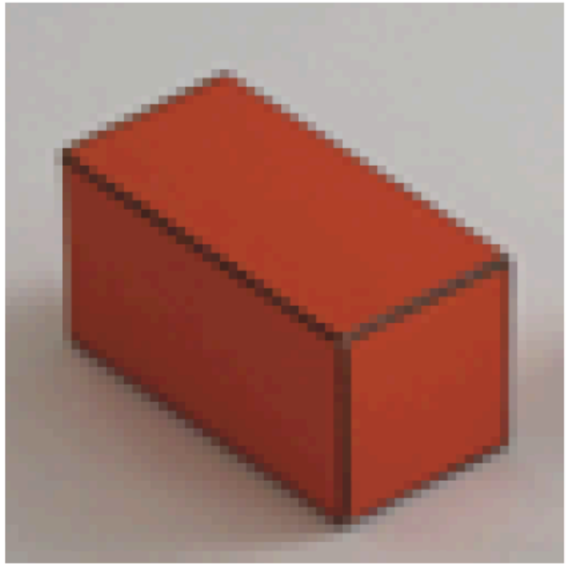
CSC496: Deep learning in computer vision
Prof. Bei Xiao

Lecture 6: Signal processing part 2: Fourier transform

For example, what is the kernel? Will check your answer next time.



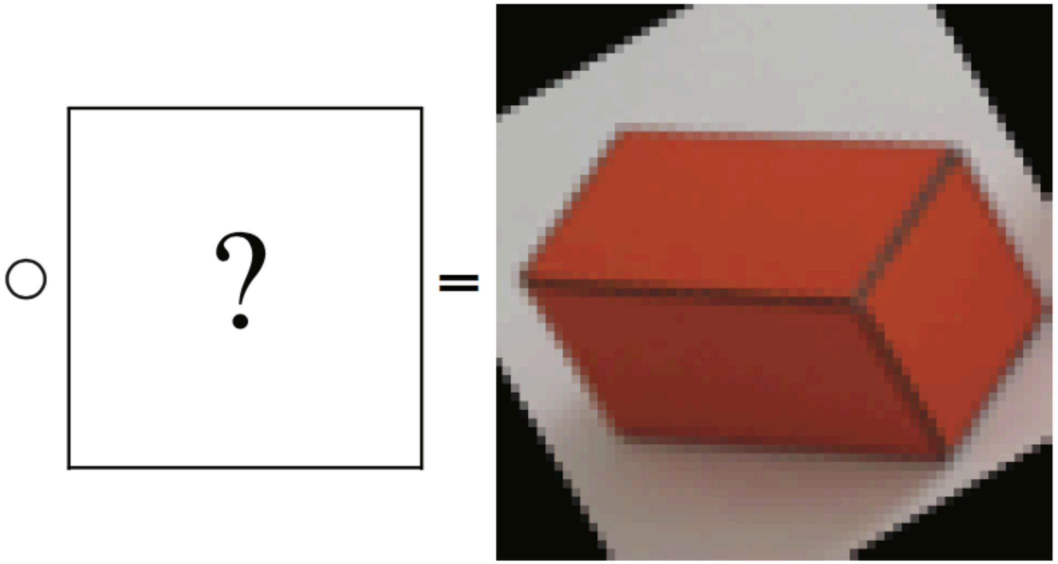
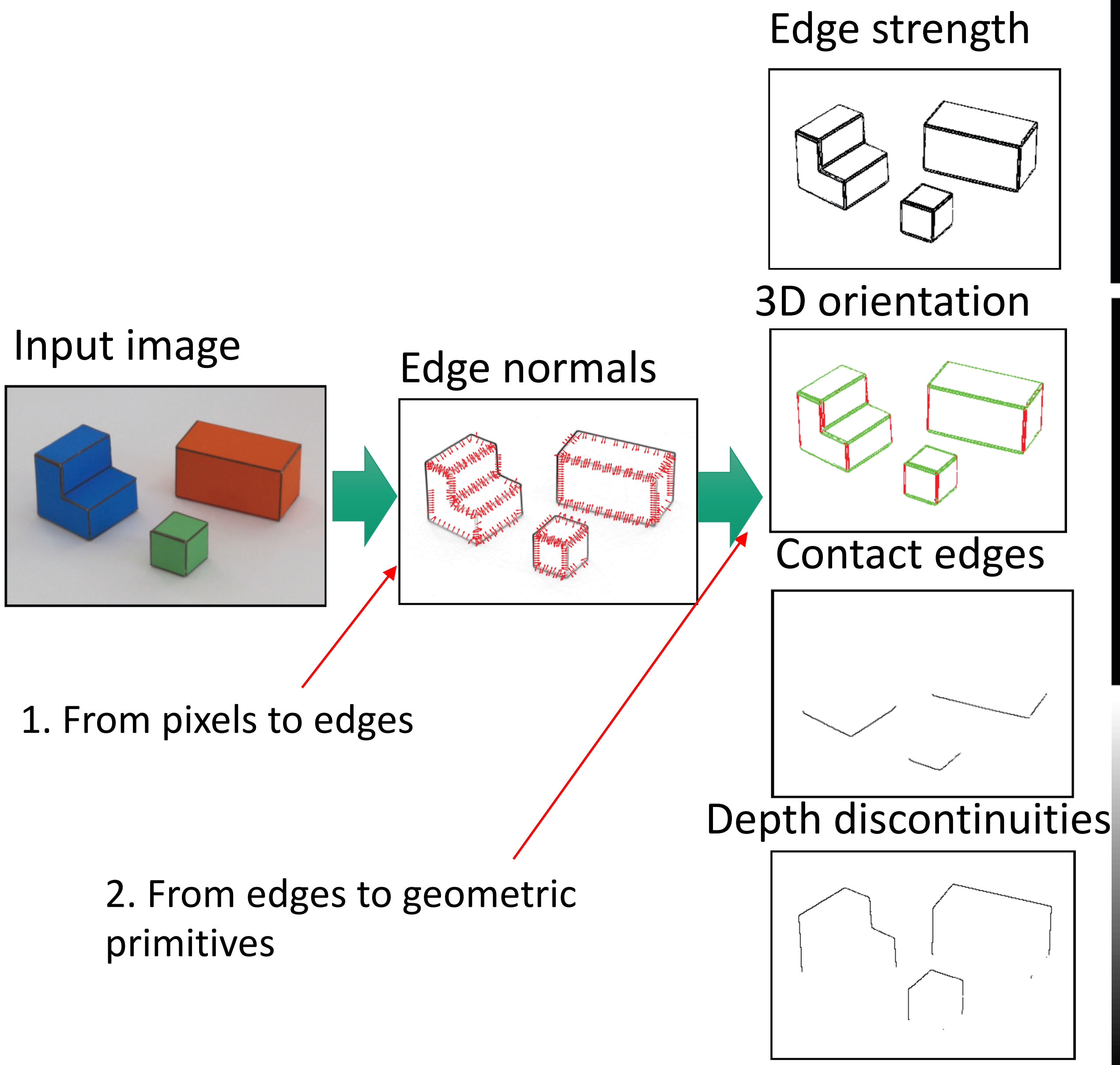
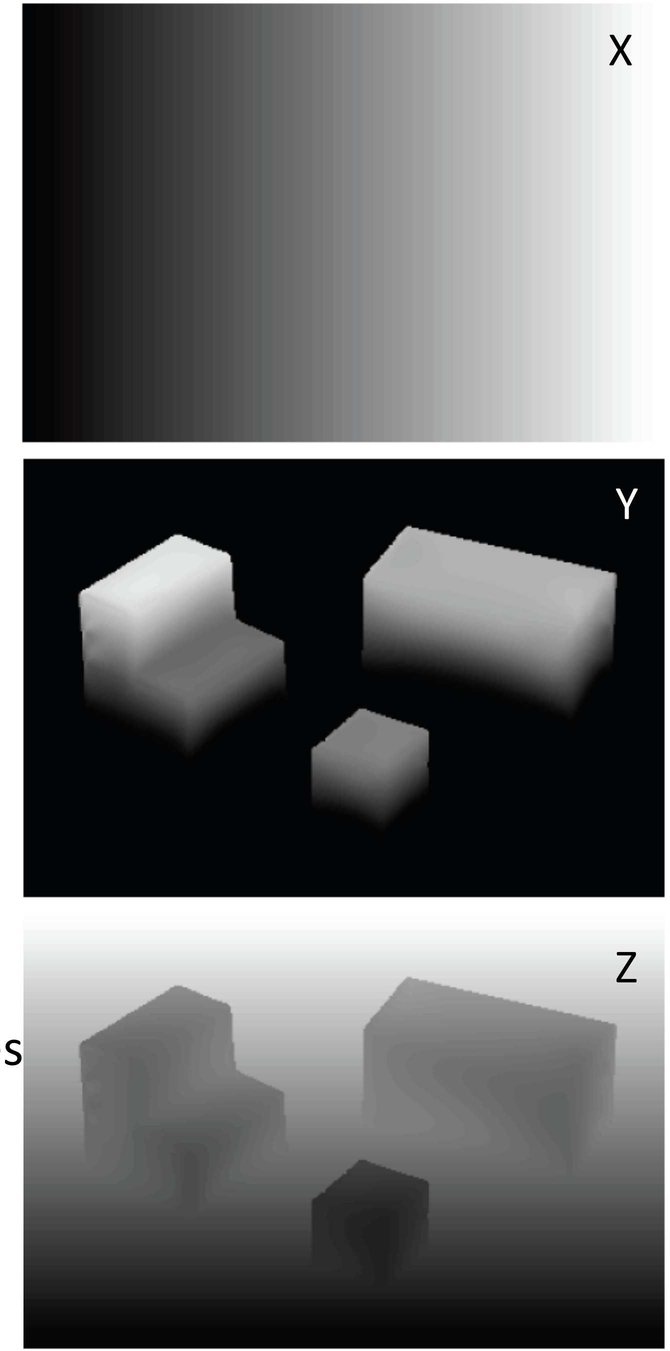


Image transformations





Linear image transformations

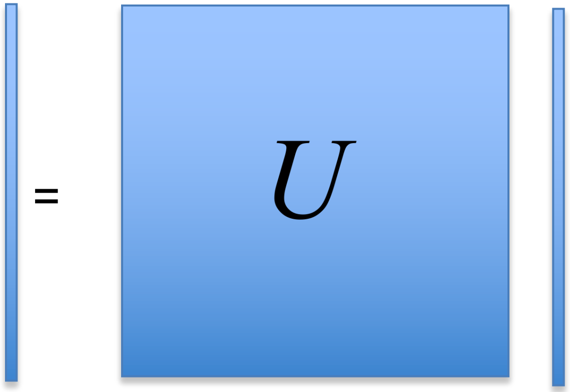
- In analyzing images, it's often useful to make a change of basis.

Transformed image

$$\vec{F} = U\vec{f}$$

Vectorized image

Fourier transform, or
Wavelet transform, or
Steerable pyramid transform



Self-inverting transforms

$$\vec{F} = U\vec{f} \longleftrightarrow \vec{f} = U^{-1}\vec{F}$$

Same basis functions are used for the inverse transform

$$\vec{f} = U^{-1}\vec{F}$$

$$= U^{+}\vec{F}$$

U transpose and complex conjugate

Important signals

Cosine and sine waves

$$s(t) = A\sin(w\,t - \theta)$$

$$s_k[n] = \sin\left(\frac{2\pi}{N}kn\right) \qquad c_k[n] = \cos\left(\frac{2\pi}{N}kn\right)$$

$k \in [1, N/2]$ denotes the number of wave cycles that will occur within the region of support

Important signals

Complex exponential

Eular's formula

$$e^{ix} = \cos x + i \sin x$$

$$s(t) = A \exp(j w t)$$

In discrete time (setting *A* = 1), we can write:

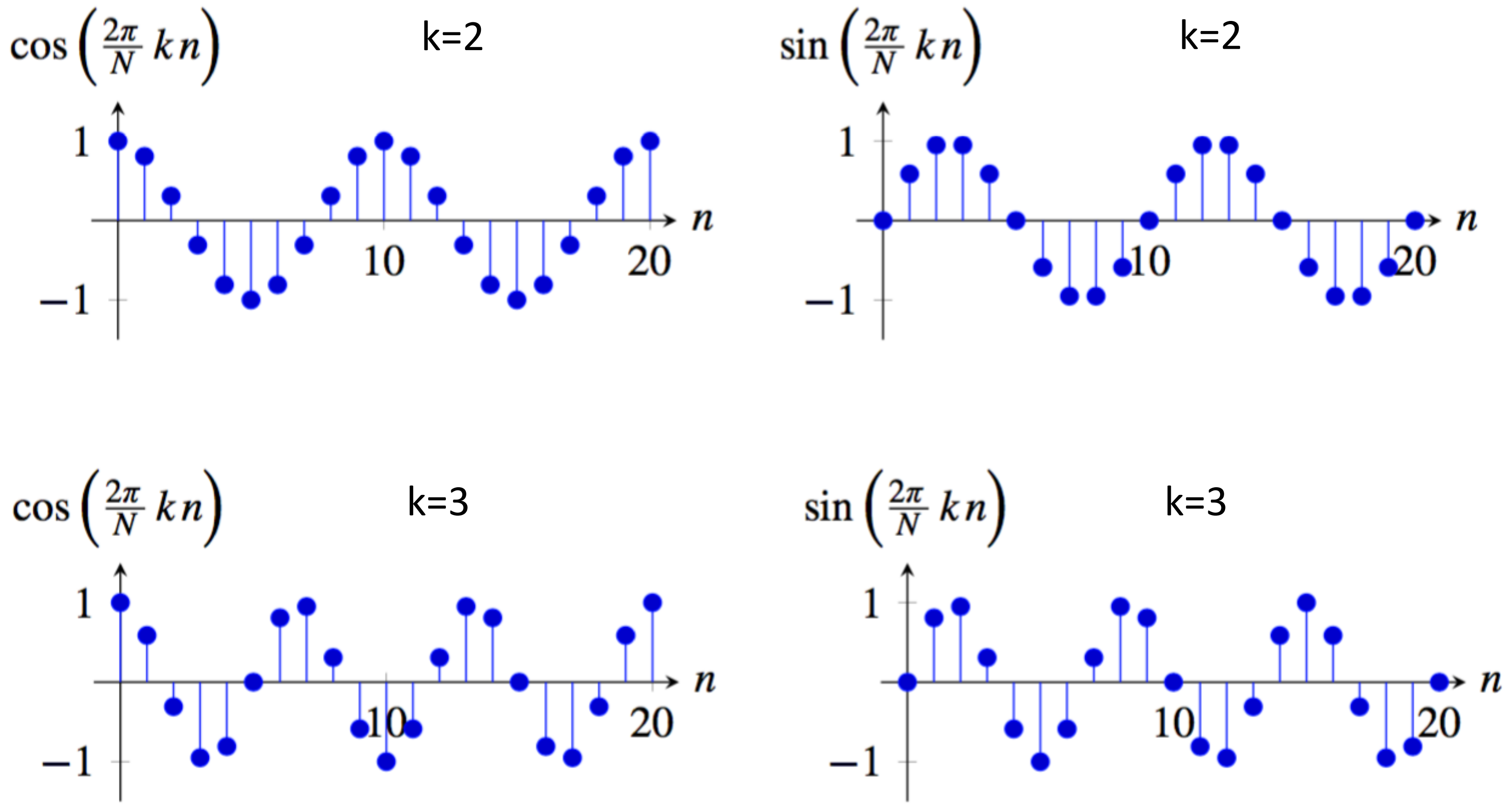
$$e_k[n] = \exp\left(j \frac{2\pi}{N} k n\right) = \cos\left(\frac{2\pi}{N} k n\right) + j \sin\left(\frac{2\pi}{N} k n\right)$$

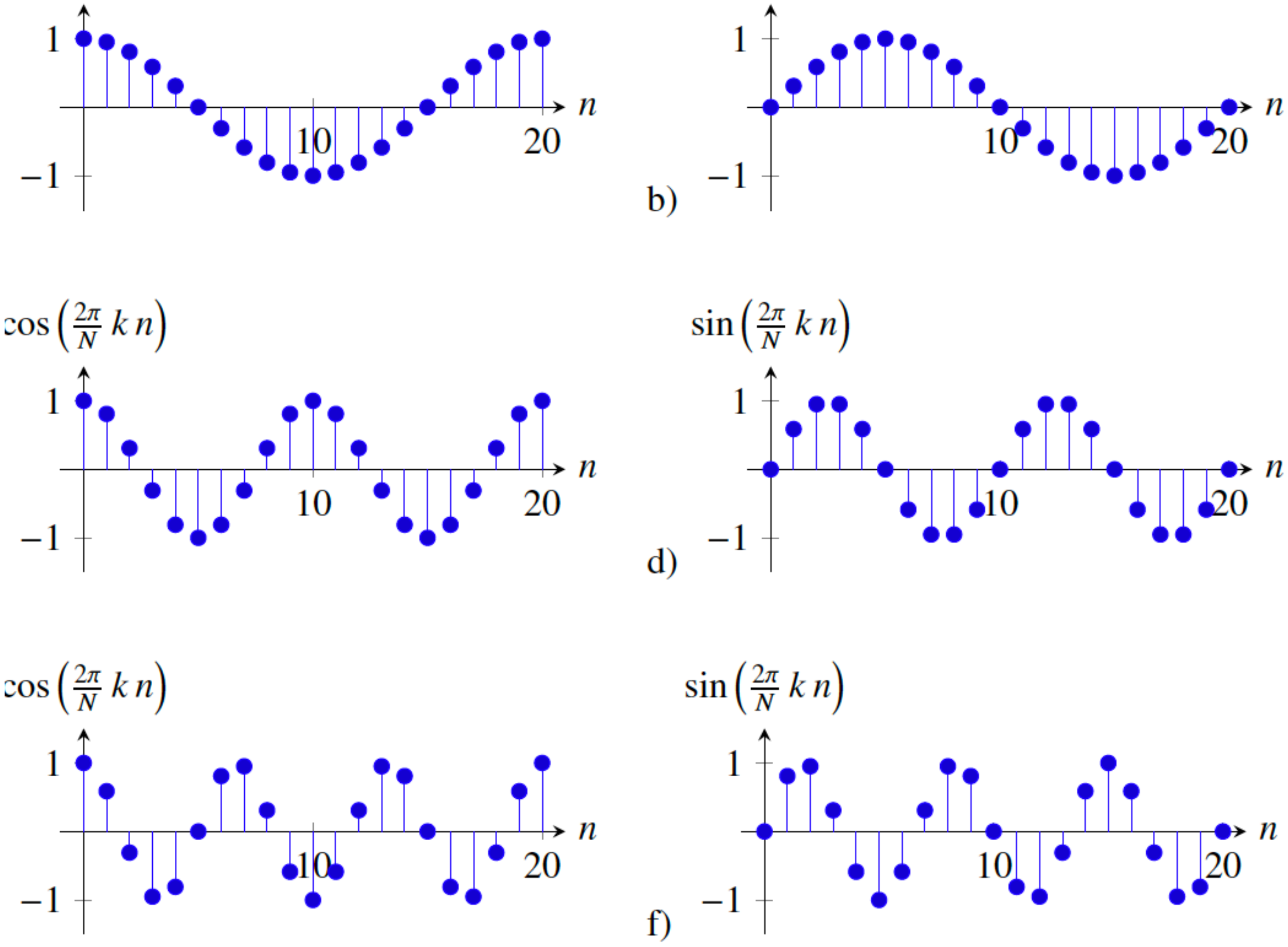
And in 2D, the complex exponential wave is:

$$e_{u,v}[n,m] = \exp\left(2\pi j \left(\frac{u n}{N} + \frac{v m}{M}\right)\right)$$

Important signals

Cosine and sine waves



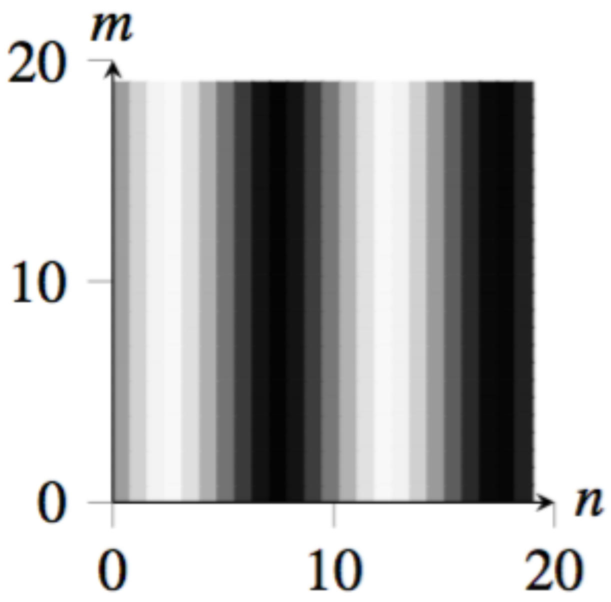


re 1.14

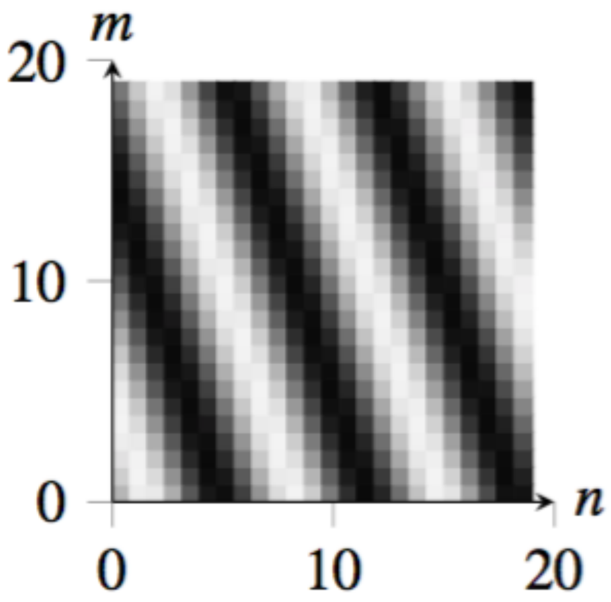
and cosine waves with $A = 1$ and $N = 20$. Each row corresponds to $k = 1$, $k = 2$ and $k = 3$. that for $k = 3$ the waves oscillates 3 times in the interval $[0, N - 1]$ but the samples in each lation are not identical and it is only truly periodic once every N samples. This is because $3/20$ irreducible fraction.

Waves in 2D

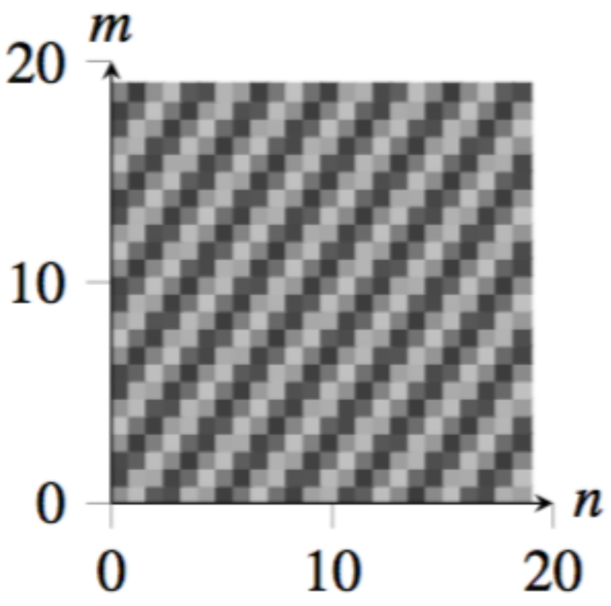
$$s_{u,v}[n,m] = A \sin\left(2\pi\left(\frac{un}{N} + \frac{vm}{M}\right)\right) \qquad c_{u,v}[n,m] = A \cos\left(2\pi\left(\frac{un}{N} + \frac{vm}{M}\right)\right)$$



u = 2, v = 0



u = 3, v = 1



u = 7, v = −5

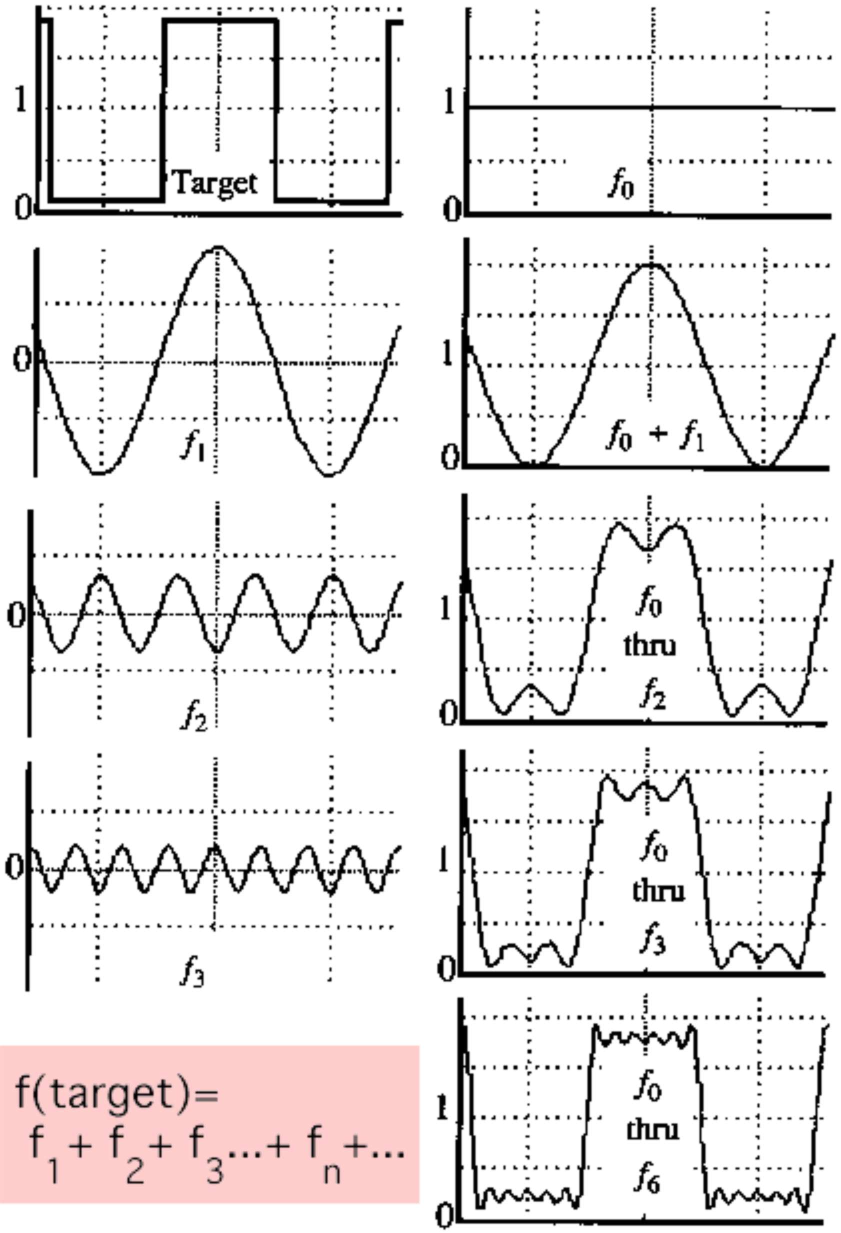
Impulses, sine and cosine waves or complex exponentials form each an orthogonal basis for signals of length N

A sum of sines

Our building block:

$$A\sin(\omega x+\phi)$$

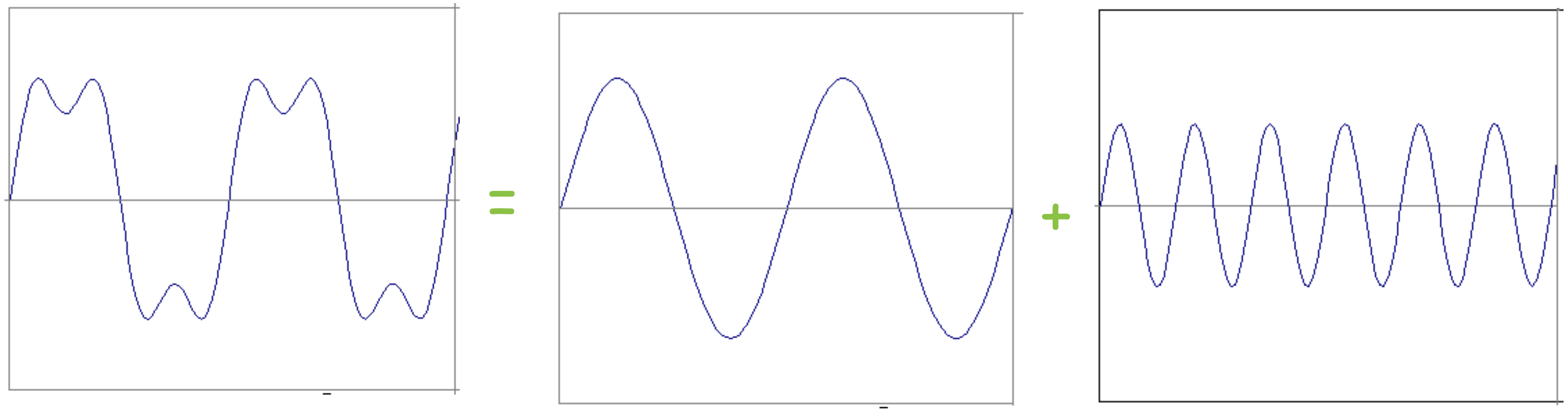
Add enough of them to get any signal *g(x)* you want!



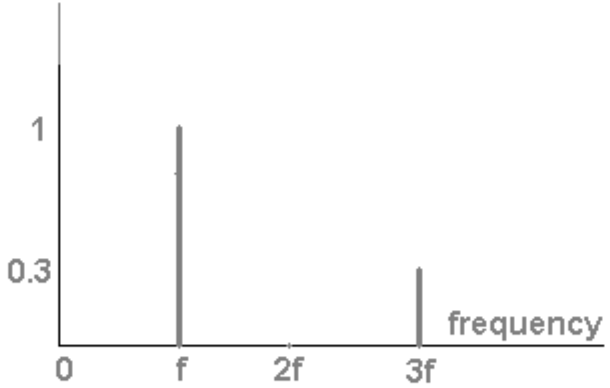
f(target)=
$f_1 + f_2 + f_3 ... + f_n + ...$

Frequency Spectra

- example : $g(t) = \sin(2\pi f\, t) + (1/3)\sin(2\pi(3f)\, t)$



amplitude



Slides: Efros

Why decompose signals into sine waves?

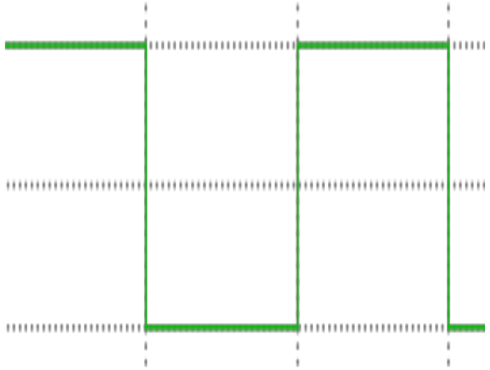
- Why don’t we represent the system with some other periodic signals?
- Sine wave is the only waveform that doesn’t change shape when subject to a linear-time-invariant (LTI) system. Input sinusoids, output will be sinusoids.
- Convolution, image domain filtering, is LTI.

We need a more precise language to talk about the effect of linear filters.

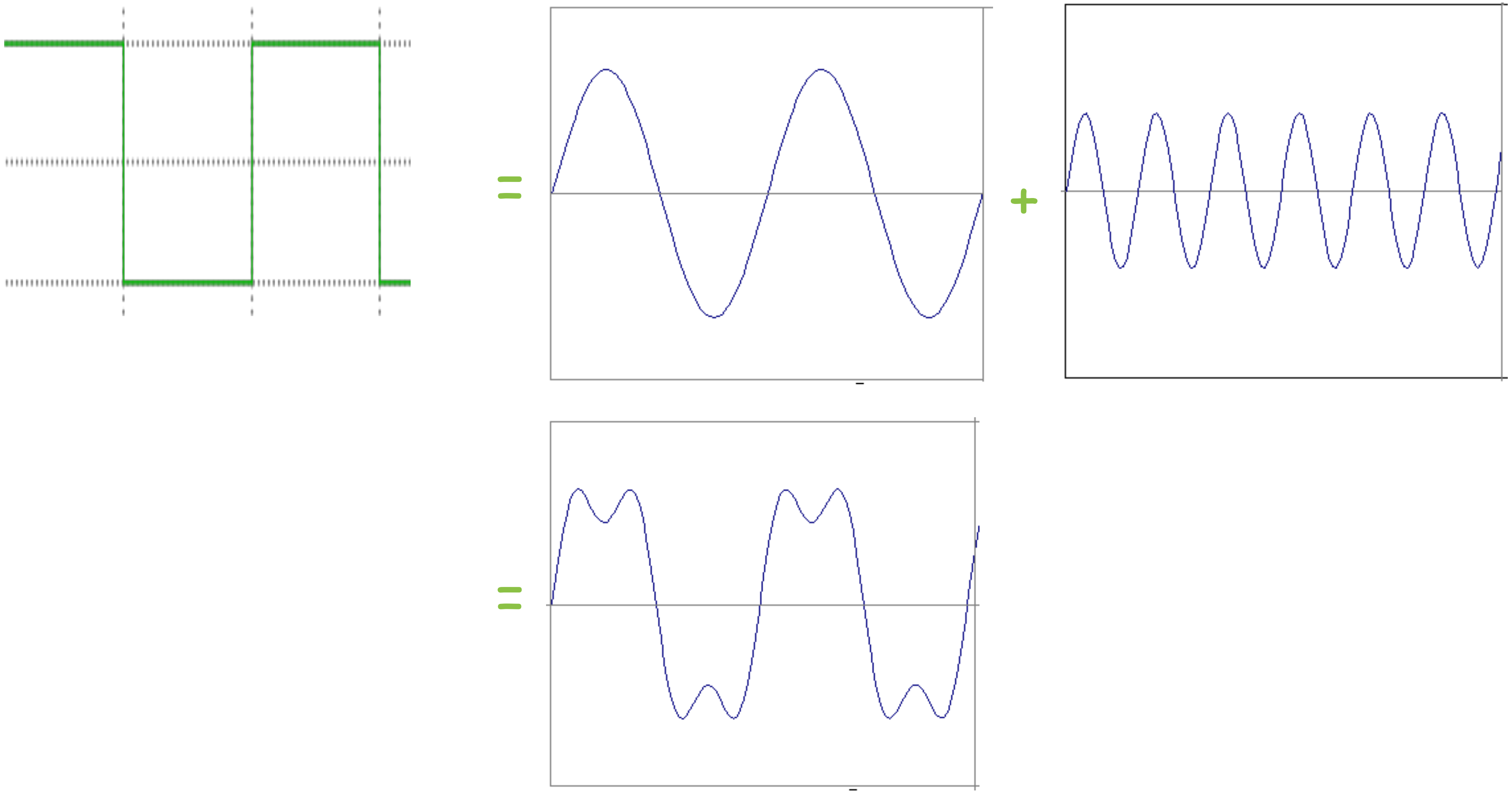
Fourier Transform provides that precision.

Frequency Spectra

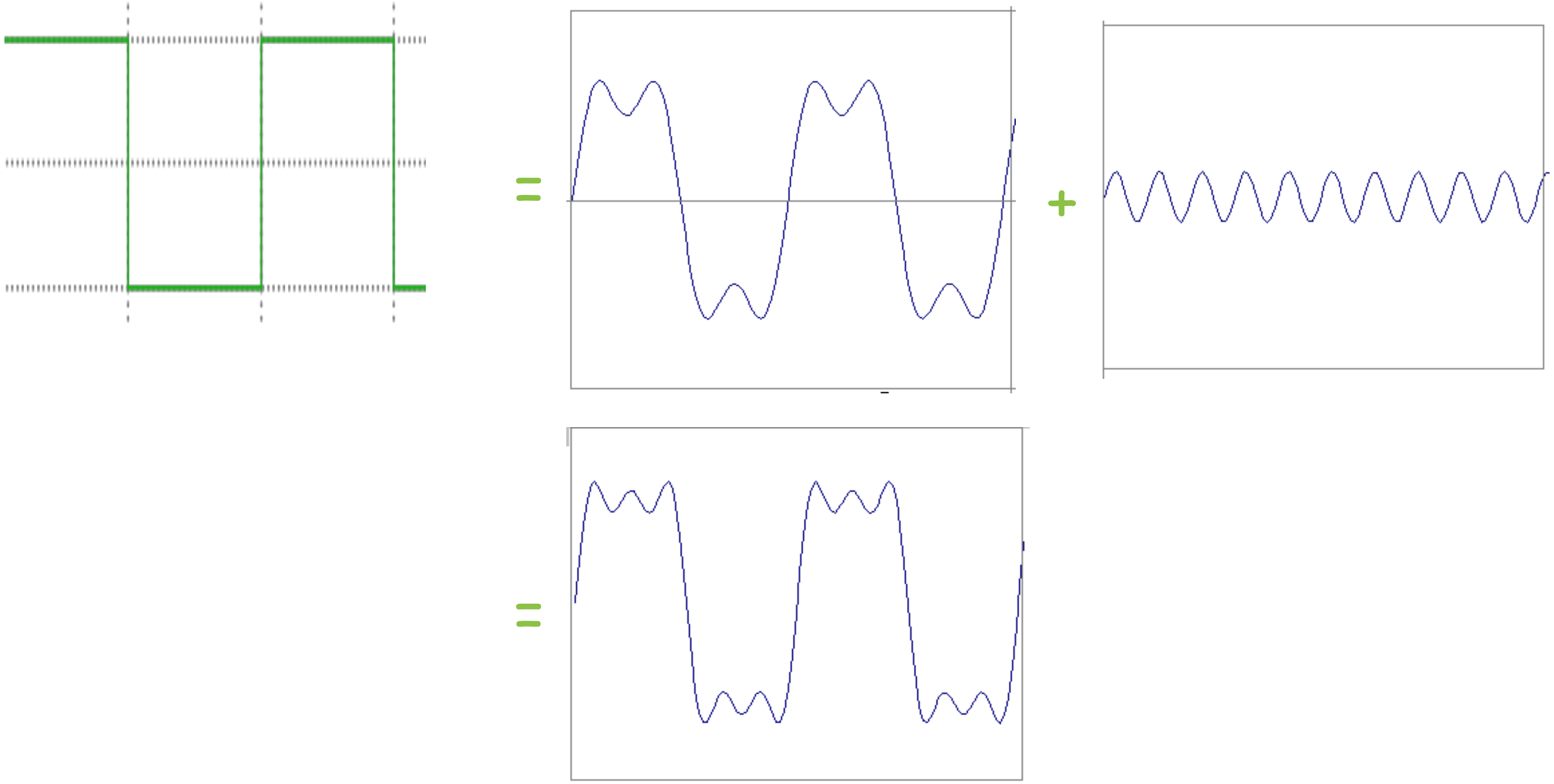
- Consider a square wave f(x)



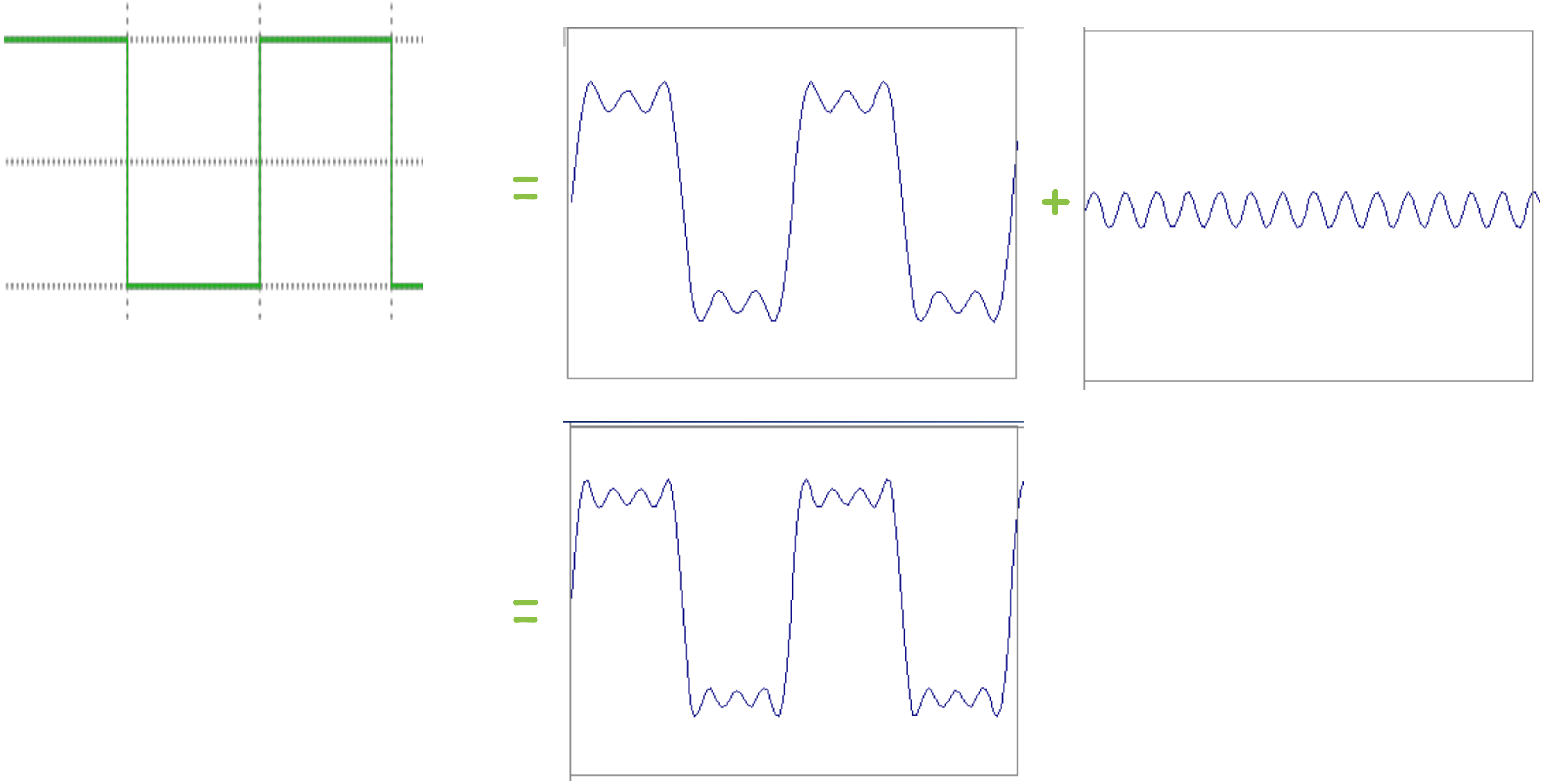
Frequency Spectra



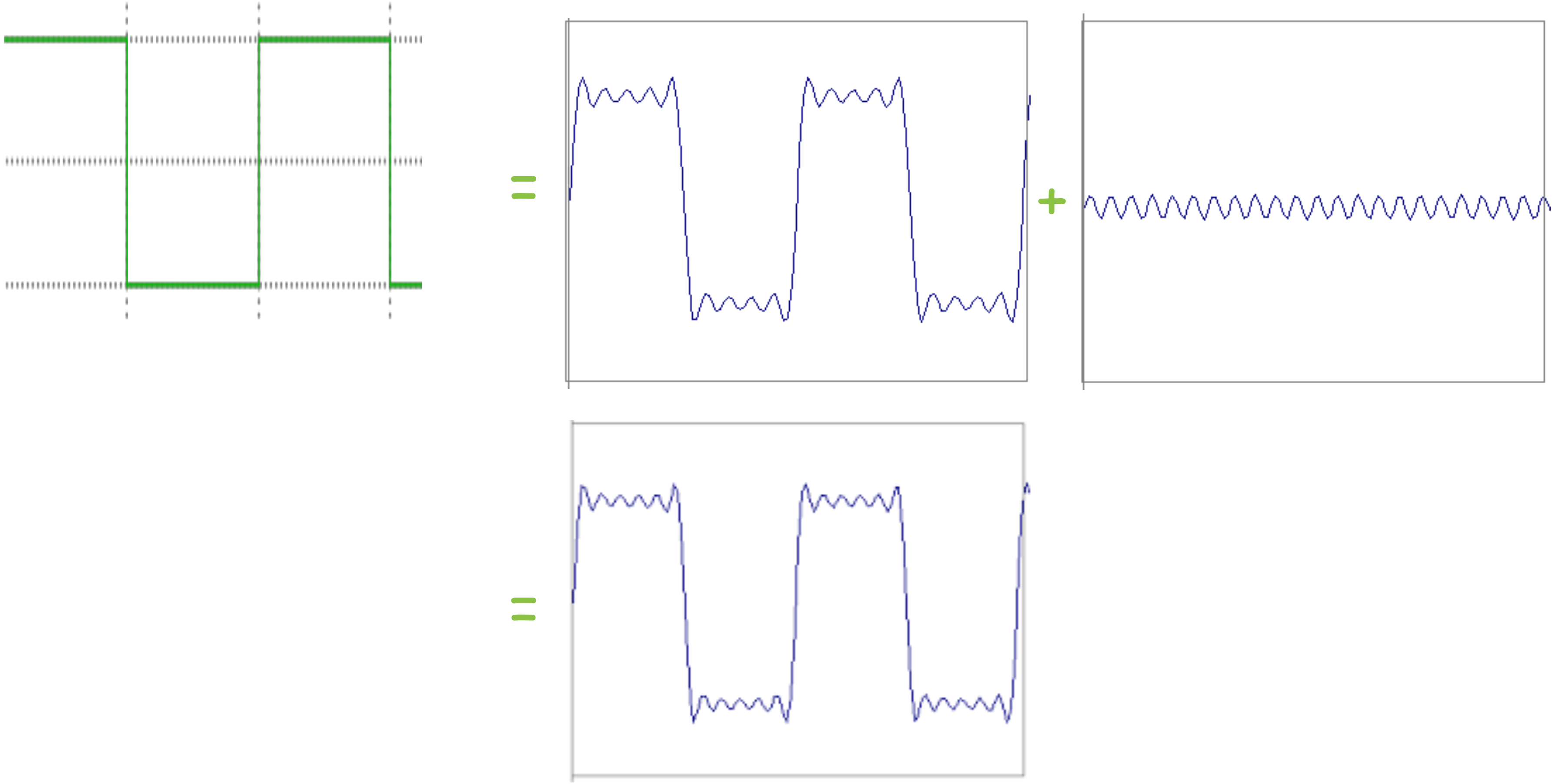
Frequency Spectra



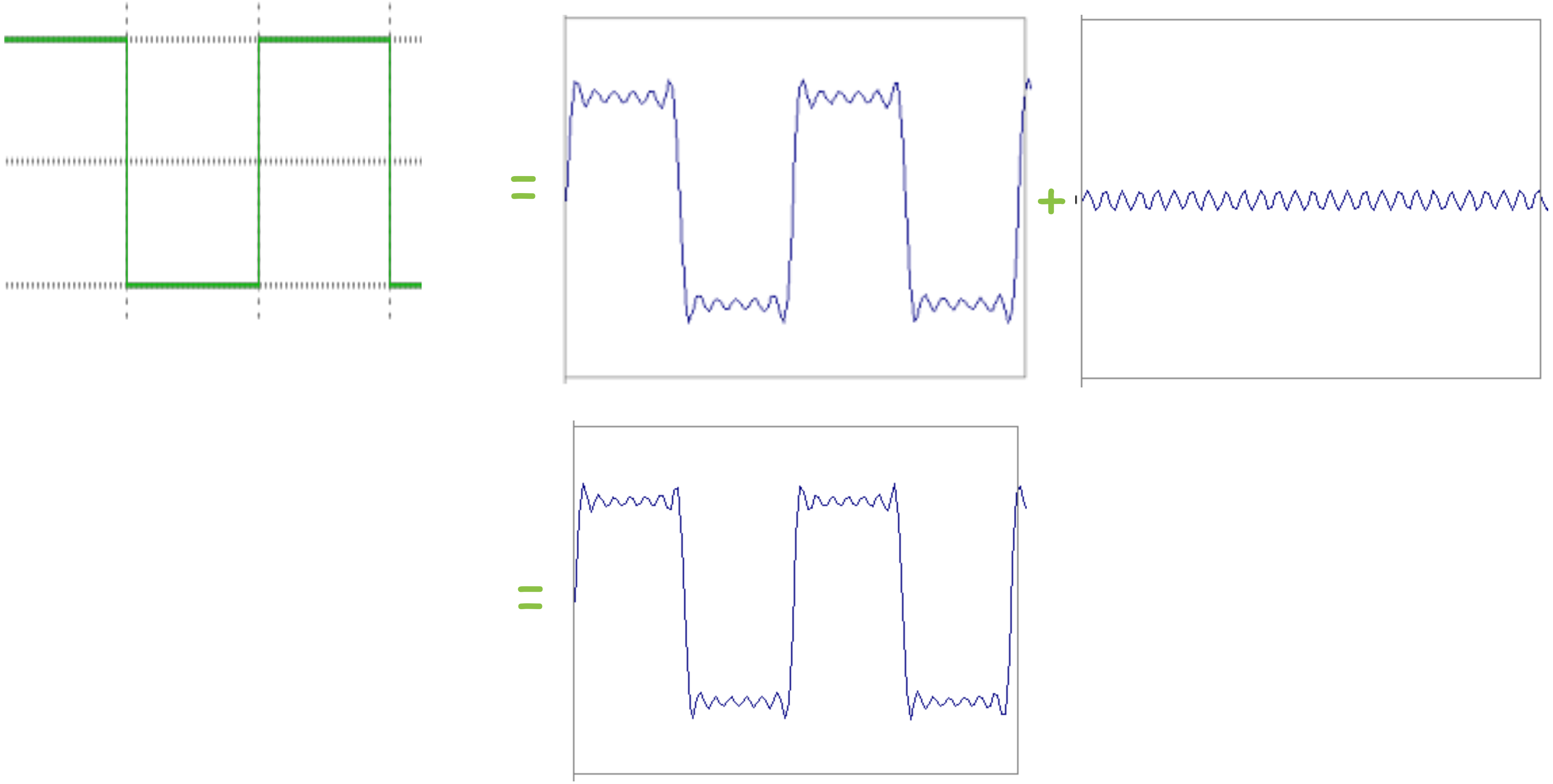
Frequency Spectra



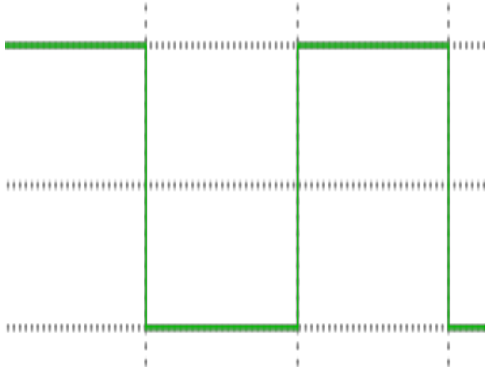
Frequency Spectra



Frequency Spectra

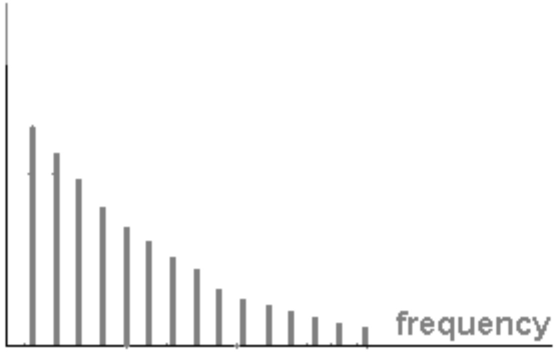


Frequency Spectra



=

$$\boxed{A\sum_{k=1}^{\infty}\frac{1}{k}\sin(2\pi kt)}$$



Fourier transform

Fourier Transform

- Fourier Transform

Inverse Fourier transform

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \qquad f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega$$

Time f(t) $\longleftrightarrow$ Frequency f(w)

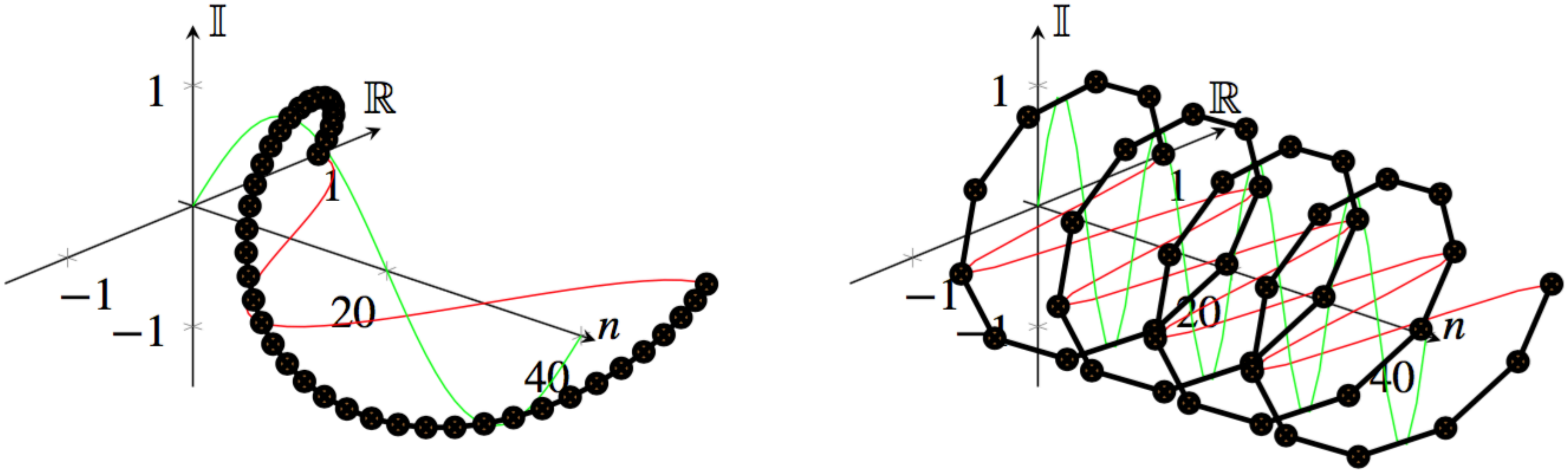
What is transformation? It is a mapping between two domains

Complex exponential

$$F[u,v] = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} f[n,m] \exp\left(-2\pi j \left(\frac{un}{N} + \frac{vm}{M}\right)\right)$$

Remember: exp (α j) = cos(α) + j sin(α)

Visualization complex exponential for two different frequencies *u*



N = 40, *u* = 1 *N* = 40, *u* = 3

Impulses, sine and cosine waves or complex exponentials form each an orthogonal basis for signals of length N

The Discrete Fourier transform

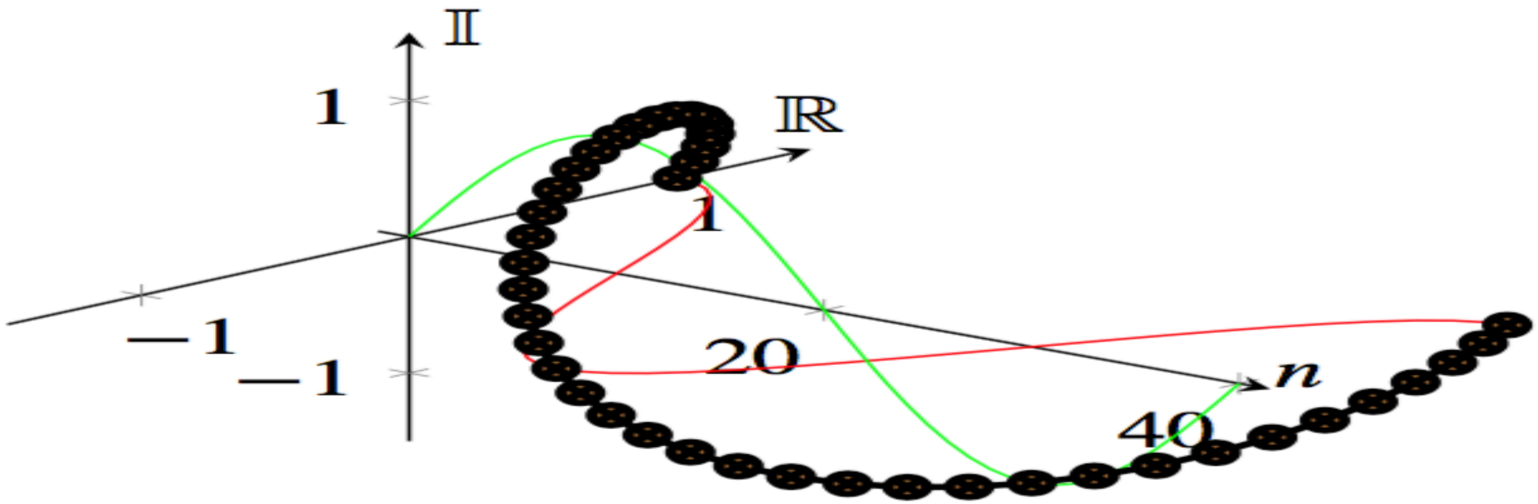
Discrete Fourier Transform (DFT) transforms an image
f [*m*, *m*] into the complex image Fourier transform *F* [*u*, *v*] as:

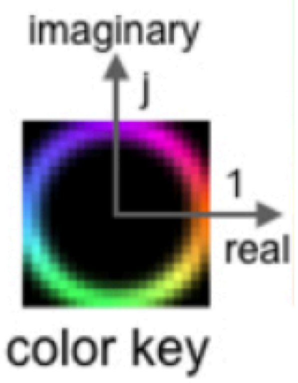
$$F[u,v] = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} f[n,m] \exp\left(-2\pi j \left(\frac{un}{N} + \frac{vm}{M}\right)\right)$$

The inverse Fourier transform is:

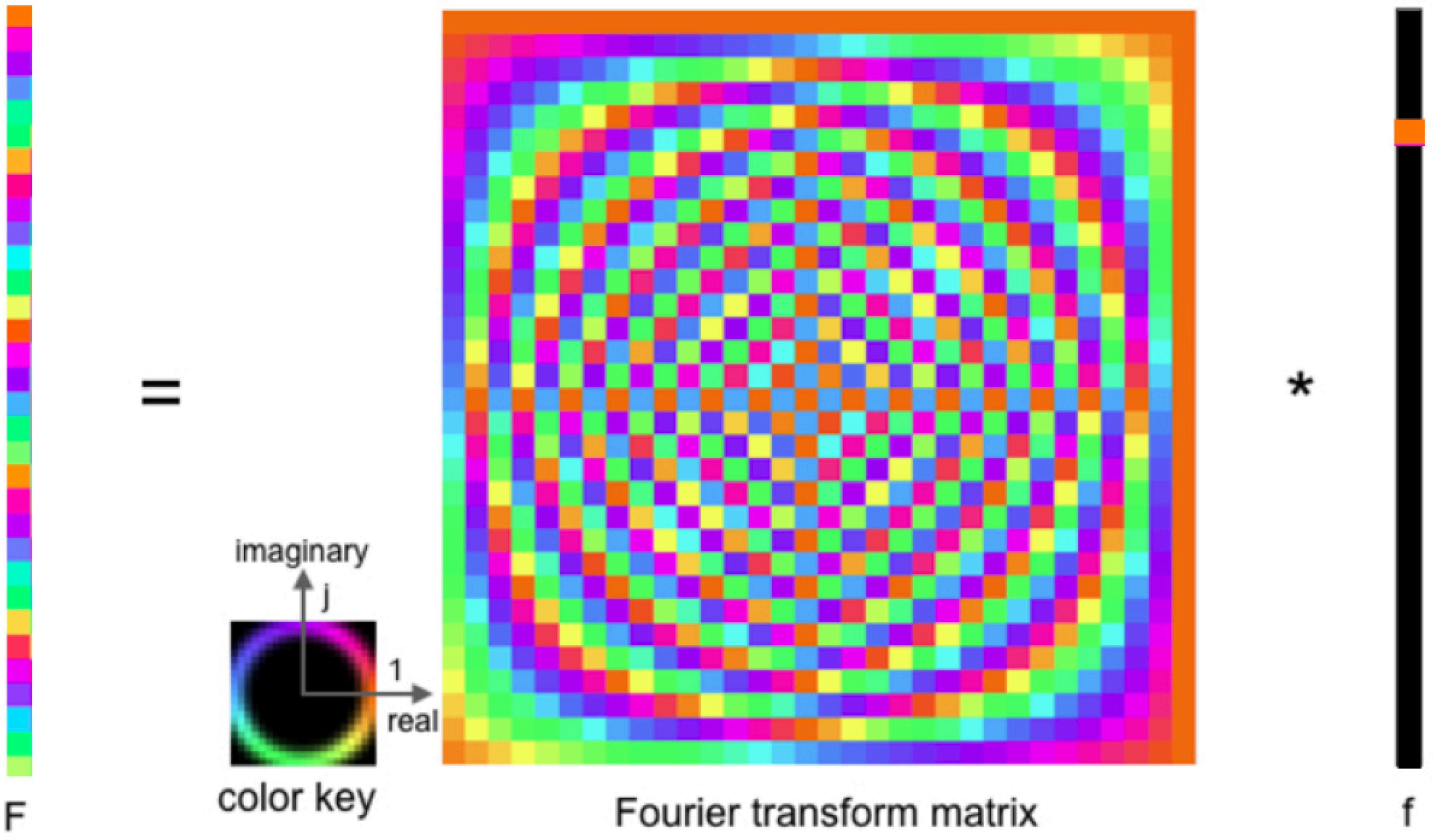
$$f[n,m] = \frac{1}{NM} \sum_{u=0}^{N-1} \sum_{v=0}^{M-1} F[u,v] \exp\left(+2\pi j \left(\frac{un}{N} + \frac{vm}{M}\right)\right)$$

Discrete Fourier transform visualization





Fourier transform visualization



Some useful transforms

Fourier transform of the Delta function δ[n ,m]**:**

$$F[u,v] = \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} \delta[n,m] \exp\left(-2\pi j\left(\frac{un}{N}+\frac{vm}{M}\right)\right) = 1$$

Observation: if we think in terms of the inverse DFT, this means that:

$$\delta[n,m] = \frac{1}{NM}\sum_{u=-N/2}^{N/2}\sum_{v=-M/2}^{M/2} \exp\left(2\pi j\left(\frac{un}{N}+\frac{vm}{M}\right)\right)$$

Some useful transforms

The Fourier transform of the cosine wave

$$\cos\left(2\pi\left(\frac{u_0 n}{N}+\frac{v_0 m}{M}\right)\right)$$

is:

$$F[u,v]=\sum_{n=0}^{N-1}\sum_{m=0}^{M-1}\cos\left(2\pi\left(\frac{u_0 n}{N}+\frac{v_0 m}{M}\right)\right)\exp\left(-2\pi j\left(\frac{un}{N}+\frac{vm}{M}\right)\right)=$$
$$=\frac{1}{2}\left(\delta[u-u_0,v-v_0]+\delta[u+u_0,v+v_0]\right)$$

Same for the sine wave:

$$\sin\left(2\pi\left(\frac{u_0 n}{N}+\frac{v_0 m}{M}\right)\right)\leftrightarrow F[u,v]=\frac{1}{2j}\left(\delta[u-u_0,v-v_0]-\delta[u+u_0,v+v_0]\right)$$

Szeliski, Computer Vision, 2010

Name	Signal		Transform
impulse	$\delta(x)$	$\Leftrightarrow$	1
shifted impulse	$\delta(x-u)$	$\Leftrightarrow$	$e^{-j\omega u}$
box filter	$\mathrm{box}(x/a)$	$\Leftrightarrow$	$a\mathrm{sinc}(a\omega)$
tent	$\mathrm{tent}(x/a)$	$\Leftrightarrow$	$a\mathrm{sinc}^2(a\omega)$
Gaussian	$G(x;\sigma)$	$\Leftrightarrow$	$\frac{\sqrt{2\pi}}{\sigma}G(\omega;\sigma^{-1})$
Laplacian of Gaussian	$(\frac{x^2}{\sigma^4}-\frac{1}{\sigma^2})G(x;\sigma)$	$\Leftrightarrow$	$-\frac{\sqrt{2\pi}}{\sigma}\omega^2 G(\omega;\sigma^{-1})$
Gabor	$\cos(\omega_0 x)G(x;\sigma)$	$\Leftrightarrow$	$\frac{\sqrt{2\pi}}{\sigma}G(\omega\pm\omega_0;\sigma^{-1})$
unsharp mask	$(1+\gamma)\delta(x) - \gamma G(x;\sigma)$	$\Leftrightarrow$	$(1+\gamma) - \frac{\sqrt{2\pi}\gamma}{\sigma}G(\omega;\sigma^{-1})$
windowed sinc	$\mathrm{rcos}(x/(aW))\ \mathrm{sinc}(x/a)$	$\Leftrightarrow$	(see Figure 3.29)

Table 3.2 Some useful (continuous) Fourier transform pairs: The dashed line in the Fourier transform of the shifted impulse indicates its (linear) phase. All other transforms have zero phase (they are real-valued). Note that the figures are not necessarily drawn to scale but are drawn to illustrate the general shape and characteristics of the filter or its response. In particular, the Laplacian of Gaussian is drawn inverted because it resembles more a "Mexican hat", as it is sometimes called.

2D Discrete Fourier Transform

$$F[u,v] = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} f[n,m] \exp\left(-2\pi j \left(\frac{un}{N} + \frac{vm}{M}\right)\right)$$

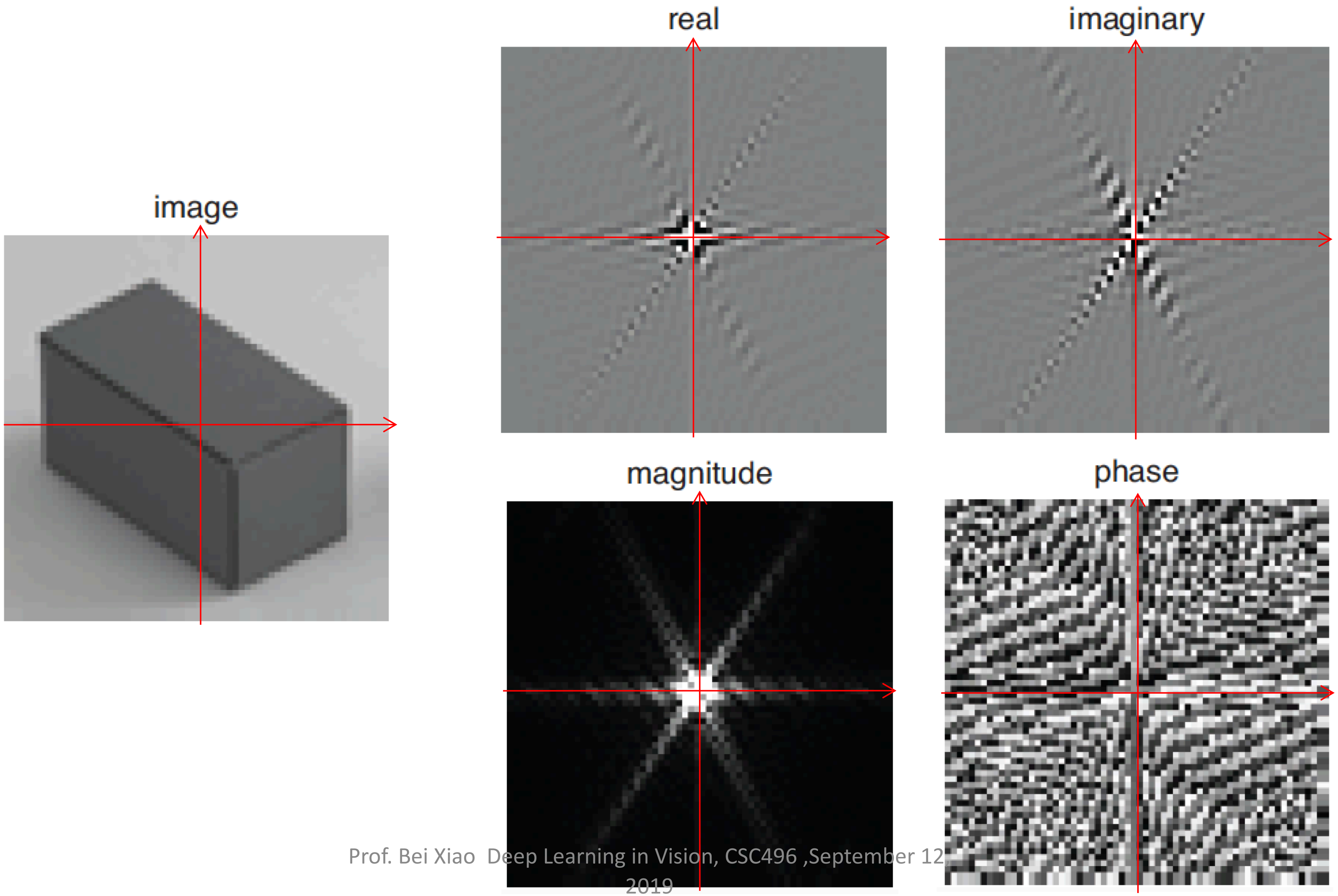
Using the real and imaginary components:

$$F[u,v] = Re\{F[u,v]\} + j\,Imag\{F[u,v]\}$$

Or using a polar decomposition:

$$F[u,v] = A[u,v] \exp(j\theta[u,v])$$

2D Discrete Fourier Transform



Properties for the DFT

$$F[u,v] = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} f[n,m] \exp\left(-2\pi j \left(\frac{un}{N} + \frac{vm}{M}\right)\right)$$

- Linearity
- Symmetry: Fourier transform of a real signal has coefficients that come in pairs, with *F* [*u*, *v*] being the complex conjugate of *F* [-*u*, -v].

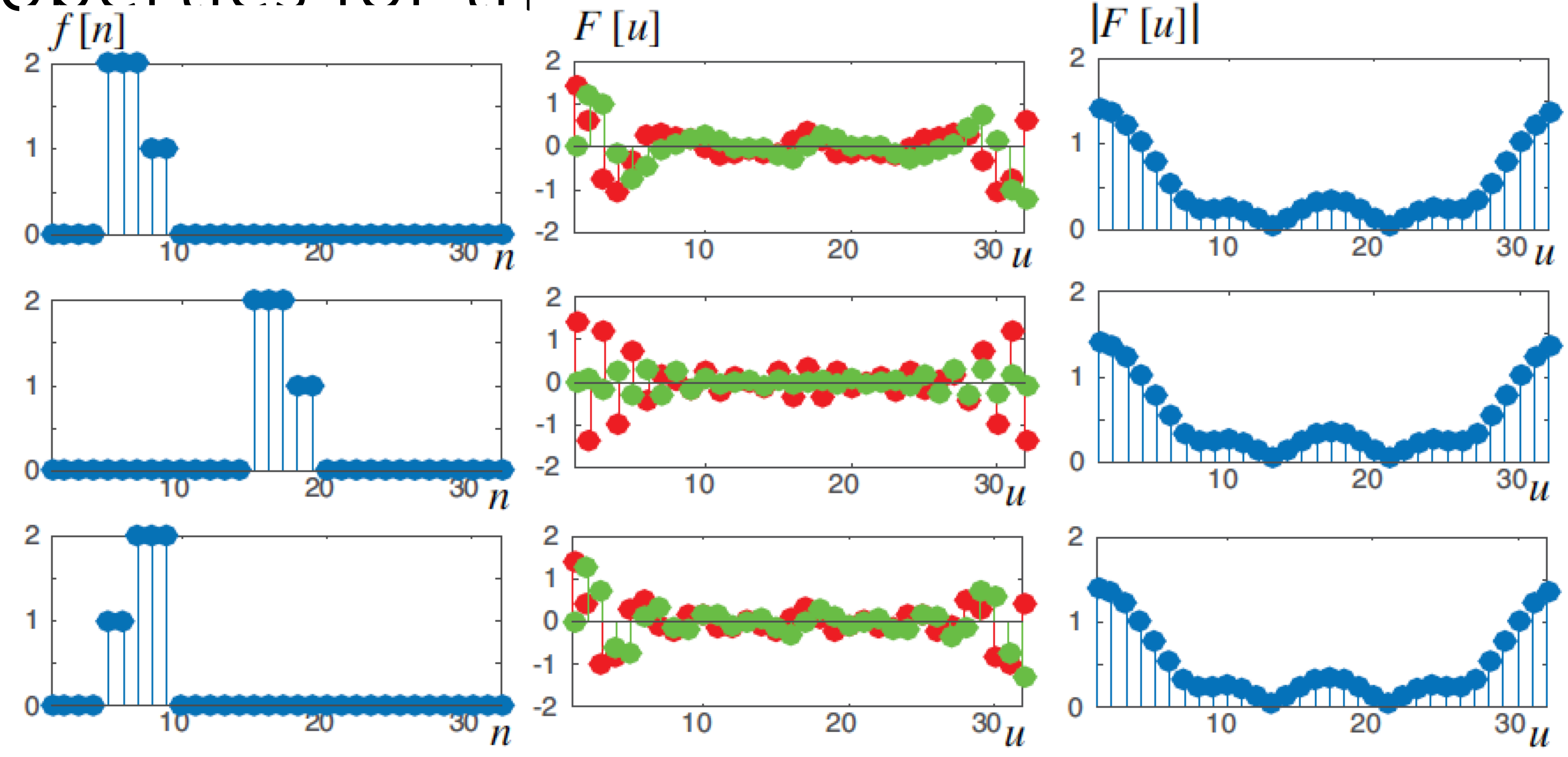
Properties for the DFT

- Shift in space

$$DFT\left\{f\left[n-n_0, m-m_0\right]\right\}=$$

$$
\begin{aligned}
& =\sum_{n=0}^{N-1} \sum_{m=0}^{M-1} f\left[n-n_0, m-m_0\right] \exp \left(-2 \pi j\left(\frac{u n}{N}+\frac{v m}{M}\right)\right)= \\
& =\sum_{n=0}^{N-1} \sum_{m=0}^{M-1} f[n, m] \exp \left(-2 \pi j\left(\frac{u\left(n+n_0\right)}{N}+\frac{v\left(m+m_0\right)}{M}\right)\right)= \\
& =\boxed{F[u, v] \exp \left(-2 \pi j\left(\frac{u n_0}{N}+\frac{v m_0}{M}\right)\right)}
\end{aligned}
$$

Properties for the DFT



Only the phase changes! The magnitude is translation invariant.

Properties for the DFT

- Modulation

$$f[n,m]\cos\left(2\pi j\left(\frac{u_0 n}{N}+\frac{v_0 m}{M}\right)\right)$$

$$f[n,m]\exp\left(-2\pi j\left(\frac{u_0 n}{N}+\frac{v_0 m}{M}\right)\right)$$

Multiplying by a complex exponential results in a translation of the DFT

$$DFT\left\{f[n,m]\exp\left(-2\pi j\left(\frac{u_0 n}{N}+\frac{v_0 m}{M}\right)\right)\right\}=F[u-u_0,v-v_0]$$

Frequencies

DFT amplitude

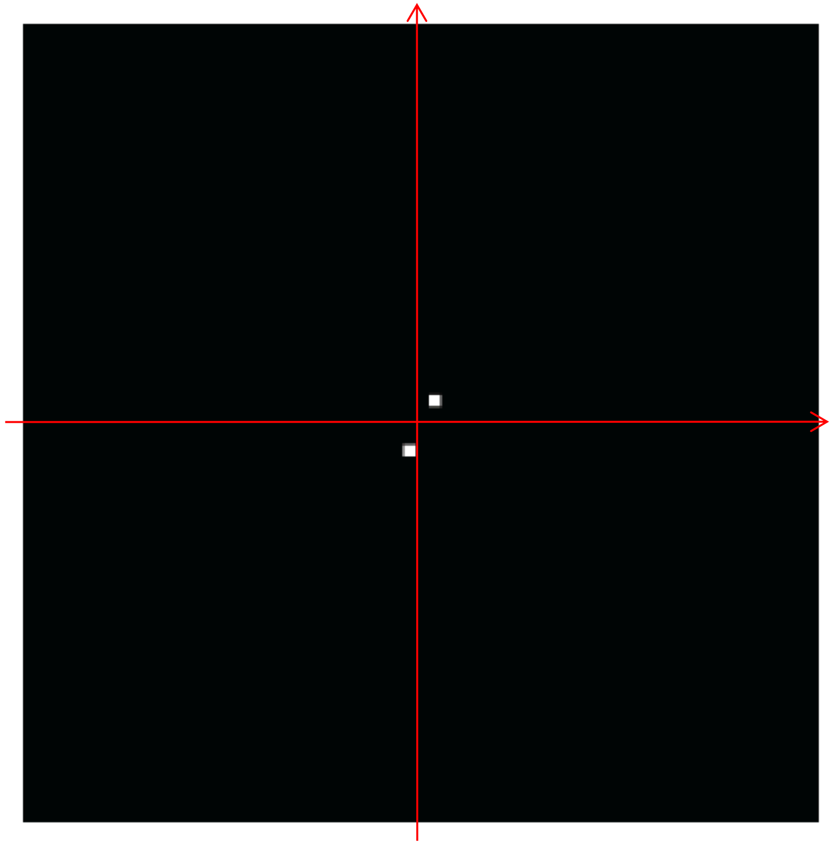
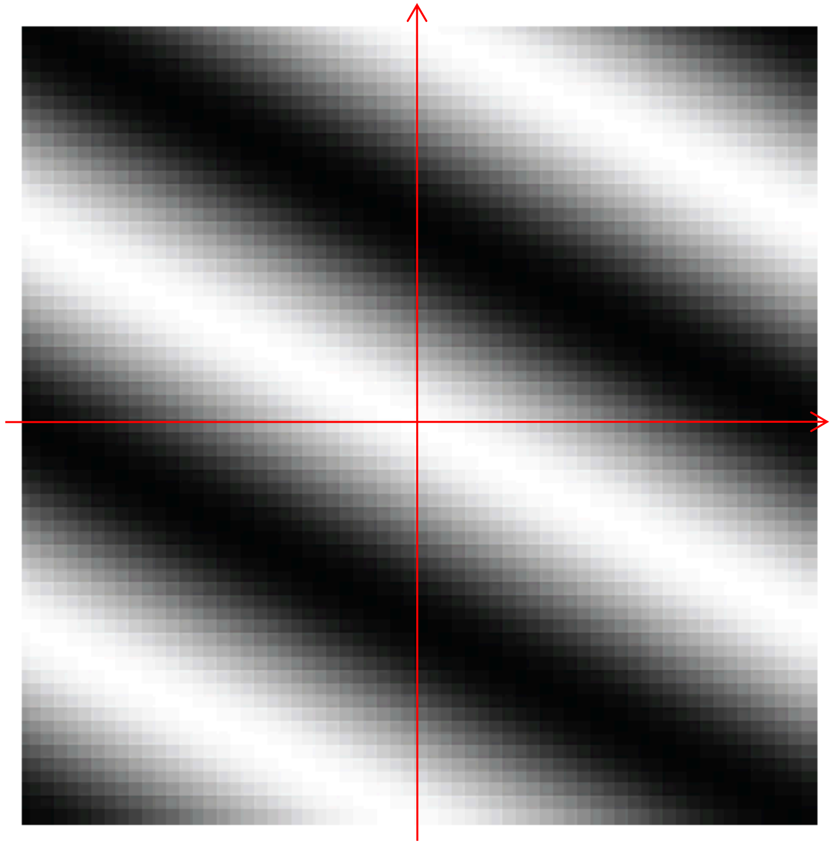


Image
(assuming zero phase)



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

Conjugate symmetric

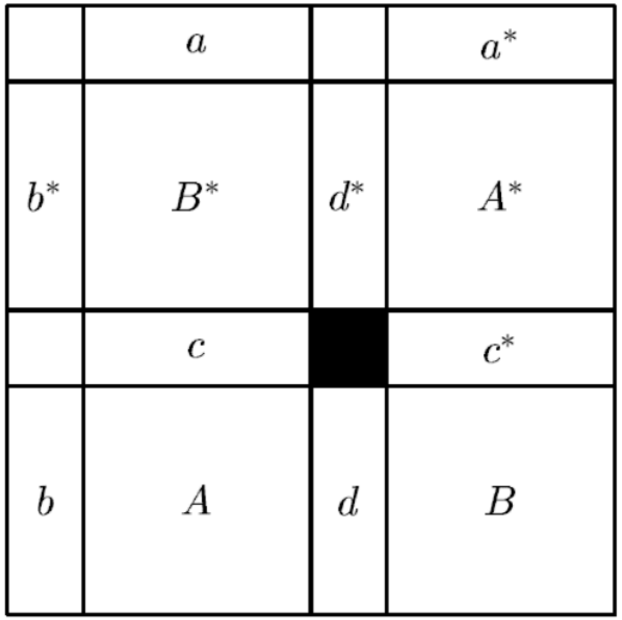
- Thus, if we put *u* = -*u*, and *v* = -*v* into the DFT equation

$$F(u,v) \quad = \quad \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\exp\left[-2\pi i\left(\frac{xu}{M}+\frac{yv}{N}\right)\right]$$

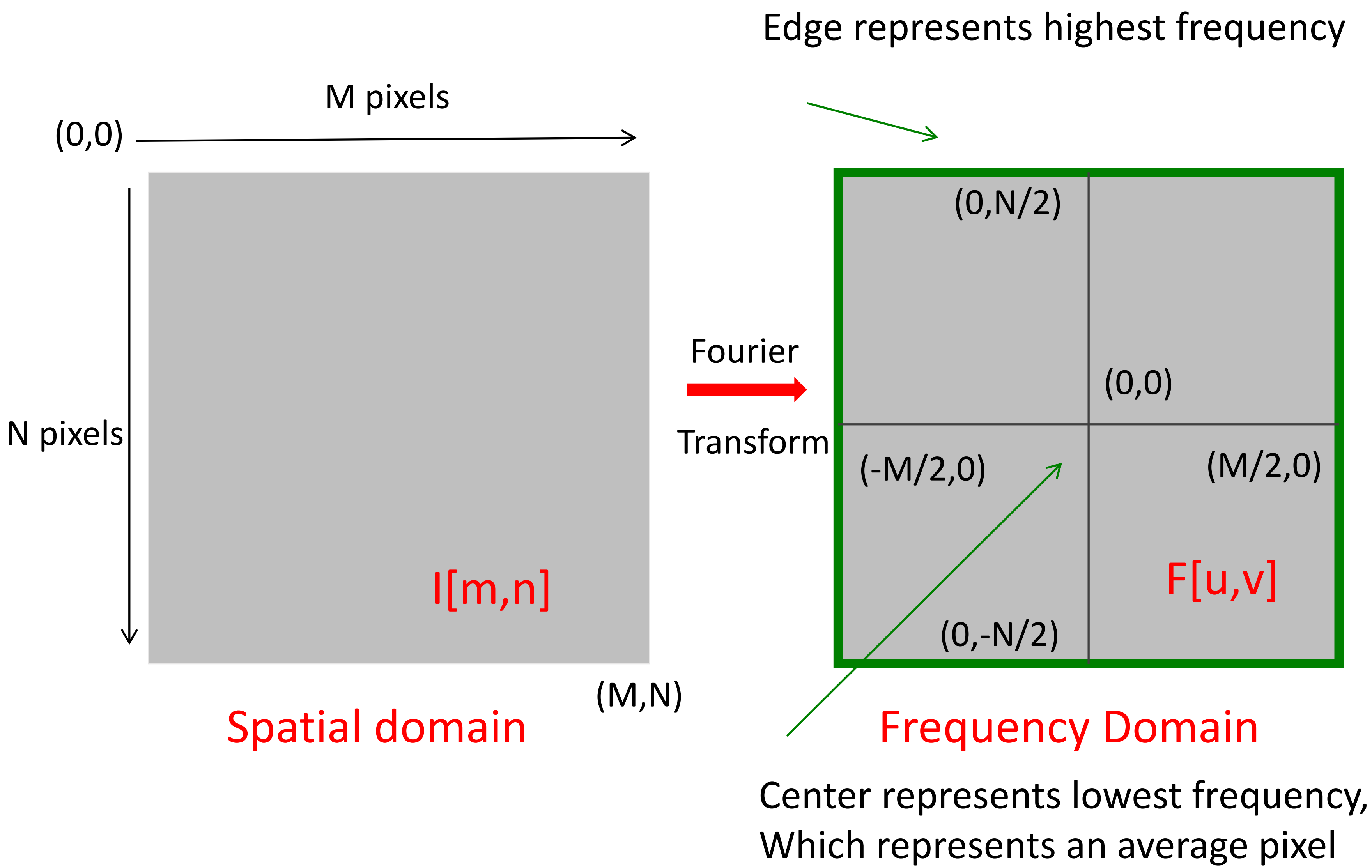
Then

$$\mathcal{F}(u,v) = \mathcal{F}^*(-u+pM, -v+qN)$$

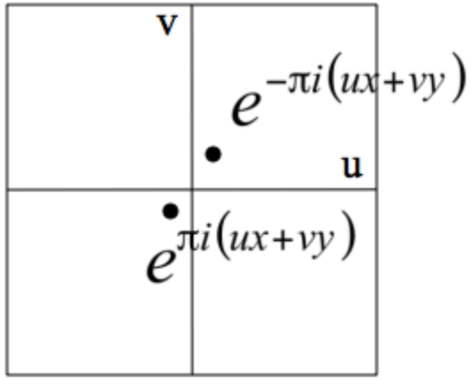
for any integers *p* and *q*

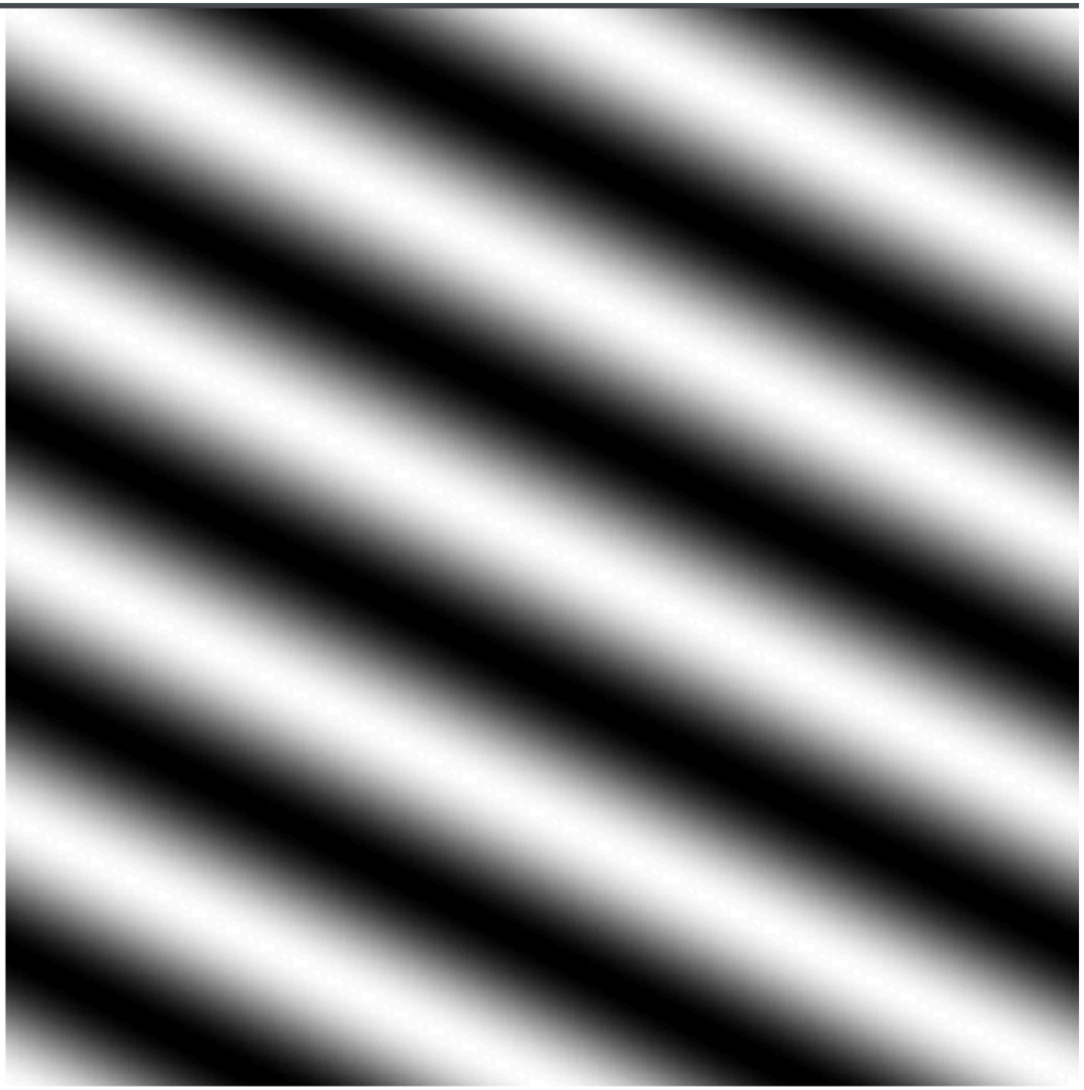


Fourier Transform of Images



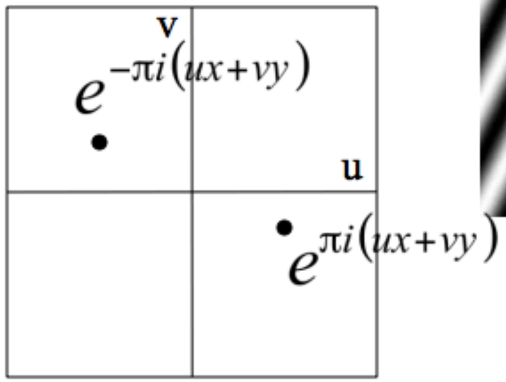
To get some sense of what basis elements look like, we plot a basis element --- or rather, its real part --- as a function of x,y for some fixed u, v. We get a function that is constant when (ux+vy) is constant. The magnitude of the vector (u, v) gives a frequency, and its direction gives an orientation. The function is a sinusoid with this frequency along the direction, and constant perpendicular to the direction.



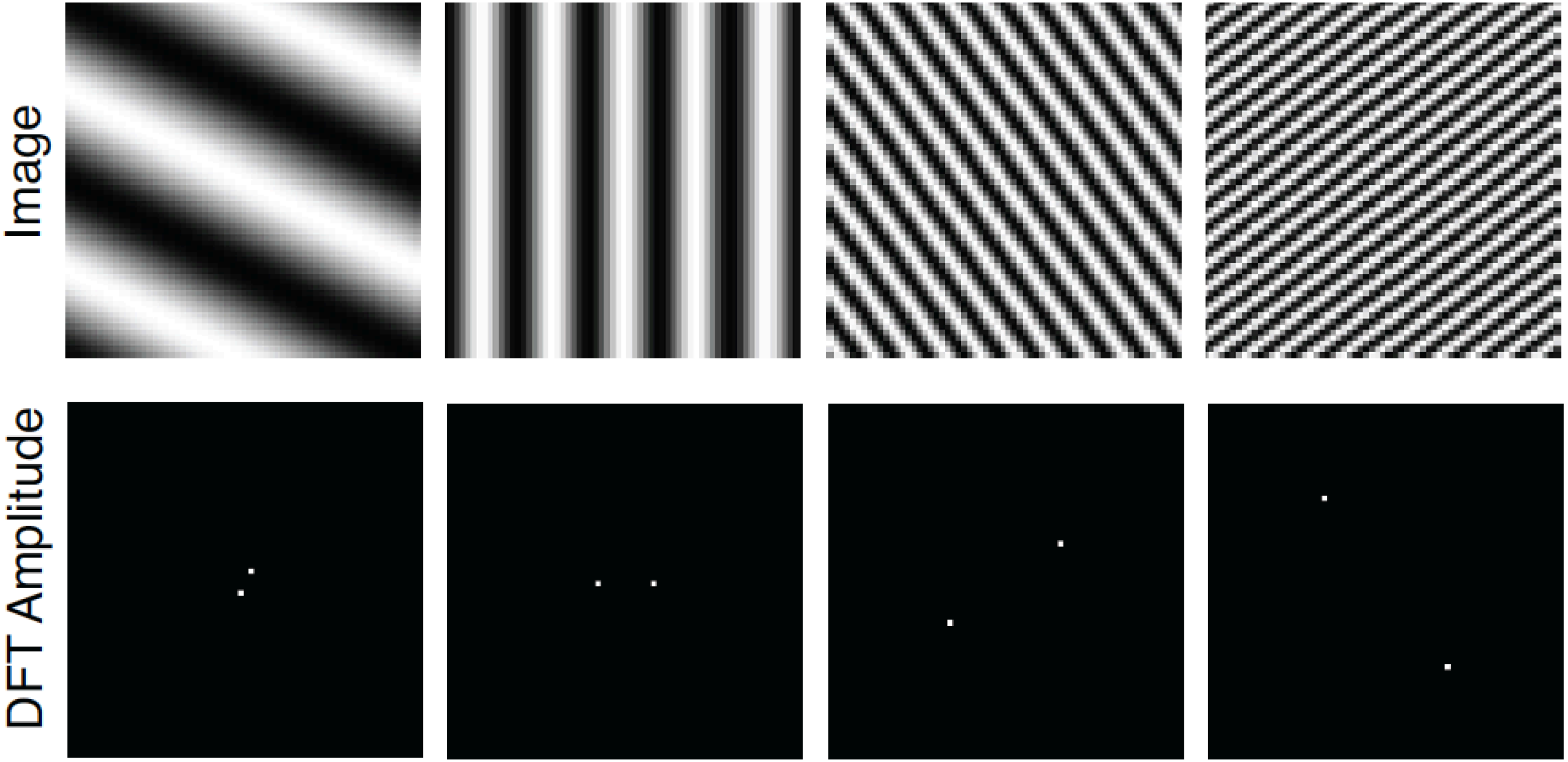


x,y plane with u,v given. Orientation tan a = v/u
Frequency of the sinusoid sqrt(u^2 + v^2).
The real component only here as intensity image.

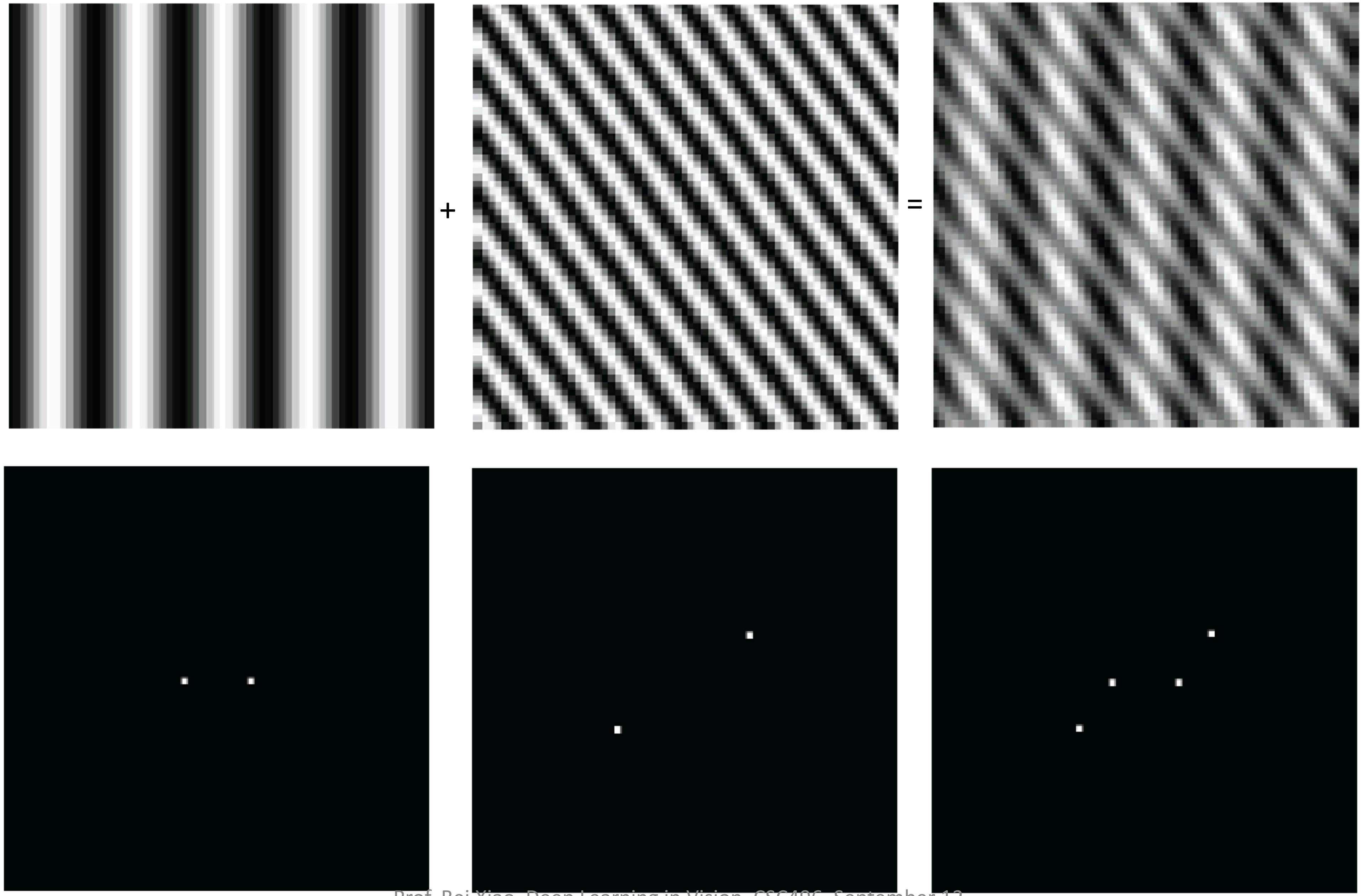
Here u and v are
larger than in
the previous
slide.



Frequencies

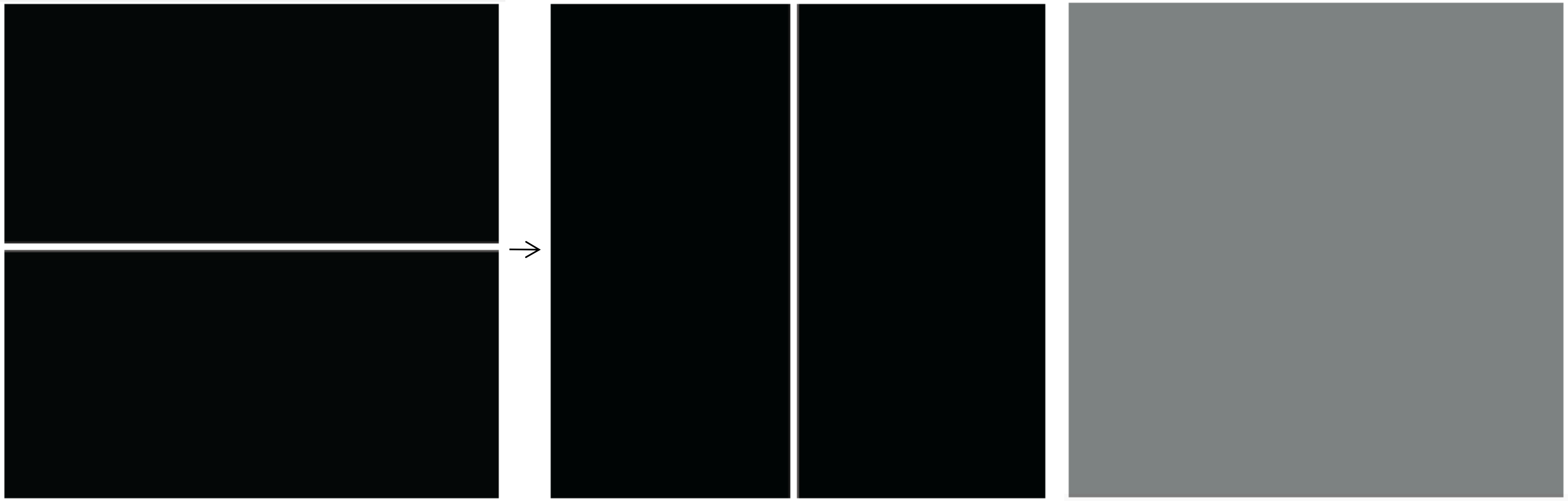


Images are 64x64 pixels. The wave is a cosine (if phase is zero).



Some important Fourier transforms

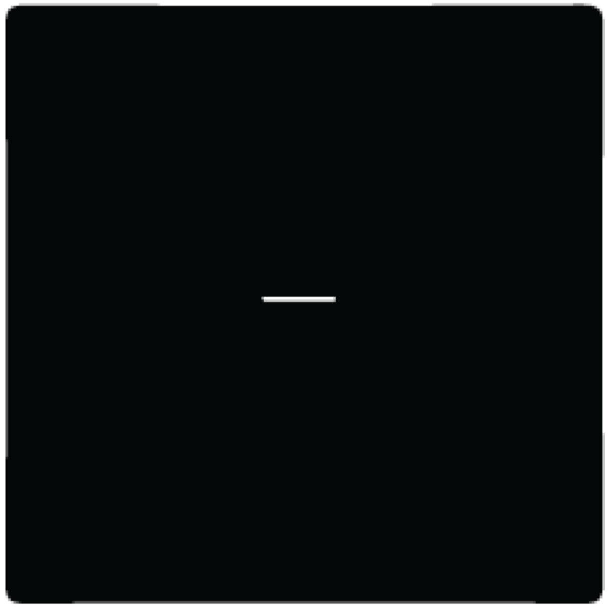
Image | Magnitude DFT | Phase DFT



Images are 64x64 pixels.

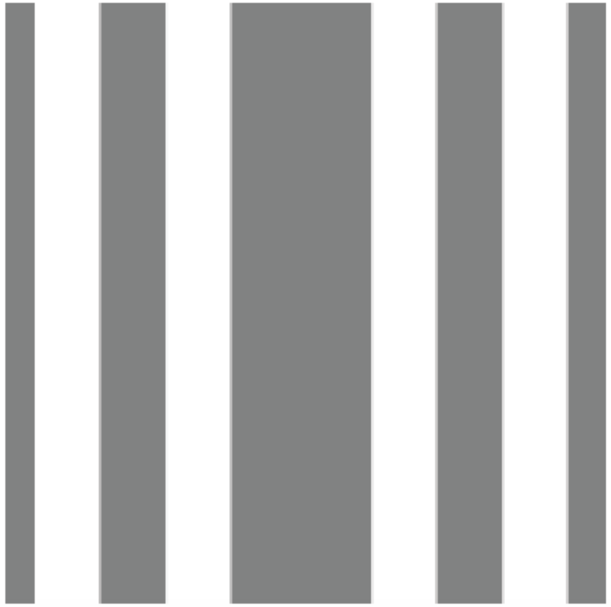
Some important Fourier transforms

Image



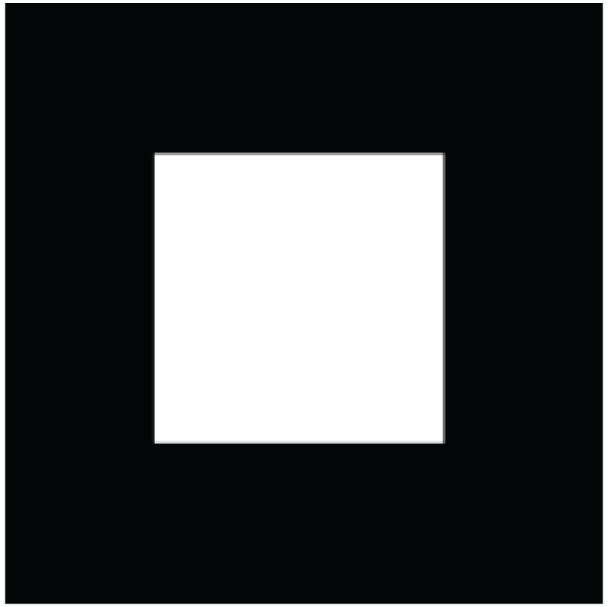
Magnitude DFT

Phase DFT

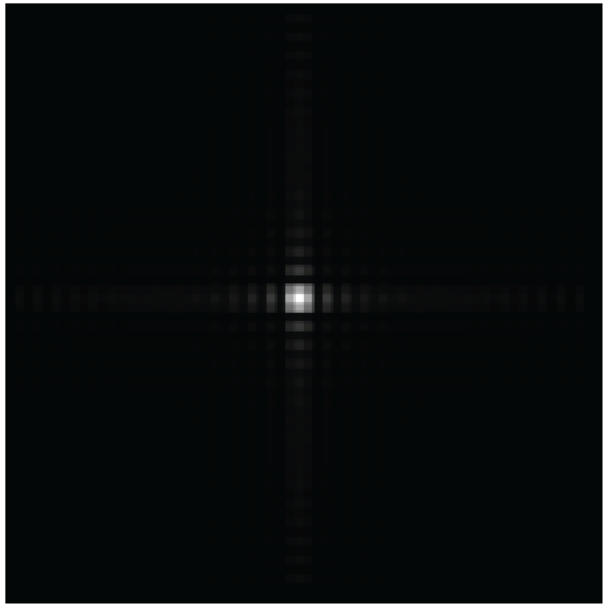


Some important Fourier transforms

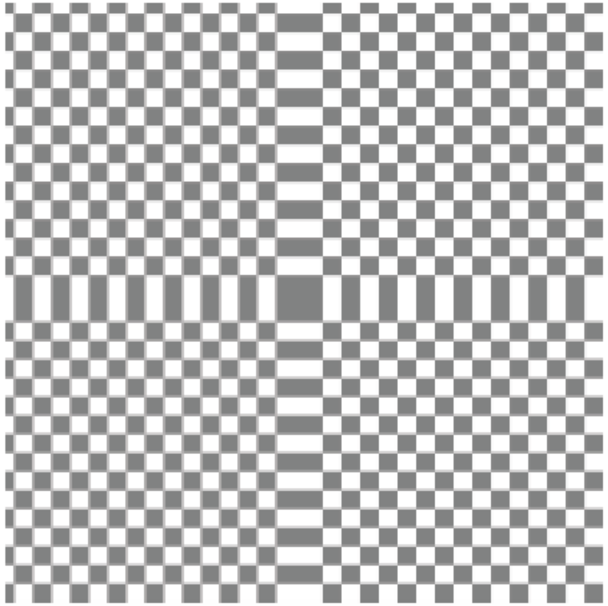
Image



Magnitude DFT

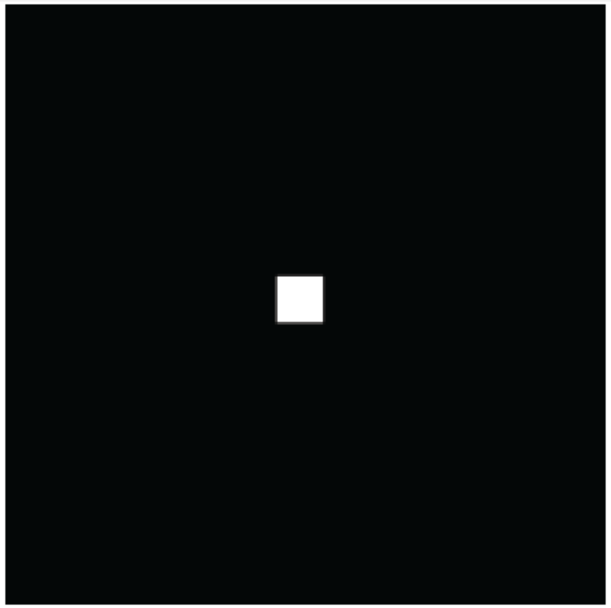


Phase DFT

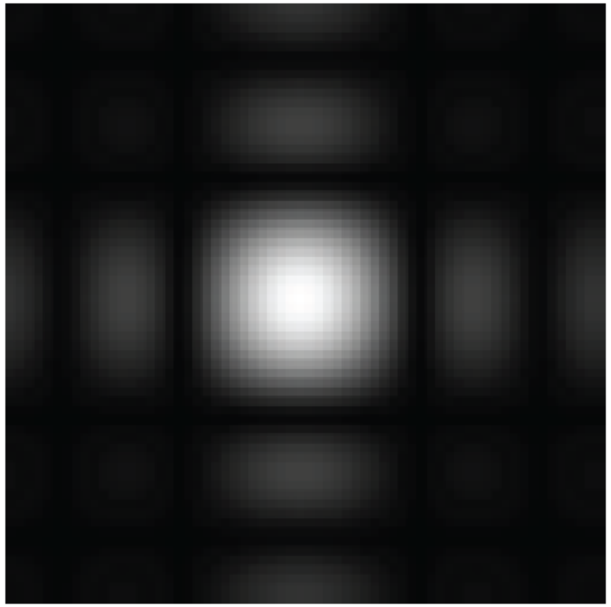


Some important Fourier transforms

Image

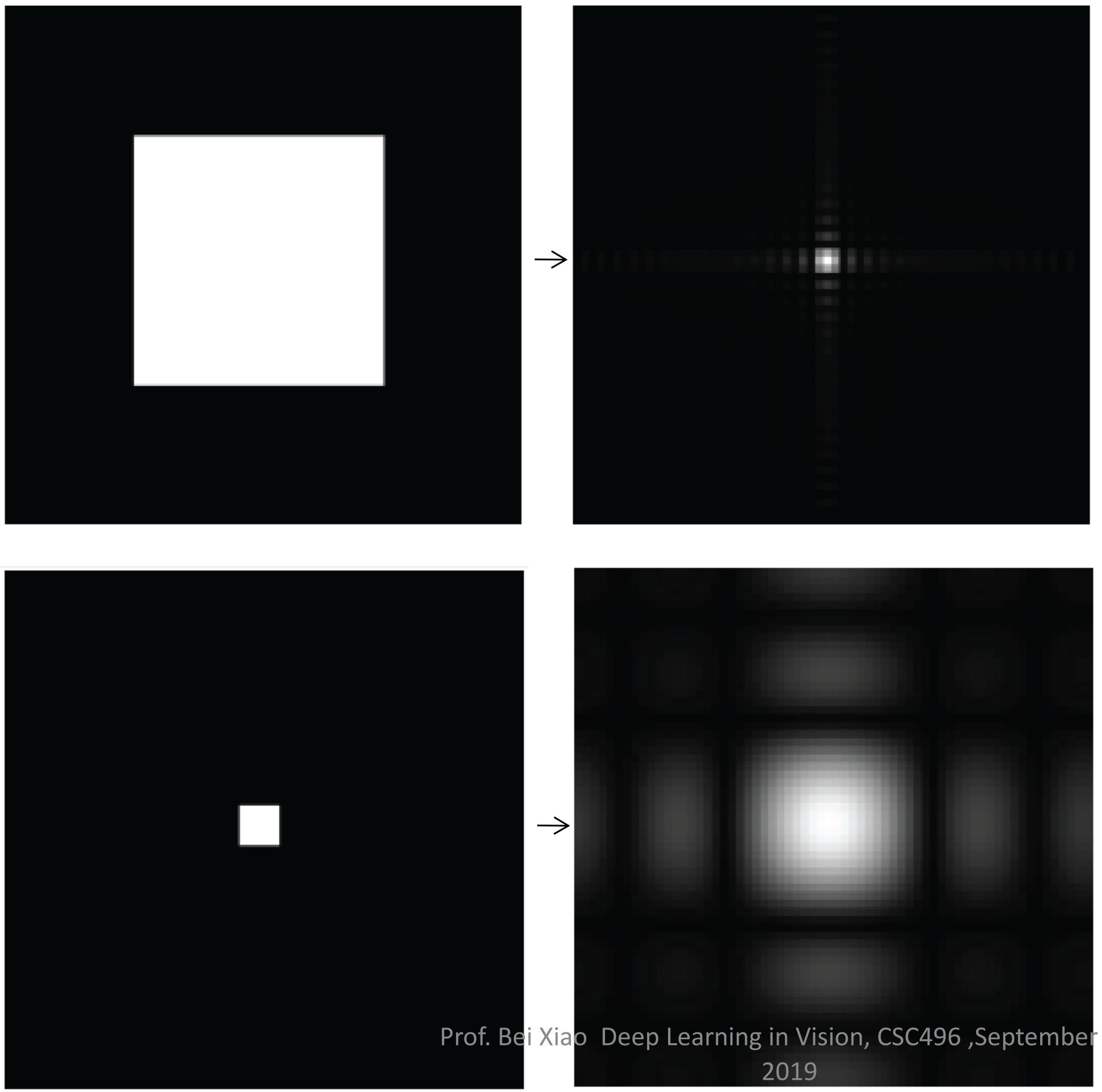


Magnitude DFT



Phase DFT

Image Magnitude DFT

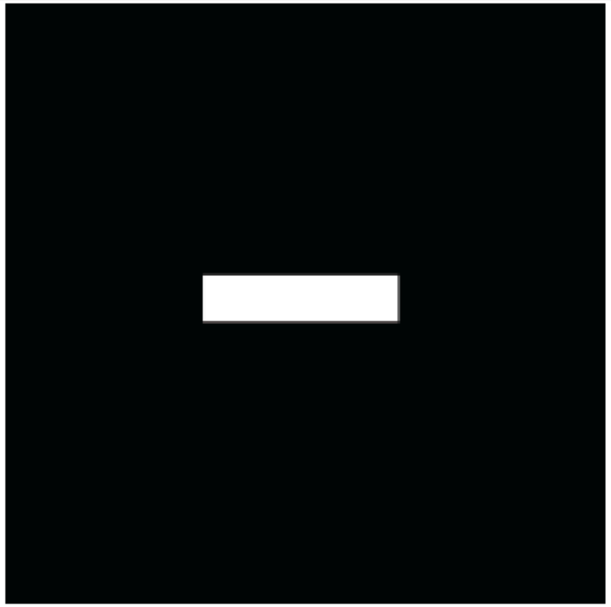


Scale

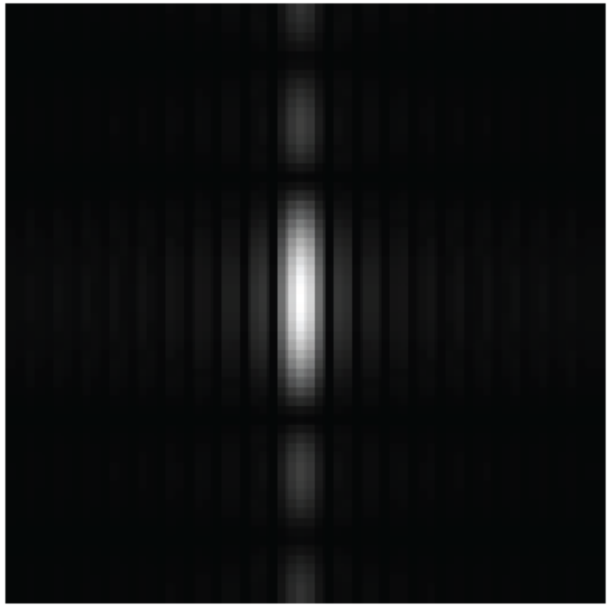
Small image
details produce content in
high spatial frequencies

Some important Fourier transforms

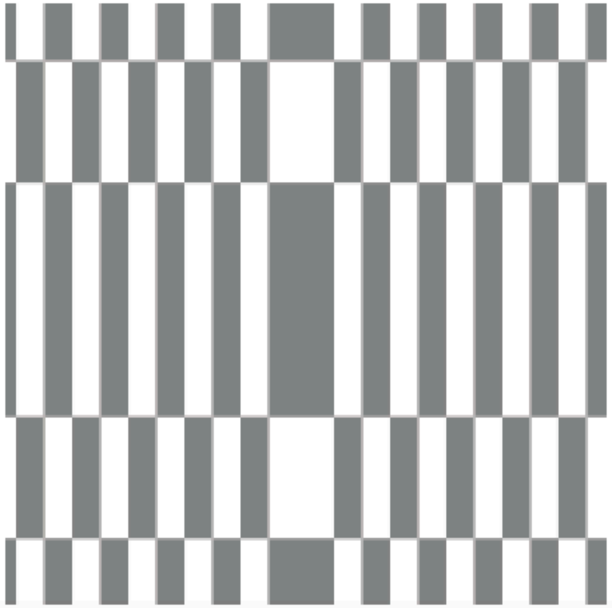
Image



Magnitude DFT

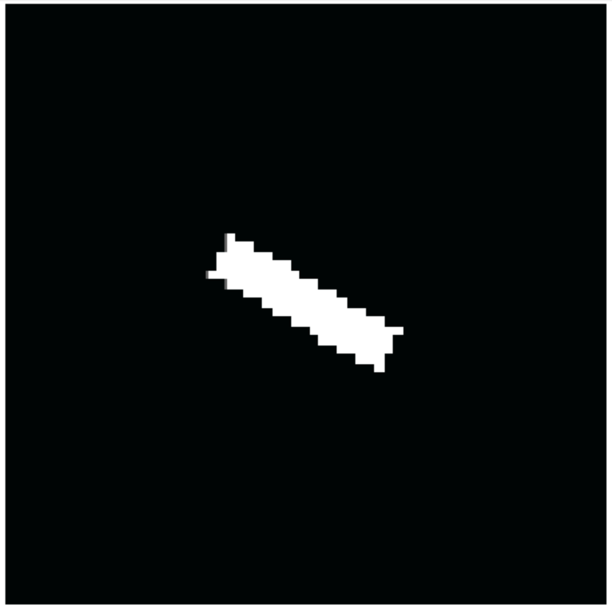


Phase DFT

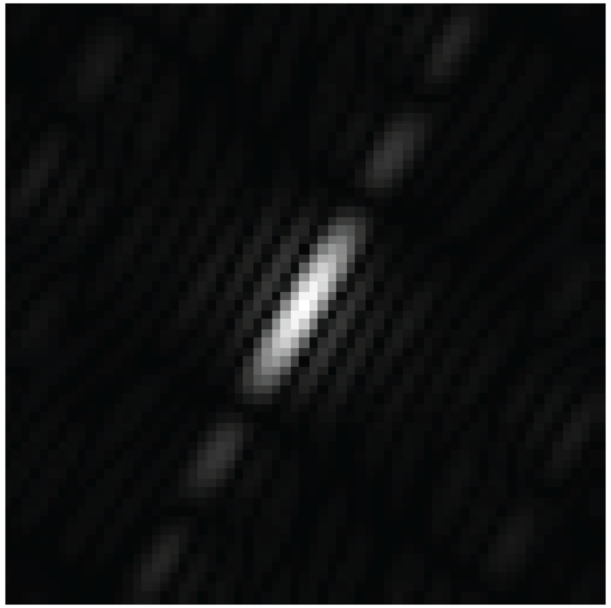


Some important Fourier transforms

Image



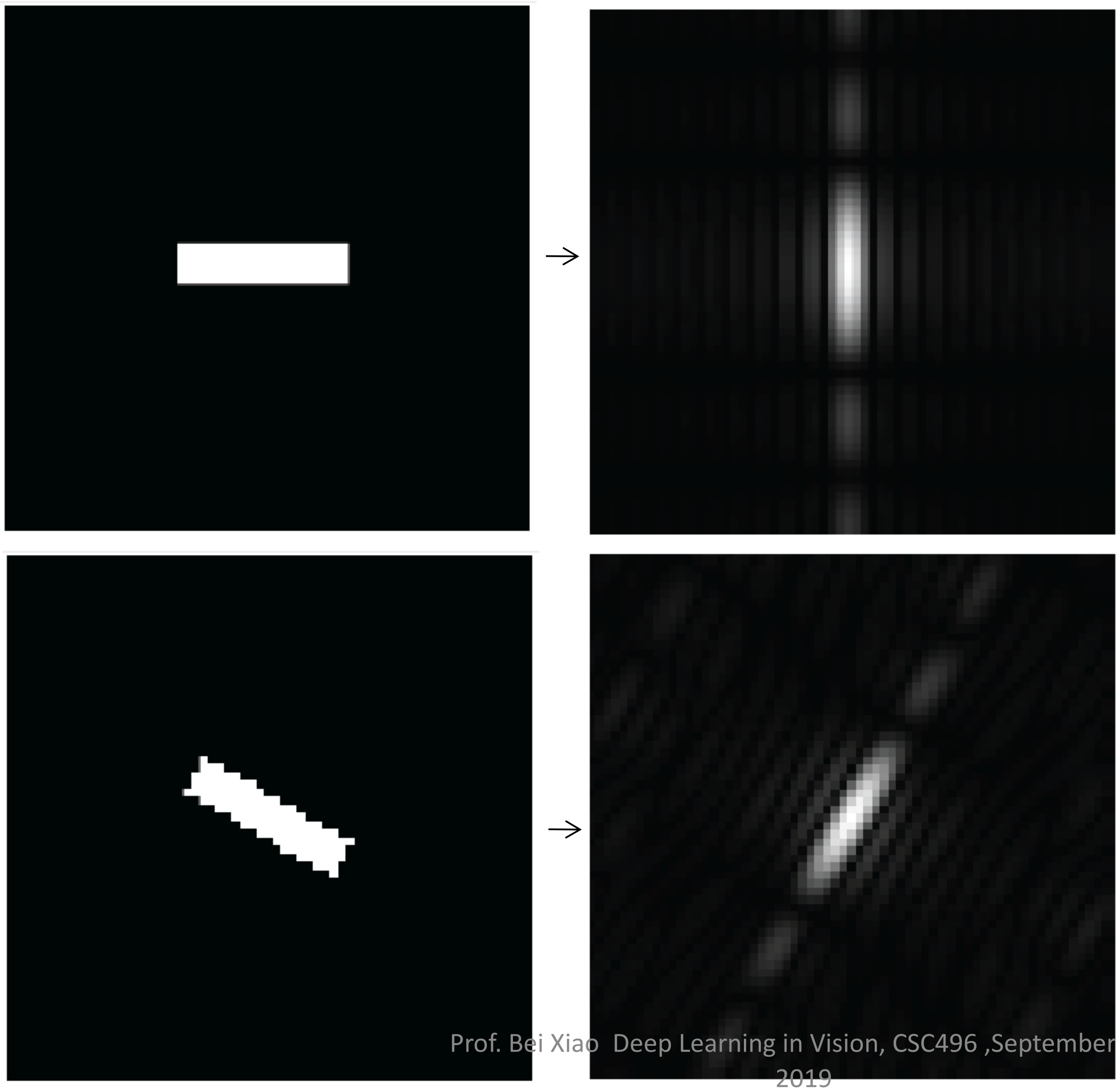
Magnitude DFT



Phase DFT

Image

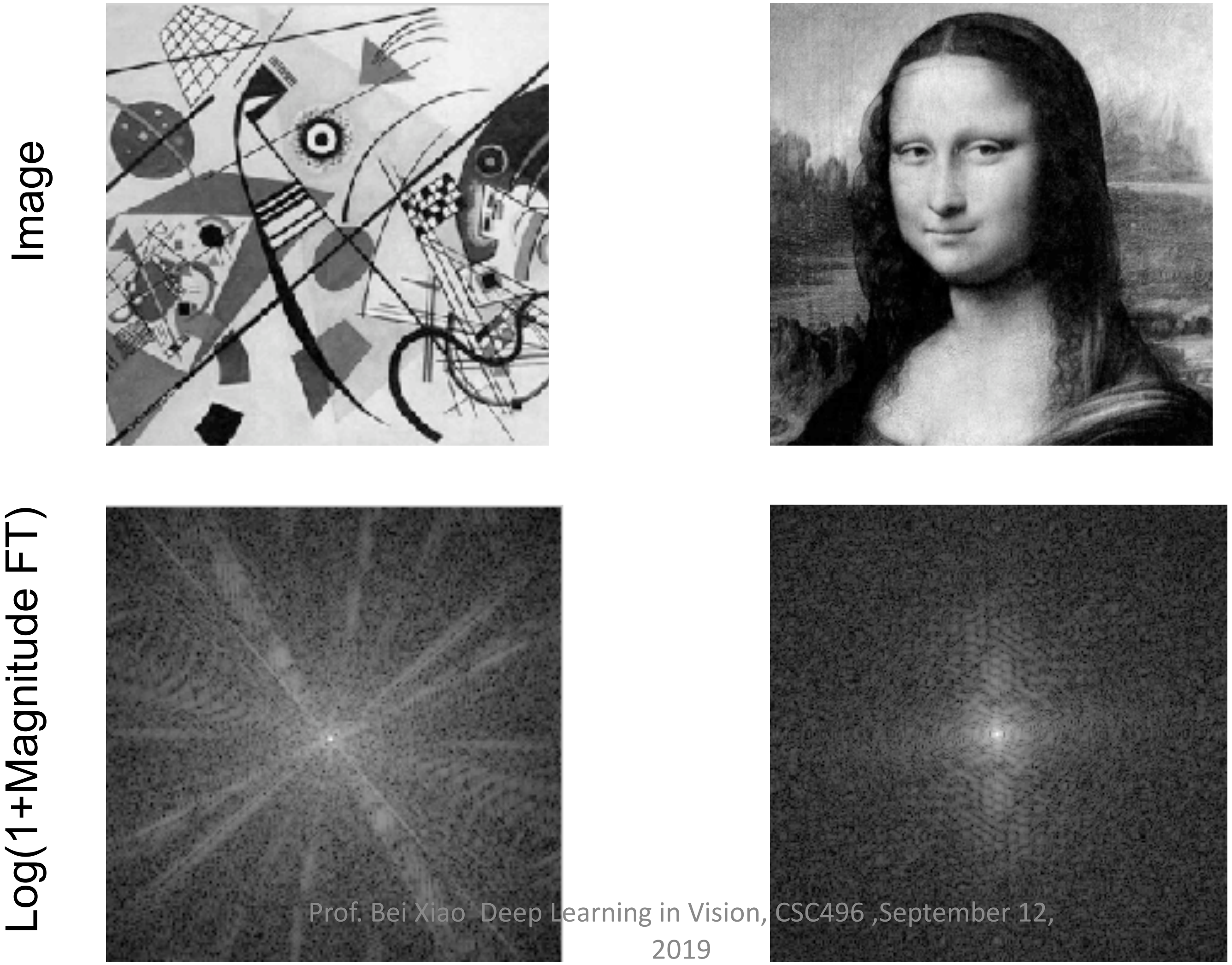
Magnitude DFT



Orientation

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

The Fourier Transform of some important images



More properties for the DFT

$$F[u,v] = \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} f[n,m]\exp\left(-2\pi j\left(\frac{un}{N}+\frac{vm}{M}\right)\right)$$

DFT of the convolution

$$f = g \circ h \longleftrightarrow F[u,v] = G[u,v]\,H[u,v]$$

$$F[u,v] = DFT\{g \circ h\}$$

$$= \sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\boxed{\sum_{k=0}^{M-1}\sum_{l=0}^{N-1} g[m-k,n-l]\,h[k,l]}\exp\left(-2\pi j\left(\frac{mu}{M}+\frac{nv}{N}\right)\right)$$

$$F[u,v] = \sum_{k=0}^{M-1}\sum_{l=0}^{N-1} h[k,l] \sum_{m'=-k}^{M-k-1}\sum_{n'=-l}^{N-l-1} g[m',n']\exp\left(-2\pi j\left(\frac{(m'+k)u}{M}+\frac{(n'+l)v}{N}\right)\right)$$

$$F[u,v] = \sum_{k=0}^{M-1}\sum_{l=0}^{N-1} G[u,v]\exp\left(-2\pi j\left(\frac{ku}{M}+\frac{lv}{N}\right)\right)h[k,l]$$

Linear filtering

$$g[m,n] \longrightarrow \boxed{h[m,n]} \longrightarrow f[m,n]$$

In the spatial domain:

$$f[m,n] = h \circ g = \sum_{k,l} h[m-k, n-l]\, g[k,l]$$

In the frequency domain:

$$F[u,v] = G[u,v]\, H[u,v]$$

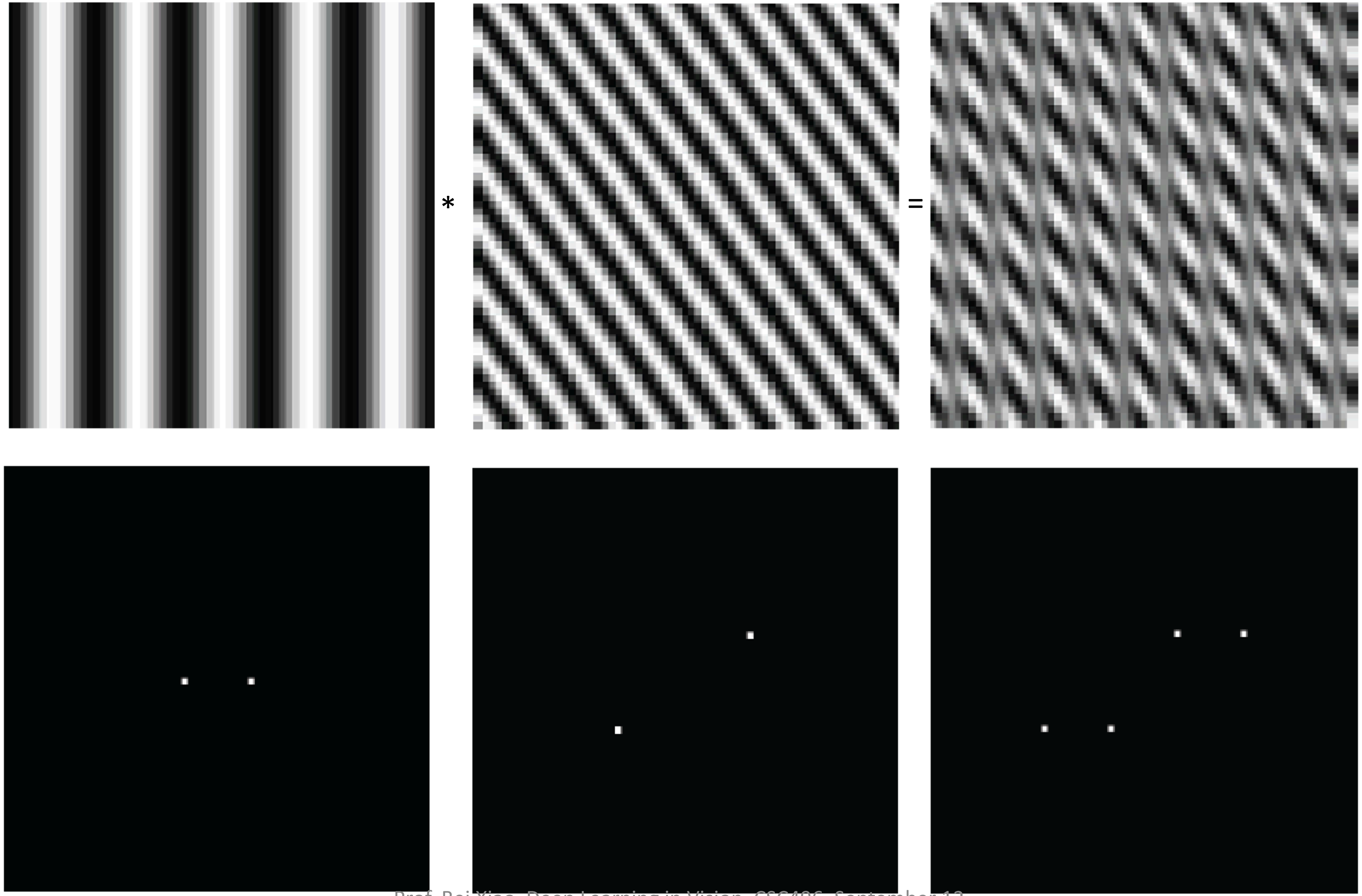
Product of images

The Fourier transform of the product of two images

$$f[n,m] = g[n,m]\, h[n,m]$$

is the convolution of their DFTs:

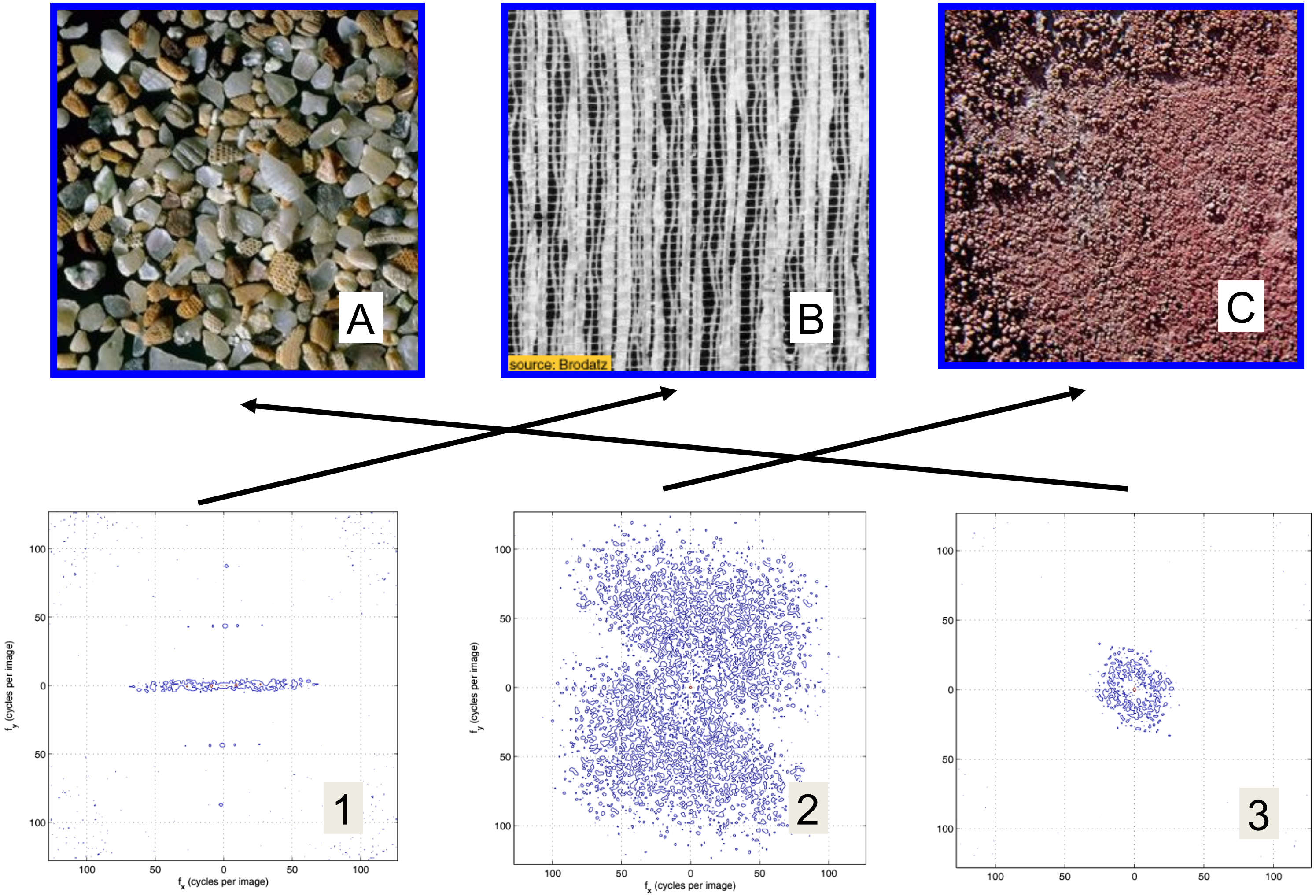
$$F[u,v] = \frac{1}{NM} G[u,v] \circ H[u,v]$$



Game: find the right pairs

Images

DFT
magnitude



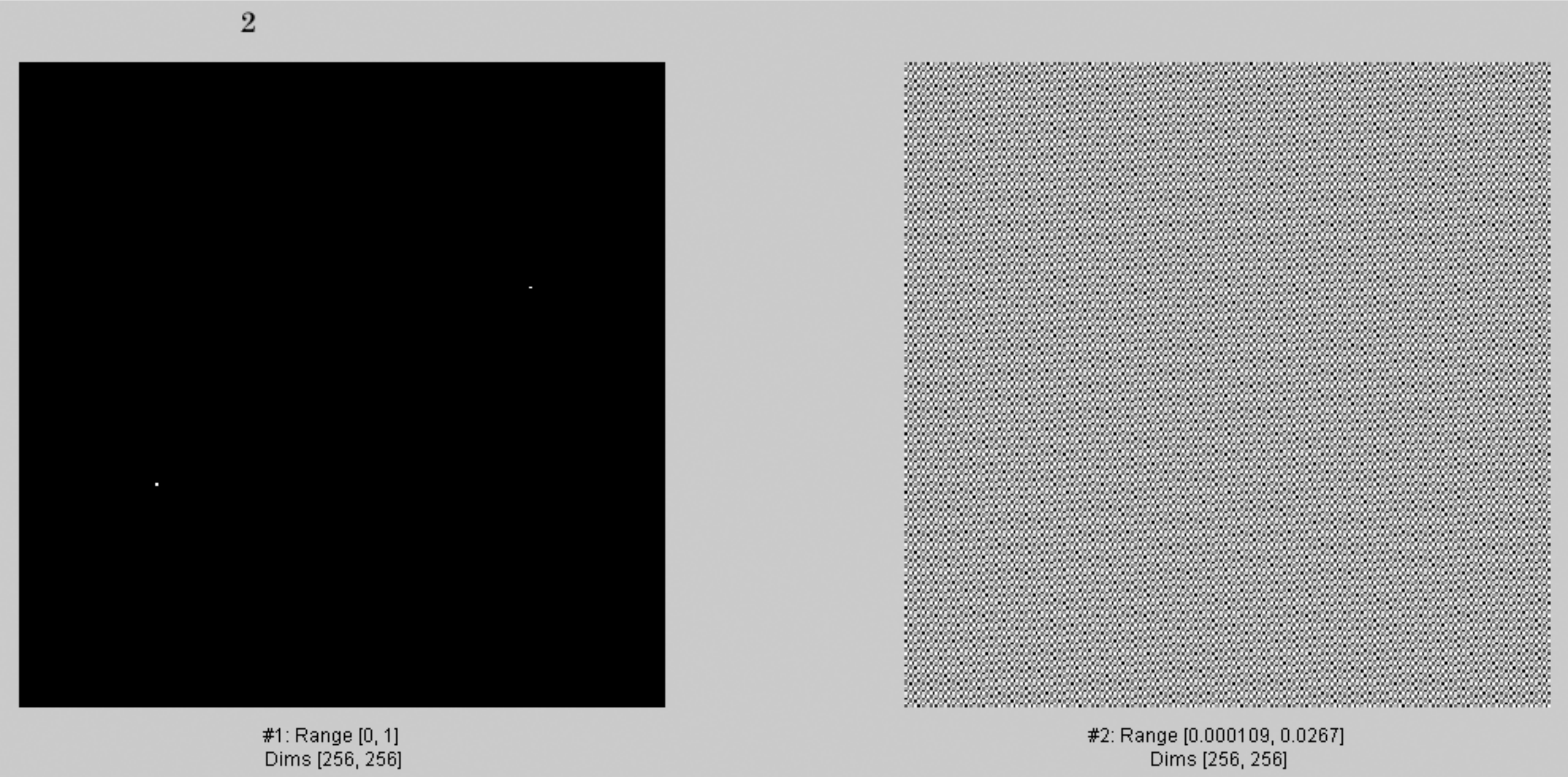
fx(cycles/image pixel size) fx(cycles/image pixel size) fx(cycles/image pixel size)

The inverse Discrete Fourier transform

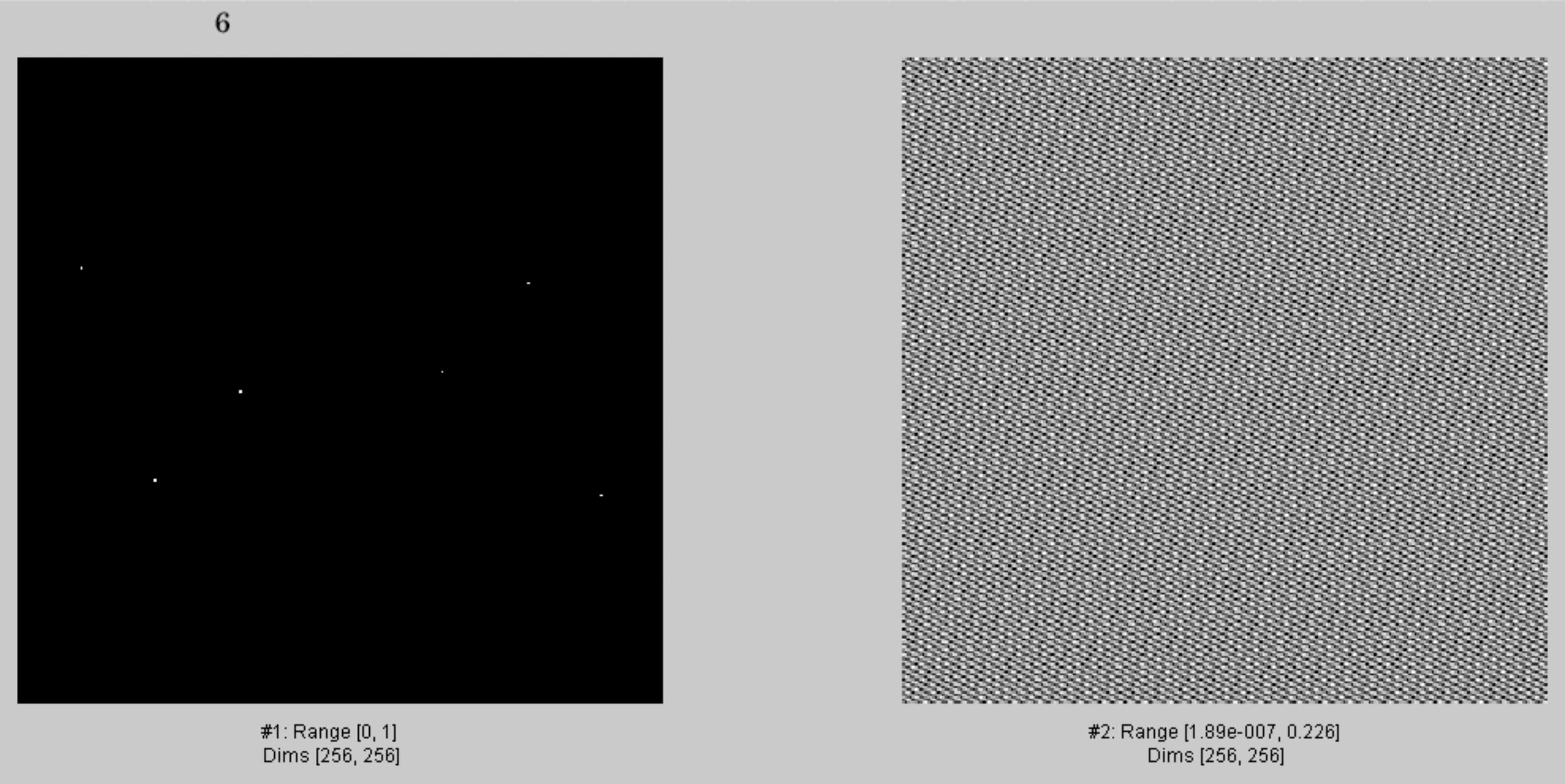
$$f[n,m] = \frac{1}{NM}\sum_{u=0}^{N-1}\sum_{v=0}^{M-1} F[u,v]\exp\left(+2\pi j\left(\frac{un}{N}+\frac{vm}{M}\right)\right)$$

How does summing waves ends up giving back a picture?

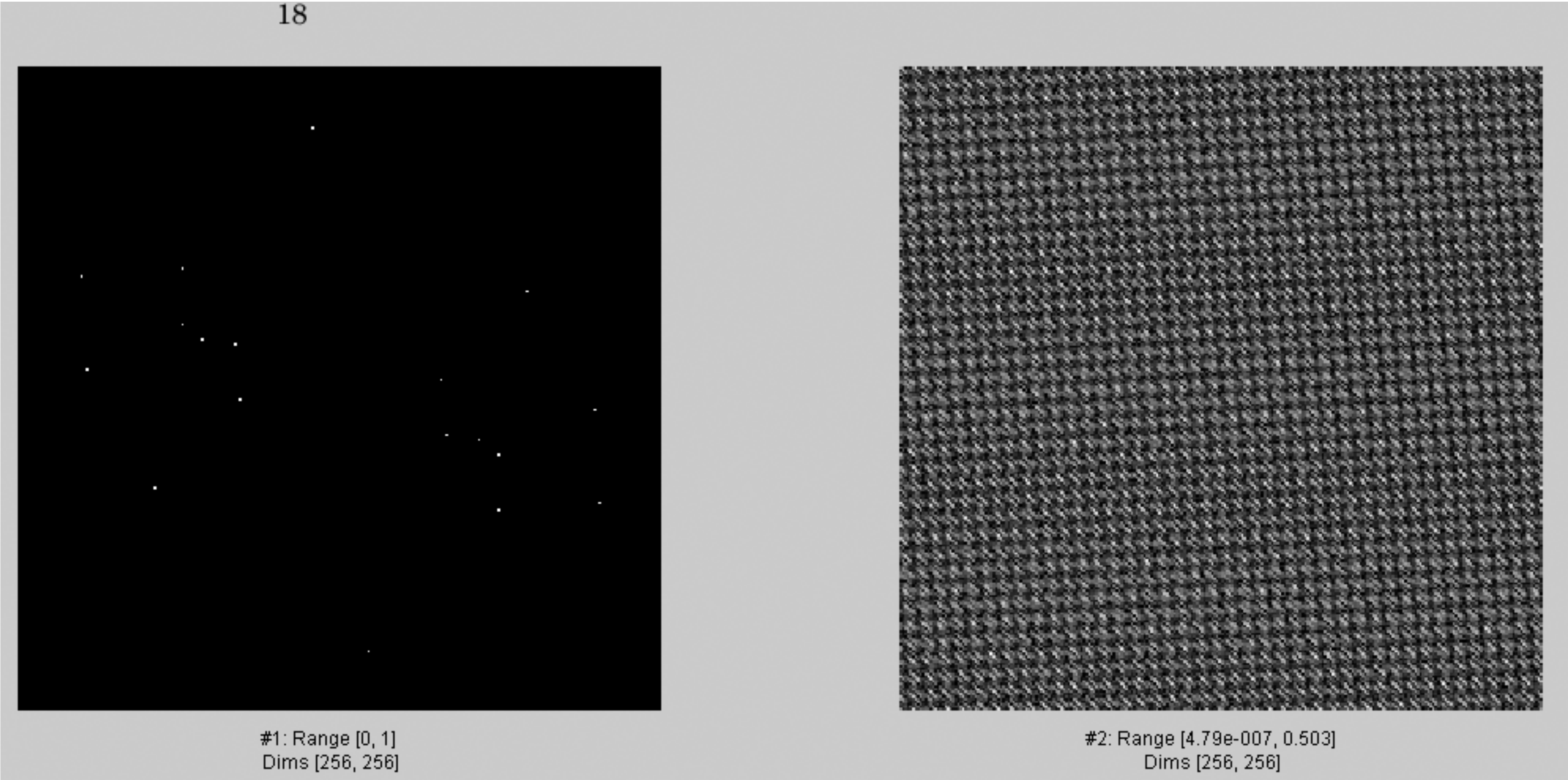
2



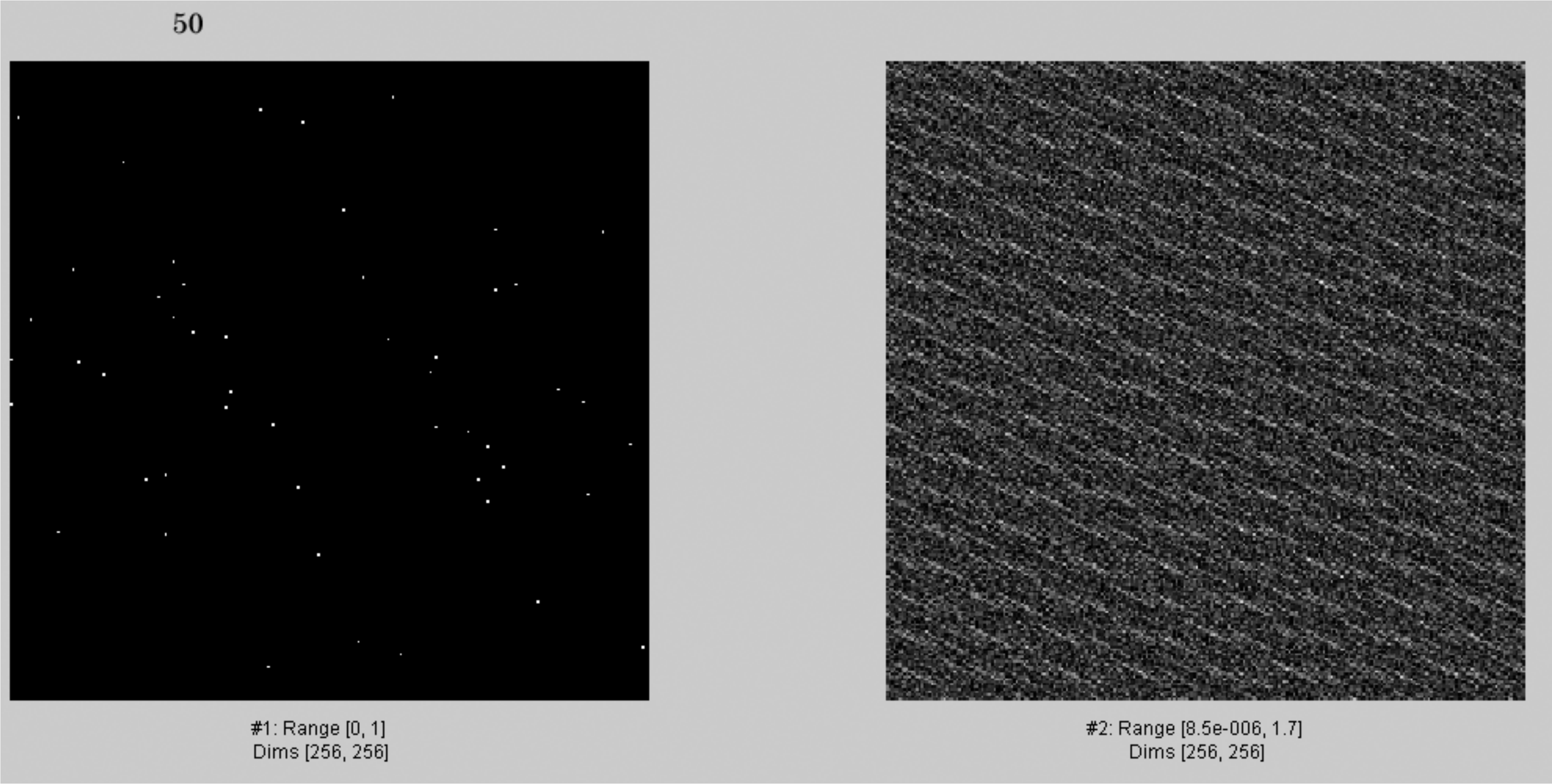
6



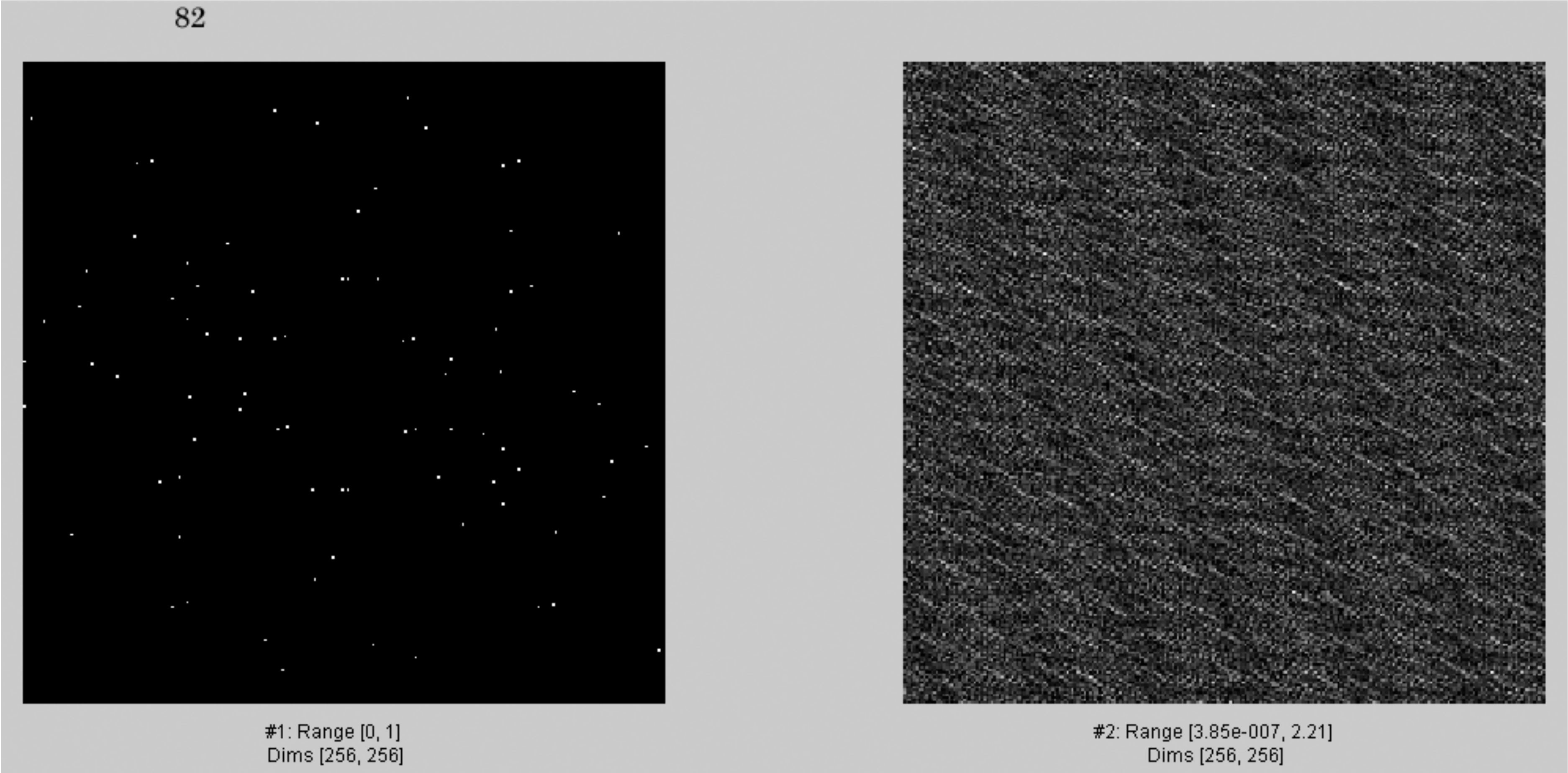
18



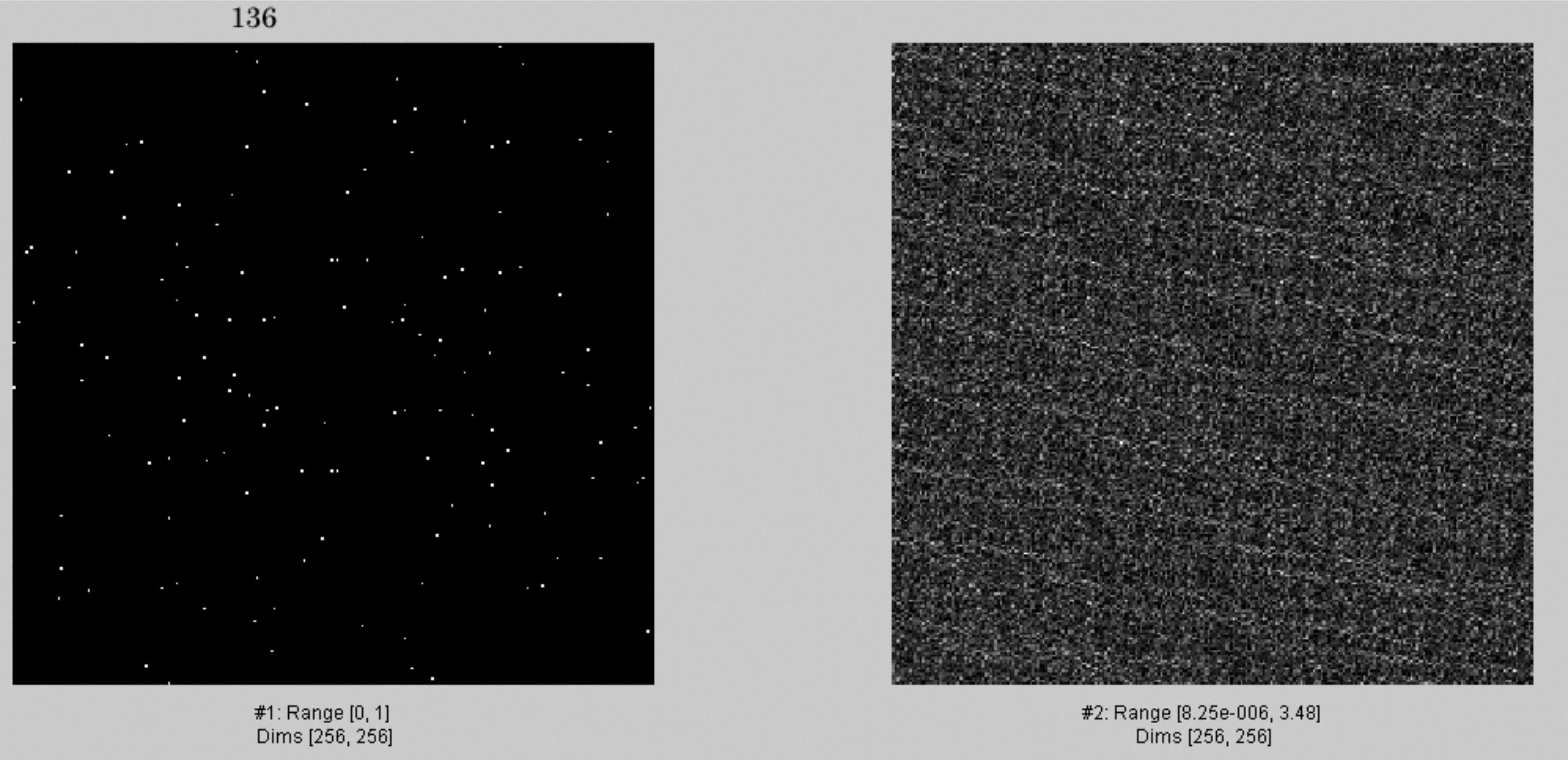
50



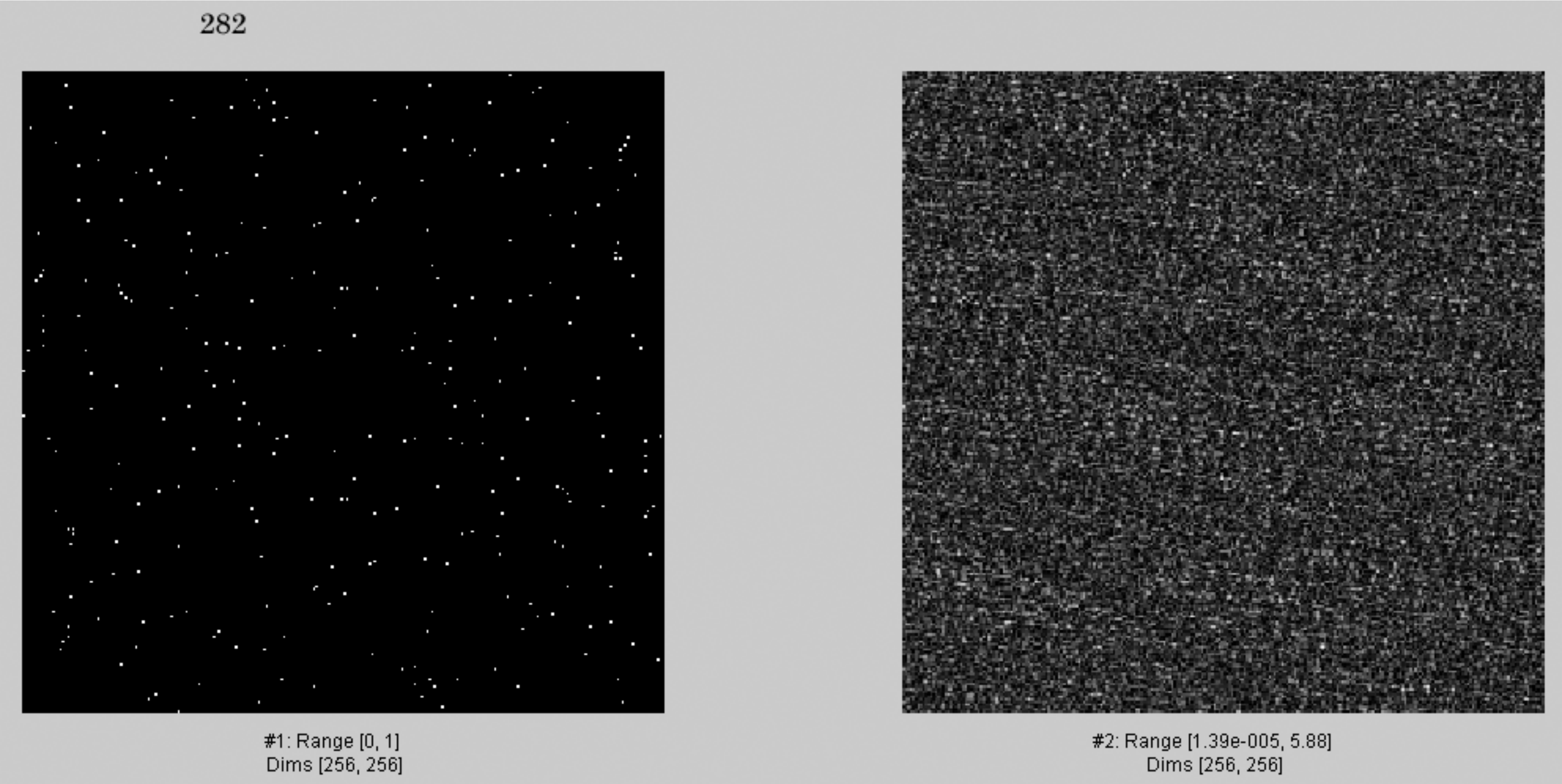
82



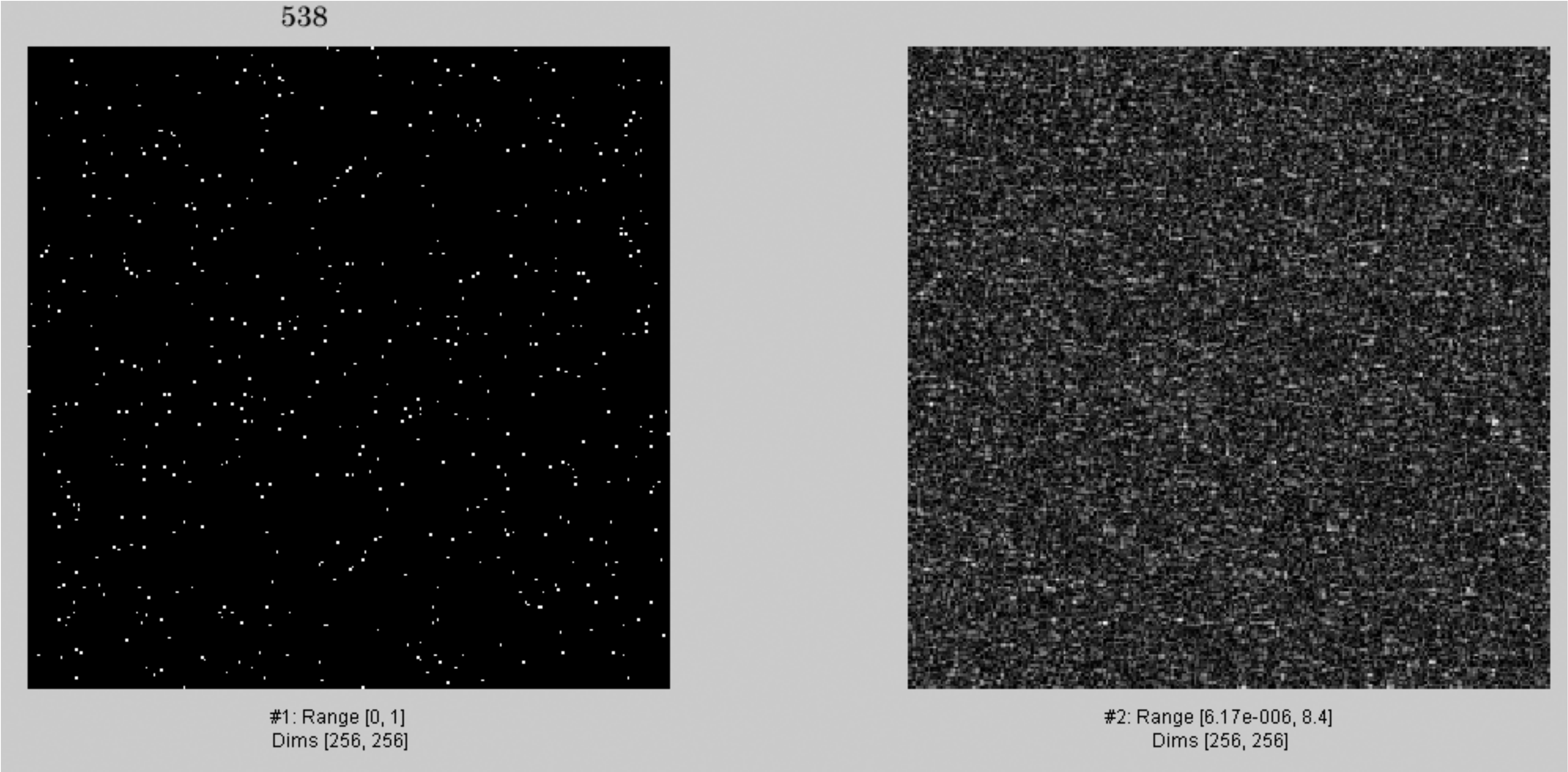
136



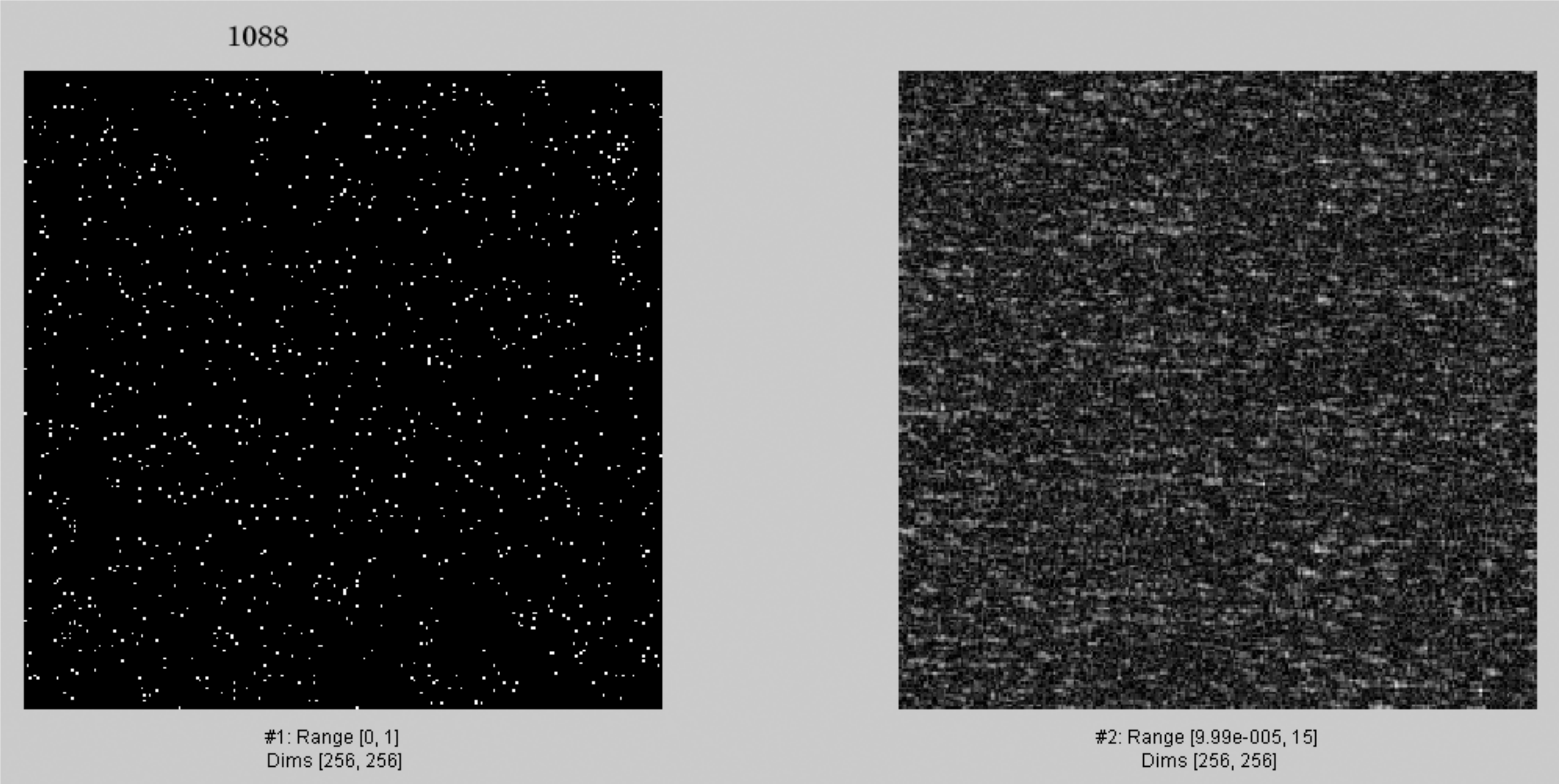
282



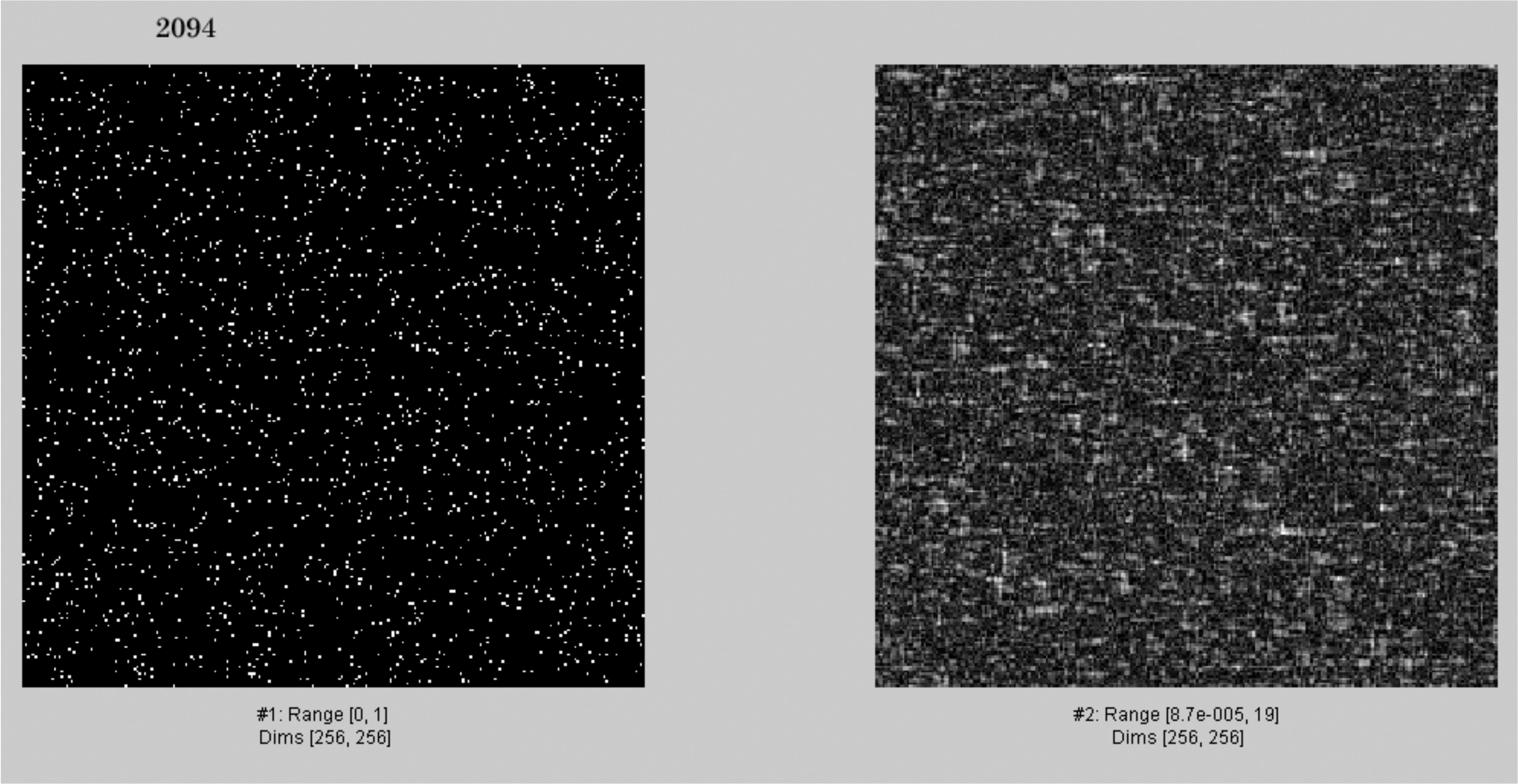
538



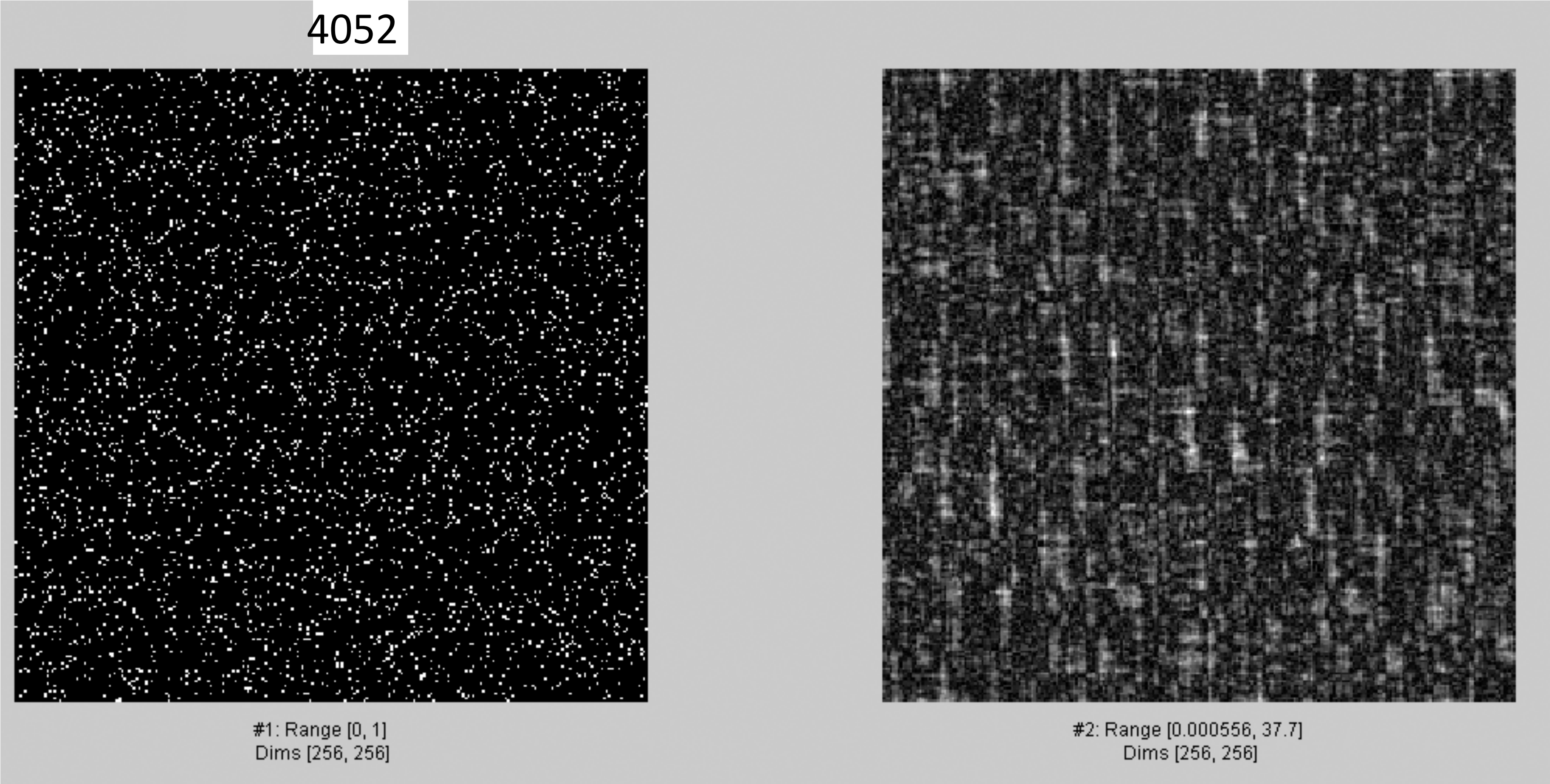
1088



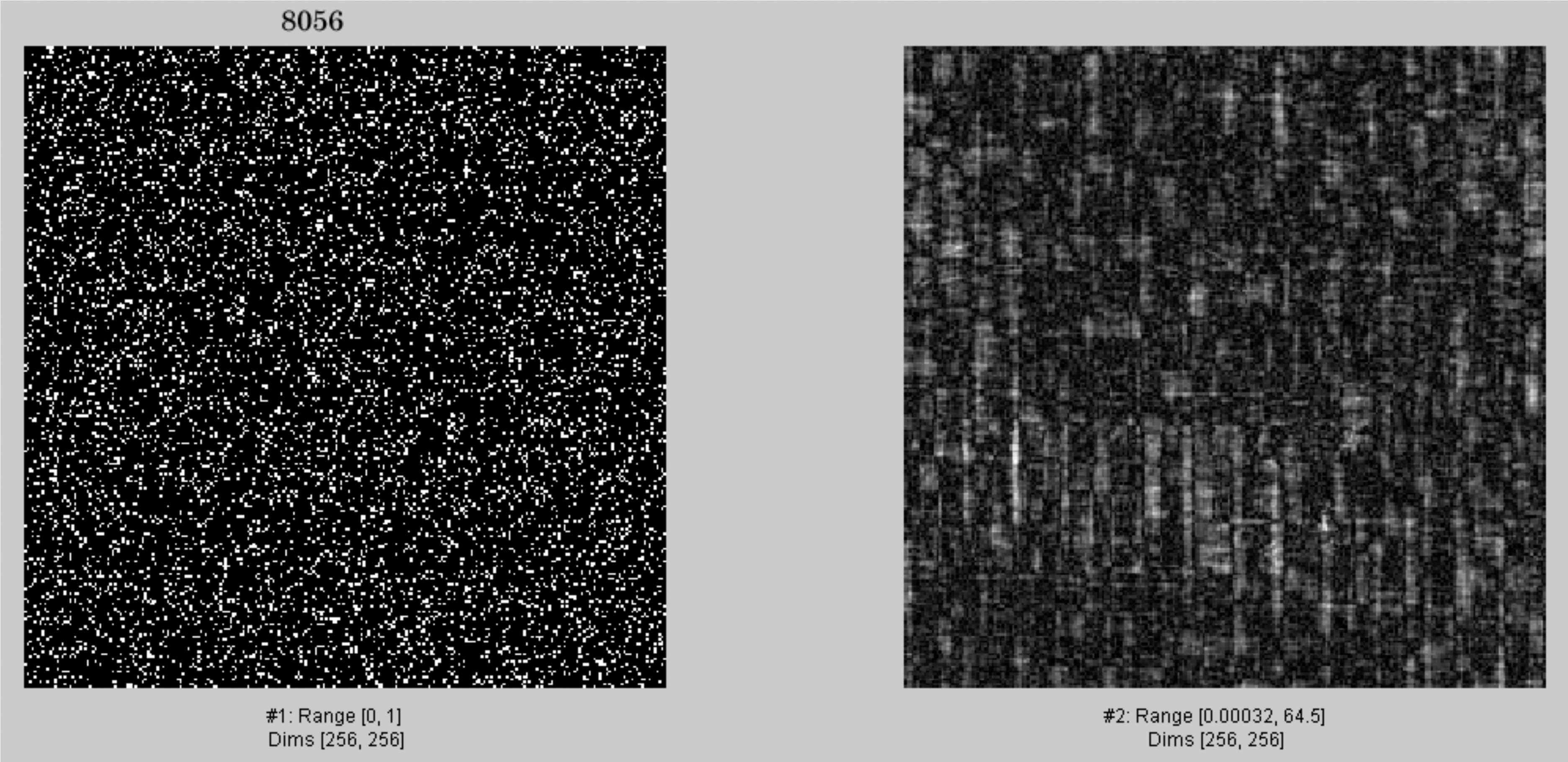
2094



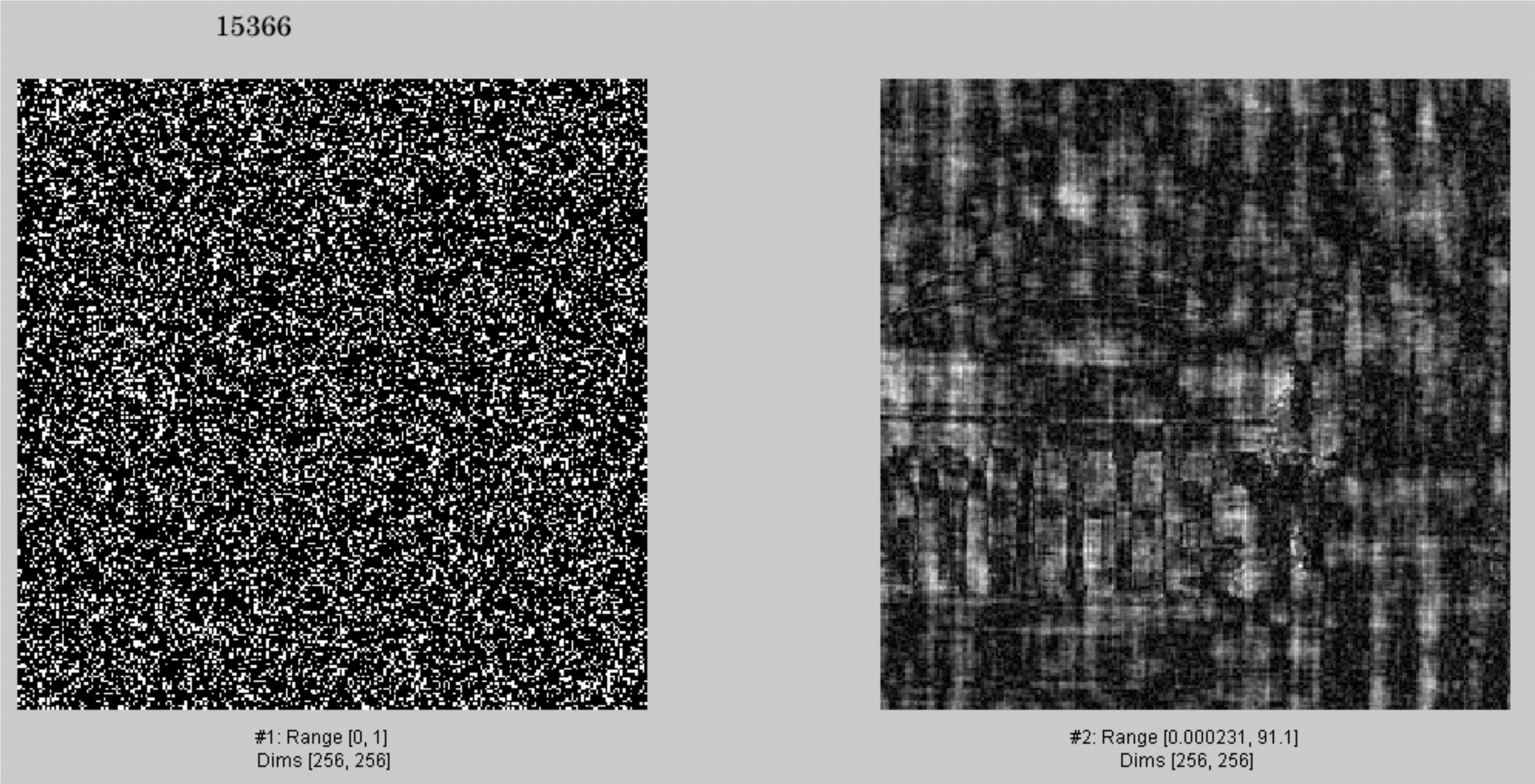
4052.



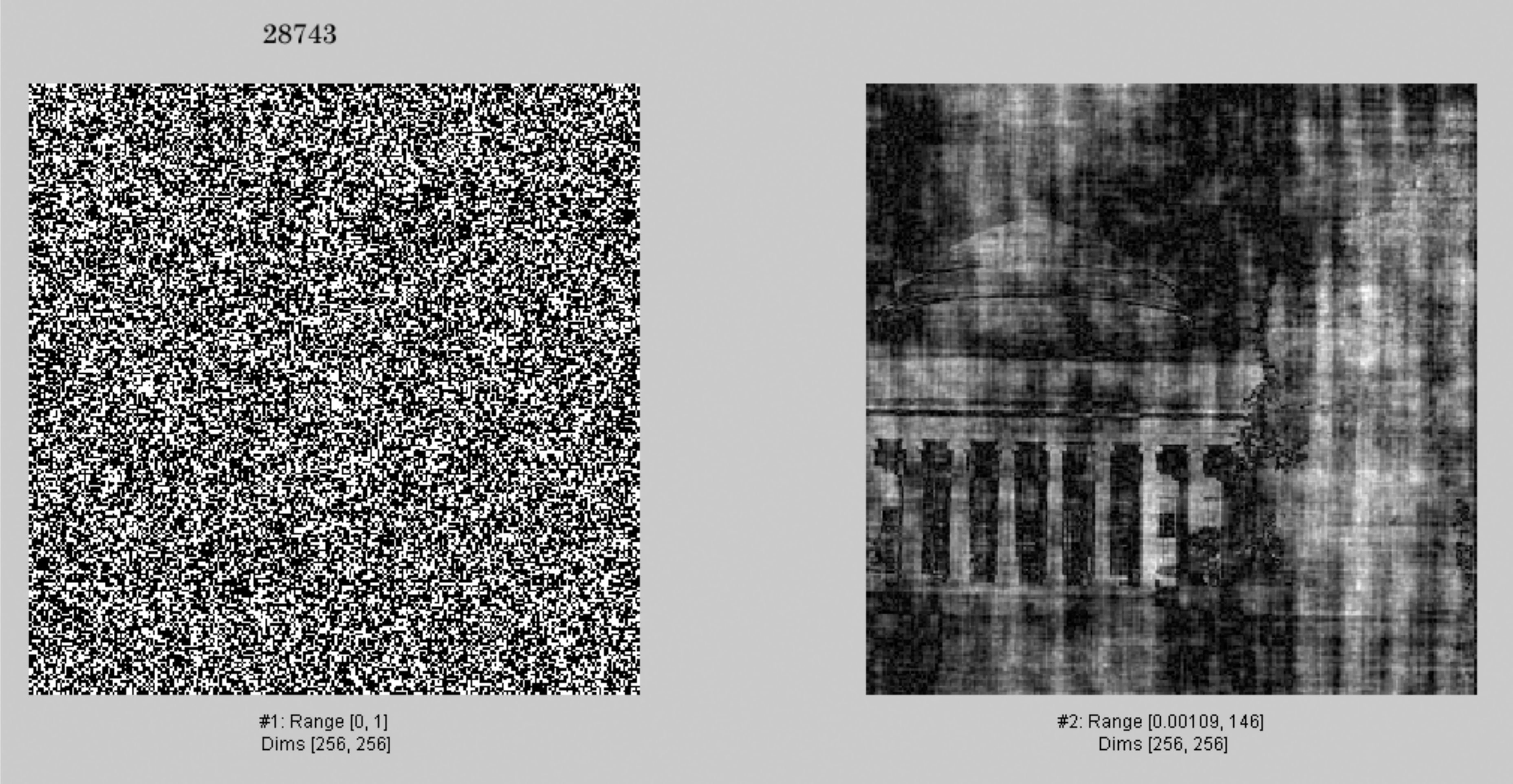
8056.



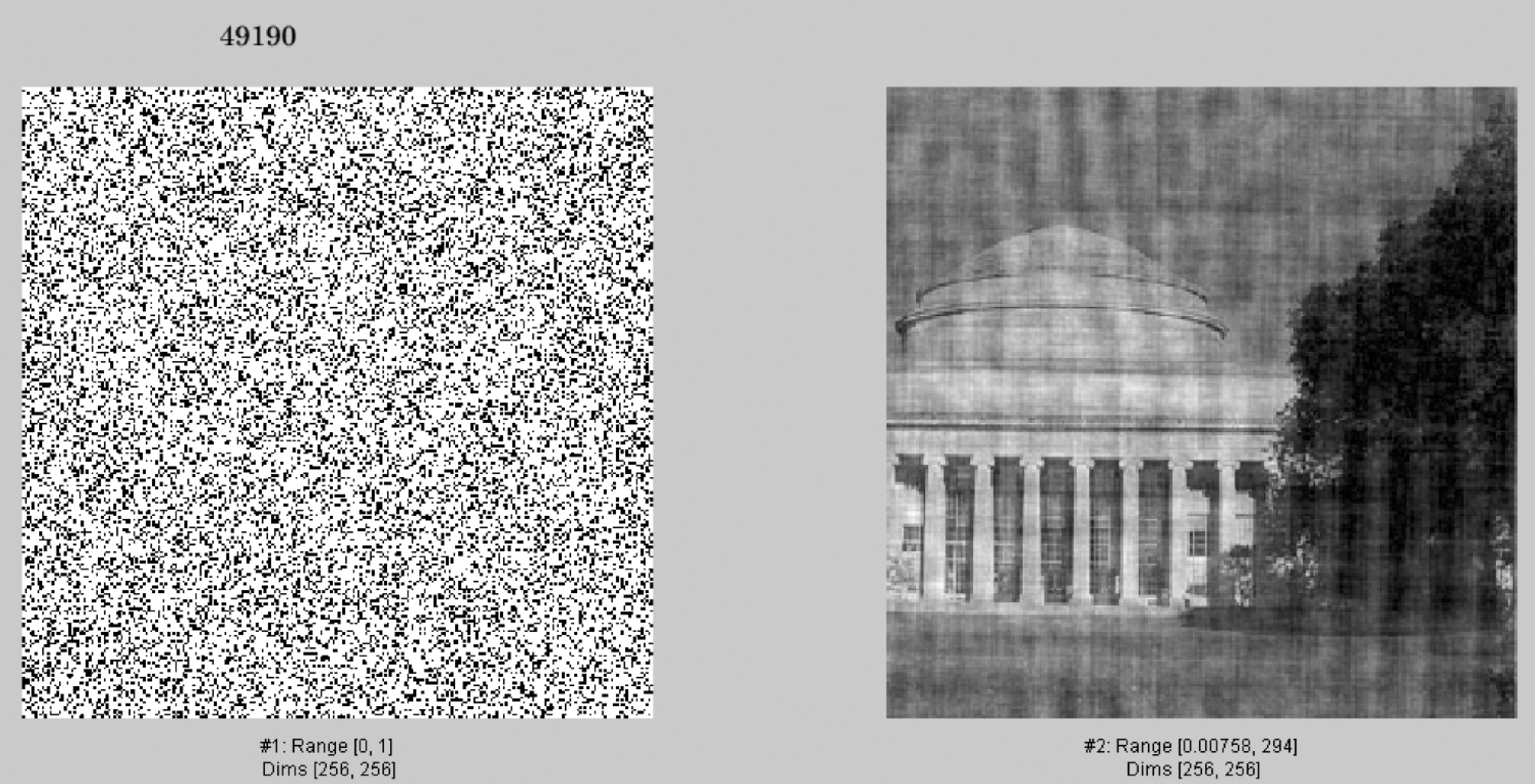
15366



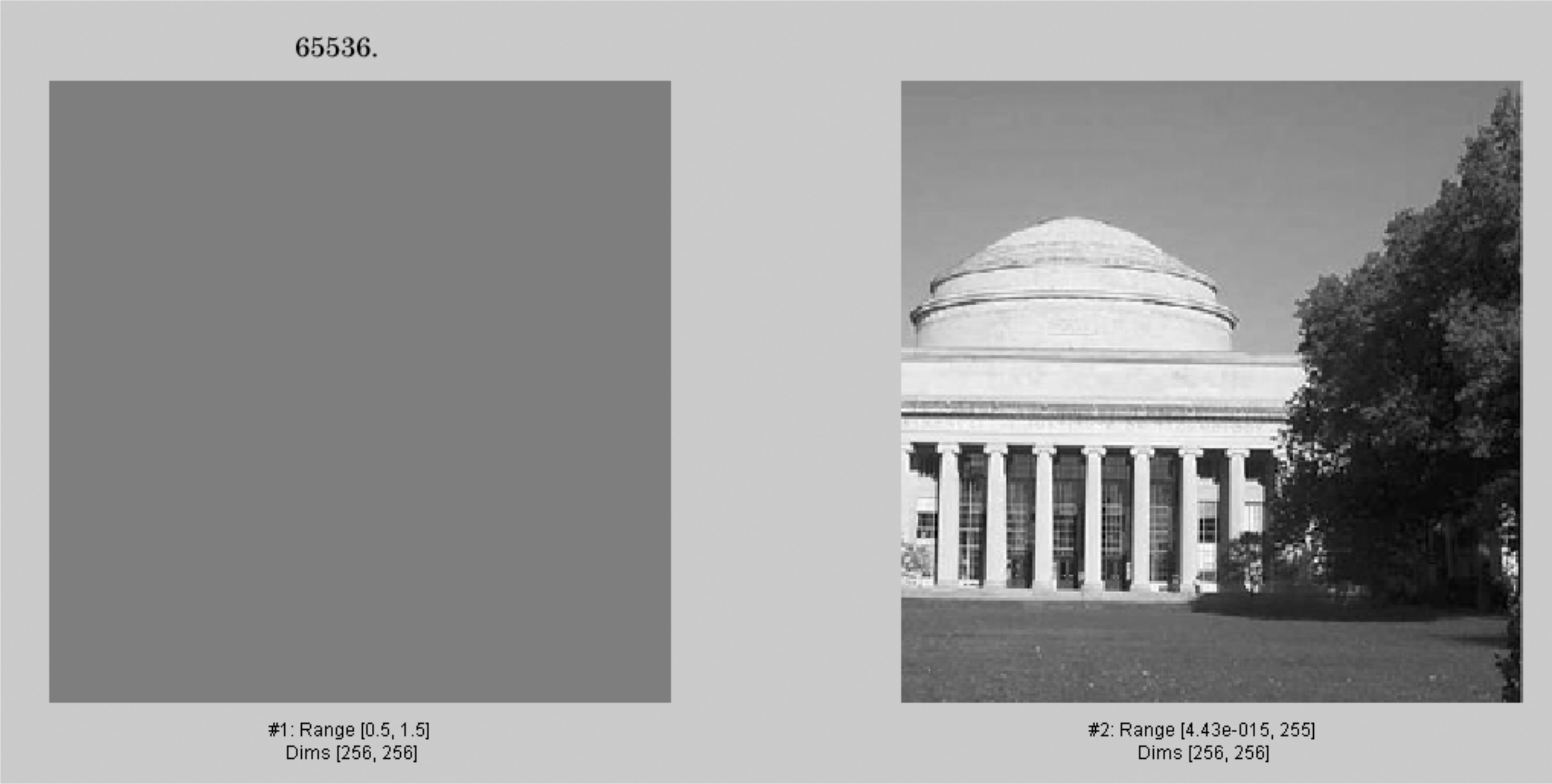
28743



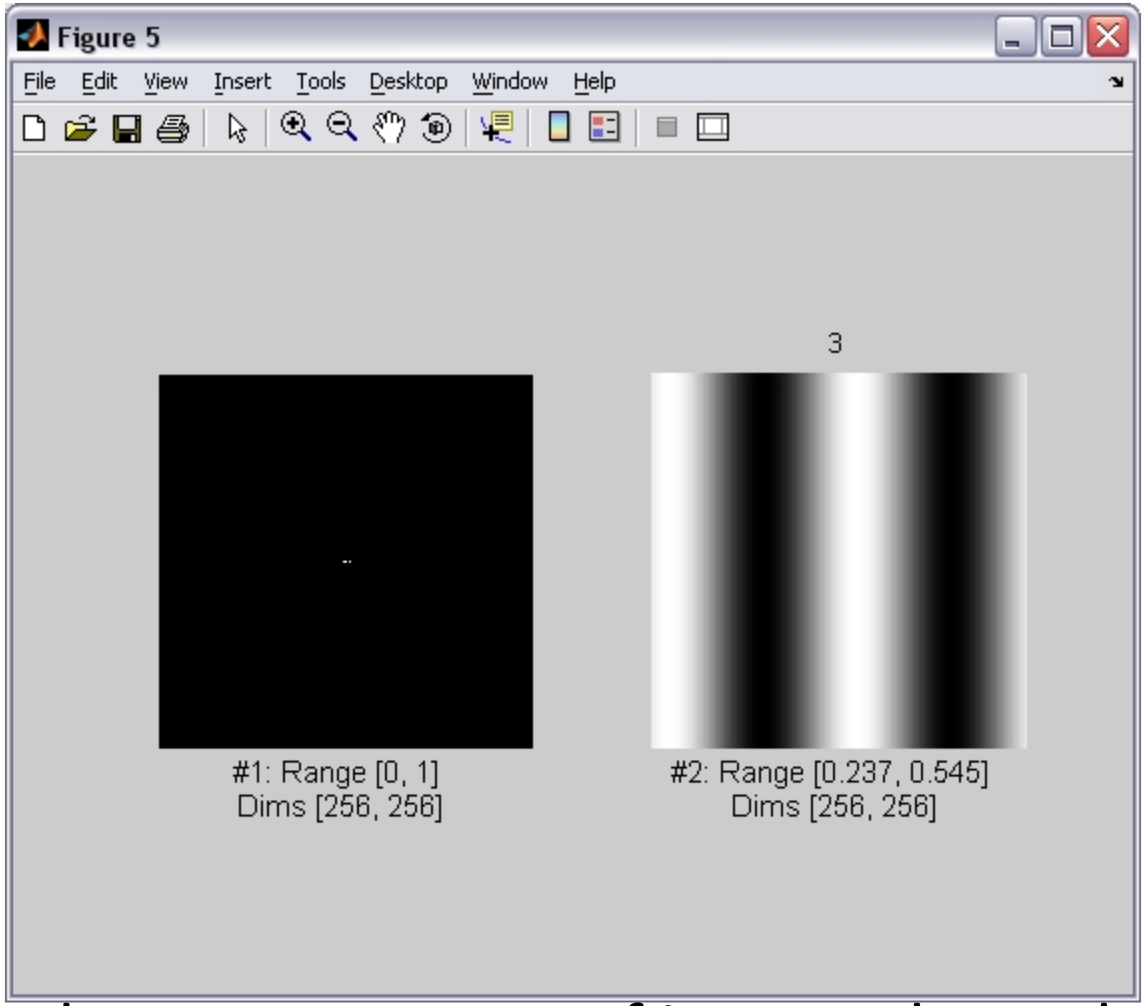
49190.



65536.


65536.
#1: Range [0.5, 1.5]
Dims [256, 256]
#2: Range [4.43e-015, 255]
Dims [256, 256]

3



Now, an analogous sequence of images, but selecting Fourier components in descending order of magnitude.

5

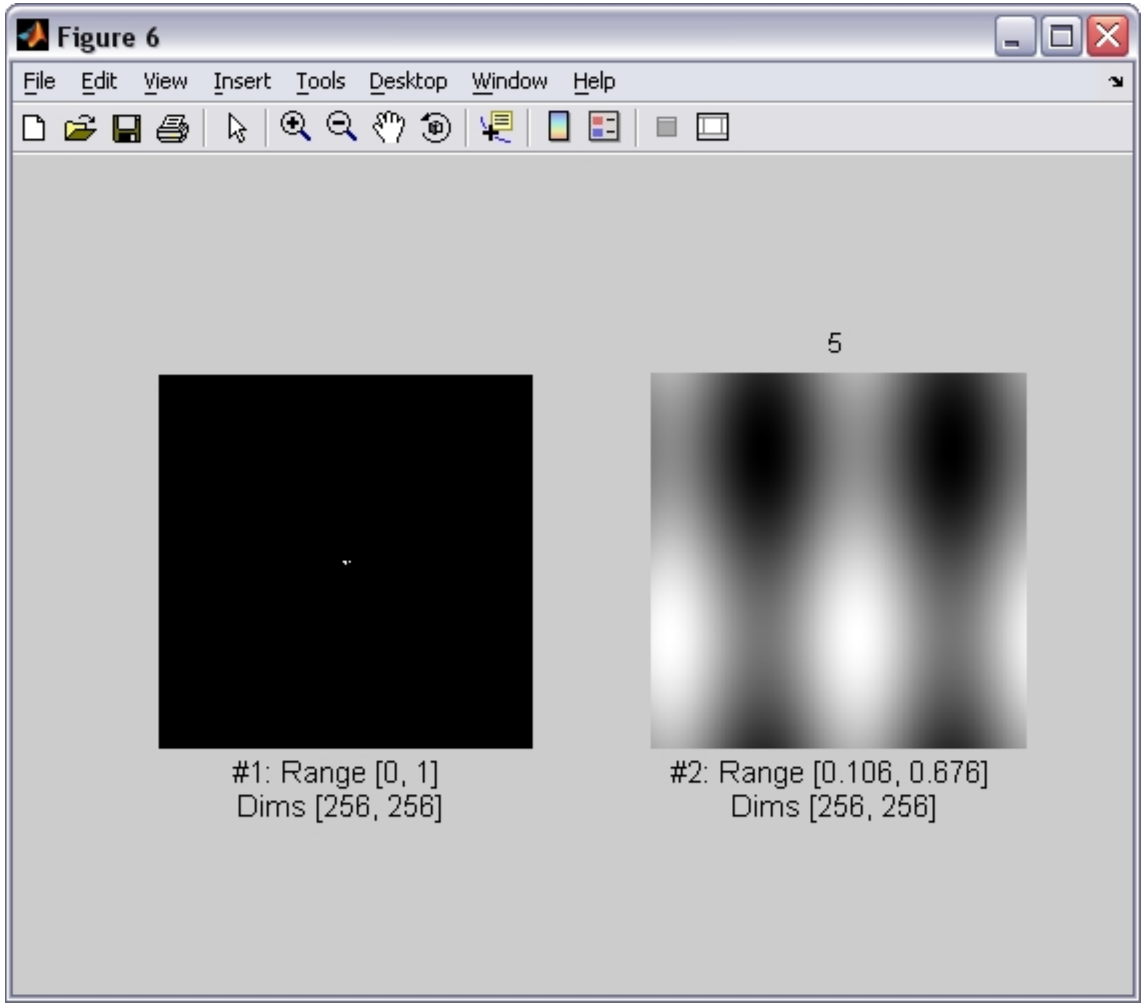


Figure 6
File
Edit
View
Insert
Tools
Desktop
Window
Help
5
#1: Range [0, 1]
Dims [256, 256]
#2: Range [0.106, 0.676]
Dims [256, 256]

Take-home reading and next class

- Multiscale pyramids

- Lecturenotes: SignalProcessing.pdf

- Linear filtering: page 111-120
- http://szeliski.org/Book/drafts/SzeliskiBook_20100903_draft.pdf