Appendix B: Exercises

B.1 Introduction to Digital Image Processing

1 Introduction to digital image processing

Q1 **Edge detection and edge enhancement**

(a) **Specify a differential operator (high-pass filter) to detect the horizontal edges in an image. Do the same for the vertical edge detection.**
A differential operator for detecting horizontal edges in an image $I(x, y)$ can be constructed by considering that an edge in an image indicates a local, strong difference (increase or decrease) in intensity. Horizontal edges are directed along the horizontal direction (i.e. the x axis) of the image, such that the local change in intensity is along the vertical direction (i.e. the y axis). Horizontal edges can thus be detected by evaluating the first derivative of the image along y (vice versa for vertical edges and the derivative along x). Due to the discrete nature of the image, a finite difference approximation for the first derivative is to be used, for instance the mean of the forward and the backward difference along y at (x, y). This can be written as a convolution with a one-dimensional filter W along y:

$$\begin{aligned}
\frac{\delta I(x,y)}{\delta y} &\approx \frac{[I(x,y) - I(x,y-1)] + [I(x,y+1) - I(x,y)]}{2} && (1)\\
&= \frac{I(x,y+1) - I(x,y-1)}{2} && (2)\\
&= \sum_{k=-1}^{k=1} w(k).I(x,y-k) && (3)\\
&= W \star I && (4)
\end{aligned}$$

with $W = 1/2.[\ 1 \quad 0 \quad -1\ ]$, indexed from -1 to +1, a 1-D filter that operates along the columns of $I(x, y)$ by convolution ($\star$). It is often more convenient to write this as a mathematical correlation ($*$) instead of a convolution:

$$\begin{aligned}
W \star I &= \sum_{k=-1}^{k=1} w(k).I(x,y-k) && (5)\\
&= \sum_{k=-1}^{k=1} w(-k).I(x,y+k) && (6)\\
&= \sum_{k=-1}^{k=1} w^*(k).I(x,y+k) && (7)\\
&= W^* * I && (8)
\end{aligned}$$

with $W^*(k) = W(-k)$ such that $W^* = 1/2.[\ -1 \quad 0 \quad 1\ ]$. Instead of a 1-D convolution, the resulting operator can be written as a 2-D filtering of the original image $I(x, y)$ with a 3×3 mask M_y:

$$\begin{aligned}
W * I &= \sum_{k=-1}^{k=1} w(k).I(x,y+k) && (9)\\
&= \sum_{k=-1}^{k=1} \sum_{l=-1}^{l=1} m(k,l).I(x+k,y+l) && (10)\\
&= M_y * I && (11)
\end{aligned}$$

with

$$M_y = [m(k,l)] = \begin{array}{|c|c|c|}\hline 0 & -1 & 0 \\ \hline 0 & 0 & 0 \\ \hline 0 & +1 & 0 \\ \hline \end{array} \tag{12}$$

and k and l indexing the columns and rows of M_y respectively, assuming that the y-axis is oriented downwards along the column direction of the image matrix, and ignoring the scale factor of 1/2. Horizontal edge responses are computed at every pixel location (x, y) by running the filter mask M_y over the image and computing the response at the central position in the mask as the linear combination $M * I$ of the original image intensities in I weighted (i.e. multiplied) by the filter weights in M. For edge detection, the sign of the contrast is of no relevance, such that horizontal edges can be detected as pixels with strong response for the absolute value of $|M_y * I|$.

Mutatis mutandis, vertical edges can be detected in the same way by switching x and y, i.e. by the 3×3 filter mask M_x which is the transpose of M_y:

$$M_x = \begin{array}{|c|c|c|}\hline 0 & 0 & 0 \\ \hline -1 & 0 & +1 \\ \hline 0 & 0 & 0 \\ \hline \end{array} \tag{13}$$

Note that these simple operators are very sensitive to noise. The Sobel edge filters try to alleviate this by averaging the edge response along y (x) by averaging over adjacent x (y):

$$\begin{aligned} \frac{\delta I(x,y)}{\delta y} &\approx \frac{1}{2}\left(\frac{(M_y * I)(x-1,y) + (M_y * I)(x,y)}{2}\right. && (14)\\ &+ \left.\frac{(M_y * I)(x,y) + (M_y * I)(x+1,y)}{2}\right) && (15)\\ &= \frac{(M_y * I)(x-1,y) + 2.(M_y * I)(x,y) + (M_y * I)(x+1,y)}{4} && (16)\\ &= \frac{1}{4}\left(\begin{array}{|c|c|c|}\hline -1 & 0 & 0 \\ \hline 0 & 0 & 0 \\ \hline +1 & 0 & 0 \\ \hline \end{array} + \begin{array}{|c|c|c|}\hline 0 & -2 & 0 \\ \hline 0 & 0 & 0 \\ \hline 0 & +2 & 0 \\ \hline \end{array} + \begin{array}{|c|c|c|}\hline 0 & 0 & -1 \\ \hline 0 & 0 & 0 \\ \hline 0 & 0 & +1 \\ \hline \end{array}\right) && (17)\\ &= \frac{1}{4}.\begin{array}{|c|c|c|}\hline -1 & -2 & -1 \\ \hline 0 & 0 & 0 \\ \hline +1 & +2 & +1 \\ \hline \end{array} && (18) \end{aligned}$$

such that the Sobel filters are given by

$$S_x = \begin{array}{|c|c|c|}\hline -1 & 0 & +1 \\ \hline -2 & 0 & +2 \\ \hline -1 & 0 & +1 \\ \hline \end{array} \quad S_y = \begin{array}{|c|c|c|}\hline -1 & -2 & -1 \\ \hline 0 & 0 & 0 \\ \hline +1 & +2 & +1 \\ \hline \end{array} \tag{19}$$

ignoring the scale factor of 1/4 (or in fact 1/8 to obtain a response of 1 for a linear edge with cross-section [−1 0 1], i.e. gradient magnitude 1).

In a similar way, 3x3 filters can be derived that approximate higher-order

derivatives of the image:

$$\begin{aligned}
\frac{\delta^2 I(x,y)}{\delta y^2} &\approx \frac{\delta I(x,y+1)}{\delta y} - \frac{\delta I(x,y)}{\delta y} && (20)\\
&= [I(x,y+1) - I(x,y)] - [I(x,y) - I(x,y-1)] && (21)\\
&= I(x,y-1) - 2.I(x,y) + I(x,y+1) && (22)\\
&= M_{y^2} \star I && (23)
\end{aligned}$$

with

$$M_{y^2} = \begin{array}{|c|c|c|}\hline 0 & +1 & 0 \\ \hline 0 & -2 & 0 \\ \hline 0 & +1 & 0 \\ \hline \end{array} \quad (24)$$

such that

$$M_{x^2} = \begin{array}{|c|c|c|}\hline 0 & 0 & 0 \\ \hline +1 & -2 & +1 \\ \hline 0 & 0 & 0 \\ \hline \end{array} \quad (25)$$

and the Laplacian $\nabla^2 I$ of I can be approximated as:

$$\nabla^2 I = \frac{\delta^2 I(x,y)}{\delta x^2} + \frac{\delta^2 I(x,y)}{\delta y^2} = M_L \star I \quad (26)$$

with

$$M_L = M_{x^2} + M_{y^2} = \begin{array}{|c|c|c|}\hline 0 & +1 & 0 \\ \hline +1 & -4 & +1 \\ \hline 0 & +1 & 0 \\ \hline \end{array} \quad (27)$$

Note that M_L has a strong positive response for a dark pixel on a bright background (i.e. a positive curvature of $I(x,y)$). It can therefore be more convenient to consider the inverse of M_L, which has a strong positive response for a bright pixel on a dark background:

$$M_L = \begin{array}{|c|c|c|}\hline 0 & -1 & 0 \\ \hline -1 & +4 & -1 \\ \hline 0 & -1 & 0 \\ \hline \end{array} \quad (28)$$

(b) **How can edges of arbitrary direction be detected using the above two operators?**

The output of both operators M_x and M_y (or S_x and S_y) that are sensitive to vertical and horizontal edges respectively can be combined by realising that they approximate the derivative of $I(x,y)$ along x and y respectively:

$$\begin{aligned}
\frac{\delta I(x,y)}{\delta x} &\approx M_x * I && (29)\\
\frac{\delta I(x,y)}{\delta y} &\approx M_y * I && (30)\\
& && (31)
\end{aligned}$$

such that the gradient magnitude $|\nabla I(x,y)|$ is given by

$$|\nabla I(x,y)| = \sqrt{\left(\frac{\delta I(x,y)}{\delta x}\right)^2 + \left(\frac{\delta I(x,y)}{\delta y}\right)^2} \approx \sqrt{(M_x * I)^2 + (M_y * I)^2} \quad (32)$$

which is independent of the orientation of the edge. Note that because of the non-linearity introduced when computing the norm $|\nabla I(x,y)|$ of the gradient vector $\nabla I(x,y)$, this operation is no longer linear and can not be written as a convolution itself.

The direction $\vec{e}$ of the edge (with $||\vec{e}|| = 1$) can be determined by considering that the gradient vector $\vec{g}$ is oriented perpendicular to the edge direction $\vec{e}$ ($\vec{g} \perp \vec{e}$), oriented at angles θ_g and θ_e respectively with respect to x (with $\theta_e - \theta_g = 90^o$):

$$\vec{g} = (\frac{\delta I}{\delta x}, \frac{\delta I}{\delta y}) \Rightarrow \vec{e} = \frac{1}{|\nabla I(x,y)|}.(\frac{\delta I}{\delta y}, -\frac{\delta I}{\delta x}) \tag{33}$$

$$\theta_g = \arctan(\frac{\delta I}{\delta y} / \frac{\delta I}{\delta x}) \tag{34}$$

$$\theta_e = \arctan(-\frac{\delta I}{\delta x} / \frac{\delta I}{\delta y}) \tag{35}$$

The directional derivative of $I(x,y)$ along any direction $\vec{n}$ (with $||\vec{n}|| = 1$) can be computed as follows:

$$\nabla_{\vec{n}} I = \frac{\delta I}{\delta \vec{n}} = \vec{n}.\nabla I, \tag{36}$$

for instance along the diagonal directions:

$$\vec{n} = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}) \Rightarrow \nabla_{\vec{n}} I = \frac{\sqrt{2}}{2}.(\frac{\delta I}{\delta x} + \frac{\delta I}{\delta y}) \tag{37}$$

$$\vec{n} = (\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}) \Rightarrow \nabla_{\vec{n}} I = \frac{\sqrt{2}}{2}.(\frac{\delta I}{\delta x} - \frac{\delta I}{\delta y}) \tag{38}$$

Let $L_x = \frac{\delta}{\delta x}$ and $L_y = \frac{\delta}{\delta y}$ denote the operators that yield the derivative along x and y respectively. Define the directions u and v as $\vec{u} = \vec{g}/||\vec{g}||$ and $\vec{v} = \vec{e}$ with $\vec{e}$ as defined above (i.e. $\vec{u} \perp \vec{v}$, with $\vec{u}$ orthogonal and $\vec{v}$ parallel to the edge). Operators for the derivatives along $\vec{u}$ and $\vec{v}$ can then be derived as follows:

$$L_u I = \nabla_{\vec{u}} I = \vec{u}.\nabla I = \frac{\vec{g}}{||\vec{g}||}.\vec{g} = ||\vec{g}|| \Rightarrow L_u = \sqrt{L_x^2 + L_y^2} \tag{39}$$

$$L_v I = \nabla_{\vec{v}} I = \vec{v}.\nabla I = \vec{e}.\vec{g} = 0 \Rightarrow L_v = 0 \tag{40}$$

The derivative along the gradient direction is maximal and equal to the gradient magnitude itself. The derivative along the edge (i.e. orthogonal to the gradient direction) is zero.

In a similar way, expressions for the second order derivatives L_{uu}, L_{uv} and L_{vv} can be derived in function of L_x,L_y,L_{xx},L_{xy} and L_{yy}.

(c) **How can these operators be exploited for edge enhancement?**

Edge enhancement can be achieved by computing the edge strength $|\nabla I|$ in every pixel and adding a scaled version of this image to the original image:

$$I_e = I + \alpha.|\nabla I| \tag{41}$$

for some value of α. An example is shown in Figure 1, where a value of $\alpha = 0.5$ and $\alpha = 1$ was selected.

Alternatively, also the Laplacian could be used for this purpose:

$$I_e = I + \alpha.\nabla^2 \tag{42}$$

for some value of α. The Laplacian (i.e. second order derivative) operator emphasises the high frequency content in the image, which is indicative for edges and line structures.

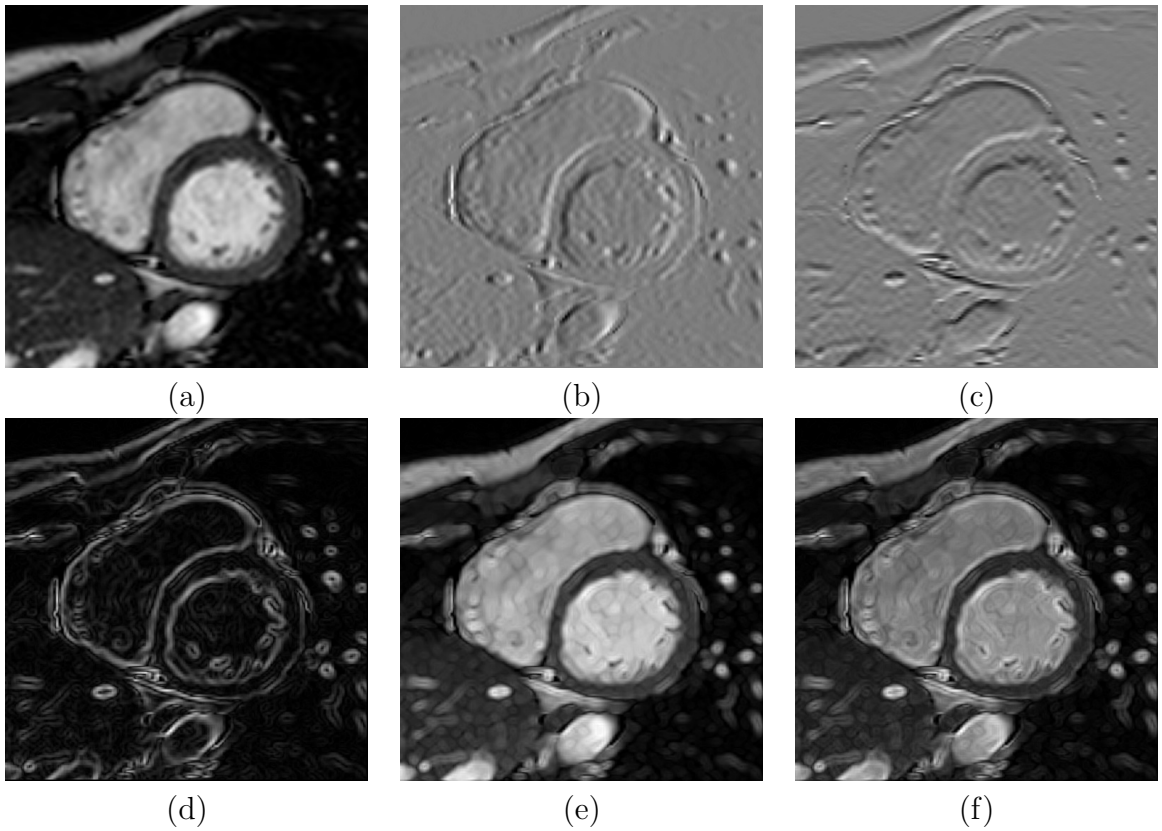


Figure 1: Example of edge enhancement. (a) Original cardiac MRI image. (b) Derivative along X (horizontal), showing vertical edges. (b) Derivative along Y (vertical), showing horizontal edges. (d) Gradient magnitude. (e) Edge enhanced image, $\alpha = 0.5$. (f) Edge enhanced image, $\alpha = 1$.

Q2 **What is the effect of a convolution with the following 3x3 masks?**

A)

$$M = \begin{array}{|c|c|c|}\hline 1 & 2 & 1 \\ \hline 2 & 4 & 2 \\ \hline 1 & 2 & 1 \\ \hline \end{array} \tag{43}$$

This mask represents a weighted local average of the image. Consider first the following filter:

$$A = \frac{1}{9}.\begin{array}{|c|c|c|}\hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array} \tag{44}$$

which computes the local average in a 3x3 region around the central pixel (x, y) of the image I whereby each neighbouring pixel in the 3x3 region is given an equal weight of $1/9$. M computes a weighted average in which pixels closer to the central pixel are given a higher weight. The effect hereof is a blurring or smoothing of the image, whereby noise is reduced and the main contrasting features are preserved.

In fact, M can be seen to approximate a 2D Gaussian kernel $G_\sigma(x, y)$ with variance σ^2:

$$G_\sigma(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{45}$$

such that

$$G_\sigma(0, 0) = \frac{1}{2\pi\sigma^2} \tag{46}$$

$$G_\sigma(\pm 1, 0) = G_\sigma(0, \pm 1) = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}} = \alpha . G_\sigma(0, 0) \tag{47}$$

$$G_\sigma(\pm 1, \pm 1) = \frac{1}{2\pi\sigma^2} e^{-\frac{2}{2\sigma^2}} = \alpha^2 . G_\sigma(0, 0) \tag{48}$$

with

$$\alpha = e^{-\frac{1}{2\sigma^2}} \tag{49}$$

For the mask M above, we have

$$M(0, 0) = 4 \tag{50}$$

$$M(\pm 1, 0) = M(0, \pm 1) = 2 = \frac{1}{2} . M(0, 0) \tag{51}$$

$$M(\pm 1, \pm 1) = 1 = \left(\frac{1}{2}\right)^2 . M(0, 0) \tag{52}$$

such that M and G_σ are equivalent (apart from a global scale factor) for $\alpha = 1/2$, which corresponds to a σ of

$$\alpha = e^{-\frac{1}{2\sigma^2}} = \frac{1}{2} \Rightarrow \sigma = \frac{1}{\sqrt{2 \ln 2}} \approx 0.85 \tag{53}$$

Note that

$$G_{\sigma=0.85}(0, 0) = \frac{1}{2\pi(0.85)^2} = 0.22 \tag{54}$$

while $M(0,0) = 4$, such that in fact all values in M should be scaled by $f = 0.22/4 = 0.055$ to be truly equivalent to $G_{\sigma=0.85}$. The scale factor $\frac{1}{2\pi\sigma^2}$ in $G_\sigma(x,y)$ originates from the fact that the integral of $G_\sigma(x,y)$ over the entire domain of all values of x and y should be 1:

$$\int_{x=-\infty}^{+\infty}\int_{y=-\infty}^{+\infty} G_\sigma(x,y)dxdy = 1 \tag{55}$$

Hence, it is more convenient to determine the scale factor f such that all values in M sum to 1 in order for M to be a discrete 3×3 approximation of G:

$$\sum_{k=-1}^{+1}\sum_{l=-1}^{+1} f.M(k,l) = f.16 = 1 \Rightarrow f = 1/16 = 0.0625 \tag{56}$$

and thus

$$M = \frac{1}{16}.\begin{array}{|c|c|c|}\hline 1 & 2 & 1 \\ \hline 2 & 4 & 2 \\ \hline 1 & 2 & 1 \\ \hline\end{array} \approx G_{0.85}(x,y) \tag{57}$$

Other Gaussian filters with different values of σ can be constructed in the same way, for instance for $\sigma = 0.5$:

$$\alpha = e^{-\frac{1}{2\sigma^2}} = e^{-2} \approx 1/7 \tag{58}$$

such that

$$M = \frac{1}{81}.\begin{array}{|c|c|c|}\hline 1 & 7 & 1 \\ \hline 7 & 49 & 7 \\ \hline 1 & 7 & 1 \\ \hline\end{array} \approx G_{0.5}(x,y) = \begin{array}{|c|c|c|}\hline 0.0113 & 0.0838 & 0.0113 \\ \hline 0.0838 & 0.6193 & 0.0838 \\ \hline 0.0113 & 0.0838 & 0.0113 \\ \hline\end{array} \tag{59}$$

whereby the values of $G_{0.5}(x,y)$ were scaled such that they sum to 1.

Note 1: In practical implementations, the filter elements are not restricted to integer values. It is more convenient to use floating point numbers instead.

Note 2: For larger values of the Gaussian kernel width σ, the 3x3 approximation no longer suffices and larger masks (5x5, 7x7...) have to be considered. As a rule of thumb, 2 or even 3 standard deviations σ are typically considered, such that the filter size w becomes: $w = 2.\text{round}(3\sigma) + 1$, i.e. 1 central element and a width of 3σ on either side. For instance, for $\sigma = 2$, this would yield $w = 13$, which results in a 13×13 convolutional filter involving $13^2 = 169$ multiplications per pixel, which can be computationally intensive for large images.

Note 3: The computational load can be reduced if the filter is separable, i.e. if the 2D filter can be written as 2 separate 1D convolutions that are applied along the rows and columns of the image respectively. This is the case for a Gaussian filter kernel. Hence, for $\sigma = 2$, the 13×13 filter operation can be reduced to 2 consecutive applications of a 1×13 filter, such that in fact only $2 \times 13 = 26$ multiplications per pixel are required.

Note 4: These 2D filters $M(k,l)$ can be extended to 3D filters $M(k,l,m)$ such that they can operate on 3D images as well.

B)

$$M = \begin{array}{|c|c|c|}\hline -1 & 0 & 1 \\ \hline -2 & 0 & 2 \\ \hline -1 & 0 & 1 \\ \hline \end{array} \tag{60}$$

Sobel edge detector filter for vertically oriented edges. This mask approximates the derivative of the image along x, while smoothing along y to reduce the effect of noise. See question Q1 above.

C)

$$M = \begin{array}{|r|r|r|}\hline 0 & -1 & 0 \\ \hline -1 & 4 & -1 \\ \hline 0 & -1 & 0 \\ \hline \end{array} \tag{61}$$

Laplacian filter. This mask approximates the Laplacian of the image, i.e. the sum of the second order derivatives along x and y. See question Q1 above.

D)

$$M = \begin{array}{|r|r|r|}\hline 0 & -1 & 0 \\ \hline -1 & 5 & 1 \\ \hline 0 & -1 & 0 \\ \hline \end{array} \tag{62}$$

This filter can be decomposed as

$$M = \begin{array}{|c|c|c|}\hline 0 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & 0 & 0 \\ \hline \end{array} + \begin{array}{|r|r|r|}\hline 0 & -1 & 0 \\ \hline -1 & 4 & 1 \\ \hline 0 & -1 & 0 \\ \hline \end{array} = M_\delta + M_L \tag{63}$$

with M_δ the identity filter (i.e. returning the original image itself) and M_L the Laplacian filter. Because convolution (or correlation) is a linear operation, we have that $M \star I = (M_\delta + M_L) \star I = M_\delta \star I + M_L \star I = I + M_L \star I$. The effect of this operation is edge enhancement. See question Q1.

E) Used in clinical practice?

No. These filters are too simplistic to be of much practical use on complex medical imaging data (see also Figure 1). They are sometimes used in low-level segmentation strategies, for instance to remove noise from the image by Gaussian smoothing or to compute the gradient magnitude to assess the location of edges in the image.

Q3 **Calculate the convolution of the following 3×3 convolution mask and the image in the demarcated rectangle:**

$$3 \times 3 \text{ mask} \qquad \begin{array}{|c|c|c|} \hline -1 & 0 & 1 \\ \hline -2 & 0 & 2 \\ \hline -1 & 0 & 1 \\ \hline \end{array} \tag{64}$$

$$\text{image} \qquad \begin{array}{|c|c|c|c|c|c|c|c|c|c|} \hline 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ \hline 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline \end{array} \tag{65}$$

The mask is the Sobel edge filter that approximates the derivative along the rows of the image. The image consists of 2 regions: a region of 0's on the left, and a region of 1's on the right. It suffices to compute the convolution only in a 3-pixel wide band along the boundary of both regions in which both values 0 and 1 are found. Elsewhere, all image values within a 3×3 region are either 0 or 1 and the convolution with the given filter evaluates to 0. The image is symmetric in vertical direction, such that only the top half needs to be computed. This gives the following result:

$$\text{filtered image} \qquad \begin{array}{|c|c|c|c|c|c|c|c|c|c|} \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 4 & 4 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 3 & 4 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 3 & 3 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 2 & 4 & 2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 3 & 3 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 3 & 4 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 4 & 4 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline \end{array} \tag{66}$$

where the values outside the demarcated rectangle were replaced by 0's to make the filtered image the same size as the original one.

Instead of a single 2D convolution, the filtering can be decomposed into 2 separate, consecutive 1D convolutions, by writing the filter as:

$$\begin{array}{|c|c|c|} \hline -1 & 0 & 1 \\ \hline -2 & 0 & 2 \\ \hline -1 & 0 & 1 \\ \hline \end{array} = \begin{array}{|c|} \hline 1 \\ \hline 2 \\ \hline 1 \\ \hline \end{array} \star \begin{array}{|c|c|c|} \hline -1 & 0 & 1 \\ \hline \end{array} = M_2 \star M_1 \tag{67}$$

M_1 differentiates along rows, while M_2 averages this response along columns. These 1D filters are more easily to compute, as each time only a single row or column is to be considered. Applying first M_1 and then M_2, we get:

$$M_1 \qquad \begin{array}{|c||c|c|c|c|c|c|c|c||c|}\hline 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline\hline 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline\hline 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline \end{array} \tag{68}$$

$$M_2 \star M_1 \qquad \begin{array}{|c||c|c|c|c|c|c|c|c||c|}\hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline\hline 0 & 0 & 0 & 4 & 4 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 3 & 4 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 3 & 3 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 2 & 4 & 2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 3 & 3 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 3 & 4 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 4 & 4 & 0 & 0 & 0 & 0 & 0 \\ \hline\hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline \end{array} \tag{69}$$

whereby M_1 and M_2 are applied everywhere were valid. In particular, M_1 need to be also applied in the first and last row, as these values impact the subsequent convolution with M_2.

Q4 **What is the effect of the following convolution operators on an image?**

A) The Laplacian of a Gaussian $\nabla^2 G_\sigma(x,y)$

The Laplacian $\nabla^2 G_\sigma(x,y)$ of the Gaussian kernel $G_\sigma(x,y)$ with width σ is given by:

$$G_\sigma(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{70}$$

$$\frac{\delta G_\sigma(x,y)}{\delta x} = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} . (-\frac{x}{\sigma^2}) = -\frac{x}{\sigma^2} . G_\sigma(x,y) \tag{71}$$

$$\frac{\delta^2 G_\sigma(x,y)}{\delta x^2} = -\frac{1}{\sigma^2} . G_\sigma(x,y) + \left(\frac{x}{\sigma^2}\right)^2 . G_\sigma(x,y) = \frac{x^2-\sigma^2}{\sigma^4} . G_\sigma(x,y) \tag{72}$$

$$\frac{\delta^2 G_\sigma(x,y)}{\delta y^2} = \frac{y^2-\sigma^2}{\sigma^4} . G_\sigma(x,y) \tag{73}$$

$$\nabla^2 G_\sigma(x,y) = \frac{\delta^2 G_\sigma(x,y)}{\delta x^2} + \frac{\delta^2 G_\sigma(x,y)}{\delta y^2} = \frac{x^2+y^2-2\sigma^2}{\sigma^4} . G_\sigma(x,y) \tag{74}$$

Note that the Laplacian of Gaussian integrates to zero:

$$\int_{x=-\infty}^{+\infty} \int_{y=-\infty}^{+\infty} \nabla^2 G_\sigma(x,y) dx dy = \frac{1}{\sigma^4} . (\sigma^2 + \sigma^2 - 2\sigma^2) = 0 \tag{75}$$

For $\sigma = 0.5$, a 3×3 filter $M(k,l)$ that approximates $\nabla^2 G_\sigma(x,y)$ is thus given by:

$$M_{\nabla^2 G}(k,l) = \nabla^2 G_{0.5}(k,l) = \frac{k^2+l^2-2.(0.5)^2}{(0.5)^4} . G_{0.5}(k,l) \tag{76}$$

which evaluates to

$$M_{\nabla^2 G} = \begin{array}{|c|c|c|} \hline 0.2798 & 0.6893 & 0.2798 \\ \hline 0.6893 & -5.0930 & 0.6893 \\ \hline 0.2798 & 0.6893 & 0.2798 \\ \hline \end{array} \tag{77}$$

which can be simplified by rounding the numbers and adjusting them slightly such that their sum is zero:

$$M_{\nabla^2 G} = \begin{array}{|c|c|c|} \hline 0.3 & 0.7 & 0.3 \\ \hline 0.7 & -4 & 0.7 \\ \hline 0.3 & 0.7 & 0.3 \\ \hline \end{array} \tag{78}$$

B) The difference of two Gaussians

The Laplacian of a Gaussian is often approximated as a difference of two Gaussians with different values of σ. This approximation follows by considering the derivative of the Gaussian kernel with respect to σ:

$$G_\sigma = \frac{1}{2\pi.\sigma^2} . e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{79}$$

$$\frac{\delta G_\sigma}{\delta \sigma} = \frac{-2}{2\pi.\sigma^3} . e^{-\frac{x^2+y^2}{2\sigma^2}} + \frac{1}{2\pi.\sigma^2} . e^{-\frac{x^2+y^2}{2\sigma^2}} . \frac{2(x^2+y^2)}{2\sigma^3} \tag{80}$$

$$= \frac{1}{2\pi.\sigma^2} . e^{-\frac{x^2+y^2}{2\sigma^2}} . \frac{x^2+y^2-2\sigma^2}{\sigma^3} \tag{81}$$

$$= \sigma . G_\sigma . \frac{x^2+y^2-2\sigma^2}{\sigma^4} \tag{82}$$

$$= \sigma . \nabla^2 G_\sigma \tag{83}$$

By considering two slightly different values σ_1 and σ_2 near σ, with $\sigma_2 > \sigma_1$, we get

$$\frac{\delta G_\sigma}{\delta \sigma} \approx \frac{G_{\sigma_2} - G_{\sigma_1}}{\sigma_2 - \sigma_1} \approx \sigma . \nabla^2 G_\sigma \tag{84}$$

or with $\sigma_1 = \sigma$ and $\sigma_2 = k.\sigma$ with $k > 1$ close to 1:

$$G_{\sigma_2} - G_{\sigma_1} \approx (k-1).\sigma^2 . \nabla^2 G_\sigma \tag{85}$$

or

$$\sigma^2 . \nabla^2 G_\sigma \sim G_{k.\sigma} - G_\sigma \tag{86}$$

if the value of k is fixed for all choices of σ.

This Difference of Gaussians (DoG) approximation of the Laplacian of Gaussian (LoG) is especially relevant as σ gets larger and a 3×3 filter no longer suffices. For instance for $\sigma = 2$ and considering 2 standard deviations σ on either side of the center, the filter extent would be $w = 1 + 2 * (2 * \sigma) = 9$. The LoG would then be a 9×9 filter:

$$M_{\nabla^2 G} = 10^{-3}. \begin{array}{|r|r|r|r|r|r|r|r|r|} \hline 1.09 & 1.86 & 2.45 & 2.67 & 2.69 & 2.67 & 2.45 & 1.86 & 1.09 \\ \hline 1.86 & 2.62 & 2.45 & 1.42 & 0.81 & 1.42 & 2.45 & 2.62 & 1.86 \\ \hline 2.45 & 2.45 & 0.00 & -3.99 & -6.03 & -3.99 & 0.00 & 2.45 & 2.45 \\ \hline 2.67 & 1.42 & -3.99 & -11.62 & -15.36 & -11.62 & -3.99 & 1.42 & 2.67 \\ \hline 2.69 & 0.81 & -6.03 & -15.36 & -19.89 & -15.36 & -6.03 & 0.81 & 2.69 \\ \hline 2.67 & 1.42 & -3.99 & -11.62 & -15.36 & -11.62 & -3.99 & 1.42 & 2.67 \\ \hline 2.45 & 2.45 & 0.00 & -3.99 & -6.03 & -3.99 & 0.00 & 2.45 & 2.45 \\ \hline 1.86 & 2.62 & 2.45 & 1.42 & 0.81 & 1.42 & 2.45 & 2.62 & 1.86 \\ \hline 1.09 & 1.86 & 2.45 & 2.67 & 2.69 & 2.67 & 2.45 & 1.86 & 1.09 \\ \hline \end{array} \tag{87}$$

Note that the convolution of this filter with the image I requires $9 \times 9 = 81$ multiplications per voxel. This LoG filter is not separable in x and y, in contrast to the Gaussian kernel itself, which can be written as the product of two functions that only depend on x or y:

$$G_\sigma = \frac{1}{2\pi.\sigma^2} . e^{-\frac{x^2+y^2}{2\sigma^2}} = \left(\frac{1}{\sqrt{2\pi}.\sigma} . e^{-\frac{x^2}{2\sigma^2}} \right) . \left(\frac{1}{\sqrt{2\pi}.\sigma} . e^{-\frac{y^2}{2\sigma^2}} \right) \tag{88}$$

The 2D Gaussian filter $M_G(k,l)$ of size $n \times n$ can thus be written as the product of two 1D kernels $M_{Gx}(k)$ and $M_{Gy}(l)$

$$M_G(k,l) = M_{Gy}(l) * M_{Gx}(k) \tag{89}$$

such that

$$M_G \star I = (M_{Gy} * M_{Gx}) \star I = M_{Gy} \star (M_{Gx} \star I) \tag{90}$$

with $M_{Gx}(k) = M_{Gy}(l)$, M_{Gx} a row vector of size $n \times 1$ and M_{Gy} a column vector of size $1 \times n$ that operate along the rows and columns of I respectively. The consecutive, separate application of M_{Gx} and M_{Gy} only requires $2n$ multiplications instead of n^2, or 18 versus 81 for a 9×9 Gaussian filter. The DoG approximation of the LoG thus only requires $2 \times 2n = 4n$ multiplications (if both Gaussians are represented as $n \times n$ filter masks) instead of n^2, or 36 versus 81 for $n = 9$.

C) The 3×3 convolution mask

$$\begin{array}{|c|c|c|}\hline 1 & 1 & 1 \\ \hline 1 & -8 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array} \tag{91}$$

This convolution mask can be decomposed as

$$\begin{array}{|c|c|c|}\hline 1 & 1 & 1 \\ \hline 1 & -8 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array} = 9 \times \left(\frac{1}{9}. \begin{array}{|c|c|c|}\hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array} - \begin{array}{|c|c|c|}\hline 0 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & 0 & 0 \\ \hline \end{array} \right) \tag{92}$$

The first mask computes a local weighted average of the image, i.e. a smoothed version of the original image, while the second mask returns the original image itself. Hence, this mask can be seen as an approximation of a DoG filter, using the averaging filter instead of a Gaussian. The DoG filter is itself an approximation of the Laplacian.

Hence, all 3 convolution operators A,B, and C above can be understood to have a similar effect.

Q5 **Unsharp masking is defined as**

$$(1+\alpha).I(x,y) - \alpha g \star I(x,y) \tag{93}$$

with $I(x,y)$ the image, g a Gaussian, and α a parameter.

The following convolution mask is an approximation of unsharp masking

$-1/8$	$-2/8$	$-1/8$
$-2/8$	?	$-2/8$
$-1/8$	$-2/8$	$-1/8$

(94)

Calculate the missing central value.

The operation above can be seen as the convolution of the image I with a mask M with

$$M = (1+\alpha).\delta - \alpha.g \tag{95}$$

with δ the identity (i.e. zero everywhere except for a central value of 1). The sum of all elements of $M(k,l)$ is

$$\begin{aligned}
\sum_{k,l} M(k,l) &= \sum_{k,l} ((1+\alpha).\delta(k,l) - \alpha.g(k,l)) && (96)\\
&= (1+\alpha)\sum_{k,l}\delta(k,l) - \alpha\sum_{k,l} g(k,l) && (97)\\
&= (1+\alpha).1 - \alpha.1 && (98)\\
&= 1 && (99)
\end{aligned}$$

as both δ and g sum to 1. Hence, for the mask above to approximate unsharp masking, it's elements should also sum to 1. With c the missing central value, the sum of all elements is:

$$c + 4.(-1/8) + 4.(-2/8) = 1 \tag{100}$$

and thus

$$c = 1 + 1/2 + 1 = 5/2 = 20/8 \tag{101}$$

B.2 Radiography

2 Radiography

Q1 X-rays

a) **What is the physical difference between X-rays, γ-rays, light and radio waves?**

These are all forms of electromagnetic (EM) radiation, but with a different frequency ν and wavelength λ. The energy of EM radiation is quantized in the form of photons, whose energy depends on the frequency ν:

$$E = h\nu = \frac{h.c}{\lambda} \tag{102}$$

with h Planck's constant and c the speed of light:

$$h = 6.62607004 \times 10^{-34}\, m^2 kg/s \; (= J.s) \tag{103}$$
$$c = 299792458\; m/s \approx 300000\; km/s \tag{104}$$
$$c = \lambda.\nu \tag{105}$$

The energy of photons is more conveniently expressed in electronvolt (eV). $1eV$ is the energy that a single electron with charge q gains by an electrical potential difference of $U = +1V$:

$$q = 1.60217662 \times 10^{-19} C \tag{106}$$
$$1\,eV = q.U = 1.60217662 \times 10^{-19} J \tag{107}$$
$$h = 4.13566766 \times 10^{-15} eV.s \tag{108}$$

Different types of EM radiation can be discriminated based on their wavelength λ, frequency ν or photon energy E, but all of these are equivalent through the above relationships. The spectrum of EM radiation is summarized in Figure 2.

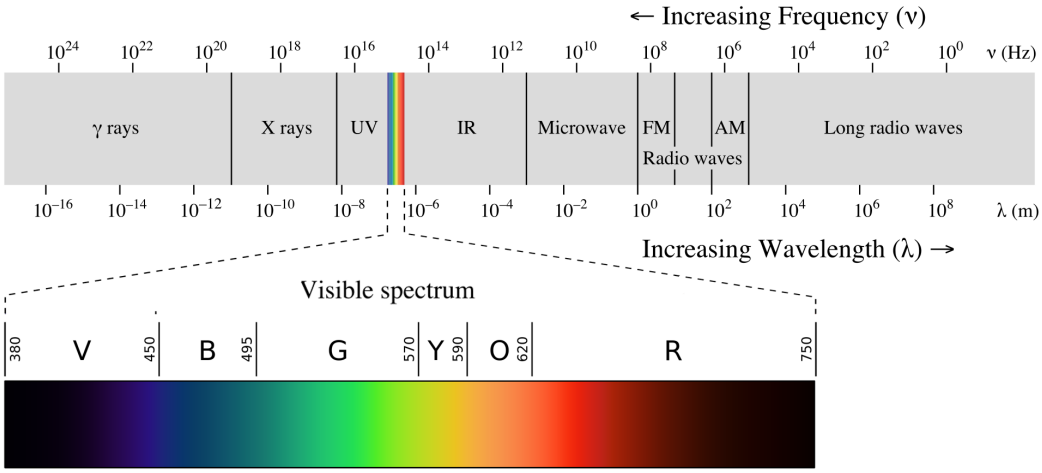


Figure 2: Spectrum of EM radiation.

In particular:

	ν (Hz)	λ	E
γ-rays	$> 3.10^{19}$	< 10 pm	> 125 keV
X-rays	$3 \times 10^{16} - 3 \times 10^{19}$	10 pm $-$ 10 nm	125 eV $-$ 125 keV
UV-light	$7.5 \times 10^{14} - 3 \times 10^{16}$	10 $-$ 400 nm	3 $-$ 125 eV
visible light	$4 \times 10^{14} - 8 \times 10^{14}$	380 $-$ 750 nm	1.6 $-$ 3.2 eV
radio waves	$< 3 \times 10^{8}$	> 1 m	$< 1.25 \times 10^{-6}$ eV

For use in imaging applications, the impact of EM radiation of a certain energy on biological tissues need to be considered. For a hydrogen atom, composed of an orbiting electron bound to a nucleus of one proton, an ionization energy of $2.18 \times 10^{-18} J = 13.6$ eV is required to force the electron from its lowest energy level entirely out of the atom. Hence, a threshold on the photon energy of around 13.6 eV is commonly used to discriminate between ionizing EM radiation (γ-rays, X-rays, highest UV energies) and non-ionizing EM radiation (radiowaves, visible light, lowest UV energies). The distinction between X-rays and γ-rays based on their energy is inadequate, as their energy ranges in fact overlap, but is made based on the specific context or their origin. For instance, EM radiation in this energy range induced by the interaction of an electron beam with a material will be termed X-rays, while photons of the same energy resulting from nuclear reactions are termed γ-rays, although both can not be distinguished from one another. Lower energy X-rays are sometimes termed soft X-rays, while higher energy X-rays are termed hard X-rays.

How do they interact with tissue (in the absence of a magnetic field)?

EM waves (like any other wave phenomenon) interact with matter by reflection, transmission, refraction, diffraction, scattering... In addition, EM radiation with photons above a certain energy (> 13.6 eV) can ionize matter through different effects:

- Rayleigh scattering: a photon is absorbed by an atom and a second photon is released by the atom with the same energy but a different direction (without ionizing the atom);
- Photo-electric absorption: a photon is absorbed by an atom and an electron of the atom is freed (thus ionizing the atom);
- Compton scattering: a photon is absorbed by an atom, an electron of the atom is freed (thus ionizing the atom) and a second photon with lower energy is released by the atom.

Depending on the energy range of the original photons, these effects will be more or less prominent. At lower energies, photo-electric absorption is more prominent, while at higher energies Compton scattering dominates.

Macroscopically, the interaction of a homogeneous material with EM radiation is described by its linear attenuation coefficient μ (in m^{-1} or cm^{-1}):

$$I_{\text{out}} = I_{\text{in}}.e^{-\mu d} \tag{109}$$

with I_{in} and I_{out} the incoming and outgoing intensity of the radiation respectively and d the material thickness (in m or cm). The attenuation coefficient μ depends on the tissue type (atomic number Z) and the energy E of the

incoming photons. An incident beam of EM radiation will typically not be mono-energetic, but rather consist of a spectrum $\sigma(E)$ of different photon energies, such that

$$I_{\text{out}} = \int_0^\infty \sigma(E).e^{-\mu(E)d} dE \tag{110}$$

For a non-homogeneous material, the contributions of each material to the overall attenuation need to be added or integrated:

$$I_{\text{out}} = \int_0^\infty \sigma(E).e^{-\int_{x_{\text{in}}}^{x_{\text{out}}} \mu(E,x)dx} dE \tag{111}$$

b) **Draw the X-ray tube spectrum, i.e., the intensity distribution of X-rays as a function of the frequency of emitted X-ray photons (1) at the exit of the X-ray tube before any filtering takes place, and (2) after the filter but before the X-rays have reached the patient.**

(1) The energy of the X-ray spectrum emitted by an X-ray tube is bounded above by $E_{\text{max}} = q.U$ with U the tube voltage (in kV) and q the electron charge. For instance, for $U = 140kV$, $E_{\text{max}} = 140keV$. The X-ray photons are emitted by the anode of the X-ray tube either as Bremsstrahlung (originating from the kinetic energy of the electrons that are "braked" by the anode material) or as characteristic radiation emitted by the anode material itself (originating from outer-shell electrons filling in the vacant positions in lower electron shells after ejection of inner-shell electrons of the anode atoms by the incident electron beam). The intensity of the Bremsstrahlung increases (from 0 to I_{max}) more or less linear with decreasing photon energy (from E_{max} to 0), while the characteristic radiation only occurs at specific energy levels depending on the anode material, resulting in peaks in the spectrum at these particular energy values.

(2) The X-rays leave the X-ray tube through a small window, typically made of a light metal such as Beryllium that is transparant to X-rays, but absorbs lower energy photons. Additional filtering of the X-ray beam is typically applied using thin Aluminum or Copper plates to remove low energy X-ray photons, as these are mainly just absorbed by the patient and thus contribute to the dose, but hardly to the image contrast. The filtering shifts the overall energy spectrum of the beam to higher energies, which is now as "beam hardening". Beam hardening occurs also during the passing of the beam through tissue, as a result of the fact that the attenuation coefficient $\mu(E)$ is typically larger for low energy photons than for high energy photons.

c) **How does the tube voltage influence the wavelength of the X-rays?**

The energy and the wavelength of a photon are related as $E = h.c/\lambda$, or $\lambda = h.c/E$. The maximal energy of X-ray photons is given by $E_{\text{max}} = q.U$ with U the tube voltage. Hence, $\lambda \geq \lambda_{\text{min}} = (h.c)/(q.U)$. Increasing the tube voltage thus decreases the minimal wavelength of the X-ray photons in the beam. However, changing the tube voltage has no impact on the energy and wavelength of the characteristic radiation emitted by the anode material, which always occurs at the same energy and wavelength. If the maximal energy E_{max}

is below the energy of a certain characteristic line of the anode material, this characteristic line will not occur in the energy spectrum of the emitted X-ray beam.

d) **Draw the linear attenuation coefficient (for an arbitrary tissue type) as a function of the energy.**

Q2 **What is the effect of the kV and mAs of an X-ray tube on**

a) **the patient dose?** The patient dose is defined as the total energy of the absorbed X-ray photons within a unit of mass (J/kg or Gray) . This depends on the beam intensity, i.e. on the number of incident photons for each energy level, and on the energy of the absorbed photons, i.e. on the energy spectrum of the beam.

The mAs of an X-ray tube is the product of the tube current (in A) and the exposure time (in ms) and a direct measure of the beam intensity or quantity, i.e. the number of photons per energy level, not their energy distribution itself. Increasing the mAs while keeping the kV fixed, increases the number of photons at each energy in the same way and proportionally to the mAs. The patient dose thus varies linearly with the mAs.

The kV, i.e. the voltage applied over the X-ray tube, controls the energy of the incident electrons and hence the energy spectrum of the resulting X-ray beam, i.e. the beam quality. However, the kV also impacts the beam intensity, as more photons are generated when the kinetic energy of the electrons hitting the anode is higher. Increasing the kV of the X-ray tube shifts the energy spectrum of the beam towards higher energies, and at the same time increases the beam intensity. The overall effect is that the incident dose increases about quadratically with the kV. The impact thereof on the patient dose depends on the attenuation coefficients $\mu(E)$ of the tissues, which are energy dependent. The attenuation coefficient μ tends to decrease for increasing photon energy in a non-linear way (lower energy photons are more likely to be absorbed by the tissues than higher energy photons). Hence, if more higher energy photons are present in the beam and absorbed by the tissues, a higher dose is deposited in the patient. On the other hand, higher energy photons are more likely to be transmitted than lower energy photons. Overall, the patient dose will increase somewhat less than quadratically with the kV setting.

Hence, low-dose imaging, especially low-dose CT, can be achieved by decreasing the mAs, but more so by decreasing kV. However, this has its limits, as image quality is of course also affected.

b) **the image quality?**

Image quality encompasses different aspects, including resolution, contrast and noise (or CNR: contrast-to-noise ratio), and artefacts.

Resolution in X-ray imaging is influenced by the focal spot size of the X-ray source and by the detector, and is not expected to change in function of the mAs or kV.

The contrast between two tissues depends on their difference in attenuation μ. Due to the shape of the curves $\mu(E)$ in function of energy for different materials, contrast tends to be higher at lower energies, thus at lower kV. The contrast itself is not influenced by the mAs.

The noise in X-ray based imaging is caused by electronic noise induced by the detector (which is independent of kV and mAs), and by fluctuations in the intensity of the X-ray beam. The generation of X-ray photons is a random process that can be modeled by a Poisson distribution. The beam intensity I is the expected value (i.e. the parameter λ) of the distribution. The standard

deviation of the Poisson distribution, which is a measure of the noise in the incident X-ray beam, is given by $\sqrt{\lambda} = \sqrt{I}$. Hence, the signal-to-noise ratio (SNR) equals $I/\sqrt{I} = \sqrt{I}$. The SNR thus improves proportionally to the square root of the beam intensity, i.e. the number of photons. The SNR increases proportionally to the square root of the mAs, and also the CNR (same contrast, improved SNR). The SNR also increases with increasing kV and more or less proportionally to kV, as the intensity varies more or less quadratic with kV. But the effect on the CNR is less or even negative due to the reduced intrinsic contrast between different materials as their μ decreases and becomes more similar for larger energies. At (too) high kV, many high-energy photons would simply be transmitted through the tissue and not contribute to the contrast.

Artefacts in 2D radiography is mainly scatter (apart from heel effect and detector defects). Scatter phenomena occur more at higher energies. To assess the impact thereof on image quality, the amount of scatter has to be compared to the primary beam intensity. The scatter will have a more pronounced effect in the image when the incident beam intensity is smaller. Hence, scatter artefacts will have a larger impact on image quality at lower mAs and lower kV.

Q3 **A digital radiograph is acquired but the CNR in the region of interest is insufficient. Calculate the relative CNR and relative dose of the following options:**

a) **double the mAs;**

The dose is proportional to the mAs, the CNR is proportional to the square root of the mAs (see Q2). Doubling the mAs doubles the dose and increases the CNR by a factor $\sqrt{2}$.

b) **take two radiographs and calculate the average image;**

By taking two radiographs with the same mAs and kV, the dose is doubled. The contrast and noise in both images are expected to be the same. Taking the average of both images, does not affect the contrast, but reduces the noise. If the SNR is sufficiently large, the noise can be modelled as Gaussian with standard deviation σ. The noise of the averaged image then has standard deviation $1/2 * \sqrt{\sigma^2 + \sigma^2} = \sigma/\sqrt{2}$. The CNR thus increases by a factor of $\sqrt{2}$.

Taking two radiographs and averaging has thus the same effect on image quality as acquiring a single radiograph with a double exposure time (mAs x 2).

c) **apply a window/level image operation;**

This is a gray level transformation that is applied to the original image intensities to improve the contrast in a certain intensity range. As this is a post-processing step, such a gray level transformation does of course not influence the dose. Because the gray level transformation applies similarly to the actual contrast as to the noise, the CNR is not affected.

d) **change the kV.**

The dose depends more or less quadratically on the kV. Increasing the kV (without changing the mAs) strongly increases the dose. The contrast tends to decrease with increasing kV (see Q2), but the SNR improves due to the larger beam intensity at higher kV. The effect on the CNR is difficult to predict a priori. At lower kV, scatter artefacts become more pronounced, whch also reduce the CNR.

Q4 **A radiograph of a structure consisting of bone and soft tissue (see Figure) is acquired by a screen-film detector. The exposure time is 1 ms. The radiographic film has a sensitometric curve $D = 2\log E$. The film-screen system has an absorption efficiency of 25%. Assume the the X-rays are monochromatic and the linear attenuation coefficients of bone, soft tissue and air are respectively 0.50 cm^{-1}, 0.20 cm^{-1}, 0.00 cm^{-1}.**

a) **Calculate the optical density D in positions A through E of the image.** Define I_0 as the incident X-ray beam intensity, $\eta = 0.25$ as the detector efficiency and $\Delta t = 1$ms as the exposure time. We then have:

$$D = 2\log E \tag{112}$$
$$E = \eta I \Delta t \tag{113}$$
$$I = I_0.e^{-\sum_{i=1}^{5}\mu_i.d_i} \tag{114}$$

Hence:

$$D = 2\log(\eta I_0 \Delta t) - 2\sum_{i=1}^{5}\mu_i.d_i \tag{115}$$

The total attenuation at the sites A to E is:

$$A, E: \quad \sum_i \mu_i.d_i = 5 \times 0.0 = 0 \tag{116}$$
$$B, D: \quad \sum_i \mu_i.d_i = 3 \times 0.5 = 1.5 \tag{117}$$
$$C: \quad \sum_i \mu_i.d_i = 2 \times 0.5 + 1 \times 0.2 = 1.2 \tag{118}$$
$$E = \eta I \delta t \tag{119}$$

Hence:

$$D_A = D_E = 2\log(\eta I_0 \Delta_t) = D_0 \tag{120}$$
$$D_0 = 2\log(0.25 \times I_0 \times 0.001) = 2\log(\frac{I_0}{4000}) \tag{121}$$
$$D_B = D_D = D_0 - 2 \times 1.5 = D_0 - 3 \tag{122}$$
$$D_C = D_0 - 2 \times 1.2 = D_0 - 2.4 \tag{123}$$

b) **Calculate the contrast, i.e., the difference in density, between positions B and C. How can this contrast be improved?**

$$|D_B - D_C| = |(D_0 - 3) - (D_0 - 2.4)| = 0.6 \tag{124}$$

Note that the contrast is independent of D_0, i.e. of I_0 or Δt, thus the mAs. The contrast between B and C only depends on the difference in μ between bone and soft tissue. This difference is energy dependent, due to the non-linear dependence of μ on the photon energy for different materials. The contrast can be improved by changing (lowering) the kV. Note that the noise depends on the mAs. Increasing the mAs (e.g. increasing the exposure time) will not affect the contrast itself, but improves the CNR, such that small contrasts become better appreciable.

Q5 **In mammography the breasts are compressed with a paddle. Explain why.**

B.3 X-ray Computed Tomography

3 X-Ray Computed Tomography

Q1 **Linear absorption coefficient.**

a) **Although the linear absorption coefficient μ depends on the energy, this dependence is not taken into account in filtered back-projection. Explain.**

b) **What is the effect of this approximation on the image quality?**

Q2 **Given the following image, consisting of two small bars, and projection $p(r,\theta)$ at angle θ. An enlargement of this projection at angle θ looks as follows and can mathematically be written as:**

$$p(r) = \Pi\left(\frac{r}{3\Delta l}\right) - \Pi\left(\frac{r}{\Delta l}\right) \tag{125}$$

a) **Calculate the Fourier transform of this projection.**
Starting from first principles (although the use of tables of Fourier transforms is also OK):
Define the rectangle function $\Pi(r)$ as:

$$\Pi(r) = 1 \text{ for } |r| < 1/2, \;\; 0 \text{ otherwise} \tag{126}$$

Its Fourier transform is given by

$$\begin{aligned}
\mathcal{F}\{\Pi(r)\} &= \int_{-\infty}^{\infty} \Pi(r) e^{-j2\pi kr} dr && (127)\\
&= \int_{-1/2}^{+1/2} \cos(2\pi kr) - j.\sin(2\pi kr) dr && (128)\\
&= \frac{1}{2\pi k} \sin(2\pi kr)|_{-\frac{1}{2}}^{+\frac{1}{2}} && (129)\\
&= \frac{\sin(\pi k)}{\pi k} && (130)\\
&= \text{sinc}(\pi k) && (131)
\end{aligned}$$

If the Fourier transform of $f(r)$ is $F(k)$, then the Fourier transform of $f(ar)$ for some constant $a \neq 0$ is:

$$\begin{aligned}
\mathcal{F}\{f(ar)\} &= \int_{-\infty}^{\infty} f(ar) e^{-j2\pi kr} dr && (132)\\
&= \int_{-\infty}^{\infty} f(ar) e^{-j2\pi \frac{k}{a}(ar)} \frac{1}{a} d(ar) && (133)\\
&= \int_{-\infty}^{\infty} f(r') e^{-j2\pi \frac{k}{a} r'} \frac{1}{a} dr' && (134)\\
&= \frac{1}{a}.F(k/a) && (135)
\end{aligned}$$

Hence:

$$\begin{aligned}
\mathcal{F}\{\Pi(\frac{r}{\Delta l})\} &= \Delta l.\text{sinc}(\pi k \Delta l) && (136)\\
\mathcal{F}\{\Pi(\frac{r}{3\Delta l})\} &= 3\Delta l.\text{sinc}(3\pi k \Delta l) && (137)
\end{aligned}$$

such that

$$\mathcal{F}\{p(r)\} = P(k) = 3\Delta l.\text{sinc}(3\pi k \Delta l) - \Delta l.\text{sinc}(\pi k \Delta l) \tag{138}$$

b) **The projection is now measured with an X-ray beam with width $\Delta s = 3\Delta l$. Assume a rectangular slice sensitivity profile (SSP).**

- **Calculate the resulting Fourier transform of this measured projection.**
  The measurement of the projection by an X-ray beam with a rectangular profile of width Δs corresponds to a convolution in the spatial domain with a rectangular function $\Pi(r/\Delta s)$. In the frequency domain, this corresponds to a multiplication of the actual projection with the Fourier transform $\mathcal{F}\{\Pi(r/\Delta s)\}$ of the beam profile. Denote the rectangular beam profile by the function $b(r)$:

$$b(r) = \Pi(\frac{r}{\Delta s}) \quad (139)$$
$$\mathcal{F}\{b(r)\} = \Delta s.\text{sinc}(\pi k \Delta s) \quad (140)$$

  and with $\Delta s = 3\Delta l$:

$$\mathcal{F}\{b(r)\} = B(k) = 3\Delta l.\text{sinc}(3\pi k \Delta l) \quad (141)$$

  The Fourier transform of the measured projection is thus:

$$B(k).P(k) = 3\Delta l.\text{sinc}(3\pi k \Delta l). \quad (142)$$
$$(3\Delta l.\text{sinc}(3\pi k \Delta l) - \Delta l.\text{sinc}(\pi k \Delta l)) \quad (143)$$
$$= 9\Delta l^2.\text{sinc}^2(3\pi k \Delta l) - \quad (144)$$
$$3\Delta l^2.\text{sinc}(\pi k \Delta l).\text{sinc}(3\pi k \Delta l) \quad (145)$$

- **Draw the measured projection (as a function of r).**
  The measured projection is the convolution of the actual projection with the rectangular beam profile of width $3\Delta l$. Convolution involves flipping, shifting, multiplication and integration of two functions. The function $p(r)$ consists of two rectangles of width Δl and the function $b(r)$ consists of a single rectangle of width $3\Delta l$. The convolution of a rectangle of width Δl with a rectangle of width $3\Delta l$ is a trapezoid of width $4\Delta l$ (with a central constant part of width $2\Delta l$) and height Δl. The final result is the sum of two such trapezoids, centered at $-\Delta l$ and $+\Delta l$, and has thus an overall width of $6\Delta l$, with values at $r = n\Delta l$ with n from -3 to +3 of 0, Δl, Δl, $2\Delta l$, Δl, Δl, and 0 respectively.
- **What is the maximal size of Δs you would recommend?**
  From the previous drawing, it is clear that the broad beam width $\Delta s = 3\Delta l$ obscures the details of width Δl, as the two individual bars in the actual profile can no longer be discriminated. Each of the two bars of width Δl, when convolved with a rectangular beam profile of width Δs, will result in a trapezoid of width $\Delta l + \Delta s$, one centered at $-\Delta l$ and the other at $+\Delta l$. For these trapezoids not to interfere with one another such that both bars remain discriminable in the measured projection, the total width of each trapezoid should not be larger than $2\Delta l$, or $\Delta s \leq \Delta l$.
- **What is the maximal sampling distance Δr you would recommend?**
  The convolution with a rectangular beam profile of width Δs (whose Fourier transform is a sinc-function with zero at frequency $k = 1/\Delta s$) limits the main frequency lobe of the measured profile to a maximal frequency of $k_{\max} = 1/\Delta s$ (as convolution corresponds to multiplication in

the frequency domain). In order to be able to reconstruct all frequency components from the sampled values, the sampling rate must be at least twice the maximal frequency (Nyquist theorem), or $1/\Delta r \geq 2 \times 1/\Delta s$ or $\Delta r \leq \Delta s/2$.

Q3 **Filtered back projection (FBP)**

a) **FBP uses the ramp filter $|k|$ cut off at $k_{\max}$. Calculate the inverse Fourier transform $q(r)$ of this filter, knowing that it can be written as the difference of a block and a triangle.**

See course notes.

b) **A sharp cut off at $k_{\max}$ is often avoided by suppressing the highest spatial frequencies. Assume the filter function as shown in the figure instead of the ramp filter. Calculate the inverse Fourier transform of this filter and compare it to $q(r)$.**

The figure consists of two triangles of width $k_{\max}$ and height $k_{\max}/2$, centered at $k = +k_{\max}/2$ and $k = -k_{\max}/2$ respectively. The figure can also be seen as a triangle of width $2.k_{\max}$ and height $k_{\max}$ centered at 0, from which the middle part is subtracted twice. The middle part of the first triangle is itself a triangle of width $k_{\max}$ and height $k_{\max}/2$, centered at 0.

Define the triangle function (in the frequency domain) as follows:

$$\Lambda(k) = 1 - |k| \text{ for } |k| < 1, \ 0 \text{ otherwise} \tag{146}$$

Its Fourier transform is given by

$$\begin{aligned}
\mathcal{F}^{-1}\{\Lambda(k)\} &= \int_{-\infty}^{\infty} \Lambda(k) e^{j2\pi kr} dk & (147)\\
&= \int_{-1}^{+1} (1-|k|).(\cos(2\pi kr) + j.\sin(2\pi kr))dk & (148)\\
&= \int_{-1}^{0} (1+k).\cos(2\pi kr)dk + \int_{0}^{+1} (1-k).\cos(2\pi kr)dk & (149)\\
&= \int_{-1}^{+1} \cos(2\pi kr)dk - 2.\int_{0}^{+1} k.\cos(2\pi kr)dk & (150)
\end{aligned}$$

Note that the imaginary part of the integral with $j.\sin(2\pi kr)$ can be dropped at it evaluates to zero, because $1 - |k|$ is an even function. The last line results from the fact that k is odd. Integrating both parts yields, using partial integration for the second part:

$$\begin{aligned}
\mathcal{F}^{-1}\{\Lambda(k)\} &= \frac{1}{2\pi r}.\sin(2\pi kr)|_{-1}^{+1} & (151)\\
&\quad -\frac{2}{2\pi r}.[k.\sin(2\pi kr)]|_{0}^{+1} - \frac{2}{(2\pi r)^2}[\cos(2\pi kr)]|_{0}^{+1} & (152)\\
&= -\frac{2}{(2\pi r)^2}.(\cos(2\pi r) - 1) & (153)\\
&= -\frac{2}{(2\pi r)^2}.(-2\sin^2(\pi r)) & (154)\\
&= \frac{\sin^2(\pi r)}{(\pi r)^2} & (155)\\
&= \text{sinc}^2(\pi r) & (156)
\end{aligned}$$

The same result would also have been obtained by realizing that the triangle function $\Lambda(k)$ defined above is in fact the convolution of the rectangular

function $\Pi(k)$ (see question Q2) with itself:

$$\Lambda(k) = \Pi(k) * \Pi(k) \tag{157}$$

such that the inverse FT of $\Lambda(k)$ is the product of the inverse FT of $\Pi(k)$, which is $\text{sinc}(\pi r)$, with itself, hence $\text{sinc}^2(\pi r)$.

If the inverse Fourier transform of $F(k)$ is $f(r)$, then the inverse Fourier transform of $F(ak)$ for some constant $a \neq 0$ is $\frac{1}{a}.f(r/a)$ by the same reasoning as in question Q2.

The given filter $F(k)$ can be written as the difference of two stretched and scaled triangles:

$$F(k) = k_{\max}.\Lambda(k/k_{\max}) - 2.\frac{k_{\max}}{2}.\Lambda(2k/k_{\max}) \tag{158}$$

Its inverse FT then becomes:

$$\mathcal{F}^{-1}\{F(k)\} = k_{\max}^2.\text{sinc}^2(\pi r.k_{\max}) - 2.\left(\frac{k_{\max}}{2}\right)^2.\text{sinc}^2(\pi r.k_{\max}/2) \tag{159}$$

Simplifying:

$$\begin{aligned}
\mathcal{F}^{-1}\{F(k)\} &= k_{\max}^2.\frac{\sin^2(\pi r.k_{\max})}{(\pi r.k_{\max})^2} - 2.\left(\frac{k_{\max}}{2}\right)^2.\frac{\sin^2(\pi r.k_{\max}/2)}{(\pi r.k_{\max}/2)^2} && (160)\\
&= \frac{1}{(\pi r)^2}.(\sin^2(\pi r.k_{\max}) - 2.\sin^2(\pi r.k_{\max}/2)) && (161)\\
&= \frac{1}{(\pi r)^2}.(4.\sin^2(\pi r.\frac{k_{\max}}{2}).\cos^2(\pi r.\frac{k_{\max}}{2}) - 2.\sin^2(\pi r.\frac{k_{\max}}{2})) && (162)\\
&= \frac{2}{(\pi r)^2}.\sin^2(\pi r.\frac{k_{\max}}{2}).(2.\cos^2(\pi r.\frac{k_{\max}}{2}) - 1) && (163)\\
&= \frac{2}{(\pi r)^2}.\sin^2(\pi r.\frac{k_{\max}}{2}).\cos(\pi r.k_{\max}) && (164)
\end{aligned}$$

Q4 **Multi-slice helical CT with 64 detector rows, 180^o interpolation, 4 rotations per second, and scan length 40 cm. A rectangular slice sensitivity profile (SSP) of width Δz was previously assumed. However, the sensitivity will be higher in the center of the slice profile. Let us assume a triangular SSP $\Lambda(z/\Delta z)$ with width $\Delta z = 0.6$ mm in the center of the field of view.**

a) **What is the effect on the resolution as compared to a rectangular SSP of width $\Delta z = 0.5$ mm. Explain.**

The SSP defines the averaging of the intensity profile that occurs along the axial (z) direction due to the finite extent Δz of the detectors (ideally, infinitely small) and the non-homogeneous response of the detector over the range Δz (ideally, flat over its full extent). Due to the cone beam geometry and the divergent X-ray beam, the SSP is specified in the center of the FOV ($\Delta z = 0.5$ mm in the center of the FOV would typically correspond to about 1 mm at the detector itself). The effect of the SSP is a convolution of the ideal projections in the spatial domain, or a multiplication with its Fourier transform in the frequency domain. The difference between a rectangular versus a triangular shaped SSP can thus be analysed by comparing their FTs.

The rectangular SSP of width Δz can be defined as

$$\text{SSP}_r = \Pi(z/\Delta z) \tag{165}$$

with $\Pi(z)$ the rectangular function extending from -1/2 to +1/2, height 1. Its Fourier transform is

$$F_r(k) = \mathcal{F}^{-1}\{\text{SSP}_r\} = \Delta z.\text{sinc}(\pi k \Delta z) \tag{166}$$

The triangular SSP of width Δz can be defined as

$$\text{SSP}_t = \Lambda(z/(\Delta z/2)) \tag{167}$$

with $\Lambda(z)$ the triangular function extending from -1 to +1, height 1 (note that $\Lambda(z)$ is sometimes defined to extend from -1/2 to +1/2). Its Fourier transform is

$$F_t(k) = \mathcal{F}^{-1}\{\text{SSP}_t\} = \frac{\Delta z}{2}.\text{sinc}^2(\pi k \frac{\Delta z}{2}) \tag{168}$$

These averaging filters behave as low-pass filters. F_r has zero's at values k for which $\pi k \Delta z = n\pi$, $n \in \mathbb{Z}_0$, thus at $k = 1/\Delta z$ and multiples. F_t has zero's at values k for which $\pi k \Delta z/2 = n\pi$, $n \in \mathbb{Z}_0$, thus at $k = 2/\Delta z$ and multiples. The cut off frequency for SSP_r, i.e. the location of the first zero, is thus at $k_{\text{max},r} = 1/\Delta z$, while for SSP_t the cut off frequency is at $k_{\text{max},t} = 2/\Delta_z$, or $k_{\text{max},t} = 2 \times k_{\text{max},r}$ in case the width Δz is the same for both. Hence, if the width Δz is the same for both SSPs, the central lobe of the averaging filter is twice as broad for the triangular SSP than for the rectangular SSP. Fine scale details at spatial frequencies around k_{max}, r and up to $2 \times k_{\text{max}}, r$ will still be included in the measured signal in case of SSP_t, but will be largely eliminated (ignoring the second and other lobes of the sinc-function for the moment) in case of SSP_r. This can be understood from the observation that the rectangular SSP weighs all information within the range Δz equally, while the triangular

SSP gives higher weight to the central part of Δz (weighted averaging instead of uniform averaging), resulting in a better ability to discriminate finer scale details.

In this particular setting, $\Delta z = 0.5$ mm for the rectangular SSP and $\Delta z = 0.6$ mm for the triangular SSP. Hence, $k_{\max,r} = 1/0.5 = 2/\text{mm}$ and $k_{\max,t} = 2/0.6 = 3.33/\text{mm}$. The spatial resolution is thus about 1.6 times better for SSP_t than SSP_r in this case.

b) **Calculate the maximal value of the table feed TF to avoid aliasing?**

The effect of the use of a detector with finite, discrete elements along the z-direction is not just signal averaging by the SSP, but also sampling of the continuous z-profile (for each particular in-slice projection direction θ and in-slice offset r) at a set of discrete z-values. To reconstruct the projections at any particular z-value, interpolation between the nearest samples along the z-direction is applied. In case of a multi-slice scanner and if no table feed was applied, the acquired samples would be simply Δz apart, with Δz the detector size (assuming no spacing between the detectors). Due to the table feed, projections that are 180^o apart can yield interleaved samples, depending on the pitch. With a proper choice of the pitch, more than one sample can be obtained per interval Δz.

Consider first the case of a single detector row, using 180^o reconstruction. Per rotation over 360^o, i.e. per table feed, two samples (from parallel but opposite directions) are obtained and the sample spacing is: $\Delta s = \text{TF}/2$. Let us calculate the number of samples needed to reconstruct the measured (= averaged) profile without aliasing. Aliasing occurs when the signal is under-sampled, such that its high-frequency content is misplaced "somewhere else" in the frequency domain, likely interfering with the low-frequency content of the signal. Assuming that the signal is band-limited with maximal (spatial) frequency $k_{\max}$, the minimal sampling rate to avoid aliasing is $2.k_{\max}$ (Nyquist criterion), i.e. the spacing Δs between samples should not be larger than $1/(2.k_{\max})$. This can be easily verified by considering that sampling in the spatial domain at sampling distance Δs corresponds to multiplication with a pulse-train of spacing Δs, which in the frequency domain is a convolution with a pulse-train of spacing $\Delta k = 1/\Delta s$, i.e. replicating the frequency spectrum $F(k)$ of the signal (with $|k| \leq k_{\max}$) at multiples of Δk. To avoid that consecutive such replications overlap, $\Delta k \geq 2.k_{\max}$, or: $\Delta s \leq 1/(2.k_{\max})$.

For the two considered SSPs above, we found that $k_{\max,r} = 1/\Delta z$ for the rectangular SSP, while $k_{\max,t} = 2/\Delta z$ for the triangular SSP. Hence:

$$\Delta s_r \leq \Delta z/2 \tag{169}$$

$$\Delta s_t \leq \Delta z/4 \tag{170}$$

Thus, while the triangular SSP offers the potential for a resolution that is twice as good as the rectangular SSP (for the same detector width Δz), in practice this potential better resolution can only be exploited if the sampling distance is only half as large, thus if twice as many non-overlapping samples per detector element are acquired.

Using the fact that $\Delta s = \text{TF}/2$ for 180^o interpolation, we have:

$$\text{TF}_r \leq \Delta z \tag{171}$$

$$\text{TF}_t \leq \Delta z/2 \tag{172}$$

for the rectangular and triangular SSP respectively. This corresponds to a pitch (= table feed over beam width, which is here equivalent to the detector size Δz) of 1 for the first and of 1/2 for the second. In this way, 2 samples are collected per detector width Δz by selecting a pitch of 1 for the rectangular SSP, and 4 samples for the triangular profile by selecting a pitch of 1/2. Of course, a pitch smaller than 1 is not optimal with respect to patient dose considerations, as the same volume is irradiated more than once.

Let us next extend this to multi-slice CT, i.e. a detector with multiple rows (64 in this example). We wish to exploit the fact that multiple samples spaced Δz apart are acquired at once, such that faster scanning should be feasible, while still achieving the required sampling rate. The definition of pitch is therefore redefined by considering the total beam width, i.e. $n \times \Delta z$, with n the number of detectors. In order to obtain 2 samples per detector width Δz, the projections at 180^o should be interleaved with those at 0 and 360 degrees. However, for a detector with an even number of rows n, note that a pitch of 1, i.e. a table feed equal to $n\Delta z$, does not yield interleaved samples at 0, 180 and 360 degrees, but identical samples. Instead, a table feed of $(n-1)\Delta z$ should be applied, such that the detector appears shifted over a distance Δz at 360^o, and hence over a distance $\Delta z/2$ at 180^o. The pitch to be applied is thus $(n-1)\Delta z/(n.\Delta z) = (n-1)/n$, or near 1 when n is large ($63/64 = 0.984$ for n=64) . Similarly, in order to obtain 4 interleaved samples with a pitch near 1/2, the detector should be shifted $(n-1)\Delta z$ during 2 rotations, or a table feed of $(n-1)/2.\Delta z$. Four interleaved samples are then obtained over 2 rotations, at 0^o, 180^o, 360^o and 540^o, at a sample distance $\Delta z/4$. The pitch is then $(n-1)/(2n)$ or 0.492 for n=64.

Conclusion: to avoid aliasing, the pitch should be 1 for the rectangular SSP and 1/2 for the triangular SSP (to exploit the potential higher resolution). To obtain interleaved samples, the pitch should not be exactly 1 or 1/2, but somewhat less (1 detector element shift per rotation or per 2 rotations). Thus:

$$\text{SSP}_r: \quad \text{TF} = (n-1).\Delta z \tag{173}$$
$$\text{SSP}_t: \quad \text{TF} = (n-1)/2.\Delta z \tag{174}$$

For this specific case, with $n = 64$:

$$\text{SSP}_r: \quad \Delta z = 0.5\text{mm} \Rightarrow \text{TF} = 31.5\text{mm} \tag{175}$$
$$\text{SSP}_t: \quad \Delta z = 0.6\text{mm} \Rightarrow \text{TF} = 18.9\text{mm} \tag{176}$$

c) **Compute the scan time**

$$\text{Scan time} = \frac{\text{Scan length}}{\text{TF}} \times \text{Time per rotation} \tag{177}$$

The time per rotation (4 rotations per second) is 0.25s.

$$\text{SSP}_r: \quad \text{Scan time} = (400/31.5) \times 0.25 = 3.17s \tag{178}$$
$$\text{SSP}_t: \quad \text{Scan time} = (400/18.9) \times 0.25 = 5.29s \tag{179}$$
$$\tag{180}$$

However, this ignores the overscan that is needed at the beginning and at the end of the scanned volume to acquire complete projection data to reconstruct the entire volume:

$$\text{Scan time} = \frac{\text{Scan length} + \text{Overscan length}}{\text{TF}} \times \text{Time per rotation} \tag{181}$$

The helical scan should start somewhat before and end somewhat further than the volume to be imaged, typically 1 total detector width on both ends. The overscan length can thus be estimated as $2 \times n\Delta z$, or 64 mm and 76.8 mm for SSP_r ($\Delta z = 0.5$ mm) and SSP_t ($\Delta z = 0.6$ mm) in this example respectively. Hence when taking the overscan into account:

$$\text{SSP}_r: \quad \text{Scan time} = (464/31.5) \times 0.25 = 3.68s \tag{182}$$

$$\text{SSP}_t: \quad \text{Scan time} = (477/18.9) \times 0.25 = 6.31s \tag{183}$$

Q5 **Given are two different tissues a and b. Two different detector sizes are used as indicated in the figure. In the first case the detector is twice as large as in the second case.**

a) **Calculate the linear attenuation coefficients μ_a, μ_b and μ_{a+b} from the input intensity I_i and the output intensities I_o, I_{oa} and I_{ob}.**
From the figure, we have:

$$I_{oa} = \frac{1}{2} I_i . e^{-\mu_a . d} \tag{184}$$
$$I_{ob} = \frac{1}{2} I_i . e^{-\mu_b . d} \tag{185}$$
$$I_o = I_{oa} + I_{ob} \tag{186}$$
$$I_o = I_i . e^{-\mu_{a+b} . d} \tag{187}$$

such that with $d = 1$:

$$\mu_a = \log \frac{I_i}{2 I_{oa}} \tag{188}$$
$$\mu_b = \log \frac{I_i}{2 I_{ob}} \tag{189}$$
$$\mu_{a+b} = \log \frac{I_i}{I_o} = \log \frac{I_i}{I_{oa} + I_{ob}} \tag{190}$$

b) **Show that μ_{a+b} is always an underestimate of the mean linear attenuation $(\mu_a + \mu_b)/2$.**

$$\frac{1}{2}(\mu_a + \mu_b) = \frac{1}{2} . (\log \frac{I_i}{2 I_{oa}} + \log \frac{I_i}{2 I_{ob}}) \tag{191}$$
$$= \frac{1}{2} . \log \frac{I_i^2}{4 I_{oa} I_{ob}} \tag{192}$$
$$= \log \frac{I_i}{2\sqrt{I_{oa} I_{ob}}} \tag{193}$$
$$\geq \log \frac{I_i}{I_{oa} + I_{ob}} = \mu_{a+b} \tag{194}$$

The last inequality can be derived as follows (with $x, y \geq 0$):

$$(x - y)^2 \geq 0 \Rightarrow x^2 + y^2 \geq 2xy \Rightarrow (x + y)^2 \geq 4xy \Rightarrow 2\sqrt{xy} \leq x + y \tag{195}$$

Another proof follows from Jenssen's inequality for concave functions f:

$$f(\alpha . x_1 + (1 - \alpha) . x_2) \geq \alpha . f(x_1) + (1 - \alpha) . f(x_2) \tag{196}$$

Because the log-function is concave, we have:

$$-\frac{1}{2}(\mu_a + \mu_b) = \frac{1}{2} . (\log \frac{2 I_{oa}}{I_i} + \log \frac{2 I_{ob}}{I_i}) \tag{197}$$
$$\leq \log \frac{1}{2} . (\frac{2 I_{oa}}{I_i} + \frac{2 I_{ob}}{I_i}) \tag{198}$$
$$= \log \frac{I_{oa} + I_{ob}}{I_i} \tag{199}$$
$$= -\mu_{a+b} \tag{200}$$

and thus

$$\mu_{a+b} \leq \frac{1}{2}(\mu_a + \mu_b) \tag{201}$$

c) **What is the influence of this underestimate on a reconstructed CT image? Explain.**

Q6 **Assume multi-slice scanning without table motion. The detector width in the center of the FOV is 0.5 mm. Assume a block-shaped SSP in the axial z-direction.**

a) **Draw schematically the Fourier transform (FT) in the z-direction of the measured (sampled) projections**

Two effects are to be considered:

- the convolution with the SSP, which in the Fourier domain corresponds to a multiplication with the FT of the SSP. The SSP smooths the measured signal, i.e. behaves as a low-pass filter in the Fourier domain. The FT of a block-shaped SSP of width Δz (width Δz the beam width, which is equivalent to the size of an individual detector element) is a sinc-function $\Delta z.\text{sinc}(\pi k \Delta z)$. This function has its first zero at $k = 1/\Delta z$ and subsequent zero's at $n/\Delta z$. Spatial details in the measured signal around $k = 1/\Delta z$, i.e. of spatial extent Δz and smaller, are thus eliminated by the SSP (only considering the main lobe of the FT of the SSP and ignoring the side lobes of the sinc-function).
- the sampling of the measured, smoothed signal in a set of discrete values spaced Δz apart. Sampling involves a multiplication of the continuous signal with a pulse-train of spacing Δz, which is a convolution in the frequency domain with a pulse-train of spacing $1/\Delta z$, i.e. the sampling frequency k_s. This convolution replicates the frequency content of the measured continuous signal (which can be considered to be band-limited to the cut off frequency $k_{\max} = 1/\Delta z$ of the SSP low-pass filter, see above) at multiples of $k_s = 1/\Delta z$.

As a result of both effects, aliasing will occur if the frequency content of the original signal to be measured contains frequencies of $k = 1/(2\Delta z)$, i.e. $k_{\max}/2$, and above. The replications of the frequency spectrum induced by the sampling at $k_s = 1/\Delta z$ will then overlap. This is due to the fact that the sampling frequency k_s and the assumed maximal frequency $k_{\max}$ in the smoothed signal are identical. To avoid aliasing, k_s should be a least twice $k_{\max}$.

Otherwise stated: without table motion, only 1 sample per detector width Δz is collected, while in fact 2 samples are needed to avoid aliasing. This can be achieved by acquiring interleaved samples at 180^o after a table shift equivalent to half a detector width $\Delta z/2$. Alternatively, focal spot wobbling along the z-direction could be applied. This means that the focal spot of the source is rapidly electro-magnetically switched during scanning between two slightly different positions along the z-direction that are $\Delta z/2$ apart, such that the source appears at a slightly different position with respect to the detector and two sets of interleaved samples are obtained for each projection.

Another option is to combine multiple samples acquired from thin slices in order to reconstruct thicker slices. This could be achieved by convolution of the samples acquired along the z-direction with a suitable filter (= z-filtering), i.e. a weighted running average of neighbouring samples. This would increase the width of the SSP while keeping the number of samples constant, such that aliasing effects are reduced.

b) **What is the maximum useful frequency of the sampled signal in the z-direction?**

As argued above, the maximal useful frequency is $k_{\max}/2$, with $k_{\max}$ the cut off frequency of the SSP, i.e. $1/(2\Delta z)$. If spatial frequencies larger than $1/(2\Delta z)$ are present in the profile to be measured along the z-direction, aliasing will occur and these spatial details will no longer be distinguishable.

c) **What is the minimum distance δ in the z-direction between small details to be distinguishable? (Represent neighboring details by a sinusoidal function).**

As argued above, the maximal useful spatial frequency is $1/(2\Delta z)$. This corresponds to a sinusoid with a period of $2\Delta z$. Different small details (corresponding to the maxima in this signal) should thus be $\delta = 2\Delta z$ apart to be distinguishable.

The detector width is specified in the center of the FOV to be 0.5 mm (note: due to the diverging nature of the beam from source to detector, this may correspond to an actual detector size of 1 mm, assuming that the center of the FOV is approximately half way the source and detector). Hence: $\Delta z = 0.5\text{mm}$, $k_{\max} = 2/\text{mm}$, $k_s = 2/\text{mm}$, $\delta = 1\text{mm}$.

Q7 **Helical CT. Assume that β is the angular position of the X-ray tube and z its axial position.**

a) **Draw the data acquisition trajectory in the (β, z) space. What is the maximal table feed to avoid aliasing in case of 360^o interpolation?**

Consider the case of a single row detector. The table feed represents the relative motion of the detector in the z-direction after a single, full 360^o rotation. In case of 360^o interpolation, only a single sample for each projection is acquired along the z-direction during 1 complete rotation, i.e. 1 sample per table feed length. The distance between consecutive samples along z thus equals the table feed (TF). To avoid aliasing, at least two samples per beam width, i.e. per detector width Δz, should be acquired. Hence, the distance between 2 samples should not be larger than $\Delta z/2$. Thus: TF $\leq \Delta z/2$, or a pitch (= TF over beam width) of 1/2. This means that the TF should be such that the beams for consecutive samples at positions $\beta + k * 360^o$ are (at least) half overlapping.

b) **What is the maximal table feed to avoid aliasing in case of 180^o interpolation? Show this in the (β, z) space.**

180^o interpolation exploits the fact that projections at β and $\beta + 180^o$ are parallel but opposite and if no table motion is applied, are expected to be identical. Due to the table motion, interleaved samples along the z-direction are obtained for the same projection directions at positions $\beta + k * 180^o$. Thus, 2 samples are obtained per full rotation, i.e. 2 samples per table feed length. The distance between consecutive samples along z thus equals half the table feed (TF/2). To avoid aliasing, at least two samples per beam width, i.e. per detector width Δz, should be acquired. Hence, the distance between 2 samples should not be larger than $\Delta z/2$. Thus: TF/2 $\leq \Delta z/2$, or a pitch (= TF over beam width) of 1. This means that the TF should be such that the beams for consecutive samples at positions $\beta + k * 360^o$ are exactly joining without gap. The dose to the patient is then only half of what would be the case in the scenario (a) above with pitch 1/2.

c) **What is the maximal table feed to avoid aliasing in case of dual source CT in which the two X-ray tubes are positioned 90^o apart (and operate simultaneously at equal kV). Draw the data acquisition trajectories in the (β, z) space for the two X-ray sources and show how the use of two sources influences the table feed.**

With two sources A and B (and two detectors) that are 90^o apart, 4 samples per rotation (i.e. per table feed length TF) are obtained for each projection direction, namely every 90 degrees at $\beta + k * 90^o$ and alternately from source A or source B. The spacing between these samples is thus TF/4. As a result, the pitch can be increased to 2 without aliasing: TF/4 $\leq \Delta z/2 \Rightarrow$ TF $\leq 2\Delta z$. Scanning can thus be twice as fast with a dual-source CT (at equal kV) without loss of axial resolution. Note that the dose to the patient remains the same as for scenario (b) above, as the X-ray intensity is doubled by the use of two sources while the scan time is halved.

The same reasoning holds for multi-slice helical CT with n detector elements, by considering the full detector width $n\Delta z$ when computing the pitch and by

making sure that the pitch values are slightly adjusted such that interleaved samples are obtained (see Q4).

Q8 **Rewrite Eqs. 3.49-3.53 for maximum-likelihood reconstruction of CT images if the Poisson distribution is replaced by a Gaussian distribution.**

Let $\mathcal{U} = \{\mu_j\}$ denote the image to be reconstructed for all voxels j, and $\mathcal{I} = \{I_i\}$ the measurements for each projection i. Assuming the variance σ to be the same for all projections:

$$\bar{I}_i = I_0 . \exp(- \sum_{j=1,M} c_{ij}\mu_j), \quad i = 1, N \tag{202}$$

$$p(I_i|\bar{I}_i) = \frac{1}{\sqrt{2\pi\sigma^2}} . \exp(-\frac{(I_i - \bar{I}_i)^2}{2\sigma^2}) \tag{203}$$

$$p(\mathcal{I}|\mathcal{U}) = \prod_i \frac{1}{\sqrt{2\pi\sigma^2}} . \exp(-\frac{(I_i - \bar{I}_i)^2}{2\sigma^2}) \tag{204}$$

$$\arg\max_{\mathcal{U}} p(\mathcal{I}|\mathcal{U}) = \arg\max_{\mathcal{U}} \ln p(\mathcal{I}|\mathcal{U}) \tag{205}$$

$$= \arg\min_{\mathcal{U}} \sum_i (I_i - \bar{I}_i)^2 \tag{206}$$

$$\frac{\delta}{\delta\mu_j} \sum_i (I_i - \bar{I}_i)^2 = 2 \sum_i (I_i - \bar{I}_i) . \bar{I}_i . c_{ij} = 0 \quad \forall j = 1, J \tag{207}$$

Q9 **Cardiac CT. The following conditions are given:**

- **A CT scanner with 128 detector rows.**
- **The detector width in the center of the FOV is 0.5 mm.**
- **A full rotation (360^o) of the X-ray tube takes 0.33 s.**
- **A full data set for reconstruction requires projection values for a range of 210^o.**
- **Maximum 1/4 of the heart cycle can be used for acquiring projection data.**
- **The heart rhythm is 72 bpm.**
- **The scan length is 20 cm.**

a) **Calculate the duration of 1/4 of heart cycle (in seconds).**

72 heart cycles per minute = 60 seconds
$\Rightarrow \Delta b$ = duration of 1 heart cycle = 60/72 seconds = 5/6 seconds
$\Rightarrow$ duration of 1/4 of a heart cycle = 5/24 seconds $\approx$ 5/25 = 0.2 s (somewhat more)

b) **Calculate (in seconds) the time needed to obtain projection values for a range of 210^o.**

360^o rotation in 0.33s = 1/3 s
$\Rightarrow 210^o$ degrees (= 180^o + fan angle) in Δt seconds
$\Rightarrow \Delta t$ = 210/360 x 1/3 = 7/36 $\approx$ 7/35 = 0.2 s (somewhat less)

c) **What can you conclude from (a) and (b)?**

The time needed to collect a full set of projection data fits within the 1/4 of the heart cycle that can be used for imaging (when cardiac motion is relatively small).

d) **Assume that the table shift per heart beat is equal to the total width of the detector rows (i.e. the total z-collimation). Calculate the acquisition time.**

Total width of the detector = 128 x 0.5 mm = 64 mm
$\Rightarrow$ table feed = 64 mm / heart beat
$\Rightarrow$ to cover 20cm, 4 heart beats in total are needed (= 200 / 64)
$\Rightarrow$ acquisition time = 4 x Δb = 4 x 5/6 s = 20/6 = 3.33 s

e) **The assumption under (d) is approximate. Explain why? How does this approximation influence the acquisition time?**

During the measurement time Δt, the detector moves relative to the table over a distance $\alpha \Delta t$, with α the table feed (in mm/s). The total length covered by the detector during collection of the projections is thus $W + \alpha.\Delta t$, with W the detector width. However, only the central part of length $W - \alpha\Delta t$ is covered by the detector during the full duration of the measurement, such that only from this part complete projection data are obtained during the time Δt. Within one heart beat of duration Δb, a slab of length $W - \alpha\Delta t$ can thus be imaged. Over a time period Δb of one heart beat, the table should thus move a distance $W - \alpha\Delta t$. This means that the table feed α should be equal to:

$$\alpha = \frac{W - \alpha.\Delta t}{\Delta b} \Rightarrow \alpha = \frac{W}{\Delta b + \Delta t} \tag{208}$$

Comparing this to the previous assumption which stated that $\alpha = W/\Delta b$, accounting for the detector motion during the measurement itself makes the table feed smaller, and hence the total acquisition time longer (in principle). It was argued above that $\Delta t \leq 1/4 \times \Delta b$. Hence:

$$\alpha \approx \frac{W}{\Delta b + \Delta b/4} = \frac{4}{5}\frac{W}{\Delta b} \tag{209}$$

i.e. 20% smaller than the assumption above would yield. The total acquisition time is thus also likely higher, if an additional heart beat is needed to cover the entire volume. For the example given above:

Table feed = $4/5 \times 64$ mm per heart beat = 52 mm / heart beat
$\Rightarrow$ to cover 20cm, 4 heart beats in total are needed (= 200 / 52)
$\Rightarrow$ acquisition time = $4 \times \Delta b = 4 \times 5/6$ s = $20/6 = 3.33$ s

In this case, there is no difference in acquisition time. However, if a volume of 22cm had to be scanned, an additional heart beat was needed and the actual scan time would be $5 \times \Delta b = 4.16s$, while the assumption under (d) above would underestimate the scan time by 1 heart beat.

Q10 **Dual-energy CT. Given**

- **Water and iodine are used as the basis materials.**
- **A patient is injected with a iodine containing fluid. The peak concentration of injected iodine in the blood for this patient is 10 mg/ml.**
- **The voxels consist of bone, soft tissue, water and/or iodine.**

a) **Draw schematically the tissue-specific coefficients of the voxels (in mg/ml).**

The attenuation coefficient of any material within the X-ray energy range of interest can be approximately written as the linear combination of two separate, energy-dependent effects (photoelectric interaction and Compton scatter), with fixed material-specific coefficients that are independent of the energy. As a result, the attenuation coefficient $\mu(E)$ for any material can be written as the linear combination of the linear attenuation coefficients $\mu_1(E)$ and $\mu_2(E)$ of two basis materials, using tissue-specific coefficients a_1 and a_2:

$$\mu(E) = a_1.\mu_1(E) + a_2.\mu_2(E) \tag{210}$$

The coefficients a_1 and a_2 are independent of E, such that the same relationship holds for all energies. Knowledge of a_1 and a_2 relative to 2 given basis materials thus fully characterizes each tissue (in the absence of K-edges within the energy range of interest !). To compute its linear attenuation coefficient μ, the energy E has to be specified such that $\mu_1(E)$ and $\mu_2(E)$ and hence $\mu(E)$ at this particular energy can be determined. Hence, once a_1 and a_2 are known in each voxel, monochromatic images $\mu(E)$ at any energy can be computed.

In order to determine $a_1(x, y)$ and $a_2(x, y)$ in every pixel (x, y) in the image plane, two measurements at (sufficiently) different X-ray tube voltages (i.e. different X-ray energy spectra, e.g. at 80 and 140 kV) are needed. This is the principle of dual-energy imaging. See the textbook for derivations.

The coefficients a_1 and a_2 can be interpreted as thicknesses. The attenuation by 1 cm of tissue is equivalent to the attenuation of a_1 cm of basis material 1 followed by a_2 cm of basis material 2:

$$\mu(E) \times 1\text{cm} = \mu_1(E) \times a_1\text{cm} + \mu_2(E) \times a_2\text{cm} \tag{211}$$

Alternatively, the coefficients can also be interpreted as a "concentration", by considering the mass attenuation coefficient $\mu_m = \mu/\rho$ of the basis materials with ρ the density:

$$\rho\mu_m = a_1.\rho_1.\frac{\mu_1}{\rho_1} + a_2.\rho_2.\frac{\mu_2}{\rho_2} = a_1^*.\mu_{m,1} + a_2^*.\mu_{m,2} \tag{212}$$

The coefficients $a_1^* = \rho_1.a_1$ and $a_2^* = \rho_2.a_2$ can be expressed in the same units as ρ, e.g. in g/ml for water ($\rho_{\text{water}} = 1$ g/ml) and mg/ml for iodine ($\rho_{\text{iodine}} = 4.93$ g/ml, but a peak concentration of 10 mg/ml is given).

Different tissues can thus be identified by their coefficients a_1^* (g/ml) and a_2^* (mg/ml) for basis materials water and iodine respectively. Obviously:

- water: $a_1 = 1$, $a_2 = 0 \Rightarrow a_1^* = 1$, $a_2^* = 0$

- undiluted iodine: $a_1 = 0$, $a_2 = 1 \Rightarrow a_1^* = 0$, $a_2^* = \rho_2 = 4930$
- solution of iodine in water at 10 mg/ml: $a_1^* \approx 1$, $a_2^* = 10 \Rightarrow a_1 \approx 1$, $a_2 \approx 0.002$

In order to characterize the materials such as soft tissue or bone, we need to know $\mu_m(E)$ for at least 2 different values of E and relate these to $\mu_{m,1}(E)$ and $\mu_{m,2}(E)$ to solve for a_1 and a_2 (or a_1^* and a_2^*):

$$\rho\mu(E_1) = a_1^*.\mu_{m,1}(E_1) + a_2^*.\mu_{m,2}(E_1) \quad (213)$$
$$\rho\mu(E_2) = a_1^*.\mu_{m,1}(E_2) + a_2^*.\mu_{m,2}(E_2) \quad (214)$$

or in matrix form:

$$[\mu_{m,1}(\boldsymbol{E}) \quad \mu_{m,2}(\boldsymbol{E})].\boldsymbol{A^*} = \rho\mu_m(\boldsymbol{E}) \quad (215)$$

with $\mu_1(\boldsymbol{E})$, $\mu_2(\boldsymbol{E})$ and $\mu(\boldsymbol{E})$ column vectors of length n_E, with n_E the number of energy values considered and $\boldsymbol{A^*} = [a_1^* \;\; a_2^*]^T$.

Values of ρ and of μ_m at different energies for different materials can be found for instance here:

`https://www.nist.gov/pml/x-ray-mass-attenuation-coefficients`

For bone, two tables can be found, namely for cortical bone and for cancellous bone (or rather for B-100, a bone-equivalent plastic, with similar attenuation properties as cancellous bone). The relevant values are summarized in the table below.

Material	Water	Undiluted iodine	Soft tissue	Cancellous bone	Cortical bone
ρ (g/ml)	1	4.93	1.06	1.45	1.92
E (keV)	$\mu_{m,1}$	$\mu_{m,2}$			
40	0.268300	22.100000	0.268800	0.517900	0.665500
50	0.226900	12.320000	0.226400	0.350600	0.424200
60	0.205900	7.579000	0.204800	0.273800	0.314800
80	0.183700	3.510000	0.182300	0.207600	0.222900
100	0.170700	1.942000	0.169300	0.179300	0.185500
150	0.150500	0.697800	0.149200	0.148200	0.148000
200	0.137000	0.366300	0.135800	0.132500	0.130900
300	0.118600	0.177100	0.117500	0.113500	0.111300

Solving for a_1^* and a_2^* using water and iodine as basis materials, yields:

Material	Water	Iodine solution	Soft tissue	Cancellous bone	Cortical bone
a_1^* (g/ml)	1	0	1.05	1.32	1.66
a_2^* (mg/ml)	0	10	0	17.6	37.0

Now look at Figure (3.19), page 49 (FMI, version 2) or Figure (3.23), page 54 (FMI, version 3). Each black dot represents the coefficients a_1^* and a_2^* for individual voxels in the reconstructed dual-energy CT image, characterizing their linear attenuation in terms of the basis materials water and iodine. Because the tissues are not homogeneous and due to measurement limitations (noise, beam hardening, partial volume effects...), the voxels in this plot form clusters that are characteristic for a certain tissue type. The tissue type corresponding

to different clusters in the plot can be identified using the values in the table above. Note that "iodine" in the plot refers to the diluted iodine of the contrast agent. Note that the maximal concentration of the diluted iodine is about 20 mg/l in the figure (while above a maximal concentration of 10 mg/l was specified). Note also that bone is located around (± 1.4, ± 17) in the plot, which is only slightly different to the value derived above for the bone-mimicking plastic. Between these clusters, such as between soft tissue and bone or soft-tissue and iodine, linear patterns of voxels appear, which can be attributed to partial volume effects, i.e. voxels containing a mixture of more than one tissue, which typically occurs at the interface between different tissues, resulting in an attenuation that is an average of the pure tissues (see question Q5).

Finally, the plot also contains lines of constant CT number, expressed in HU. The CT number is defined as:

$$HU = 1000.\frac{\mu - \mu_{\text{water}}}{\mu_{\text{water}}} \tag{216}$$

Consider first the horizontal axis along which $a_2^* = 0$. For the values along this axis, we have that $\mu = a_1^*.\mu_{m,\text{water}}$ or $\mu = a_1^*.\mu_{\text{water}}$ as $\rho_{\text{water}} = 1$ g/ml. Hence:

$$HU(a_1^*|a_2^* = 0) = 1000.\frac{a_1^*.\mu_{\text{water}} - \mu_{\text{water}}}{\mu_{\text{water}}} = 1000.(a_1^* - 1) \tag{217}$$

Thus: $a_1^* = 1 \Rightarrow$ HU $= 0$, $a_1^* = 1.4 \Rightarrow$ HU $= 400$, $a_1^* = 1.7 \Rightarrow$ HU $= 700$ Note that the values for HU (the CT number) along this axis are independent of the energy E at which $\mu(E)$ is considered, because the values are relative to $\mu_{\text{water}}(E)$ and as E changes, the reference value for water also changes. This defines one point of the lines drawn, we need a second point.

Consider next the vertical axis along which $a_1^* = 0.8$ (note: not 0!). In order to find the value of a_2^* that corresponds to a material with attenuation of 0 HU, we should equate its attenuation to that of the corresponding point on the horizontal axis, i.e. $a_1^* = 1, a_2^* = 0$:

$$0 \text{ HU}: \quad 0.8.\mu_{m,1} + a_2^*.\mu_{m,2} = 1.\mu_{m,1} + 0.\mu_{m,2} \tag{218}$$

Hence:

$$a_2^* = 0.2\frac{\mu_{m,1}}{\mu_{m,2}} \tag{219}$$

in the same units as a_1^* (g/ml), or

$$a_2^* = 200\frac{\mu_{m,1}}{\mu_{m,2}} \tag{220}$$

with a_2^* in mg/l. However: $\mu_{m,1}$, $\mu_{m,2}$ and their ratio are dependent on the energy E, such that also the searched for coefficient a_2^* is energy dependent:

$$a_2^*(E) = 200\frac{\mu_{m,1}(E)}{\mu_{m,2}(E)} \tag{221}$$

The line of 0 HU (and similarly for the other lines of 400 HU and 700 HU) can thus only be drawn if the energy is specified. Referring back to the table above, we have:

E (keV)	$\mu_{m,1}$	$\mu_{m,2}$	$\frac{\mu_{m,1}}{\mu_{m,2}}$	$a_2^* = 200.\frac{\mu_{m,1}}{\mu_{m,2}}$
40	0.268300	22.100000	0.012140	2.428
50	0.226900	12.320000	0.018417	3.683
60	0.205900	7.579000	0.027167	5.433
80	0.183700	3.510000	0.052336	**10.467**
100	0.170700	1.942000	0.087899	17.580
150	0.150500	0.697800	0.215678	43.136
200	0.137000	0.366300	0.374010	74.802
300	0.118600	0.177100	0.669678	133.936

The line that is drawn in the figure corresponding to 0 HU intersects the vertical axis in $a_2^* \approx 10$. This is close to the value of 10.467 in the table above corresponding to $E = 80$ keV. We can thus deduce that the plot is valid for an energy of 80 keV. This can be verified for the other specified HU values as well:

$$\text{HU} = 1000.(a_1^* - 1) \Rightarrow a_1^* = 1 + \frac{\text{HU}}{1000} \tag{222}$$

$$0.8.\mu_{m,1} + a_2^*.\mu_{m,2} = (1 + \frac{\text{HU}}{1000}).\mu_{m,1} + 0.\mu_{m,2} \tag{223}$$

$$a_2^*(\text{HU}) = (0.2 + \frac{\text{HU}}{1000}).\frac{\mu_{m,1}}{\mu_{m,2}} \text{ (in g/ml)} \tag{224}$$

$$a_2^*(\text{HU}) = (200 + \text{HU}).\frac{\mu_{m,1}}{\mu_{m,2}} \text{ (in mg/ml)} \tag{225}$$

and specifying $E = 80$ keV and selecting the corresponding value in the table above:

$$\frac{\mu_{m,1}}{\mu_{m,2}}(E = 80\text{keV}) \approx 0.05 \tag{226}$$

$$a_2^*(\text{HU}) \approx 0.05 \times (200 + \text{HU}) = 10 + \frac{\text{HU}}{20} \tag{227}$$

Thus: 0 HU $\Rightarrow a_2^* = 10$; 400 HU $\Rightarrow a_2^* = 30$; 700 HU $\Rightarrow a_2^* = 45$. (Note that the line for 700 HU is not accurately drawn in the figure).

For larger energies, the difference in attenuation between water and iodine becomes smaller, such that a higher concentration of iodine is needed to be equivalent to water. Vice versa for lower energies. At larger energies, the lines of constant HU thus become steeper (more vertical); at smaller energies, they become flatter (more horizontal).

b) **How can we obtain a precontrast (unenhanced) and a contrast scan?**

With a_1 and a_2 determined for every voxel, a water image ($a_1(x, y)$) and an iodine image ($a_2(x, y)$) can be obtained. However, because all materials denser than water also have a certain iodine component (i.e. a non-zero a_2 coefficient), the a_2 image will contain more than just the contrast agent and the a_1 image will not correspond to a pre-contrast scan. To obtain an image of the contrast itself, the voxels that are affected by the presence of the contrast agent need to be identified and visualized separately. The contrast agent is diluted in the blood. Hence, voxels containing contrast agent can thus assumed to be water-like (i.e. $a_1 \approx 1$) and to have a significantly large iodine content ($a_2 > 0$). This corresponds to the "iodine" cluster depicted in green in the figure. Denoting the voxels in this clusters by the set C_{iodine}, a mono-chromatic precontrast

scan $I_{\text{pre}}(E)$ and a contrast scan I_C can be obtained as follows for any selected energy E and each voxel j:

$$
\begin{aligned}
I_{\text{pre}}(E)(j) &= a_1(j).\mu_1(E) + a_2(j).\mu_2(E) \text{ if } j \notin C_{\text{iodine}} && (228)\\
I_{\text{pre}}(E)(j) &= a_1(j).\mu_1(E) + 0.\mu_2(E) \text{ if } j \in C_{\text{iodine}} && (229)\\
I_{\text{C}}(E)(j) &= 0 \text{ if } j \notin C_{\text{iodine}} && (230)\\
I_{\text{C}}(E)(j) &= 0.\mu_1(E) + a_2(j).\mu_2(E) \text{ if } j \in C_{\text{iodine}} && (231)
\end{aligned}
$$

The water content and the iodine content over the voxels identified as belonging to the iodine cluster are thus separated over the two images.

Q12 **CT of the lungs on a 64-row scanner, 120 kV, 90 mA s, pitch 1, 360^o rotation time 0.33 s, detector width 0.60 mm, slice thickness 1 mm, scan length 38.4 cm, $\mathbf{CTDI_{vol}}$ 10 mGy.**

a) **Calculate the scan time.**

$$\text{Scan time} = \frac{\text{Scan length}}{\text{Table feed}} \times \text{Rotation time} \tag{232}$$

$$\text{Table feed} = \text{Pitch} \times \text{Total detector width} \tag{233}$$

$$\text{Total detector width} = \text{Number of detector rows} \times \text{Detector width} \tag{234}$$

$$\Rightarrow \text{Scan time} = \frac{\text{Scan length} \times \text{Rotation time}}{\text{Pitch} \times \text{Number of detector rows} \times \text{Detector width}} \tag{235}$$

Note:

- The scan time is independent of kV and mA s.
- The slice thickness refers to the reconstruction into image slices of the acquired samples. With a pitch of 1, 2 samples per detector width Δz are acquired from opposite projections (180^o interpolation), spaced $\Delta z/2$ apart (in this case 0.3 mm). These samples can be combined during reconstructing of the image to yield thicker slices (1 mm in this example) by weigthed averaging (=convolution with some filter) of the samples along the z-direction. This is called z-filtering. Reconstructing thicker slices from finer sections reduces partial volume effects compared to measuring thicker sections directly (see Q5).
- To acquire complete date over the entire scan length, typically an overscan at both ends is needed of about 1 time the full detector width. Hence:

$$\text{Actual scan length} = \text{Imaged scan length} + \text{Overscan length} \tag{236}$$

$$\text{Actual scan length} \approx \text{Imaged scan length} + 2 \times \text{Total detector width} \tag{237}$$

$$\Rightarrow \text{Total scan time} = \frac{\text{Actual scan length} \times \text{Rotation time}}{\text{Pitch} \times \text{Total detector width}} \tag{238}$$

$$\Rightarrow \text{Total scan time} = \text{Scan time without overscan} + \frac{2}{\text{Pitch}} \times \text{Rotation time} \tag{239}$$

In the example above, the detector width corresponds to 10% of the specified scan length, such that the actual scan length is 20% larger than the imaged scan length and an increase in scan time of 20% is to be expected when the overscan is also accounted for, thus 4 s (instead of 3.3 s).

b) **Calculate the estimated effective dose. Certain organs are only partially and/or indirectly (scatter) irradiated. The following table gives for each of the irradiated organs the percentage of irradiated tissue and the tissue weighting factor w_T.**

The specified CTDI_{vol} value represents an estimate of the average dose (in Gray, Gy = J/kg) deposited by this particular scan protocol (with the specified kV, mAs, beam width...) during a single 360^o rotation at each location of the imaged volume. CTDI_{vol} accounts for inhomogeneities in the dose distribution along z and within the axial xy plane, and also accounts for the pitch used for helical scanning. Higher pitches spread the dose per rotation over a larger volume, hence the mean dose is reduced in proportion to the pitch (provided that the mAs is not adjusted to maintain SNR in case of faster scanning). The CTDI is derived based on measurements in a standard phantom. It thus only gives an approximate estimate of the dose absorbed by the patient. Moreover, the CTDI is independent of the scan length. Scanning a larger volume of course involves a larger amount of radiation and dose. To take this into account, the dose-length product is defined as $\text{DLP} = \text{CTDI}_{\text{vol}} \times \text{L}$ with L the scan length.

The effective dose is a measure for the risk of radiation induced damage, taking into account the nature of the radiation (X-ray photons, α particles, β particles...) and the organ or tissue that is being irradiated, as some tissues are more sensitive to radiation induced damage than others. To this end, tissue weighting factors w_T are defined, for instance 0.12 for lungs, 0.04 for liver and 0.01 for skin. These should be interpreted as follows: a homogeneous dose of 1 Gy to the entire lung volume is equivalent w.r.t. the risk of radiation induced damaged as a dose of 3 Gy to the entire liver volume or a dose of 12 Gy to the entire skin volume. In practice, only part of an organ is irradiated, hence the irradiated volume fractions f_V have to be considered as well.

The effective dose (in Sievert, Sv) can thus be computed as:

$$E = CTDI_{\text{vol}} \times \sum_{j=1}^{N_T} f_{V,j}.w_{T,j} \tag{240}$$

with j each of the N_T different tissue classes considered. A homogeneous dose of 1 Gy to the entire liver volume thus corresponds to the same effective dose as a dose of 1 Gy to 1/3 of the lung volume. This can be rewritten as follows:

$$E = \left(\frac{1}{L} \sum_{j=1}^{N_T} f_{V,j}.w_{T,j} \right) \times (CTDI_{\text{vol}} \times L) = k.\text{DLP} \tag{241}$$

with k the average regional conversion factor. Standard values for k can be defined for each CT examination protocol (head, thorax, abdomen...), based on the typical volume fractions (f_V) and radiation sensitivity (w_T) of the tissues within the particular anatomical region that is being scanned.

In this example, values for $f_{V,j}$ and $w_{T,j}$ are given in the table for 12 tissues. Note that the w_T values do not sum to 1, but to 0.9. This is because three tissues are not included in the table: gonads with $w_T = 0.08$, brain with $w_T = 0.01$, and salivary glands with $w_T = 0.01$. These tissues therefore are considered to receive only negligible radiation when scanning the thorax ($f_V = 0$). Given $L = 38.4$ cm, we can compute k and DLP using the expressions above: $k = 0.0114$ mSv/(mGy.cm) and DLP = 384 mGy.cm, such that $E = k.DLP = 4.385$ mSv.

Q13 **CT of the lungs. The following table shows the irradiated organs, the irradiated portion of the organs (in %), and the tissue weighting factor w_T. The average regional conversion factor $k = 0.016$ mSv/(mGy.cm). Calculate the scan length.**

See Q12:

$$k = \frac{1}{L}\left(\sum_{j=1}^{N_T} f_{V,j}.w_{T,j}\right) \Rightarrow L = \frac{1}{k}\left(\sum_{j=1}^{N_T} f_{V,j}.w_{T,j}\right) \tag{242}$$

Using the values in the table:

$$\begin{aligned}
\sum_{j=1}^{N_T} f_{V,j}.w_{T,j} &= 1 \times 0.12 + 1 \times 0.12 + 0.65 \times 0.12 + 0.3 \times 0.12 + && (243)\\
&\quad 0.5 \times 0.04 + 1^{(*)} \times 0.04 + 0.3 \times 0.12 && (244)\\
&= 3.25 \times 0.12 + 1.5 \times 0.04 && (245)\\
&= (9.75 + 1.5) \times 0.04 = 11.25 \times 0.04 = 0.45 && (246)
\end{aligned}$$

(*) The table gives a value of 00% for esophagus. This is likely an error and should be 100%, cfr. question Q12.

$$L = \frac{0.45}{0.016} = \frac{450}{16} = \frac{480 - 32 + 2}{16} = 30 - 2 + \frac{1}{8} = 28.125\text{cm} \tag{247}$$

Q16 **Assume a dedicated maxillofacial cone-beam CT scanner operating at 100 kV and 40 mAs. The beam thickness in the center of the FOV is 60 mm. $\mathbf{CTDI}_w$ = 90 mGy and the average regional conversion factor $k = 2.0 \times 10^{-3}$ mSv/(mGy cm).**

a) **Calculate the effective dose. Show the details of your calculations**

$$E = k \times \text{DLP} = k \times L \times \text{CTDI} \tag{248}$$

For a maxillofacial cone-beam CT scanner, a single circular rotation is performed (axial scan). The scan length L is equivalent to the beam width in the FOV, i.e. L = 60 mm. CTDI_w is used to account for the fact that the dose is not the same in the periphery of the FOV versus the center of the FOV. Hence:

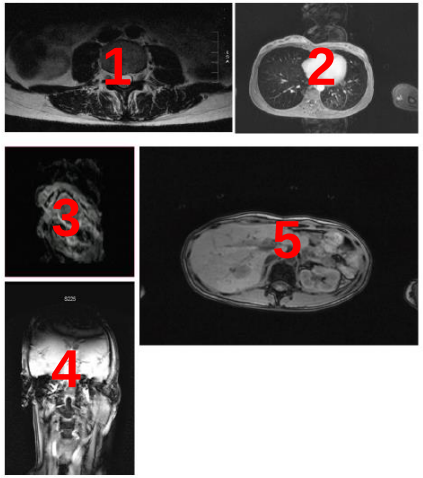
$$E = 2 \times 10^{-3} \frac{\text{mSv}}{(\text{mGy cm})} \times 6\text{cm} \times 90\text{mGy} = \frac{1080}{1000} = 1.08\text{mSv} \tag{249}$$

B.4 Magnetic Resonance Imaging

Medical Imaging and Analysis: MRI

Solutions to Exercises on
Magnetic Resonance Imaging

Exercise 1



1: - increased noise anterior (no coil active)

- chemical shift artefact on spine (fat displaced right)

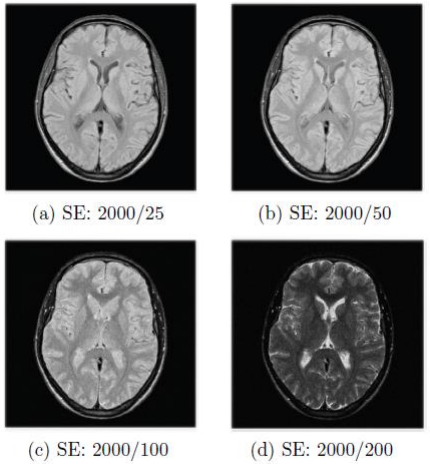
2: cardiac motion/pulsation (ghosting artefact)

3: bad shim/metal artefact (geometric distortion)

4: metal artefact (susceptibility artefact)

5: opposed (out-of-)-phase imaging (phase cancellation artefact)

Exercise 2



(a) SE: 2000/25 (b) SE: 2000/50 (c) SE: 2000/100 (d) SE: 2000/200

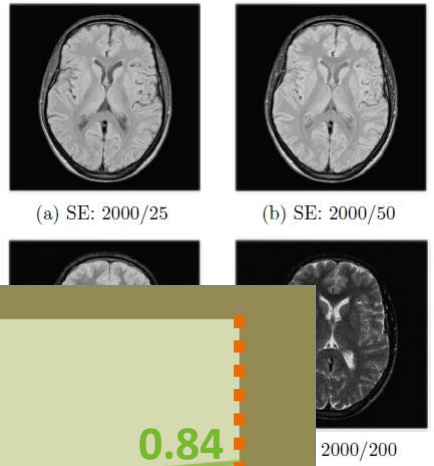
Figure 1.7

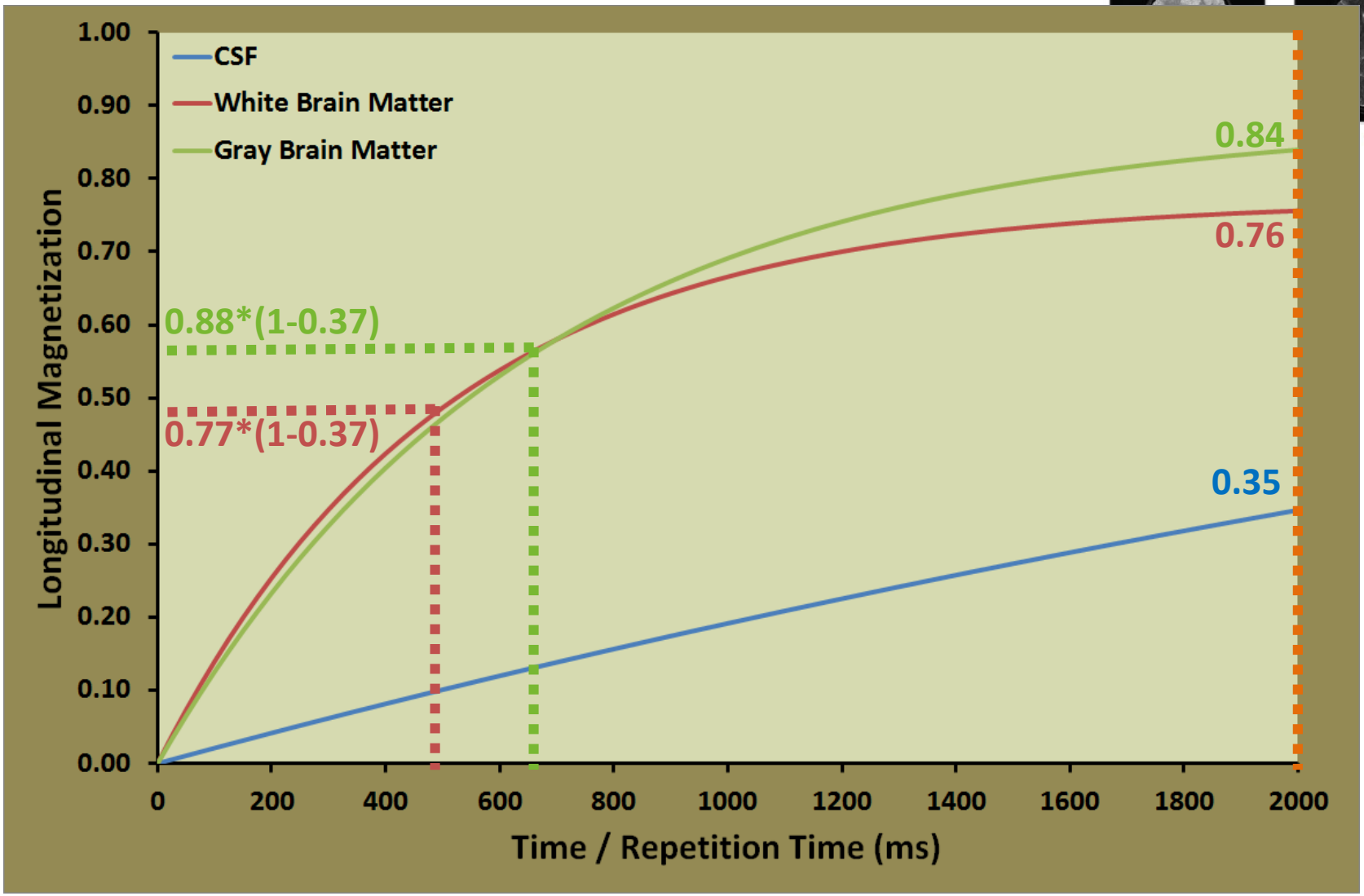
- Assumptions:
  - TR = 2000 ms, TE = 25/50/100/200 ms

	T1 (ms)	T2 (ms)	PD (relative)
White Brain Matter	500	90	0.77
Gray Brain Matter	650	100	0.88 (14% higher than WM)
CSF (fluid)	4700	2200	1

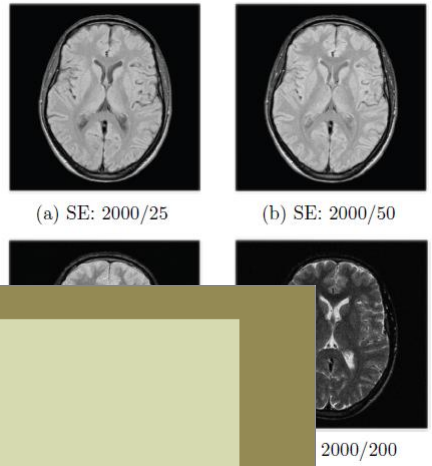
- Formulas:
  - Longitudinal relaxation **(4.26):**
    - $M_l(t) = M_0 \cos\alpha \, e^{-t/T_1} + M_0(1 - e^{-t/T_1})$
    - with α=90° and M_0 = relative proton density
  - Transverse relaxation **(4.31):**
    - $s(t) = M_0(1 - e^{-TR/T_1})e^{-t/T_2}$

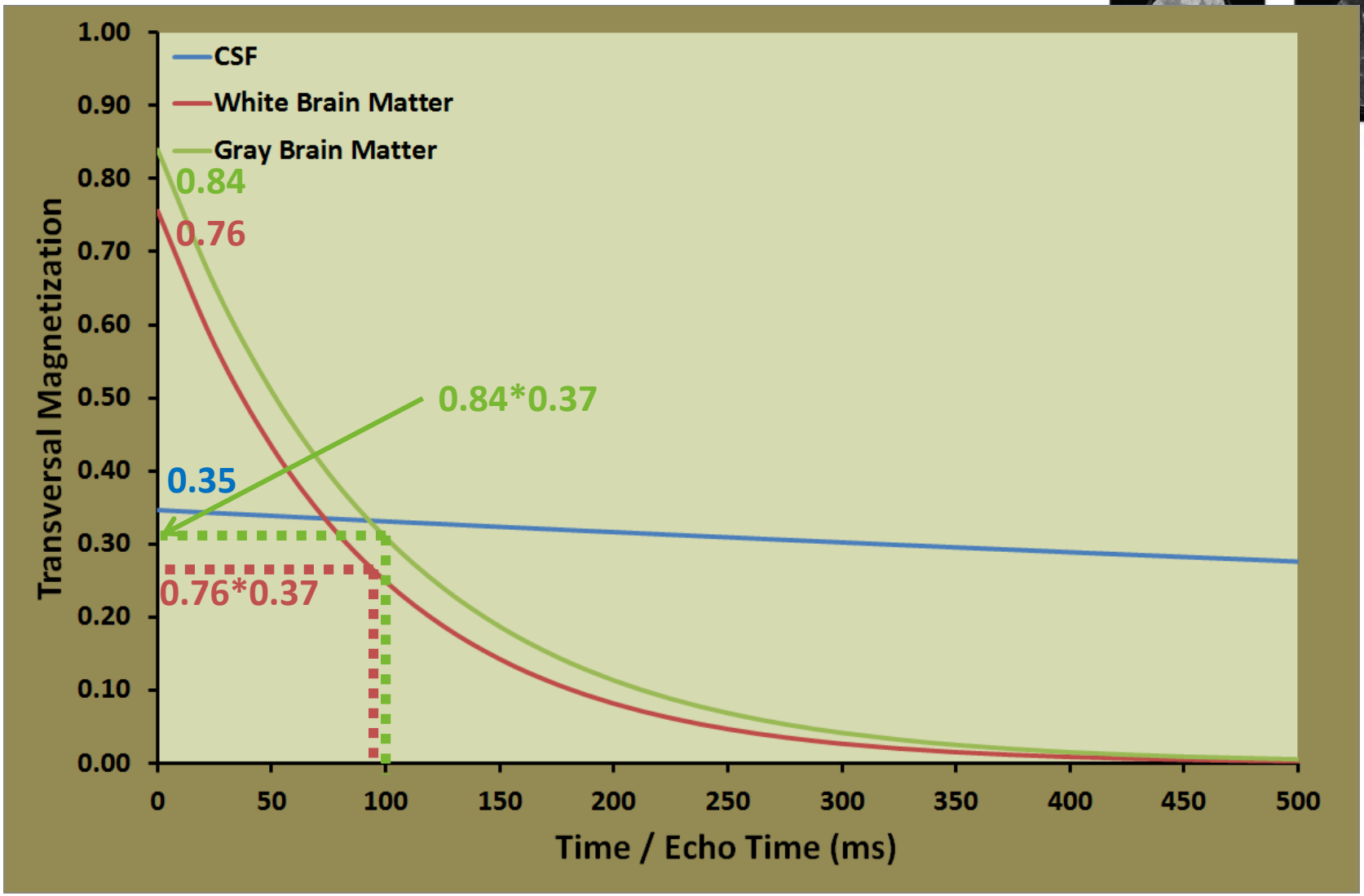
Exercise 2


(a) SE: 2000/25
(b) SE: 2000/50
2000/200

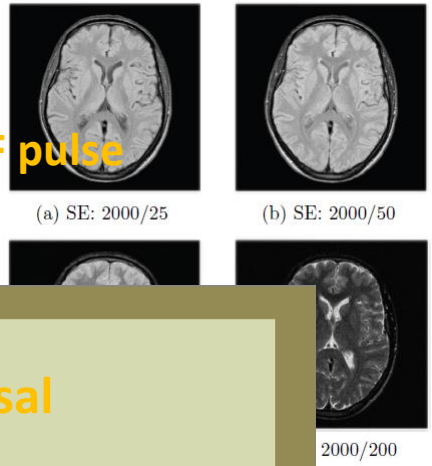

CSF
White Brain Matter
Gray Brain Matter
0.84
0.76
0.35
0.88*(1-0.37)
0.77*(1-0.37)
Longitudinal Magnetization
Time / Repetition Time (ms)
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0
200
400
600
800
1000
1200
1400
1600
1800
2000

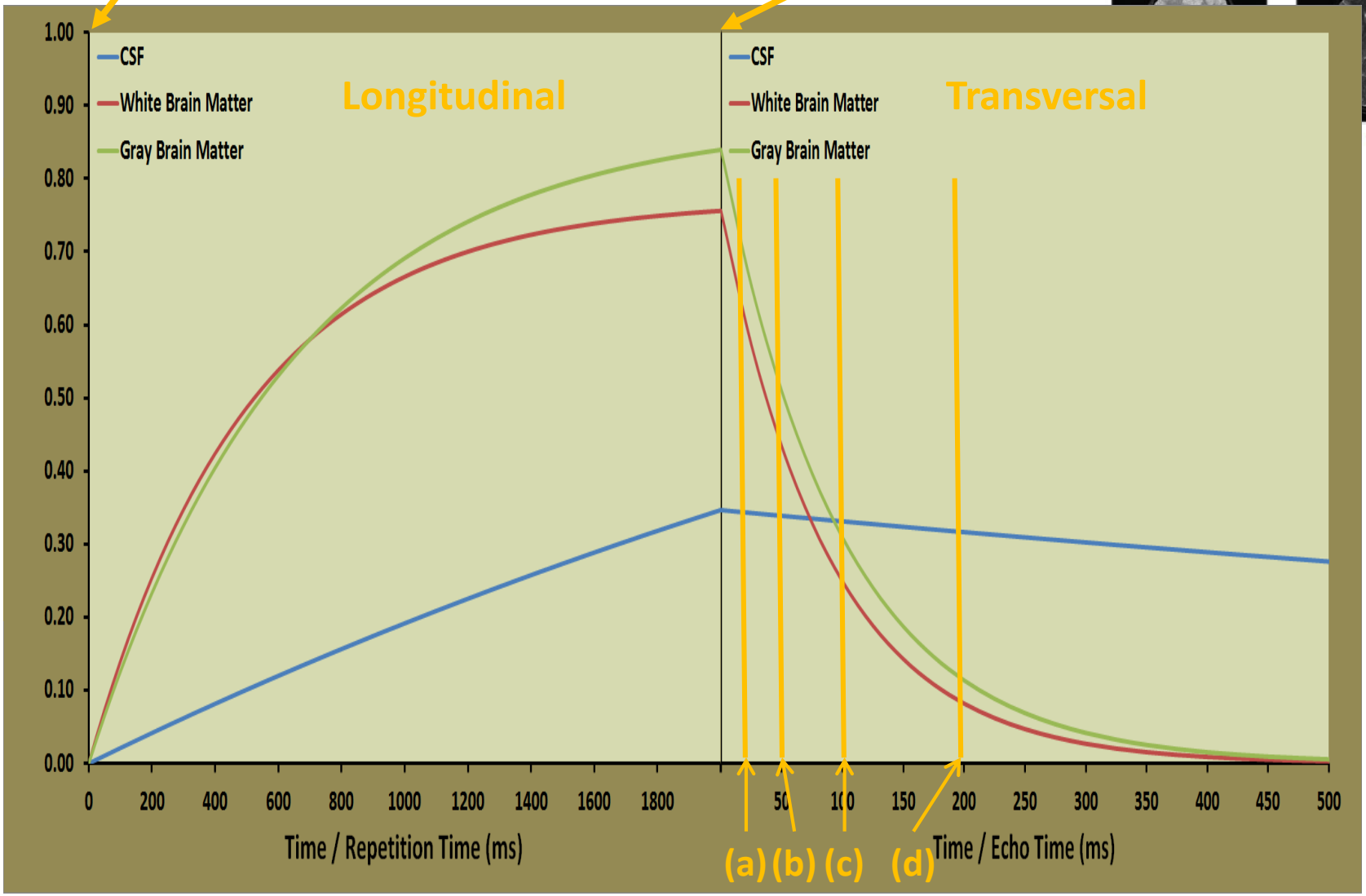
Exercise 2


(a) SE: 2000/25
(b) SE: 2000/50
2000/200


CSF
White Brain Matter
Gray Brain Matter
0.84
0.76
0.84*0.37
0.35
0.76*0.37
Transversal Magnetization
Time / Echo Time (ms)
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0
50
100
150
200
250
300
350
400
450
500

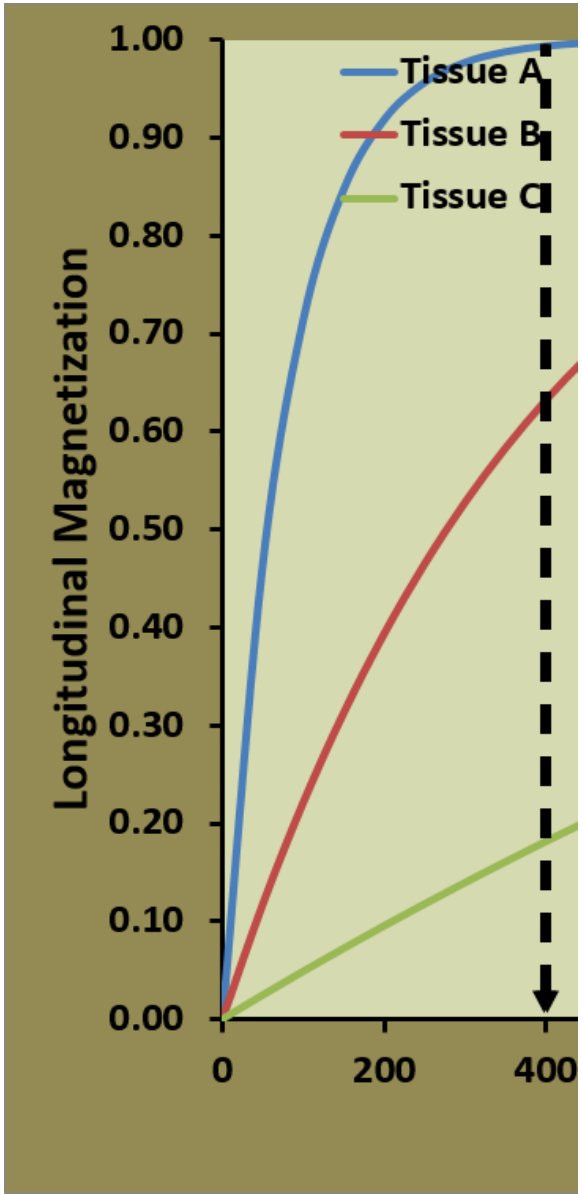
Exercise 2


(a) SE: 2000/25
(b) SE: 2000/50
2000/200

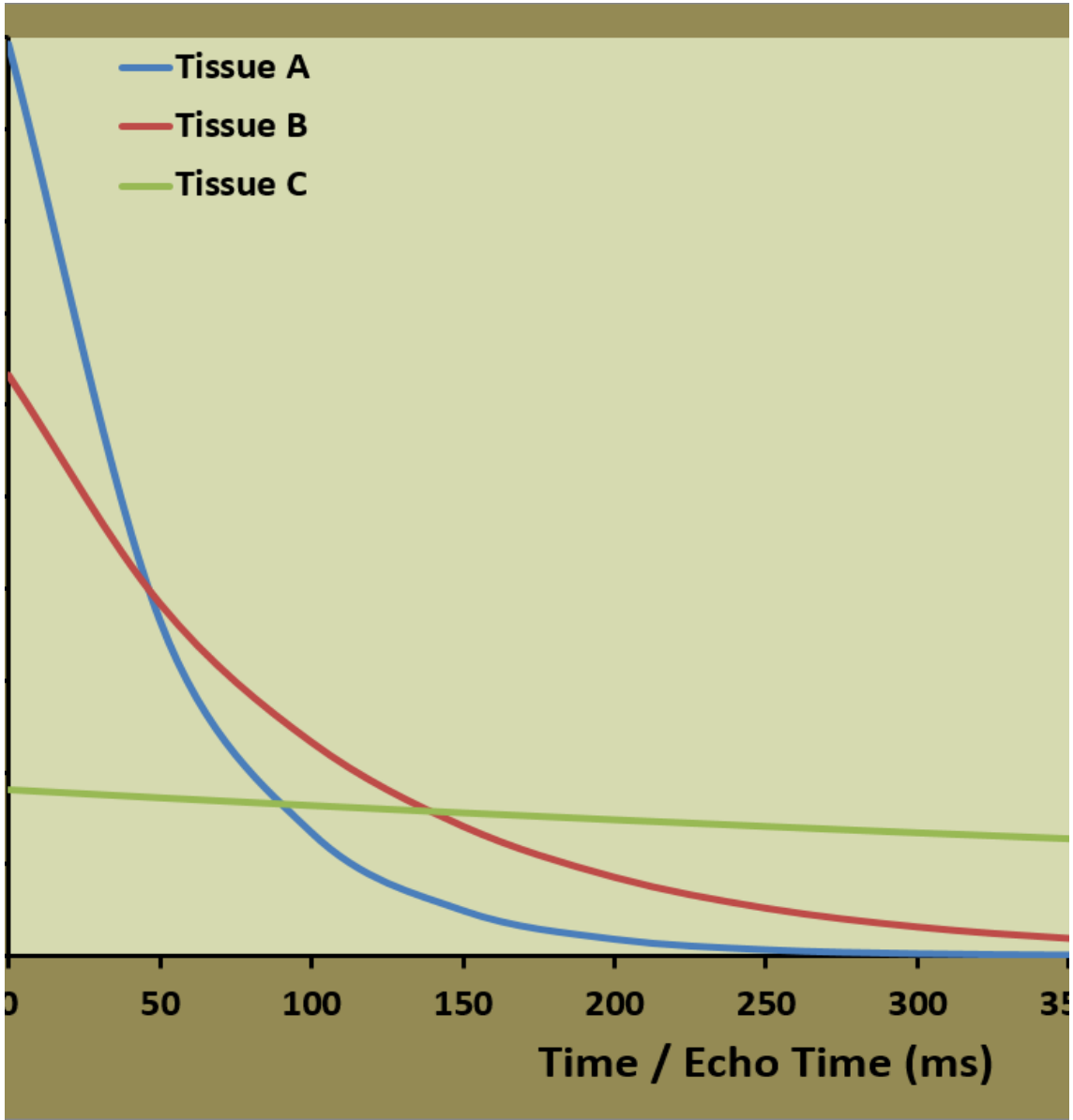

90° RF pulse
90° RF pulse
CSF
White Brain Matter
Gray Brain Matter
Longitudinal
Transversal
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0 200 400 600 800 1000 1200 1400 1600 1800
Time / Repetition Time (ms)
50 100 150 200 250 300 350 400 450 500
Time / Echo Time (ms)
(a) (b) (c) (d)

Exercise 3

(a)



(b)



Exercise 3

(c): No T2-weighting, or TE = 0 ms

Formulas:

- Longitudinal relaxation **(4.26):**
  - $M_l(t) = M_0 \cos\alpha \, e^{-t/T1} + M_0(1 - e^{-t/T1})$
  - with α=90° and M_0 = relative proton density
- Acquisition Time **(4.68)**:
  - $TA_{2D} = \frac{N_{ph} * TR}{ETL}$

-Tissue A: 0.99*M0; Tissue B: 0.63*M0; Tissue C: 0.18*M0

- Acq time = 240 * 0.4 s = 96 s

Exercise 3

(d): TE = 50 ms

Formulas:

- Transverse relaxation **(4.31):**
  - $s(t) = M_0(1 - e^{-TR/T1})e^{-t/T2}$
- Acquisition Time **(4.68)**:
  - $TA_{2D} = \frac{N_{ph} * TR}{ETL}$

-Tissue A: 0.37*M0; Tissue B: 0.38*M0; Tissue C: 0.17*M0

- Acq time = 240 * 0.4 s = 96 s

Exercise 4

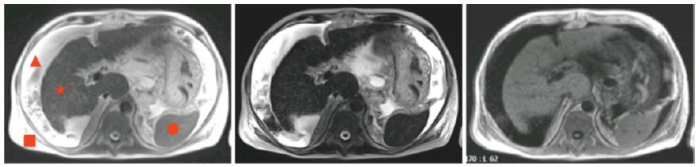


Figure 1.8

- Assumptions:

	T1 (ms)	T2 (ms)	PD (relative)
Liver	550	50	1
Spleen	750	80	1
Fat	200	100	1
Peritoneal Water	3500	2000	1

- Formulas:
  - Longitudinal relaxation **(4.26):**
    - $M_l(t) = M_0 \cos\alpha \, e^{-t/T1} + M_0(1 - e^{-t/T1})$
    - with α=90° and M_0 = relative proton density
  - Transverse relaxation **(4.31):**
    - $s(t) = M_0(1 - e^{-TR/T1})e^{-t/T2}$

Exercise 4

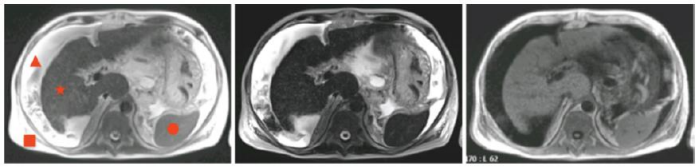
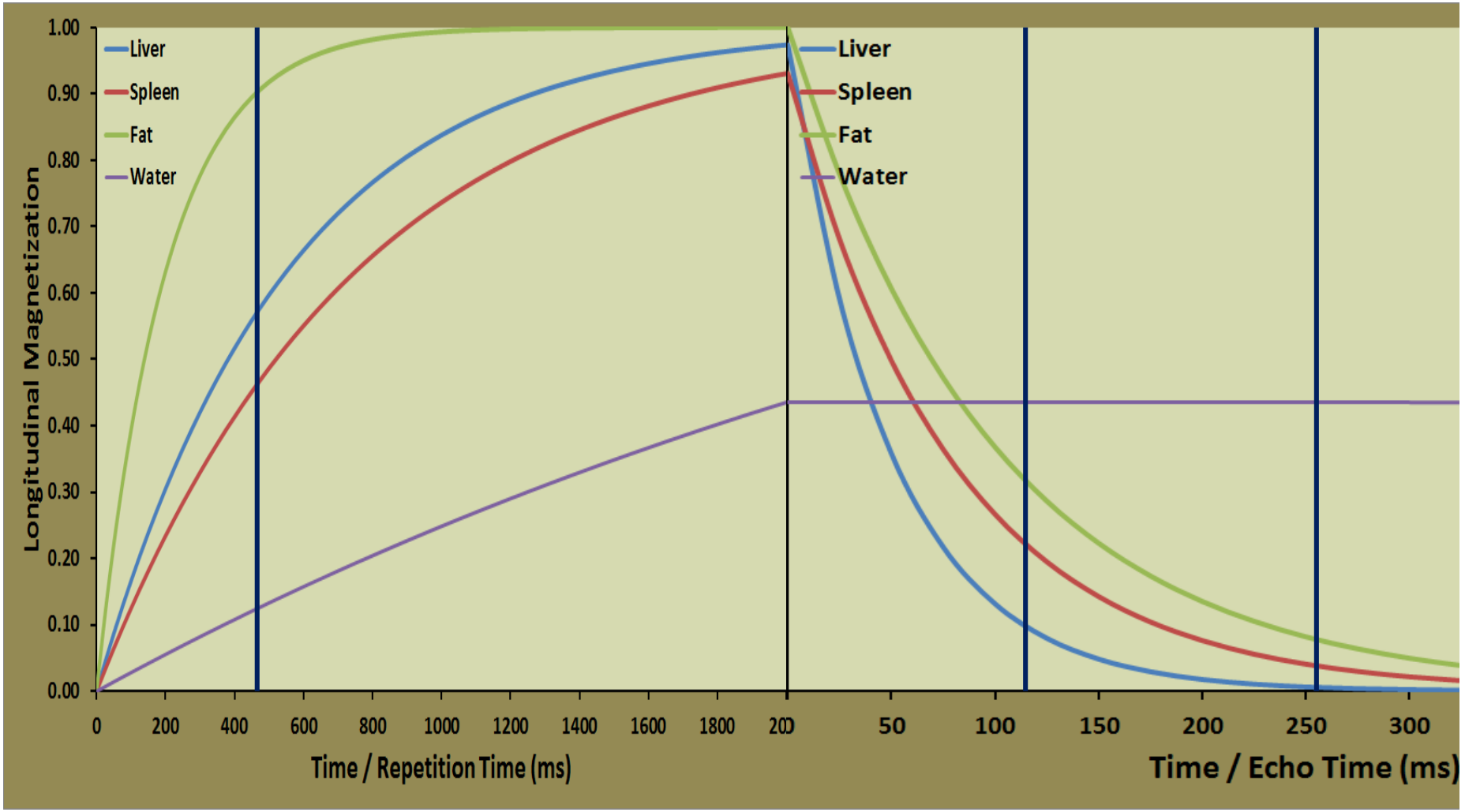


Figure 1.8

c

a

b



Exercise 4

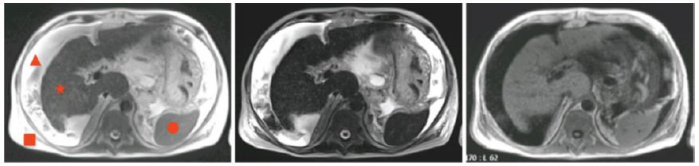


Figure 1.8

➔Approximate values!!

- Image a: TR=2000ms, TE=120ms (Early T2-weighted)
- Image b: TR=2000ms, TE=280ms (Late T2-weighted)
- Image c: TR=450ms, TE=10ms (as short as possible) (T1-weighted)

Exercise 5

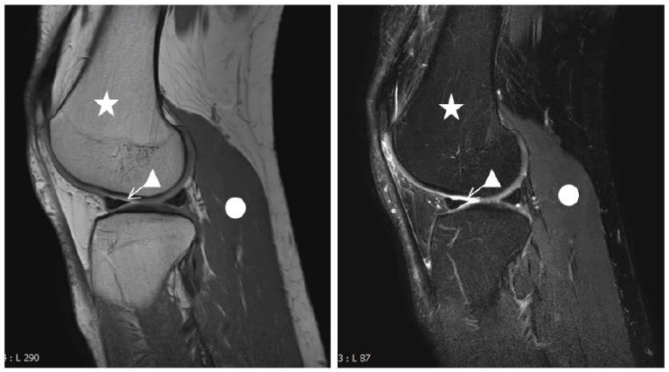


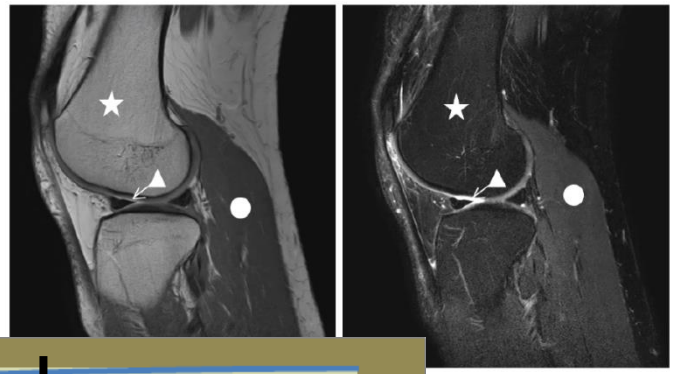
Figure 1.9

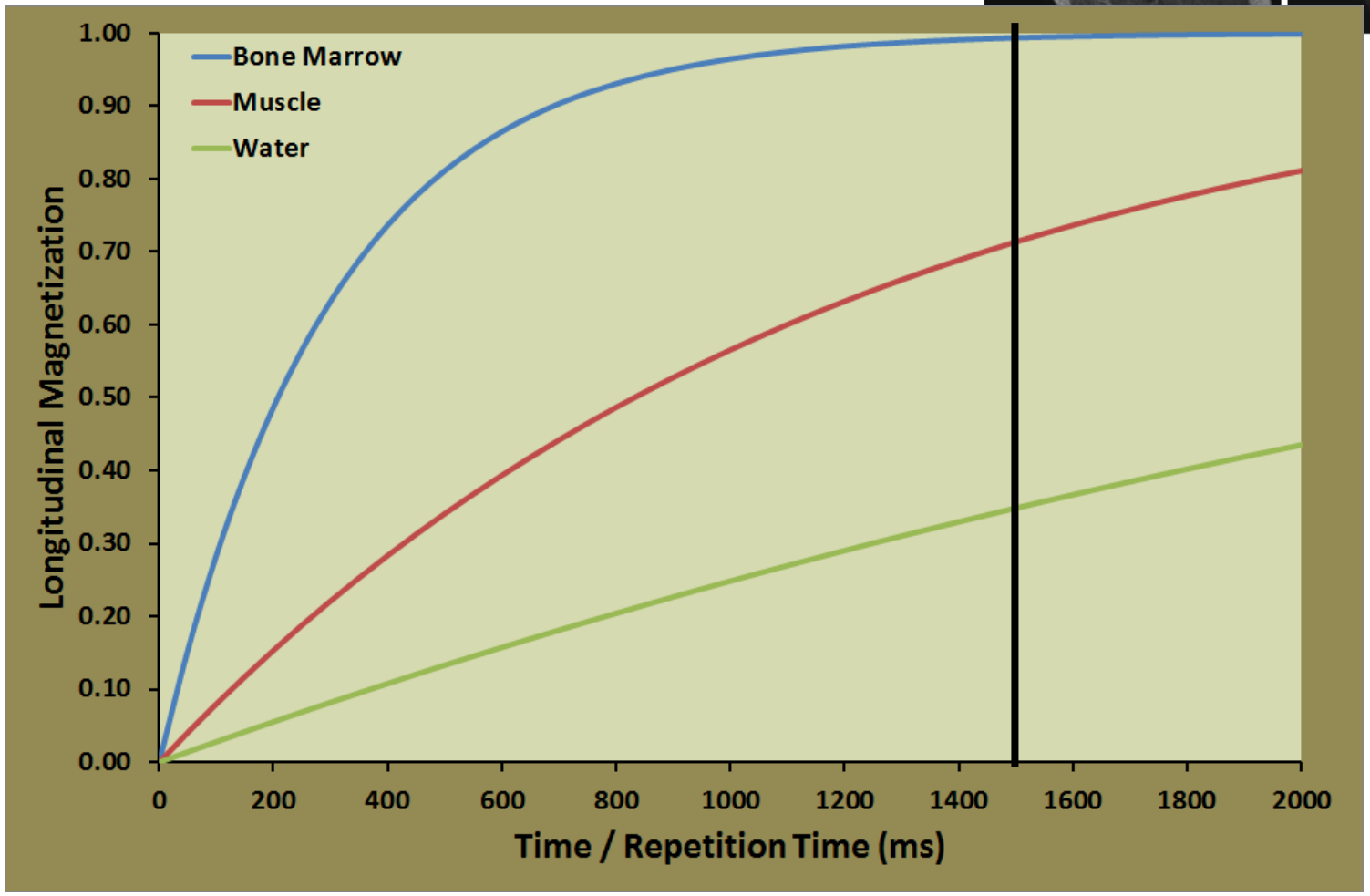
- Assumptions:
  - TR = 1500 ms, 90° excitation pulse

	T1 (ms)	T2 (ms)	PD (relative)
Bone Marrow	300	100	1
Muscle	1200	120	1
Water	3500	2000	1

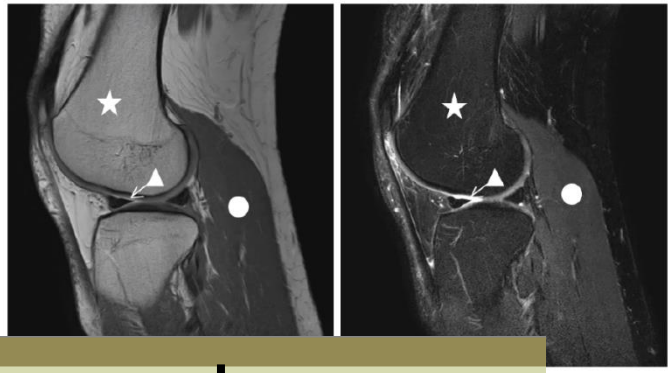
- Formulas:
  - Longitudinal relaxation **(4.26):**
    - $M_l(t) = M_0 \cos\alpha \, e^{-t/T_1} + M_0(1 - e^{-t/T_1})$
    - with α=90° and M_0 = relative proton density
  - Transverse relaxation **(4.31):**
    - $s(t) = M_0(1 - e^{-TR/T_1})e^{-t/T_2}$

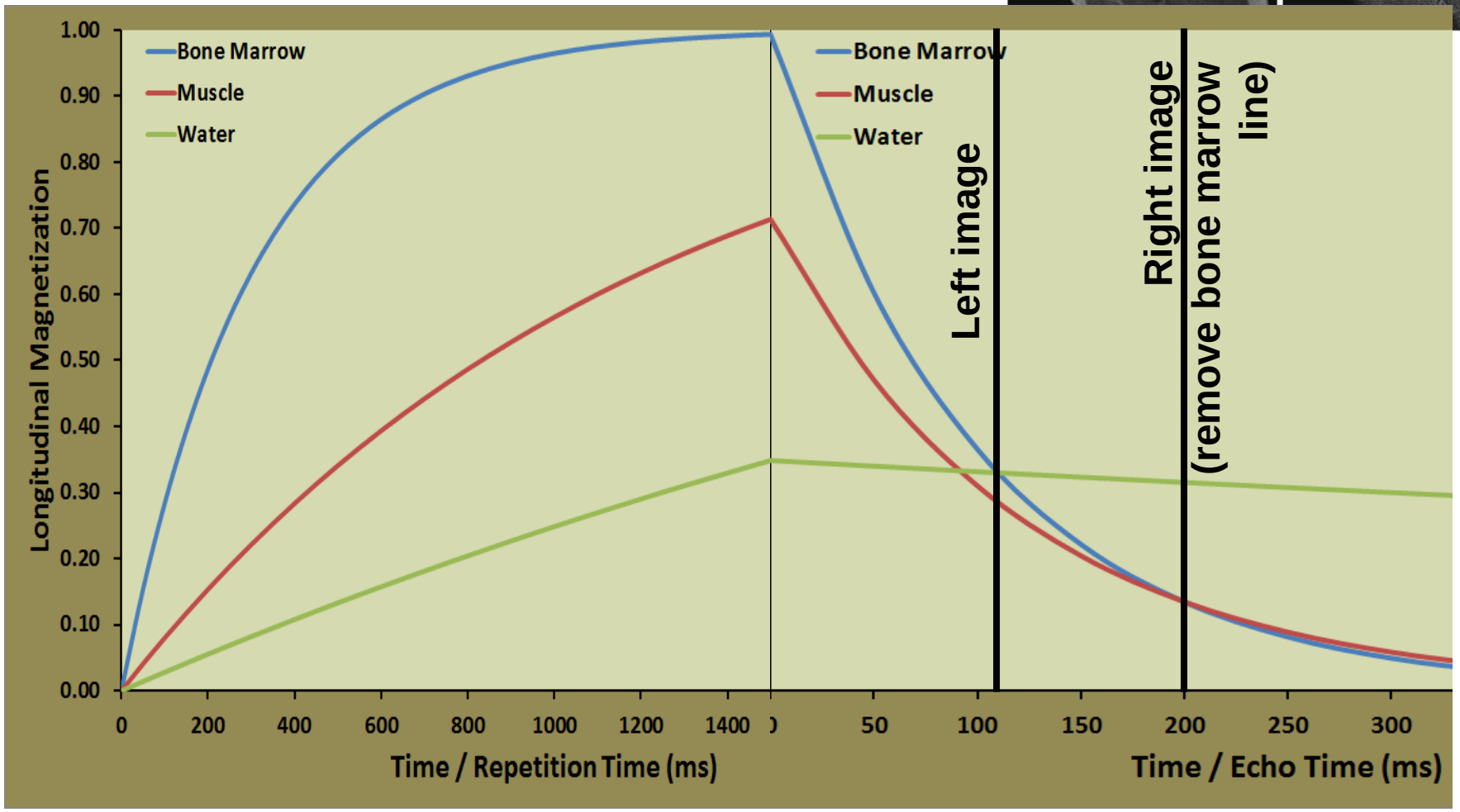
Exercise 5




Bone Marrow
Muscle
Water
Longitudinal Magnetization
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0
200
400
600
800
1000
1200
1400
1600
1800
2000
Time / Repetition Time (ms)

Exercise 5




Bone Marrow
Muscle
Water
Longitudinal Magnetization
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0
200
400
600
800
1000
1200
1400
Time / Repetition Time (ms)
Bone Marrow
Muscle
Water
Left image
Right image
(remove bone marrow line)
0
50
100
150
200
250
300
Time / Echo Time (ms)

Exercise 6

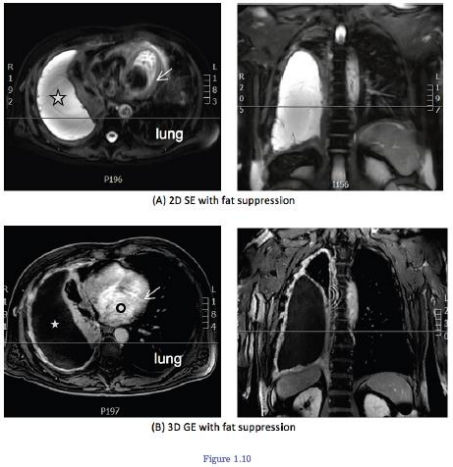
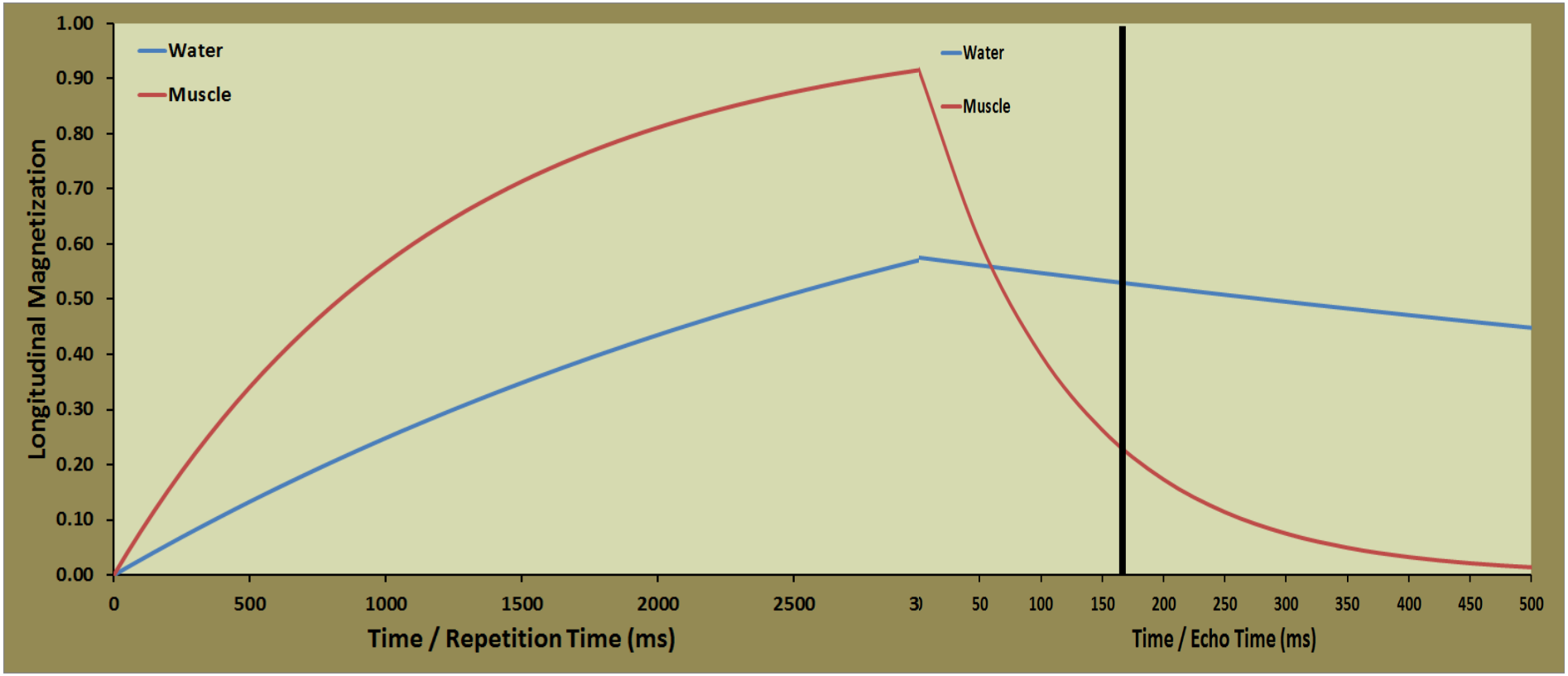


Figure 1.10

(a) Water bright, muscle grey (T2w)

(normal lungs is mostly air → no signal)



Exercise 6

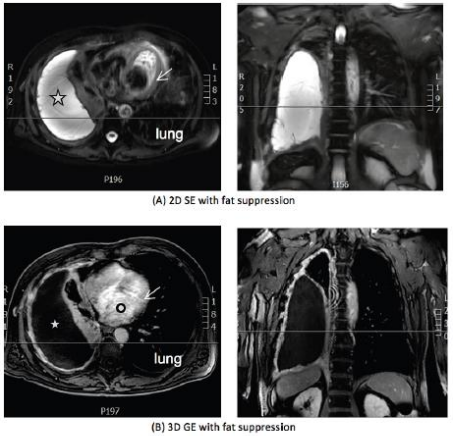
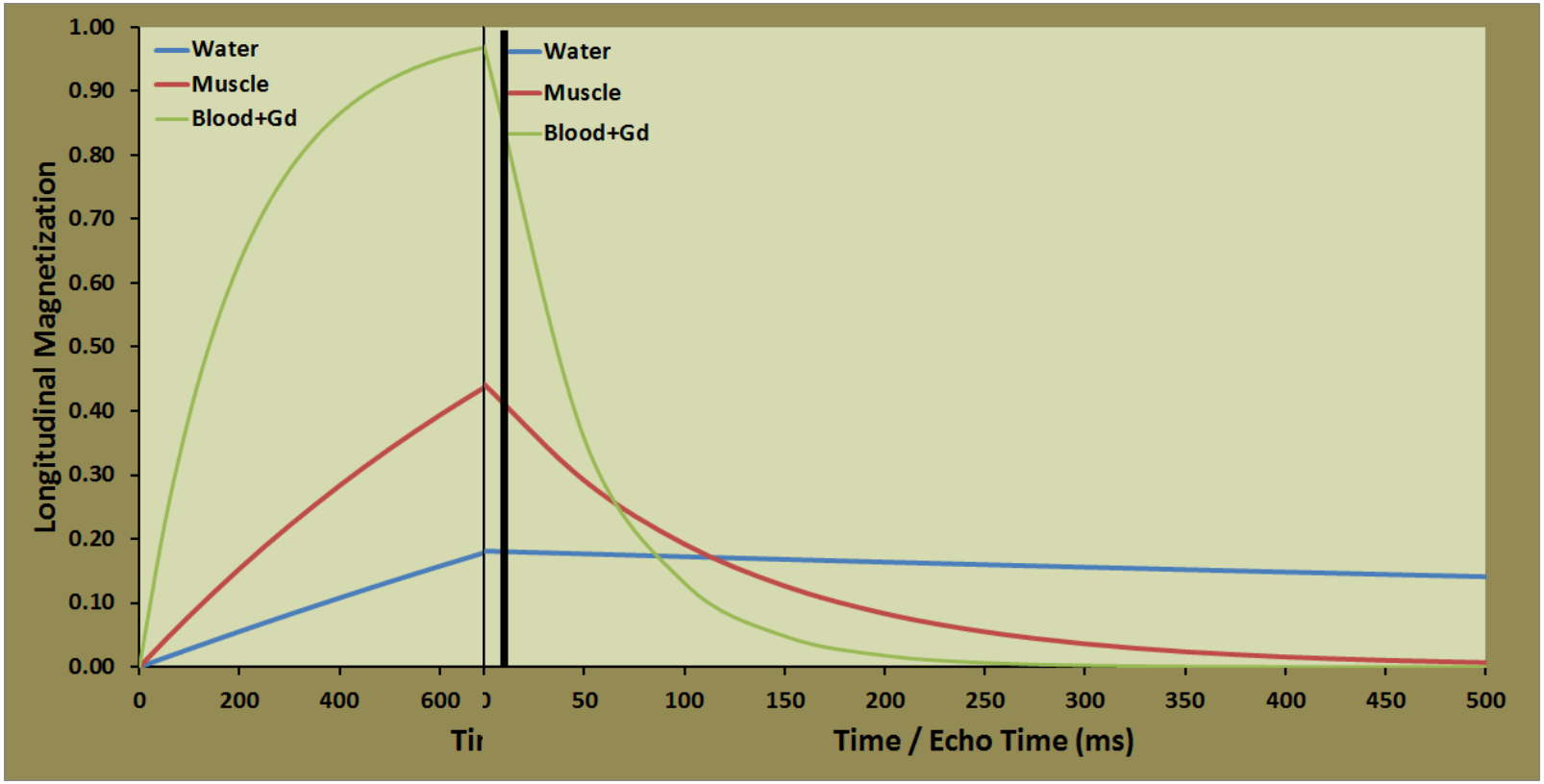


Figure 1.10

(b) Water dark, muscle hyperintense (T1w)

(blood with Gadolinium has very short T1 and T2!)



Exercise 7

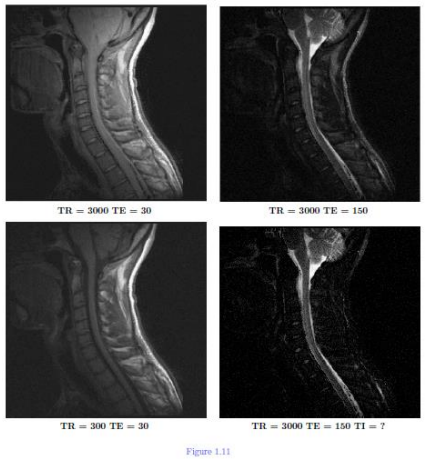


Figure 1.11

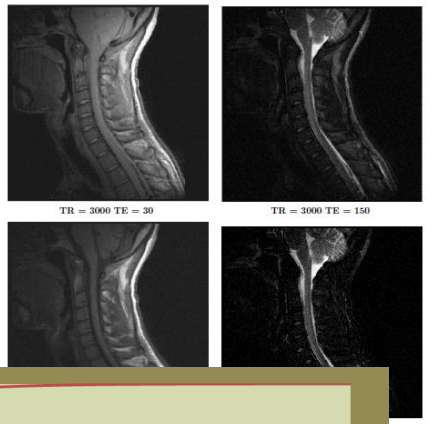
- Assumptions:
  - Fat: T1 = 200 ms, T2 = 100 ms, relative proton density 1.
  - CSF: T1 = 4700 ms, T2 = 2200 ms, relative proton density 1.
- Formulas

  Longitudinal relaxation:

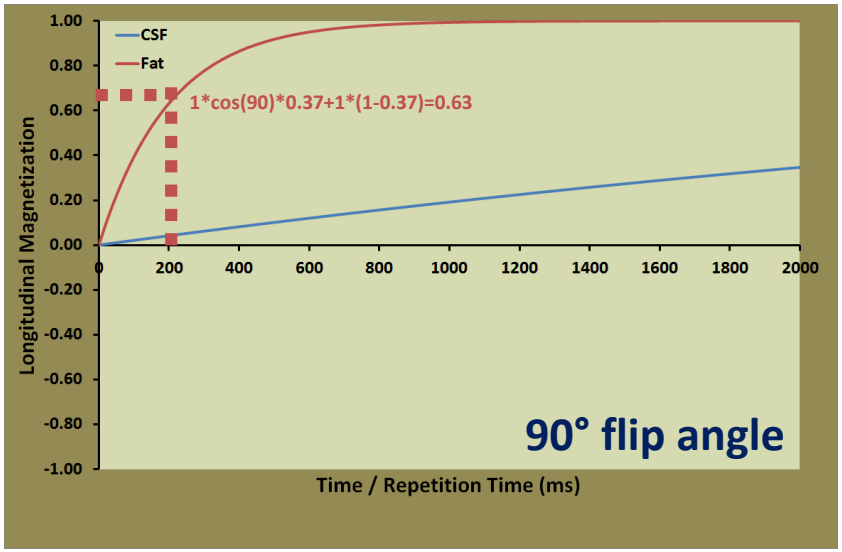
  (4.26) $M_l(t) = M_0 \cos\alpha \, e^{-t/T1} + M_0(1 - e^{-t/T1})$

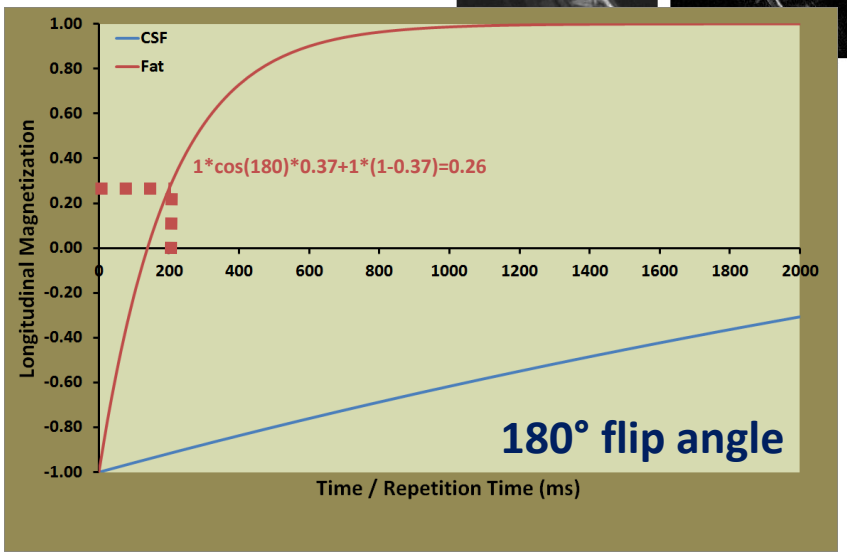
  with α=90°/180° and M_0 = relative proton density

Exercise 7



a)

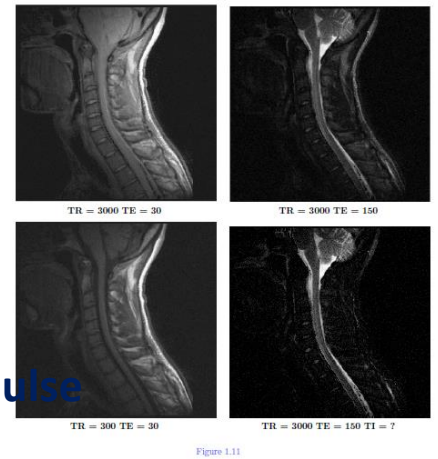




b) $M_l(TI) = 1\cos 180° \, e^{-TI/200} + 1\left(1 - e^{-TI/200}\right) = 0$

➔ $\left(1 - 2e^{-TI/200}\right) = 0$ ➔ TI = 138.63 ms

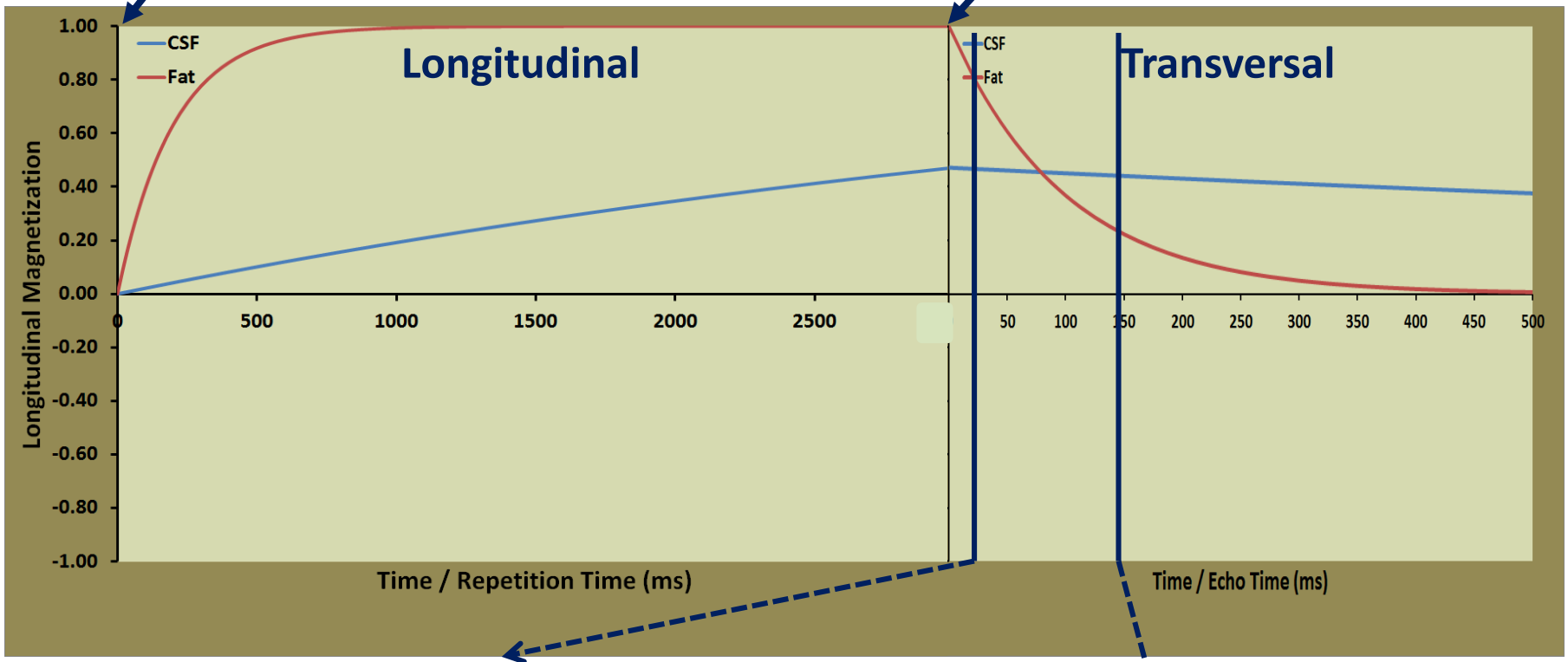
Exercise 7



c)-d)

90° RF pulse

90° RF pulse



TR 3000, TE 30, Fat brighter than Water)

TR 3000, TE 150, Water brighter than Fat)

Exercise 7

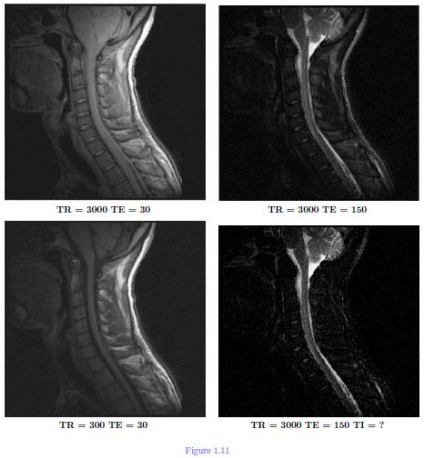
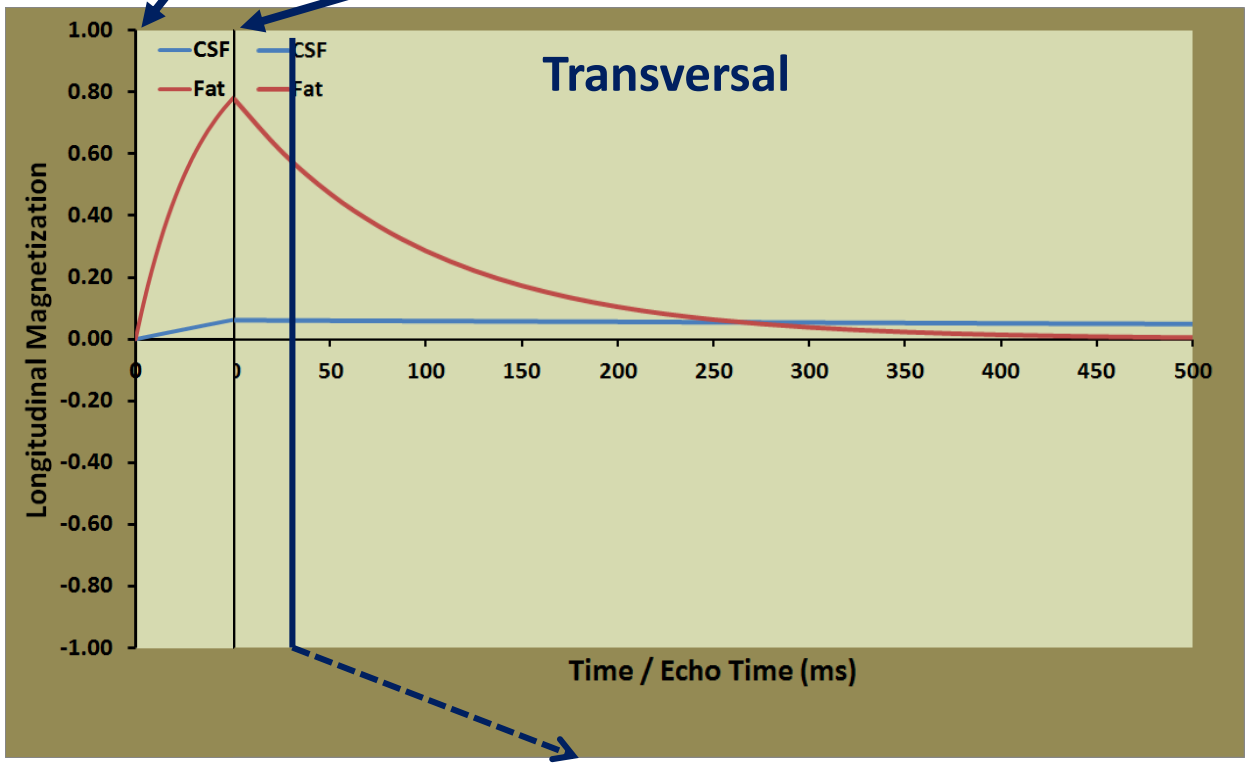


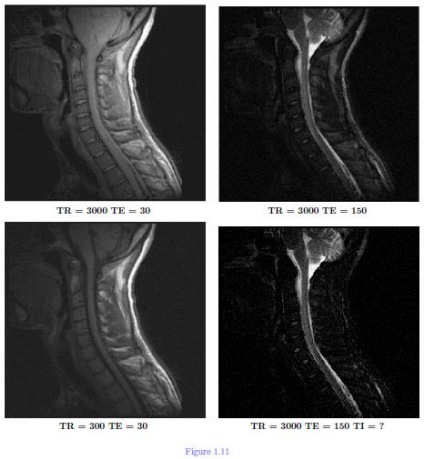
Figure 1.11

c)-d) **90° RF pulse** **90° RF pulse**

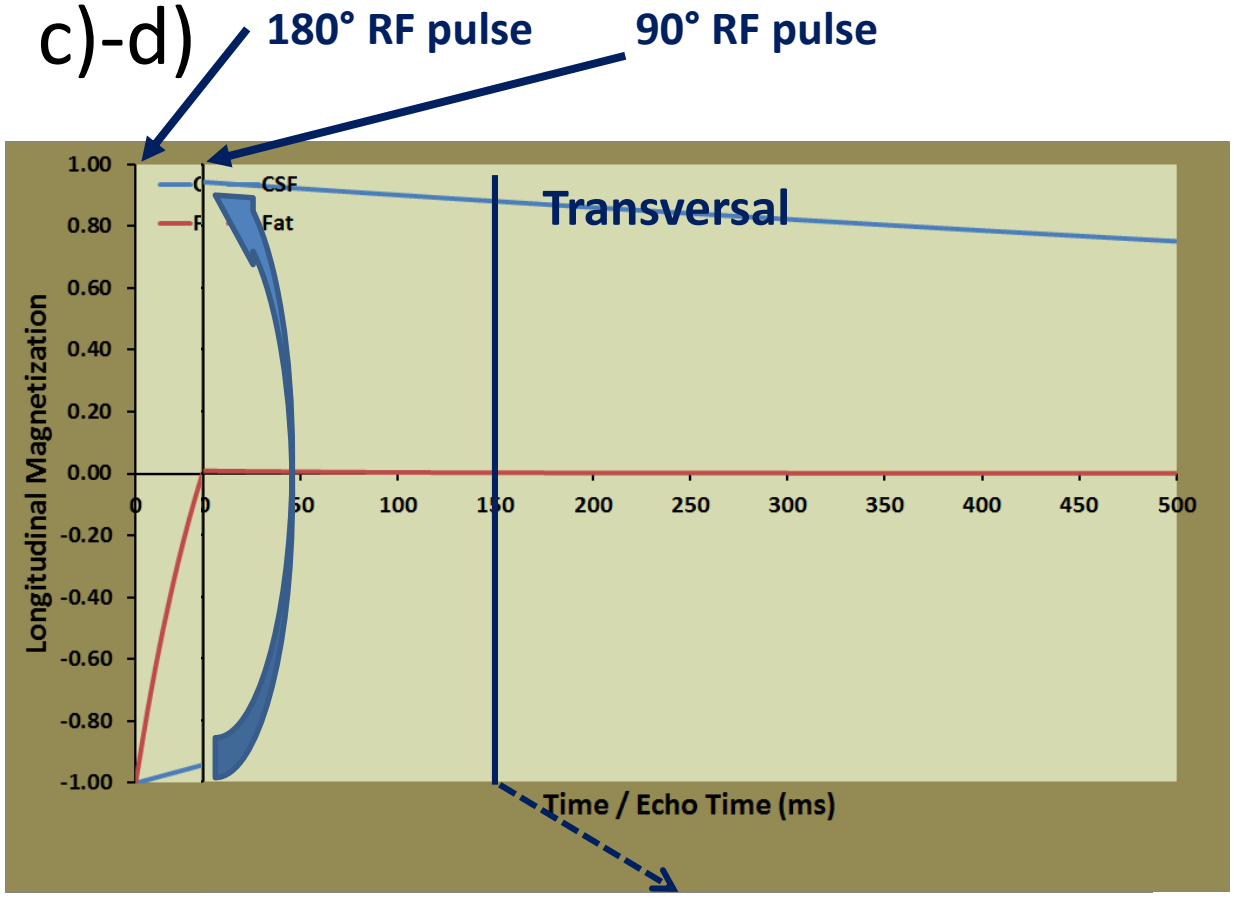


TR 300, TE 30, Fat bright, Water very dark)

Exercise 7

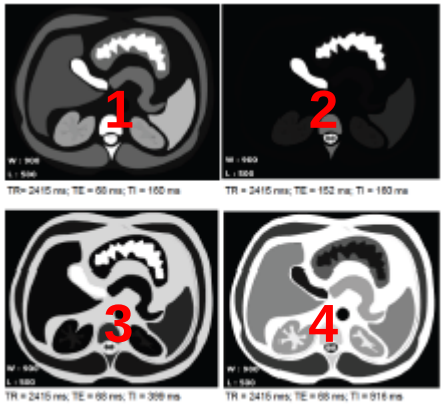


c)-d)



TR 3000, TE 150, TI 138.6, Water very bright, no Fat visible)

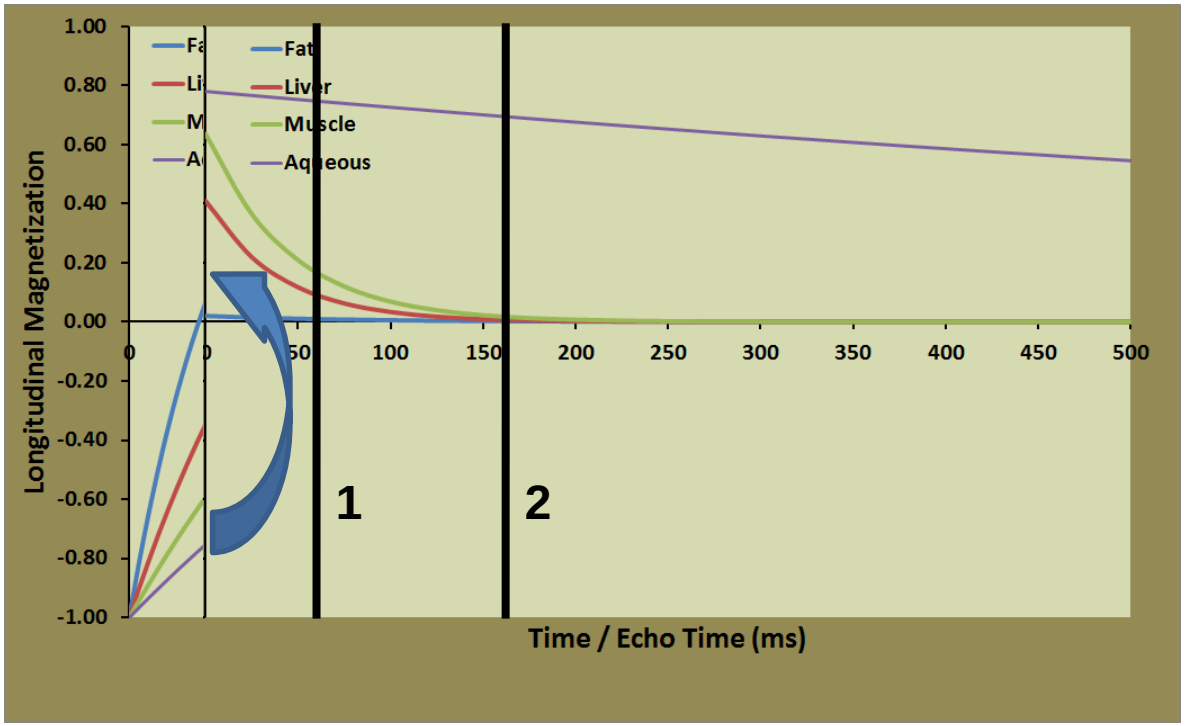
Exercise 8



a) $M_l(TI) = 1 \cos 180° \, e^{-TI/260} + 1\left(1 - e^{-TI/260}\right) = 0$ Figure 1.12

➔ $\left(1 - 2e^{-TI/260}\right) = 0$ ➔ TI = 180.22 ms

b)-c)



(Liver: 346.6 ms
Muscle: 603 ms
Aqueous: 1039.7 ms)

Exercise 8

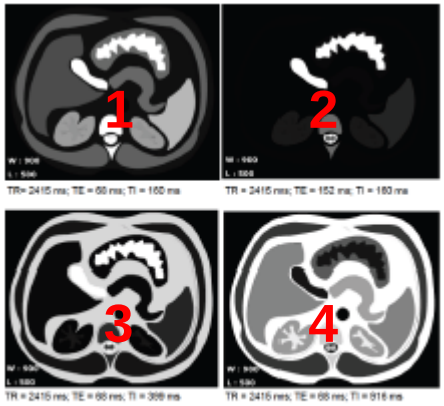
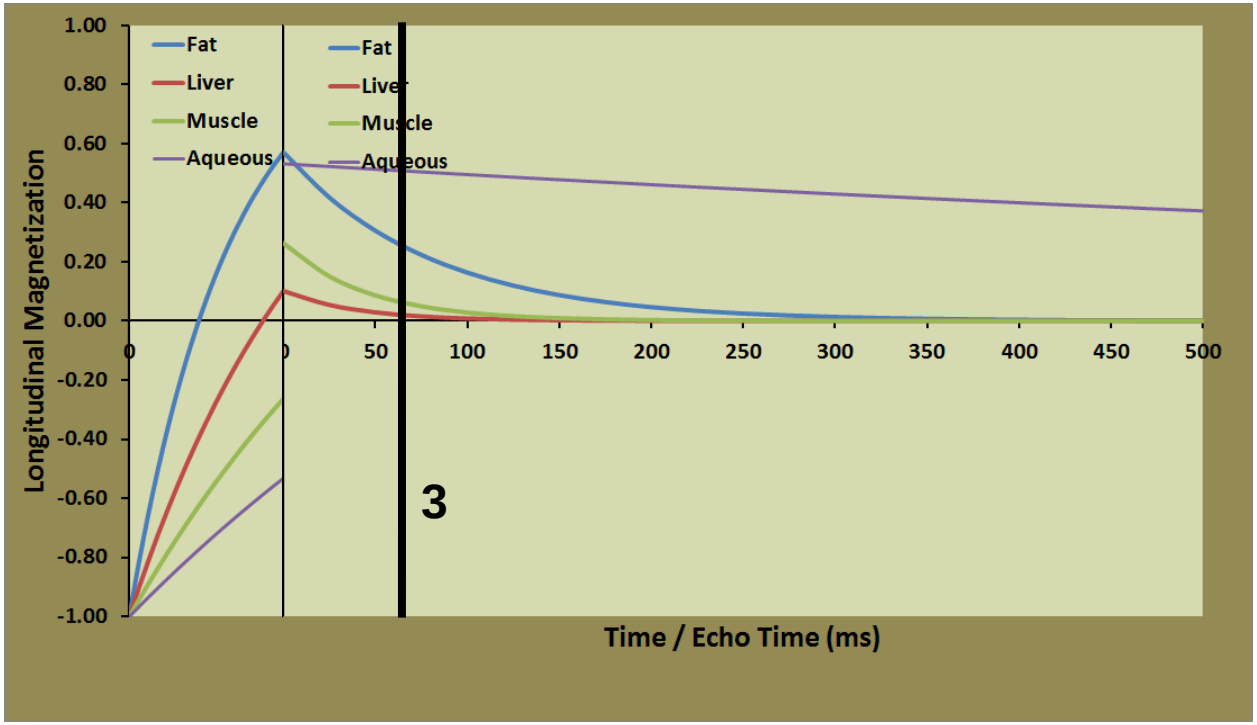


Figure 1.12

b)-c)



Exercise 8

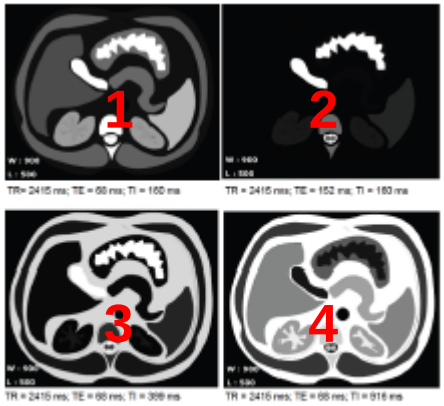
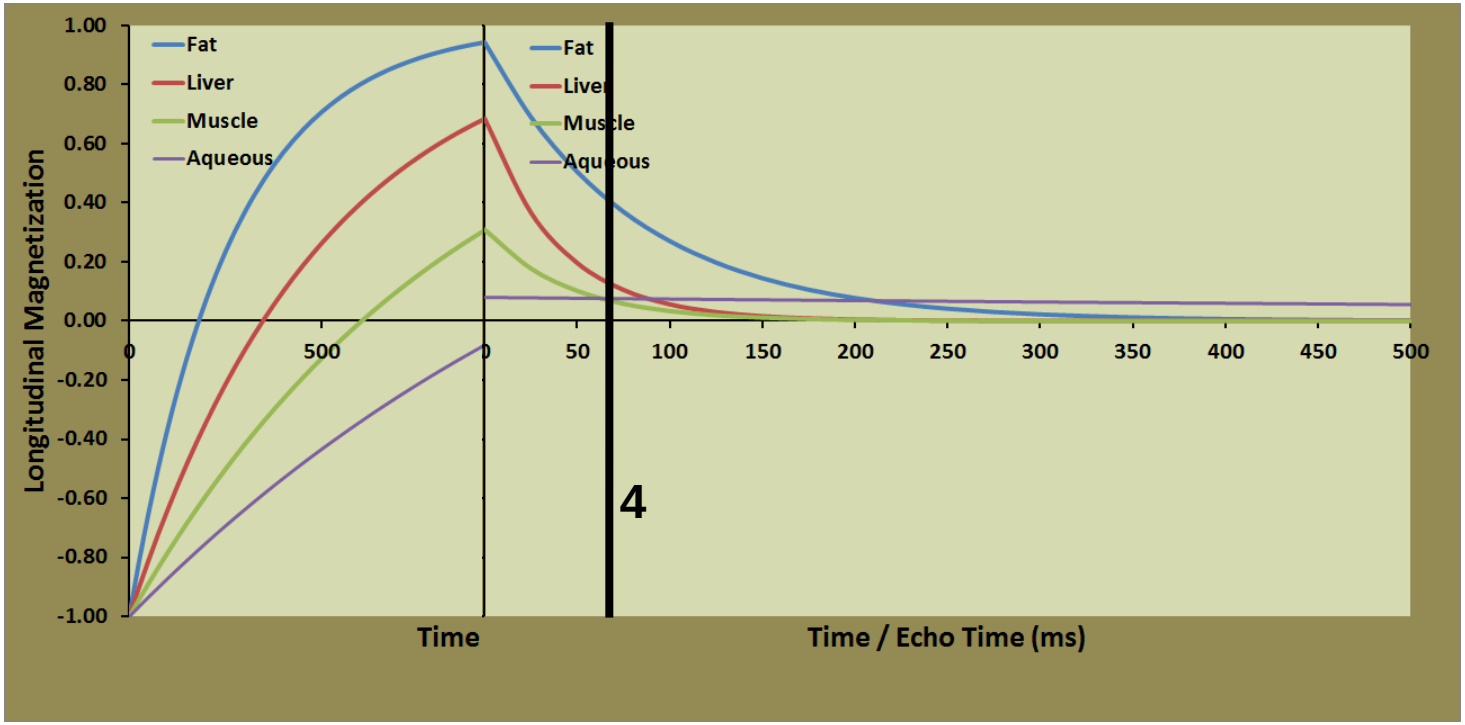


Figure 1.12

b)-c)



Exercise 8

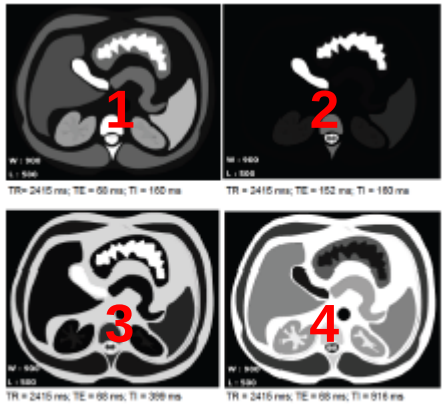


Figure 1.12

d) Changing 1.5T to 3T:

- Increased SNR (approx. 70%-100%)
- Increased T1 (contrasts change, so sequences need to be adapted)
- Higher RF deposition needed for excitation/refocusing
- Larger rotational differences between fat and water → chemical shift artefacts are larger
- …

Exercise 9

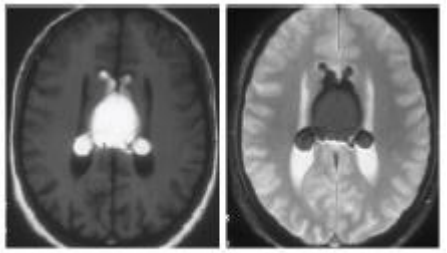


Figure 1.13

- Image on the left:
  - Water is black, subcutaneous fat is bright, so the image is T1-weighted; lesion is bright
- Image on the right:
  - Water is bright, subcutaneous fat is dark, so the image is T2-weighted; lesion is dark

➔ Lesion behaves like fat in both images, so will probably contain a lot of fat (lipoma, teratoma, dermoid cyst, …)

Exercise 10

(a) Formula:

(4.6) $\omega = \gamma B$, with γ/2π=42.57 MHz/T

➔ω=42.57 MHz/T * 1.5 T = <u>63.855 MHz</u>

(b) Gradients apply differences in magnetic field strength according to position, but the magnetic field itself is always in the Z direction (feet-head direction)!

➔$B = (0,0,B_0 + x*G_x + y*G_y + z*G_z)$

Exercise 11

- Formula:

  (4.6) $\omega = \gamma B$, with γ/2π=42.57 MHz/T

  B = B_0 + 0.1 (m) * 10 (mT/m) = 1.5 T + 1 mT = 1.501 T

  ➔frequency = 42.57 * 1.501 = <u>63.9 MHz</u>

- Formula **(4.35)**:

  $BW = \gamma G_z \Delta z$

  ➔BW = 42.57 (MHz/T) * 10 (mT/m) * 2 (mm)

  = <u>851.4 Hz</u>

Exercise 12

(a)-(c): See text book.

For more information, you can also read:

http://mriquestions.com/slice-selective-excitation.html

https://www.math.upenn.edu/~cle/notes/selpls.pdf

Exercise 13

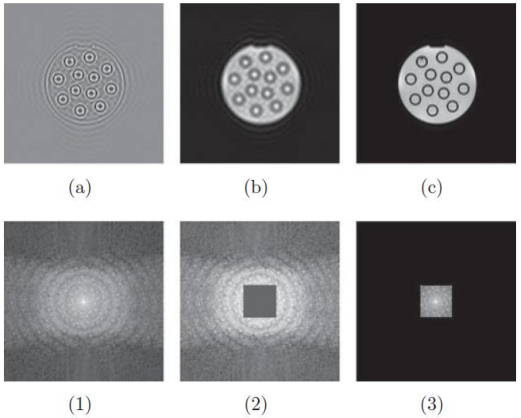


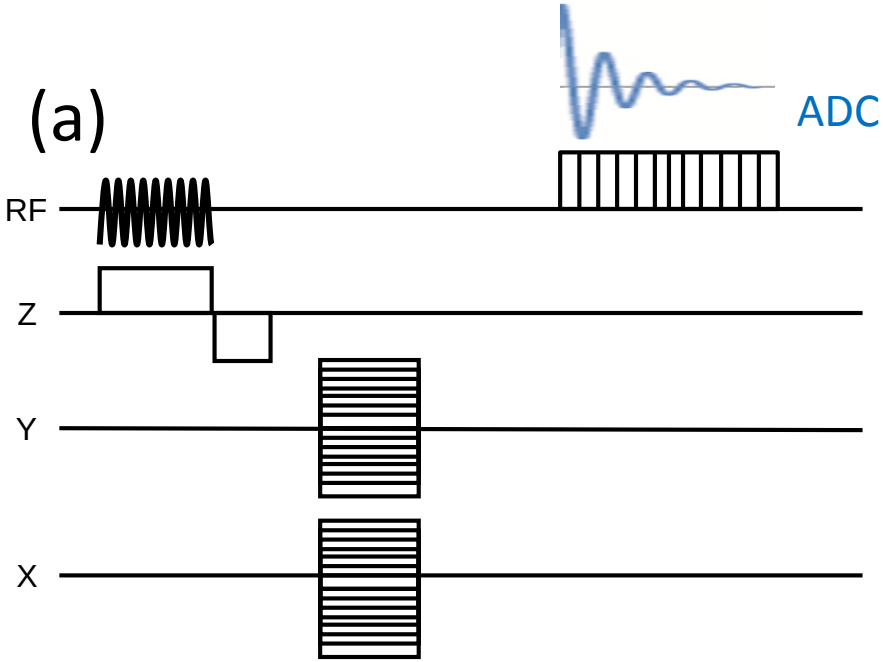
Figure B.17

(a)

- (c) + (1): Complete K-space filling gives the best image quality.
- (b) + (3): Filled centre of K-space provides the good image contrast, but high frequency lines are missing, causing blurring.
- (a) + (2): K-space centre not filled, so bad image contrast, but high frequencies are filled in, causing sharp edges.

Exercise 14

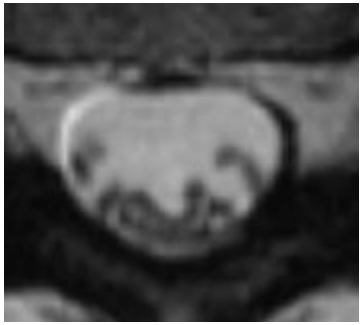
(a)



ADC readout of the FID

One iteration for each x,y pixel

(b) Chemical shift artefact: misregistration due to different rotation speed of different tissues.



Exercise 14

(c) As each point in K-space is measured by one choice of x and y gradients, followed by a complete readout (without frequency gradient), they are all phase encodings, so:

(4.68): $TA_{2D} = \frac{N_{ph} * TR}{ETL}$

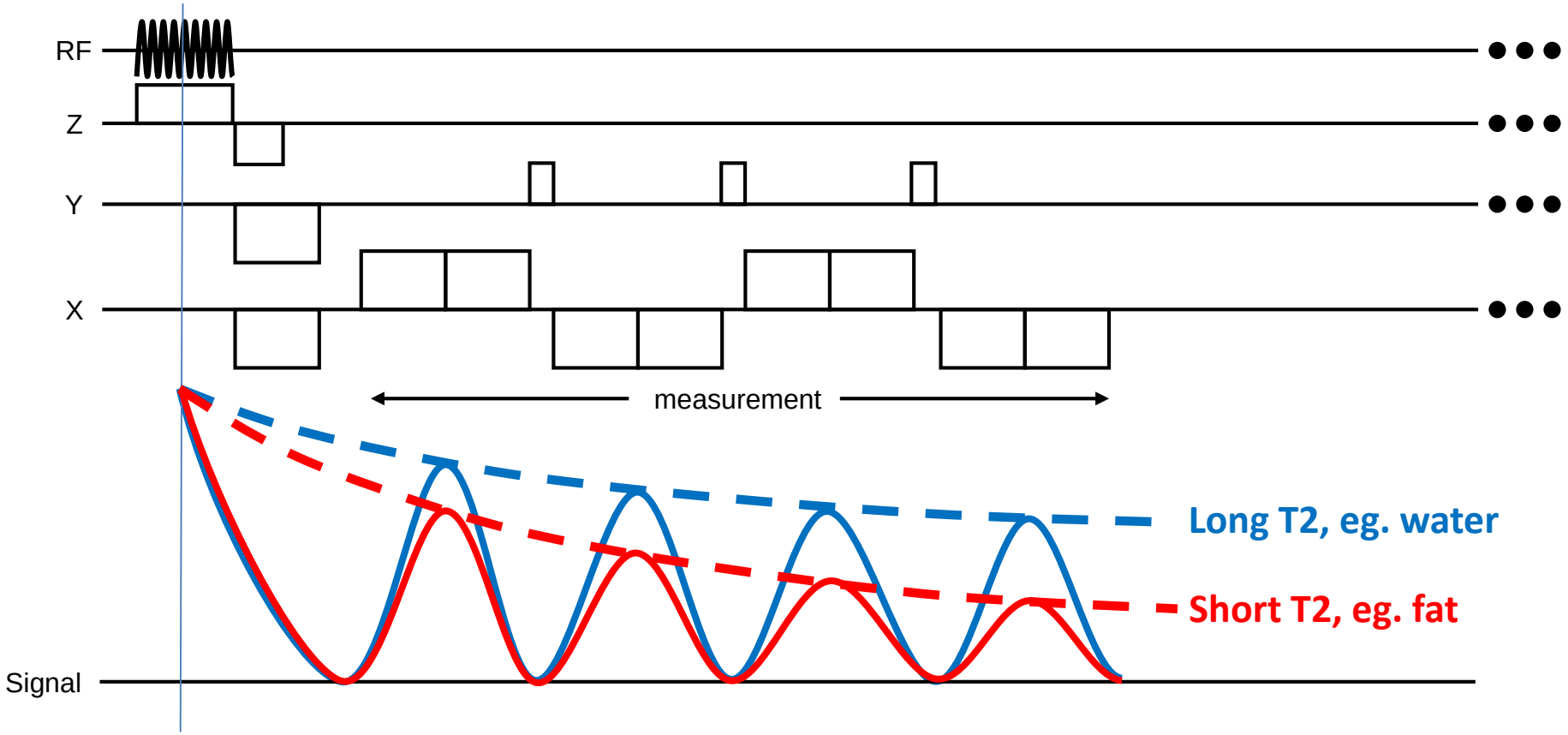
becomes:

$$TA_{CS} = \frac{N_x * N_y * TR}{ETL}$$

Therefore, for a single slice of a basic 2D CSI with 16 x 16 pixels (common choice), and a TR of at least 2s (no T1 or T2 contrast should be inside), and an ETL of 1, the TA = 16 * 16 * 2s = 512s (approx 8.5 min)

Exercise 15

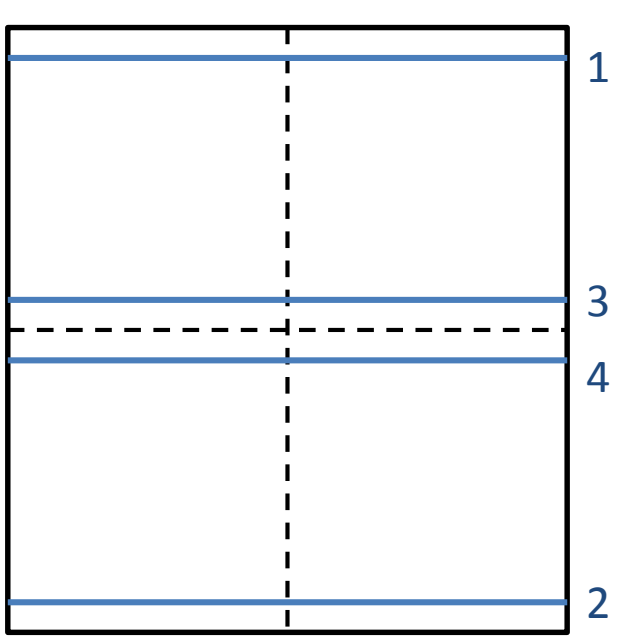
(a) 4 lines in K-space measured instead of one

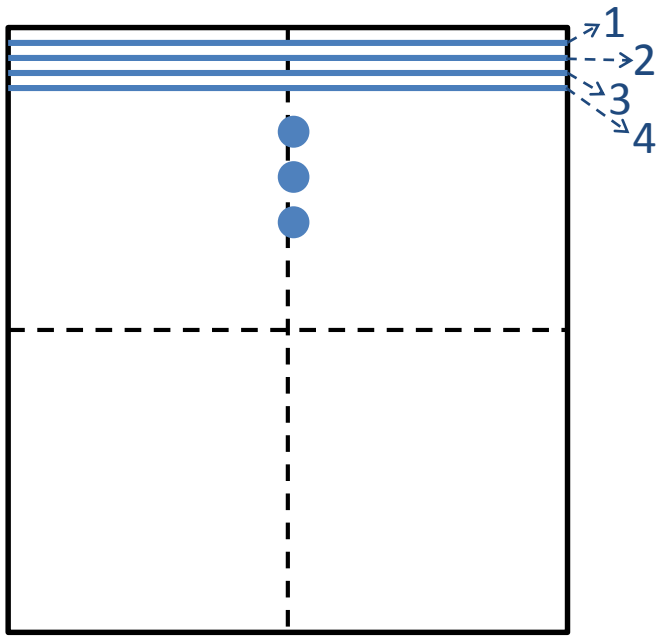


➔different signal intensities and slightly different contrasts combined into one image.

Exercise 15

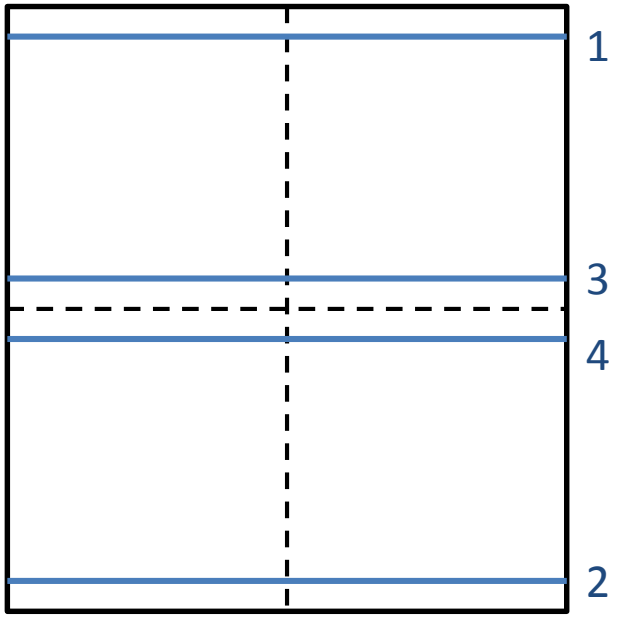
(a) Many possible k-space fill schemes are possible, such as:





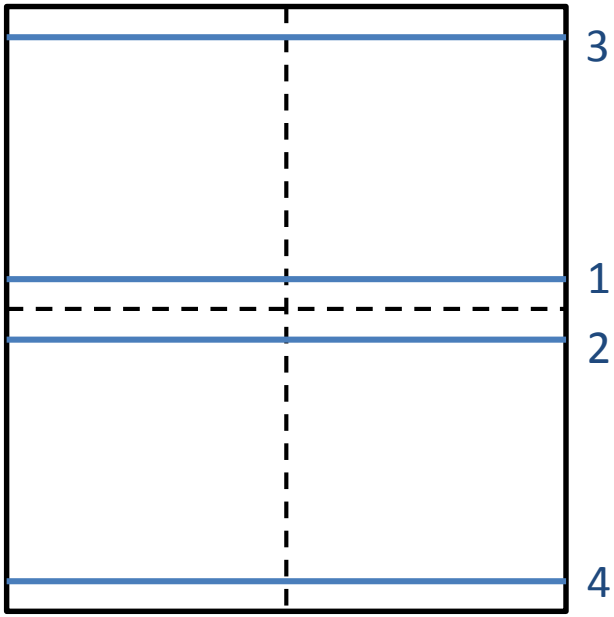
Exercise 15

(b) Effect of different filling:



Highest signal lines on the outside

→good resolution
→less contrast



Highest signal lines in the center

→good contrast
→lower resolution (“blurring”)

Exercise 16

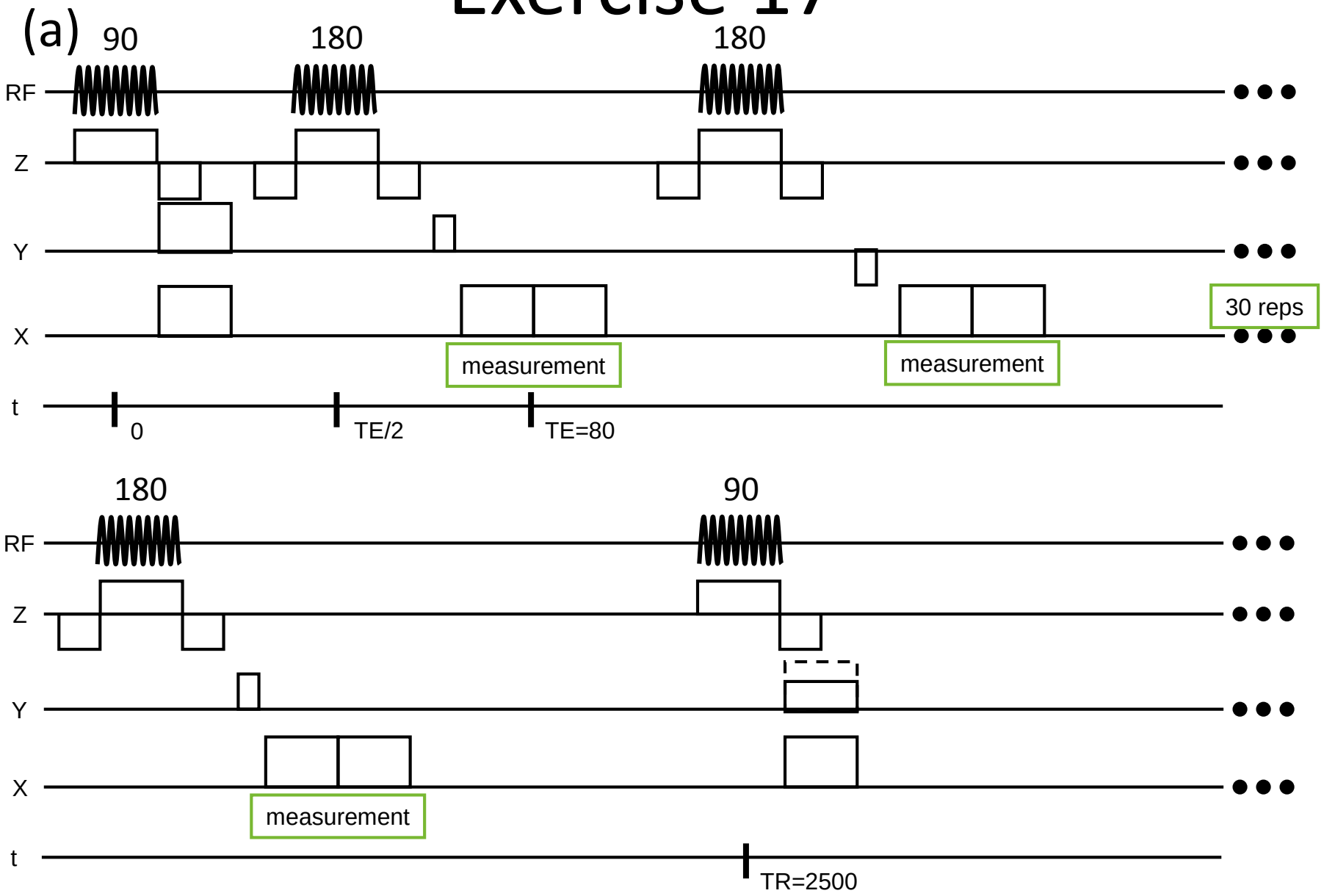
(a) - x = 200/384, y = 150/245, z = 5mm, so pixel resolution will be 0.52 x 0.61 x 5 mm

- acq time = 245 x 0.522s = 127.89s

(b) T1-weighted sequence, so water=dark, fat=bright (see Ex 4-c).

(c) Alternate slices within the same TR (principle of concatenations, see exercise session)

Exercise 17


(a)
90
180
180
RF
Z
Y
X
t
30 reps
measurement
measurement
0
TE/2
TE=80
180
90
RF
Z
Y
X
t
measurement
TR=2500

Exercise 17

(b) - pixel size = 200/128 x 200/256 x 5 = 1.56 x 0.78 x 5mm

- acq time = (128 x 2.5s)/30 = 10.67s (in practice 12.5s)

(c) Images will have some contrast blurring, so CNR will be lower. Also signal blurring so lower apparent resolution.

(d) acq time = **4**x 10.67 = 42.67s (in practice 4x 12.5=50s)

SNR multiplied by sqrt(**4**), so double SNR (CNR should be same effect)

Exercise 18

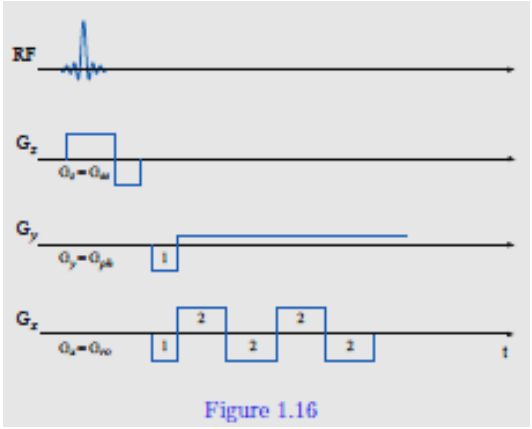
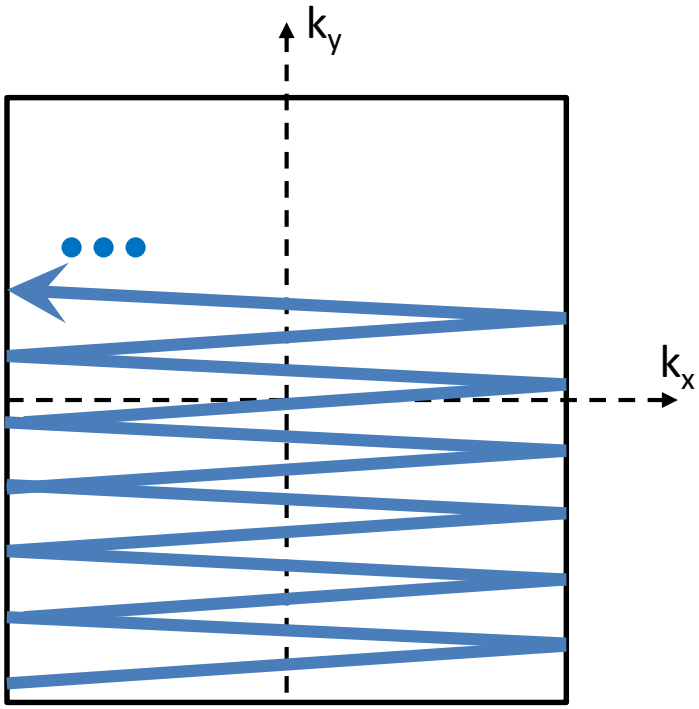


Figure 1.16



Exercise 19

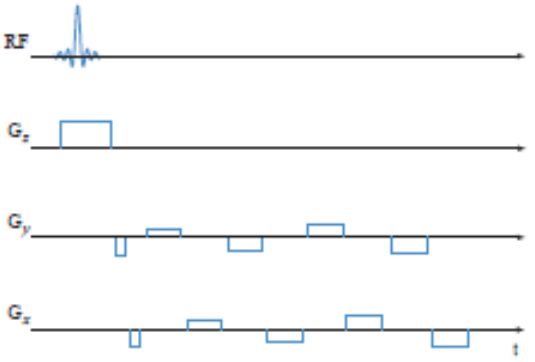
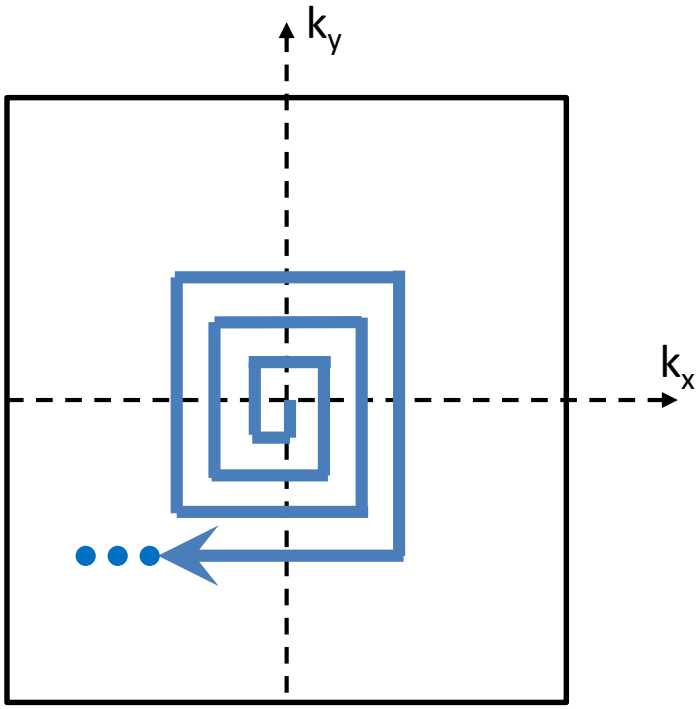


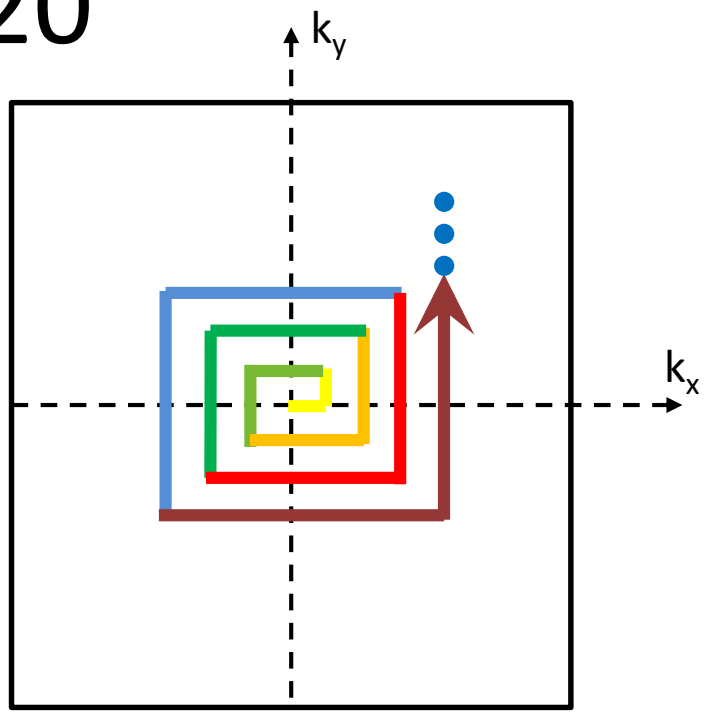
Figure 1.17

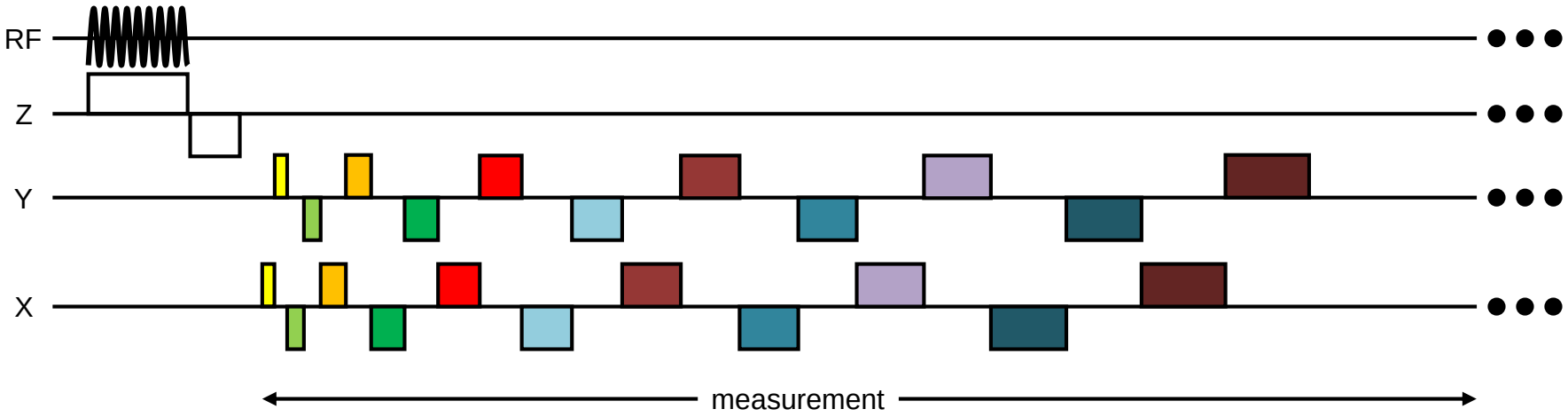
-



- Premature stop would mean only the centre of k-space is filled in, so good contrast but bad resolution (blurred!)

Exercise 20





Exercise 21

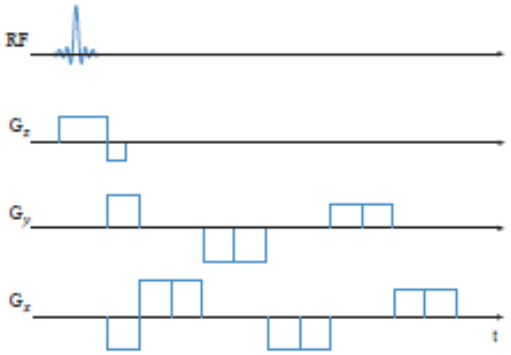
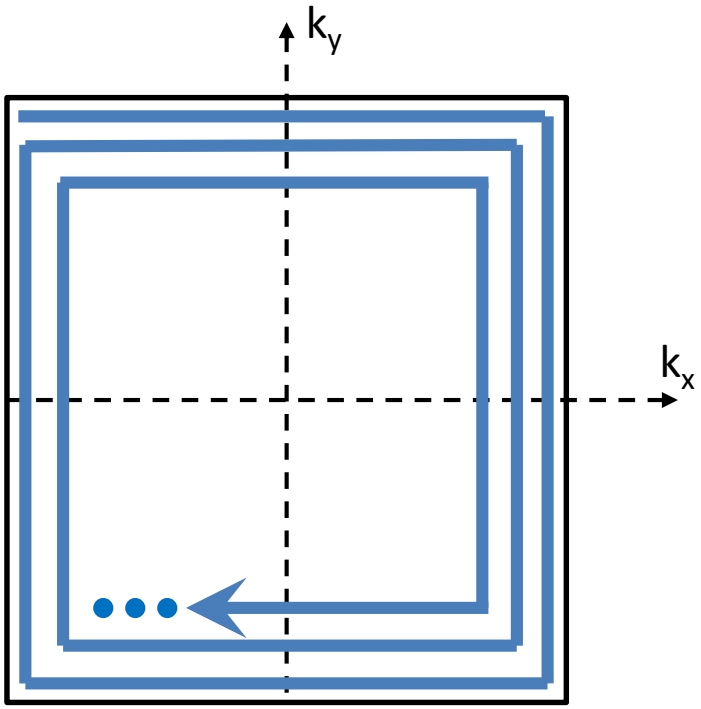


Figure 1.19

-



- Premature stop would mean only the edge of k-space is filled in, so good resolution (sharp), but the contrast will be gone!

Exercise 22

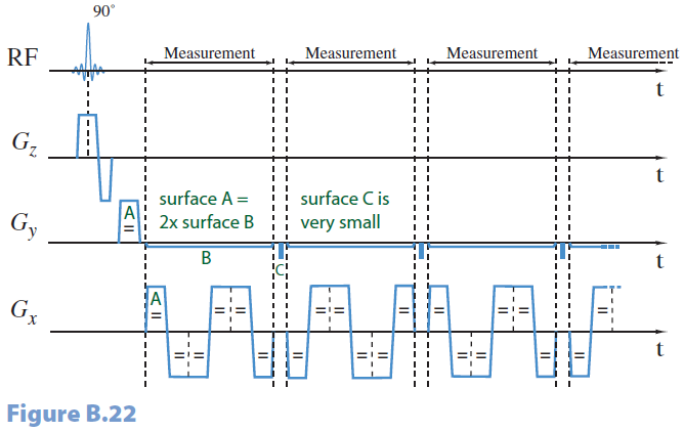
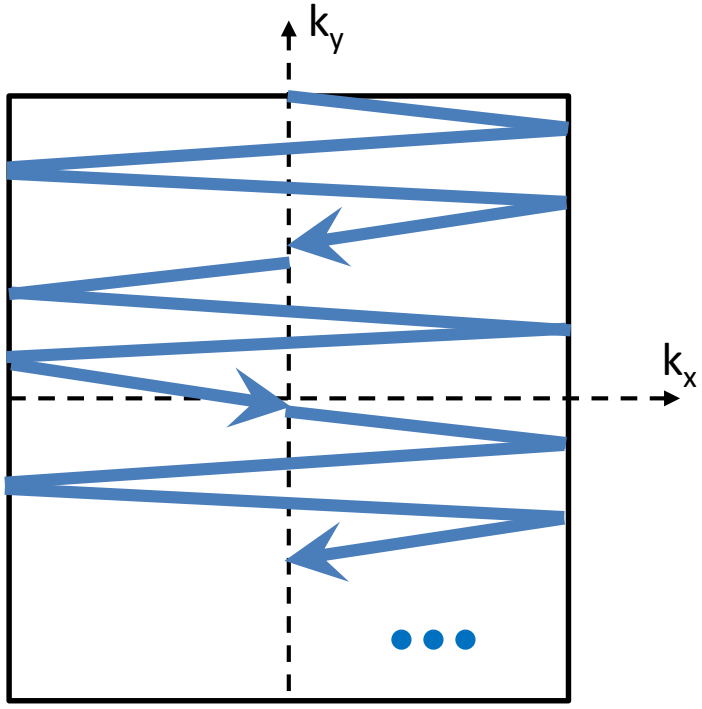


Figure B.22

(a)

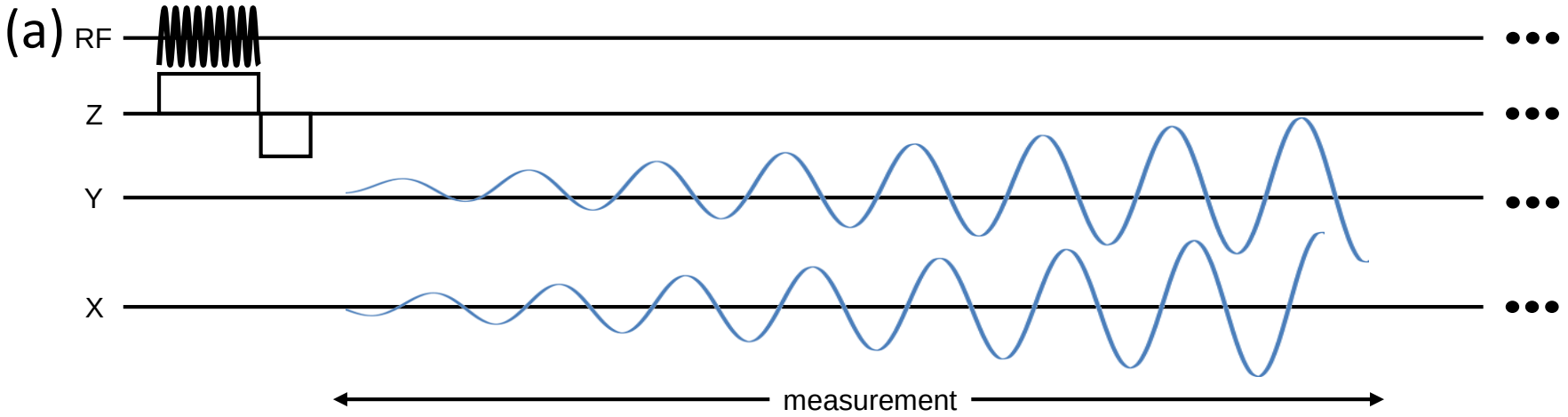


(b) Any application needing very fast imaging (f.i. DWI, BOLD, …)

(c) Contrast blurring + Signal blurring

(d) Low resolution in the phase encode direction.

Exercise 23



(b) Premature stop would mean only the centre of k-space is filled in, so good contrast but bad resolution (blurred!)

(c) First the outside of K-space, so high SNR measurements on the outside (good resolution), and lower SNR with higher TE in the centre (lower SNR and strongly T2*-weighted).

(d) Higher sample rate in the centre, leads to images with higher SNR.

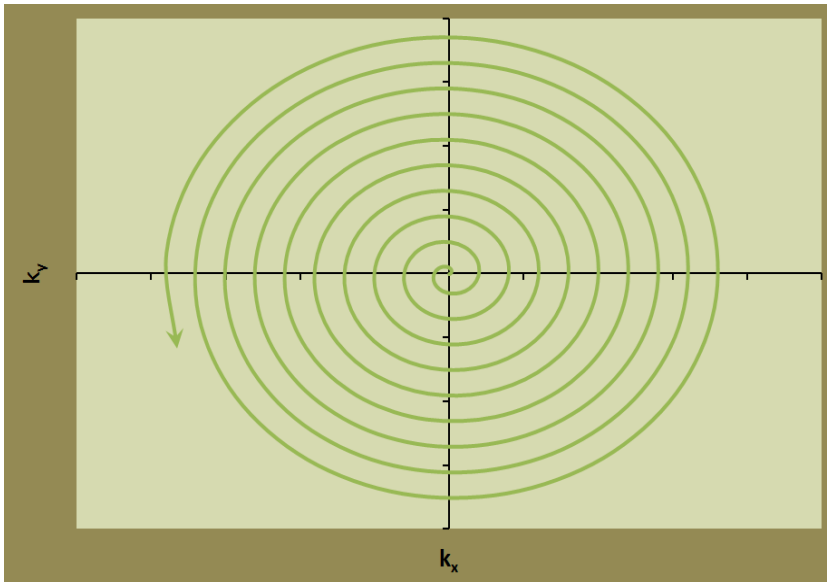
Exercise 24

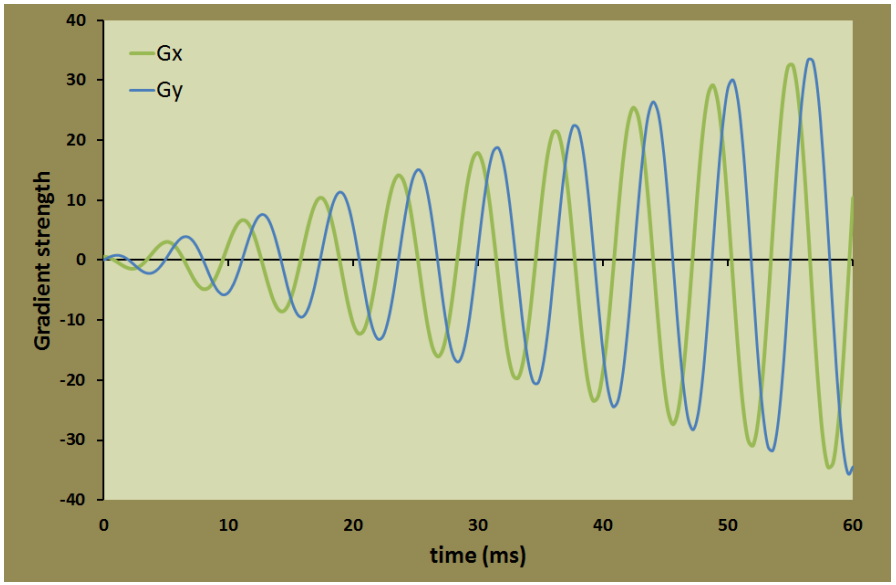
- **(4.49)** $k(t) = \frac{\gamma}{2\pi}\int_0^t G(\tau)d\tau \rightarrow G(t) = \frac{2\pi}{\gamma}\frac{dk(t)}{dt}$

a=amplitude increase over time
b=speed of radial rotations

$$k_x(t) = \frac{\gamma}{2\pi} at\cos(bt) \qquad G_x(t) = \frac{2\pi}{\gamma}\frac{dk_x(t)}{dt} = a[\cos(bt) - bt\ sin(bt)]$$

$$k_y(t) = \frac{\gamma}{2\pi} at\sin(bt) \qquad G_y(t) = \frac{2\pi}{\gamma}\frac{dk_y(t)}{dt} = a[\sin(bt) + bt\ cos(bt)]$$





Exercise 25

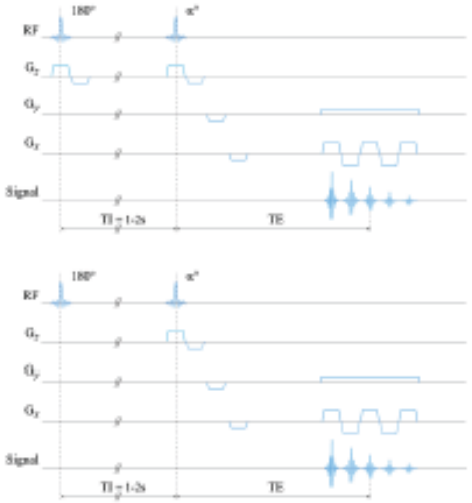
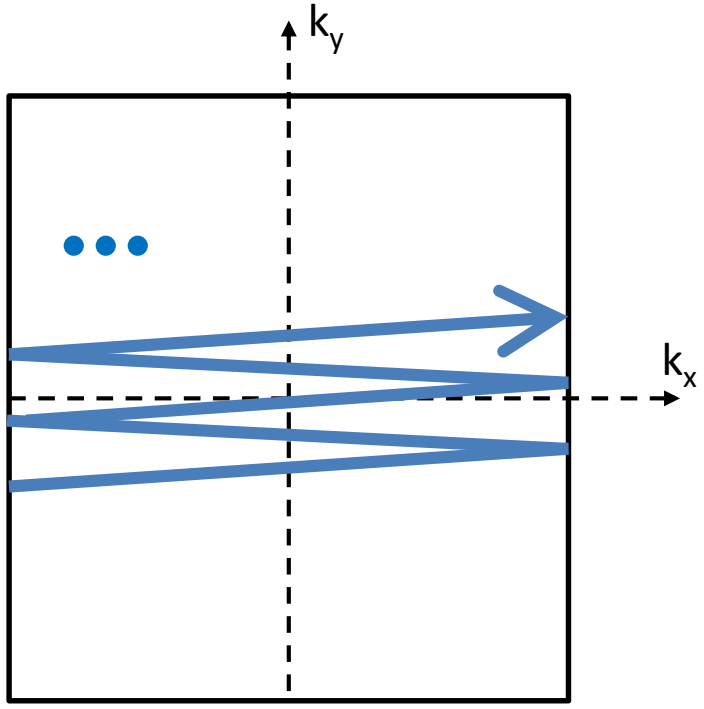


Figure 1.21

(a)



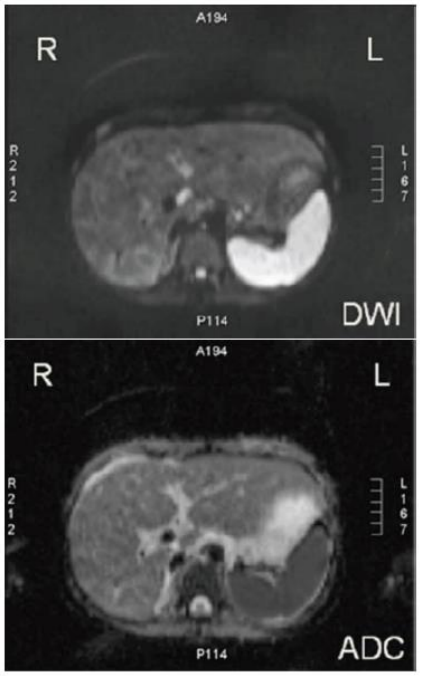
(b) Slice-selective vs non-selective 180° RF pulses are identical for all stationary tissue, but have inversed phase histories for inflowing blood.

(c) Difference between the two images provides image of inflowing blood – ‘phase contrast imaging’

Exercise 26

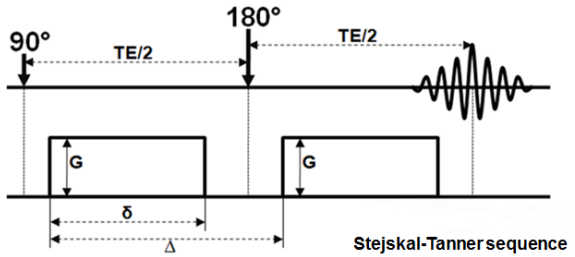
(a) – (c) See Slide 49 (MRA) of the MRI slides and pages 30-32 of Chapter 4.

Exercise 27



- Liver vs Spleen ADC
  - Native DW image:
    - Spleen bright
    - Liver iso-intense
  - ADC image:
    - Spleen dark
    - Liver iso-intense

Figure 1.22



➔ Spleen has higher restriction to movement than Liver (water molecules in liver more mobile).

Exercise 28

Ferritin (iron-based, ferromagnetic) causes local magnetic field inhomogeneities, yielding dephasing and a hypointense signal.

Any sequence that is sensitive to local magnetic field inhomogeneities can be used, most often gradient-echo sequences.

Exercise 29

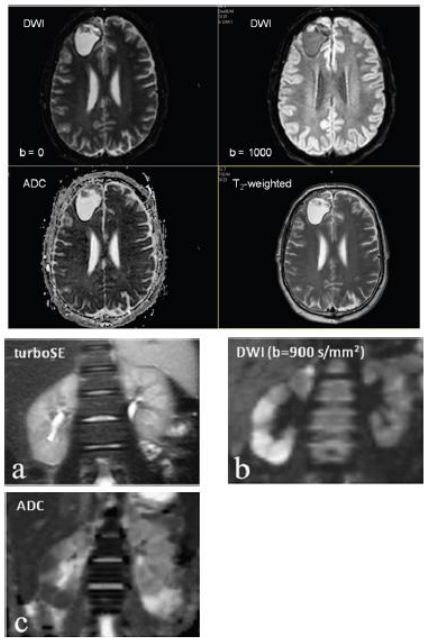


Figure 1.23

- Above: cystic lesion
  - Bright on b0/T2, dark on b1000
  - Bright on ADC
  - → High mobility (benign cyst)

- Below:
  - Iso-intense on T2 TSE (a)
  - DWI b900 (b) bright spots corresponding to low ADC (c)
  - →Low mobility (malignant lesion / focal pyelonephritis)

Exercise 30

(a) All spins present outside the boundary of the FOV will be folded into the image at the opposite border (phase difference of 361° is identical to that of 1°). Hence the nose will be visible at the back of the neck.

(b) You need to subdivide the max phase shift of 360° into a double amount of voxels, therefore double the sampling frequency.

(c)

- As you need double the amount of voxels in the phase direction, the Nph doubles, which means the acquisition time doubles.
- Doubling the sampling in the frequency encode direction does not (necessarily) have any effect on the TR of the sequence, or on the Nph, so no effect on acquisition time. (In practice, all MRI sequences are oversampled in the frequency encode direction as it has no downsides, which means you will nearly never see foldover artefacts in the frequency encode direction).

Exercise 31

a) Avoiding aliasing using Nyquist criteria (slide 64)

$$\Delta k_x \le \frac{1}{2 * x_{\max}} \quad \text{and} \quad x_{\max} = \frac{FOV_x}{2} \quad \text{and} \quad \Delta k_x = \frac{\gamma}{2\pi} * G_x * \Delta t$$

So, $G_x.\Delta t \le \frac{2\pi}{\gamma} . \frac{1}{FOV_x}$ with FOVx = L

b) $BW = \frac{\gamma}{2\pi} G_{x_{\max}} .L \rightarrow G_{x_{\max}} = \frac{2\pi}{\gamma} \frac{BW}{L}$

c) $G_x.\Delta t \le \frac{2\pi}{\gamma} . \frac{1}{FOV_x}$ and $BW \le \frac{\gamma}{2\pi} G_x L$

So, $\Delta t \le \frac{1}{BW}$ if FOVx = L

Exercise 32

- Formula

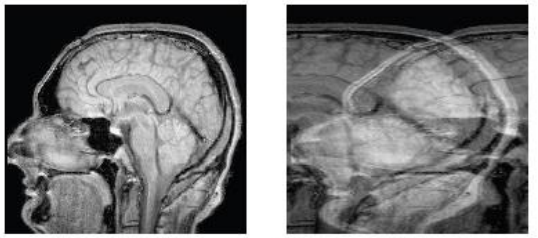
  (4.97) $G_y \Delta t \leq \frac{2*\pi}{\gamma * FOV_y}$

  with Gy=0.010T/m, FOVy=0.08m and γ/2π = 42.57 MHz/T

$$\rightarrow \Delta t \leq \frac{2*\pi}{\gamma} \frac{1}{FOV_y * G_y} = \frac{1}{42.57*10^6*0.08*0.010} = 29 * 10^{-6} s$$

(= highest resolution without aliasing)

Exercise 33



(a) TR = 3000, TE = 10 without artifact

(b) TR = 3000, TE = 10 with artifact

Figure 1.24

- Image on the right:
  - undersampling of k-space in the k_x direction (i.e., Δt too large).
  - Nyquist criteria was not respected.

Exercise 34

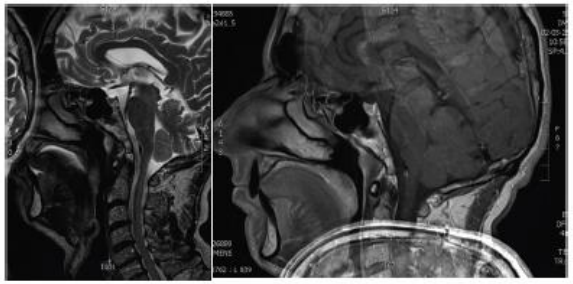


Figure 1.25

- Aliasing / fold-over artefacts
  - Anteroposterior in the left image, head-feet direction in the right.
  - Nyquist criteria not respected (max frequency is larger than expected).
  - Avoiding:
    - Increase FOV (but resolution decreases)
    - SAT blocks to destroy signal outside FOV (scan time increases)
    - Oversampling (scan time increases)
    - Faster sampling, decrease Δt (increase BW, more noise)

Exercise 35

turboSE

a) liver: TA = 1 x 160/10 x 0.5s = 8s

b) heart TR = RR: TA = 1 x 160/10 x 1s = 16s

B.5 Nuclear Medicine Imaging

Exercises

Nuclear Medicine

1.

- (a) if N(t) and $N(t_0)$ are known

$$\rightarrow N(T_{\frac{1}{2}}) = N(0)e^{-T_{\frac{1}{2}}/\tau} = \frac{1}{2}N(0)$$

$$\rightarrow T_{\frac{1}{2}} = \ln 2 / \alpha$$

- (b) $T_{1/2}(^{99m}Tc)$ = 6h

$T_{1/2}(^{18}F)$ = 109min

1.

- (c)

❑ keep distance to other people

❑ restrict the time close to other people

❑ duration of these recommendations depends on $T_{1/2}$

❑ Most tracers are excreted via the kidneys and the bladder → drink a lot and empty the bladder regularly

❑ use a separate toilet

2.

- a correction for the attenuation in PET can be performed by multiplying the emission measurement with the factor $N_0/N(d_1,d_2)$. Consequently, FBP can be applied. For SPECT, however, this is not possible.
- Unlike in iterative reconstruction, Poisson noise cannot be modeled.
- most modern PET scanners contain gaps between detector modules (incomplete angular sampling), which leads to streak artifacts in FBP
- Resolution effects due to finite detector size and a-collinearity of the photon pairs in PET cannot be modeled in FBP Consequence: image artifacts.

3.

- Anterior and posterior whole body (^{99m}Tc-MDP) image acquired with a dual-head gamma camera. The two images look different as a result of the attenuation.

4.

- (a) 0.79cm
- (b) 134.5kg
- (c) too heavy to use in practice.

5.

- (a) $N_A = Ne^{-\mu(L+x)} + \frac{N}{4}e^{-\mu(L-x)}$

  $SD_A = \sqrt{N_A}$

- (b) $N_B = \frac{N}{4}e^{-2\mu L}$

  $SD_B = \sqrt{N_B}$

6.

- A single photon is recorded with an energy of 280 keV and the position is the average position of the two incident photons.

- Probability = $e^{-A\Delta T}\,\frac{(A\Delta T)^2}{2}$

Note: the expected number of photons is $\varepsilon * A * \Delta T$ where ε is the (geometric + absorption) efficiency.
→ factor ε is missing in front of $A*\Delta T$

8. $$\lambda = \frac{y_1 + y_2 + y_3}{\alpha\, e^{-2\mu r} + 2\alpha\, e^{-\mu r}}$$

9

Scatter broadens the PSF and, hence, reduces the resolution. However, the situation for SPECT and PET is slightly different.

SPECT: a photon scattered by the attenuating object will be detected only if it enters the detector perpendicularly (because of the parallel hole collimator). Consequently, the backprojection during reconstruction will assign this activity to a (wrong) location inside the object boundary.

PET: a photon scattered by the attenuating object will yield a wrong line of response that may lie outside the attenuating object.

Note: in PET and SPECT there are methods to correct for scattered events. In case these corrections are 100% correct, there is no influence on the reconstructed image!
However, 100% accurate scatter correction is challenging!

10. $\mu_1 = \ln \frac{N_3}{N_2}$

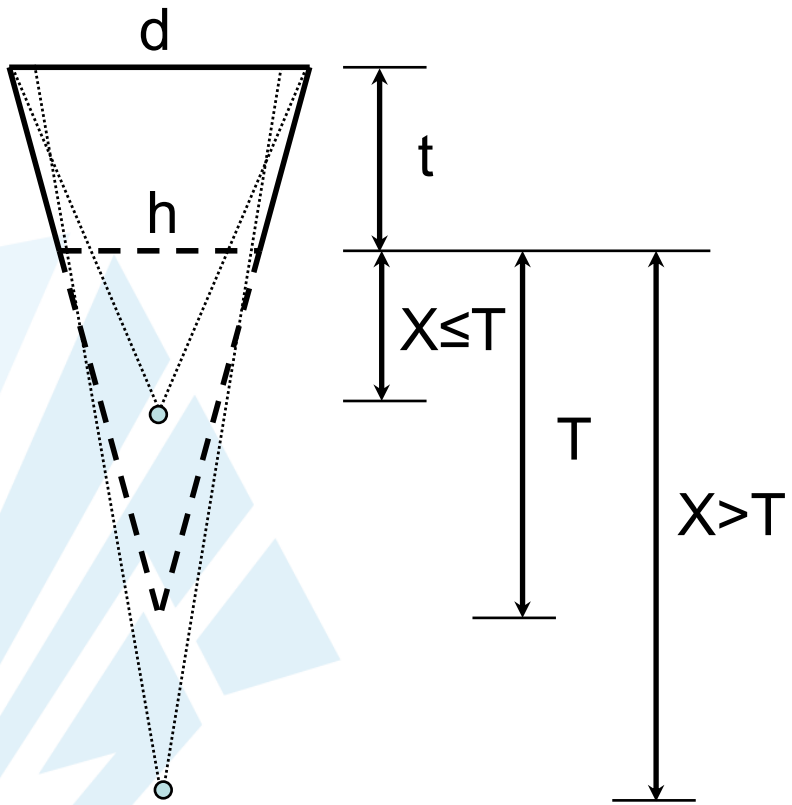
$\mu_2 = \ln \frac{N_1 N_2}{N_3{}^2}$

11.

- Use Poisson distribution

-> Probability = $\frac{1}{24e^2}$

12(a).

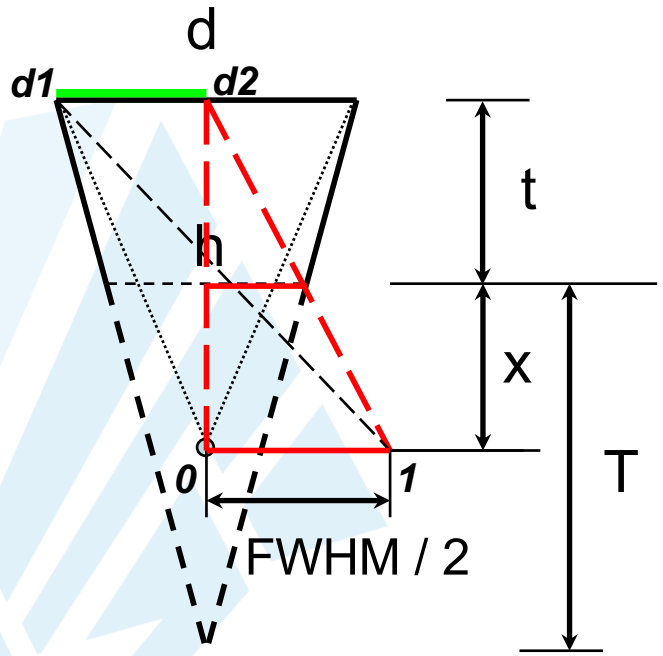


- **X≤T**
- $$S = \frac{d^2}{4\pi(x+t)^2}$$
- **T<X**

$$S = \frac{h^2}{4\pi x^2}$$

12(b) — x≤T

point 1 is perceived by half of the detectors that perceive point 0 (i.e., $|d_1 d_2|$)



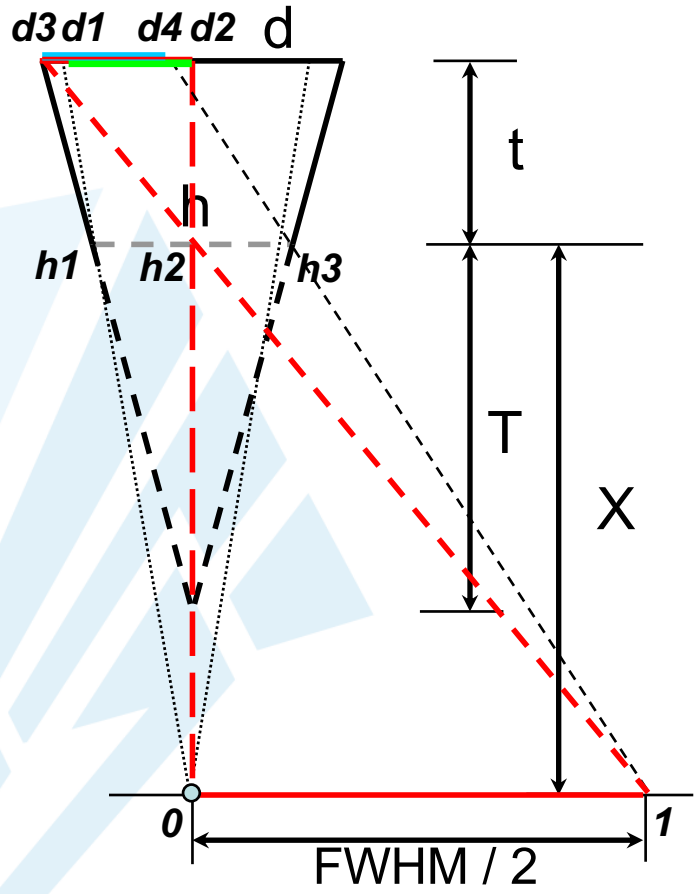
In red triangles:

$$\frac{FWHM/2}{h/2} = \frac{x+t}{t}$$

$$FWHM = \frac{x+t}{t} \cdot h$$

12(b) — X>T

point 1 is perceived by half of the detectors that perceive point 0 (i.e., $|d_1 d_2|$)



$$\frac{|d_1 d_2|}{|h_1 h_2|} = \frac{|d_3 d_4|}{|h_2 h_3|}$$

$$|h_1 h_2| = |h_2 h_3| \Rightarrow |d_1 d_2| = |d_3 d_4|$$

In red triangles:

$$\frac{\mathrm{FWHM}/2}{d/2} = \frac{x}{t}$$

$$\mathrm{FWHM} = \frac{x}{t} \cdot d$$

13.

The maximum likelihood of the total number of emitted photons $\lambda = 1500$

14.

The expected number of measured photons N for

- single photon detection:

$$N = S\int_0^L \lambda e^{-\mu x} dx = \frac{S\lambda}{\mu}\left(1 - e^{-\mu L}\right)$$

- coincidence detection:

$$N = S^2 \lambda L e^{-\mu L}$$

Notes:

- The absorption efficiency and the geometrical efficiency have a different influence. This answer is correct if S is the absorption efficiency of single detector block and the geometrical efficiency is 100%.
- In the coincidence case both photons must be absorbed (thus S^2). It should be "number of measured photon pairs" in the coincidence case.

15.

- (a) $\frac{1}{2}\frac{a^2}{2\pi R^2}$
- (b) $\frac{5}{12}\frac{a^2}{2\pi R^2}$

16.

(a) fraction = $3/5 \cdot 1/(900\pi)$

(b) fraction = $4/5 \cdot 1/(900\pi)$

(c) fraction = $1/(4\pi) \cdot 2/(5L) \cdot \varepsilon_1 \cdot \varepsilon_2 \cdot \lambda \cdot e^{-2L\mu} (4a^2/3 + b^2/4)$

17.

- (a) $\varepsilon_1 A e^{-1.5}$
- (b) $A\,\varepsilon_1 \varepsilon_2 e^{-6}$

Note:
The absorption efficiency and the geometrical efficiency have a different influence in case of coincidence measurement, therefore we need to introduce additional parameters to describe them separately.

If the geometrical efficiency is 100%, then the above solutions are correct.

Suppose ε_1 ε_2 are the absorption efficiency, and g_1 g_2 are the geometrical efficiency (and $g_2 <= g_1$), then the answer should be (a) $A\varepsilon_1 g_1 e^{-3/2}$ (b) $2A\varepsilon_1\varepsilon_2 g_2 e^{-6}$

18.

a) $N_1 = \frac{3}{4} \cdot a^2/(4\pi R^2)$

$N_2 = \frac{1}{2} \cdot a^2/(9\pi R^2)$

b) $N_1 = \frac{3}{4} \cdot a^2/(4\pi R^2) \cdot e^{-3/2}$

$N_2 = \frac{1}{2} \cdot a^2/(9\pi R^2) \cdot e^{-2}$

c) $N = 2\, a^2/(9\pi R^2) \cdot \frac{3}{4} \cdot \frac{1}{2} \cdot e^{-7/2}$

μ is function of the energy E and the material

usually $\mu(E=511keV) < \mu(E=140keV)$

19.

- (a) $\frac{1}{e}$
- (b) $\frac{1}{2e}$

20.

Assume a positron emitting source -> photon energy = 511 keV and $T_{1/2}$ = 7200 s (^{18}F)

The absorbed dose =

$$(1\times 3.7\times 10^{7}\times \int_{0}^{\infty} e^{-\frac{\ln 2}{3600.T_{1/2}}t}\,dt)\times \frac{h^{2}}{2\pi L^{2}}\times (1-e^{-\mu h})\times 511000\times 1.6\times 10^{-19}\,/\,0.125 = 0.437 mGy$$

B.6 Ultrasound Imaging

1. (a)

Reflection is the change in direction of a wavefront at an interface between two different media so that the wavefront returns into the medium from which it originated.

During *refraction* the wavefront passes from one medium to the next but changes the angle between the direction of sound propagation and the interface between both media.

Scattering is the process in which acoustic energy is re-directed due to local inhomogeneities in mass density or compressibility.

Absorption refers to the phenomenon of dissipation of acoustic energy by the tissue causing tissue heating.

Both scattering en reflections are essential to conventional ultrasound imaging as they form the basis for the signal that is received upon transmission of an ultrasound wave. Although reflections typically show the borders between different tissues, the scatter reflections will typically generate a typical 'speckle' pattern that can help differentiate tissue types. Although refraction effects can typically be neglected in ultrasound imaging, they could result in a spatial shift and morphing of structures that lay behind an interface. Finally, absorption will result in a loss of signal with distance limiting the penetration depth of the ultrasound wave and thereby the field-of-view of the ultrasound image.

1. (b)

The difference in acoustic impedance between two media will determine both the amplitude and the phase of the reflected wave.

1. (c)

As air and soft tissue have a dramatic difference in acoustic impedance (air: 420 Pa·s/m while water, i.e. soft tissue: 1.5 MPa·s/m) air reflects ultrasound strongly. As such, a layer of air between the transducer and the patient would result in almost no acoustic energy being transmitted into the patient and only an image of low SNR (or no image at all) may be obtained. As such, a coupling layer (water-based gel) between the ultrasound transducer and the patient typically needs to be used.

1. (d)

Constructive interference is the phenomenon in which waves from multiple sources add together in a constructive manner as to amplify each other (in contrast to destructive interference where the waves from multiple sources cancel one another out). By using multiple transducers (i.e. multiple pieze-electric elements embedded in a single device) and firing them in such a way that their individual wavefronts

travel equal distances up to a point in space, constructive interference of these wavefronts in this spatial point will result in an ultrasound focus. As such, ultrasound beams can be both steered and focussed by introducing proper time delays between the transmission of the individual wavelets.

3.

Continuous wave Doppler, Pulsed-wave Doppler and Color Flow mapping. See chapter on ultrasound imaging for details.

4. (a)

One image line can be formed every 1/3000 s. As one image consists of 60 scan lines, the generation of 1 frame takes 60/3000 s (20 ms). Per second, 1 / 0.02 images can thus be made implying a frame rate of 50 Hz.

4. (b)

The wave will travel 2 * 10cm = 20 cm to generate a single line in the image resulting in 0.2m / (1530 m/s) * 20e6 samples/second = 2.6144 ksamples/line. As every sample is 16 bit (i.e. 2 byte), every scan line will thus take 5.2288 kbytes. A total of 256 Mb / (60 lines/image * 5.2288 kb/line) = 1632 images can thus be stored. At a frame rate of 50 Hz this corresponds to 32.64 seconds.

Note that the above reasoning implies that samples are only stored when strictly required for image display.

4. (c)

Cf. 4 (b)

7. (a)

The period of the sampled signal is 5 samples (every 5^{th} sample we observe a multiple of 2*pi). As such, the period of the Doppler signal corresponds to 5 times the PRF or the Doppler frequency to the PRF divided by 5 being 2.4 kHz. Simply filling out all values in the Doppler equation then results in a blood velocity of 0.7333 m/s.

7. (b)

According to the Nyquist criterion, the sampling frequency of the signal (i.e. the PRF) must be larger than twice the maximum frequency in the signal. As such, in this example, the PRF must remain larger than twice the Doppler frequency: PRF > 2.fD = 4 v fT / c. The maximal velocity that can be measured without artefact is thus: vmax = c PRF / 4fT = 1.836 m/s

7. (c)

The PRF is limited by the fact that the ultrasound wave needs to be able to propagate to-and-from the sample volume: Tprf > 2d / c. This combined with the sampling criterion mentioned above leads to the relationship that vmax = c^2 / (8 d fT). In other words, this maximal distance equals 1530^2 / (8 * 2.5e6 * 1.836) m = 6.38cm.

7. (d)

The sampling process can be seen as the multiplication of the (continuous) sinusoidal Doppler signal (with frequency fD) with a Dirac comb with spacing (1/PRF). The result can thus been seen as the convolution of the Fourier transform of this sinusoidal signal (a Dirac pulse at +fD and –fD) with the Fourier transform of the Dirac comb (being another Dirac comb with spacing PRF). The Dirac pulses at –fD and +fD will thus be positioned around the Dirac's of the Dirac comb. A simple graphical representation of this convolution process tells us that a Doppler frequency of 9.268kHz (corresponding to a velocity of 2.836 m/s) will fold into an apparent Doppler frequency of 9.268kHz – 6kHz = 3.268 kHz. This corresponds to a velocity of 1m/s.

A simpler way of seeing this is by realizing that vmax will result in a Doppler frequency aliasing into 0kHz. As such, the apparent Doppler frequency will simply correspond to the part of the velocity above vmax, i.e. 1m/s.

9. (a)

The pulse length (in this case 2*T*c) must be smaller than twice the distance between the structures that need to be resolved (dmin). Thus: 2Tc < 2dmin. Filling in the proper numbers results in a fmin = 3.06MHz.

9. (b)

The attenuation coefficient at is 1 dB/cm.MHz. As we need to propagate 5cm, the total attenuation will be 5 dB/MHz. As 100dB is the limit to still obtain adequate SNR, the maximally allowed frequency is 100dB / (5dB/MHz) = 20 MHz.

9. (c)

A compromise between SNR and resolution would be required and an intermediate frequency of e.g. 12MHz could be recommended.

B.7 Medical Image Computing

1. Dynamic programming

Gradient

0	1	1	1	1	0	0	0	1	3	3
0	3	3	5	4	3	2	3	3	5	4
0	4	5	4	4	5	4	5	5	2	2
0	3	3	2	2	5	5	3	3	0	0
0	1	2	1	3	1	2	4	1	4	4

Cost

0	4	7	7	7	9	10	11	10	8	8
0	2	4	2	3	5	7	7	7	5	6
0	1	1	2	3	3	4	4	4	7	9
0	2	4	5	6	4	4	6	7	10	12
0	4	6	7	6	9	8	6	10	7	8

Path

0	1	2	3	2	2	3	3	3	3	2
0	2	2	3	2	2	3	3	3	3	2
0	3	3	3	3	3	3	3	3	3	2
0	4	3	3	3	3	3	4	3	3	2
0	5	3	3	3	3	3	4	3	3	5

2. SSD/MI

- SSD(A,B): 1495/N, MI(A,B): 2.1250
- SSD(A,B1): 1351/N, MI(A,B1): 1.1569

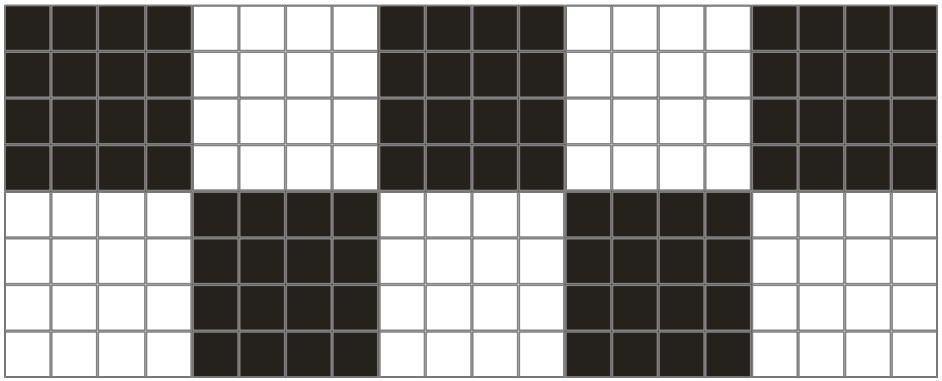
3. MI = 2.3750

4. T1 patient, T2 atlas + segmentation

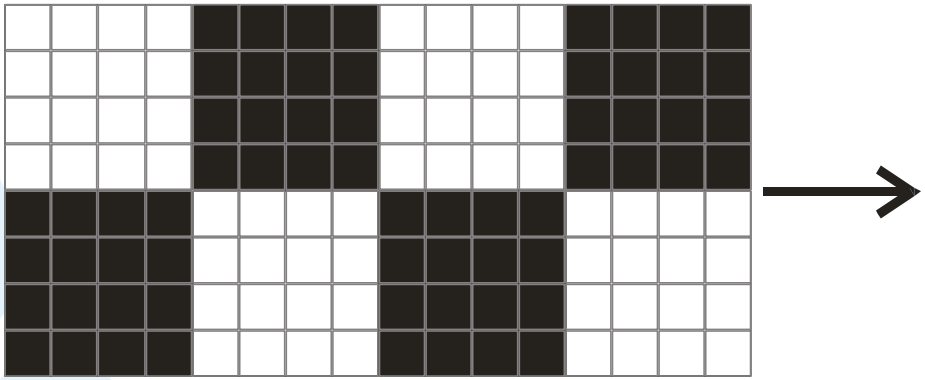
- Registration only: register T2 → T1, segmentation by atlas segmentation
- Registration & segmentation: register T2 → T1. Use atlas segmentation as prior (supervised segmentation: learn mu, sigma) to segment T1

5.

R



F



7. Binary MI

P0	Wit	Zwart
Wit	0	1/2
Zwart	1/2	0

1

P1	Wit	Zwart
Wit	1/8	3/8
Zwart	3/8	1/8

0.185

P2	Wit	Zwart
Wit	1/4	1/4
Zwart	1/4	1/4

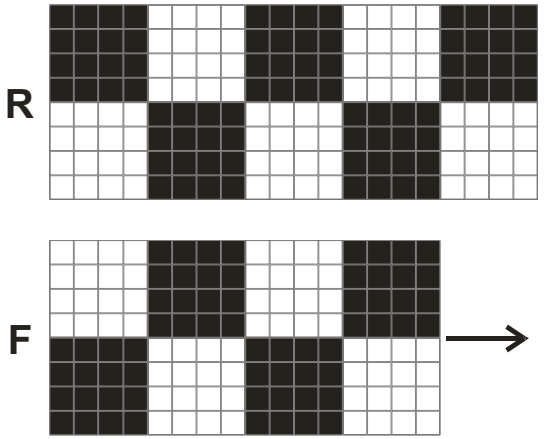
0

P3	Wit	Zwart
Wit	3/8	1/8
Zwart	1/8	3/8

0.185

P4	Wit	Zwart
Wit	1/2	0
Zwart	0	1/2

1



P0: no translation,

P1: translation over 1 pixel, …

P0 and P4: MI = 1;

P2: MI = 0;

P1 and P3: MI = 1+1–2(1/8*3+3/8*1.42)=0.185

B.8 Visualization for Diagnosis and Therapy

6. Augmented reality

Calibration = we know projection matrix $A_{camera \rightarrow endo}$. Estimate $A_{pre\text{-}op \rightarrow camera}$ from (manually indicated) corresponding points.

$$\begin{bmatrix} u_{endo} \\ v_{endo} \\ 1 \end{bmatrix} = A_{camera \rightarrow endo} A_{pre-op \rightarrow camera} \begin{bmatrix} x_{pre-op} \\ y_{pre-op} \\ z_{pre-op} \\ 1 \end{bmatrix}$$

Moving endoscope: add tracker to continuously re-estimate $A_{world \rightarrow camera}$:

$$\begin{bmatrix} u_{endo} \\ v_{endo} \\ 1 \end{bmatrix} = A_{camera \rightarrow endo} A_{world \rightarrow camera} A_{pre-op \rightarrow world} \begin{bmatrix} x_{pre-op} \\ y_{pre-op} \\ z_{pre-op} \\ 1 \end{bmatrix}$$

7. Image guided surgery

Find markers in each radiograph, thus obtaining for each marker (x_w, y_w, z_w), (u_1, v_1), (u_2, v_2). Estimate $A_{1,2\rightarrow world}$ and apply to (u_{1l},v_{1l},u_{2l},v_{2l}).

$$\begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = A_{1,2\rightarrow world} \begin{bmatrix} u_1 \\ v_1 \\ u_2 \\ 1 \end{bmatrix}$$

We need to estimate 3x4 matrix (one of (u_{1l},v_{1l},u_{2l},v_{2l}) is redundant), thus have 12 degrees of freedom. Ergo, 4 points are needed.

9. Maxillo-facial

a) Segment skin (e.g. threshold at -500 HU) + marching cubes

b) Align/estimate photograph pose (e.g. using landmarks such as the eyes), project 3D surface on photograph.