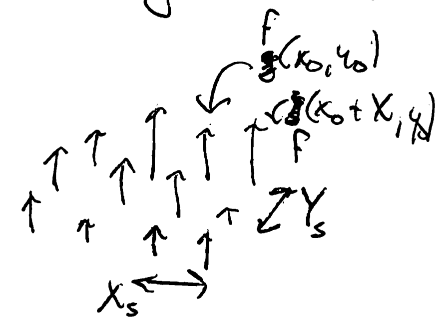
WHITTAKER SHANNON SAMPLING THEOREM

Often, we don't know the full continuous function $f(x,y)$, but instead only have a sampled version of it $f_s(x,y)$.

A 1D example is measuring a time signal at uniformly spaced time intervals. A 2D example would be a digital image from your cell phone camera, where each pixel in the sensor samples the local irradiance of the scene imaged onto it.

A simplified model of this process is

$$f_s(x,y) = \mathrm{comb}\left(\frac{x}{X_s}\right)\mathrm{comb}\left(\frac{y}{Y_s}\right) f(x,y)$$

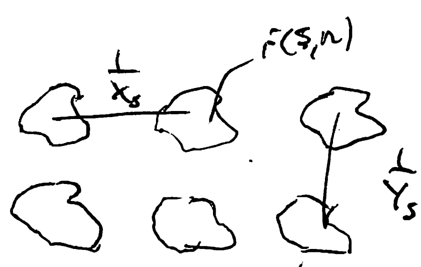


This is just an array of delta functions space by X and Y in the x and y directions, respectively. The areas of each delta function corresponds to the local value of $f(x,y)$.

Fourier transforming gives

$$F_s(\xi,\eta) = X_s Y_s\, \mathrm{comb}(X_s \xi)\, \mathrm{comb}(Y_s \eta) ** F(\xi,\eta)$$

and from the convolution theorem

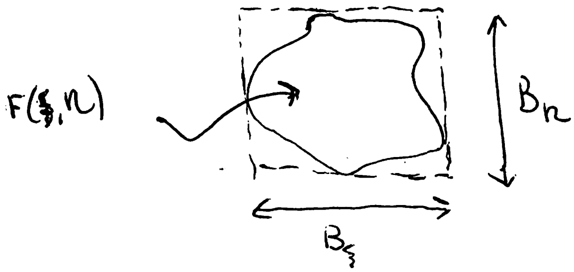


$$F_s(\xi,\eta) = \sum_{N,m} F\left(\xi - \frac{N}{X_s}, \eta - \frac{m}{Y_s}\right)$$

so the spectrum of a sampled signal looks like a bunch of shifted versions of the continuous spectrum $F(\xi,\eta)$, with spacings of $\frac{1}{X_s}$ and $\frac{1}{Y_s}$ in the ξ and η directions, respectively.

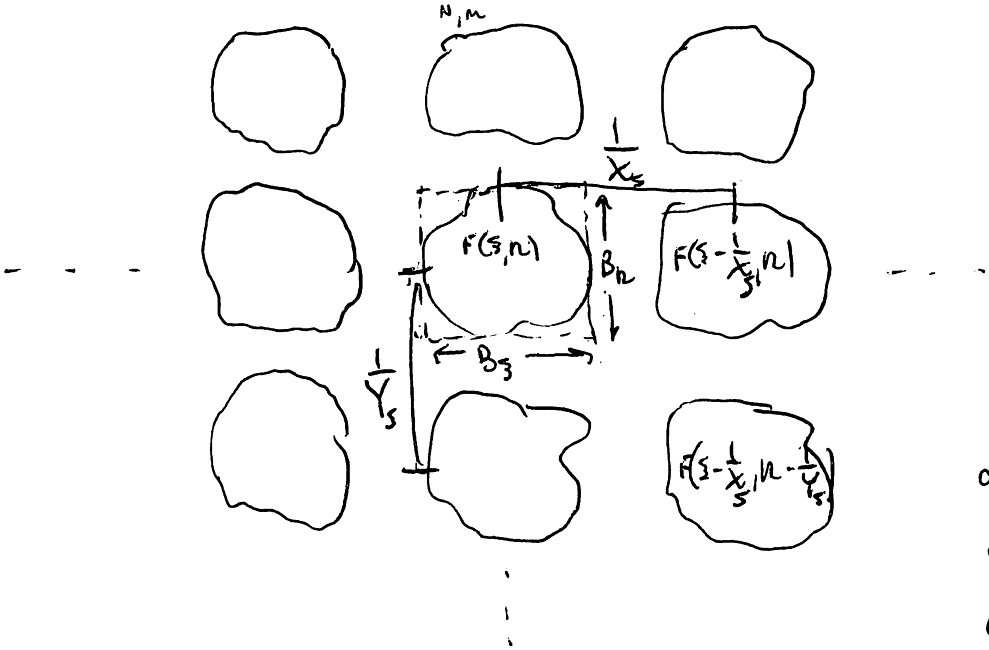
Band-limited Functions

The spectrum of a band-limited function is only non-zero over a finite region of Fourier space.



The spectrum of the sampled version looks like

$$F_s(\xi,\eta) = \sum_{n,m} F\left(\xi - \frac{n}{X}, \eta - \frac{m}{Y}\right)$$



Let's now multiply the sampled spectrum by a mask that blocks all the outside spectra and passes just the center one. $\text{Rect}\left(\frac{\xi}{B_\xi}, \frac{\eta}{B_\eta}\right)$ will work

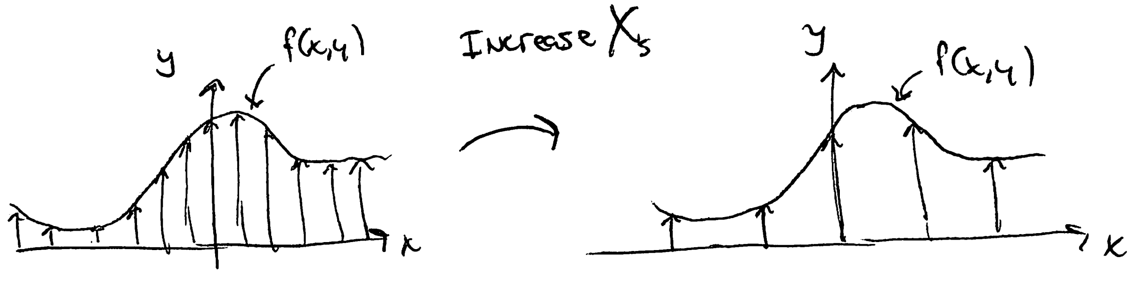
$$F_s(\xi,\eta)\,\text{rect}\left(\frac{\xi}{B_\xi}, \frac{\eta}{B_\eta}\right) = F(\xi,\eta)$$

Inverse transform both sides

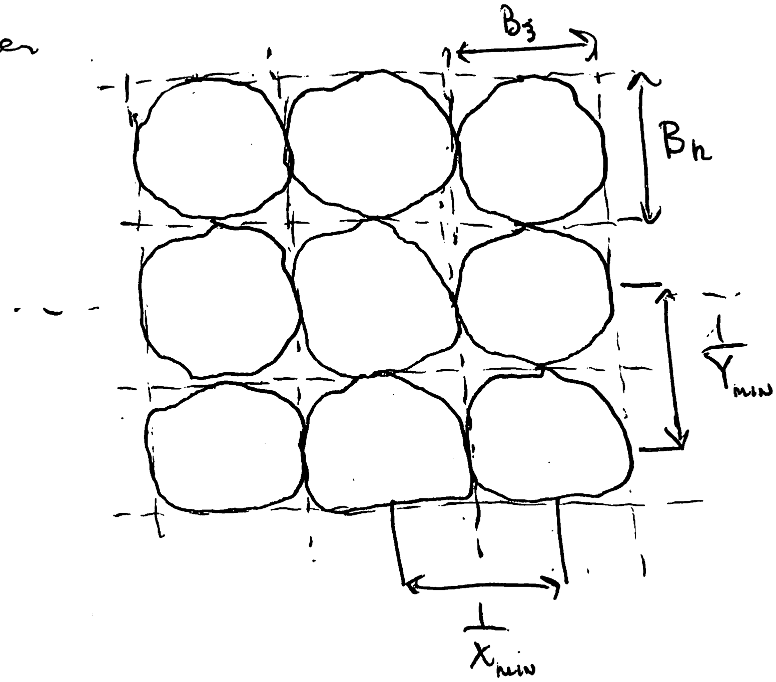
$$f(x,y) = \mathcal{F}^{-1}\left\{F_s(\xi,\eta)\,\text{rect}\left(\frac{\xi}{B_\xi}, \frac{\eta}{B_\eta}\right)\right\}$$

Perfect reproduction of continuous object $f(x,y)$

In general, the mask can have any shape as long as it passes the center spectrum and blocks the remaining spectra, What happens when X and Y are ~~not~~ increased? First, this means that the separation between samples in (x,y) space becomes larger.



In Fourier space, increased X_s and Y_s means spectra get closer together



There is a minimum sampling frequency $\frac{1}{X_{min}}$ and $\frac{1}{Y_{min}}$ where the spectra do <u>not</u> overlap

$$\frac{1}{X_{min}} = B_\xi \qquad \frac{1}{Y_{min}} = B_\eta$$

~~These are called the Nyquist frequencies~~

So to avoid overlap for a band-limited function

$$X_s \leq \frac{1}{B_\xi} \quad \text{and} \quad Y_s \leq \frac{1}{B_\eta}$$

The <u>Nyquist</u> Frequencies are defined as

$$N_\xi = \frac{B_\xi}{2} \qquad N_\eta = \frac{B_\eta}{2}$$

which are just the maximum <u>absolute</u> frequencies of the spectra.

i.e. spectra goes from $-\frac{B_\xi}{2} \le \xi \le \frac{B_\xi}{2}$, so $|\xi| \le \frac{B_\xi}{2} \le N_\xi$

which means

$$\frac{1}{X_{min}} = 2N_\xi \qquad \frac{1}{Y_{min}} = 2N_\eta$$

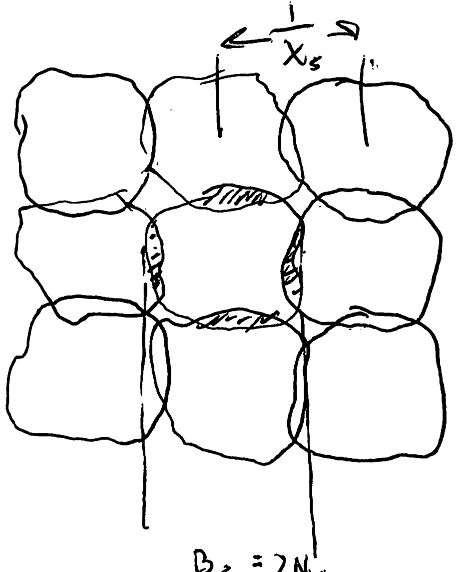
So, you will often hear

$$\frac{1}{X} \ge 2N_\xi \quad \text{and} \quad \frac{1}{Y} \ge 2N_\eta$$

as "the sampling frequency needs to be at least twice the Nyquist frequency."

<u>ALIASING</u> - what happens when the sampling frequency is less than twice the Nyquist frequency?

i.e. $\frac{1}{X_s} < 2N_\xi$ and $\frac{1}{Y_s} < 2N_\eta$

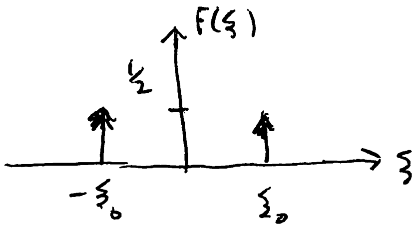


The spectra begin to overlap, so there is no mask that can isolate just one of the spectra. Consequently, the original function $f(x,y)$ cannot be perfectly recovered.

Artifacts in the recovered function will appear and these artifacts are referred to as aliasing.

Example $\quad f(x) = \cos(2\pi \xi_0 x)$ start with 1D example

$$F(\xi) = \frac{1}{2}\left(\delta(\xi - \xi_0) + \delta(\xi + \xi_0)\right)$$



Band-limited since $F(\xi) = 0$ for $|\xi| > \xi_0$

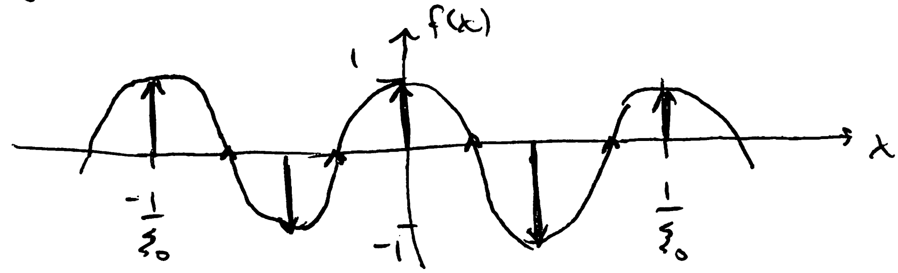
Nyquist frequency $= \xi_0$

$B_\xi = 2\xi_0$

Let's start of sampling at $4\xi_0$ four times the Nyquist frequency

so $\dfrac{1}{X_s} = 4\xi_0$

$X_s = \dfrac{1}{4\xi_0}$



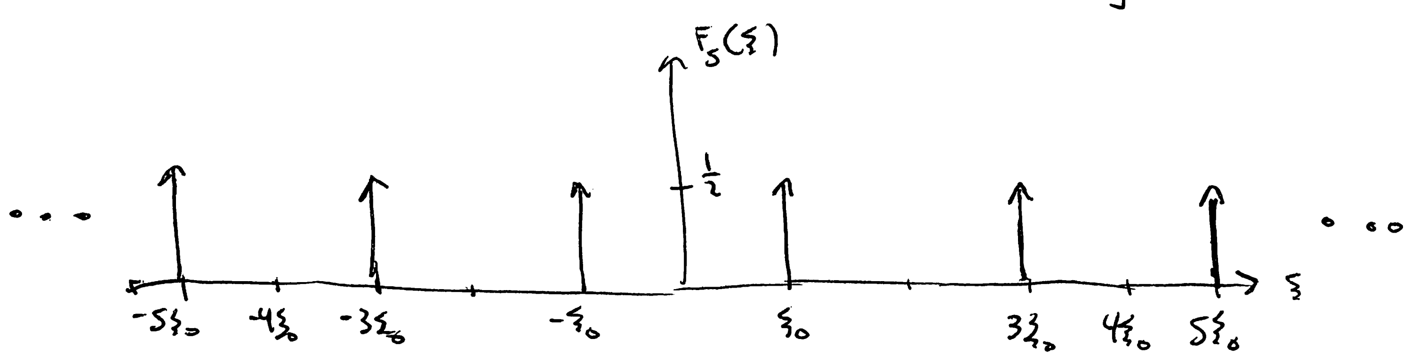
$$f_s(x) = \text{comb}(4\xi_0 x)\cos(2\pi \xi_0 x)$$

$$F_s(\xi) = \frac{1}{4\xi_0}\,\text{comb}\left(\frac{\xi}{4\xi_0}\right) * \frac{1}{2}\left[\delta(\xi - \xi_0) + \delta(\xi + \xi_0)\right]$$

$$F_s(\xi) = \sum_{N=-\infty}^{\infty} \frac{1}{2}\left[\delta(\xi - \xi_0) + \delta(\xi + \xi_0)\right] * \delta(\xi - 4N\xi_0)$$

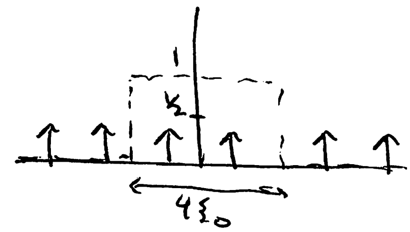
$$F_s(\xi) = \sum_{N=-\infty}^{\infty} \frac{1}{2}\left[\delta(\xi - 4N\xi_0 - \xi_0) + \delta(\xi - 4N\xi_0 + \xi_0)\right]$$

$$F_s(\xi) = \sum_{N=-\infty}^{\infty} \frac{1}{2}\left[\delta(\xi - (4N+1)\xi_0) + \delta(\xi - (4N-1)\xi_0)\right]$$



Now let's multiply $F_s(\xi)$ by a mask to capture the center spectra. In general, we don't know the maximum ~~spectr~~ frequency ξ_0 of the signal. Instead we only know the sampling frequency $4\xi_0$. If the signal is band-limited, and we're sampling at $\frac{1}{x_s} \geq N_\xi$, then multiplying be $\mathrm{rect}\left(\frac{\xi}{4\xi_0}\right)$ the best option to recover the original signal $f(x)$.

$$F_s(\xi)\,\mathrm{rect}\left(\frac{\xi}{4\xi_0}\right) = \frac{1}{2}\left[\delta(\xi-\xi_0) + \delta(\xi+\xi_0)\right]$$

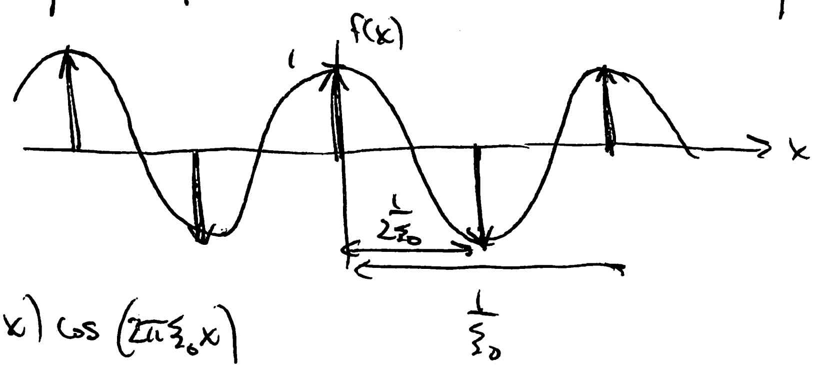


$$f(x) = \mathcal{F}^{-1}\left\{ F_s(\xi)\,\mathrm{rect}\left(\frac{\xi}{4\xi_0}\right)\right\} = \cos(2\pi\xi_0 x)$$ Perfect recovery

Let's reduce the sampling frequency to $2\xi_0$ or twice the Nyquist frequency.

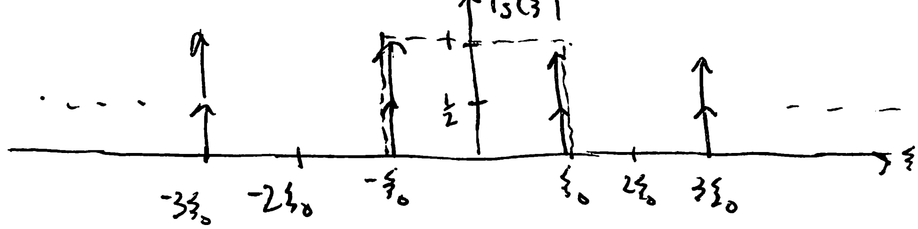
So $\frac{1}{X_s} = 2\xi_0$

or $X_s = \frac{1}{2\xi_0}$



$$f_s(x) = \mathrm{comb}(2\xi_0 x)\cos(2\pi\xi_0 x)$$

$$F_s(\xi) = \sum_{N=-\infty}^{\infty} \frac{1}{2}\left[\delta(\xi-(2N+1)\xi_0) + \delta(\xi-(2N-1)\xi_0)\right]$$ similar steps as before



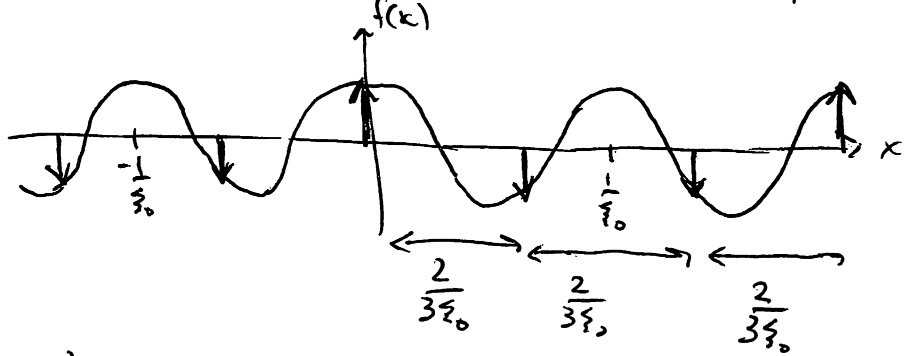
$$F_s(\xi)\,\mathrm{rect}\left(\frac{\xi}{2\xi_0}\right) = \frac{1}{2}\left[\delta(\xi-\xi_0) + \delta(\xi-\xi_0)\right]$$ ← Remember $\mathrm{rect}\left(\frac{\pm\xi_0}{2\xi_0}\right) = \frac{1}{2}$

$$f(x) = \mathcal{F}^{-1}\left\{F_s(\xi)\,\mathrm{rect}\left(\frac{\xi}{2\xi_0}\right)\right\} = \cos(2\pi\xi_0 x)$$ Perfect recovery

Let's reduce the sampling frequency to $\frac{3}{2}\xi_0$ or 1.5x the Nyquist frequency

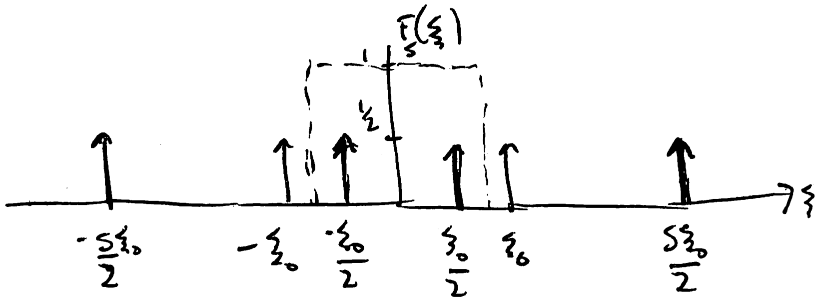
so $\frac{1}{X_s} = \frac{3\xi_0}{2}$

$X_s = \frac{2}{3\xi_0}$



$$f_s(x) = \mathrm{comb}\left(\frac{3\xi_0}{2}x\right)\cos(2\pi\xi_0 x)$$

$$F_s(\xi) = \sum_{n=-\infty}^{\infty}\frac{1}{2}\left[\delta\left(\xi - \left(\tfrac{3}{2}n+1\right)\xi_0\right) + \delta\left(\xi - \left(\tfrac{3}{2}n-1\right)\xi_0\right)\right]$$



$$F_s(\xi)\,\mathrm{rect}\left(\frac{\xi}{\frac{3\xi_0}{2}}\right) = \frac{1}{2}\left[\delta\left(\xi - \frac{\xi_0}{2}\right) + \delta\left(\xi + \frac{\xi_0}{2}\right)\right]$$

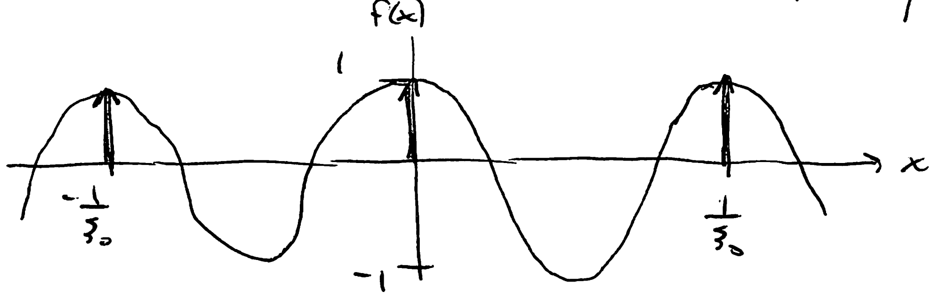
$$f(x) = \mathcal{F}^{-1}\left\{F_s(\xi)\,\mathrm{rect}\left(\frac{\xi}{\frac{3\xi_0}{2}}\right)\right\} = \cos\left(2\pi\left(\frac{\xi_0}{2}\right)x\right)$$

aliased signal maps to lower spatial frequency

Finally, reduce the sampling frequency to ξ_0 i.e. the Nyquist frequency.

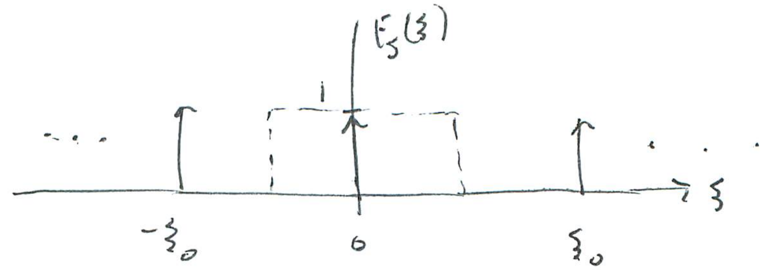
$\frac{1}{X_s} = \xi_0$

$X_s = \frac{1}{\xi_0}$



$$f_s(x) = \mathrm{comb}(\xi_0 x)\cos(2\pi\xi_0 x) = \mathrm{comb}(\xi_0 x)$$

$$F_s(\xi) = \frac{1}{\xi_0}\,\mathrm{comb}\left(\frac{\xi}{\xi_0}\right)$$



$$F_s(\xi)\,\mathrm{rect}\left(\frac{\xi}{\xi_0}\right) = \delta(\xi)$$

$$f(x) = \mathcal{F}^{-1}\left\{F_s(\xi)\,\mathrm{rect}\left(\frac{\xi}{\xi_0}\right)\right\} = 1 \qquad \text{NO COSINE MODULATION}$$

Aliasing is an ~~sampling~~ error where high frequency patterns get mapped to low frequency patterns. For the examples above, when

$\frac{1}{X_s}$	output	
$4N_\xi$	$\cos(2\pi\xi_0 x)$	Perfect recovery
$2N_\xi$	$\cos(2\pi\xi_0 x)$	
$\frac{3}{2}N_\xi$	$\cos\left(2\pi\left(\frac{\xi_0}{2}\right)x\right)$	← half the real frequency
N_ξ	1	← no modulation

SHOW IMAGES OF ALIASED PATTERNS

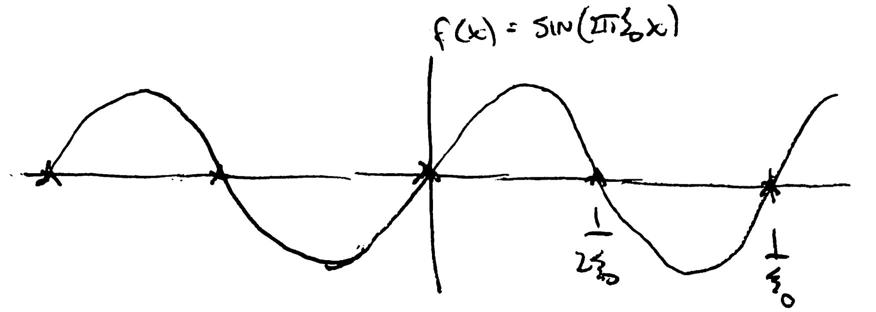
aliasing occurs when
1. $f(x)$ is band-limited by X_s too large
2. $f(x)$ is not band-limited.

Sampling with $\frac{1}{X_s} = 2N_\xi$ does not ensure perfect recovery. This is just a fuzzy boundary between aliasing and non-aliasing

Example sampling frequency is $2\xi_0$ for an input $\sin(2\pi\xi_0 x)$

$$\frac{1}{X_s} = 2\xi_0 \quad \text{twice the Nyquist frequency}$$

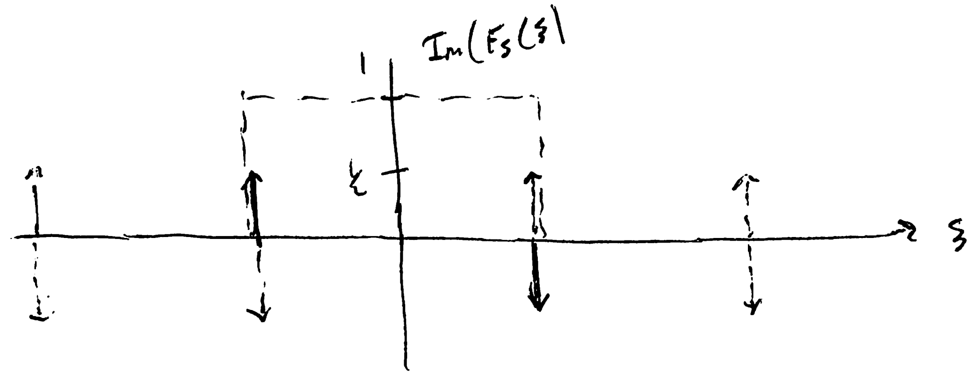
$$X_s = \frac{1}{2\xi_0}$$



$$f_s(x) = \text{comb}(2\xi_0 x)\,\sin(2\pi\xi_0 x)$$

sample points occur at zeros of $\sin(2\pi\xi_0 x)$

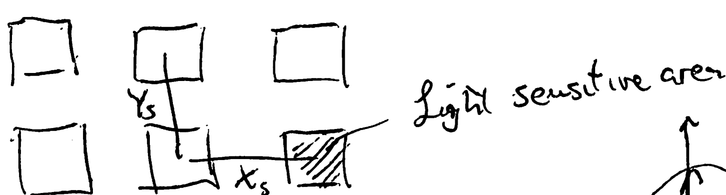
$$F_s(\xi) = \sum_{n=-\infty}^{\infty} \frac{1}{2i}\Big[\delta(\xi - (2n+1)\xi_0) - \delta(\xi - (2n-1)\xi_0)\Big]$$

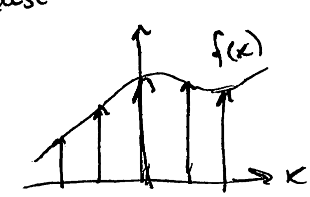


$$F_s(\xi)\,\text{rect}\left(\frac{\xi}{2\xi_0}\right) = 0$$

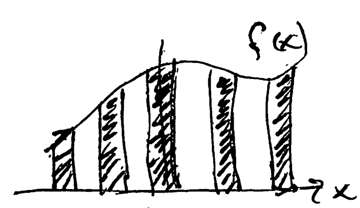
$$f(x) = \mathcal{F}^{-1}\left\{F_s(\xi)\,\text{rect}\left(\frac{\xi}{2\xi_0}\right)\right\} = 0$$

CAMERA PIXELS - Up to now, we have used simple δ functions to sample points on $f(x,y)$. Camera pixels are slightly different in that the value they supply is an integration of the irradiance over the aperture of the pixel. However, the same sampling issues arise.

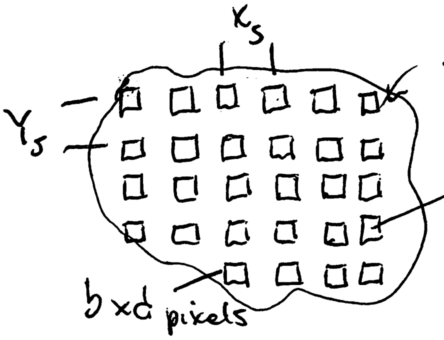




delta functions



pixels



pixel records (integrates) over its area

$$\iint_{-\infty}^{\infty} f(\alpha,\beta)\, \mathrm{rect}\left(\frac{\alpha - x}{b}, \frac{\beta - y}{d}\right) d\alpha\, d\beta$$

This has the form of a convolution. $f(x,y) ** \mathrm{rect}\left(\frac{x}{b}, \frac{y}{d}\right)$

The sampled signal now looks like

$$f_s(x,y) = \left[f(x,y) ** \mathrm{rect}\left(\frac{x}{b}, \frac{y}{d}\right)\right] \mathrm{comb}\left(\frac{x}{X_s}\right) \mathrm{comb}\left(\frac{y}{Y_s}\right)$$

Recall that convolution is a smoothing operation, so our original object $f(x,y)$ is "pre-smoothed" before it is sampled now.

$$F_s(\xi,\eta) = bd\, X_s Y_s \left[F(\xi,\eta)\, \mathrm{sinc}(b\xi, d\eta)\right] ** \mathrm{comb}(X_s\xi)\, \mathrm{comb}(Y_s\eta)$$

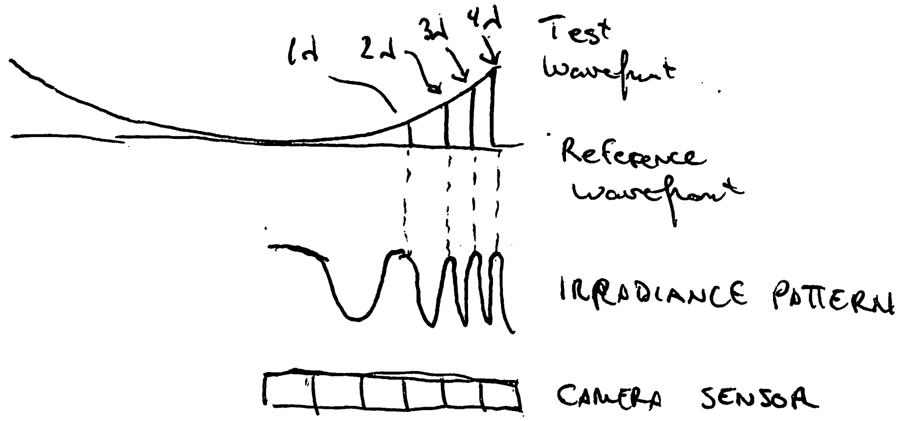
$$F_s(\xi,\eta) = bd \sum_{N,M} F\left(\xi - \frac{N}{X_s}, \eta - \frac{m}{Y_s}\right) \mathrm{sinc}\left(b\left(\xi - \frac{N}{X_s}\right), d\left(\eta - \frac{m}{Y_s}\right)\right)$$

sinc function causes a roll off in the high frequencies of $F(\xi,\eta)$. Where the sinc function is zero, the spectrum is lost.

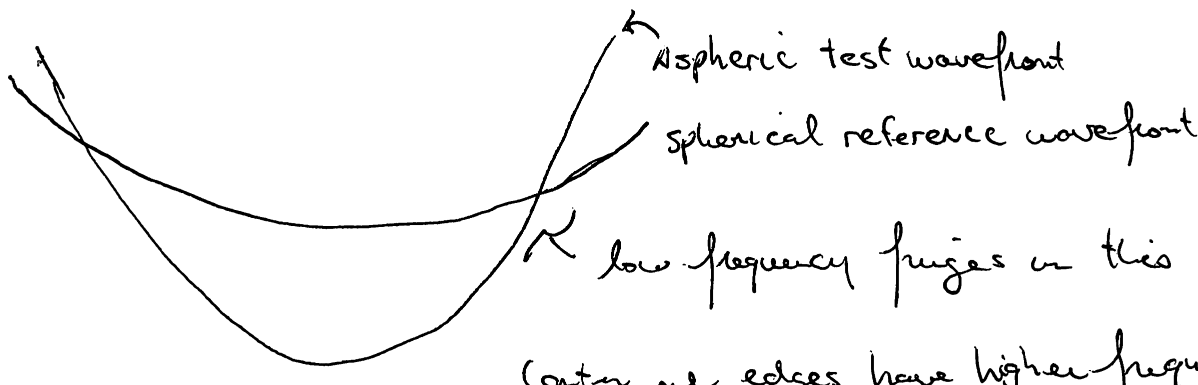
Pixels also record (integrate) in time domain so we can extend all these sampling concepts to inputs $f(x,y,t)$ (e.g. image on a sensor changes with time). Video imaging is a good example. Temporal aliasing occurs when the frame rate of the recording is slower than the temporal variations in the scene. This leads to the "wagon wheel" effect.

OPTICAL METROLOGY AND ALIASING

Interferometry is routinely used for testing the shape of an optical surface. This is typically used to compare a surface being fabricated to a known surface to look for imperfections in shape. Two wavefronts, one from a known reference surface and one from the test surface are combined. Anywhere the two wavefronts differ by an integer multiple of the wavelength, a fringe appears. Fringe patterns can be related to surface shape.



If the fringes get too close together compared to the size of the pixel in the sensor, then aliasing will occur. This is often the case when testing highly aspheric surfaces where the reference wavefront is typically a "best-fit" spherical shape.



Center and edges have higher frequency fringes.