Image Processing and Computer Vision

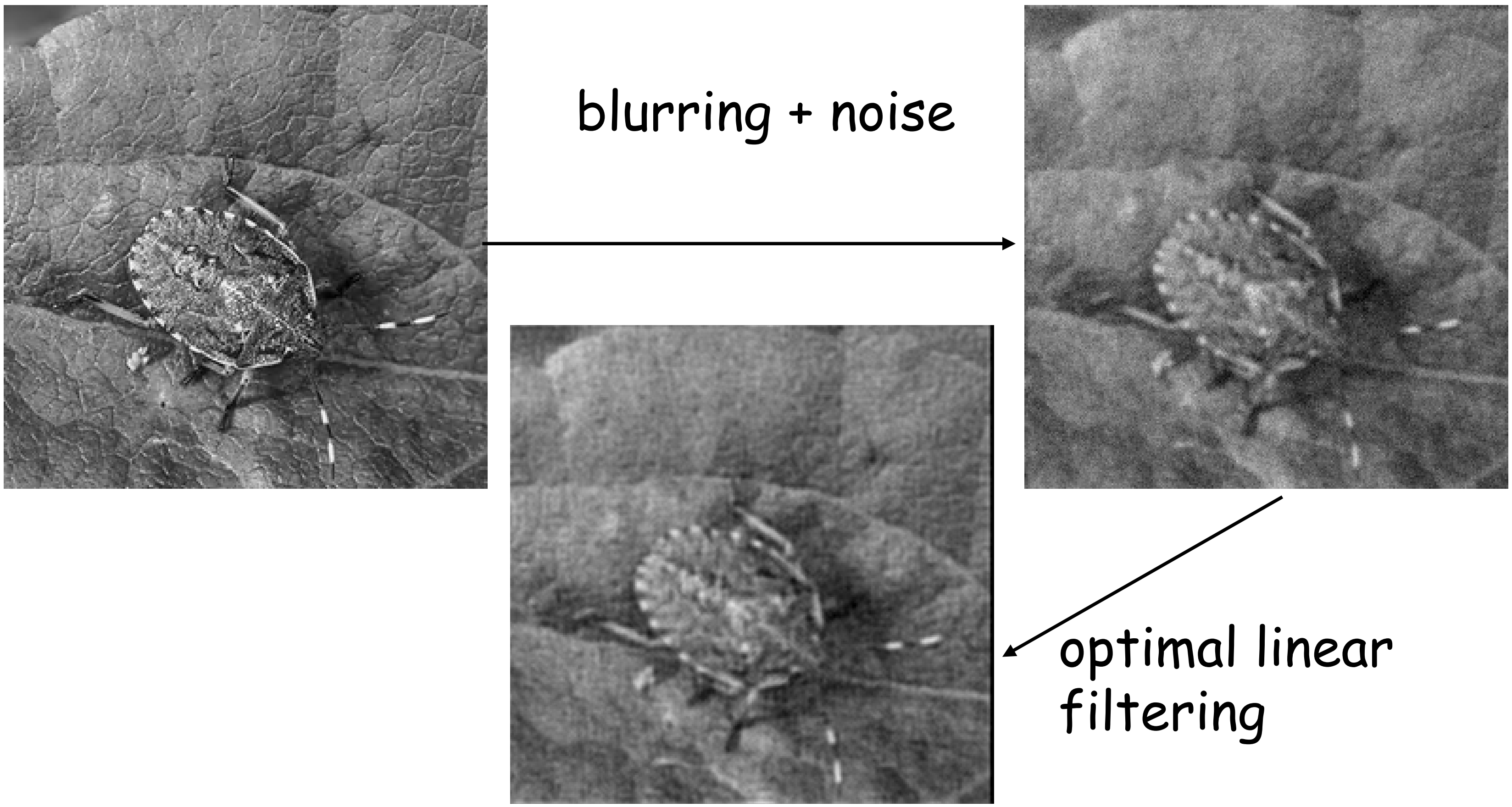
Computer Vision

What is computer vision? **Interpreting images!**

The computer sees $100111010010101000000001110101\ldots$

Image Processing

What is image processing? **Restoring images without extraction of semantic information!**



Some Topics of Computer Vision: Quantization

Image Quantization in space and intensity/color.

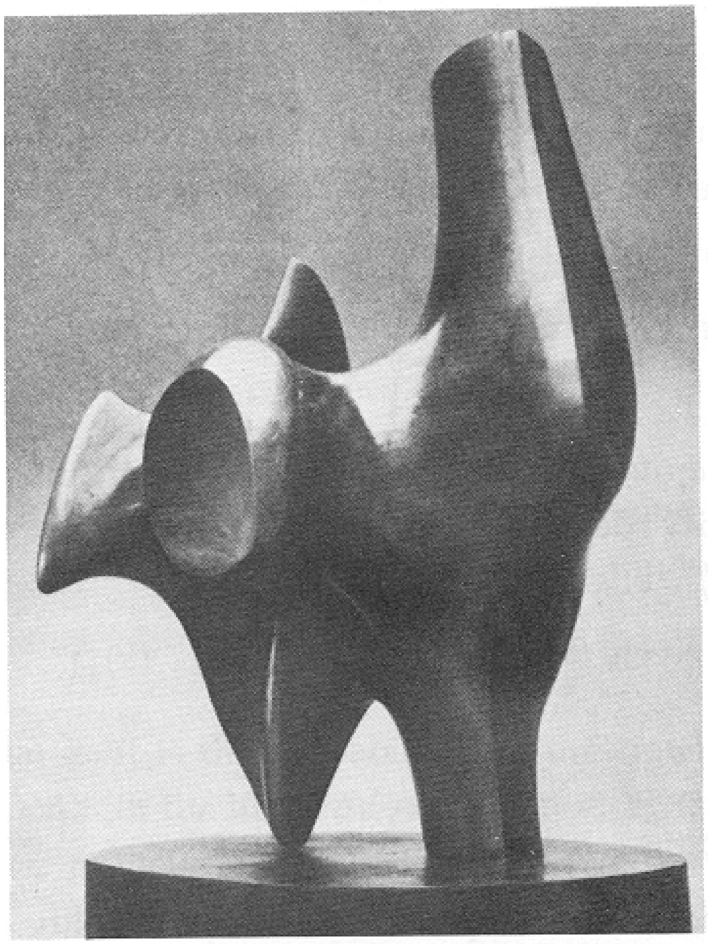
Sampling the image intensity at discrete positions is called spatial quantization which requires to limit the high frequency content of images to avoid aliasing.
Intensity quantization is achieved by scalar quantization of grey values or vector quantization (clustering) of color spaces.
Below is an example for spatial color quantization which combines spatial quantization and dithering with error diffusion.

b)

c)

Edge Detection

Goal: reduce the image content to semantically informative parts -¿ edges.



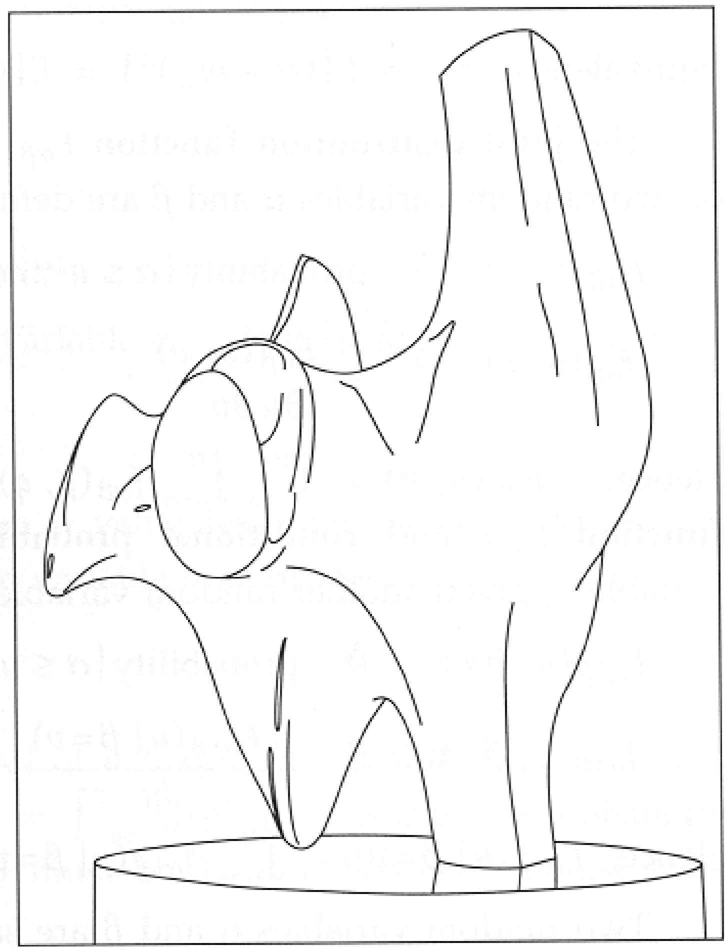
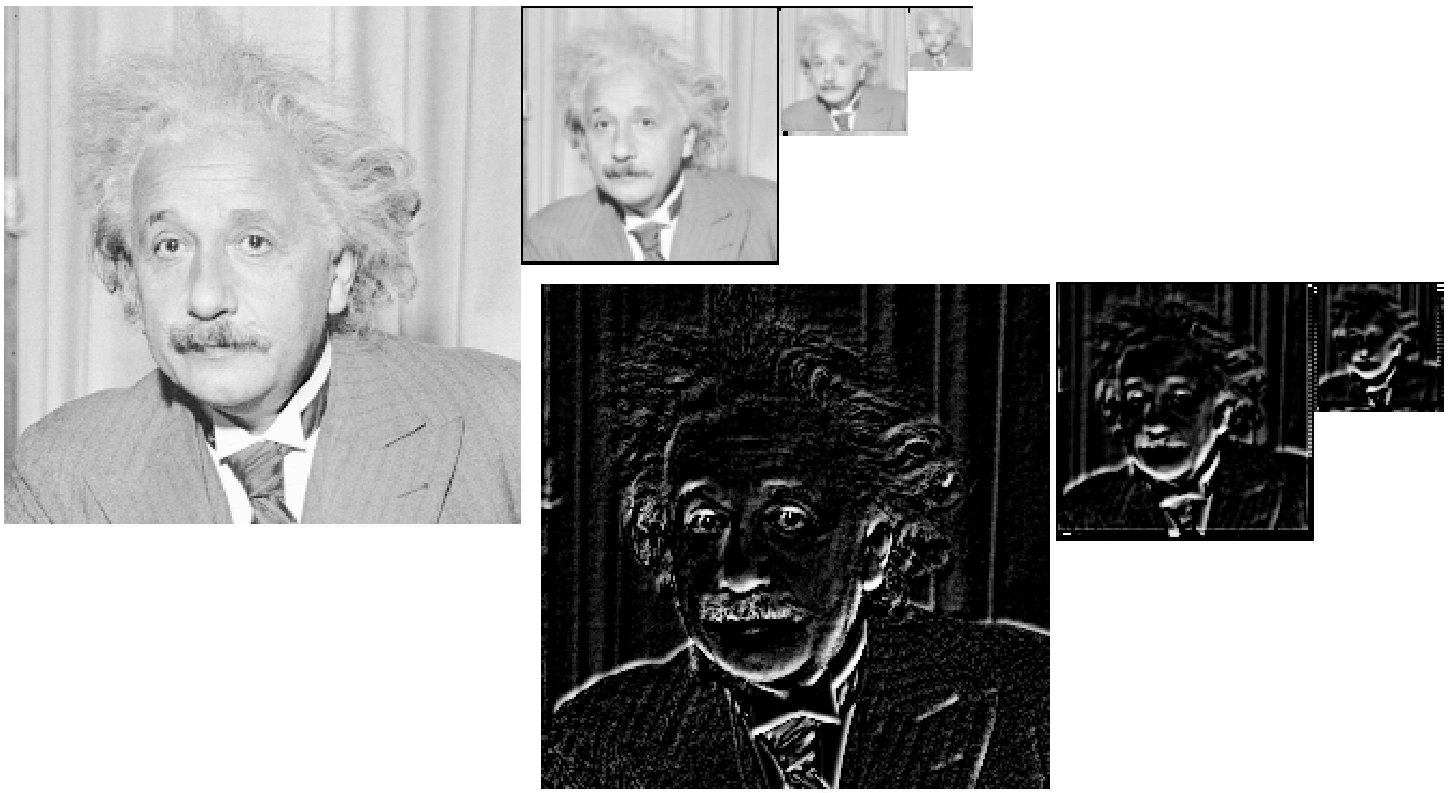


Image Formats: Gauss-Laplace Pyramid

Wavelets or Gauss-Laplace pyramids recode the image to decorrelate pixels by exploiting the self-similar nature of natural images. Thereby, we achieve significantly higher compression rates.



Stereo vision: Extract depth information from pairs of camera images.

Motion estimation: Estimate the movement of objects in the image by estimating their apparent motion from optical flow.

Shape from shading: Estimate the shape of objects from their appearence and their shading.

Shape from motion: Use motion parallax to estimate depth.

Shape from texture: Exploit texture variations as changes of surface normal w.r.t. line of sight.

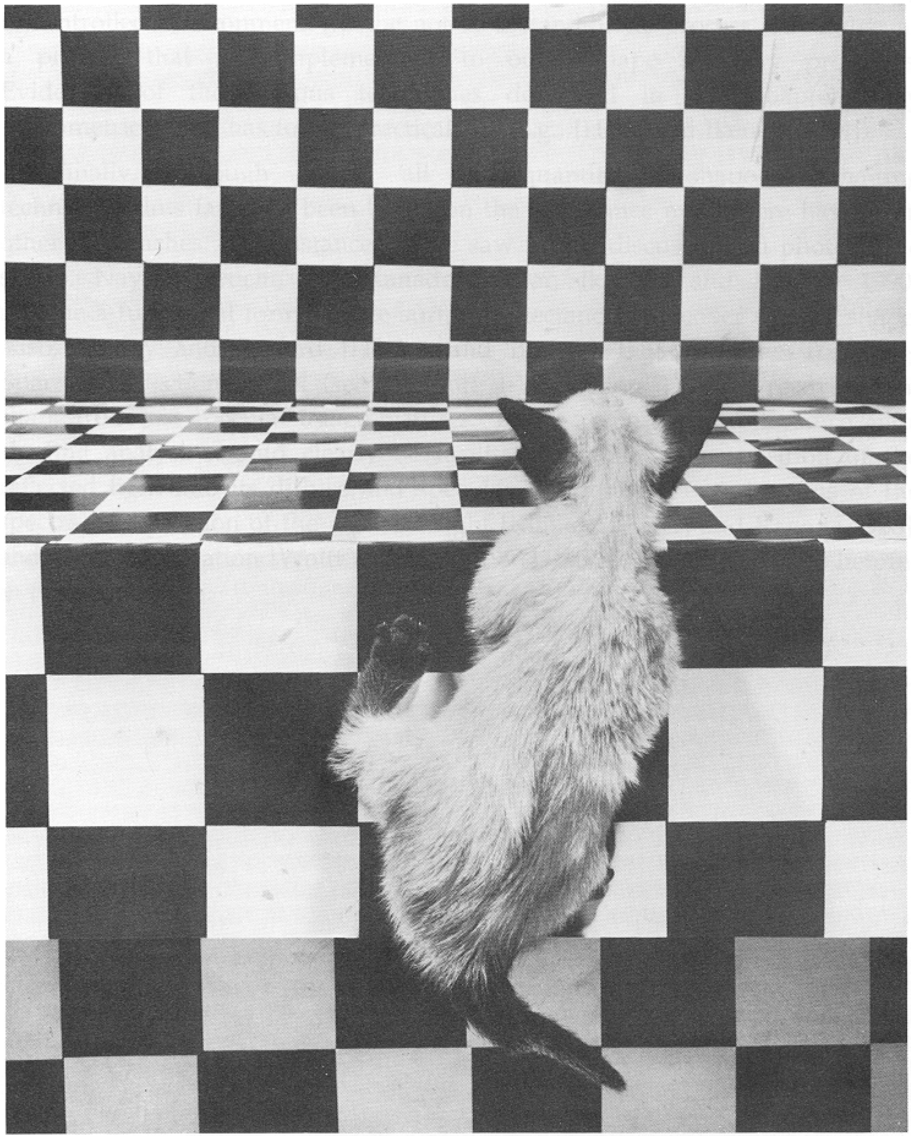
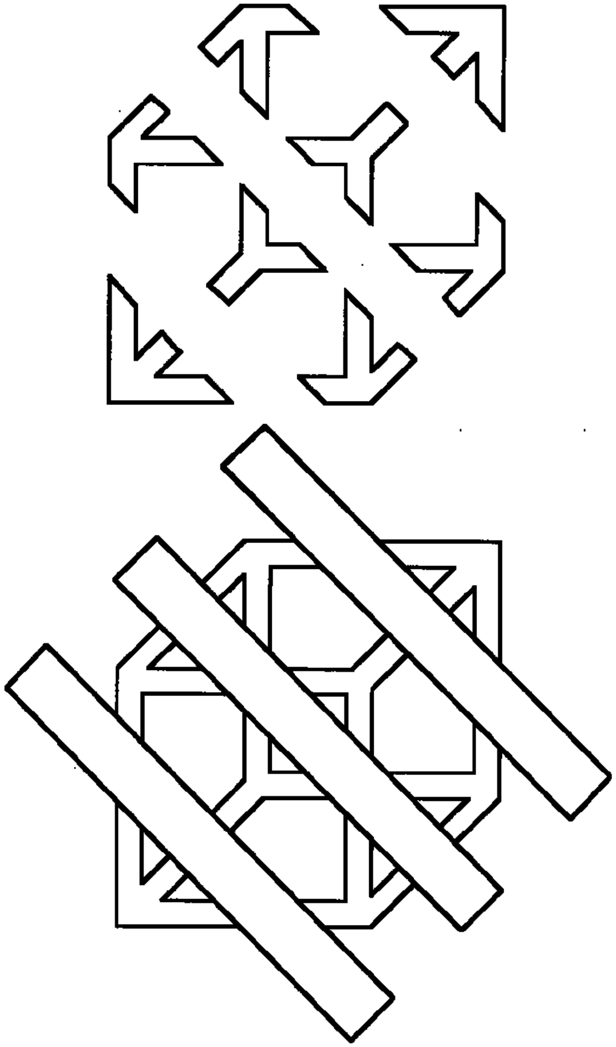


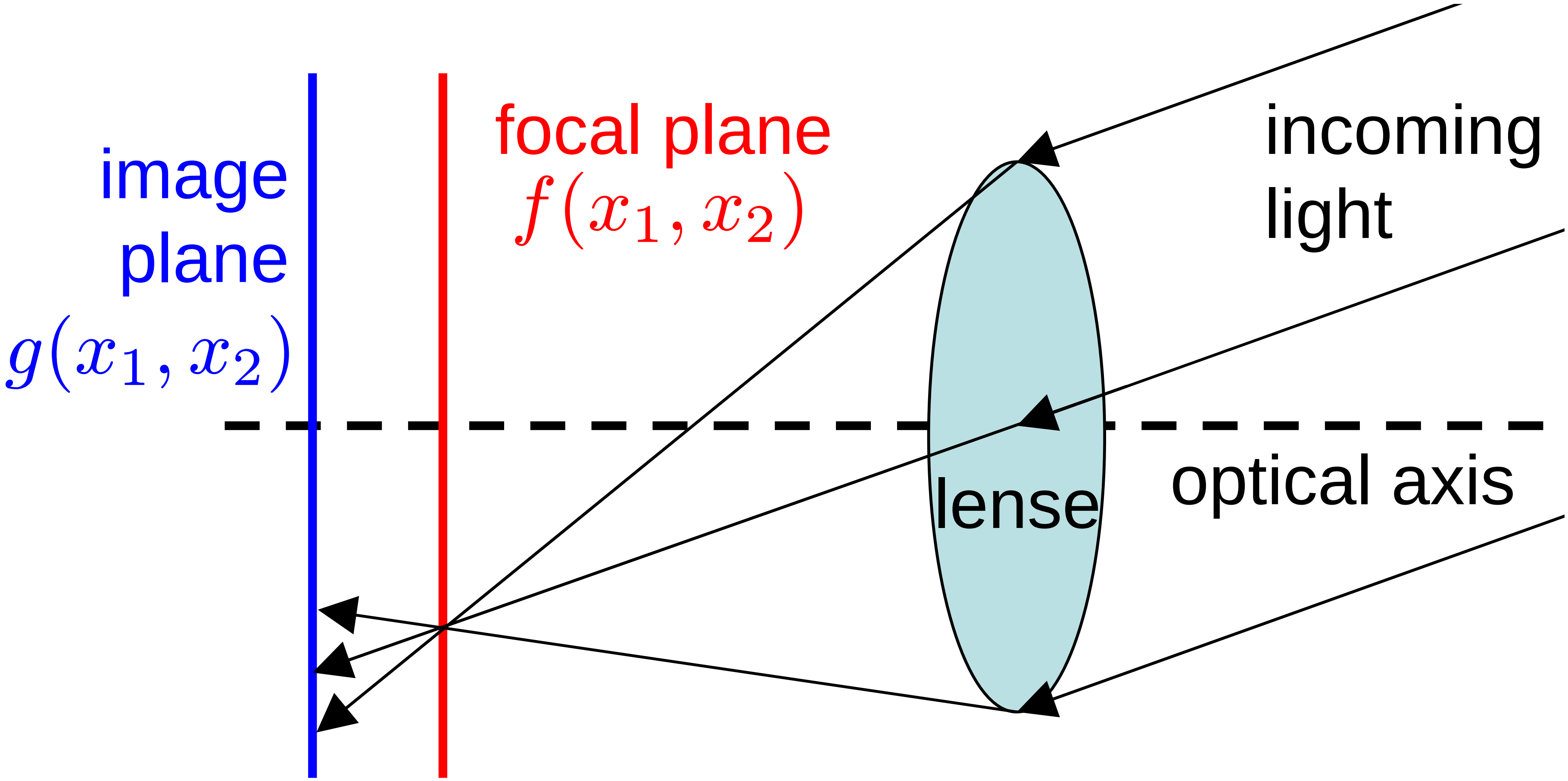
Image Understanding

Figure-ground segmentation

Perceptual Grouping



The Image Formation Process



Mathematical Modelling of Image Processing

Def.: An image is a continuous, two-dimensional function of the light intensity or color ($\mathbf{x} = (x_1, x_2)$)

$$\begin{aligned} f \;:\; \mathbb{R}^2 &\rightarrow \mathbb{R}_+ \quad \text{or } \mathbb{R}^3_+ \\ \mathbf{x} &\mapsto f(\mathbf{x}) \quad \text{or } (f_1, f_2, f_3)^T(\mathbf{x}) \end{aligned}$$

Question: How can we compensate an image deformation, e.g., defocussing?

Goal: reconstruct $f(\mathbf{x})$ from $g(\mathbf{x})$ in the presence of noise!

Model assumption:

1) When $f(\mathbf{x})$ is shifted then $g(\mathbf{x})$ is shifted as well.

2) Doubling the incoming light intensity will double the brightness $g(\mathbf{x})$.

Linear Shift-Invariant Systems

Strategy for restauration: invert the transformation $\mathcal{T}$ which maps the original image $f(\mathbf{x})$ to the defocussed image $g(\mathbf{x})$.

Linearity: (assumption) restrict $\mathcal{T}$ to linear operators!

$$f_1 \longrightarrow \boxed{\text{transform } \mathcal{T}} \longrightarrow g_1$$

$$f_2 \longrightarrow \boxed{\text{transform } \mathcal{T}} \longrightarrow g_2$$

$$\alpha f_1 + \beta f_2 \longrightarrow \boxed{\text{transform } \mathcal{T}} \longrightarrow \alpha g_1 + \beta g_2 \; \forall \, \alpha, \beta \in \mathbb{R}$$

- Linearity is typically only in the low intensity range fulfilled since physical systems tend to saturate.
- f_i, g_i are intensities $\equiv$ power per area with $f_i, g_i \geq 0$ in the full domain.
- Often we experience non-linear imaging errors!

Shift invariance: (assumption)

$$f(x_1, x_2) \longrightarrow \boxed{\text{transform } \mathcal{T}} \longrightarrow g(x_1, x_2)$$

$$f(x_1 - a, x_2 - b) \longrightarrow \boxed{\text{transform } \mathcal{T}} \longrightarrow g(x_1 - a, x_2 - b)$$

- Shift invariance holds only in a limited range since images are finite objects.

Remarks: The assumption of linearity is a significant limitation but it gives the advantage that the linear filter theory is completely developed.

- An analogous one-dimensional theory applies to passive electrical circuits, although there time is the essential dimension and causality constraints the signal.

How Can We Identify a Transformation?

Dirac's δ-function (1D): $\int\limits_{-\infty}^{\infty} \delta(x-a) f(x) dx = f(a)$

- Integration with the δ-function "samples" the function $f(x)$ at the position $x_0 = a$.
- The δ-function is a "generalized function".

- Regularization of the Dirac "function":

$$\delta(x) = \lim_{\epsilon \to 0} \begin{cases} \frac{1}{\epsilon} & |x| \leq \frac{\epsilon}{2} \\ 0 & \text{else} \end{cases} \qquad \text{or}$$

$$\delta(x) = \lim_{\epsilon \to 0} \frac{1}{\sqrt{2\pi\epsilon}} \exp(-\frac{x^2}{2\epsilon^2})$$

Convolution and the Point Spread Function

Assumption: $\delta(x_1, x_2) \quad \longrightarrow \boxed{\quad \mathcal{T} \quad} \longrightarrow \quad h(x_1, x_2)$

With linearity and shift invariance it holds:

$$
\begin{aligned}
g(x_1, x_2) \quad &= \quad \mathcal{T} f(x_1, x_2) \\
&= \quad \mathcal{T} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\xi, \eta) \delta(x_1 - \xi, x_2 - \eta) d\xi d\eta \\
&\overset{\text{linearity}}{=} \quad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\xi, \eta) \underbrace{\left[\mathcal{T} \delta(x_1 - \xi, x_2 - \eta)\right]}_{\substack{h(x_1 - \xi, x_2 - \eta) \\ \text{[shift inv.]}}} d\xi d\eta \\
&= \quad (f * h)(x_1, x_2)
\end{aligned}
$$

Linear, shift invariant systems can be written as convolutions!

Identification of the Kernel

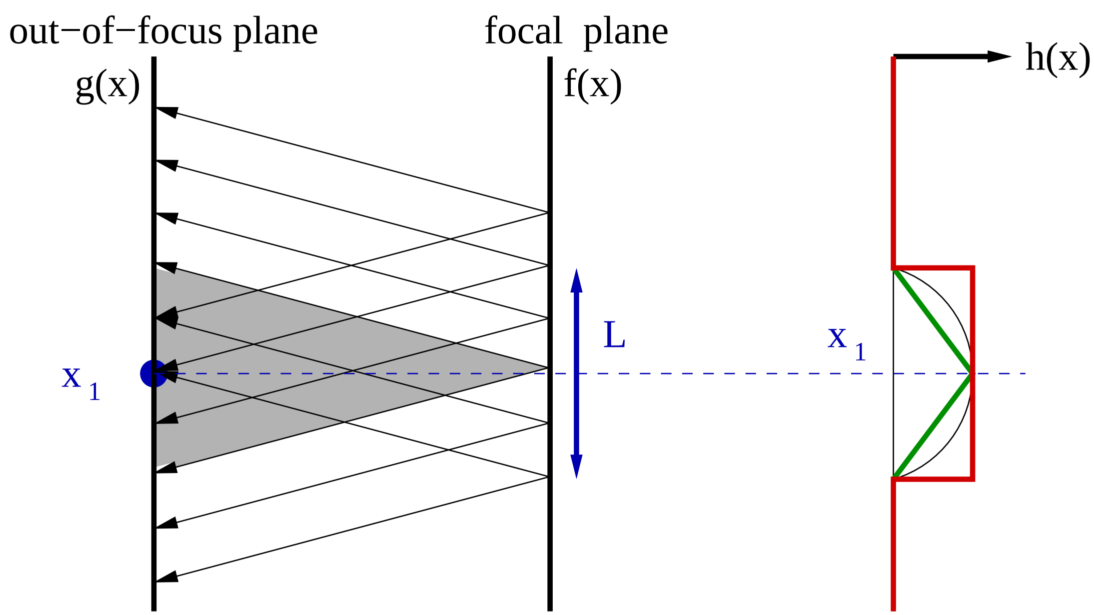
Let $f(x_1, x_2) = \delta(x_1, x_2)$, i.e., the image is a white dot with "infinite" intensity. Then the measured image $g(x_1, x_2)$ is given by

$$
\begin{aligned}
g(x_1, x_2) &= (\delta * h)(x_1, x_2) \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(\xi, \eta) h(x_1 - \xi, x_2 - \eta) d\xi d\eta \\
&= h(x_1, x_2)
\end{aligned}
$$

$$
\Rightarrow \quad \mathcal{T}\delta(x_1, x_2) = h(x_1, x_2)
$$

$\Rightarrow$ testing the linear shift-invariant system with a δ-peak will reveal the convolution kernel $h(x_1, x_2)$ of the system.

Schematic View of a Convolution

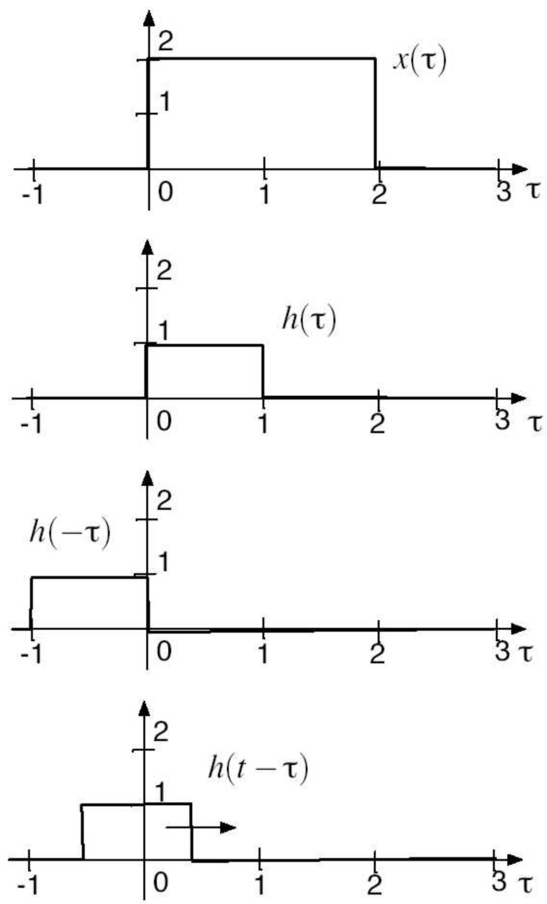


- $g(x_1)$ depends on $f(x)$ for all $x \in [x_1 - \frac{L}{2}, x_1 + \frac{L}{2}]$.
- convolution kernel $h_{x_1}(x)$ describes the influence of $f(x)$ onto $g(x_1)$.
- shift invariance of $h_{x_1}(x)$ results in cumulative influence:

$$
\begin{aligned}
g(x_1) &= \int_{-L/2}^{L/2} f(x)h(x_1 - x)dx = \int_{-L/2}^{L/2} f(x_1 - x)h(x)dx \\
&\approx f(0)h(x_1)\Delta + f(\Delta)h(x_1 - \Delta)\Delta + f(2\Delta)h(x_1 - 2\Delta)\Delta + \ldots \\
&\qquad + f(-\Delta)h(x_1 + \Delta)\Delta + f(-2\Delta)h(x_1 + 2\Delta)\Delta + \ldots
\end{aligned}
$$

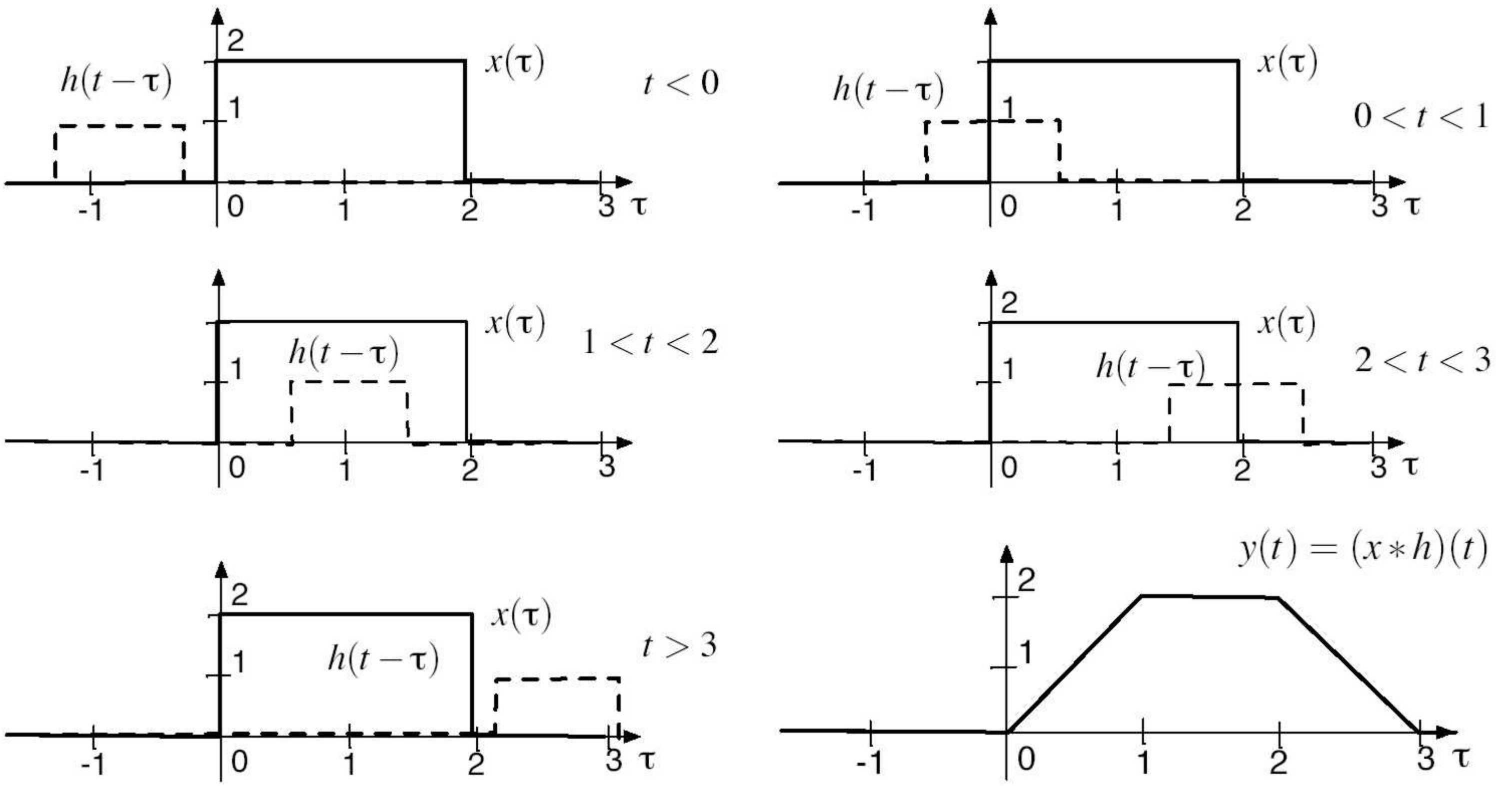
Convolution: 1D-Example

$$y(t) = (x * h)(t) = \int x(\tau) h(t - \tau) d\tau$$



Convolution: 1D-Example (cont'd)

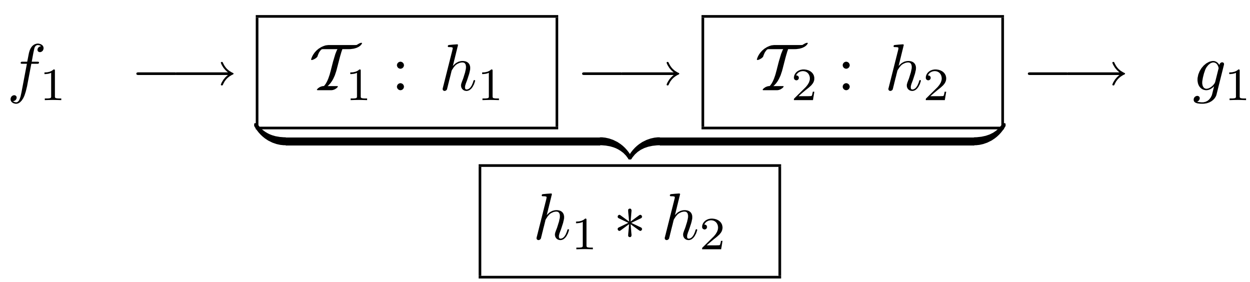
$$y(t) = (x * h)(t) = \int x(\tau)h(t - \tau)d\tau$$



Facts about Convolution

- Linear shift-invariant (LSI) systems can be written as convolutions.
- The convolution kernel h characterizes the LSI system uniquely.
- Cascades of LSI systems: the convolution is commutative and associative:

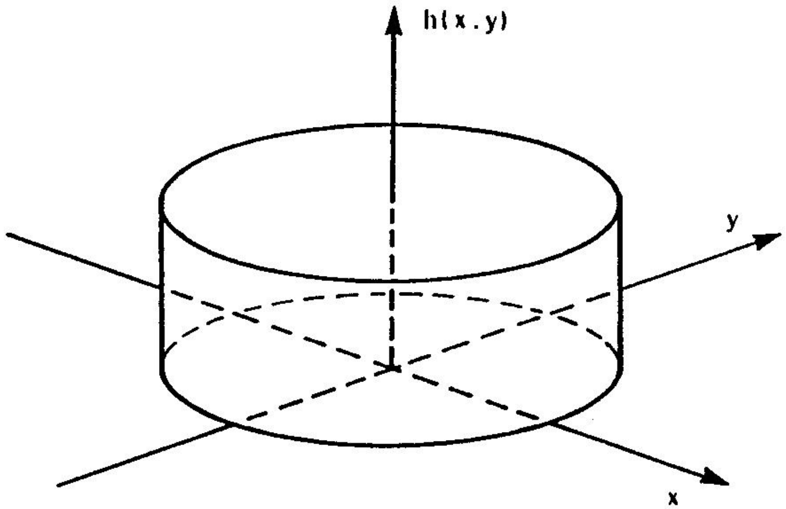
$$
\begin{aligned}
g * h &= h * g \\
(f * g) * h &= f * (g * h)
\end{aligned}
$$



$\Rightarrow$ one of the most important operations in signal processing

Convolution Kernel for Image Defocussing

Defocussing an image amounts to convolving it with a 'pillbox':



$$h(x_1, x_2) = \begin{cases} \frac{1}{\pi R^2} & x_1^2 + x_2^2 \leq R^2 \\ 0 & \text{otherwise} \end{cases}$$

Note: this convolution kernel is normalized: $\int \int h(x_1, x_2) dx_1 dx_2 = 1$

Convolution Kernel for Image Defocussing

original image

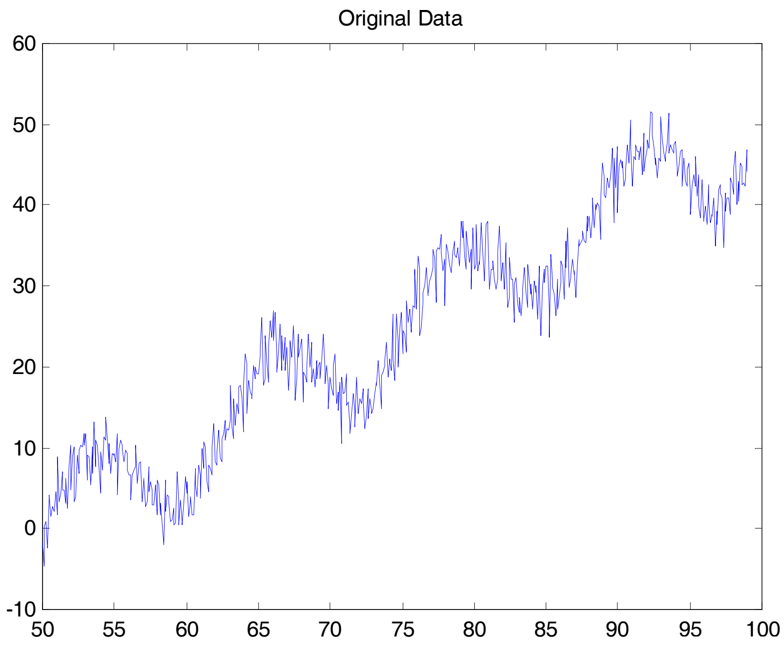
convolved with pillbox kernel

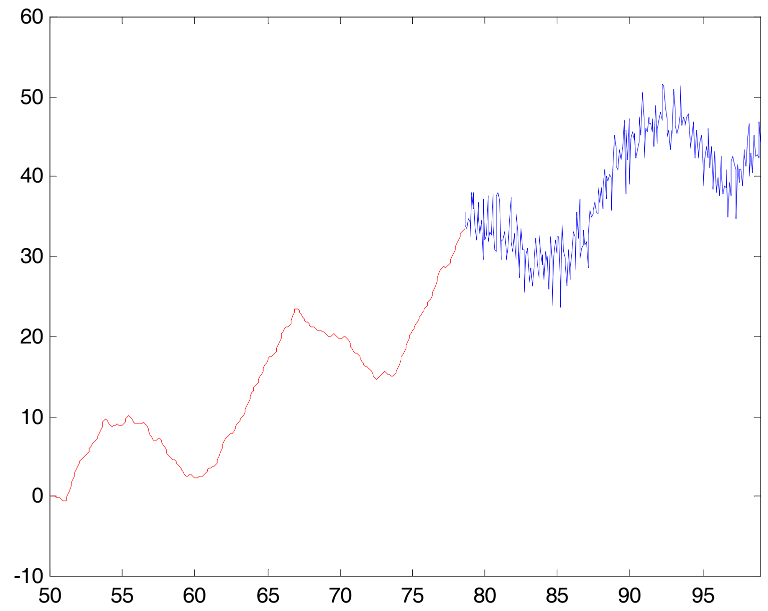
A Motion Kernel

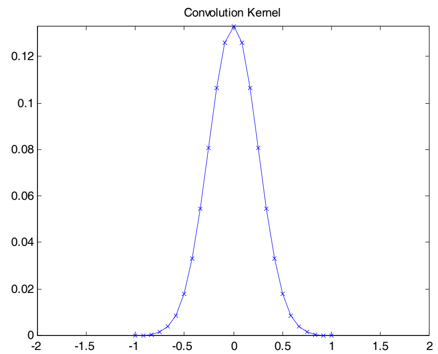
Each light dot is transformed into a short line along the x_1-axis:

$$h(x_1, x_2) = \frac{1}{2l}\left[\theta(x_1 + l) - \theta(x_1 - l)\right]\delta(x_2)$$

Denoising Time Series





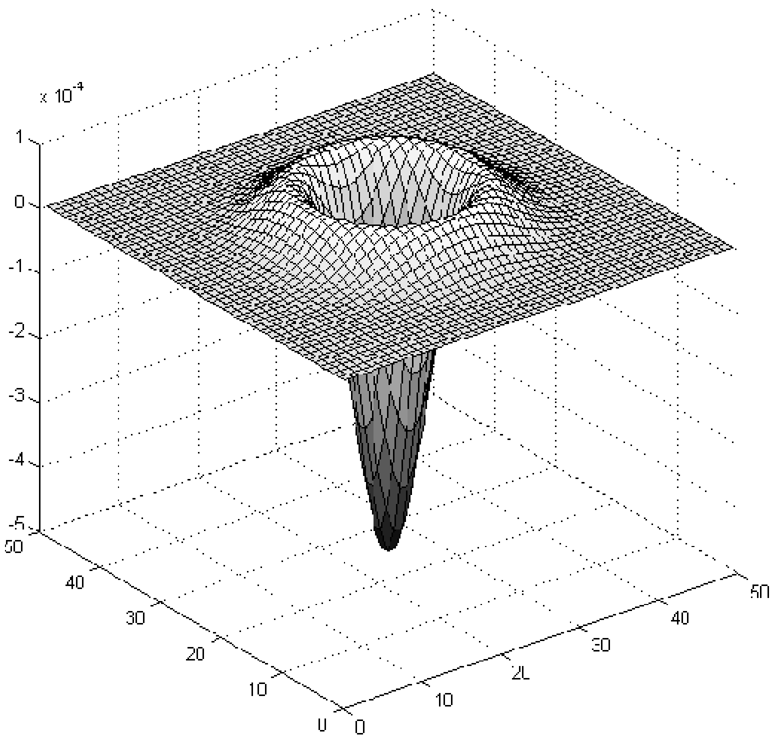


Lena with Gaussian Blurring and Noise

Gaussian blurring kernel:

$$h(x_1, x_2) = \frac{1}{2\pi\sigma^2} \exp(-\frac{x_1^2 + x_2^2}{2\sigma^2})$$

Lena Convolved with a Laplacian Filter



Laplacian filter:

$$h(x_1, x_2) = \nabla^2 \exp(-\frac{x_1^2 + x_2^2}{2\sigma^2}) = -\frac{1}{\sigma^2}\left(1 - \frac{r^2}{\sigma^2}\right)\exp(-\frac{r^2}{2\sigma^2}) \quad r^2 := x_1^2 + x_2^2$$

Note: here the normalization is $\int\int h(x_1, x_2)dx_1 dx_2 = 0$.

The Fourier Transformation: Basic Facts

Def.: Let f be an absolutely integrable function over $\mathbb{R}$. The Fourier transformation of f is defined as

$$\hat{f}(u) \equiv \mathcal{F}[f(x)] = \int_{-\infty}^{+\infty} f(x) \exp(-\imath 2\pi u x) dx.$$

The inverse Fourier transformation is given by the formula

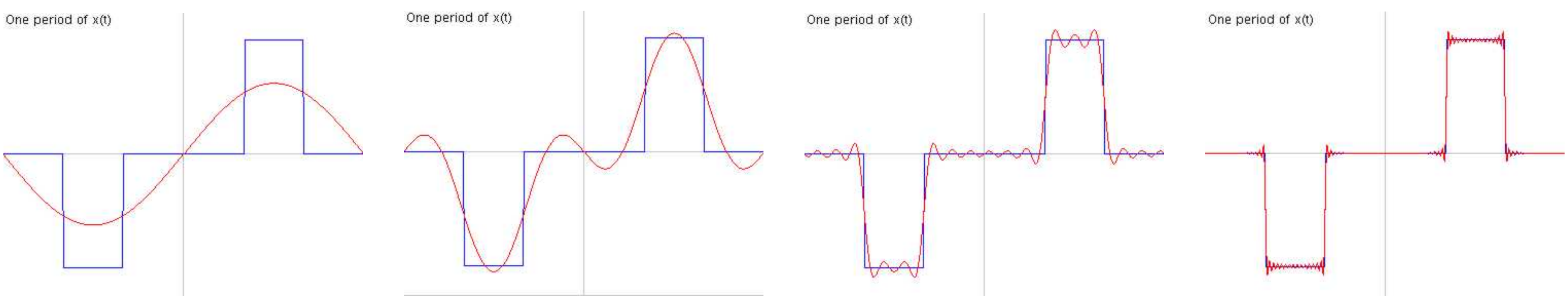
$$f(x) \equiv \mathcal{F}^{-1}[\hat{f}(u)] = \int_{-\infty}^{+\infty} \hat{f}(u) \exp(\imath 2\pi u x) du.$$

Note: while $f(x)$ is always real, $\hat{f}(u)$ is typically complex.

- $\hat{f}(u)$ is also called the **continuous spectrum** of $f(x)$.
- If x is a space coordinate, then u is called the **spatial frequency**.

Inversion formula: $f(x)$ is represented as a continuous superposition of waves with amplitude $\hat{f}(u)$.

Example of an odd function approximated by sinus waves (Remember: $\exp(\imath x) = \cos(x) + \imath \sin(x)$):



$$f(x) \approx \hat{f}(u_0)\sin(2\pi u_0 x) + \hat{f}(u_1)\sin(2\pi u_1 x) + \hat{f}(u_2)\sin(2\pi u_2 x) + \ldots$$

Fourier Transformation: Example 1 (box)

Given the box function

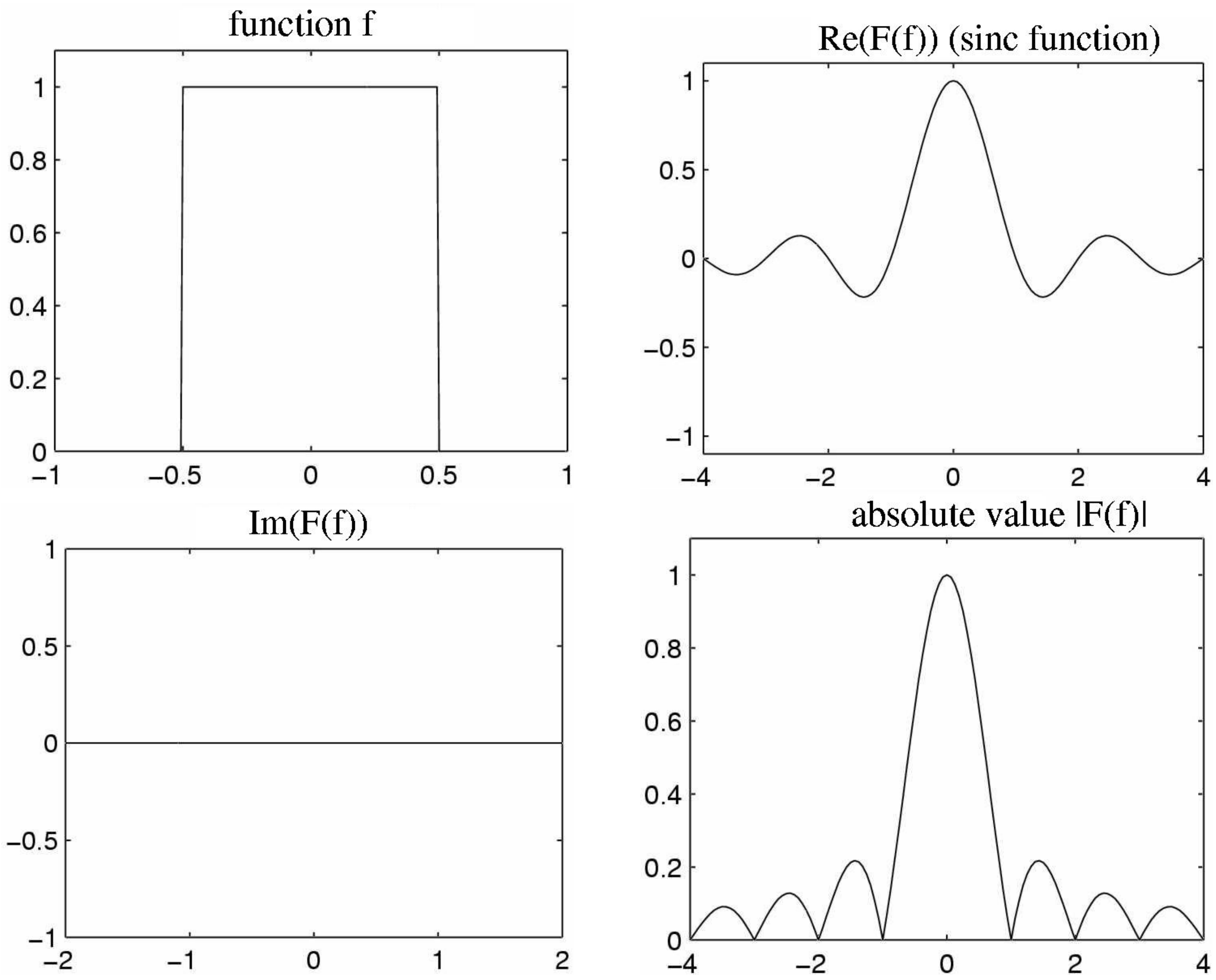
$$f(x) = \frac{1}{2l}\left(\theta(x+l) - \theta(x-l)\right) = \begin{cases} \frac{1}{2l} & \text{if} |x| \le l \\ 0 & \text{otherwise} \end{cases}$$

the Fourier transform is

$$\begin{aligned} \hat{f}(u) \equiv \mathcal{F}[f(x)] &= \int_{-\infty}^{+\infty} f(x) \exp(-\imath 2\pi u x) dx \\ &= \int_{-l}^{l} \frac{1}{2l} \cdot \left(\cos(2\pi u x) - \imath \underbrace{\sin(2\pi u x)}_{\int \to 0}\right) dx \\ &= \frac{\sin(2\pi u l)}{2\pi u l} \equiv \operatorname{sinc}(2\pi u l) \end{aligned}$$

Fourier Transformation: Example 1 (box)

Graphs of box and sinc-function for $l = \frac{1}{2}$:



Fourier Transformation: Example 2 (Gauss)

Given the function

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_x} \exp(-\frac{x^2}{2\sigma_x^2})$$

the Fourier transform is

$$\begin{aligned}\hat{f}(u) &\equiv \mathcal{F}[f(x)] = \int_{-\infty}^{+\infty} f(x) \exp(-\imath 2\pi u x) dx \\ &= \frac{1}{\sqrt{2\pi}\sigma_x} \int_{-\infty}^{\infty} \exp(-\frac{x^2}{2\sigma_x^2}) \cdot (\cos(2\pi u x) - \imath \underbrace{\sin(2\pi u x)}_{\int \to 0} dx \\ &=^{\dagger} \exp(-\frac{u^2}{2\sigma_u^2}) \qquad \text{where} \quad \sigma_u = \frac{1}{2\pi\sigma_x}\end{aligned}$$

† [Abramowitz, Stegun: Handbook of Mathematical Functions, 1972]

$\Rightarrow$ the Fourier transform of a Gaussian is a (unnormalized) Gaussian!

The larger the variance σ_x^2, the smaller the variance σ_u^2: $\quad \sigma_x \cdot \sigma_u = \frac{1}{2\pi}$

Fourier Transformation: Example 3 (Dirac's δ)

The Fourier transform of Dirac's δ-function is

$$\begin{aligned}\hat{\delta}(u) \equiv \mathcal{F}[\delta(x)] &= \int_{-\infty}^{+\infty} \delta(x) \exp(-\imath 2\pi u x) dx \\ &= \exp(-\imath 2\pi u \cdot 0) \\ &= 1\end{aligned}$$

$\Rightarrow$ the Fourier transform of the δ-function equals 1 for *all* frequencies u.

Properties of the Fourier Transformation

Linearity: If $\mathcal{F}[f(x)] = \hat{f}(u)$ and $\mathcal{F}[g(x)] = \hat{g}(u)$ then it holds for all complex numbers $a, b \in \mathbb{C}$

$$\mathcal{F}[af(x) + bg(x)] = a\hat{f}(u) + b\hat{g}(u)$$

Shift: If $\mathcal{F}[f(x)] = \hat{f}(u)$ then it holds for $c \in \mathbb{R}$

$$\mathcal{F}[f(x - c)] = \hat{f}(u) \exp(-\imath 2\pi cu)$$

Modulation: If $\mathcal{F}[f(x)] = \hat{f}(u)$ then it holds for $c \in \mathbb{R}$

$$\mathcal{F}[f(x) \exp(\imath 2\pi cx)] = \hat{f}(u - c)$$

Scaling: If $\mathcal{F}[f(x)] = \hat{f}(u)$ and $c > 0$

$$\mathcal{F}[f(cx)] = \frac{1}{c}\hat{f}(\frac{u}{c})$$

Differentiation: Let f be piecewise continuous and absolutely integrable. If the function $xf(x)$ is absolutely integrable then the Fourier transform $\hat{f}$ is continuous and differentiable. It holds

$$\begin{aligned} \mathcal{F}[xf(x)] &= \frac{\imath}{2\pi}\frac{d}{du}\hat{f}(u) \\ \mathcal{F}[\frac{d}{dx}f(x)] &= \imath 2\pi u\hat{f}(u) \end{aligned}$$

Parseval's Equality: Let f be piecewise continuous and absolutely integrable. Then the Fourier transform $\hat{f}(u) = \mathcal{F}[f(x)]$ satisfies:

$$\int_{-\infty}^{\infty} |f(x)|^2 dx = \int_{-\infty}^{\infty} |\hat{f}(u)|^2 du$$

Power Spectrum: Considering the *auto-correlation function* $\Phi_{ff}(x)$ of a complex function f for $x \in \mathbb{R}$,

$$\Phi_{ff}(x) = \int_{-\infty}^{\infty} \bar{f}(\xi - x) f(\xi) d\xi \ .$$

The Fourier transform is given by

$$\hat{\Phi}_{ff}(u) \equiv \mathcal{F}[\Phi_{ff}(x)] = |\hat{f}(u)|^2.$$

($\bar{f}(x)$ is the conjugate complex function of $f(x)$)

Fourier Transform of Convolution

Given: convolution $g(x) = (f * h)(x) = \int f(\xi)h(x-\xi)d\xi$

Calculate Fourier transform of g:

$$\begin{aligned}
\hat{g}(u) &\equiv \mathcal{F}[g(x)] = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} f(\xi)h(x-\xi)d\xi \right] \exp(-\imath 2\pi u x)dx \\
&= \int_{-\infty}^{+\infty} f(\xi) \left[\int_{-\infty}^{+\infty} h(x-\xi)\exp(-\imath 2\pi u x)dx \right] d\xi \\
&= \int_{-\infty}^{+\infty} \hat{h}(u) f(\xi) \exp(-\imath 2\pi u \xi)d\xi \\
&= \hat{h}(u)\hat{f}(u)
\end{aligned}$$

$\Rightarrow$ Convolution in spatial domain becomes multiplication in Fourier space.

Modulation Transfer Function

System Behavior in Fourier Space: How is a harmonic oscillation transformed by convolution kernel h?
$\Rightarrow$ amplitude modulation $A(u)$:

$$\exp(\imath 2\pi u x) \quad \longrightarrow \quad \boxed{\text{kernel } h(x)} \quad \longrightarrow \quad A(u)\exp(\imath 2\pi u x)$$

Eigenfunction of the convolution with eigenvalue $A(u)$ is the oscillation $f(x) = \exp(\imath 2\pi u x)$.

$$\begin{aligned} \text{Output } g(x) &= (f * h)(x) = \int \exp(\imath 2\pi u \xi) h(x-\xi) d\xi \\ &= \exp(\imath 2\pi u x) \int \exp(-\imath 2\pi u \xi) h(\xi) d\xi \;=\; \hat{h}(u) \exp(\imath 2\pi u x) \end{aligned}$$

Note: the eigenvalue $A(u)$ equals $\hat{h}(u) = \mathcal{F}[h](u)$.

Image Filtering in the Frequency Domain

2D Fourier transformation of an image $f(\mathbf{x})$, $\mathbf{x} := (x_1, x_2)$:

$$\hat{f}(u,v) \equiv \mathcal{F}[f(\mathbf{x})] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x_1,x_2)\exp(-\imath 2\pi(ux_1 + vx_2))d\mathbf{x}$$

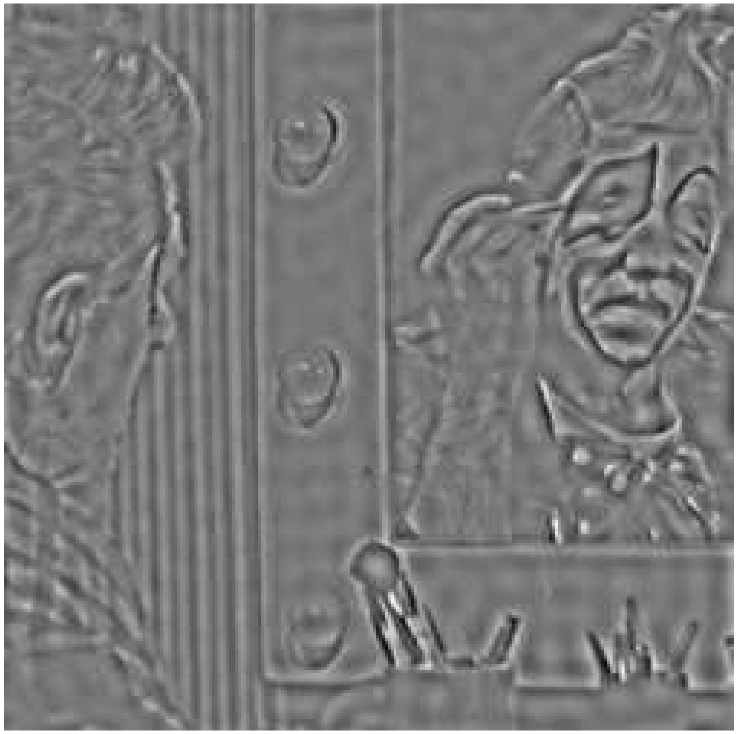
High-pass filtering: remove low frequencies, for example choose maximum value B:

$$\hat{f}_{\mathrm{hp}}(u,v) = \begin{cases} \hat{f}(u,v) & \text{if } u^2 + v^2 > B^2 \\ 0 & \text{otherwise} \end{cases}$$

Inverse Fourier transformation yields high-pass-filtered image $f_{\mathrm{hp}}(\mathbf{x}) = \mathcal{F}^{-1}[\hat{f}_{\mathrm{hp}}(u,v)]$

Example of Image Filtering

original image

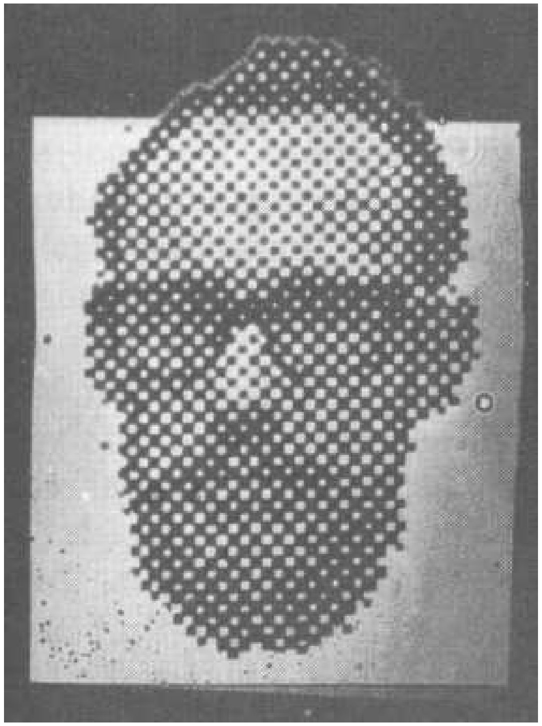


high-pass-filtered

$\Rightarrow$ edge detection

Low-pass filtering: analogous to high-pass filter, but remove high frequencies

Example:



original image

low-pass-filtered

⇒ removing noise

The Image Restoration Problem

$$f(\mathbf{x}) \longrightarrow \boxed{h(\mathbf{x})} \longrightarrow g(\mathbf{x}) \longrightarrow \boxed{\tilde{h}(\mathbf{x})} \longrightarrow f(\mathbf{x})$$

The 'inverse' kernel $\tilde{h}(\mathbf{x})$ should compensate the effect of the image degradation $h(\mathbf{x})$, i.e.,

$$(\tilde{h} * h)(\mathbf{x}) = \delta(\mathbf{x})$$

$\tilde{h}$ may be determined more easily in Fourier space:

$$\mathcal{F}[\tilde{h}](u,v) \cdot \mathcal{F}[h](u,v) = 1$$

To determine $\mathcal{F}[\tilde{h}]$ we need to estimate

1. the distortion model $h(\mathbf{x})$ (point spread function) or $\mathcal{F}[h](u,v)$ (modulation transfer function)
2. the parameters of $h(\mathbf{x})$, e.g. r for defocussing.

Image Restoration: Motion Blur

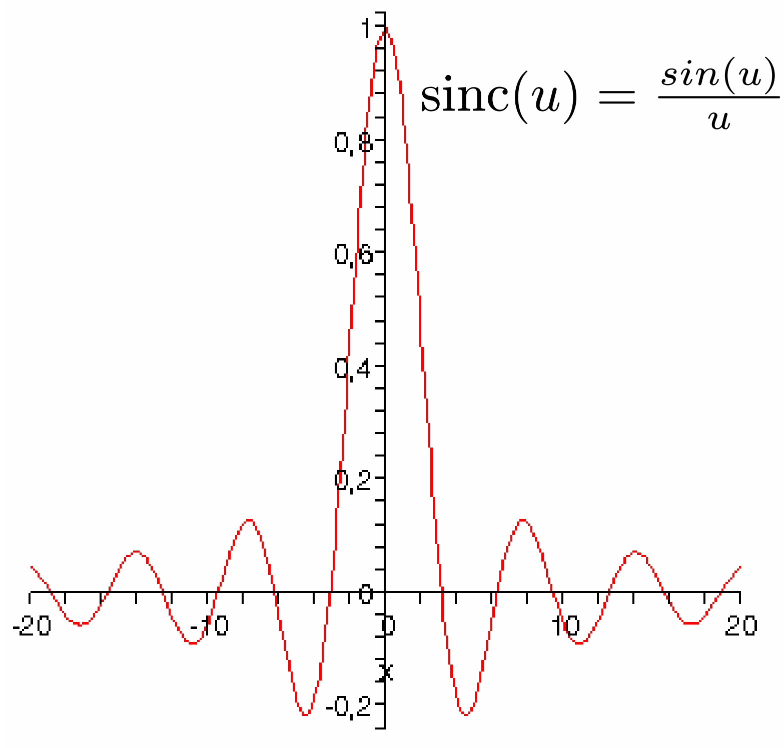
Kernel for motion blur $h(\mathbf{x}) = \frac{1}{2l}\big(\theta(x_1 + l) - \theta(x_1 - l)\big)\delta(x_2)$

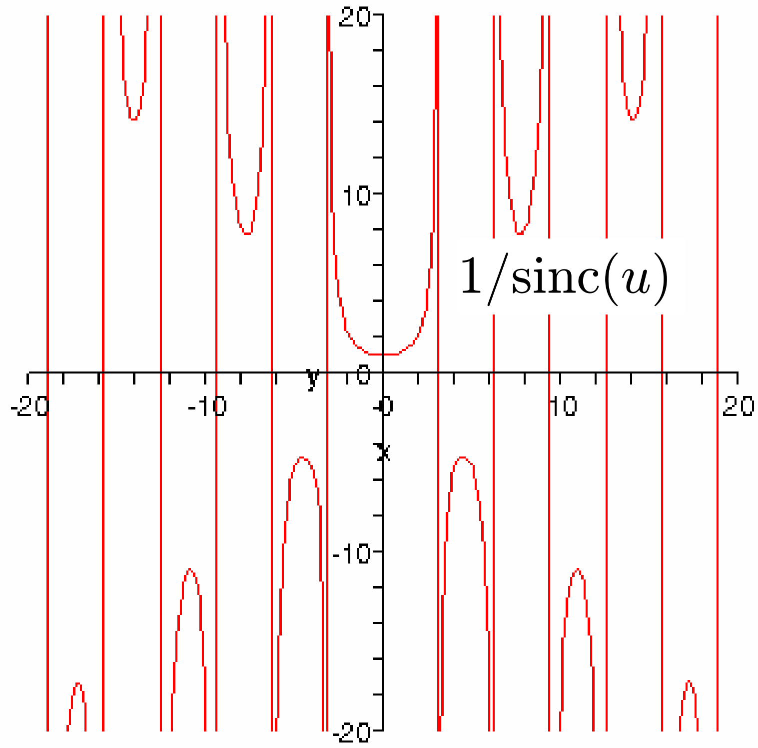
(a light dot is transformed into a small line in x_1 direction).

Fourier transformation:

$$\mathcal{F}[h](u,v) = \frac{1}{2l}\int_{-l}^{+l} \exp(-\imath 2\pi u x_1) \underbrace{\int_{-\infty}^{+\infty} \delta(x_2)\exp(-\imath 2\pi v x_2) dx_2}_{=1} dx_1$$

$$= \frac{\sin(2\pi u l)}{2\pi u l} =: \operatorname{sinc}(2\pi u l)$$





$$\hat{h}(u) = \mathcal{F}[h](u) = \mathrm{sinc}(2\pi u l) \qquad \mathcal{F}[\tilde{h}](u) = 1/\hat{h}(u)$$

Problems:

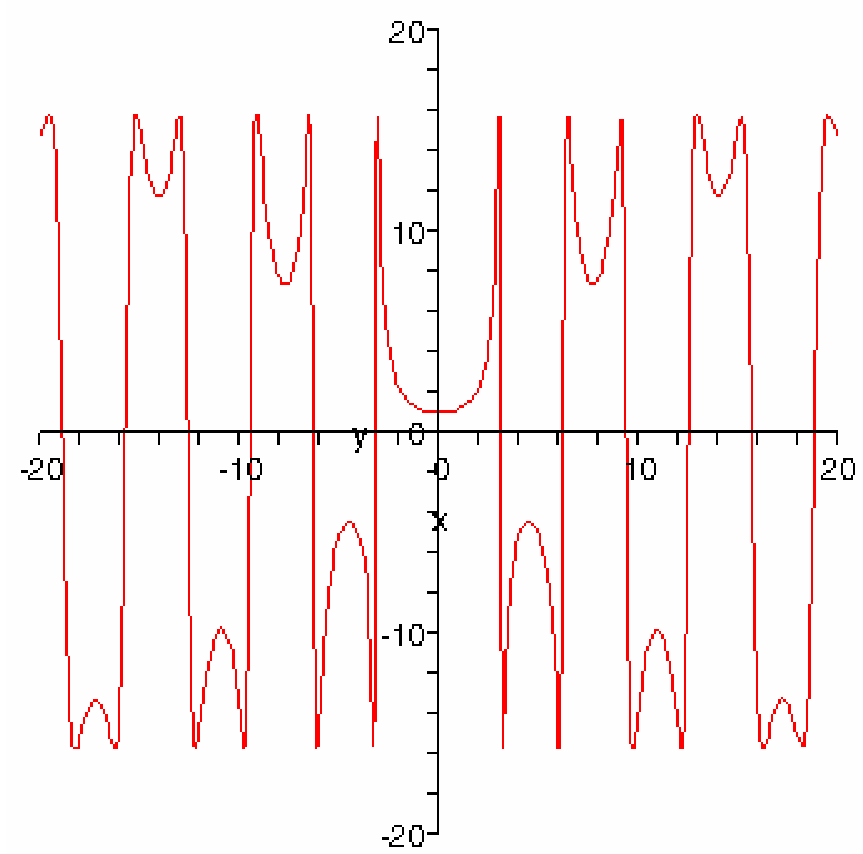
- Convolution with the kernel h completely cancels the frequencies $\frac{\nu}{2l}$ for $\nu \in \mathcal{Z}$. Vanishing frequencies cannot be recovered!
- Noise amplification for $\mathcal{F}[h](u, v) \ll 1$.

Avoiding Noise Amplification

Regularized reconstruction filter:

$$\tilde{\mathcal{F}}[\tilde{h}](u,v) = \frac{\mathcal{F}[h]}{|\mathcal{F}[h]|^2 + \epsilon}$$

Singularities are avoided by the regularization ϵ.



The size of ϵ implicitly determines an estimate of the noise level in the image, since we discard signals which are dampened below the size ϵ.