ENGINEERING TRIPOS PART IIB 2012
4B20 DISPLAY TECHNOLOGY

Dr T D Wilkinson

Q1 a) Radiometry and photometry are, both, systems, terminology, mathematical relationships, measurement instruments and units of measure devised to describe and measure electromagnetic radiation and its interaction with matter. Radiometry is the pure measurement of radiation with respect to wavelength, whereas photometry is the measurement of radiation as seen through the response of the human eye. There are many different units in radiometry and photometry, some of which are now purely historical, but still cause considerable confusion. There is now a serious move all over the world to standardise these units under the SI system. It is important to note that the calculation of the radiation from complex shaped source such as a tungsten filament light bulb can only be approximated without the use of numerical simulation of the radiometric properties.

It is important that the nomenclature of radiometry is standardised, which seems to be the case in modern literature. For radiometric units, it is quite common to use the subscript *e* (Φ_e), to define the *energy* properties, whereas photometric units often bare the subscript *v* (Φ_v) to indicate the *visual* dependence of the units. This is a matter of choice. In the case of wavelength, the unit is standardised through the use of subscripts. If the radiometric quantity Φ is used without reference to wavelength then it has no subscript. If Φ is to be referenced to a wavelength λ. For every radiometric unit, there is a corresponding photometric unit, usually denoted by the use of the subscript *v* for visual. The difference between the two, is that the photometric quantities have been corrected for the wavelength response of the human eye.

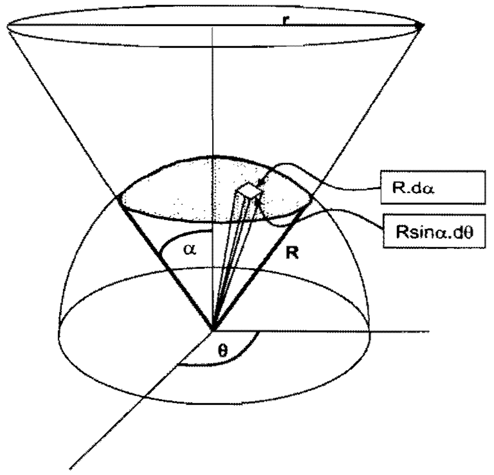
In order to under stand the definitions of radiometric units, it is important to first understand the **Radiant Energy**, Q_e. The radiant energy is the quantity of energy propagating onto, through or emerging from a specified surface of given area in a given period of time. The radiant energy is measured in **Joules** and is only useful as a starting definition for the more common radiometric units.

The **irradiance** E_e at a point on a surface is defined as the radiant flux $d\Phi_e$ falling on the surface, per unit area *dA* of that surface. $E_e = d\Phi_e / dA$. The unit of irradiance is watts per metre squared ($W\ m^{-2}$)

Thus the irradiance produced by a point source is inversely proportional to the square of the distance; this is known as the inverse square law. Furthermore, the irradiance is proportional to the cosine of the angle between the direction of irradiation and the normal to the surface; this is known as the cosine

The **candela** is the luminous intensity, in a given direction, of a source that emits monochromatic radiation of frequency 540×10^{12} hertz and that has a radiant intensity in that direction of 1/683 watt per steradian. In other words, this definition states that at a specific frequency of green light, a monochromatic radiant intensity of $1/683\ W\ sr^{-1}$ will produce a luminous intensity of 1 candela. Therefore, at this specific frequency of green light, 1 watt of radiant flux is directly equal to 683 lumens.

b) Project circle of radius, r, onto a sphere of radius R. From spherical co-ordinate geometry. Surface area element, $dS = R\sin\alpha.d\theta * R.d\alpha = R^2\sin\alpha.\ d\theta.\ d\alpha$



Then, for $d\alpha = 0$ to α, and $d\theta = 0$ to 2π, then area of sphere segment, $S = \int_0^{2\pi} \int_0^{\alpha} R^2 \sin\alpha \, d\theta d\alpha$

$= 2\pi R^2(1-\cos\alpha)$

Solid angle $\Omega = A/R^2$, therefore $\Omega = 2\pi R^2(1-\cos\alpha) / R^2 \quad = 2\pi (1-\cos\alpha)$

c) When a scene is viewed in good lighting conditions (illuminance levels > 50 lux) the spectral response of the eye is not affected by the actual level of illumination. These are the conditions for photopic vision, in which the visual process is entirely governed by cone receptors. There are three different types of cone receptor in the normal human eye, which have three distinct visual pigments, and these provide the basis for trichromatic colour vision, i.e. under photopic conditions, we have colour vision. The photopic condition is the most common one, the condition under which most observations are made and, most important from the photometric point of view, the eye is in a stable state.

Under very dim lighting conditions (illuminance levels < 0.05 lux) only the rods are sensitive and the eye operates in quite a different mode. This is the region of scotopic vision and extends down to the visual threshold which corresponds to an illuminance in the region of three microlux. Since there is only one type of visual pigment associated with rods (rhodopsin), there is no colour perception and the world is grey. However, this is also a situation in which the eye is operating in a stable condition. One of the anomalies of the scotopic range of vision is that an object may be visible 'out of the corner of they eye' (in the periphery) but invisible when viewed directly. This is a direct consequence of the absence of rods in the foveal region.

The ability of optical radiation to produce a visual response is termed the luminous efficacy of radiation, denoted *K*. It is defined as the quotient of luminous flux by the corresponding radiant flux, measured in lumen per watt (lm W^{-1}). $K = \Phi_v / \Phi_e$ There are two scales of luminous efficacy of radiation, one for photopic adaption *K* and one for scotopic adaption *K'*. Both *K* and *K'* are functions of wavelength; figure15 shows these functions, $K(\lambda)$ and $K'(\lambda)$, plotted against wavelength.

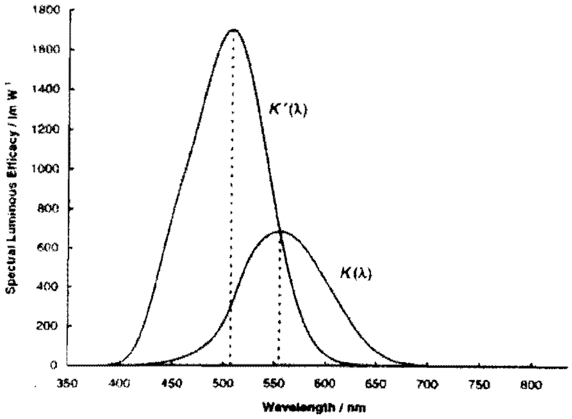


Figure . Spectral Luminous efficacy functions for photopic and scotopic vision.

To calculate the luminous flux in lumen produced by a radiant flux measured in watts, the radiant flux is weighted using the appropriate spectral luminous efficacy function $K(\lambda)$ or $K'(\lambda)$. Thus for photopic conditions:

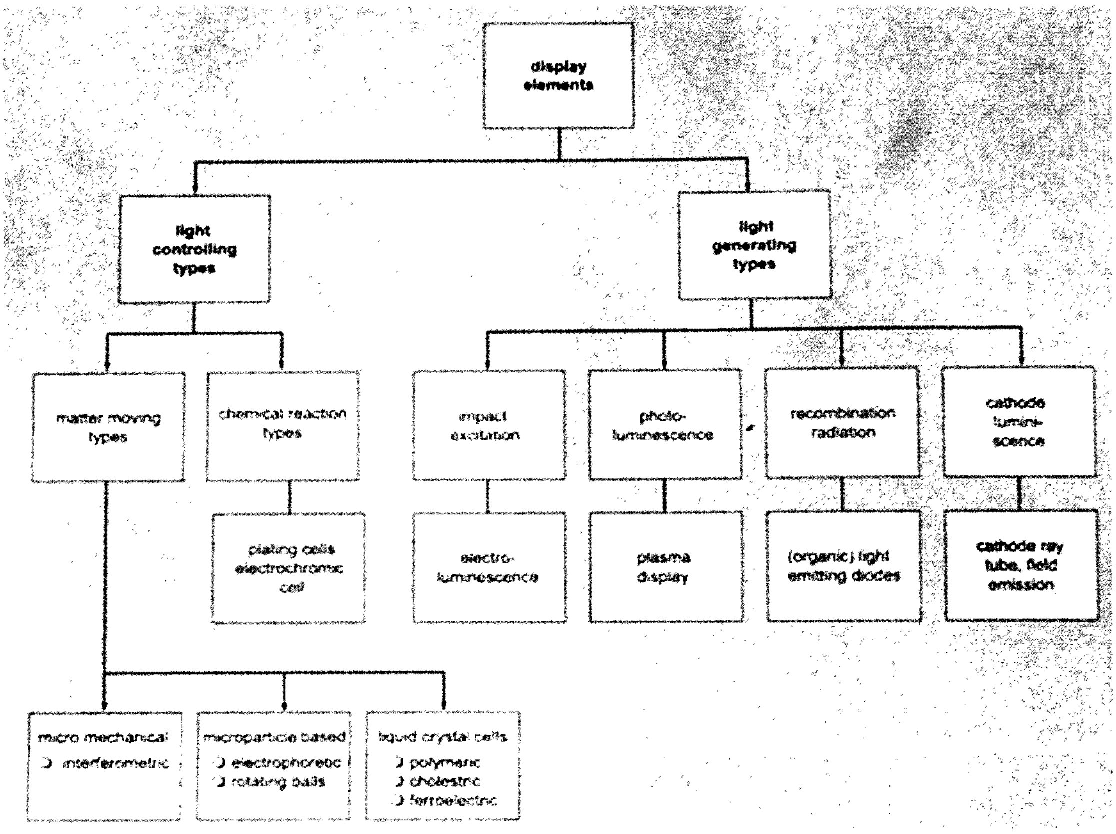
$$\Phi_v = \int_{360nm}^{830nm} \Phi_{e\lambda}(\lambda) K(\lambda) d\lambda$$

And for scotopic vision:

$$\Phi'_v = \int_{380nm}^{780nm} \Phi_{e\lambda}(\lambda) K'(\lambda) d\lambda$$

This means that there are two scales of luminous flux and hence two values of lumen, the photopic lumen and the scotopic lumen. It is not always realised that these are two totally different units, and are as different as a degree Fahrenheit is from a degree Celsius. For lighting applications it is usually only the photopic scale that is of interest, and this is the scale that we shall concentrate on during this course. The scotopic scale is generally restricted for special situations where the eye is dark adapted, and so the interest in scotopic measurements is very limited.

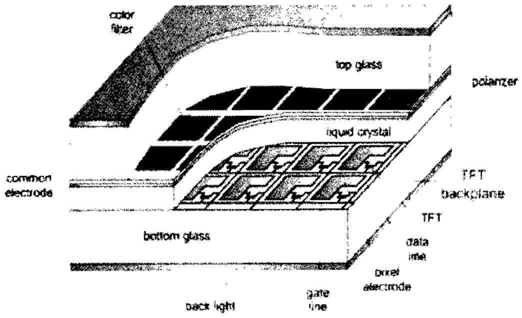
Q2 a) The differences between light controlling and light generating displays can be classified by their physical mechanisms used to convert the electronic energy to an optical signal as shown in the diagram below:



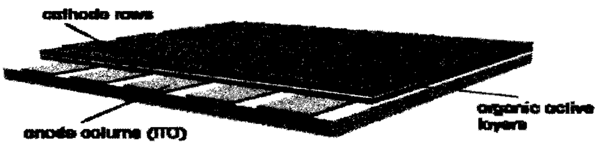
Light generating displays (e.g. CRTs, OLEDs, plasmas) generate photons as part of the inherent display process. Ie the modulation of the light is done at in the same process as the generaton.

Light controlling displays (e.g. LCDs, MEMs, electrophoretic, interferometric) interact (diffuse, polarize, absorb, or depolarize) with ambient or other light. Ie they modulate a source of photons which is not part of the modulation process itself.

b) A display consists of a matrix of picture elements or pixels. Each pixel can be addressed by a passive or active matrix architecture. In a passive matrix display, the entire row and column of the matrix is switched, in which the overlapping electrode areas define the pixel. This can result in smearing of the picture, which serves to undermine the resolution. In an active matrix, each pixel is uniquely turned on and off by a transistor. Active matrix displays are essential for high information content.

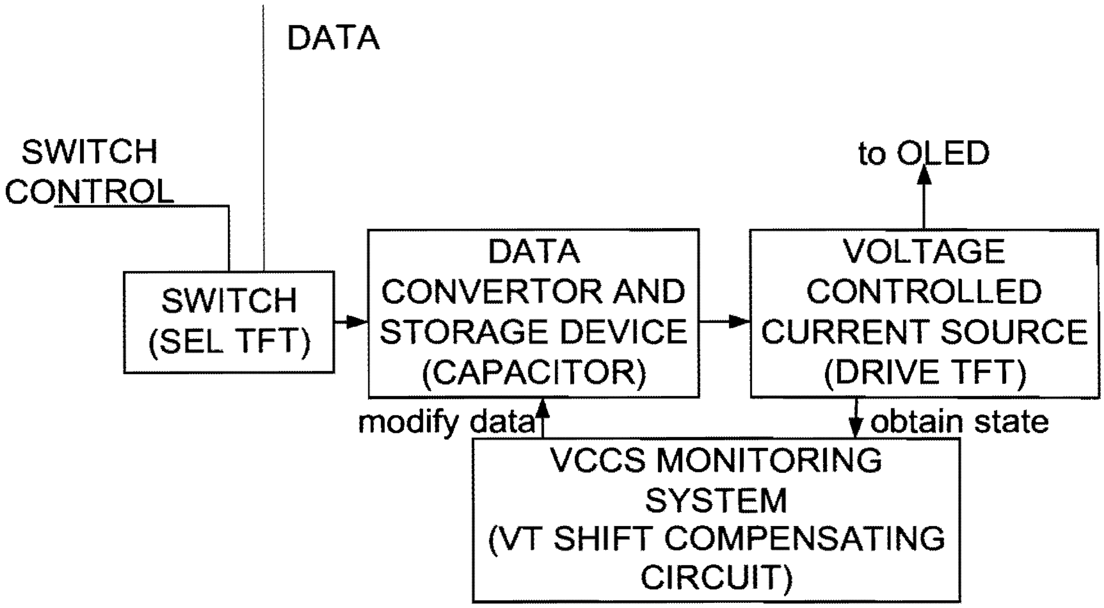


Active matrix illustrated with the liquid crystal display.



Passive matrix illustrated with organic light emitting diode) display.

In addition to the switching transistor used in the AMLCD pixel, the AMOLED pixel requires a current driving stage (since the OLED is a current-driven device) and a stage to compensate for V_T-shift in the TFT.



c) Requirements are:

- Low processing temperature: ~300°C glass, ~350°C metal foils, ~150°C plastic.
- TFTs with low leakage current, high ON/OFF ratio, low voltage operation, and small area.

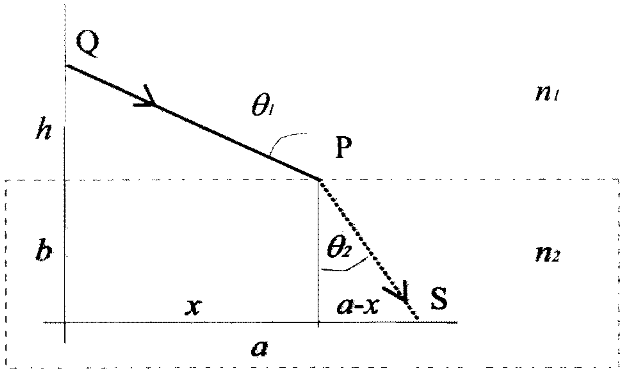
- High uniformity/ device matching, high stability and lifetime.

Summary of key performance attributes:

Attribute	a-Si:H	poly-Si	uc/nc-Si:H	organic	oxides
TFT type	*n-type*	*n-type/p-type*	*n-type/p-type*	*p-type*	*n-type*
Process Temp (°C)	*150-300*	*>250*	*150-300*	*100*	*<100*
Mobility *(cm²/Vs)*	*< 1*	*10~100*	*1~100*	*~ 1*	*1~100*
Temporal Stability (ΔV_T)	*can be managed*	*more stable than a-Si:H*	*more stable than a-Si:H*	*improving*	*research*
Initial Uniformity	*high*	*low*	*potentially high*	*low*	*research*
Scalability	*promising*	*issue*	*research*	*promising*	*promising*
Cost	*low*	*high*	*low*	*low*	*low*

Students should elaborate on the entries they provide for each technology.

Q3 a) Pure ray theory only contains information about the direction the ray is travelling in, to solve even simple problems like reflection or refraction more information such as time and velocity are required. A ray propagating through a vacuum (which is approximately the same for air) will travel at the speed of light c, however a ray passing through any other medium will travel at a slower speed such that $c_2 = c / n_2$ where n_2 is referred to as the refractive index of that medium. A ray travelling from one medium will be refracted as demonstrated in a way that is described by Snell's law. The proof comes from Fermat's principle which states the ray shout transit through the system in the shortest possible time.



$$t = \frac{\overline{QP}}{v_1} + \frac{\overline{PS}}{v_2}$$

Which can be translated into physical co-ordinates.

$$t = \frac{\sqrt{h^2 + x^2}}{v_1} + \frac{\sqrt{b^2 + (a-x)^2}}{v_2}$$

This shortest time will be when t is minimised with respect to x, hence we set $dt/dx = 0$

$$\frac{dt}{dx} = \frac{x}{v_1\sqrt{h^2 + x^2}} + \frac{-(x-a)}{v_2\sqrt{b^2 + (a-x)^2}} = 0$$

Which can be re-written to give.

$$\frac{\sin\theta_1}{v_1} = \frac{\sin\theta_2}{v_2}$$

Which is in fact Snell's law $n_1 \sin\theta_1 = n_2 \sin\theta_2$

b) This is one of the fundamental properties used to describe ray propagation through optical systems as the interaction between media such as glass and plastic surfaces can be used to control and focus the direction of the light in displays. The principle of Snell's law can be used to solve the optical problem of light propagation through a thin wedge shaped prism. The refraction of the rays at each surface dictate the how light will pass through the thin prism.

This is the basic principle used in all most geometric ray problems. Deviation from small values of apex and entrance angles lead to aberration in the optical system, hence these principles form a solid basis for good lens design and minimisation or potential aberrations. They do, however limit what can be done in an optical system, especially if size is a constraint. A good example of how this property can be is when a thin lens is made from a series of thin prism sections.

If we consider the simplest thin lens interaction we must make some assumptions in order to simplify the mathematics. As has already been stated from the prism analysis, the rays must enter the prism at a shallow angle in order for the approximation of $\sin x \approx x$ to hold. This is often referred to as the *first order* analysis or *the paraxial ray* approximation as it only applies to rays which are close or parallel to the optical axis. A more detailed analysis is the *third order* approach, which assumes that $\sin x \approx x - x^3/3!$.

The paraxial approximation is very useful for setting up basic systems, but it will also lead to aberrations. Hence any optical system must undergo a series of optimisations after the initial approximation to minimise the desired aberrations. No lens system will be perfect, so compromises must be made during the optimisation procedure, based on which aberrations will most effect the image quality of the lens system. This is usually done through a series of trial and error simulations using ray tracing software such as ZEMAX and CODE V. The main types of aberration are either point based aberrations such as spherical, coma and astigmatism or image based aberrations such as Petzval field curvature and distortion.

The analysis made so far has been using the paraxial ray approximation and that the lens will only operate at one wavelength, however this is rarely the case in real optical applications. Most lenses operate over range of wavelengths, which means that dispersion will have an effect, as each glass element will have a different refractive index at each wavelength. This is usually referred to as chromatic aberration.

c) The choice of modulation technology is a critical design step in any image projection system as it will limit performance and also other technology choices. The main two technologies for low cost projection are liquid crystal displays or MEMs devices such as the DLP. The simplest way analyse these choices uis to list each one:

LCD panels are cheap so 3 displays can be used, one foe each colour. This means that a simple while light source or LEDs can be used and split into 3 channels. The main drawback of using an LCD is that they rely on polarisation to modulate the light hence they need polarisors (capable of taking high power) and ideally a polarised light source (not trivial). The brightness will always be limited due to the absorptive nature of the modulation, however greyscale is inherently possible. The colour gamut is set by the source and the dichroic mirrors used to split the colours. Can use LED sort lasers to improve this gamut. There will be few motion artefacts, but they are too slow for full frame video. They will also produce polarised images unless diffusers are used. There is not much impact on the projection lens.

MEMs such as the DLP are more optically efficient (mirrors) and therefore brighter. There is less pressure on the choice of optical source as the effects of polarisation are negligible. The main limitation is that multichip systems are expensive and complicated in terms of electronics, hence with the single chip version a colour wheel is required. This limits the brightness and the colour gamut and also creates colour based motion artefacts. Solid state sources are possible for improved colour gamut. Another problem is that greyscale must be time sequentially dithered which adds to the problems of motion artefacts. It has little impact on the projection lens.

If 3D movies are to be projected, then the issue of polarisation becomes important. With systems like Real-D an add on polarisation switch is used to create stereoscopic views and a polarisation marinating screen is required. This works very well with MEMs system but not with LCDs. Other 3D systems do not use polarisation and do not require the fancy screens, and can work with either LCDs or MEMs but require more complicated viewing glasses.

4B20 Exam question – crib

Describe, with the aid of diagrams, the operating principles and basic construction of a twisted nematic (TN) liquid crystal display (LCD) device.

(a) A schematic of the twisted nematic device is shown in the figure. Illustrations should show, correctly, the arrangement of the polarizers, glass substrates, ITO coating, alignment layer, and colour filters.

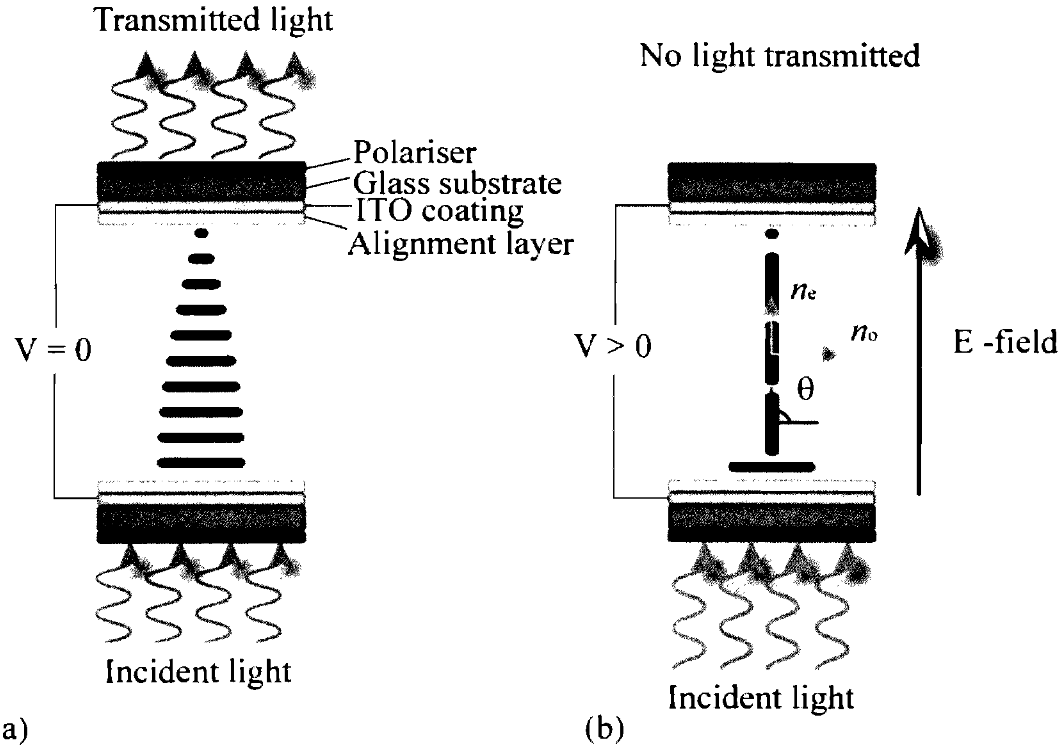


Figure. Cell construction consists of polarizers laminated on to the glass substrates.

The important features are that the rubbing directions are twisted relative to each other and that, as a result, the liquid crystal has a helical configuration. The propagation through helical birefringent medium can be described by two elliptically polarized eigenmodes. The twisted nematic device can operate in one of two modes: the first is the *normally white mode* whereby the transmission direction of the polarizer is parallel to the rubbing direction on the respective glass substrate. Without an electric field, the twisted structure cause the plane of polarization of the incident light that has passed through the first polarizer to rotate so that it passes through the second polarizer (analyzer). However, when an electric field is applied across the cell, the LC molecules reorient to form a predominantly homeotropic alignment (with the exception of the molecules in the vicinity of the substrates), which is optically inactive. The device can also function in the *normally black mode* whereby the transmission directions of the polarizers are parallel with respect to each other. In the absence of an electric field no light is transmitted but the device becomes optically active with the application of an electric field. For the TN-LCDs, liquid crystals with a positive dielectric anisotropy are required to ensure that a homeotropic alignment is formed with an applied electric field. A sketch of the transmission as a function of voltage to show grayscale capability should accompany the diagrams explaining the operating principle of the TN device. [*Marks should be awarded for an accurate illustration of the device and a description of the two possible modes of operation. Marks should also be awarded for the requirements on the liquid crystal formulation, e.g. a positive dielectric anisotropy*].

Explain the meaning of the terms, threshold voltage, Mauguin limit, voltage holding ratio, and isocontrast curve.

The Mauguin limit is the condition for waveguiding along a twisted structure. This can be expressed as,

$$\phi << \frac{2\pi}{\lambda}\Delta nd,$$

where ϕ is the twist angle, Δn is the birefringence, and d is the film thickness. For a TN-LCD with $\phi = 90$ the condition reduces to

$$\frac{\lambda}{4} << \Delta nd.$$

When this condition is not satisfied the output beam is elliptically polarized. [*Marks should be awarded for the correct definition and for writing down the relationship for the twist angle*].

The voltage holding ratio is important for active matrix LCDs and represents the ability of the device to hold the charge. It is defined as the ratio of the voltage (root-mean-square) at a pixel within one frame period to the initial voltage. This can be written as

$$VHR = \frac{1}{V_0}\left[\frac{1}{T_f}\int_0^{T_f}\left(V_0 e^{-\frac{t}{RC}}\right)^2 dt\right]^{\frac{1}{2}},$$

where V_0 is the voltage applied of the addressing and RC is the time constant of the pixel. Active matrix LCDs store charges at the pixel until that pixel is addressed in the next frame. The VHR depends upon a number of factors including the alignment layer and the polarity of the liquid crystal. Generally, the large values of the dielectric anisotropy result in low VHRs. Impurities in the LC mixture can significantly affect the VHR. [*Marks should be awarded for an accurate description of the VHR and why it is important*].

The isocontrast curve is a means with which to visualise the viewing angle dependency of a device. It is a polar plot consisting of both the polar angle, defined as the angle between the observation direction and the direction normal to the device, and the azimuth angle, which is defined as the angle between the transmission axis of the first polarizer and the projection of the observation direction onto the cell.

The redistribution of the director in the TN device occurs above a particular value of the voltage, known as the threshold voltage, and is related to the liquid crystal properties.

$$V_{th} = \pi\sqrt{\frac{K_{11}}{\Delta\varepsilon\varepsilon_0}\left[1+\left(\frac{K_{33}-2K_{22}}{4K_{11}}\right)\right]}^{\frac{1}{2}}$$

Where K_{11}, K_{22}, and K_{33}, represent the splay, twist, and bend deformations of the liquid crystal, respectively. [*Marks should be awarded for correct definition and any pictures/graphs which illustrate the threshold effect. Do not need to know the exact form of the equation but should know what LC parameters are important*].

How does the super-twisted nematic device differ from the TN device and what are the benefits of this mode?

The supertwisted nematic device has a greater twist angle which results in a steeper electro-optic curve making it more suitable for multiplexing compared to the TN device. However, grayscale can be more difficult for the STN mode. [*The steeper curve could be illustrated by a sketch on a graph of the transmission as a function of the voltage comparing it with the TN case*].

(b) ***Sketch the Gooch-Tarry curve for the transmission as a function of the optical retardation and find the value of u for the 1^{st} three successive minima.***

The Gooch-Tarry curve

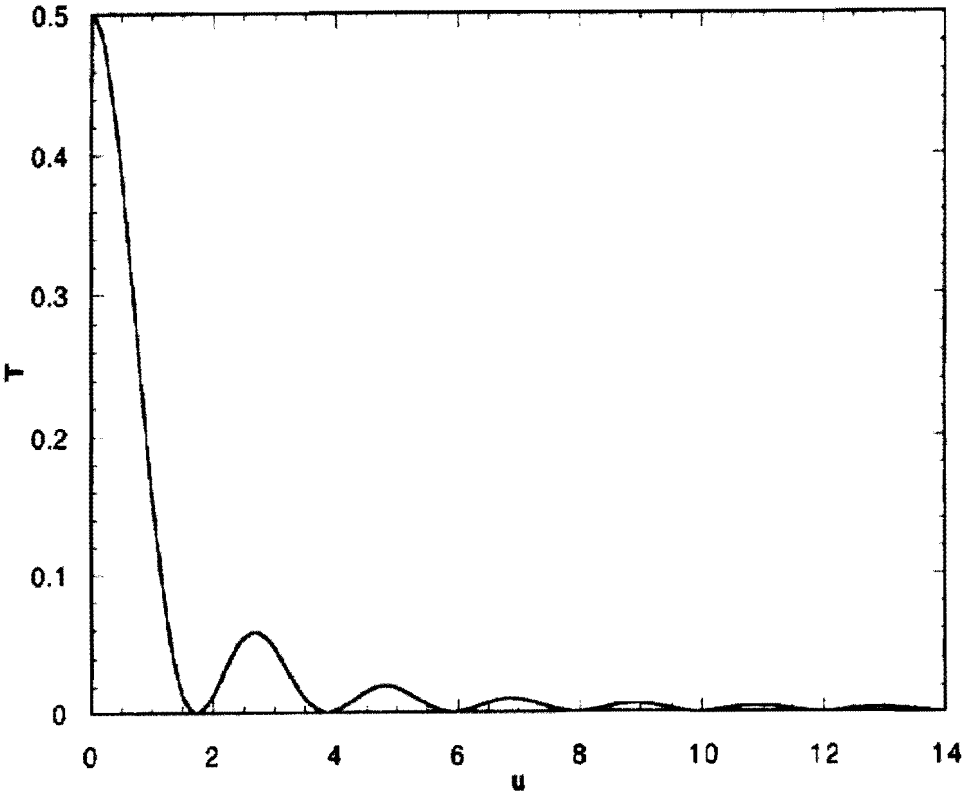


Figure. The Gooch-Tarry curve, transmission as a function of the optical retardation.

To find the values of the optical retardation, u, at the 1st three minima, the transmission is given in the equation:

$$T = \frac{1}{2}\frac{\sin^2\left(\frac{\pi}{2}\sqrt{1+u^2}\right)}{1+u^2}$$

For the minima, $T = 0$, therefore, $\sin^2\left(\frac{\pi}{2}\sqrt{1+u^2}\right) = 0$, and so $n\pi = \frac{\pi}{2}\sqrt{1+u^2}$, where n is an integer. It is then straightforward to show that,

1st minimum, $u = \sqrt{3}$,

2nd minimum, $u = \sqrt{15}$,

3rd minimum, $u = \sqrt{35}$.

If the birefringence of the liquid crystal is $\Delta n = 0.10$ and the incident wavelength is $\lambda = 550$ nm, what film thickness is required to operate at either the first or second minima?

To find film thickness, it is necessary to use the expression for the optical retardation, u, in terms of the birefringence. [*Marks are awarded for writing the expression for the optical retardation in terms of d and Δn and then correctly calculating the value of the film thickness using the information given*].

1st minimum

$u = \sqrt{3} = \frac{2d\Delta n}{\lambda}$. Therefore, $d = 4.8$ μm

2nd minimum

$u = \sqrt{15} = \frac{2d\Delta n}{\lambda}$. Therefore, $d = 10$ μm.

When constructing a TN-LCD, which minimum is preferred and why?

The first minimum is generally preferred due to the smaller viewing angle dependence and the faster response time.

(c) ***Describe, with the aid of diagrams, in-plane switching and vertically aligned nematic display devices.***

Illustrations of the basic operating principles of the in-plane switching and vertically aligned nematic devices should resemble something similar to the figures shown.

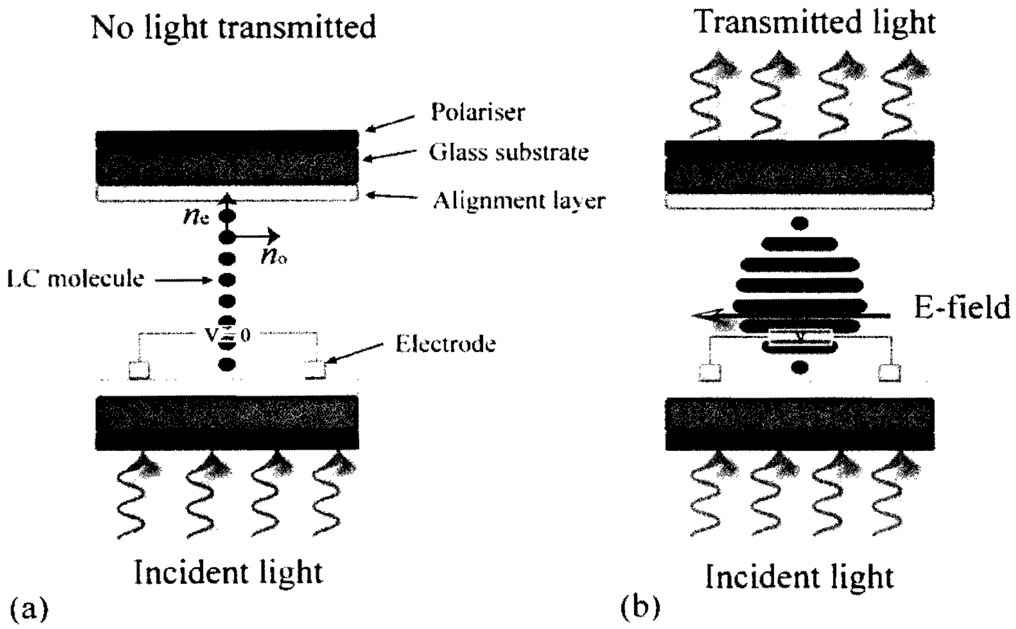


Figure. Illustration of an in-plane switching device.

For the in-plane switching mode, the electrodes are placed on one glass substrate. Between the electrodes, the electric field is, to a first order approximation, parallel to the surface of the substrate. As a result, the LC molecules are switched in the plane of the device, hence its name. In the absence of an electric field, light is blocked by the analyzer as the homogeneous alignment of the LC results in no optical activity. However, with the application of an electric field, the molecules reorient in the direction of the field to form a twisted structure as shown in the figure. Consequently, light is transmitted through the analyzer. The main benefit of the IPS mode is the viewing angle and this mode can provide a very wide viewing angle without additional optical compensation films. [*Marks should be awarded for accurate illustration of the devices*].

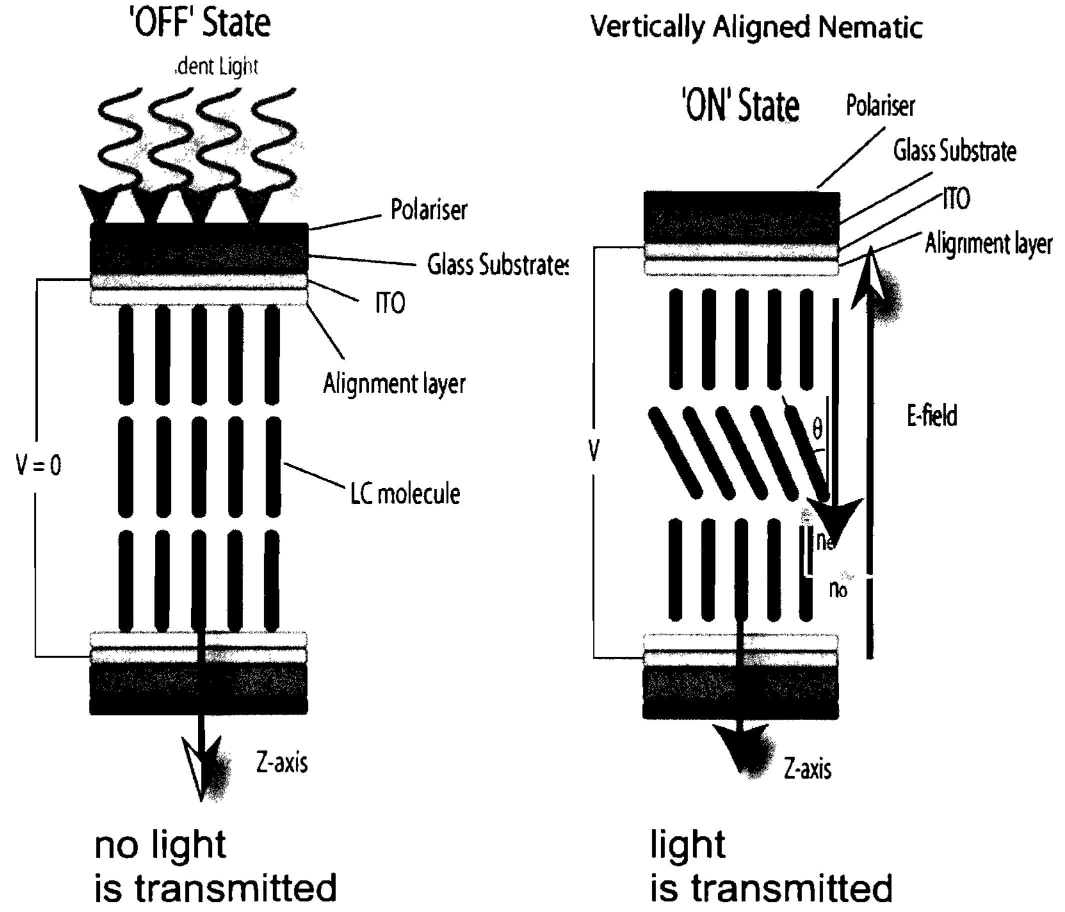


Figure. Illustration of the vertically aligned nematic device.

For a vertically aligned nematic device, incoming light that is polarized parallel to the normal of the cell does not encounter a birefringence and therefore its polarization state is unchanged at the second polarizer. These devices are typically operated in the normally black mode and therefore the polarizers are crossed. This results in a dark state for all wavelengths and is independent of the film thickness. VAN devices exhibit an excellent black state which leads to a very high contrast ratio. For an electric field applied across the cell, negative dielectric anisotropy LCs are used and the molecules reorient to align perpendicular to the field. The device then becomes optically active and light is transmitted through the analyzer. To avoid reverse tilt domains, the homeotropic alignment requires a preferred direction of switching. This is often achieved using multidomain vertically aligned LCD. The VAN-LCD has excellent contrast ratios.

Specify whether positive or negative dielectric anisotropy liquid crystals are required for each display mode?

In both cases nematic liquid crystals are used. Typically, the IPS mode uses positive dielectric anisotropy materials whereas the VAN mode uses negative dielectric materials.

(d) ***Sketch the equivalent circuit diagrams of a twisted nematic and in-plane switching display device.***

The equivalent circuits for the TN device and IPS device are shown in the following figures.

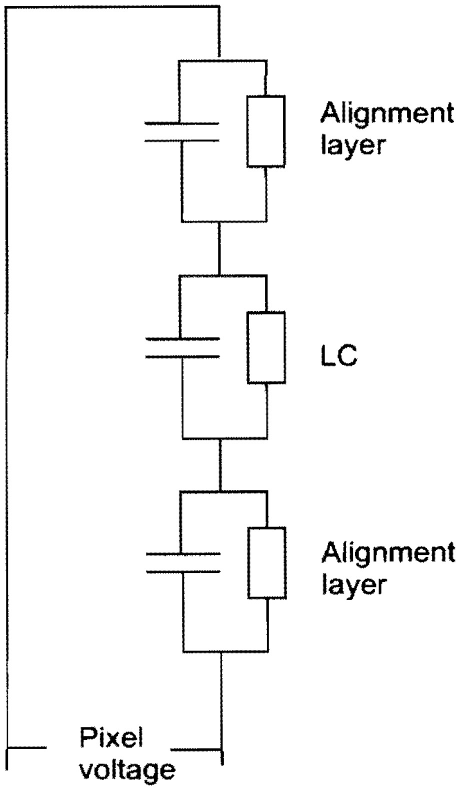


Figure. Equivalent circuit of a TN device. Series connection of the alignment layers with the LC.

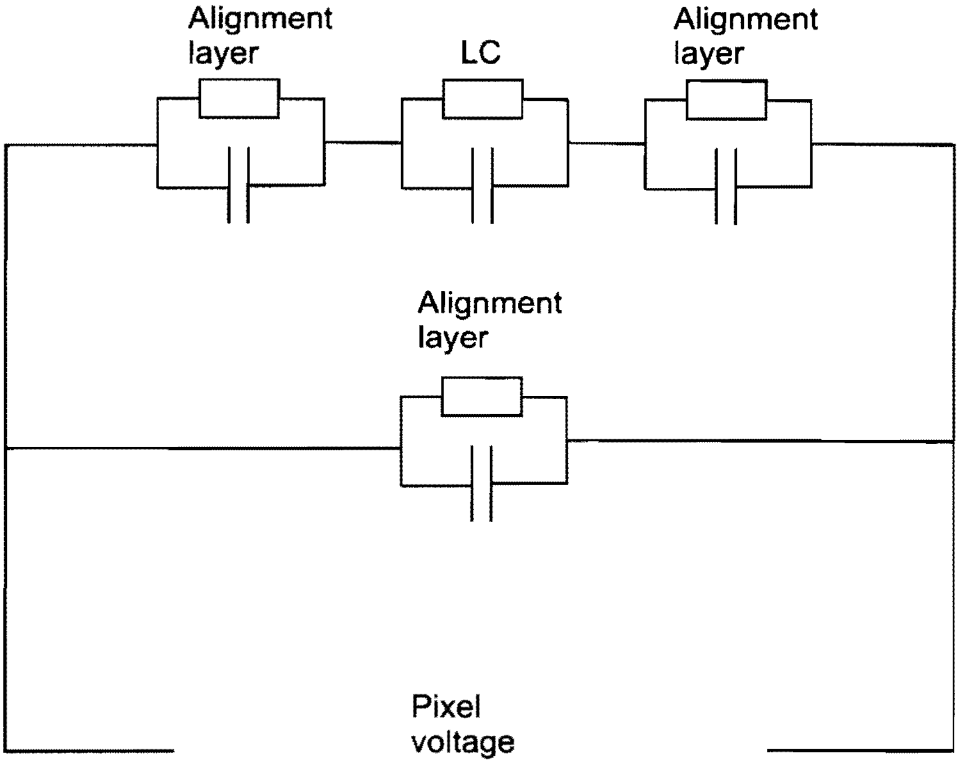


Figure. Equivalent circuit of an IPS device. Parallel connection of the alignment layers and the LC.

For which of the two display modes does the resistance of the liquid crystal have a greater impact on the voltage holding ratio?

The resistance of the LC is more important for the TN mode than the IPS mode because of the series connection for the former. The total resistance of the IPS cell is less dependent upon the resistance of the LC than is the case for the TN cell because of the parallel connection. It is then possible to use LCs with a lower resistance for the IPS mode but still maintain a sufficient VHR.

Specify which mode can use highly polar cyano-based liquid crystal structures.

The IPS mode can use highly polar cyano liquid crystals which are not compatible with TN-TFT devices. This is because lower resistance LCs can be used.

Examiner's comments:

Question 1:

Well answered. Basic book work with a few twists. Most understood the radiometric units and the limits of visual acuity.

Question 2:

This question was avoided by all candidates and was probably a little too vague when compared with the others in the exam. The material was taught in a more unusual way which will be revised for next year's course.

Question 3:

Very well answered question with good understanding of classical optics and its limits. Some of the later modulation comparisons were not a clear as should be but on average ok.

Question 4:

Well answered questions and a good understanding of LCDs. The last section was not so clear and only a few got there..

T.D. Wilkinson