Mathematics for Economists, Fourth Edition

Malcolm Pemberton and Nicholas Rau

SOLUTIONS TO PROBLEMS

Revised December 31, 2018

If you find any errors in these solutions,
please notify the authors at n.rau@ucl.ac.uk

1 LINEAR EQUATIONS

1–1. Let demand and supply schedules be respectively

$$q^D = ap + b, \quad q^S = cp + d.$$

Then $a = \dfrac{11-31}{8-4} = -5$, $b = 31 + 5\times 4 = 51$, $c = \dfrac{15-3}{12-8} = 3$ and $d = 3 - 3\times 8 = -21$, so

$$q^D = -5p + 51, \quad q^S = 3p - 21.$$

Therefore equilibrium price p^* and quantity q^* are given by

$$p^* = \frac{51+21}{3+5} = 9, \quad q^* = 3\times 9 - 21 = 6.$$

1–2. Gaussian elimination reduces the system to

$$\begin{array}{rcrcrcr} x & + & 3y & - & 2z & = & 2 \\ & & -11y & + & 5z & = & -4 \\ & & -11y & + & 5z & = & k-6 \end{array}$$

When $k = -4$, the last two equations are inconsistent and the system has no solution. When $k = 2$, the last two equations are identical and hence the third equation can be dropped. Then assigning $z = s$ and solving for y and then x gives the solution as

$$x = \frac{7s+10}{11}, \quad y = \frac{5s+4}{11}, \quad z = s.$$

1–3. Substitute the expression for T into that for C and the resulting expression for C into that for Y. Solving the resulting equation for Y gives

$$Y = 3.33 + 2.78(I+G), \quad C = 3.33 + 1.78(I+G), \quad T = 1.67 + 0.56(I+G).$$

If G increases by x units, Y, C and T increase by $2.78x$, $1.78x$ and $0.56x$ respectively.

1–4. The gross outputs x, y of X, Y satisfy

$$x - 0.1x - 0.2y = a, \quad y - 0.7x - 0.4y = b.$$

These equations lead to $x = 1.5a + 0.5b$, $y = 1.75a + 2.25b$. Since a and b are assumed to be positive numbers, and their coefficients in the equations for x and y are all positive, x and y are positive.

2 LINEAR INEQUALITIES

2–1. Let

$$A = c_0 - c_1 t_0, \quad B = I + G, \quad c = c_1(1 - t_1).$$

Substituting the expression for T into that for C gives $C = A + cY$. Substituting this expression for C into that for Y gives $Y = A + B + cY$. Hence

$$Y = \frac{A+B}{1-c}, \quad C = A + \frac{c}{1-c}(A+B) = \frac{A+cB}{1-c}, \quad T = t_0 + \frac{t_1}{1-c}(A+B).$$

The answers to the last two parts are Yes and No. Since c is the product of two numbers which are strictly between 0 and 1, $0 < c < 1$, whence $\dfrac{1}{1-c} > 1$. If G increases by x units, where $x > 0$, B increases by x units with A remaining unchanged, so Y increases by $\dfrac{x}{1-c} > x$ units. C increases by $\dfrac{cx}{1-c}$ units, which is less than x if and only if $c < 1 - c$, i.e. $c < \frac{1}{2}$.

2–2. The budget line has equation $p_1x_1 + p_2x_2 = m$ when $x_1 \leq z$. Hence the budget line has slope $-p_1/p_2$ to the left of z. Since the price of good 1 is $p_1 + t$ for all consumption in excess of z, the budget line has slope $-(p_1 + t)/p_2$ to the right of z.

(i) When $t < 0$, the budget line is less steep to the right of z than to the left.

(ii) When the consumption of good 1 is rationed at z, the budget line becomes vertical at z.

2–3. (i) Total usage of labour is $7x + 6y$, where x and y are the gross outputs of of X and Y respectively. Using the expressions for x and y given in the answer to Problem 1–4, total usage of labour is

$$7(1.5a + 0.5b) + 6(1.75a + 2.25b) = 21a + 17b.$$

Similarly, total usage of land is

$$3(1.5a + 0.5b) + 2(1.75a + 2.25b) = 8a + 6b.$$

(ii) Using the answer to (i), the conditions are the labour constraint $21a + 17b \leq 800$, the land constraint $8a + 6b \leq 300$ and the non-negativity constraints $a \geq 0$, $b \geq 0$. The corners of the feasible set in the ab–plane are $(0, 0)$, $(37.5, 0)$, $(30, 10)$ and $(0, 47.06)$.

2–4. Let x and y be the amounts of FB and KC consumed each day by Oleg. Then the cost minimisation programme is to *minimise* $2x + y$ subject to

$$10x + 4y \geq 20, \quad 5x + 5y \geq 20, \quad 2x + 6y \geq 12, \quad x \geq 0, \quad y \geq 0.$$

Once the feasible set has been drawn, it is clear that costs are minimised at the intersection of the calcium and protein borders. Here the slope of the isocost lines (-2) lies between the slopes of the two borders ($-\frac{5}{2}$ and -1). The required point of intersection is $(\frac{2}{3}, \frac{10}{3})$ and the least cost is

$$2\times\frac{2}{3} + \frac{10}{3} = \frac{14}{3}.$$

(i) The slope of the isocost lines is still -2 so the optimal combination is still $(\frac{2}{3}, \frac{10}{3})$; the least cost is now $\frac{28}{3}$.

(ii) The slope of the isocost lines is now $-\frac{3}{2}$ which still lies between $-\frac{5}{2}$ and -1. So the optimal combination is still $(\frac{2}{3}, \frac{10}{3})$; the least cost is now $\frac{26}{3}$.

(iii) The slope of the isocost lines is now -3, so these lines are now steeper than the calcium border and the optimal combination is $(0, 5)$. The least cost is now 5.

The solution is not unique when the isocost lines are parallel to one of the borders. Denoting the prices of FB and KC by p_1 and p_2 respectively, this will happen when p_1/p_2 is $\frac{5}{2}$, 1 or $\frac{1}{3}$.

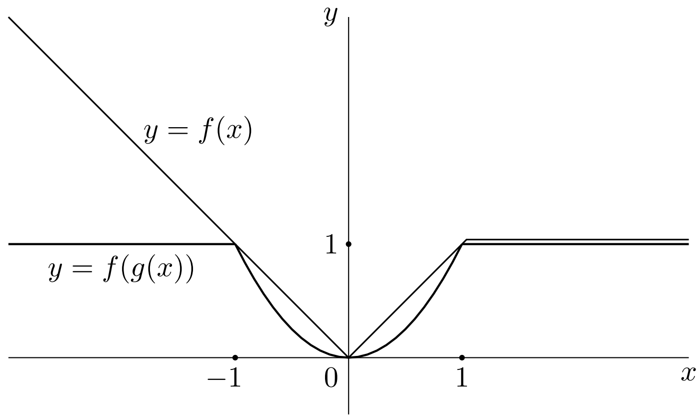
3 SETS AND FUNCTIONS

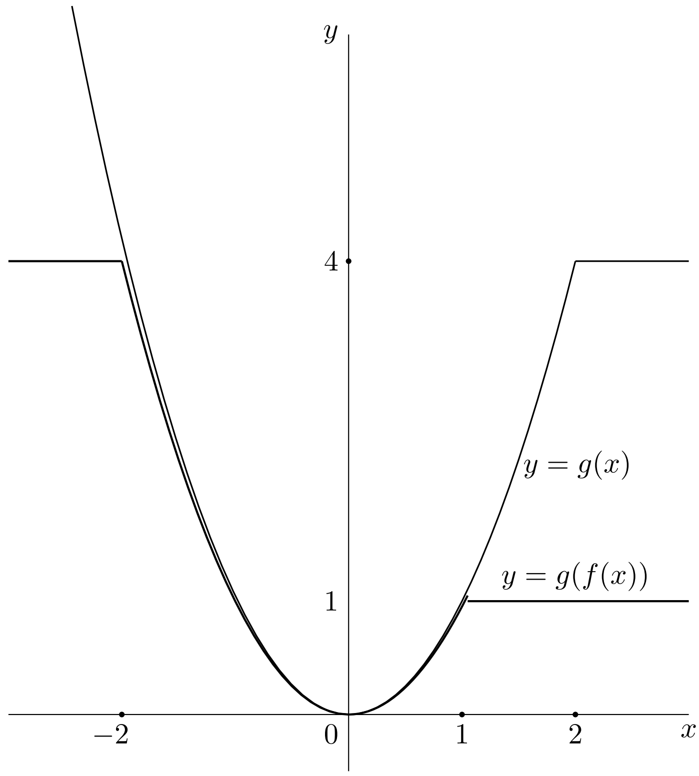
3–1. $f(g(x)) = 1$ if $g(x) \geq 1$, which happens if and only if $|x| \geq 1$. If $|x| < 1$ then $g(x) = x^2 < 1$, so $f(g(x)) = |x^2| = x^2$.

$g(f(x)) = 4$ if $f(x) \geq 2$, which happens if and only if $x \leq -2$. If $-2 < x < 1$ then $f(x) = |x| < 2$, so $g(f(x)) = |x|^2 = x^2$. If $x \geq 1$ then $f(x) = 1$, so $g(f(x)) = 1^2 = 1$.

Summarising,

$$f(g(x)) = \begin{cases} x^2 & \text{if } |x| < 1, \\ 1 & \text{if } |x| \geq 1; \end{cases} \qquad g(f(x)) = \begin{cases} 4 & \text{if } x \leq -2, \\ x^2 & \text{if } -2 < x < 1, \\ 1 & \text{if } x \geq 1. \end{cases}$$





3–2. The function takes the form $Y = X^2$. Hence the graph in the XY–plane is U-shaped with the bottom of the U at the origin. When $X = 0$, $x = 3$; when $Y = 0$, $y = 2$. Therefore $(3, 2)$ in the xy–plane corresponds to $(0, 0)$ in the XY–plane, and the graph in the xy–plane is U-shaped with the bottom of the U at $(3, 2)$.

A similar argument shows that the second graph is V-shaped with a right angle at $(2, 5)$.

3-3. (i) When $E > P$,

$$F(X) = \begin{cases} X + s(P - X) & \text{if } 0 \leq X < P, \\ X & \text{if } P \leq X < E, \\ X - t(X - E) & \text{if } X \geq E. \end{cases}$$

The graph consists of three line segments: the first has slope $1 - s$ and intercept sP, the second starts at the right end of the first and has slope 1, the third starts at the right end of the second and has slope $1 - t$.

(ii) When $E < P$ and $s + t < 1$,

$$F(X) = \begin{cases} X + s(P - X) & \text{if } 0 \le X < E, \\ X - t(X - E) + s(P - X) & \text{if } E \le X < P, \\ X - t(X - E) & \text{if } X \ge P. \end{cases}$$

The graph consists of three line segments: the first has slope $1 - s$ and intercept sP, the second starts at the right end of the first and has slope $1 - s - t$, the third starts at the right end of the second and has slope $1 - t$.

(iii) When $E < P$ and $s + t > 1$, the graph is similar to that in (ii) except that the middle segment now has negative slope.

3–4. (i) $f(v_3, v_4, v_5, v_6) = v_3 + v_4 + v_5 + v_6$, 4 variables;

(ii) $f(v_2, v_3, v_4, v_5, v_6) = v_2 - v_3 - v_4 - v_5 - v_6$, 5 variables;

(iii) $f(v_1, v_2) = v_2/v_1$, 2 variables;

(iv) $f(v_1, v_4) = v_4/v_1$, 2 variables;

(v) $f(v_1, v_2, v_3, v_4, v_5, v_6) = (v_2/v_1,\ (v_3 + v_4 + v_5 + v_6)/v_1)$, 6 variables;

(vi) $f(v_1, v_2, v_3, v_4, v_5, v_6) = (v_2/v_1,\ (v_2 - v_3 - v_4 - v_5 - v_6)/v_1,\ v_5/v_1)$, 6 variables.

4 QUADRATICS, INDICES AND LOGARITHMS

4–1. The graph of the first equation is $\bigcap$-shaped with vertex at $(0, 5)$; the graph of the second equation is a straight line with slope 2 and intercept -3. Eliminating q between the two equations gives $5 - p^2 = 2p - 3$, i.e. $p^2 + 2p - 8 = 0$. This factorises to $(p + 4)(p - 2) = 0$ so p is -4 or 2. When $p = -4$, substituting back into either of the original two equations equation gives $q = -11$; similarly, when $p = 2$, $q = 1$.

The equilibrium price and quantity are 2 and 1 respectively.

4–2. (i) $f(x) = \left(\sqrt{ax} - \sqrt{c/x}\right)^2 + b + 2\sqrt{ac}$. Hence $f(x)$ is minimised when $ax = c/x$, i.e. when $x = \sqrt{c/a}$. The minimum value of $f(x)$ is $b + 2\sqrt{ac}$.

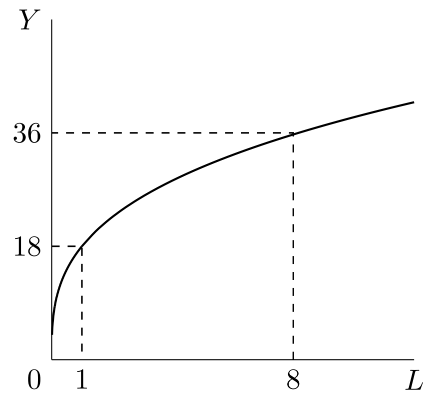
(ii) Average cost is $0.08x + 2 + 50/x$. From (i), this is minimised when $x = \sqrt{50/0.08} = 25$ and its minimum value is $2 + 2\sqrt{50 \times 0.08} = 6$.

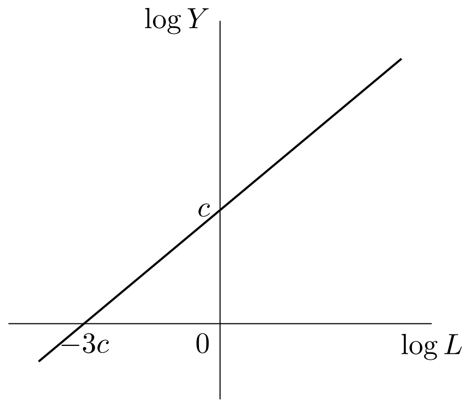
4-3. (i) Suppose K and L both increase by 1%. Let old value of Y be Y_0, new value Y_1. Then

$$Y_1 = 2(1.01\,a)^{2/3}(1.01\,b)^{1/3} = 2a^{2/3}b^{1/3}(1.01)^{2/3}(1.01)^{1/3} = Y_0 \times 1.01,$$

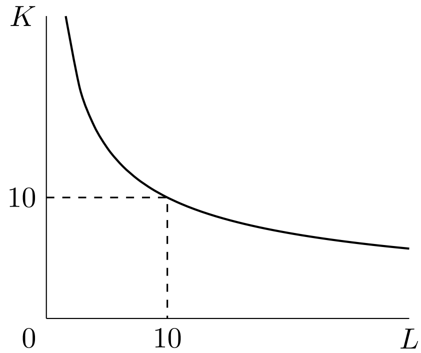
so Y increases by 1%. Similar argument holds if 1.01 is replaced by 1.10 or by $1 + \dfrac{x}{100}$ for any $x > 0$ (or, more generally, any $x > -100$). Thus if K and L both increase by 10% (or x%), then Y increases by 10% (resp. x%).

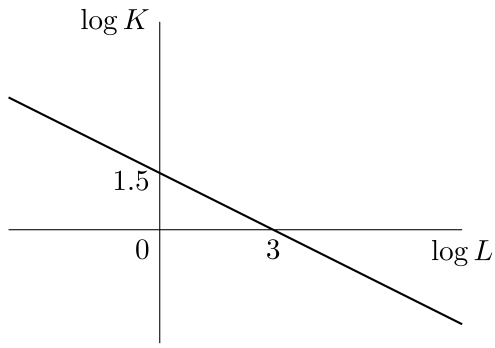
(ii) $Y = 18L^{1/3}$, $\log Y = c + \frac{1}{3}\log L$, where $c = \log 18 = 1.255$ to 3 decimal places. Here, as in (iii) below, logarithms are to base 10.





(iii) $K = \sqrt{1000/L}$, $\log K = g - \frac{1}{2}\log L$, where $g = \frac{1}{2}\log 1000 = 1.5$.





4-4. The graph of $q = \frac{1}{4}p^4$ has a U shape with the bottom of the U at the origin. The part in the positive quadrant is the graph of the supply function. Let $p > 0$; as p increases, $1/p$ decreases, so $q = 8p^{-1}$ decreases as p increases. The equilibrium occurs where $\frac{1}{4}p^4 = 8p^{-1}$, i.e. $p^5 = 32$; thus the equilibrium price and quantity are 2 and 4.

The supply and demand functions in log-linear form are

$$\log q = -\log 4 + 4\log p, \quad \log q = \log 8 - \log p.$$

Solving these linear equations for $\log p$ and $\log q$ gives

$$\log p = \tfrac{1}{5}(\log 8 + \log 4) = \tfrac{1}{5}(\log 32) = \tfrac{1}{5}\log 2^5 = \log 2$$

and $\log q = \log 8 - \log 2 = \log 4$. Hence $p = 2$ and $q = 4$.

5 SEQUENCES, SERIES AND LIMITS

5–1. $u_n = 3 - \dfrac{9}{n+3}$. As $n \to \infty$, $\dfrac{9}{n+3} \to 0$, so $u_n \to 3$. Since $\dfrac{9}{n+3} < \dfrac{10}{n}$, it is clear that for $p = 10^{-2}$, $N = 10^3$ satisfies (5.1). [This of course is not the smallest value but you are not asked for that.] The same argument gives the following table of values of N satisfying (5.1) for various values of p.

p	10^{-2}	10^{-3}	10^{-4}	10^{-5}
N	10^3	10^4	10^5	10^6

$v_n = \dfrac{3}{n + 3n^{-2}}$. When n is large, the term n^{-2} is negliglible; so for large n the sequence $\{v_n\}$ behaves similarly to $\{3/n\}$. Therefore $v_n \to 0$ as $n \to \infty$.

By similar arguments $3x/(x+3) \to 3$ and $3x^2/(x^3+3) \to 0$ as $x \to \infty$.

5-2. (i) Flat rate is $100r\%$ where $r = 12s$.

(ii) APR is $100r'\%$ where

$$r' = \left(1 + \frac{12s}{12}\right)^{12} - 1 = (1+s)^{12} - 1.$$

When $r' = 0.20$, $s = 0.0153$ and when $r' = 0.25$, $s = 0.0188$. So the interest rate per month increases from 1.53% to 1.88%.

5-3. (i) From the second equation, the further increase of c_1 units in Y causes a further increase of c_1^2 units in C which then, by the first equation, causes a further increase of c_1^2 units in Y. Continuing in this way, the total increase in Y is $1 + c_1 + c_1^2 + \ldots$ Since $0 < c_1 < 1$, we may apply the geometric series formula: the sum is $1/(1 - c_1)$.

(ii) Substituting the expression for C into that for Y and solving the resulting equation for Y, we have

$$Y = \frac{c_0 + I + G}{1 - c_1}, \quad C = \frac{c_0 + c_1(I+G)}{1 - c_1}.$$

From the expression for Y, if G increases by one unit, Y will increase by $1/(1 - c_1)$, which agrees with the answer obtained in (i).

(iii) Same answer as in (ii).

5-4. (i) The profit obtained at time T is $pf(T)$, so the value of the forest at time 0 is $\dfrac{pf(T)}{(1+r)^T}$.

(ii) The value of the forest at time 0 is

$$\frac{pf(T)}{(1+r)^T} + \frac{pf(T)}{(1+r)^{2T}} + \frac{pf(T)}{(1+r)^{3T}} + \ldots$$

This is a geometric progression with first term $xpf(T)$ and common ratio x, where $x = (1+r)^{-T}$. Since $0 < x < 1$, the sum is $xpf(T)/(1-x)$. The value of the forest at time 0 is therefore $pf(T)\big/\big((1+r)^T - 1\big)$.

6 INTRODUCTION TO DIFFERENTIATION

6–1. $\dfrac{dy}{dx} = 3x^2$. If $x = 2$ then $y = 8$ and $\dfrac{dy}{dx} = 12$, so the equation of the tangent is

$$y - 8 = 12(x - 2),$$

or $y = 12x - 16$.

By the small increments formula,

$$(2+h)^3 - 8 \approx 12h.$$

To verify that this is a good approximation if h is small, notice that

$$\begin{aligned} \text{LHS} - \text{RHS} &= 2^3 + 3\times 2^2 \times h + 3 \times 2 \times h^2 + h^3 - 8 - 12h \\ &= 4(2+3h) + h^2(6+h) - 4(2+3h) \\ &= h^2(6+h). \end{aligned}$$

If h is small then h^2 is very small and $6 + h \approx 6$, so LHS − RHS is indeed very small.

If $x = 2 + h$, the value of the function is $(2 + h)^3$ and the value of y given by the tangent is $8 + 12h$. The error of approximation is the same as RHS − LHS in the small increments formula and is therefore equal to $-h^2(6 + h)$. The ratio of the absolute value of the error to the true value of the function is

$$\frac{h^2(6+h)}{(2+h)^3},$$

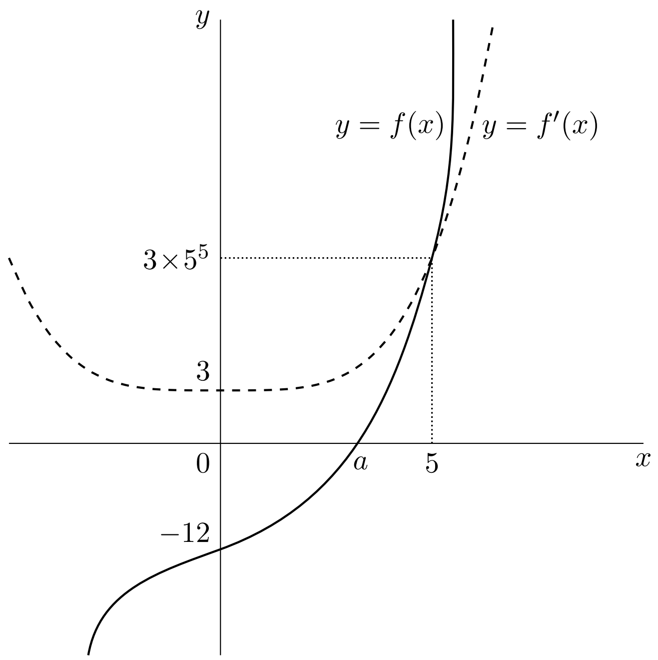
which is

$$\frac{6.01}{2.01^3} \times 10^{-4} \text{ if } h = 0.01, \quad \frac{7}{27} \text{ if } h = 1.$$

The required percentages are (i) 0.0074%, (ii) 25.93%.

6–2. Let $f(x) = x^5 + 3x - 12$. Then $f'(x) = 5x^4 + 3$, which is positive for all x. The curve $y = f(x)$ is therefore upward-sloping, with slope 3 at the point $(0, -12)$ and the slope increasing as we move away from the y–axis in either direction. Thus the curve cuts the x–axis exactly once, at a point $(a, 0)$ such that $a > 0$. One can see without using a calculator that a is slightly less than 1.5: for if $x = 1.5$ then $x^5 = 243/32$, which is slightly greater than 7.5, while $3x - 12 = -7.5$, so $f(x)$ is positive but very small.

The curve $y = f'(x)$ has the same general U-shape as the curve $y = x^4$, but with the vertex at the point $(0, 3)$. [The shapes of the power functions were introduced without explanation in Chapter 4, but we can now see why the curve $y = x^4$ looks as it does. Obviously the curve passes through the origin. Also, since the slope is $4x^3$, the curve is downward-sloping where $x < 0$, upward-sloping where $x > 0$, and the absolute value of the slope increases as we move away from the origin in either direction.]



$f(0) < 0 < f'(0)$ and $f(x) > f'(x)$ if x is positive and very large, so the curves $y = f(x)$ and $y = f'(x)$ cross at least once. Since $f(x) \leq 0 < f'(x)$ if $x \leq a$, the crossing-point(s) must be such that $x > a$. In fact,

$$f(x) - f'(x) = x^5 - 5x^4 + 3x - 15 = (x^4 + 3)(x - 5),$$

which is zero if *and only if* $x = 5$. Thus there is exactly one crossing-point, namely the point $(5,\ 5^5 + 3)$.

6–3. Revenue $R(x)$ and marginal revenue $R'(x)$ are given by

$$R(x) = px = \tfrac{1}{3}(10-x)x = \tfrac{1}{3}(10x - x^2)$$

and $R'(x) = \frac{2}{3}(5-x)$. If sales increase from 3 to 4 then revenue increases from 7 to 8, an increase of 1 unit. By contrast, $R'(3) = 4/3$. That this is not a good approximation to the true increase shows that, in this case, $h = 1$ is not small enough for the small increments formula to be accurate.

When sales increase from 3 to 3.1, the approximate change in revenue is $\frac{2}{3} \times 2 \times 0.1 = 0.133$ to three decimal places. [The true increase is $\frac{1}{3}(1 - 6.1 \times 0.1) = 0.13$, so in this case the approximation is good.]

6–4. Suppose p increases from p_0 by Δp and the corresponding increase in q from q_0 is Δq. By the small increments formula, $\Delta q \approx f'(p_0)\Delta p$. Denote the elasticity of demand at (p_0, q_0) by ε, so that $\varepsilon = p_0 f'(p_0)/q_0$. Substituting $f'(p_0) = \varepsilon q_0/p_0$ into the approximation for Δq gives $\Delta q \approx \varepsilon q_0 \Delta p/p_0$. Therefore,

$$\text{if } \frac{\Delta p}{p_0} = \frac{1}{100}, \text{ then } \frac{\Delta q}{q_0} \approx \frac{\varepsilon}{100}:$$

the approximate percentage change in quantity demanded is $\varepsilon\%$.

(i) The small increments formula is exact and takes the form $\Delta q = -b\Delta$. The percentage change in quantity demanded is exactly $\varepsilon\%$ where $\varepsilon = -bp_0/q_0$ notice that ε depends on p_0 and q_0.

(ii) The elasticity of demand is $-n$ and is therefore independent of the initial point (p_0, q_0). Thus, for any 1% change in p, the corresponding percentage change in q is approximately $-n\%$. Notice that this is an approximation, not an exact formula; it is not hard to see that the percentage error is independent of the initial point.

7 METHODS OF DIFFERENTIATION

7–1. Substituting for K and L gives

$$Q = (5+2t)^{1/2}(2+t)^{1/3}.$$

Let $u = (5+2t)^{1/2}$, $v = (2+t)^{1/3}$. By the composite function rule,

$$\frac{du}{dt} = \frac{1}{2}(5+2t)^{-1/2} \times 2 = \frac{u}{5+2t}, \quad \frac{dv}{dt} = \frac{1}{3}(2+t)^{-2/3} \times 1 = \frac{v}{3(2+t)}.$$

Hence by the product rule,

$$\frac{dQ}{dt} = \frac{vu}{5+2t} + \frac{uv}{3(2+t)} = \frac{uv(6+3t+5+2t)}{3(5+2t)(2+t)}.$$

Simplifying,

$$\frac{dQ}{dt} = \frac{11+5t}{3(5+2t)^{1/2}(2+t)^{2/3}}.$$

7–2. (i) $y^{1/3} = cx^{-1/2}$, so $y = c^3x^{-3/2}$. $dy/dx = -bx^{-5/2}$, where b is the positive constant $3c^3/2$.

(ii) The equation of a typical isoquant is

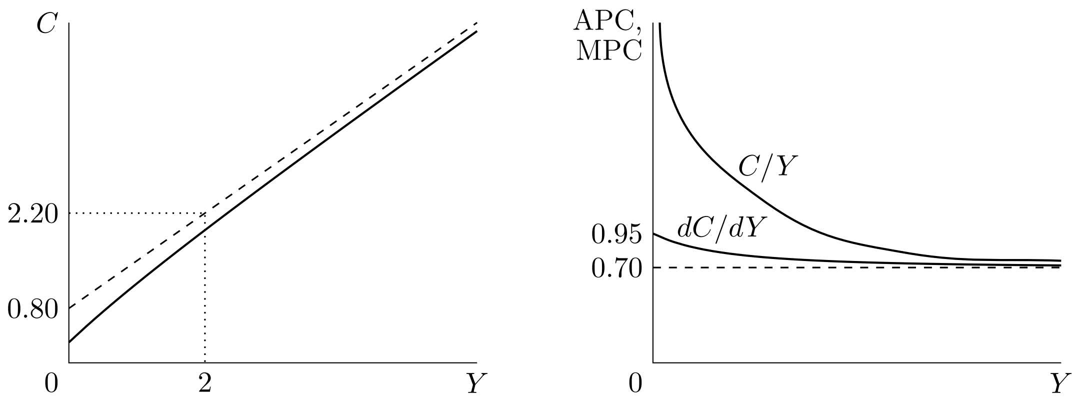
$$K^{1/2}L^{1/3} = c \qquad (K, L > 0)$$

where c is a positive constant. From (i) $dL/dK = -bK^{-5/2}$, where b is a positive constant. Hence $dL/dK < 0$.

7–3. $dC/dY = 0.7 + (Y+2)^{-2} > 0$, so C is monotonic increasing in Y. dC/dY decreases from 0.95 where $Y = 0$, tending to the limit 0.7 as $Y \to \infty$. Graph of C against Y has asymptote $C = 0.8 + 0.7Y$, the broken line in the left-hand panel of the diagram.

$$\frac{C}{Y} = 0.7 + \frac{0.8}{Y} - \frac{1}{Y(Y+2)} = 0.7 + \frac{0.3}{Y} + \frac{0.5}{Y+2},$$

using the Hint, which enables us to check that C/Y is a monotonic function of Y without further messy differentiation. In fact, the graph against Y of C/Y (the APC), like that of dC/dY (the MPC), is downward-sloping, with the same horizontal asymptote: both APC and MPC tend to 0.7 as $Y \to \infty$. The vertical axis is also an asymptote of APC.



7–4. The inverse function of the function $y = f(x)$ is

$$x = \begin{cases} \left(1 - \sqrt{1+y^2}\right) y^{-1}, & \text{if } y \neq 0; \\ 0, & \text{if } y = 0. \end{cases}$$

The negative square root of $1 + y^2$ is chosen, so as to ensure that x and y have opposite signs.

$$\frac{dx}{dy} = -\frac{1}{\sqrt{1+y^2}} - \frac{1-\sqrt{1+y^2}}{y^2} = -\frac{\left(x^2-1\right)^2}{2\left(x^2+1\right)},$$

where the first expression (call it a) is found by direct calculation and the second (b) by the inverse function rule. To verify that $a = b$, note that

$$-a = \frac{1}{1-xy} + \frac{x}{y} = \frac{x^2-1}{x^2-1-2x^2} + \frac{x^2-1}{2} = \frac{x^2-1}{2}\left(1 - \frac{2}{x^2+1}\right) = \frac{\left(x^2-1\right)^2}{2(x^2+1)} = -b.$$

8 MAXIMA AND MINIMA

8–1. $dy/dx = 5x^4(2-x)^4 - 4x^5(2-x)^3 = x^4(2-x)^3(10-9x)$; this is zero if x is 0, 10/9 or 2. So the critical points are $(0,0)$, (1.111, 1.057) and $(2,0)$.

When $x = 0-$, $dy/dx = (+)(+)(+) = +$; when $x = 0+$, $dy/dx = (+)(+)(+) = +$; hence $(0,0)$ is a critical point of inflexion. Similarly (1.111, 1.057) is a maximum point and $(2,0)$ is a minimum point.

$$\begin{aligned} d^2y/dx^2 &= 20x^3(2-x)^4 - 40x^4(2-x)^3 + 12x^5(2-x)^2 \\ &= x^3(2-x)^2\left(20(4-4x+x^2) + (-80x + 40x^2 + 12x^2)\right). \end{aligned}$$

Simplifying, $d^2y/dx^2 = 8x^3(2-x)^2(9x^2-20x+10) = 8x^3(2-x)^2([3x-\frac{10}{3}]^2 - \frac{10}{9})$. It follows that d^2y/dx^2 is 0 at four values of x, namely 0, $(10-\sqrt{10})/9$, $(10+\sqrt{10})/9$ and 2, but changes sign only at the first three. So the points of inflexion are $(0,0)$, (0.760, 0.599) and (1.462, 0.559). The function is

(i) convex for $0 < x < 0.760$ and $x > 1.462$,

(ii) concave for $x < 0$ and $0.760 < x < 1.462$.

The information needed to sketch the curve is completed by noting that as $x \to -\infty$, $y \to -\infty$; and as $x \to \infty$, $y \to \infty$.

8-2 Let O denote the origin, and let P be some other point on the curve $y = f(t)$; then $f(t)/t$ is the slope of OP. Once the diagram showing the shape of $f(t)$ has been drawn, it is clear that the slope of OP is at its global maximum when OP is tangential to the curve; further, there is only one point on the curve with this property.

The tangent to the curve at P has slope $f'(t)$. Since this tangent and the line OP both pass through P, they are identical if and only if they have the same slope. Thus the optimal t is given uniquely by the equation $f'(t) = f(t)/t$.

The same equation can be obtained by noting that

$$\text{if } z = \frac{f(t)}{t} \text{ then } \frac{dz}{dt} = \frac{tf'(t) - f(t)}{t^2},$$

which is zero when $tf'(t) = f(t)$. However, the geometric argument given above provides the easiest way of seeing that the critical point must be the global maximum.

8–3. From the production function,

$$L(Q-K) = KQ.$$

Hence the isoquant $Q = Q_0$ may be written

$$L = \frac{KQ_0}{K-Q_0}.$$

Along the isoquant,

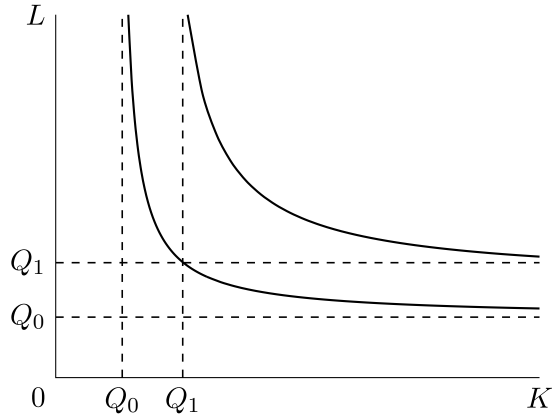
$$\frac{dL}{dK} = \frac{Q_0}{K-Q_0} - \frac{KQ_0}{(K-Q_0)^2} = -\left[\frac{Q_0}{K-Q_0}\right]^2 < 0$$

and

$$\frac{d^2L}{dK^2} = \frac{2Q_0^2}{(K-Q_0)^3}.$$

From the production function, Q must be less than K if $L > 0$. Hence $K - Q_0 > 0$, so $d^2L/dK^2 > 0$.

Since $dL/dK < 0$ and $d^2L/dK^2 > 0$, the isoquant is downward-sloping and convex. If K is very large and $Q = Q_0$ then L is close to Q_0. Similarly, if L is very large and $Q = Q_0$ then K is close to Q_0. Hence the asymptotes of the isoquant are the lines $K = Q_0$ and $L = Q_0$. Note also that the isoquant crosses the line $L = K$ at the point $(2Q_0, 2Q_0)$ and its slope at that point is -1. The diagram shows two isoquants, $Q = Q_0$ and $Q = Q_1$, where $Q_0 < Q_1 < 2Q_0$.



8–4. (i) Differentiating the average revenue function,

$$\begin{aligned}\frac{dp}{dx} &= -1800 + 100x - \tfrac{3}{2}x^2 \\ &= -\tfrac{3}{2}\left[\left(x - \tfrac{100}{3}\right)^2 - \left(\tfrac{100}{3}\right)^2 + 1200\right] \quad \text{by completing the square} \\ &\le 100\left(\tfrac{100}{6} - 18\right) < 0,\end{aligned}$$

so AR is monotonic.

(ii) $\text{MR} = \dfrac{d}{dx}(px) = 36000 - 3600x + 150x^2 - 2x^3$, so

$$\frac{d}{dx}\text{MR} = -3600 + 300x - 6x^2 = -6(x-20)(x-30).$$

Thus MR is not monotonic, being a decreasing function of x for $0 \le x < 20$ and $x > 30$, and an increasing function for $20 < x < 30$.

(iii) Both graphs meet the vertical axis at 36000. AR is monotonic decreasing with a point of inflexion at $x = 100/3$, where the curve changes from convex to concave. MR has a minimum at $x = 20$ and a maximum at $x = 30$. AR is always above MR: this follows from the fact that $dp/dx < 0$. Both AR and MR are negative for all sufficiently large x: AR $= 0$ when $x = 60$, MR is positive when $x = 40$ but negative when $x = 45$.

9 EXPONENTIAL AND LOGARITHMIC FUNCTIONS

9–1. (i) $dy/dx = -2ax\exp(-ax^2) = -2axy$, which always has the opposite sign to x. Thus the only critical point is $(0, 1)$, and this is the global maximum.

(ii) By the product rule,

$$\frac{d^2y}{dx^2} = -2a\left(y + x\frac{dy}{dx}\right) = -2a(y - 2ax^2y) = 2ay(2ax^2 - 1),$$

which always has the same sign as $2ax^2 - 1$. Let $b = (2a)^{-1/2}$: the function is convex for $|x| > b$, concave for $|x| < b$, and the points of inflexion occur where $x = \pm b$. Since $ab^2 = \frac{1}{4}$, the points of inflexion are $((2a)^{-1/2}, e^{-1/4})$ and $(-(2a)^{-1/2}, e^{-1/4})$.

(iii) Bell-shaped, asymptotic to the x–axis, with global maximum given by (i) and points of inflexion given by (ii).

9–2. (i) Suppose the original sum of money is P. Then T is such that $Pe^{rT} = 2P$. Therefore $rT = \ln 2$, and

$$T = \frac{\ln 2}{r} = \frac{100 \ln 2}{R} \approx \frac{69}{R},$$

since $\ln 2 = 0.69315$ to 5 decimal places.

(ii) Let the APR be $S\%$, and let $s = 0.01S$. By the answer to Exercise 9.2.1, $\ln(1+s)$ is what we called r in part (i) of this problem. Hence

$$T = \frac{\ln 2}{\ln(1+s)} = \frac{\ln 2}{S} \times \frac{S}{\ln(1+0.01S)} = \frac{A}{S}, \quad \text{where } A = \frac{S \ln 2}{\ln(1+0.01S)}.$$

The values of A for different values of S are given in the following table:

S	2	4	6	8	10
A	70.0	70.7	71.4	72.1	72.7

9–3 (i) Let $f(x) = e^{ax}$. Then $f(0) = 1$, so

$$\lim_{x \to 0} \frac{e^{ax} - 1}{x} = \lim_{x \to 0} \frac{f(x) - f(0)}{x - 0} = f'(0) = ae^{a \times 0} = a.$$

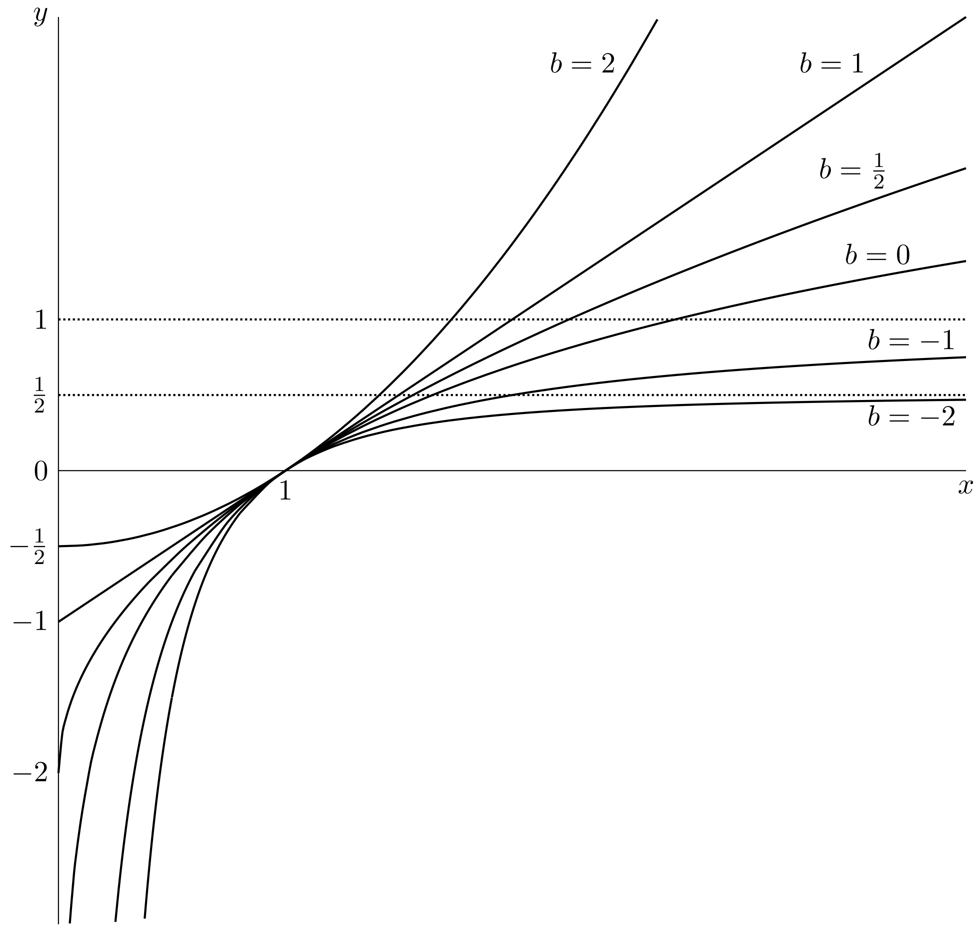
The second result follows by interchanging a and x.

(ii) $f_a{}'(x) = e^{ax}$, $f_a{}''(x) = ae^{ax}$, $f_a(0) = 0$, $f_a{}'(0) = 1$.
All curves are upward-sloping. For $a = 0$ the curve is the straight line $y = x$; the curves for $a \neq 0$ are all tangential to that line at the origin. For $a = 1$ and $a = 5$, the curves are convex and hence lie above $y = x$, the graph of f_5 being more curved than the graph of f_1. For $a = -1$ and $a = -5$, the curves are concave and hence lie below $y = x$, the graph of f_{-5} being more curved than the graph of f_{-1}.

(iii) For $x > 0$, $x^b = e^{b \ln x}$. Hence, by the second result of (i),

$$\lim_{b \to 0} g_b(x) = \lim_{b \to 0} \frac{e^{b \ln x} - 1}{b} = \ln x = g_0(x).$$

$g_b{}'(x) = x^{b-1}$, $g_b{}''(x) = (b-1)x^{b-2}$, $g_b(1) = 0$, $g_b{}'(1) = 1$.
All curves are upward-sloping. For $b = 1$ the curve is the straight line $y = x - 1$; the curves for $b \neq 1$ are all tangential to that line at the point $(1, 0)$. For $b > 1$, and in particular for $b = 2$, the curve is strictly convex. For $b < 1$, the curve is strictly concave, being more curved the lower is b. For all $b > 0$, $g_b(0) = -1/b$ and $g_b(x) \to \infty$ as $x \to \infty$. If $b = 0$ then $g_b(x) = \ln x$ for all $x > 0$; 0 is the only value of b such that $g_b(x) \to -\infty$ as $x \to 0$ and $g_b(x) \to \infty$ as $x \to \infty$. If $b < 0$, and in particular for $b = -1$ and $b = -2$, $g_b(x) \to -\infty$ as $x \to 0$ and $g_b(x) \to -1/b$ as $x \to \infty$.



9–4. (i) Let the value of the forest at time 0 be $v(T)$. Since the profit obtained at time T is $pf(T)$, $v(T) = pf(T)e^{-rT}$. Therefore

$$v'(T) = pf'(T)e^{-rT} - pf(T)re^{-rT} = (f'(T) - rf(T))pe^{-rT},$$

which is zero when $f'(T)/f(T) = r$.

(ii) In this case, let the value of the forest at time 0 be $V(T)$. Then

$$V(T) = pf(T))e^{-rT} + pf(T)e^{-2rT} + pf(T)e^{-3rT} + \dots\,,$$

the sum of the geometric progression whose first term is $v(T)$ (as defined in (i)) and whose common ratio is e^{-rT}. Since $0 < e^{-rT} < 1$,

$$V(T) = \frac{v(T)}{1 - e^{-rT}} = \frac{pf(T)}{e^{rT} - 1}.$$

To find $V'(T)$ it is easiest to take logs and then differentiate:

$$\frac{V'(T)}{V(T)} = \frac{d}{dT}\big(\ln p + \ln f(T) - \ln(e^{rT} - 1)\big) = \frac{f'(T)}{f(T)} - \frac{re^{rT}}{e^{rT} - 1}.$$

Thus $V'(T) = 0$ when (9.7′) holds.

(iii) The right-hand side of (9.7′) may be written $\dfrac{a}{e^{aT} - 1}$, where $a = -r$. The required result now follows from the second result of Problem 9–3, part (i).

(iv) Differentiating $v'(T)$ gives $v''(T) = (f''(T) - 2rf'(T) + r^2 f(T))pe^{-rT}$. Using condition (9.7) we see that, at the critical point, $v''(T) = (f''(T) - r^2 f(T))pe^{-rT}$. Therefore, if the value of T for which $v''(T) = 0$ is in the region for which f is concave, the critical value is a local maximum. If we assume further that there is only one critical point, it must be the global maximum.

Similarly, differentiating $V'(T)/V(T)$ gives

$$\frac{V''(T)}{V(T)} - \left[\frac{V'(T)}{V(T)}\right]^2 = \frac{f''(T)}{f(T)} - \left[\frac{f'(T)}{f(T)}\right]^2 - r\frac{d}{dt}\left(1 + (e^{rT} - 1)^{-1}\right).$$

Using condition (9.7$'$) we see that at the critical point

$$\frac{V''(T)}{V(T)} = \frac{f''(T)}{f(T)} - \left[\frac{re^{rT}}{e^{rT} - 1}\right]^2 + \frac{r^2 e^{rT}}{(e^{rT} - 1)^2},$$

whence

$$V''(T) = \frac{V(T)}{f(T)}f''(T) - \frac{r^2 e^{rT} V(T)}{e^{rT} - 1} = \frac{p}{e^{rT} - 1}f''(T) - \frac{pr^2 e^{rT}}{(e^{rT} - 1)^2}f(T).$$

The argument concerning the global maximum is similar to that for $v(T)$.

10 APPROXIMATIONS

10–1. (i) $f(x) = 0$ has at most one root. For if there were two distinct roots, say a and b where $a < b$, then by Rolle's theorem there would be a real number c, with $a < c < b$, which is a root of $f'(x) = 0$.

(ii) $f(x) = 0$ has at most two roots. For if there were three distinct roots, say a, b, c where $a < b < c$, then by Rolle's theorem there would be real numbers p and q, with $a < p < b < q < c$, which are roots of $f'(x) = 0$.

(iii) $f(x) = 0$ has at most three roots, by a similar argument to (i) and (ii).

General result: if $f'(x) = 0$ has n distinct roots, then $f(x) = 0$ has at most $n + 1$ roots.

10–2. (i) Let $f(x) = x^5 - 5x + 2$; then $f'(x) = 5x^4 - 5$ and $f''(x) = 20x^3$. $f'(x) = 0$ when $x = \pm 1$; using $f''(x)$, it follows that $(-1, 6)$ is a maximum point and $(1, -2)$ is a minimum point. As $x \to -\infty$, $f(x) \to -\infty$; as $x \to \infty$, $f(x) \to \infty$.

(ii) Since $f(-2) = -20$ and $f(-1.5) \approx 2$, x_1 is between -2 and -1.5. Since $f(1) = -2$ and $f(1.5) \approx 2$, x_3 is between 1 and 1.5. Taking -1.6 as the initial approximation to x_1 and using Newton's method, $V(-1.6) = -0.01749$ and so the second approximation is -1.58251. Applying the method again, $V(-1.58251) = -0.00047$, so $x_1 = -1.582$ to 3 decimal places. Similarly, taking 1.2 as the initial approximation to x_3 and carrying out 4 iterations yields the successive approximations 1.48161, 1.39093, 1.37258 and 1.37188; thus $x_3 = 1.372$ to 3 decimal places.

(iii) The equation $f'(x) = 0$ has 2 roots at $x = \pm 1$. Hence, by the result of Problem 10–1, the equation $f(x) = 0$ has *at most* 3 roots. On the other hand, $f(-2) = -20$, $f(-1) = +6$, $f(1) = -2$ and $f(2) = +24$. Hence, by the intermediate value theorem, the equation $f(x) = 0$ has at least one root between -2 and -1, at least one between -1 and 1, at least one between 1 and 2 and therefore *at least* three roots in all. It follows that $f(x) = 0$ has exactly 3 roots.

10–3. (i) Suppose $g'(x) = c$ for all x, where c is a constant. Let $f(x) = g(x) - cx$; then $f'(x) = 0$ for all x, so $f(x)$ is a constant, say b. Hence $g(x) = b + cx$ for all x.

(ii) Let $\ln y = g(t)$. By assumption, $g'(t)$ is a constant, say c. By the result of (i) with x replaced by t, there is a constant b such that $g(t) = b + ct$ for all t. Let $A = e^b$. Then for all t,

$$y(t) = \exp g(t) = \exp(b + ct) = Ae^{ct}.$$

10–4. $D_a = x/(1-x)$, $D_c = -\ln(1-x)$. Since $0 < x < 1$, the expressions given for D_a and D_c follow from the power series expansions for $(1-x)^{-1}$ and $\ln(1-x)$ respectively. Since $x > 0$, the fact that $D_b < D_c < D_a$ can be read from the coefficients in these series.

11 MATRIX ALGEBRA

11–1. (i) (a) Suppose there are scalars $\alpha_1, \alpha_2, \alpha_3$ such that

$$\alpha_1\mathbf{a}^1 + \alpha_2\mathbf{a}^2 + \alpha_3\mathbf{a}^3 = \mathbf{0}.$$

By equating components, $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Therefore $\mathbf{a}^1$, $\mathbf{a}^2$, $\mathbf{a}^3$ are linearly independent.

(b) Let $\mathbf{x}$ be any n–vector; denote its components by x_1, x_2, x_3. Then

$$\mathbf{x} = x_1\mathbf{a}^1 + x_2\mathbf{a}^2 + x_3\mathbf{a}^3.$$

(ii) (a) Suppose there are scalars $\alpha_1, \alpha_2, \alpha_3$ such that

$$\alpha_1\mathbf{a}^1 + \alpha_2\mathbf{a}^2 + \alpha_3\mathbf{a}^3 = \mathbf{0}.$$

By equating components,

$$\alpha_1 + \alpha_2 + \alpha_3 = 0, \quad \alpha_2 + \alpha_3 = 0, \quad \alpha_3 = 0.$$

By back-substitution, $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Therefore $\mathbf{a}^1$, $\mathbf{a}^2$, $\mathbf{a}^3$ are linearly independent.

(b) Let $\mathbf{x}$ be any n–vector; denote its components by x_1, x_2, x_3. We wish to find scalars $\lambda_1, \lambda_2, \lambda_3$ such that

$$\mathbf{x} = \lambda_1\mathbf{a}^1 + \lambda_2\mathbf{a}^2 + \lambda_3\mathbf{a}^3.$$

By equating components

$$x_1 = \lambda_1 + \lambda_2 + \lambda_3, \quad x_2 = \lambda_2 + \lambda_3, \quad x_3 = \lambda_3.$$

We may now solve for $\lambda_1, \lambda_2, \lambda_3$ by back-substitution, getting $\lambda_1 = x_1 - x_2$, $\lambda_2 = x_2 - x_3$, $\lambda_3 = x_3$.

11–2. (i) Let $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then $\mathbf{A}^2 = \begin{bmatrix} a^2 + bc & (a+d)b \\ (a+d)c & d^2 + bc \end{bmatrix}$.

To make the off-diagonal entries of $\mathbf{A}^2$ equal to 0, we need either $a + d = 0$ or $b = c = 0$. In the latter case, $\mathbf{A}^2$ cannot be $-\mathbf{I}$ since a and d are real numbers. Thus $\mathbf{A}^2 = -\mathbf{I}$ requires $a + d = 0$ and $a^2 + bc = -1$. The required matrices are therefore those of the form

$$\begin{bmatrix} a & b \\ c & -a \end{bmatrix}, \text{ where } bc \le -1 \text{ and } a = \pm\sqrt{-bc-1}.$$

Examples are $\begin{bmatrix} 2 & -1 \\ 5 & -2 \end{bmatrix}$ and $\begin{bmatrix} 0 & -2 \\ \frac{1}{2} & 0 \end{bmatrix}$.

(ii) Let $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then $\mathbf{A}^2$ is as in the first line of (i).

To make the diagonal entries of $\mathbf{A}^2$ equal to a^2 and d^2, we need $bc = 0$; thus at least one of b and c must be 0. If $b = 0$, then making the off-diagonal entries of $\mathbf{A}^2$ equal to b^2 and c^2 requires that $(a + d)c = c^2$, which means that c is either 0 or $a + d$. Similarly, if $c = 0$ then the property in question requires that b is either 0 or $a + d$. Hence the required matrices are those of the form

$$\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}, \quad \begin{bmatrix} a & 0 \\ a+d & d \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} a & a+d \\ 0 & d \end{bmatrix},$$

where a and d are any real numbers.

11–3. (i) By back-substitution, $x_3 = \frac{1}{2}y_3$, $x_2 = \frac{1}{2}y_3 - y_2$, $x_1 = \frac{1}{2}(y_1 - y_2 - y_3)$.

(ii) Here we can use **forward substitution**: solve for x_1, then for x_2 and finally for x_3. We have $x_1 = y_1/4,\ x_2 = (3y_1 - 4y_2)/8,\ x_3 = (5y_1 - 4y_2 - 8y_3)/8$.

(iii) $x_1 = -y_1/3,\ x_2 = y_2/2,\ x_3 = y_1$.

In (iii), the only arithmetical operation required is division.

11–4. (i)

$$\begin{aligned}
y_1 &= x_1 - a_{11}x_1 - a_{12}x_2 - \ldots - a_{1n}x_n \\
y_2 &= x_2 - a_{21}x_1 - a_{22}x_2 - \ldots - a_{2n}x_n \\
&\ldots \\
y_n &= x_n - a_{n1}x_1 - a_{n2}x_2 - \ldots - a_{nn}x_n
\end{aligned}$$

(ii) $\mathbf{y} = \mathbf{x} - \mathbf{A}\mathbf{x} = (\mathbf{I} - \mathbf{A})\mathbf{x}$, so $\mathbf{B} = \mathbf{I} - \mathbf{A}$.

12 SYSTEMS OF LINEAR EQUATIONS

12–1. (i) Two Gaussian elimination steps give

$$\left[\begin{array}{cccc|c} 2 & 1 & 5 & 2 & t \\ 0 & -1/2 & 3/2 & 2 & t/2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

The system has been reduced to one in which the coefficient matrix is a Type 4 echelon matrix. The third equation has been reduced to $0 = 0$ and hence may be ignored. The system has a solution for all values of t.

The rank of $\mathbf{A}$ is the number of nonzero rows in the echelon form and hence is 2. Similarly, the rank of $[\mathbf{A}\ \ \mathbf{b}]$ is 2: notice that $\mathbf{A}$ and $[\mathbf{A}\ \ \mathbf{b}]$ have the same rank.

(ii) Three Gaussian elimination steps give

$$\left[\begin{array}{ccccc|c} 1 & 6 & -7 & 3 & 5 & 1 \\ 0 & 3 & 1 & 1 & 4 & 1 \\ 0 & 0 & 0 & 2 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & t-7 \end{array}\right].$$

The system has been reduced to one in which the coefficient matrix is a Type 4 echelon matrix. The third equation has been reduced to $0 = t - 7$, so it may be ignored if $t = 7$ and is absurd if $t \neq 7$. The system has a solution if $t = 7$ and no solution otherwise.

The rank of $\mathbf{A}$ is the number of nonzero rows in the echelon form and hence is 3. When $t = 7$, the rank of $[\mathbf{A}\ \ \mathbf{b}]$ is also 3; otherwise the rank of $[\mathbf{A}\ \ \mathbf{b}]$ is 4. Notice that the rank of $\mathbf{A}$ is equal to the rank of $[\mathbf{A}\ \ \mathbf{b}]$ if the system has a solution, and less than the rank of $[\mathbf{A}\ \ \mathbf{b}]$ if it does not.

12–2. Suppose $\mathbf{A}_1$ is $k\times k$ and $\mathbf{A}_2$ is $\ell\times\ell$. Then

$$\begin{bmatrix} \mathbf{A}_1 & \mathbf{O} \\ \mathbf{O} & \mathbf{A}_2 \end{bmatrix}\begin{bmatrix} \mathbf{A}_1^{-1} & \mathbf{O} \\ \mathbf{O} & \mathbf{A}_2^{-1} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_k & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_\ell \end{bmatrix} = \mathbf{I}.$$

By Fact 4, $\mathbf{A}$ is invertible with inverse as stated.

For the second part, denote the two matrices by $\mathbf{B}$ and $\mathbf{C}$. We may write $\mathbf{B} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{O} \\ \mathbf{O} & \mathbf{B}_2 \end{bmatrix}$ where $\mathbf{B}_1$ and $\mathbf{B}_2$ are 2×2 matrices. Hence $\mathbf{B}^{-1} = \begin{bmatrix} \mathbf{B}_1^{-1} & \mathbf{O} \\ \mathbf{O} & \mathbf{B}_2^{-1} \end{bmatrix}$, where $\mathbf{B}_1^{-1}$ and $\mathbf{B}_2^{-1}$ are calculated by the inversion formula. Also,

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{0} \\ \mathbf{0} & 2 \end{bmatrix}$$

where $\mathbf{C}_1$ is a 3×3 matrix; hence

$$\mathbf{C}^{-1} = \begin{bmatrix} \mathbf{C}_1^{-1} & \mathbf{0} \\ \mathbf{0} & \frac{1}{2} \end{bmatrix},$$

where $\mathbf{C}_1^{-1}$ is obtained from $\mathbf{C}$ by Gauss–Jordan. Performing the calculations,

$$\mathbf{B}^{-1} = \begin{bmatrix} 3/19 & 2/19 & 0 & 0 \\ -2/19 & 5/19 & 0 & 0 \\ 0 & 0 & 1/3 & -2/3 \\ 0 & 0 & 1/6 & 1/6 \end{bmatrix}, \quad \mathbf{C}^{-1} = \begin{bmatrix} -1 & 9 & -4 & 0 \\ 2 & -15 & 7 & 0 \\ 2 & -17 & 8 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{bmatrix}.$$

12–3. (i) Since $\mathbf{B}$ is invertible, the only 2–vector $\mathbf{z}$ with the given property is $\mathbf{B}^{-1}\begin{bmatrix} a_3 \\ b_3 \end{bmatrix}$.

(ii) If we assign the arbitrary value $-\lambda$ to x_3, the given system can be written as

$$\mathbf{B}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \lambda a_3 \\ \lambda b_3 \end{bmatrix}.$$

This holds if and only if

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{B}^{-1}\begin{bmatrix} \lambda a_3 \\ \lambda b_3 \end{bmatrix} = \lambda\mathbf{B}^{-1}\begin{bmatrix} a_3 \\ b_3 \end{bmatrix} = \lambda\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}.$$

Hence the given vector equation holds if and only if

$$\mathbf{x} = \lambda\begin{bmatrix} z_1 \\ z_2 \\ -1 \end{bmatrix} \quad \text{for some scalar } \lambda. \tag{†}$$

(iii) $\mathbf{A}$ is singular if and only if there is a non-zero $\mathbf{x}$ such that $\mathbf{Ax} = \mathbf{0}$. From (ii), the first two equations of this system hold if and only if (†) is true; and, for $\mathbf{x}$ to be non-zero, λ must be non-zero. It follows that $\mathbf{A}$ is singular if and only if this $\mathbf{x}$ satisfies the third equation of the system. Therefore, $\mathbf{A}$ is singular if and only if

$$c_1\lambda z_1 + c_2\lambda z_2 - c_3\lambda = 0 \quad \text{for some non-zero } \lambda,$$

which happens if and only if $c_1z_1 + c_2z_2 = c_3$.

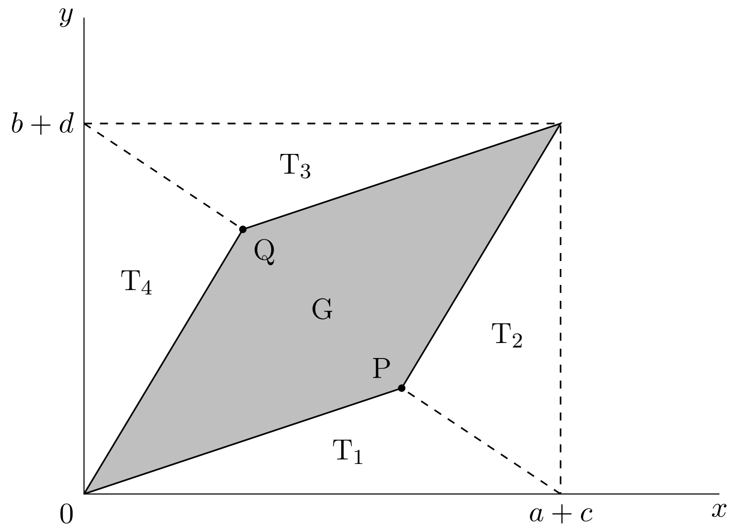
(iv) If $\mathbf{A}$ is singular then c_3 must be as in (iii). So if we replace the $(3,3)$ entry of $\mathbf{A}$ by any number other than c_3, then $\mathbf{A}$ becomes invertible.

12–4. From Problem 11–4, $\mathbf{Bx} = \mathbf{y}$, where $\mathbf{B} = \mathbf{I} - \mathbf{A}$. If there is to be a unique $\mathbf{x}$ for any $\mathbf{y}$, then $\mathbf{B}$ must be invertible and $\mathbf{x} = \mathbf{B}^{-1}\mathbf{y}$.

In addition, it is given that $\mathbf{y}$ has non-negative components. To ensure that $\mathbf{x}$ has non-negative components for every such $\mathbf{y}$ it is necessary that all entries of $\mathbf{B}^{-1}$ be non-negative. For suppose that $\mathbf{B}^{-1}$ had some negative entry, say the $(2,3)$ entry. By taking $\mathbf{y}$ to be the vector with third component 1 and zeros elsewhere, we see that the second component of $\mathbf{x}$ is negative.

13 DETERMINANTS AND QUADRATIC FORMS

13–1. Let P be the point (a,b), Q the point (c,d).



In the diagram, the area of each of the triangles T_1 and T_3 is $\frac{1}{2}(a+c)b$ by the half-base-times-height formula. Similarly, the area of each of the triangles T_2 and T_4 is $\frac{1}{2}(b+d)c$. Hence the area of G is

$$\begin{aligned}(a+c)(b+d) - (a+c)b - (b+d)c &= (a+c)d - (b+d)c \\ &= ad - bc.\end{aligned}$$

If we exchange the positions of P and Q, the area of G becomes $cb - da$, which equals $-(ad-bc)$. Thus the general formula for the area of G is $|ad - bc|$.

13-2. (i) $\det \mathbf{A} = c_1 \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} - c_2 \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} + c_3 \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}.$

Let $\mathbf{C}$ be the matrix obtained from $\mathbf{A}$ by replacing its $(3,3)$ entry by $c_3 + \delta$. Replacing c_3 by $c_3 + \delta$ in the expression just given for $\det \mathbf{A}$, we see that $\det \mathbf{C} = \det \mathbf{A} + \delta \det \mathbf{B}$. By our assumptions about $\mathbf{A}$ and $\mathbf{B}$, $\det \mathbf{A} = 0$ and $\det \mathbf{B} \neq 0$. Hence $\det \mathbf{C} \neq 0$ if $\delta \neq 0$.

(ii) Let $\mathbf{A}$ be a singular 2×2 matrix and let $\mathbf{C}$ be the matrix obtained by adding x to each of its diagonal entries. Since $\det \mathbf{A} = 0$, $\det \mathbf{C} = tx + x^2$, where t is the sum of the diagonal entries of $\mathbf{A}$. If $t = 0$, $\det \mathbf{C} > 0$ for any non-zero x; if $t \neq 0$, $\det \mathbf{C} > 0$ whenever x has the same sign as t; in each case, $|x|$ can be as small as we please.

Now suppose we have a singular 3×3 matrix $\mathbf{A}$. As in (i), we denote the 2×2 leading principal submatrix of $\mathbf{A}$ by $\mathbf{B}$. If $\mathbf{B}$ is invertible then, as in (i), we can make $\mathbf{A}$ invertible by an arbitrarily small change to its $(3, 3)$ entry. If $\mathbf{B}$ is singular we can apply the proposition in the 2×2 case, making $\mathbf{B}$ invertible by arbitrarily small changes to its diagonal entries; we can then use (i) as before. This proves the proposition for 3×3 matrices.

For the 4×4 case, if necessary we apply the proposition for the 3×3 case to ensure that the leading principal submatrix of order 3 is nonsingular. Then, by a similar argument to (i), the 4×4 matrix can be made invertible by an arbitrarily small change to its $(4, 4)$ entry. In the same way, the proposition for the 4×4 case can then be used to prove it for the 5×5 case, and so on.

(iii) It is easy to see from the expansion formulae that small changes in the entries of a matrix cause only small changes in the determinant. Therefore, arbitrarily small changes in diagonal entries are not enough to transform a matrix with nonzero determinant into a singular matrix.

13–3. The cost of producing each unit of gross output of good j is

$$c_j + p_1 a_{1j} + p_2 a_{2j} + \ldots + p_n a_{nj}.$$

If all industries exactly break even, then this expression must be equal to p_j for all j. Hence we may write the break-even condition for all industries as the single vector equation $\mathbf{c} + \mathbf{A}^{\mathrm{T}}\mathbf{p} = \mathbf{p}$, or $(\mathbf{I} - \mathbf{A}^{\mathrm{T}})\mathbf{p} = \mathbf{c}$.

Now observe that $\mathbf{I} - \mathbf{A}^{\mathrm{T}} = (\mathbf{I} - \mathbf{A})^{\mathrm{T}}$. Denoting $\mathbf{I} - \mathbf{A}$ by $\mathbf{B}$ as in Problems 11–4 and 12–4, we may write the break-even condition as $\mathbf{B}^{\mathrm{T}}\mathbf{p} = \mathbf{c}$. If there is to be a unique $\mathbf{p}$ for any $\mathbf{c}$, then $\mathbf{B}^{\mathrm{T}}$ must be invertible and $\mathbf{p} = (\mathbf{B}^{\mathrm{T}})^{-1}\mathbf{c}$. In addition, it is given that $\mathbf{c}$ has non-negative components. To ensure that $\mathbf{p}$ has non-negative components for every such $\mathbf{c}$, it is necessary that $(\mathbf{B}^{\mathrm{T}})^{-1}$ has non-negative entries. This follows from an argument similar to that given in Problem 12–4.

Finally, observe that $\mathbf{B}^{\mathrm{T}}$ is invertible if and only if $\mathbf{B}$ is invertible. If $\mathbf{B}$ is invertible then $(\mathbf{B}^{\mathrm{T}})^{-1} = (\mathbf{B}^{-1})^{\mathrm{T}}$; in particular, all entries of $(\mathbf{B}^{\mathrm{T}})^{-1}$ are non-negative if and only if all entries of $\mathbf{B}^{-1}$ are non-negative. Thus $\mathbf{A}$ has the properties required here if and only if it has the properties required in Problem 12–4.

13–4. (i) The ith component of $\mathbf{y} - \mathbf{X}\mathbf{b}$ is $y_i - b_1 x_{1i} - b_2 x_{2i}$. The result follows.

(ii) $\mathbf{y} - \mathbf{X}\mathbf{b} = \mathbf{y} - \mathbf{X}\mathbf{b}^* + \mathbf{X}(\mathbf{b}^* - \mathbf{b}) = \mathbf{p} + \mathbf{q}$ where $\mathbf{p} = \mathbf{X}(\mathbf{b}^* - \mathbf{b})$, $\mathbf{q} = \mathbf{y} - \mathbf{X}\mathbf{b}^*$ and $\mathbf{p}^{\mathrm{T}}\mathbf{q} = 0$. The result then follows from that of Exercise 13.3.1.

(iii) $(*)$ can be written as $\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)\mathbf{b}^* = \mathbf{X}^{\mathrm{T}}\mathbf{y}$. Since $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ is invertible, there is only one vector $\mathbf{b}^*$ which satisfies $(*)$; this is given by $\mathbf{b}^* = \left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y}$.

(iv) The answer to (ii) expresses $Q(\mathbf{b})$ as the sum of two terms, only the second of which depends on $\mathbf{b}$. Since $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ is positive definite, this second term is positive if $\mathbf{b} \neq \mathbf{b}^*$, zero if $\mathbf{b} = \mathbf{b}^*$. Hence $Q(\mathbf{b})$ is minimised when $\mathbf{b} = \mathbf{b}^*$.

14 FUNCTIONS OF SEVERAL VARIABLES

14–1. (i) $\partial z/\partial x = y$ and $\partial z/\partial y = x$, so the equation of the tangent plane is

$$z = 12 + 3(x - 4) + 4(y - 3).$$

When $x = 4 + h$ and $y = 3 + k$, the value of z given by the tangent plane is $12 + 3h + 4k$.

The small increments formula gives $\Delta z \approx y\Delta x + x\Delta y = 3h + 4k$, so the value of z given by that formula is also $12 + 3h + 4k$.

(ii) When $x = 4 + h$ and $y = 3 + k$,

$$f(x, y) = (4 + h)(3 + k) = 12 + 3h + 4k + hk.$$

Therefore, the error when the surface $z = xy$ near the point $(4, 3, 12)$ is approximated by the tangent plane at that point is hk.

(a) The error as a percentage of the true value is

$$\frac{0.01 \times 0.01 \times 100}{4.01 \times 3.01} = 0.000828\%.$$

(b) The error as a percentage of the true value is $\dfrac{1\times1\times100}{5\times4} = 5\%$.

14–2.

$$(1)\ \frac{\partial H}{\partial i} = \frac{\partial f}{\partial Y}\frac{\partial g}{\partial i} + \frac{\partial f}{\partial i}, \quad (2)\ \frac{\partial H}{\partial u} = \frac{\partial f}{\partial Y}\frac{\partial g}{\partial u}.$$

In the particular case given, $H(i, u) = ABe^{-(a+b)i}u^c$. Then
$\partial H/\partial i = -(a + b)H(i, u) = -(a + b)M$ and $\partial H/\partial u = (c/u)H(i, u) = cM/u$.
Also $\partial f/\partial Y = f(y, i)/Y = M/Y$, $\partial f/\partial i = -af(y, i) = -aM$,
$\partial g/\partial i = -bg(i, u) = -bY$ and $\partial g/\partial u = (c/u)g(i, u) = cY/u$. Hence

$$\text{RHS}(1) = \frac{M}{Y}\times(-bY) - aM = -(a+b)M = \text{LHS}(1), \quad \text{RHS}(2) = \frac{M}{Y}\times\frac{cY}{u} = \frac{cM}{u} = \text{LHS}(2).$$

14–3. (i) In this special case,

$$\frac{\partial F}{\partial K} = A(\alpha K^{\alpha-1})L^{\beta}e^{\mu t} = \frac{\alpha}{K}F(K, L, t) = \frac{\alpha Q}{K}.$$

Similarly, $\dfrac{\partial F}{\partial L} = \dfrac{\beta Q}{L}$ and $\dfrac{\partial F}{\partial t} = \mu Q$. By assumption, $\dfrac{dK}{dt} = mK$ and $\dfrac{dL}{dt} = nL$.
Hence by equation (14.10) in the text,

$$\frac{dQ}{dt} = \frac{\alpha Q}{K}(mK) + \frac{\beta Q}{L}(nL) + \mu Q = (\alpha m + \beta n + \mu)Q,$$

so the rate of growth of output is $\alpha m + \beta n + \mu$.

(ii) Using (14.10) and the assumptions about the rates of growth of K and N,

$$\frac{dQ}{dt} = \frac{\partial H}{\partial K}\frac{dK}{dt}e^{\mu t} + \frac{\partial H}{\partial L}\frac{dL}{dt}e^{\mu t} + \mu H(K, L)e^{\mu t} = \left[nK\frac{\partial H}{\partial K} + nL\frac{\partial H}{\partial L}\right]e^{\mu t} + \mu Q.$$

By Euler's theorem, the expression in square brackets is equal to $nrH(K, L)$, Hence

$$\frac{dQ}{dt} = nrH(K, L)e^{\mu t} + \mu Q = (nr + \mu)Q.$$

The rate of growth of output is $nr + \mu$.

14–4. (i) Since $F(K, L)$ is homogeneous of degree 1, $F(K, L) = LF(K/L, 1)$. Obviously $F(K/L, 1)$ depends only on K/L: call it $f(K/L)$. Letting $k = K/L$, we have $Q = F(K, L) = Lf(k)$.

(ii) Using the fact that $F(K, L) = Lf(k)$,

$$\frac{\partial F}{\partial K} = Lf'(k)\frac{\partial k}{\partial K} = Lf'(k)\times\frac{1}{L} = f'(k),$$

$$\frac{\partial F}{\partial L} = f(k) + Lf'(k)\frac{\partial k}{\partial L} = f(k) + Lf'(k)\times\left(-\frac{K}{L^2}\right) = f(k) - kf'(k).$$

(iii) Multiplying K and L by λ leaves k unchanged; hence by (i) and (ii), the average and marginal products of labour and capital are left unchanged.

15 IMPLICIT RELATIONS

15–1. Let $c > 0$. Since $V(x, y) = \ln U(x, y)$, any point on the indifference curve $U(x, y) = c$ satisfies $V(x, y) = \ln c$. Conversely, since $U(x, y) = \exp V(x, y)$, any point satisfying $V(x, y) = \ln c$ must lie on the indifference curve $U(x, y) = c$. The curve $U(x, y) = c$ is therefore identical to the curve $V(x, y) = \ln c$. Similarly, the curve $V(x, y) = k$ is the same as the curve $U(x, y) = e^k$. Thus V and U give rise to the same indifference curve diagrams. Since the natural logarithm is a monotonic increasing function, the ordering of the curves is also the same.

The indifference curve diagram for W is the same as that for V except that the lines $x = a$ and $y = b$ correspond to the axes.

15–2. (i) Using the results of Exercise 14.3.2(a),

$$\frac{dL}{dK} = -\frac{\delta A^\gamma (Q/K)^{1-\gamma}}{(1-\delta)A^\gamma (Q/L)^{1-\gamma}} = -\frac{\delta}{1-\delta}\left[\frac{L}{K}\right]^{1-\gamma}.$$

Since $0 < \delta < 1$ the isoquants are negatively sloped. Now consider moving along an isoquant in the direction of $K\uparrow$ and $L\downarrow$. Then $L/K\downarrow$; since $\gamma < 1$, it follows that $|dL/dK|$ decreases. Hence the isoquants are convex.

(ii) Let $\gamma < 0$. We may write the the equation of the isoquant $Q = \overline{Q}$ in the form

$$\delta K^\gamma + (1-\delta)L^\gamma = (\overline{Q}/A)^\gamma. \qquad (*)$$

As $L \to \infty$, $L^\gamma \to 0$ (since $\gamma < 0$), so $\delta K^\gamma \to (\overline{Q}/A)^\gamma$; therefore, $K \to d_1\overline{Q}/A$, where $d_1 = \delta^{-1/\gamma}$. It follows that the line $K = d_1\overline{Q}/A$ is an asymptote. Similarly, setting $d_2 = (1-\delta)^{-1/\gamma}$, we see that the line $L = d_2\overline{Q}/A$ is also an asymptote.

(iii) Let $0 < \gamma < 1$. The equation of the isoquant $Q = \overline{Q}$ is still $(*)$. Since we now have $\gamma > 0$, $L^\gamma = 0$ when $L = 0$; thus the isoquant meets the K–axis where $\delta K^\gamma = (\overline{Q}/A)^\gamma$. Hence the isoquant $Q = \overline{Q}$ meets the K–axis at the point $(d_1\overline{Q}/A,\ 0)$, where d_1 is defined as in (ii). Similarly, the isoquant $Q = \overline{Q}$ meets the L–axis at the point $(0,\ d_2\overline{Q}/A)$, where d_2 is defined as in (ii).

From the formula for dL/dK, the slope of the isoquant is 0 at the first point and $-\infty$ at the second. The isoquant therefore meets the two axes tangentially.

(iv) $\ln\dfrac{Q}{A} = \dfrac{m(\gamma)}{\gamma}$, where

$$m(\gamma) = \ln[\delta K^\gamma + (1-\delta)L^\gamma].$$

Since $m(0) = 0$, we infer from l'Hôpital's rule (or the definition of a derivative) that $\ln(Q/A) \to m'(0)$ as $\gamma \to 0$. Using the fact that $\dfrac{d}{dx}(a^x) = a^x \ln a$, we have

$$m'(\gamma) = \frac{\delta K^\gamma \ln K + (1-\delta)L^\gamma \ln L}{\delta K^\gamma + (1-\delta)L^\gamma}.$$

If $\gamma = 0$, the numerator of this expression is $\delta \ln K + (1-\delta)\ln L$ and the denominator is 1. Hence

$$\lim_{\gamma\to 0} \ln(Q/A) = m'(0) = \delta \ln K + (1-\delta)\ln L = \ln\big(K^\delta L^{1-\delta}\big)$$

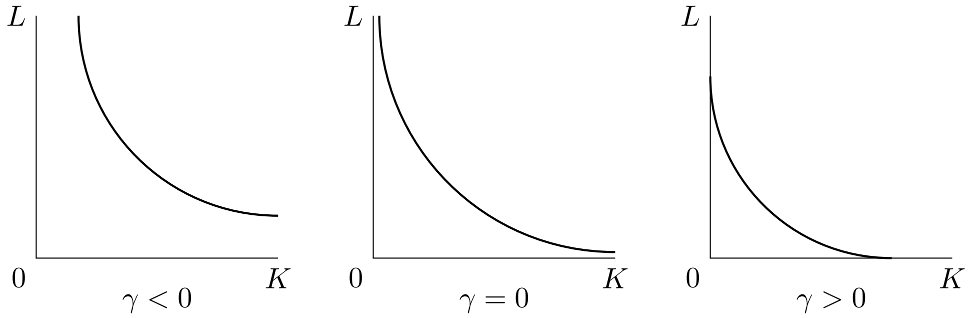
and

$$\lim_{\gamma\to 0} Q = AK^\delta L^{1-\delta}.$$

Notice that this is a Cobb–Douglas production function, with isoquants asymptotic to the axes.

(iv) Results (i)–(iii) on the shape of the isoquants remain unchanged. The limiting form (v) of the production function as $\gamma \to 0$ is the Cobb–Douglas $AK^\alpha L^\beta$, where $\alpha = \nu\delta$ and $\beta = \nu(1-\delta)$.

The diagram shows a typical isoquant in each of the three cases $\gamma < 0$, $\gamma = 0$ (Cobb–Douglas) and $0 < \gamma < 1$



15–3. (i) $dr/dT = \left(f(T)f''(T) - [f'(T)]^2\right) / [f(T)]^2$. So, by the inverse function rule,

$$dT/dr = [f(T)]^2 \,/\left(f(T)f''(T) - [f'(T)]^2\right) .$$

So long as the optimal value of T is in the region for which f is concave, then $dT/dr < 0$.

(ii) Define the function $F(r,T) = f'(T)(e^{rT}-1) - f(T)re^{rT}$. The Faustmann rule can be written in the form $F(r,T) = 0$, so by implicit differentiation

$$\frac{dT}{dr} = -\frac{\partial F}{\partial r} \Big/ \frac{\partial F}{\partial T} .$$

By definition of the function F,

$$\frac{\partial F}{\partial r} = [Tf'(T) - (1+rT)f(T)]e^{rT}, \quad \frac{\partial F}{\partial T} = f''(T)(e^{rT}-1) - r^2 f(T)e^{rT}.$$

It follows that

$$\frac{dT}{dr} = \frac{[Tf'(T)/f(T)] - 1 - rT}{r^2 + [-f''(T)/f(T)](1-e^{-rT})}. \tag{$\dagger$}$$

So long as the optimal value of T is in the region for which f is concave, the denominator on the right-hand side of (†) is positive. By the Faustmann rule, the numerator on the right-hand side of (†) can be written as $\dfrac{1+rT-e^{rT}}{e^{rT}-1}$, which is easily seen to be negative (use the series for e^x). Hence $dT/dr < 0$.

Part (ii) of this problem could be solved by the method given in the text for comparative statics of optima, which directly exploits the second order condition at the optimum. The above method, which starts by transforming the rule into a form not involving quotients, is simpler in this case.

15–4. (i) $F(Y,r) = Y - f(Y) - \phi(r)$.

(ii) The Jacobian matrix $\mathbf{J}$ of (Φ, Ψ) with respect to (Y, r) is

$$\begin{bmatrix} \partial\Phi/\partial Y & \partial\Phi/\partial r \\ \partial\Psi/\partial Y & \partial\Psi/\partial r \end{bmatrix} = \begin{bmatrix} 1-f'(Y) & -\phi'(r) \\ \partial L/\partial Y & \partial L/\partial r \end{bmatrix}.$$

Notice that $\det \mathbf{J} = (1-f'(Y))\dfrac{\partial L}{\partial r} + \phi'(r)\dfrac{\partial L}{\partial Y} < 0$, so $\mathbf{J}$ is invertible.

(iii) The slope of the IS relation in the (Y, r) plane is

$$-\frac{\partial \Phi}{\partial Y} \Big/ \frac{\partial \Phi}{\partial r} = \frac{1 - f'(Y)}{\phi'(r)} < 0.$$

The slope of the LM relation in the (Y, r) plane is

$$-\frac{\partial \Psi}{\partial Y} \Big/ \frac{\partial \Psi}{\partial r} = -\frac{\partial L}{\partial Y} \Big/ \frac{\partial L}{\partial r} > 0.$$

A typical diagram of the IS–LM model in the non-negative quadrant of the (Y, r) plane shows the graphs of the two relations sloping as we have just indicated and intersecting at a unique point.

Since $\mathbf{J}$ is invertible at the given equilibrium, there is a unique local solution for Y and r in terms of G and M which may be differentiated as follows:

$$\begin{bmatrix} \partial Y/\partial G \\ \partial r/\partial G \end{bmatrix} = -\mathbf{J}^{-1} \begin{bmatrix} \partial \Phi/\partial G \\ \partial \Psi/\partial G \end{bmatrix}, \quad \begin{bmatrix} \partial Y/\partial M \\ \partial r/\partial M \end{bmatrix} = -\mathbf{J}^{-1} \begin{bmatrix} \partial \Phi/\partial M \\ \partial \Psi/\partial M \end{bmatrix}.$$

Since $\begin{bmatrix} \partial \Phi/\partial G & \partial \Phi/\partial M \\ \partial \Psi/\partial G & \partial \Psi/\partial M \end{bmatrix} = -\mathbf{I}$,

$$\begin{bmatrix} \partial Y/\partial G & \partial Y/\partial M \\ \partial r/\partial G & \partial r/\partial M \end{bmatrix} = \mathbf{J}^{-1} = \frac{1}{\det \mathbf{J}} \begin{bmatrix} \partial L/\partial r & \phi'(r) \\ -\partial L/\partial Y & 1 - f'(Y) \end{bmatrix}.$$

Since $\det \mathbf{J} < 0$, the partial derivatives $\dfrac{\partial Y}{\partial G}$, $\dfrac{\partial Y}{\partial M}$ and $\dfrac{\partial r}{\partial G}$ are all positive, while $\dfrac{\partial r}{\partial M} < 0$.

(iv) Using the expressions for $\det \mathbf{J}$ and $\dfrac{\partial Y}{\partial G}$ given in the answers to (i) and (iii) respectively, we see that

$$\frac{\partial Y}{\partial G} = \frac{1}{1 - f'(Y) - s\phi'(r)}, \quad \text{where } s = -\frac{\partial L}{\partial Y} \Big/ \frac{\partial L}{\partial r}.$$

As we saw in the answer to (ii), s is positive, and is in fact the slope of the LM relation. If s is small, or if $\phi'(r) \approx 0$, then $\partial Y/\partial G \approx [1 - f'(Y)]^{-1}$: this is the expression for dY/dI in Exercise 15.2.2, and is known as the Keynesian multiplier. Notice that in this case $\partial Y/\partial G > 1$. In the general case, where $\phi'(r) < 0 < s$, $\partial Y/\partial G$ is less than the Keynesian multiplier and may be less than 1.

16 OPTIMISATION WITH SEVERAL VARIABLES

16–1. The contours $f(x, y) = k$ where $k = 0, 1, 2, 3, 4, 5$ are respectively the origin and circles with centre the origin and radius 1, $\sqrt{2}$, $\sqrt{3}$, 2, $\sqrt{5}$.

The contour $g(x, y) = k$ can be expressed as $x^2 + y^2 = 2 + k$. The contours are therefore also circles with centre the origin and, as k increases from -2, the same contour diagram is obtained as for $f(x, y)$, except that the contour for $g(x, y) = k$ is $f(x, y) = k + 2$.

(i) Provided $k \geq 0$, the contour $(x^2 + y^2 - 2)^2 = k$ can be expressed as $x^2 + y^2 - 2 = 2 \pm \sqrt{k}$. So, for $0 \leq k \leq 4$, each contour has two branches consisting of circles with centre the origin and radii $(2 \pm \sqrt{k})^{1/2}$. Note the extreme cases: when $k = 0$, the two circles coincide; when $k = 4$, one of the circles reduces to the origin. For $k > 4$, each contour has one branch consisting of the circle with centre the origin and radius $(2 + \sqrt{k})^{1/2}$. The global minimum of $(x^2 + y^2 - 2)^2$ is 0 which occurs at each point of the circle $x^2 + y^2 = 2$.

(ii) The contour $(x^2+y^2-2)^3 = k$ can be expressed as $x^2+y^2 = 2+k^{1/3}$ where, for $k<0$, $k^{1/3}$ is interpreted as $-|k|^{1/3}$. For $k=-8$, the contour is the origin. For $k>-8$, the contour is a circle with centre the origin. The global minimum of $(x^2+y^2-2)^3$ is -8 which occurs at the origin.

In case (ii), unlike case (i), the global minimum occurs at the same point as the global minimum of x^2+y^2-2. The reason for the difference is that $H(w) = w^3$ is a strictly increasing function, whereas $H(w) = w^2$ is not. This difference is also illustrated by the ordering of the contours.

The contour $(x^2+y^2-2)^{-1} = k$ can be expressed as $x^2+y^2 = 2+k^{-1}$. For $k>0$, the contours are circles with centre the origin and radius $\sqrt{(2k+1)/k}$. As k increases the radius decreases, approaching $\sqrt{2}$ as $k\to\infty$. Now consider the case where $k\leq 0$. For $-\frac{1}{2}<k\leq 0$, the contour is undefined; for $k=-\frac{1}{2}$, the contour is the origin; and for $k<-\frac{1}{2}$, the contour is a circle with centre the origin. As k decreases from $-\frac{1}{2}$ the radius increases, approaching $\sqrt{2}$ as $k\to-\infty$.

16–2. The firm's profit is

$$\Pi(K,L) = pF(K,L) - rK - wL, \quad \text{where } F(K,L) = AK^\alpha L^\beta.$$

Therefore, $D\Pi(K,L) = pDF(K,L) - \begin{bmatrix} r \\ w \end{bmatrix}$ and $D^2\Pi(K,L) = pD^2F(K,L)$. It follows that Π is concave if and only if F is concave; as shown in the answer to Exercise 16.2.6, this is so if and only if $\alpha+\beta\leq 1$.

From now on, assume that $\alpha+\beta\leq 1$. The firm's problem is to maximise $\Pi(K,L)$ subject to $K\geq 0$, $L\geq 0$. Since $\Pi(K,L)$, a solution to the the first order conditions, *if it exists*, will give a global maximum.

The first-order conditions for a solution with $K>0$ and $L>0$ are

$$p\alpha AK^{\alpha-1}L^\beta = r, \quad p\beta AK^\alpha L^{\beta-1} = w.$$

Taking natural logarithms and rearranging, we may write this pair of equations as

$$\begin{bmatrix} \alpha-1 & \beta \\ \alpha & \beta-1 \end{bmatrix} \begin{bmatrix} \ln K \\ \ln L \end{bmatrix} = \begin{bmatrix} \ln(r/\alpha) - \ln(pA) \\ \ln(w/\beta) - \ln(pA) \end{bmatrix}. \qquad (*)$$

If $\alpha+\beta<1$, the coefficient matrix is invertible, and $(*)$ has the unique solution

$$\begin{bmatrix} \ln K \\ \ln L \end{bmatrix} = \frac{1}{1-\alpha-\beta} \begin{bmatrix} \beta-1 & -\beta \\ -\alpha & \alpha-1 \end{bmatrix} \begin{bmatrix} \ln(r/\alpha) - \ln(pA) \\ \ln(w/\beta) - \ln(pA) \end{bmatrix}.$$

Hence the profit-maximising inputs are

$$K = \left[pA\left(\frac{\alpha}{r}\right)^{1-\beta} \left(\frac{\beta}{w}\right)^{\beta} \right]^{1/(1-\alpha-\beta)}, \quad L = \left[pA\left(\frac{\alpha}{r}\right)^{\alpha} \left(\frac{\beta}{w}\right)^{1-\alpha} \right]^{1/(1-\alpha-\beta)}.$$

Setting

$$Z = \left[pA\left(\frac{\alpha}{r}\right)^{\alpha} \left(\frac{\beta}{w}\right)^{\beta} \right]^{1/(1-\alpha-\beta)},$$

it is not hard to see that, at the optimum, $rK = \alpha Z$, $wL = \beta Z$ and $pAK^\alpha L^\beta = Z$. Hence the maximal profit is $(1-\alpha-\beta)Z$.

If $\alpha + \beta = 1$, the coefficient matrix in $(*)$ is singular. Before exploring the consequences of this, consider the expression for profit: if $\alpha + \beta = 1$, multiplying the two inputs by any positive number λ also multiplies the profit by λ. This suggests three possible cases: (a) profit is positive for some $K^* > 0$, $L^* > 0$ and can be made arbitrarily large by letting $K = MK^*$ and $L = ML^*$ where M is very large; (b) profit is negative for all positive K and L, and is therefore maximised uniquely at 0 when $K = L = 0$; (c) profit is non-positive for all positive K and L, but there are positive (K, L) pairs for which profit is zero. Returning to $(*)$, with $\alpha + \beta = 1$, there are infinitely many solutions if $\alpha \ln(r/\alpha) + (1 - \alpha)\ln(w/(1 - \alpha)) = \ln(pA)$, and no solution otherwise. Setting

$$\bar{p} = \frac{1}{A}\left(\frac{r}{\alpha}\right)^{\alpha}\left(\frac{w}{1-\alpha}\right)^{1-\alpha},$$

we see that case (c) occurs if and only if $p = \bar{p}$, in which case the (K, L) pairs which maximise profit (at zero) are given by $\dfrac{rK}{\alpha} = \dfrac{wL}{1-\alpha}$. Case (a) occurs if $p > \bar{p}$, and case (b) if $p < \bar{p}$; the reasons for this will become clear when you have read Section 17.3 and done Exercise 17.3.2.

16–3 Denote the expression to be minimised by $Q(b_1, b_2)$. Then

$$\frac{\partial Q}{\partial b_1} = \sum_{i-1}^{n}(-2x_{1i})(y_i - b_1x_{1i} - b_2x_{2i}), \quad \frac{\partial Q}{\partial b_2} = \sum_{i-1}^{n}(-2x_{2i})(y_i - b_1x_{1i} - b_2x_{2i}).$$

It follows that

$$D^2Q(b_1, b_2) = \begin{bmatrix} 2\sum_{i=1}^{n} x_{1i}^2 & 2\sum_{i=1}^{n} x_{1i}x_{2i} \\ 2\sum_{i=1}^{n} x_{1i}x_{2i} & 2\sum_{i=1}^{n} x_{2i}^2 \end{bmatrix} = 2\mathbf{X}^{\mathrm{T}}\mathbf{X}.$$

Since the columns of $\mathbf{X}$ are linearly independent, $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ is positive definite. This shows that the function, $Q(b_1, b_2)$ has positive definite Hessian and is therefore convex.

Now $DQ(b_1, b_2) = \mathbf{0}$ when

$$\sum_{i=1}^{n} x_{1i}^2 + \sum_{i=1}^{n} x_{1i}x_{2i} = \sum_{i=1}^{n} x_{1i}y_i, \quad \sum_{i=1}^{n} x_{1i}x_{2i} + \sum_{i=1}^{n} x_{2i}^2 = \sum_{i=1}^{n} x_{2i}y_i.$$

This may be written as $(\mathbf{X}^{\mathrm{T}}\mathbf{X})\mathbf{b} = \mathbf{X}^{\mathrm{T}}\mathbf{y}$, where $\mathbf{y}$ is the n–vector whose ith component is y_i. Since $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ is positive definite, it is invertible. It follows that

$$\mathbf{b} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y}.$$

Since Q is convex, this gives the global minimum.

16–4. Denote the given utility function by $W(c, y)$. The first-order conditions are

$$\frac{\partial W}{\partial c} = \frac{\partial U}{\partial c} - (1 + r)V'(p) = 0, \quad \frac{\partial W}{\partial y} = \frac{\partial U}{\partial y} + (1 + r)V'(p) = 0,$$

where $p = (1 + r)(y - c)$. [p stands for 'pension'.]

The Jacobian matrix $\mathbf{J}$ of $(\partial W/\partial c,\ \partial W/\partial y)$ with respect to (c, y) is the Hessian matrix $D^2W(c, y)$; therefore

$$\mathbf{J} = \begin{bmatrix} \dfrac{\partial^2 U}{\partial c^2} + s & \dfrac{\partial^2 U}{\partial c\,\partial y} - s \\[2ex] \dfrac{\partial^2 U}{\partial c\,\partial y} - s & \dfrac{\partial^2 U}{\partial y^2} + s \end{bmatrix}, \quad \text{where } s = (1+r)^2 V''(p).$$

Assume that for a given value of r there is a unique pair of optimal values c and y which satisfy the first-order conditions. If in addition $\mathbf{J}$ is invertible at the given optimum, then c and y may be differentiated with respect to r using the implicit function theorem;

$$\begin{bmatrix} dc/dr \\ dy/dr \end{bmatrix} = \mathbf{J}^{-1} \begin{bmatrix} t \\ -t \end{bmatrix},$$

where t is the partial derivative with respect to r of $(1+r)V'((1+r)(y-c))$, considered as a function of c, y, and r. Calculating $\mathbf{J}^{-1}$ by the inversion formula for 2×2 matrices, and t by partial differentiation, we see that

$$\frac{dc}{dr} = \frac{t}{\det \mathbf{J}} \left[\frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial c\,\partial y}\right], \quad \frac{dy}{dr} = -\frac{t}{\det \mathbf{J}} \left[\frac{\partial^2 U}{\partial c^2} + \frac{\partial^2 U}{\partial c\,\partial y}\right],$$

where $t = V'(p) + pV''(p)$.

To discuss the signs of dc/dr and dy/dr, notice that $\det \mathbf{J} \geq 0$ by the second-order conditions for a maximum. These second-order conditions will be met, with $\mathbf{J}$ invertible, if at the optimum $0 < c < y$ (so that $p > 0$), U is concave and $V''(p) < 0$. From now on, assume these further conditions are met; in particular $\det \mathbf{J} > 0$. Let

$$\varepsilon = -\frac{pV''(p)}{V'(p)} > 0,$$

and let A, B denote the expressions in square brackets in the solutions just given for dc/dr and dy/dr respectively. Then dc/dr has the sign of $(1-\varepsilon)A$, and dy/dr has the sign of $(\varepsilon - 1)B$. By the concavity of U, A and B cannot both be positive. If $\dfrac{\partial^2 U}{\partial c\,\partial y} \leq 0$ both A and B will be non-positive; in this case

$$\frac{dc}{dr} \leq 0 \leq \frac{dy}{dr} \text{ if } \varepsilon \leq 1, \quad \frac{dc}{dr} \geq 0 \geq \frac{dy}{dr} \text{ if } \varepsilon \geq 1.$$

If $\dfrac{\partial^2 U}{\partial c\,\partial y} > 0$ it is possible, but not inevitable, that dc/dr and dy/dr have the same sign.

Finally,

$$\frac{ds}{dr} = \frac{dy}{dr} - \frac{dc}{dr} = -\frac{t}{\det \mathbf{J}} \left[\frac{\partial^2 U}{\partial c^2} + 2\frac{\partial^2 U}{\partial c\,\partial y} + \frac{\partial^2 U}{\partial y^2}\right].$$

Since U is concave, the term in square brackets is non-positive, regardless of the sign of $\dfrac{\partial^2 U}{\partial c\,\partial y}$. Also, we are assuming that $\det \mathbf{J} > 0$. Hence ds/dr has the same sign as t, so

$$\frac{ds}{dr} \geq 0 \text{ if } \varepsilon \leq 1, \quad \frac{ds}{dr} \leq 0 \text{ if } \varepsilon \geq 1.$$

17 PRINCIPLES OF CONSTRAINED OPTIMISATION

17–1. (i) The least-cost combination (K, L) occurs where the isoquant corresponding to the given output level q is tangential to an isocost line. The result follows from the fact that this point, and hence also the capital-labour ratio, depend only on q and the slope of the isocost lines.

(ii) If s increases (i.e. the price of capital increases relative to that of labour) we would expect the capital-labour ratio to decrease. Thus $\partial g/\partial s < 0$ and therefore $\sigma > 0$.

(iii) The first-order conditions are $r = \mu\partial F/\partial K,\ w = \mu\partial F/\partial L$. Now

$$\frac{\partial F}{\partial K} = AZ^{(1/\gamma)-1}\delta K^{\gamma-1}, \quad \frac{\partial F}{\partial L} = AZ^{(1/\gamma)-1}(1-\delta)L^{\gamma-1},$$

where $Z = \delta K^{\gamma} + (1-\delta)L^{\gamma}$. Substituting these into the first-order conditions and dividing, we obtain

$$\frac{r}{w} = \frac{\delta}{1-\delta}\left(\frac{K}{L}\right)^{\gamma-1}.$$

It follows that $g(s, q) = (s(1-\delta)/\delta)^{1/(\gamma-1)}$, so $\sigma = (1-\gamma)^{-1}$.

(iv) Proceed as in (iii). Although the expressions for $\partial F/\partial K$ and $\partial F/\partial L$ are different, division of one first-order condition by the other yields the same result as in (iii). Hence $g(s, q)$ and σ are as in (iii).

(v) In both (iii) and (iv), $\partial g/\partial q = 0$. More generally, let $F(K, L)$ be any homogeneous function of degree $\nu > 0$. Then $\partial F/\partial K$ and $\partial F/\partial L$ are homogeneous of degree $\nu - 1$ so their ratio is homogeneous of degree 0. Therefore, r/w is a function of K/L alone, so K/L depends only on s. The same argument applies to the still more general case where the production function is $G(K, L) = H(F(K, L))$, where $F(K, L)$ is homogeneous of degree $\nu > 0$ and H is a monotonic increasing transformation. [A function which is a monotonic transformation of a homogeneous function is said to be **homothetic**.]

17–2. (i) w can be interpreted as the wage rate and t as the firm's fixed cost per worker.

(ii) The Lagrangian is

$$L(h, N, \mu) = whN + tN - \mu(F(h, N) - q),$$

so the first-order conditions are

$$wN = \mu\frac{\partial F}{\partial h}, \quad wh + t = \mu\frac{\partial F}{\partial N}.$$

Now $\partial F/\partial h = bN^2(ah + bN)^{-2}$ and $\partial F/\partial N = ah^2(ah + bN)^{-2}$. Substituting these into the first-order conditions and dividing the second condition by the first gives

$$\frac{wh + t}{wN} = \frac{ah^2}{bN^2},$$

whence $\dfrac{ah}{bN} = 1 + \dfrac{t}{wh}$. But $\dfrac{ah}{bN} = \dfrac{h}{bq} - 1$ by the output constraint. Equating our two expressions for $\dfrac{ah}{bN}$ and rearranging, we see that

$$wh^2 - 2bqwh - bqt = 0.$$

Solving for h by the quadratic formula and taking the positive root gives

$$h = \left(1 + \left[1 + \frac{t}{bqw}\right]^{1/2}\right) bq.$$

N is now found by substituting the solution for h into the output constraint.

(iii) The isoquant $F(h, N) = q$ is a negatively sloped convex curve lying in the positive quadrant with asymptotes $h = bq$ and $N = aq$. The isocost curves have equations of the form $N = k/(wh + t)$ for different values of the total cost k. The economically meaningful parts of these curves lie in the non-negative quadrant, are negatively sloped, convex, have the h–axis as an asymptote and meet the N–axis at $(0, k/t)$. The answer to (ii) lies at the point of tangency of the isoquant and an isocost curve. This isocost curve corresponds to the lowest intercept on the N–axis, and hence the lowest value of k, consistent with being on the isoquant.

17–3. Two goods: $p_1x_1(p_1, p_2, m) + p_2x_2(p_1, p_2, m) = m$. Differentiating with respect to m, p_1, p_2 respectively gives the results as stated.

n goods: $\sum_{i=1}^{n} p_i x_i(p_1, \ldots, p_n, m) = m$. Differentiating with respect to m gives

$$\sum_{i=1}^{n} p_i \frac{\partial x_i}{\partial m} = 1,$$

while differentiating with respect to p_j gives

$$x_j + \sum_{i=1}^{n} p_i \frac{\partial x_i}{\partial p_j} = 0.$$

17–4. (i) Samantha's problem is to

$$\text{maximise } x^\alpha \ell^\beta \text{ subject to } px = w(T - \ell) + N.$$

Thus the Lagrangian for the problem is $u - \lambda(px + w\ell) + \lambda(wT + N)$, where $u = x^\alpha \ell^\beta$. The first-order conditions are $\alpha u/x = \lambda p$, $\beta u/\ell = \lambda w$, whence $px = (\alpha/\beta)w\ell$. This, together with the constraint, implies that

$$x = \frac{a}{p}(wT + N), \quad \ell = \frac{1-a}{w}(wT + N), \quad \text{where } a = \frac{\alpha}{\alpha + \beta}.$$

Hence $h = T - \ell = aT - (1 - a)(N/w)$. From this expression, it is clear that as N increases h decreases. As w increases, h increases or decreases according as N is positive or negative: notice in particular that h is independent of w if $N = 0$.

(ii) Samantha's problem is to

$$\text{maximise } x_1^\alpha x_2^\beta \text{ subject to } p_1x_1 + p_2x_2 = w(T - t_1x_1 - t_2x_2) + N.$$

This is similar to the problem in (i) with x, ℓ, p and w replaced by x_1, x_2, $p_1 + wt_1$ and $p_2 + wt_2$ respectively. Hence the optimal values of x_1 and x_2 are

$$x_1 = \frac{a(wT + N)}{p_1 + wt_1}, \quad x_2 = \frac{(1-a)(wT + N)}{p_2 + wt_2},$$

where $a = \alpha/(\alpha + \beta)$ as before. It follows that

$$h = aT - t_1x_1 + (1-a)T - t_2x_2 = \frac{a(p_1T - t_1N)}{p_1 + wt_1} + \frac{(1-a)(p_2T - t_2N)}{p_2 + wt_2}.$$

As in (i), $\partial h/\partial N < 0$ and $\partial h/\partial w$ may have either sign, but here the criterion for the sign of the latter derivative is much more complicated.

(iii) The model in (ii), with $p_1 = p$, $t_1 = 0$, $p_2 = 0$ and $t_1 = 1$, reduces to that in (i).

18 FURTHER TOPICS IN CONSTRAINED OPTIMISATION

18–1. (i) Since $\Phi(z_0) = \Pi(p_0, z_0)$,

$$D = [\Phi(z_0 + h) - \Phi(z_0)] - [\Pi(p_0, z_0 + h) - \Pi(p_0, z_0)].$$

Dividing through by h, letting $h \to 0$ and using the subscript notation for partial derivatives,

$$\lim_{h\to 0} \frac{D}{h} = \Phi'(z_0) - \Pi_2(p_0, z_0).$$

The right-hand side is zero by the envelope theorem; therefore $D/h \approx 0$ if $|h|$ is small.

(ii) $\Pi(p, z) = -bp^2 + (bc + z)p - (cz + k)$. Completing the square,

$$\Pi(p, z) = -b(p - t)^2 + bt^2 - (cz + k),$$

where $t = \frac{1}{2}(c + [z/b])$. It is clear that t is the optimal value of p, given z. Hence

$$\Phi(z_1) = bp_1^2 - (cz_1 + k), \quad \Pi(p_0,\, z_1) = -b(p_0 - p_1)^2 + bp_1^2 - (cz_1 + k),$$

and $p_i = \frac{1}{2}(c + [z_i/b])$ for $i = 1, 2$. It follows that

$$D = b(p_0 - p_1)^2 = b\left(\frac{z_0}{2b} - \frac{z_1}{2b}\right)^2 = \frac{h^2}{4b}.$$

[A slightly different way of answering part (i) is to approximate D for small $|h|$ by a quadratic function of h, using the method of Section 10.3; then use the envelope theorem to show that only the term in h^2 does not vanish. Part (ii) exhibits the case where the approximation is exact.]

18–2. The Lagrangian for the problem is

$$\begin{aligned}\mathcal{L}(K_1,&K_2, L_1, L_2, \lambda, \mu, p_1, p_2, K, L)\\ &= p_1F_1(K_1, L_1) + p_2F_2(K_2, L_2) - \lambda(K_1 + K_2 - K) - \mu(L_1 + L_2 - L).\end{aligned}$$

(i) For $i = 1, 2$, $\partial V/\partial p_i = \partial\mathcal{L}/\partial p_i$ by the envelope theorem, and $\partial\mathcal{L}/\partial p_i = F_i(K_i, L_i)$.

(ii) $\partial V/\partial K = \partial\mathcal{L}/\partial K$ by the envelope theorem, and $\partial\mathcal{L}/\partial K = \lambda$. It remains to show that $\lambda = \partial F_i/\partial K_i$ for $i = 1, 2$. But this follows from the first-order conditions $\partial\mathcal{L}/\partial K_i = 0$ for $i = 1, 2$.

(iii) Similar to (ii).

18–3. (i) By Roy's identity,

$$\frac{x_1}{x_2} = \frac{\partial V}{\partial p_1} \bigg/ \frac{\partial V}{\partial p_2} = \frac{a\alpha m^\alpha p_1^{-\alpha-1}}{b\beta m^\beta p_2^{-\beta-1}}.$$

Setting $A = a\alpha(m/p_1)^\alpha,\ B = b\beta(m/p_2)^\beta$, we obtain

$$\frac{p_1x_1}{p_2x_2} = \frac{A}{B}.$$

(ii) By (i), $s_1 = A/(A + B)$ and $s_2 = B/(A + B)$.

(iii) $\ln x_1 - \ln x_2 = \ln(x_1/x_2) = \ln A - \ln B - \ln(p_1/p_2)$. Hence

$$\frac{\partial}{\partial m}(\ln x_1 - \ln x_2) = \frac{\partial}{\partial m}(\ln A) - \frac{\partial}{\partial m}(\ln B) = \frac{\alpha}{m} - \frac{\beta}{m}.$$

Denoting the income elasticities of demand by η_1 and η_2, this gives

$$\eta_1 - \eta_2 = \alpha - \beta. \tag{1}$$

The Engel aggregation condition is $p_1 \dfrac{\partial x_1}{\partial m} + p_2 \dfrac{\partial x_2}{\partial m} = 1$. This can be written in terms of s_1, s_2, η_1, η_2 as

$$s_1\eta_1 + s_2\eta_2 = 1. \tag{2}$$

Solving (1) and (2) simultaneously for η_1 and η_2 and remembering that $s_1 + s_2 = 1$ gives

$$\eta_1 = 1 + (\alpha - \beta)s_2, \quad \eta_2 = 1 - (\alpha - \beta)s_1.$$

(iv) Differentiating the expression for $\ln x - \ln x_2$ stated in (iii) with respect to p_1,

$$\frac{\partial}{\partial p_1}(\ln x_1 - \ln x_2) = \frac{\partial}{\partial p_1}(\ln A) - \frac{\partial}{\partial p_1}(\ln p_1) = -\frac{\alpha}{p_1} - \frac{1}{p_1}.$$

Denoting the two own-price elasticities by $\varepsilon_{11}, \varepsilon_{22}$ and the two cross-price elasticities by $\varepsilon_{12}, \varepsilon_{21}$, we have

$$\varepsilon_{11} - \varepsilon_{21} = -(\alpha + 1). \tag{3}$$

Similarly

$$\varepsilon_{22} - \varepsilon_{12} = -(\beta + 1). \tag{4}$$

The Cournot aggregation conditions are

$$p_1 \frac{\partial x_1}{\partial p_i} + p_2 \frac{\partial x_2}{\partial p_i} = -x_i \quad \text{for } i = 1, 2.$$

These can be written in elasticity form as

$$s_1\varepsilon_{11} + s_2\varepsilon_{21} = -s_1, \tag{5}$$

$$s_1\varepsilon_{12} + s_2\varepsilon_{22} = -s_2. \tag{6}$$

Solving (3) and (5) simultaneously for ε_{11} and ε_{21}, remembering that $s_1 + s_2 = 1$, gives

$$\varepsilon_{11} = -(1 + \alpha s_2), \quad \varepsilon_{21} = \alpha s_1.$$

A similar argument using (4) and (6) shows that

$$\varepsilon_{12} = \beta s_2, \quad \varepsilon_{22} = -(1 + \beta s_1).$$

Since α and β are positive, the own-price elasticities are negative and greater than 1 in absolute value, and the cross-price elasticities are positive.

18–4. (i) The firm's problem is to

$$\text{minimise} \;\; w_1x_1 + w_2x_2 \;\; \text{subject to} \;\; F(x_1, x_2) \geq q, \;\; \phi_1x_1 + \phi_2x_2 \leq E.$$

The feasible set lies above the negatively sloped convex isoquant $F(x_1, x_2) = q$ and below the straight line $\phi_1x_1 + \phi_2x_2 = E$. Suppose the points of intersection of the isoquant and straight line are A and B, and the absolute values of the slopes of the isoquant at these points are a b respectively, with $a < b$. Then there are three possibilities:

(1) if $a < w_1/w_2 < b$, the optimum is at a point of tangency of the isoquant and a member of the family of isocost lines $w_1x_1 + w_2x_2 = k$;

(2) if $w_1/w_2 \le a$, the optimum is at A;

(3) if $w_1/w_2 \ge b$, the optimum is at B.

(ii) The Lagrangian is

$$L(x_1, x_2, \lambda, \mu) = w_1x_1 + w_2x_2 - \lambda[F(x_1, x_2) - q] + \mu[\phi_1x_1 + \phi_2x_2 - E].$$

The Kuhn–Tucker conditions are

(a1) $w_1 - \lambda(\partial F/\partial x_1) + \mu\phi_1 = 0$;
(a2) $w_2 - \lambda(\partial F/\partial x_2) + \mu\phi_2 = 0$;
(b) $\lambda \ge 0,\ \ F(x_1, x_2) \ge q$, with complementary slackness;
(c) $\mu \ge 0,\ \ \phi_1x_1 + \phi_2x_2 \le E$, with complementary slackness.

(iii) (a1) and (a2) give

$$w_1/w_2 = \left(\lambda\frac{\partial F}{\partial x_1} - \mu\phi_1\right) \Big/ \left(\lambda\frac{\partial F}{\partial x_2} - \mu\phi_2\right),$$

which rearranges to

$$\lambda\left(\frac{1}{w_1}\frac{\partial F}{\partial x_1} - \frac{1}{w_2}\frac{\partial F}{\partial x_2}\right) = \mu\left(\frac{\phi_1}{w_1} - \frac{\phi_2}{w_2}\right). \qquad (*)$$

In case (1), $\phi_1x_1 + \phi_2x_2 < E$, so the complementary slackness condition of (c) gives $\mu = 0$ and $(*)$ reduces to tangency of the isoquant and an isocost line.

In case (2), the isocost lines must be less steep than the line $\phi_1x_1 + \phi_2x_2 = E$, so $w_1/w_2 < \phi_1/\phi_2$. Then $(*)$ confirms that the isoquant is at least as steep as the isocost line. Similarly, in case (3), $(*)$ confirms that isocost line is at least as steep as the isoquant.

(iv) The firm's problem now is to

minimise $w_1x_1 + \ldots + w_nx_n$ subject to $F(x_1, \ldots, x_n) \ge q$, $\phi_1x_1 + \ldots + \phi_nx_n \le E$.

The Lagrangian is

$$L(x_1, \ldots, x_n, \lambda, \mu) = (\textstyle\sum_{i=1}^n w_ix_i) - \lambda[F(x_1, \ldots, x_n) - q] + \mu[(\textstyle\sum_{i=1}^n \phi_ix_i) - E].$$

The Kuhn–Tucker conditions are

(a) $w_i - \lambda(\partial F/\partial x_i) + \mu\phi_i = 0$ for $i = 1, \ldots, n$;
(b) $\lambda \ge 0,\ \ F(x_1, \ldots, x_n) \ge q$, with complementary slackness;
(c) $\mu \ge 0,\ \ \phi_1x_1 + \ldots \phi_nx_n \le E$, with complementary slackness.

19 INTEGRATION

19–1. The first part is simple algebra. Using that result,

$$\int_3^4 \frac{3x-1}{x^2+x-6}\,dx = \int_3^4 \frac{1}{x-2}\,dx + \int_3^4 \frac{2}{x+3}\,dx = \Big[\ \ln(x-2)\Big]_3^4 + \Big[\,2\ln(x+3)\Big]_3^4,$$

which is evaluated as $\ln 2 - \ln 1 + 2\ln 7 - 2\ln 6 = \ln\frac{49}{18}$.

(i) The more general version of Rule 2 gives the required integral as

$$\Big[\ln|x-2|\Big]_{-5}^{-4} + 2\Big[\ln|x+3|\Big]_{-5}^{-4} = \ln 6 - \ln 7 + 2\ln 1 - 2\ln 2 = -\ln\tfrac{14}{3}.$$

(ii) Similarly to (i), the required integral is

$$\Big[\ln|x-2|\Big]_{-1}^{1} + 2\Big[\ln|x+3|\Big]_{-1}^{1} = \ln 1 - \ln 3 + 2\ln 4 - 2\ln 2 = \ln\tfrac{4}{3}.$$

(iii) Since 1 and 4 are on opposite sides of 2, the integral is not defined.

(iv) Since -4 and 0 are on opposite sides of -3, the integral is not defined.

(v) In this case there are two reasons why the integral is undefined!

19–2. The gross consumer's surplus is $\displaystyle\int_0^{q_0} f(q)\,dq$. The net consumer's surplus is the area bounded by the p–axis, the inverse demand function and the horizontal line $p = f(q_0)$ and is given by

$$\int_0^{q_0} f(q)\,dq - q_0 f(q_0).$$

When $f(q) = 30 - q^2$, the gross consumer's surplus is

$$\int_0^{q_0} (30 - q^2)\,dq = 30q_0 - \tfrac{1}{3}q_0^3$$

and the net consumer's surplus is

$$30q_0 - \tfrac{1}{3}q_0^3 - q_0(30 - q_0^2) = \tfrac{2}{3}q_0^3.$$

19–3. (i) The present value at time 0 of the profit gained during the short time interval $[t,\, t+h]$ is approximately $e^{-rt}g(t)h$. If we split $[0, T]$ into a large number of small sub-intervals, the present value at time 0 of the profit stream up to T can be approximated by a sum of terms of the above form. Passing to the limit as $h \to 0$, we get

$$V(T) = \int_0^T e^{-rt}g(t)\,dt.$$

(ii) $V'(T) = e^{-rT}g(T)$.

(iii) Let $f(t) = e^{-t/20}/(1+\sqrt{t})$. Then Simpson's rule with 5 ordinates gives

$$V(12) \approx 60[f(0) + 4f(3) + 2f(6) + 4f(9) + f(12)] \approx 207.0.$$

(iv) When T increases from 12 to 12.5, then, by (ii) and the small increments formula,

$$\Delta V \approx 60e^{-0.6} \times (1+\sqrt{12})^{-1} \times 0.5 \approx 3.7,$$

so $V(12.5) \approx 210.7$.

19–4. (i) (a) The value of the investment at time $t + \Delta t$ is equal to the value at time t plus the interest gained in the time interval $[t,\, t+\Delta t]$. Approximating this interest by that on $A(t)$ at the rate $r(t)$ gives the result as stated.

(b) Rearranging the result of (a) gives

$$\frac{A(t+\Delta t)-A(t)}{A(t)\,\Delta t}=r(t).$$

Taking the limit as $\Delta t\to 0$, we obtain $A'(t)/A(t)=r(t)$, as required.
Integrating the result of (b) over the interval $[0,T]$ gives

$$\Big[\ln A(t)\Big]_0^T=\int_0^T r(t)\,dt. \tag{$*$}$$

Since $A(0)=P$, the left-hand side of $(*)$ is $\ln(A(T)/P)$, and the result follows.

(ii) Let $P(t,h)$ be the present value at time 0 of the income received during the short time interval $[t,t+h]$. By the final result of (i), with T replaced by t and t by s,

$$f(t)\,h\approx P(t,h)\,\exp\left[\int_0^t r(s)\,ds\right].$$

Hence $P(t,h)\approx e^{-R(t)}f(t)h$, where $R(t)=\displaystyle\int_0^t r(s)\,ds$. If we split $[0,T]$ into a large number of small sub-intervals, the present value at time 0 of the income stream up to T can be approximated by a sum of terms of the form $P(t,h)$. Passing to the limit as $h\to 0$, the present value of the stream is $\displaystyle\int_0^T e^{-R(t)}f(t)\,dt$.

20 ASPECTS OF INTEGRAL CALCULUS

20–1. (i) Putting $t=1-x^2$ changes the integral to

$$\int_1^0 t^{1/2}\left(-\frac{1}{2}dt\right)=\frac{1}{2}\int_0^1 t^{1/2}\,dt=\frac{1}{2}\left[\frac{2}{3}t^{3/2}\right]_0^1=\frac{1}{3}.$$

(ii) Putting $t=1-x$ changes the integral to

$$\int_1^0 (1-t)t^{1/2}(-dt)=\int_0^1 (t^{1/2}-t^{3/2})\,dt=\frac{2}{3}-\frac{2}{5}=\frac{4}{15}.$$

(iii) Putting $t=1-x$ changes the integral to

$$\int_1^0 (1-t)^2t^{1/2}(-dt)=\int_0^1 (t^{1/2}-2t^{3/2}+t^{5/2})\,dt=\frac{2}{3}-\frac{4}{5}+\frac{2}{7}=\frac{16}{105}.$$

20–2. $\displaystyle\int_{-A}^A f(x)\,dx=I+J$, where $I=\displaystyle\int_{-A}^0 f(x)\,dx$ and $J=\displaystyle\int_0^A f(x)\,dx$. Making the substitution $y=-x$,

$$I=-\int_A^0 f(-y)\,dy=\int_0^A f(-y)\,dy.$$

In case (i), $I=J$ and the result follows. In case (ii), $I=-J$ and the result follows.

(iii) Denote the required integral by K. By result (i),

$$K=2\int_0^1 e^{-x^2/2}\,dx.$$

Using Simpson's rule with 5 ordinates, $K\approx 1.49$ to 2 decimal places.

(iv) The function $y = xe^{-x^2/2}$ is odd. Hence by (ii), $\int_{-A}^{A} xe^{-x^2/2}\,dx = 0$ for all A, and the result follows by letting $A \to \infty$.
This is a special case of the Example in Section 20.2, because we are forcing the limits of integration to tend to infinity together; in the original example, they tend to infinity independently.

(v) By a similar argument to (iv), $\int_{-A}^{A} x^3\,dx = 0$ for all A, so the integral remains zero when we let $A \to \infty$. It is not correct to infer that $\int_{-\infty}^{\infty} x^3\,dx = 0$, because the integrals $\int_0^{\infty} x^3\,dx$ and $\int_{-\infty}^{0} x^3\,dx$ diverge.

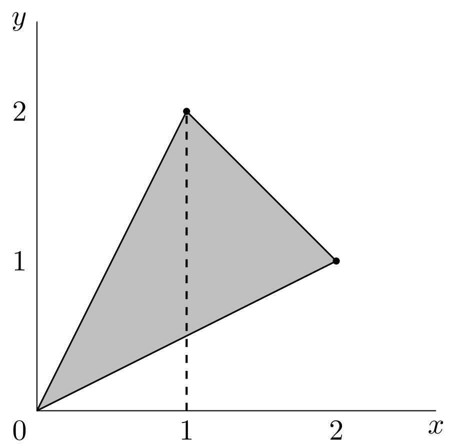
20–3. $V(s) = \int_0^s (-ce^{-rt})\,dt + \int_s^T f(s,t)e^{-rt}\,dt$. The first integral on the right-hand side is easily evaluated as $(c/r)(e^{-rs} - 1)$. Thus

$$V'(s) = -ce^{-rs} + \frac{\partial}{\partial s}\int_s^T f(s,t)e^{-rt}\,dt = -ce^{-rs} - f(s,s)e^{-rs} + \int_s^T \frac{\partial f}{\partial s}e^{-rt}\,dt$$

by Leibniz's rule. Therefore, the value of s which maximises $V(s)$ must satisfy the equation

$$f(s,s) + c = \int_s^T e^{r(s-t)}\frac{\partial f}{\partial s}\,dt.$$

20–4. The equations of the sides are $y = 2x$, $y = \frac{1}{2}x$ and $x + y = 3$.



Let the required integral be I. Dividing the region of integration as in the diagram,

$$\begin{aligned}
I &= \int_0^1 \left[\int_{x/2}^{2x} (x+y)\,dy\right] dx + \int_1^2 \left[\int_{x/2}^{3-x} (x+y)\,dy\right] dx \\
&= \int_0^1 \left[\frac{3}{2}x^2 + \frac{1}{2}(2x)^2 - \frac{1}{2}(x/2)^2\right] dx + \int_1^2 \left[3x\left(1 - \frac{x}{2}\right) + \frac{1}{2}(3-x)^2 - \frac{1}{8}x^2\right] dx \\
&= \frac{27}{8}\int_0^1 x^2\,dx + \frac{9}{8}\int_1^2 (4 - x^2)\,dx \\
&= \frac{9}{8} + \frac{9}{2} - \frac{3}{8}(8-1) \\
&= 3.
\end{aligned}$$

21 PROBABILITY

21–1. (i) We number the arriving couples $1, 2, \ldots, n$, where couple j consists of Mr j and Ms j. Let A be the event that at least one of these couples leaves together, B the event that Ms 1 leaves with Mr 1, C the event that Ms 1 leaves with Mr 2. Then

$$p_n = P(A) = P(B) + (n-1)P(A \cap C).$$

Since $P(B) = P(C) = 1/n$, it follows that

$$p_n = \frac{1}{n}\left[1 + (n-1)P(A|C)\right].$$

Those present at the party other than Ms 1 and Mr 2 are Mr 1, Ms 2 and couples $3, \ldots, n$. Hence

$$P(A|C) = p_{n-1} - P(D|C),$$

where D is the event that Mr 1 leaves with Ms 2 and **none** of the couples $3, \ldots, n$ leave together. Now

$$P(D|C) = \frac{1}{n-1}(1 - p_{n-2}),$$

so

$$p_n = \frac{1}{n}(1 + (n-1)p_{n-1} - 1 + p_{n-2}) = p_{n-1} - \frac{1}{n}(p_{n-1} - p_{n-2}).$$

(ii) $p_2 - p_1 = -\dfrac{1}{2}$. Hence by (i), $p_3 - p_2 = \left(-\dfrac{1}{3}\right) \times \left(-\dfrac{1}{2}\right) = \dfrac{1}{3!}$, $p_4 - p_3 = -\dfrac{1}{4} \times \dfrac{1}{3!}$ and so on. Thus for all $n > 1$,

$$p_n - p_{n-1} = \frac{(-1)^{n-1}}{n!},$$

whence

$$\begin{aligned} p_n &= p_1 + (p_2 - p_1) + \ldots + (p_n - p_{n-1}) \\ &= 1 - \frac{1}{2} + \ldots + \frac{(-1)^{n-1}}{n!}. \end{aligned}$$

It follows that

$$1 - p_n = 1 - \frac{1}{1!} + \frac{1}{2!} - \ldots + \frac{(-1)^n}{n!},$$

which approaches e^{-1} as $n \to \infty$. Thus

$$\lim_{n \to \infty} p_n = 1 - e^{-1} = 0.632 \text{ to 3 decimal places.}$$

21–2. $P(A) = P(B) = \frac{1}{2}$. For C you need either Heads at 9am and 11am or Tails at 9am and 10am, so $P(C) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$. $A \cap B$ is the event that you get Heads at 9am and 10am, while $A \cap C$ is the event that you get Heads at 9am and 11am. $B \cap C$ is the same event as $A \cap B \cap C$, namely Heads at 9am, 10am and 11am. Therefore

$$P(A) = P(B) = P(C) = \frac{1}{2}, \quad P(A \cap B) = P(A \cap C) = \frac{1}{4}, \quad P(B \cap C) = P(A \cap B \cap C) = \frac{1}{8}.$$

In particular (i) $P(B \cap C) \neq P(B)P(C)$, (ii) $P(A \cap B \cap C) = P(A)P(B)P(C)$. (i) says that B and C are not independent; hence the three events A, B, C are **not** independent. The example shows that equation (ii) alone is not enough to ensure that three events A, B, C are independent.

21–3. (i)

$$f_X(x) = \begin{cases} 0 & \text{if } x \le k, \\ \alpha k^\alpha x^{-\alpha-1} & \text{if } x > k. \end{cases}.$$

(a) If $x \le K$, $P(Y > x) = \dfrac{P(X > K)}{P(X > K)} = 1$. If $x > K$,

$$P(Y > x) = \frac{P(X > x)}{P(X > K)} = \frac{(k/x)^\alpha}{(k/K)^\alpha} = (K/x)^\alpha.$$

Thus Y is Pareto with parameters α and K.

(b)

$$P(Z > z) = P(X/k > e^z) = \begin{cases} 1 & \text{if } e^z \le 1, \\ e^{-\alpha z} & \text{if } e^z > 1. \end{cases}$$

Hence

$$F_Z(z) = \begin{cases} 0 & \text{if } z \le 0, \\ 1 - e^{-\alpha z} & \text{if } z > 0. \end{cases}$$

Z is exponential with parameters α.

(ii) In each case, we denote the median by ν.

(a) Let X be Pareto with parameters α and k. Then $(k/\nu)^\alpha = \frac{1}{2}$, so $\nu = 2^{1/\alpha}k$.
(b) Let X be exponential with parameter α. Then $e^{-\alpha\nu} = \frac{1}{2}$, so $\nu = \alpha^{-1}\ln 2$.
(c) $\nu^3 = \frac{1}{2}$, so $\nu = 2^{-1/3}$.

21–4. Let $p_r(u)$ be the probability that there are r calls in an interval of length u. This is the probability that a Poisson variate with parameter λu takes the value r, so

$$p_r(u) = \frac{(\lambda u)^r}{r!} e^{-\lambda u}. \tag{$*$}$$

$P(t < T_k \le t+\delta)$ is the probability that the kth call takes place between time t and time $t+\delta$; it is therefore the probability that for some $j = 1, \ldots, k$ there are $k - j$ calls in an interval of length t followed by j calls in an interval of length δ. Thus

$$P(t < T_k \le t + \delta) = \sum_{j=1}^{k} p_{k-j}(t)\, p_j(\delta).$$

If δ is small then $p_1(\delta) \approx \lambda\delta$ and $p_j(\delta)$ is negligible for $j > 1$; hence

$$P(t < T_k \le t + \delta) \approx p_{k-1}(t)\lambda\delta,$$

and the required result follows from $(*)$.

Denoting the density function of T_k by f, we have

$$f(t) = \lim_{\delta \downarrow 0} \frac{P(t < T_k \le t + \delta)}{\delta} = \lambda p_{k-1}(t)$$

for all $t > 0$. Hence by $(*)$,

$$f(t) = \begin{cases} 0 & \text{if } t \le 0, \\ \dfrac{\lambda^k t^{k-1}}{(k-1)!}\, e^{-\lambda t} & \text{if } t > 0; \end{cases}$$

22 EXPECTATION

22–1. To simplify notation, we denote the function F_X by F and the associated density function by f.

(i) If the rth raw moment exists it is

$$\alpha k^\alpha \int_k^\infty x^{r-\alpha-1}\,dx.$$

The integral converges if and only if $r-\alpha-1<-1$, which happens if and only if $r<\alpha$. In that case

$$E(X^r) = \alpha k^\alpha \left[0 - \frac{k^{r-\alpha}}{r-\alpha}\right] = \frac{\alpha}{\alpha-r}k^r.$$

(ii) Let total population be N. Average wealth per person is EX. If x and h are positive, then the proportion of total wealth held by persons whose wealth is between x and $x+h$ is

$$\frac{w[F(x+h)-F(x)]N}{NEX},$$

where w is some number between x and $x+h$. If h is small, this proportion is approximately $xf(x)h/EX$. Hence, for $x>1$,

$$V(x) = \frac{1}{EX}\int_1^x uf(u)\,du = \frac{\alpha-1}{\alpha}\int_1^x \alpha u^{-\alpha}\,du = 1 - x^{1-\alpha}.$$

(iii) For $x>1$, $p = 1-x^{-\alpha}$ and $q = 1-x^{-(\alpha-1)}$. Thus

$$L(p) = 1-(1-p)^\beta,$$

where $\beta = (\alpha-1)/\alpha$. In a notation similar to that of Section 19.4, the Gini coefficient G is $A/(A+B)$ where $A+B=\frac{1}{2}$ and $B=\int_0^1 L(p)\,dp$. Using the substitution $t=1-p$,

$$B = \int_0^1 \left(1-(1-p)^\beta\right)dp = 1 - \int_0^1 t^\beta\,dt = 1-\frac{1}{\beta+1} = \frac{\alpha-1}{2\alpha-1}.$$

Hence

$$G = 1-2B = \frac{1}{2\alpha-1}.$$

22–2. (i) By symmetry, each of X and Y has the same c.d.f. F and therefore the same density function f. Since $P(X\le x) = P(X\le x \text{ and } Y\le 1)$ for all x, $F(x)=\frac{1}{2}x(x+1)$ if $0\le x\le 1$. Hence

$$F(x) = \begin{cases} 0 & \text{if } x\le 0,\\ \frac{1}{2}x(x+1) & \text{if } 0<x<1,\\ 1 & \text{if } x\ge 1,\end{cases} \qquad f(x) = \begin{cases} 0 & \text{if } x\le 0,\\ x+\frac{1}{2} & \text{if } 0<x<1,\\ 0 & \text{if } x\ge 1.\end{cases}$$

(ii) Suppose that $0<x-\delta<x<x+\delta<1$. If $0\le y\le 1$, then

$$\begin{aligned} &P(Y\le y \text{ and } x-\delta<X<x+\delta)\\ &\qquad = \tfrac{1}{2}(x+\delta)y(x+\delta+y) - \tfrac{1}{2}(x-\delta)y(x-\delta+y)\\ &\qquad = \tfrac{1}{2}y((x+\delta)^2-(x-\delta)^2+2y\delta)\\ &\qquad = (2x+y)y\delta.\end{aligned}$$

Setting $y = 1$, we see that $P(x-\delta < X < x+\delta) = (2x+1)\delta$. Hence for any y such that $0 \le y \le 1$,

$$P(Y \le y \,|\, x-\delta < X < x+\delta) = \frac{2xy+y^2}{2x+1}.$$

This remains true as $\delta \downarrow 0$, so

$$P(Y \le y \,|\, X = x) = \frac{2xy+y^2}{2x+1}.$$

The corresponding density function is $(2x+1)^{-1}(2x+2y)$, whence

$$E(Y \,|\, X=x) = \frac{1}{2x+1}\int_0^1 y(2x+2y)\,dy = \frac{1}{2x+1}\left[xy^2 + \frac{2}{3}y^3\right]_{y=0}^{y=1} = \frac{3x+2}{3(2x+1)}.$$

Hence

$$E(Y|X) = \frac{3X+2}{3(2X+1)}.$$

(iii) Using (ii) and the law of iterated expectations, $EY = \int_0^1 g(x)f(x)\,dx$, where

$$g(x) = \frac{3x+2}{3(2x+1)} = \frac{3x+2}{6f(x)}.$$

Hence

$$EY = \int_0^1 \left(\frac{x}{2}+\frac{1}{3}\right)\,dx = \frac{1}{4}+\frac{1}{3} = \frac{7}{12}.$$

The same value of EY is obtained by direct calculation:

$$\int_0^1 yf(y)\,dy = \int_0^1 \left(y^2+\frac{y}{2}\right)\,dy = \frac{1}{3}+\frac{1}{4} = \frac{7}{12}.$$

22–3. (i) Since the pair (X,Y) is bivariate normal, there is an invertible matrix $\mathbf{A}$, a vector $\mathbf{b}$ and a pair of independent, standard normal r.v.s (U,V) such that

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \mathbf{A}\begin{bmatrix} U \\ V \end{bmatrix} + \mathbf{b}.$$

Hence

$$\begin{bmatrix} W \\ Z \end{bmatrix} = \mathbf{BA}\begin{bmatrix} U \\ V \end{bmatrix} + \ (\mathbf{Bb}+\mathbf{c}).$$

Since $\mathbf{A}$ and $\mathbf{B}$ are invertible, so is $\mathbf{BA}$: the pair (W,Z) is bivariate normal.

(ii) Let $\theta = \operatorname{var} X$, $\lambda = \operatorname{cov}(X,Y)$. Then

$$\begin{bmatrix} X \\ Z \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -\lambda & \theta \end{bmatrix}\begin{bmatrix} X \\ Y \end{bmatrix}.$$

The pair (X,Z) is bivariate normal by (i), so to prove that X and Z are independent it suffices to show that $\operatorname{cov}(X,Z) = 0$. In fact,

$$\operatorname{cov}(X,Z) = \theta\operatorname{cov}(X,Y) - \lambda\operatorname{var} X = \theta\lambda - \lambda\theta = 0.$$

Since X and Z are independent, $E(Z|X) = EZ$. Therefore

$$\theta E(Y|X) - \lambda X = \theta EY - \lambda EX$$

and

$$E(Y|X) = EY + \frac{\lambda}{\theta}(X-EX) = EY + \frac{\operatorname{cov}(X,Y)}{\operatorname{var} X}(X-EX).$$

22–4. As usual, Φ and ϕ denote the standard normal c.d.f. and density function.

(i) Let $Y = \ln X$. If $x > 0$, then

$$P(X \leq x) = P(Y \leq \ln x) = \Phi\left(\frac{\ln x - \mu}{\sigma}\right).$$

The c.d.f. of X is therefore

$$F_X(x) = \begin{cases} 0 & \text{if } x \leq 0, \\ \Phi\left(\dfrac{\ln x - \mu}{\sigma}\right). & \text{if } x > 0. \end{cases}$$

Since $X = e^Y$, $EX = M_Y(1)$. Using the formula for the moment generating function of a normal r.v. at the end of Section 22.2, we see that

$$EX = \exp\left(\mu + \tfrac{1}{2}\sigma^2\right).$$

Similarly

$$E(X^2) = M_Y(2) = \exp(2\mu + 2\sigma^2) = (EX)^2 \exp(\sigma^2),$$

whence

$$\operatorname{var} X = (EX)^2\,[\exp(\sigma^2) - 1].$$

(ii) $\ln X$ is the sum of the n independent, identically distributed r.v.s $\ln Y_1, \ldots, \ln Y_n$. Thus if n is large, $\ln X$ has an approximately normal distribution by the central limit theorem, so X has an approximately log-normal distribution.

In the numerical example, let $W_i = \ln Y_i$ $(i = 1, \ldots, 100)$, $\theta = 0.1$. Then

$$EW_i = \frac{1}{2\theta}\int_{-\theta}^{\theta} \ln(1+t)\,dt.$$

Writing the integrand as $\ln(1+t) \times \dfrac{d}{dt}(1+t)$ and integrating by parts, we see that

$$EW_i = \frac{1}{2\theta}\left\{(1+\theta)\ln(1+\theta) - (1-\theta)\ln(1-\theta)\right\} - 1.$$

A similar integration by parts shows that

$$E(W_i^2) = \frac{1}{2\theta}\left\{(1+\theta)[\ln(1+\theta)]^2 - (1-\theta)[\ln(1-\theta)]^2\right\} - 2EW_i.$$

Since $\theta = 0.1$, $EW_i = -1.6717 \times 10^{-3}$ and $E(W_i^2) = 3.3517 \times 10^{-3}$, so $\operatorname{var} W_i = 3.3490 \times 10^{-3}$. It follows from the central limit theorem that $\ln X$ is approximately normally distributed with mean -0.16717 and variance 0.33490. Therefore

$$P(X \leq 1) = P(\ln X \leq 0) \approx \Phi(z),$$

where $z = 0.16717/\sqrt{0.33490} = 0.289$. From tables, the required probability is approximately 0.61.

23 INTRODUCTION TO DYNAMICS

23–1. (i) All solutions of the differential equation satisfy $dy/dt = 0$ when $t = 0$ and $t = 2$. It is also clear from the differential equation that, as t increases from $0-$ to $0+$, dy/dt changes from negative to positive. Hence there is a minimum at $t = 0$. Similarly, as t increases from $2-$ to $2+$, dy/dt changes from positive to negative; there is therefore a maximum at $t = 2$.

From the differential equation, the slopes of the solution curves when $t = -1, 1, 3$ are $-3, 1, -3$ respectively. The directions of the tangents to the curves when $t = -1, 0, 1, 2, 3$ can be shown as sets of parallel line segments of slopes $-3, 0, 1, 0, -3$ cutting the lines $t = -1, 0, 1, 2, 3$ respectively. This enables us to draw the family of solution curves, each with a minimum at $t = 0$, a maximum at $t = 2$ and with directions of tangents as just stated at $t = -1, 1, 3$. It is then clear that, as t increases beyond 2, dy/dt becomes more negative and y decreases. As $t \to \infty$, $y \to -\infty$.

(ii) Along the line $y = 0$, y is constant, so $dy/dt = 0$; also $y(2-y) = 0$; hence $y = 0$ is a solution curve. By a similar argument, $y = 2$ is also a solution curve. The directions of the tangents to the curves when $y = -1$, -0.5, 0.5, 1, 1.5, 2.5, 3 can be shown as sets of parallel line segments of slopes -3, -1.25, 0.75, 1, 0.75, -1.25, -3 cutting the lines $y = -1$, -0.5, 0.5, 1, 1.5, 2.5, 3 respectively. This enables us to draw the family of solution curves. Along the solution curves below $y = 0$, y decreases as t increases and, as $t \to \infty$, $y \to -\infty$. Along the solution curves between $y = 0$ and $y = 2$, y increases as t increases and, as $t \to \infty$, $y \to 2$. Along the solution curves above $y = 2$, y decreases as t increases and, as $t \to \infty$, $y \to 2$.

(iii) This problem corresponds to the case $a = 2$, $b = 1$ of Exercise 21.1.5 (see also Exercise 21.3.4). There you were asked to find the solution of the differential equation which satisfies $y = y_0$ when $t = 0$ where $0 < y_0 < a/b$. The solution obtained satisfies $0 < y < a/b$ for all t and, as $t \to \infty$, $y \to a/b$. This confirms the behaviour of the solution curves lying between $y = 0$ and $y = 2$.

23–2. (i) Separating the variables and integrating, $\int z^{-1}\,dz = \int r\,dt$. Hence $\ln z = rt + B$, which can be arranged in the form $z = Ce^{rt}$.

(ii) Since extraction costs are zero, (i) gives $p = Ce^{rt}$. Assuming the market for the resource clears at each instant of time, we have $Ce^{rt} = q^{-\alpha}$, which can be arranged in the form $q = Ae^{(-r/\alpha)t}$.

(iii) Since the total amount of mineral to be extracted is S, $\displaystyle\int_0^\infty q(t)\,dt = S$. But

$$\int_0^\infty q(t)\,dt = \lim_{T\to\infty}\int_0^T Ae^{-\gamma t}dt = \lim_{T\to\infty}\left[(A/\gamma)(1-e^{-\gamma T})\right] = A/\gamma,$$

where $\gamma = r/\alpha$. It follows that $A = \gamma S$. Summarising,

$$q(t) = \frac{rS}{\alpha}e^{(-r/\alpha)t}, \quad p(t) = \left(\frac{\alpha}{rS}\right)^\alpha e^{rt}.$$

23–3. (i) K satisfies the differential equation

$$\frac{dK}{dt} + \delta K = sAK^\alpha L^{1-\alpha}.$$

Hence

$$\frac{1}{L}\,\frac{dK}{dt} = sAk^\alpha - \delta k,$$

where $k = K/L$. Now

$$\frac{dk}{dt} = \frac{1}{L}\frac{dK}{dt} - \frac{K}{L^2}\frac{dL}{dt} = \frac{1}{L}\frac{dK}{dt} - nk,$$

since $L(t) = L(0)e^{nt}$. It follows that

$$\frac{dk}{dt} + (\delta + n)k = sAk^{\alpha}.$$

This is similar to (21.17) but with K replaced by k and δ by $\delta + n$. The general solution for $k = K/L$ is therefore

$$k(t) = \left[\frac{sA}{\delta + n} + ce^{-(1-\alpha)(\delta+n)t}\right]^{1/(1-\alpha)},$$

where c is a constant. Hence

$$K(t) = L(0)e^{nt}k(t) = L(0)\left[\frac{sA}{\delta + n}e^{(1-\alpha)nt} + ce^{-(1-\alpha)\delta t}\right]^{1/(1-\alpha)}.$$

Notice that c is related to the value of K/L at time 0 by the equation

$$c = k_0^{1-\alpha} - \frac{sA}{\delta + n},$$

where $k_0 = K(0)/L(0)$. Thus c has the same sign as $k_0 - k^*$, where

$$k^* = \left[\frac{sA}{\delta + n}\right]^{1/(1-\alpha)}.$$

Also notice that $k(t) \to k^*$ as $t \to \infty$.

The graph of K/L against t meets the vertical axis at $(0, k_0)$. If $k_0 < k^*$, the graph is upward-sloping and concave, approaching the horizontal line through $(0, k^*)$ from below as $t \to \infty$. If $k_0 > k^*$, the graph is downward-sloping and convex, approaching the same horizontal line from above as $t \to \infty$.

The graph of $\ln K$ against t meets the vertical axis at $(0, \ln K(0))$ and is asymptotic (as $t \to \infty$) to a straight line S of slope n. If $k_0 < k^*$, the graph is upward-sloping and concave, approaching S from below as $t \to \infty$. If $k_0 > k^*$, the graph is convex, approaching S from above as $t \to \infty$. In the latter case the graph is downward-sloping for large negative t and may slope up or down for small $|t|$.

(ii) K satisfies the differential equation

$$\frac{dK}{dt} + \delta K = sAK^{\alpha}L^{\beta}.$$

Setting $N = L^{\beta/(1-\alpha)}$ we obtain the differential equation

$$\frac{dK}{dt} + \delta K = sAK^{\alpha}N^{1-\alpha}.$$

We may therefore proceed as in (i), with L replaced by N; notice that the rate of growth of L must be replaced by the rate of growth of N, so n is replaced by $n\beta/(1-\alpha)$. The solution for K is therefore

$$K(t) = \left[L(0)^{\beta}\left(\frac{(1-\alpha)sA}{(1-\alpha)\delta + \beta n}e^{\beta nt} + ce^{-(1-\alpha)\delta t}\right)\right]^{1/(1-\alpha)},$$

where c is a constant.

23–4. The discrete-time analogue is $\Delta y_t + ay_t = b$, or

$$y_{t+1} + (a-1)y_t = b.$$

Letting $\bar{y} = b/a$, we may write the general solution of the differential equation as $y(t) = \bar{y} + Ae^{-at}$, and the general solution of the difference equation as $y_t = \bar{y} + A(1-a)^t$. Both for the differential equation and for the difference equation, the stationary solution is $y = \bar{y}$.

For the differential equation, all solutions approach the stationary solution as $t \to \infty$, provided $\lim_{t\to\infty} e^{-at} = 0$; this occurs if and only if $a > 0$. For the difference equation, all solutions approach the stationary solution as $t \to \infty$, provided $\lim_{t\to\infty}(1-a)^t = 0$; this occurs if and only if $|1-a| < 1$, i.e. $0 < a < 2$. The general solution of the difference equation exhibits alternating behaviour if $a > 1$.

Notice that the variety of possible behaviour is greater for the difference equation than for the differential equation, in that alternating behaviour is possible for the latter. The qualitative behaviour of the discrete-time analogue is not necessarily the same as that of the differential equation for the same parameter values.

24 THE CIRCULAR FUNCTIONS

24–1. (i) $y = e^{-3x}\sin 4x$ cuts the x–axis where $\sin 4x = 0$: $x = k\pi/4$ for $k = 0, \pm1, \pm2, \ldots$ $dy/dx = e^{-3x}(4\cos 4x - 3\sin 4x)$. Hence $dy/dx = 0$ if and only if $\tan 4x = \frac{4}{3}$. Setting $\beta = \arctan\frac{4}{3}$ ($= 0.927$ to 3 decimal places), we see that $dy/dx = 0$ if and only if $4x = \beta + k\pi$ for some integer k. Hence the critical points are $\{\,(x_k, y_k) : k = 0, \pm1, \pm2, \ldots\}$, where $x_k = \frac{1}{4}(\beta + k\pi)$ and

$$y_k = \exp(-\tfrac{3}{4}(\beta + k\pi))\sin(\beta + k\pi) = e^{-3\beta/4}e^{-3k\pi/4}(-1)^k\sin\beta.$$

Thus $y_k = au^k$ for all k, where $a = e^{-3\beta/4}\sin\beta$ and $u = -e^{-3\pi/4}$: notice that $-1 < r < 0$. It is clear from the above that the positive critical values are maxima and the negative critical values are minima; for a rigorous demonstration of this, see Problem 26–2. As $x \to -\infty$, $|y| \to \infty$; as $x \to \infty$, $y \to 0$. The graph is that of a damped oscillation.

(ii) As in (i), the graph cuts the x–axis where $x = k\pi/4$ for $k = 0, \pm1, \pm2, \ldots$ Also as in (i), let $\beta = \arctan\frac{4}{3}$. The critical points are now $\{\,(X_k, Y_k) : k = 0, \pm1, \pm2, \ldots\}$, where $X_k = \frac{1}{4}(-\beta + k\pi)$ and

$$Y_k = \exp(\tfrac{3}{4}(-\beta + k\pi))\sin(-\beta + k\pi) = e^{-3\beta/4}e^{3k\pi/4}(-1)^{k+1}\sin\beta.$$

Thus $Y_k = bv^k$ for all k, where $b = -e^{-3\beta/4}\sin\beta$ and $v = -e^{3\pi/4}$: notice that $s < -1$. The graph is that of an explosive oscillation.

24–2. (i) The graph of $3t + 1$ is a straight line of slope 3 and intercept 1. The graph of $2\sin 6t$ is like that of $\sin t$ but magnified by a factor of 2 and with period $\pi/3$. The graph of $\ln Y$ is the sum of these two and is thus an oscillation of period $\pi/3$ about $3t + 1$.

(ii) The graph of $2t + 5$ is a straight line of slope 2 and intercept 5. If $\alpha > 0$, the graph of $3\sin(6t + \alpha)$ is like that of $\sin t$ but magnified by a factor of 3, with period $\pi/3$ and shifted to the left through $\alpha/6$ (since $6t + \alpha = 6[t + \frac{1}{6}\alpha]$). The graph of $\ln Z$ is the sum of these two and is thus an oscillation of period $\pi/3$ about $2t + 5$. Since, for example, the maximum and minimum points of $3\sin(6t + \alpha)$ occur at a time $\alpha/6$ earlier than the corresponding points of $2\sin 6t$, the former periodic function is said to **lead** the latter by $\alpha/6$. Similarly, if $\alpha < 0$, the maximum and minimum points of $3\sin(6t + \alpha)$ occur at a time $|\alpha|/6$ later than the corresponding points of $2\sin 6t$; the former periodic function is then said to **lag behind** the latter by $|\alpha|/6$.

(a) If $\alpha = \pi$, the periodic component of $\ln Z$ leads the corresponding component of $\ln Y$ by $\pi/6$, i.e. half a period. When the periodic component of $\ln Y$ is at a maximum, the corresponding component of $\ln Z$ is at a minimum and vice versa.

(b) If $\alpha = -\pi$, the periodic component of $\ln Z$ lags behind the corresponding component of $\ln Y$ by $\pi/6$, i.e. half a period. Again, when the periodic component of $\ln Y$ is at a maximum, the corresponding component of $\ln Z$ is at a minimum and vice versa.

(c) If $\alpha = 3\pi$, the periodic component of $\ln Z$ leads the corresponding component of $\ln Y$ by $\pi/2$, i.e. 1.5 periods. As in (a), when the periodic component of $\ln Y$ is at a maximum, the corresponding component of $\ln Z$ is at a minimum and vice versa.

(d) If $\alpha = -3\pi$, the periodic component of $\ln Z$ lags behind the corresponding component of $\ln Y$ by $\pi/2$, i.e. 1.5 periods. As in (b), when the periodic component of $\ln Y$ is at a maximum, the corresponding component of $\ln Z$ is at a minimum and vice versa.

(e) If $\alpha = 6\pi$, the periodic component of $\ln Z$ leads the corresponding component of $\ln Y$ by π, i.e. 3 periods. When the periodic component of $\ln Y$ is at a maximum or minimum, the corresponding component of $\ln Z$ is at a similar point.

(f) If $\alpha = -6\pi$, the periodic component of $\ln Z$ lags behind the corresponding component of $\ln Y$ by π, i.e. 3 periods. When the periodic component of $\ln Y$ is at a maximum or minimum, the corresponding component of $\ln Z$ is at a similar point.

(iii) Yes they can. Denote the periodic components of $\ln Y$ and $\ln P$ by y, p respectively. Suppose for example that

$$y = a \sin 6t, \quad p = -b \sin 6t,$$

where a and b are positive constants. Then A is obviously right. The periodic component of the inflation rate at time t is

$$dp/dt = -6b \cos 6t = 6b \sin\left(6t - \tfrac{\pi}{2}\right).$$

Hence the periodic component of the inflation rate at time $t + \frac{\pi}{12}$ is $6b \sin 6t$, so B is also correct.

24–3. Let (R, α) be the polar coordinates of the point with Cartesian coordinates (A, B). Then

$$\begin{aligned} A\cos\theta + B\sin\theta &= R\cos\alpha\cos\theta + R\sin\alpha\sin\theta \\ &= R(\cos\theta\cos\alpha + \sin\theta\sin\alpha) \\ &= R\cos(\theta - \alpha). \end{aligned}$$

In case (i), $R = 3\sqrt{2}$, $\alpha = \pi/4$; in case (ii), $R = 4$, $\alpha = \pi/2$; in case (iii), $R = 13$, $\alpha = \arctan 2.4 = 1.176$ to 3 decimal places.

From (iii), we may write the function as $y = 13\cos(x - \alpha)$, where $\alpha = \arctan 2.4$. Thus the graph is like that of $y = \cos x$ but magnified by a factor of 13 and shifted to the right by approximately 1.176 radians. In particular, the maximum and minimum values of y are ± 13.

24–4. (i) As $x \to 0$, $\arctan([m/x]^\alpha) \to \pi/2$, so $G(x) \to 1$. As $x \to \infty$, $\arctan([m/x]^\alpha) \to 0$, so $G(x) \to 0$.

(ii) $G(m) = \dfrac{2}{\pi} \arctan 1 = \dfrac{2}{\pi} \times \dfrac{\pi}{4} = \dfrac{1}{2}$.

(iii) Since

$$\lim_{x\to\infty} \frac{\arctan([m/x]^\alpha)}{[m/x]^\alpha} = \lim_{y\to 0} \frac{\arctan y}{y} = 1,$$

$\lim_{x\to\infty} G(x)/[m/x]^\alpha = 2/\pi$. Therefore $x^\alpha G(x) \to 2m^\alpha/\pi$ as $x \to \infty$.

(iv)

$$f(x) = -G'(x) = -\frac{2}{\pi}\left(-\frac{\alpha}{x}\left[\frac{m}{x}\right]^{\alpha}\right)\Bigg/\left(1+\left[\frac{m}{x}\right]^{2\alpha}\right)$$

Simplifying,

$$f(x) = \frac{2\alpha/\pi}{m^{-\alpha}x^{1+\alpha} + m^{\alpha}x^{1-\alpha}}. \qquad (*)$$

Since m and α are positive numbers, $f(x) > 0$ for all $x > 0$. Also, since $\alpha > 1$, $x^{1+\alpha} \to 0$ and $x^{1-\alpha} \to \infty$ as $x \to 0$, while $x^{1+\alpha} \to \infty$ and $x^{1-\alpha} \to 0$ as $x \to \infty$. Thus the denominator on the right-hand side of $(*)$ becomes very large both as $x \to 0$ and as $x \to \infty$. It follows that $\lim_{x\to 0} f(x) = \lim_{x\to\infty} f(x) = 0$.

(v) Let $h(x) = \alpha/f(x) = \frac{1}{2}\pi(m^{-\alpha}x^{1+\alpha} + m^{\alpha}x^{1-\alpha})$. Then

$$h'(x) = \tfrac{1}{2}\pi\left((1+\alpha)(x/m)^{\alpha} + (1-\alpha)(x/m)^{-\alpha}\right).$$

Since $\alpha > 1$, we may define the positive number

$$\beta = \left[\frac{\alpha-1}{\alpha+1}\right]^{1/(2\alpha)}.$$

Then $\beta < 1$ and $h'(\beta m) = 0$. Also, $h'(x) < 0$ if $0 < x < \beta m$ and $h'(x) > 0$ if $x > \beta m$. Since $f'(x)$ always has the opposite sign to $h'(x)$, the required properties are satisfied by $x^* = \beta m$. Since $\beta < 1$, $x^* < m$.

25 COMPLEX NUMBERS

25–1. $1 + i = \sqrt{2}(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4})$ and $1 - i = \sqrt{2}(\cos\frac{\pi}{4} - i\sin\frac{\pi}{4})$. By De Moivre's theorem and its corollary, $(1+i)^t = 2^{t/2}(\cos\frac{\pi t}{4} + i\sin\frac{\pi t}{4})$ and $(1-i)^t = 2^{t/2}(\cos\frac{\pi t}{4} - i\sin\frac{\pi t}{4})$. Hence

$$y_t = 2^{t/2}\left((A+B)\cos(\pi t/4) + i(A-B)\sin(\pi t/4)\right). \qquad (*)$$

(i) If A and B are real, the real part of y_t is $(A+B)2^{t/2}\cos\frac{\pi t}{4}$ and the imaginary part of y_t is $i(A-B)2^{t/2}\sin\frac{\pi t}{4}$.

(ii) If A and B are conjugates, $A+B$ and $i(A-B)$ are real numbers, so y_t is real. Conversely, if y_t is real for all t, then $A+B$ and $i(A-B)$ are real numbers, say C and D respectively. Then $2iA = iC + D$, whence $A = \frac{1}{2}(C - iD)$ and

$$B = C - A = \tfrac{1}{2}(C + iD) = \bar{A}.$$

(iii) Here $2 = y_0 = A+B$ and $5 = y_1 = A+B+i(A-B)$. Hence $A+B = 2$ and $i(A-B) = 3$. From $(*)$,

$$y_t = 2^{t/2}\left(2\cos(\pi t/4) + 3\sin(\pi t/4)\right) \quad \text{for all } t.$$

To find A and B, proceed as in (ii) with $C = 2$ and $D = 3$: $A = 1 - \frac{3}{2}i$ and $B = 1 + \frac{3}{2}i$.

25–2. Since $e^{(-2+5i)t} = e^{-2t}(\cos 5t + i\sin 5t)$ and $e^{(-2-5i)t} = e^{-2t}(\cos 5t - i\sin 5t)$,

$$y = e^{-2t}((A+B)\cos 5t + i(A-B)\sin 5t).$$

(i) If A and B are real, the real part of y is $(A+B)e^{-2t}\cos 5t$ and the imaginary part of y is $i(A-B)e^{-2t}\cos 5t$.

(ii) Similar to Problem 25–1, part (ii).

(iii) Since $y = 0$ when $t = 0$, $A + B = 0$. Also,

$$dy/dt = -2y - 5e^{-2t}((A + B)\sin 5t - i(A - B)\cos 5t):$$

setting $t = 0$ we have $1 = 0 - 5(0 - i[A - B])$. Thus $B = -A$ and $A - B = -i/5$. It follows that $A = -i/10$, $B = i/10$ and

$$y = \frac{\sin 5t}{5e^{2t}}.$$

25–3. (i) Let $f(z) = z^3 - 2z^2 - 2z - 3$. Then $f(3) = 0$, so $z - 3$ is a factor of $f(z)$. By inspection, $f(z) = (z - 3)(z^2 + \lambda z + 1)$ for some λ; equating coefficients of z (or z^2), we see that $\lambda = 1$. It follows that $f(z) = (z-3)(z^2+z+1) = (z-3)(z-u)(z-v)$ where u, v are the roots of $z^2 + z + 1 = 0$. By the quadratic formula, we may set $u = \frac{1}{2}(-1 + \sqrt{3})$, $v = \bar{u}$.

(ii) Let $g(z) = z^3 - 4z^2 + 14z - 20$; also let $u = 1 - 3i$. Then $u^2 = -8 - 6i$ and $u^3 = 1 - 9i - 27 + 27i = -26 + 18i$, so

$$g(u) = -26 + 18i + 32 + 24i - 6 - 42i = 0.$$

Thus u is a root; since the polynomial g has real coefficients, $\bar{u} = 1 + 3i$ is also a root. Denoting the third root by v, we see that

$$g(z) = (z - u)(z - \bar{u})(z - v) = ([z - 1]^2 - [3i]^2)(z - v).$$

Thus $z^3 - 4z^2 + 14z - 20 = (z^2 - 2z + 10)(z - v)$ for all z. Putting $z = 0$ we see that $-20 = -10v$, so $v = 2$.

25–4. (i) $\displaystyle\int_0^{\pi/2} e^{it}dt = \Big[e^{it}/i\Big]_0^{\pi/2} = -i(e^{i\pi/2} - 1) = i(1 - i) = 1 + i.$

(ii) Integrating by parts,

$$\int_0^{\pi/2} te^{it}dt = \Big[(t/i)e^{it}\Big]_0^{\pi/2} - (1/i)\int_0^{\pi/2} e^{it}dt.$$

Hence, using the result of (i),

$$\int_0^{\pi/2} te^{it}dt = (-i\pi/2)e^{i\pi/2} + i(1 + i) = \tfrac{1}{2}(\pi - 2) + i.$$

Again by integration by parts,

$$\int_0^{\pi/2} t^2e^{it}dt = \Big[(t^2/i)e^{it}\Big]_0^{\pi/2} - (2/i)\int_0^{\pi/2} te^{it}dt.$$

Hence, using the result above,

$$\int_0^{\pi/2} t^2e^{it}dt = (-i\pi^2/4)e^{i\pi/2} + 2i(\tfrac{1}{2}[\pi - 2] + i) = \tfrac{1}{4}(\pi^2 - 8) + i(\pi - 2).$$

(iii) Denote the integrals by I and J. Then iI is the imaginary part of the first integral in (ii), so $I = 1$. J is the real part of the second integral in (ii), so $J = \frac{1}{4}(\pi^2 - 8)$.

26 FURTHER DYNAMICS

26–1. Denote the differential equations by (i) and (ii) and observe that (ii) may be rearranged as follows:

$$\text{(ii}'\text{)}\quad \frac{dz}{dt} + 2z = y.$$

Adding $\dfrac{d}{dt}$(i) to $2\times$(i) and using (ii$'$),

$$\frac{d^2y}{dt^2} + 2\frac{dy}{dt} = 4\left(\frac{dy}{dt} + 2y\right) - 5y.$$

Rearranging,

$$\frac{d^2y}{dt^2} - 2\frac{dy}{dt} - 3y = 0.$$

The characteristic equation is $x^2 - 2x - 3 = 0$, which has roots 3 and -1. The general solution for y is therefore

$$y = Ae^{3t} + Be^{-t},$$

where A and B are constants. Substituting in (i) gives

$$z = \tfrac{1}{5}Ae^{3t} + Be^{-t}.$$

(a) Here $A + B = 1$ and $(A/5) + B = -3$; hence $A = 5,\ B = -4$.

(b) When $t = 0$, $y = 1$ and $y - 2z = 7$, so $z = -3$. Therefore the boundary conditions are equivalent to those of (a), and we have the same solution.

(c) When $t = 0$, $y = 1$ and $4y - 5z = -1$, so $z = 1$. Therefore $A + B = (A/5) + B = 1$; hence $A = 0,\ B = 1$.

(d) Since $\lim_{t\to\infty} e^{-t} = 0$, $y \approx Ae^{3t}$ when t is large and positive. Hence the boundary condition $\lim_{t\to\infty} y = 0$ implies that $A = 0$. Since $A + B$ is again equal to 1, the solution is as in (c).

The key feature which leads to (d) completely determining the solution is that the characteristic equation has real roots of opposite sign. In this case, the condition $y \to 0$ as $t \to \infty$ means that the coefficient of the component of the solution corresponding to the positive root must be zero.

26–2. (i) (a) Integrating, $dy/dt = -\frac{1}{6}(2t-1)^3 + A$. Integrating again,

$$y = -\frac{1}{48}(2t-1)^4 + At + B.$$

Since $dy/dt = 0$ when $t = 0$, $A = -1/6$.
From the differential equation, $d^2y/dt^2 < 0$ for all $t \neq \frac{1}{2}$ and $d^2y/dt^2 = 0$ if $t = \frac{1}{2}$. Hence the function is concave, with its only critical point at $t = 0$. Therefore the graph is $\bigcap$-shaped with vertex at $(0, 1)$.

(b) The equation can be written as $d^2y/dt^2 = 3(t-1)^2 + 1$. Integrating twice as above gives

$$y = \tfrac{1}{4}(t-1)^4 + At + B.$$

Imposing the initial conditions, we obtain the required solution:

$$y = \tfrac{1}{4}\left((t-1)^4 + 4t + 3\right).$$

From the differential equation, $d^2y/dt^2 > 0$ for all t. Hence the function is strictly convex, with its only critical point at $t = 0$. Therefore the graph is U-shaped with vertex at $(0, 1)$.

(c) Integrating twice,

$$y = \tfrac{1}{3}t^3 - \tfrac{1}{2}t^2 + At + B.$$

Imposing the initial conditions, we obtain the required solution:

$$y = \tfrac{1}{3}t^3 - \tfrac{1}{2}t^2 + 1.$$

From the differential equation, $d^2y/dt^2 < 0$ for $t < \frac{1}{2}$ and $d^2y/dt^2 > 0$ for $t > \frac{1}{2}$. Hence the function is strictly concave for $t < \frac{1}{2}$ and strictly convex for $t > \frac{1}{2}$. The critical points occur where t is a root of the equation $t^2 - t = 0$, i.e. where t is 0 or 1; there is therefore a local maximum at $(0, 1)$ and a local minimum at $(1, \frac{5}{6})$.

(ii) The characteristic equation is $p^2 + 2ap + (a^2 + b^2) = 0$, which has roots $-a \pm ib$. The general solution of the differential equation is therefore

$$y = Ce^{-ax}\cos(bx + \theta),$$

where C and θ are arbitrary constants.

From the differential equation, d^2y/dx^2 has the opposite sign to y if $dy/dx = 0$. Therefore, any critical point (X, Y) such that $Y > 0$ is a maximum, and any critical point (X, Y) such that $Y < 0$ is a minimum.

The function $y = e^{-3x}\sin 4x$ of Problem 24–1 is the special case of the general solution with $a = 3$, $b = 4$ and the constants C and θ put equal to 1 and $-\pi/2$ respectively. Therefore, the function satisfies the differential equation and has the above property concerning critical values. Similarly, $y = e^{3x}\sin 4x$ has the same property.

26–3. First look for a particular solution of the form $y = At+B$. Substituting this into the differential equation gives

$$0 + bA + c(At + B) = kt + \ell.$$

Hence $cA = k$ and $bA + cB = \ell$, so $A = c^{-1}k$ and $B = c^{-2}(\ell c - bk)$. The complementary solution is oscillatory (**O**) when the roots of the characteristic equation are complex, i.e. when $b^2 < 4c$. Otherwise the complementary solution is non-oscillatory (**N**). The complementary solution tends to 0 as $t \to \infty$ (**S**) when both roots of the characteristic equation have negative real parts, the criterion for which is obtained in the text as $b > 0$, $c > 0$. Otherwise the complementary solution does not tend to 0 as $t \to \infty$ (**U**).

We may therefore classify the possible forms taken by the general solution as follows:

SO The general solution oscillates about the particular solution and tends to the particular solution as $t \to \infty$.

SN The general solution is non-oscillatory and tends to the particular solution as $t \to \infty$.

UO The general solution oscillates about the particular solution but does not tend to the particular solution as $t \to \infty$.

UN The general solution is non-oscillatory and and does not tend to the particular solution as $t \to \infty$.

26–4. The discrete-time analogue of the differential equation is

$$\Delta^2 y_t + b\Delta y_t + cy_t = u,$$

which may be written as

$$(y_{t+2} - 2y_{t+1} + y_t) + b(y_{t+1} - y_t) + cy_t = u,$$

or more simply as

$$y_{t+2} + fy_{t+1} + gy_t = u$$

where $f = b - 2$, $g = 1 - b + c$. For the differential equation, the stationary solution occurs if $d^2y/dt^2 = dy/dt = 0$ and is therefore $y = u/c$. For the difference equation, the stationary solution occurs if $\Delta^2 y_t = \Delta y_t = 0$ and is therefore $y_t = u/c$.

The differential equation exhibits oscillatory behaviour if its characteristic equation has complex roots, i.e. if $b^2 < 4c$. The difference equation exhibits oscillatory behaviour if *its* characteristic equation has complex roots, i.e. if $f^2 < 4g$: this inequality may be written

$$(b-2)^2 < 4(1-b+c)$$

and therefore reduces to $b^2 < 4c$.

For the differential equation, the condition for the stationary solution to be stable has been obtained in the text as $b > 0$, $c > 0$. For the difference equation, the condition for the stationary solution to be stable is that the roots of the characteristic equation are are both < 1 in absolute value (or modulus, if the roots are complex). In the case of real roots, it is therefore necessary that $x^2 + fx + g > 0$ at $x = \pm 1$. This ensures that one of the following three cases occurs: (a) both roots between -1 and 1, (b) both roots < -1, (c) both roots > 1. But the product of the roots is g, so if we assume that $g < 1$ then cases (b) and (c) are eliminated and we are left with (a). In the case of complex roots, we have $x^2 + fx + g > 0$ for all x, and in particular at $x = \pm 1$. Also the roots are complex conjugates, so the product of the roots is r^2 where r is the common modulus: to ensure that $r < 1$ we must therefore impose the condition $g < 1$.

To summarise, the criterion ensuring stability in the cases of both real and complex roots is

$$1 + f + g > 0, \quad 1 - f + g > 0, \quad g < 1.$$

In terms of b and c these three conditions may be written respectively as $c > 0$, $c > 2b-4$, $c < b$ and therefore reduce to the chain of inequalities

$$0 < c < b < 2 + \tfrac{1}{2}c.$$

As in the first-order case, we note that the qualitative behaviour of the discrete-time analogue is not necessarily the same as that of the differential equation for the same parameter values.

27 EIGENVALUES AND EIGENVECTORS

27–1. (i) The characteristic polynomial of $\mathbf{A}$ is $\lambda^2 - \lambda - 20$, so the eigenvalues are -4 and 5.
$\mathbf{Ax} = -4\mathbf{x}$ if and only if

$$2x_1 + 6x_2 = -4x_1, \quad 3x_1 - x_2 = -4x_2.$$

Each of these equations simplifies to $x_1 + x_2 = 0$: the eigenvectors corresponding to the eigenvalue -4 are the non-zero multiples of $[1 \quad -1]^{\mathrm{T}}$.
$\mathbf{Ax} = 5\mathbf{x}$ if and only if

$$2x_1 + 6x_2 = 5x_1, \quad 3x_1 - x_2 = 5x_2.$$

Each of these equations simplifies to $x_1 = 2x_2$: the eigenvectors corresponding to the eigenvalue 5 are the non-zero multiples of $[2 \;\; 1]^{\mathrm{T}}$.

The characteristic polynomial of $\mathbf{A}^{\mathrm{T}}$ is, as for $\mathbf{A}$, $\lambda^2 - \lambda - 20$. Thus the eigenvalues are also -4 and 5.

$\mathbf{A}^{\mathrm{T}}\mathbf{x} = -4\mathbf{x}$ if and only if

$$2x_1 + 3x_2 = -4x_1, \quad 6x_1 - x_2 = -4x_2.$$

Each of these equations simplifies to $2x_1 + x_2 = 0$: the eigenvectors corresponding to the eigenvalue -4 are the non-zero multiples of $[1 \;\; -2]^{\mathrm{T}}$.

$\mathbf{A}^{\mathrm{T}}\mathbf{x} = 5\mathbf{x}$ if and only if

$$2x_1 + 3x_2 = 5x_1, \quad 6x_1 - x_2 = 5x_2.$$

Each of these equations simplifies to $x_1 = x_2$: the eigenvectors corresponding to the eigenvalue 5 are the non-zero multiples of $[1 \;\; 1]^{\mathrm{T}}$.

(ii) (a) True for all $\mathbf{A}$. Let λ be any scalar: since $\lambda\mathbf{I} - \mathbf{A}^{\mathrm{T}}$ is the same matrix as $(\lambda\mathbf{I} - \mathbf{A})^{\mathrm{T}}$, it has the same determinant as $\lambda\mathbf{I} - \mathbf{A}$.

(b) True for all $\mathbf{A}$. By (a), the eigenvalues of $\mathbf{A}$ and $\mathbf{A}^{\mathrm{T}}$ are the roots of the same polynomial equation.

(c) False, in general. Suppose λ is an eigenvalue of $\mathbf{A}$ and hence, by (b), of $\mathbf{A}^{\mathrm{T}}$. The systems of equations $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ and $\mathbf{A}^{\mathrm{T}}\mathbf{x} = \lambda\mathbf{x}$ are in general different and hence the corresponding eigenvectors are not, in general, the same. See (i) above for an example.

27–2. (i) $\mathbf{z}^{\mathrm{H}}\mathbf{w} = \bar{z}_1 w_1 + \ldots + \bar{z}_1 w_1$ and $\mathbf{w}^{\mathrm{H}}\mathbf{z} = \overline{w}_1 z_1 + \ldots + \overline{w}_1 z_1$. In particular,

$$\mathbf{z}^{\mathrm{H}}\mathbf{z} = \bar{z}_1 z_1 + \ldots + \bar{z}_1 z_1 = |z_1|^2 + \ldots + |z_n|^2.$$

For $j = 1, \ldots, n$, $\bar{z}_j w_j$ is equal to $w_j \bar{z}_j$, whose conjugate is $\overline{w}_j z_j$. Therefore $\mathbf{z}^{\mathrm{H}}\mathbf{w}$ and $\mathbf{w}^{\mathrm{H}}\mathbf{z}$ are complex conjugates. For each $j = 1, \ldots, n$, $|z_j| \geq 0$, with equality if $z_j \neq 0$. Therefore $\mathbf{z}^{\mathrm{H}}\mathbf{z} > 0$ provided at least one of $z_1, \ldots, z_n$ is non-zero, i.e. provided $\mathbf{z} \neq \mathbf{0}$.

(ii) Since $\bar{\mathbf{A}}^{\mathrm{T}}$ has the same diagonal entries as $\bar{\mathbf{A}}$, the diagonal entries of $\mathbf{A}^{\mathrm{H}}$ are the complex conjugates of those of $\mathbf{A}$. Thus, if $\mathbf{A}$ is Hermitian, the diagonal entries must be equal to their conjugates and hence must be real. When all the entries are real numbers, $\bar{\mathbf{A}} = \mathbf{A}$ and the condition for $\mathbf{A}$ to be Hermitian reduces to $\mathbf{A}^{\mathrm{T}} = \mathbf{A}$.

(iii) The general forms are

$$\begin{bmatrix} a & u \\ \bar{u} & b \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} a & u & v \\ \bar{u} & b & w \\ \bar{v} & \overline{w} & c \end{bmatrix},$$

where a, b, c are real and u, v, w are complex.

(iv) Let $\mathbf{A}$ be the 2×2 Hermitian matrix given in (iii) and let $\mathbf{z} \in \mathbb{C}^2$. Multiplying out the expression for $\mathbf{z}^{\mathrm{H}}\mathbf{A}\mathbf{z}$, we have

$$\mathbf{z}^{\mathrm{H}}\mathbf{A}\mathbf{z} = P + Q,$$

where $P = a\bar{z}_1 z_1 + b\bar{z}_2 z_2$ and $Q = u\bar{z}_1 z_2 + \bar{u}\bar{z}_2 z_1$. Now a is real and $\bar{z}_1 z_1 = |z_1|^2$, so $a\bar{z}_1 z_1$ is real. A similar argument using the fact that b is real shows that $b\bar{z}_2 z_2$ is real, and hence that P is real. Also, since $u\bar{z}_1 z_2$ and $\bar{u} z_1 \bar{z}_2$ are complex conjugates, their sum Q is real. Hence $P + Q$ is real, as required.

Now let $\mathbf{A}$ be the 3×3 Hermitian matrix given in (iii) and let $\mathbf{z} \in \mathbb{C}^3$. Reasoning as in the 2×2 case, we may write $\mathbf{z}^{\mathrm{H}}\mathbf{A}\mathbf{z} = P + Q$, where P is the real number

$$a|z_1|^2 + b|z_2|^2 + c|z_3|^2$$

and Q is twice the real part of

$$u\bar{z}_1 z_2 + v\bar{z}_1 z_3 + w\bar{z}_2 z_3.$$

Hence $P + Q$ is real, as required. The general case is similar.

(v) Suppose λ is an eigenvalue. Then there is a non-zero vector $\mathbf{z}$ such that $\mathbf{A}\mathbf{z} = \lambda\mathbf{z}$. Therefore

$$\mathbf{z}^{\mathrm{H}}\mathbf{A}\mathbf{z} = \mathbf{z}^{\mathrm{H}}\lambda\mathbf{z} = \lambda\mathbf{z}^{\mathrm{H}}\mathbf{z}.$$

Since $\mathbf{z} \neq \mathbf{0}$, $\mathbf{z}^{\mathrm{H}}\mathbf{z} > 0$ by (i), so we may define the positive real number $p = (\mathbf{z}^{\mathrm{H}}\mathbf{z})^{-1}$. But then $\lambda = p\mathbf{z}^{\mathrm{H}}\mathbf{A}\mathbf{z}$, which is real by (iv).

(vi) The $(1,1)$ entry of $\mathbf{A}$ is the value of $\mathbf{z}^{\mathrm{H}}\mathbf{A}\mathbf{z}$ when $\mathbf{z}$ is the first column of the identity matrix $\mathbf{I}$; the $(2,2)$ entry of $\mathbf{A}$ is $\mathbf{z}^{\mathrm{H}}\mathbf{A}\mathbf{z}$ when $\mathbf{z}$ is the second column of $\mathbf{I}$; and so on. Hence by (iv), the diagonal entries of a Hermitian matrix are real.

Let

$$\mathbf{A} = \begin{bmatrix} 1 & w \\ \overline{w} & 1 \end{bmatrix}, \quad \mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}.$$

Multiplying out as in (iv), we have

$$\mathbf{z}^{\mathrm{H}}\mathbf{A}\mathbf{z} = |z_1|^2 + |z_2|^2 + Q,$$

where the real number q is twice the real part of $w\bar{z}_1 z_2$. On the other hand,

$$|z_1 + wz_2|^2 = (\bar{z}_1 + \overline{w}\bar{z}_2)(z_1 + wz_2) = |z_1|^2 + Q + |w|^2|z_2|^2.$$

Subtracting and rearranging,

$$\mathbf{z}^{\mathrm{H}}\mathbf{A}\mathbf{z} = |z_1 + wz_2|^2 + (1 - |w|^2)|z_2|^2.$$

This is clearly positive if $|w|^2 < 1$ and at least one of z_1 and z_2 is not zero; while if $|w|^2 \leq 1$, we can make $\mathbf{z}^{\mathrm{H}}\mathbf{A}\mathbf{z}$ non-positive by setting $z_1 = w$, $z_2 = -1$. Thus $\mathbf{A}$ is positive definite if and only if $|w|^2 < 1$, which happens if and only if $\det \mathbf{A} > 0$.

The general result for $n = 2$ is that a Hermitian matrix is positive definite if and only if its diagonal entries and its determinant are all positive. The *really* general result is that a Hermitian matrix is positive definite if and only if all its principal minors are positive. [As in the case of real symmetric matrices, it is also true that a Hermitian matrix is positive definite if and only if all its leading principal minors are positive.]

(vii) When all the entries of a matrix are real numbers, the Hermitian transpose reduces to the ordinary transpose. So a unitary matrix whose entries are all real numbers has as its inverse its transpose and hence is the same thing as an orthogonal matrix.

Denoting the given matrix by $\mathbf{S}$,

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix}$$

By straightforward matrix multiplication, $\mathbf{S}^{\mathrm{H}}\mathbf{S} = \mathbf{I}$, so $\mathbf{S}^{\mathrm{H}} = \mathbf{S}^{-1}$ as required.

27–3. (i) By inspection, the given matrix is equal to its transpose. By straightforward matrix multiplication, we see that the square of the given matrix is equal to the matrix itself.

(ii) Let λ is an eigenvalue of a projection matrix $\mathbf{P}$, and let $\mathbf{x}$ be a corresponding eigenvector: $\mathbf{x} \neq \mathbf{0}$ and $\mathbf{Px} = \lambda\mathbf{x}$. Since $\mathbf{P} = \mathbf{P}^2$, $\mathbf{Px} = \mathbf{P}(\mathbf{Px})$: hence

$$\lambda\mathbf{x} = \mathbf{P}(\lambda\mathbf{x}) = \lambda\mathbf{Px} = \lambda^2\mathbf{x}.$$

Since $\mathbf{x} \neq \mathbf{0}$, it follows that $\lambda = \lambda^2$, so λ is either 1 or 0.

(iii) By (ii), the characteristic polynomial of $\mathbf{P}$ is $(\lambda - 1)^r\lambda^{n-r}$ for some r.

Let $\mathbf{D}$ be the diagonal matrix whose first r diagonal entries are equal to 1 and whose remaining diagonal entries are all zero. Hence, by Theorem 1 of Section 25.3, there is an orthogonal matrix $\mathbf{S}$ such that $\mathbf{S}^{\mathrm{T}}\mathbf{PS} = \mathbf{D}$. Since $\mathbf{S}$ is an orthogonal matrix, $\mathbf{SS}^{\mathrm{T}} = \mathbf{I}$, whence $\mathbf{P} = \mathbf{SDS}^{\mathrm{T}}$.

Partition $\mathbf{S}$ as $(\mathbf{ZY})$, where $\mathbf{Z}$ consists of the first r columns. Then the equation $\mathbf{SS}^{\mathrm{T}} = \mathbf{I}_n$ may be written

$$\begin{bmatrix} \mathbf{Z}^{\mathrm{T}}\mathbf{Z} & \mathbf{Z}^{\mathrm{T}}\mathbf{Y} \\ \mathbf{Y}^{\mathrm{T}}\mathbf{Z} & \mathbf{Y}^{\mathrm{T}}\mathbf{Y} \end{bmatrix} \begin{bmatrix} \mathbf{I}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_{n-r} \end{bmatrix}.$$

In particular, $\mathbf{Z}^{\mathrm{T}}\mathbf{Z} = \mathbf{I}_r$. Also, the equation $\mathbf{P} = \mathbf{SDS}^{\mathrm{T}}$ may be written

$$\mathbf{P} = \begin{bmatrix} \mathbf{Z} & \mathbf{Y} \end{bmatrix} \begin{bmatrix} \mathbf{I}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix} \begin{bmatrix} \mathbf{Z}^{\mathrm{T}} \\ \mathbf{Y}^{\mathrm{T}} \end{bmatrix}.$$

Hence

$$\mathbf{P} = \begin{bmatrix} \mathbf{Z} & \mathbf{Y} \end{bmatrix} \begin{bmatrix} \mathbf{Z}^{\mathrm{T}} \\ \mathbf{O} \end{bmatrix} = \mathbf{ZZ}^{\mathrm{T}},$$

as required.

(iv) Let $\mathbf{Z}$ be an $n\times r$ matrix such that $\mathbf{Z}^{\mathrm{T}}\mathbf{Z} = \mathbf{I}_r$. Then $\mathbf{ZZ}^{\mathrm{T}}$ is $n\times n$ and

$$(\mathbf{ZZ}^{\mathrm{T}})^{\mathrm{T}} = \mathbf{Z}^{\mathrm{TT}}\mathbf{Z}^{\mathrm{T}} = \mathbf{ZZ}^{\mathrm{T}}.$$

Also

$$(\mathbf{ZZ}^{\mathrm{T}})^2 = \mathbf{ZZ}^{\mathrm{T}}\mathbf{ZZ}^{\mathrm{T}} = \mathbf{ZI}_r\mathbf{Z}^{\mathrm{T}}.$$

Hence $\mathbf{ZZ}^{\mathrm{T}}$ is a projection matrix.

27–4. (i) Let the characteristic polynomial of $\mathbf{A}$ be $f(\lambda)$. Then

$$f(\lambda) = \begin{vmatrix} \lambda{-}2 & -1 & 2 \\ -1 & \lambda{-}2 & 2 \\ 2 & 2 & \lambda - 5 \end{vmatrix}.$$

Subtracting the first row from the second,

$$f(\lambda) = \begin{vmatrix} \lambda - 2 & -1 & 2 \\ 1 - \lambda & \lambda{-}1 & 0 \\ 2 & 2 & \lambda - 5 \end{vmatrix}.$$

It is now easy to expand by the second row:

$$\begin{aligned} f(\lambda) &= (\lambda - 1)\begin{vmatrix} -1 & 2 \\ 2 & \lambda - 5 \end{vmatrix} + (\lambda - 1)\begin{vmatrix} \lambda - 2 & 2 \\ 2 & \lambda - 5 \end{vmatrix} \\ &= (\lambda - 1)(1 - \lambda + \lambda^2 - 7\lambda + 10 - 4) \\ &= (\lambda - 1)(\lambda^2 - 8\lambda + 7) \\ &= (\lambda - 1)^2(\lambda - 7). \end{aligned}$$

(ii) $\mathbf{Ax} = \mathbf{x}$ if and only if $(\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$. By inspection, each row of $\mathbf{I} - \mathbf{A}$ is a multiple of $(1\ \ 1\ \ -2)$. Thus $\mathbf{Ax} = \mathbf{x}$ if and only if $x_1 + x_2 = 2x_3$. We may therefore choose $\mathbf{x} = [1\ \ -1\ \ 0]^{\mathrm{T}}$; $\mathbf{y}$ is then a vector such that $y_1 + y_2 = 2y_3$ and $y_1 - y_2 = 0$. We may therefore choose $\mathbf{y} = [1\ \ 1\ \ 1]^{\mathrm{T}}$.

(iii) $\mathbf{Az} = 7\mathbf{z}$ if and only if $(7\mathbf{I} - \mathbf{A})\mathbf{z} = \mathbf{0}$. By inspection,

$$7\mathbf{I} - \mathbf{A} = \begin{bmatrix} 5 & -1 & 2 \\ -1 & 5 & 2 \\ 2 & 2 & 2 \end{bmatrix}.$$

Solving the system by Gaussian elimination, we see that the solution consists of all vectors for which $z_1 = z_2 = -\frac{1}{2}z_3$. We may therefore choose $\mathbf{z} = [1\ \ 1\ \ -2]^{\mathrm{T}}$.

(iv) Let the vectors $\mathbf{x}, \mathbf{y}, \mathbf{z}$ be as in (ii) and (iii). It is easy to see that these vectors are linearly independent. Arguing as in the proof of Proposition 1 in Section 25.1, we set $\mathbf{D} = \operatorname{diag}(1, 1, 7)$ and let $\mathbf{S}$ be a matrix whose columns are scalar multiples of $\mathbf{x}, \mathbf{y}, \mathbf{z}$. To ensure that $\mathbf{S}^{\mathrm{T}}\mathbf{S} = \mathbf{I}$, each column of $\mathbf{S}$ must have length 1; we therefore define the columns of $\mathbf{S}$ to be

$$(1+1)^{-1/2}\mathbf{x},\ \ (1+1+1)^{-1/2}\mathbf{y},\ \ (1+1+4)^{-1/2}\mathbf{z}.$$

Then

$$\mathbf{S} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6} \\ -1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6} \\ 0 & 1/\sqrt{3} & -2/\sqrt{6} \end{bmatrix}.$$

28 DYNAMIC SYSTEMS

28–1. (i) Expanding the determinant by its first row,

$$\det(\lambda\mathbf{I} - \mathbf{A}) = (\lambda - 1)(\lambda^2 - 3\lambda + 1) - (\lambda - 1) = (\lambda - 1)(\lambda^2 - 3\lambda).$$

Therefore, the characteristic polynomial of $\mathbf{A}$ is $\lambda(\lambda - 1)(\lambda - 3)$, and the eigenvalues are 0, 1 and 3. It is easy to show that

$$\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

are corresponding eigenvectors. Therefore, the general solution is

$$\mathbf{y}(t) = 0^t c_1\mathbf{u} + c_2\mathbf{v} + 3^t c_3\mathbf{w}.$$

Here and below, 0^0 should be interpreted as 1.

(ii) Using the given initial condition and our convention that $0^0 = 1$,

$$c_1 + c_2 + c_3 = 1, \quad c_1 - 2c_3 = 2, \quad c_1 - c_2 + c_3 = 4.$$

Solving these equations simultaneously gives $c_1 = 7/3$, $c_2 = -3/2$, $c_3 = 1/6$. The solution is therefore

$$\mathbf{y}(t) = \alpha_t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{3}{2}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + \frac{3^{t-1}}{2}\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix},$$

where $\alpha_0 = 7/3$ and $\alpha_t = 0$ for all $t > 0$.

(iii) Since 1 is an eigenvalue, $\mathbf{I}-\mathbf{A}$ is not invertible and we cannot use the formula $\mathbf{x}^* = (\mathbf{I}-\mathbf{A})^{-1}\mathbf{b}$ for the particular solution. However, all we need for a constant particular solution is *some* vector $\mathbf{x}^*$ such that $(\mathbf{I}-\mathbf{A})\mathbf{x}^* = \mathbf{b}$, and it is easy to see that, for example, $\mathbf{x}^* = [1\ 1\ 0]^{\mathrm{T}}$ does the trick. The general solution is

$$\mathbf{x}(t) = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 0^t c_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + 3^t c_3 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}.$$

(iv) In this case, the system of linear equations $(\mathbf{I}-\mathbf{A})\mathbf{x} = \mathbf{b}$ has no solution, so the system of difference equations does not have a constant particular solution. We therefore look for a solution of the form $\mathbf{x}^*(t) = t\mathbf{p} + \mathbf{q}$, where $\mathbf{p}$ and $\mathbf{q}$ are constant vectors. Then

$$(t+1)\mathbf{p} + \mathbf{q} = t\mathbf{A}\mathbf{p} + \mathbf{A}\mathbf{q} + \mathbf{b}$$

for all t. Equating coefficients of t, we have $\mathbf{A}\mathbf{p} = \mathbf{p}$; equating constant terms, we have $\mathbf{A}\mathbf{q} + \mathbf{b} = \mathbf{p} + \mathbf{q}$. Hence $\mathbf{p}$ is an eigenvector of $\mathbf{A}$ corresponding to the eigenvalue 1; we may therefore set $\mathbf{p} = \beta\mathbf{v}$, where $\mathbf{v} = [1\ 0\ -1]^{\mathrm{T}}$ as in (i) and β is a constant. It follows that

$$(\mathbf{I}-\mathbf{A})\mathbf{q} = \mathbf{b} - \beta\mathbf{v} = \begin{bmatrix} 1-\beta \\ 0 \\ 2+\beta \end{bmatrix}.$$

Hence $q_2 = 1-\beta = \beta+2$ and $q_1 + q_3 = q_2$. Therefore $\beta = -\frac{1}{2}$, $q_2 = \frac{3}{2}$ and we are at liberty to let $q_1 = \frac{3}{2}$, $q_3 = 0$. Our particular solution is

$$\mathbf{x}^*(t) = \frac{3}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \frac{t}{2}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix},$$

and the general solution is

$$\mathbf{x}(t) = \frac{3}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 0^t c_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \left(c_2 - \frac{t}{2}\right)\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + 3^t c_3 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}.$$

28–2. (i) The characteristic polynomial of $\mathbf{A}$ is $(\lambda-1)^2 - \frac{4}{9}$, so the eigenvalues are $\frac{1}{3}$ and $\frac{5}{3}$. It is easy to show that $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ are corresponding eigenvectors.

Thus the general solution is

$$\mathbf{x}(t) = \mathbf{x}^*(t) + \frac{c_1}{3^t}\begin{bmatrix} -1 \\ 2 \end{bmatrix} + \frac{5^t c_2}{3^t}\begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

where $\mathbf{x}^*(t)$ is a particular solution. To find this, try $\mathbf{x}^*(t) = t\mathbf{p} + \mathbf{q}$, where $\mathbf{p}$ and $\mathbf{q}$ are constant vectors. Then

$$(t+1)\mathbf{p} + \mathbf{q} = t\mathbf{A}\mathbf{p} + \mathbf{A}\mathbf{q} + \mathbf{b}(t)$$

for all t. Equating coefficients of t and then equating constant terms, we have

$$(\mathbf{I}-\mathbf{A})\mathbf{p} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad (\mathbf{I}-\mathbf{A})\mathbf{q} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} - \mathbf{p}.$$

Hence $p_2 = -3$, $p_1 = -9/4$, $q_2 = -3(2 + \frac{9}{4}) = -51/4$ and $q_1 = -\frac{3}{4}(0+3) = -9/4$. Thus our particular solution is

$$\mathbf{x}^*(t) = -\frac{3t}{4}\begin{bmatrix} 3 \\ 4 \end{bmatrix} - \frac{3}{4}\begin{bmatrix} 3 \\ 17 \end{bmatrix}$$

and the general solution is

$$\mathbf{x}(t) = -\frac{3t}{4}\begin{bmatrix} 3 \\ 4 \end{bmatrix} - \frac{3}{4}\begin{bmatrix} 3 \\ 17 \end{bmatrix} + \frac{c_1}{3^t}\begin{bmatrix} -1 \\ 2 \end{bmatrix} + \frac{5^t c_2}{3^t}\begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

(ii) The characteristic polynomial of $\mathbf{A}$ is $\lambda^2 + 8\lambda + 15$, so the eigenvalues are -5 and -3. It is easy to show that $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$ are corresponding eigenvectors. Thus the general solution is

$$\mathbf{x}(t) = \mathbf{x}^*(t) + c_1 e^{-5t}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 e^{-3t}\begin{bmatrix} 3 \\ 1 \end{bmatrix},$$

where $\mathbf{x}^*(t)$ is a particular solution, depending on $\mathbf{b}(t)$. It remains to find $\mathbf{x}^*(t)$ in each of the cases (a), (b) and (c).

(a) For a particular solution, try $\mathbf{x}^*(t) = e^{-t}\mathbf{p}$, where $\mathbf{p}$ is a constant vector. Then

$$-e^{-t}\mathbf{p} = \frac{d}{dt}(e^{-t}\mathbf{p}) = \mathbf{A}(e^{-t}\mathbf{p}) + \mathbf{b}(t) = e^{-t}\mathbf{A}\mathbf{p} + e^{-t}\begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Multiplying by e^t and rearranging, we obtain

$$(\mathbf{I} + \mathbf{A})\mathbf{p} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}.$$

Solving for $\mathbf{p}$, we have $p_1 = 5/8$, $p_2 = 1/8$. Hence

$$\mathbf{x}^*(t) = e^{-t}\begin{bmatrix} 5/8 \\ 1/8 \end{bmatrix}.$$

(b) For a particular solution, try $\mathbf{x}^*(t) = e^{-2t}\mathbf{p}$, where $\mathbf{p}$ is a constant vector. Reasoning as in (a), we obtain

$$(2\mathbf{I} + \mathbf{A})\mathbf{p} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$

Solving for $\mathbf{p}$, we have $p_1 = -1$, $p_2 = 0$. Hence

$$\mathbf{x}^*(t) = -e^{-2t}\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

(c) Let $\mathbf{u}$ and $\mathbf{v}$ satisfy

$$\frac{d\mathbf{u}}{dt} = \mathbf{A}\mathbf{u} + \begin{bmatrix} e^{-t} \\ 0 \end{bmatrix}, \quad \frac{d\mathbf{v}}{dt} = \mathbf{A}\mathbf{v} + \begin{bmatrix} 0 \\ e^{-2t} \end{bmatrix}. \qquad (*)$$

Then

$$\frac{d}{dt}(4\mathbf{u} + 3\mathbf{v}) = 4\mathbf{A}\mathbf{u} + 3\mathbf{A}\mathbf{v} + \begin{bmatrix} 4e^{-t} \\ 3e^{-2t} \end{bmatrix} = \mathbf{A}(4\mathbf{u} + 3\mathbf{v}) + \mathbf{b}(t),$$

so $4\mathbf{u}+3\mathbf{v}$ is a solution to our differential equation. But we showed in (a) and (b) that $(*)$ is satisfied if

$$\mathbf{u} = e^{-t}\begin{bmatrix} 5/8 \\ 1/8 \end{bmatrix}, \quad \mathbf{v} = -e^{-2t}\begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Hence a particular solution is

$$\mathbf{x}^*(t) = 4\mathbf{u}+3\mathbf{v} = \tfrac{1}{2}e^{-t}\begin{bmatrix} 5 \\ 1 \end{bmatrix} - 3e^{-2t}\begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

28–3. For the given functional forms, the system of equations is

$$\dot{x} = rx(1-k^{-1}x) - qxy, \quad \dot{y} = \eta(pqxy - cy)$$

There are therefore three fixed points: $(0,0)$, $(k,0)$ and the point given by

$$x = \frac{c}{pq}, \quad y = \frac{r}{q}\left(1-\frac{x}{k}\right).$$

Setting $\theta = \dfrac{c}{kpq}$, we may write the coordinates of this third point as $\left(k\theta,\, \dfrac{r}{q}(1-\theta)\right)$. This is the only fixed point that could be in the positive quadrant: it will be in the positive quadrant if $\theta < 1$, i.e. if $c < kpq$. From now on we assume that this condition holds.

Writing the system of differential equations as

$$\dot{x} = f(x,y), \quad \dot{y} = g(x,y),$$

we have

$$\frac{\partial f}{\partial x} = \frac{f(x,y)}{x} - \frac{rx}{k}, \quad \frac{\partial f}{\partial y} = -qx, \quad \frac{\partial g}{\partial x} = \eta pqy, \quad \frac{\partial g}{\partial y} = \frac{g(x,y)}{y}.$$

Thus at the fixed point in the positive quadrant,

$$\frac{\partial f}{\partial x} = -r\theta, \quad \frac{\partial f}{\partial y} = -kq\theta, \quad \frac{\partial g}{\partial x} = \eta pr(1-\theta), \quad \frac{\partial g}{\partial y} = 0.$$

The associated linear sytem is therefore

$$\dot{x} = -r\theta(x-k\theta) - k\theta(qy - r(1-\theta)), \quad \dot{y} = \eta pr(1-\theta)(x-k\theta).$$

Let

$$vA = \begin{bmatrix} -r\theta & -kq\theta \\ \eta pr(1-\theta) & 0 \end{bmatrix},$$

the Jacobian matrix of the linearised system. Then

$$\operatorname{tr}\mathbf{A} = -r\theta < 0, \quad \det\mathbf{A} = \eta kpqr\theta(1-\theta) = \eta rc(1-\theta) > 0.$$

It follows that both eigenvalues of $\mathbf{A}$ have negative real parts: the fixed point is locally stable. Also, the characteristic polynomial of $\mathbf{A}$ is

$$\begin{vmatrix} \lambda + r\theta & kq\theta \\ -\eta pr(1-\theta) & \lambda \end{vmatrix} = \lambda^2 + r\theta\lambda + \eta rc(1-\theta).$$

Hence the eigenvalues of $\mathbf{A}$ are real and distinct if $(r\theta)^2 > 4\eta rc(1-\theta)$; they are complex conjugates if $(r\theta)^2 < 4\eta rc(1-\theta)$. Thus the fixed point is a stable node of the linearised system if $\eta < \eta_0$ and a spiral sink if $\eta > \eta_0$, where

$$\eta_0 = \frac{r\theta^2}{4c(1-\theta)} = \frac{rc}{4kpq(kpq-c)}.$$

28–4. (i) Writing $k = \ln K$ and $h = \ln H$, we have

$$\frac{dk}{dt} = \frac{1}{K}\frac{dK}{dt} = s_1 A K^{\alpha-1} H^{\gamma} - \delta_1, \quad \frac{dh}{dt} = \frac{1}{H}\frac{dH}{dt} = s_2 A K^{\alpha} H^{\gamma-1} - \delta_2.$$

Since $k = \ln K$, $K = e^k$, whence $K^\theta = e^{\theta k}$ for any constant θ. Similarly, $H^\theta = e^{\theta h}$ for any constant θ. We therefore have the following autonomous system in k and h:

$$\dot{k} = s_1 A \exp([\alpha - 1]k + \gamma h) - \delta_1, \quad \dot{h} = s_2 A \exp(\alpha k + [\gamma - 1]h) - \delta_2.$$

(ii) Let $b_i = \ln(s_i A/\delta_i)$ for $i = 1, 2$. Then the set of points in the kh–plane for which $\dot{k} = 0$ is the straight line

$$(\alpha - 1)k + \gamma h + b_1 = 0.$$

Since $\alpha < 1$ and $\gamma > 0$, this is an upward-sloping line of slope $(1 - \alpha)/\gamma$. By a similar argument, the set of points in the kh–plane for which $\dot{h} = 0$ is an upward-sloping straight line of slope $\alpha/(1 - \gamma)$. Since $\alpha + \gamma < 1$,

$$\frac{1-\alpha}{\gamma} > 1 > \frac{\alpha}{1-\gamma}.$$

Thus the line $\dot{k} = 0$ is steeper than the line $\dot{h} = 0$, so the two lines intersect at exactly one point (k^*, h^*).

(iii) The Jacobian of the autonomous system is

$$\begin{bmatrix} (\alpha-1)s_1 e^{-k}Q & \gamma s_1 e^{-k}Q \\ \alpha s_2 e^{-h}Q & (\gamma-1)s_2 e^{-h}Q \end{bmatrix},$$

where $Q = A\exp(\alpha k + \gamma h)$. For all k and h, both diagonal entries are negative and the determinant is

$$(1 - \alpha - \gamma)s_1 s_2 e^{-(k+h)} Q^2 > 0.$$

Therefore, conditions (i) and (iia) of Olech's theorem hold and (k^*, h^*) is globally stable.

29 DYNAMIC OPTIMISATION IN DISCRETE TIME

29–1. (i) The problem may be written

$$\text{maximise} \quad -\sum_{t=0}^{9}(3x_t^2 + 4y_t^2) \ \text{ subject to } \ y_{t+1} - y_t = x_t \ \ (t = 0, 1, \ldots, 9)$$

and the given endpoint conditions. The Hamiltonian is $-3x_t^2 - 4y_t^2 + \lambda_t x_t$, the control condition is $\lambda_t = 6x_t$ and the costate equation is $-8y_t = \lambda_{t-1} - \lambda_t$.

(ii) Eliminating x_t between the state equation and the control condition, we see that $\lambda_t = 6(y_{t+1} - y_t)$. Hence from the costate equation,

$$8y_t = 6(y_{t+1} - 2y_t + y_{t-1}),$$

which simplifies to the required equation.

(iii) The characteristic equation of the second-order difference equation has roots 3 and 1/3, so we can write the solution as $3^t A + 3^{-t} B$ where A and B are constants. From the left-endpoint condition, $A + B = 1$, so

$$y_t = 3^t A + 3^{-t}(1 - A) \quad (t = 0, 1, \ldots, 10). \qquad (*)$$

From the right-endpoint condition, $A = \left(100 - 3^{-10}\right) / \left(3^{10} - 3^{-10}\right)$.

(iv) The solution is again (∗) but with a different value of A. Instead of a right-endpoint condition, we have the transversality condition $\lambda_9 = 0$. As in (ii), $\lambda_t = 6(y_{t+1} - y_t)$ for $t = 0, 1, \ldots, 9$. Therefore $y_{10} = y_9$, or

$$\left(3^{10} - 3^9\right)A + \left(3^{-10} - 3^{-9}\right)\left(1 - A\right) = 0.$$

Hence $A = \left(3^{19} + 1\right)^{-1}$.

29–2. For this problem, it is helpful to use the subscript notation for partial derivatives:

$$\pi_1(h, s) = \frac{\partial \pi}{\partial h}, \quad \pi_2(h, s) = \frac{\partial \pi}{\partial s}.$$

To simplify further, let

$$\pi_{1t} = \pi_1(h_t, s_t), \quad \pi_{2t} = \pi_2(h_t, s_t).$$

(i) The resource manager's optimisation problem is to

$$\text{maximise} \quad \sum_{t=0}^{T}(1+\rho)^{-t}\pi(h_t, s_t)$$

subject to

$$h_t \geq 0, \quad s_{t+1} - s_t = g(s_t) - h_t \quad (t = 0, 1, \ldots, T)$$

and given s_0.

(ii) The Hamiltonian for period t is

$$(1+\rho)^{-t}\pi(h_t, s_t) + \lambda_t(g(s_t) - h_t).$$

The control condition is

$$(1+\rho)^{-t}\pi_{1t} \leq \lambda_t, \ \text{ with equality if } h_t > 0,$$

and the costate equation is

$$(1+\rho)^{-t}\pi_{2t} + \lambda_t g'(s_t) = \lambda_{t-1} - \lambda_t.$$

Setting $\mu_t = (1+\rho)^t\lambda_t$, we may write the costate equation in the form

$$\pi_{2t} = (1+\rho)\mu_{t-1} - (1 + g'(s_t))\mu_t.$$

(iii) The control condition for the steady state says that μ_t is constant over time and equal to $\pi_1(\bar{h}, \bar{s})$. Substituting this into the costate equation,

$$\pi_2(\bar{h}, \bar{s}) = (\rho - g'(\bar{s}))\pi_1(\bar{h}, \bar{s}),$$

which rearranges to the required equation.

29–3. (i) It is convenient to use $z = I/K$ as the control variable; the state variable is again K. The state equation is $K_{t+1} = (z_t - \delta)K_t$. Denoting the costate variable by λ, the Hamiltonian for period t is

$$H_t(K_t, z_t, \lambda_t) = (1+r)^{-t}(\pi_t - az_t - bz_t^{1+\gamma})K_t + \lambda_t(z_t - \delta)K_t.$$

The control condition is

$$(1+\gamma)bz_t^{\gamma} = (1+r)^t\lambda_t - a$$

if the right-hand side is positive, $z_t = 0$ otherwise. The costate equation is

$$(1+r)^{-t}(\pi_t - az_t - bz_t^{1+\gamma}) + \lambda_t(z_t - \delta) = \lambda_{t-1} - \lambda_t.$$

The transversality condition is $\lambda_T = 0$.

(ii) Let $\mu_t = (1+r)^t\lambda_t$ for $t = 0, 1, \ldots, T$; also let $\nu_t = \max(\mu_t - a,\, 0)$. Then the control condition may be written

$$z_t = \left[\frac{\nu_t}{(1+\gamma)b}\right]^{1/\gamma}.$$

From the costate equation and the control condition,

$$(1+r)\mu_{t-1} = \pi_t + (1-\delta)\mu_t + \frac{\gamma}{1+\gamma}\nu_t z_t. \qquad (\dagger)$$

To find the optimal path of investment, begin by noting that $\mu_T = 0$ by the transversality condition; therefore ν_T and z_T are both zero. We can then calculate μ_{T-1} using (†) with $t = T$, ν_{T-1} and z_{T-1} in the obvious way, μ_{T-2} using (†) with $t = T-1$, and so on back to z_0. The state equation and the fact that $I_t = z_t K_t$ then give I_t and K_t for all t.

To solve the problem using dynamic programming, one may try a very simple form for the value function, namely $V_t(K) = (1+r)^{-t}\xi_t K$, and solve for ξ_t for each t. Obviously there is no investment in period T, so $\xi_T = \pi_T$. For $t < T$, the Bellman equation is

$$(1+r)\xi_t = \pi_t + (1-\delta)\xi_{t+1} + \max_{z\geq 0}\left\{(\xi_{t+1} - a)z - bz^{1+\gamma}\right\}.$$

This gets us back to (†), with $\mu_{t-1} = \xi_t$.

29–4. (i) Because the only place to go from (i, n) is $(i+1, n)$.

(ii) Because the only place to go from (m, j) is $(m, j+1)$, $v_{mj} = a_{mj} + v_{m,j+1}$.

(iii) By Bellman's principle of optimality, $v_{ij} = a_{ij} + \max(v_{i,j+1},\, v_{i+1,j})$.

(iv)

$$\begin{bmatrix} \underline{12} & 8 & 12 & 7 & 4 & 1 \\ \underline{10} & \underline{7} & \underline{9} & 3 & 5 & 1 \\ 3 & 6 & \underline{5} & \underline{7} & \underline{4} & \underline{-1} \end{bmatrix}$$

Underlined entries show the optimal path.

(v)

$$\begin{bmatrix} \underline{14} & 10 & 14 & 10 & 4 & 1 \\ \underline{12} & \underline{9} & \underline{11} & 3 & 5 & 1 \\ 3 & 6 & 5 & \underline{7} & \underline{4} & \underline{-1} \end{bmatrix}$$

Underlined entries show the optimal path.

30 DYNAMIC OPTIMISATION IN CONTINUOUS TIME

30–1. (i) The Hamiltonian is

$$H(h, s, \lambda, t) = e^{-\rho t}R(h(t), t) + \lambda(g(s) - h(t)).$$

The control condition is $e^{-\rho t}\partial R/\partial h = \lambda$ and the costate equation is $\lambda g'(s) = -\dot{\lambda}$.

(ii) In the special case, the control condition becomes

$$e^{-\rho t}f'(h)e^{\alpha t} = \lambda,$$

i.e. $f'(h) = \mu$. Since f is strictly concave, f' is a decreasing function; let its inverse be the decreasing function ϕ. Then the control condition may be written $h = \phi(\mu)$, and the state equation then has the required form. By definition of μ,

$$\frac{\dot{\mu}}{\mu} = \frac{\dot{\lambda}}{\lambda} + \rho - \alpha,$$

so the costate equation may be written $\dot{\mu} = [\rho - \alpha - g'(s)]\mu$. This and the reformulated state equation form an autonomous system in s and μ.

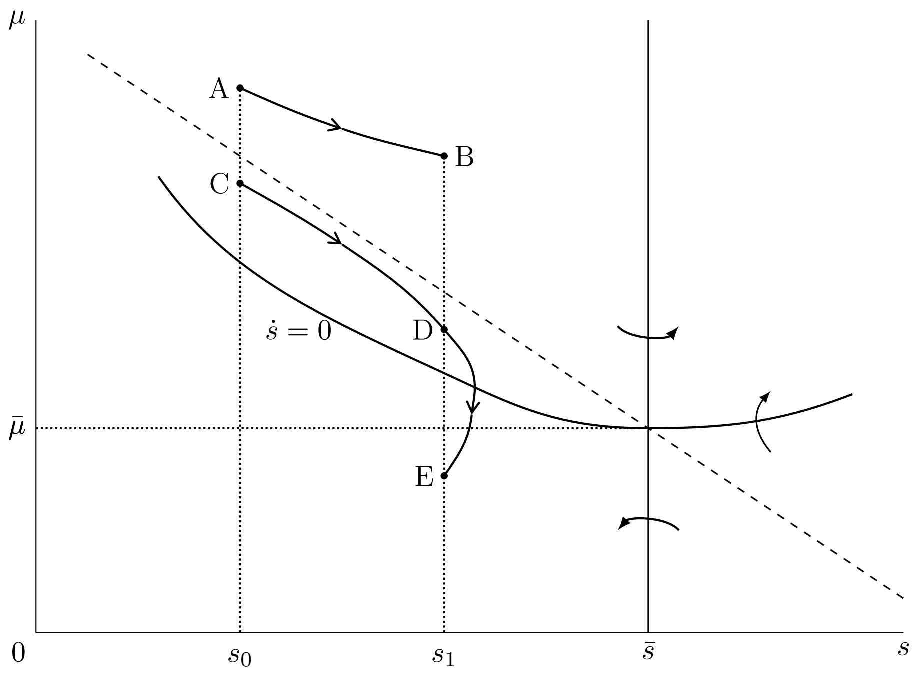
(iii) The Jacobian matrix of the autonomous system, evaluated at the unique fixed point $(\bar{s}, \bar{\mu})$, is

$$\begin{bmatrix} g'(\bar{s}) & -\phi'(\bar{\mu}) \\ -\bar{\mu}g''(\bar{s}) & -g'(\bar{s}) \end{bmatrix}$$

The determinant is

$$-\left[g'(\bar{s})\right]^2 - \bar{\mu}\phi'(\bar{\mu})g''(\bar{s}).$$

This expression is negative because g is concave and ϕ is decreasing. It follows that the fixed point is a saddle point.



In the phase diagram, the broken curve is the stable branch. To get an idea of what optimal paths look like, consider for definiteness the case where $s(0) = s_0$, $s(T) = s_1$ and $s_0 < s_1 < \bar{s}$. Then the solution may look like AB or CD or CDE in the diagram, depending on the value of T. There is a unique value of T (say T^*) such that the solution path is along the stable branch. There is also a unique value of T (say T^{**}, where $T^{**} > T^*$) such that the solution path has the property that $\dot{s}(T) = 0$. If $T < T^*$ and a solution exists, the solution path looks like AB. If $T^* < T < T^{**}$, the solution path looks like CD; and if $T > T^{**}$, like CDE.

30–2. (i) If the given inequality did not hold, then Mark's initial debt would be at least as great as the present value of his labour income, discounted at the borrowing rate. Hence it would be impossible for him to have positive consumption at each moment and die solvent.

An alternative way of making the same point is as follows. Suppose Mark's problem has a solution. If $a_0 \geq 0$ the required inequality obviously holds, so suppose $a_0 < 0$. By the right-endpoint condition, $a(t) \geq 0$ for some $t \leq T$; let τ be the smallest such t. Then $a(\tau) = 0$ (since a cannot jump), $0 < \tau \leq T$ and $a(t) < 0$ for all t such that $0 \leq t < \tau$. For such t, multiplication of the state equation by e^{-it} gives

$$\frac{d}{dt}\left(e^{-it}a(t)\right) = e^{-it}(w(t) - c(t)).$$

Setting

$$v(t) = a(t) + \int_t^\tau e^{-is} w(s)\, ds \quad (0 \leq t \leq \tau)$$

we see that

$$\dot{v}(t) = -e^{-it} c(t) < 0 \quad (0 < t < \tau).$$

Hence $v(0) > v(\tau)$. But $v(\tau) = a(\tau) = 0$. Therefore $v(0) > 0$, whence

$$a_0 + \int_0^T e^{-it} w(t)\, dt = v(0) + \int_\tau^T e^{-it} w(t)\, dt > 0.$$

(ii) The current-value Hamiltonian is $\ln c + \mu(f(a) + w - c)$. The control condition is $c = 1/\mu$. Since

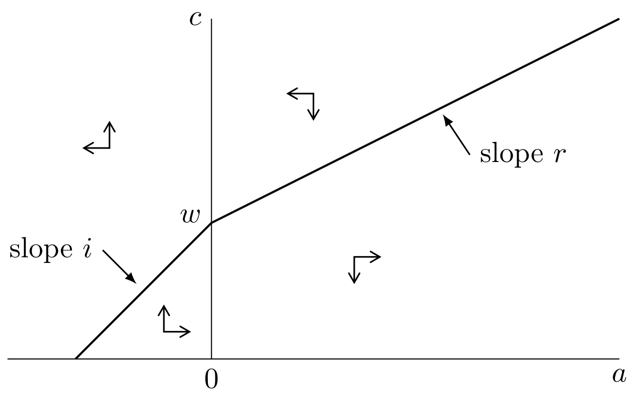
$$f'(a) = \begin{cases} r & \text{if } a > 0, \\ i & \text{if } a < 0, \end{cases}$$

the costate equation is $\dot{\mu} = (\rho - r)\mu$ if $a > 0$, $\dot{\mu} = (\rho - i)\mu$ if $a < 0$. Since $f'(a)$ is not defined if $a = 0$, nor is the costate equation. This means that μ is allowed to jump when $a = 0$, though it may not be optimal for it to do so.

(iii) From the control condition and the costate equation,

$$\dot{c} = \begin{cases} (r - \rho)c & \text{if } a > 0, \\ (i - \rho)c & \text{if } a < 0. \end{cases}$$

This and the state equation form an autonomous system in a and c. It follows from our assumptions about i, r and ρ that $\dot{c} > 0$ if $a < 0$, $\dot{c} < 0$ if $a > 0$. The phase diagram is as follows.



(iv) Consider first the case where $a_0 > 0$. We see from the phase diagram that $a(t) \geq 0$ for all t; and for as long as $a > 0$, $c > w$ and c is falling at the proportional rate $\rho - r$. Thus there are two possibilities:

(a) $c(t) = c_0 e^{(r-\rho)t}$ $(0 \leq t \leq T)$, where c_0 is chosen such that $c(T) \geq w$ and $a(T) = 0$. To satisfy the latter condition, the present value at time 0 of consumption less labour income must equal initial assets:

$$\int_0^T c_0 e^{-\rho t} dt - \int_0^T w e^{-rt} dt = a_0.$$

Hence

$$c_0 = \frac{\rho a_0 + (\rho w/r)\left[1 - e^{-rT}\right]}{1 - e^{-\rho T}}.$$

This will be the solution if a_0 is sufficiently large that $c_0 e^{(r-\rho)T} \geq w$. Specifically, this will be so if and only if $a_0 \geq a_+$, where

$$a_+ = we^{-rT}\left[\frac{e^{\rho T}-1}{\rho} - \frac{e^{rT}-1}{r}\right].$$

(b) If $0 < a_0 < a_+$, the consumption path is given by $c(t) = c_0 e^{(r-\rho)t}$ $(0 \leq t < \tau)$ and $c(t) = w$ if $\tau \leq t \leq T$. Here c_0 and τ are chosen such that $a(\tau) = 0$, and $\lim_{t\uparrow\tau} c(t) = w$. The latter condition is required because, with a concave utility function, jumps in consumption are undesirable. [The same conclusion is reached by considering the free-time problem of maximising $\int_0^\tau (\ln c - \ln w)e^{-\rho t}dt$ and using the free-terminal-time optimality condition to obtain τ.] Hence $c_0 = we^{(\rho-r)\tau}$ and τ is chosen such that

$$\int_0^\tau w\left[e^{(r-\rho)(t-\tau)} - 1\right]e^{-rt}dt = a_0.$$

Simplifying, τ is given by the equation

$$we^{-r\tau}\left[\frac{e^{\rho\tau}-1}{\rho} - \frac{e^{r\tau}-1}{r}\right] = a_0.$$

Now suppose $-\frac{w}{i}\left(1-e^{-it}\right) < a_0 < 0$. From the phase diagram, $a(t) \leq 0$ for all t; and for as long as $a < 0$, $c < w$ and c is rising at the proportional rate $i - \rho$. There are two possibilities:

(e) $c(t) = c_0 e^{(i-\rho)t}$ $(0 \leq t \leq T)$, where c_0 is chosen such that $c(T) \leq w$ and $a(T) = 0$. To satisfy the latter condition we must have

$$c_0 = \frac{\rho a_0 + (\rho w/i)\left[1 - e^{-iT}\right]}{1 - e^{-\rho T}}.$$

This will be the solution if $-a_0$ is sufficiently large that $c_0 e^{(i-\rho)T} \leq w$. This will be so if and only if $a_0 \leq a_-$, where

$$a_- = we^{-iT}\left[\frac{e^{\rho T}-1}{\rho} - \frac{e^{iT}-1}{i}\right] < 0.$$

(d) If $a_- < a_0 < 0$, the consumption path is given by $c(t) = c_0 e^{(i-\rho)t}$ $(0 \leq t < \tau)$ and $c(t) = w$ if $\tau \leq t \leq T$. Here c_0 and τ are chosen such that $a(\tau) = 0$, and $\lim_{t\uparrow\tau} c(t) = w$. Hence $c_0 = we^{(\rho-i)\tau}$ and τ is chosen such that

$$we^{-i\tau}\left[\frac{e^{i\tau}-1}{i} - \frac{e^{\rho\tau}-1}{\rho}\right] = -a_0.$$

Finally, if $a_0 = 0$ then $a(t) = 0$ and $c(t) = w$ for all t. This is case (c).

(v) See (iv) above, parts (a) and (e).

30–3. (i) Recall from Problem 9–3 that this is the limiting case of (ii) as $\gamma \downarrow 1$. See below.

(ii) The dynamic system corresponding to (28.12) in the text is

$$\dot{K} = (b+\rho)K - C, \quad \dot{C} = bC/\gamma,$$

where $b = A - \delta - \rho$. Therefore

$$\frac{d}{dt}\ln\frac{C}{K} = \frac{\dot{C}}{C} - \frac{\dot{K}}{K} = \frac{C}{K} - \theta,$$

where $\theta = \left(1-\gamma^{-1}\right) b+\rho$. Hence our dynamic system has a solution where C/K remains unchanged over time and C and K grow at the same constant rate:

$$C = \theta K, \quad \frac{\dot{C}}{C} = \frac{\dot{K}}{K} = \frac{A-\delta-\rho}{\gamma}$$

for all $t \geq 0$. By our definitions of b and θ,

$$\theta = \frac{(\gamma-1)(A-\delta)+\rho}{\gamma}. \tag{†}$$

Because $\gamma > 1$ and $A > \delta+\rho$, θ and the common growth rate of C and K are positive. To show that this solution satisfies the transversality condition — (28.9) in the text — note that $KU'(C) = C^{1-\gamma}/\theta$, which grows at the negative rate $(\gamma^{-1}-1)(A-\delta-\rho)$. Since also $\rho > 0$, (28.9) is satisfied. Since the utility function is concave and the differential equations are linear, the path is optimal. On the optimal path, output AK grows at the same rate as K, namely $(A-\delta-\rho)/\gamma$, which is constant, positive and decreasing in ρ for given A, δ, γ. Also $C/K = \theta$ which, by (†), is increasing in ρ for given A, δ, γ.

For part (i) we have the same dynamic system, but with $\gamma = 1$. A feasible path is given by $C = \theta K$ where $\theta = \rho$. On this path, C and K grow at rate $A-\delta-\rho$ and $KU'(C)$ is the constant $1/\rho$; the transversality condition now follows from the fact that $\rho > 0$. The growth rate is again decreasing in ρ for given A and δ.

(iii) If $0 < \gamma < 1$ we may start as above, but now we need extra assumptions to ensure that $\theta > 0$ and the transversality condition is satisfied. For the former, we need the right-hand side of (†) to be positive, i.e.

$$\gamma > 1 - \frac{\rho}{A-\delta}. \tag{††}$$

Given (††), $KU'(C)$ grows at the positive rate $(\gamma^{-1}-1)(A-\delta-\rho)$, and for the transversality condition we need this to be less than ρ. In fact,

$$\gamma\rho - (1-\gamma)(A-\delta-\rho) = \rho - (1-\gamma)(A-\delta),$$

which is positive by (††). Thus if (††) holds, the path along which $C/K = \theta$ for all t is feasible and satisfies the transversality condition; it is therefore optimal.

30–4. We assume throughout that

$$\frac{\psi A}{1+\psi} > \delta + \rho. \tag{$*$}$$

In particular, $A > \delta$. Let ζ, c be constants such that $\delta/A < \zeta \leq 1$ and $0 < c < 1$. If we set $z(t) = \zeta$, $C(t) = (A\zeta-\delta)cK(t)$ for all t, then C and K grow at the positive rate $(1-c)(A\zeta-\delta)$. Thus endogenous growth is possible. To show that it is not desirable, we derive the optimal path.

The current-value Hamiltonian is

$$\ln C - bAKz^{1+\psi} + \mu(AKz - C - \delta K).$$

The control conditions, taking into account the constraint $z \leq 1$, are

$$\mu = C^{-1}, \quad (1+\psi)bz^{\psi} = \min(\mu, [1+\psi]b).$$

Setting $B = \dfrac{1}{(1+\psi)b}$, we see that

$$z = \begin{cases} 1 & \text{if } C \leq B; \\ (B/C)^{1/\psi} & \text{if } C > B. \end{cases}$$

Thus if $C \le B$, the costate equation implies that

$$-\frac{1}{C^2}\frac{dC}{dt} = \frac{\rho}{C} + \left(\frac{1}{(1+\psi)B} - \frac{\rho}{C}\right)A + \frac{\delta}{C} = \frac{A}{(1+\psi)B} - \frac{A-\delta-\rho}{C},$$

while if $C > B$,

$$-\frac{1}{C^2}\frac{dC}{dt} = \frac{\rho}{C} - \frac{\psi}{1+\psi}A\mu z + \frac{\delta}{C} = -\frac{A\psi}{1+\psi}\frac{(B/C)^{1/\psi}}{C} + \frac{\delta+\rho}{C}.$$

We therefore have the following autonomous system:

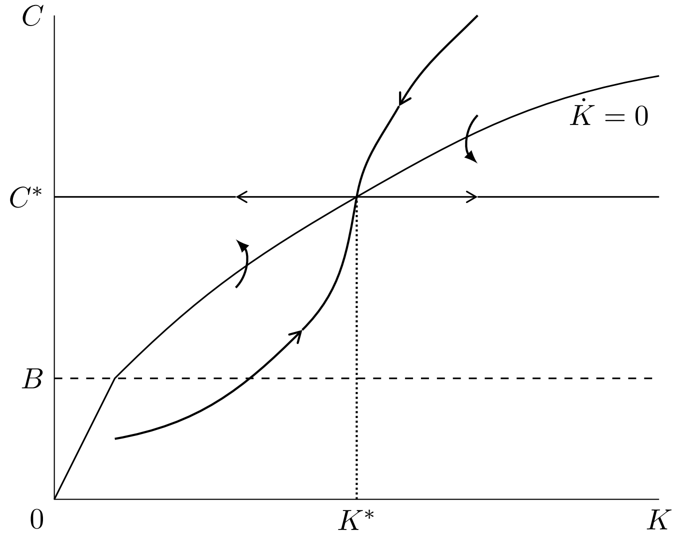
$$\text{if } C \le B, \quad \frac{\dot K}{K} = A - \delta - \frac{C}{K} \quad \text{and} \quad \frac{\dot C}{C} = \left[1 - \frac{C}{(1+\psi)B}\right]A - (\delta+\rho);$$

$$\text{if } C > B, \quad \frac{\dot K}{K} = A\left[\frac{B}{C}\right]^{1/\psi} - \delta - \frac{C}{K} \quad \text{and} \quad \frac{\dot C}{C} = \frac{A\psi}{1+\psi}\left[\frac{B}{C}\right]^{1/\psi} - (\delta+\rho).$$

Let

$$C^* = \left[\frac{\psi A}{(1+\psi)(\delta+\rho)}\right]^{1/\psi} B, \quad K^* = \frac{\psi C^*}{\delta + (1+\psi)(\delta+\rho)}.$$

By $(*)$, $C^* > B$ and (K^*, C^*) is the unique fixed point of the autonomous system.



;

It is clear from the phase diagram that the fixed point is a saddle point. The optimal policy is: given $K(0)$, choose $C(0)$ so that $(K(0), C(0))$ is on the stable branch. the optimal path then follows the stable branch, which converges to (K^*, C^*). Since $C^* > B$, $z(t) < 1$ for all sufficiently large t.

How do we know that this path is optimal? If we use the control conditions to express C and z in terms of K and μ, we see that they depend only on μ and not on K. Hence the maximised Hamiltonian, like the ordinary one, is linear in the state variable. Thus the standard concavity condition is satisfied; the transversality condition is satisfied because $\rho > 0$.

31 INTRODUCTION TO ANALYSIS

31–1. (i) Let $y = \frac{1}{2}v + v^{-1}$, $z = \frac{1}{2}v - v^{-1}$; then $y + z = v$ and $y - z = 2v^{-1}$, so $y^2 - z^2 = 2$. Our assumptions about v imply that y and z are positive. Since $y = v - z$ and $y^2 = 2 + z^2$, the inequalities $y < v$ and $y^2 > 2$ follow from the fact that $z > 0$.

(ii) Let $x \in A$. Then $y + x$ is the sum of two positive numbers, while $y^2 - x^2$ is the sum of the two positive numbers $y^2 - 2$ and $2 - x^2$. Thus, $y + x > 0$ and $y^2 - x^2 > 0$; but $y^2 - x^2 = (y + x)(y - x)$; hence $y - x > 0$. Since this argument is valid for any x in A, y is an upper bound for A. In particular, $y \geq s$; hence $v > s$.

(iii) By assumption, $2/u$ is a positive number such that $(2/u)^2 > 2$. Arguing as in (i), we obtain a real number t such that $0 < t < 2/u$ and $t^2 > 2$. Then $w = 2/t$ satisfies our requirements.

(iv) By the axiom of Archimedes, we may choose a natural number $N > (w - u)^{-1}$. By the same axiom, the set of non-negative integers k such that $k \leq Nu$ is finite; let K be the greatest member of this set. Then

$$Nu < K + 1 < Nu + N(w - u) = Nw,$$

so we may let x be the rational number $(K + 1)/N$. To show that u is not an upper bound for A, it suffices to show that $x \in A$. We have already ensured that x is positive and rational, so it remains to prove that $x^2 < 2$. This is so because $2 - x^2$ is the sum of the positive numbers $2 - w^2$ and $(w + x)(w - x)$.

(v) The proof is by contraposition. Let c be a real number such that $c > 0$ and $c^2 \neq 2$; we wish to show that $c \neq s$. If $c^2 > 2$ we may apply (ii) with $v = c$, inferring that $c > s$. If $c^2 < 2$ we may apply (iv) with $u = c$, inferring that c is not an upper bound for A; in particular, $c \neq s$.

(vi) Let a be a real number such that $a > 1$, and let $A = \{\, x \in \mathbb{Q} : x > 0 \text{ and } x^2 < a \,\}$. Then $1 \in A$, so A is non-empty. By the axiom of Archimedes there is a positive integer M such that $M > a > 1$; hence $M^2 > M > a$. Thus if $x \in A$ then

$$(M - x)(M + x) = M^2 - x^2 > a - x^2 > 0,$$

so $x < M$. Therefore A is bounded above by M, and we may define the positive real number $s = \sup A$. We shall show that $s^2 = a$.

Let v be a positive real number such that $v^2 > a$. Let

$$y = \tfrac{1}{2}(v + v^{-1}a), \quad z = \tfrac{1}{2}(v - v^{-1}a).$$

Our assumptions about v imply that y and z are positive, and it is easy to show that $y = v - z < v$ and $y^2 = a + z^2 > a$. This proves the analogue of (i) when 2 is replaced by a (except where 2 is used as an index); the analogues of (ii)–(v) are proved exactly as above. Hence $s^2 = a$.

Finally, let b be a real number such that $0 < b \leq 1$; we wish to prove that there is a positive real number σ such that $\sigma^2 = b$. If $b = 1$ then $\sigma = 1$. If $0 < b < 1$ then, for reasons just explained, there is a positive real number s such that $s^2 = 1/b$; we may therefore let $\sigma = 1/s$.

31–2. (i) Let $a_n = n^{1/n} - 1 \;\forall\, n \in \mathbb{N}$. Since $n^\alpha > 1$ if $n > 1$ and $\alpha > 0$, $a_n > 0 \;\forall\, n > 1$. Let $n > 1$ and apply the given inequality with $a = a_n$:

$$n > \frac{n(n-1)}{2}a_n^2, \text{ whence } a_n^2 < \frac{2}{n-1}.$$

If $\varepsilon > 0$ and N is a natural number greater than $2/\varepsilon^2$, then $0 < a_n < \varepsilon \ \forall n > N$. Therefore $\lim_{n\to\infty} a_n = 0$, as required.

(ii) Let $a = b^{-1} - 1 > 0$. For all n, $b^{-n} > \frac{1}{2}n(n-1)a^2$ by the given inequality. Therefore $0 < nb^n < (2/a^2)(n-1)^{-1} \ \forall n > 1$, and $nb^n \to 0$ by **SQ1** in Section 31–3.

To prove the last part, let $c = b^{1/k}$. Then $0 < c < 1$ and $n^k b^n = (nc^n)^k$ for all n. By (ii), with b replaced by c, $\lim_{n\to\infty} nc^n = 0$. But then $\lim_{n\to\infty}(nc^n)^k = 0$ by repeated application of **SQ2**, part (b), in Section 31–3.

31–3. (i) Let

$$s_n = \sum_{r=1}^{n} x_r, \quad t_n = \sum_{r=1}^{n} u_r.$$

$\{t_n\}$ is a convergent sequence and therefore a Cauchy sequence. But if $m > n$,

$$|s_n - s_m| = \left|\sum_{r=m+1}^{n} x_r\right| \le \sum_{r=m+1}^{n} |x_r| \le \sum_{r=m+1}^{n} u_r = |t_n - t_m|.$$

Hence $\{s_n\}$ is also a Cauchy sequence and is therefore convergent.

(ii) Since $\lim_{n\to\infty}(x_n/u_n) = 1$, we may choose an integer k such that $0 < x_n/u_n < 2 \ \forall n > k$. Since $u_n > 0$, $0 < x_n < 2u_n \ \forall n > k$, and hence $|x_n| < 2u_n$ for such n. We may therefore repeat the argument of (i), replacing u_n by $2u_n$ and assuming $m > n > k$ in the chain of inequalities.

For the last part, let $x_n = \dfrac{2^n + 3}{3^n + 2}$, $u_n = \dfrac{2^n}{3^n}$. Then

$$\frac{x_n}{u_n} = \frac{1 + 3\times 2^{-n}}{1 + 2\times 3^{-n}} \to 1 \ \text{ as } n \to \infty.$$

Now $\sum_{r=1}^{\infty} u_r = \frac{2}{3}\big/\left(1 - \frac{2}{3}\right) = 2$. Hence by (ii), the the series $\sum_{r=1}^{n} x_r$ is convergent.

31–4. (i) Let $x_0 \in I$, $\varepsilon > 0$. By definition of uniform continuity, $\exists\ \delta > 0$ such that $x \in I$ and $|x - x_0| < \delta$ imply $|f(x) - f(x_0)| < \varepsilon$. Since this argument is valid for all $\varepsilon > 0$, f is continuous at x_0. Since this is so for all $x_0 \in I$, f is continuous on I.

(ii) Suppose $f\colon I \to \mathbb{R}$ is not uniformly continuous. Then $\exists\ \alpha > 0$ with the following property: for any $\delta > 0$, however small, there exist $u, v \in I$ such that $|u - v| < \delta$ and $|f(u) - f(v)| \ge \alpha$. Using this fact for $\delta = 1, \frac{1}{2}, \frac{1}{3}, \ldots$, we obtain sequences $\{u_n\}$ and $\{v_n\}$ in I such that $|u_n - v_n| < 1/n$ and $|f(u_n) - f(v_n)| \ge \alpha$ for every natural number n. Since $a \le v_n \le b$ for all n, it follows from **SQ6** that there is a subsequence $\{v_{n_k}\}$ of $\{v_n\}$ which converges to a real number $\bar{x}$ with the same property; thus $\bar{x} \in I$. [This is where we use the assumption that I contains its endpoints. Without that assumption, the proof doesn't work and the conclusion may be false: see (iii) below.] We shall prove that f is not continuous at $\bar{x}$.

Since $u_n - v_n \to 0$ as $n \to \infty$, $\lim_{k\to\infty} u_{n_k} = \lim_{k\to\infty} v_{n_k} = \bar{x}$. But $|f(u_n) - f(v_n)| \ge \alpha$ for all n, so the sequences $\{f(u_{n_k})\}$ and $\{f(v_{n_k})\}$ cannot both converge to $f(\bar{x})$. The required result now follows from **SQ8**.

(iii) It suffices to show that for any $\delta > 0$ there exist real numbers u, v satisfying the conditions

$$0 < u < v < 1, \quad v < u + \delta, \quad u^{-1} - v^{-1} \ge 1.$$

Let v be a real number such that $0 < v < \min(1, 2\delta)$ and let $u = v/2$. Then the first of the three conditions is obviously satisfied, $v - u = \frac{1}{2}v < \delta$ and $u^{-1} - v^{-1} = v^{-1} > 1$.

32 METRIC SPACES AND EXISTENCE THEOREMS

32–1. (i) $\langle \mathbf{a}+\mathbf{b}\rangle = |a_i + b_i|$ for some i, say $i = k$. Hence

$$\langle \mathbf{a}+\mathbf{b}\rangle \leq |a_k| + |b_k| \leq \langle \mathbf{a}\rangle + \langle \mathbf{b}\rangle.$$

Setting $\mathbf{a} = \mathbf{x} - \mathbf{z}$ and $\mathbf{b} = \mathbf{z} - \mathbf{y}$, we see that $\tilde{d}$ satisfies **M4**. Also $|x_i - y_i| \leq 0$ for $i = 1, \ldots, m$ if and only if $\mathbf{x} = \mathbf{y}$; thus $\tilde{d}$ satisfies **M2**. **M1** and **M3** are obvious.

(ii) Let $1 \leq i \leq m$; then $|a_{in}| \leq \langle \mathbf{a}_n\rangle$ for all n; therefore, if the sequence $\{\langle \mathbf{a}_n\rangle\}$ converges to 0, so does $\{a_{in}\}$. Conversely, suppose $a_{in} \to 0$ for $i = 1, \ldots, m$. Let $\varepsilon > 0$. Then we may choose natural numbers $N_1, N_2, \ldots, N_m$ such that

$$|a_{in}| < \varepsilon \;\forall\, n > N_i \quad (i = 1, \ldots, m).$$

Let $N = \max(N_1, N_2, \ldots, N_m)$; then $\langle \mathbf{a}_n\rangle < \varepsilon \;\forall\, n > N$. Hence $\langle \mathbf{a}_n\rangle \to 0$.

(iii) Let $\mathbf{a} \in \mathbb{R}^m$, and let the integer k be such that $1 \leq k \leq m$ and $|a_k| = \langle \mathbf{a}\rangle$. Then

$$a_k^2 \leq a_1^2 + \ldots + a_m^2 \leq m a_k^2,$$

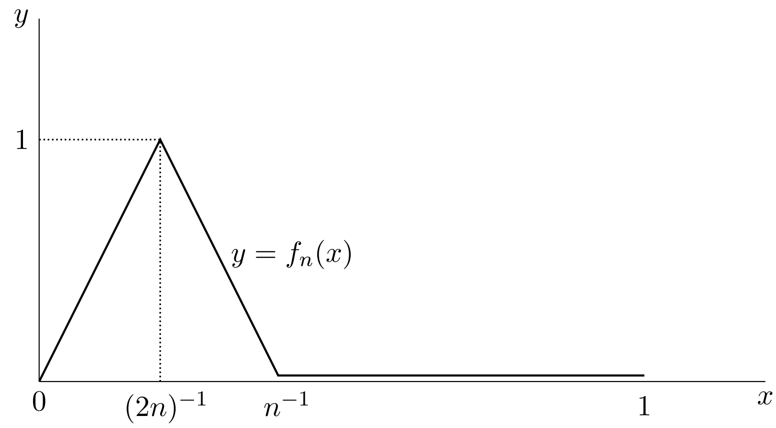
Taking square roots, we obtain the required result.

(iv) By a basic property of sequences of real numbers (**SQ1** in Section 31.3), it suffices to find positive numbers K and L such that $\|\mathbf{x}_n - \mathbf{x}\| \leq K\langle \mathbf{x}_n - \mathbf{x}\rangle$ for all n and $\langle \mathbf{x}_n - \mathbf{x}\rangle \leq L\|\mathbf{x}_n - \mathbf{x}\|$ for all n. By (iii), we may set $K = \sqrt{m}$ and $L = 1$.

(v) By (ii) with $\mathbf{a}_n = \mathbf{x}_n - \mathbf{x}$, the sequence $\{\mathbf{x}_n\}$ converges componentwise to $\mathbf{x}$ if and only if $\langle \mathbf{x}_n - \mathbf{x}\rangle \to 0$. By (iv), this happens if and only if $\|\mathbf{x}_n - \mathbf{x}\| \to 0$.

32–2. For each $n \in \mathbb{N}$, let $x_n = (2n)^{-1}$. Then $f_n(x_n) = 1$ and $f_m(x_n) = 0 \;\forall\, m \geq 2n$.

If $n \in \mathbb{N}$ and $x \in [0,1]$, then $|f_n(x) - f(x)| = f_n(x) \leq 1$, with equality if $x = x_n$. Thus $d(f_n, f) = 1$. If $m, n \in \mathbb{N}$ and $x \in [0,1]$, then $0 \leq f_m(x) \leq 1$ and $0 \leq f_n(x) \leq 1$, so $|f_n(x) - f_m(x)| \leq 1$. If $m \geq 2n$, $f_n(x_n) - f_m(x_n) = 1$, whence $d(f_n, f_m) = 1$.



(i) Since $d(f_n, f) = 1 \;\forall\, n \in \mathbb{N}$, the sequence $\{f_n\}$ does not converge uniformly to f. To prove pointwise convergence, let $x \in [0,1]$; we show that $f_n(x) = 0$ for all sufficiently large n. If $x = 0$ then $f_n(x) = 0$ for all n. If $0 < x \leq 1$, we can choose a natural number N such that $N > x^{-1}$; then If $n \geq N$, $x > n^{-1}$, so $f_n(x) = 0$.

(ii) Since $d(f_m, f_n) = 1$ whenever $m \geq 2n$, the sequence $\{f_n\}$ has no subsequence that is a Cauchy sequence, and hence no convergent subsequence. The sequence $\{f_n\}$ is bounded because $d(f_n, f) = 1 \;\forall\, n \in \mathbb{N}$.

(iii) Immediate from the proof of (ii) and the definition of compactness.

(iv) (i)–(iii) remain true when $B[0,1]$ is replaced by $C[0,1]$ because each f_n is a continuous function, as is f.

32–3. (i) Let $x \in I$. Then $|f_m(x) - f_n(x)| \le d(f_m, f_n)$ for all m and n, where d is the metric of $B[a,b]$. Hence $\{f_n(x)\}$ is a Cauchy sequence of real numbers. By the completeness of $\mathbb{R}$, the sequence $\{f_n(x)\}$ converges, say to $f(x)$. This defines a function $f\colon I \to \mathbb{R}$ such that the sequence $\{f_n\}$ converges pointwise to f.

(ii) Fix $n > N$ and $x \in I$; then $|f_m(x) - f_n(x)| < \varepsilon \;\forall\, m > N$. Let $\alpha > 0$. By definition of the function f we may choose a positive integer M such that $|f(x) - f_m(x)| < \alpha \;\forall\, m > M$. If $m > \max(M, N)$, then

$$|f(x) - f_n(x)| \le |f(x) - f_m(x)| + |f_m(x) - f_n(x)| < \alpha + \varepsilon.$$

Since this reasoning is valid for any positive number α, however small, $|f(x) - f_n(x)| \le \varepsilon$. This is so for all $x \in I$ and all $n > N$.

(iii) Since the argument of (ii) is valid for every positive number ε, the sequence $\{f_n\}$ converges uniformly to f. We may therefore choose a positive integer k with the property that $|f(x) - f_k(x)| < 1 \;\forall\, x \in I$. Since f_k is a bounded function, we may choose a positive number K such that $|f_k(x)| < K \;\forall\, x \in I$. Then $|f(x)| < K + 1 \;\forall\, x \in I$.

(iv) Completeness of $B[a,b]$ follows immediately from (i) and (iii). $C[a,b]$ is a closed set in $B[a,b]$ (Exercise 32.2.7); the completeness of the metric space $C[a,b]$ now follows from the completeness of $B[a,b]$ and the result of Exercise 30.2.8.

32–4. (i) Since we are using 'd' in the conventional manner of integral calculus, we denote the standard metric on $C[a,b]$ by δ. We shall show that for all $V_1, V_2 \in C[a,b]$,

$$\delta(\widetilde{V}_1, \widetilde{V}_2) \le \beta\delta(V_1, V_2). \tag{1}$$

The required result then follows from the fact that $0 < \beta < 1$.
Let $V_1, V_2 \in C[a,b]$ and let $y \in [a,b]$. Then we may choose $x_1 \in X$ such that

$$\widetilde{V}_1(y) = f(x_1, y) + \beta \int_a^b \psi(x_1, y, z)\, V_1(z)\, dz.$$

By definition of the function $\widetilde{V}_2$,

$$\widetilde{V}_2(y) \ge f(x_1, y) + \beta \int_a^b \psi(x_1, y, z)\, V_2(z)\, dz.$$

Hence

$$\widetilde{V}_1(y) - \widetilde{V}_2(y) \le \beta \int_a^b \psi(x_1, y, z)\, [V_1(z) - V_2(z)]\, dz. \tag{2}$$

For all $z \in [a,b]$, $V_1(z) - V_2(z) \le \delta(V_1, V_2)$ and $\psi(x_1, y, z) \ge 0$, so the integrand on the right-hand side of (2) cannot exceed $\delta(V_1, V_2)\psi(x_1, y, z)$. Thus (2), together with the fact that $\int_a^b \psi(x_1, y, z)\, dz = 1$, implies that

$$\widetilde{V}_1(y) - \widetilde{V}_2(y) \le \beta\delta(V_1, V_2). \tag{3}$$

A similar argument with the roles of V_1 and V_2 reversed shows that (3) remains true when its left-hand side is replaced by its absolute value. Since this is so for all y in $[a,b]$, (1) is also true.

(ii) There is a unique function $V \in C[a,b]$ such that

$$V(y) = \max_{x \in X} \left\{ f(x, y) + \beta \int_a^b \psi(x, y, z)\, V(z)\, dz \right\} \;\; \forall\, y \in [a,b].$$