ENGINEERING TRIPOS PART IB, 2003

P8:

Introductory Business Economics
Question Cribs

1.

(a) Explain why a firm will continue producing in the short run even if it is making a loss, providing it can cover its variable costs. How long will it be willing to continue to make a loss? [6]

Fixed costs (rent, rates etc) have to be paid even if the firm is producing nothing at all. Providing, therefore, that the firm is more than covering its variable costs it can pay some of its fixed costs and will therefore continue to produce. It will shut down if it cannot cover its variable costs – the short-run shut down point. In the long run the firm will close down if the firm cannot cover its long-run average costs – the long-run shut-down point.

(b) For what reasons would you expect a monopoly to charge (i) a higher price and (ii) a lower price than if the industry was operating under perfect competition. [8]

(i) A monopolist will normally produce a lower output at a higher price than under perfect competition. This is because a monopolist is a price maker with a downward sloping demand curve. Conversely, in perfect competition each firm is a price taker with a perfectly elastic demand curve.

(ii) A monopoly may be able to exploit increasing returns to scale. If this produces in a marginal cost curve substantial below that of an industry under perfect competition, the monopoly may produce a higher output at a lower price.

(c) Explain why competition between oligopolists may reduce total industry profits. [6]

Oligopolists will want to maximize their joint profits. This will tend to make them collude to keep prices high. On the other hand they will want the biggest share of industry profits for themselves. This will tend to make them compete. Whether they collude or compete will depend upon the conditions in the industry. They are more likely to collude if there are few of them; if they are open with each other; if they have similar products and cost structures; if there is a dominant firm; if there are significant barriers to entry; if the market is stable; and if there is no government legislation to prevent collusion.

2.

(a) What factors may influence the level of investment in the macroeconomy? [10]

Rate of interest - investment is negatively related to the rate of interest (the marginal efficiency of capital) and savings are positively related to the rate of interest. Accelerator theory – which shows that investment depends on the level of economic activity. There are limitations to accelerator models (eg they ignore the role of the price mechanism and assume a constant capital-output ratio). Other factors include: uncertainty, profitability, technological change and government policy.

(b) What factors may influence the level of consumption in the macroeconomy? [10]

The Keynesian consumption function focuses on current income. Alternative theories (Life Cycle hypothesis and Permanent Income hypotheses) focus on 'normal' income and stress an important role for human and non-human wealth. Where as the Keynesian consumption function suggests that government can influence the level of economic activity through fiscal policy, the alternative theories suggest that such policies will have little impact.

Part IB, Paper 8, Selected Topics, Section B, Civil and Structural Engineering

Question 3 crib (note: principally essay question)

(a) NATM involves using sprayed concrete for temporary tunnel linings. It can be used to line any shaped tunnel (not just circular). Excavation sequences and face areas excavated can be continuously varied, according to the ground conditions and depending on measurements made.

NATM is most appropriate in ground that can remain temporarily stable when a hole is excavated in it without support, i.e. in soils such as stiff clays.

Potential risks centre around construction practice: if the workmanship is bad, the thickness of the sprayed concrete and competence of joints can be inadequate, thereby threatening structural stability of the lining – leading to risk of complete collapse (eg Heathrow tunnels). [3]

(b) Stability ratio, $N = (\sigma_v - \sigma_t)/ C_u$

where σ_v = total vertical stress at tunnel axis level
σ_t = temporary support pressure (if any) in the tunnel heading
C_u = undrained shear strength of the clay at tunnel axis level

For tunnel in stiff clay, $N = (15 \times 20)/ 150 = 2.0$

For tunnel in soft clay, $N = (15 \times 17)/30 = 8.5$

Face support is needed if N exceeds about 5. Hence it would be acceptable to construct the tunnel in stiff clay without any face support, ie in open face mode. But it would be necessary to provide face support in the case of the soft clay, and the tunnel would have to be constructed using closed face mode.

[4]

(c) Groundwater can be controlled when tunnelling in sands and gravels by the following alternative techniques (two only needed for question):

(i) *Compressed air*. The pressure is adjusted to balance the water pressure and prevent it flowing into the tunnel face.
(ii) *Ground treatment by dewatering*. This involves installing pumps in the soil and lowering the water pressures in the ground prior to tunnelling.
(iii) *Ground treatment by grouting*. This involves injecting the ground with cement or chemicals to fill the pores and reduce the permeability, so that water flow ceases to be a problem.
(iv) *Closed face tunnelling machines* – either slurry or earth balance pressure shields. In both cases there is a bulkhead in front of which the pressurized soil and water is contained, and from this the spoil is removed either by slurry circulation or by means of a screw auger that allows the pressure to drop to atmospheric.

(d) Settlement potentially causes damage to masonry buildings because *differential* settlement causes the building to experience curvature, which in turn leads to tensile strain. Masonry is vulnerable to tensile strain – this causes it to crack. The magnitude of the *total* settlement is not in itself significant, especially in cases where the settlement is uniform, because then differential settlements and curvatures are small and hence tensile strains are small. [3]

Possible protective measures are (two only needed for question):

(i) *Strengthening of the ground by grouting by chemical grouting.* This can only be achieved in sands or gravels, because to be effective the chemical must permeate through the pores of the ground. It is not applicable to fine grained soils such as silts or clays, because the pore size is far too small. It would only be applicable, therefore, to the third area of ground considered for the metro project (the sands and gravels).

(ii) *Strengthening of the building.* This involves installing tie rods to reduce the effects of tensile strains potentially being induced in the masonry. The applicability of this technique is *independent* of the ground type – and hence could be applied to the tunnels being considered for *all three ground conditions.*

(iii) Structural jacking. This involves isolating a building from its foundations, inserting jacks, and then using them to control potential differential settlement of the building. The applicability of this technique is also *independent* of the ground type – and hence it too could be applied to the tunnels being considered for *all three ground conditions.*

(iv) *Compensation grouting.* This involves the injection of grout into the ground at some point between the tunnel and the foundations of the building, in order to *compensate* for the ground movements being caused by tunnel excavation. There is no need for the grout to permeate the pores – it can simply displace the ground. Hence it can be used in *stiff clays or sands and gravels.* There are problems with soft clays, because undesirably high ground displacements are likely to be caused by the grouting.

[4]

(e) Instrumentation and monitoring are important for this metro scheme, because (three reasons only needed for question):

- Forewarning of unexpected events can be provided
- Tunnel construction procedures can be adjusted in a safe and systematic manner (the Observational Method)
- The quality of the work can be controlled
- Records of ground and building movements are needed to determine what, if any, buildings may need repairs

- The measurements are an important source of data for improvements in design and construction of future projects.

Examples of instrumentation that might be installed (three examples only needed for question):

- Settlement stations on ground surface and on buildings
- Subsurface extensometers to measure subsurface ground settlements
- Subsurface inclinometers to measure subsurface horizontal ground movements
- Piezometers to measure pore pressure changes in the ground (usually only in clays)
- Pressure cells to measure total stress changes in the ground (usually only in clays)

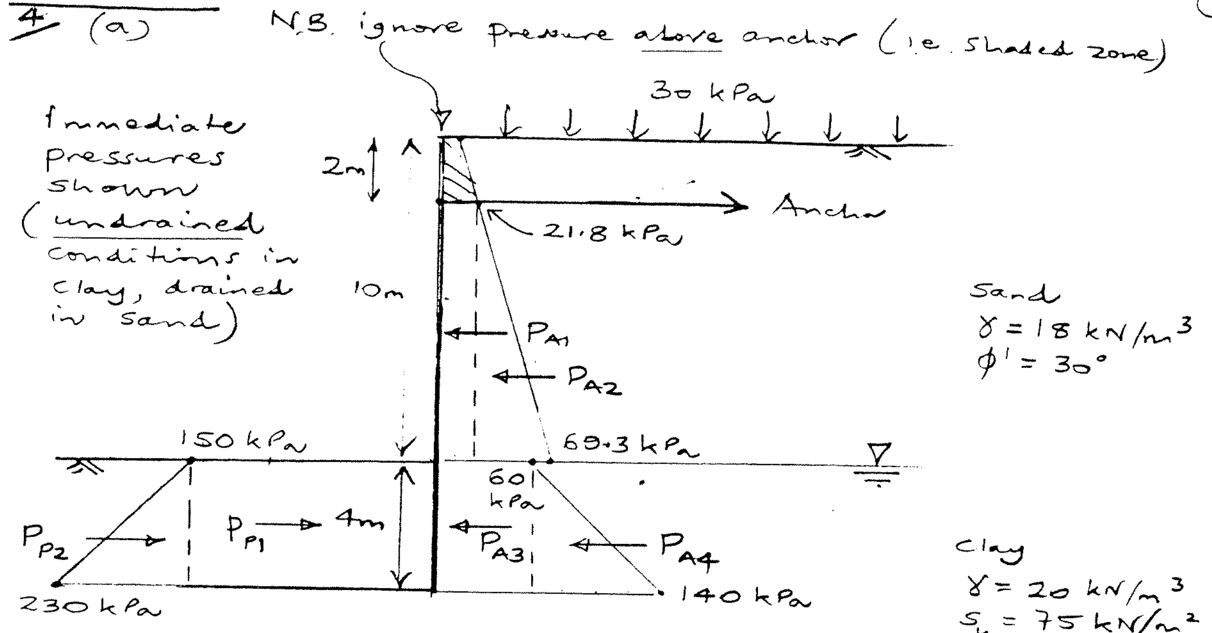
[6]

Comments from Examiner

The parts of the question relating to the New Austrian Tunnelling Method (NATM) and its risks, stability of tunnels and groundwater control were all generally well answered. Reasonably good answers were also provided to the part of the question relating to differential settlement, potential damage to buildings and protective measures. The final part relating to instrumentation and monitoring was not well answered. In general the standard of english was poor, with answers usually given in rather rough bullet point form (and often with spelling mistakes).

Question 4

4 (a)



Active side: $K_a = \dfrac{1-\sin\phi'}{1+\sin\phi'}$

$\phi' = 30°$ $\therefore$ $K_a = \dfrac{1-0.5}{1+0.5} = 0.33$

- At level of anchor $\sigma_v = 30 + 2\times18 = 66$ kPa

  $\therefore \sigma_h' = \sigma_h = K_a\sigma_v' = 0.33\times66 = $ 21.8 kPa

- At bottom of sand,

  $\sigma_v = 66 + 8\times18 = 210$ kPa

  $\sigma_h' = 0.33\times210 = 69.3$ kPa

- At top of clay,

  total horizontal stress $\sigma_h = \sigma_v - 2S_u = 210 - (2\times75) = 60$ kPa

- At bottom of wall (in clay)

  $\sigma_v = 210 + 4\times20 = 290$ kPa

  $\therefore \sigma_h = \sigma_v - 2S_u = 290 - (2\times75) = 140$ kPa

Passive side $\sigma_h = \sigma_v + 2S_u$

- At top of clay, $\sigma_v = 0$

$\therefore \sigma_h = 2\times75 = 150$ kPa

- At bottom of wall (in clay) ②

$\sigma_v = 4 \times 20 = 80$ kPa

$\therefore \sigma_h = 80 + (2 \times 75) = \underline{230 \text{ kPa}}$ [8]

(ii) • Total active force from sand $= P_{A1} + P_{A2}$ (see diagram)

$P_{A1} = 8 \times 21.8 = 174.4$ kN/m

$P_{A2} = \frac{1}{2} \times 8(69.3 - 21.8) = 190$ kN/m

- Total active force from clay $= P_{A3} + P_{A4}$

$P_{A3} = 60 \times 4 = 240$ kN/m

$P_{A4} = \frac{1}{2} \times 4(140 - 60) = 160$ kN/m

- Total passive force from clay $= P_{P1} + P_{P2}$

$P_{P1} = 4 \times 150 = 600$ kN/m

$P_{P2} = \frac{1}{2} \times 4(230 - 150) = 160$ kN/m

$\therefore$ Resolving horizontally,

anchor force required $T = (P_{A1} + P_{A2} + P_{A3} + P_{A4}) - (P_{P1} + P_{P2})$

$= (174.4 + 190 + 240 + 160) - (600 + 160)$

$= 764.4 - 760$

$= 4.4$ kN/m [4]

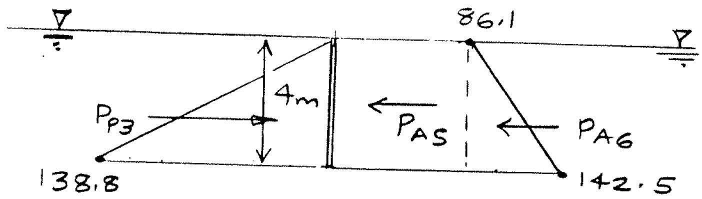
(very small anchor force required to maintain equilibrium)

(b) Long term condition

(i) • Both active and passive pressures in clay should be calculated in terms of effective stress parameters (using $\phi' = 25°$) and water table at top of clay on both sides of wall

• Active pressures in sand remain unchanged

• Modified pressures on wall in clay only :



Active side : top of clay

$\sigma_v = 210 \text{ kPa} = \sigma_v' \quad (u = 0)$

$\phi' = 25° \quad \therefore K_a = \frac{1 - \sin 25°}{1 + \sin 25°} = 0.41$

$\therefore \sigma_h' = 0.41 \times 210 = \underline{86.1 \text{ kPa}}$

bottom of clay, $\sigma_v = 210 + 4 \times 20 = 290 \text{ kPa}$

Pore pressure, $u = 4 \times 10 = 40 \text{ kPa}$

$\sigma_h' = 0.41 \times \sigma_v' = 0.41(290 - 40) = 102.5 \text{ kPa}$

$\therefore \sigma_h = \sigma_h' + u = 102.5 + 40 = \underline{142.5 \text{ kPa}}$

Passive side : top of clay

$\sigma_v' = \sigma_h' = 0 \quad (u = 0)$

bottom of clay : $\sigma_v = 4 \times 20 = 80 \text{ kPa}$

Pore pressure, $u = 4 \times 10 = 40 \text{ kPa} \quad \therefore \sigma_v' = 80 - 40 = 40 \text{ kPa}$

$K_p = \frac{1 + \sin 25°}{1 - \sin 25°} = 2.47$

$\therefore \sigma_h = 2.47\,\sigma_v' + 40 = (2.47 \times 40) + 40 = \underline{138.8 \text{ kPa}}$

[4]

(ii) Total active force in clay = $P_{A5} + P_{A6}$ (4)

$P_{A5} = 86.1 \times 4 = 344.4$ kN/m

$P_{A6} = \frac{1}{2} \times 4(142.5 - 86.1) = 112.8$ kN/m

Total passive force in clay = P_{P3}

$P_{P3} = \frac{1}{2} \times 4 \times 138.8 = 277.6$ kN/m

Required anchor force T = Σ active forces − Σ total passive forces

$\therefore T = (P_{A1} + P_{A2} + P_{A5} + P_{A6}) - P_{P3}$

$= (174.4 + 190 + 344.4 + 112.8) - 277.6$

$= 821.6 - 277.6 = \underline{544 \text{ kN/m}}$ [4]

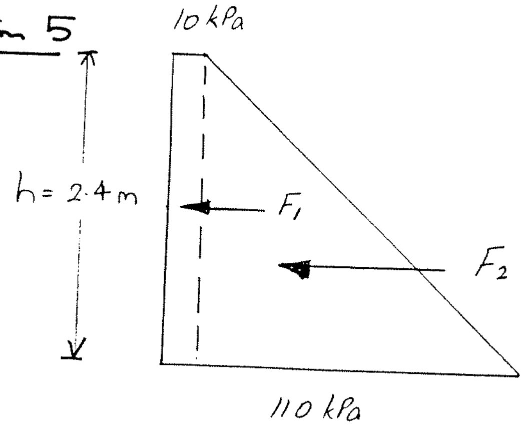
(much larger anchor force than required in short term)

Part IB, Paper 8, Selected Topics, Civil and Structural Engineering, 2003

Comments from Examiner on Question 4

Many candidates did not appreciate that the undrained shear strength of the clay was to be used to assess the short term soil pressures, and of those that did appreciate this a significant number got into difficulty in applying the undrained shear strength (and tried to introduce effective stresses). Effective stresses and water pressures are not relevant if the undrained shear strength is used – but this is only applicable in the short term. In the second part of the question, candidates were asked to calculate the long term active and passive pressures, and the associated anchor force required for equilibrium. This part was generally answered in a more satisfactory way – here the effective stresses and water pressures should be considered. In both parts of the question, many missed the point that the anchor force could be simply calculated by resolving horizontally (ie the difference between the total active and passive forces) – instead they got into much more difficult calculations involving moment equilibrium.

Question 5



Consider 1m length of wall.
$f_{cu} = 40$ MPa
$f_y = 460$ MPa

(a) $F_1 = 10 \times 2.4 \times 1 = 24$ kN @ $\frac{h}{2} = 1.2$ m from base.
$F_2 = \frac{1}{2} \times 100 \times 2.4 \times 1 = 120$ kN @ $\frac{h}{3} = 0.8$ m from base.

Total shear force $F_1 + F_2 = 144$ kN at base of wall.

Moment $M_{max} = F_1 \times \frac{h}{2} + F_2 \times \frac{h}{3} = 24 \times 1.2 + 120 \times 0.8 = 28.8 + 96 = 124.8$

$M_{max} = 124.8$ kNm at base of wall. (≈ 125 kNm)

[4]

(b) OPTION 1 (Constant wall thickness)

Wall thickness

$M = 0.15 f_{cu} b d^2$ from datasheet (N.B. γ_m's already in eqn's.)

$\Rightarrow d = \sqrt{\frac{M}{0.15 f_{cu} b}} = \sqrt{\frac{125 \times 10^6}{0.15 \times 40 \times 10^3}} = 144$ mm. (minimum)

Select $d = 150$ mm assume approx. 50mm cover (here assume 50mm from centroid of r/f to surface)

$\therefore$ Adopt $t = 200$ mm (thickness of wall for option 1)

Steel

$M = 0.87 f_y A_s d (1 - x/2)$ from datasheet.

$\therefore A_s = \frac{M}{0.87 f_y d (1 - x/2)}$

Guess neutral axis depth $x = 0.5$

$$\therefore A_s = \frac{125 \times 10^6}{0.87 \times 460 \times 150 \times (1 - 0.25)} = 2776 \text{ mm}^2/\text{m of wall.}$$

$$x = 2.175 \left(\frac{f_y}{f_{cu}}\right)\left(\frac{A_s}{bd}\right) \quad \text{from data sheet.}$$

$$\therefore x = 2.175 \times \frac{460}{40} \times \frac{2776}{1000 \times 150} = 0.463$$

Iterate

$$A_s = 2776 \times \frac{(1 - 0.25)}{(1 - \frac{0.463}{2})} = 2709 \text{ mm}^2$$

$$\therefore x = 0.463 \times \frac{2709}{2776} = 0.452$$

Iterate

$$A_s = 2709 \frac{(1 - \frac{0.463}{2})}{(1 - \frac{0.452}{2})} = 2709 \times \frac{0.769}{0.774} = 2690$$

$$x = 0.452 \times \frac{2690}{2709} = 0.449$$

Close enough i.e. $x = 0.45$ < 0.5 OK

Require 2690 mm^2

(1 extra iteration gives $A_s = 2685 \text{ mm}^2$, $x = 0.451$)

from table

20 @ 100 → 3140 mm²
25 @ 175 → 2810 mm² *
32 @ 275 → 2920 mm²

Option 1

Adopt 25 @ 175 → 2810 mm² for main tension steel.

[5]

(CHECK: ACTUAL $A_s = 2810 \text{ mm}^2$

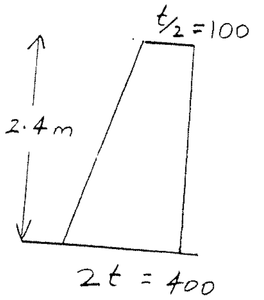
$$x = 2.175 \times \frac{460}{40} \times \frac{2810}{1000 \times 150} = 0.469 < 0.5 \text{ (OK)}$$

$$M = 0.87 f_y A_s d(1 - \tfrac{x}{2}) = 0.87 \times 460 \times 2810 \times 150 \underbrace{(1 - \tfrac{0.469}{2})}_{0.766} \times 10^{-6}$$

$$= 129 \text{ kNm/m} > M_{max} \text{ OK.}$$

Q 3 (cont.)

(c) $t = 200mm$.



Assume distance from surface to ℄ of primary bars is 50mm.

$\Rightarrow$ At base of wall $d = 350mm$

Assume $x = 0.5$

$$\therefore A_s = \frac{125 \times 10^6}{0.87 \times 460 \times 350 \times (1 - 0.25)} = 1190 \text{ mm}^2$$

$$x = 2.175 \left(\frac{f_y}{f_{cu}}\right)\left(\frac{A_s}{bd}\right) = 2.175 \times \frac{460}{40} \times \frac{1190}{1000 \times 350} = 0.08$$

Iterate

$$A_s = 1190 \times \frac{(1 - 0.25)}{(1 - \frac{0.08}{2})} = 930 \text{ mm}^2$$

$$x = 0.08 \times \frac{930}{1190} = 0.063$$

Iterate $A_s = 930 \dfrac{(1 - \frac{0.08}{2})}{(1 - \frac{0.063}{2})} = 922 \text{ mm}^2$

$$x = 0.063 \times \frac{922}{930} = 0.062 \quad \text{Close enough.}$$

Require $A_s = 922 \text{ mm}^2$ /m of wall)

(1 extra iteration gives $A_s = 922 \text{ mm}^2$)

12 @ 100 → 1130 mm²
16 @ 200 → 1010 mm² *
20 @ 300 → 1050 mm²

Option 2

Adopt 16 @ 200 → 1010 mm² for main tension steel.

[5]

(CHECK : ACTUAL $A_s = 1010 \text{ mm}^2$

$$x = 2.175 \times \frac{460}{40} \times \frac{1010}{1000 \times 350} = 0.0722 < 0.5 \text{ (ok)}$$

$$M = 0.87 f_y A_s d (1 - \tfrac{x}{2}) = 0.87 \times 460 \times 1010 \times 350 \times 10^{-6} \underbrace{(1 - \tfrac{0.0722}{2})}_{0.964}$$

$$= 136 \text{ kNm} > M_{required} \therefore \text{ ok})$$

Q3 (cont.)

(d) COST

Option 1

Concrete wall Volume = $2.4 \times 0.2 = 0.48\,m^3/m$ @ £200/m^3 = £96

R/F steel. 25 @ 175 → 2810 mm^2/m length of wall $h = 2.4\,m$.

$\rho_{steel} = 7840\,kg/m^3$

Volume of steel = $2810 \times 10^{-6} \times 2.4 = 6.74 \times 10^{-3}\,m^3$.

Mass of steel = $7840 \times 6.744 \times 10^{-3} = 52.8\,kg$

∴ Cost of steel = $\frac{52.8}{1000} \times 1200$ = £63.40

Option 1 ∴ Total cost of wall = 96 + 63.40 = £159.40

i.e. £159/m length

Option 2

Concrete wall Volume $2.4 \times \left(\frac{100+400}{2}\right) \times 10^{-3} = 0.6\,m^3$ @ £200/m^3 = £120

R/F steel 16 @ 200 → 1010 mm^2

Vol. steel = $1010 \times 10^{-6} \times 2.4 = 2.42 \times 10^{-3}\,m^3$

Mass steel = $7840 \times 2.42 \times 10^{-3} = 19\,kg$

Cost steel = $\frac{19}{1000} \times 1200$ = £22.80

Option 2 ∴ Total cost of wall = 120 + 22.80 = £142.80

i.e. £143/m length.

⇒ OPTION 2 most economical option [4]

(e) Shear (option 2) $v_c = 0.68 \left(\frac{100 A_s}{bd}\right)^{0.33} \left(\frac{400}{d}\right)^{0.25}$

$$v_c = 0.68 \times \left(\frac{100 \times 1010}{1000 \times 350}\right)^{0.33} \left(\frac{400}{350}\right)^{0.25} = 0.47\ MPa.$$

$V_{cap} = 0.47 \times 350 \times 1 = 164.5\,kN$ > $V_{applied} = 144\,kN$ OK [2]

no steel required.

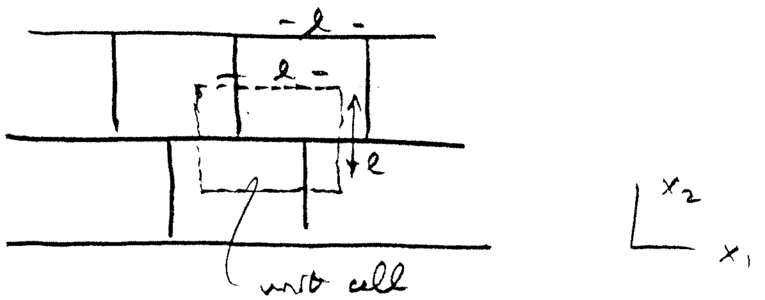
Part IB, Paper 8, Selected Topics, Civil and Structural Engineering, 2003

Comments from Examiner on Question 5

Many candidates were unable to calculate the maximum shear force and bending moment – instead of undertaking a simple two line calculation, they tried to write an equation for the loading and then integrate this twice, often making mistakes in so doing. Most candidates found the calculation of reinforcement steel required to be reasonably straightforward (some neglected to do the necessary iteration to arrive at a reasonable value); the majority missed the need to have satisfactory cover to the steel. Candidates were asked to undertake a simple cost comparison between the two walls; surprisingly this part of the question was not well answered. Very few candidates were able to answer the last part of the question on shear capacity, even though this was straightforward (using the formula given in the Data Sheet).

6. (a)

[4]



[4] (b) $\bar{\rho} = \dfrac{2lt}{l^2} = \dfrac{2t}{l}$

(c) $E_1 = ?$ Apply $\overline{\sigma}_{11}$ (macroscopic)
The bars along the x_1 direction stretch, whilst the bars along the x_2 direction carry no load.

$\Rightarrow \quad E_1 = \dfrac{\bar{\rho}}{2} \cdot E_s$

[6]

check: bar force f in horizontal bar is

$f = l\,\bar{\sigma}_{11}$

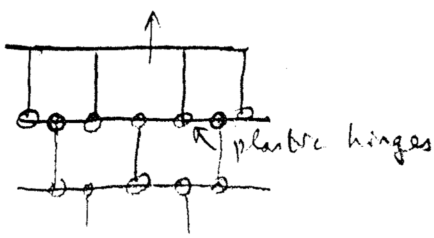
$\Rightarrow$ strain in bar is $\varepsilon = \dfrac{f}{E_s t} = \dfrac{l\,\bar{\sigma}_{11}}{E_s t}$

$\Rightarrow \quad E_1 = \dfrac{\bar{\sigma}_{11}}{\varepsilon} = \dfrac{E_s t}{l} = \underline{\tfrac{1}{2}\bar{\rho}\,E_s}$

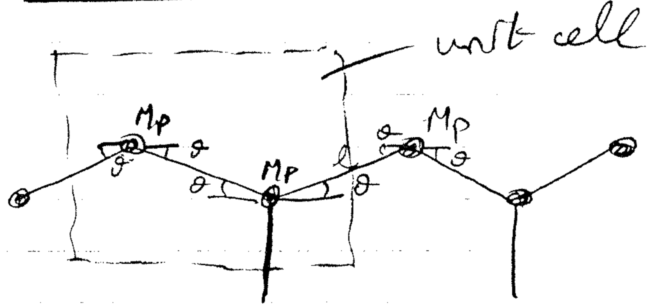
When the bar tension attains $f = \sigma_{YS}\,t$, the material yields.

$\bar{\sigma}_{11} = \dfrac{f}{l} = \sigma_{YS}\dfrac{t}{l} = \underline{\tfrac{1}{2}\bar{\rho}\,\sigma_{YS}}$

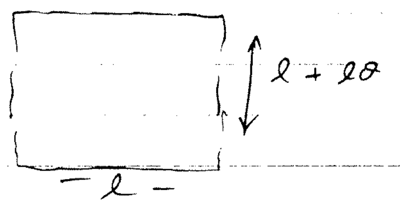
(d) Collapse mode is:



Upper bound calculation



[6] Unit cell extends along the x_2 direction by $l\theta$



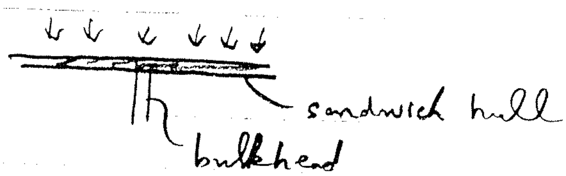
2 hinges in unit cell :

Work rate $\delta W = 2 \cdot (M_p \, 2\theta) = \bar{\sigma}_{22} \cdot l \, (l\theta)$

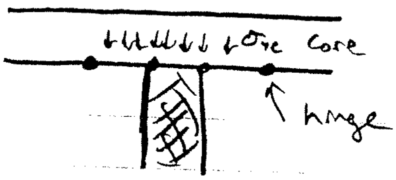
$\Rightarrow \quad \bar{\sigma}_{22} = \dfrac{4 M_p}{l^2} \qquad M_p = \frac{1}{4} t^2 \sigma_{ys}$

$\Rightarrow \quad \bar{\sigma}_{22} = \left(\dfrac{t}{l}\right)^2 \sigma_{ys} = \frac{1}{4} \bar{\rho}^2 \sigma_{ys}$

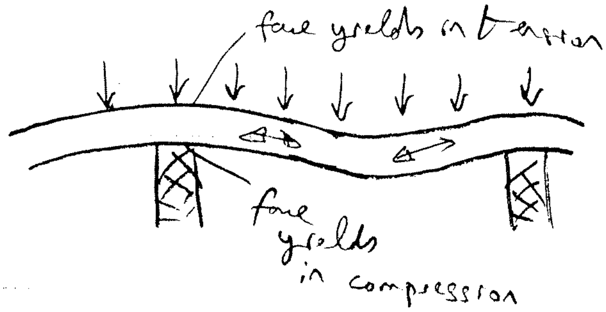
7. (a)



Modes : – indentation of inner face sheet

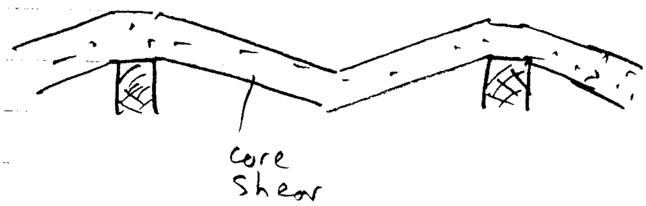


– face yield



6]

– core shear



– face wrinkling, perhaps, of the compressive face sheet.

(b) delamination will reduce the core shear strength
-] to zero. Face wrinkling is replaced by face sheet buckling at low loads.

(c) Minimum weight design.

Design constraints : minimum mass for a given
5] strength (or stiffness) of structure, without exceeding any of the ~~the~~ failure criteria for

7.(c) Minimum weight design (contd.)

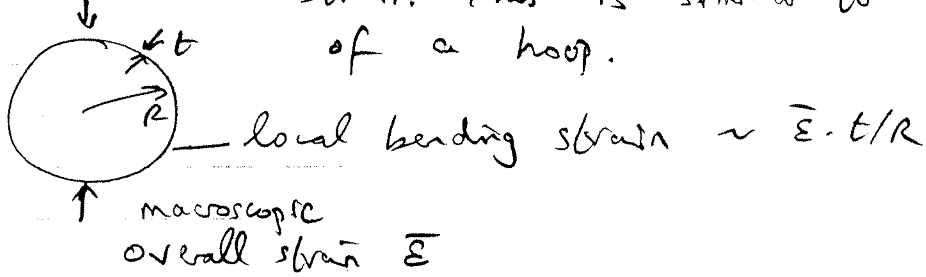
each individual mechanism. Also, faces need to be sufficiently thick to resist perforation, and bulkheads adequately strong.

Procedure. Given a set of material parameters, and assumed geometries (core and face thickness, prescribed length), calculate the mass m and failure loads P_f for each mechanism. The weakest mechanism is operative. Then, repeat for other geometries at fixed mass, to determine ~~that~~ geometry which maximises P_f. Impose the design constraints to limit the choice of geometries, as necessary.

(d) Superior energy absorption can be achieved by

5] core crush, without perforation of the inner face sheet. The load can be spread over a large number of transverse bulkheads.

8. (a) Cast alloys contain brittle phases and have lower intrinsic ductility than forged alloys. However, foams have a bending microstructure with lower local strains than the macroscopic imposed strain. This is similar to bending of a hoop.



[4]

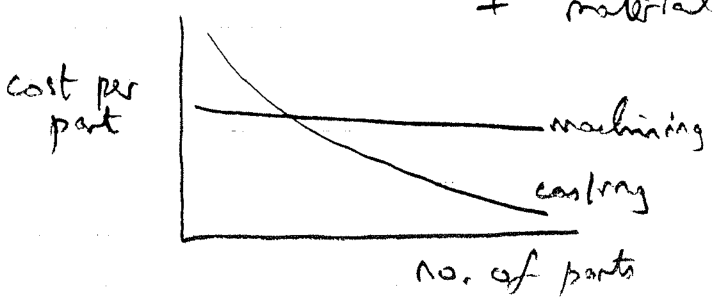
So the macroscopic ductility of foams can be large.

Lattice material: local stretching strain is of similar magnitude to the macroscopic strain $\bar{\varepsilon}$.

(b) Capital costs vary – e.g. the setting-up costs of a casting facility are much greater than a milling machine cost.

$$\text{Cost per part} \approx \frac{\text{Capital cost}}{\text{no. of parts}} + \frac{\text{Cost per hour}}{\text{production rate}}$$

$$+ \text{ material costs}$$



8. (c) 'Plastic Poisson ratio' for a foam is almost zero, while = ½ for a fully dense solid.

4]

Uniaxial compression on foam ⇒ no lateral expansion ⇒ response is the same as for indentation.

For a fully dense solid, plastic constraint in indentation leads to redundant plastic work and indentation pressure ≈ 3 uniaxial yield strength.

(d) the heavy buffer plate of mass m impulse I from the blast and acquires a velocity V, such that

$$mV = I$$

4] Kinetic energy $KE = \frac{1}{2} mV^2$, is dissipated by the foam.

$$K.E. = \frac{1}{2} mV^2 = \frac{1}{2} \frac{I^2}{m}$$

I is prescribed, so minimise KE by maximising the mass m.

(e) A perfect foam is a stretching structure under hydrostatic loading with strength $\propto \bar{\rho}\, \sigma_{ys}$.

4] Imperfections induce bending and reduce the strength (hydrostatic and deviatoric) to that of a bending structure, $\propto \bar{\rho}^{3/2} \sigma_{ys}$.

A perfect foam bends under deviatoric loading, $\propto \bar{\rho}^{3/2} \sigma_{ys}$.

ENGINEERING TRIPOS, PART IB, 2003.

PAPER 8: AEROTHERMAL ELECTIVE

Q9 (a) $V = M\sqrt{\gamma R T_1} = 0.8 \times \sqrt{1.4 \times 287 \times 226} = \underline{\underline{241.1 \text{ m/s}}}$

$P_{02} = P_1 \left(1 + \frac{\gamma-1}{2} M^2\right)^{\frac{\gamma}{\gamma-1}} = \underline{\underline{44.2 \text{ kPa}}}$

$T_{02} = T_1 \left(1 + \frac{\gamma-1}{2} M^2\right) = \underline{\underline{254.9 \text{ K}}}$

GROSS THRUST, $F_g = \dot{m}_c (1+\beta) V_j$
$= 60 \times 6 \times 350 = \underline{\underline{126 \text{ kN}}}$

NET THRUST, $F_N = F_g - \dot{m}_c (1+\beta) V = \underline{\underline{39.24 \text{ kN}}}$

(b)

$$\Pi_1 = \frac{\dot{m}_a \sqrt{c_p T_{02}}}{A P_{02}} \equiv \frac{\text{MOMENTUM FLUX}}{\text{FORCE.}}$$

$$\Pi_2 = \frac{\dot{m}_f \, LCV}{A P_{02} \sqrt{c_p T_{02}}} \equiv \frac{\text{FUEL ENERGY FLUX}}{\text{FORCE} \times \text{VELOCITY.}}$$

where $\sqrt{c_p T_{02}}$ gives a characteristic velocity proportional to the stagnation sound speed. The forms are different because $\dot{m}_a$ relates to momentum flux, whereas $\dot{m}_f$ determines the energy flux from the fuel.

(c) Nozzle remains choked but some expansion occurs downstream of exit plane. $\therefore P_9 > P_a$

At new altitude (denoted 'H'):

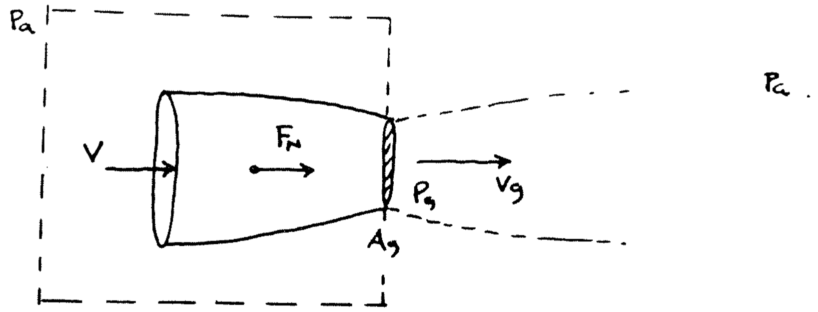
$T_{02}^H = T_1^H \left(1 + \frac{\gamma-1}{2} 0.9^2\right) = 251 \text{ K}$

$P_{02}^H = P_1^H \left(1 + \frac{\gamma-1}{2} 0.9^2\right)^{3.5} = 30.44 \text{ kPa}$

$\frac{\dot{m}_a \sqrt{T_{02}}}{P_{02}} = \text{const.} \Rightarrow \dot{m}_a^H = \left(\frac{30.44}{44.2}\right)\left(\frac{254.9}{251}\right)^{1/2} \times 360 = \underline{\underline{249.5 \text{ kg/s}}}$

$V^H = 0.9\sqrt{\gamma R T_H} = 265.1 \text{ ms}^{-1}$

Q9 (CONT)



SFME: $F_N + (P_a - P_g)A_g = \dot{m}_a (V_g - V)$

$\therefore \quad F_N + \dot{m}_a V + P_a A_g = \dot{m}_a V_g + P_g A_g$

RHS = function on engine operating point (since V_g, P_g are sonic)

$$\therefore \quad \frac{F_G + P_a A_g}{A_g P_{02}} = fn\left(\frac{\dot{m}_a \sqrt{c_p T_{02}}}{A P_{02}}\right) = \text{CONST.}$$

$$\therefore \quad (F_G + P_a A_g)^H = (126 \times 10^3 + 29 \times 10^3 \times 3) \times \frac{30.44}{44.2} = 146.67 \text{ kN.}$$

$$\therefore \quad F_N = F_G - \dot{m}_a V = 146.67 - 18 \times 3 - 249.5 \times 265.1 \times 10^{-3} \text{ kN}$$

$$= \underline{\underline{26.5 \text{ kN}}}$$

Examiner's note: most difficulty was encountered in part (b) where the majority of students failed to introduce any physics into the dimensional analysis.

(a) The 'adverse' pressure gradient in compressors increases the likelihood of flow separation and so the pressure rise per stage must be limited, giving a large number of stages. In turbines the pressure gradient is favourable.

(i) Too few turbine stages means stages are overloaded, leading to a low isentropic efficiency.

(ii) Overloading the compressor stages may lead to rotating stall or surge and the engine will cease to function.

(Note also: the pressure ratio is less across the turbine).

(b) Euler Work: $\Delta h_0 = U_2 V_{\theta 2} - U_1 V_{\theta 1}$

but $U_2 = U_1 = U$ and $V_{\theta 2} = V_x \tan\alpha_2$, $V_{\theta 1} = V_x \tan\alpha_1$

$$\therefore \quad \Delta h_0 = U V_x (\tan\alpha_2 - \tan\alpha_1)$$

$$= U^2 \phi (\tan\alpha_2 - \tan\alpha_1)$$

$$\therefore \quad \psi = \frac{\Delta h_0}{U^2} = \phi (\tan\alpha_2 - \tan\alpha_1).$$

(c) $$\Delta T_0 = \frac{T_{01}}{\eta_c}\left(r^{\frac{\gamma-1}{\gamma}} - 1\right) = \frac{288}{0.88}\left(25^{1/3.5} - 1\right)$$

$\Delta T_0 = 493.7\ \text{K}$, overall.

$U_m = r_m \Omega = 0.5 \times 5730 \times \frac{2\pi}{60} = 300\ \text{ms}^{-1}$.

$$\frac{\Delta h_0}{U^2} = \frac{c_p \Delta T_0^{ST}}{U^2} \leq 0.4$$

$$\therefore \quad \Delta T_0 \leq \frac{0.4 \times 300^2}{1010} \leq 35.65 \quad \text{PER STAGE.}$$

(cont.)

$\therefore$ No. STAGES $\geq \frac{493.7}{35.65} \geq 13.84$

i.e., <u>14 STAGES are needed.</u>

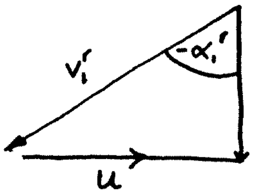
For 14 Stages, $\psi = \frac{13.84}{14} \times 0.4 = \underline{\underline{0.396}}$

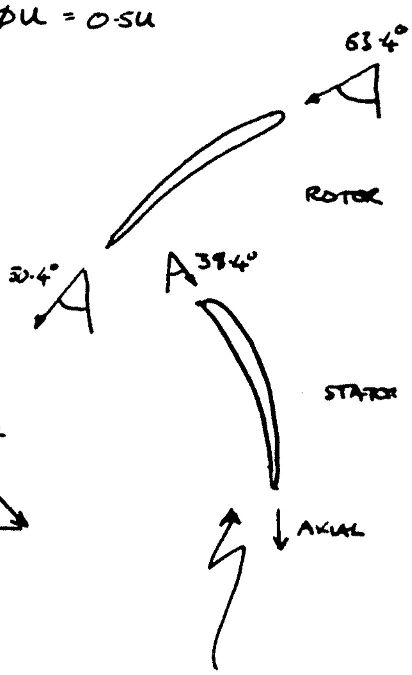
(d) Zero inlet swirl $\Rightarrow \alpha_1 = 0$.

$\therefore \tan\alpha_2 = \frac{\psi}{\phi} = \frac{0.396}{0.5} = 0.792$

$\Rightarrow \underline{\alpha_2 = 38.4^\circ}$

<u>ROTOR INLET</u>:

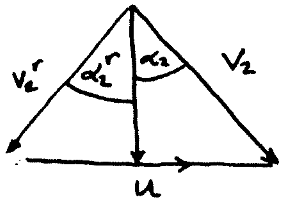




$\alpha_1^r = -\tan^{-1}(1/\phi) = -\tan^{-1} 2$

$\underline{\underline{\alpha_1^r = -63.4^\circ}}$

<u>ROTOR EXIT</u>:



$V_x \tan\alpha_2^r = -U + V_x \tan\alpha_2$

$\tan\alpha_2^r = -2 + 0.792$

$\therefore \underline{\underline{\alpha_2^r = -50.38^\circ}}$.

Sketch assumes zero incidence and deviation!

Examiner's note: part (d) was not done well: most students drew blades and velocity triangles for a turbine not a compressor.

(a) Propulsive efficiency represents the effectiveness of the propulsive system at flight speed V, but does not account for enthalpy rise through engine.

$$\eta_P = \frac{\text{Power to Aircraft}}{\text{Mech. Power to jet}} \approx \frac{\text{Net thrust} \times V}{\dot{m}_a \times \Delta KE \text{ per kg}}.$$

The thermal efficiency represents the effectiveness of the engines in delivering mech. power to the jet and so accounts for enthalpy rise.

$$\eta_{TH} = \frac{\text{Mech. Power to Jet}}{\text{Rate of fuel energy supply}} \approx \frac{\dot{m}_a \times \Delta KE \text{ per kg}}{\dot{m}_f \times LCV}.$$

Overall efficiency represents overall effectiveness in converting fuel energy to <u>useful</u> mechanical energy

$$\eta_o = \frac{\text{Power to Aircraft}}{\text{Rate of fuel energy supply}} = \frac{\text{Net thrust} \times V}{\dot{m}_f \, LCV}$$

$$\eta_o = \eta_P \, \eta_{TH}$$

and

$$\eta_o = \frac{\text{Net thrust}}{\dot{m}_f} \times \frac{V}{LCV}$$

$$= \frac{1}{SFC} \times \frac{V}{LCV}.$$

(b) As fuel is burnt, weight of aircraft and hence required lift decreases. The aircraft moves to higher altitude, where the air density is lower, so as to keep the lift coefficient at its optimum value, whilst maintaining approximately constant speed.

$$C_L = \frac{L}{\frac{1}{2}\rho V^2 A} = \frac{W}{\frac{1}{2}\rho M^2 \frac{\gamma p}{\rho} A} = \frac{W}{\frac{1}{2}\gamma M^2 p A}.$$

$$\therefore A = \frac{Mg}{\frac{1}{2}\gamma M^2 C_L p}$$ from graph, $C_L^{opt} = 0.5$ & $p = 28.74$

$$\Rightarrow A = \frac{600 \times 10^3 \times 9.81}{\frac{1}{2} \times 1.4 \times 0.86^2 \times 0.5 \times 28.7 \times 10^3} = \underline{\underline{792.3 \text{ m}^2}}$$

(c) To maintain constant C_L, $\frac{W}{\frac{1}{2}\rho V^2 A} = \frac{W}{\frac{1}{2}\rho M^2 \frac{\gamma p}{\rho} A} = \text{const}.$

$\therefore \quad W/p = \text{const}.$

$\therefore \quad \frac{p_2}{p_1} = \frac{W_2}{W_1} = \frac{2}{3} \Rightarrow p_2 = 19.13$ kPa.

Interpolation in table gives altitude = $\underline{\underline{39{,}630 \text{ ft}}}$

(Air traffic control restrictions would probably give 39,000 ft).

(d) $$S = \frac{V}{SFC} \times \frac{(L/D)}{g} \times \ln\left(\frac{W_1}{W_2}\right) = \frac{\eta_o \times LCV \times (L/D)}{g} \times \ln\left(\frac{W_1}{W_2}\right)$$

from $(M L/D)^{opt} = 15 \Rightarrow L/D = \frac{15}{0.86} = 17.44$

$$\therefore S = \frac{0.35 \times 43 \times 10^6 \times 17.44}{9.81} \times \ln 1.5$$

$\underline{\underline{S \simeq 10{,}850 \text{ km}}}$

Examiner's note: students seemed to find this question easy. Most attempted it.

SECTION D (Electrical Engineering)

12 (a) The carrier velocity , v, is given by

$$v = \mu E_d$$

Therefore, if the channel length is L, then the transit time is

$$\tau = \frac{L}{\mu E_d}$$ [1]

This equation is valid if the velocity of the electron is significantly lower than the scattering velocity. [1]

(b) In order to turn the device OFF, we need to deplete the doped channel layer. We assume that there are no free charges in the depleted layer, and that all donors have been positively ionized, so the charge density is eN_D. [1]

Hence, from the Gauss Law of Electrostatics,

$$E_g WL = \frac{eN_D xWL}{\varepsilon_0 \varepsilon_r}$$

where x is the distance into the depletion region from the interface with the silicon substrate. So,

$$E_g = \frac{eN_D x}{\varepsilon_0 \varepsilon_r}$$ [2]

The threshold voltage is then given by

$$V_T = -\int E\mathrm{d}x$$
$$= -\int_D^0 \frac{eN_D x}{\varepsilon_0 \varepsilon_r}$$ [2]
$$V_T = \frac{eN_D D^2}{2\varepsilon_0 \varepsilon_r}$$

The maximum gate field at the interface is then

$$E_{g\max} = \frac{eN_D D}{\varepsilon_0 \varepsilon_r}$$ [1]

(c) We know from the expression for the gate threshold voltage that

$$D = \left(\frac{2V_T \varepsilon_0 \varepsilon_r}{eN_D} \right)^{1/2}$$

$$= \left(\frac{2.1.8 \cdot 854 \times 10^{-12}.11 \cdot 8}{1 \cdot 602 \times 10^{-19}.10^{20}} \right)^{1/2}$$

$$\underline{D = 3 \cdot 6\ \mu\text{m}}$$

[2]

The drain source electric field is

$$E = \frac{V_d}{L} = \frac{0 \cdot 5}{10 \times 10^{-6}} = 5 \times 10^4\ \text{V m}^{-1}$$

[1]

From the graph, the carrier velocity is then ~7000 m s^{-1}, so

$$\tau = \frac{L}{v}$$

$$= \frac{10 \times 10^{-6}}{7000}$$

$$\underline{\tau = 1 \cdot 4\ \text{ns}}$$

[2]

(d) From the expression for E_{gmax}, it is clear that for the electric field to be a constant, then the product $N_D D$ must also be a constant. Therefore, if D is reduced by a factor of R by reducing the device dimensions, then N_D must increase by R so

$$\underline{N_{D2000} = N_{D1970}.R = 10^{20}.100 = 10^{22}\ \text{m}^{-3}}$$

[2]

(e) The new channel length is 0·1 μm, and the minimum transit time is achieved when the carrier is travelling at the scattering velocity. From the graph, the scattering velocity is 10^5 m s^{-1} and the minimum field required to achieve this is ~10^7 V m^{-1}. Hence,

$$\tau = \frac{L}{v}$$

$$= \frac{0 \cdot 1 \times 10^{-6}}{10^5}$$

$$\underline{\tau = 1 \cdot 0\ \text{ps}}$$

[1]

The minimum drain voltage is then

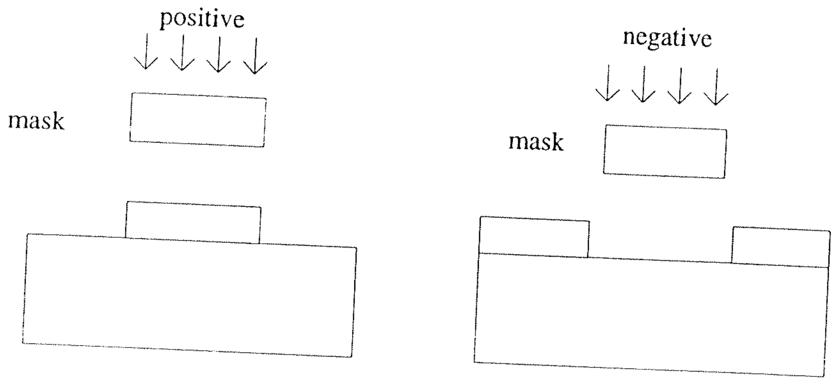
$$V_{d\min} = E_{d\min} L = 10^7.0 \cdot 1 \times 10^{-6} = 1\ \text{V}$$

[1]

(f) Limits to further device size reduction include

- Capacitance effects increase with reducing size, and this puts a limit on the switching speed. [1]
- Electron tunnelling across small dimensions increases leakage currents. [1]
- In reality, as device dimensions decrease, the electric field inside the device also increase and this can lead to problems in terms of both breakdown and a reduced effective carrier mobility caused by the high gate field. [1]
- *Other sensible limits are also acceptable with one mark for each up to a maximum of three.*

13 (a)



+ve resist the exposed regions become more soluble and thus are more easy to remove during development. Net result is that the patterns formed in the +ve resist are the same as those on the mask.

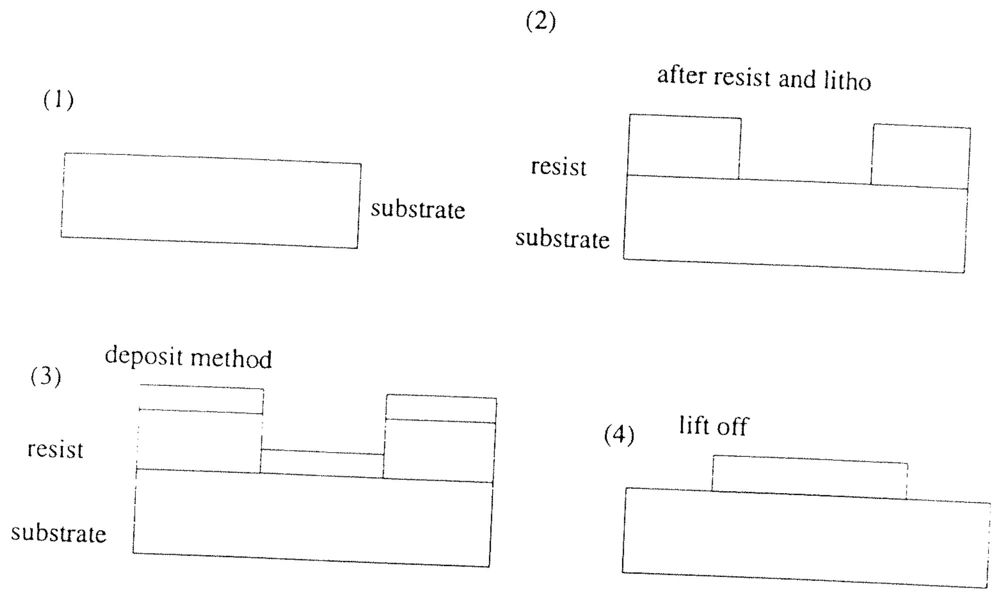
For –ve the exposed regions become less soluble and the patterns formed on the substrate are the reverse of the mask pattern.

[4]

Two methods of pattern tranfer - substration and lift off

to make metal dot with lift off

[3]



(b) Silicon Dioxide

Successful MOSFET operation depends upon a good gate insulator. Silicon dioxide is the standard insulator and it is also used in many of the different process steps as a barrier layer. It is therefore vital to have a reliable method of producing such material.

Oxidation of silicon is achieved by heating silicon wafers in a furnace, shown in the following diagram, in the presence of an oxidising atmosphere such as oxygen or water vapour. The two common approaches are:

Wet oxidation where the oxidising atmosphere in the furnace contains water vapour. Temperature is approx 900-1100 °C. This leads to a fast growth rate – depends on temperature but is of order 0.5 microns per hour at 1100 °C. The net reaction for the oxidation process is

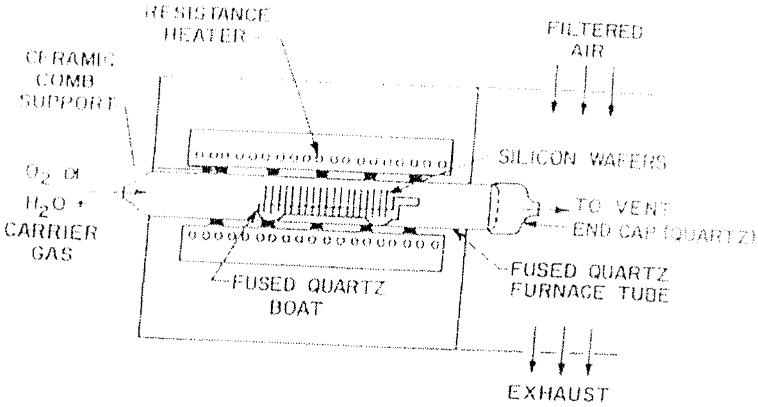
$$Si(solid) + 2H_2O \rightarrow SiO_2(solid) + 2H_2$$

[2]

Dry oxidation where the oxidising atmosphere in the furnace is pure oxygen. Growth temperatures have to be in the region of 1200 °C to achieve acceptable growth rates of 0.1 microns/hour.

$$Si(solid) + O_2 \rightarrow SiO_2(solid)$$

[2]



[1]

As device dimensions shrink, the oxide layer becomes thinner and electrons may quantum mechanically tunnel cross the oxide barrier between the gate and the source-drain channel.

[1]

(c) (i) For a system where the boundary is at a constant concentration (in this case the solubility limit of P in Si) then the concentration of dopant as a function of distance into the solid is given by

$$C = C_0 \mathrm{erfc}\left(\frac{x}{2\sqrt{Dt}}\right)$$

From the figure, at 1100 °C, C_0 is 10^{27} m^{-3}. To dope the p-type silicon substrate with phosphorous (n-type), we need to have at least compensated for the p-type dopant concentration of 4×10^{22} m^{-3}. Therefore, we need to know the depth, d, at which the concentration of P has fallen to this value after a time t. From the graph, the diffusion coefficient for P in Si at 1100 °C is 3×10^{-17} m^{-3}. Now, from above

$$d = 2\sqrt{Dt}\,\mathrm{erfc}^{-1}\left(\frac{C_{\min}}{C_0}\right)$$

$$= 2\sqrt{3\times10^{-17}(30.60)}\,\mathrm{erfc}^{-1}\left(\frac{4\times10^{22}}{10^{27}}\right)$$

$$d = 2\sqrt{3\times10^{-17}.1800}\,\mathrm{erfc}^{-1}\left(4\times10^{-5}\right)$$

From the graph, then

$\underline{d \sim 1{\cdot}3\ \mu\mathrm{m}}$

[4]

(ii) For the oxide, we will assume the same solubility limit. This is more likely to be an overestimate than under, as we are approaching the atomic density! The oxide is 100 nm thick so the concentration of dopants at the interface with the channel is

$$C = C_0 \text{erfc}\left(\frac{x}{2\sqrt{Dt}}\right)$$

$$= 10^{27} \text{erfc}\left(\frac{100\times10^{-9}}{2\sqrt{1\times10^{-19}.1800}}\right)$$

$$= 10^{27} \text{erfc}(3\cdot72)$$

$$C = 10^{27}.10^{-7}$$

$$C = 10^{20}\ m^{-3}$$

This is small compared with the p-type dopant density, so the thickness of oxide is sufficient to protect the channel.

[3]

14 (a) Early experimental evidence that electrons have like waves came from electron diffraction patterns. Davisson and Germer used a beam of monoenergetic electrons to irradiate a nickel crystal. The electrons were found to be scattered from the crystal planes, rather like light from a diffraction grating, such that there were preferred scattering directions. The Bragg criterion could be used to predict these scattering directions if the electrons were behaving like waves with the de Broglie wavelength. More recently, quantum mechanical electron tunnelling through a barrier provides further evidence. It is found that electrons can tunnel through a thin barrier, such as an insulating layer, in a way that can be predicted by the Schrodinger equation, which requires the electron to have a wavelength. The fact that electrons exist in well-defined energy states inside the atom provides further evidence and hence photon emission lines are observed when an electron 'falls' from one energy state to a lower one. [1 mark for stating each experiment with an extra 1 mark for a brief description that shows understanding of why this is evidence that the electron is a wave up to a maximum of 4 marks].

(b) From de Broglie,

$$\lambda = \frac{h}{p}$$

where h is the Plank constant. Now, from kinetic energy,

$$E = \frac{p^2}{2m^*}$$

and hence

$$\underline{\lambda = \frac{h}{(2m^* E)^{1/2}}}$$

[3]

(c) In the expression given, k is the wavenumber (or strictly speaking the complex component of it). Hence,

$$k = \frac{2\pi}{\lambda} = \frac{(2m^*|E-V|)^{1/2}}{\hbar}$$

[2]

where $|E-V|$ is the modulus of the difference between the energy of the electron and the barrier height. Hence, by substitution

$$J = J_0 \exp\left(\frac{-2d(2m^*|E-V|)^{1/2}}{\hbar}\right)$$

[2]

(d) If the capacitance is to remain constant, then

$$\frac{\varepsilon_0 \varepsilon_{r1} A}{d_1} = \frac{\varepsilon_0 \varepsilon_{r2} A}{d_2}$$

$$d_2 = \frac{\varepsilon_2 d_1}{\varepsilon_{r1}}$$ [4]

(e) The increased oxide thickness will decrease the tunnelling current in the new device, as will raising either the barrier height or the effective mass. [2]

The new tunnelling current may be related to that in the oxide as follows,

$$J_h = J_0 \exp\left(\frac{-2d_h \left(2m_2^* |E - V_2|\right)^{1/2}}{\hbar} \right)$$

$$J_h = J_0 \exp\left(\frac{-2d_{ox} \varepsilon_{r2} \left(2m_2^* |E - V_2|\right)^{1/2}}{\varepsilon_{r1x} \hbar} \right)$$ [2]

The important parameters are those that are dependent on the new dielectric. Hence, a good figure of merit will be given by

$$FOM = \varepsilon_{r2} \left(m_2^* V_2\right)^{1/2}$$

[1]

ENGINEERING TRIPOS, PART IB, 2003
PAPER 8 - INFORMATION ELECTIVE

P8 Q15 Architecture for Mobile Computing 2003

You are required to design a low power network suitable for mobile computers and other devices. The system will be used to provide location information as well as communication facilities.

(a) Outline the main characteristics of the Active Badge Infra-red network in terms of the following:
1) Physical layer protocol
2) The data link layer protocol
3) Capability of generating location information [8]

(b) Outline the main characteristics of the Bluetooth radio network in terms of the following
1) Physical layer protocol
2) The data link layer protocol
3) Capability of generating location information [8]

(c) Comment on a hybrid system, which has sufficient performance for transmission of audio data, yet can accurately locate geometrically proximate objects in different rooms (for example on either side of a wall). [4]

Active Badge (infra-red)

IR devices are used in home controllers on a huge scale hence are cheap
IR transmitters and receivers are well characterised
IR spectrum is not regulated
Interference in the band is not a huge problem
Baseband modulation can be used and is simple
Pulse position modulation is appropriate and is power efficient
Data rate is likely to be low (in the order of 1-10 K bits per second) for a diffuse system

For a very simple system we can transmit and repeat if there has been a clash
Detecting the clash at the transmitter may be possible much of the time
Another approach is to use acknowledgements from the receiver
Because the data rates are so low the end-to-end propagation delay is not a problem

Rooms and opaqueness of walls can be used to generate location information
Such a location system will work better if the transmissions are diffuse
Randomizing transmissions is a good medium access scheme if only location is required
Another approach is to measure received signal strength to obtain range information
This will require many more receivers but will give more precise location information
Bluetooth (radio)

Operates in Industrial, Scientific and Medical Band (ISM) at 2.4 GHz

This band is not regulated
Microwave ovens are in the same band!
Interference can be a problem due to multi-path fading, movement of mobile devices
A frequency hopping (FH) technique is used with a hop rate of 1600/sec (625μsec)
Packets are short to allow fast hopping
Binary FSK modulation is used
Transmitter power can be increased from 0dBm (10 meters) to 20dBm (100 meters)
Packet headers are protected by redundant error correction scheme
Acknowledgement with Automatic Repeat Request (ARQ) is used
Special coding for voice traffic where retransmission and ARQ is not suitable

Time division duplex (TDD) is used on each channel
Each device transmits in alternate slots and receives in the others
Synchronous Connection Oriented links are used for voice
Asynchronous Connection Less links are used for data transfer
Data rates up to 1Mbps are possible

Received Signal Strength Indication (RSSI) from multiple neighboring devices can be used to compute location but is not accurate

Hybrid

It would be very difficult to separate (resolve) two devices either side of a wall with radio
It is difficult to design a high speed diffuse infra-red communication system
A combination of radio for communication and infra-red for location is promising
Future "sentient" applications with computing embedded in the physical world would take advantage of such systems

P8 Q16 Architecture for Mobile Computing 2003

(a) For a virtual memory system:

(i) Briefly describe the structure and function of a virtual memory system. [2]

(ii) Explain the terms "page", "page table", "page fault" and indicate the criteria for choosing a page size. [2]

(iii) What is the purpose and function of the TLB? [2]

(iv) Draw a diagram of a virtual memory system, naming all the fields of all the data structures involved. [2]

(b) After briefly pointing out the structural similarities between virtual memory systems and cache systems, explain the differences in greater detail. [6]

(c) Describe and compare three page replacement strategies, including the "optimal page replacement algorithm". Describe Belady's anomaly and say which of the strategies you cite exhibit it. [6]

(a)

(i) A virtual memory system gives the illusion of having much more main memory than there really is. Infrequently used portions of this memory are stored on disk, which is very slow in comparison. However, by careful management, it is on average possible to keep the access speed close to that of the main memory.

Each process is allocated its own virtual address space, which may be much larger than the available main memory. Processes cannot access addresses outside their own space, so this protects from mutual interference: a process cannot read or overwrite the data of another.

(ii) Memory is split into equal-sized allocation units called pages, usually a few kilobytes each.

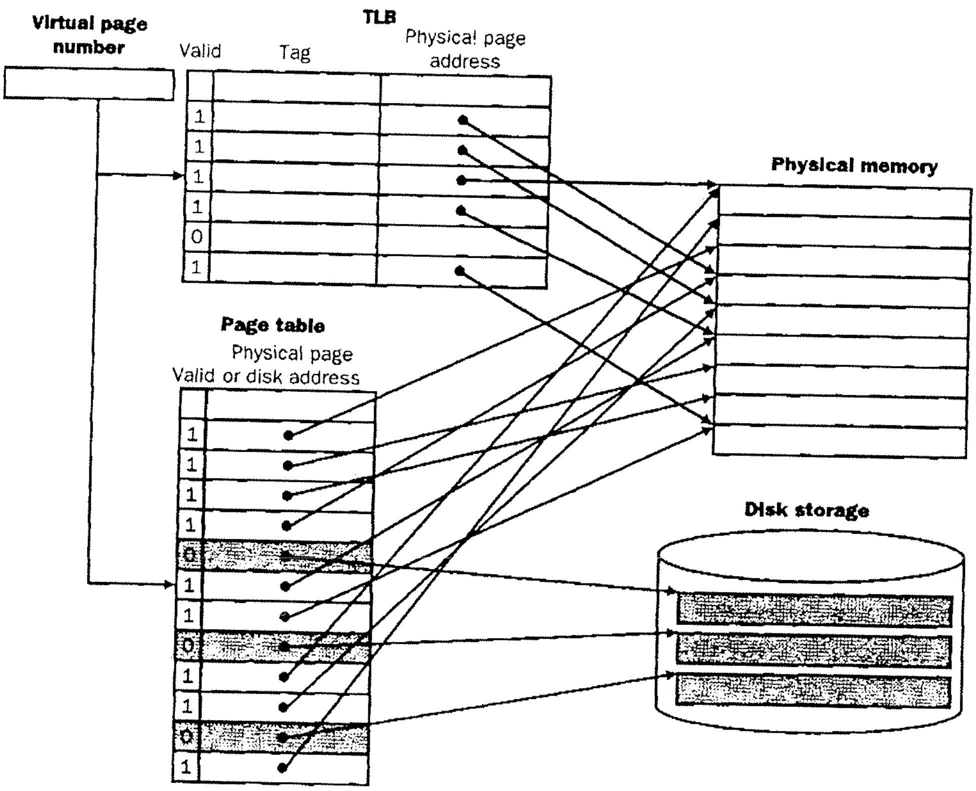
Each process has a page table that, for every virtual page number, stores a "valid" bit (indicating whether the page is in memory) and the corresponding physical page number (if the bit is set) or disk address (if not). Accessing a virtual page in the second case will cause a "page fault" (i.e. having to fetch the page from disk into memory, replacing one of the existing pages).

Page faults are very expensive in terms of performance, since disk is orders of magnitude slower than memory. Therefore, the choice of a good strategy for replacing pages (trying to avoid replacing pages that are going to be accessed again soon) affects performance dramatically. For this reason page faults are handled in software rather than hardware, so that more complex strategies can be implemented. On page fault, the processor invokes the operating system.

The choice of page size is a trade-off. Small pages imply less wastage at the end of the memory area used by each process and give finer granularity when deciding what to page out. Large pages can be transferred more efficiently to and from disk, and require a smaller page table. On the whole, over the years, typical page sizes have been growing steadily as memory becomes cheaper.

(iii) With the scheme so far described every memory access by a program, even when no page faults are involved, requires another preliminary memory access to consult the page table. The TLB (Translation Lookaside Buffer) is introduced to avoid this degradation of performance: it is basically a cache for the page table. A virtual page number is first looked up in the fast TLB. If present, the TLB entry gives the physical page number. If not present, this may be either a TLB miss, in which case the physical page number will be in the page table, or a page fault.

(iv) See figure: vm-picture.png (or .tif or .jpg)
Source: fig 7.24 on page 591 of Patterson/Hennessy.



(b)
In both cases there is a small, fast, expensive memory which stores a dynamic subset of a large, slow, cheap memory. With careful tuning, the combined system approaches the speed performance of the fast memory while offering the capacity of the large memory.

The main difference is that the ratio between the speeds of the slow and fast memory is quite different in the two cases. A cache may be about 10 times faster than main memory, but main memory may be about 1000 times faster than disk.

This means that, in comparison, a page fault is a much more catastrophic event, performance-wise, than a cache miss. So for virtual memory it pays to adopt a fully associative structure and very elaborate page replacement algorithms, whereas sophisticated cache block replacement algorithms would consume more time than they'd save.

For the reason just mentioned, caches are rarely fully associative, whereas virtual memory always is. The difficulty of a fully associative scheme is in finding the sought entry, since an exhaustive search is usually too expensive. For virtual memory, this is solved by building an index of the fast memory: this index is the page table. In a cache system the equivalent data structure does not exist. Fully associative caches implement the search with parallel hardware (expensive).

In caches, both write-through and write-back strategies are possible. In virtual memory, though, the speed ratio makes write-through very unfavourable, so write-back is the chosen solution. All changes to a page are made to the in-memory version and the write to disk need only happen when the dirty page needs to be replaced.

(c)
The "optimal page replacement algorithm" assumes knowledge of the future portion of the "reference string", i.e. the sequence of pages that the process will access. Whenever a page needs to be discarded to make room for a new one, it selects the one that will not be used for the longest period of time. This minimizes the number of page faults.

This ideal algorithm cannot be implemented in practice, but it is useful as a reference against which the performance of more practical algorithms can be compared.

For a given page replacement algorithm and for a given reference string, if the number of available pages of main memory is increased, one would expect a reduction in the number of page faults. Belady's anomaly is the paradoxical situation in which page faults instead increase. The optimal algorithm never exhibits it.

The "FIFO" (first in, first out) algorithm is very simple: it always discards the oldest page, i.e. the one that was loaded before all the others. Easy to implement but it may suffer from Belady's anomaly. It may also behave pathologically when the working set is larger than the available physical memory.

The "LRU" (least recently used) algorithm discards the page that was least recently used. It corresponds to the optimal algorithm but looking backward in time rather than forward. It usually offers good performance and does not suffer from Belady's anomaly, but it is tricky to implement. One way is to keep a timestamp for each page, updated on each access. Replacement involves finding the page with the smallest time.

P8 Q17 Architecture for Mobile Computing 2003

The Auto-ID tagging system can be loosely described as a radio-based bar code.

(a) Describe the general architecture of the system and the main technical challenges involved. [6]

(b) Discuss its technical advantages and disadvantages compared to the optical bar code. [4]

(c) Discuss the security risks of this technology. [4]

(d) Describe the implementation of an access control scheme under which tags can be identified by their owner but not by unauthorized readers. What resources are needed in the tag to support this mode of operation? [6]

(a)
RFID tag, or transponder, is affixed to the object. Can be active (has mini-battery) or passive (powered by the radio signal). Can have onboard memory behaving as ROM, WORM, rewritable. Can have circuitry to do inexpensive computations. Tx range: couple of metres for passive, maybe 10 m if active.

RFID reader, or transceiver, can read or write tags. Tx range: around 100 metres. Connected to a back-end database.

Reader must be able to "singulate" a tag (address it among the many in range). Anti-collision algorithm required. Probabilistic (eg Aloha) or deterministic (eg binary tree walking).

Tight cost requirements, else not ubiquitous. So tags limited to passive power, minimal memory (a few 100s bits), minimal circuitry (a few 1000s gates).

(b)
Auto-id code space is much larger -> can tag instances, not just classes

Auto-ID tag is costly whereas bar code tag practically free

But auto-id can be scanned without line-of-sight requirement

Scanning of auto-id is much faster (hundreds of reads / second)

Optical scanning cumbersome, so only done at POS; auto-id scanning could be ubiquitous, done many times over lifetime and usage pattern of object

(c)
Corporate espionage: retail inventory and sales data tracked by competitors via surreptitious scans

Personal privacy: you can be "X-rayed" by someone scanning what you're carrying and wearing (could also be a thief screening potential victims)

Location privacy: you can be recognized and tracked by the combination of objects you carry

Denial of service: making it impossible to read tags

Tag spoofing: pretending it's Coca Cola instead of Moet & Chandon at the check-out

Tag killing: breaking the tag so the removal (theft) of the tagged object will go undetected

(d)
The tag has an ID i, which is kept secret. The owner of the object knows i (e.g. it was inside the packaging when the item was bought).

When queried, the tag generates a random number r. It then transmits the pair (r, h(i|r)) where h() is hash and | is concatenation.

Message transmitted by tag is different every time, so eavesdropper can't identify the tag from the transmission.

The owner tries the i of all the objects she has (brute force, but then she doesn't have that many). When one matches, she knows which object transmitted.

Eavesdroppers can't do the same because they don't know the ids of that person's objects, so all they could do is to brute-force the whole code space of i, which we assume is infeasible.

Tag must have a random number generator and a hash function. Currently, this is not feasible on a 0.05 $ budget; however it is conceivable that we might get there in a few years.

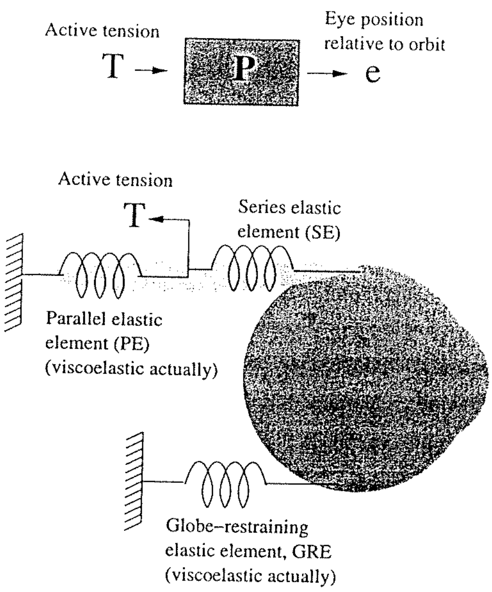
Method won't work for people (eg retailers) owning huge numbers of objects.

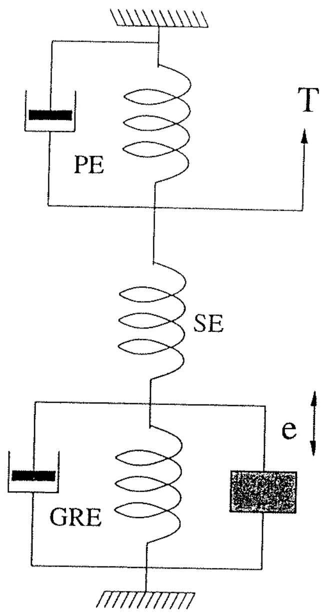
PAPER 8 SECTION G: Biological and Medical Engineering

Solutions to 2003 Tripos Paper

18. Oculomotor control

(a)





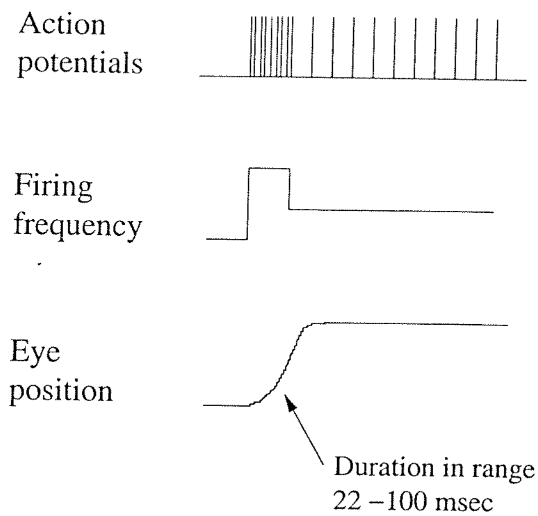
[4]

(b) The effect of inertia is negligible because the system is heavily over-damped, dominated by a low-frequency pole. Experimentally, the inertia has to be increased by a factor of around 100 before there is any sign of transient oscillation in a saccade. The visual effect of frequencies beyond some 30 Hz is negligible, because of the low-pass filtering exerted by the early stages of visual processing.

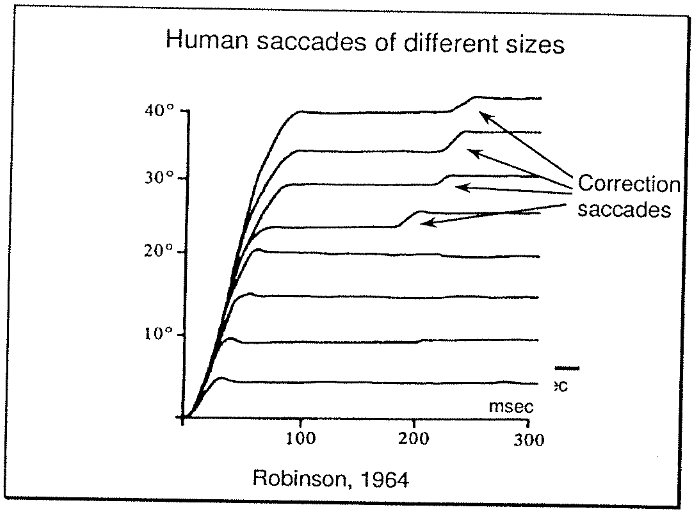
$$\frac{a}{\mathrm{D}+k} \quad \text{or} \quad \frac{a}{\tau s+1}$$

The characteristic frequency is about 0.5 Hz in Man (ω_c around 3, τ around 0.3), 1–2 Hz in cat. Candidates are not expected to know values for particular species: anything in this general region would be fine. [3]

(c)



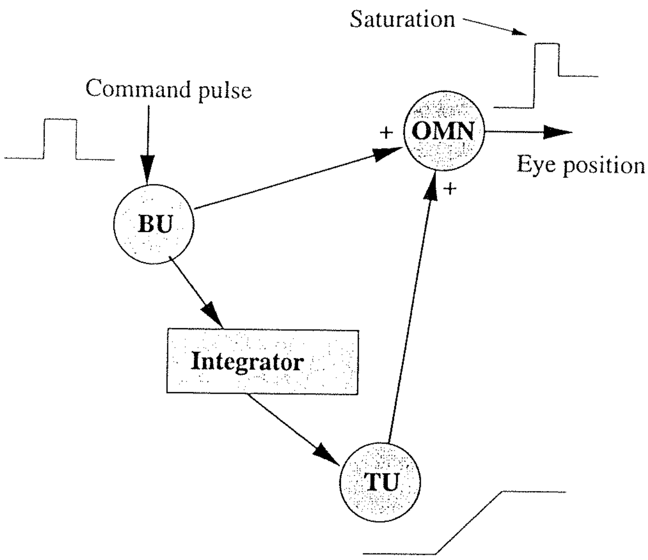
In more detail, (though not required).



[4]

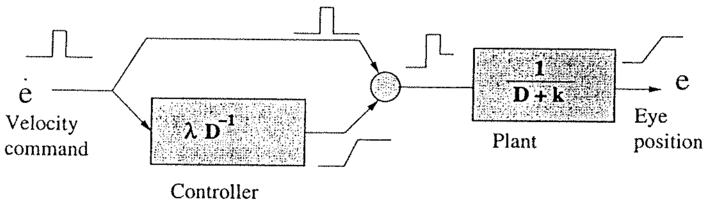
(d) The motor neurons receive inputs from two kinds of interneurons in the brainstem reticular formation: burst units (BU), that appear to receive the primary command signal, fire in a burst during a saccade; tonic units (TU) appear to be driven by burst units through an integrator, creating a pattern of firing similar to the time-course of the eye movement itself. The sum of these two signals creates the pulse-step frequency profile characteristic of the oculomotor motor neurons.

160



[3]

(e) The ratio of the gains of the two pathways, corresponding to the integrator route and the direct route (TU and BU) should ideally correspond with the characteristic frequency of the dominant low-pass behaviour of the plant. In other words, if the plant is $a/(\mathrm{D}+k)$, and we take the gain of the direct pathway as 1, and the gain of the integrator as λ, then λ must be equal to k.



The value of λ appears to be kept in step with any changes in k that may occur through, age, fatigue, disease, or growth, through a process of parametric feedback using visual information derived from retinal slip velocity, which is an error signal. [4]

(f) Experiments in monkeys have shown that if the movement of the eye in a saccade is unexpectedly hindered, the command firing pattern of the oculomotor neurons is unaffected; in other words, there is no evidence for immediate feedback during a saccade as to whether it is proceeding according to plan. [2]

Assessor's comments: This question tested the candidates understanding of the structure and function of the human oculomotor control system. It was the most popular question in the section, but resulted in a wide range of marks. A small number of students had a clear understanding of material, and were able to reproduce the block diagrams from the lectures accurately. As a result, they scored highly. The majority of candidates, however, produced surprisingly garbled answers resulting in the low average mark, despite a generous marking strategy.

161

19. **Bacterial navigation**

(a) In constructing an artefact such as a house or an engine, one has an overall plan to enable one to assemble the various components in the proper way. Such a plan is *external* to the components themselves.

In living organisms, by contrast — whether they be animals, insects, plants, fungi, bacteria — a prescription for the construction of the entire organism is contained in every individual cell; and when a cell divides into two cells, the full information is passed on to each.

In this way living organisms can reproduce into more-or-less identical offspring; and indeed it is difficult to see how life itself could occur without such a mechanism. It may seem extravagant to have the full genetic information in every single cell — compared to a single external plan — but this is how life works.

(After the first few rounds of cell division from a fertilised egg, cells begin to specialise into muscle, liver, eye or whatever, by "expressing" only the necessary genes; but all the other genes are still present, in an inactive form.) [4]

(b) Proteins consist of strings of amino acids attached to an unbranched polypeptide chain. (The sequence of amino acids is determined by the code in a portion of DNA which is "read" as the protein chain is assembled.)

There are 20 different kinds of amino acid; and the way in which the protein chain folds up depends precisely on the sequence of amino acids. Some amino acids are large and some are small; some are charged (+/−) while others are not; some are water repellent. The whole mass packs together, leaving no gaps, with the water-repellent (hydrophobic) amino acids buried in the interior, hidden from the surrounding water.

Proteins come in many sizes (typically 50 to 500 amino acids) and shapes (globular, tubular, stringy), with geometry allowing the construction of rings with 11 or 17, etc. members. Some act as electrical switches or gates. Some are transparent, and make lenses. And some are particularly strong because one kind of amino acid is able to form covalent bonds, thereby cross-linking the underlying flexible chain. [4]

(c) Bacteria such as *E. coli* propel themselves by rotating corkscrew-like "flagella"; each flagellum has a tiny rotary motor at its base, embedded in the cell wall. There are, typically, six flagella.

When the bacterium is swimming forwards, smoothly, the flagella associate to form a single bundle, which pushes the cell forwards.

Periodically, the motors switch into reverse rotation. It is easy to show that if the flagella were rigid, like wire corkscrews, this would soon lead to the tangling of the bundle and the stalling of the motors.

Nature has provided an ingenious way out of this problem. Although the flagellar filaments appear to be rigid in smooth swimming, they *switch* into helices of opposite handedness when driven in reverse; and so they are able to screw themselves out of the bundle when the motors reverse, and then wave around independently until the motors revert to their "normal" direction of rotation.

Bacteria depend on this remarkable ability to switch their sense of helical-handedness under mechanical torque in order to be able to navigate towards their nutrients. Such a switching is due to a very subtle form of construction from identical, mechanically bistable, protein molecules — far more subtle than any "smart structures" devised by human engineers. [4]

(d) The application of mechanical torque is one way in which bacteria flagellar filaments switch from one helical form to another: see (c).

Other ways are:

- different strains of bacteria (having slightly different amino acid sequences).
- change of pH (acidity).
- change of temperature.
- co-polymerisation of protein monomer (from different strains of bacteria).

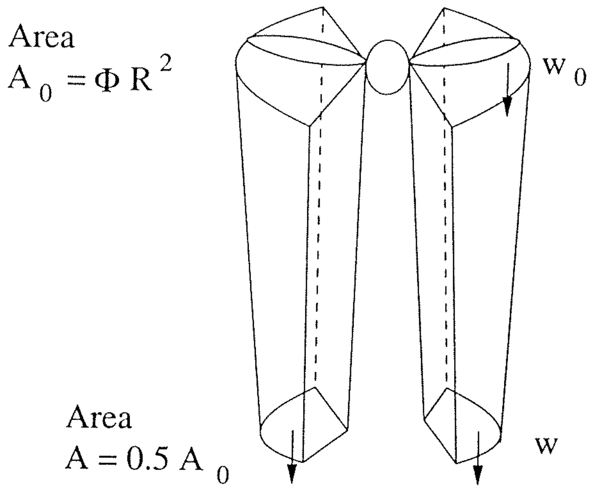
The key to understanding all of these changes is that each produces a small (tiny) change in the conformation of the protein building-block, which then assembles into a different member of the 12 different helical forms that can be constructed by the bistable subunits. [4]

(e) The amino-acid sequence of a protein is determined by the corresponding base-sequence in the DNA: each of the 20 types of amino acid is specified by a 3-letter sequence of DNA bases. If a single DNA code-letter (A, T, G or C) is changed for another — eg. by a random mutation or a copying error — then the triplet in which it occurs is changed, and a different amino acid may be specified. When the protein folds up, such a substitution may well be accommodated by a relatively small geometrical distortion of the conformation of the protein.

In the case of the *flagellin* protein, which is the building-block of the flagellar filament, such a minor change in geometry of the building-block can predispose the assembly of adopt another member of the 12 discrete helical waveforms that may occur. Thus a single-letter substitution in the flagellin's DNA code can (and does) cause a major change of conformation in the assembly of building blocks, viz — the flagellar filament. (This is a most subtle and unusual feature for a protein assembly: bacterial flagellar are very sophisticated structures.) [4]

Assessor's comments: The question covered the elementary structure of proteins and its application to the transport of bacteria. It was the least popular question in the section, but those students who attempted it had clearly revised carefully from the detailed notes provided and produced answers of a high calibre. The problem with "essay type" questions is often that the marks are all close together, and it may be difficult for good students to score highly. It was therefore pleasing to see a reasonable range of marks, many of which were high.

20. **Hovering animal and swimming fish**



(a)(i) Because $A_0 = 2A$, continuity requires $w_0 = \frac{1}{2}w$.

Find the downwash such that the jet force balances the weight: $F = mg$.

F is given by the momentum flux of the jet:

$$\begin{aligned} F &= \text{mass flux} \times \text{velocity} \\ &= (\rho A w)w = \rho A w^2 \\ \Rightarrow w &= \sqrt{\frac{mg}{\rho A}} \end{aligned}$$

Substitute expressions for w_0 and A_0:

$$\begin{aligned} 2w_0 &= \sqrt{\frac{mg}{\frac{1}{2}\rho A_0}} \\ w_0 &= \sqrt{\frac{mg}{2\rho A_0}} = \sqrt{\frac{mg}{2\rho \Phi R^2}} \end{aligned}$$

[4]

(a)(ii) Power can be calculated either as the rate of working by the 'actuator disc' $P = F.w_0$, or by the kinetic energy flux in the far wake

$$\begin{aligned} P &= \frac{1}{2} \times \text{mass flux} \times \text{velocity}^2 = \frac{1}{2}(\rho A w)w^2 \\ &= \frac{1}{2}F.w = F.w_0 \end{aligned}$$

Either way, $P/F = P/mg = w_0$. [3]

For $m = 75$ kg, $R = 2$ m, $\Phi = 2$ radians.

$$\begin{aligned} w_0 &= \sqrt{\frac{75 \text{ kg} \times 9.8 \text{ m/s}^2}{2 \times 1.2 \text{ kg/m}^3 \times 2 \text{ rad} \times 4 \text{ m}^2}} \\ &= 6.19 \text{ m/s} \end{aligned}$$

$$\begin{aligned} P &= F.w_0 = 75 \text{ kg} \times 9.8 \text{ m/s}^2 \times 6.19 \text{ m/s} \\ &= 4548 \text{ W} \end{aligned}$$

[3]

(a)(iii) These facts suggest that $\frac{P}{mg} \propto \frac{P}{m}$ must be relatively constant. Therefore w_0 is constant, which requires $R^2 \propto m$. So $R \propto m^{\frac{1}{2}}$, instead of $R \propto m^{\frac{1}{3}}$ as would be expected for isometric scaling. Therefore the wings of larger animals are relatively longer. [3]

(b)(i) Two distinct swimming modes are anguilliform and carangiform swimming motions. Snakes and eels swim in the anguilliform mode, with a travelling wave of nearly constant amplitude propagating aft at a certain speed. The wave length of the body wave is usually less than a body length. In carangiform type swimming the front portion of the body does not undulate, only the aft half or third of the body undulates. Most of these swimmers have a distinct foil-like tail that is used to augment thrust production. [4]

(b)(ii) Non-dimensional parameters: f = frequency of tail oscillation, A = amplitude of tail, U = swimming speed, L = fish length, ρ = fluid density, μ = fluid viscosity, λ = wavelength of swimming motion.

$$\begin{aligned} \text{Stroudhal number } St &= \frac{f(2A)}{U} \\ \text{Reynolds number } Rl &= \frac{\rho U L}{\mu} \\ \text{Length scaling} &= \frac{A}{L}, \frac{\lambda}{L} \end{aligned}$$

Stroudhal number characterises the wake dynamics and compares the oscillatory motion with the mean velocity scale. For different species of fish, the tail amplitude, and frequency will change, as will their optimal cruising speed. These non-dimensional numbers allow us to compare various swimmers in a uniform fashion. For example, the Stroudhal number for many fish falls around 0.3, however each species may have a very different size, swimming parameters, and average speed. [3]

Assessor's comments: This question led the candidates through the analysis of hovering using the momentum jet approach. The second part asked for some more general

165

comments about the ways in which different types of fish swim. It was generally well answered, although a significant number of candidates tried to reproduce the algebra from the lectures verbatim, and then got muddled when they were required to develop the argument in a slightly different way to that given in their notes. Some candidates were clearly using knowledge from the second year fluids course, rather than the approach given in the elective lectures, to enable them to solve the quantitative part of the question. This was pleasing, as it demonstrated that the elective linked into the general fluids lectures at an appropriate level.