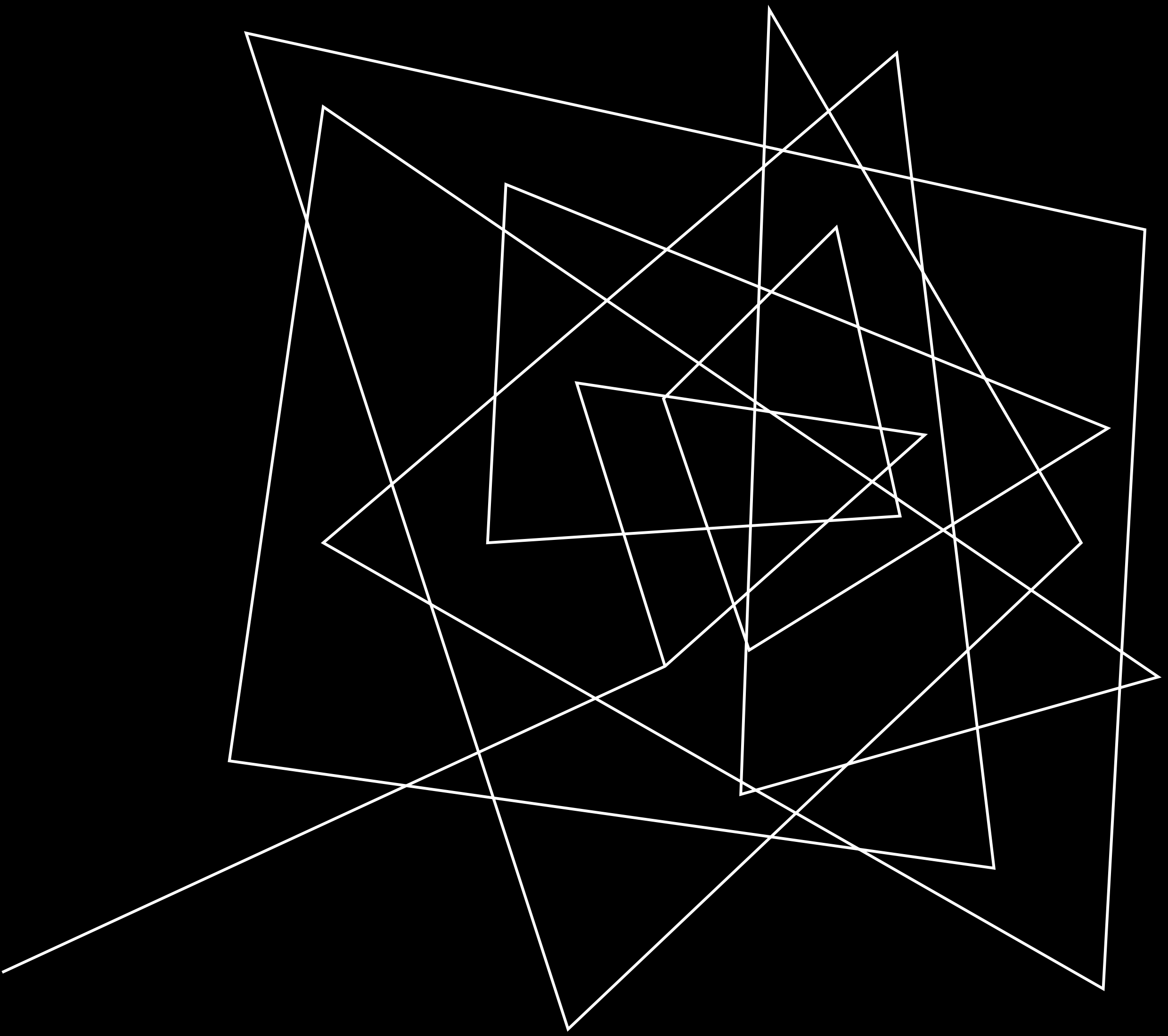
REPRESENTATION OF IMAGES IN THE FREQUENCY DOMAIN

BY

SAIFUL BARI IFTU [SDQ]
LECTURER, DEPT. OF CSE, BRAC UNIVERSITY

CONTENTS



THE FREQUENCY DOMAIN

SIGNAL (DATA)

A **Signal** is a function that **conveys information**. It represents how a quantity (e.g., brightness, sound intensity, or temperature) changes over time, space, or another independent variable. Signals can represent physical phenomena, like sound, light, temperature, or even abstract data like stock prices or sensor readings.

Types of Signals

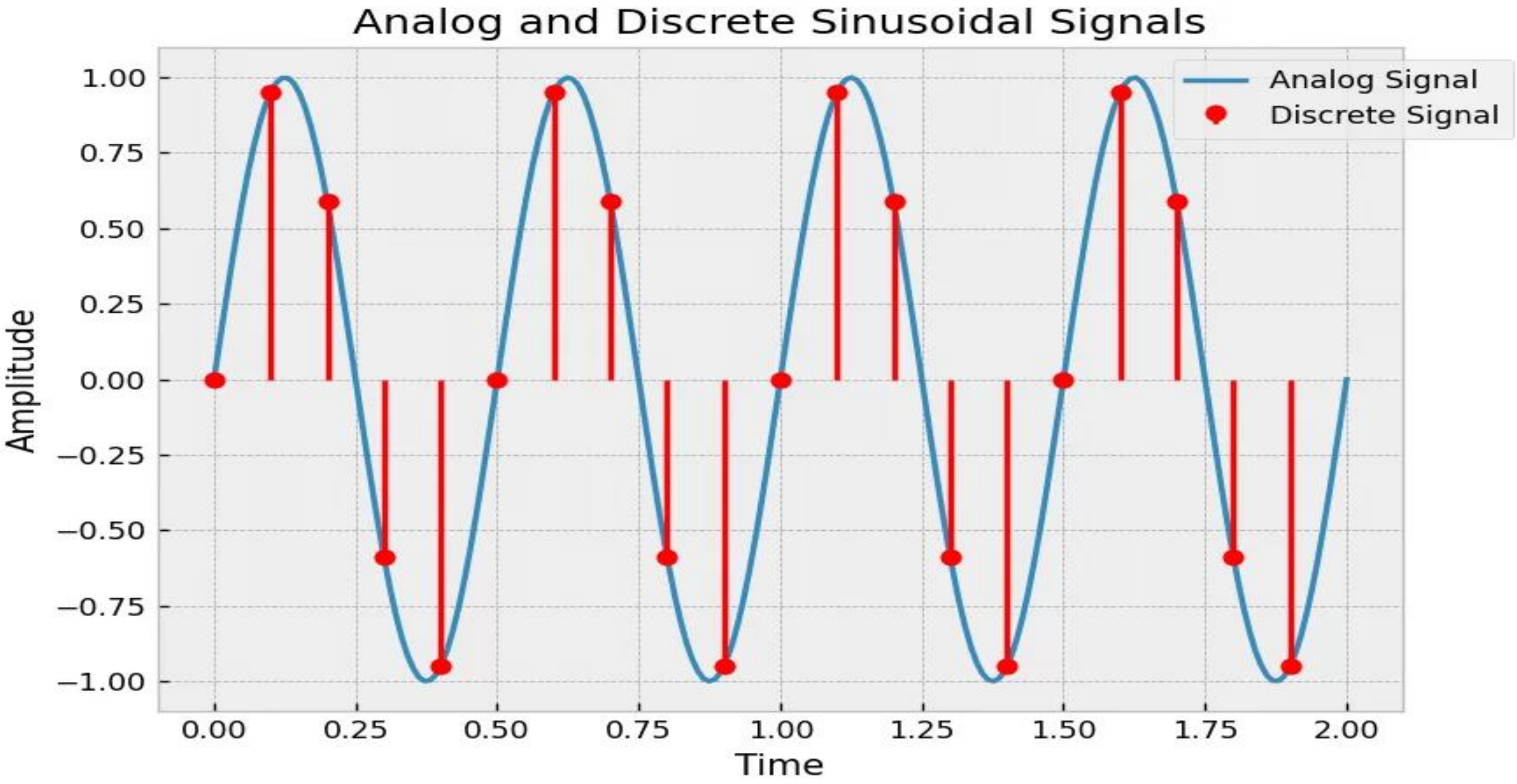
Continuous Signals: Defined for all values of the independent variable. Example: A sine wave, temperature changes over time.

Discrete Signals: Defined only at specific points (usually sampled from a continuous signal). Example: Digital audio (a series of amplitude samples).

1D Signals: Depend on a single independent variable (e.g., time). Example: Audio signals.

Multi-Dimensional Signals: Depend on multiple independent variables (e.g., x and y in an image). Example: Images (brightness at each pixel).

CONTINUOUS (ANALOG) VS DISCRETE SIGNALS



WHAT IS A DOMAIN?

The **Domain** is that **independent variable** or the "dimension" in which a signal is represented or analyzed, for example, **Time** or **Space**. A signal is defined by its values and the domain in which it's analyzed.

Common Domains

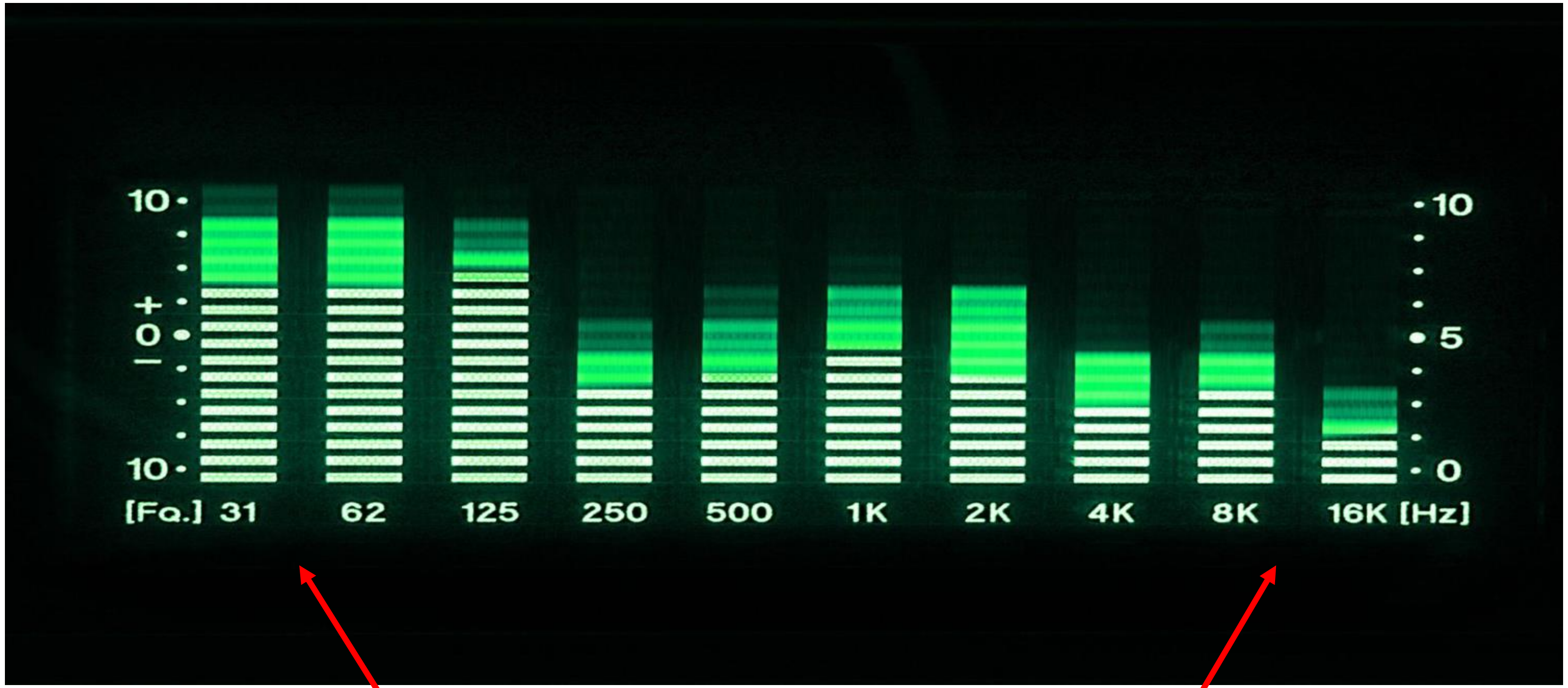
Time Domain or Spatial Domain: The natural representation of a signal. Example: An audio waveform shows amplitude over time. An image shows brightness at spatial coordinates along vertical and horizontal axes.

Frequency Domain: Represents the signal in terms of its frequency components rather than time or space. Example: The Fourier Transform converts a signal from time/spatial to frequency domain.

Different domains highlight different aspects of a signal. The time/spatial domain is best for understanding ***when*** or ***where*** something happens. The frequency domain shows ***what*** components make up the signal.

MULTIPLE REPRESENTATIONS OF A SIGNAL

Have you seen **Equalizers**? On a music player? Basically, It separates the song into its constituent frequencies, showing how much bass, treble, or mid-range exists at any moment.

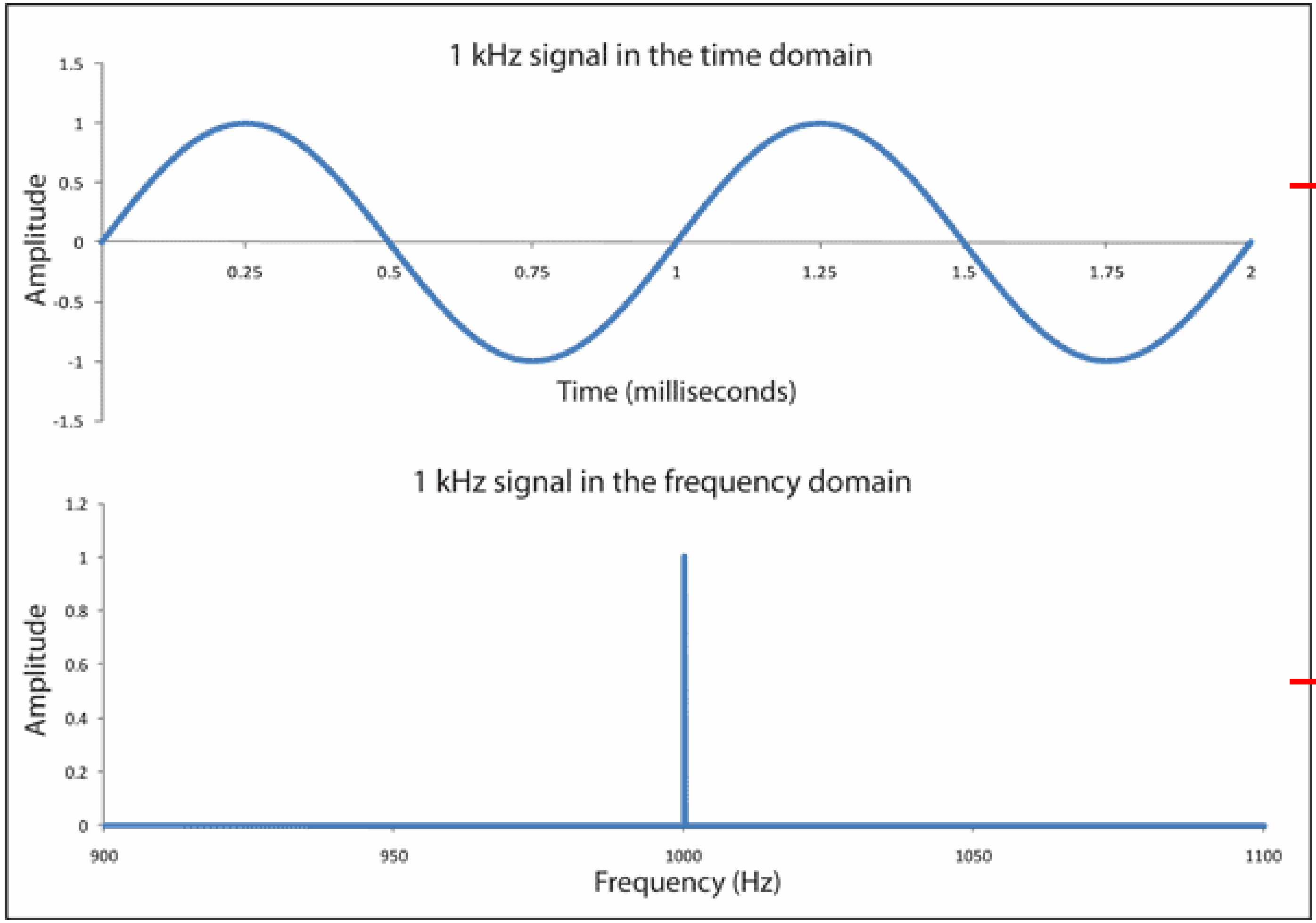


Low-Frequency Components of the Song

High-Frequency Components of the Song

MULTIPLE REPRESENTATIONS OF A SIGNAL

So, it's possible to divide a signal into multiple components, each of which represents a different frequency. Let's take a look at a simple sinusoid.

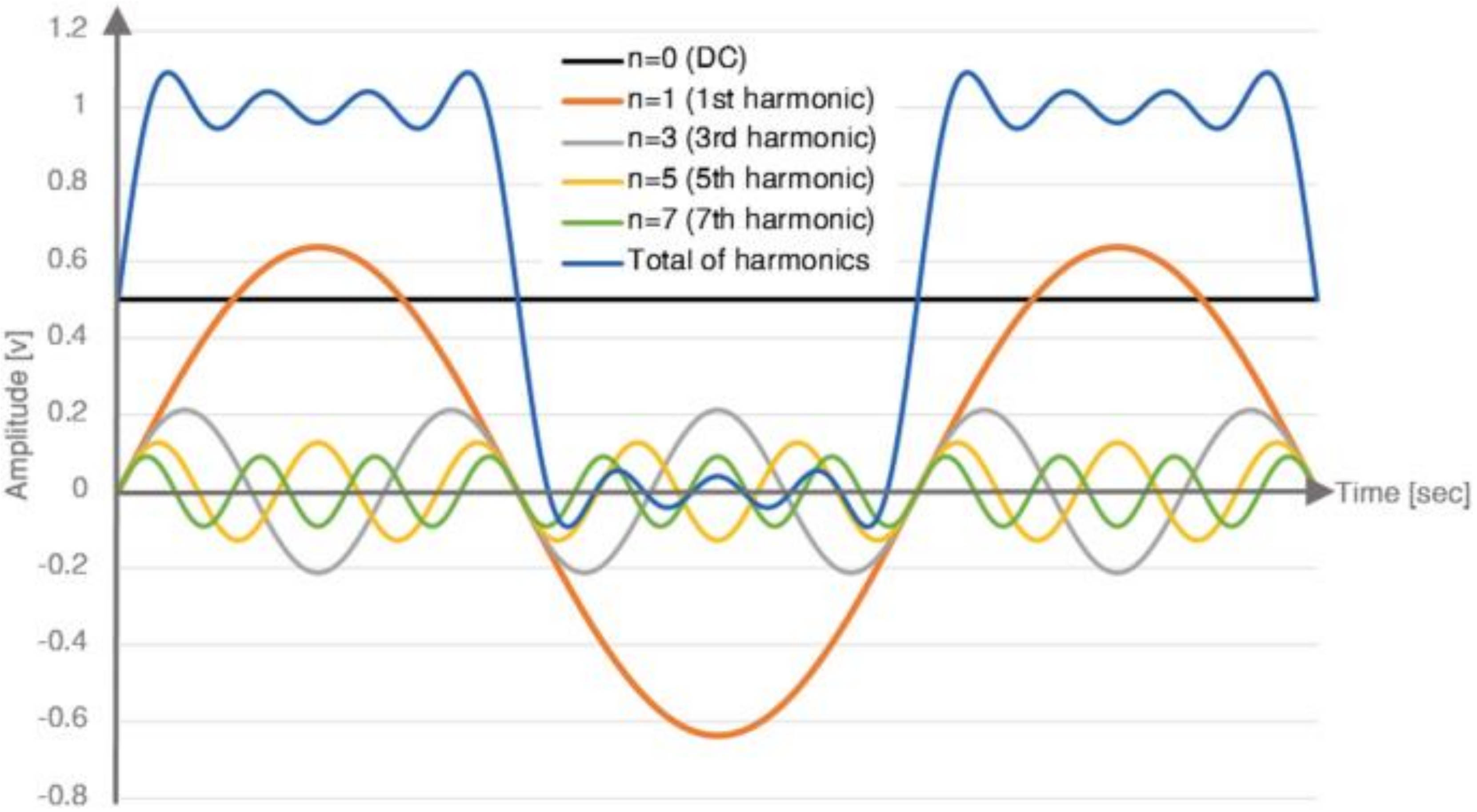


A sine wave oscillates up and down over time, and we can see its amplitude at every moment.

This same sine wave is represented as a single spike at its frequency (e.g., 1000 Hz if it oscillates 1000 times per second).

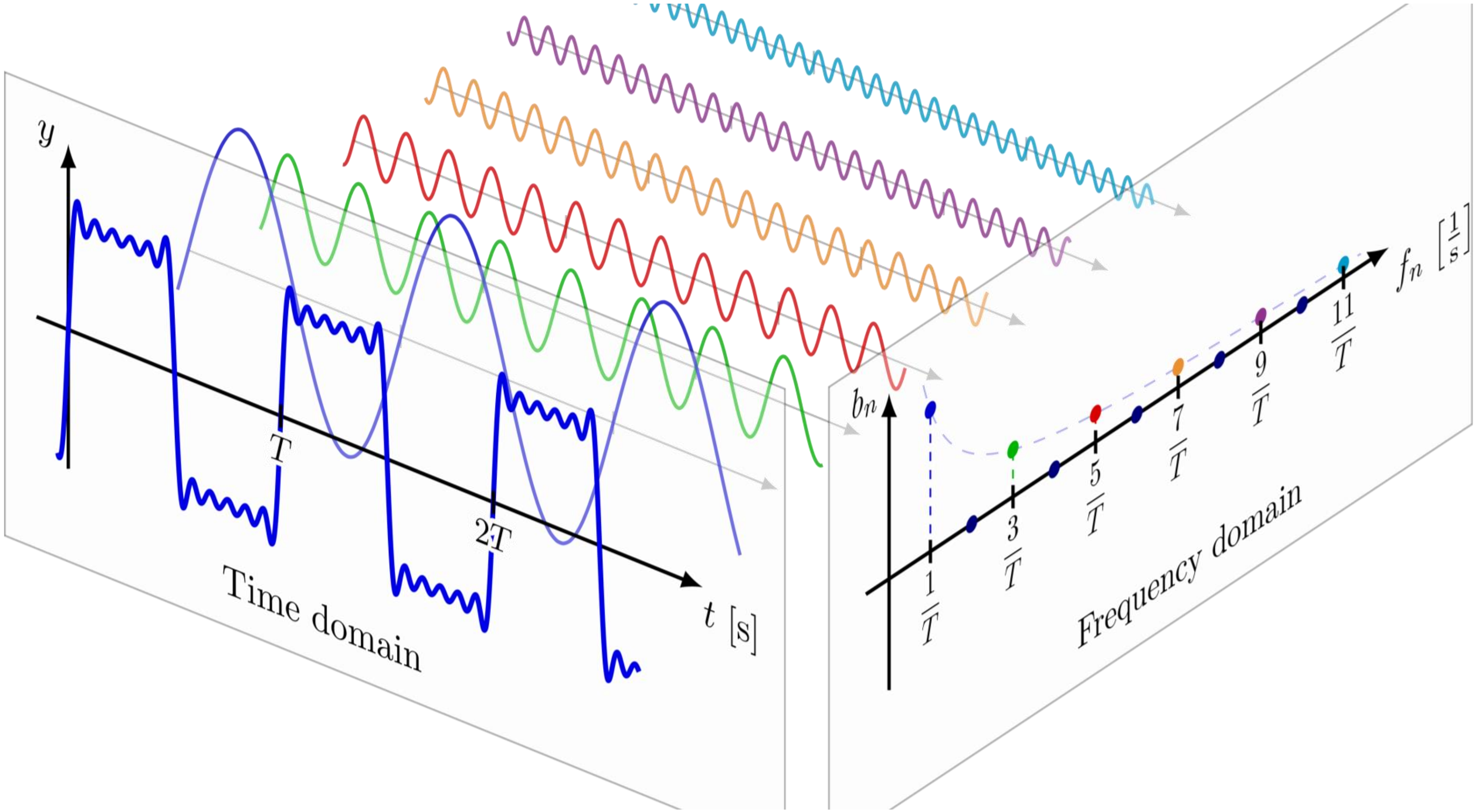
MULTIPLE REPRESENTATIONS OF A SIGNAL

The blue line in the figure below represents a signal that is the summation of multiple sine waves of different frequencies as shown in the image, creating a more complex waveform in the time domain.



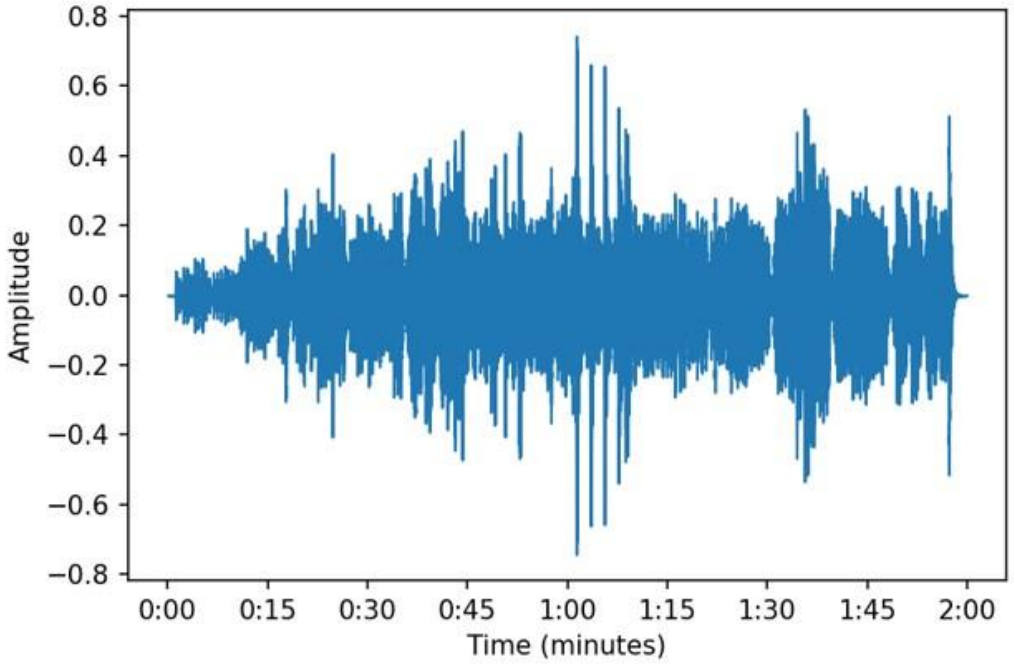
MULTIPLE REPRESENTATIONS OF A SIGNAL

In the frequency domain, this appears as multiple spikes, one for each frequency component.

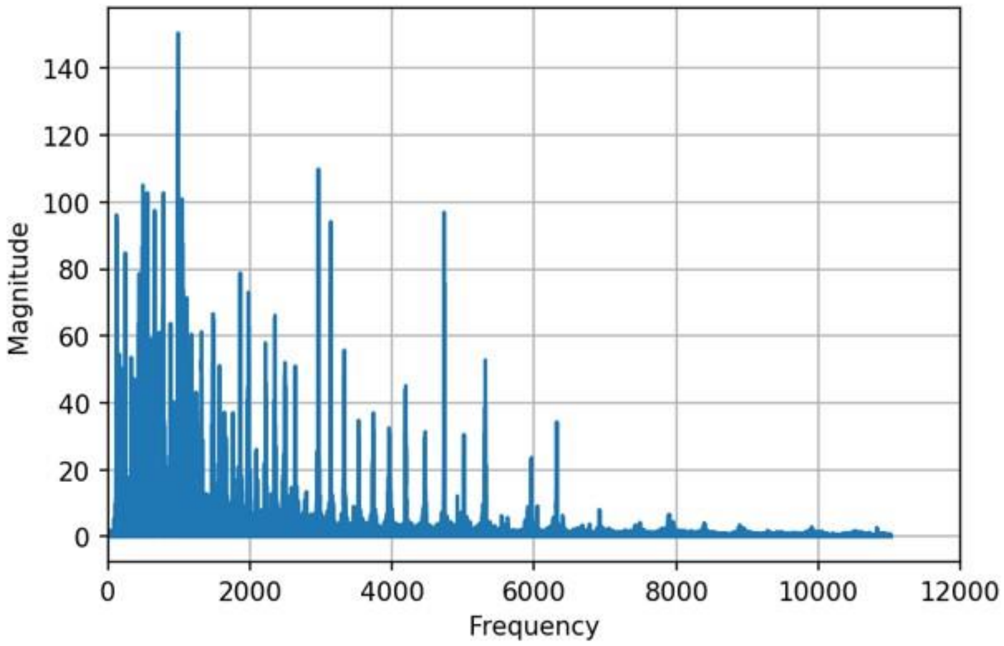


MULTIPLE REPRESENTATIONS OF A SIGNAL

Most real-world signals, like images or sound, are made up of a combination of many frequencies. So, their frequency domain representation often shows components spanning a significant portion of the frequency spectrum.



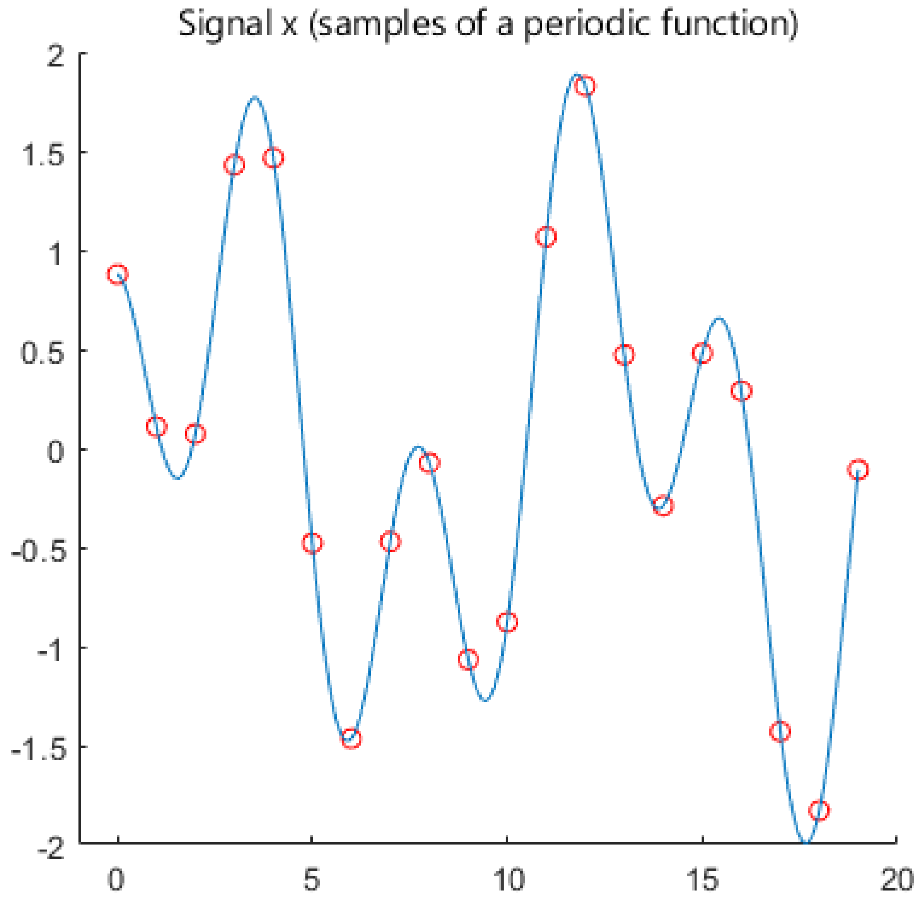
Time Domain Representation of an Audio Signal

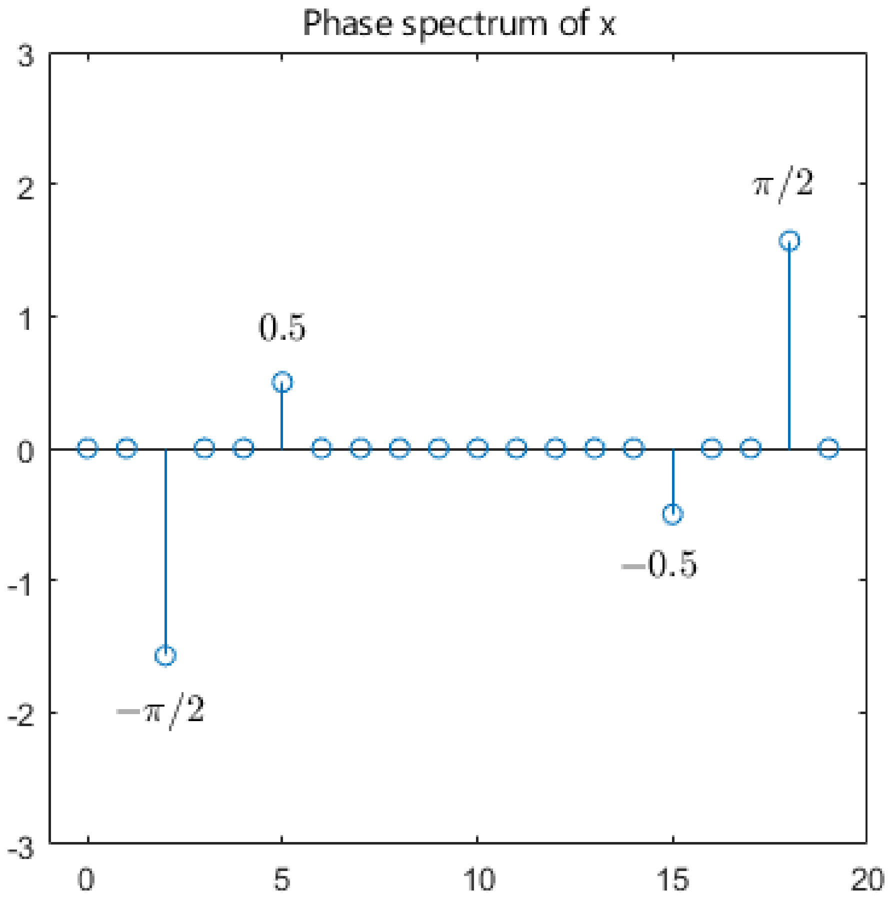


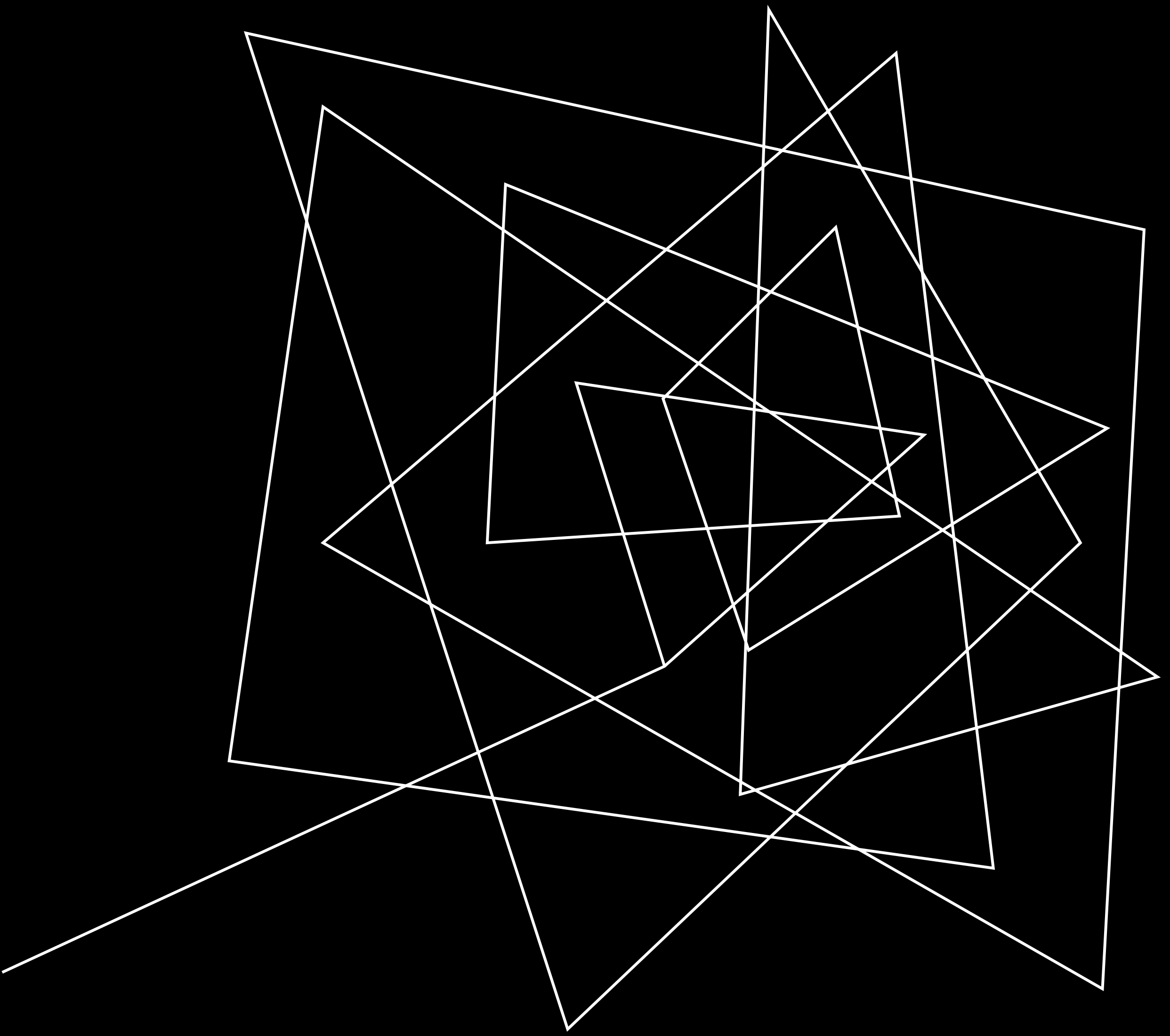
Frequency Domain Representation of the Audio Signal

MAGNITUDE & PHASE SPECTRUMS

However, the frequency domain representations we have seen so far carry information on the **Magnitude** for each **Frequency**, but not the **Phase**. These are called **Magnitude Spectrums.** Similarly, there is a **Phase Spectrum** that carries information on the **Phase** for each **Frequency.**







FOURIER SERIES
&
FOURIER TRANSFORM

THE FOURIER THEOREM

As we saw, we can decompose complex signals into multiple sinusoids with different frequencies if the signal is a summation of the same. But is this generalizable? Can we represent all signals in the frequency domain? Even the aperiodic ones?

The answer is **Yes**. The idea that all signals can be represented in the frequency domain is grounded in fundamental mathematical principles, particularly the **Fourier Theorem**.

Fourier Theorem

"Any signal, whether continuous or discrete, periodic or aperiodic, can be represented as a sum of sinusoidal components (sines and cosines) with different frequencies."

These sinusoidal components form a **"basis"** for the frequency domain.

Why Sinusoids?

Sinusoids are the natural building blocks of periodic functions and appear in solutions to many physical systems (like vibrating strings or electrical circuits). They are **mathematically complete**, meaning **they can combine to represent any signal**.

FOURIER SERIES

The **Fourier Series** is a specific method for representing periodic signals as a sum of sinusoidal components (sines and cosines). These sine and cosine waves, called **Harmonics**, have frequencies that are integer multiples of the fundamental frequency of the signal. It is, basically, a special application of the Fourier Theorem, **restricted to periodic signals**.

Fourier Series works under the assumption that the signal is periodic and can be expressed as:

$$f(t) = a_0 + \sum_{n=1}^{\infty} [a_n cos(\frac{2\pi nt}{T}) + b_n sin(\frac{2\pi nt}{T})]$$

Where,

$$a_k = \frac{1}{T}\int_0^T f(t) cos(\frac{2\pi kt}{T})\, dt \qquad \& \qquad b_k = \frac{1}{T}\int_0^T f(t) sin(\frac{2\pi kt}{T})\, dt$$

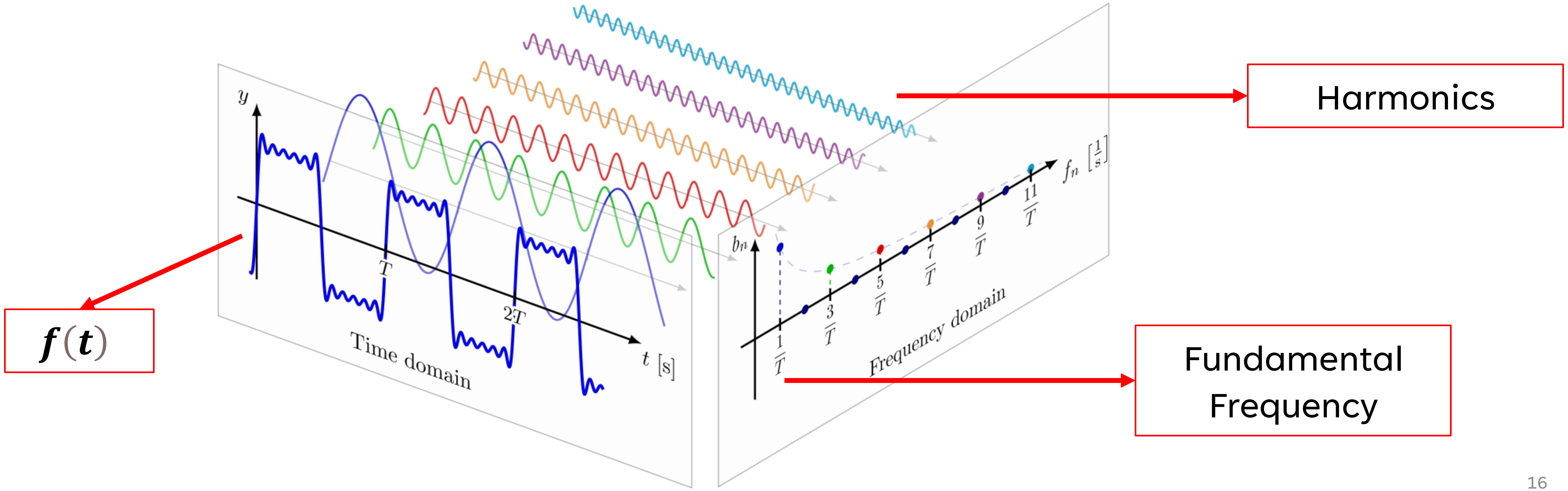
Here,
a_0: The average value (DC component) of the signal, $f(t)$.
a_n: Amplitudes of cosine components.
b_n: Amplitudes of sine components.
n: The harmonic number (integer multiple of the fundamental frequency).

FOURIER SERIES

So, this means any periodic function can be reconstructed using a weighted sum of sinusoidal components.

Components of a Fourier Series

- **Fundamental Frequency:** The signal's base frequency, $\boldsymbol{f_0} = \frac{\mathbf{1}}{\mathbf{T}}$, where $\mathbf{T}$ is the period.
- **Harmonics:** Sinusoidal components at frequencies $\boldsymbol{f_0}$, $\boldsymbol{2f_0}$, $\boldsymbol{3f_0}$, … (multiples of the fundamental frequency). These harmonics capture finer details of the signal.
- **Coefficients $\boldsymbol{a_n}$** and $\boldsymbol{b_n}$: Represent how much of each sine and cosine wave is needed to reconstruct the signal.



FOURIER TRANSFORM

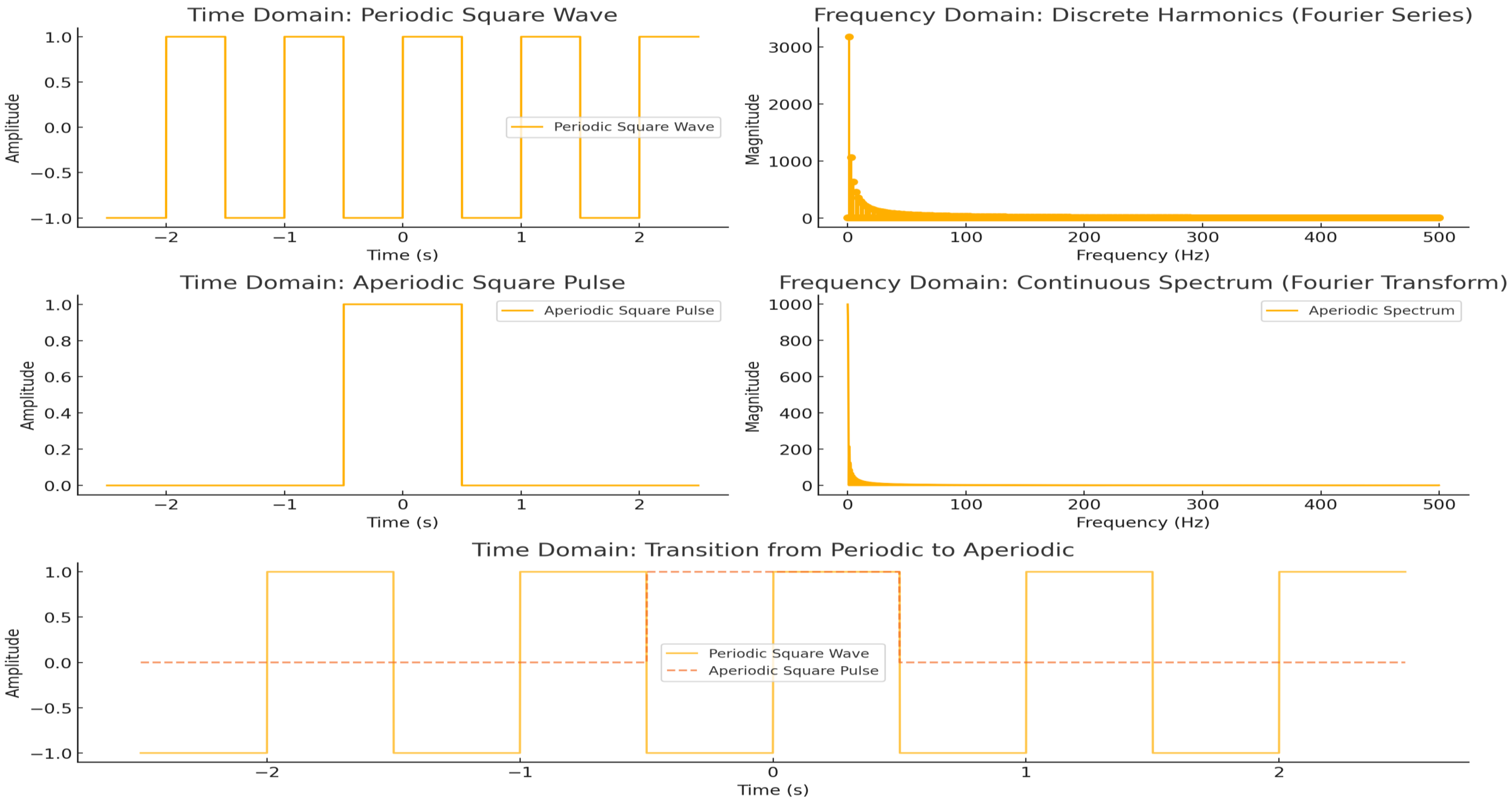
The **Fourier Transform** (FT) is a mathematical operation that transforms a signal from its original domain (typically time or space) to the frequency domain. In the frequency domain, a signal is represented as a combination of sinusoidal functions (sines and cosines) of different frequencies, amplitudes, and phases. The Fourier Transform generalizes the Fourier Series by accommodating signals **without periodicity**, resulting in a **continuous** frequency spectrum. **It states that,**

$$\boldsymbol{X(f) = \int_{-\infty}^{\infty} x(t)\mathrm{e}^{-j^{2\Pi ft}}\,\mathrm{d}t}$$

Where, $\boldsymbol{x(t)}$ is the original signal in the **time (or spatial) domain**, and $\boldsymbol{X(f)}$ is its representation in the **frequency domain**. $\boldsymbol{t}$ represents **time (or space,** for spatial signals**).** $\boldsymbol{f}$ represents **frequency.**

The frequency domain describes what frequencies are present in the signal and their magnitudes. The Fourier Transform analyzes a signal as a sum of **infinitely** many (hence, continuous) sinusoidal waves (of different **frequencies** and **amplitudes**). FT basically assumes **T** → ∞. The magnitude of $\boldsymbol{X(f)}$ tells us how much of the signal's **energy** is concentrated at a particular frequency. So, FT also gives us the Energy Distribution of the signal.

FOURIER SERIES TO FOURIER TRANSFORM



INVERSE FOURIER TRANSFORM

The **Inverse Fourier Transform** reconstructs the original signal from its frequency components. **It states that,**

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{j2\Pi ft} \, df$$

Where, $x(t)$ is the original signal in the **time (or spatial) domain**, and $X(f)$ is its representation in the **frequency domain**. t represents **time (or space,** for spatial signals**).** f represents **frequency.**

Fourier transforms are often used in audio processing (e.g., speech recognition) and image processing (e.g., texture analysis). The frequency components of the signal help reveal periodic patterns that can be crucial features for models. **FFT**, **STFT**, and **DCT** are different optimized versions of Fourier Transform that are essential in machine learning for extracting meaningful patterns, reducing noise, and providing a more compact and informative representation of data. **Inverse Fourier Transform** is necessary in all these cases to get back the original, interpretable signal.

Why is the basis $e^{j2\Pi ft}$ here? Think! *{Hint: Euler's Formula}*

DISCRETE FOURIER TRANSFORM (DFT)

The **Discrete Fourier Transform (DFT)** is a mathematical technique used to analyze **finite sequences of data** and transform them from the time (or spatial) domain into the frequency domain. It's particularly useful for **Digital Signal/Image Processing** because it operates on **discrete data sampled from continuous signals/Pixel values**. Given a discrete sequence of $\mathbf{N}$ samples $\boldsymbol{f}[\boldsymbol{n}]$ (where $\boldsymbol{n}$ = 0, 1, 2, ... , N−1), the DFT is defined as:

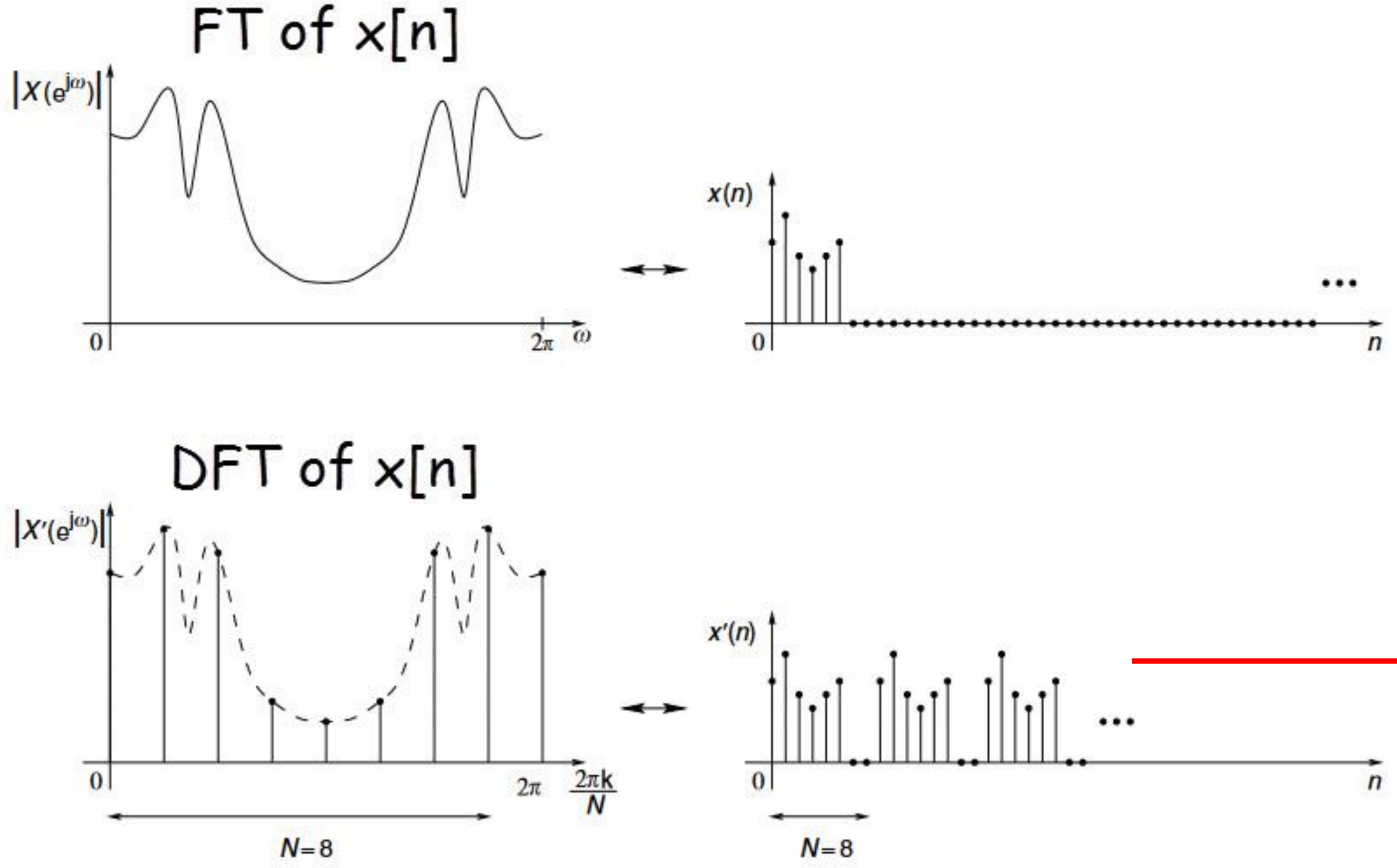
$$\boldsymbol{F}[\boldsymbol{k}] = \sum_{\boldsymbol{n}=\mathbf{0}}^{\mathbf{N}-\mathbf{1}} \boldsymbol{f}[\boldsymbol{n}]\mathbf{e}^{\frac{-\mathbf{j}2\pi \boldsymbol{k}\boldsymbol{n}}{\mathbf{N}}}, \qquad \boldsymbol{k} = \mathbf{0, 1, 2, \ldots, N\text{-}1}$$

Where, $\boldsymbol{f}[\boldsymbol{n}]$ is the original signal in the **time (or spatial) domain**, and $\boldsymbol{F}[\boldsymbol{k}]$ is The DFT of the signal $\boldsymbol{f}[\boldsymbol{n}]$, representing its frequency components. $\boldsymbol{n}$ represents the **nth** sample along the **time** axis **(or space,** for spatial signals**).** $\boldsymbol{k}$ represents the **frequency index.** N is the total number of samples.

The **Inverse DFT (IDFT)** reconstructs the original signal from its frequency components:

$$\boldsymbol{f}[\boldsymbol{n}] = \frac{\mathbf{1}}{\mathbf{N}} \sum_{\boldsymbol{k}=\mathbf{0}}^{\mathbf{N}-\mathbf{1}} \boldsymbol{F}[\boldsymbol{k}]\mathbf{e}^{\frac{\mathbf{j}2\pi \boldsymbol{k}\boldsymbol{n}}{\mathbf{N}}}$$

DISCRETE FOURIER TRANSFORM (DFT)



Why Repeating? This repetition in the frequency domain is the consequence of **sampling** in the time domain. And it depends on the sampling rate!

Source: https://www.ee-diary.com/2023/05/what-is-discrete-fourier-transformdft.html

FREQUENCY IN IMAGES

FREQUENCY IN SPATIAL DOMAIN

Let us consider a simple Sinusoid in the Time Domain:

Time, the independent variable, is the dimension in which the signal is being represented.

$$\mathrm{A}\bullet\sin(2\pi \boldsymbol{f}_0 \boldsymbol{t} + \boldsymbol{\theta})$$

Amplitude Frequency Phase

An alternative representation in the frequency domain is valid if All three measurable quantities can be represented in the magnitude and phase spectrums. For example, the **magnitude spectrum** of the sinusoid here is,

Amplitude

Frequency

$$\mathrm{X}(\boldsymbol{f}) = \frac{\mathrm{A}}{2}\bullet \boldsymbol{\delta}(\boldsymbol{f} - \boldsymbol{f}_0) + \frac{\mathrm{A}}{2}\bullet \boldsymbol{\delta}(\boldsymbol{f} + \boldsymbol{f}_0)$$

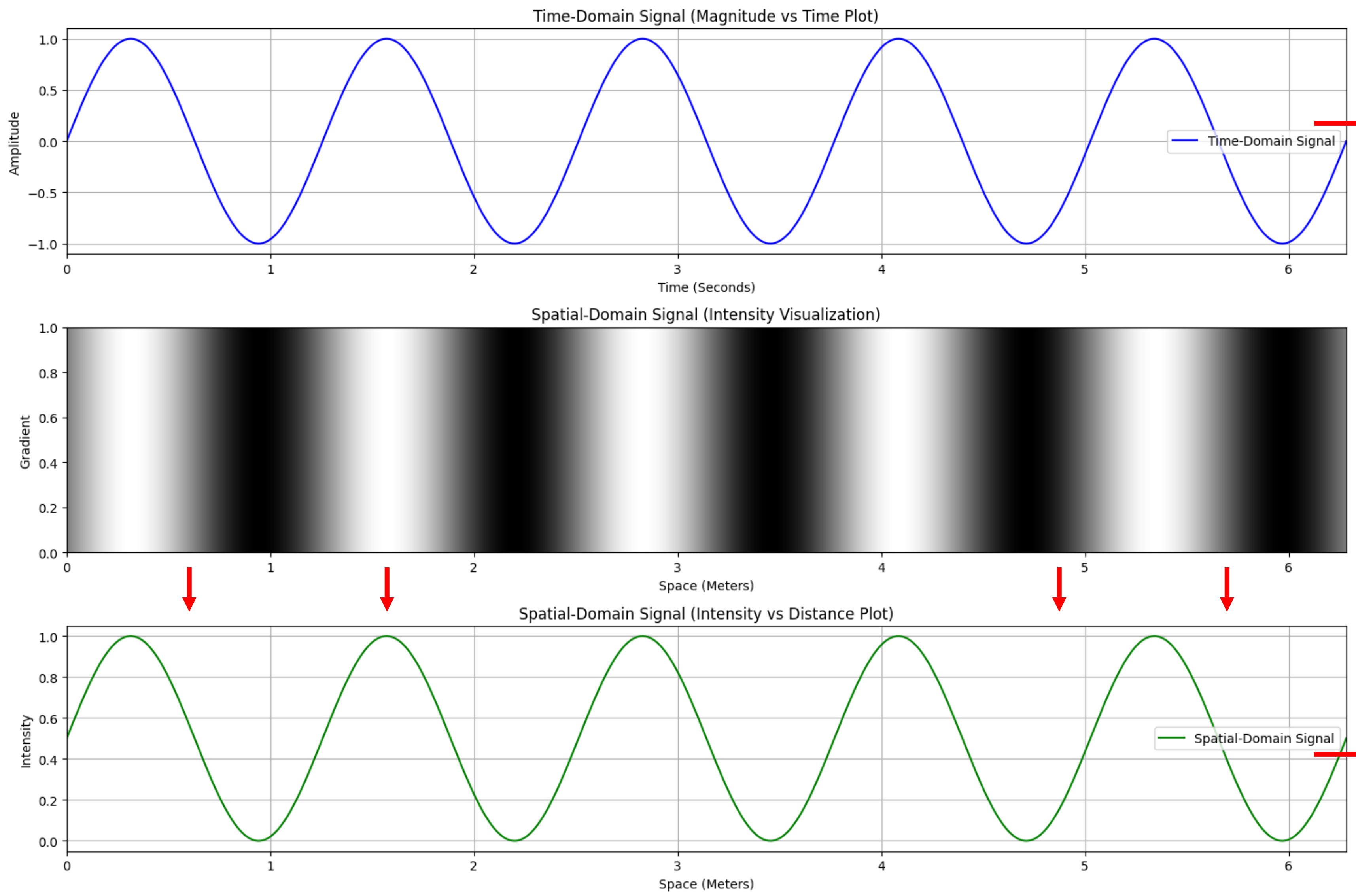
It contains both the **amplitude** and the **frequency**.

Sinusoid in Spatial Domain (1D): There is no difference in the structure of the sine function. The only difference is the independent variable, which is **space/distance/displacement** in this case.

$$\mathrm{A}\bullet\sin(2\pi \boldsymbol{f}_0 \boldsymbol{x} + \boldsymbol{\theta})$$

Amplitude Frequency Phase

FREQUENCY IN SPATIAL DOMAIN

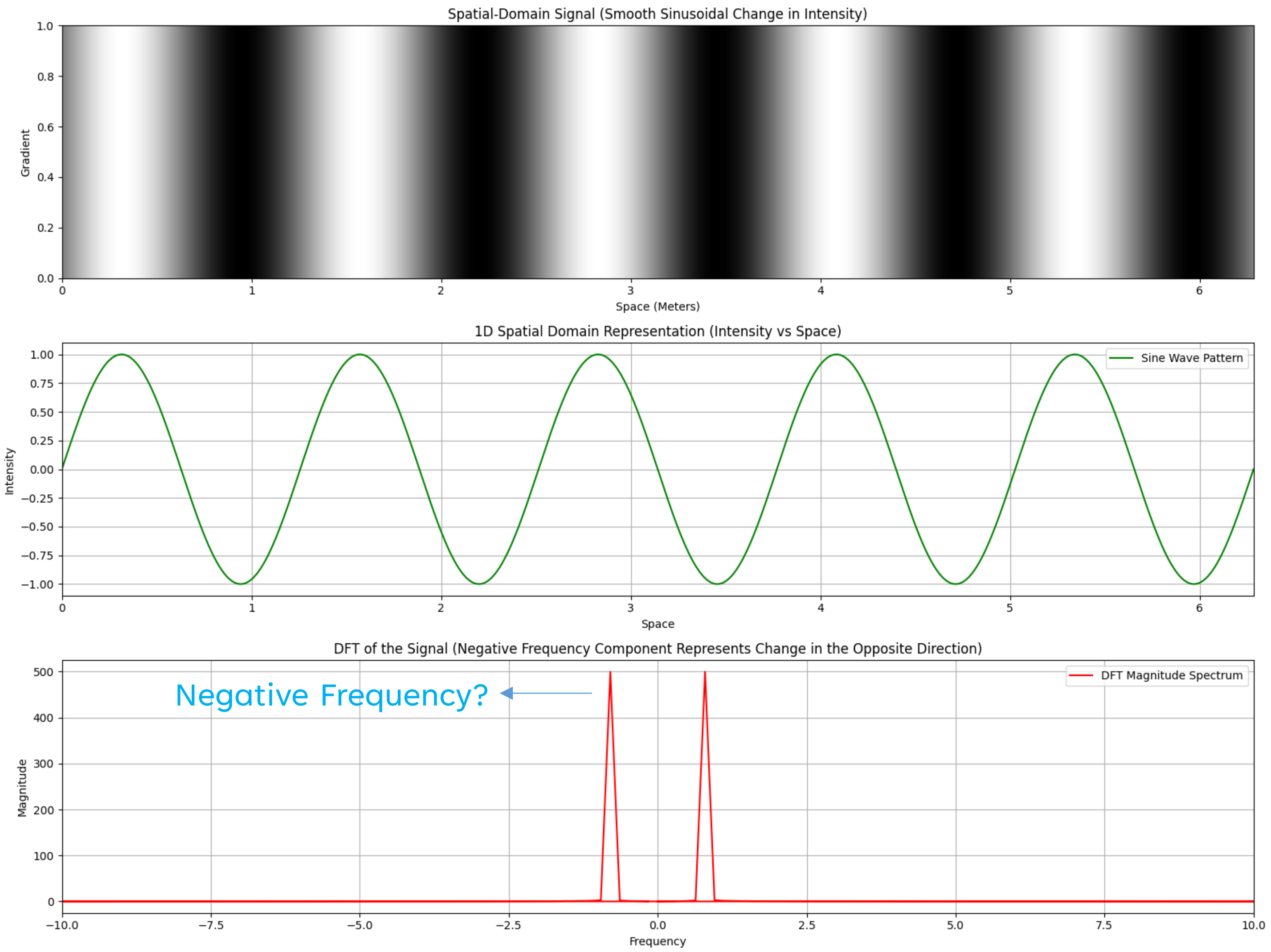


Time (Seconds) plotted along the X-axis. So, **Frequency = 1/T [Sec^{-1}]**

Clearly, Intensity in different locations (**spatial coordinates**) can also be represented using sinusoids. Hence, spatial domain signals have **frequency** components.

Distance (Meters) along the X-axis. So, **Frequency = 1/(Period in Space) [Meter^{-1}]**

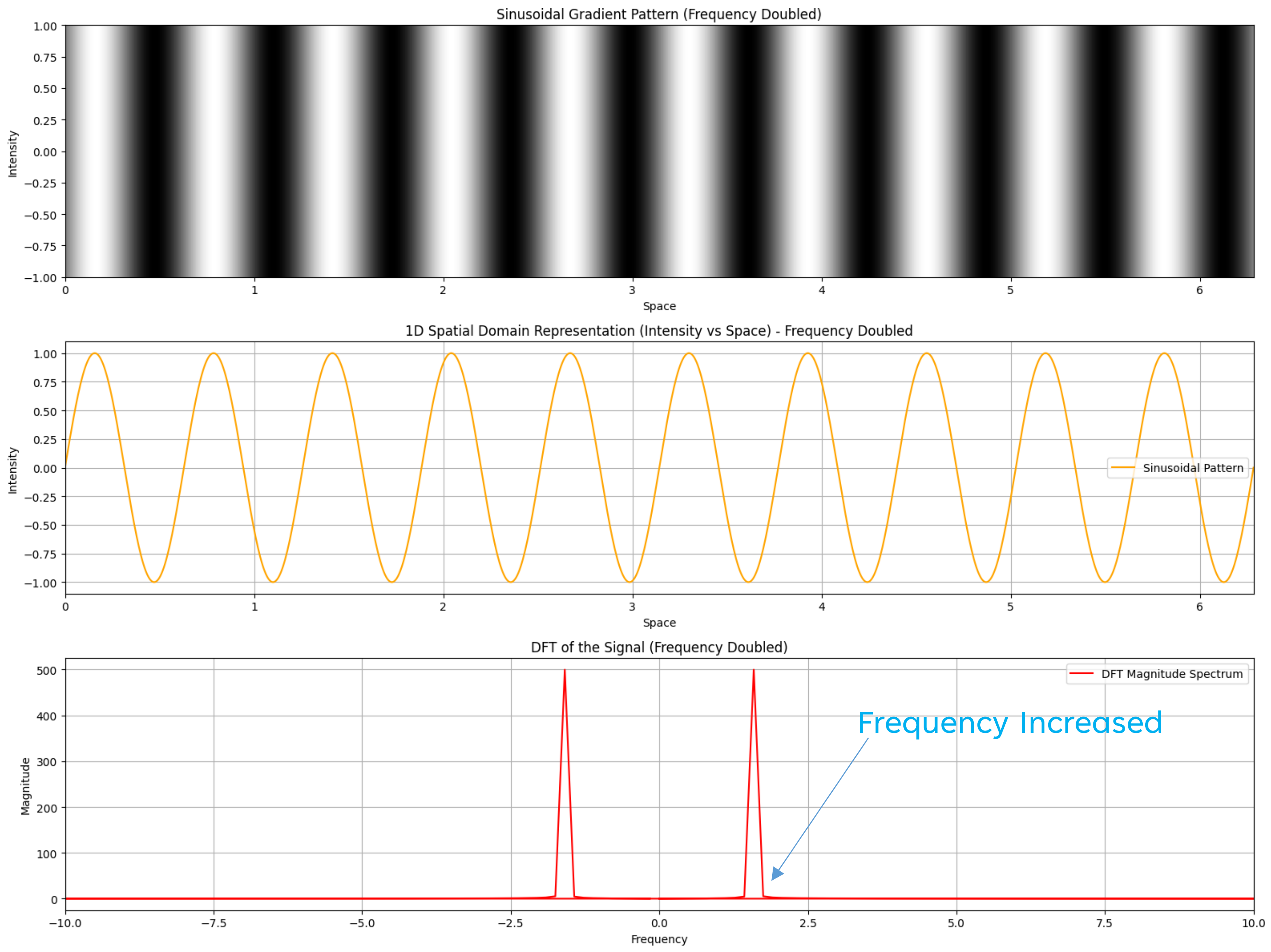
FREQUENCY IN SPATIAL DOMAIN: 1D



The image here shows the change in intensity along one direction following a sinusoidal pattern. Expectedly, in the DFT, there is a spike at around **0.8,** which is the frequency of the original sinusoid.

The spike at around **-0.8** is due to due to the **complex conjugate symmetry** of the transform for real-valued signals (remember the basis?). These negative frequencies represent the same information as the positive ones but in the opposite direction. *For real signals, only the **positive frequencies** are typically considered for analysis.*

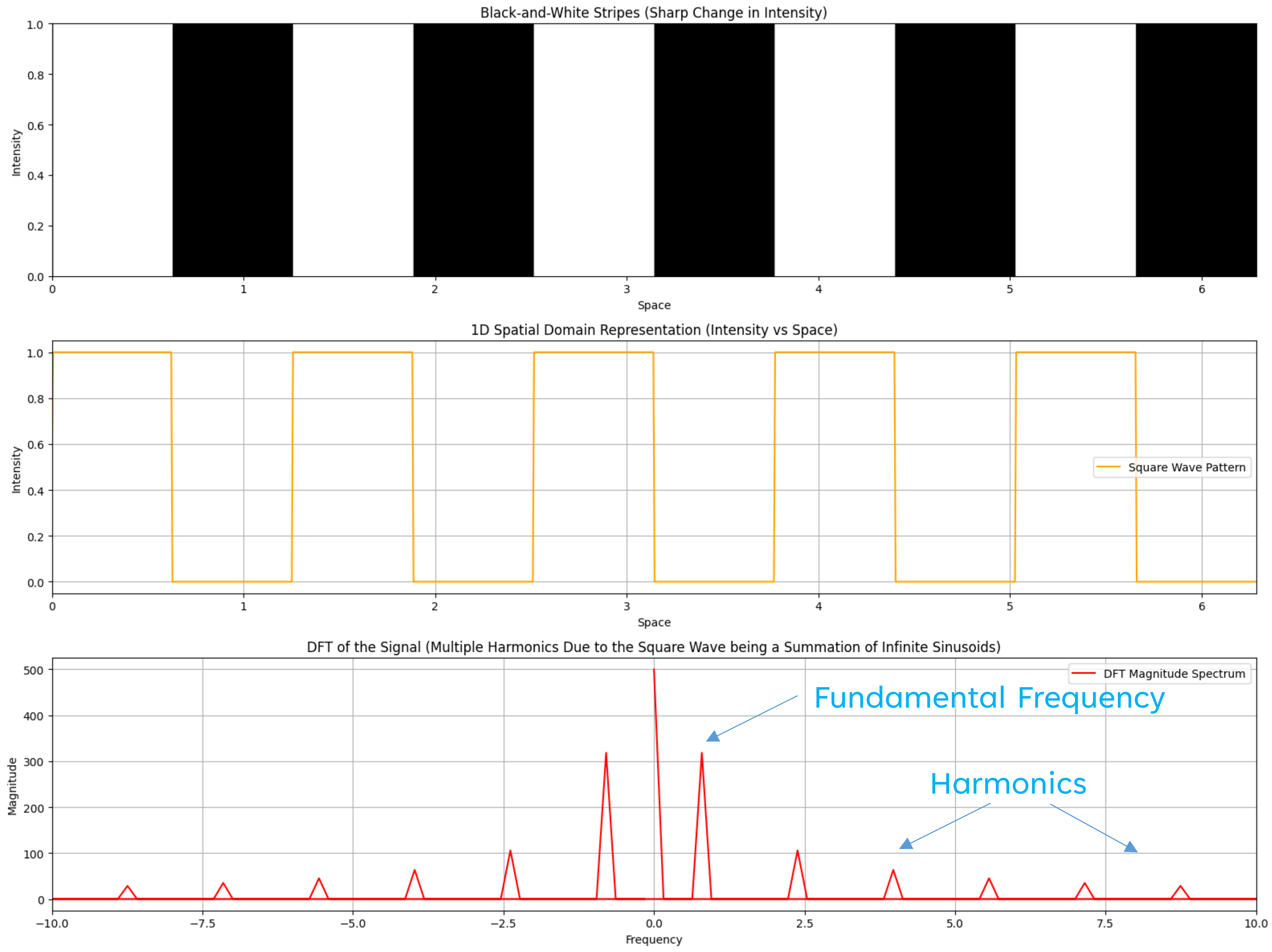
FREQUENCY IN SPATIAL DOMAIN: 1D



The image here shows the aftermath of doubling the rate of change in intensity along the spatial axis. Expectedly, in the DFT, the spike is now at around **1.6,** which is double the frequency of the original sinusoid.

Notice that while the signal is compressing in the time domain, it is stretching in the frequency domain (and vice versa for a lower-frequency signal).

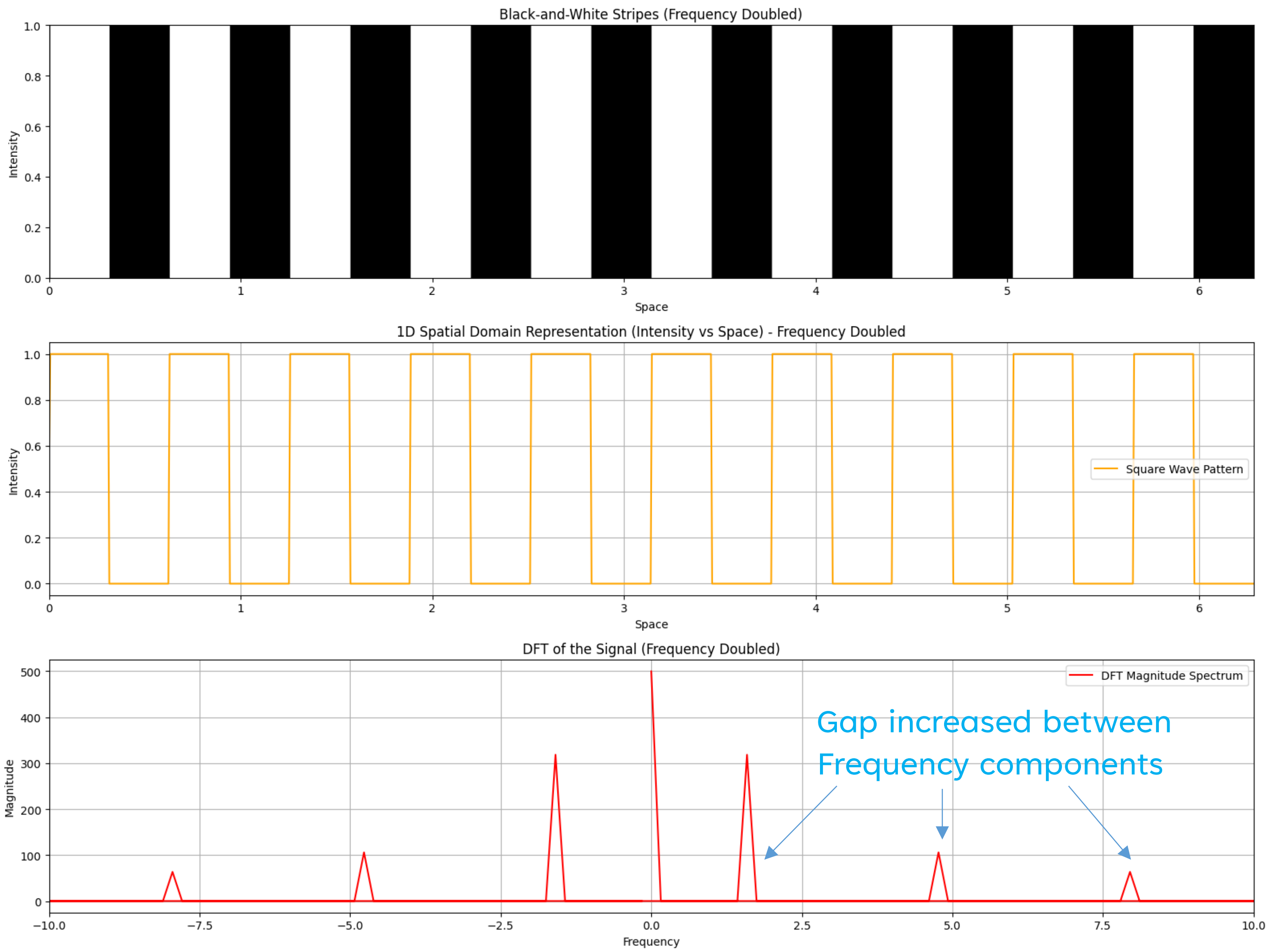
FREQUENCY IN SPATIAL DOMAIN: 1D



In this case, The change in intensity along the spatial axis does not follow sinusoidal functions. Rather, the intensity is changed abruptly from high to low periodically. The 1D spatial domain representation of this is a **square wave**.

Here, apart from the fundamental frequency $\boldsymbol{f_0}$ at around **0.8**, we can see the harmonics too at $\boldsymbol{3f_0}$, $\boldsymbol{5f_0}$, etc. This is because **a perfect square wave can be represented as a summation of infinite sinusoids.** That's why we can see the frequency components of all those sinusoids here.

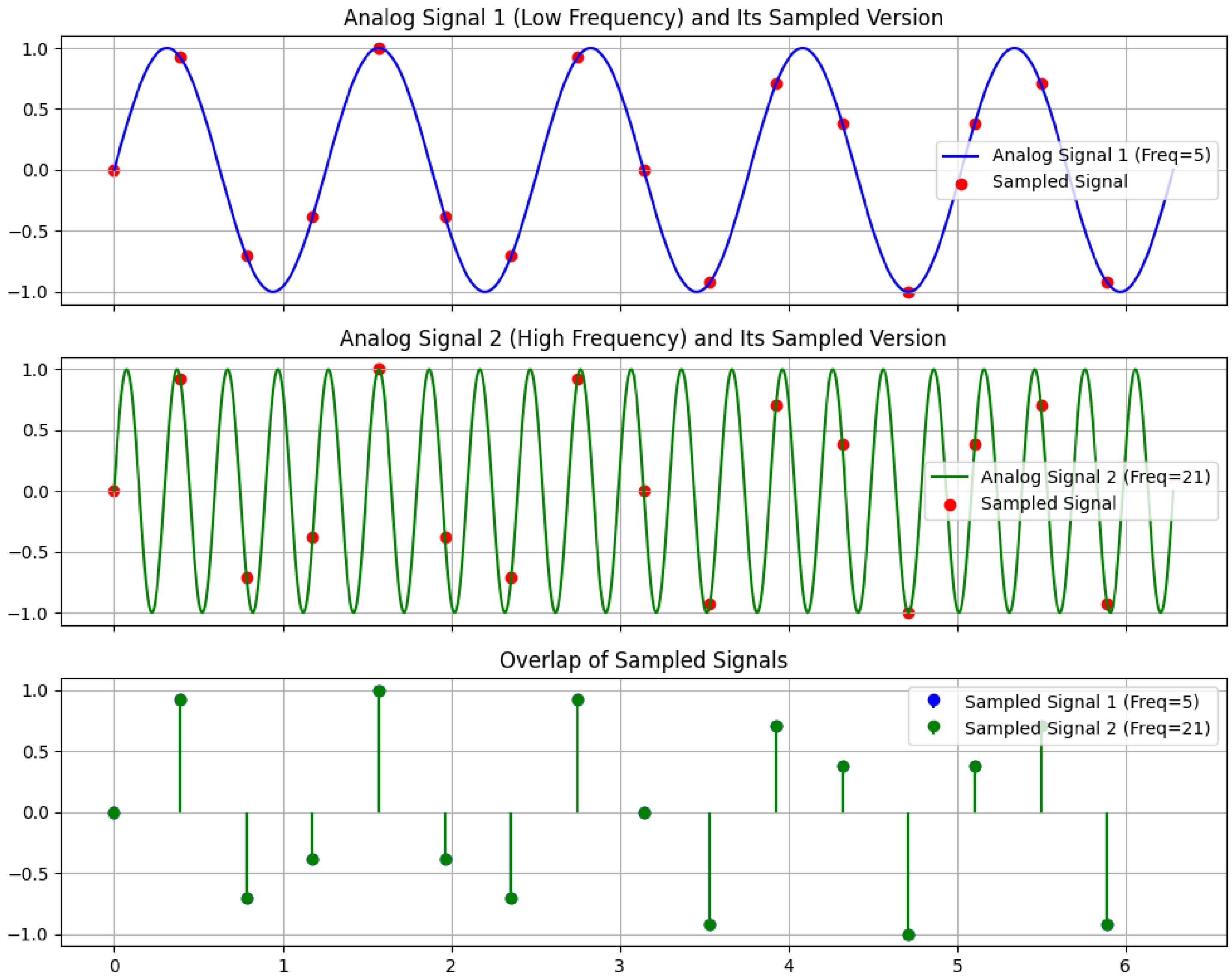
FREQUENCY IN SPATIAL DOMAIN: 1D



Once again, we can see the aftermath of doubling the rate of change in intensity along the spatial axis. Expectedly, in the DFT, the spike for the fundamental frequency is now at around **1.6,** which is double the frequency of the previous square wave.

As we can see, higher spatial frequencies correspond to closer spacing in the pattern itself, yet their Fourier Transform produces spikes that are further apart. It is because If the spatial pattern has a high frequency (denser lines, closer together), the wavelength λ of the pattern is smaller.

FREQUENCY IN SPATIAL DOMAIN: SAMPLING



DFT can't capture every frequency component in the spectrum. The reason behind that is the **Sampling Rate**. Observe the sampled signals here (spanning 1 second/meter).

Despite the difference in frequency and values, the sampled outputs for both signals 1 and 2 are identical. Hence, it's impossible to recover both signals. That's why the sampling rate must be at least **twice** the highest frequency present in the signal, as per the **Nyquist Theorem**. **Frequencies above half the sampling rate get folded back into the range below it**, which is what happens to signal 2 here (it gets lost), and signal 1 is reconstructed.

KEY INSIGHTS

Before moving on to images, let's summarize the key takeaways from what we've seen so far:

- A negative frequency spike appears due to the complex conjugate symmetry inherent in the DFT of real-valued signals. Negative frequencies convey the same information as positive frequencies but in the opposite direction. So, they are often ignored in analysis.
- The relationship between spatial frequency and spacing in the pattern is direct: closer lines (**frequent changes in intensity**, higher spatial frequency) result in **larger values in the frequency domain**. For the same reason, **sharp edges represent higher frequencies**.
- The harmonics arise because a square wave is composed of an infinite summation of sinusoids with odd multiples of the fundamental frequency. **Most real-life signals are composed of many sinusoids, hence several frequency components**.
- Frequent changes in intensity mean Higher spatial frequencies. It corresponds to denser patterns in the spatial domain, which translate into **spikes further apart in the frequency domain**.
- DFT can only extract frequencies up to half the sampling rate, as higher frequencies fold back into this range due to **aliasing**. After that, the frequency spectrum produced by the DFT repeats every **2π** units (Mapping the highest representable frequency = Sampling frequency/2, called the Nyquist frequency, to **π**). So, the spectrum will be periodic.

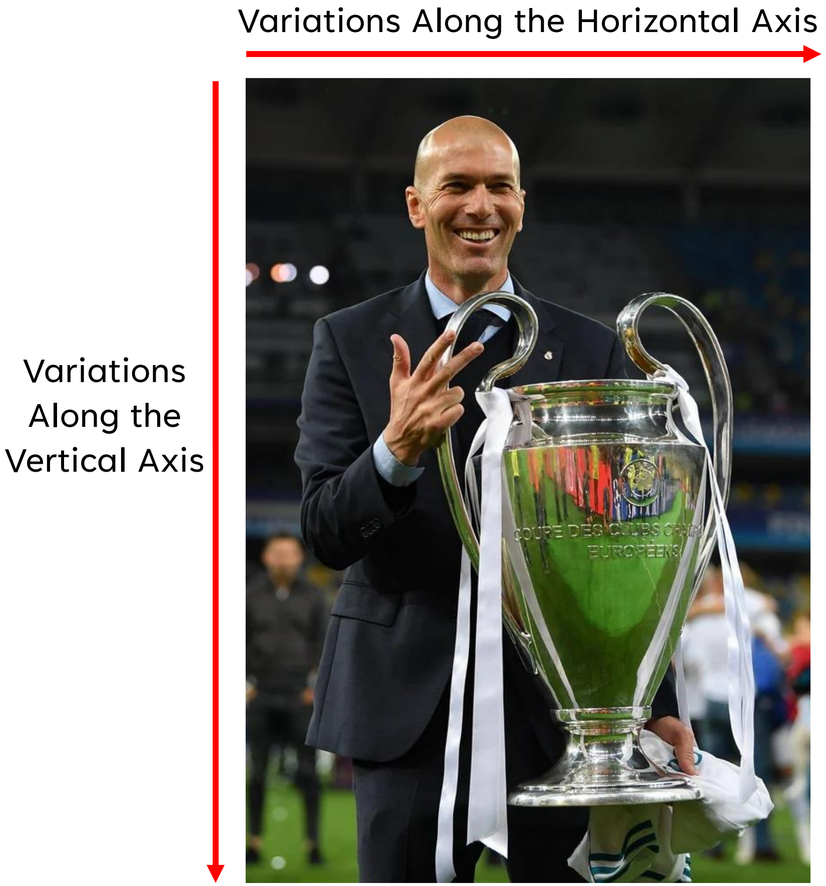
FREQUENCY IN 2D IMAGES

A 1D signal (e.g., a sine wave) varies in one direction (time or space). A 2D signal (e.g., an image) varies in two directions (rows and columns of pixels in the spatial domain). In 1D, we deal with variations over time/space, and in 2D, these variations occur along horizontal (x) and vertical (y) axes (both spatial in the case of images).

In 1D, frequency measures how rapidly the signal oscillates over time/space. In 2D, frequency now measures variations in both horizontal and vertical directions:

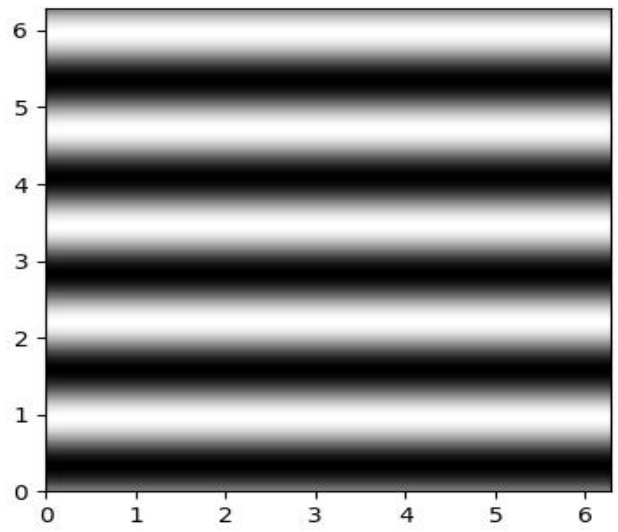
- **Horizontal frequency:** How rapidly pixel intensity changes horizontally (e.g., vertical stripes in an image).
- **Vertical frequency:** How rapidly pixel intensity changes vertically (e.g., horizontal stripes in an image).

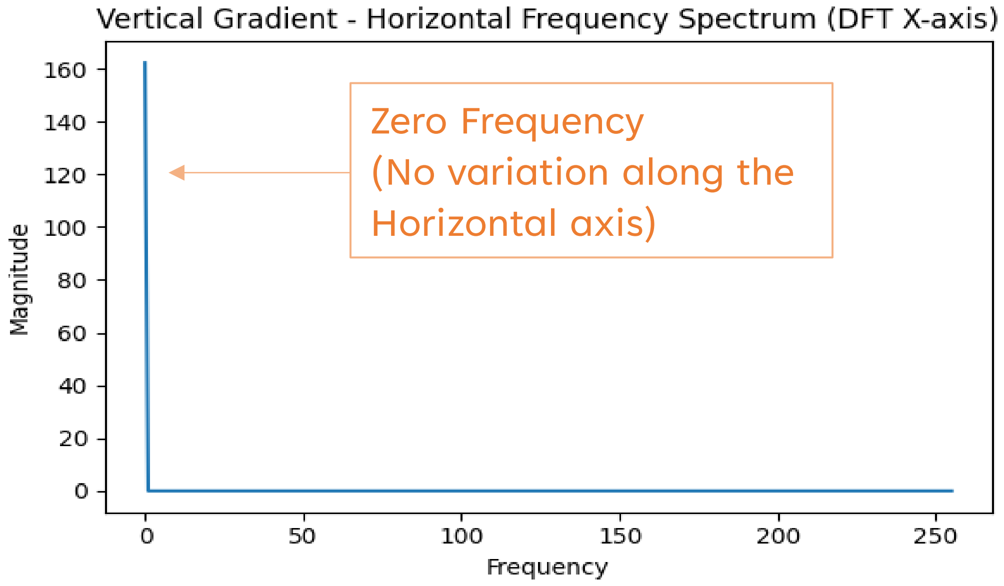
The horizontal and vertical components together can capture variations along any direction in a 2D plane.

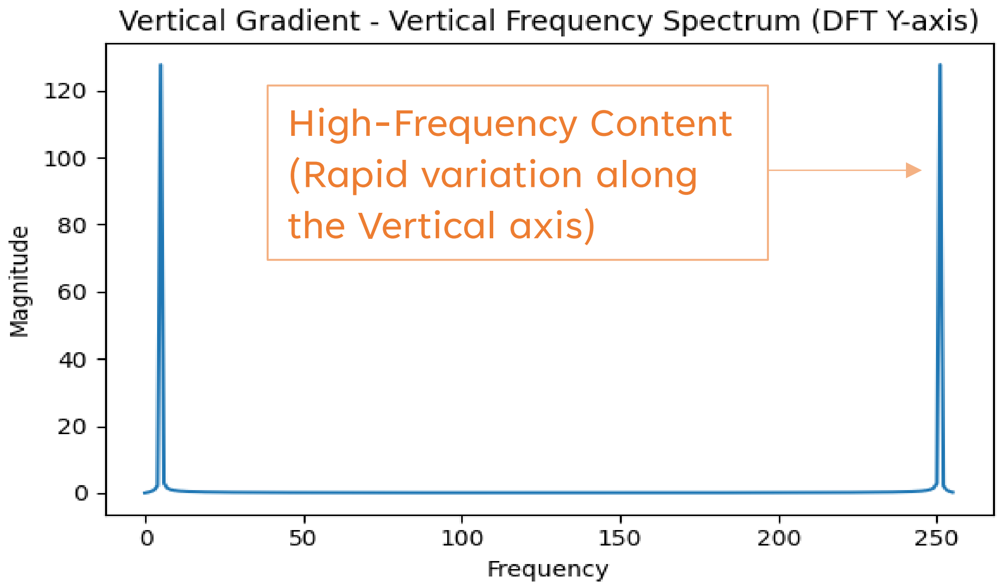


FREQUENCY IN 2D IMAGES

Let us take a look at the image here and the DFTs along the horizontal (X) and Vertical (Y) axis below. As we can see, there is no frequency component (apart from the DC component) in the horizontal frequency spectrum (since there's no change in intensity along the X-axis). But there's a high-frequency component in the vertical frequency spectrum as intensity changes periodically along the Y-axis.

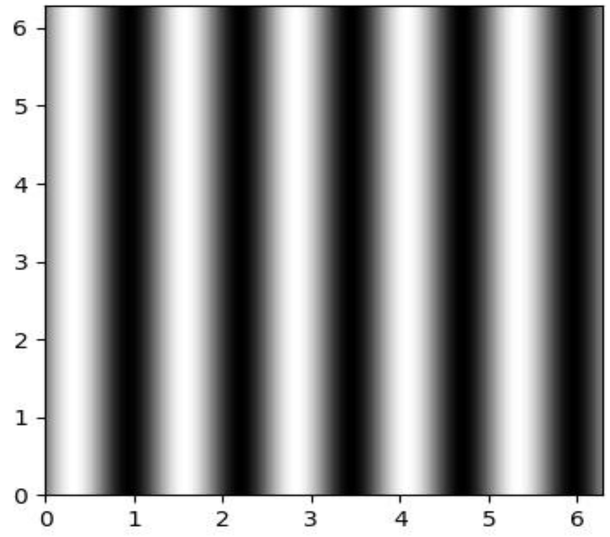


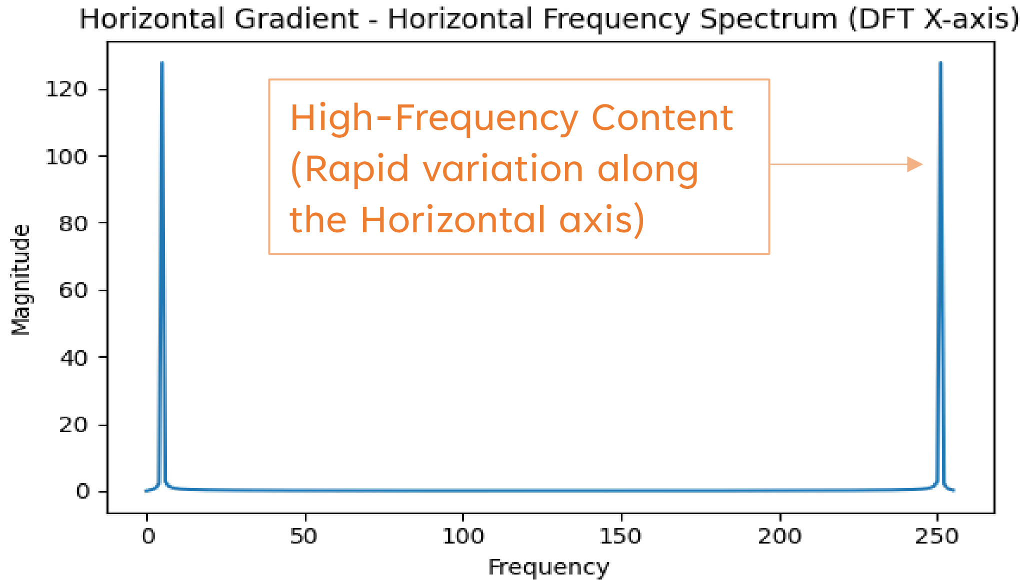


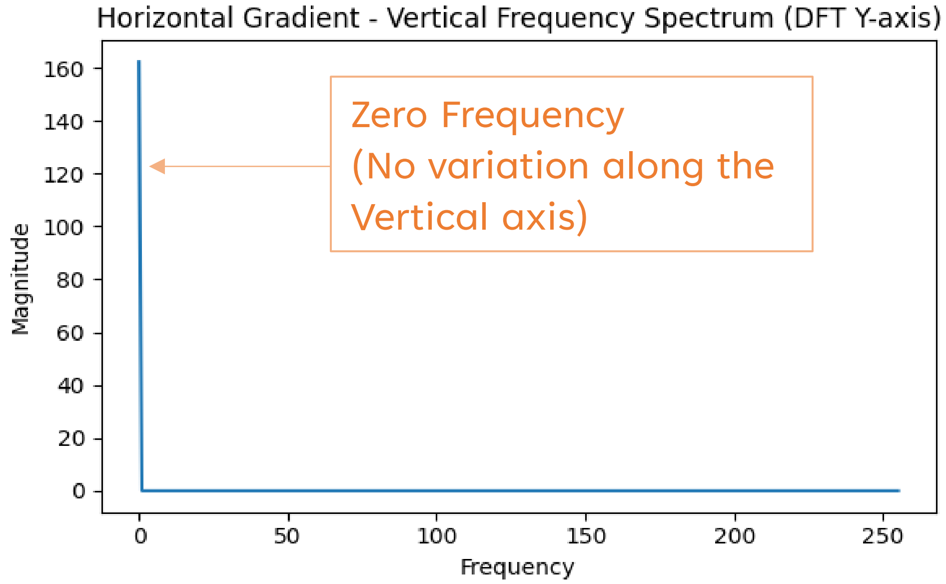


FREQUENCY IN 2D IMAGES

Now take a look at the image here and the DFTs along the horizontal (X) and Vertical (Y) axis below. As we can see, there is no frequency component (apart from the DC component) in the vertical frequency spectrum (since there's no change in intensity along the Y-axis). But there's a high-frequency component in the horizontal frequency spectrum as intensity changes periodically along the X-axis.

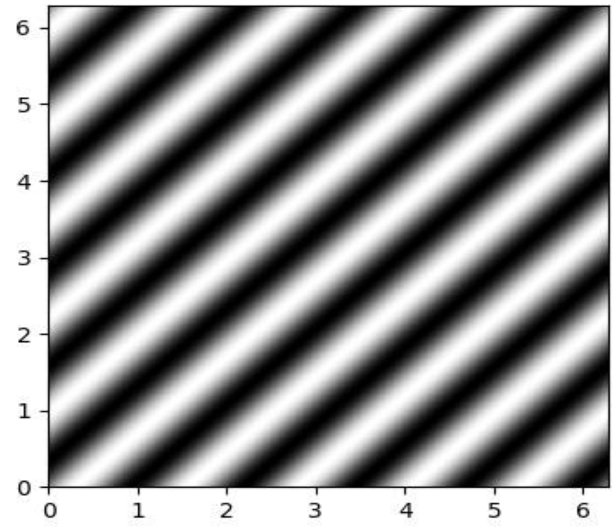


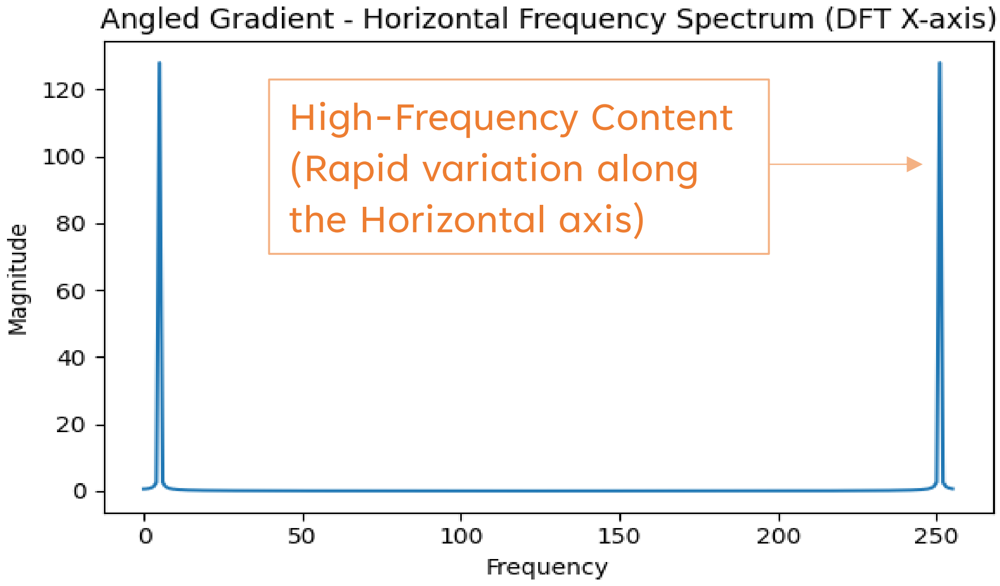


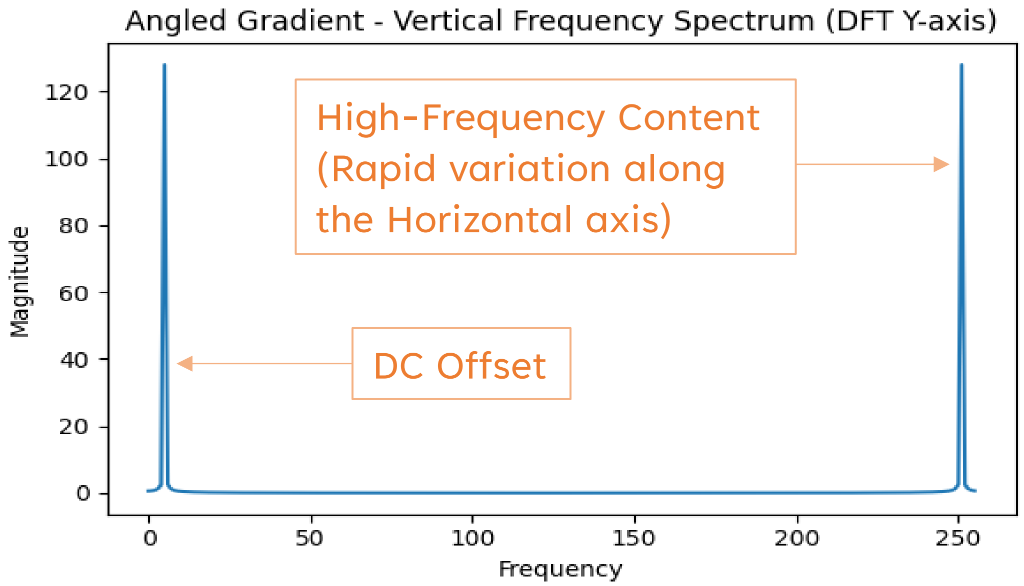


FREQUENCY IN 2D IMAGES

Now take a look at the image here and the DFTs along the horizontal (X) and Vertical (Y) axis below. As we can see, there's a high-frequency component in both cases, the horizontal frequency spectrum, and the vertical frequency spectrum, as intensity changes periodically along both the X-axis and the Y-axis.

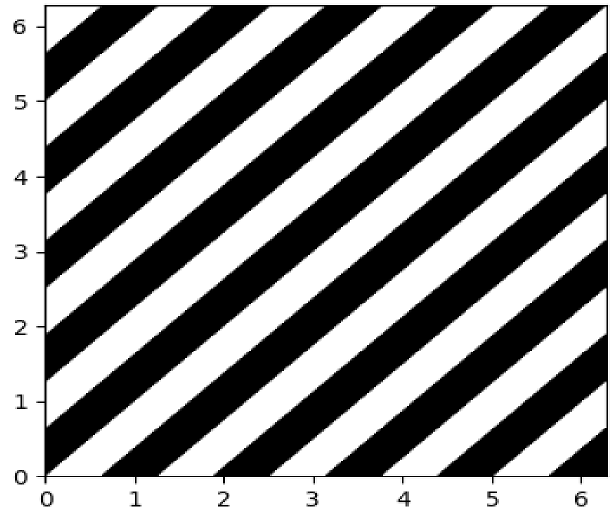


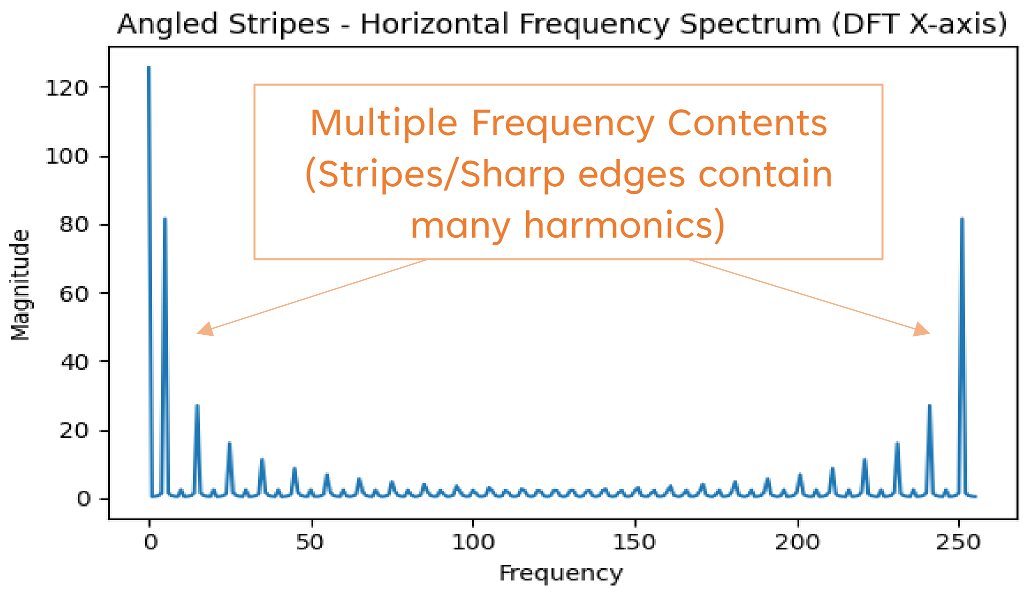


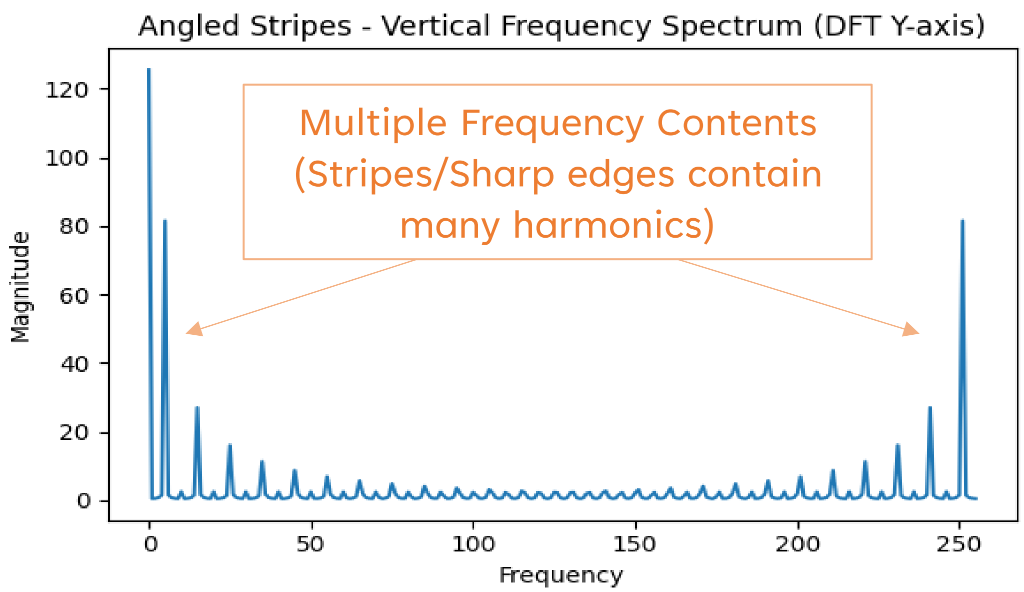


FREQUENCY IN 2D IMAGES

Finally, take a look at the image here and the DFTs along the horizontal (X) and Vertical (Y) axis below. As we can see, there are many frequency components in both cases, the horizontal frequency spectrum, and the vertical frequency spectrum, as intensity changes periodically along both the X-axis and the Y-axis. The various frequencies are due to the sharp edges in this image that contain many frequencies.

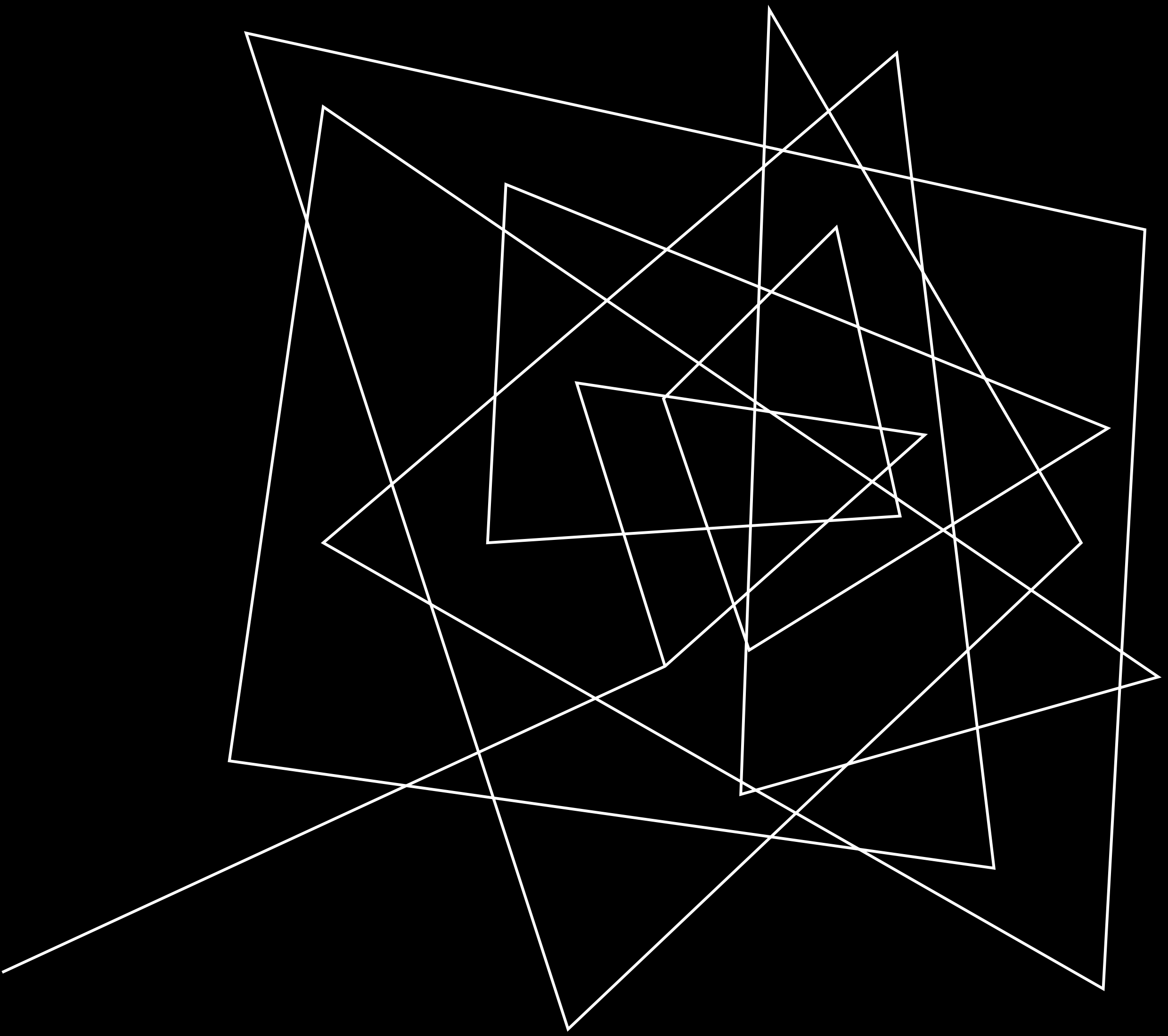






HOW MANY DFTS?

So, we can plot the frequency contents along both axes of an image. However, there are many horizontal and vertical lines within an image. It is unrealistic to plot so many 1D DFTs to analyze the frequency contents of an image. Then how can we represent an image in the frequency domain properly? The answer is in **2D Discrete Fourier Transform (DFT).**



2D DISCRETE FOURIER TRANSFORM (DFT)

2D DISCRETE FOURIER TRANSFORM (DFT)

The **2D Discrete Fourier Transform (DFT)** is an extension of the 1D DFT to two dimensions, typically applied to 2D signals like images. It transforms a spatial-domain representation (pixel intensities) into the frequency domain, capturing the frequency components in both horizontal and vertical directions. For a 2D signal $\boldsymbol{f}(\boldsymbol{x},\boldsymbol{y})$ of size $\mathbf{M} \times \mathbf{N}$, the **2D DFT** is given by:

$$\boldsymbol{F}(\boldsymbol{u},\boldsymbol{v}) = \sum_{\mathrm{x}=0}^{\mathrm{M}-1} \sum_{y=0}^{\mathrm{N}-1} \boldsymbol{f}(\boldsymbol{x},\boldsymbol{y})\mathbf{e}^{-j2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)}$$

Where,

- $\boldsymbol{f}(\boldsymbol{x},\boldsymbol{y})$ is the Intensity at spatial position $(\boldsymbol{x},\boldsymbol{y})$.
- $\boldsymbol{F}(\boldsymbol{u},\boldsymbol{v})$ is Frequency component at $(\boldsymbol{u},\boldsymbol{v})$.
- $\boldsymbol{u},\boldsymbol{v}$ are the horizontal and vertical frequency indices, respectively.
- $\mathbf{M},\mathbf{N}$ represent the dimensions of the image (Number of pixels along the $\boldsymbol{x}$ & $\boldsymbol{y}$ axis).

2D DISCRETE FOURIER TRANSFORM (DFT)

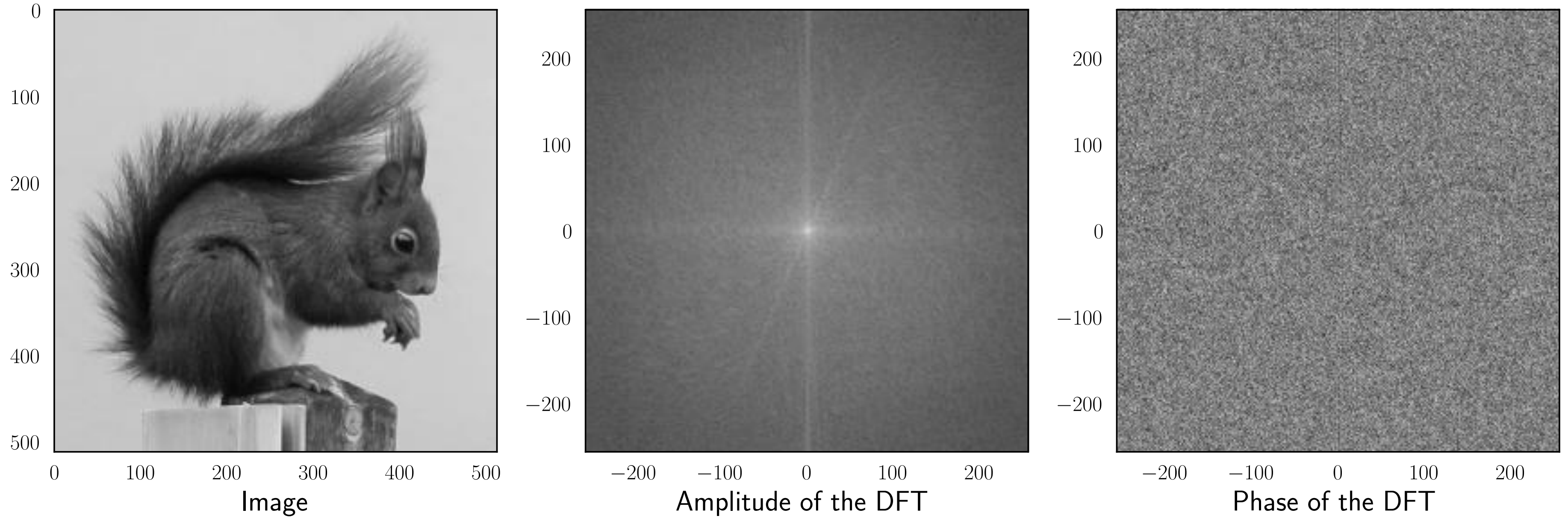
Clearly, the value of 2D DFT $\boldsymbol{F(u,v)}$ at each frequency index $\boldsymbol{u,v}$ is a complex number. So,

$$\boldsymbol{F(u,v) = a + bj = r\angle\theta}$$

Where,

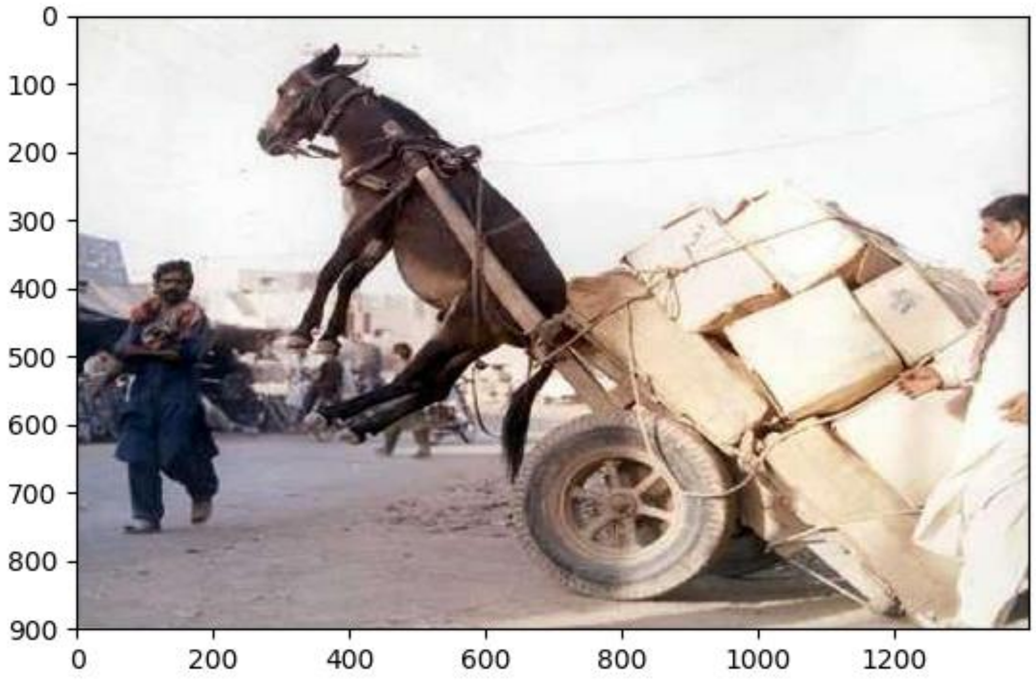
- $\boldsymbol{r} = |\boldsymbol{F(u,v)}|$ is the magnitude/amplitude.
- $\boldsymbol{\theta}$ represents the phase.

These two quantities for all frequency indices correspond to **Magnitude Spectrum** and **Phase Spectrum** respectively. These 2D image-like spectrums show higher intensities for larger values.

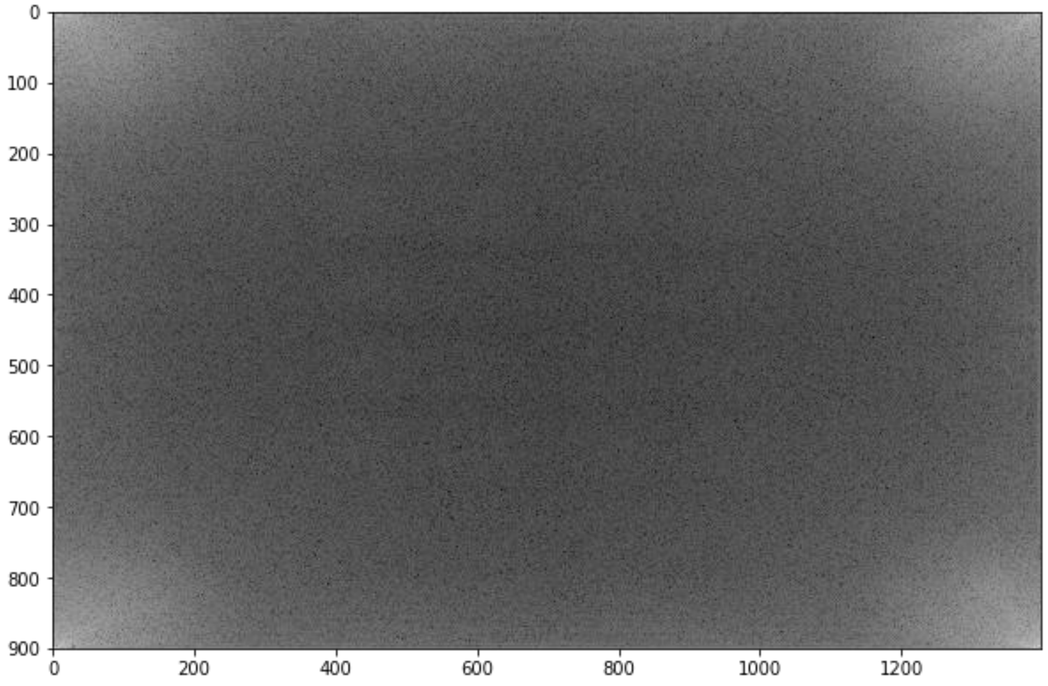


2D DFT: SHIFTING THE ORIGIN

In 2D DFT, by default, the low-frequency components (representing overall trends or smooth variations) are at the corners of the spectrum, while the high-frequency components (representing fine details or sharp edges) are distributed outward from the origin. This default layout can be unintuitive for human analysis.



Original Image

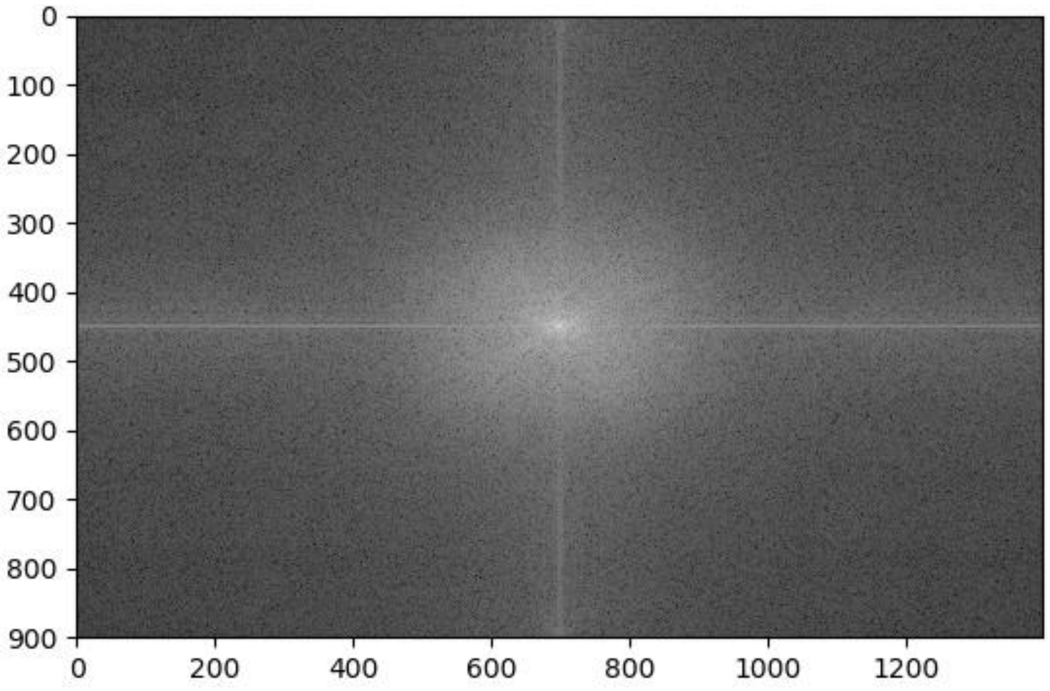


DFT

2D DFT: SHIFTING THE ORIGIN

In practice, we shift the origin to the center of the frequency spectrum in the 2D DFT for visualization and interpretation purposes. Placing the zero frequency (DC component) in the middle makes it easier to identify and analyze low frequencies. These represent the most prominent or general features in the image, like overall brightness. High-frequency components, which correspond to fine details or noise, are placed outward from the center. This arrangement aligns better with how we interpret spatial and frequency information.

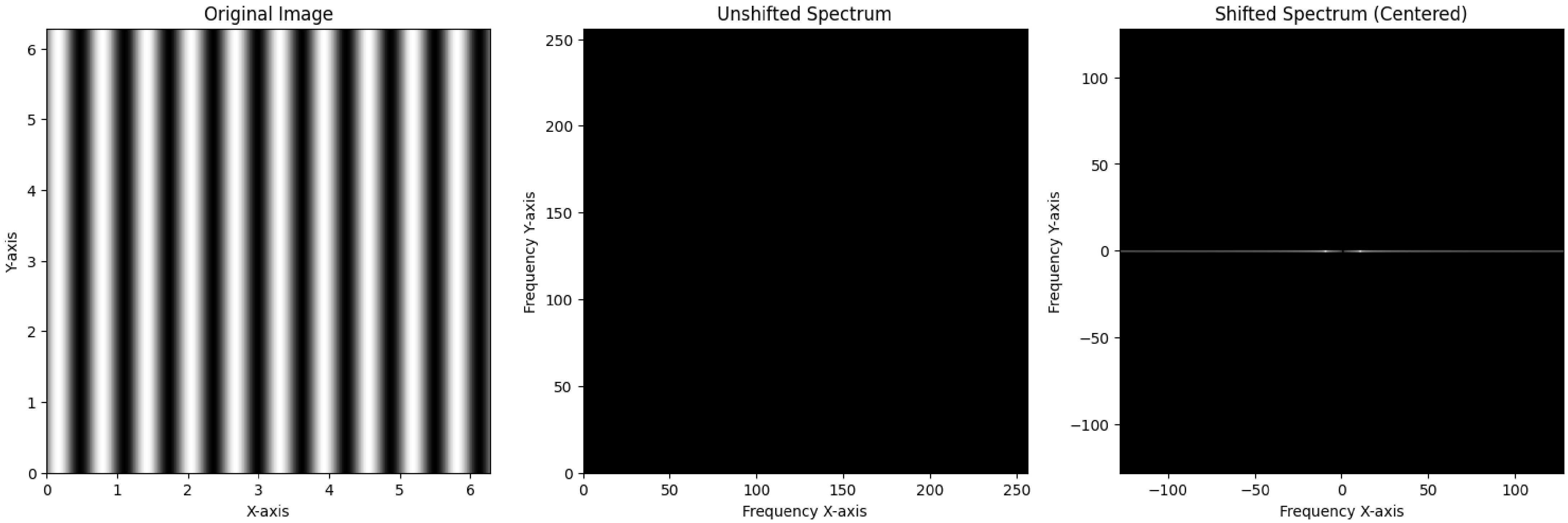
Original Image



Shifted DFT

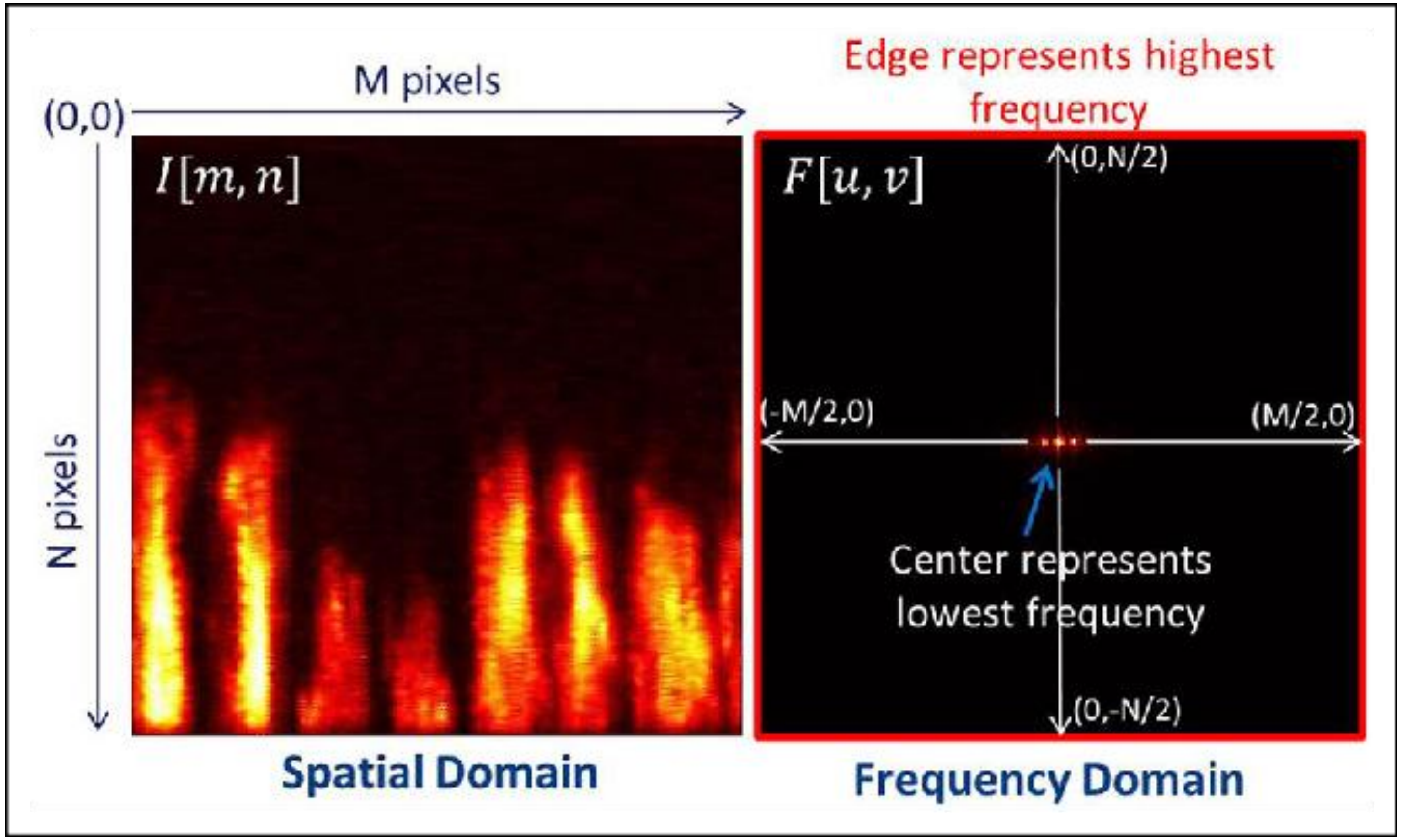
2D DFT: SHIFTING THE ORIGIN

The Picture below clearly shows why shifting is crucial to make the frequency domain representation interpretable. The higher frequency contents in the horizontal axis due to variations along the horizontal direction are clearly visible in the shifted spectrum, but they were hard to visualize in the unshifted one.



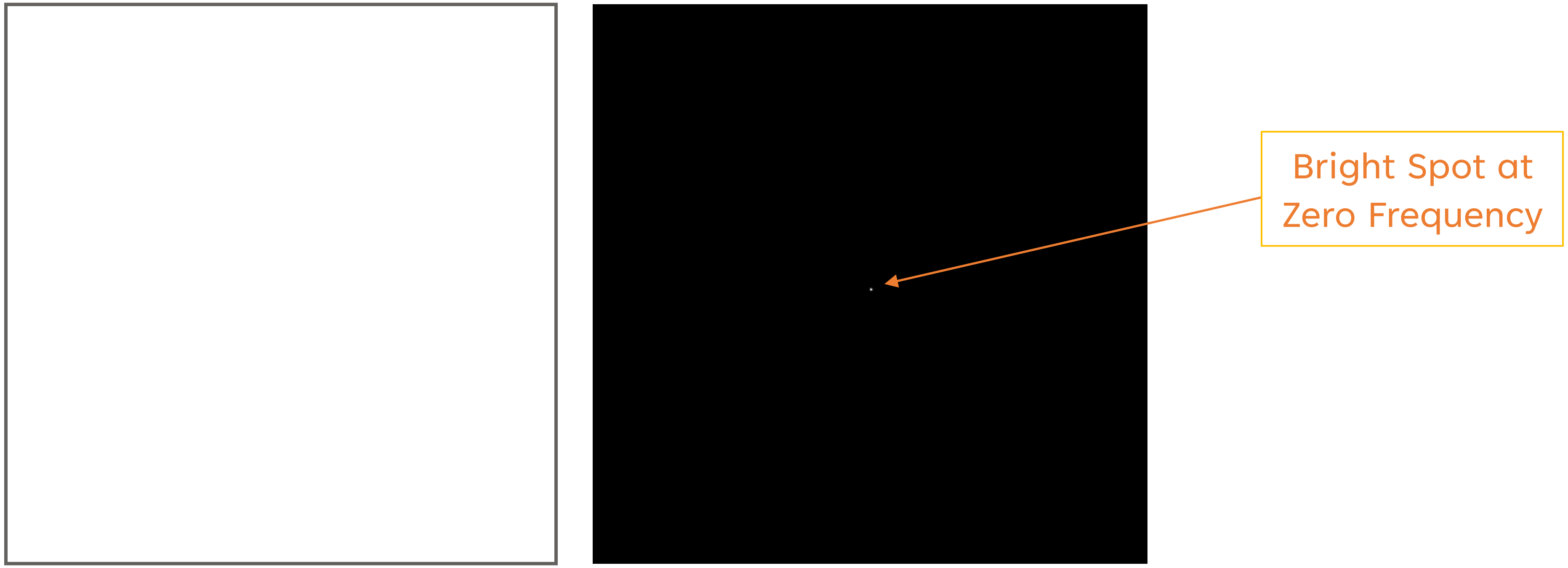
2D DFT: INTERPRETING THE SPECTRUM

This image illustrates the transition from the spatial domain to the frequency domain. On the left is the spatial representation of an image, I, where pixel values correspond to intensity variations. The right side shows its 2D frequency spectrum, F after applying the DFT. The symmetric nature of the frequency spectrum is a property of real-valued spatial data.



2D DFT: EXAMPLES

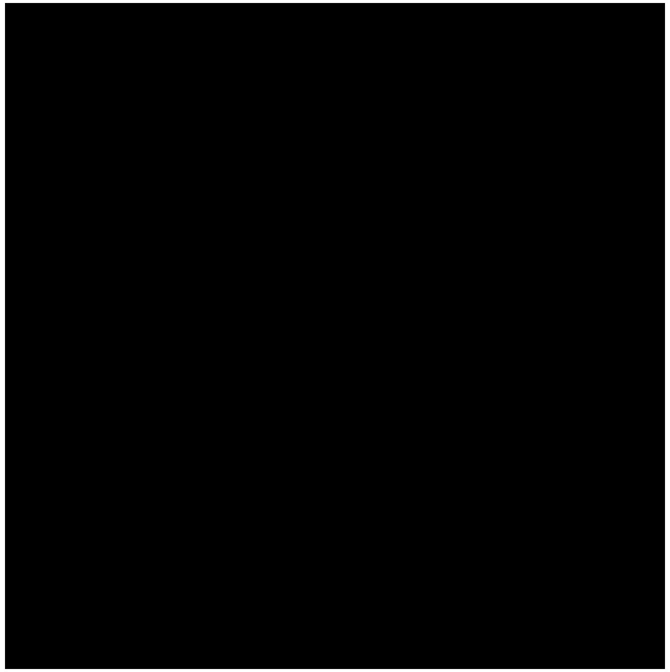
The DFT of a fully white image is a single bright spot at the center (**DC component**), indicating a uniform, zero-frequency intensity across the entire image. No other frequency component is visible, as there is no variation in the image at all.



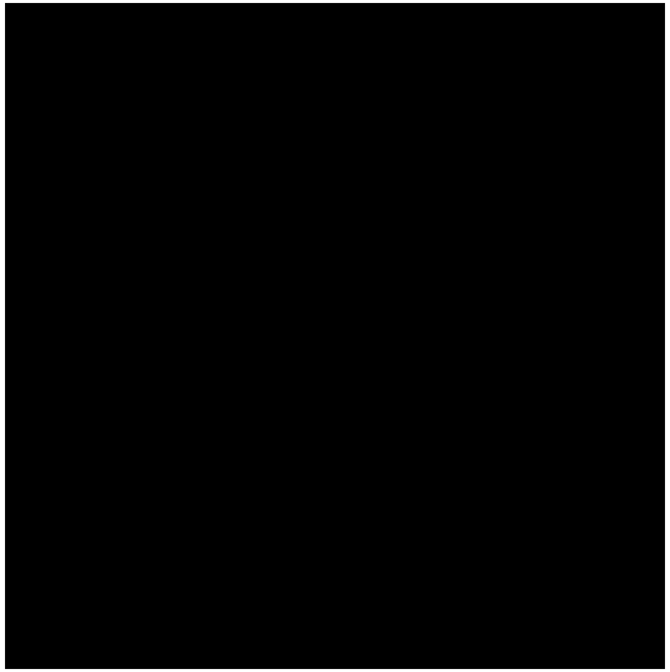
Original Image **DFT**

2D DFT: EXAMPLES

The DFT of a fully black image is completely dark because there is no signal (intensity is zero everywhere), and hence no frequency components. There is no value even at the center (average brightness is 0) and no frequency components as there are no variations in the image.



Original Image

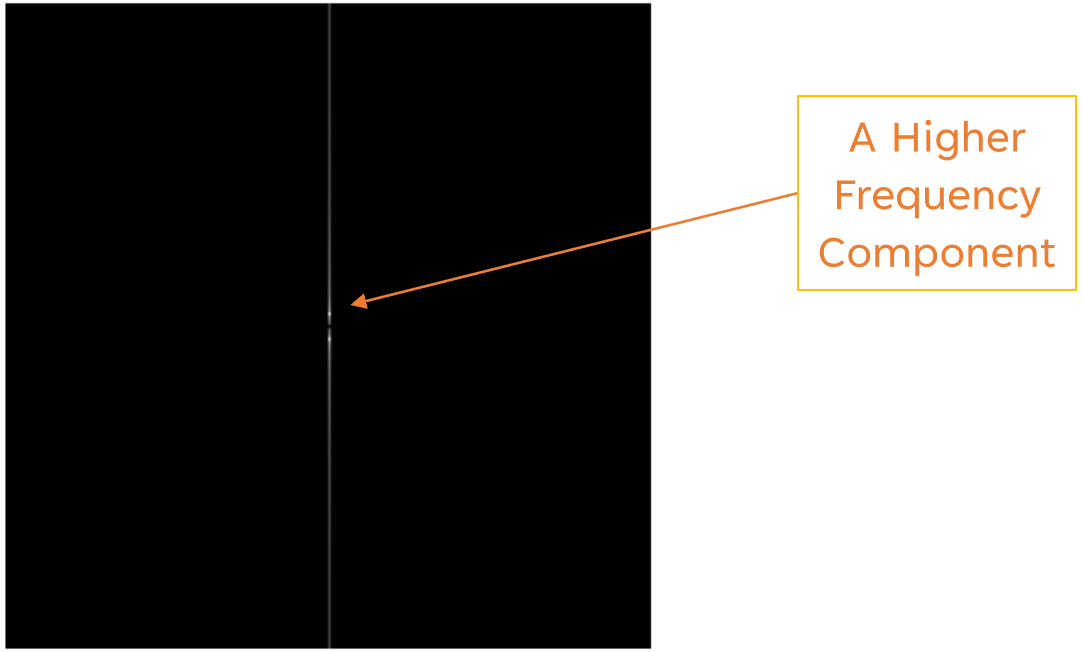


DFT

2D DFT: EXAMPLES

The DFT of this image shows high-frequency components along the vertical axis (y-direction), as the intensity varies vertically. As the variation follows a sinusoid pattern, there is a **peak**, a single bright higher frequency content (and its negative counterpart). The line that appears apart from the peaks arises due to boundary discontinuities in the image. If there is a sharp discontinuity between the edges of the image (boundaries), this creates artificial high-frequency components in the DFT (**Spectral Leakage**), manifesting as lines in the frequency spectrum, oriented perpendicularly to the direction of the variation (vertical in this case).

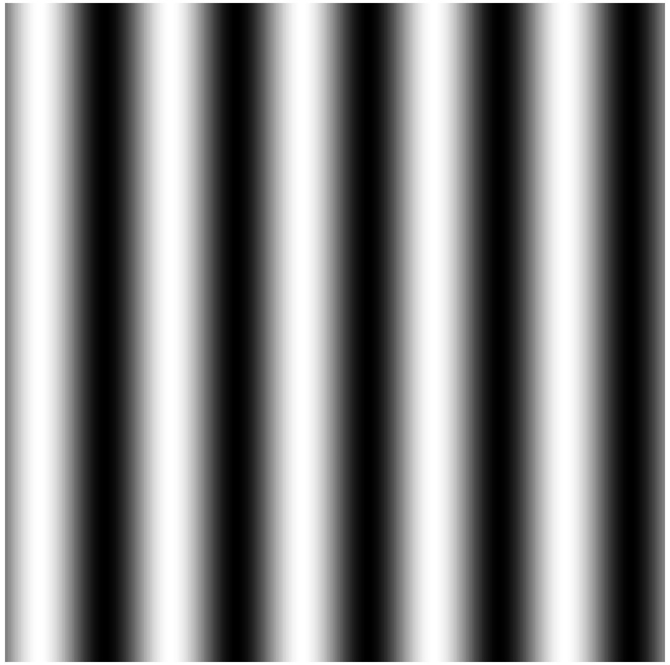
Original Image



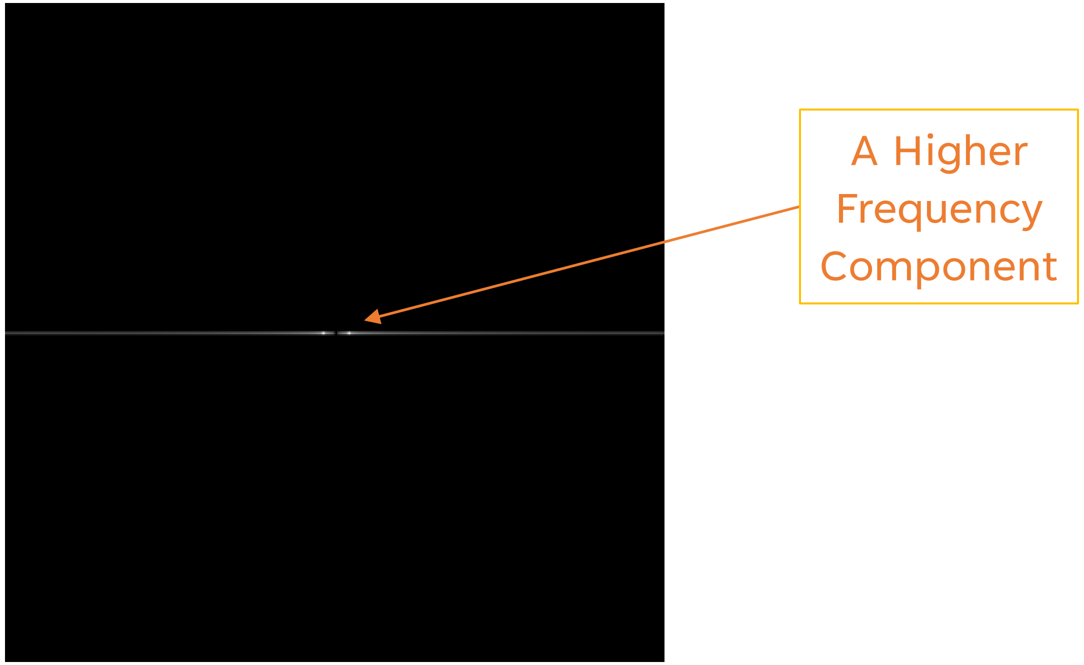
DFT

2D DFT: EXAMPLES

The DFT of this image shows high-frequency components along the horizontal axis (x-direction), as the intensity varies horizontally. As the variation follows the pattern of a sinusoid, there is a single bright higher frequency content (and its negative counterpart). Again, the spectral leakage that appears apart from the peaks arises due to boundary discontinuities in the image. The sharp discontinuity at the boundary of the image creates artificial high-frequency components in the DFT, manifesting as lines in the frequency spectrum.



Original Image



DFT

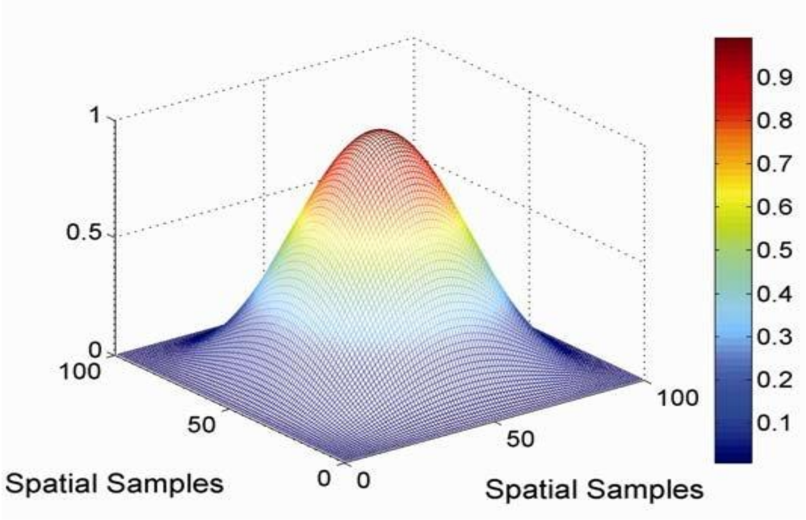
2D DFT: REDUCING SPECTRAL LEAKAGE (OPTIONAL)

So, when analyzing a finite signal (e.g., images) in the frequency domain, the boundaries of the signal introduce discontinuities. These discontinuities result in spectral leakage, where energy from one frequency spreads into others. The **Hamming Window** reduces these discontinuities by tapering the signal's edges smoothly to zero (or nearly zero), rather than abruptly cutting it off. The Hamming window is defined as:

$$w[n] = 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right) \quad , \quad 0 \le n \le N-1$$

$$w[m] = 0.54 - 0.46\cos\left(\frac{2\pi m}{M-1}\right) \quad , \quad 0 \le m \le M-1$$

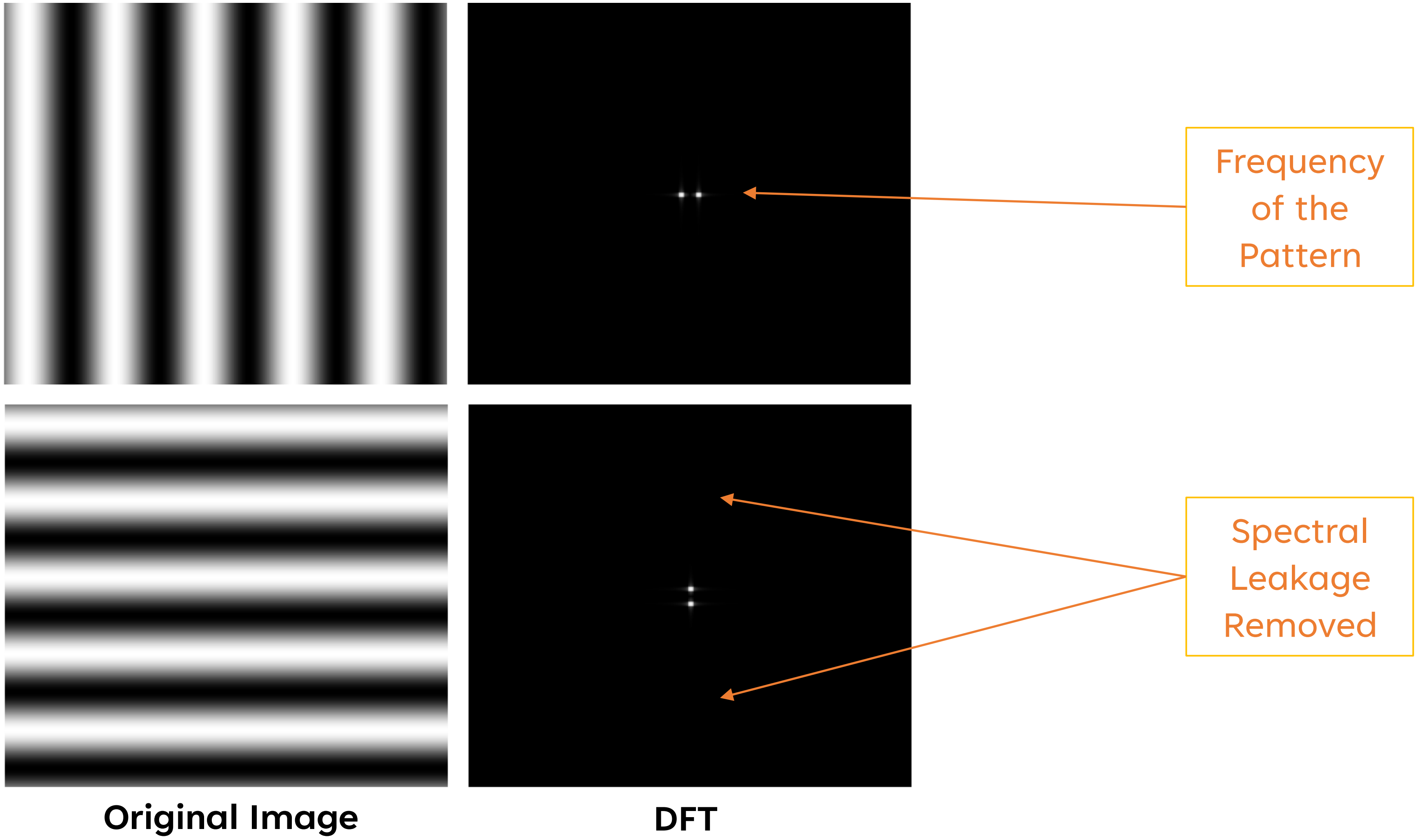
Where, $\mathbf{M}$ and $\mathbf{N}$ are the dimensions of the image and $\mathbf{m}$ and $\mathbf{n}$ are indices for rows and columns, respectively.



A 2D Hamming Window

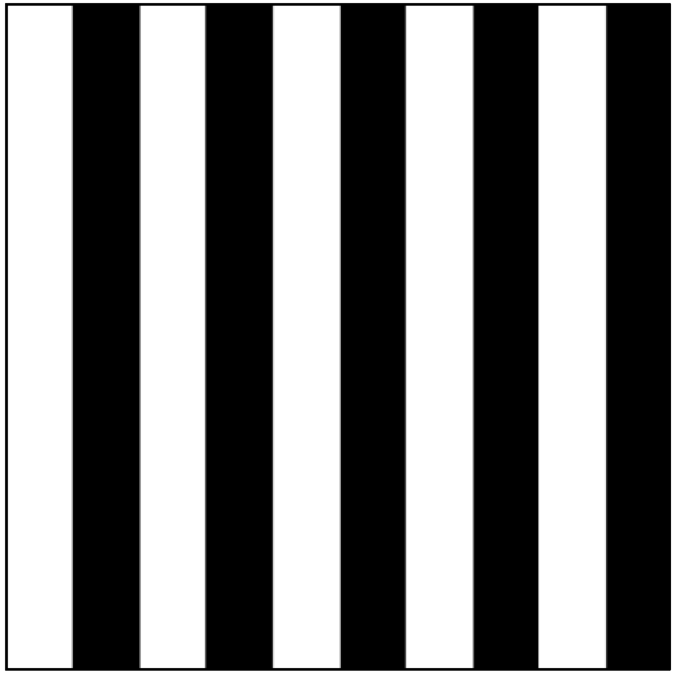
EFFECT OF HAMMING WINDOW (OPTIONAL)

Look at the DFTs below after applying the hamming window. The frequencies of the pattern within the images are much clearer now, and spectral leakage has been reduced significantly.

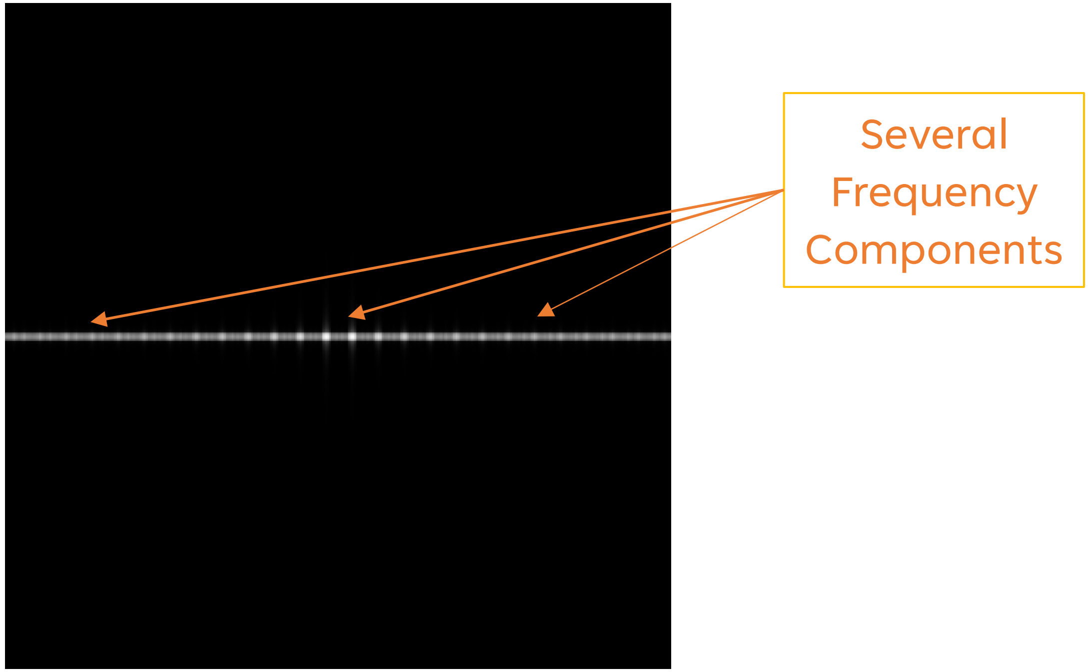


2D DFT: EXAMPLES

The DFT of this image shows high-frequency components along the horizontal axis (x-direction), as the intensity varies horizontally. As the image contains sharp changes in intensity (from fully black to white), there are several bright higher frequency contents (represents **peaks** along the vertical axis frequency spectrum) here. Sharp edges correspond to high frequencies, that is why we can see bright spots near the edge of the spectrum. There are no frequency contents bar zero along the vertical axis, due to no variations along the y-axis.



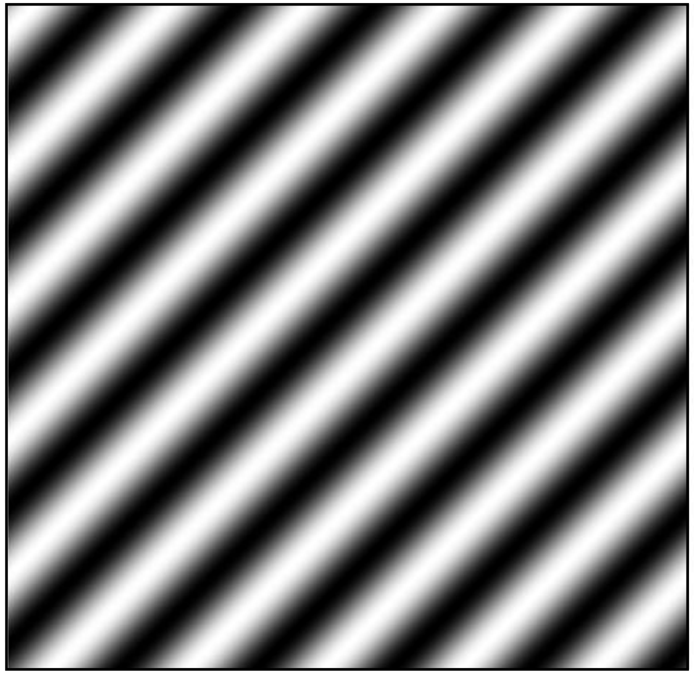
Original Image



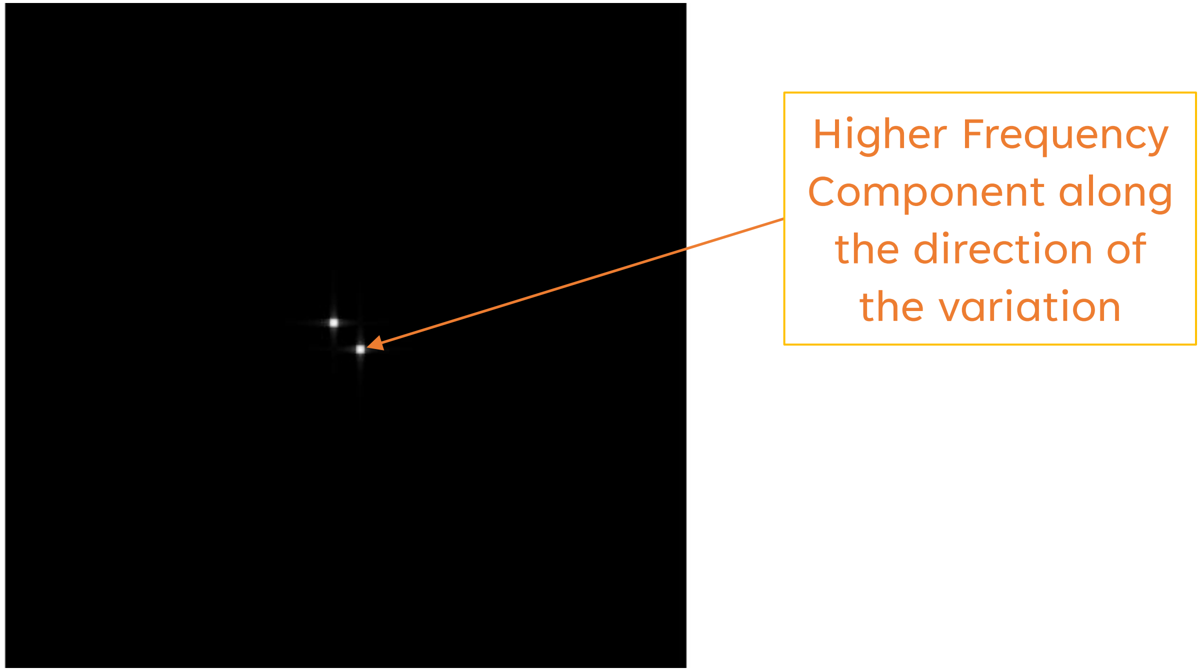
DFT

2D DFT: EXAMPLES

The DFT of this image displays a diagonal pattern of frequency components, corresponding to the direction and rate of intensity change along the angle. It contains a single **peak** (and its negative) along the angled direction of variation, corresponding to the frequency of the variation. As we can see, this frequency has both vertical and horizontal components, which is due to the fact that the intensity varies to some extent along both these directions.



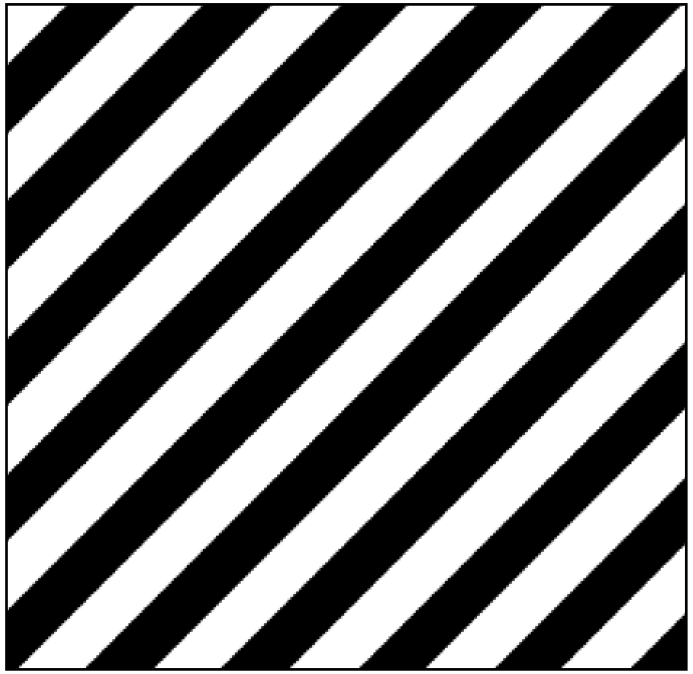
Original Image



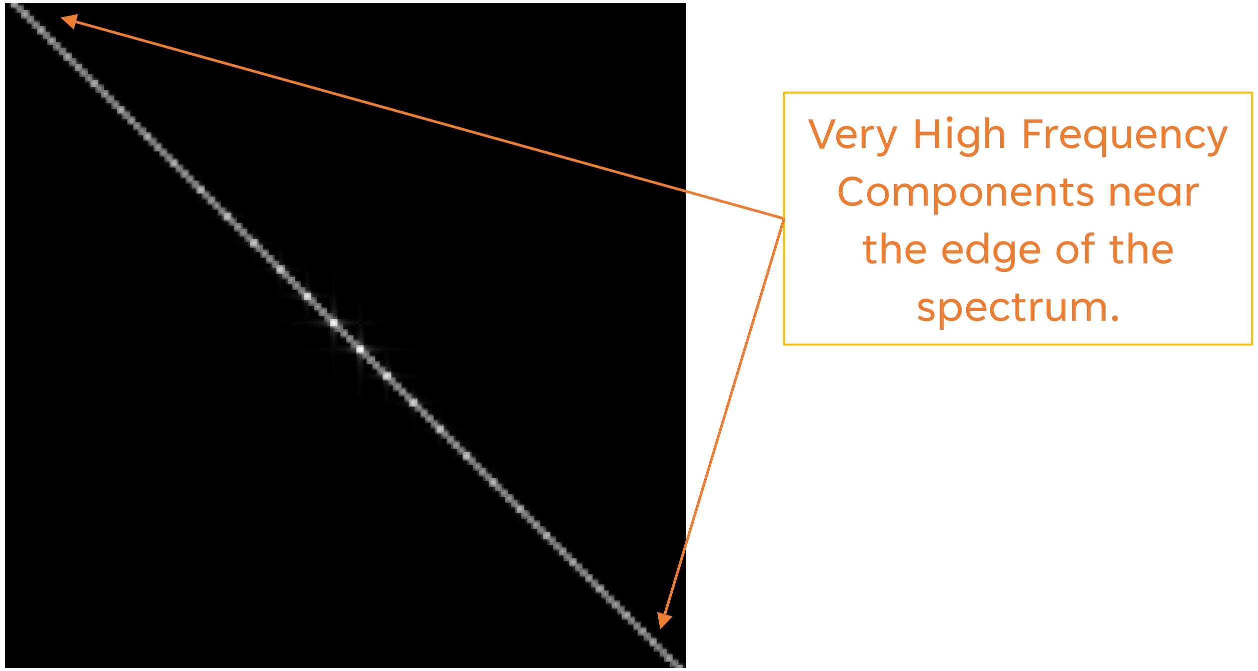
DFT

2D DFT: EXAMPLES

The DFT of this image shows sharp peaks along the diagonal axis, indicating strong periodic frequencies in the angled direction. As there are sharp edges along that angled direction, the DFT contains multiple **peaks** along the angled direction of variation, corresponding to the many frequencies the edges represent. Bright peaks near the corners of the spectrum show that edges contains quite strong high frequency components.



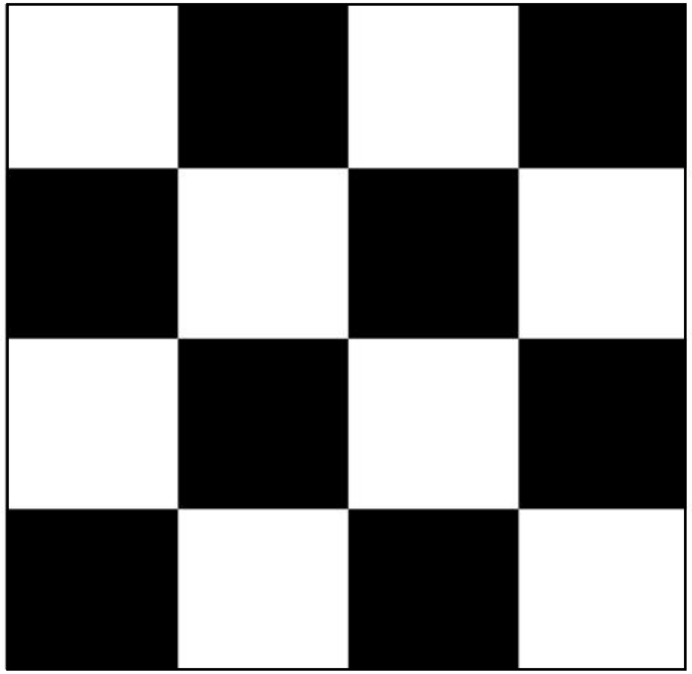
Original Image



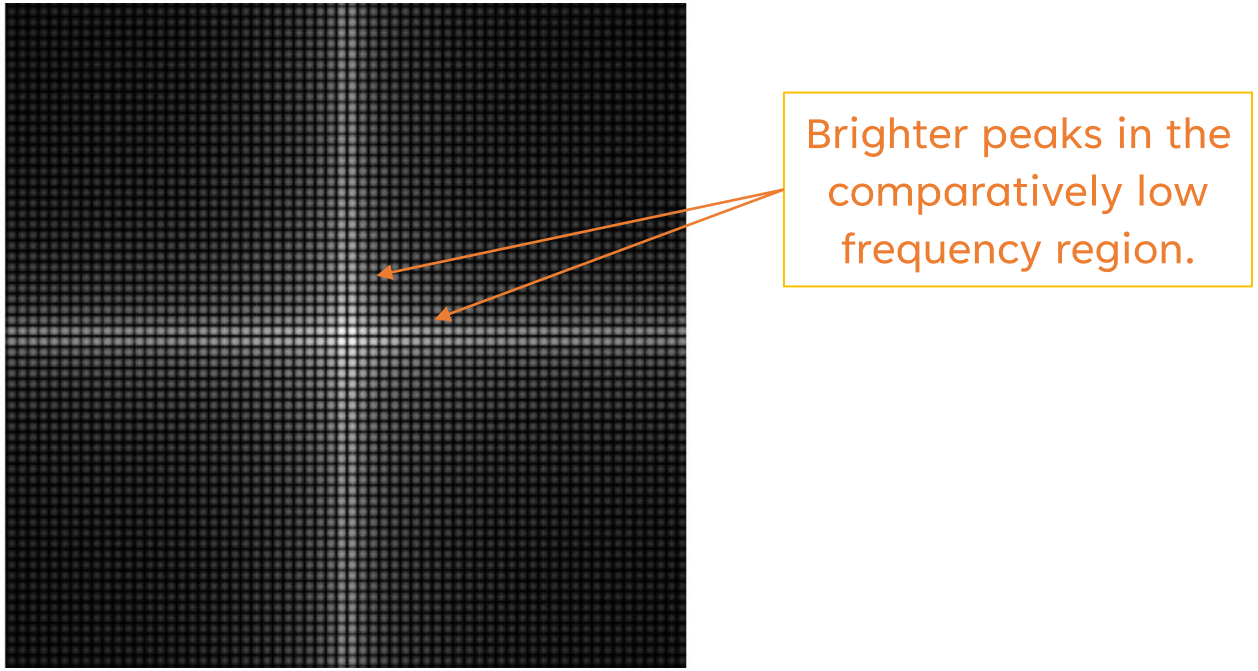
DFT

2D DFT:EXAMPLES

The DFT of this checkerboard contains symmetric peaks in both horizontal and vertical axes, representing periodic patterns in all directions. We can also notice that the **peaks** corresponding to comparatively lower frequencies are brighter, which is expected because the rate of variation in the checkerboard is not too much. Also, the adjacent peaks in the frequency domain are quite close to each other.



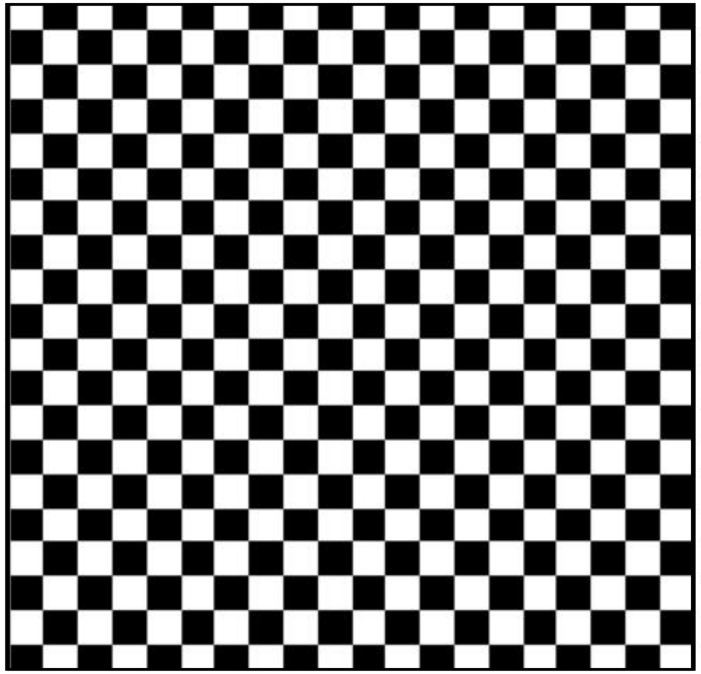
Original Image



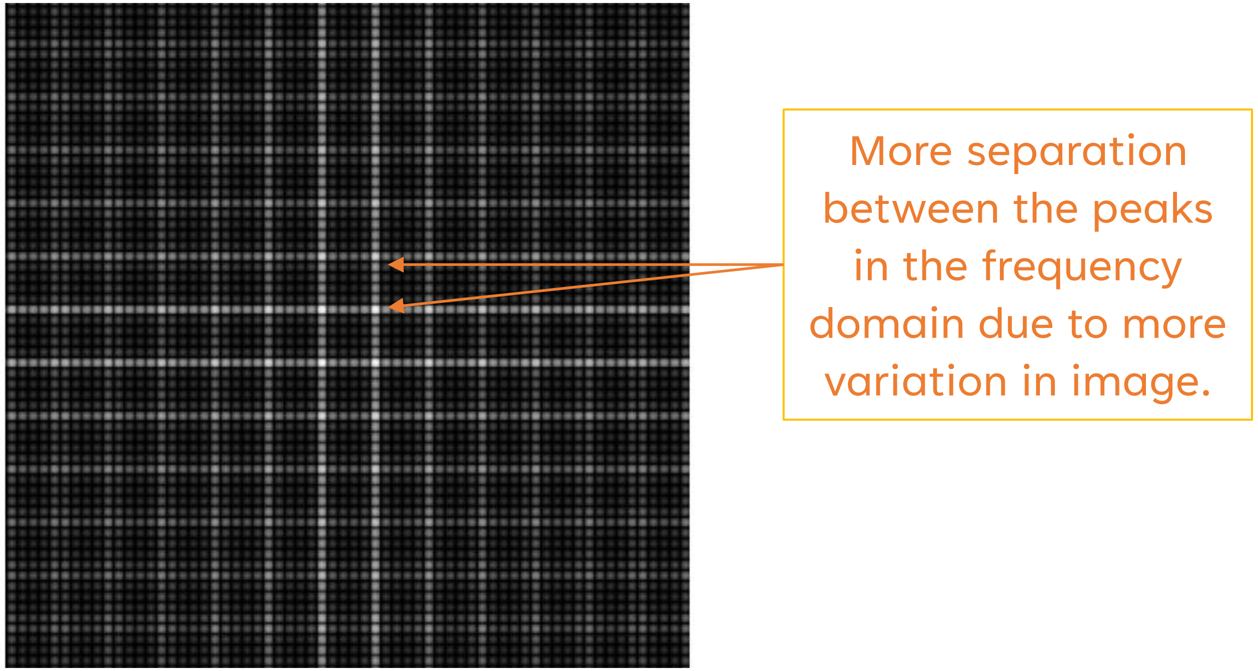
DFT

2D DFT: EXAMPLES

The DFT of this checkerboard again contains symmetric peaks in both horizontal and vertical axes, representing periodic patterns in all directions almost creating another checkerboard. The **peaks** corresponding to higher frequencies are now brighter, which is expected because the rate of variation in the checkerboard is much more in this case. Also, the distance between adjacent peaks in the frequency domain has increased due to the higher fundamental frequency and its harmonics (multiples of the fundamental frequency).



Original Image

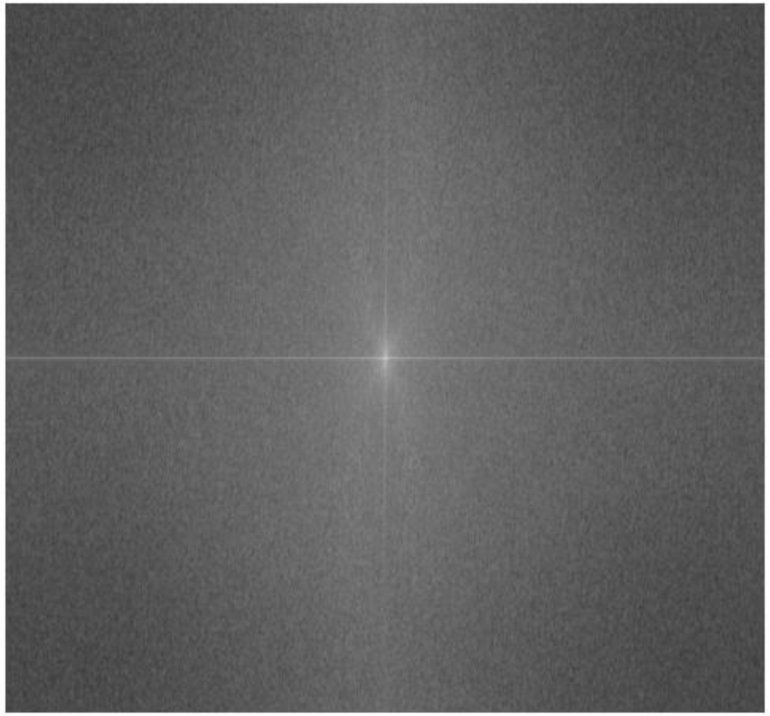


DFT

2D DFT: REAL LIFE IMAGES

The DFT of this image below shows that the image mainly contains low frequency components, which is expected as there is little change in the skyline covering large portions of the picture. The **peak** is at a low frequency and there is a slightly bright region along the vertical axis, which is due the fact that the transition from sky to mountain to water happens in this direction.

Original Image

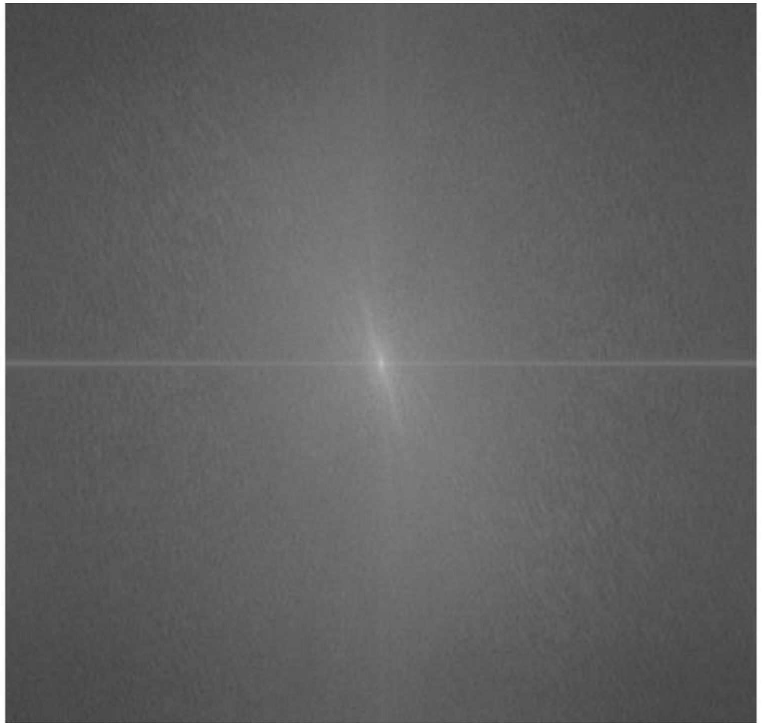


DFT

2D DFT: REAL LIFE IMAGES

The DFT of this image below once again shows that the image mainly contains low frequency components, which is obvious because there is hardly any variation throughout most of the picture. The **peak** is at a low frequency and there is a slightly bright region along the vertical axis at a small angle, which is again due the fact that the transition from sky to snowy mountain happens in this direction.

Original Image

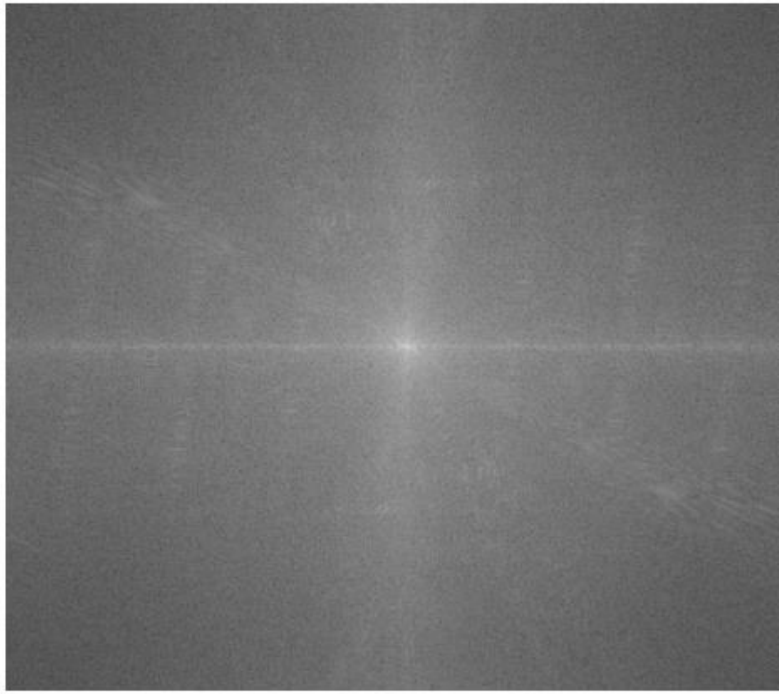


DFT

2D DFT: REAL LIFE IMAGES

The DFT of this image below once again shows that the image contains both low & high frequency components. This is because the nighttime cityscapes with sharp lights from windows and neon signs shows strong high frequencies. There are many bright regions throughout the frequency spectrum, even far from the center, which implies the existence of high-frequency components in the image.

Original Image

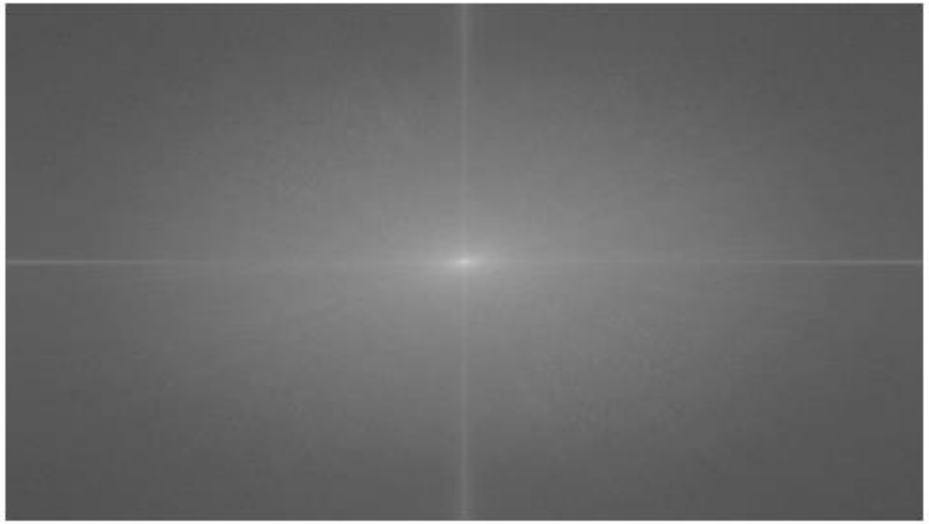


DFT

2D DFT: REAL LIFE IMAGES

The DFT of another image containing both low & high frequency components. This is because this image also features lots of variations in intensity. There are bright regions near the edge of the frequency spectrum, which implies the existence of high-frequency components in the image. Also notice how the dimensions of the DFT follows the dimensions of the original image *(Why? Notice the 2D DFT formula!)*.

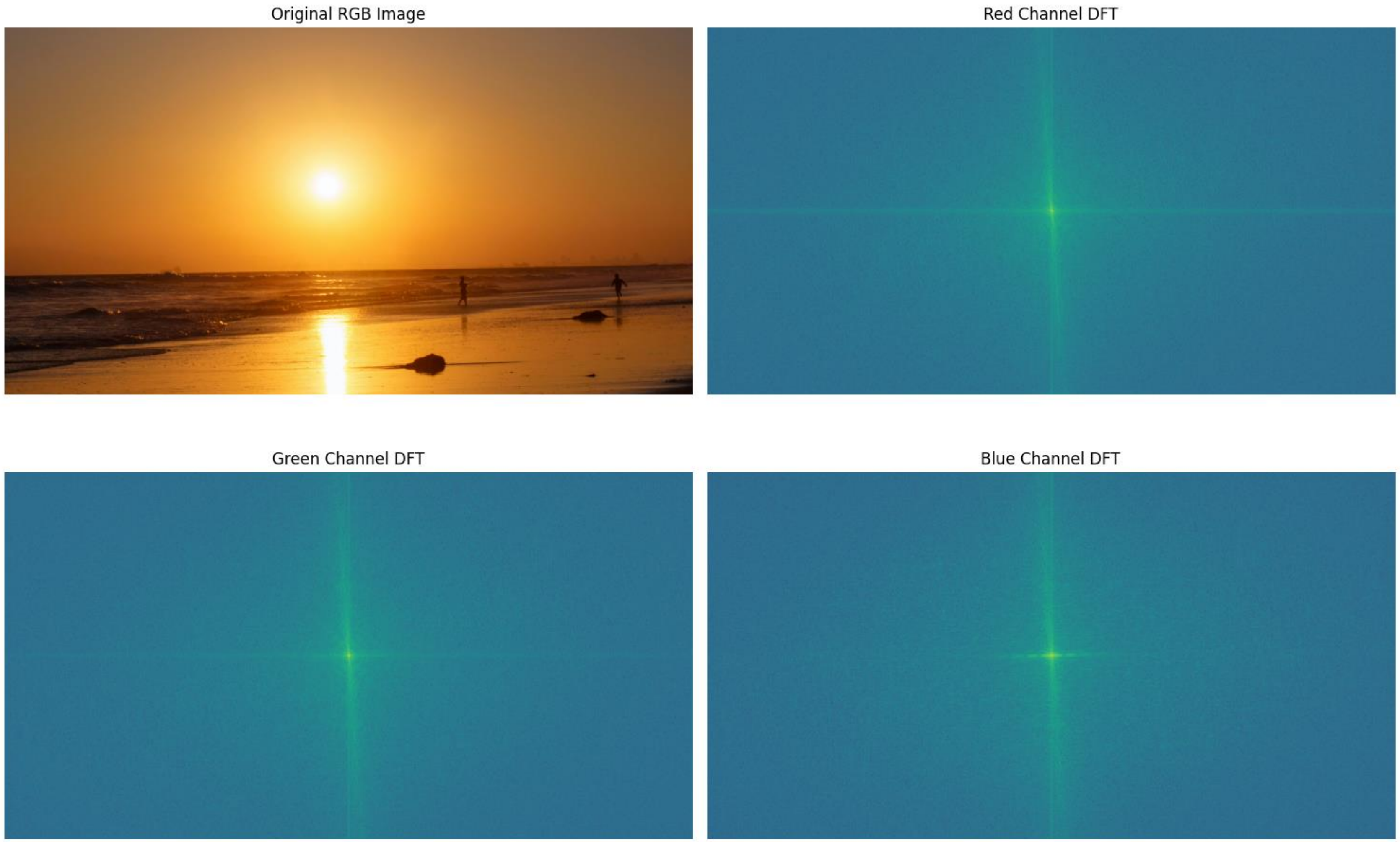
Original Image



DFT

DFT OF RGB IMAGES

To compute the Discrete Fourier Transform (DFT) of an RGB image, one way could be applying the DFT to each of the 3 color channels (red, green, and blue) separately.

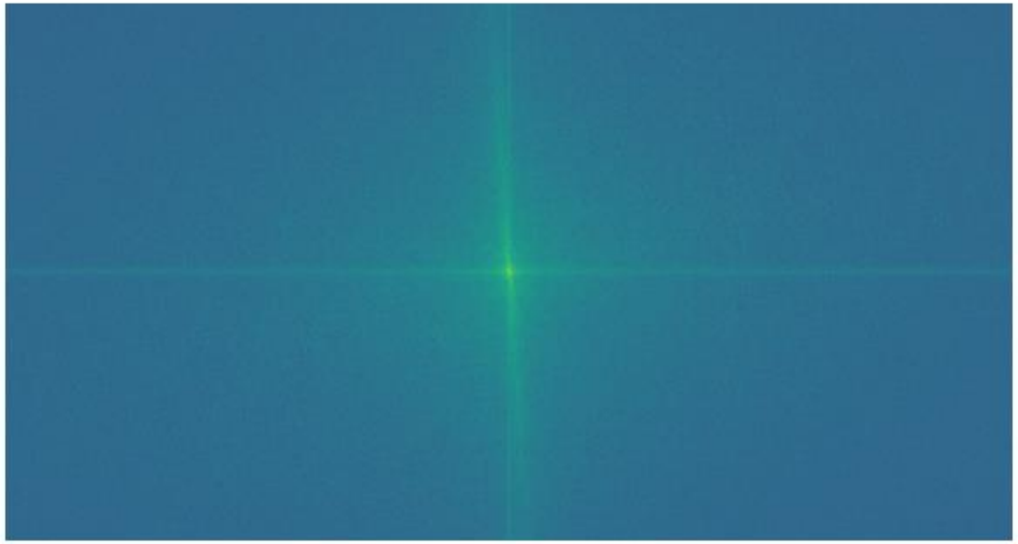


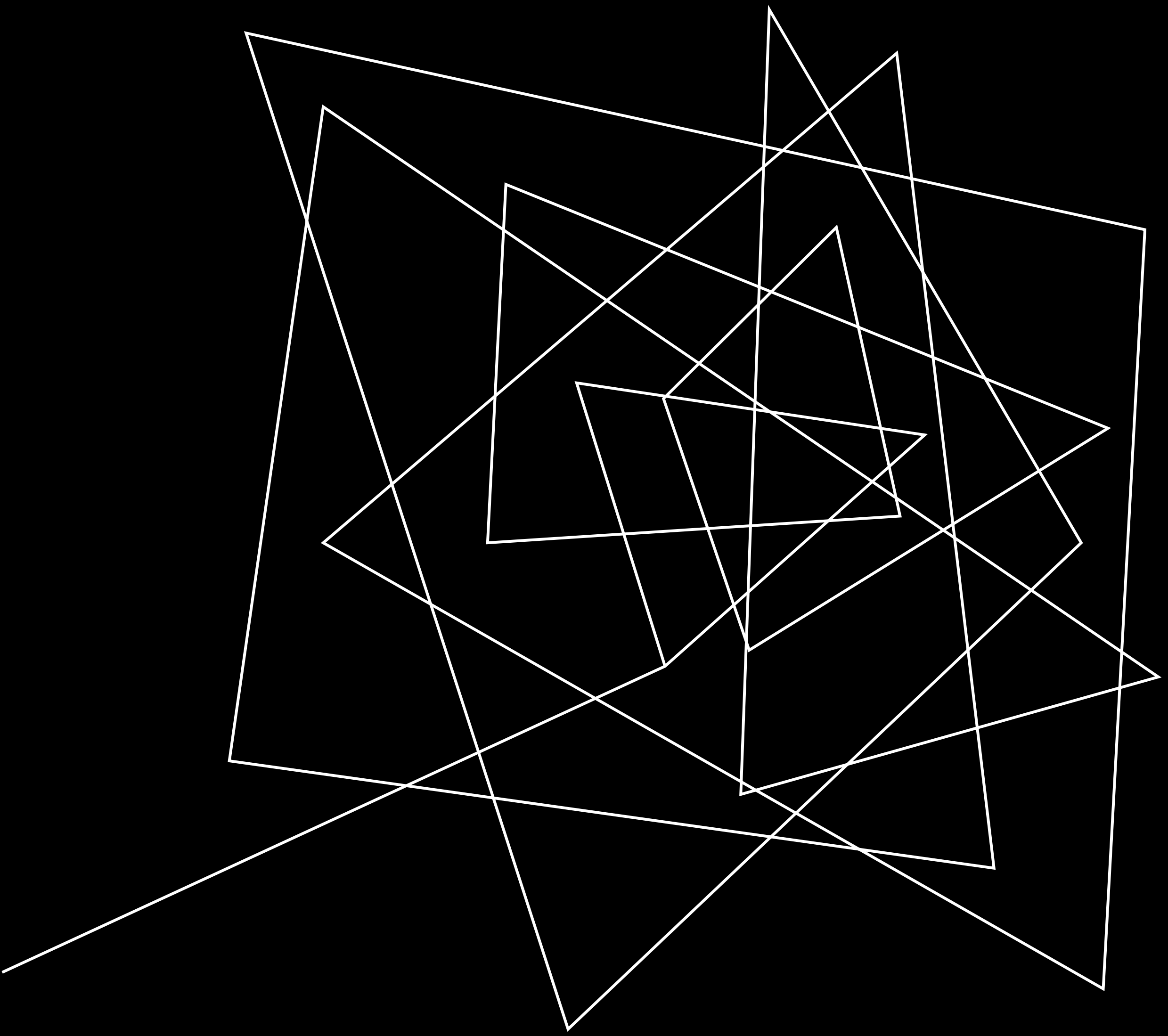
DFT OF RGB IMAGES

However, in RGB, the intensity and color information are mixed across all three channels. This makes it less straightforward to perform frequency-domain operations specific to brightness or color. That's why **RGB images are often converted into alternate color spaces like HSV or LAB before applying a DFT.** After conversion, operations like filtering or DFT are applied specifically to the channel specific to intensity (brightness), which is typically more important for frequency-domain analysis. For **HSV**, that channel is the **Value** channel which represents Brightness. Also, since color channels are untouched (**Hue** in HSV), the original colors remain intact even if intensity transforms are applied using the DFT.

Value Channel (Brightness)

DFT of Value Channel





INVERSE FOURIER TRANSFORM (IDFT)

2D INVERSE FOURIER TRANSFORM (IDFT)

The **2D Inverse Fourier Transform (IDFT)** is the reverse process of the 2D Discrete Fourier Transform (DFT). While the DFT converts an image from the spatial domain (pixel values) into the frequency domain (frequency components), the IDFT converts it back from the frequency domain to the spatial domain. The 2D IDFT is vital for converting processed frequency data back into interpretable images in the spatial domain. Given the 2D DFT of an image, $\boldsymbol{F(u,v)}$, of size $\mathbf{M \times N}$, the **2D IDFT** is given by:

$$\boldsymbol{f(x,y)} = \frac{\mathbf{1}}{\boldsymbol{M \bullet N}} \sum_{\boldsymbol{u}=\mathbf{0}}^{\mathbf{M-1}} \sum_{\boldsymbol{v}=\mathbf{0}}^{\mathbf{N-1}} \boldsymbol{F(u,v)} \mathbf{e}^{\boldsymbol{j}2\pi\left(\frac{\boldsymbol{ux}}{\boldsymbol{M}} + \frac{\boldsymbol{vy}}{\boldsymbol{N}}\right)}$$

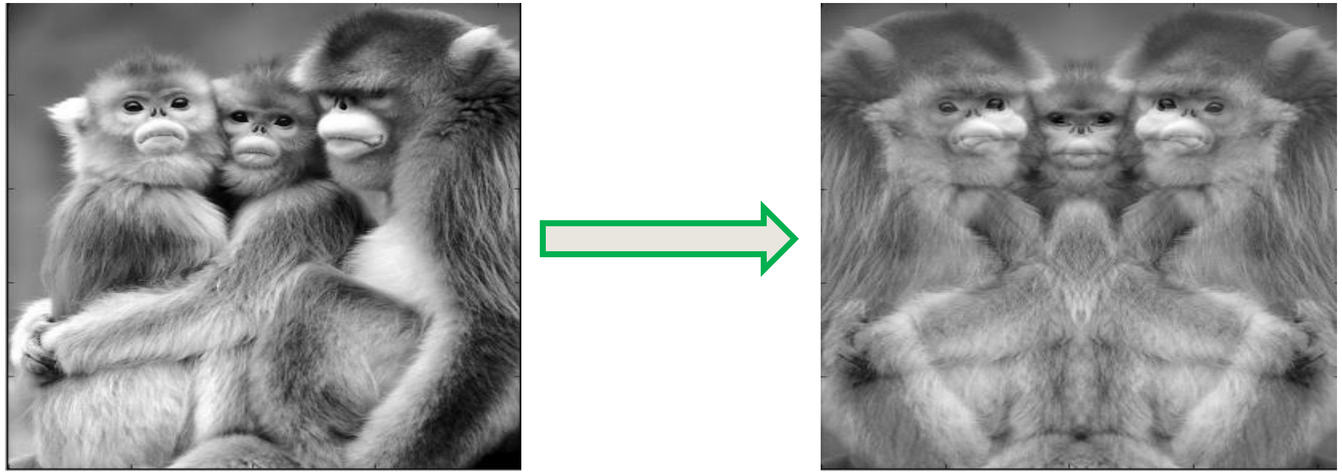
Where,

- $\boldsymbol{f(x,y)}$ is the Intensity at spatial position $(\boldsymbol{x,y})$.
- $\boldsymbol{F(u,v)}$ is Frequency component at $(\boldsymbol{u,v})$.
- $\boldsymbol{u,v}$ are the horizontal and vertical frequency indices, respectively.
- $\mathbf{M,N}$ represent the dimensions of the image (Number of pixels along the $\boldsymbol{x}$ & $\boldsymbol{y}$ axis).

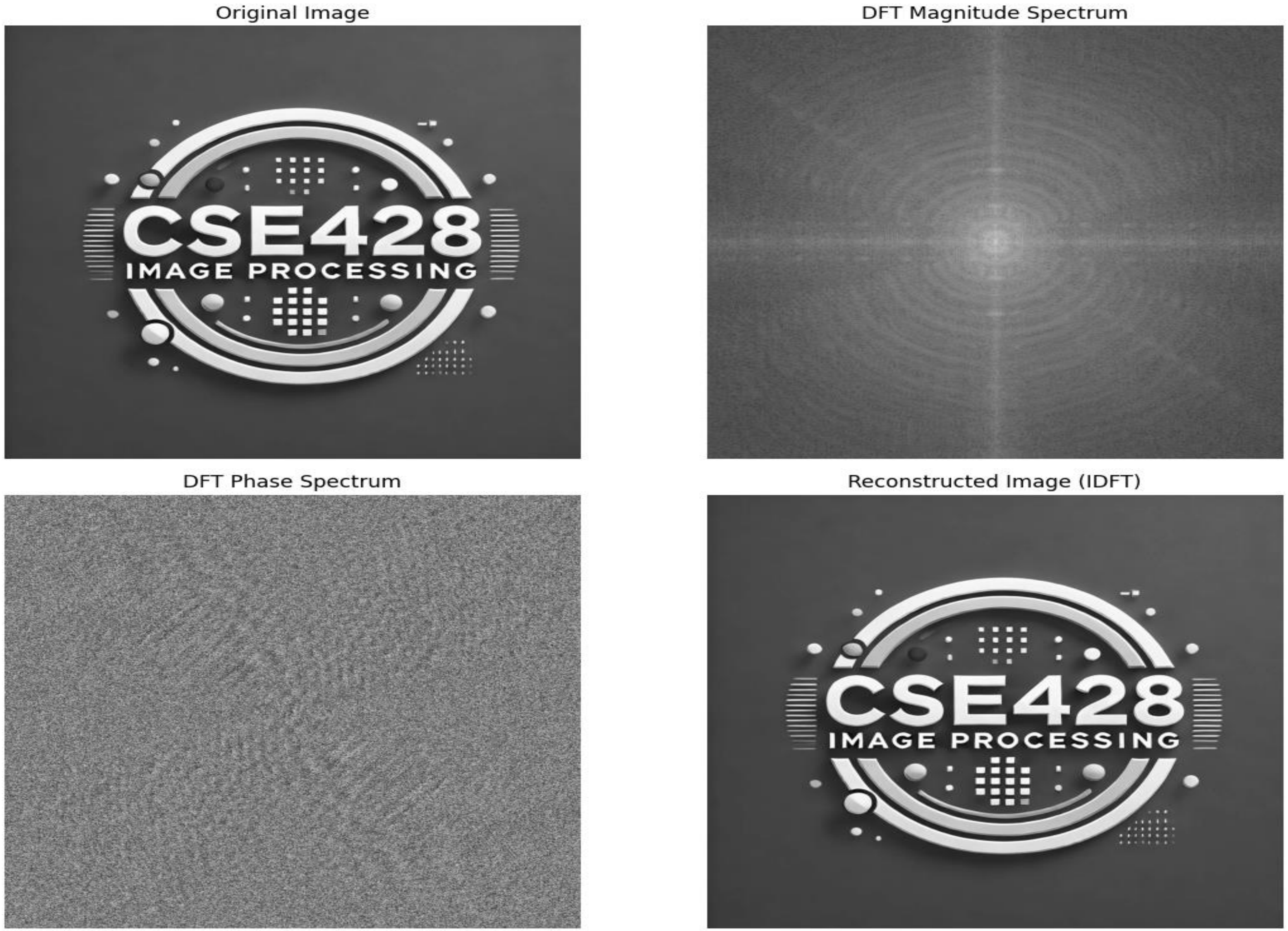
2D INVERSE FOURIER TRANSFORM (IDFT)

Clearly, like the value of 2D DFT $\boldsymbol{F(u,v)}$, the value of **2D IDFT** at each pixel location $\boldsymbol{x, y}$ is also a complex number. However, if the DFT was derived from a real valued signal (e.g., images), then the IDFT should approach a real value too, as the imaginary parts of the reconstructed spatial domain tend to be near zero, leaving the real part dominant.

It should be noted that even though when working with the DFT in image processing, we often work with the magnitude of the DFT and discard the phase information for visualization, **both magnitude and phase are needed for the full reconstruction applying IDFT.** Also, **if the zero-frequency component was shifted to the center it needs to be shifted back to its original position.** Otherwise, you may find out that your reconstructed image has been contaminated like this..



2D INVERSE FOURIER TRANSFORM (IDFT)



SAMPLE CODING IMPLEMENTATION OF DFT & IDFT

CODE SNIPPET TO COMPUTE DFT

```
import numpy as np
import cv2
import matplotlib.pyplot as plt

# Load the image in grayscale
image = cv2.imread('path_to_image.jpg', cv2.IMREAD_GRAYSCALE)

# Compute the DFT
dft = np.fft.fft2(image)  # Compute the 2D Fourier Transform
dft_shift = np.fft.fftshift(dft)  # Shift zero frequency to center

# Compute magnitude spectrum
magnitude_spectrum = 20 * np.log(np.abs(dft_shift) + 1)  # Adding 1 to avoid log(0)

# Plot the original image, magnitude spectrum, and reconstructed image
plt.figure(figsize=(12, 8))
plt.subplot(121), plt.imshow(image, cmap='gray'), plt.title('Original Image')
plt.subplot(122), plt.imshow(magnitude_spectrum, cmap='gray'), plt.title('Magnitude Spectrum')
plt.show()
```

CODE SNIPPET TO COMPUTE IDFT

```
import numpy as np
import cv2
import matplotlib.pyplot as plt

# Compute the IDFT
idft_shift = np.fft.ifftshift(dft_shift)  # Shift zero frequency back to original
idft = np.fft.ifft2(idft_shift)  # Compute the Inverse Fourier Transform
idft_image = np.abs(idft)  # Take the magnitude to get the image

# Plot the original image, magnitude spectrum, and reconstructed image
plt.figure(figsize=(12, 8))
plt.subplot(131), plt.imshow(image, cmap='gray'), plt.title('Original Image')
plt.subplot(132), plt.imshow(magnitude_spectrum, cmap='gray'), plt.title('Magnitude Spectrum’)
plt.subplot(133), plt.imshow(idft_image, cmap='gray'), plt.title('Reconstructed Image’)
plt.tight_layout()
plt.show()
```

EXAMPLE PROBLEMS

EXAMPLE 1

Compute the **2D Discrete Fourier Transform (DFT)** for the following $3X3$ images. Present the results as matrices of complex numbers and provide the **magnitude spectrum** of the transformed images for visualization. **Comment** about the Frequency contents of these images. Can you relate them to any of the spatial image processing techniques?

2.

Image 1

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

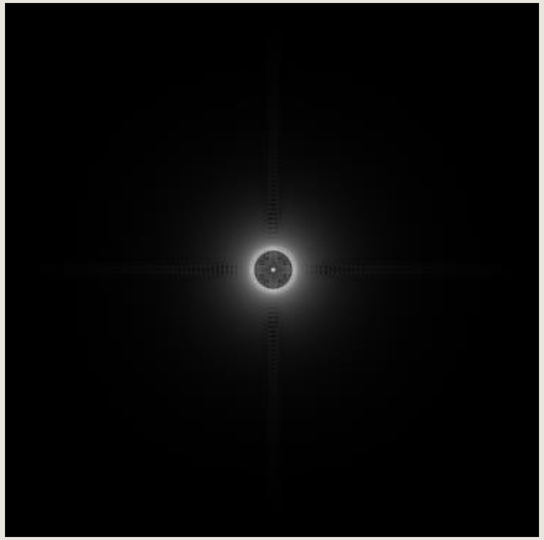
Image 2

$$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$

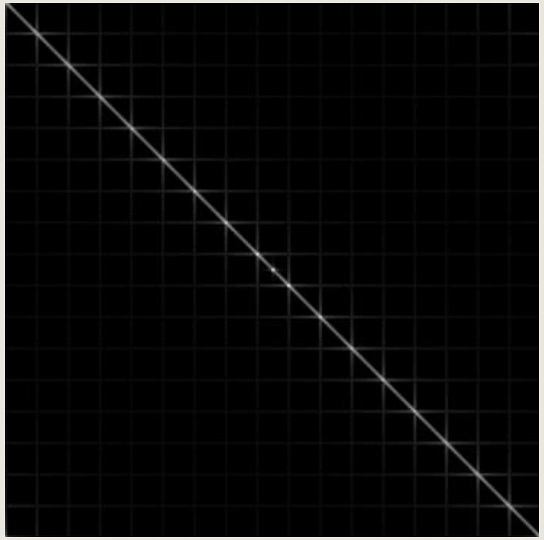
EXAMPLE 2

Observe the image below. The radial gradients follow sinusoidal pattern. Which DFT on the right do you think is the frequency domain representation of this image?

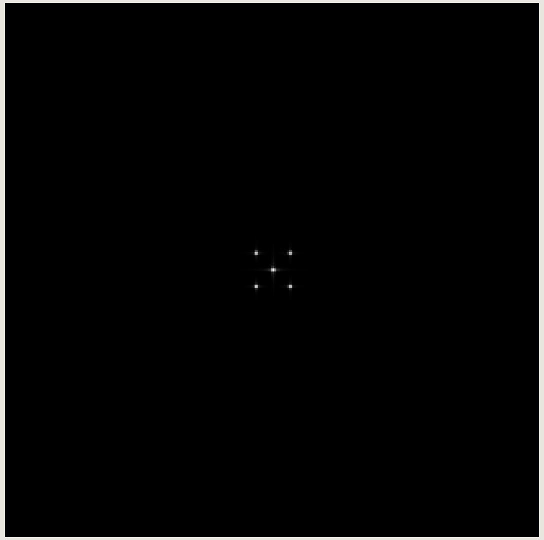
A



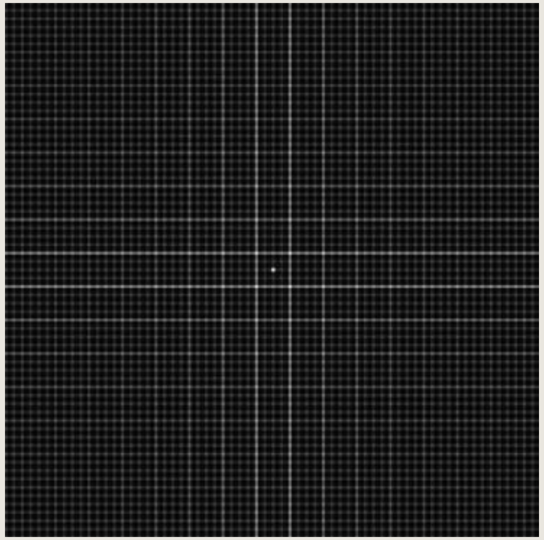
B



C

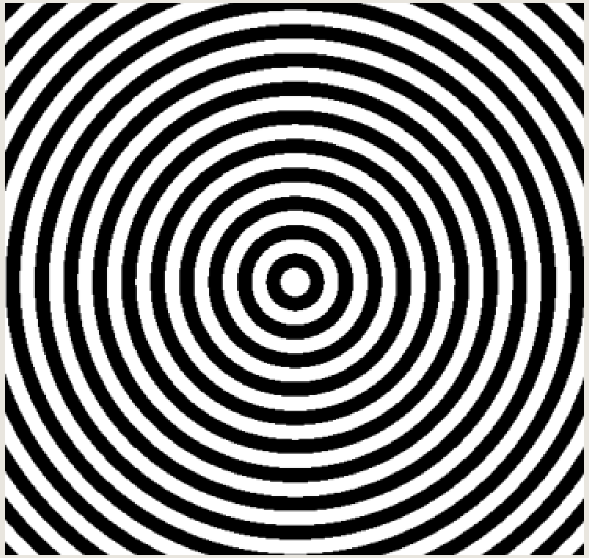


D

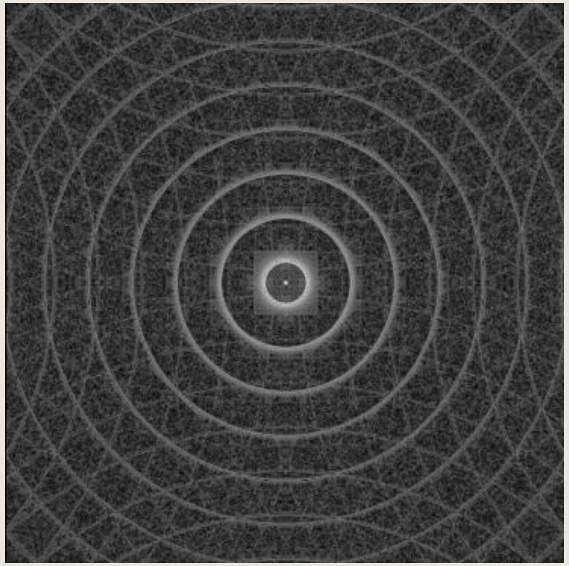


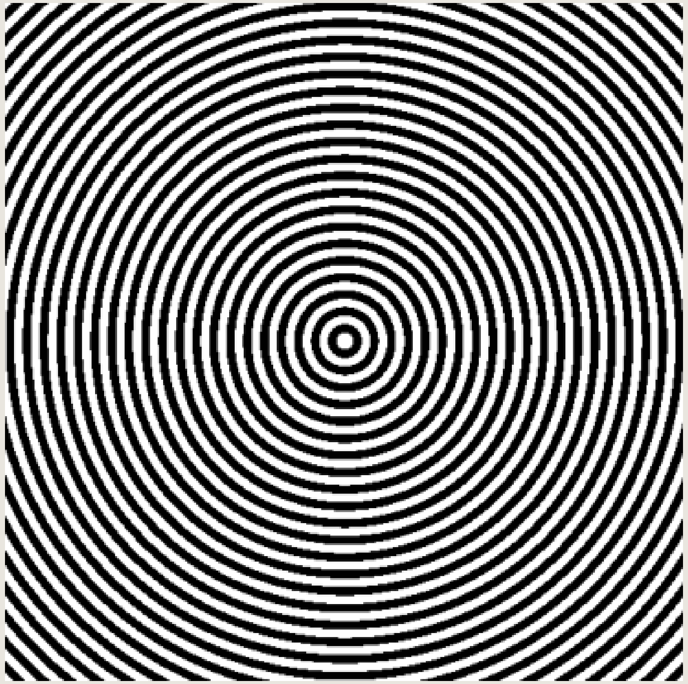
EXAMPLE 3

Observe the image **on the right** and its **DFT**. Which DFT below them (**A** or **B**) do you think is the frequency domain representation of the image **below this text**?

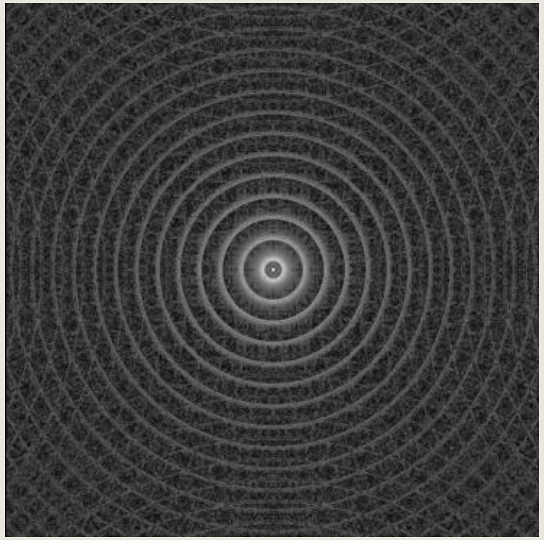


DFT

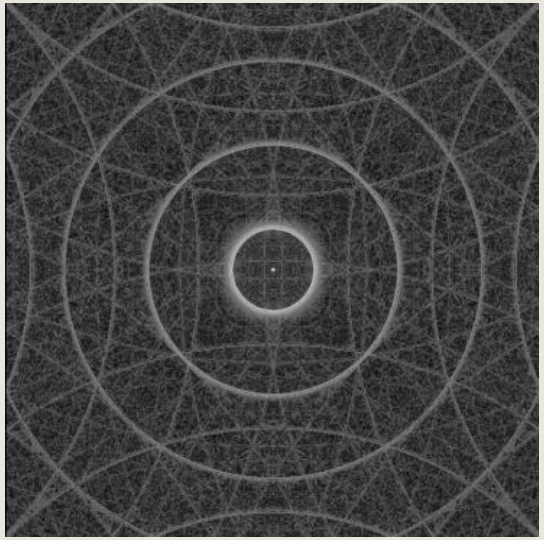




A



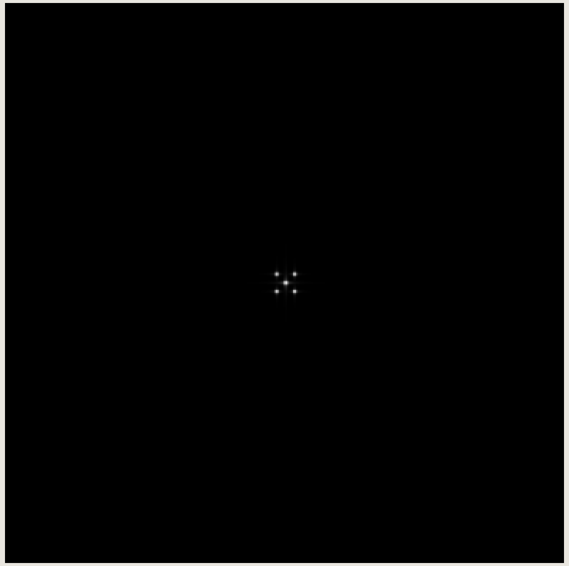
B

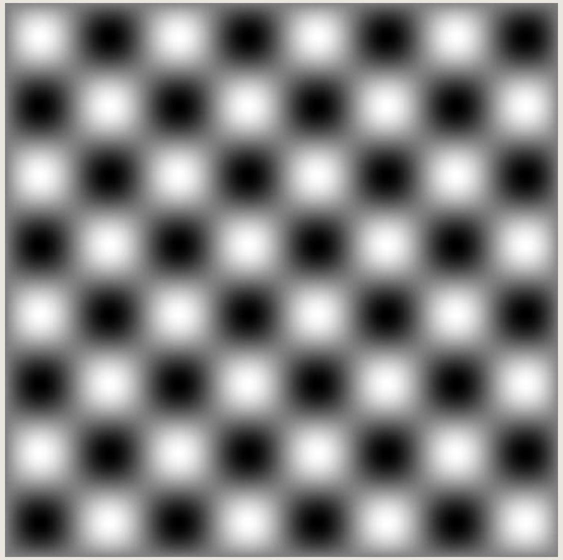


EXAMPLE 4

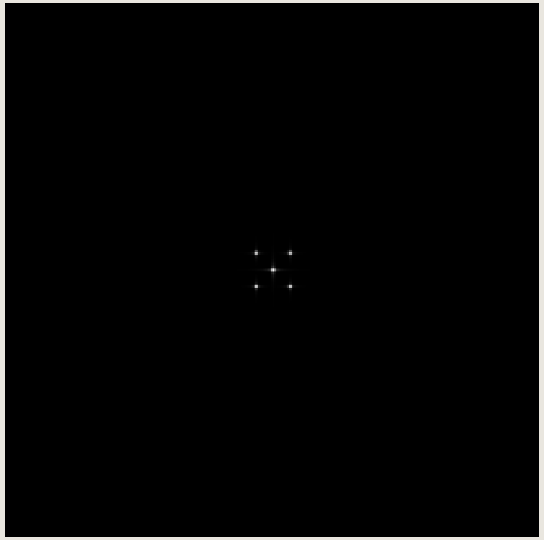
Observe the image **on the right** and its **DFT**. Which DFT below them (**A** or **B**) do you think is the frequency domain representation of the image **below this text**?

DFT

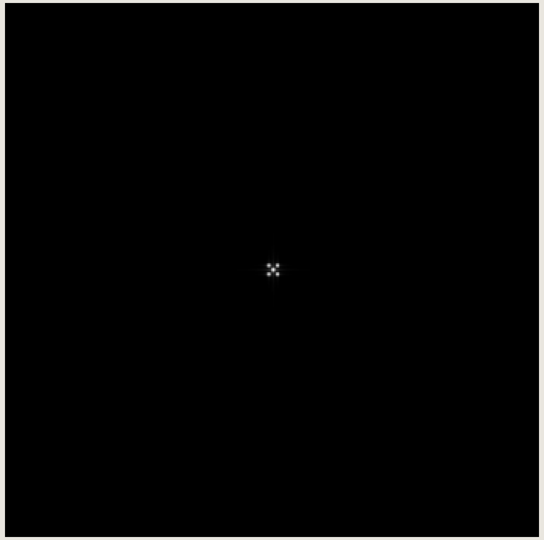




A



B

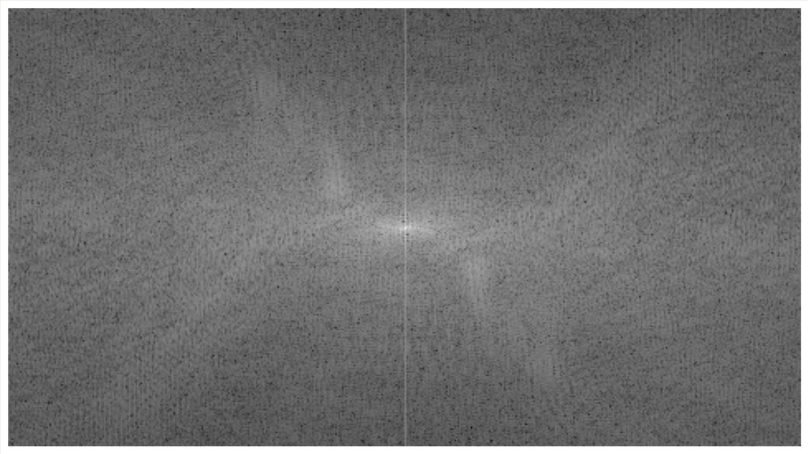


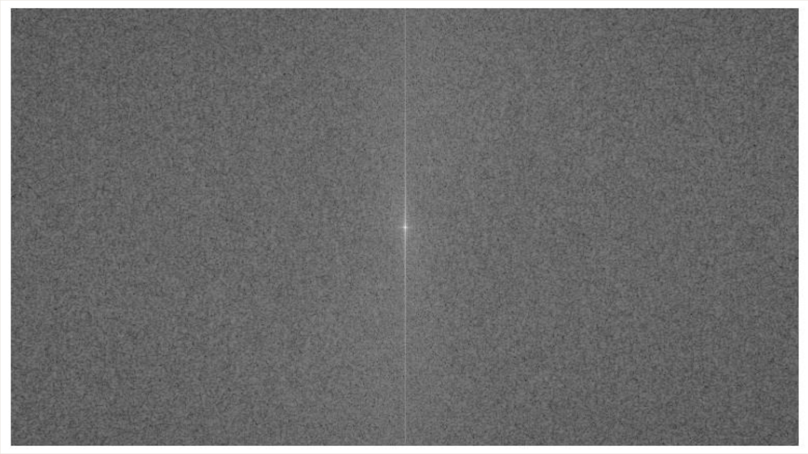
EXAMPLE 5

Observe the images **on the left.** Their **DFTs** are **on the right** column but we don't know which DFT corresponds to which image. Map them to their DFTs on the right using an arrow and explain your choices.

Images

DFTs





REFERENCES

1. Image Processing Handbook – J. Russ, CRC Press, 2007.
2. Digital Image Processing – R. C. Gonzalez and R. E. Woods, Pearson, 2008.
3. OpenCV Documentation on Color Spaces: https://docs.opencv.org/2.4/modules/imgproc/doc/miscellaneous_transformations.html
4. https://sbme-tutorials.github.io/2018/cv/notes/3_week3.html

THANK YOU