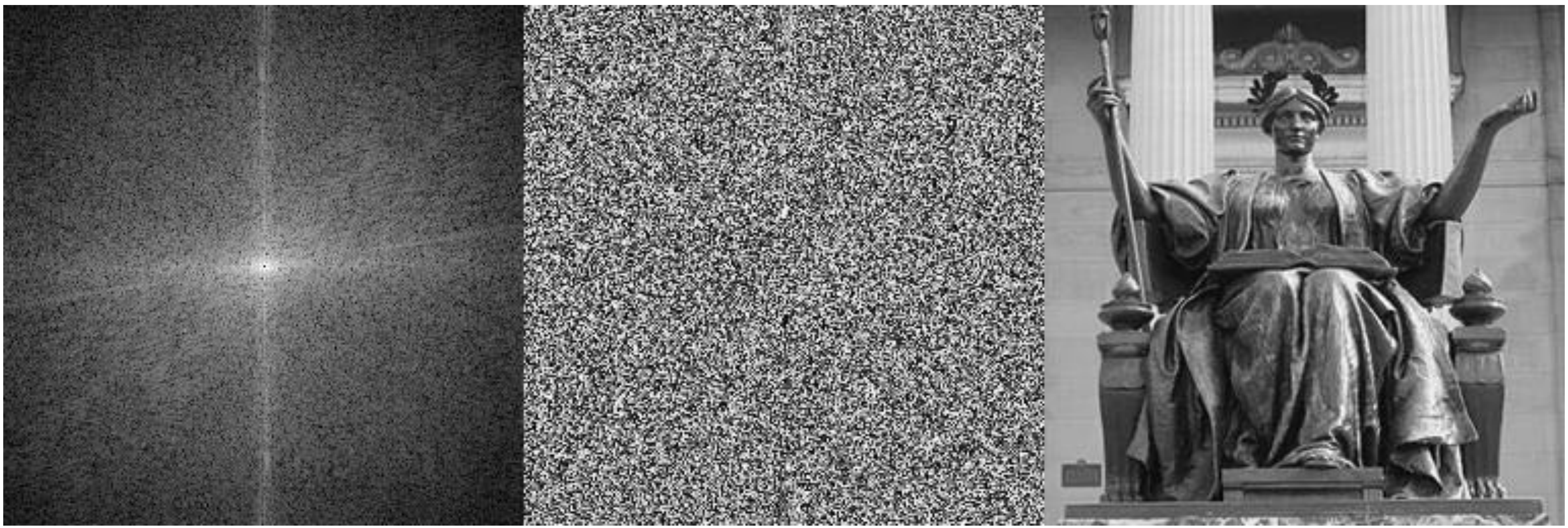


Removing High Frequencies

Magnitude | Phase | Reconstruction

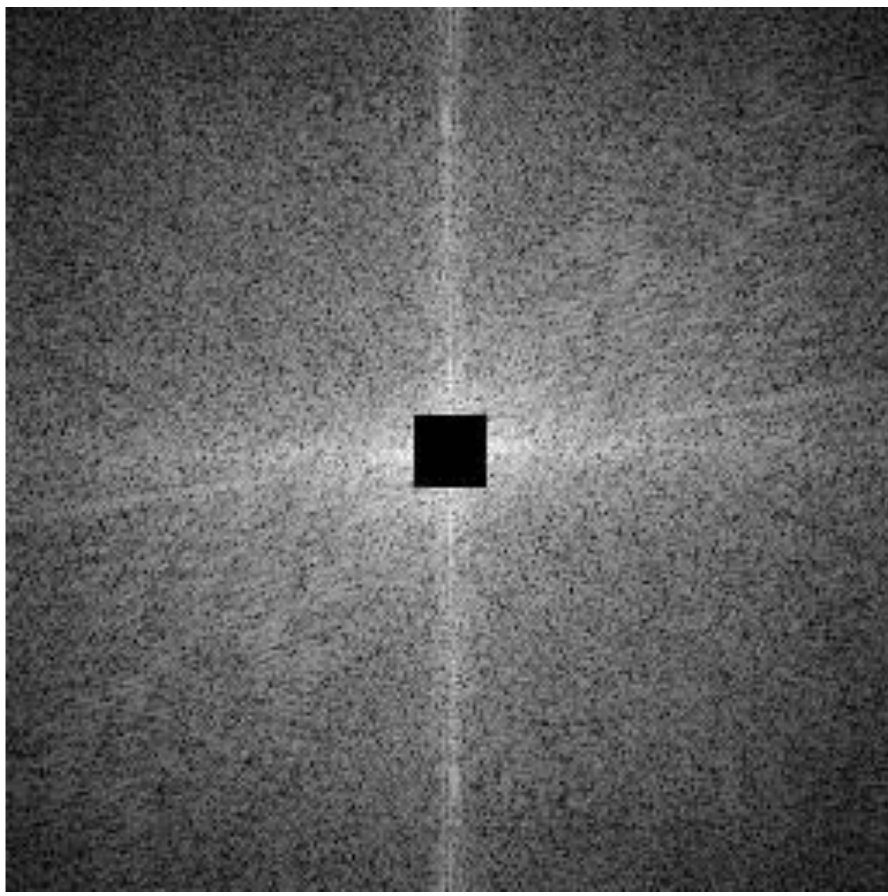


Magnitude Phase Reconstruction

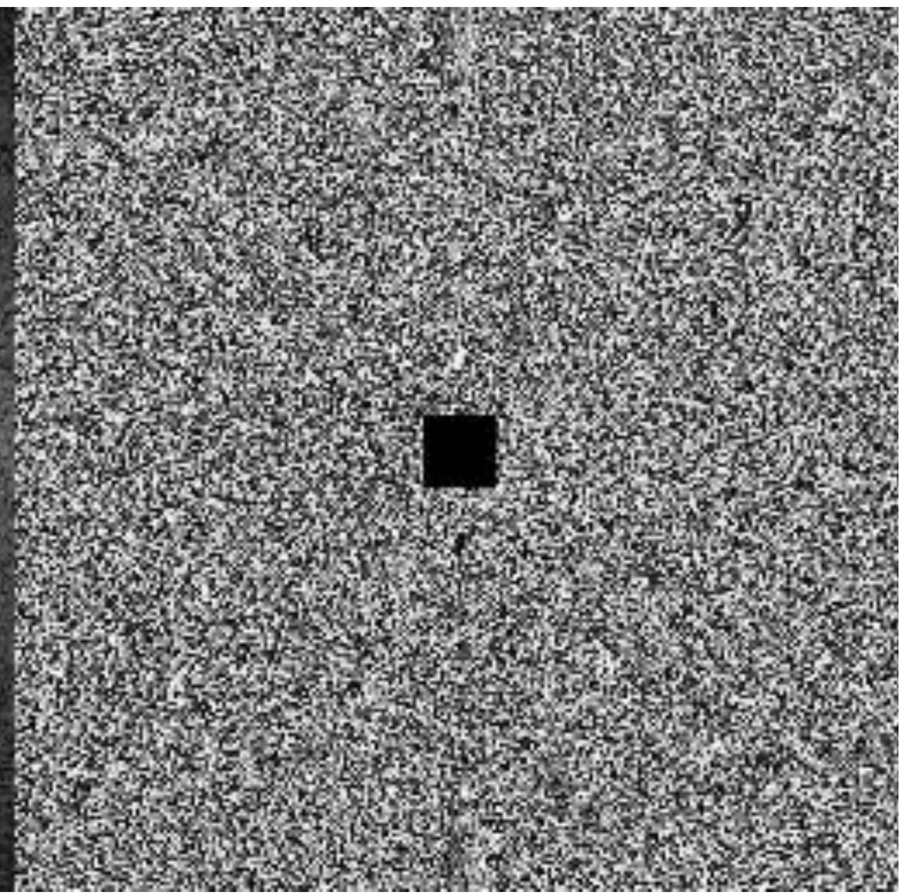


What will happen?

Magnitude



Phase

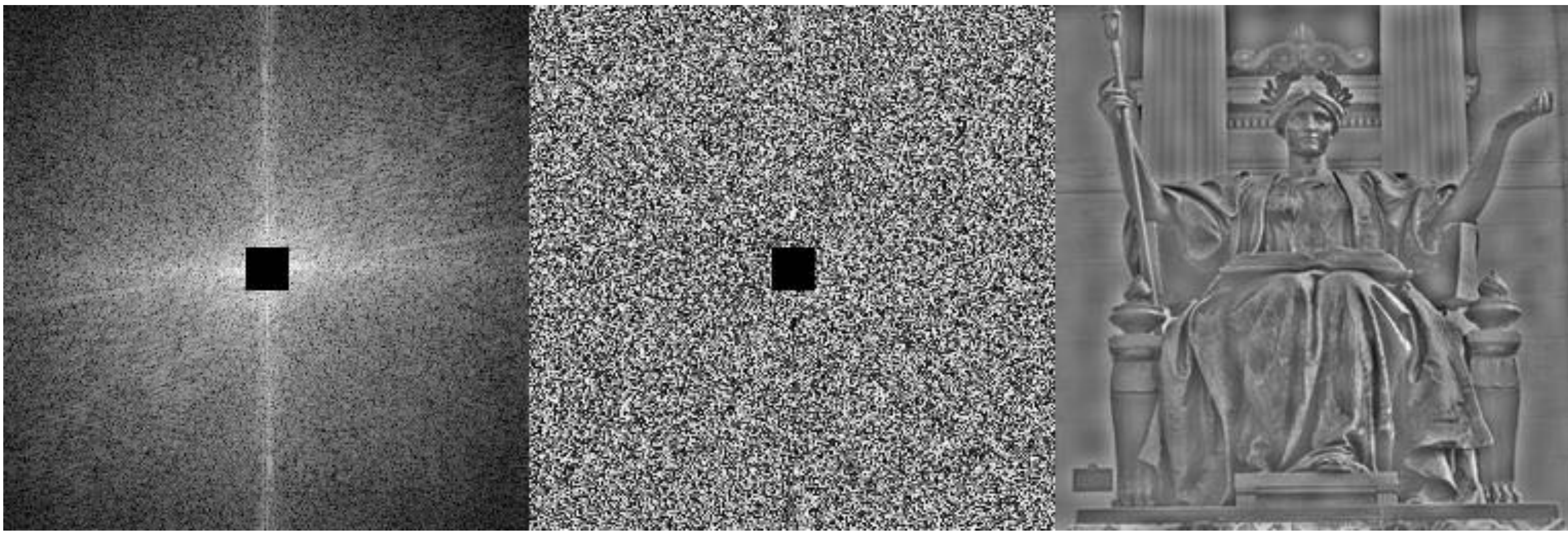


Reconstruction

?

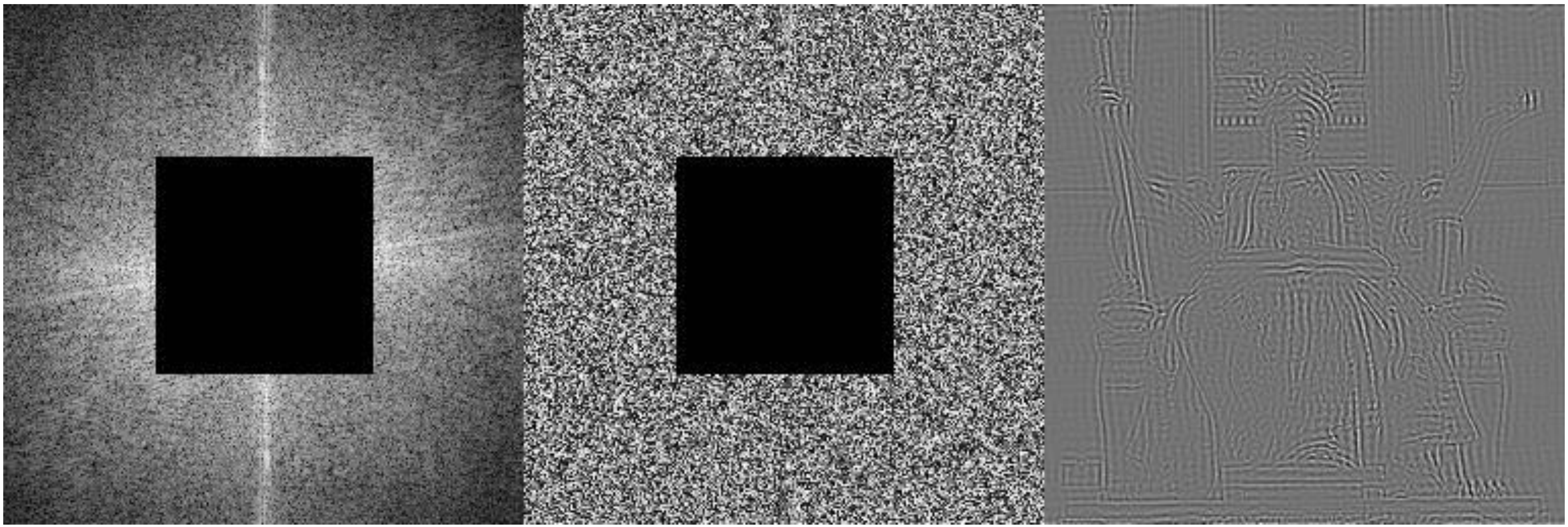
Removing Low Frequencies

Magnitude Phase Reconstruction

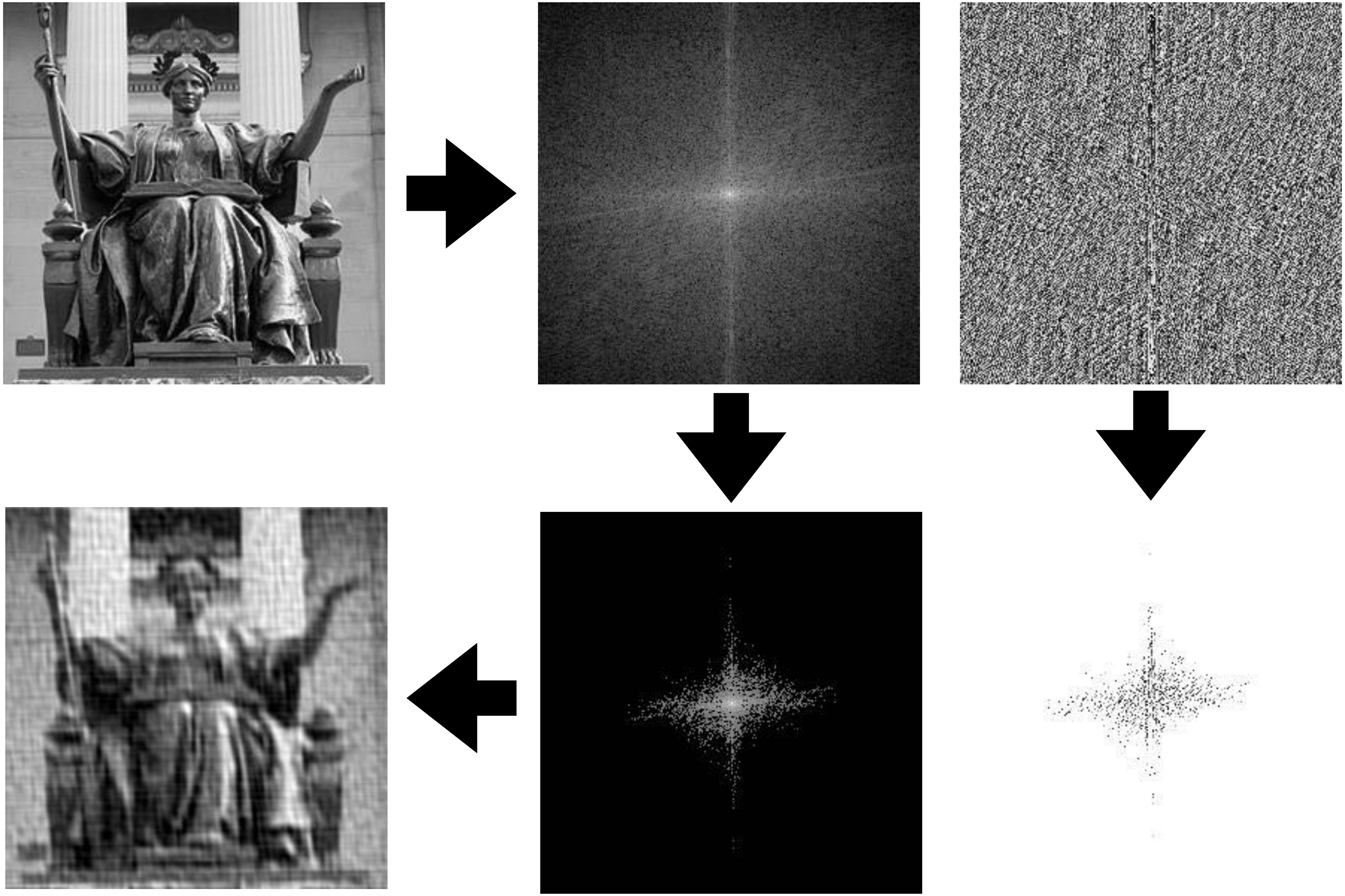


Removing Low Frequencies

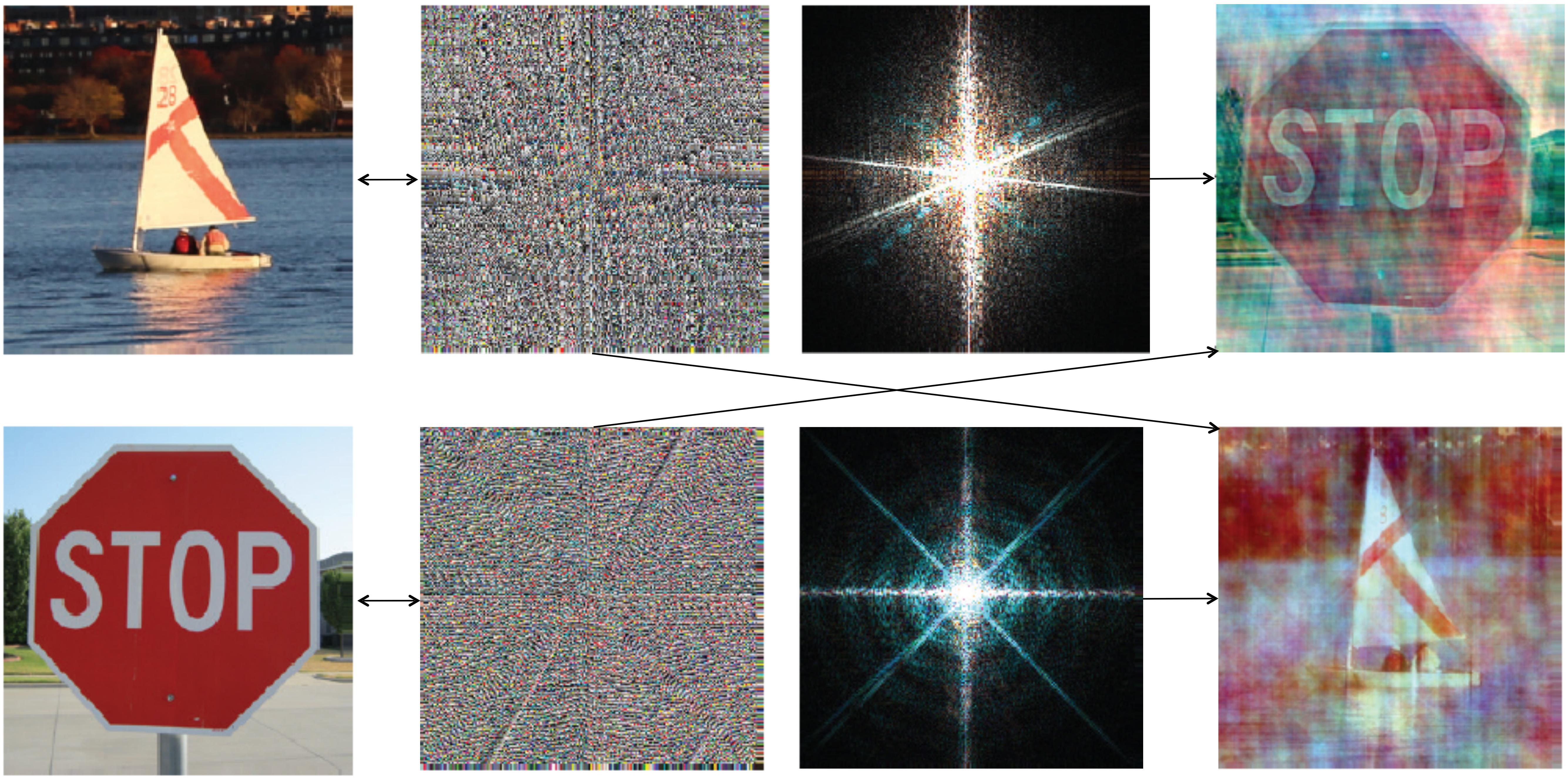
Magnitude Phase Reconstruction



Compression



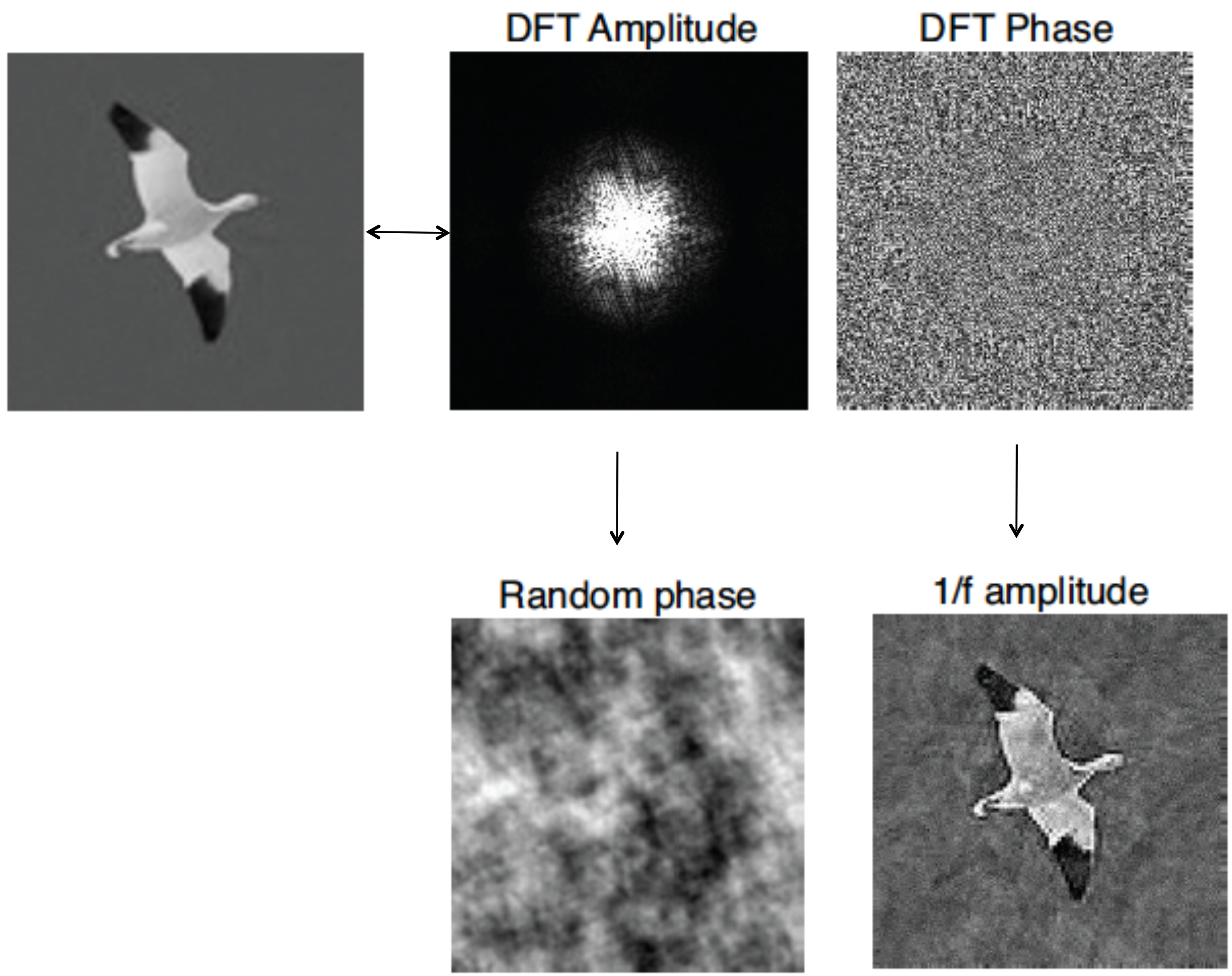
Phase and Magnitude



Each color channel is processed in the same way.

Source: Bill Freeman

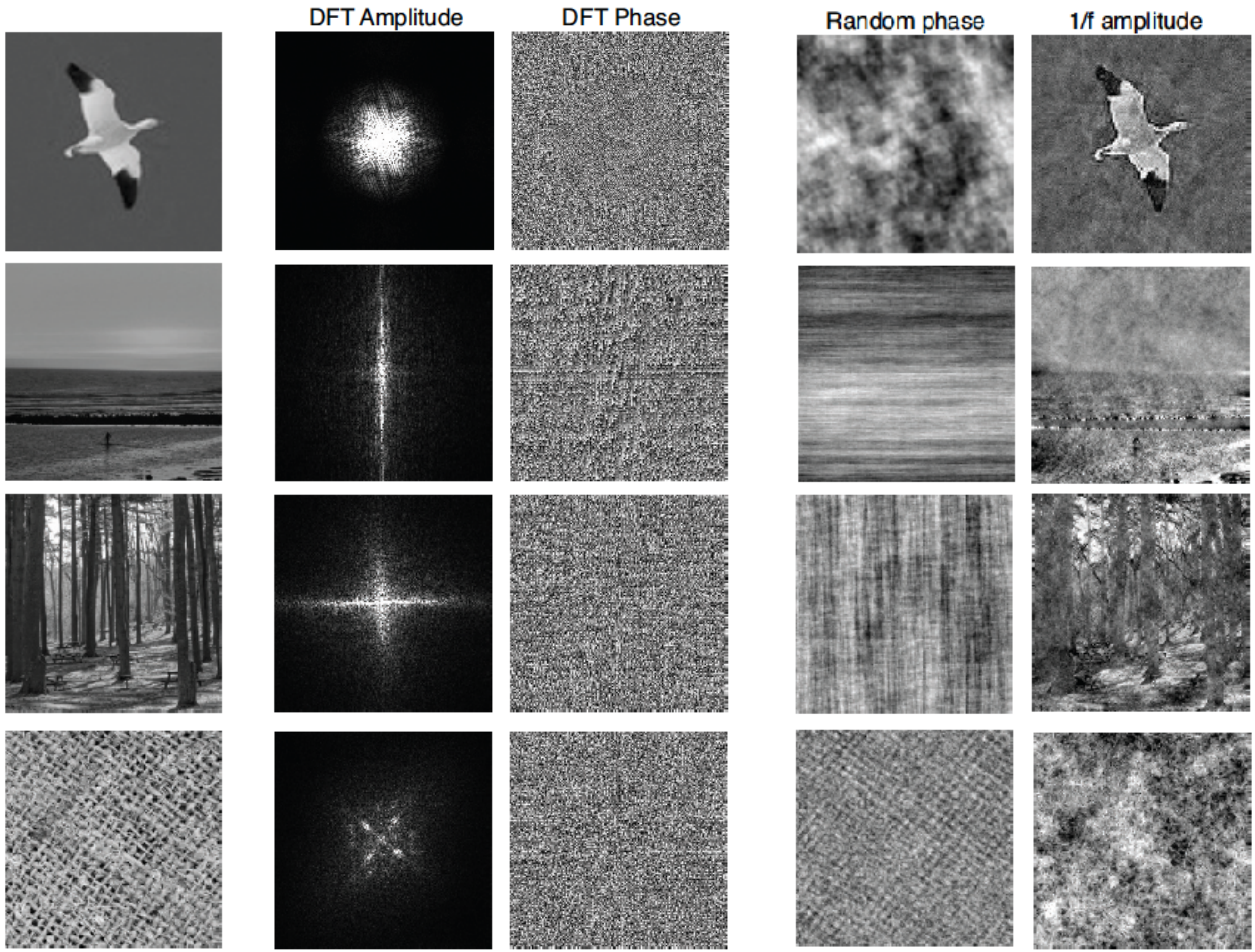
Phase and magnitude



Using random amplitude does not look good.

Source: Bill Freeman

Does phase always win?



Convolution

- Convolution in time space is multiplication in frequency space:

$$g(x) * h(x) = \mathcal{F}^{-1}\left[\mathcal{F}[g(x)] \cdot \mathcal{F}[h(x)]\right]$$

- Convolution in frequency space is multiplication in time space:

$$\mathcal{F}[g(x)] * \mathcal{F}[h(x)] = \mathcal{F}\left[g(x) \cdot h(x)\right]$$

Which is more computationally efficient?

FFT in Practice

Fast Training of Convolutional Networks through FFTs

Michael Mathieu
Courant Institute of Mathematical Sciences
New York University
mathieu@cs.nyu.edu

Mikael Henaff
Courant Institute of Mathematical Sciences
New York University
mbh305@nyu.edu

Yann LeCun
Courant Institute of Mathematical Sciences
New York University
yann@cs.nyu.edu

Convolutional Neural Networks

Correlation Filters (600 FPS on CPU)

Source: Deva Ramanan

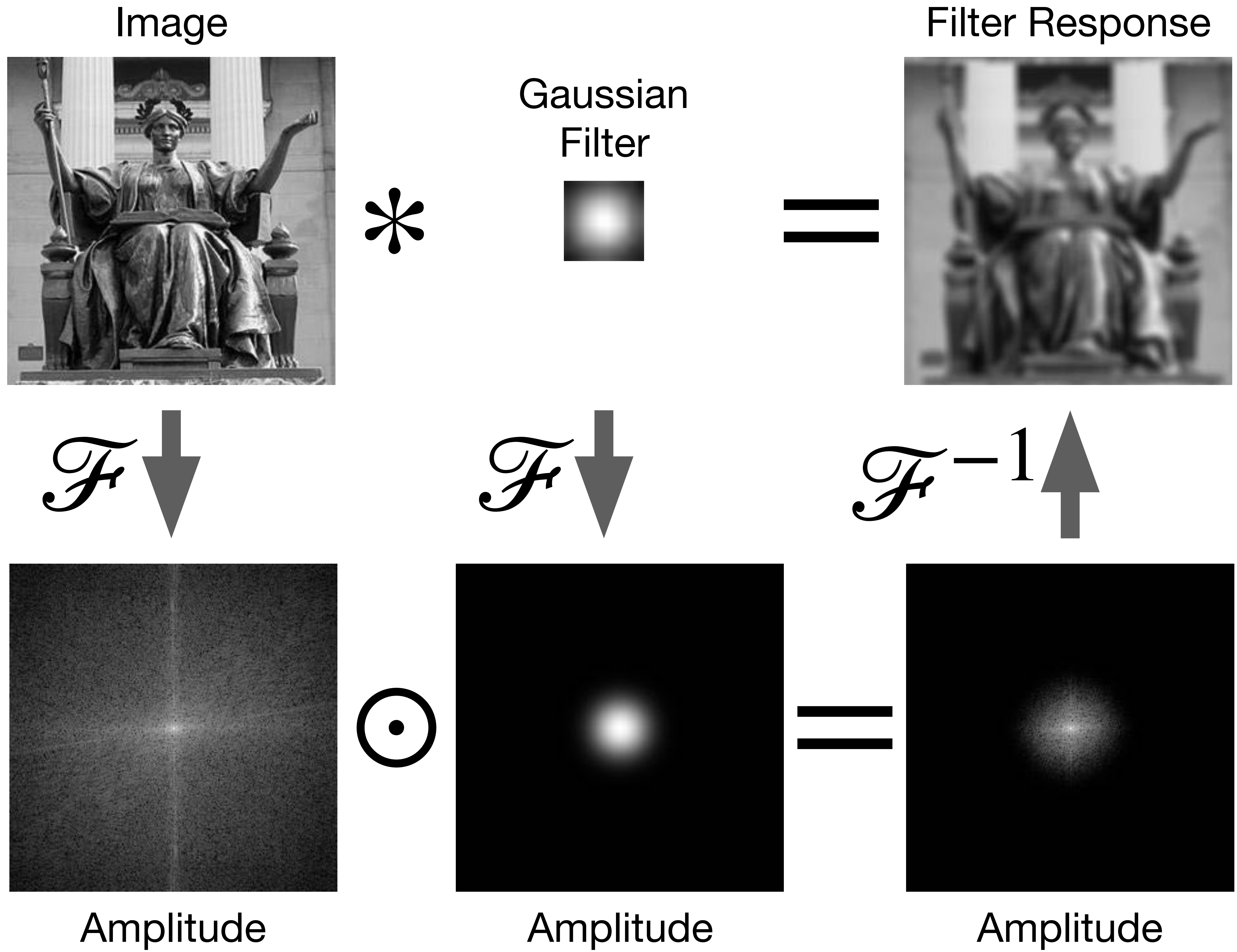
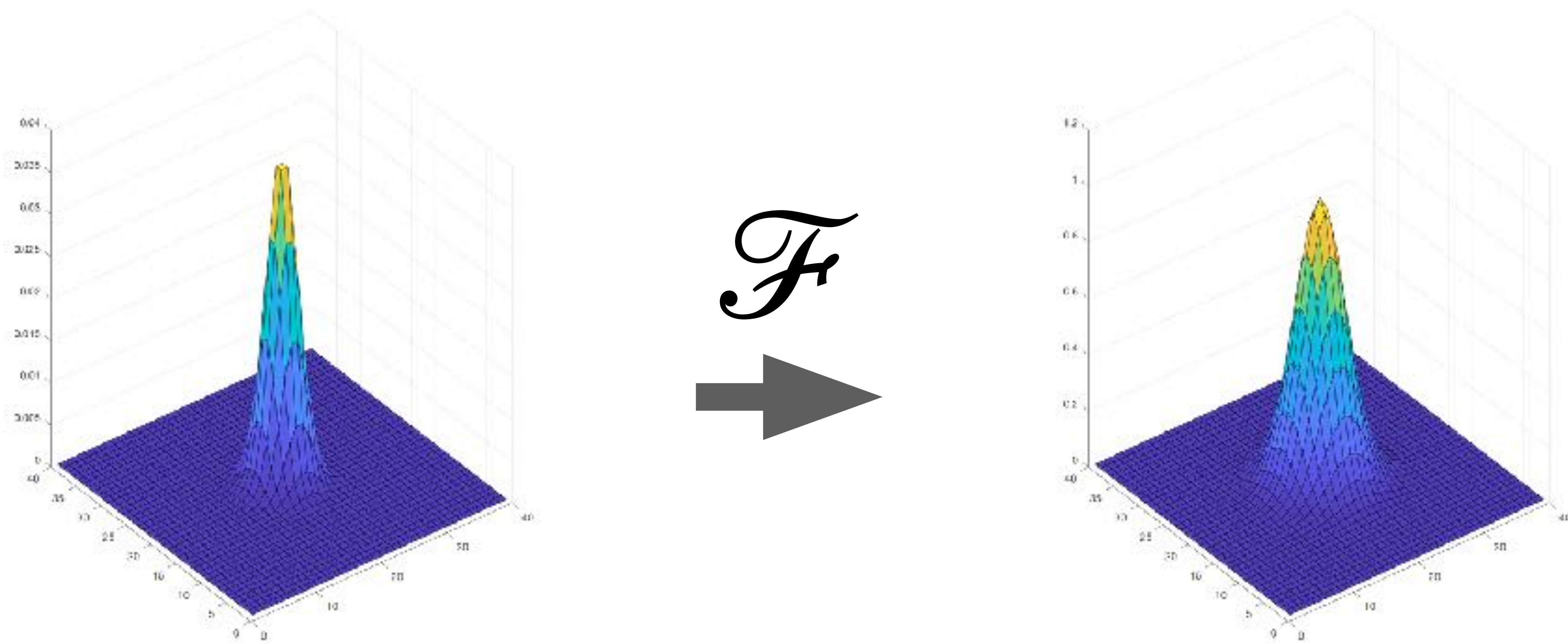


Image
Gaussian
Filter
Filter Response
$\mathcal{F}$
$\mathcal{F}$
$\mathcal{F}^{-1}$
Amplitude
Amplitude
Amplitude

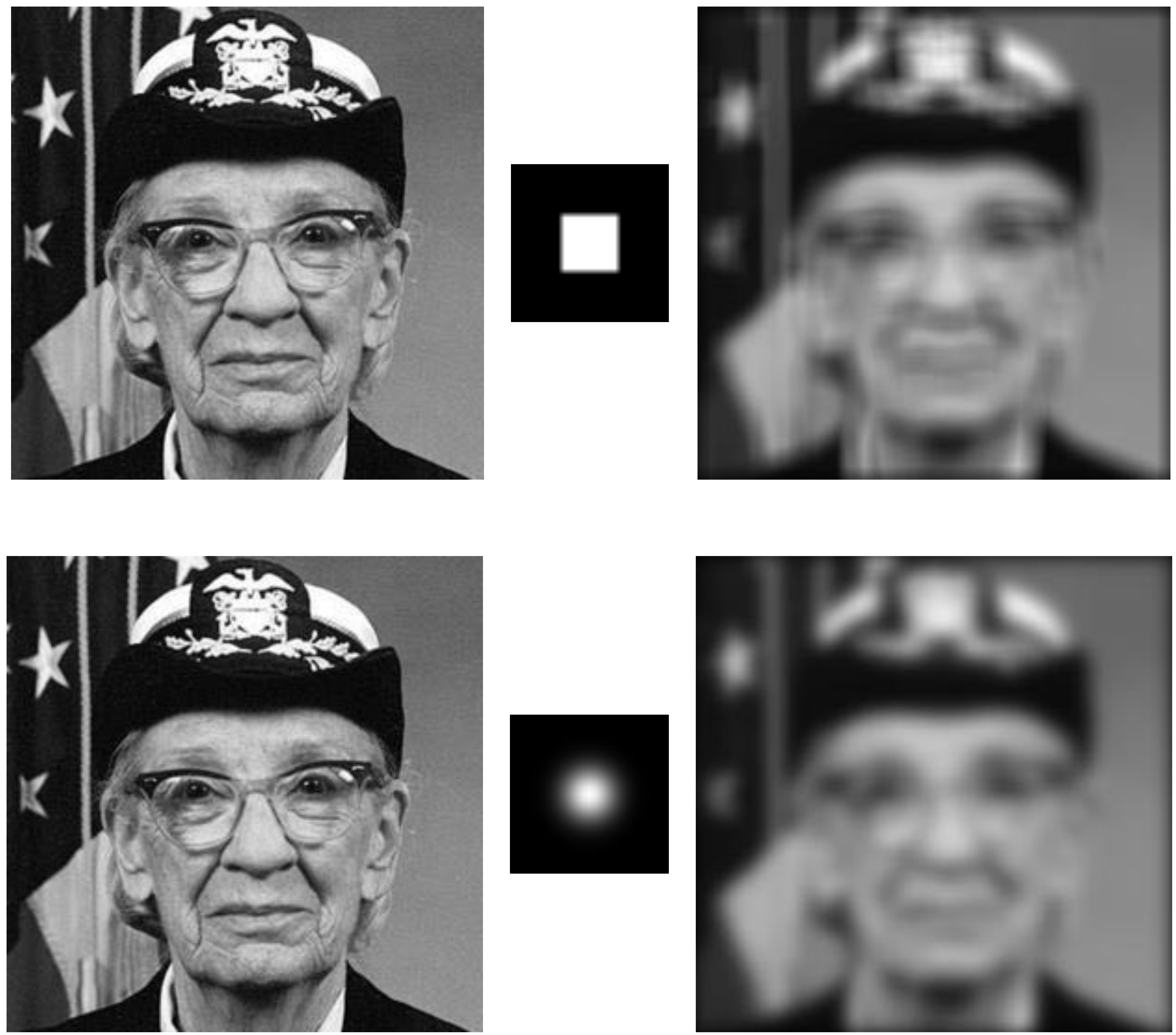
FT of Gaussian is Gaussian



Gaussian (sigma = 2)

Gaussian

Remember box filters?



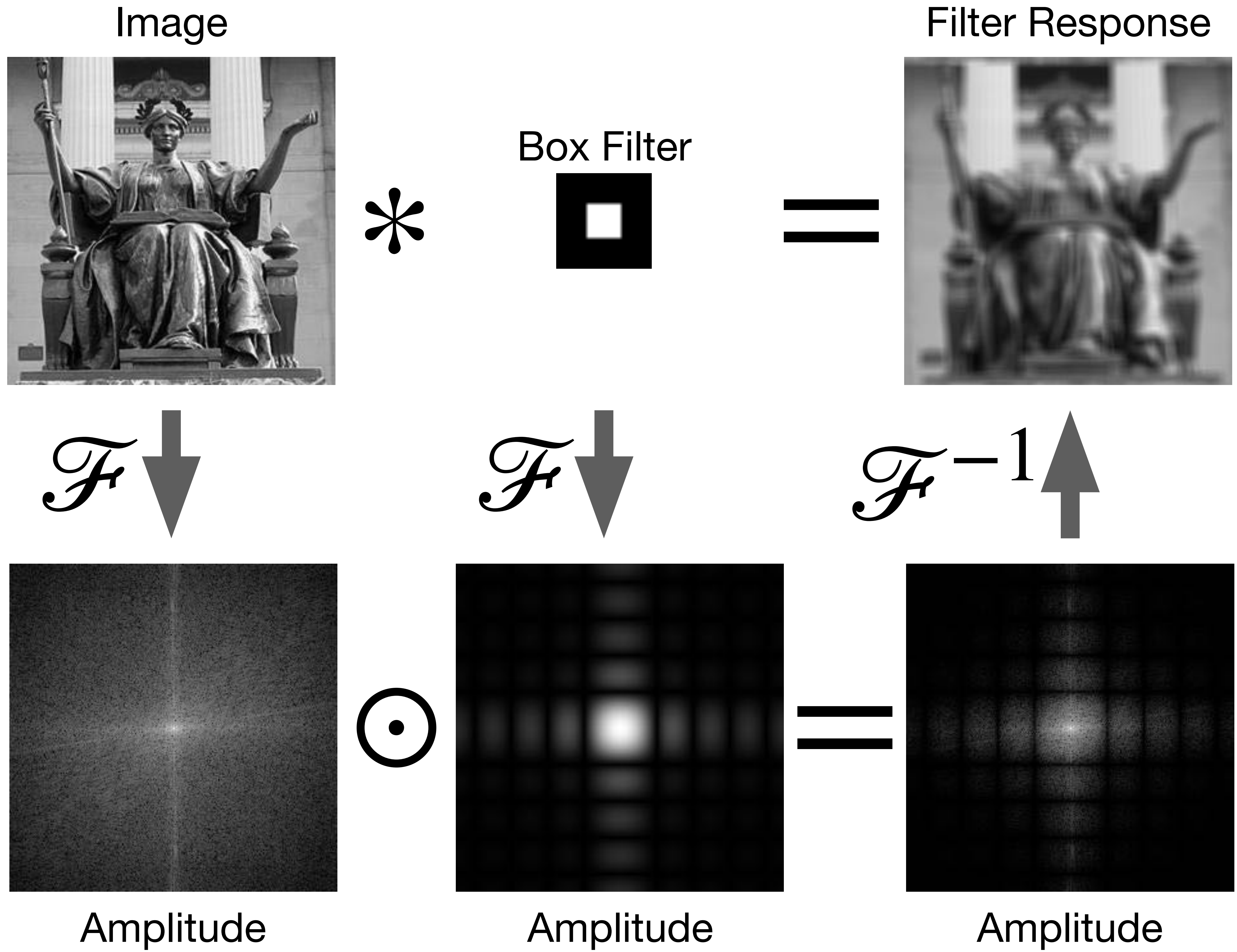
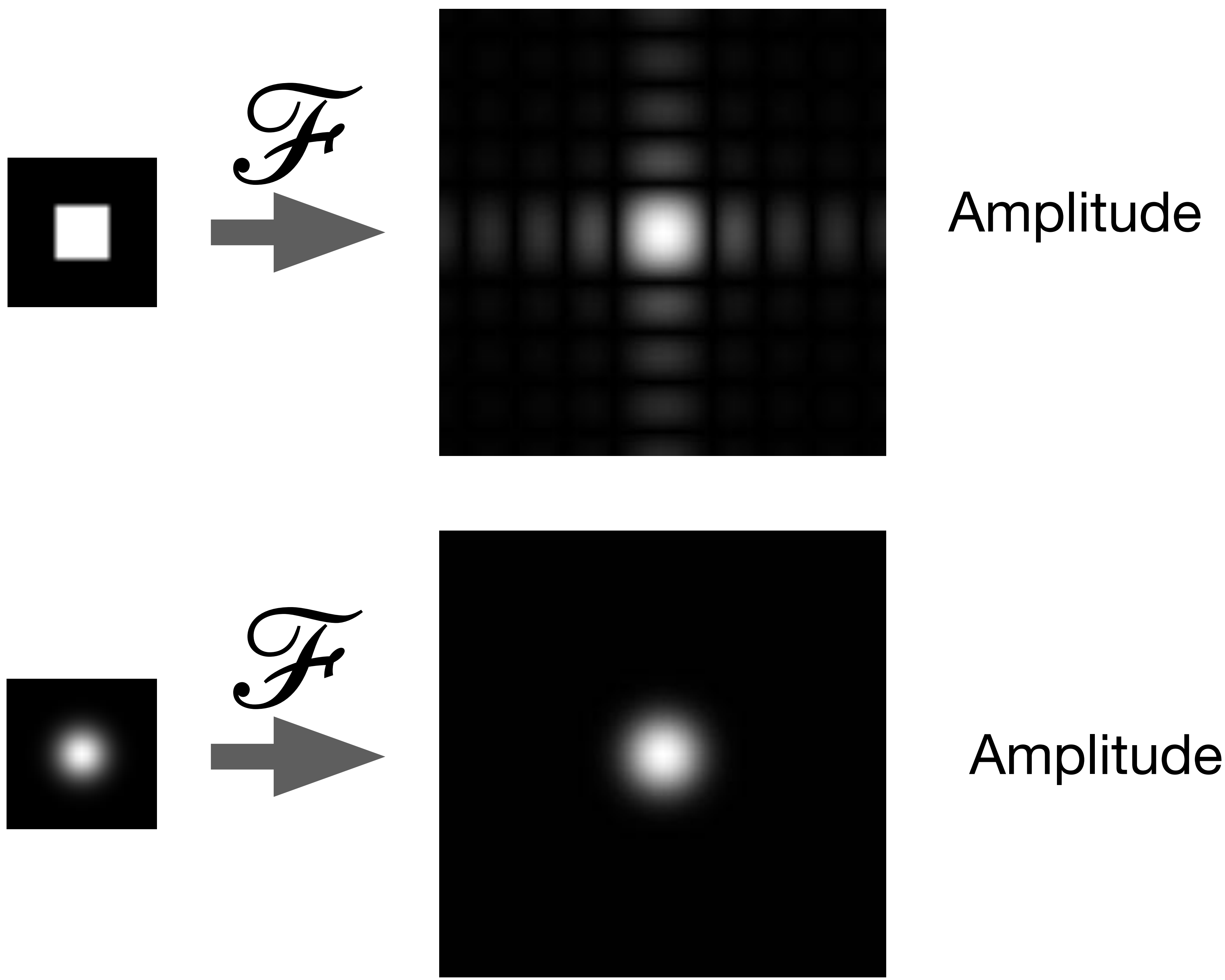


Image
Box Filter
Filter Response
Amplitude
Amplitude
Amplitude

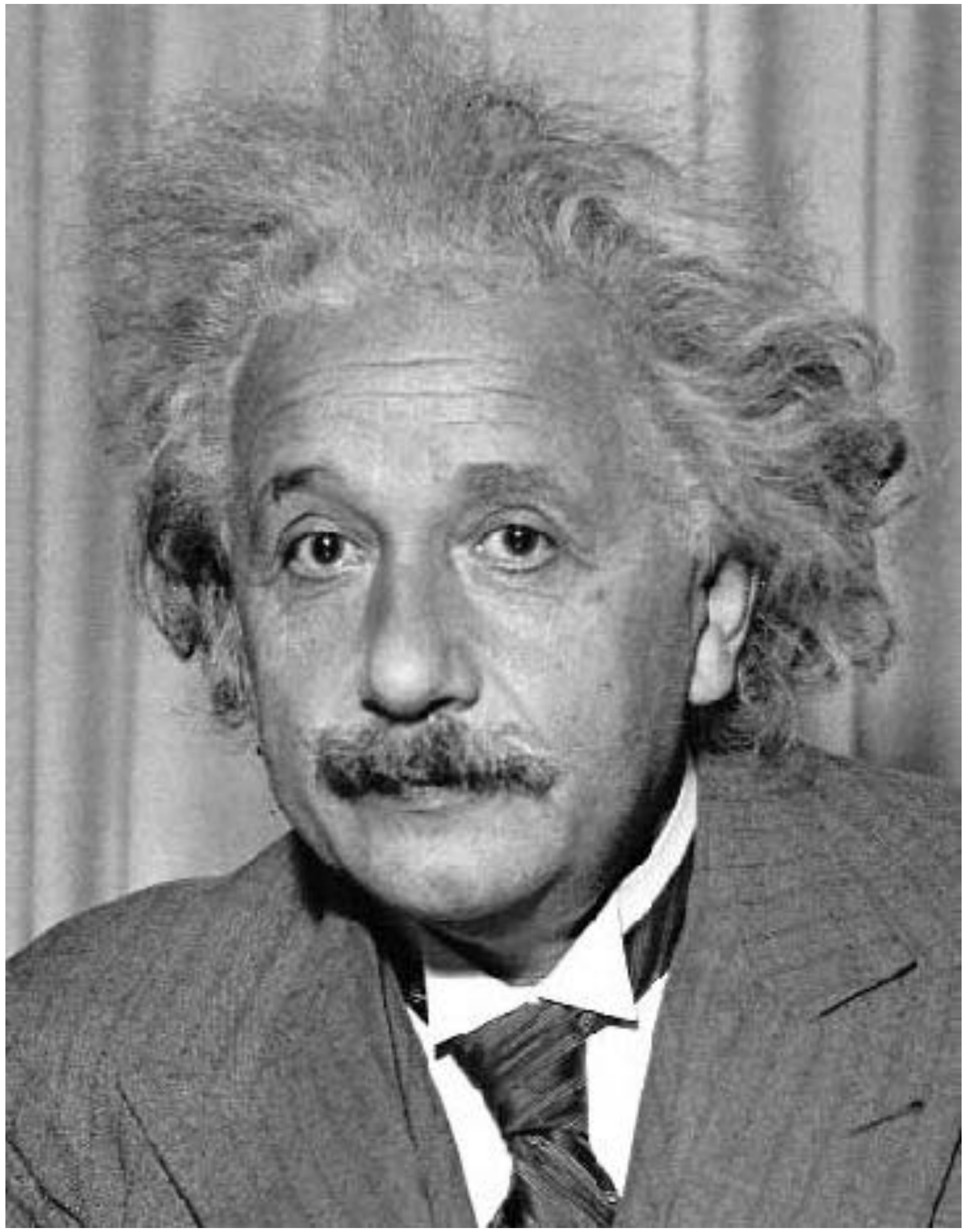
FT of Box Filters

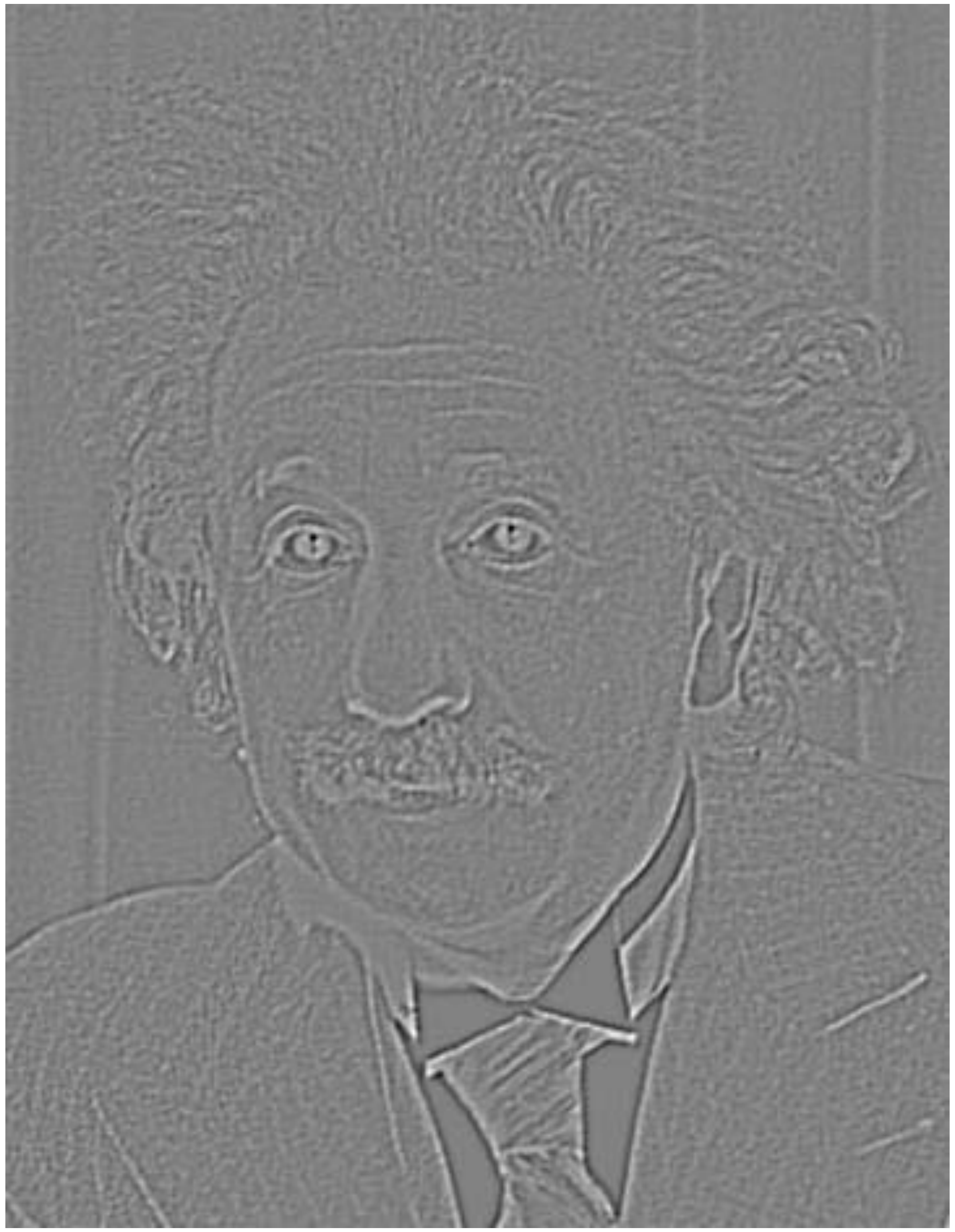


Remember Laplacian filters?

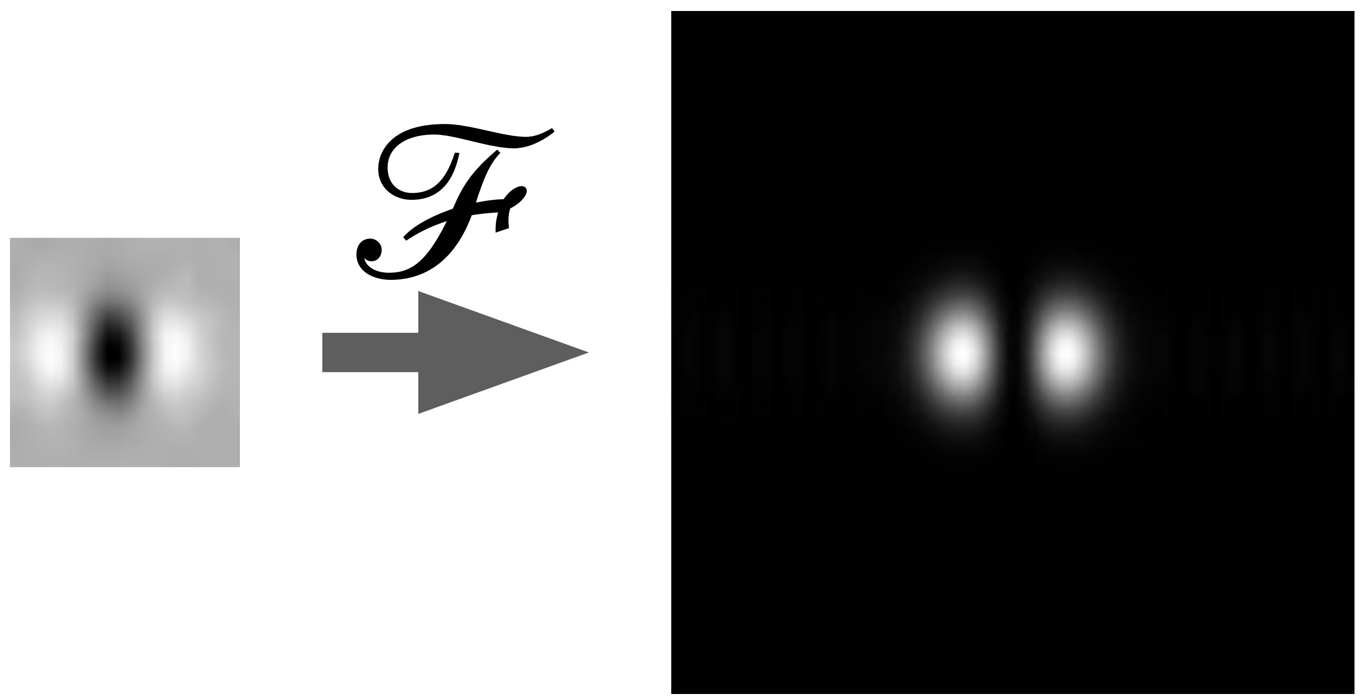
f

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$



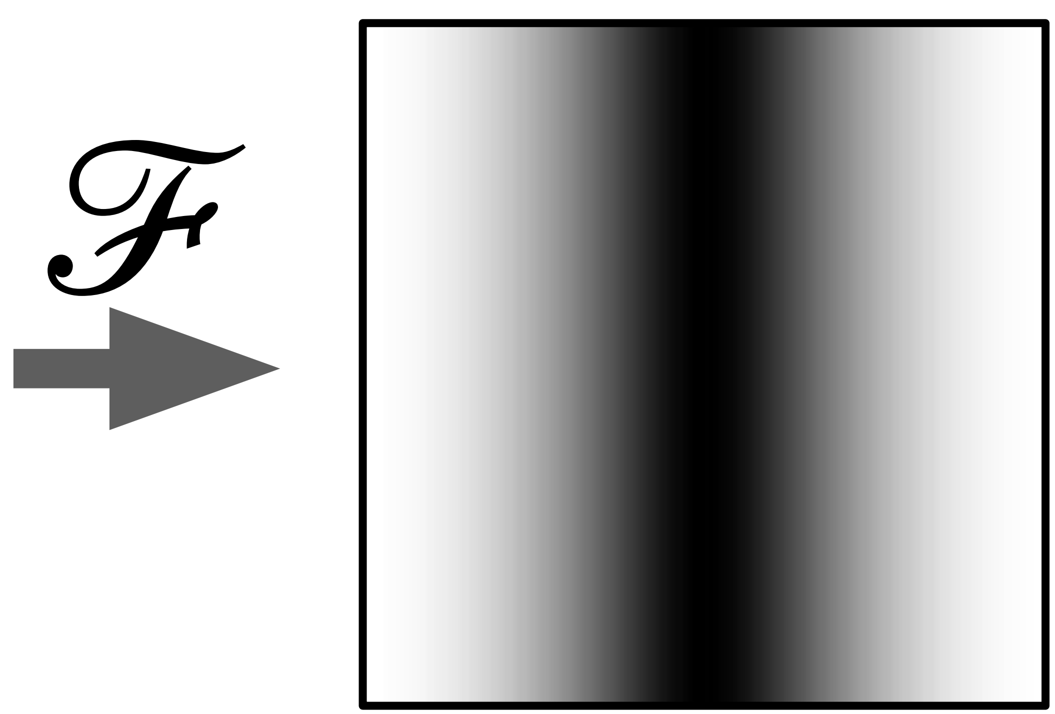


FT of Laplacian Filters



Amplitude

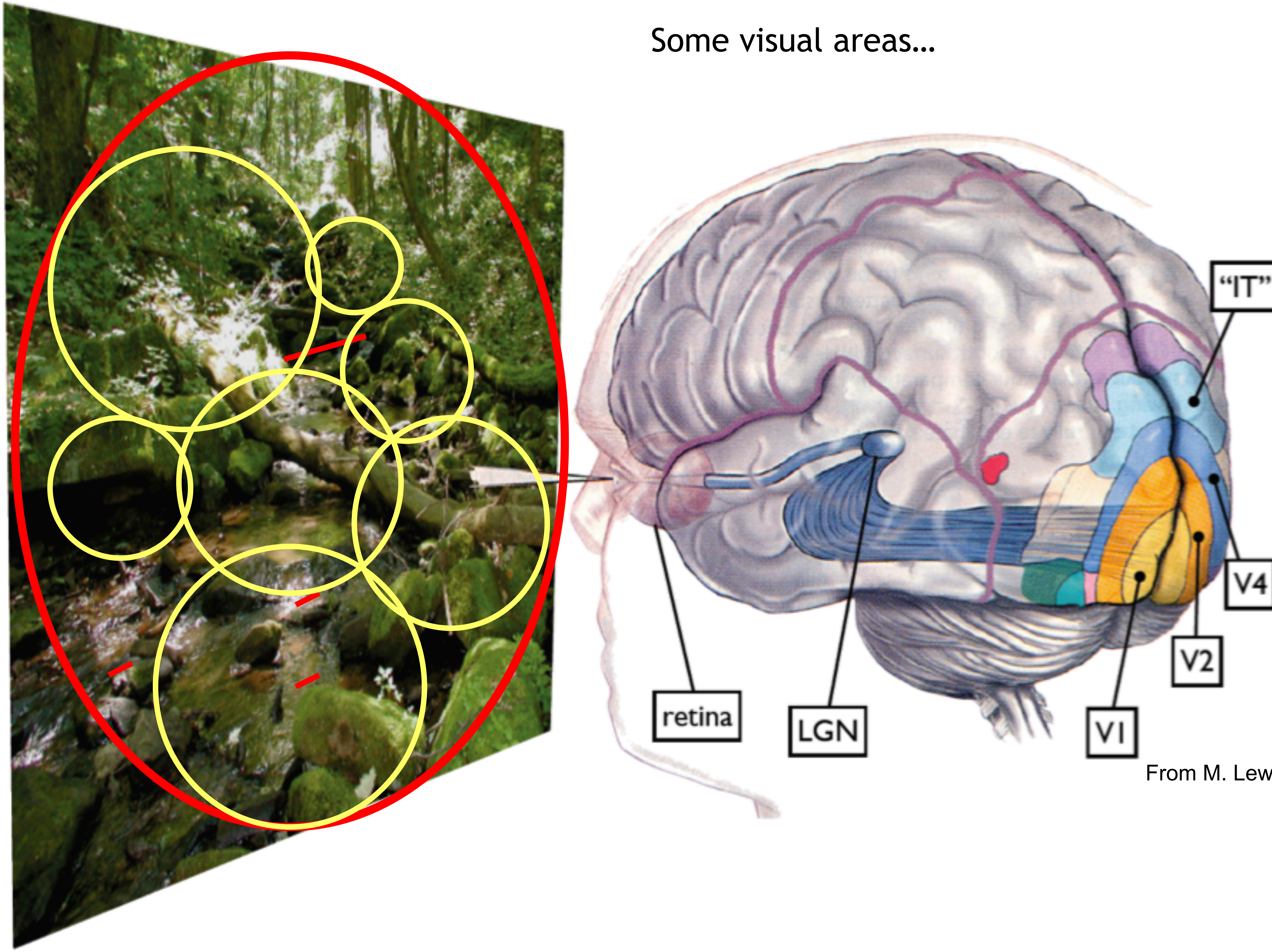
$[-1,2,1]$



Amplitude

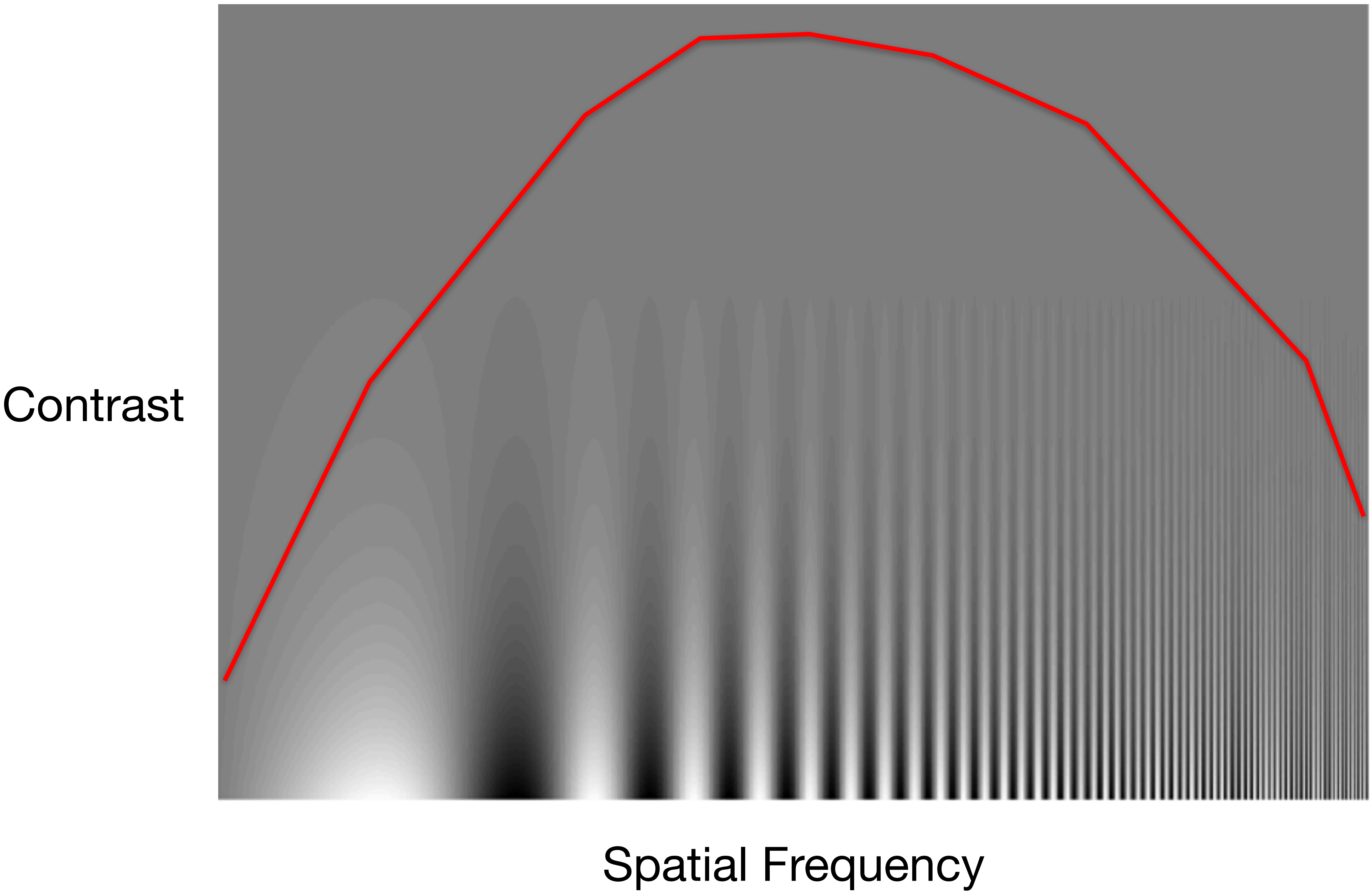
Creating an Illusion

Some visual areas…

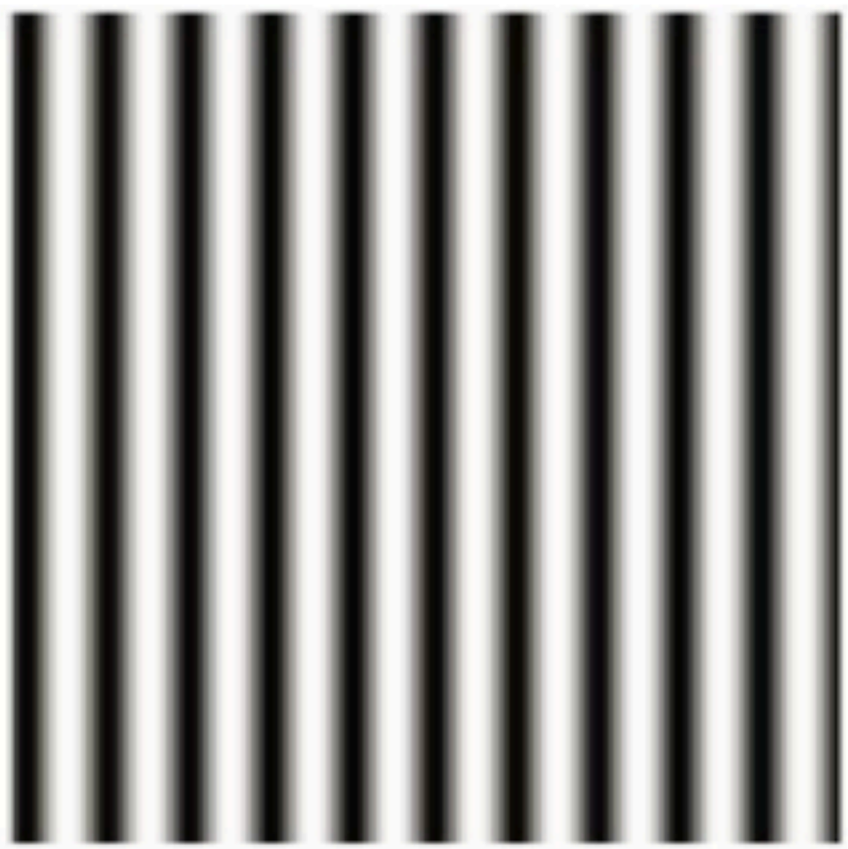


From M. Lewicky

Source: Aude Oliva



Source: Aude Oliva



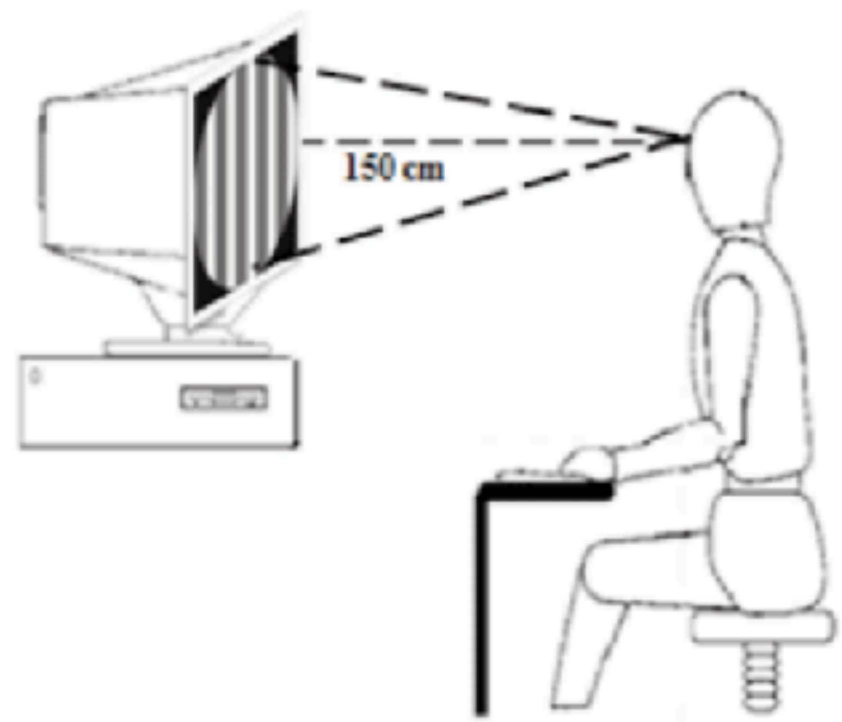
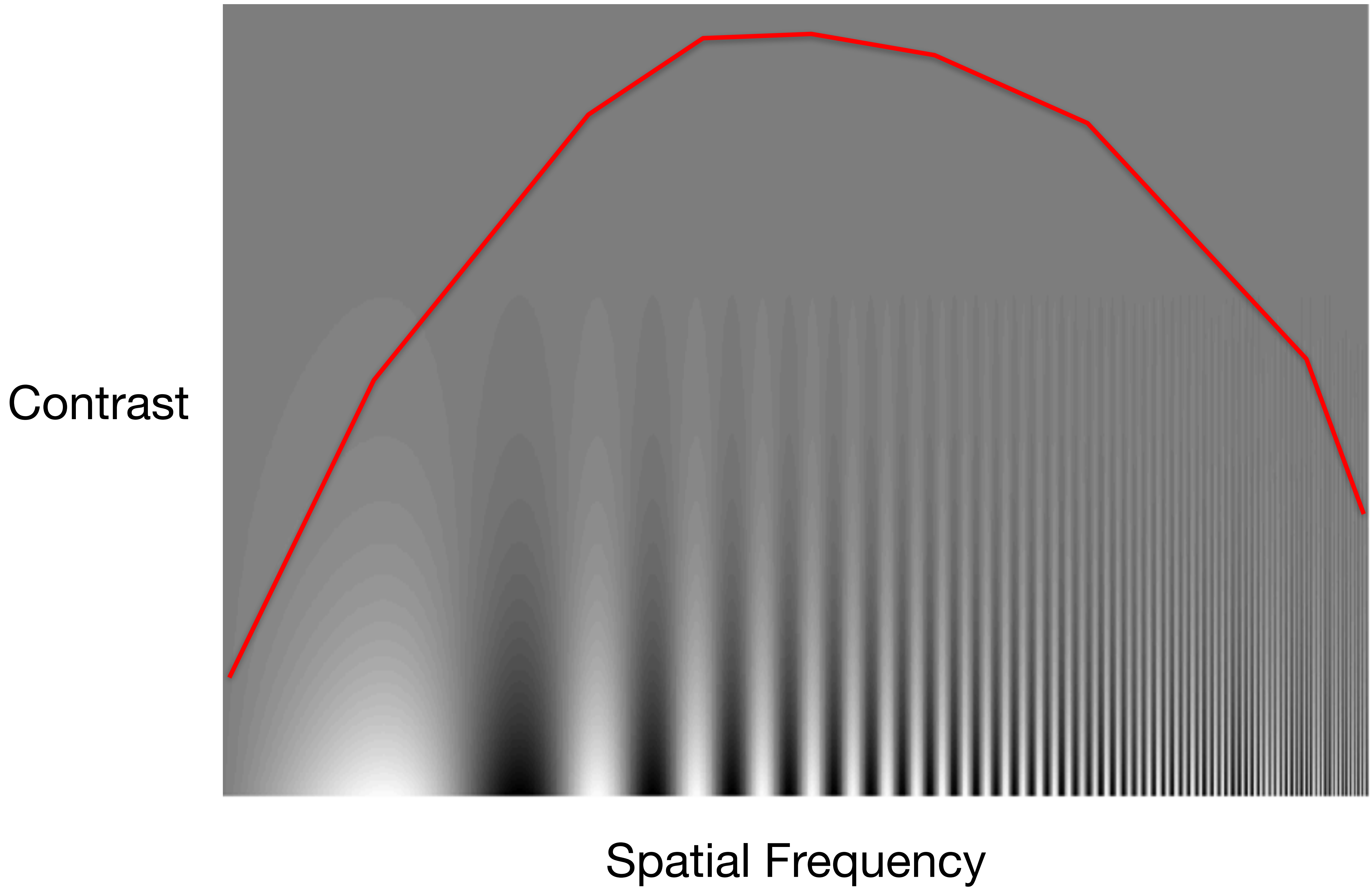


Figure 1. Stimulus presentation scheme. The stimuli were originally calibrated to be seen at a distance of 150 cm in a 19" display.

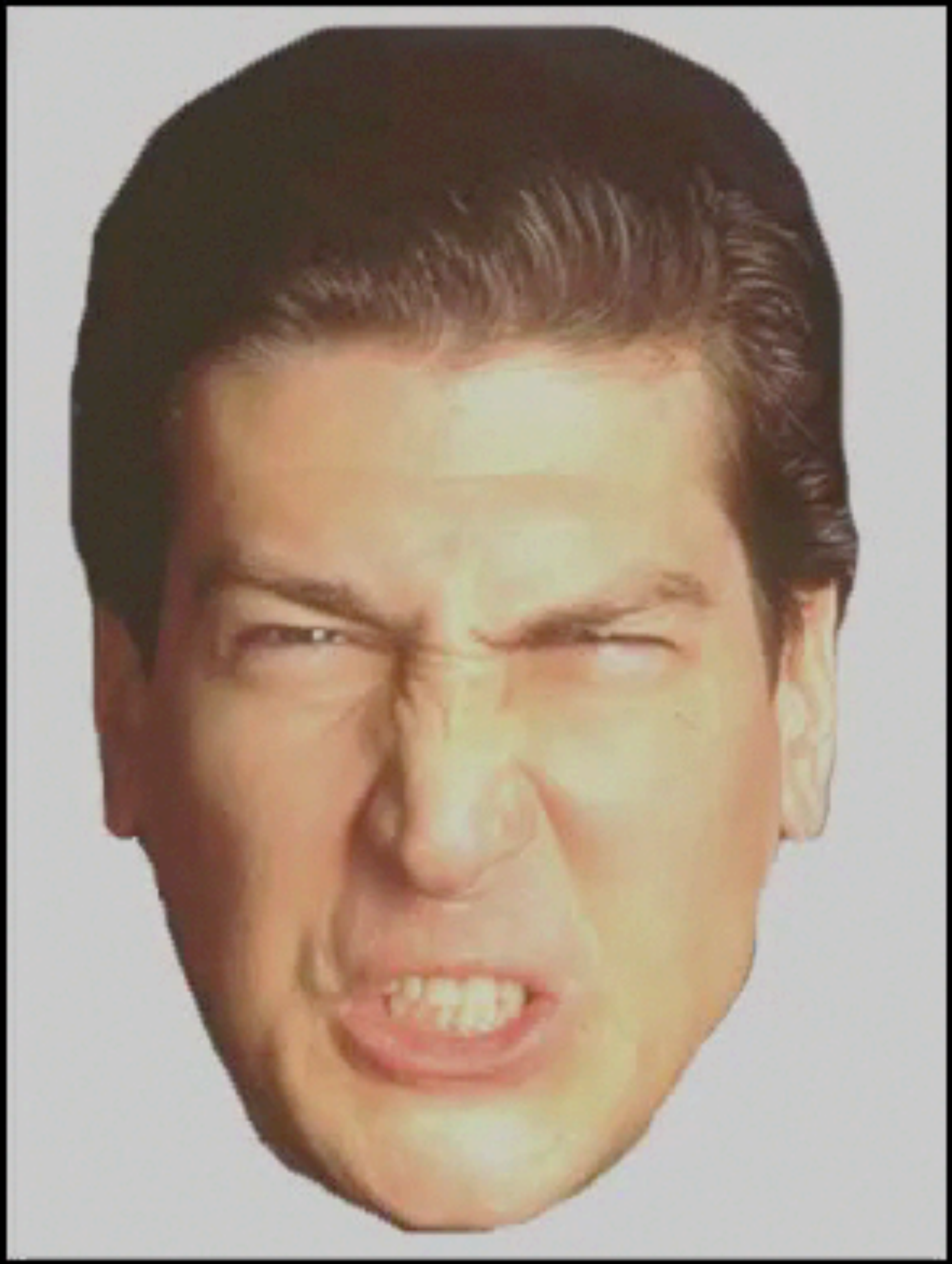
Source: Aude Oliva

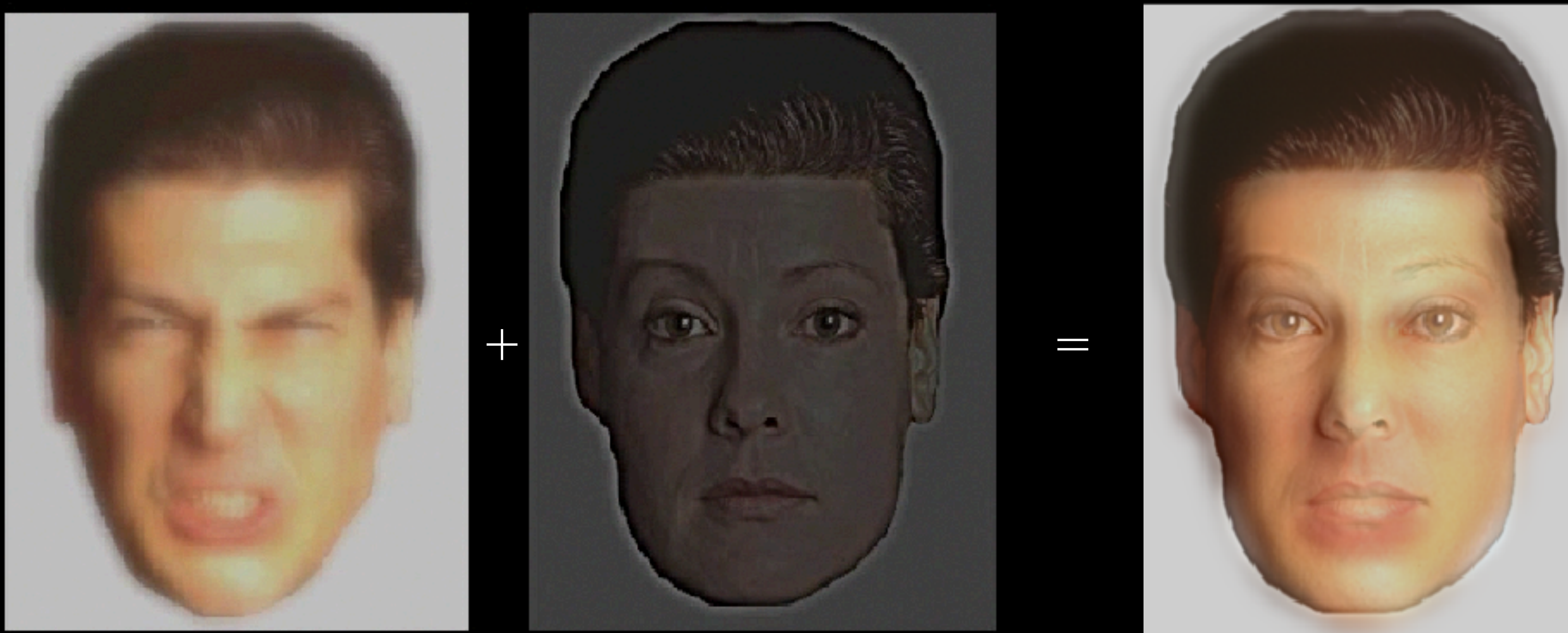


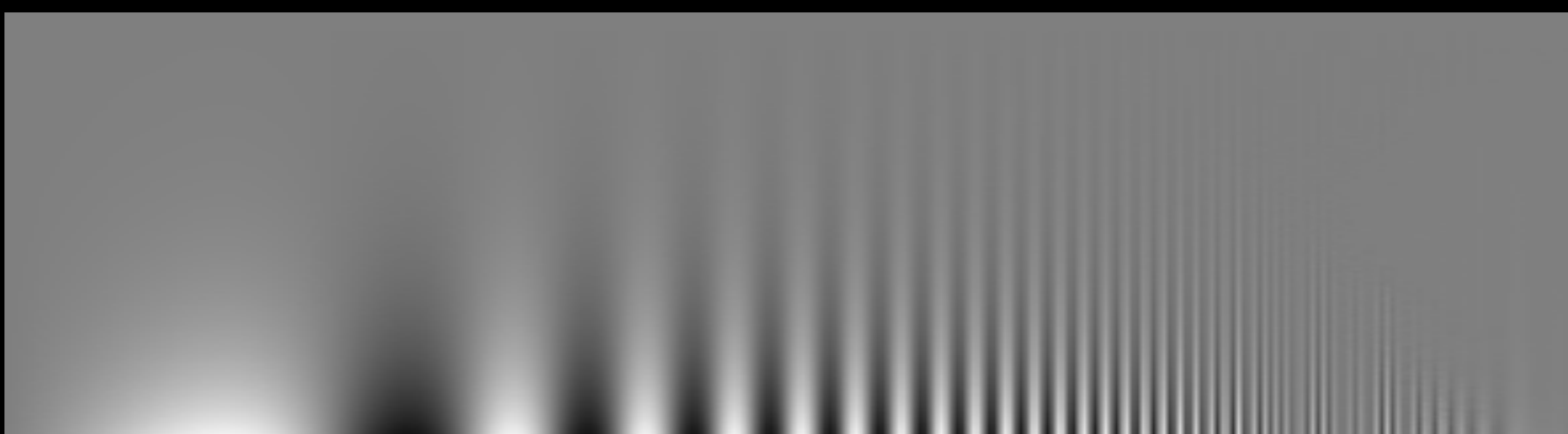
Source: Aude Oliva

Hybrid Images

Oliva & Schyns

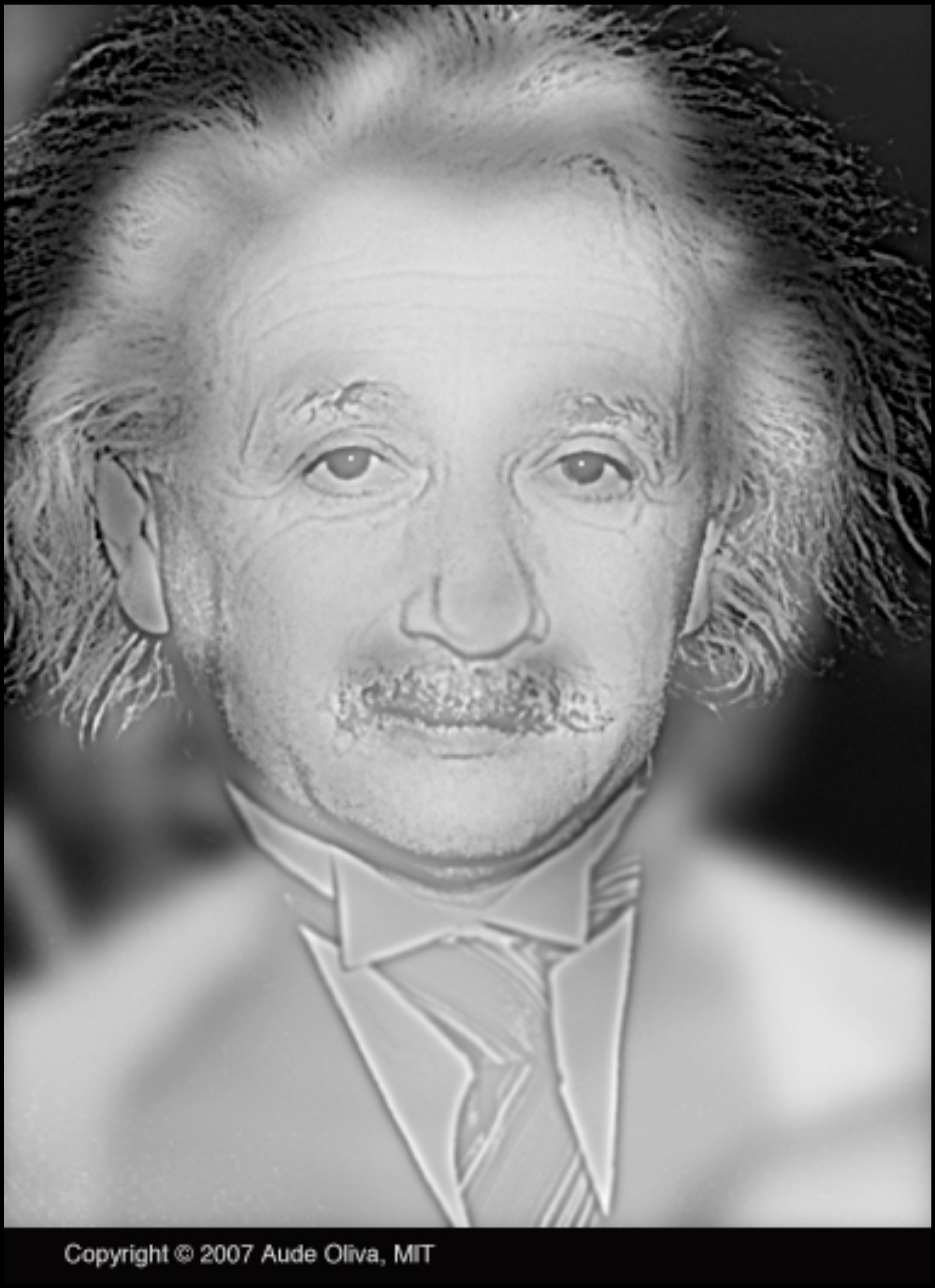




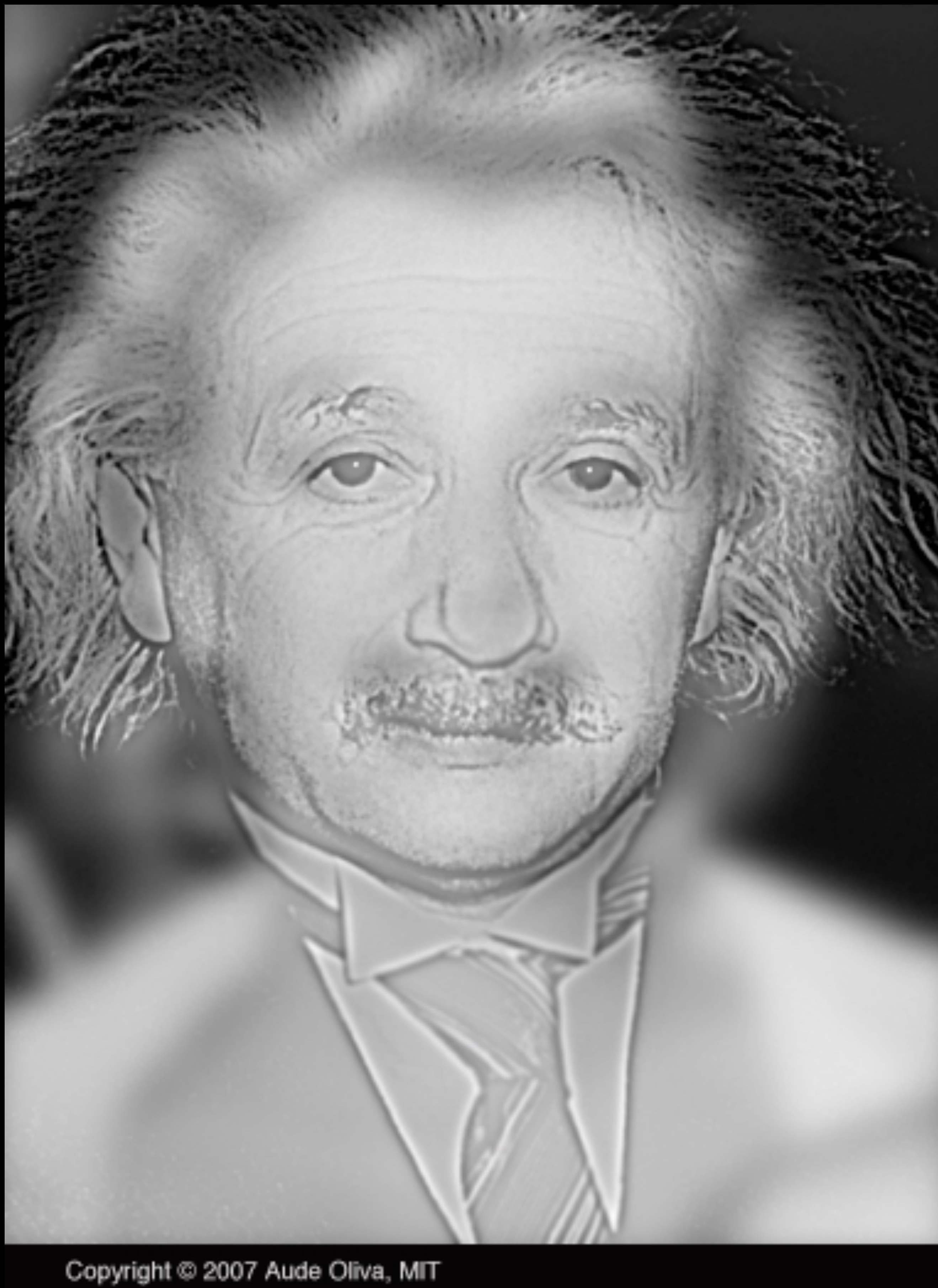


Hybrid Images

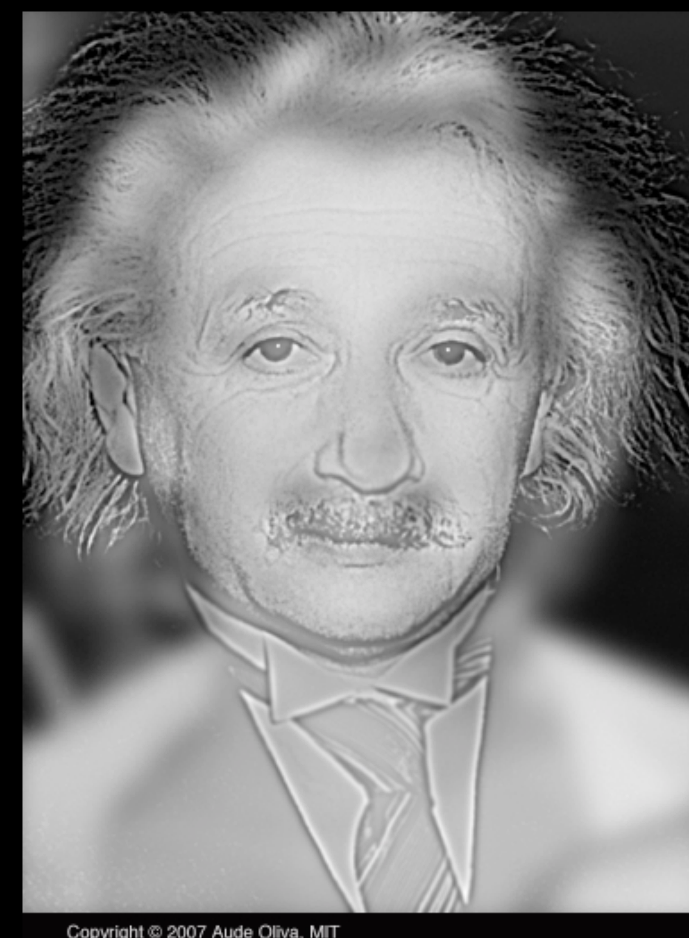
Hybrid Images

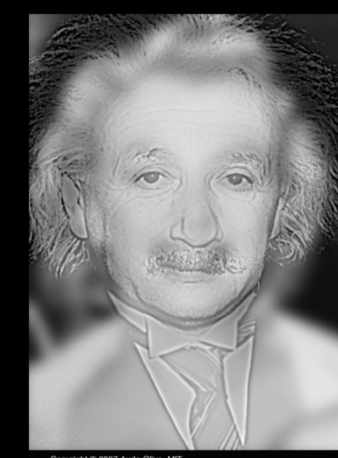


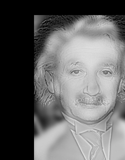
Source: Aude Oliva



http://cvcl.mit.edu/
hybrid_gallery/gallery.html







Source: Aude Oliva

Next Class:
Image Formation