Module 3G4: Medical Imaging & 3D Computer Graphics

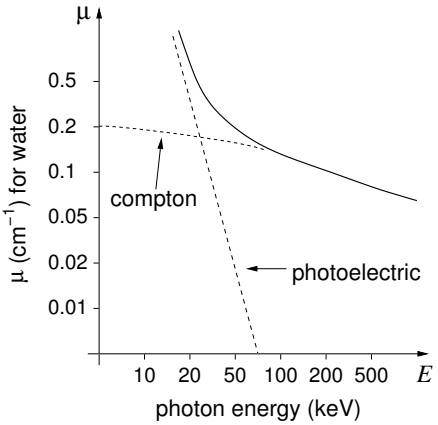
Solutions to 2019 Tripos Paper

1. **Attenuation in ultrasonic and nuclear imaging**

(a) When ultrasound propagates through soft tissue, it is attenuated. This is mainly due to the medium's viscosity, which causes acoustic energy to be converted into heat. The result is an exponential decay in the amplitude of the pressure wave as it propagates. If the wave has amplitude A_z and A_0 at propagation distances z and 0 respectively, then $A_z = A_0 e^{-\alpha z} = e^{-\alpha_0 f^n z}$, where f is the wave's frequency. Rearranging, we can write the attenuation coefficient as $\alpha = \frac{-\ln(A_z/A_0)}{z}$. The units of α are therefore Np/cm, where a neper (Np) is a dimensionless unit defined as $\ln(A_z/A_0)$.

At imaging frequencies, most tissues have an α that is proportional to frequency ($n = 1$). Hence the higher the frequency, the higher the attenuation and the lower the penetration. But axial resolution improves for shorter wavelengths, so we get better image quality at higher frequencies. A compromise has to be struck between resolution and penetration. [30%]

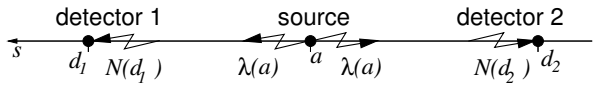
(b) As a stream of X-ray photons of intensity I and energy E passes through a homogeneous medium of thickness x, they are attenuated according to $I = I_0 e^{-\mu x}$.



The linear attenuation coefficient μ depends on the photon energy (E), the atomic number of the material (Z) and the density of the material (ρ). At diagnostic X-ray wavelengths, attenuation is mostly due to photoelectric absorption and Compton scattering.

The energy of X-ray radiation is proportional to its frequency. Hence we can see from the graph that the attenuation is reduced at higher frequencies. This reduction is not linear, even on log-log axes, because it is caused by two distinct mechanisms. [10%]

(c) In PET imaging, we need to detect both photons. Each photon has an independent probability of being attenuated between the source and the detector.



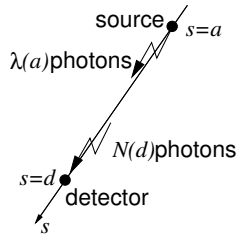
If $\lambda(a)$ pairs are emitted along the s-axis, the number detected is

$$N(d_1, d_2) = \lambda(a) \exp\left[-\int_{d_1}^{a} \mu(s)\, ds\right] \exp\left[-\int_{a}^{d_2} \mu(s)\, ds\right]$$
$$= \lambda(a) \exp\left[-\int_{d_1}^{d_2} \mu(s)\, ds\right]$$

Notice that the received signal is a function of the rate at which *pairs* of photons are received, and hence it depends on the attenuation along the complete path between the two detectors that receive the two photons. If we knew the total attenuation along this path through the subject, then we could correct for it. The attenuation along complete paths through the subject is exactly what is measured by a CT scanner.

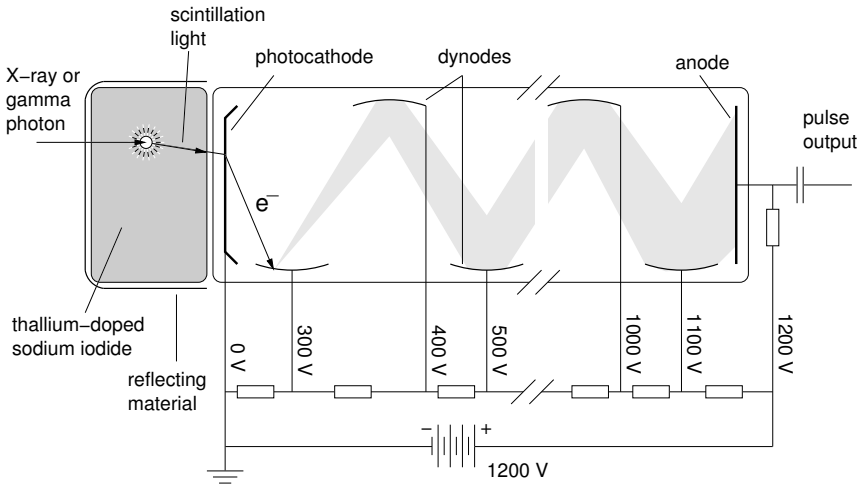
Attenuation correction can therefore be achieved by building a CT scanner into the same housing as the PET scanner. Reconstructing the CT data provides the PET attenuation correction factors $\exp\left[-\int_{d_1}^{d_2} \mu(\nu)\, d\nu\right]$. The emission λ, as a function of location, can then be found using Fourier reconstruction or filtered backprojection.

A similar approach does not work for SPECT data because the attenuation during the acquisition process only relates to the shorter path from the point of emission to the detector.



The attenuation along this shorter path cannot be measured directly by a CT scanner. [30%]

(d)



The scintillation material produces flashes of light as X-rays are absorbed by Compton scattering and eventual photoelectric absorption. This results in an electron being emitted from the photocathode, travelling towards the first dynode. As the electron hits the first dynode, it causes the emission of a larger number of electrons which are attracted towards the second dynode that is at a higher voltage. As the electrons hit each dynode, their numbers are multiplied. Hence the weak signal from the scintillation flashes is amplified millions of times by the photomultiplier tube to produce a current output. Importantly, the magnitude of the output voltage is proportional to the energy of the incoming X-ray photon. Detectors of this type are fast and have high efficiency, however they have low packing density. They are used in current nuclear medicine scanners.

These detectors are good for nuclear medicine imaging because they are very sensitive and can therefore detect low count-rates. They also measure the energy of the incoming photons: this enables energy filtering to eliminate multiple photon events and events with too low an energy resulting from photons that have been subject to Compton scattering. [30%]

Assessors' remarks: This question tested candidates' understanding of attenuation effects in ultrasound, nuclear and X-ray imaging. It was a popular question. A large number of candidates lost marks because they reproduced a section of the notes that they believed to be relevant rather than actually answering the question. Knowledge of the attenuation of X-rays and ultrasound was generally good. Understanding of the attenuation of the PET signal was generally weak. Some candidates failed to explain how the attenuation information from a CT scan can be used to correct for the attenuation of the PET signal. Very few candidates were able to produce a good labelled diagram of the scintillation crystal photomultiplier-coupled detector.

2. **Magnetic resonance imaging**

(a) Magnetic resonance imaging has a number of advantages:

- non-hazardous — subject exposed only to a (non-ionising) magnetic field,
- produces high-definition three-dimensional data,
- images anything containing hydrogen atoms — particularly water and soft tissue,
- can be used for functional as well as structural imaging.

However, it has some disadvantages:

- It is inherently slower than X-ray CT.
- The scanners are expensive to buy and to run.
- An MRI scanner applies a strong magnetic field (up to a few Teslas). Ferromagnetic materials must therefore be kept well clear.
- MRI is not good for imaging bone or air, so CT is preferred for the skeleton, the lungs and when looking for calcifications.

An MRI system images a (variable) combination of three nuclear magnetic properties that relate to the way in which the direction of the nuclear-magnetic spin returns to its equilibrium orientation after excitation. The three properties are T_1 ("spin-lattice") relaxation, T_2 ("spin-spin") relaxation and proton density. [20%]

(b) The main fixed magnetic field is a strong constant field along the principal axis of the scanner. It can be anything from 0.3 Tesla up to 7 Tesla. The nuclear magnetic spins seek to line up with the field, as far as quantum-mechanical constraints allow. In the quiescent state, the nuclear spins are, on average, aligned with the main field.

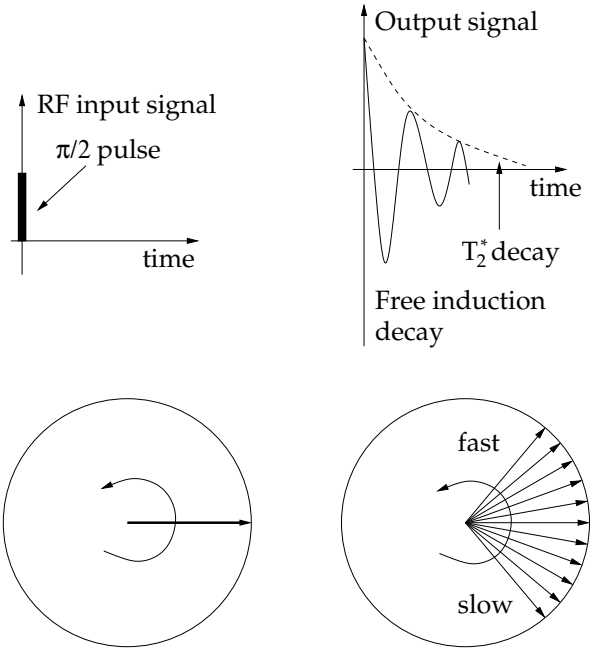
The radio frequency transmitter is used to disturb some of the nuclear spins from alignment with the main field. Once they are out of alignment, they begin to precess around the main field direction, emitting a signal that can be detected.

The gradient fields are used to vary the strength of the overall magnetic field (which is always in the same direction as the main field) as a function of spatial location. This variation in field strength causes similar variations in the rate at which nuclear spins precess. This means that the output signal will vary as a function of spatial location. [30%]

(c) (i) In a T1-weighted image, the output relates to the rate at which the magnetisation can recover between cycles. Hence a shorter T1 produces a brighter image. Of the materials listed, fat has the shortest T1 and so will appear brightest. [15%]

(ii) In a T2-weighted image, the brightness is proportional to the signal and so a slow decay constant will produce a brighter image. A longer T2 thus produces a brighter image. Of the materials listed, water has the longest T2 and so will be the brightest. [15%]

(d)



Net magnetisation in the x–y plane after the π/2 pulse

Inhomogeneities in the magnetic field cause rapid T_2*dephasing which reduces the signal.

To initiate a free induction decay, it is necessary to apply an RF pulse of sufficient intensity and length to swing the net magnetisation vector round by an angle of $\pi/2$, into the x–y plane. It is then necessary to apply a constant gradient field and record the high frequency signal emitted by the precessing nuclear magnetic spins. In theory, the gradient fields can be used in various combinations to make different points in space respond at different frequencies. In practice, the free induction decay signal is rarely measured, as it decays very quickly because of an effect called T_2^* whereby the rotating spins rapidly move out of phase alignment and the net magnetization vector in the x–y direction is consequently reduced in magnitude. This is caused by effects that are not random (in time) such as: small scale inaccuracy in the field from the gradient coils, imperfections in the main magnet, and variations in the magnetic susceptibility of the patient. These effects are not significantly dependent on material properties. There is therefore no output that is of interest in the imaging process. The rapid decay of the signal, before it can be detected or influenced by material properties, is the reason why free induction decay, on its own, is no use for imaging. [20%]

Assessors' remarks: This question tested candidates' understanding of Magnetic Resonance Imaging (MRI). It was a popular question and was, in general, very well answered. Candidates showed good knowledge of the advantages and disadvantages of MRI. Many also understood what physical properties were being imaged. The descriptions of the components of the MRI system were good, apart from for the gradient coil. Many candidates failed to make it clear that the various gradient fields are aligned with, and hence vary the strength of, the main field, without altering its direction. The question about T1-weighted imaging and T2-weighted imaging resulted in a mixture of both strong and weak answers; candidates had either learnt and understood the required material or they had not. The final part of the question about free induction decay produced a surprisingly wide variety of different levels of understanding. Some candidates reproduced the notes on the spin-echo sequence rather than answering the question.

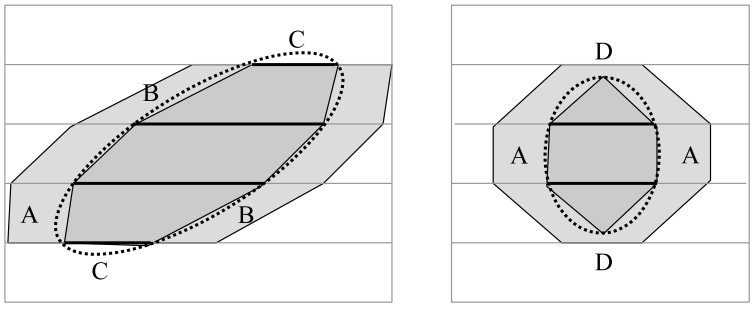
3. **Shape-based interpolation**

(a) Shape-based interpolation is a technique for interpolating between contours rather than data values. First, each cross-section is converted to a binary image, with ones inside and zeros outside the shape. Each of these images is turned into distance image using a distance transform, which replaces each pixel in the image with a distance estimate (usually a chamfer estimate) to the closest part of the contour. Negative distances are used outside of the cross-section and positive distances inside the cross-section. The distance transform is an in-place transform which can be performed very efficiently with only two passes over each image. These distance transforms are then linearly interpolated between the original cross-sections, to generate intermediate distance transforms between the original cross-sections. This creates a more dense volume of data which can compensate for slice spacing larger than the in-slice pixel size. An iso-surfacing algorithm like Marching Cubes can then be used to select the zero-isosurface from this interpolated volume of distance data: this will estimate the surface corresponding to the original cross-sections. [20%]

(b)(i) The further apart the CT slices are, the fewer cross-sections there will be through the tumour. This creates two problems. Firstly, shape-based interpolation is essentially a linear interpolation, and hence the shape will only vary linearly between slices. If the real tumour is either convex or concave, this linear variation will not accurately represent the real shape, and this effect is greater with fewer cross-sections. Secondly, shape-based interpolation cannot represent the shape beyond the first and last slices, so greater slice separation will risk missing more of the tumour at the top and the bottom. This applies equally to both the tumour and planning volumes. [10%]

(ii) Again, there are two problems when the tumour surface is at an acute angle to the slices. Firstly, if the cross-sections are in very different positions in each slice, shape-based interpolation will not be able to estimate the real tumour surface in between these slices very well, and may produce incorrect concave sections. Secondly, the planning volume is created by extending the 2D distance transforms, and hence the 1 cm gap is only added in the horizontal (CT slice) direction. If the tumour surface is acute to this, then the actual 3D gap between tumour and planning volumes will be considerably less than 1 cm. [10%]

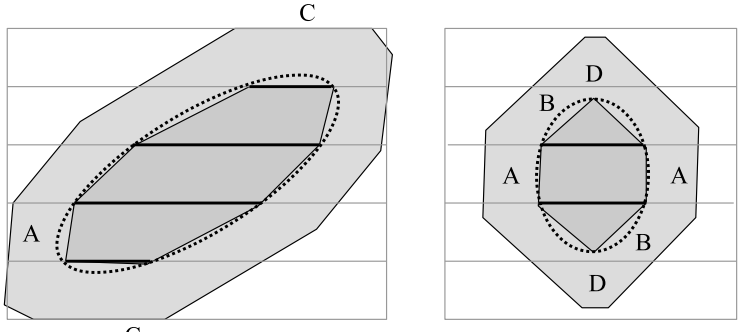
(c)



Plane A Plane B

The tumour volume is mid-grey and the planning volume light-grey. Areas where the planning volume is fine are marked 'A'. 'B' are areas where the 3D distance is actually much less than 1 cm as a result of the acute scanning angle. 'C' are areas where the top and bottom of the tumour have been missed off, and the planning volume cannot extend beyond the original cross-sections. 'D' is where the planning volume extends slightly beyond the tumour volume but not beyond the neighbouring CT slices [20%]

(d)(i)



Plane A Plane B

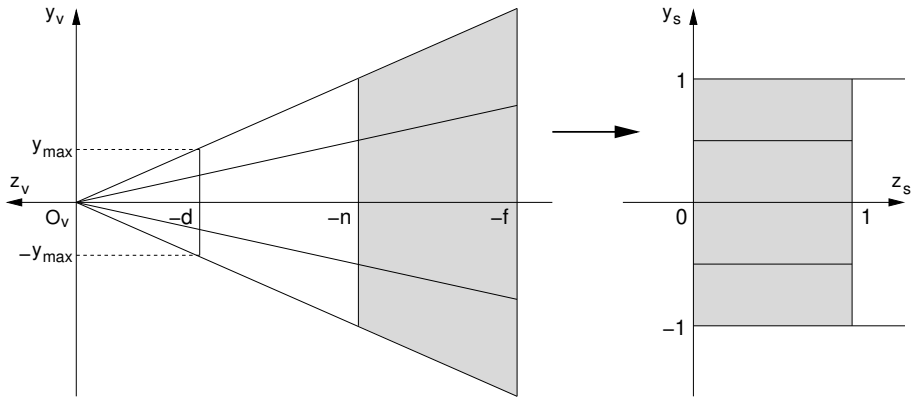
This time the areas 'A' where the planning volume is fine are much more widespread, since the extra 1 cm was added in 3D. There are still some issues with the convexity of the tumour not being represented by the linear interpolation ('B'). At 'C', the planning volume is 1 cm above and below the first and last cross-sections, but will be less than 1 cm from the actual tumour which was beyond these CT slices. At 'D', the planning volume will be correct since the full tumour volume is accounted for in this plane. [20%]

(ii) This new method largely corrects the difficulties of tumour orientation, at least those which are related to using a horizontal 2D offset for the planning volume rather than a full 3D distance. It will not, however, correct for problems in the shape-based representation of the initial tumour volume. It does not correct for large CT spacing missing highly convex or concave tumour regions, but it does deal better with the missing parts of tumour at the top and bottom of the scan: whilst the planning volume is still too small in these areas, it at least goes well beyond the tumour volume. [20%]

Assessors' remarks: This question tested the candidates' knowledge of shape-based interpolation using distance transforms. The book work in (a) was generally answered very well. (b) was also answered well, with nearly all students noting that larger spacing of slices would lead to poorer shape representation, and also that changing the orientation would have a shape-dependent effect on the result. Sketches in (c) were of varying quality. Most noted the linear interpolation leading to a piecewise linear shape, but only a few realised that the planning volume would be completely absent at the top and the bottom due to the use of 2D transforms in each plane. As a result, the repeated sketches for the 3D transform in (d) were often very similar to (c), though there were some answers that correctly noted both the additional top and bottom regions, and also the correct 3D planning volume gap.

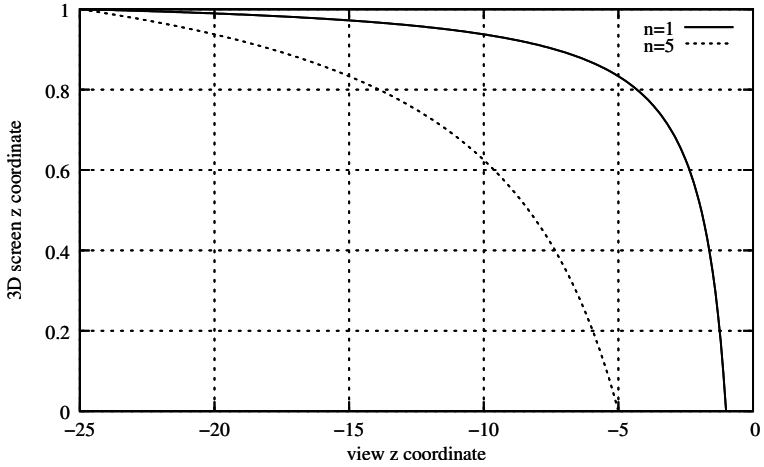
4. **View volumes and clipping**

(a) (i)



[20%]

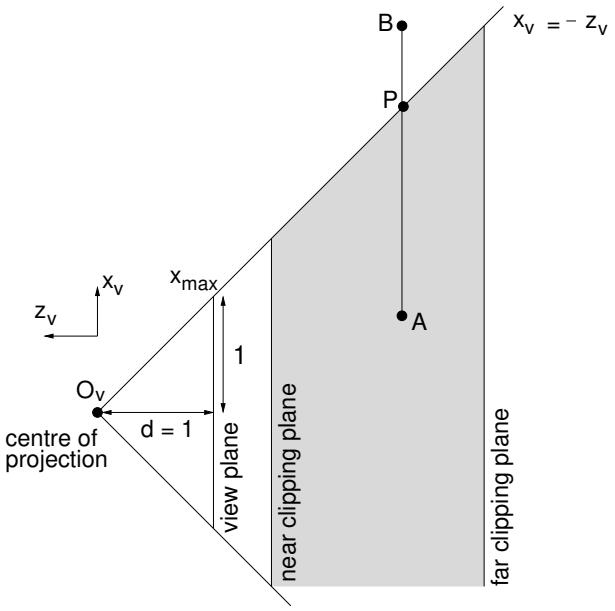
(ii)



The curves above are for $f = 25$, $n \in \{1, 5\}$. [10%]

(iii) The nonlinear definition of z_s normalises depth values to the range 0 to 1. More importantly, it is designed to ensure that lines map to lines and planes map to planes. If this were not the case, we would not be able to use linear interpolation in 3D screen space for clipping and depth interpolation. [10%]

(b) (i)



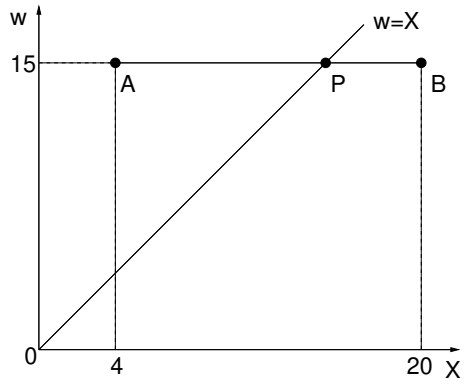
It is easily seen that A and B lie on the line $y_v = 0, z_v - 15$. By examination of the diagram (which is not drawn to scale), it is also easily seen that one of the boundary planes of the

view volume is $x_v = -z_v$. The line joining A and B intersects this plane at the point P, which has view coordinates $(15, 0, -15)$. The Sutherland-Hodgman clipper will cull vertex B, replacing it with P. [20%]

(ii) Substituting values, the perspective projection matrix is

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -5/4 & -25/4 \\ 0 & 0 & -1 & 0 \end{bmatrix}$$

Multiplying by the view coordinates of A and B, we find their homogeneous 3D screen coordinates to be $(4, 0, 12.5, 15)$ and $(20, 0, 12.5, 15)$ respectively.



In homogeneous 3D screen coordinates, the clipping limits are $-w \leq X \leq w$, so the clipping plane of interest in this instance is $w = X$. By inspection of the diagram, the homogeneous 3D screen coordinates of P are $(15, 0, 12.5, 15)$. [30%]

(iii) Multiplying the view coordinates found in (i) by the perspective projection matrix also gives $(15, 0, 12.5, 15)$. Hence, clipping in homogeneous 3D screen coordinates gives the same result as clipping in view coordinates, for the reasons identified in (a)(iii). [10%]

Assessors' remarks: This question tested the candidates' understanding of 3D screen coordinates for calculating perspective projections and clipping to a view volume. The book work in (a) was very well answered, with most candidates demonstrating a good working knowledge of the basic principles. In (b), almost all candidates were able to clip in view coordinates, though many missed the point of the follow-up question asking them to clip in *homogeneous* 3D screen coordinates. There was a tendency to convert to Cartesian 3D screen coordinates, which is an expensive and avoidable operation. Of those candidates who did work in homogeneous coordinates, many did not realise that the relevant clip boundary is the straight line $X = w$.

Andrew Gee, Richard Prager & Graham Treece
May 2019