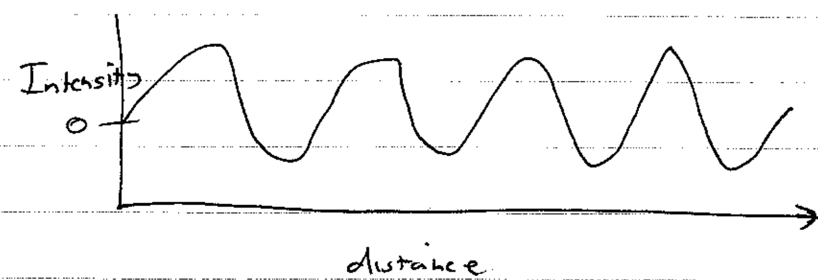
The Frequency Domain and Fourier Transforms

We have represented all of the images we have seen so far as intensity functions of location, $f(x, y)$. But there is an equally useful point of view, that of viewing images as a function of spatial frequency instead of position, that will prove to be very useful for many image processing applications.

Let's begin by looking at one-dimensional functions, and then extend the results to 2-D images, to keep things simple at first. The multidimensional extension will be ~~be~~ easy!

Time vs. frequency

Suppose we want to describe an intensity that varies periodically with distance (or time), that we might plot as follows:



In fact the equation for this curve might be

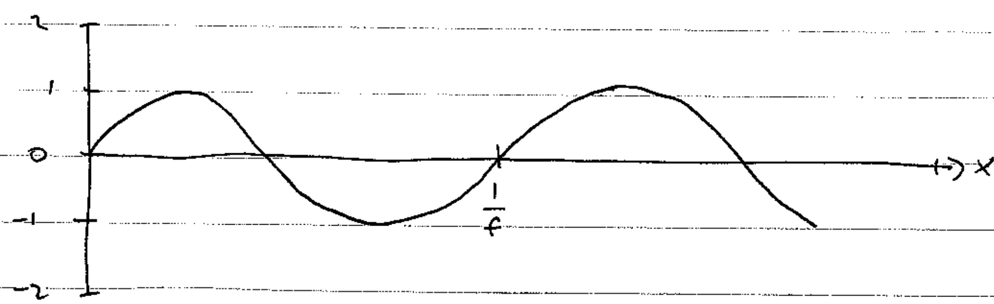
$$I(x) = A \cdot \cos(2\pi f x + \phi)$$

where A is the amplitude (or maximum value) of the "wave", f is the frequency in cycles per unit distance (or time), x is the location along the axis, and ϕ is the starting phase, that is where in the cycle the curve starts.

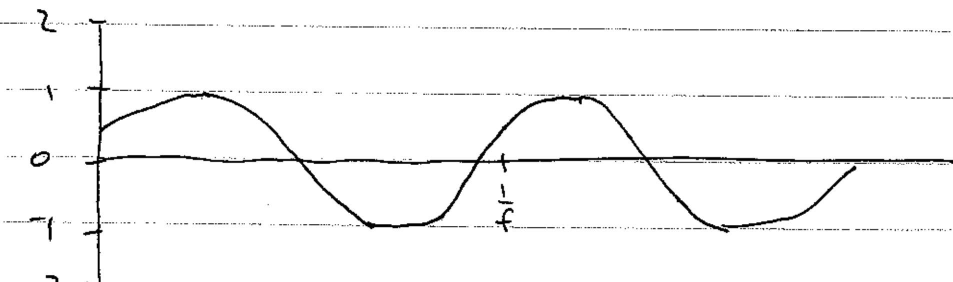
The above equation gives the magnitude of the signal as a function of several parameters. Suppose we want to represent numerically for a computer how the intensity changes at each position. We could form a vector $X = (x_0, x_1, x_2, \ldots)$ that gives the magnitude of the signal at each point, and this vector requires entries at each sample position. Alternatively, we could assume that the function has the cosinusoidal form of the above equation, and then we would instead need to know the values of A, f, and ϕ. With knowledge of these we can calculate the magnitude without having to know the value of the function at all locations.

Hence we have two equivalent methods of representing a cosine signal, either a list of the function at each position in space or time, or the amplitude, frequency, and phase of the signal. We call the first method a time-domain representation, as it lists values as a function of time (space), and the second is a frequency-domain representation, as it requires knowledge of the frequency of the signal.

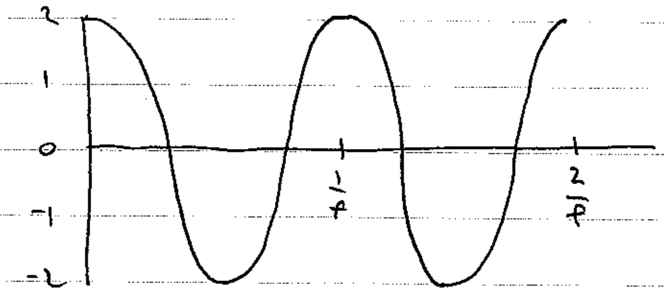
Let's familiarize ourselves with several signals of differing amplitude, frequency, and phase:



$y = \sin(2\pi f x)$ [amplitude = 1, freq = f, $\phi = 0$]



$y = \sin(2\pi f x + \frac{\pi}{8})$ [$\phi = \frac{\pi}{8} = 45°$]



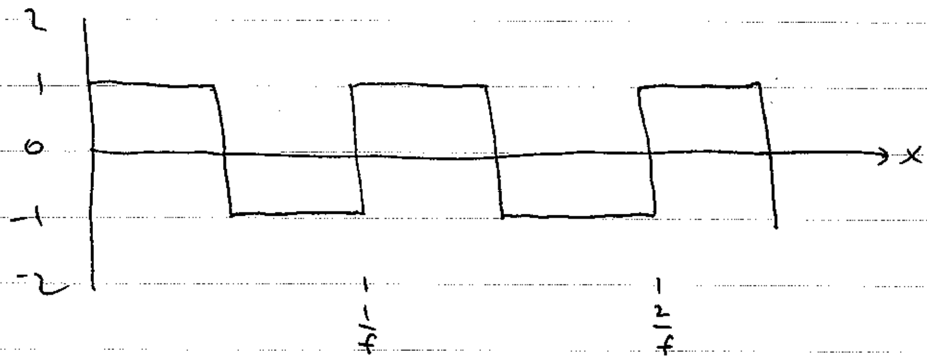
$$y = 2\sin\left(2\pi f x + \frac{\pi}{2}\right) = 2\cos(2\pi f x)$$

Power Spectrum

Now, most images are more complicated than a single sine or cosine wave. So, what is the relevance of frequency-domain representations for general images? We can represent any image in either the time or frequency domain, and sometimes it is very helpful to use the frequency domain viewpoint to discern or enhance image features.

The usefulness of frequency representations can be first appreciated by examining the decomposition of a signal or image into its frequency components. Because a complete set of cosines and sines of various frequency are sufficient to span space, and because each is orthogonal to the others, there is a unique frequency mapping that goes back and forth between the two domains. That is, an arbitrary image may be described as a matrix of points in the space domain, or as an ordered list of frequency amplitude and phase coefficients.

One example of this is to consider just the amplitude of the frequency coefficients, which we denote the power spectrum of a signal (again, or image). Consider a repeating square wave signal as in the following graph:



This signal mimics the sine signal we looked at earlier, but has sharp edges. So what is the frequency domain representation of the square wave?

We can calculate the frequency-domain coefficients of the square wave by deriving its Fourier series, which gives the magnitude and phase of each frequency component. For the square-wave signal $s(x)$, each coefficient f_i is found by

$$F_i = 2f\int_0^{1/f} s(x)\sin(i2\pi fx)\,dx$$

$$= 2f\int_0^{1/2f} \sin(i2\pi fx)\,dx - 2f\int_{\frac{1}{2f}}^{1/f} \sin(i2\pi fx)\,dx$$

(For F_0 we ~~don't~~ use ~~the~~ factor of 2!) The first few terms are

$F_1 = 1.27324$

$F_3 = 0.42420$

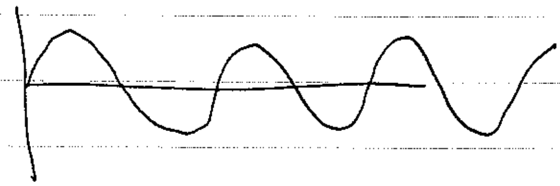
$F_5 = 0.\cancel{1819}\ 0.2546$

$F_7 = 0.1819$, and so on.

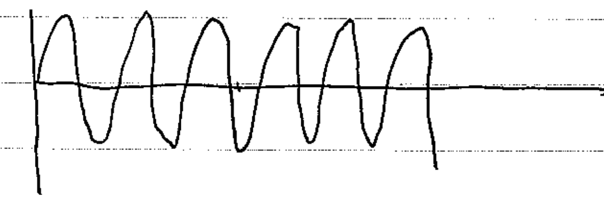
All even coefficients are zero.

What does this do for us? Consider that if we know the coefficients, we can reconstruct the original signal. Let's take the set of sine waves and add them, with each weighted by its coefficient:

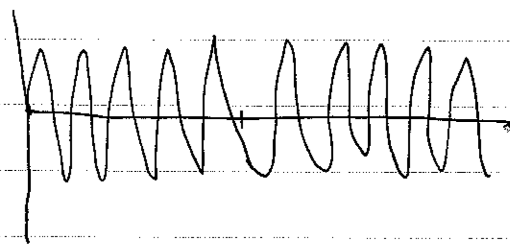
$F_1 = 1.273$, $\sin(2\pi f x) =$



$F_3 = 0.424$, $\sin(3 \cdot 2\pi f x) =$



$F_5 = 0.254$, $\sin(5 \cdot 2\pi f x) =$

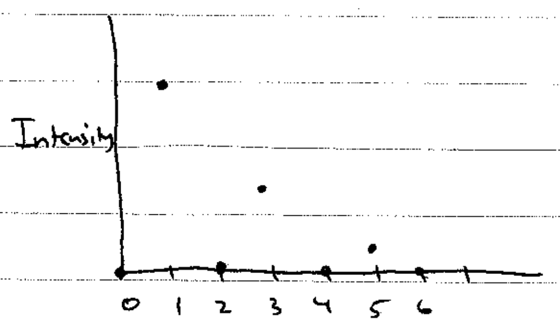


and so on. But if we form the sum

$$F_1 \cdot \sin(2\pi f x) + F_3 \cdot \sin(3 \cdot 2\pi f x) + F_5 \cdot \sin(5 \cdot 2\pi f x) + \cdots$$

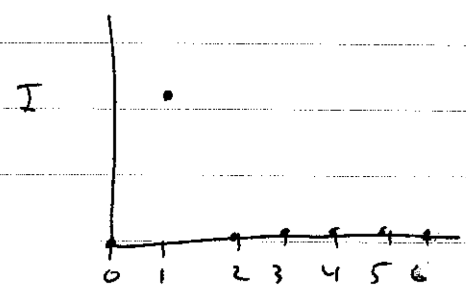
we see that we obtain our original square wave.

Now, we can plot these coefficients as a function of frequency:



This plot is called the power spectrum of the square wave, showing the relative intensity of each frequency coefficient.

Contrast this with the power spectrum of ~~of~~ the single sine wave:

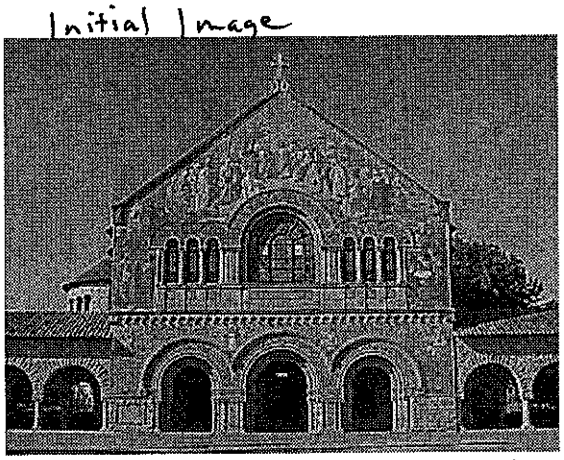


Here every coefficient but the first is zero. We can see that the addition of the higher order coefficients transforms the smoothly-varying sine function into the very sharp-edged square wave. The low frequency terms give the overall shape of the function, while the sharpness and predominance of edges are dependent on the high-frequency components.

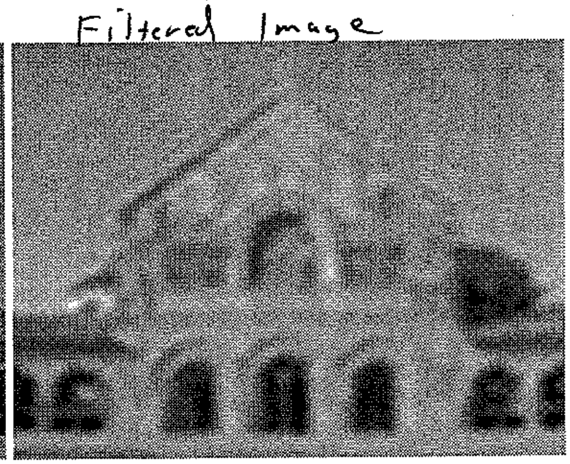
Let's apply this generalization to an image: Start with an image with a lot of fine detail, and remove the high-frequency parts by

1) Transform the image to its frequency-domain representation
2) Zero out coefficients greater than some frequency
3) Reconstruct the image from only the low-frequency parts.

Initial Image



Filtered Image



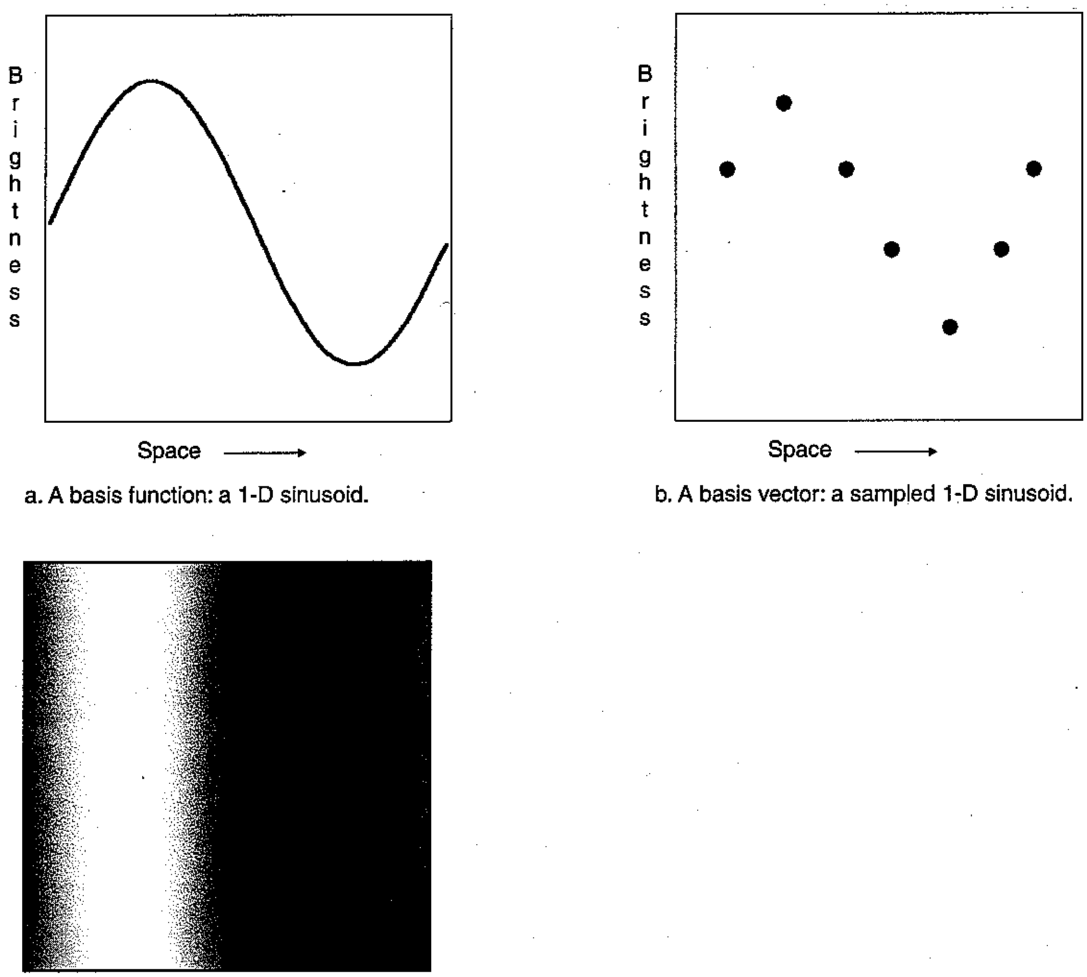
Power Spectrum of an Image

How do we present a power spectrum for a two-dimensional image, where the input data form a matrix rather than a 1-D sequence? We can easily generalize the one-d set of sines and cosines to a set of 2-D functions. Begin with a 2-D function that varies only in one direction, so that we have a function such as

$$f(x,y) = \sin(2\pi f x + \phi)$$

Here the brightness changes sinusoidally left to right but not at all as a function of y (up and down), such as the following:

Figure 2.5-2 Basis Vectors and Images



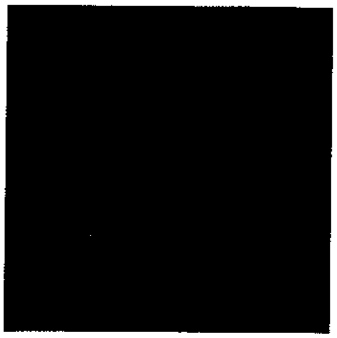
a. A basis function: a 1-D sinusoid.

b. A basis vector: a sampled 1-D sinusoid.

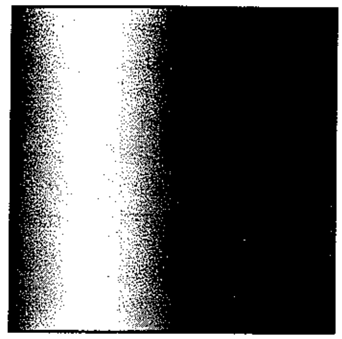
c. A basis image: a sampled sinusoid shown in 2-D as an image. The pixel brightness in each row corresponds to the sampled values of the 1-D sinusoids, which are repeated along each column.

Then we can look at variations in frequency or even variations in the sharpness of the edges similarly to the way did for a sequence:

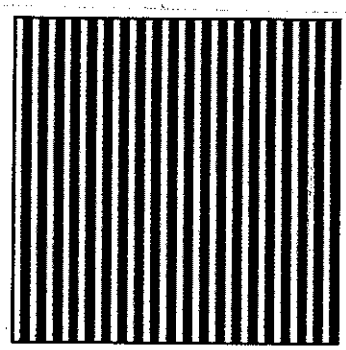
Figure 2.5-3 Spatial Frequency



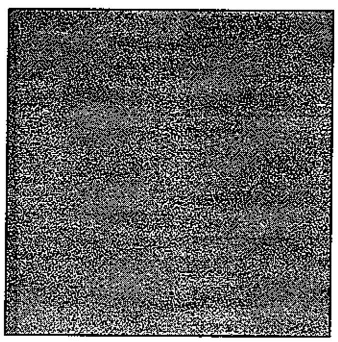
a. Frequency = 0, gray level = 54.



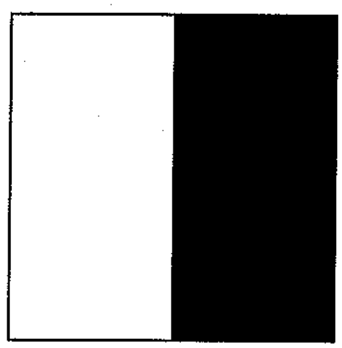
c. Frequency = 1, horizontal sine wave.



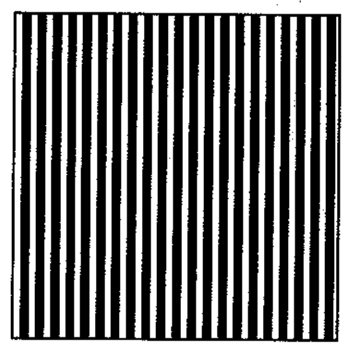
e. Frequency = 20, horizontal sine wave.



b. Frequency = 0, gray level = 202.



d. Frequency = 1, horizontal square wave.

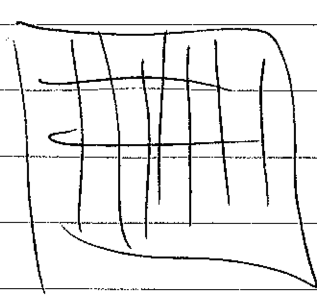


f. Frequency = 20, horizontal square wave.

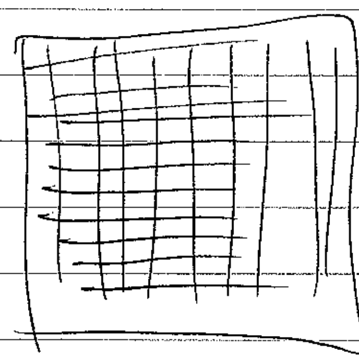
Finally, we can let the function vary in both directions at once:

$$f(x,y) = \sin(2\pi f_x x + \phi_x) \cdot \sin(2\pi f_y y + \phi_y)$$

and we might see functions that look like:



or



Now, we can derive a 2-D Fourier series for each of these simple, repeating patterns. To be completely general, we want to calculate the coefficient for each possible combination of frequencies in both directions, so that

$$F_{ij} = f^2 \int_0^{1/f} \int_0^{1/f} f(x,y) \sin(2\pi i f_x x) \sin(2\pi j f_y y)\, dx\, dy$$

might give ~~us~~ a set of coefficients for a set of sine waves, with a similar definition for repetitive waveforms for cosines. In fact, to be completely general, we usually use exponentials for the basis vectors and can then express arbitrary waveforms as expansions using the following

$$\boxed{F_{u,v} = \int_0^{X_{max}} \int_0^{Y_{max}} f(x,y)\, e^{-j2\pi(ux+vy)}\, dx\, dy}$$

where now we have explicitly introduced variables u and v for frequencies in the x and y directions, respectively. Note also the use of the complex exponential kernel function $e^{-2\pi j(ux+vy)} = \cos(2\pi ux + 2\pi vy) + j \sin(2\pi ux + 2\pi vy)$ That is now a complex quantity, so that any real or complex quanity can be represented. Note also that we are no longer restricted to repeating waveforms, although the specific form we have written above implicitly assumes that the function repeats with x and y cycle lengths X_{max} and y_{max}, respectively.

This particular form of generating frequency coefficients is called a Fourier transform, because it transforms a time- (or space) domain signal into its frequency components.

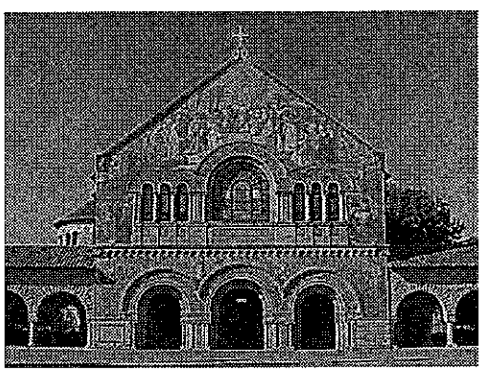
Thus, we can plot a new image consisting of how intensity of an image is distributed in frequency by creating a new image as a function of the frequency variables u and v, and plot either the magnitude or phase of the frequency coefficients as a new image. Because of the complex kernel, even real functions have both real and imaginary components for every frequency coefficient $F_{u,v}$. If $R(u,v)$ is the real part of $F(u,v)$ and $I(u,v)$ is the imaginary part, then we define the magnitude as

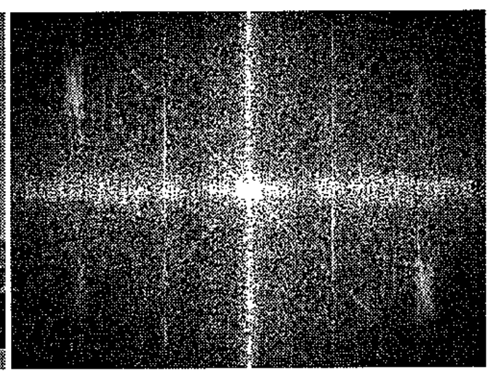
$$|F(u,v)| = \sqrt{R(u,v)^2 + I(u,v)^2}$$

and the phase Φ as

$$\Phi(u,v) = \tan^{-1} \frac{I(u,v)}{R(u,v)}$$

A sample image and its transform magnitude are





where the brightness at each location u,v is the magnitude of the frequency coefficient for frequencies u,v. Note that the greatest power is found in frequencies near zero, plotted in the center of this plot. Because frequencies can be either positive or negative, we put zero in the middle for now. More on this later.