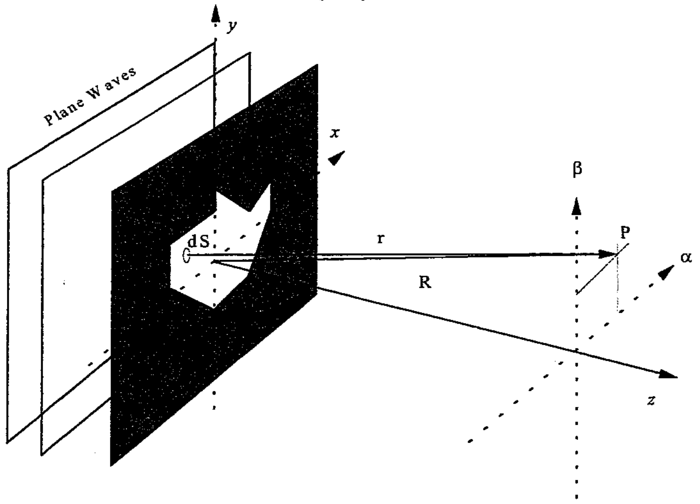
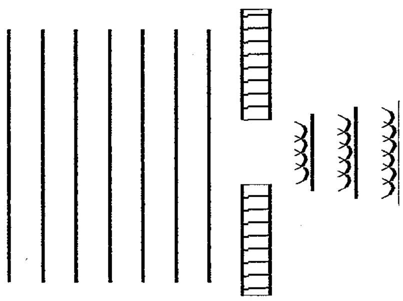
IIB EXAMS 2011

PHOTONIC SYSTEMS

DR TIM WILKINSON

4B11 exam cribs 2011: Q1 a)





If we consider an infinitely small differential of the aperture, *dS*, we can model this as a point source of light emitting spherical 'Huygens' wavelets with an amplitude of *A(x,y)dS*. The wavelet acts as a radiating point source, so we can calculate its field at the point *P*, a distance *r* from *dS*. The point source *dS* can be considered to radiate a spherical wave front of frequency ω. As soon as the wavelets propagate from the aperture, the edges (where adjacent wavelets are missing) begin to propagate with reduced energy. This changes the summation of the phases and after a few wavelengths of propagation, the plane waves are distorted and wrinkled. This is the near field region.

Assumptions: Plane wave coherent (laser) illumination of the aperture. Only consider forward propagating waves. Aperture is larger than wavelength x2.

b) In order to understand and analyse the propagating wavelets, a series of approximations and assumptions must be made. If we consider only the part of the wavelets which are propagating in the forward (+*z*) direction and are contained in a cone of small angles away from the z axis, then we can evaluate the change in field *dE* at the point *P*, due to *dS*. As the wavelet *dS* acts as a point source, we can say that the power radiated is proportional to $1/r^2$ (spherical wavefront), hence the field *dE* will be proportional to $1/r$. We can see that for a real propagating wave of frequency ω and wave number *k*, ($k = 2\pi/\lambda$) we have the cosine component of a complex wave. The full complex field radiating from the aperture can be written in terms of exponentials as the cosine is just the real part of the complex exponential.

$$dE = \frac{A(x,y)}{r} e^{j\omega t} e^{-jkr} dS$$

Now, we need to change coordinates to the plane containing the point *P*, which are defined as [α,β].

$$r = R\sqrt{1 - \frac{2\alpha x + 2\beta y}{R^2} + \frac{x^2 + y^2}{R^2}}$$

The final full expression in terms of *x* and *y* ($dS = dxdy$) for *dE* will now be.

$$dE = \frac{A(x,y)\, e^{j\omega t} e^{-jkR\sqrt{1 - \frac{2\alpha x + 2\beta y}{R^2} + \frac{x^2 + y^2}{R^2}}}}{R\sqrt{1 - \frac{2\alpha x + 2\beta y}{R^2} + \frac{x^2 + y^2}{R^2}}} dxdy$$

Such an expression can only be solved directly for a few specific aperture functions. To account for an arbitrary aperture, we must approximate, simplify and restrict the regions in which we evaluate the diffracted pattern.

If the point P is reasonably coaxial (close to the z axis, relative to the distance R) and the aperture $A(x,y)$ is small compared to the distance R, therefore $r = R$. The similar expression in the exponential term in the top line of the original equation is not so simple. It can not be considered constant as small variations are amplified through the exponential. To simplify this section we must consider only the far field or Fraunhofer region where.

$$R^2 \gg x^2 + y^2$$

In this case, the final term in the exponential ($(x^2 + y^2)/R^2$) can be considered negligible. To further simplify, we use the binomial expansion,

$$\sqrt{1-d} = 1 - \frac{d}{2} - \frac{d^2}{8} \ldots$$

and keep the first two terms only to further simplify the exponential expression.

$$dE = \frac{A(x,y)}{R} e^{j(\omega t - kr)} e^{jk\left(\frac{\alpha x + \beta y}{R}\right)} dx dy$$

The total effect of the dS wavelets can be integrated across dE to get an expression for the far field or Fraunhofer diffraction pattern.

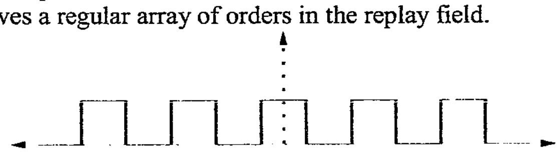
$$E(\alpha,\beta) = \frac{1}{R} e^{j(\omega t - kR)} \iint_{Aperture} A(x,y) e^{jk(\alpha x + \beta y)/R} dx dy$$

The initial exponential term $e^{j(\omega t - kR)}$ refers the wave to an origin at $t = 0$, but we are only interested in the scaling of relative points at P with respect to each other, so it is safe to normalise this term to 1. Thus, our final expression for the far field diffraction pattern becomes:

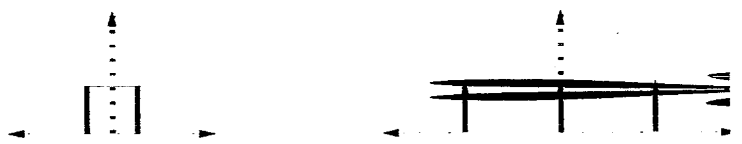
$$E(\alpha,\beta) = \iint_{A} A(x,y) e^{jk(\alpha x + \beta y)/R} dx dy$$

Hence the far field diffraction pattern at the point P is related to the aperture function $A(x,y)$, by the Fourier transform.

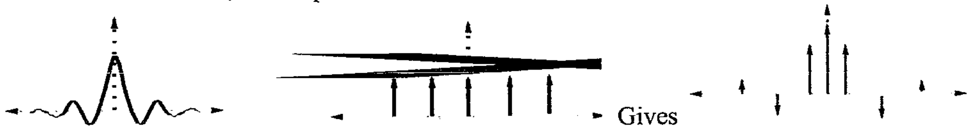
c) The pixel pitch and shape governs of the hologram defines the envelope function of the replay field and therefore define its overall physical size. The shape of the envelope is related to the shape of the pixel via the FT, hence a square pixel gives a sinc envelope. The pixel pitch also means that the hologram is effectively sampled, hence there will be an ordered harmonic structure via the FT. Ie a regular array of pixels gives a regular array of orders in the replay field.



Which can be expressed as a convolution of two functions, pixel shape and its pitch.



After the Fourier transform, the shape is a sinc with orders.

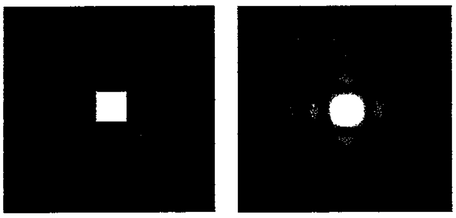


d) The sidelobes of the replay field have two fundamental origins. 1) The shape of the pixel is square, therefore the RPF is within a sinc envelope (with sidelobes). If we change the shape of the pixel to one which has no sidelobes in the Fourier domain then we remove them from the final hologram RPD. This could be done using pixels that have a Gaussian profile as the FT of a Gausian is another Gaussian function. 2) The repetition of the central order comes from the sampling of the pixels in the hologram. They are effectively sampled on a train of delta functions which FT's to a similar train of orders in the

RPF. This could be avoided by removing this structured nature of the sampling in the hologram. If the pixels were positioned at random intervals, then the sidelobes of the sinc function would contain no energy. Neither of these solutions are terribly practical however.

b) The FT can be found by finding the FT of a binary amplitude grating and then convolving a horizontal and a vertical pattern. This can then be converted to binary phase.

A single pixel has the FT of a sinc function.



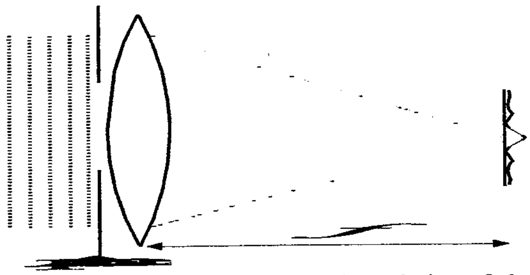
If we extend the pixel to infinity then is reduces the sinc width to a delta function. If we space the pixels on a regular grid, the we introduce odd harmonic orders.

Spacing is two times half of the sinc envelope first order $= \frac{f\lambda}{\Delta}$

Examiner's comment:

Well answered. Basic book work with a few twists. All coped with diffraction but a lot glossed over r = R. Virtually nobody go the origins of the outer orders.

Q2 a) Such a far field distance is clearly difficult to achieve in practical terms, so a means of shortening the distance is needed. If a positive focal length lens is included directly after the aperture, the far field pattern appears in the focal plane of the lens. A positive lens performs a Fourier transform of the aperture placed behind it.



The application of Snell's law at the spherical lens/air boundaries of the lens shows that the lens converts plane waves incident upon it into spherical waves convergent on the focal plane. For this reason, the diffraction to the far field pattern now occurs at in the focal plane of the lens. In order to display a grating or computer generated hologram (CGH) we need to be able to understand the effects that optical components will have on the replay field. The spatial coordinates (u,v) are related to the original absolute coordinates (α,β) by the relations.

$$u = \frac{k\alpha}{2\pi f} \qquad v = \frac{k\beta}{2\pi f}$$

Hence, the scaling of the FT is inversely proportional to f, and as the focal length shortens, the FT shrinks in dimension. Also, the replay field is a function of the wavelength λ, hence different wavelengths will produce different size replay fields. We can use the above relationships to directly calculate the positions of peaks in the hologram's replay field.

A two dimensional grating or hologram comprises of NxN square apertures (pixels) with a pixel pitch Δ having an amplitude A given by either amplitude modulation [0,1] or phase modulation [+1,-1]. The two dimensional envelope due to the fundamental pixel which covers the far field diffraction pattern (or FT) of the hologram is just a 2-D sinc function (where $a = \Delta$).

$$A\Delta^2 \mathrm{sinc}(\Delta u)\,\mathrm{sinc}(\Delta v)$$

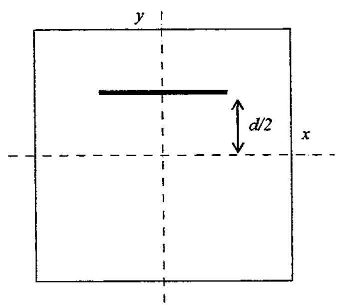
The useful information of the replay field is contained in the central first lobe of the sinc function, so we can calculate the width of the replay field as where the first zero of the sinc function occurs ($\pi\Delta u = \pi$, $\pi\Delta v = \pi$). We want the coordinates in terms of $[\alpha,\beta]$, so we use the above transformation to get.

$$\alpha_M = \frac{f\lambda}{\Delta} \qquad \beta_M = \frac{f\lambda}{\Delta}$$

The choice of modulation of the CGH has a very strong effect on the RPF. If a static hologram is made, then binary phase is the simplest and cheapest option, but that limits the RPF to a 180 degree symmetry and will have higher noise. Multi-level phase helps to break this symmetry but is harder to fabricate. A LC device could also be used – either nematic (multilevel but slow) or a FLC (binary but fast).

b) The target function in Fig 2 is 180 degree rotationally symmetric, hence we can use a binary phase scheme to create the CGH. If the CGH was to be a static or fixed design then we could use a patterned thin film layer or step function created by photolithography. Many different photoresists or polymers could be used as long as they were sufficiently transparent and the correct thickness for a pi phase shift between a step and a gap. If the CGH was to be dynamic, then an FLC SLM could be used, optimised for binary phase at that wavelength. It could be used without polarisors by using the right FLC material and pi retardance.

The target function does not need to have both lines due to the symmetry. It should include some of the background areas as well when calculated to help reduce the background noise as close to zero as possible.

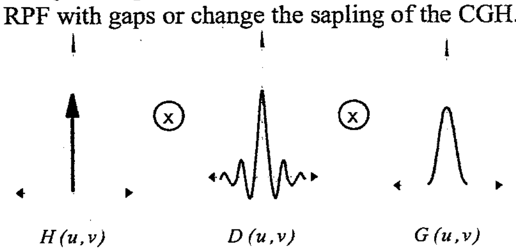


c) From this we can assume that an NxN pixel hologram will generate NxN spatial frequency 'pixels' in the replay field with the pitch.

$$\alpha_0 = \frac{f\lambda}{N\Delta} \qquad \beta_0 = \frac{f\lambda}{N\Delta}$$

Hence $\alpha_0 = \beta_0 = 20.8$um. For a 512x512 pixel CGH (or any even number of pixels), the central point of the plane can be set at either (256,256) or (255,255) as the actual centre lies between those points. Hence for the d = 100 pixel spacing it is possible to locate the bars exactly 50 pixels either side of this central point Hence it is possible to locate an exact 180 degree symmetrical position for the target, so when the CGH is calculated the spacing will be 100x20.8 = 2.08mm. In the case of d = 101 pixels, it is not possible to maintain the symmetry by ½ a pixel, hence the spacing will either have to be 100 or 102 pixels.

d) There are several factors that will lead to imperfections in the RPF generated. 1)The number of pixels and the pitch fix the aperture of the hologram, this then defines the shape of the spatial frequency 'pixels' in the replay field. This is apodisation and the delta functions above become a convolution of the FT of the aperture and for adjacent spots in the RPF, the sidelobes of D will interfere. Possible solution: space spots in the RPF with gaps or change the sapling of the CGH.

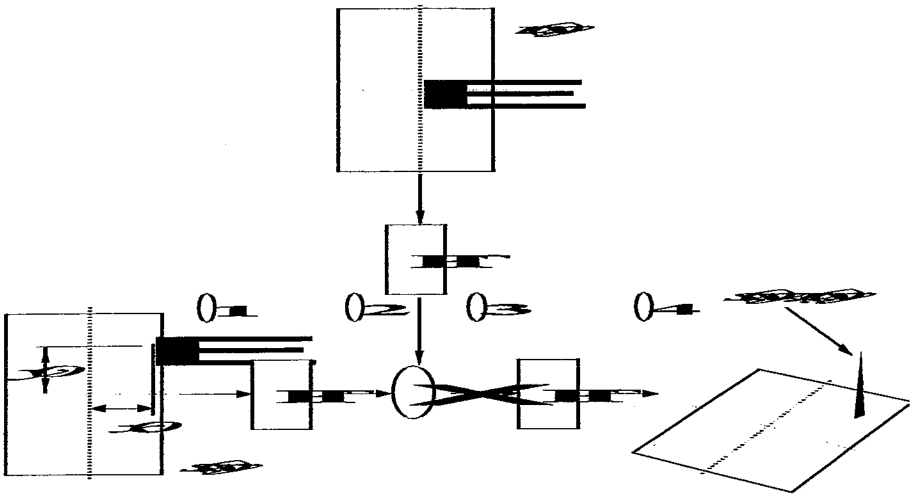


2) The central order of the CGH is the central lobe of the overall sinc envelope due to the square pixels. This means that even if the simulated CGH has a perfectly even amplitude for the bar, the actual bar will follow the profile of the sinc in the generated RPF. Solution: sinc compensate the target profile of the bar with the reciprocal values of the sinc envelope.
3) No lens is perfect and there will be aberrations. Solution : use a doublet lens or custom optics.

Examiner's comment:

A new question with most understanding the ideas behind CGHs well, but a few could not tell a technology from an algorithm. No one spotted the fact that a separation of 101 pixels was not possible with N even..

Q3 a) The matched filter architecture is laid out in a linear fashion.



The input image *s(x,y)* is displayed in plane 1 on SLM#1 before the FT into plane 2.

$$S(u,v)e^{-j2\pi(x_0u+y_0v)}$$

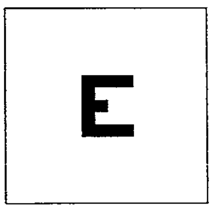
The FT of *s(x,y)* is then multiplied by the FT of the reference *r(x,y)*.

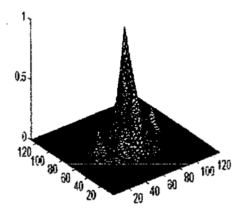
$$R(u,v)S(u,v)e^{-j2\pi(x_0u+y_0v)}$$

The FT of the reference is done off line on a computer and is defined as the matched filter *R(u,v)* for that particular reference *r(x,y)*. The filter is displayed on SLM#2. In fact, the generation of the filter may be more complicated (to include invariances).

The product of the input FT and the filter then undergoes a further FT to give the correlation in plane 4. The object in the reference *r(x,y)* is centred in the process of generating the filter *R(u,v)*, so that if a correlation peak occurs, its position is directly proportional to the object in the input image, with no need for any decoding. Unlike in the JTC, there is only one correlation peak and there are no DC terms to degrade the correlator output.

The best test for any matched filter is to perform an autocorrelation with the filter that has been generated. The reference image *r(x,y)* is used as the input to the correlator to judge its performance and the autocorrelation will have optimum SNR. The MF autocorrelation peak is very broad and has a huge SNR, as there is no appreciable noise in the outer regions of the correlation plane. Such a filter is not very useful for pattern recognition as such a broad peak could lead to confusion when the position of the peak is to be determined. Also, similar shaped objects (such as the letter F) will correlate well with the filter leading to incorrect recognition. Another identical E which is placed in the input along with the original one will also cause problems as the correlation peak will take an extremely complex structure. Finally, the filter is a complex function and there is no technology available to display the filter in an optical system.

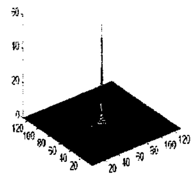




b) Great improvements can be made to the usefulness of the correlation peak, by using a phase only matched filter (POMF). The matched filter *F(u,v)* is stripped of its phase information (i.e. the phase angle of the complex data at each pixel) and this is used as the filter in the correlator.

$$F(u,v) = F_{amp}(u,v)e^{i\phi(u,v)} \qquad F_{POMF}(u,v) = e^{i\phi(u,v)}$$

The autocorrelation for the POMF is much more desirable even though there is a reduction in the SNR due to the increase in the background noise. The correlation peak is much narrower which is due to the information which is stored in the phase of the matched filter. The POMF is the most desirable filter to use as it has good narrow peaks but



still remains selective of similar structured objects (like Fs). The POMF is also a complex light modulation scheme, so the problems associated with binary phase (180° symmetry) will not occur. However, the continuous phase structure of $\phi(u,v)$ means that it cannot easily be displayed in an optical system. Twisted nematic displays are capable of multilevel phase modulation, but the quality is poor, difficult to control and they are slow. There is the possibility that new FLC crystals and phases may allow the implementation of four level phase modulation in the near future.

The penalties associated with going to binary phase are greatly out weighed by the advantages gained by using FLC SLMs in the optical system.

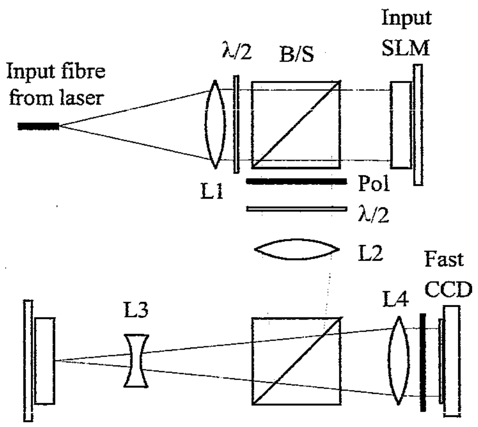
1) The SNR is up to 6dB worse than the case of the POMF.
2) The filter cannot differentiate between an object and the same object rotated by 180° (due to the fact that the BPOMF is a real function).
3) The BPOMF is not as selective as the POMF due to the loss of information in the thresholding.

The binary phase is selected from the POMF by two thresholds δ_1 and δ_2. These the thresholding is done such that.

$$F_{BPOMF} = \begin{cases} 0 & \delta_1 \le \phi(u,v) \le \delta_2 \\ \pi & \text{Otherwise} \end{cases}$$

The selection of the two boundaries is by exhaustive searching, as it depends on the shape and structure of the reference used to generate the filter. The benefits of this process are not high and it is only likely to improve the SNR by a few percent. A safe threshold to get consistent results is $\delta_1 = -\pi/2$, $\delta_2 = \pi/2$.

c) The BPOMF is made as follows:



The modulated light passes through lens f_0 which performs the FT of the input image. The FT is formed in the focal plane of the lens and will have a finite resolution (or 'pixel' pitch) given by.

$$\Delta_0 = \frac{f_0 \lambda}{N_1 \Delta_1}$$

There are N_1 'pixels' in the FT of the input image on SLM1, hence the total size of the FT will be $N_1\Delta_0$. The BPOMF is displayed on SLM2 in binary phase mode. SLM2 is also a FLC device with $N_2 x N_2$ pixels of pitch Δ_2. The FT of SLM1 must match pixel for pixel with the BPOMF on SLM2 in order for the correlation to occur. For this reason we must choose f_0 such that.

$$N_1 \Delta_0 = N_2 \Delta_2$$

Hence we can say.

$$f_0 = \frac{N_2 \Delta_2 \Delta_1}{\lambda}$$

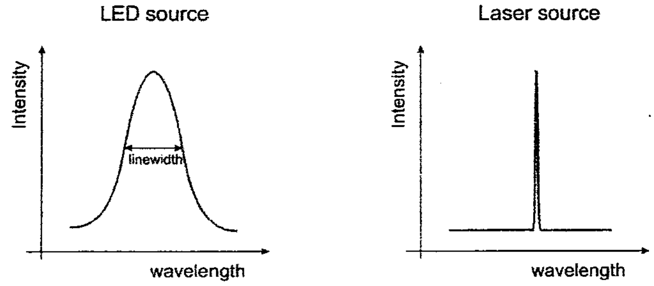
The required focal length to match the input FT to the BPOMF nearly always impractical as an experimental system as it would be physically too large. It is possible to shorten the actual length of the optical transform whilst still keeping the effective focal length that is desired by including further lenses in a combination lens. One technique is to combine a positive lens with a negative lens to make a two lens composite. This gives a length compression of around $f_0/5$ which in the example above is still 2m and impractical. Furthermore, the two lenses combine in aberrations which leads to poor correlations due to optical quality, hence a 3rd lens is needed at the filter SLM surface.

Examiner's comment:

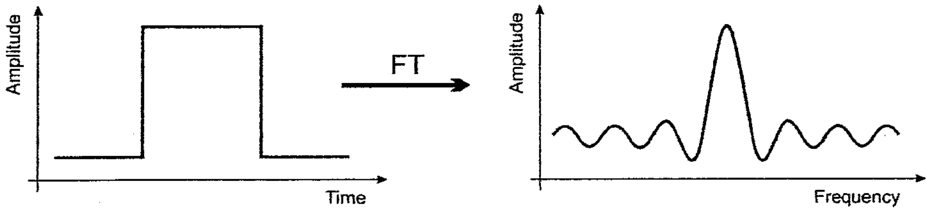
Standard bookwork answered well on the origins of patched filtering. A few missed the role of the POF in the evolution. Not many spotted that an LCOS device must be reflective when designing the last part.

Q4 a) The question specifies a monomode fibre. Therefore we only need consider intramodal dispersion (intermodal dispersion cannot occur). Dispersion can be thought of as the temporal broadening of optical pulses.

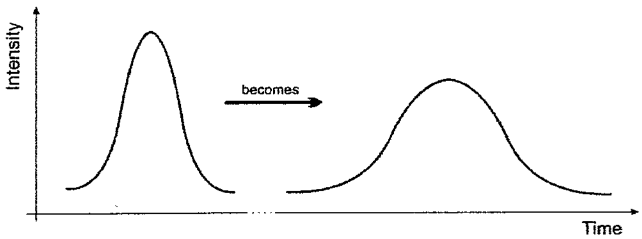
Intramodal dispersion occurs due to a range of frequencies being present in the optical source. This may be due to either a broad linewidth source (such as an LED, as opposed to a laser):

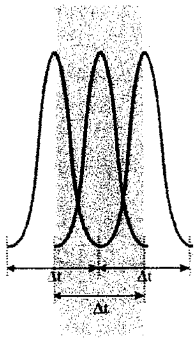


... or due to the frequency content of pulsed data packets:



These frequencies travel at different speeds down the fibre, and therefore arrive at different times. The pulse therefore spreads out temporally. If optical pulses spread too much, then sequential signals may overlap with one another and be lost. This limits time-division multiplexing, and puts a maximum limit to the data carrying capacity of the fibre.





b) (Eq.61) Bandwidth-Length product: $BL\,(\text{bits/s}\cdot\text{m}) = \dfrac{c}{2(n_1 - n_2)}$

Therefore, the bit-rate is simply: $B\,(\text{bits/s}) = \dfrac{c}{2L(n_1 - n_2)}$

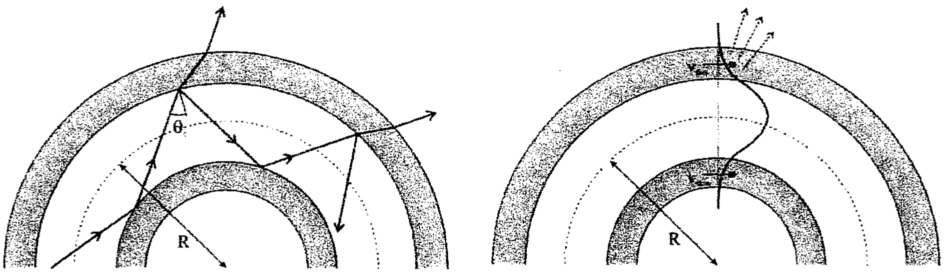
$c = 3\text{x}10^8$ m/s, L = 100 km, n_1 = 1.5000 (core), n_2 = 1.4999 (cladding)

Substituting values gives: B = 1.5 Mbits/s
(1 Byte = 8 bits) : = 187.5 kB/s
= 675 MB/hour

Therefore they can send **675 documents** (each of 1MB size) in 1 hour.

c) Bending of optical waveguides can cause the leakage of light (and therefore also data) out through the cladding. The competitor could therefore simply bend the fibre and "read" the data as it leaks out at that point.

To explain this, the simplest way is to use the ray model of light propagation within waveguides: Bending of the fibre causes the incident light within the core to strike the cladding interface at angles that exceed the critical angle for total internal reflection. The light is therefore partially reflected, and partially refracted out through the cladding.



Ray model EM model

Students may instead also provide a similar explanation using the EM model of wave propagation, whereby light at the outside edge of the bend, has to travel at a faster speed (exceeding the speed of light) in order to keep up with light on the inside. It cannot do this, and so is lost to the cladding.

We find the critical bending radius, R_c from the following (eq. 64): $R_c \approx \frac{a}{(NA)^2}$

Where, $NA = \sqrt{n_1^2 - n_2^2}$ (Eq. 13), so that: $R_c \approx \frac{a}{n_1^2 - n_2^2}$

Substituting the following values: a = 15 μm, n_1 = 1.500 (core), n_2 = 1.499 (cladding), gives: R_c = **5.002 mm.**

d) The initial experiment can be used to characterise the bending loss in the fibre and then that value used to estimate the loss of the 90degree bend.

(Eq. 51): Loss in optical power characterised by: $\frac{P(z)}{P(0)} = \exp(\alpha z)$

(Eq. 62): Bending loss coefficient: $\alpha_B = C \exp\left(-\frac{R}{R_c}\right)$

Combining these equations gives: $\frac{P(z)}{P(0)} = \exp\left[-zC \exp\left(\frac{-R}{R_c}\right)\right]$

We first substitute values for the "controlled experiment", where:
$P(z)/P(0) = ½$, R = 10 cm, $z = 10 \times 2\pi R$, R_c = 5.002 mm (from part (e)).

Rearrange for C gives: $C = \frac{\ln(P(z)/P(0))}{20\pi R . \exp(R/R_c)} = 5.3096 \times 10^7 \text{ m}^{-1}$

Now we have a value for C, we can use it in the 90° bend situation, where:
R = 7 cm, $z = ¼ \times 2\pi R$

Substituting these values: $\frac{P(z)}{P(0)} = \exp\left[-\frac{\pi R}{2} C \exp\left(\frac{-R}{R_c}\right)\right]$

Gives us the result: $\frac{P(z)}{P(0)} = 7.58 \times 10^{-3}$

Therefore the power loss is 0.758 %

Examiner's comment:

This question was still a little too easy but it did encourage people not to bin this section of the course. Not many got the final section right but there were a few solid attempts to put together the loss and bending ratio...