Imaging using ionizing radiations

II-SPECT

Ziad El Bitar
ziad.elbitar@iphc.cnrs.fr
Institut Pluridisciplinaire Hubert Curien, Strasbourg

ESIPAP, March 2017

Slides are courtesy of David Brasse

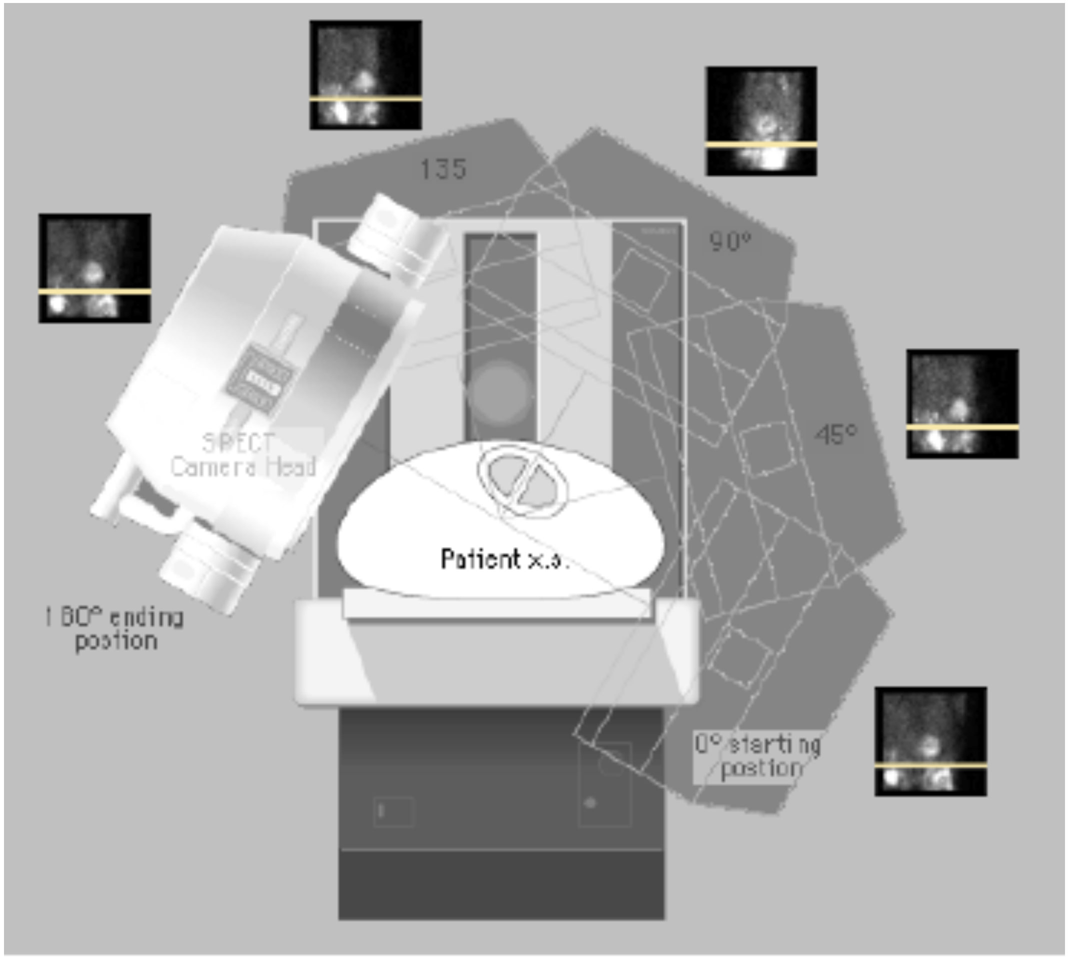
Tomographic acquisition

- More than a single projection is required in order to obtain the radiotracer distribution.
  - Many possibilities for the solution

↓

- Increasing the number of projections
  - Reduce the number of possibilities

Uniqueness of the solution for an infinity of projections

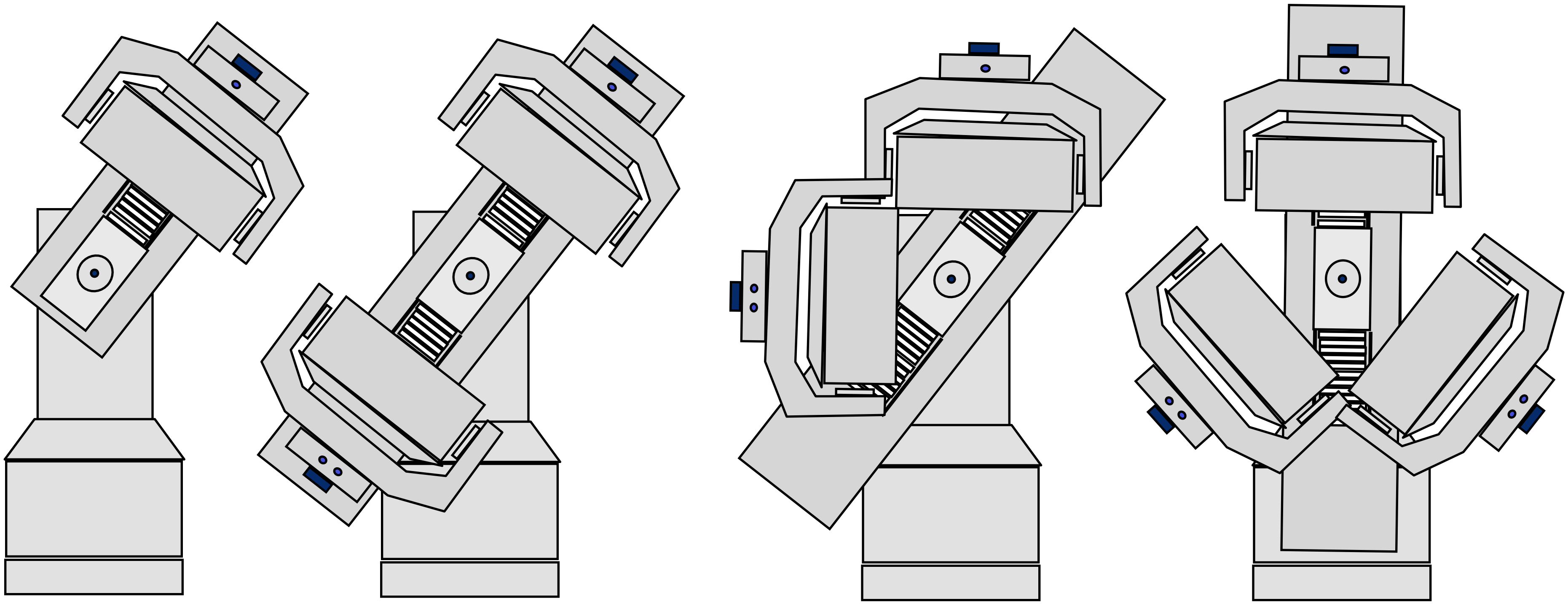


2D projection

Several angles

Reconstruction

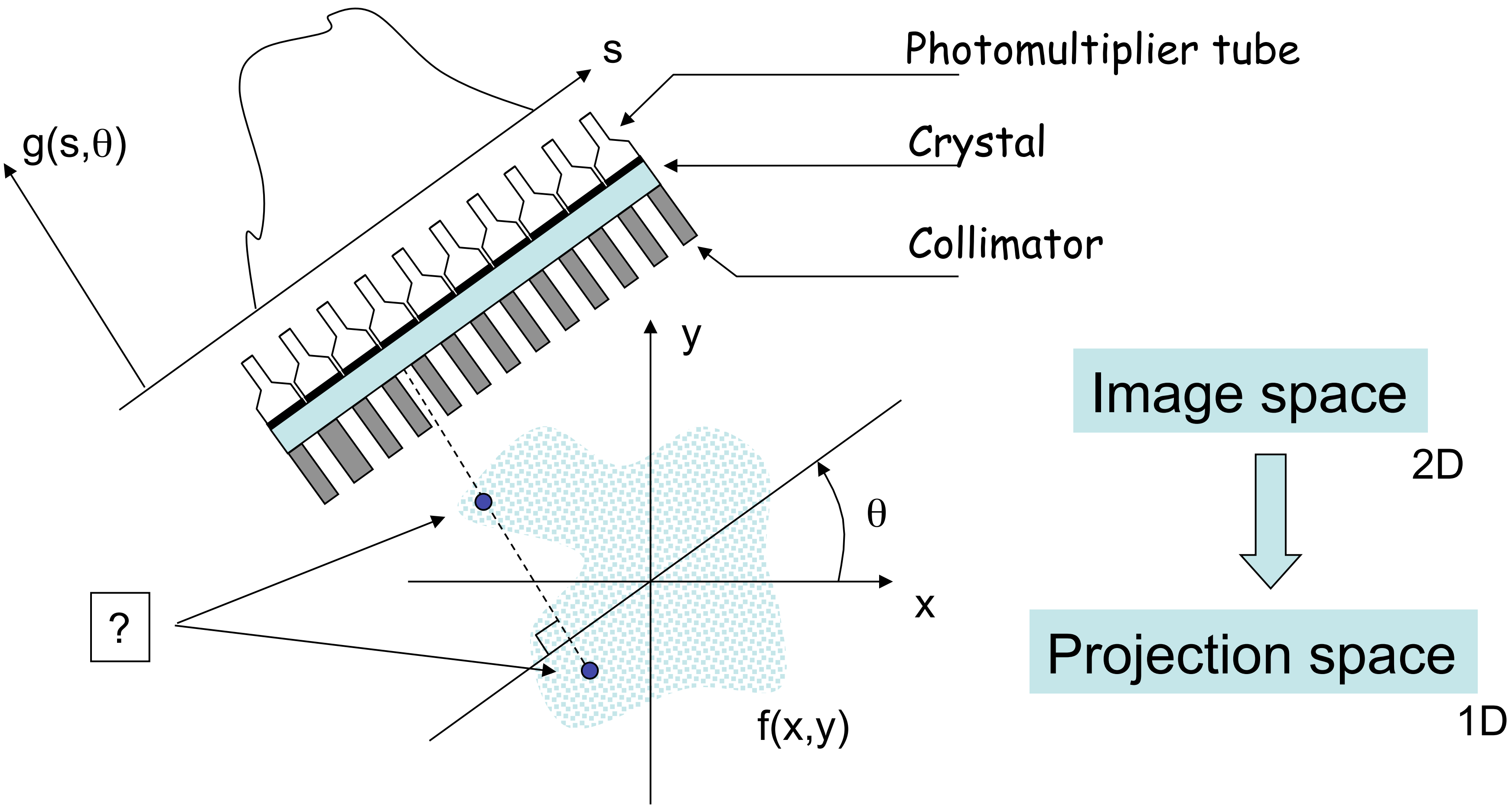
3D Volume



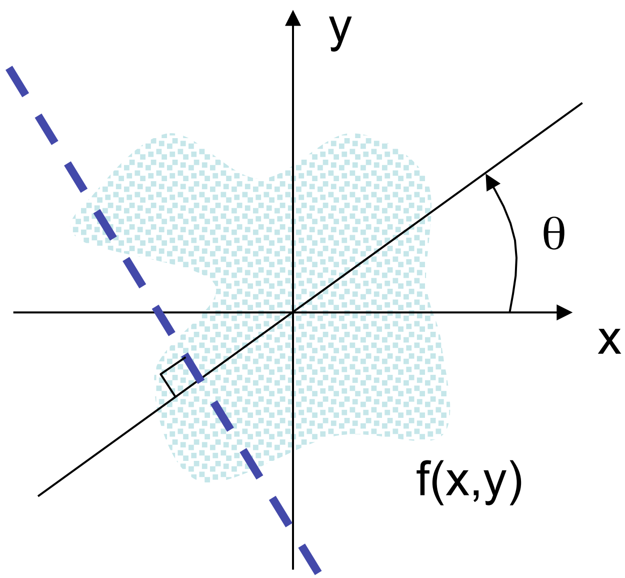
Positionning the problem

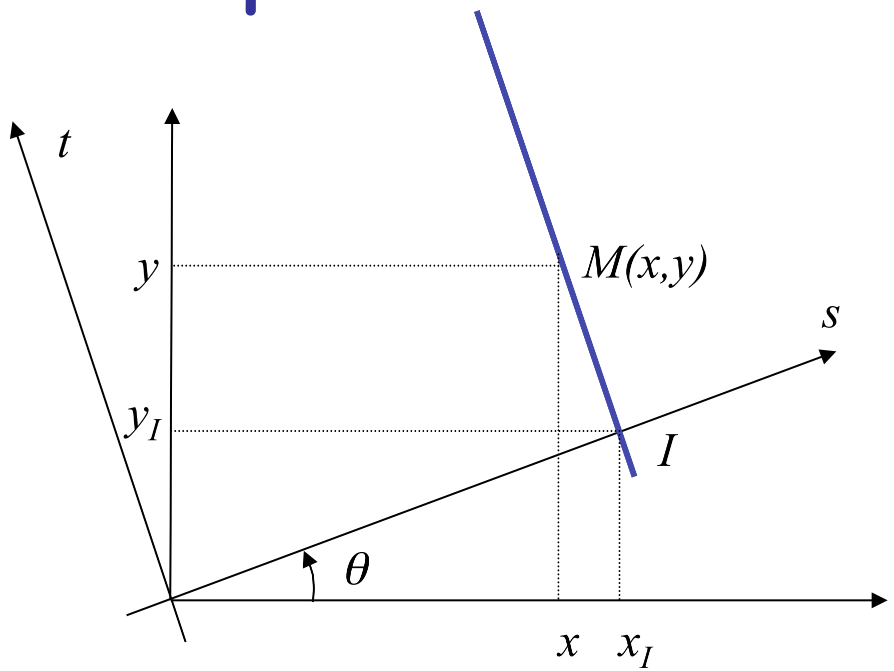
- Emission imaging
  - Injection of a radiopharmaceutical
  - Marker/tracer coupling
- Observables
  - Distribution of γ emitters in the field of view of the camera.
  - $f(x,y)$: Estimation of the number of photons emitted at (x,y).
- Hypothesis:
  - $f(x,y)$ is proportional to the concentration of the injected product.

Principle of acquisition



Projection operation





$$x_I = s\cos\theta$$
$$y_I = s\sin\theta$$

$$x_I - x = t\sin\theta$$
$$y_I - y = t\cos\theta$$

$$x = s\cos\theta - t\sin\theta$$
$$y = s\sin\theta + t\cos\theta$$

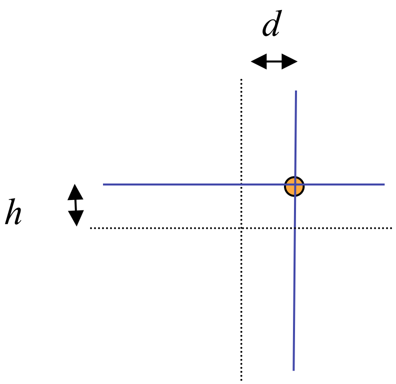
$$s = x\cos\theta + y\sin\theta$$
$$t = -x\sin\theta + y\cos\theta$$

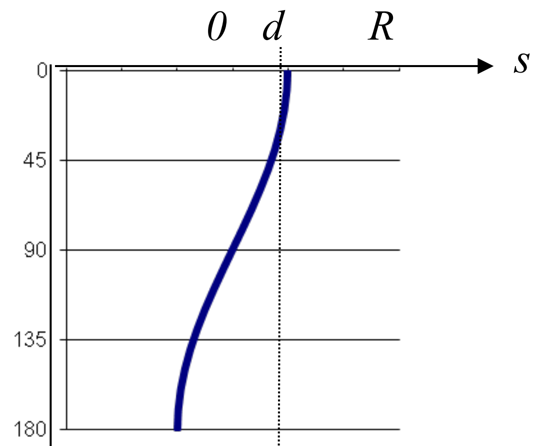
$$s = x\cos\theta + y\sin\theta$$

All points *M(x,y)* describing the LOR

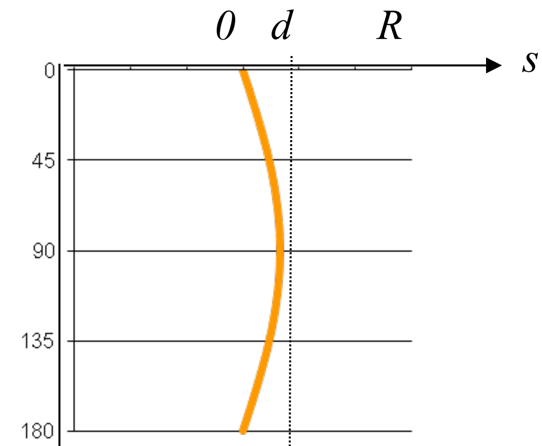
Sinogramme

$$s = x\cos\theta + y\sin\theta$$

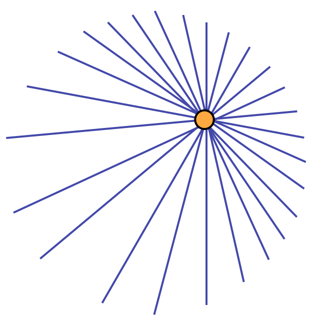


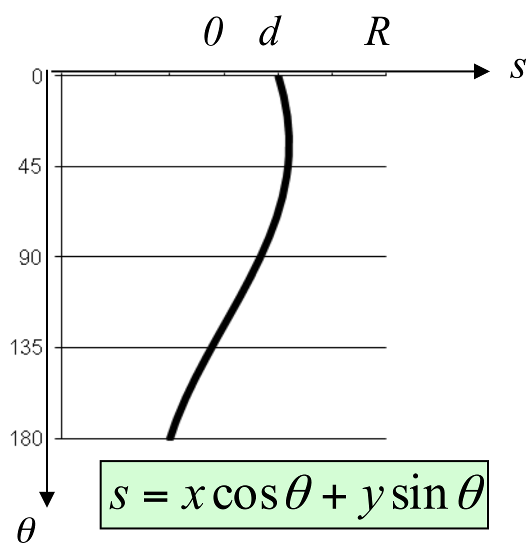


$$s = d\cos\theta$$

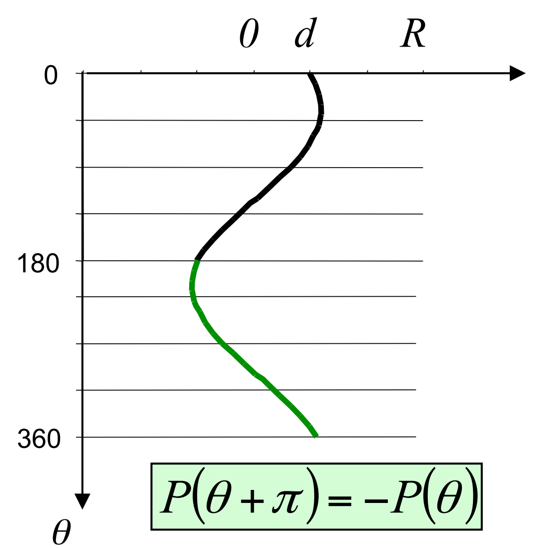


$$s = h\sin\theta$$

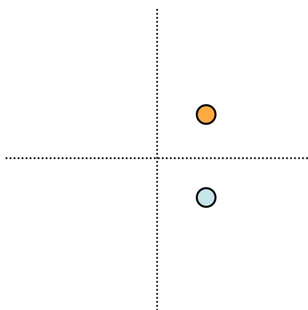


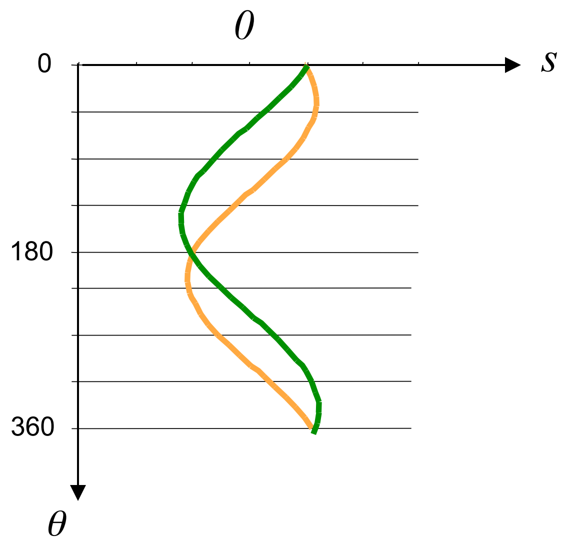


$$s = x\cos\theta + y\sin\theta$$

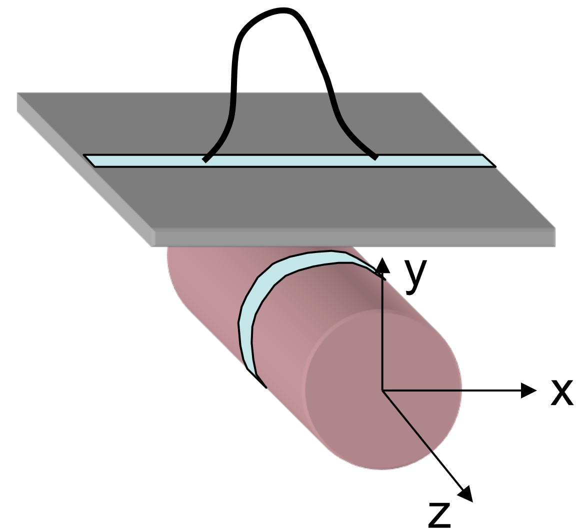


$$P(\theta+\pi) = -P(\theta)$$

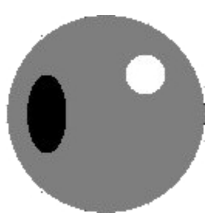


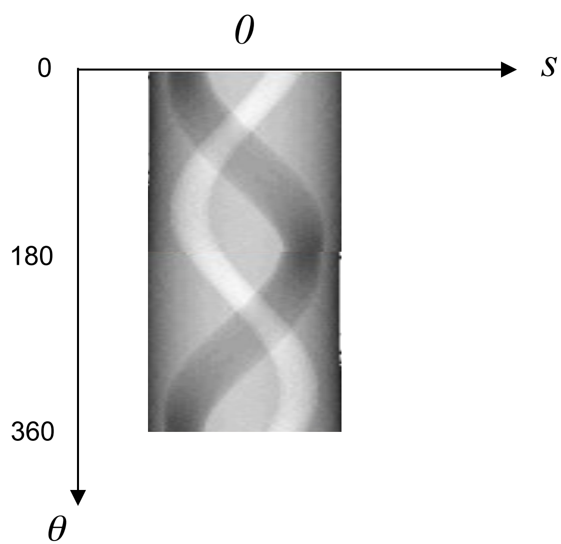


0
0
s
180
360
θ

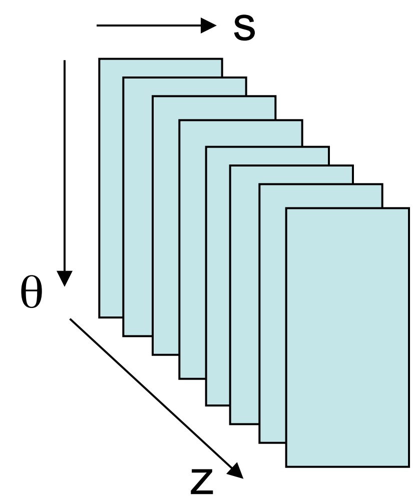


y
x
z



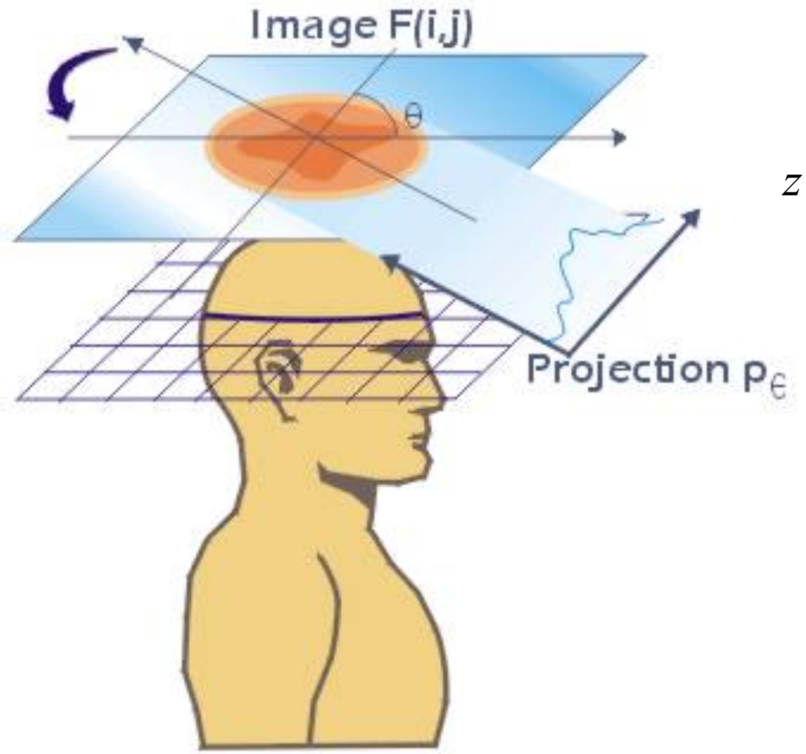


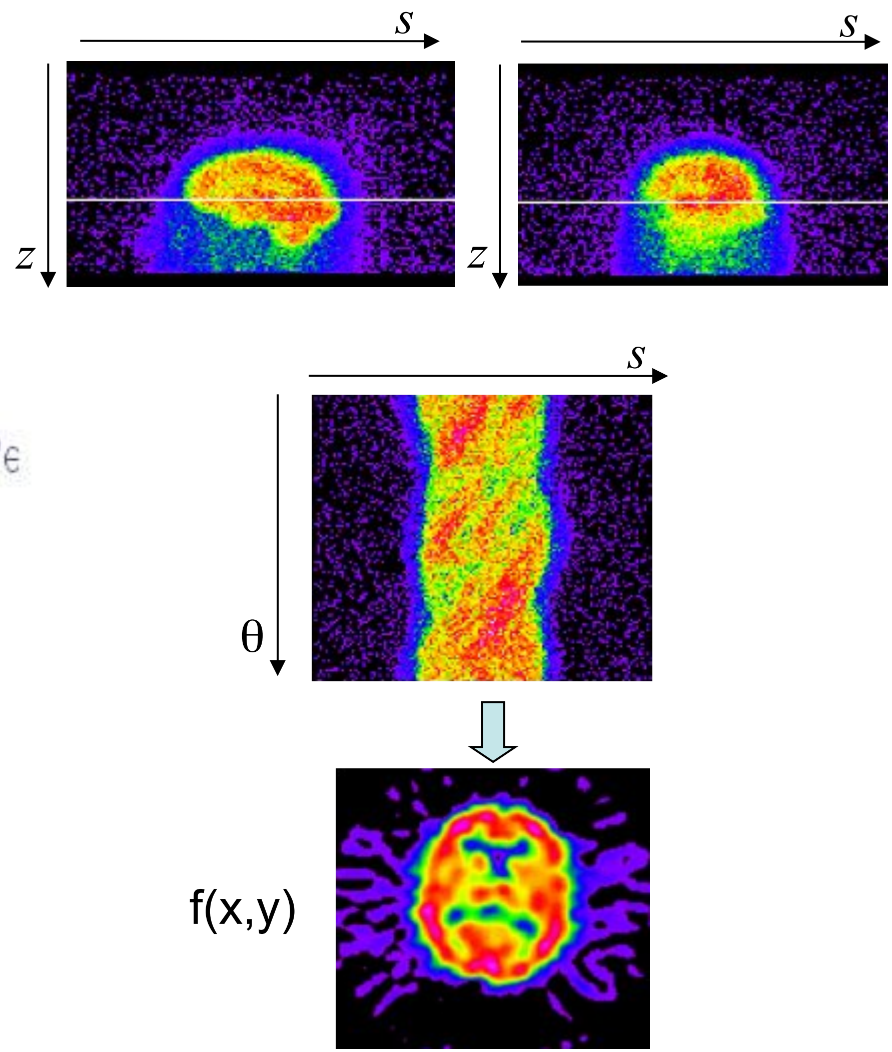
0
0
s
180
360
θ



s
θ
z

Illustration 2D





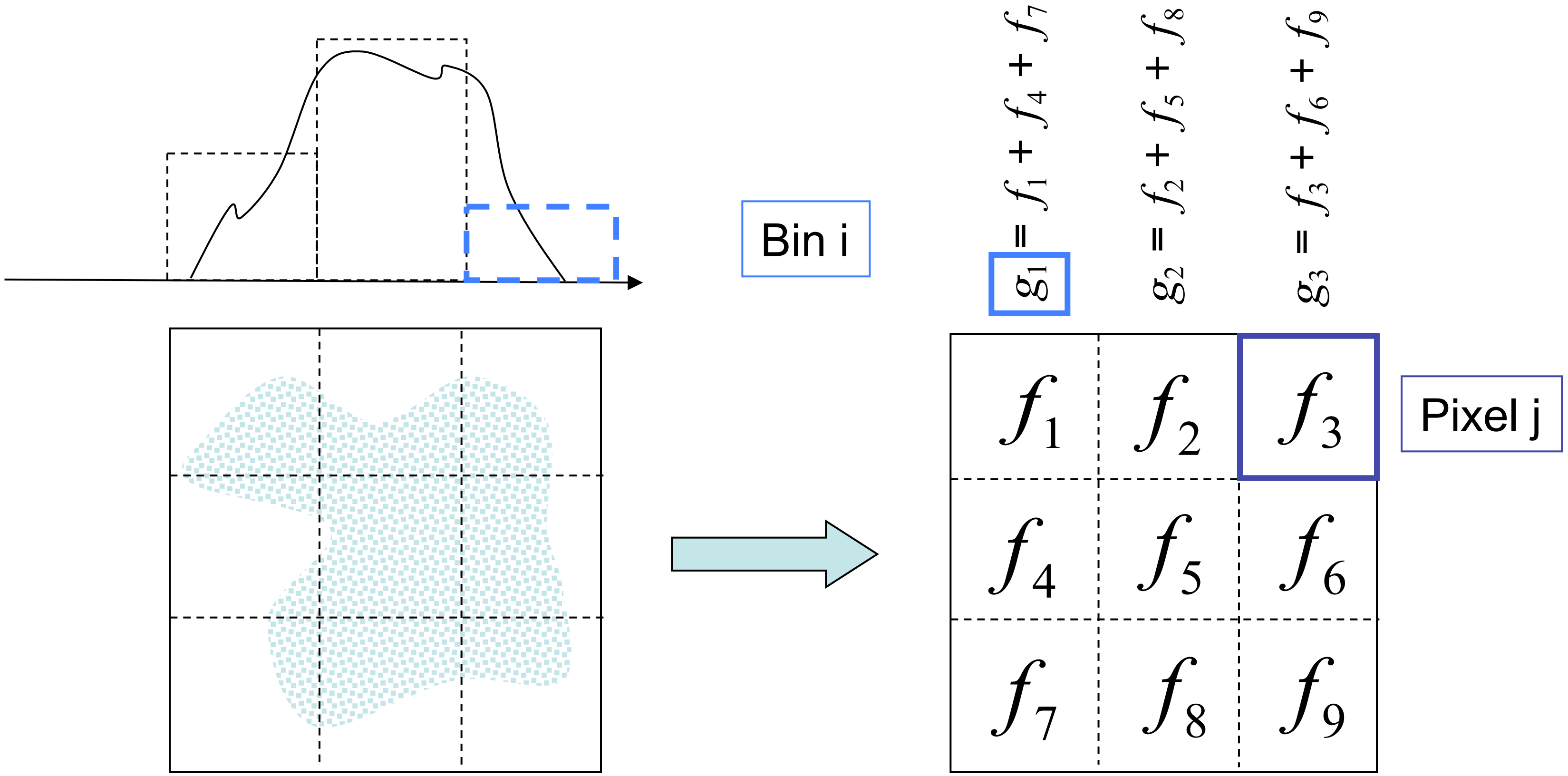
Radon Transform

In mathematics, the projection operation is defined by the Radon transform

Radon transform g(s,θ) =
Integral of f(x,y) along the line D'

$$g(s,\theta) = \int_{-\infty}^{+\infty} f(s\cos\theta - t\sin\theta, s\sin\theta + t\cos\theta)dt$$

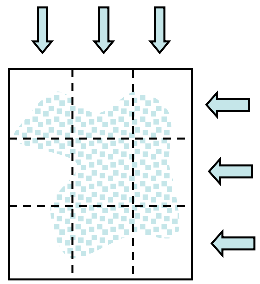
Continuous to discrete



$$g_i = a_{i1} f_1 + a_{i2} f_2 + \cdots + a_{im} f_m = \sum_{j=1}^{m} a_{ij} f_j$$

Matrix representation

$$g = Af$$



$$\begin{bmatrix} g_1 \\ g_2 \\ g_3 \\ g_4 \\ g_5 \\ g_6 \end{bmatrix} = \begin{bmatrix} 100100100 \\ 010010010 \\ 001001001 \\ 111000000 \\ 000111000 \\ 000000111 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \\ f_7 \\ f_8 \\ f_9 \end{bmatrix}$$

$$\underbrace{\qquad}_{g} \quad \underbrace{\qquad\qquad}_{A} \quad \underbrace{\qquad}_{f}$$

$$\begin{aligned}
g_1 &= f_1 + f_4 + f_7 \\
g_2 &= f_2 + f_5 + f_8 \\
g_3 &= f_3 + f_6 + f_9 \\
g_4 &= f_1 + f_2 + f_3 \\
g_5 &= f_4 + f_5 + f_6 \\
g_6 &= f_7 + f_8 + f_9
\end{aligned}$$

Definitions

- *A: Projection operator*
- *a_{ij}: Weight factor representing the contribution of pixel j to the number of counts detected in bin i.*
- *In other words: probability that a photon emitted from pixel j is detected in bin i.*

Problem inversion

In theory, direct methods exist to solve the equation:

$$g = Af$$

These methods, called direct inversion consist in finding A^{-1}

$$f = A^{-1}g$$

Many difficulties

- Inversion of A
- A^{-1} does not exist
- A^{-1} is not unique

In practice: inverse problem are badly conceived

- Solution is not unique and *A* is unstable:
  - Data contamination by noise
  - Finite number of projections
- Approached solution

Problem:

Knowing the sinogramm,
What is the radioactive distribution f(x,y)?

Image reconstruction

Analytical

- FBP
- BPF
- “Gridding”
- …

Iterative

Algebraic

Least squares

- CG
- CD
- ISRA
- …

Statistical

Poisson

- EM
- OSEM
- SAGE
- CG
- …

Analytic algorithms

- Backprojection operation
- Central slice theorem
- Backprojection + filtering
- Backprojection of filtered projections
- Fourier domain (space)

Operator: Back Projection

Inverse operrator

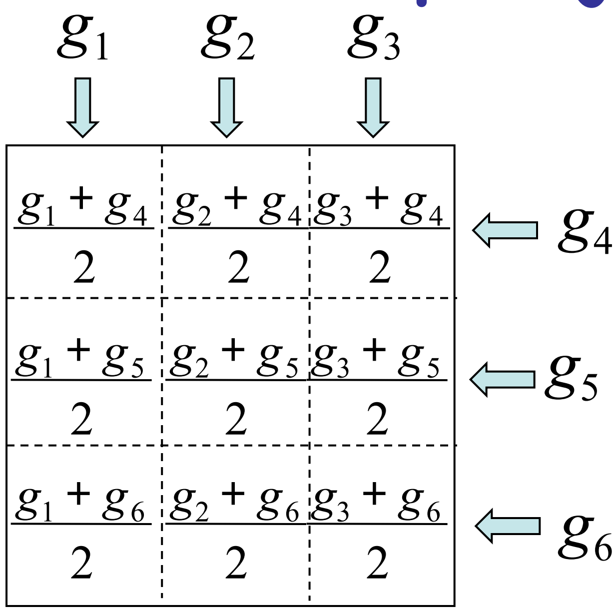
$$b(x,y)=\int_0^{\pi} g(s,\theta)\,\mathrm{d}\theta$$

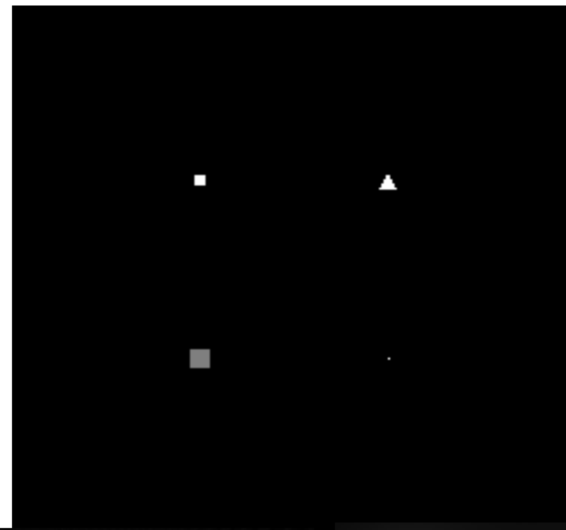
$$\widetilde{b}(x,y)=\sum_{k=1}^{p} g(s_k,\theta_k)\,\Delta\theta$$

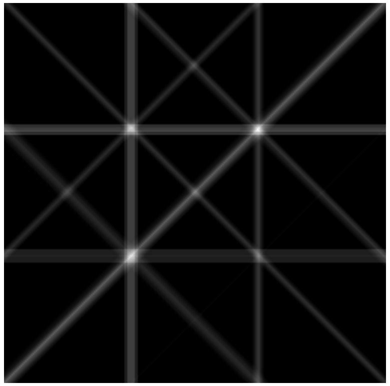
p: number of projections

$\Delta\theta$: Sampling (π/p)

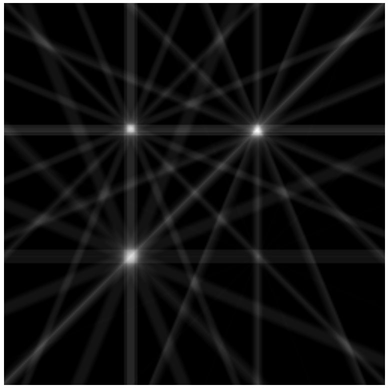
Back projection: Artifacts



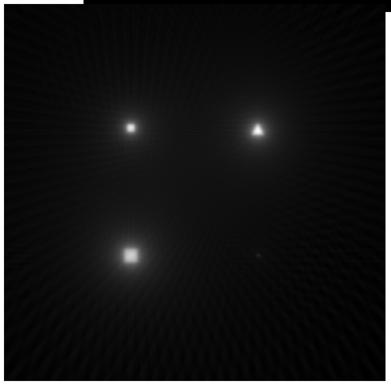




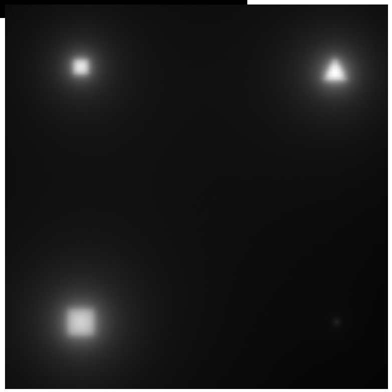
p=4



p=8



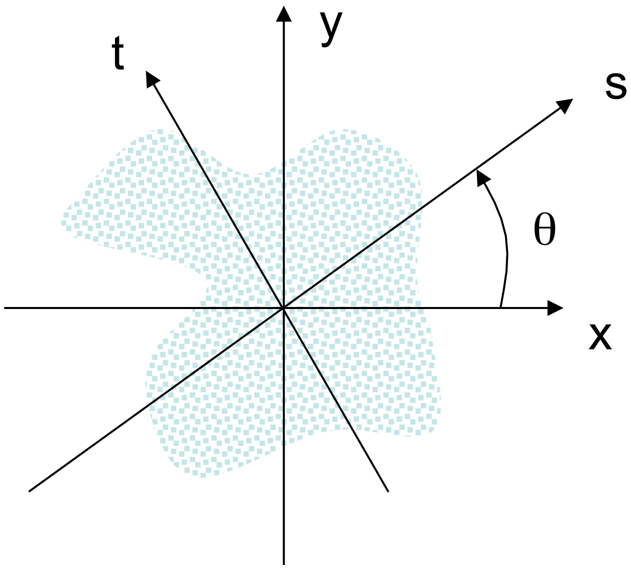
p=64



p=256

Central slice theorem

$$g(s,\theta) = \int_{-\infty}^{+\infty} f(x,y)\,dt$$



TF (g)

$$G_{10}(\upsilon_s,\theta) = \int_{-\infty}^{+\infty} g(s,\theta)\,\mathrm{e}^{-2i\pi\upsilon_s s}\,\mathrm{d}s$$

$$G_{10}(\upsilon_s,\theta) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,\mathrm{e}^{-2i\pi\upsilon_s s}\,\mathrm{d}s\mathrm{d}t$$

$$s = x\cos\theta + y\sin\theta$$

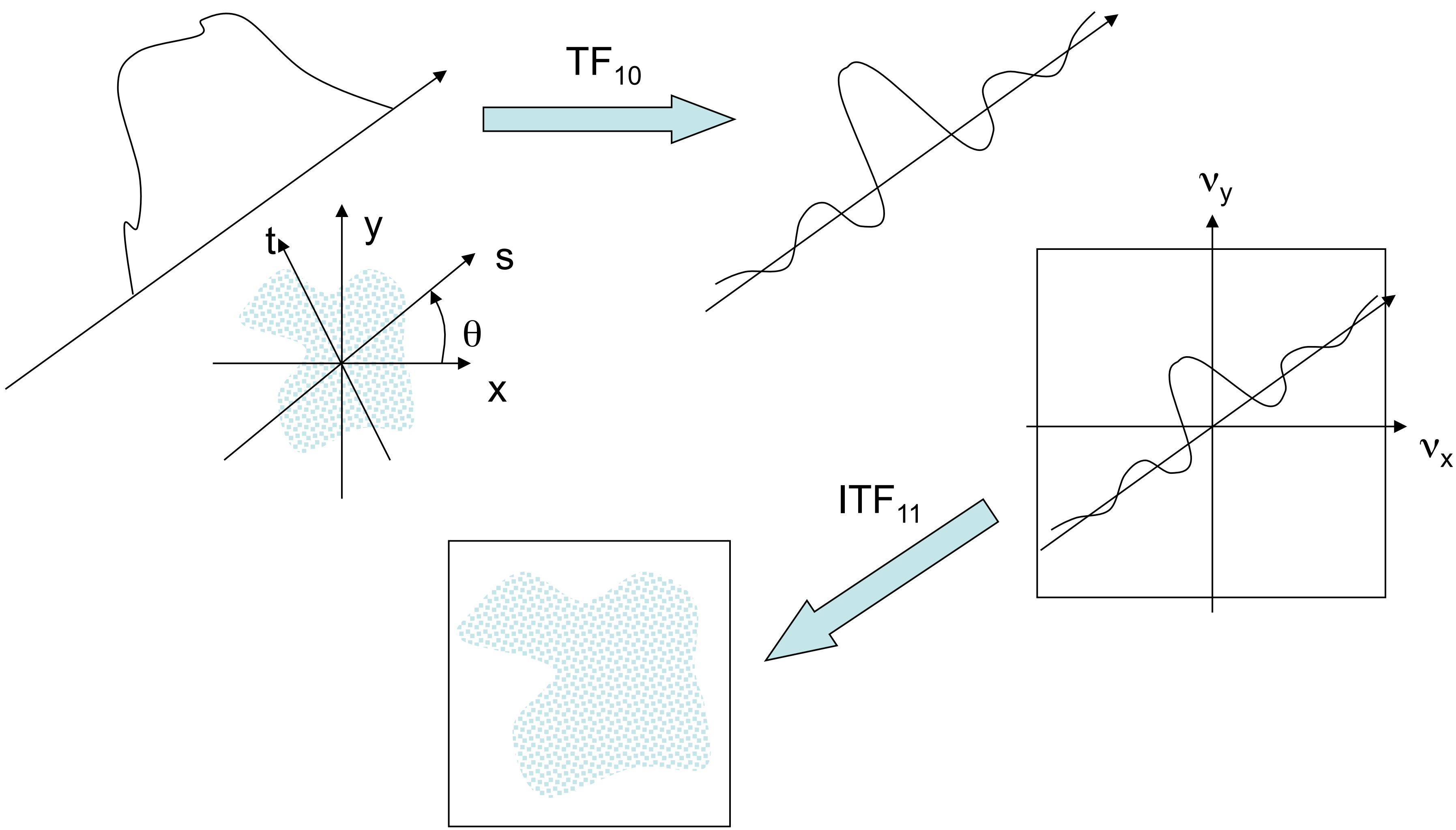
$$\upsilon_x = \upsilon_s \cos\theta$$

$$\upsilon_y = \upsilon_s \sin\theta$$

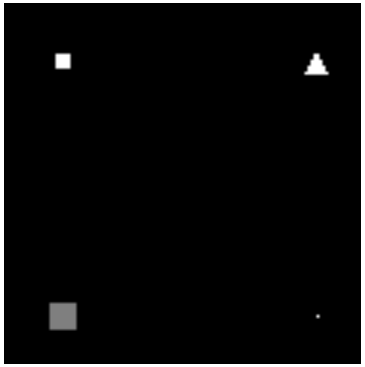
$$G_{10}(\upsilon_s,\theta) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,\mathrm{e}^{-2i\pi(x\upsilon_x + y\upsilon_y)}\,\mathrm{d}x\mathrm{d}y$$

$$F_{11}(\upsilon_x,\upsilon_y)\Big|_{\upsilon_t=0} = G_{10}(\upsilon_s,\theta)$$

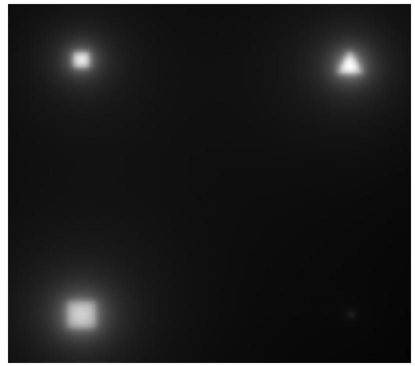
Graphical illustration



Sampling



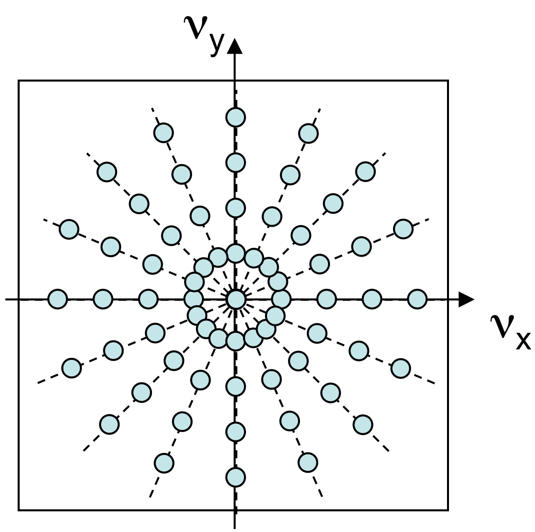
?

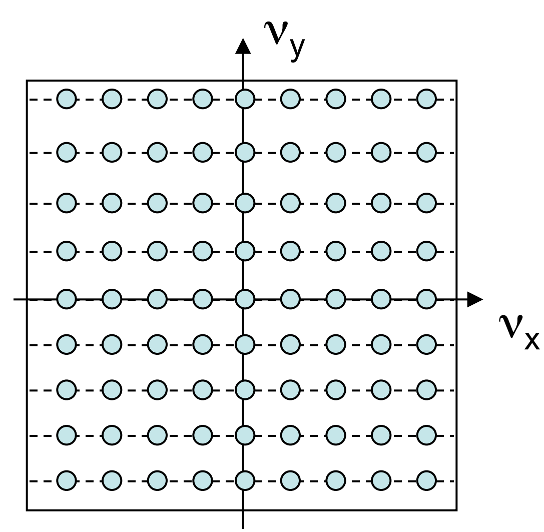


Low frequencies upsampling

Raising high frequencies

Polar -> Cartesien





Proof

$$f(x,y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}F_{11}(\upsilon_x,\upsilon_y)\,\mathrm{e}^{2i\pi(x\upsilon_x+y\upsilon_y)}\,\mathrm{d}\upsilon_x\mathrm{d}\upsilon_y$$

TF 2D

$$f(x,y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}G_{10}(\upsilon_s,\theta)\,\mathrm{e}^{2i\pi(x\upsilon_x+y\upsilon_y)}\,\mathrm{d}\upsilon_x\mathrm{d}\upsilon_y$$

$$F_{11}(\upsilon_x,\upsilon_y)=G_{10}(\upsilon_s,\theta)$$

Changement of variables:

$$(\upsilon_x,\upsilon_y)\rightarrow(\upsilon_s,\theta)$$

$$\upsilon_x=\upsilon_s\cos\theta$$
$$\upsilon_y=\upsilon_s\sin\theta$$

$$s=x\cos\theta+y\sin\theta$$

$$\mathrm{d}\upsilon_x\mathrm{d}\upsilon_y=|\upsilon_s|\,\mathrm{d}\upsilon_s\mathrm{d}\theta$$

$$f(x,y)=\int_{0}^{\pi}\underbrace{\int_{-\infty}^{+\infty}G_{10}(\upsilon_s,\theta)\,|\upsilon_s|\,\mathrm{e}^{2i\pi\upsilon_s s}\,\mathrm{d}\upsilon_s}\,\mathrm{d}\theta$$

$$f(x,y)=\int_{0}^{\pi}g'(s,\theta)\,\mathrm{d}\theta$$

Filtering

- Exact inversion is not possible for two reasons:
  - Discrete sampling -> Limited space
    - Shannon: fréquency max. reconstructed : Nyquist=$1/2\Delta s$
  - Presence of statistical noise
    - Utilization of « ramp » filter -> Noise amplification

→ Utilisation of an apodisation window

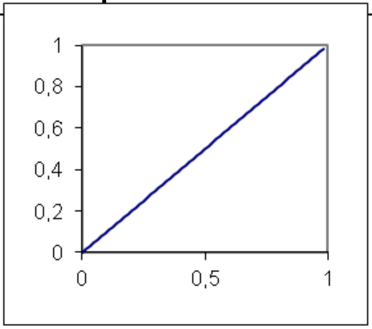
Apodisation window

- Cut-off frequency influence:
  - The resolution of the reconstructed image
  - Noise properties

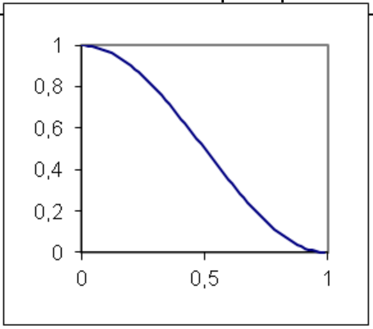
Hann Filter

$$W(\upsilon_s) = \begin{cases} 0{,}5\left(1+\cos\left(\dfrac{\pi \upsilon_s}{\upsilon_c}\right)\right), & |\upsilon_s| < \upsilon_c \\ 0 & |\upsilon_s| \geq \upsilon_c \end{cases}$$

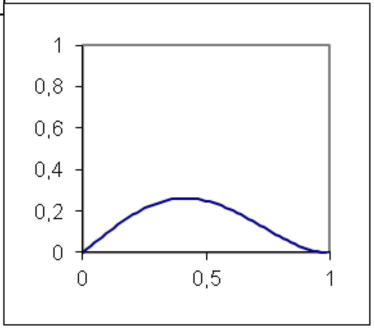
$$0 \leq \alpha \leq 1$$



$|\upsilon_s|$

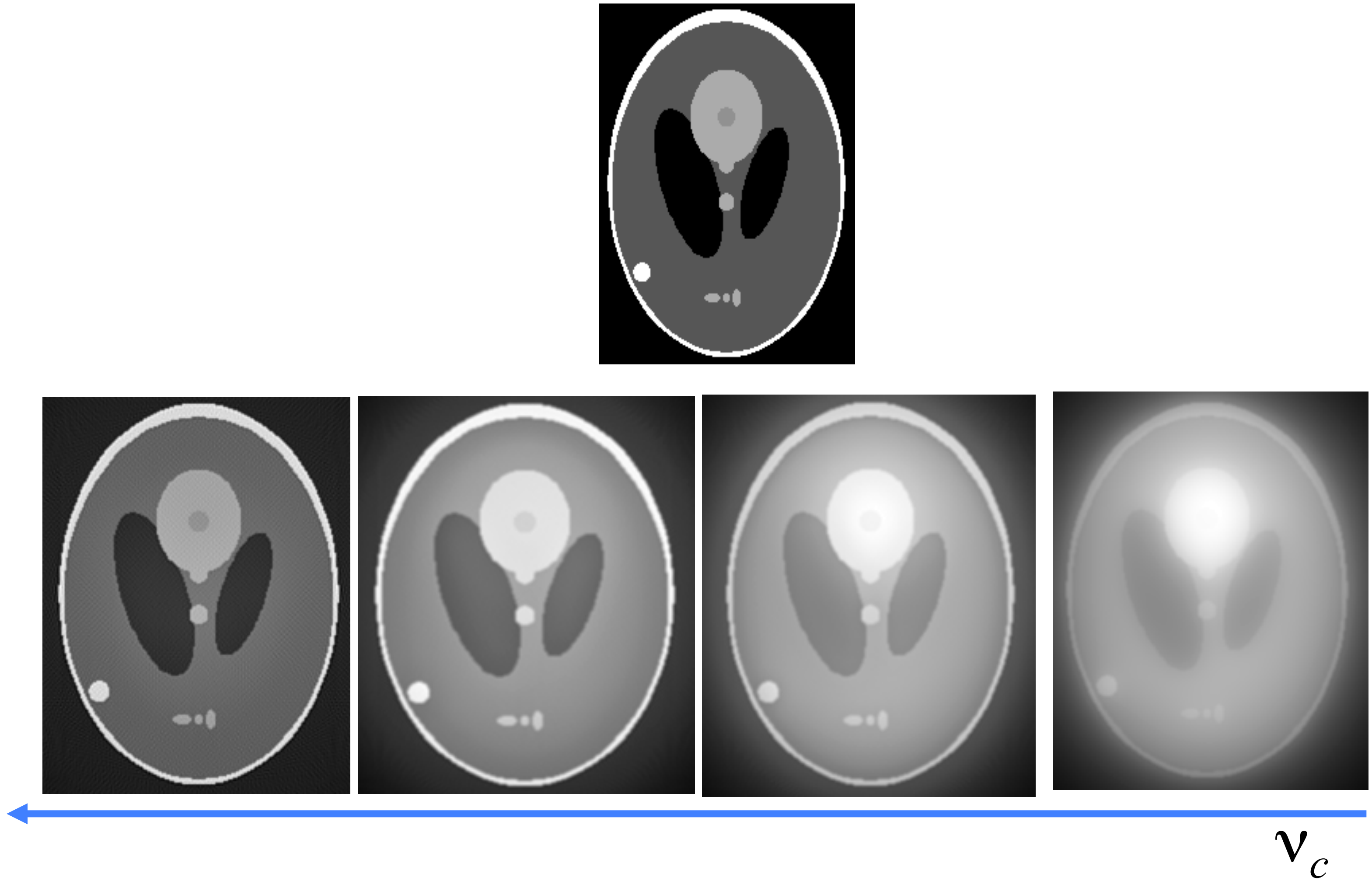


$W(\nu_s)$

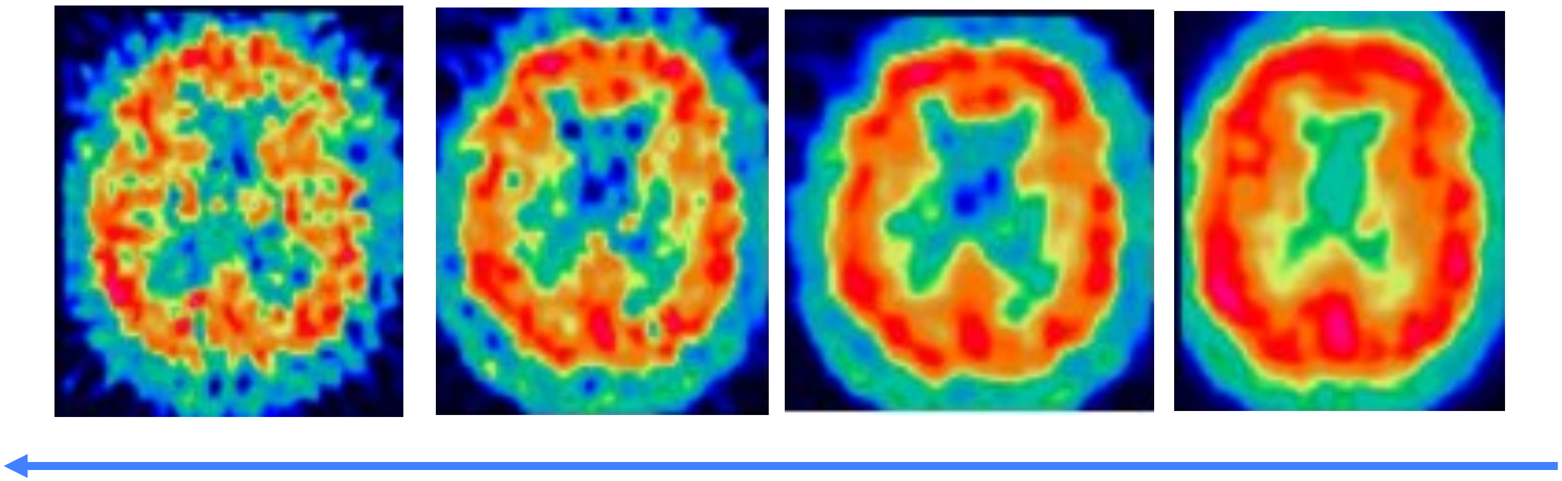


$|\upsilon_s|$ x $W(\nu_s)$

Cut-off frequency: Resolution



Cut-off frequency: Noise

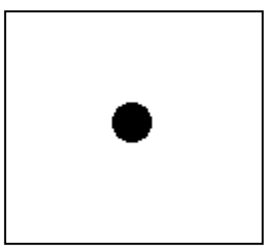


ν_c

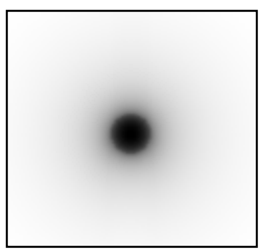
Back projection + Filtering

$$f(x,y) = b(x,y) \otimes psf(x,y)$$

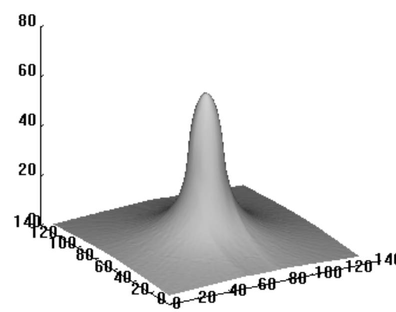
- Drawback: requires an entire huge back projection matrix *b(x,y)*
  - Otherwise, tuncation artifacts

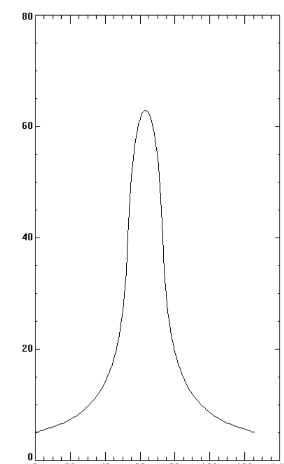


f (x,y)

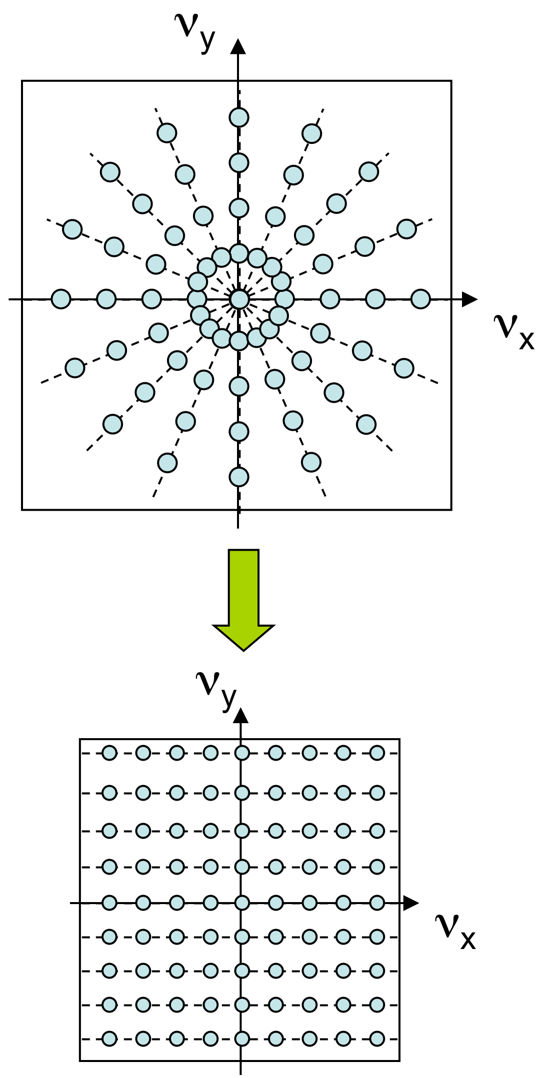


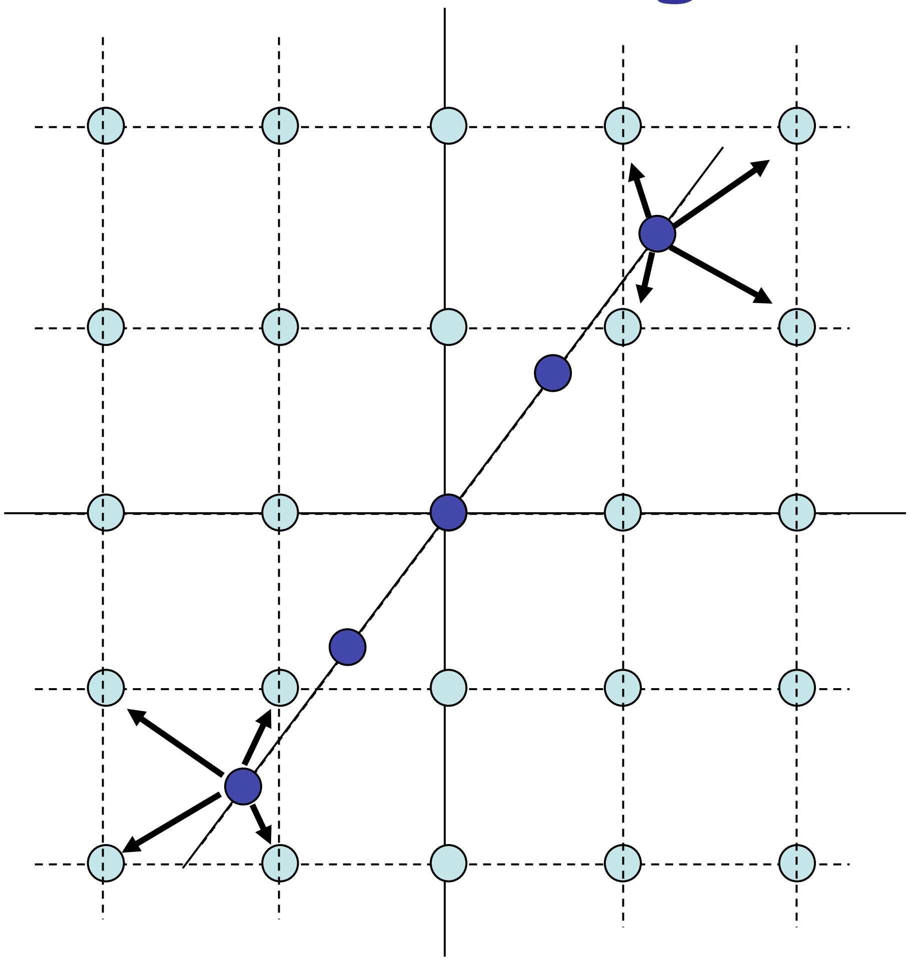
b(x,y)





Fourier space: Gridding

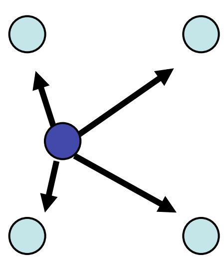




Drawback of fourier methods

- Sensitivity to noise
- Interpolation kernel in Fourier

Multiplication in real space



$$F_{11}(\upsilon_x,\upsilon_y)\otimes W(\upsilon_x,\upsilon_y)\rightarrow f(x,y)\cdot w(x,y)$$

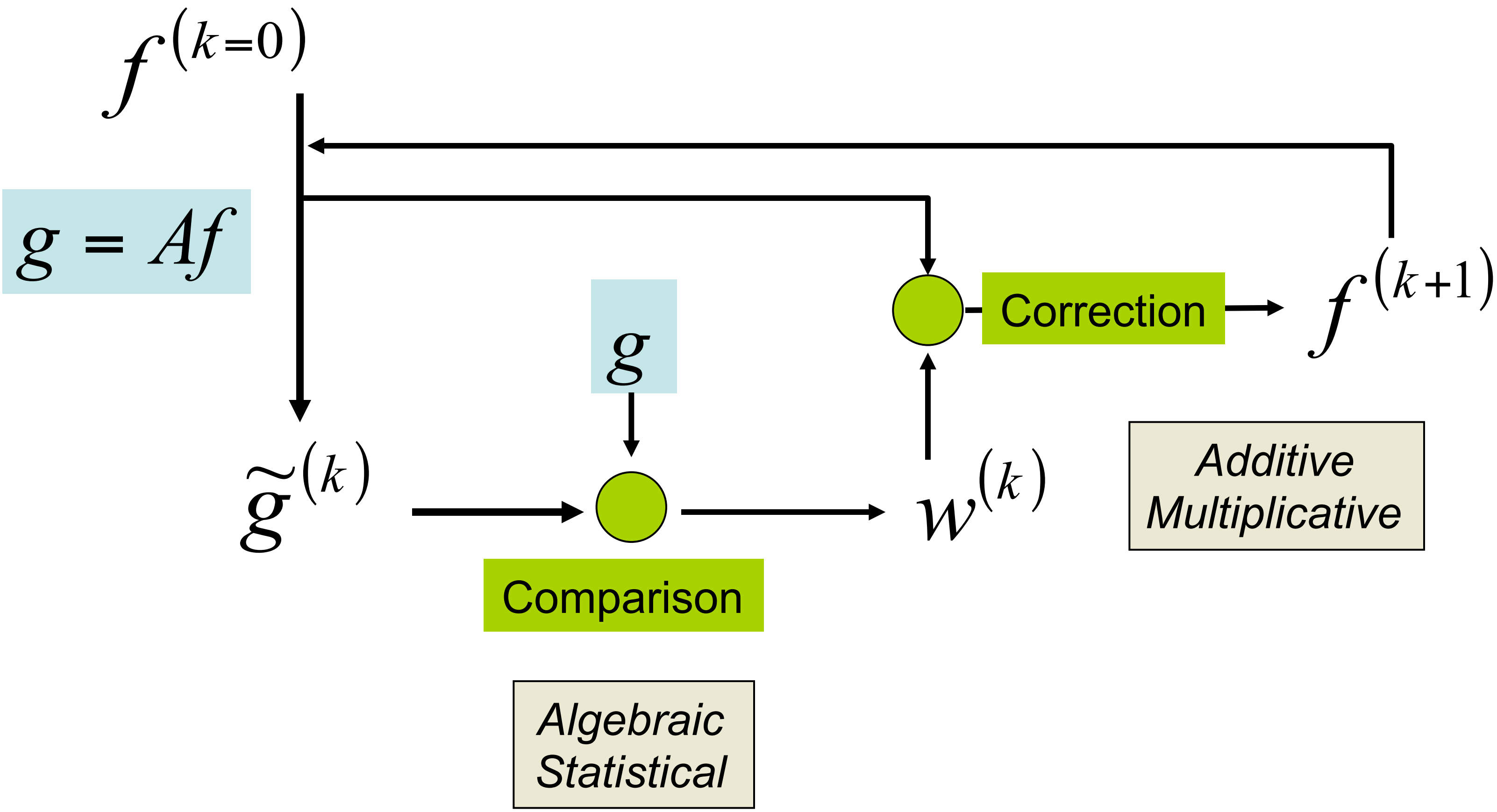
Iterative algorithms

- Find the f vector, solution of the equation

$$g = Af$$

- Iterative algorithms are based on the principle of finding a solution by successive estimations.
- The projection corresponding to the current estimation is compared to the acquired projections.
- The comparison result is used to modify the estimation and create a new one.

Principle



$f^{(k=0)}$
$g = Af$
g
Correction
$f^{(k+1)}$
$\tilde{g}^{(k)}$
$w^{(k)}$
Comparison
Additive
Multiplicative
Algebraic
Statistical

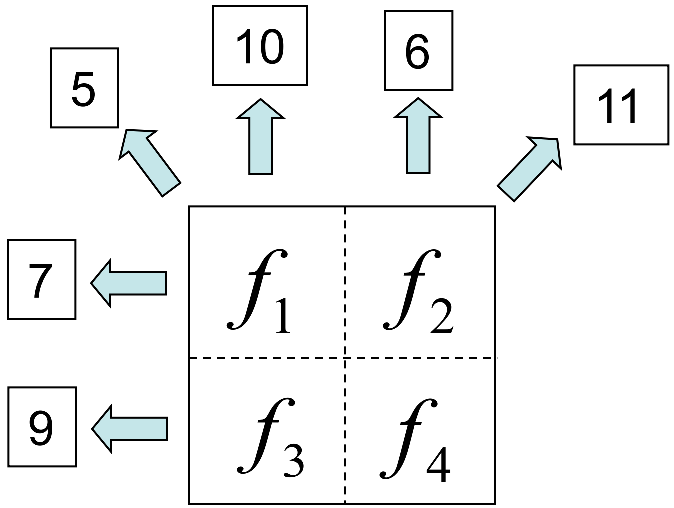
Algebraic methods: ART

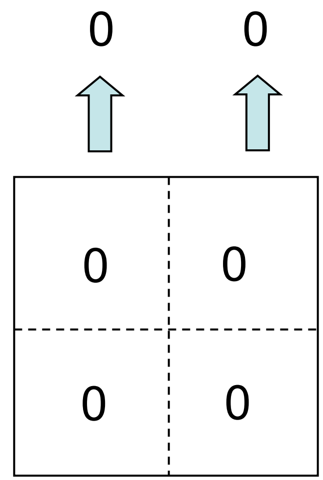
- ART: « Algebraic Reconstruction Technique »

$$f_j^{(k+1)} = f_j^{(k)} + \frac{g_i - \sum_{j=1}^{N} f_{ji}^{(k)}}{N}$$

Example ART-1

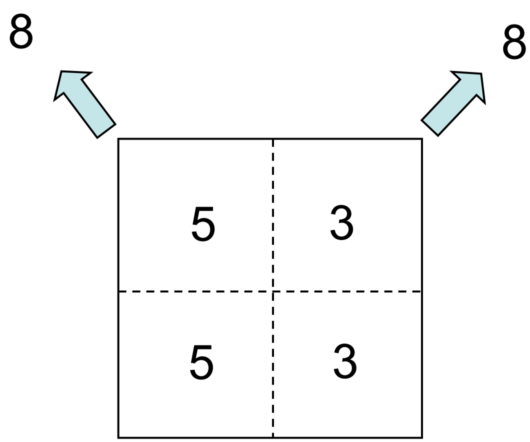
$$f_j^{(k+1)} = f_j^{(k)} + \frac{g_i - \sum_{j=1}^{N} f_{ji}^{(k)}}{N}$$





$$f_1^{(1)} = f_3^{(1)} = 0 + \frac{10-0}{2} = 5$$

$$f_2^{(1)} = f_4^{(1)} = 0 + \frac{6-0}{2} = 3$$



$$f_1^{(2)} = 5 + \frac{5-8}{2} = 3{,}5$$

$$f_2^{(2)} = 3 + \frac{11-8}{2} = 4{,}5$$

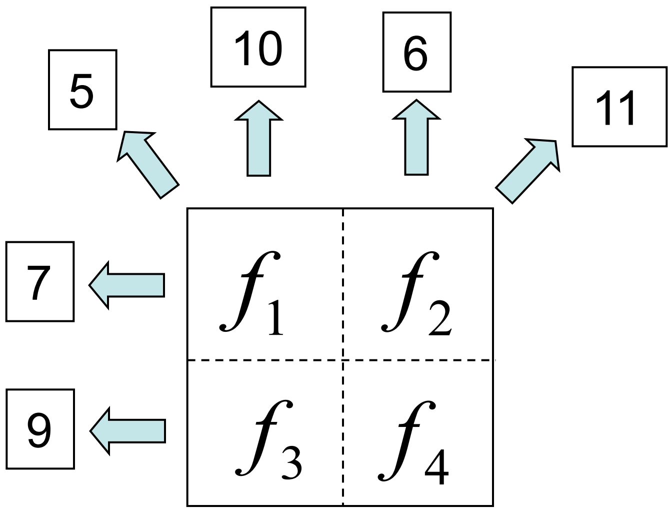
$$f_3^{(2)} = 5 + \frac{11-8}{2} = 6{,}5$$

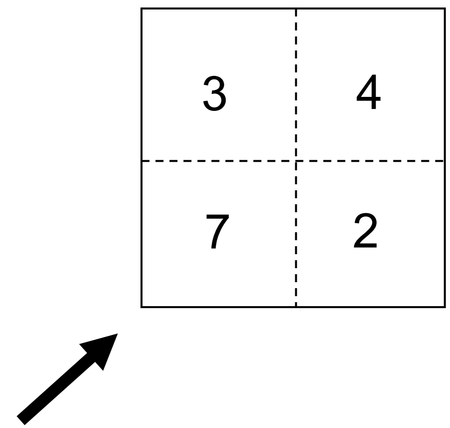
$$f_4^{(2)} = 3 + \frac{5-8}{2} = 1{,}5$$

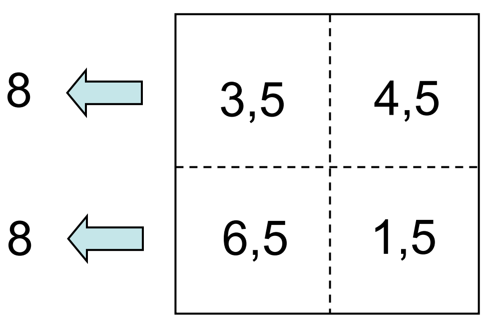
3,5	4,5
6,5	1,5

Example ART-2

$$f_j^{(k+1)} = f_j^{(k)} + \frac{g_i - \sum_{j=1}^{N} f_{ji}^{(k)}}{N}$$







$$f_1^{(3)} = 3{,}5 + \frac{7-8}{2} = 3$$

$$f_2^{(3)} = 4{,}5 + \frac{7-8}{2} = 4$$

$$f_3^{(3)} = 6{,}5 + \frac{9-8}{2} = 7$$

$$f_4^{(3)} = 1{,}5 + \frac{9-8}{2} = 2$$

Why using statistical methods?

Advantages:

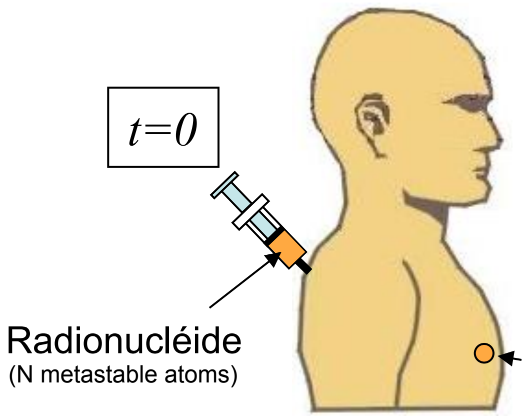
- Constraints on the object: non negativity, support…
- Incorporation of physics models
  - Photons transportation / geometrical efficiency…
- Appropriated statistical model (Noise reducing).
- Flexibility regarding geometry.
- Incorporation of anatomical information.

Drawbackss:

- Computation time.
- Complex model.
- Tedious implementation.

Investigating the object

- Realize the image of the radiotracer distribution



$X_k(t)$: position of the atom @ time t

Imaging system: provide $X_k(t)$?

<u>First hypothesis</u>: $X_k(t)$ Independent random variables distributed according to the same probability density function $f_{\vec{X}(t)}(\vec{x})$

Imaging system:provide $f_{\vec{X}(t)}(\vec{x})$!!

<u>Secone hypothesis</u>: Atoms distribution follows a Poisson law

Radioactive decay

An atom can only be observed when it deexcitates and emits photons. The deexcitation time af an atom kth is a random variable T_k .

Third hypothesis: The T_k are independent random variables

Fourth hypothesis: Each T_k has an exponential distribution whose mean is $\mu_T = t_{1/2}/\ln 2$

$t_{1/2}$ = « half-life »

The photon emission is a statistical processus that follows a Poisson law

Statistics of an ideal counter

$K(t,V)$: number of atomes @ time t, located in a volume V
$K(t,V)$: counting process following a Poisson law with a mean

$$E\left[K\left(t,V\right)\right]=\int_0^t\int_V \lambda\left(\vec{x},\tau\right)\mathrm{d}\vec{x}\,\mathrm{d}\tau$$

with

$$\lambda\left(\vec{x},t\right)=\mu_N\,\frac{\mathrm{e}^{-t/\mu_T}}{\mu_T}\cdot f_{\vec{X}(t)}\left(\vec{x}\right)$$

Detection element

For example: one element of the sinogram (bin)

does not correspond necessarily to a physical element of the detector

Fifth hypothesis: Each desintegration produce a detected event in one bin at least.

If a fraction of the event is attributed to 1 bin
$\Rightarrow$ Counting statistics follows a law different than Poisson.

Detection efficiency

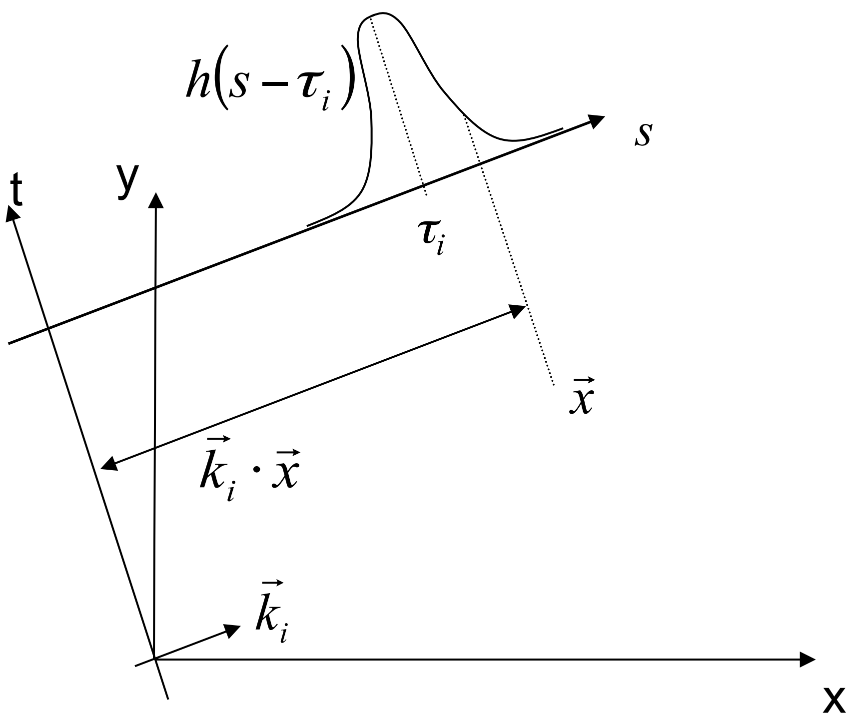
$s_i(\vec{x})$ Probability of detecting in bin i, an event coming from position x

PSF: Impulse response of the detection system
« Point Spread Function »

$$s_i(\vec{x}) = h\left(\vec{k}_i \cdot \vec{x} - \tau_i\right)$$

$$h(s) = \delta(s)$$

Ideal detector



Detection efficiency

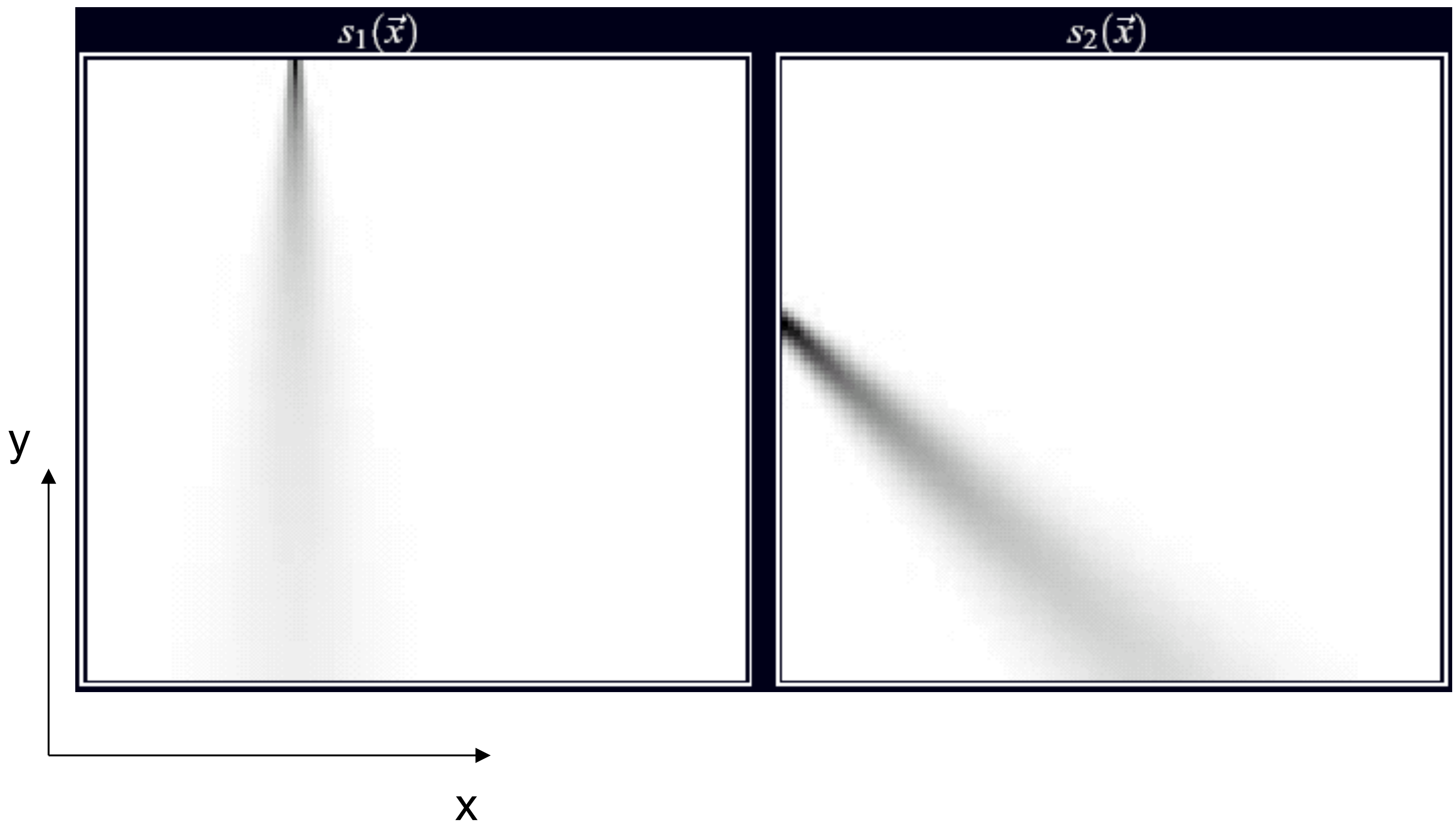
$$s_i(\vec{x})$$

Including:

- The geometry / solid angle of detection
- The collimation
- The scatter
- The attenuation
- Detector response
- Detection efficiency of the detector
- Positon range, acolinearity, etc…

Examples

Detection efficiency for an Anger gamma camera



Acquisition

- Register events
  - for t between t_1 and t_2
- Y_i: number of events registered by the ith detector element
- $\{Y_i : i=1,\ldots,n_d\}$ represents the sinogram data.

In summary,

$$Y_i \sim Poisson\left\{\int s_i(\vec{x})\lambda(\vec{x})\mathrm{d}\vec{x}\right\}$$

With

$$\lambda(\vec{x},t) = \mu_N \frac{\mathrm{e}^{-t/\mu_T}}{\mu_T} \cdot f_{\vec{X}(t)}(\vec{x})$$

$$\lambda(\vec{x}) = \mu_N \int_{t_1}^{t_2} f_{\vec{X}(t)}(\vec{x}) \frac{\mathrm{e}^{-t/\mu_T}}{\mu_T}\mathrm{d}t$$

= Emission density

Poisson Statistical Model

Measurements = real events + noise

Sources of noise:

- cosmic rays
- ambient (surrounding) noise
- All counts not taken into account in $s_i(x)$

$$Y_i \sim Poisson\left\{\int s_i(\vec{x})\lambda(\vec{x})\,\mathrm{d}\vec{x} + r_i\right\}, \quad i = 1,\ldots,n_d$$

Mean number of events originated from a noise source in bin i

Problem posed by reconstruction

Estimate the emission density λ using:

$$Y_i \sim Poisson\left\{\int s_i(\vec{x})\lambda(\vec{x})\mathrm{d}\vec{x} + r_i\right\}, \quad i = 1,\ldots,n_d$$

$\{Y_i = y_i\}_{i=1}^{n_d}$	Events collected in bin i
$s_i(\vec{x})$	Detection efficiency in bin i
r_i	Noise source in bin i

In summary: five multiple choice parts

- -1- Object description $\lambda(\vec{x})$
- -2- Physical model of the system $s_i(\vec{x})$
- -3- Statistical model of the measurement Y_i
- -4- Optimization criteria
- -5- Used algorithm

-1- Object description

$\lambda(\vec{x})$ is a continuous function

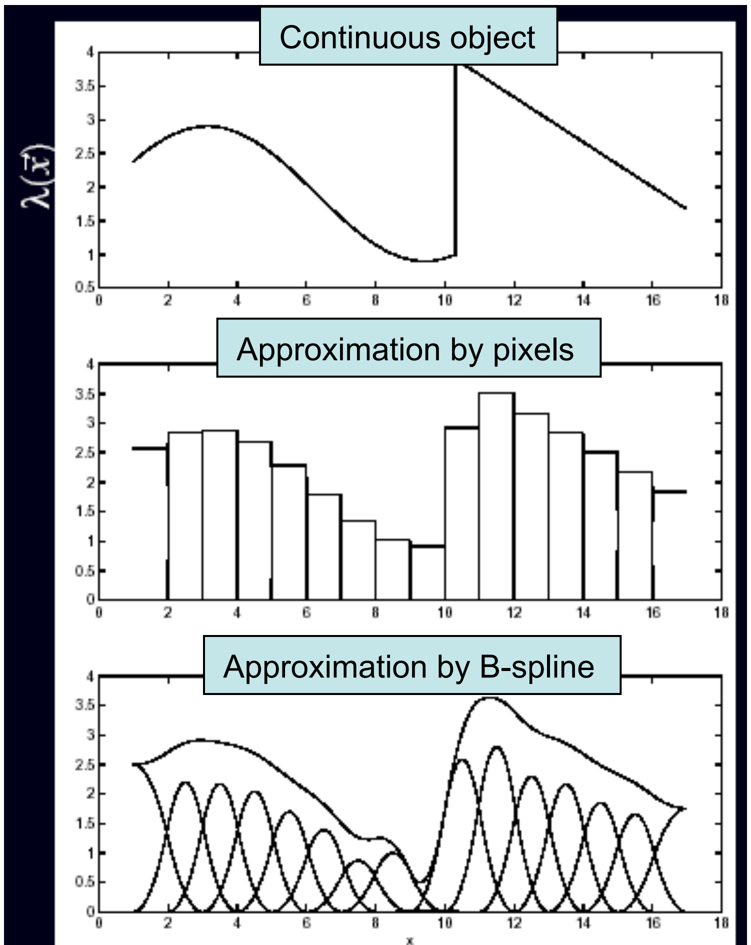
Replaced by $\lambda = (\lambda_1, \ldots, \lambda_{n_p})$

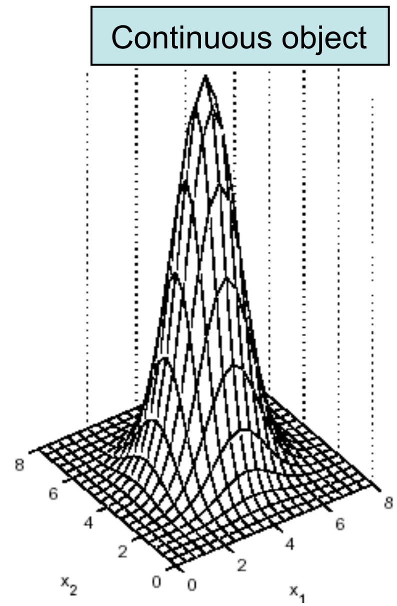
With $$\lambda(\vec{x}) \approx \sum_{j=1}^{n_p} \lambda_j \boxed{b_j(\vec{x})}$$ basis function

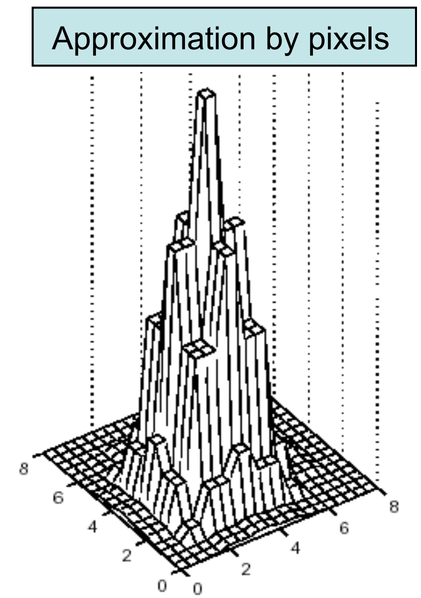
- Fourier series
- Wavelette
- Kaiser-Bessel
- B-splines

- Rectangular pixels
- Basis on the organes
- …

Examples


Continuous object
Approximation by pixels
Approximation by B-spline


Continuous object


Approximation by pixels

-1- Projection algorithm

$$g(s,\theta)=\int_{-\infty}^{+\infty} f(x,y)\,dt$$

$$g_i = a_{i1}f_1 + a_{i2}f_2 + \cdots + a_{im}f_m = \sum_{j=1}^{m} a_{ij}f_j$$

$$g(s,\theta)=\int s_i(\vec{x})\lambda(\vec{x})\,\mathrm{d}\vec{x} = \int s_i(\vec{x})\left[\sum_{j=1}^{n_p}\lambda_j b_j(\vec{x})\right]\mathrm{d}\vec{x}$$

$$= \sum_{j=1}^{n_p}\left[\int s_i(\vec{x})b_j(\vec{x})\,\mathrm{d}\vec{x}\right]\lambda_j = \sum_{j=1}^{n_p} a_{ij}\lambda_j$$

-1- Discrete Reconstruction

$$Y_i \sim Poisson\left\{\int s_i(\vec{x})\lambda(\vec{x})\,\mathrm{d}\vec{x} + r_i\right\}, \quad i = 1,\ldots,n_d$$

$$\Downarrow$$

$$Y_i \sim Poisson\left\{\sum_{j=1}^{n_p} a_{ij}\lambda_j + r_i\right\}, \quad i = 1,\ldots,n_d$$

-2- Physical model of the Système

$$a_{ij} = \int s_i(\vec{x}) b_j(\vec{x}) \mathrm{d}\vec{x}$$

- The geometry / solid angle of detection
- The collimation
- The scatter
- The attenuation
- Detector response
- Detection efficiency of the detector
- Positon range, acolinearity, etc…

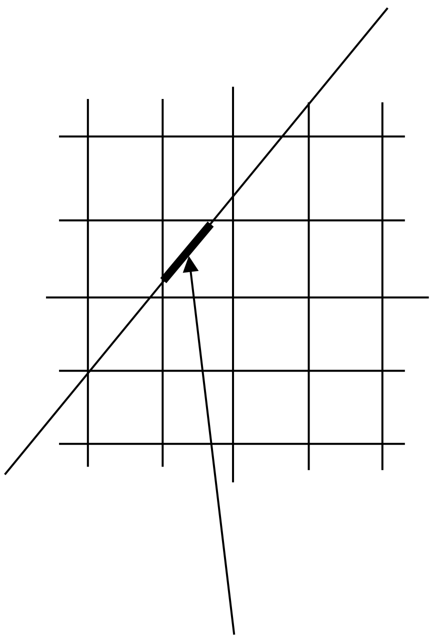
Improving the physical model enables:

Better quantification results
Better spatial resolution
…

Model measuring:

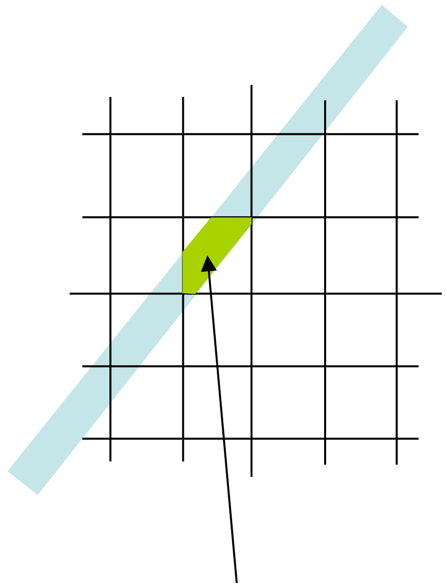
No approximation in the analytical calculation
Long time acquisition
Storage
…

-2- Integration line



a_{ij} = intersection length

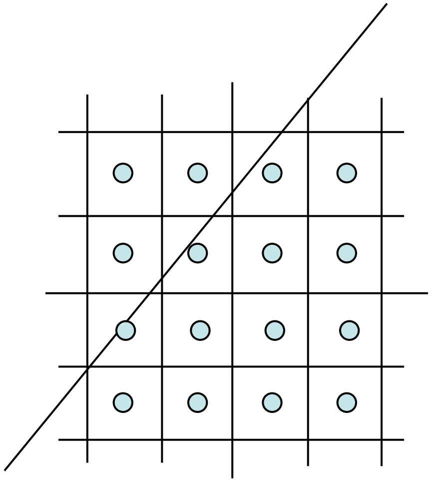
$$s_i(\vec{x}) = \delta\left(\vec{k}_i \cdot \vec{x} - \tau_i\right)$$



a_{ij} = area of intersection

$$s_i(\vec{x}) = \mathrm{rect}\left(\frac{\vec{k}_i \cdot \vec{x} - \tau_i}{w}\right)$$

Detector size



a_{ij} = interpolation

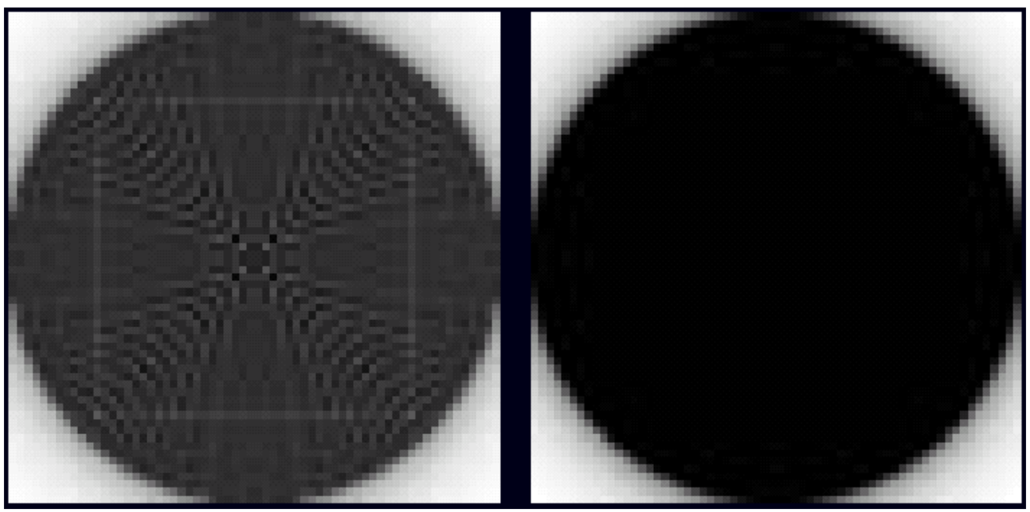
$$s_i(\vec{x}) = \mathrm{tri}\left(\frac{\vec{k}_i \cdot \vec{x} - \tau_i}{\Delta r}\right)$$

Distance between lines

-2- Examples…

Uniforme Sinogram

Back projection

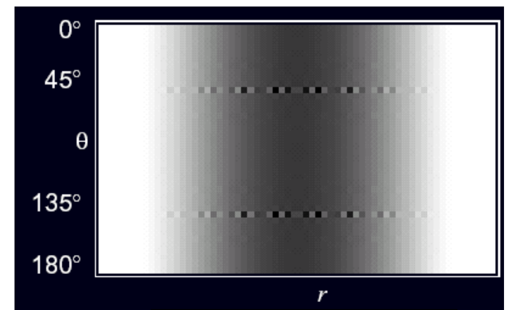


line

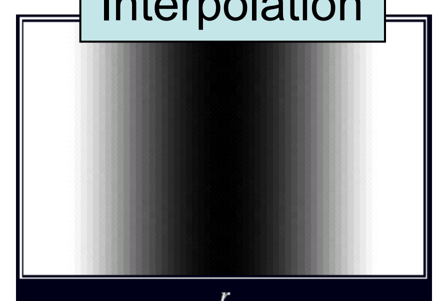
Area

Uniform object

Projection

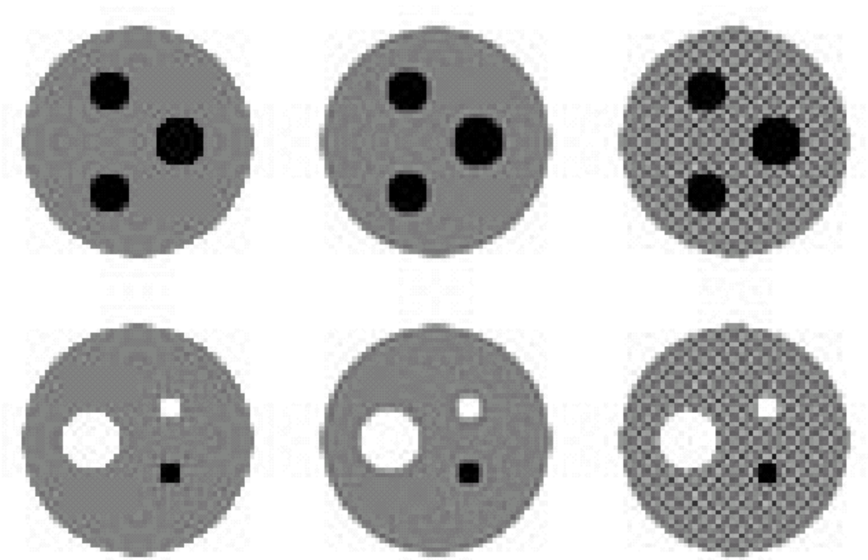


Interpolation



Area

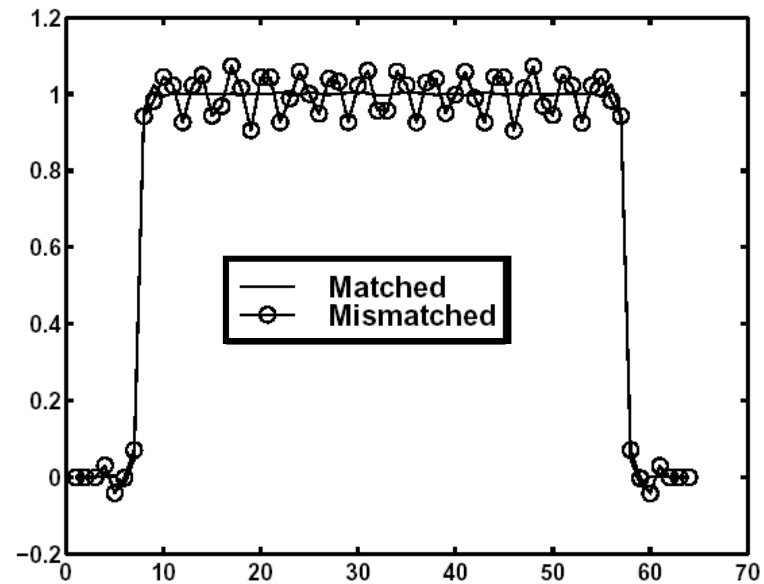
-2- Matched/Mismatched Projector/BackProjector operators



object

Matched operators

Mis-Matched operators



-3- Statistical mode of measurement

$$Y \approx A\lambda + r$$

Statistical model (→ A)

- Good model:
  - Variance reduction in image
  - Increasing computing time
  - Algorithm complexity
- Incorrect model
  - Statistics (dead time)
  - Model (transmission log)

-3- Choice of the Statistical model

- No model: $Y - r = A\lambda$ Resolve algebraically in order to find λ.

- Uniform gaussian noise: Least squares method, minimize $\left\| Y - A\lambda \right\|^2$

- Not uniform gaussian noise: Weighted least squares method, minimize

$$\left\| Y - A\lambda \right\|_w^2 = \sum_{i=1}^{n_d} w_i \left(y_i - [A\lambda]_i \right)^2, \quad [A\lambda]_i \cong \sum_{j=1}^{n_p} a_{ij}\lambda_j$$

- Poisson Model: $Y_i \sim Poisson\{[A\lambda]_i + r_i\}$

-4 & 5- Optimized criteria and used algorithm

Example:

Most used method: ML-EM

« *Maximum Likelihood – Expectation Maximisation* »

Optimization criteria — Algorithm

Two steps per iteration

- 1st Step: E (« Expectation »)
  - Calculate the likelihood expectation
- 2nd Step: M (« Maximisation »)
  - Maximise the expectation.

Mathematics derivation

Definition:

λ_j Mean number of desintegration in pixel j

a_{ij} Probability that a photon emitted in pixel j is detected in bin i

$a_{ij}\lambda_j$ Mean number of photons emitted from pixel j and detected in bin i

$\bar{g}_i = \sum_{j=1}^{m} a_{ij}\bar{f}_j$ Mean number of photons detected in bin i

Mathematics derivation

We have proved that g_i is a variable which statistics follows the Poisson law

→ The probability of detecting g_i photons is:

$$P(g_i) = \frac{e^{-\overline{g}_i}\,\overline{g}_i^{\,g_i}}{g_i!}$$

Example: Probability of detecting 5 with 3 as mean number of events

$$P(5) = \frac{e^{-3}3^5}{5!} \approx 0.101$$

Mathematics derivation

- Hypothesis on acquired data
  - The variables i are independents

$P(g|\lambda)$ Probability of observing the vector g when the emission vector is λ

$$=$$

$\prod_{i=1}^{n} P(g_i)$ Product of the individual probabilities

$$=$$

$L(\lambda)$ Likelihood function

Mathematics derivation

- Find the maximum value $\Rightarrow L(\lambda)$
- Calculate its derivative
- In order to maximize the likelihood, we use the following algorithm

$$l(\lambda) = \ln(L(\lambda))$$

$$l(\lambda) = \ln\left(\prod_{i=1}^{n} \frac{e^{-\bar{g}_i}\bar{g}_i^{g_i}}{g_i!}\right) = \sum_{i=1}^{n} \ln\left(\frac{e^{-\bar{g}_i}\bar{g}_i^{g_i}}{g_i!}\right) = \sum_{i=1}^{n}\left(-\bar{g}_i + g_i \ln(\bar{g}_i) - \ln(g_i!)\right)$$

Mathematics derivation

$$l(\lambda) = \sum_{i=1}^{n} \left(-\bar{g}_i + g_i \ln(\bar{g}_i) - \ln(g_i!) \right)$$

avec

$$\bar{g}_i = \sum_{j=1}^{m} a_{ij} \lambda_j$$

$$l(\lambda) = \sum_{i=1}^{n} \left(-\sum_{j=1}^{m} a_{ij}\lambda_j + g_i \ln\left(\sum_{j=1}^{m} a_{ij}\lambda_j \right) - \ln(g_i!) \right)$$

$\Rightarrow$ Probability to observe a projection from a mean image.

We want the image with the maximum probability of having g

In other words, the vector λ for which $l(\lambda)$ is maximum and considered as the best estimation of the solution.

Mathematics derivation

- It was proved that $l(\lambda)$ has a unique maximum.
- $\dfrac{\partial l(\lambda)}{\partial \lambda_j} = 0 \qquad \Rightarrow$ maximum

$$\frac{\partial l(\lambda)}{\partial \lambda_j} = -\sum_{i=1}^{n} a_{ij} + \sum_{i=1}^{n} \frac{g_i}{\sum_{j'=1}^{m} a_{ij'}\lambda_{j'}} a_{ij} = 0$$

$$l(\lambda) = \sum_{i=1}^{n} \left(-\sum_{j=1}^{m} a_{ij}\lambda_j + g_i \ln\left(\sum_{j=1}^{m} a_{ij}\lambda_j \right) - \ln(g_i!) \right)$$

Mathematics derivation

$$\frac{\partial l(\lambda)}{\partial \lambda_j} = -\sum_{i=1}^{n} a_{ij} + \sum_{i=1}^{n} \frac{g_i}{\sum_{j'=1}^{m} a_{ij'}\lambda_{j'}} a_{ij} = 0$$

$$\boxed{\lambda_j} \frac{\partial l(\bar{f})}{\partial \bar{f}_j} = -\boxed{\lambda_j} \sum_{i=1}^{n} a_{ij} + \boxed{\lambda_j} \sum_{i=1}^{n} \frac{g_i}{\sum_{j'=1}^{m} a_{ij'}\lambda_{j'}} a_{ij} = 0$$

$$\lambda_j = \frac{\lambda_j}{\sum_{i=1}^{n} a_{ij}} \sum_{i=1}^{n} \frac{g_i}{\sum_{j'=1}^{m} a_{ij'}\lambda_{j'}} a_{ij}$$

Iterative form

$$\lambda_j = \frac{\lambda_j}{\sum_{i=1}^{n} a_{ij}} \sum_{i=1}^{n} \frac{g_i}{\sum_{j'=1}^{m} a_{ij'} \lambda_{j'}} a_{ij}$$

$$\lambda_j^{(k+1)} = \frac{\lambda_j^{(k)}}{\sum_{i=1}^{n} a_{ij}} \sum_{i=1}^{n} \frac{g_i}{\sum_{j'=1}^{m} a_{ij'} \lambda_{j'}^{(k)}} a_{ij}$$

Description

Measured Projection

$$\lambda_j^{(k+1)} = \frac{\lambda_j^{(k)}}{\sum_{i=1}^{n} a_{ij}} \sum_{i=1}^{n} \frac{g_i}{\sum_{j'=1}^{m} a_{ij'} \lambda_{j'}^{(k)}} a_{ij}$$

Normalization factor

Estimated projection

$$\text{Image}^{(k+1)} = \text{Image}^{(k)} \text{ x back projection normalized with } \frac{\text{Measured projection}}{\text{Estimated projection}}$$

ML-EM Algorithm

- Multiplicative method
- Positive or null solution
  - Initial values at 0 remain 0
  - Positive initial value remain positive
- Conservation of the global activity in the image
- Slow convergence
- For a low number of iterations
  - Cold zones: excellent reconstruction
  - Hot zones: reconstruction < FBP
- For a high number of iterations
  - Cold zones: excellent reconstruction
  - Hot zones: Noisy images (bias near to 0)

Noise at Convergence

- At convergence,
  - « Perfect » reconstruction of the counts number in each pixel
- However,
  - No correlation between neighbouring pixels.
  - High Poisson noise level ⇒ Chessboard effect
- Corrections
  - Stop the iterations (need to define a stop criteria…)
  - Penalization function

Research domains

- Convergence acceleration
- Problem regularization
  - Penality function
  - Introduction of an A PRIORI knowledge

Convergence acceleration

- Algorithm OS-EM

$$\lambda_j^{(k+1)} = \frac{\lambda_j^{(k)}}{\sum_{i=1}^{n} a_{ij}} \sum_{i=1}^{n} \frac{g_i}{\sum_{j'=1}^{m} a_{ij'} \lambda_{j'}^{(k)}} a_{ij}$$

OS-EM = ML-EM applied on a subset S

$S=1 \Rightarrow$ ML-EM

Convergence has not been proved but seems to be similar to that of ML-EM.

$$\lambda_j^{(k+1)} = \frac{\lambda_j^{(k)}}{\sum_{i\in S} a_{ij}} \sum_{i\in S} \frac{g_i}{\sum_{j'=1}^{m} a_{ij'} \lambda_{j'}^{(k)}} a_{ij}$$

Acceleration factor ~ S

Adequate choice of subsets

Regularization

- Criteria:
  - Estimate projection ~ measured projection.
- Replaced by:
  - (a) Estimated projection ~ measured projection.
  - (b) Low noise obtained image.
- The introduction of an a priori knowledge on the image = regularization
  - Promote convergence!

Find λ to maximize (a) and (b)

Mathematics derivation

Bayes theorem:

Probability of observing the vector g when the emission vector is f

A priori knowledge on the image

$$P(\lambda|g) = \frac{P(g|\lambda)P(\lambda)}{P(g)}$$

A posteriori probability

A priori knowledge on the projections

Algorithm MAP: Maximum a Posteriori

Consider the logarithm:

$$P(\lambda|g) = \frac{P(g|\lambda)P(\lambda)}{P(g)} \Rightarrow \ln P(\lambda|g) = \ln P(g|\lambda) + \ln P(\lambda) - \ln P(g)$$

$$\underbrace{\ln P(\lambda|g)}_{\text{A posteriori probability}} = \underbrace{\ln P(g|\lambda)}_{\text{Likelihood}} + \underbrace{\ln P(\lambda)}_{\text{a priori}} - \underbrace{\ln P(g)}_{\text{Constant}}$$

a priori ⇓ « prior »

MAP = ML if no a priori information ⇒ ML = special case of MAP

MAP = ML penalized, the penality being the a priori knowledge.

A priori example

Gibbs a priori ⇒ local image smoothing

$$P(\lambda) = C\,\mathrm{e}^{-\beta U(\lambda)}$$

U : Energy function of λ
β : A priori weighting
C : Normalization constant

$$\ln P(\lambda|g) = \sum_{i=1}^{n}\left(-\sum_{j=1}^{m} a_{ij}\lambda_j + g_i \ln\left(\sum_{j=1}^{m} a_{ij}\lambda_j\right) - \ln(g_i!)\right) - \beta U(\lambda) + K$$

$K = \ln C - \ln P(g)$: Constant independent of λ

Maximize likelihood

- Derive the likelihood in order to maximize λ

$$\frac{\partial P(\lambda|g)}{\partial \lambda_j} = -\sum_{i=1}^{n} a_{ij} + \sum_{i=1}^{n} \frac{g_i}{\sum_{j'=1}^{m} a_{ij'}\lambda_{j'}} a_{ij} - \beta \frac{\partial}{\partial \lambda_j} U(\lambda_j) = 0$$

$$\lambda_j^{(k+1)} = \frac{\lambda_j^{(k)}}{\sum_{i=1}^{n} a_{ij} + \beta \frac{\partial}{\partial f_j} U\left(\lambda_j^{(k)}\right)} \sum_{i=1}^{n} \frac{g_i}{\sum_{j'=1}^{m} a_{ij'}\lambda_{j'}^{(k)}} a_{ij}$$

Example of function *U*

- A quadratic a priori:

$$\frac{\partial}{\partial \lambda_j^{(k)}} U\left(\lambda_j^{(k)}\right) = \sum_{b \in N_j} w_{jb}\left(\lambda_j^{(k)} - \lambda_b^{(k)}\right)$$

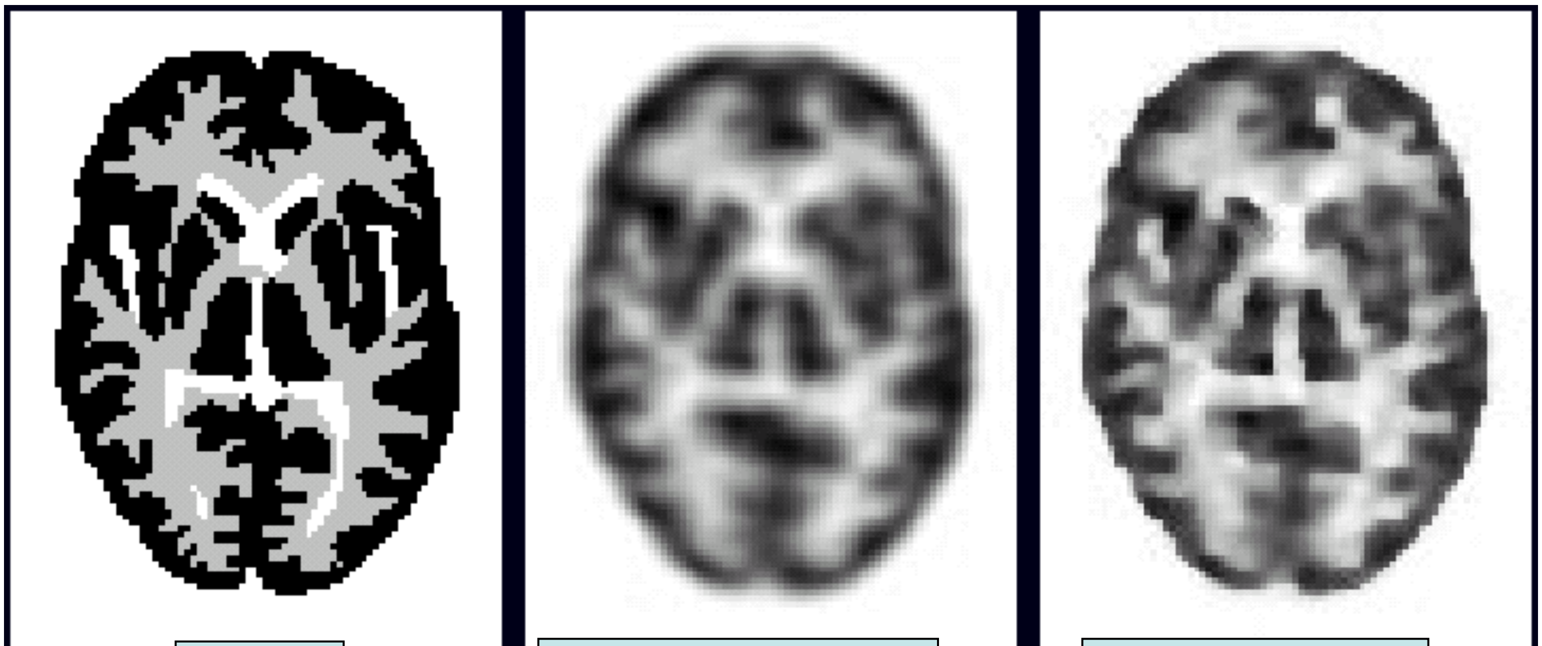
N_j : set of points neighbouring pixel j

Si $b \sim j \Rightarrow$ the term is zero $\Rightarrow \lambda^{(k+1)}$ same to ML-EM

Si $j > b \Rightarrow$ the term >0 $\Rightarrow \lambda^{(k+1)}$ < a ML-EM

Si $j < b \Rightarrow$ the term is <0 $\Rightarrow \lambda^{(k+1)}$ > a ML-EM

Examples



object | quadratic a priori | Hubert a priori

Some remarks

$$\lambda_j^{(k+1)} = \frac{\lambda_j^{(k)}}{\sum_{i=1}^{n} a_{ij} + \beta \frac{\partial}{\partial f_j} U\left(\lambda_j^{(k)}\right)} \sum_{i=1}^{n} \frac{g_i}{\sum_{j'=1}^{m} a_{ij'} \lambda_{j'}^{(k)}} a_{ij}$$

Possibility of negativity

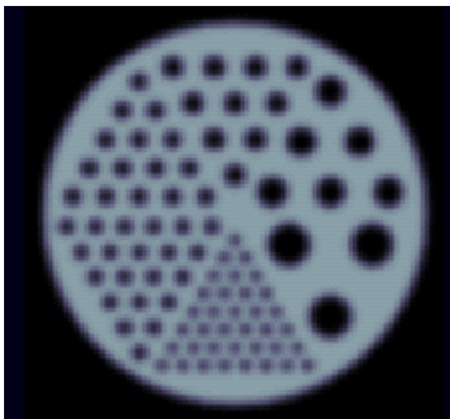
⇨ Keep a low value of β so that values remain positive

The a priori smoothes also the edges ⇨ Loss of resolution

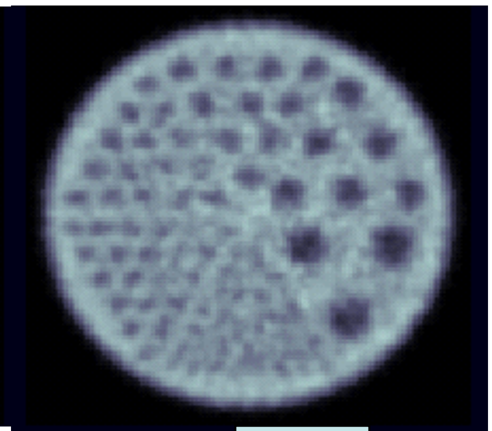
⇩

Modify the a priori
Introduce anatomical information

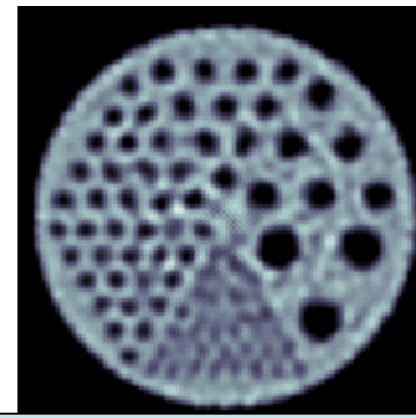
Illustrations



object

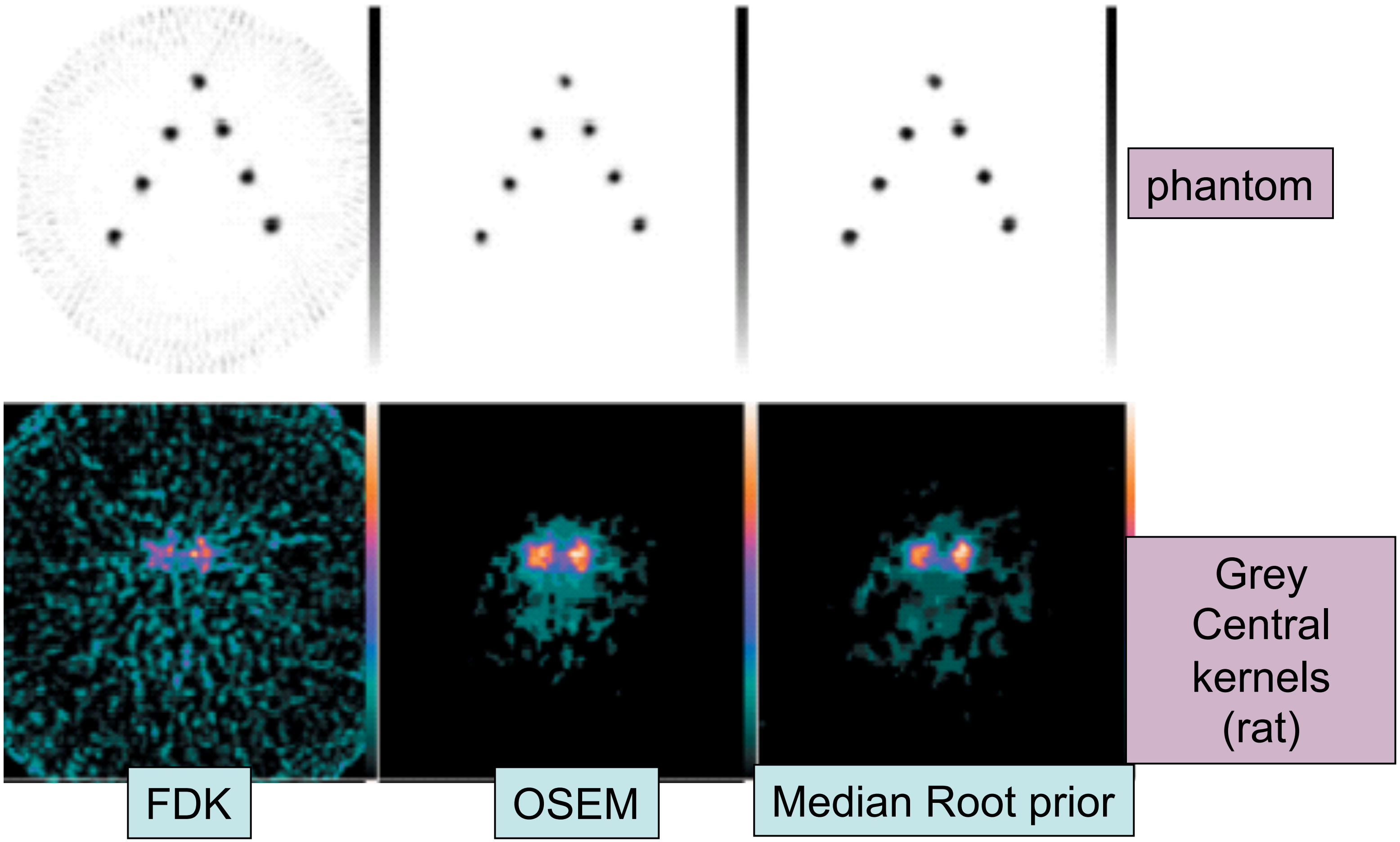


FBP



Iterative + regularisation

Illustrations



A. Sohlberg et al, Eur J Nucl Med (2004) 31:986-994

References

- F. Beekman, *Discrete Reconstruction Methods*, NSS-MIC 2000.
- J. A. Fessler, *Statistical Method for Image Reconstruction*, NSS-MIC 2001.
- P.P. Bruyant, *Analytic and Iterative Reconstruction Algorithms in SPECT*, JNM 2002.

Type of tomographic studies

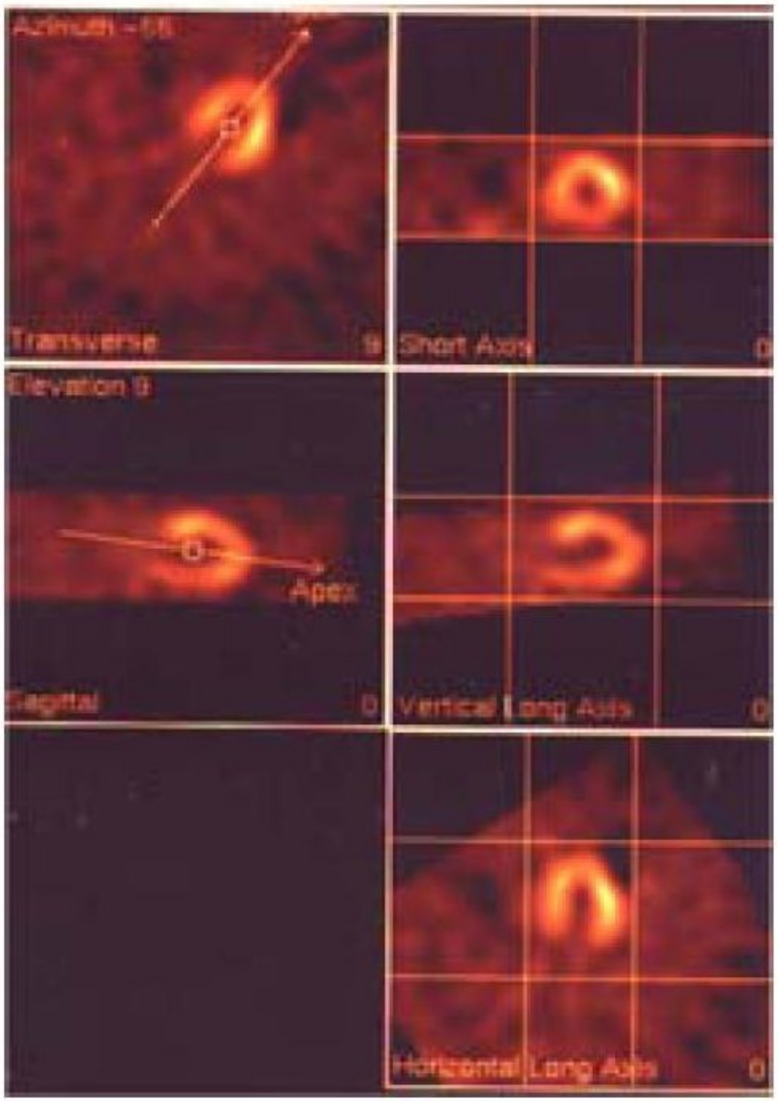
- Heart
  - Myocarde (thallium, MIBI…)
  - Ventricular cavities
  - Gated SPECT
- Brain
- Lungs
- Bone
- Others (peptides, antibodies…)

Slices plans in myocardiac imaging

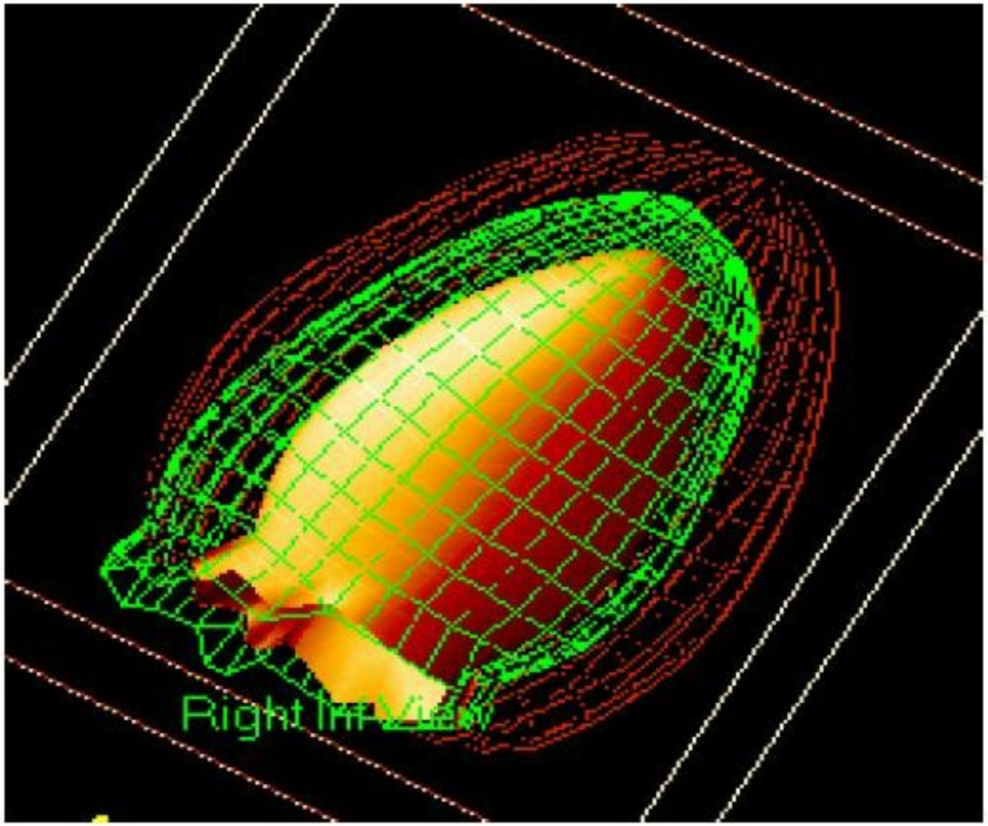
Small axis

Big axis

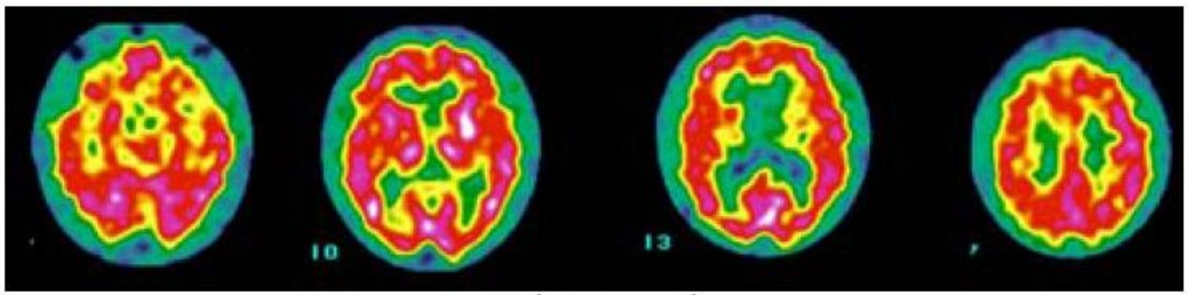
Horizontal



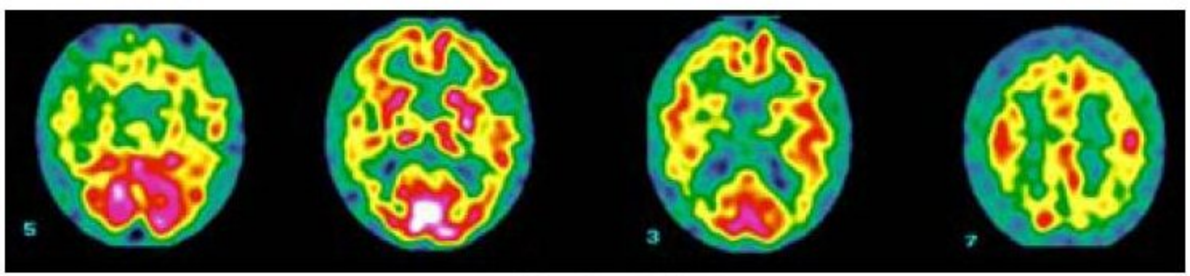
Representation of 3D contours



Cerebral tomography



Sujet normal



Maladie d'Alzheimer

Endocrine tumor in the liver