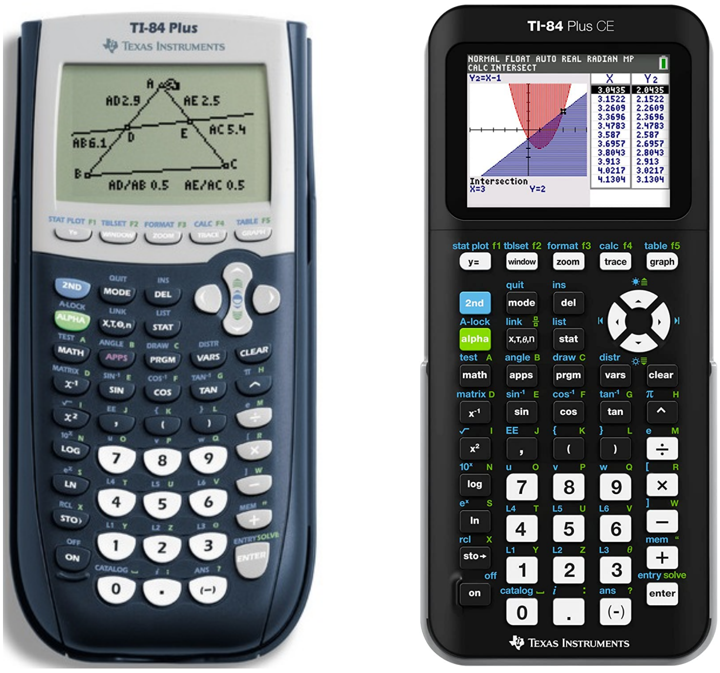
Mathematics on a TI-84/CE

Volume 4

Mathematics Labs



Peter McIntyre

School of Science

UNSW Canberra

2022

Contents

Volume 1 of this book contains the basic topics: Graphics Calculators and Mathematics; Getting Started; Coordinate Geometry; Inequalities and Linear Programming; Fitting Curves to Data 1 – Calculator Functions; Population Modelling 1 – Exponential Growth; Financial Mathematics 1 – Compound Interest; and Probability and Statistics 1 – Descriptive Statistics. The Volume 1 Supplement contains extra activities for Coordinate Geometry and Probability and Statistics 1.

Volume 2 of this book contains topics directly relevant to Calculus and its applications, although the first chapter, *Functions and their Graphs*, is of more general relevance. The topics in Volume 2 are: Functions and their Graphs; Graph and Calculus Operations; Numerical Integration; Taylor Series; Differential Equations; Population Modelling 2 – Logistic and Epidemic Models; and Multivariable Calculus.

The last chapter in Volume 2 gives a list of the programs cited in all volumes of the book, and full information on copying and using these programs.

Volume 3 of this book contains more advanced topics, relevant to students and teachers of Specialist Mathematics and first-year university Mathematics courses. The topics in Volume 3 are: Sequences and Series; Probability and Statistics 2 – Probability Distributions and Hypothesis Testing; Matrices and Vectors; Population Modelling 3 – Matrix Models; Fitting Curves to Data 2; Financial Mathematics 2 – TVM Calculations; Complex Numbers; and Programming.

Calculator versions

Currently (early 2022), TI-84 calculators come in two versions: the TI-84Plus and the more recent TI-84CE. The main difference is that the CE screen has much higher resolution. It also has colour but I have done most of the screens in black and white to avoid the need for colour printers or photocopiers. Calculations, screenshots and figures were done on a TI-84CE in CLASSIC mode.

Some programs have had to be changed for the CE because of the different screen: I usually append 'CE' to the program name to indicate this.

All the programs here are available at *www.XXX*.

1 Introduction

This document contains twenty-eight labs used in a first-year university Mathematics course, the first twenty-two of which require a TI-84/CE graphics calculator. Most of these labs are also suitable as short projects for good students in senior secondary school. They cover topics in Calculus, Linear Algebra, Vectors, Probability, Complex Numbers and introductory Discrete Dynamics.

As well as a spread in topics, there is a spread in how the labs/projects are presented, from the quite prescriptive with step-by-step instructions to less-well-defined problems in which working out the question is as important as finding the answer.

Each lab/project is accompanied by an Instructors' Guide containing solutions to the questions, together with suggestions on running the lab and the equations/techniques used.

Our students all had TI-84Plus programmable graphics calculators. The use of such calculators not only allows a broader range of approaches to problems — graphical and numerical approaches are available — but also means that the problems can be more realistic in that the numbers don't have to come out 'nicely'. Calculator screens are often used in the solutions to show students (and instructors) what they should see on their calculator.

In some labs, a calculator program is used.[1] These avoid the need for lengthy hand calculations so students can explore the model they are using. The equations used in these programs and instructions for their use are given in the Instructors' Guide for each lab. Section 5 gives details on copying and manipulating programs. The programs and instructions can be modified for other graphics calculators or computers.

The labs here were run with groups of mostly four students over a double period (110 minutes in total). Each group was required to hand in a lab report at the end of the lab period, with each student in the group receiving the group mark. More details are provided in the Lab Manual (which students received) in Section 4. Also discussed there are some benefits of groupwork, which we talk about with the students, and the groupwork evaluation scheme. These can be modified if the labs are run as projects.

Students need lots of time to do these problems properly, so they are particularly suited to projects. The main aim is to get students to think about, discuss, formulate and solve some problems as a group.

Some labs contain supplementary questions; these can be used as bonus questions for good students, alternative questions to those in the lab/projects or as a basis for further short projects for very good students.

A second aim is to get students to communicate mathematics well on paper. They need help with this in the form of *guidelines* (see the Lab Manual), *examples* such as the Snow White lab in the Lab Manual, and *feedback*, which they get from the mark they receive, the comments of the marker and oral feedback in the following lab.

We did the *Balloons, Submarines and Drag* lab first, over two lab periods, as a good ice-breaker both for the students and staff, and for initial guidance in writing up a lab. If you use another lab first, consider incorporating in it the second part of the *Balloons, Submarines and Drag* lab on writing up and the *Snow White* exercise.

[1]these programs are available from *www.XXX*

A comment on student approaches to the labs. We want our students to work as a group, because we think that groupwork provides significant benefits. Students, particularly those with an eye on the clock, are inclined to divide up the problems amongst the group members, in the belief that they will finish sooner. To counter this, we require that the lab report be written by only one person, the scribe, and in some labs we give them only one problem, e.g. Labs 2.12 – 2.15 or give them the problems one at a time, as in Lab 2.7.

Finally a note on scenarios. Many of the labs incorporate some sort of scenario relevant to the background and interests of the students (UNSW Canberra is located at the Australian Defence Force Academy, hence the number of labs here with a military flavour). It should not be too difficult to adapt the scenarios to your own environment. Having a relevant scenario seems to help engage students in a problem ('not just another Maths exercise') and at least provides a basis for some discussion, even if it is not directly related to the mathematics. In some labs, we even ask the groups to provide an alternative scenario. This too helps the students identify with the problem and has uncovered a wealth of creative talent.

Labs 2.1 – 2.16 are based on Calculus, ranging from basic derivatives and integrals to solution of simple, first-order differential equations. Lab 2.1 requires a balance capable of measuring to tenths of a gram.

Lab 2.17 requires a basic knowledge of the solution of (linear) simultaneous equations.

Lab 2.18 uses matrices at an advanced level (eigenvalues and eigenvectors).

Labs 2.19 and 2.20 require a knowledge of mean and variance — the notation could be changed here to make it more accessible to secondary students. Notes covering the material are available.

Lab 2.21 introduces difference equations but does not assume much, if any, previous knowledge.

Lab 2.22 is a little more complicated than the others, as it involves coupled difference equations. However, the complications are mostly overcome by using calculator operations (sequence graphing) to plot the solutions.

Acknowledgements

Ruth Hubbard, ex Queensland University of Technology — for getting us started and some of the labs.

Colin Pask — for Labs 2.17, 2.18 and 3.1 – 3.5.

David Rowland, now University of Queensland — for *Trial of the Session* (a novel setting for Newton's Law of Cooling), for his contributions to the groupwork section in the Lab Manual and for lots of other good ideas.

Barbara Catchpole and *Zlatko Jovanoski* — as lab instructors over the years, they have contributed something to all of these labs and much to many of them.

Leesa Sidhu — for the tennis lab.

Mark Collins — for the ideas and content for the two gambling labs.

All the others, some acknowledged, some undoubtedly not, whose ideas were an important part of many of the labs.

Annabelle Boag — for LaTeX typing, advice and many of the figures.

2 Labs requiring a TI-84/CE

2.1 Balloons, Submarines and Drag

Based on *A Balloon Experiment in the Classroom* by T. Gruszka, The College Mathematics Journal 25, 442–444 (1994). See also *Modelling Air Resistance in the Classroom* by A. Battye, Teaching Mathematics and its Applications 10, 32–34 (1991).

Aims

- To carry out an experiment and compare the data with several theoretical models.
- To look at the process of mathematical modelling.
- To introduce some of the uses of your calculator.
- To introduce you to the benefits of groupwork.

Introduction

In this lab, we try to fit several mathematical models to experimental data. In constructing a mathematical model, we usually start with a model based on simple assumptions and try the model against the data. If the model does not explain the data, we build in greater and greater complexity until, hopefully, we find a model that does explain the data. We would then test this model against other data to try to confirm or reject the model.

Note that you are not required to solve any equations in this lab, just to read it carefully and to put together all the information.

The Scenario

You have graduated from the Naval Academy and, in view of your excellent results in Mathematics Honours, have been invited to join the Higher Mathematics Corps.

The Navy has asked the Higher Maths Corps to help in the redesign of the shape of the new Collins-class submarine to make it faster and quieter. There is a need to minimise the drag exerted by the water on the submarine as it moves forward. Your knowledge of motion under a drag force is restricted to some work on air resistance that you did in first-year Maths, so you decide to go back to the basics you did there. Defence money for equipment is tight, so you start with some balloons left over from your birthday, a 1m length of string and your trusty TI-84 graphics calculator.

The Problem

You know that the drag force D acting on a body moving through air (or water) increases as the velocity v increases, but when is the drag force important? By finding out how long it takes a balloon to fall from several different heights and comparing your data with theory, you should be able to find out whether drag is important for the balloon.

What do you predict?

The Experiment

You will write a full report in the second week of this lab. For the first week, the scribe should keep a detailed log of what you did and your results. Don't forget to record the units of all quantities. Work as a group.

1. Collect your kit, consisting of a balloon and a 2m-long piece of string. Blow up your balloon.

2. Drop the inflated balloon so that it falls a distance (s) of 2 m; record the time it takes to fall this distance. Repeat this a few times and work out the average time. Call this time T_2.

Use the STOPWTCH/STPWCHCE program to time the balloon. Press prgm, then the number against the program name to copy the name to the Home screen and enter to execute it. Alternatively, highlight the program name with the cursor and press enter twice.

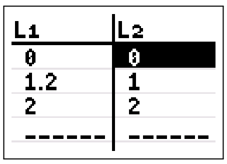
Calibrate the timer first for at least 1 minute using your watch or phone.

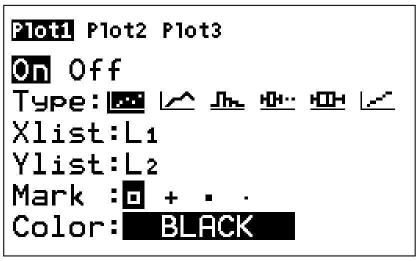
3. Repeat Step 2 with the balloon falling a distance of 1 m. Call this average time of fall T_1.

4. Graph your 3 data points $(t, s) = (0, 0)$ (*why is this a data point?*), $(T_1, 1)$ and $(T_2, 2)$ on a graph of s vs t on your calculator, as described below.

Entering and plotting data on a TI calculator

Entering: Press stat EDIT to display lists L1 – L3.[2] If you have data in lists L1 and L2, move the cursor to the heading, press clear and move the cursor down again. Store the independent variable for our graph, time t, in L1 (*value* enter), the dependent variable, distance fallen s, in L2 (figure below left).





Plotting: Press 2nd stat plot (top left key on the calculator). With enter, select *Plot1*, *On*, scatter plot (the first one), set Xlist to L1 (on the 1 key), Ylist to L2 and select the box for a marker (figure above right). On a CE, you can also choose the colour of the marker.

Press zoom 9 (ZoomStat), which automatically sets a window and graphs the data. If necessary, change the window so that the graph fills the screen.

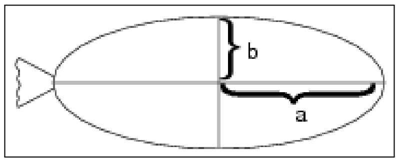
[2] If lists L1 – L3 don't appear, press stat 5 enter and try again.

To fit two of the models to your data, you will need to know the mass and volume of your inflated balloon.

5. Weigh your inflated balloon on the balance (kg).

6. Estimate the volume of the balloon by assuming it to be an ellipsoid of revolution.

Use a ruler to estimate the radii a and b (m). The volume of an ellipsoid of revolution is $\frac{4}{3}\pi ab^2$. Be careful not to confuse a and b.



Model 1 *Gravity Only*

Newton's Second Law of Motion tells us that

$$\textit{mass} \times \textit{acceleration} = \textit{sum of forces}.$$

The force due to gravity is mg, where m is the mass of the inflated balloon, and acts in the downwards direction.

We write acceleration as the first derivative of velocity with respect to time t and take velocity as positive in the downwards direction. The differential equation (equation of motion) for the velocity of the balloon according to the gravity-only model is then

$$m\frac{dv}{dt} = mg.$$

Take $g = 9.8\,\mathrm{m/s^2}$. Note units.

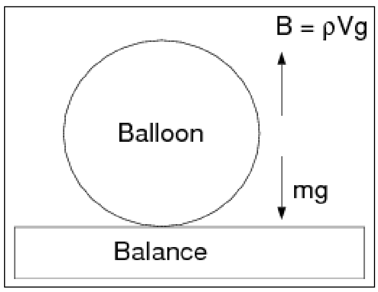
We can solve this differential equation (*you don't have to do this now*[3]) to find first $v(t)$, the velocity as a function of time, and then $s(t)$, the distance fallen as a function of time, given that the balloon is released from rest.[4]

$$s(t) = \frac{1}{2}gt^2. \tag{1}$$

Put Eq. (1) (our first model) into Y1 on your calculator: press [y=] and set[5] Y1 = $4.9X^2$. Press [graph]. *Does this model fit the data? What do you conclude about the first model?*

Model 2 *Gravity + Buoyancy*

When we weigh an inflated balloon, what we measure is less than the actual weight, because the balloon actually floats in the air. It experiences a buoyancy force B equal to the weight of air displaced by the balloon and acting upwards. The weight of the air displaced is ρVg, where ρ (Greek rho) is the density of air and V is the volume of the balloon.



[3]We would integrate both sides with respect to t to give the velocity $v(t)$ and integrate again to give $s(t)$, not forgetting the two constants of integration.

[4]Initial velocity $v(0) = 0$; initial distance fallen $s(0) = 0$ — these initial values give the constants of integration.

[5]The calculator always uses X as the independent variable in function graphs.

The effective mass M, the value you read from the balance, is given by

$$M = m - \rho V,$$

where m is the actual mass. At 20°C, the density of air is approximately 1.204 kg/m^3.

Knowing M, ρ and V, we can work out m, the actual mass of the inflated balloon, which we need for our theoretical curves. Watch units — mks system recommended.

When we drop the balloon, according to this model there are two forces, gravity and buoyancy, acting on the balloon, giving the equation of motion as

$$m\frac{dv}{dt} = mg - \rho V g = Mg.$$

Again we can solve this differential equation by integrating twice to give the distance fallen as

$$s(t) = \frac{M}{2m}\, gt^2. \tag{2}$$

Put Eq. (2) (our second model) into Y2 on your calculator: set Y2 = *coefficient* X^2, where *coefficient* is your numerical value for $Mg/(2m)$.

Does this model fit the data? Is it a better model than the first model? What do you conclude about the second model?

Model 3 *Gravity + Buoyancy + Linear Drag*

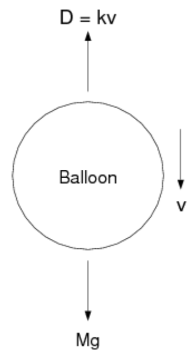
When we drop the balloon, apart from gravity and buoyancy, we also have drag (air resistance) acting to slow the balloon down. We assume the drag force D to be proportional to v (called *linear drag*), so that $D = kv$, where k is a constant which we have to find. The drag force acts in the opposite direction to the velocity.

The total force acting on the balloon is $mg - \rho Vg - kv$. The first term is the force due to gravity, the second term is the buoyancy force and the third term is the drag force.

The equation of motion according to this model is therefore

$$m\frac{dv}{dt} = mg - \rho Vg - kv = Mg - kv,$$

where M is the effective mass of the inflated balloon.



We can solve this differential equation — find the velocity function $v(t)$ and the distance-fallen function $s(t)$ — algebraically. Later on in the course you will actually do this calculation. Here we just write down $s(t)$, given that the balloon is released from rest:

$$s(t) = \frac{Mg}{k}\left(t - \frac{m}{k}(1 - e^{-kt/m})\right). \tag{3}$$

Put Eq. (3) (our third model) into Y3 on your calculator.

Tip: To make plotting easier, let k be the letter K (alpha K) when you enter Eq. (3) into Y3. Store a value for k in memory K using the sto key[6] and graph the function.

[6] *value* sto alpha K enter .

Repeat with different values for k.[7] *What should happen to the graph of Y3 as the drag coefficient k gets closer and closer to 0?* Use this to check your third model.

Does this model fit the data? Is it a better model than the first two models? What do you conclude about the third model?

Before Writing Your Report

- Read through the section on writing lab reports in the Lab Manual (page 225).
- Complete, as a group, the Exercise in the Lab Manual (page 227). Hand this in with your report.

Writing Your Report (*a group activity*)

Your report should describe in sentences (and maths where appropriate) what you did, what you discovered and what you concluded. **It should be possible for a fellow student who hasn't done the lab to follow your report without looking at this lab sheet. This is what the person who marks the lab will be looking for regarding presentation.**

A suitable format for your report is as follows.

- **Introduction:** outline the problem and reasons for doing it.
- **Methods and Results:** how you carried out each of your measurements, the data you acquired and any subsequent processing and fitting of the data.
- **Discussion:** interpret your results and talk about any questions that arise or observations you have made.
- **Conclusion:** what you, the group, concluded as a result of your measurements, calculations and deliberations.

Supplementary Questions

Model 4 *Gravity + Linear Drag*

What if we leave out the buoyancy force? Can we explain our experimental results with just gravity and linear drag?

Hint: If there is no buoyancy force, $m = M$, the measure mass of the balloon. Change Y3 to plot this model curve.

Model 5 *Gravity + Buoyancy + Quadratic Drag*

Under some circumstances, the drag force is found to be proportional to the square of the velocity, that is $D = qv^2$, where q is the quadratic drag coefficient.

The equation of motion in this case is

$$m\frac{dv}{dt} \;=\; mg - \rho V - qv^2 \;=\; Mg - qv^2,$$

where M is the effective mass of the inflated balloon.

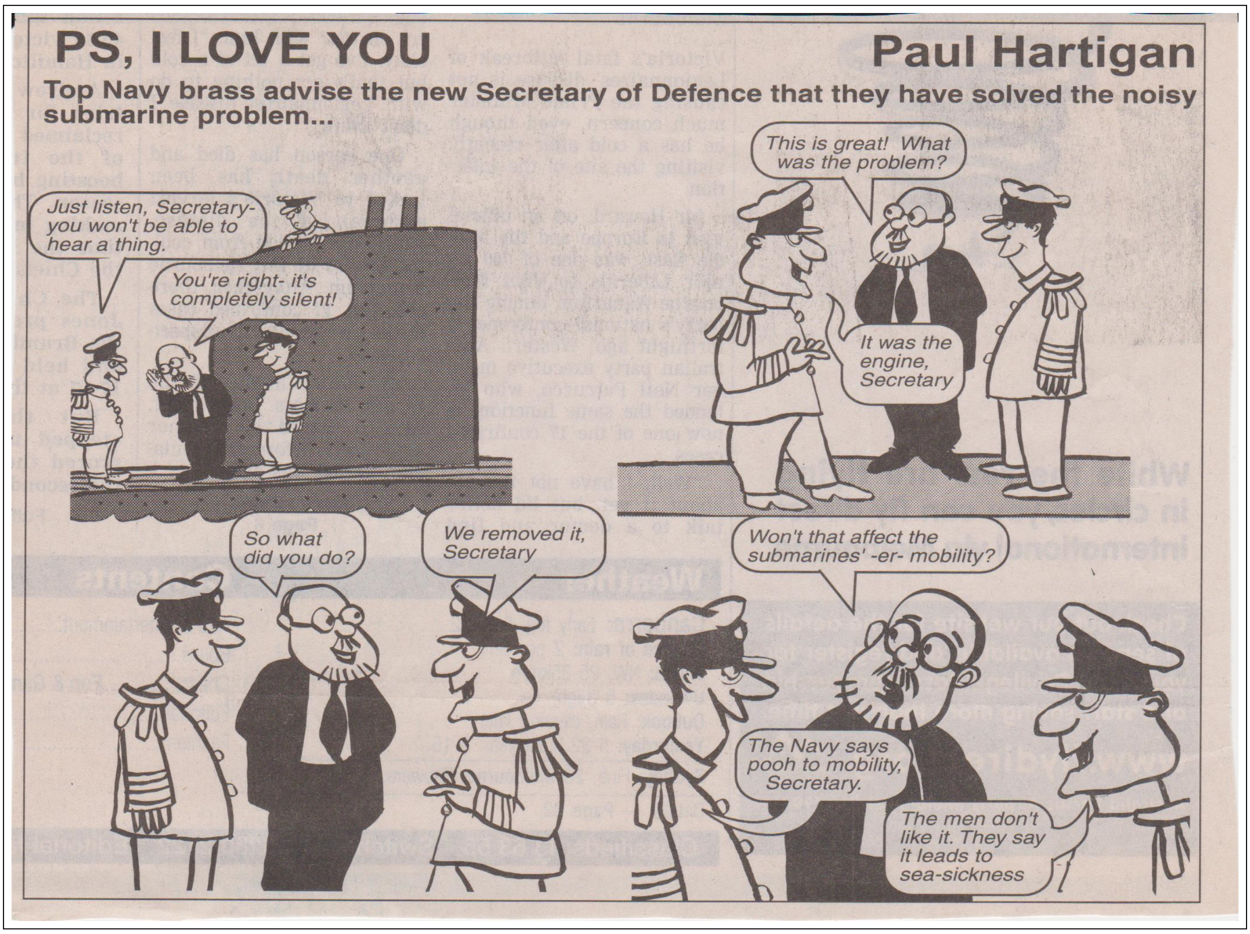
[7] You may want to turn off Y1 and Y2 while experimenting with values of k. To turn off a function, press [y=], move the cursor over the appropriate = sign and press [enter]. The same process turns the function back on again, giving a highlighted = sign. You can turn plots on/off in the same way.

This equation can be solved algebraically to give

$$s(t) = \frac{m}{q} \ln\left(\cosh\left(\frac{\sqrt{(Mgq)}}{m}\, t\right)\right). \tag{4}$$

Plot this function as Y4 on your calculator[8] and experiment to find the best value of the quadratic drag coefficient q.

Is this a better model than Model 3?



Cartoon from *The Canberra Times*

[8]The hyperbolic cosine function is defined by $\cosh(x) = \frac{1}{2}(e^x + e^{-x})$. You'll find it in the catalog menu of the TI-84/CE — press the C key, then scroll down to cosh. Press enter to select it. Then be careful with brackets.

Instructors' Guide

This lab is the first lab given to our students. It takes place over two 110-minute classes, with the first class devoted to carrying out the experiments and graphing the data. In the second class, students first read and discuss the two sample lab reports in the Lab Manual (Section 4) as an introduction to writing a lab report before writing up the balloon lab.

For the lab, you will need balloons, something (e.g. a 2m length of string) to measure 1 m and 2 m for each group and a balance that can measure masses around 2.0 g (from the Science Department?). If you can't source a balance, use a value of 2 g for the mass M of the inflated balloon.

The lab is a good 'icebreaker', as students have to carry out the relatively simple experiments together. There is also plenty of scope for questions, so the lab instructors soon become involved.

The scenario is clearly set up for students at the Australian Defence Force Academy who are all members of the Australian Defence Force. However, it should not be too difficult to adapt the scenario to the local situation.

Solutions

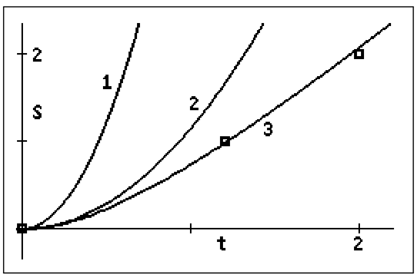
The data generated may vary depending on the type of balloon and how much it is inflated. Below is a summary of a set of data obtained by one of our lab groups. In our experience, most of the numbers seem to come out roughly the same.

Timing the balloon falling from 1 m and 2 m is not an exact science, but with averaging over a few trials, reasonable results can be obtained. In our example, the times were $T_1 = 1.2$ s, $T_2 = 2.0$ s. The STOPWTCH/STPWCHCE program for the TI-84/CE is available at *www.XXX*. It is an easy way to provide stopwatches for every group without having to raid the Physics Department.

The inflated balloon gave a reading of $M = 2.0$ g on an electronic balance; this is 0.002 kg. The measured radii of the balloon were $a = 13.5$ cm or 0.135 m, and $b = 9.75$ cm or 0.0975 m, giving a calculated volume of $V = 5.38 \times 10^{-3}$ m^3 using the ellipsoid method. Be very careful of units here.

The actual mass of the balloon $m = M + \rho V = 0.002 + 1.204 \times 5.38 \times 10^{-3} = 0.0085$ kg.

Putting these numbers into the three models and plotting distance fallen s versus time t and the data points gives the figure below.



Clearly neither Model 1 nor Model 2 fits the data, so drag (air resistance) is important, as we might expect. In Model 3, the drag parameter k was chosen to give the best fit to the data; $k = 0.013$ gave a good fit here.

We conclude that air resistance is important in modelling a falling balloon and that Model 3, incorporating linear drag, provides a good fit to the experimental data.

The equation for $s(t)$ for Model 4 is just that of Model 3, but with $m = M$, i.e.

$$s(t) \;=\; \frac{Mg}{k}\left(t - \frac{M}{k}(1 - e^{-kt/M})\right).$$

Plotting this with $k = 0.0195$ gives a reasonable fit to the data, not quite as good as that of Model 3 but not significantly worse.

Model 5 with drag coefficient $q = 0.013$ gives a very good fit to the data in this case.

The data allow us to determine that Models 1 and 2 (no drag) are not valid. However, the data do not allow us to conclude that any of Models 3, 4 or 5 is significantly better than the others, although in this case Model 5 gave the best fit, followed by Model 3 and then Model 4. More data, perhaps measured more accurately, are needed to decide which of Models 3, 4 or 5 is best. Perhaps you could try dropping the balloon from 2.5 m and/or 3 m as well.

2.2 Understanding the Derivative

Based on a lab by Ruth Hubbard from the Queensland University of Technology.

Aims

- To show that the *difference quotient*, giving the average rate of change over a small interval, is an approximation to the *derivative*, giving the instantaneous rate of change.
- To demonstrate that when a small section of most curves is magnified, the curve looks like a straight line.
- To use the numerical derivative, *nDeriv* on a TI-84/CE, to calculate the *symmetric difference quotient*, a better approximation to the derivative than the difference quotient.

Introduction

The idea of derivative is central to Calculus. The derivative of a function f at a particular point $\big(x, f(x)\big)$ can be thought of as the *slope of the tangent* to the curve at that point,[9] as the (instantaneous) *rate of change* of the function at that point or as the value $f'(x)$, where f' is the function obtained by differentiating the original function f.

Only by the *algebraic process of differentiation* can we obtain *the exact value of the derivative*, and we can only carry out differentiation if we have a formula for the function. If we do not have a formula (for example, if the function is given by a graph or a table) or if we want to estimate the derivative at a point *graphically or numerically*, we can only find *an approximation to the derivative.* This lab is devoted to methods of approximating the derivative, numerically and graphically; to understanding the nature and accuracy of such approximations; and, through such methods, to helping you gain a better understanding of the derivative.

In this lab, you will estimate the *derivative / instantaneous rate of change / slope of the tangent* of the function $f(x) = x^x$ at the point on the graph of f for which $x = 2$.

Question 1 *Using the difference quotient*

The expression for the *average* rate of change of a function f over the interval from x to $x+h$ is the difference quotient

$$\frac{f(x+h) - f(x)}{h}.$$

(a) Plot the function $f(x) = x^x$ in Y1 using [window] parameters $[0, 3, 1] \times [0, 6, 1]$. Zoom in several times near the point on the graph for which $x = 2$ until the graph looks like a straight line $\big($[trace] to X = 2 and [Zoom] [2] [enter] several times$\big)$.

Use [trace] to find the coordinates of two points on either end of this 'line' and hence calculate the slope of the straight line through these two points. Zoom in again near the point $(2, 4)$ and repeat the calculation. Do this a third time. A table in your report — *zoom number* vs *slope* — might be a good idea here. *Based on the table values*, what is your best estimate for the derivative / instantaneous rate of change / slope of the tangent of the function $f(x) = x^x$ at $x = 2$, and to how many significant digits is your answer accurate?

[9] or we just say the *slope of the curve* at that point

(b) **(i)** On the graph of f provided (page 14), show how the difference quotient is represented. Relate this to what you measured in (a).

(ii) Use the difference quotient to estimate numerically the derivative of f at $x = 2$, starting with $h = 0.1$.

One way to do this is to evaluate $\big(\text{Y}_1(2+0.1) - \text{Y}_1(2)\big)/0.1$, then use the [entry] key to recall the expression, change the value of h (in two places) and re-evaluate.[10]

Decrease h successively by a factor of 10 ($h = 0.01$, 0.001, 0.0001, etc) until your difference-quotient approximation no longer changes in the fifth *significant digit*[11] (after rounding).

In your report, draw up a table with a column for h, starting at 0.1, and a column for the approximation to the the derivative using the difference quotient. Allow space for a third column which you will need in Question 2.

Question 2 *Using the symmetric difference quotient*

The TI-84/CE has a built-in operation *nDeriv* to calculate the average rate of change of a function f over the interval from $x-h$ to $x+h$ (compare with the difference quotient):

$$\text{average rate of change} = \frac{f(x+h) - f(x-h)}{(x+h) - (x-h)} = \frac{f(x+h) - f(x-h)}{2h}.$$

This expression, called the *symmetric difference quotient*, usually gives a more accurate approximation to the instantaneous rate of change at the point $\big(x, f(x)\big)$ than does the difference quotient (for the same h), because it is equal to the mean of the average rates of change on either side of the point.

(a) On the graph on page 14, show how the symmetric difference quotient is represented. Also show how the derivative or gradient, to which both the difference quotient and the symmetric difference quotient are approximations, is represented.

(b) *nDeriv* takes four arguments: a *function*; the *variable to differentiate with respect to* (X here); the *value of this variable* to estimate the derivative at; and a *value for h*.

Use *nDeriv* ([math] [8]) to estimate the derivative of $f(x) = x^x$ at $x = 2$, as follows.

- The function to use in *nDeriv* is Y_1 or X^X, and our starting value for h is 0.1. Evaluate $\text{nDeriv}(\text{Y}_1, \text{X}, 2, 0.1)$ or $\text{nDeriv}(\text{X\^{}X}, \text{X}, 2, 0.1)$.
- In the table you started in Question 1(b)(ii), add a column for the approximation to the derivative using the symmetric difference quotient.
- Now decrease h successively by a factor of 10 until your approximation no longer changes in the fifth significant digit. Again, [entry] after each calculation will save retyping the nDeriv expression each time you change h. You don't need to put in the closing bracket.

[10] Y_1 is [vars] Y-VARS [1] [1].
[11] not fifth decimal place.

Question 3 *Reflect and report*

(a) What particular quantity have you *estimated* in Question 1 and 2? What is your best estimate for this quantity? How accurate is your estimate? How do you know?

(b) It should be clear that the quantities you found in 1(b) and 2 are approximate, but why is it approximate in 1(a)? How could you make the approximation in 1(a) more accurate?

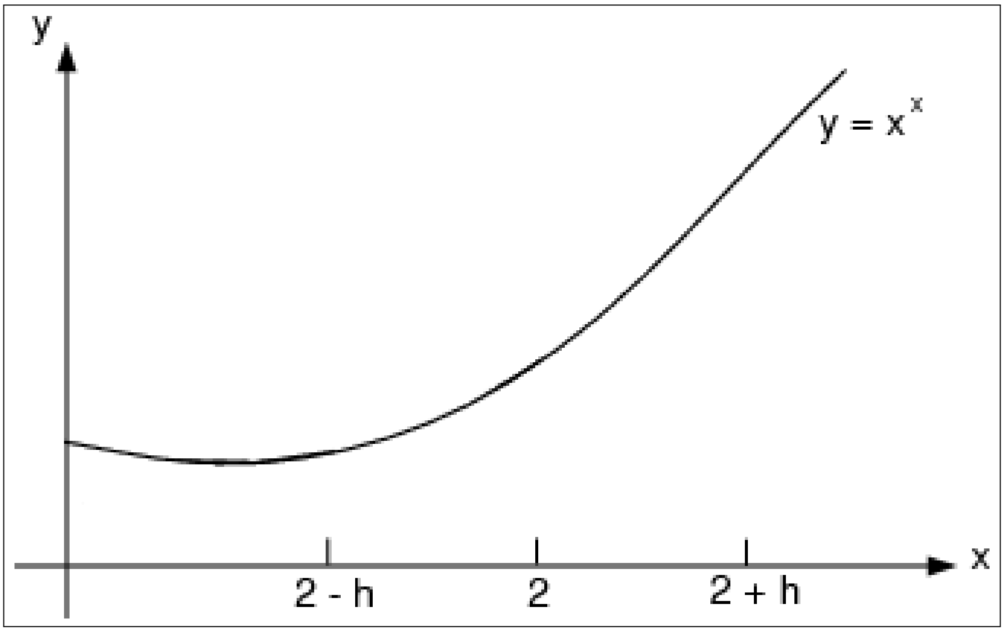
(c) Comment on the relative accuracies of the difference quotient and symmetric difference quotient for a given value of h.

(d) Given that the *exact* derivative of f is $f'(x) = x^x\big(1 + \ln(x)\big)$,[12] check the accuracy of your approximations in Questions 1 and 2 and comment.

[12] You might like to show this when you've done the rest of the lab. Remenber that any positive function f can be written as $f(x) = e^{\ln(f(x))}$.

Representing Difference Quotients and the Derivative

Please hand in one copy of this page with your report.



Instructors' Guide

Solutions

1. **(a)** With ZOOM FACTORS 4 (the default setting), we obtained the following table for the slope of a line between two points roughly at the top and bottom of each screen respectively. All values are rounded to 4 significant digits. Clearly the numbers you obtain will vary, but they should all tend to the same limit.

ZOOM #	Slope
3	6.758
4	6.765
5	6.771
6	6.772
7	6.773
8	6.773

We conclude that $f'(2)=6.773$, accurate to four significant digits.

Hint: In carrying out the calculations here, we can make use of the fact that the cursor coordinates are stored in memories X and Y. Trace to the first point, quit and execute the command X→U: Y→V. Trace to the second point, quit and execute the command (Y−V)/(X−U) to calculate the slope. Now zoom in, trace to the first point, quit and use [2nd] [entry] twice to recall the first command. Execute it, trace to the second point, quit and use [2nd] [entry] twice to recall the second command. Execute it to give the second slope value. Keep repeating the these steps to calculate subsequent slopes.

(b) **(i)** The difference quotient we want is the *slope* of the line between the points $\big(2, f(2)\big)$ and $\big(2+h, f(2+h)\big)$. In (a), we measured slopes of (secant) lines between two points on the curve, but $\big(2, f(2)\big)$ was probably not one of the points. However, the slopes of the secant lines that we measured will be very close to the difference quotient we want, providing h is small. The smaller h, the more nearly parallel the secant lines is to the tangent line.

(ii) The values for the difference quotient (DQ) and symmetric difference quotient (SDQ) $\big($Question 2(b)$\big)$ are given below.

	$f'(2) \approx$	
h	DQ	SDQ
0.1	7.4964	6.8203
0.01	6.8404	6.7731
0.001	6.7793	6.7726
0.0001	6.7732	6.7726
0.00001	6.7727	—
0.000001	6.7726	—
0.0000001	6.7726	—

We conclude that $f'(2) = 6.7726$, accurate to five significant digits. See the hint on the next page.

Hint: A quick way to carry out these calculations is to execute the command 0.1 → H: (Y1(2+H)−Y1(2))/H. Use [2nd] [entry] to recall the command, change the value for H and re-execute.

2. **(a)** The symmetric difference quotient is the *slope* of the line between the points $\big(2-h, f(2-h)\big)$ and $\big(2+h, f(2+h)\big)$.

(b) See the table above.

3. **(a)** We have estimated the derivative of f at $x=2$ using the difference quotient and the symmetric difference quotient. The best estimate, obtained in this case with all three methods, is $f'(2)=6.773$. This is probably accurate to this number of digits because we obtain the same value from two successive approximations from each method.

(b) The estimate in 1(a) is approximate because we actually calculate the slopes of secant lines (joining the two chosen points) as approximations to the slope of the tangent line. To improve accuracy, zoom in more and more to make the portion of the curve shown on the screen straighter and straighter, that is to make the secant lines closer and closer to the tangent line.

(c) The symmetric difference quotient is more accurate than the difference quotient for a given h value.

(d) $f'(2)=4\big(1+\ln(2)\big)=6.7726$ to 5 significant digits, in agreement with the results from our approximate methods.
Hint: To differentiate x^x, rewrite it as $e^{\ln(x^x)}=e^{x\ln(x)}$.

2.3 The Derivative as a Function

Based on a lab by Ruth Hubbard from the Queensland University of Technology.

Aims

- To use the difference quotient and the symmetric difference quotient to plot approximations to the derivative of a given function.
- To understand how the behaviour of a function is reflected in the behaviour of its derivative.
- To investigate a case in which the derivative does not exist.

Questions 2 and 3 are on the basics of the relationship between a function and its derivative; these are fundamental.

Question 1 *The derivative as a function: three methods*

In the lab *Understanding the Derivative*, we investigated the instantaneous rate of change of a function, that is the rate of change **at a particular value** of the independent variable (t or x), using the difference quotient and the symmetric difference quotient as approximations.

In this question we will look at the instantaneous rate of change **as a function itself**, defined for each value of the independent variable in its domain, a subset of the domain of the original function.

We shall start by using the difference quotient to approximate the instantaneous-rate-of-change function. We have to write the formula for the difference quotient at an arbitrary value x, instead of at a particular value such as $x=0.5$ or $x=2$. Thus, for a function f, the **difference-quotient function** $d(x)$ (which calculates the average rate of change of f in the interval from x to $x+h$) is given by

$$d(x) = \frac{f(x+h) - f(x)}{h}.$$

Previously, we chose a value for x (we called it a) and varied h. Now we will fix h at 0.001 and let x vary, so that d will become

$$d(x) = \frac{f(x+0.001) - f(x)}{0.001}.$$

With this value of h, we expect $d(x)$ to be a reasonable approximation to the instantaneous rate of change of f at x (but beware nasty functions!).

As you know, the algebraic process of finding the instantaneous rate of change is called **differentiation**, and the instantaneous-rate-of-change function is called the **derivative**.

PTO

Let $f(x) = x^3 - x + 1$.

(a) Write down the difference-quotient function $d(x)$ for this function, program $d(x)$ into Y2 and plot its graph for $-2 < x < 2$.

- Don't try to expand $d(x)$ out and simplify it — this takes time and can lead to errors. Program it just as you first wrote it down. Better still, program f into Y1 and use Y1 in the expression for $d(x)$ that you program into Y2.[13]
- Choose window parameters so that the graph fills most of the screen and so that the cursor coordinates and function formula don't obscure the graph when you are using trace.

(b) Write down the expression for the **symmetric-difference-quotient function** for f, call it $s(x)$, corresponding to the difference-quotient function $d(x)$ above.

Use *nDeriv* to plot the symmetric difference quotient for f as follows. Set Y3 = nDeriv(X³–X+1, X, X, 0.001), or nDeriv(Y1, X, X, 0.001) if you have f in Y1. Graph $d(x)$ and $s(x)$ together.

(c) Finally an algebraic approach. Differentiate f algebraically to find the **derivative function** f'. Write down the expression for $f'(x)$, program it into Y4 and graph it with $d(x)$ and $s(x)$.

(d) Describe what you see when you plot the three graphs together, and explain it. Use trace to explore the graphs.

(e) Repeat the process from (a) down with a larger value of h, say 0.5. Again describe what you see and explain it.

Question 2 *The relationship between a function and its derivative*

- Program the function $f(x) = e^{\sin(x)}$ into Y1 (Radian mode!).
- Set Y2 = nDeriv(Y1, X, X, 0.0001) to give a numerical approximation to f'.
- Plot Y1 and Y2 together using window parameters $[0, 2\pi, \frac{\pi}{2}] \times [-2, 4, 1]$.
- Sketch both functions.

Answer the following questions by looking carefully at these two graphs. You might find *zero*, *maximum* and *minimum* in the calc menu useful.[14]

(a) Over what intervals is the graph of f **increasing**, that is rising as x increases?

(b) Over what intervals is the graph of f' **positive**?

(c) Over what intervals is the graph of f **decreasing**, that is falling as x increases?

(d) Over what intervals is the graph of f' **negative**?

[13] Y2 = (Y1(X+0.001)−Y1(X))/0.001.

[14] These all work the same way: specify a left bound for X by moving the cursor along the graph to the appropriate place or by typing in a value and pressing enter; then a right bound and a guess in the same way. *How accurate?* — see the TI Guidebook.

(e) What are the coordinates of the high point (**local maximum** of f) and the low point (**local minimum** of f) of the graph of f?

(f) At what values of x does the graph of f' **cross the x axis**?

Question 3 *Putting the relationship into words*

(a) On the basis of your results in Question 2, write a statement describing the relationship between f and f'. Does this apply to any function and its derivative?

(b) **(i)** In terms of *increasing* and *decreasing*, what characterises a local maximum and what characterises a local minimum of a function f?

(ii) What is the corresponding behaviour of f' in each case that allows you to distinguish a local maximum of f from a local minimum?

Supplementary Questions

Question 4 *Do derivatives always exist?*

So far in the course, we have considered functions and points for which the derivative was defined. Sometimes, however, this is not the case.

Plot the function $f(x)=|x|$ (Y1 = abs(X)[15]) using a [window] of $[-1,1,1] \times [-0.2,1,0.2]$. We are interested in the derivative of f or, equivalently, in the slope of the tangent to the graph of f, at values of x around $x=0$. Turn off your axes ([2nd] [format]), so that they do not obscure the graph at this point. We will use a graphical approach, a numerical approach and finally an algebraic approach. *Make sure you read right through each part before answering the questions.*

(a) **(i)** Zoom In a number of times on f near or at $x=0$. What should happen when you Zoom In on a point on a graph at which a derivative exists? What happens with f? What do you conclude?

(ii) Now compare the behaviour of f with that of $g(x) = \sqrt{x^2+0.00001}$. Plot f and g together, starting with the original [window] $[-1,1,1] \times [-0.2,1,0.2]$. Find both functions using [trace]. Then Zoom In as you did in (i). Describe what happens. What do you conclude about the derivatives of f and g at $x=0$?

(b) **(i)** Use the symmetric difference quotient (*nDeriv*) to *calculate* an approximation to the derivative of f at $x = -1$, 0, 1. Comment on what you find and how it fits in with your results in (a).

- Remember to try decreasing h values in each case to determine accuracy. Draw up the usual table. Think here! Comment.
- For each x value, draw a sketch of f and show what the symmetric difference quotient is calculating.[16] Do some more thinking.

[15] Where is *abs* on the TI-84/CE? In the [math] NUM menu.

[16] Remember, the symmetric difference quotient calculates the slope of a particular line. Draw the line.

(ii) Then use *nDeriv* to *plot* an approximation to the derivative. Set Y2 = nDeriv (Y1, X, X, 0.0001). Think about mode settings. What should you do when you see what looks like a vertical line on a graph? Sketch f and its derivative. Comment on what you find and how it fits in with your results in (a) and (b)(i).

(c) Write down the difference quotient $\big(f(x+h)-f(x)\big)/h$ for $f(x)=|x|$ at $x=0$ and with a general h, i.e. just leave it as h. Simplify your expression as much as possible. To get rid of the absolute-value signs, you will have to consider two cases, h positive and h negative.

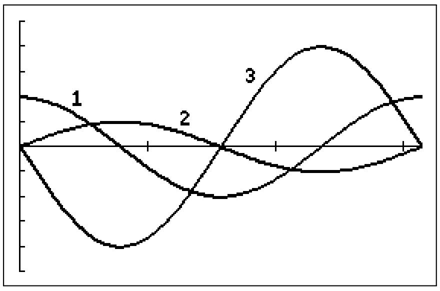
- Does this result support what you found in (b)(i)?
- The derivative at $x=0$ is given by the limit as $h \to 0$ of the difference quotient. What is the value of this limit here? For a limit to exist, it must be unique. Does this result support your conjectures in (a) and (b)?

(d) What are your final conclusions regarding the derivative of f, particularly at $x=0$?

Question 5 *Which is the derivative?*

In the figure below:

- identify which is the function, which is the derivative and which is the second derivative, explaining your reasoning;
- prove that it could not be another way around.



Note that it is *not* sufficient just to say something like “when f' is positive, f is increasing” since this is *always* true. You have to pin such assertions to specific intervals relevant to the given function.

Question 6 *Do derivatives always exist #2?*

(a) As a preliminary to (b), let's think about how to calculate on the calculator a function in which there is a one-third power or cube root. Take the simplest example: plot using a window of $[-3,3,1]\times[-2,2,1]$ the following functions one at a time.

(i) $f(x)=x^{1/3}$ X^(1/3)

(ii) $f(x)=\sqrt[3]{x}$ the cube root is math 4

(iii) $f(x)=x^{0.333333333333}$ (twelve 3s)

(iv) $f(x)=x^{0.3333333333333}$ (thirteen 3s).

What do you observe? Which graphs are identical (use trace)? Which are correct? Can you think why (iii) and (iv) are the way they are?

For thinking about and experimenting some other time: Does it happen with other powers of the form 1/(an integer), 2/(an integer), . . . ?

(b) Let $f(x) = x^{1/3}$. Discuss the behaviour of the derivative of f. Use the same tools you used in Question 4, that is a graphical approach (with axes turned off), a numerical approach and an algebraic approach. In particular,

- for what values of a does $f'(a)$ exist and for what values does it fail to exist?
- if it fails to exist, why does it fail to exist?

(c) Describe and explain the behaviour of the second derivative f'' of $f(x) = x^{1/3}$ (see the procedure below) in terms of the behaviour of

(i) f (concave up/down)

(ii) f' (increasing/decreasing).

Set up the function and its derivatives in your calculator as follows: f in Y1; (an approximation to) its derivative f' as Y2 = nDeriv(Y1, X, X, 0.0001); and (an approximation to) its second derivative f'' (derivative of f') as Y3 = nDeriv(Y2, X, X, 0.0001). Plotting the second derivative in this way is rather slow as the calculator has to evaluate the function f four times for each point.

Alternatively, you may find f' and f'' algebraically, then plot them (plots much faster). Be careful with programming them. You should check your algebraic results either graphically or numerically with nDeriv.

Question 7 *Calculator-aided delusions*

Consider the function $f(x) = |x|^x, \quad x \neq 0$.

Plot the function on your calculator using a [window] of $[-2, 2, 1] \times [0, 4, 1]$. Make sure you have brackets in the right places in the function on your calculator. Check a couple of points by hand to make sure.

(a) "Evaluate" the function at $x = 0$ using your calculator. What value should you put in for $f(0)$ to make the function continuous, i.e. what is $\lim_{x \to 0} f(x)$?

(b) Investigate the slope at $x = 0$ graphically and numerically.

(c) Do the problem algebraically (differentiate the function) and show that your assertions in (b) were correct.

Note that for any positive function f, $f(x) = e^{\ln(f(x))}$ and that $\dfrac{d}{dx} \ln|x| = \dfrac{1}{x}$.

Instructors' Guide

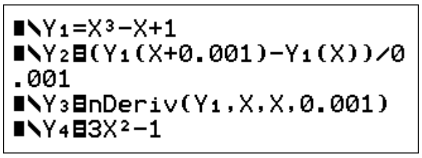
Solutions

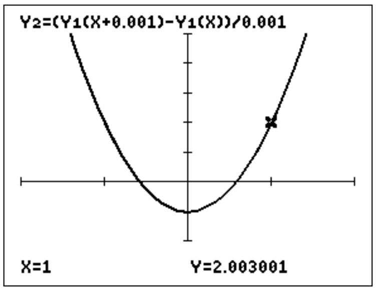
1. **(a)** $d(x) = \dfrac{(x+0.001)^3 - (x+0.001+1 - (x^3 - x + 1)}{0.001}$.

(b) $s(x) = \dfrac{(x+0.001)^3 - (x+0.001) + 1 - \big((x-0.001)^3 - (x-0.001) + 1\big)}{0.002}$.

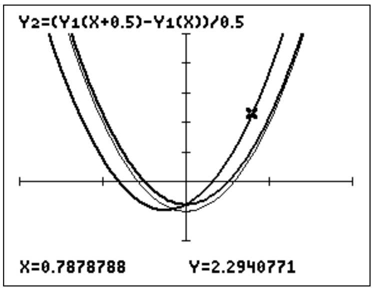
(c) $f'(x) = 3x^2 - 1$.

(d) The three graphs are indistinguishable on the screen when $h = 0.001$, although trace reveals small differences. With a window of $[-2,2,1]\times[-2,5,1]$, we get



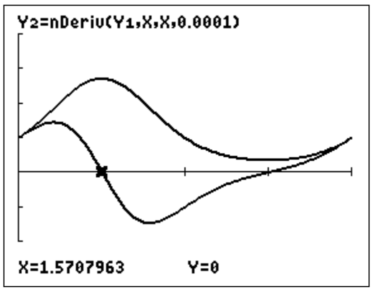


(e) With a relatively large h, say $h = 0.5$, we expect the difference quotient and the symmetric difference quotient to be less accurate approximations to the derivative than in (d). As it turns out (see the figure below), the symmetric difference quotient is still a reasonable approximation (the thin line in the figure is the graph of the derivative), but the difference quotient (with the cursor on its graph) is noticeably less accurate.



PTO

2. Graphing the two functions:



(a) The function f is increasing for $0 < x < 1.57$ $(\pi/2)$ and for $4.71\,(3\pi/2) < x < 2\pi$. Use *maximum/minimum* on the calculator.

(b) The graph of f' is positive over the same intervals. Use *zero* on the calculator.

(c) The function f is decreasing for $1.57 < x < 4.71$.

(d) The graph of f' is negative over the same intervals.

(e) The local maximum of f is at $(1.57, e)$, the local minimum of f at $(4.71, e^{-1})$.

(f) The graph of f' crosses the axis at $x = 1.57$ and $x = 4.71$.

3. **(a)** When f is increasing, f' is positive; when f is decreasing, f' is negative. This applies to any function which is differentiable.

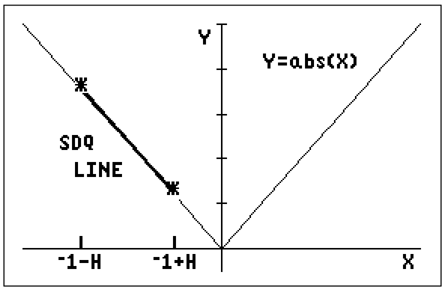
(b) **(i)** A local maximum is characterised by the function first increasing, then decreasing as x increases. A local minimum is characterised by the function first decreasing, then increasing as x increases.

(ii) As x increases through a local maximum, f' is positive, then negative. As x increases through a local minimum, f' is negative, then positive.

4. **(a)** **(i)** If the derivative of a function exists at a point, the graph of the function will become straight if you zoom in a sufficient number of times. In the case of f here, the corner in the graph remains, no matter how many times you zoom in. This suggests the derivative of f does not exist at $x = 0$.

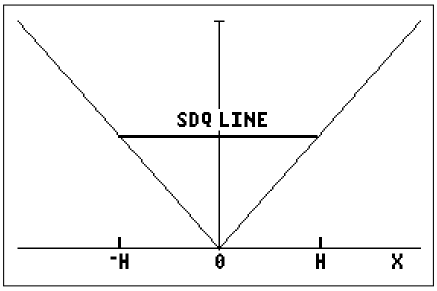
(ii) The graphs of f and g coincide, with the given window. After zooming in a few times, the graph of g becomes rounded at $x = 0$ and eventually straight, whereas the corner remains in the graph of f. It looks like $g'(0)$ exists, but $f'(0)$ does not.

(b) **(i)** The symmetric difference quotient (*nDeriv*) gives $f'(-1) = -1$ for all values of h down to 10^{-12}. For $h = 10^{-13}$, *nDeriv* gives 0, because of roundoff error.
The symmetric difference quotient calculates the slope of the straight line through the points $(-1-h, |-1-h|)$ and $(-1+h, |-1+h|)$. For $h \leq 1$, this line coincides with the graph of f. Hence, the SDQ value for $f'(-1)$ is exact.



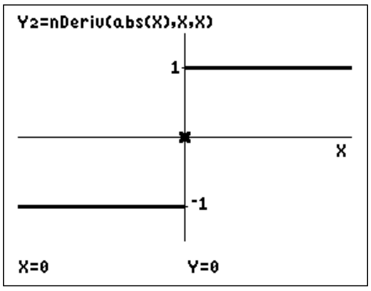
The same considerations apply to the calculation of $f'(1)$, for which we obtain the value 1.

The SDQ gives $f'(0) = 0$, for all values of h. Does this make sense? The symmetric difference quotient calculates the slope of the straight line through the points $(-h, |-h|) = (-h, |h|)$ and $(h, |h|)$. Because the two y values are the same (independent of the value of h), the line is horizontal and hence its slope is 0.



Although the SDQ does give a value for $f'(0)$, it is clear from the sketch why and it is also clear that this value is not necessarily correct at $x = 0$.

(ii) A vertical line usually means a jump or discontinuity in the graph of the function: the calculator joins points on either side of the discontinuity to try to draw a connected graph. Change to Dot mode to see just the calculated points.



The derivative clearly has a discontinuity at $x = 0$, at which it jumps from -1 for $x < 0$ to 1 for $x > 0$. This is consistent with the graph of the function and our deliberations in (b)(i). trace shows a calculated f' value of 0 at $x = 0$, but as the graph is drawn using values from *nDeriv*, this is not surprising. The existence or otherwise of $f'(0)$ is still uncertain.

(c)

$$\begin{aligned} \frac{f(x+h) - f(x)}{h} &= \frac{|h| - |0|}{h} \\ &= \frac{|h|}{h} \\ &= \begin{cases} 1 & h > 0 \\ -1 & h < 0 \end{cases} . \end{aligned}$$

Now $f'(0) = \lim_{h \to 0} \big(|h|/h\big)$. If we approach $x = 0$ from below ($h < 0$), we get -1 for the limit; if we approach $x = 0$ from above ($h > 0$), we get $+1$. Hence we get different values for the limit, depending on from which direction we approach $x = 0$. The limit is not unique, and therefore does not exist.

We conclude that $f'(0)$ does not exist, consistent with (a) and our suspicions in (b).

(d) The derivative of $f(x) = |x|$ exists for $x \neq 0$: $f'(x) = -1$ for $x < 0$ and $f'(x) = 1$ for $x > 0$.

5. Consider Curves 1 and 2. Curve 1 initially has a negative slope, so Curve 2, which is positive, cannot be the derivative of Curve 1. Curve 1 could be the derivative of Curve 2 because it is positive where Curve 2 has a positive slope, negative where Curve 2 has a negative slope and zero at the maximum and minimum of Curve 2.

Now consider Curves 1 and 3. By the same considerations as above, Curve 1 cannot be the derivative of Curve 3, but Curve 3 could be the derivative of Curve 1 and therefore the second derivative of Curve 2.

Looking at Curves 2 and 3, we see that Curve 2 cannot be the derivative of Curve 3, because Curve 2 is positive initially whereas Curve 3 has a slope that is initially negative. The same argument about initial slope shows that Curve 3 cannot be the derivative of Curve 2.

We conclude that Curve 1 is the derivative of Curve 2 and that Curve 3 is the derivative of Curve 1, and therefore the second derivative of Curve 2.

The actual function is $f(x) = \sin(2x)$, plotted for $0 < x < \pi$, $-5 < y < 5$.

6. **(a)** On a TI-84/CE, graphs (i), (ii) and (iv) are the same — they all plot the function for negative values of x. Graph (iii) plots the function only for non-negative values of x.

The calculator is smart enough to work out that $\sqrt[3]{x}$ is defined for negative values of x. $x^{1/3}$ is just another way to represent the cube root and 0.3333333333333 (thirteen 3s) is interpreted by the calculator as 1/3. However, it takes 0.333333333333 (twelve 3s) as being different to 1/3 and needs to use logs to work out the function values, thus excluding negative exponents.

(b) Zooming in on the origin on the graph of $f(x) = x^{1/3}$ eventually produces a vertical line, suggesting the derivative at $x = 0$ is infinite. For other values of x, after zooming in we obtain (approximately) straight lines, with slope depending on x.

Using the symmetric difference quotient produces larger and larger numbers as h becomes smaller and smaller.

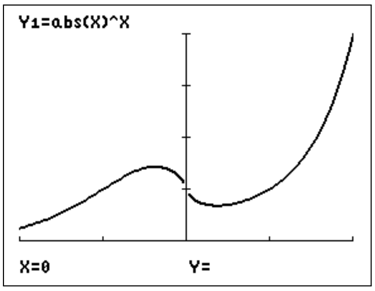
The derivative $f'(x) = 1/3x^{2/3}$, which goes to ∞ as x approaches 0 from above or below.

We conclude that the derivative of $f(x) = x^{1/3}$ does not exist at $x = 0$ because it is infinite, but exists for all other values of x.

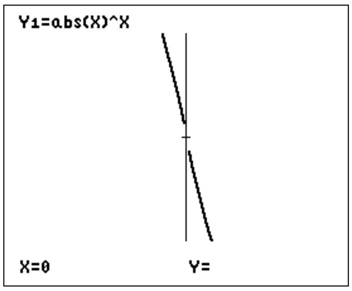
(c) Graphing the second derivative shows that $f''(x) > 0$ for $x < 0$, corresponding to the function f being concave up and the function f' increasing; for $x < 0$, $f''(x) < 0$, f is concave down and f' is decreasing. The second derivative changes sign at $x = 0$ (although $f''(0)$ is not defined), so that $x = 0$ is a point of inflection of f.

PTO

7. **(a)** The calculator gives no value for $f(0)$. However, it is clear from the graph that $\lim_{x \to 0} f(x) = 1$.



(b) The graph looks quite well-behaved, except for the missing value at $x=0$. However, after zooming in a few times, we see that the curve becomes steeper and steeper near $x=0$ and suspect that the slope may become infinite as $x \to 0$.



Using the symmetric difference quotient, we find that the slope 'at' $x=0$ slowly becomes more and more negative as h is decreased, until eventually round-off error gives a value of 0. There is clearly no convergence to a finite value for the slope 'at' $x=0$. In the figure below, the first column contains the H values, the second column the corresponding symmetric difference quotients.

0.1	-2.323
0.01	-4.607
0.001	-6.908
1E-4	-9.21
1E-5	-11.51
1E-6	-13.82
1E-7	-16.12
1E-8	-18.42
1E-9	-20.72
1E-10	-23.03
1E-11	-25.33

(c) $f'(x) = |x|^x \big(1+\ln(|x|)\big)$. Therefore, $f'(0)$ is not defined, as both $|x|^x$ and $\ln(|x|)$ are not defined at $x=0$. Because of the $\ln(|x|)$ term, $f'(x) \to -\infty$ as $x \to 0$, as we suspected from (a) and (b).

2.4 Graphics and Calculus: Families of Curves

Aim: To use both graphics and Calculus to study the behaviour of families of curves.

Our study of Calculus has given us experience in using the first and second derivatives to study the qualitative properties of a function. In this lab, we study two *families of functions*, one the Normal Distribution, which arises frequently in Probability and Statistics, the other used in modelling damped oscillatory motion.

In modelling some phenomenon, a crucial first step involves recognising which families of functions might fit the available data. You can think of our work in this lab as an introduction to some of the skills needed in mathematical modelling, without much discussion of the phenomena being modelled.

Here we will be looking at families of functions, initially using the FAMILY2/FAMLY2CE program in the calculator to see how changing parameters in a function changes the form of the function's graph. However we will also be using Calculus to draw some general conclusions about the entire family of such functions. Calculus is most useful when applied to families rather than specific functions; specific functions can be analysed individually using the calculator graphics, but to understand dependence on parameters you need general laws.

A. Curves of the form $f(x)=e^{-(x-a)^2/b}$

This is a constant multiple of the *normal density* function or *Normal Distribution* used in Probability and Statistics, and the Gaussian curve of Physics and Mathematics. The parameter a corresponds to the mean μ and b to twice the square of the standard deviation σ of the distribution. This function has wide application and was probably used, for example, in scaling your Year 12 results.[17]

Here we will consider the function as a two-parameter family, depending on a and b, and we will assume $b>0$.

Question 1 *Doing it graphically*

(a) Graph some representative members of this family using the FAMILY2/FAMLY2CE program, first by keeping $b=1$ and varying a (try say 0, ±1, ±2).

Set Y1 = e^(⁻(X−A)²/B). Use a [window] of $[-5,5,1]\times[0,1.2,1]$.

Press [on] [1] to stop the program.

What is the effect on the graph of varying a? Explain this by referring to the formula for f.

(b) Now keep $a=0$ and vary b (>0). What is the effect of varying b? Explain this too by referring to the formula for f.

(c) What is the limit of $f(x)$ as $x\to\pm\infty$? Does it depend on a or b?

[17] In applications, f is usually multiplied by a positive constant such that the total area under the graph of $f(x)$ from $x=-\infty$ to $x=\infty$ is equal to 1.

Question 2 *Now we go algebraic*

(a) First the critical points.

(i) Find the critical points of f in terms of a and b.

(ii) Classify each critical point as a local minimum, local maximum or neither. The first-derivative test is straightforward here.

(iii) Does your result agree with what you found graphically in Question 1?

(b) What is the global maximum of f on $(-\infty, \infty)$ in terms of a and b? the global minimum?

(c) Find, in terms of a and b, the x coordinates of the points where $f''(x)=0$.

- A point at which the graph of a function f changes concavity is called a *point of inflection* of f. Equivalently, a point of inflection is where the first derivative has a local maximum or minimum, and therefore where the second derivative f'' **changes sign**. Mostly (but not always) this is where the second derivative is zero.

  Are the points you just found points of inflection? *Hint*: What does the graph of the quadratic that arises here look like?

(d) Sketch a representative graph from the family and mark on it all the points you have found in (a) to (c).

B. Curves of the form $g(t)=e^{-at}\cos(bt)$

These functions represent *damped harmonic motion*, such as a pendulum oscillating with friction, the shock absorber/spring system of a car's suspension and many other forms of vibration. $g(t)$ gives the displacement from equilibrium at time t.

Question 3

(a) Graph the function $\cos(2t)$ on $[0, 2\pi, \pi/2]\times[-1, 1, 0.5]$ Y1 = cos(2X).

(i) What is its period? What is the period of $\cos(bt)$? The function $\cos(bt)$ represents an undamped vibration.

(ii) Graph $\cos(2t)$ and $e^{-0.1t}$. The exponential curve represents damping or absorption of energy.

(iii) Now graph $e^{-0.1t}\cos(2t)$. What is the 'period' of this oscillatory function? What is its 'amplitude'? It might help to graph the functions $e^{-0.1t}$ and $-e^{-0.1t}$ as well.

(b) Graph some members of the family of functions $g(t) = e^{-at}\cos(bt)$. Take $a, b > 0$ and consider especially the cases $a << b$ (weak damping) and $a >> b$ (strong damping). You may need to change Xmax.

Describe the behaviour of the functions in these two limiting cases.

(c) What is the limit of g as $t \to \infty$? Does it depend on a or b?

PTO

(d) You have to design a car suspension whose response to a bump (which displaces the system from the equilibrium position $g(t) = 0$) is modelled by the family of curves $g(t) = e^{-at}\cos(bt)$.

(i) Explain the sort of characteristics you would look for in such a suspension system. Think about the behaviour of the different curves of g you examined in (b) — which one would you want to describe the suspension of the car you are in?

(ii) Sketch a member of the family that gives these characteristics. Don't forget to label the axes and give some sort of scale for each axis.

(e) What can you say about the *zeros* of g relative to those of $\cos(bt)$? Use an algebraic approach here.

(f) Now think about the *critical points* of g relative to those of $\cos(bt)$? If b is fixed, what happens to the critical points as a goes from 0 to ∞? Again, do this algebraically. *Hint*: Look at what happens to one critical point, say the first local minimum as a is increased from 0. For $a=0$, this is at the first local minimum of $\cos(bt)$, $bt=\pi$.

Check your result graphically by plotting g' for different values of a starting at 0, keeping $b=1$.

Supplementary Question

Question 4

(a) Just by looking at the graph of f (i.e. without using the calculator or differentiating the function), sketch the first derivative of your representative curve in Question 2(d). Explain your reasoning in arriving at this sketch. Mark on the sketch the critical points and points of inflection of f (**not of** f'). You may use the calculator (plot f' worked out algebraically or use *nDeriv*) to **check** your answer.

(b) Similarly, sketch the second derivative of your representative curve in Question 2(d). Again explain your reasoning. Mark on this sketch the points of inflection of f (**not of** f''). Again you may use the calculator to check your answer.

(c) Sketch a function F whose derivative is the function f in Question 2(d). In other words, what does the graph of the function that we differentiate to give f look like? Use what you know about the relationship between a function and its derivative. Don't forget the explanation.

- What you are sketching here is the function

$$F(x) = \int_C^x f(t)\,dt,$$

where C is some constant. Take C to be a value at or to the left of the x interval over which you sketched the curve in 2(d).

- Note that $\int_{-\infty}^{\infty} f(t)\ dt = \sqrt{\pi b}$, i.e. it is finite.

Instructors' Guide

The FAMILY2/FAMLY2CE program (or SPGRAPH/SPGRPHCE)[18] allows students to input values for the parameters a and b, with the resulting curve plotted. Successive curves are superimposed. Some of the more recent graphics calculators have a similar program built in.

Solutions

1. (a) Varying a causes the graph to be translated along the x axis. The basic curve ($a=0$) is centred on $x=0$ because of the $x-a$ in the exponent. With a non-zero, the graph is centred on $x=a$.

(b) The effect of increasing b is to broaden the curves — they all pass through the point $(0,1)$. We are effectively changing the x scaling when we change b — the exponent is x/b.

(c) In all the curves, $f(x) \to 0$ as $x \to \pm\infty$. This behaviour does not depend on a or b.

2. (a) (i) $f'(x) = -2(x-a)e^{-(x-a)^2/b}/b$.
$f'(x)=0$ when $x=a$. f' is defined for all x. Therefore, the only critical point is $x=a$.

(ii) From (i), when $x<a$, $f'(x)>0$; when $x>a$, $f'(x)<0$. The function therefore changes from increasing to decreasing as we pass through the critical point, showing that $x=a$ is a local maximum.

(iii) In the graphs, there was a local maximum at $x=a$.

(b) As there is only one critical point of the continuous function f and it is a local maximum, it must also be the global maximum.
There is no global minimum — $f(x)>0$ for all x, but never attains the value 0.

(c) $f''(x) = \dfrac{4}{b^2}\left((x-a)^2 - \dfrac{b}{2}\right)e^{-\frac{(x-a)^2}{b}} = \dfrac{4}{b^2}\left(x-a+\sqrt{\dfrac{b}{2}}\right)\left(x-a-\sqrt{\dfrac{b}{2}}\right)e^{-\frac{(x-a)^2}{b}}$.

Therefore, $f''(x)=0$ when $x=a\pm\sqrt{b/2}$.

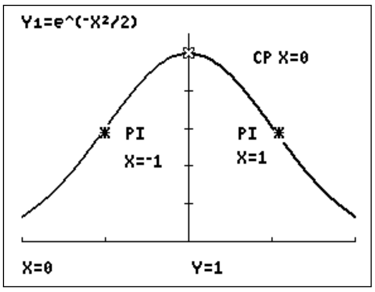
The quadratic $(x-a)^2 - b/2$ is a heads-down parabola, with zeros at $x=a\pm\sqrt{b/2}$.

From this or from the factorisation of the quadratic above, we see that $f''(x)>0$ when $x<a-\sqrt{b/2}$ and when $x>a+\sqrt{b/2}$; the function f is concave up in these intervals. Similarly, $f''(x)<0$ when $a-\sqrt{b/2}<x<a+\sqrt{b/2}$; the function f is concave down in this interval.

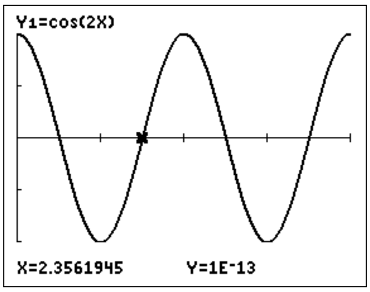
Therefore, $f''(x)$ changes sign at $x=a\pm\sqrt{b/2}$, so that there are points of inflection at these x values. This agrees with the behaviour of the function as seen on the graphs — concave down in the middle, concave up to the left and right.

[18]available from *www.XXX*

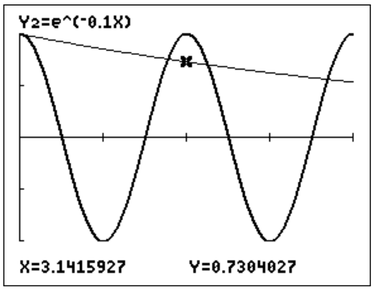
(d) The graph of $f(x) = e^{-x^2/2}$ $(a=0,\ b=2)$ for $-2 < x < 2$ and $0 < y < 1.1$ is shown below. It has a global maximum at $x=0$ and points of inflection (PI) at $x = \pm 1$.

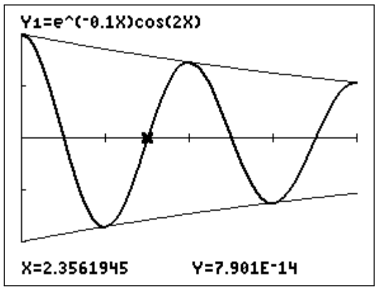


3. **(a)** **(i)** The period of $\cos(2t)$ is π (below) — there are two cycles on the interval $[0, 2\pi]$. The period of $\cos(bt)$ is $2\pi/b$.



(ii) Below left.





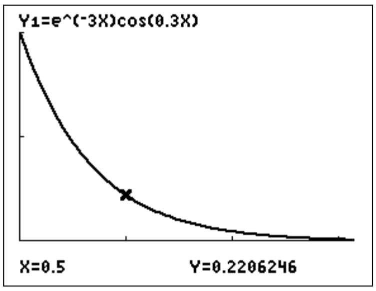
(iii) This function (above right) has a period π (from the $\cos(2t)$), but an amplitude now of $e^{-0.01t}$. The curves $e^{-0.01t}$ and $-e^{-0.01t}$ provide an envelope for the oscillations.

(b) In the case $a << b$, there are many oscillations before the amplitude decreases to near 0. For $a >> b$, there are few or no oscillations before the amplitude decreases to near 0.

(c) $\lim_{t\to\infty} g(t) = 0$. This does not depend on a (provided $a > 0$) or b. The parameter a determines how quickly the function decreases, but not the limiting value.

(d) **(i)** You would like a suspension set up so that the amplitude decreases to near zero reasonably quickly (to be ready for the next bump), but without oscillations (other than possibly when the amplitude is close to 0 — such low-amplitude oscillations would not be noticed).

(ii) Values of $a=3$, $b=0.3$ give these characteristics.



window $[0, \pi/2, 0.5] \times [0, 1, 0.5]$

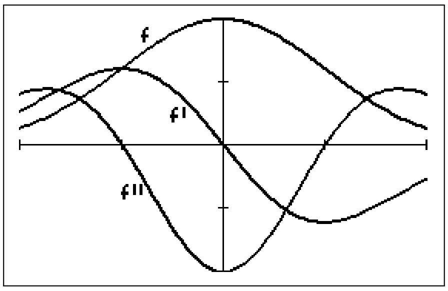
(e) The zeros of g are the same as those of $\cos(bt)$, because g is just $\cos(bt)$ multiplied by a function, $e^{-0.01t}$, which is never 0.

(f) The critical points of g are given by $g'(t)=0$, i.e.

$$-\big(a\cos(bt) + b\sin(bt)\big)e^{-at} = 0.$$

Therefore, $a\cos(bt)+b\sin(bt) = 0$ or $bt = \arctan(-a/b)$.
If $a = 0$, this gives $bt = 0, \pi, 2\pi, \ldots$, the critical points of $\cos(bt)$. As a increases, the value of bt at the critical point decreases following the arctan curve.
The value of bt at the first local minimum is given by $bt=\pi+\arctan(-a/b)$, which starts at π when $a=0$ and tends to $\pi/2$ as $a\to\infty$ $\big(\arctan(-x)\to-\pi/2$ as $x\to\infty\big)$.

4. The three functions f, f' and f'' plotted on $-2<x<2$ and $-1<y<1$ are shown below.



(a) In sketching the first derivative of f from its graph, we take into account the following features.

- f' is zero at the critical point $x=0$.
- To the left of the critical point, f is increasing at an increasing rate, then at a decreasing rate, so f' is positive and increasing, then positive and decreasing, thus passing through a local maximum at the point of inflection of f at $x=-1$.
- Similarly, to the right of the critical point, f' passes through a local minimum at $x=1$. This too is a point of inflection of f.

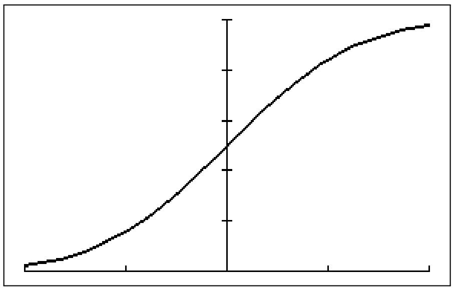
(b) In sketching the second derivative of f, we use (a) and the fact that f'' is the derivative of f'.

- f'' is zero at the critical points (local maximum/minimum) $x = \pm1$ of f', corresponding to the points of inflection of f.
- f'' has local maxima at $x=\pm\sqrt{3}$ and a local minimum at $x=0$, corresponding to the points of inflection of f'.

(c) The function $F(x) = \int_C^x f(t)\,dt$ has the following properties.

- It is 0 at $x = C$ (property of the definite integral).
- It is increasing for all $x > C$, because $f' > 0$.
- It has a point of inflection at $x = 0$, changing from concave up (f' increasing) when $x < 0$ to concave down (f' decreasing) when $x > 0$.

The function $F(x) = \int_{-5}^x e^{-x^2/2}\,dt$, plotted on $-2 < x < 2$ and $0 < y < 2.5$, is shown below.



2.5 Newton's Method: A Classic Problem

Based partly on material from the book *Exploring Calculus with a Graphing Calculator* by C.E. Beckmann and T.A. Sundstrom, Addison-Wesley, 1992.

Aims

- To show *graphically* how Newton's Method approximates the solution of a particular equation and to relate this to the corresponding *algebraic* steps.
- To become proficient in using the NEWTON/NEWTONCE program as a *numerical* method to solve equations of the form $f(x)=0$ and $f(x)=g(x)$.

Introduction

When there are no algebraic methods that will give an exact solution of an equation $f(x)=0$, numerical methods are used to approximate the solution. There is a built-in root finder on the TI-84/CE, but how does it work?

Newton's Method (sometimes called the Newton-Raphson Method or the Newton-Raphson-Simpson Method or even the Newton-Raphson-Simpson-Fourier Method uses tangent-line (linear) approximations to the function f to successively approximate the solution of the equation.

Question 1 *The process done graphically and algebraically using a specific function*

Let $f(x) = x^2 - 10$. Carry out the graphical construction of Newton's Method and the corresponding algebraic steps for solving the equation $f(x) = 0$, using an initial approximation $x_0 = 1$, **as follows**.

(a) Write down the equation of the tangent line to the graph of f at $x = x_0 = 1$. You may differentiate f algebraically to find the slope of this line.

(b) Show *algebraically* that the zero or x intercept, x_1, of the tangent line is given by $x_1 = 5.5$. x_1 is then the next approximation to the solution of the equation $f(x) = 0$, that is we approximate the zero of the function by the zero of the tangent-line approximation to the function.

Use the points $(1, -9)$ and $(5.5, 0)$ to draw the tangent line at x_0 on the graph of f in Figure 1 (page 35); label x_1.

(c) Repeat steps (a) and (b) using the new approximation x_1 (instead of x_0) to find the next approximation x_2. Draw the tangent line at x_1 on Figure 1 (using the two points you know) and label x_2.

(d) Then use x_2 to find x_3 graphically, i.e. this time, just draw the appropriate tangent line by eye. There is no need to calculate things algebraically. Label x_3.

(e) Estimate from the graph of f the value of x for which $f(x) = 0$. Comment on the sequence of values x_1, x_2, x_3.

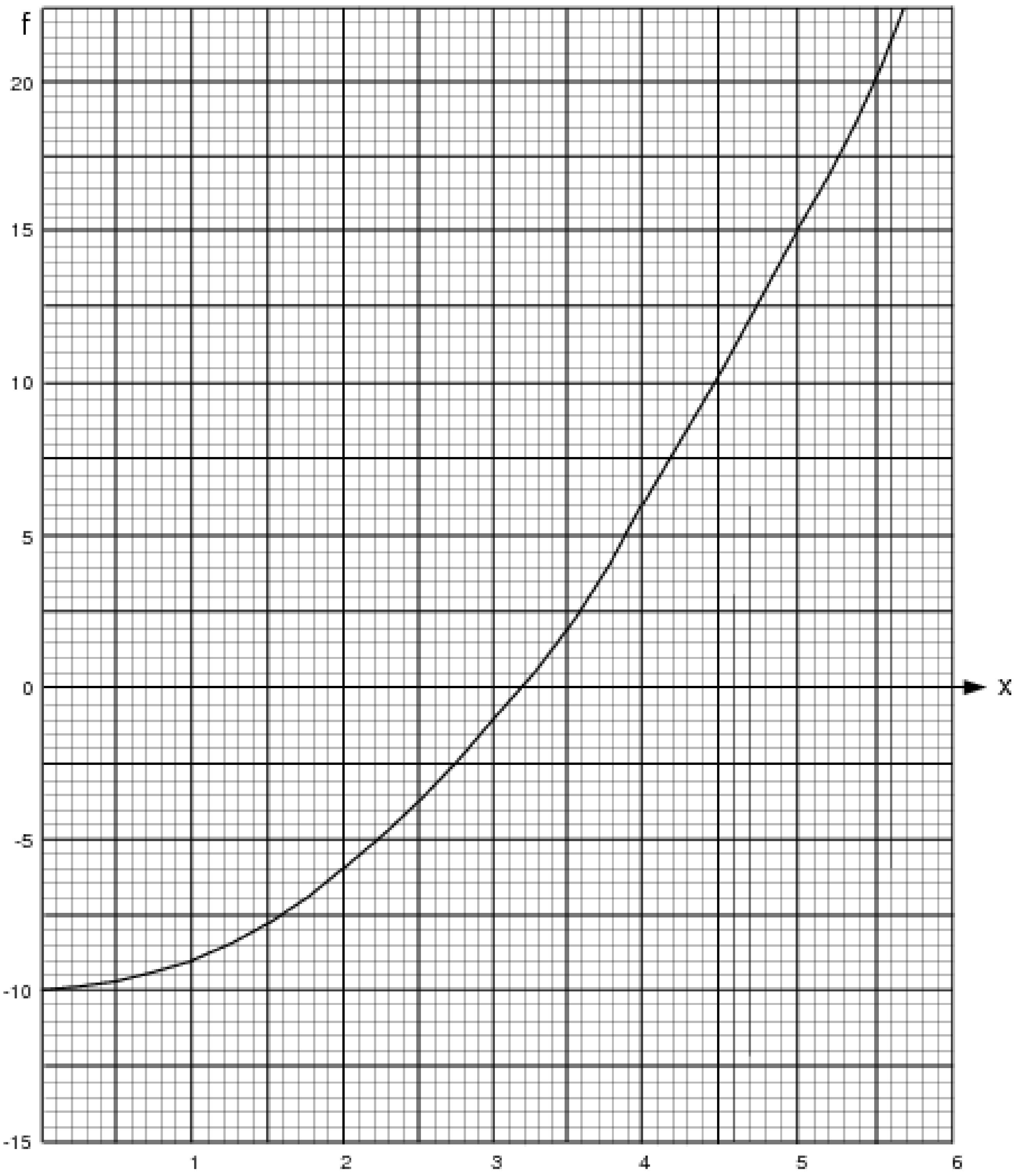


Figure 1: $f(x) = x^2 - 10$.

Newton's Method in Recursive Form

The overall procedure of Newton's Method is usually stated in recursive form as follows.

(1) Start with an initial approximation x_0 for the solution of $f(x)=0$.

(2) To go from the n^{th} approximation x_n to the next approximation x_{n+1}, use the formula

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

That is, we start with x_0, use the formula to find x_1, then use it again to find x_2, etc. This process is called *iteration.* In most cases, the values of x_n $(n = 0, 1, 2, \dots)$ will converge rapidly to the solution, and so give decimal approximations to the solution of the equation $f(x)=0$. From the work done in Question 1, you should be able to see how the above formula is derived.

Question 2 *Using the NEWTON/NEWTONCE program*

Read the instructions regarding use of the NEWTON program in the Appendix, especially on the two ways of entering the initial approximation and on how to stop the program.

(a) Solve $f(x) = x^2 - 10 = 0$ as follows:

- set Y1 = $X^2 - 10$;
- store the initial approximation $x_0=1$ in X: [1] [sto ▶] [X,T,θ,n] [enter];
- run the NEWTON/NEWTONCE program, choosing Y1 as the function, through three steps to give x_1, x_2, x_3.

Write down x_1, x_2, x_3 and compare them with your answers to Question 1.

(b) Find an approximation to the solution *accurate* to 9 significant digits by continuing to run the program.

- We assume that a result here is *accurate to n* significant digits if the first n significant digits are the same in two successive answers (after rounding).
- How many iterations (key presses) did it take to obtain this accuracy?
- Check your answer by squaring it (quit the program and evaluate X^2). Comment.

(c) Plot the graph of $f(x)=x^2–10$ on your calculator using the [window] $[0, 6, 1]\times[-15, 20, 5]$. This time generate the initial guess x_0 to the zero of f by tracing along the curve to a point close to the zero. Then run the program until you have an answer accurate to 9 significant digits. What do you notice about the convergence here compared to (b)? Explain why.

(d) What is the other solution to the equation $x^2 - 10 = 0$?

Note: The program uses the calculator routine *nDeriv*, a numerical approximation to the derivative, to estimate f'. This may cause some inaccuracy, particularly when we get close to the solution, but does not seem to be a real problem with most of the functions we use. A more accurate alternative is to work out the derivative algebraically, put it into Y2 say and replace *nDeriv* command in the program with Y2.

Question 3 *A classic problem*

A hare and tortoise compete in a one-kilometre race. The distance each competitor has travelled from the starting point is given by a formula. In time t **minutes**, the distance in **metres** travelled by the hare is given by $H(t)=\frac{500}{3}\left(2\sqrt{t}+\sqrt[3]{t}\right)$, while the distance in **metres** travelled by the tortoise is given by $T(t)=100t+250\sqrt{t}$.

Plot graphs of H and T as functions of time on your calculator[19] using a **suitable** domain and range, i.e. so that the two graphs go from the bottom left to the top right of the screen (*how far is the race?*). If you select Simul in the mode menu of your calculator before graphing, you will get a real-time view of the race.

Write a commentary of the race, including at least the following features:

- which competitor gets to the halfway point first and how long it takes;[20]
- the time and distance (after the start) at which the two competitors are neck and neck;
- the winner of the race, the time margin and distance margin by which it wins.
- the payout on a $1 bet on the winner.

In a *Technical Appendix*, give details of how you worked out these features: the equations you solved; how you solved them (Newton's Method should be prominent here); and the accuracy of your answers. You don't have to describe how the program works, just the inputs and outputs. Document your steps so that someone else with a TI-84/CE could repeat your calculations.

[19] The cube root is math 4 on a TI-84/CE.

[20] *Hint*: An equation $f(x)=g(x)$ can be written in the form $f(x)-g(x)=0$. Set Y3 = Y1−500 and use Newton's Method on Y3, etc.

Supplementary Questions

Question 4 *The general case done algebraically*

Here we essentially repeat the process of Question 1, but with a general function f and general points x_0, x_1, x_2, ...

In Figure 2 below, the solution of $f(x)=0$ corresponds to the x intercept r of the graph of f. The exact value of r is not known, and we wish to approximate it using Newton's method. Let x_0 be the initial approximation to r.

(a) Write down the equation of the tangent line to the graph of f at $x=x_0$ in terms of its slope $f'(x_0)$ and x_0, $f(x_0)$. Then write this equation in the form $y=T(x)$, i.e. put y on the left-hand side of the equation and everything else on the right-hand side.

(b) This tangent line is a **linear approximation** to the function f, accurate for x near x_0. As it is not possible to solve the equation $f(x)=0$, we solve the equation $T(x)=0$ to obtain the next approximation to the solution of $f(x)=0$.

- **(i)** On Figure 2, locate the point on the graph of $y=T(x)$ for which $T(x)=0$. Label this point x_1.
- **(ii)** From the equation of $y=T(x)$ found in (a), determine an expression for x_1 in terms of x_0, $f(x_0)$ and $f'(x_0)$. Simplify this expression as much as possible.

(c) x_1 is the next approximation to the solution of $f(x)=0$. The procedure can then be repeated to find the x-intercept x_2 of the tangent line to the graph of f at $x=x_1$.

- **(i)** Mark the point $(x_1, f(x_1))$ on Figure 2 and sketch the tangent line to the graph of f through this point.
- **(ii)** Label the x-intercept of this second tangent line x_2. Is x_2 a better approximation for r than x_1?

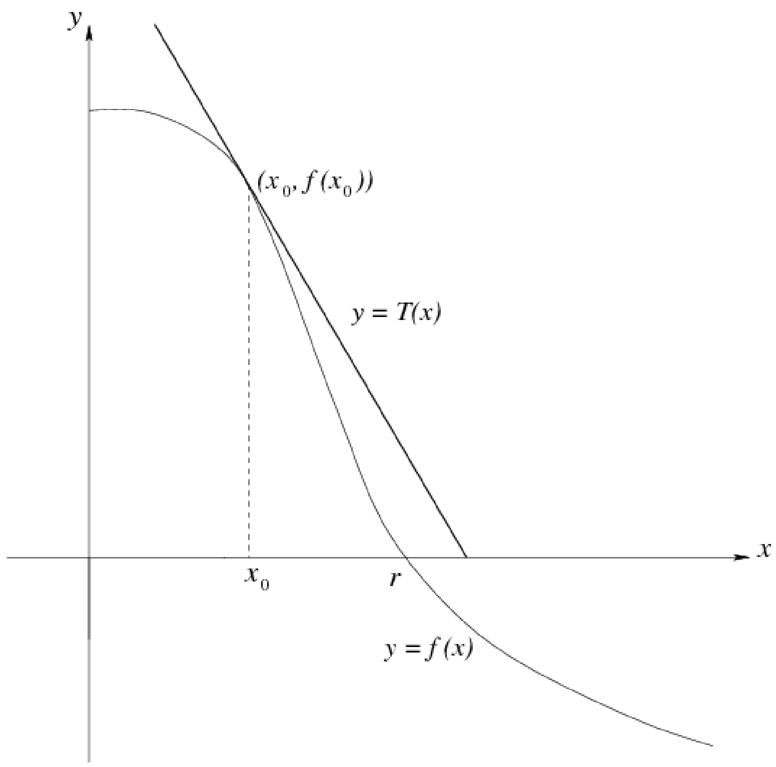


Figure 2

Question 5 *Solving $f(x)=g(x)$*

Rewrite the equation $\cos(x)=0.5x$ in a form suitable for Newton's Method.

(a) Find the solution to the equation, accurate to 10 significant digits, by

- graphing the appropriate function using a window of $[-5,5,1]\times[-5,5,1]$. *Did you get a straight line?* If so, you are not in Radian mode.
- using the cursor to provide the initial approximation.
- What is the solution?
- Check your answer by stopping the program and evaluating Y1 (X). (The last approximation in the program is stored in X.)

(b) Now start with an initial approximation of 3.67, i.e. store 3.67 in X and run the program through a few iterations.

- What do you observe?
- Was 3.67 a good choice for the initial approximation? Explain in terms of the tangent line at $x=3.67$, using a sketch.

Question 6 *Sensitivity to the initial approximation*

Let $f(x)=x^3-x$. The equation $x^3-x=0$ has three solutions: -1, 0 and 1.

(a) Sketch f and show **algebraically** that a local minimum occurs at $x=1/\sqrt{3}$ and a local maximum at $x=-1/\sqrt{3}$.

(b) Argue from the graph of f (you might like to consider its slope and concavity) that if $x_0>1/\sqrt{3}$, Newton's Method will converge to the solution 1. Therefore, by symmetry, if $x_0<-1/\sqrt{3}$, Newton's Method will converge to the solution -1. Discuss what happens if $x_0=\pm 1/\sqrt{3}$.

(c) Demonstrate **algebraically** that if we start Newton's Method with $x_0=\pm 1/\sqrt{5}$, then $x_1=\mp 1/\sqrt{5}$ and $x_2=\pm 1/\sqrt{5}$. Therefore, if we start with $x_0=\pm 1/\sqrt{5}$, we do not converge to a solution. In this case, $\pm 1/\sqrt{5}$ are called *Period 2 points*.

(d) Interesting chaotic behaviour occurs when $1/\sqrt{5}<x_0<1/\sqrt{3}$ or, by symmetry, when $-1/\sqrt{3}<x_0<-1/\sqrt{5}$. Complete the table below and comment on the sensitivity of Newton's Method to the choice of x_0.

x_0	**Solution found**
0.4656	
0.4657	
0.44721	
0.44722	
0.44723	

Appendix

The NEWTON/NEWTONCE program finds a zero of any of $Y_1 - Y_5/Y_7$, equivalently a solution X of the corresponding equation $Y_n(X) = 0$) using Newton's Method. The derivative of the function is estimated numerically using the symmetric difference quotient with an H/tolerance of 10^{-6}.

The program requires an initial guess for the zero to be stored in memory X. One way to generate this guess is to graph the function and move the cursor approximately to the zero. The X co-ordinate of the cursor is automatically stored in X. Then run the program. This should also almost guarantee that the zero the program finds is the one you want. If you don't want to graph the function, just store the initial guess in X.

You could also use the calculator's built-in numerical or graphical solver to find zeros of functions. The nice thing about the Newton's Method program is that you can see the method converging (hopefully).

Use: Type the function into any of $Y_1 - Y_5/Y_7$.

Select the initial guess for X by tracing along the graph of the appropriate function or by storing a value in X. Run the program.

Choose which Y function to use.

Press [enter] to generate successive values.

Press [on] and select Quit to stop the program.

The NEWTON/NEWTONCE program is available at *www.XXX*.

Instructors' Guide

General comments

While a program is useful for this lab, Newton's Method can be achieved with a one-line command. Put the function in Y1 and its derivative (either numerical using *nDeriv* or exact) in Y2. Store the initial guess in X. Then repeatedly execute (keep pressing enter) the command X − Y1/Y2 → X to give successive approximations to the solution. You could also use *zero* in the CALC menu but you then have to accept the calculator accuracy (more than adequate here) and skip the lesson on convergence.

Solutions

1. **(a)** $f'(x)=2x$, so that $f'(1)=2$. The tangent at $x=1$ is therefore a straight line of slope 2, passing through $\big(1, f(1)\big)=(1,-9)$. Its equation is then $y+9=2(x-1)$ or $y=2x-11$.

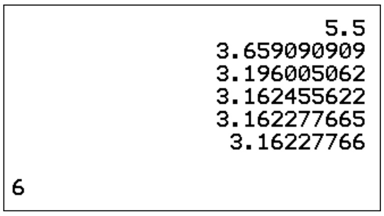
(b) The x intercept is given by $y=0$, so that $x=x_1=5.5$.

(c) $f'(5.5)=11$, so that the tangent at $x=5.5$ is a straight line of slope 11, passing through $\big(5.5, f(5.5)\big)=(5.5, 20.25)$. Its equation is then $y-20.25=11(x-5.5)$ or $y=11x-40.25$. The x intercept of this line is $x_2=40.25/11\approx 3.659$.

(d) Graphical estimates should put x_3 within about 0.1 or 0.2 of the zero of f and certainly significantly closer than x_2. Algebraically, $x_3\approx 3.196$.

(e) From the graph, the zero of f is about 3.18. The sequence of values x_1, x_2, x_3 is getting closer and closer to this value.

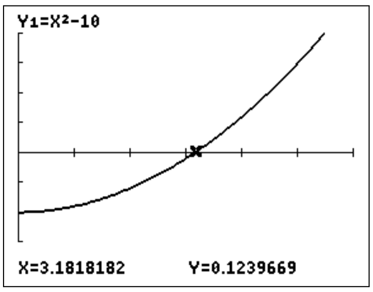
2. **(a)** $x_1=5.5$, $x_2\approx 3.659$ and $x_3\approx 3.196$, as we found in Question 1.

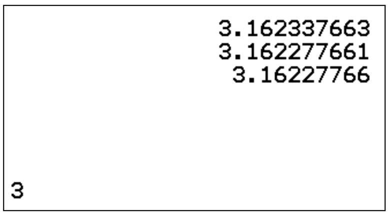


(b) The fifth iteration gives the solution to $f(x)=0$ as $x=3.1622777$, rounded to 8 significant digits. The sixth iteration confirms the accuracy of this answer.

Squaring the sixth iteration gives (exactly, according to the calculator) 10, so that our solution looks to be very accurate.

(c) This time we obtain the same answer as in (b), but after only 3 iterations. The convergence is faster because we started with a better initial guess.





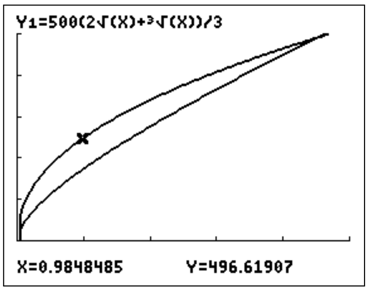
(d) The exact solution to $f(x)=0$ is $x=\pm\sqrt{10}$.

We have found an approximation to $\sqrt{10}$, so the other solution is $x=-\sqrt{10}=-3.16227766$, accurate to 9 significant digits.

3. The *Technical Appendix* should contain the following information.

The equation for the hare was put in Y_1, that for the tortoise in Y_2.

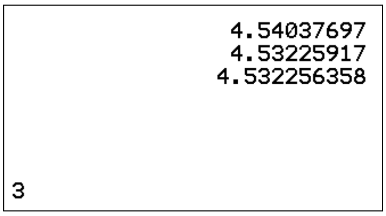
- From the graph, the hare clearly reached the halfway point (500 m) first.



To find how long it took, we solve $H(t) = 500$ for t. The easiest way to do this on the calculator is to set $Y_3 = Y_1 - 500$ and find the zero of Y_3 using Newton's Method with any reasonable initial guess. The value for t quickly converges to $t = 1$ minute, a value we can confirm algebraically to be exact.

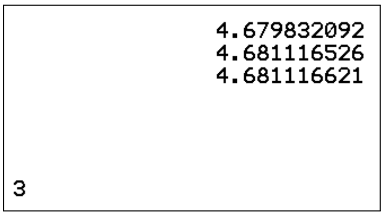
The hare reaches the halfway point first in a time of 1 minute.

- To find when they are neck and neck, we have to solve $H(t) = T(t)$ or $H(t) - T(t) = 0$. Set $Y_3 = Y_1 - Y_2$ and again find the zero of Y_3 using Newton's Method with any reasonable initial guess. We obtain $t = 4.53$ minutes, accurate to 3 significant digits.

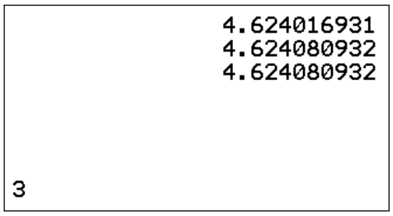


The hare and tortoise are neck and neck after about 4.53 minutes or about 4 minutes 32 seconds.

- To find the winner, we have to determine the time at which each competitor reaches the finish (1000 m). Setting $Y_3 = Y_1 - 1000$, we find the hare finishes at $t = 4.681$ minutes. Similarly, with $Y_3 = Y_2 - 1000$, we find the tortoise finishes at $t = 4.624$ minutes.



Hare



Tortoise

To find the distance margin, calculate $H(4.624)$, the position of the hare when the tortoise finishes: $H(4.624) = 994.45$ m, rounded to 5 significant digits.

The tortoise wins the race by a margin of 0.057 minutes or 3.42 seconds. The distance margin is 5.55 m.

- A short discussion on possible odds, given the historical setting?

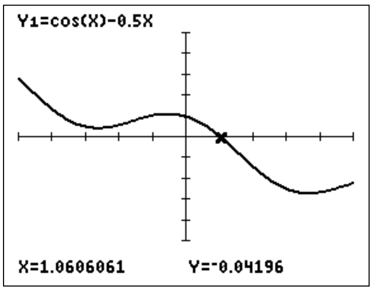
4. **(a)** The equation of the tangent is $y - f(x_0) = f'(x_0)\big(x - x_0\big)$ or $y = f(x_0) + f'(x_0)\big(x - x_0\big)$, giving $T(x) = f(x_0) + f'(x_0)\big(x - x_0\big)$.

(b) **(ii)** $T(x_1) = 0 \;\Rightarrow\; f(x_0) + f'(x_0)\big(x_1 - x_0\big) = 0 \;\Rightarrow\; x_1 = x_0 - \dfrac{f(x_0)}{f'(x_0)}$.

(c) From the graph, x_2 is clearly a better approximation to r than x_1.

5. $\cos(x) = 0.5x \;\Rightarrow\; \cos(x) - 0.5x = 0$. Set $f(x) = \cos(x) - 0.5x$ and put in Y1.

(a) Following the usual procedure with the NEWTON/NEWTONCE program, we obtain a solution $x = 1.029866529$, accurate to 10 significant digits. Y1(X) = 0, confirming our solution.




```
                1.030037679
                1.029866535
                1.029866529
                1.029866529
4
```

(b) Starting with $x_0 = 3.67$, the sequence of approximations shows no sign of converging to a particular value. Clearly, $x = 3.67$ is not a good choice for the first approximation.

```
        652.6007849
        1946.880496
        5329.145934
        3347.049506
        7066.273845
        21836.64101
        11327.34105
9       24012.63095
        48440.80508
```

The reason is that $x = 3.67$ is very close to a local minimum of f, so that the first tangent is almost horizontal. The next approximation x_1 is therefore a very long way away from the zero of f and never returns (*using the numerical derivative for* f') or takes a long time to return (*using the exact derivative*) to the right value.

PTO

6. **(a)** $f(x) = x^3 - x$, so that $f'(x) = 3x^2 - 1$ and $f''(x) = 6x$.
f has critical points when $f'(x) = 0$, giving $x = \pm 1/\sqrt{3}$, or when f' is undefined — there are no values of x for which f' is undefined.
When $x = 1/\sqrt{3}$, $f''(x) > 0$, so that $x = 1/\sqrt{3}$ is a local minimum.
When $x = -1/\sqrt{3}$, $f''(x) < 0$, so that $x = -1/\sqrt{3}$ is a local maximum.

(b) If $x_0 > 1/\sqrt{3}$, the tangents will have positive slopes and will lie beneath the graph of f because f is concave up. All tangents will therefore cross the x axis to the right of $x = 1$; successive approximations will therefore tend to 1.
If $x_0 = \pm 1/\sqrt{3}$, the method will fail because the initial tangent, being horizontal, will never cross the x axis. On the calculator, this will give a divide-by-zero error message.

(c) If $x_0 = \pm 1/\sqrt{5}$, $f(x_0) \;=\; \pm\dfrac{1}{\sqrt{5}^3} \mp \dfrac{1}{\sqrt{5}} \;=\; \mp\dfrac{4}{5\sqrt{5}}$ and $f'(x_0) = -2/5$.

Therefore,

$$\begin{aligned} x_1 &= \pm 1/\sqrt{5} \pm \frac{4/5\sqrt{5}}{-2/5} \\ &= \pm\left(\frac{1}{\sqrt{5}} - \frac{2}{\sqrt{5}}\right) \\ &= \mp\frac{1}{\sqrt{5}} \\ &= -x_0. \end{aligned}$$

The same calculation gives $x_2 = \pm 1/\sqrt{5} = x_0$.
Therefore, if we start with either value for x_0, we will return to this value after two iterations; successive iterations will oscillate between $\pm x_0$.

(d) Using the NEWTON/NEWTONCE program with the given value of x_0 stored in X, we obtain the following results.

x_0	**Solution found**
0.4656	1
0.4657	−1
0.44721	0
0.44722	1
0.44723	−1

Clearly, near particular values of x_0, Newton's Method is very sensitive to small changes in x_0.

2.6 Optimisation

Aims

- To look at different approaches to finding optimum values.
- To develop skills in mathematical modelling.

Optimisation problems are usually regarded as applications of differentiation. In reality, they are examples of the common modelling question *What's the best way to ...?*. If you are interested in the modelling aspect, having to refind critical points every time you change a parameter gets in the way of exploring the model. This is where graphics calculators are really handy. The process of finding a maximum or minimum is simple, once you have plotted the function to be optimised. Having a graph of the function you are optimising also makes it easy to see which value you are trying to find, valuable when there are multiple possibilities. It also allows you to check an algebraic result.

One aspect of modelling beyond finding the optimum is just how critical is the optimum value, often called sensitivity analysis. Algebraically this can be difficult or impossible but on a graphics calculator it is just an additional simple step.

Procedure for Optimisation Problems

In most problems (not just optimisation problems) it helps to break down the problem into parts (mathematical thinking). Here is a procedure for optimisation problems.

1. **Define variables with units.** Drawing a sketch is always a good idea too. At this stage you should also think about the problem: what should the answer be?

   It is also useful at this stage to write down the information given in terms of your variables.

2. **Formulate the equations.** Write down an equation for the variable to be maximised/minimised as a function of the other variables. Use other equations to rewrite this equation in terms of one independent variable.

3. **Determine the domain of the function to be maximised/minimised.** Usually follows from the nature of the problem. Is the domain open or closed? If there is a choice, make it closed (endpoints included), because then there will always be a global maximum/minimum (provided the function is continuous).

4. **Find the GLOBAL maximum/minimum.** Do this graphically or algebraically using the methods we have developed. If you find a **local** maximum/minimum, you have to argue in some way that it is also the **global** maximum/minimum.

5. **Answer the question.** As the question was in words, your answer should be too. Don't forget the units. Does your answer make sense?

Global Maxima and Minima

1. A global maximum and a global minimum always exist for a **continuous** function with a **closed** domain. They occur either at a critical point of the function or at an endpoint of the domain.

2. A global maximum and a global minimum **may not exist** if the **domain is open** (or if the function is not continuous). If they exist, they occur at critical points of the function (or at a point of discontinuity).

Question 1 *Global maxima/minima*

(a) Write down what you understand by a global maximum and a global minimum of a function on a given domain.

(b) Sketch a continuous function on a closed domain for which the global maximum occurs at a critical point and the global minimum occurs at an endpoint (you don't need to draw specific functions, just general shapes).

(c) Sketch a continuous function on an open domain which has two critical points, but neither a global maximum nor a global minimum.

(d) Sketch a continuous function on an open domain which has a global maximum, but no global minimum. Where does the global maximum occur here?

(e) Sketch a continuous function on an closed domain which has a local minimum, but no global minimum.

Question 2 *A warm-up problem*

What is the maximum area possible for a rectangle of perimeter 16m? Follow the steps on page 45 to answer this question.

- Show that a suitable domain in this problem is $0 \leq x \leq 8$, where x is one side of the rectangle. Note the $=$ part of $\leq$.
- Do Step 4 graphically first, accurate to 4 decimal places, then algebraically — don't forget you are looking for a *global* maximum.
- What shape is the rectangle of maximum area?

Question 3 *A bit harder*

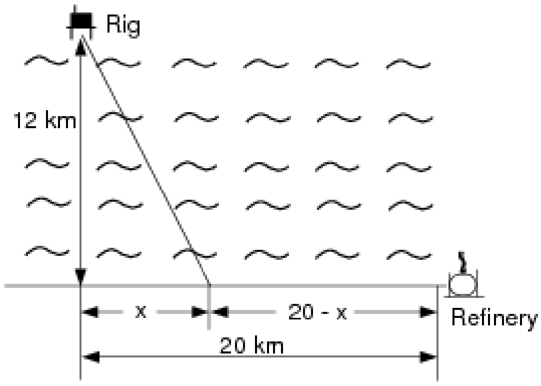
What dimensions of a rectangle of area A give **maximum** perimeter? Again you should follow the steps of the procedure on page 45.

- Because of the parameter A in the equation to be maximised, you will need to find the global maximum, if it exists, algebraically. You could check your answer graphically by choosing different values for A and plotting perimeter as a function of side length.

Note to Instructor: Choose (one or more) of Questions 4 − 9 to include in the lab. All of these questions can be put in the short form of Question 4 or be set out in steps like Question 5 to aid in understanding the process. One question in each form is a good idea. Questions 10 and 11 are supplementary (harder) problems.

Question 4 *Oil pipeline*

You are drawing up plans for the piping that will connect a drilling rig 12 km offshore to a refinery onshore 20 km down the coast (see the diagram).



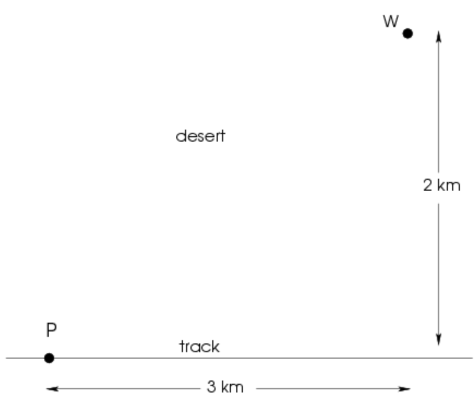
(a) What value of x will give you the least expensive connection if underwater pipe costs \$50,000 per km and pipe on land costs \$30,000 per km? What is the total cost of this connection? Make sure you explain what you are doing and how you know your solution is the least-expensive solution.

(b) It turns out that there is a large mass of rock right at the optimum point for the pipeline to come onshore. You have to relocate this point either side of the rock, but you only have \$5,000 to do this. How far either side of the optimum point can you go and still stay within budget? An aproximate answer is sufficient here, but say how you found it and comment on the accuracy.
Hint: An algebraic solution may *not* be the smartest way to try to solve this question!

(c) If you did (a) algebraically, now do it graphically. If you did it graphically, now do it algebraically.

Question 5 *Survival*

A thirsty soldier out on a survival exercise is at point P on a straight track through the desert. She desperately wants to reach a waterhole at point W. She can walk at 8 km/h along the track, but only at 3 km/h through the soft sand off the track. Determine, **using the steps below**, how far along the track from P she should walk before setting off in a straight line for the waterhole, so that she reaches the waterhole in minimum time.



(a) Define suitable variables and units.

(b) Derive, **with reasons**, an equation for the quantity to be minimised, remembering that time = distance/speed.

(c) Determine a sensible domain for the independent variable and explain why it is sensible.

(d) Draw a sketch of the function in (b) with this domain and a suitable range for the dependent variable.

(e) Determine **graphically** the position of the point (to within 10 m) at which she should leave the track to reach the waterhole in minimum time. State precisely how you found your answer and achieved the desired accuracy. What is the minimum time?

(f) Repeat (e) algebraically.

(g) She decides that it doesn't matter if she takes up to 5 minutes longer than this minimum time and wants to know how accurate she has to be in determining where to leave the track.

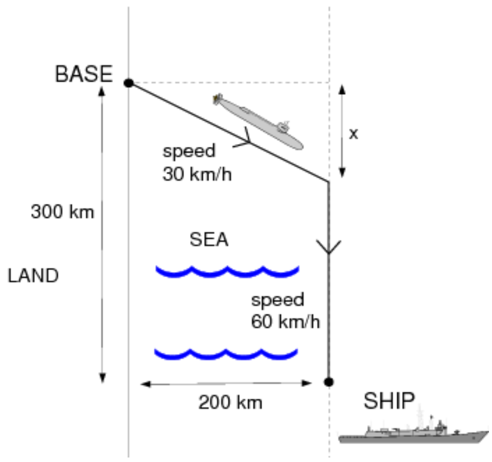
- **(i)** Outline briefly three possible approaches to this problem: graphical, numerical and algebraic. Which approach is likely to be easiest here if an approximate quantitative answer is required?
- **(ii)** Use one of the approachs to answer the question.
- **(iii)** Is the decision on where to leave the track crucial?

(h) In (a)–(d) you have constructed a mathematical model for a particular situation. What assumptions have been made in the model that may not apply in the real situation and that might affect the minimum-time calculations? **Explain how the calculations would be affected.**

Question 6 *Submarine navigation*

A submarine can travel 30 km/h submerged and 60 km/h on the surface. The submarine must stay submerged if within 200 km of the coast. The submarine has to go to the aid of a surface ship 200 km offshore. The submarine leaves from a base 300 km along the coast from the surface ship.

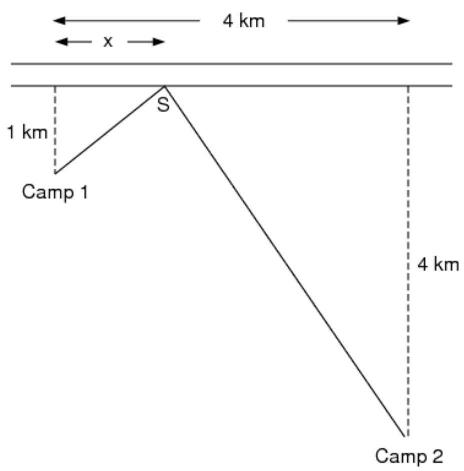
(a) Find *algebraically* the value of x (see the diagram) that will minimise the time for the sub to reach the ship. *Hint*: Remember that time = distance/velocity. What is the time to reach the ship in terms of x?
Check your answer graphically.



(b) The submarine commander is told that the sub has at most 11 hours to reach the ship. What range of x values does this allow? Use a calculator graph to obtain an approximate answer.

Question 7 *A supply problem*

You have to establish a supply depot at some point S on the banks of a straight river to supply troops in Camps 1 and 2, as shown in the figure below. Supply trucks can drive in straight lines across the desert from S to each of the camps.



(a) To make the total travel time to the camps as short as possible, you want to minimise the total distance from the depot to the camps. Where should the supply depot be located? What is the minimum total distance?

(b) You are allowed a tolerance of 200m in the total distance to the camps. What tolerance does this give you in locating the supply depot?

Question 8 *Running a boat*

The cost of fuel to propel a boat through the water (in dollars per hour) is proportional to the cube of the speed. A certain ferry uses \$100 worth of fuel per hour when cruising at 10 km/h. Apart from fuel, the cost of running this ferry (labour, maintenance, etc) is \$675 per hour. At what speed should it travel so as to minimize the cost *per kilometre* travelled?

PTO

Question 9 *Running a truck*

You, the owner of a trucking company, have to send a truck across the Nullarbor Plain.

- The total trip is 1000 kilometres.
- Running the truck (except for the cost of fuel) and paying the driver together cost $60 per hour.
- If the truck is driven at a constant 100 kilometres per hour, the fuel consumption is 2 kilometres per litre. For each kilometre per hour increase in speed above 100 kilometres per hour, the fuel efficiency decreases by 0.02 kilometres per litre.
- Fuel in the outback costs $1 per litre.

(a) Show that the total cost (in $) for the trip if the truck is driven at a constant speed of v kilometres per hour is

$$C(v) = \frac{60\,000}{v} + \frac{1000}{4-0.02v}.$$

Hint: Look carefully at the units of each quantity.

(b) What speed minimises the overall cost of the trip? Your answer should be accurate to at least the nearest kilometre per hour.

(c) What is the minimum cost for the trip (to the nearest dollar)?

(d) You know that driving at a constant speed is difficult and tiring. You are therefore prepared to allow the cost of the trip to go up to 5% above the minimum cost. What is the lowest constant speed and the highest constant speed at which the driver could travel and stay within budget? Again, your answer should be accurate to at least the nearest kilometre per hour.

Question 10 *Converting a try*

In rugby, how far back from the tryline should a kicker take the ball to have the best chance of converting a try? How critical is the choice of distance? You should think initially about the apparent width of the goalposts, rather than the height of the crossbar. Take the width of the goalposts to be 5.6 m.

There are a number of variations on this problem, such as how far back should you be to get the best view of the whiteboard/the TV/your favourite public monument.

Question 11 *A fun problem*

Find (algebraically) the global maximum and global minimum of $f(x) = x - 3\lambda x^{1/3}$ on the interval $[-1, 1]$, where $\lambda \geqslant 0$ is a parameter.

- *Hint*: Your answer will depend on the value of λ. The usual algebraic approach should give you some inequalities in λ to solve to find the global maximum and mimimum. The simplest way to do this is graphically.
- Check your algebraic answer by graphing f for $-1 < x < 1$ and with $\lambda = 0.15, 0.5, 1.1$. Note that just choosing some values for λ and using a graphical or numerical method does not answer this question fully.

Instructors' Guide

Solutions

1. **(a)** A variety of possibilities, but plain English is probably the best.

The global maximum of a function on a given domain is the largest value of the function on that domain. The global minimum is the smallest value of the function on that domain.

(b) For example, a head-up parabola on a closed domain.

(c) For example, a generic cubic shape with a local minimum and a local maximum, but defined on an open domain such that the endpoints, if they were included, would be the global maximum/minimum.

(d) For example, a head-up parabola on an open domain. The global maximum occurs at the local maximum.

(e) This is not possible because of the Extreme-Value Theorem.

2. Following the suggested steps ...

Define variables with units

Let x and y be the lengths of the sides of the rectangle in metres and let A be the area in square metres.

Formulate the equations

We have $2x+2y=16$ from the perimeter or $y=8-x$. The area of the rectangle is $A=xy$.

Putting $y=8-x$ in the expression for A, we have as the function to be maximised

$$A(x) = x(8-x).$$

Determine the domain of the function to be maximised/minimised

The area cannot be negative, so that $0 \leqslant x \leqslant 8$. Graph $A(x)$ if you are not sure.

Find the GLOBAL maximum/minimum

As the function is continuous and defined on a closed domain, the global maximum exists and must occur at either a critical point of A or at an endpoint of the domain.

$A'(x)=8-2x$, so that the only critical point is $x=4$.

Evaluating A at the critical point and at the endpoints gives

$A(4)=16$

$A(0)=A(8)=0$.

The global maximum therefore occurs at $x=4$. The corresponding y value is also 4.

Answer the question

The maximum area of a rectangle of perimeter 16 m is 16 m^2; the rectangle of maximum area is a square of side 4 m.

3. Following the same steps as Question 2, but with perimeter P, we have to maximise

$$P(x) = 2\left(x + \frac{A}{x}\right)$$

with respect to x on an open domain $0 < x < \infty$. The global maximum, if it exists, will occur at a critical point.

$$P'(x) = 2\left(1 - \frac{A}{x^2}\right).$$

The critical points are $x = \pm\sqrt{A}$ ($P'(x) = 0$), of which $x = \sqrt{A}$ lies in the domain, and $x = 0$ ($P'(x)$ undefined), which does not lie in the domain.

Use either the first-derivative test (P' changes from negative to positive as we pass through $x = \sqrt{A}$) or the second-derivative test ($P''(\sqrt{A}) < 0$) to show that $x = \sqrt{A}$ is a local minimum. As the function is continuous on its domain and there is only one critical point, this must also be a global minimum.

That there is no global maximum can be seen from the fact that $P(x) \to \infty$ as $x \to 0$ or as $x \to \infty$. This is also evident in a graph of $P(x)$ for any value of A.

Therefore, there is no rectangle of given area that has maximum perimeter, something that makes sense when we think about it.

4. Follow the usual scheme. (c) is included in (a).

(a) Define variables and units

Let x km be the distance shown in the figure. Let C be the total cost of the pipeline in thousands of dollars.

Formulate the equation

By Pythagoras, the distance from the rig to land is $\sqrt{x^2 + 144}$ km, and the distance from there to the refinery is $20 - x$ km.

Therefore, the total cost $C(x) = 50\sqrt{x^2 + 144} + 30(20 - x)$ thousand dollars.

The problem: find the global minimum of C with respect to the variable x.

Find the domain

A *sensible* domain here is $0 \leqslant x \leqslant 20$. $x = 0$ means the pipe goes straight to the land by the shortest route; $x = 20$ means the pipeline goes direct to the refinery, all underwater. Choose the domain to be closed.

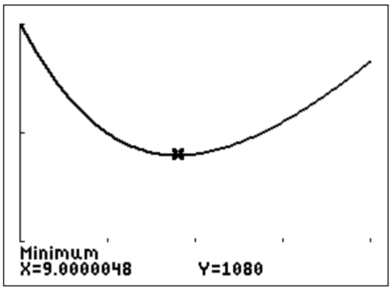
Find the global minimum

The function C is continuous and has a closed domain. The global minimum therefore exists and must occur at an endpoint of the domain or at a critical point.

PTO

Graphically

Plot $C(x)$ on $[0, 20]$. Using *minimum* on the calculator, the global minimum is at the local minimum $x = 9$, giving a total cost of $C = 1080$ (claimed accuracy for *minimum* is at least 5 significant digits).



window $[0, 20, 5] \times [1000, 1200, 100]$

Algebraically

$$\frac{dC}{dx} = \frac{50x}{\sqrt{x^2+144}} - 30.$$

The derivative is defined for all x, so that it is defined for all x in the domain.
$C'(x) = 0$ when $50x = 30\sqrt{x^2+144}$.
After squaring both sides and carrying out some algebra, we find $C'(x) = 0$ when $x^2 = 81$ or $x = \pm 9$. $x = -9$ does not lie in our domain, so the only critical point is $x = 9$.

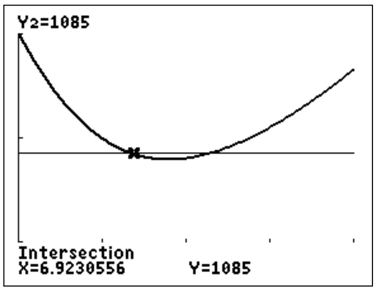
$C(0) = 1200$ endpoint
$C(9) = 1080$ critical point
$C(20) = 1166.2$ (5SD) endpoint

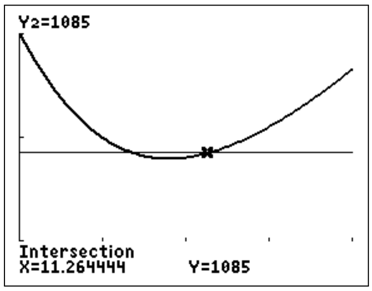
Therefore, the global minimum occurs at $x = 9$.

Answer the question

The pipeline should come ashore 9 km towards the refinery from the point on land directly opposite the rig. The total cost is $1,080,000.

(b) The total cost can rise to $1,085,000. To find the corresponding x values, we have to solve $C(x) = 1085$. The easiest way to do this is to plot the line $y = 1085$ and find where it cuts the graph of C using *intersect*.





window $[0, 20, 5] \times [1000, 1200, 100]$

To 5 significant digits, $6.9231 \leqslant x \leqslant 11.264$ (claimed accuracy for *intersect* is at least 5 significant digits).

A sensible answer to the problem is that the pipeline can come ashore anywhere between 6.9 km and 11.3 km towards the refinery from the point on land directly opposite the rig and still come within a budget of $1,085,000.

5. (a) Define variables

Let x km be the distance along the track from P to where she leaves the track. Let T hours be the total time taken to reach W.

(b) Formulate the equation

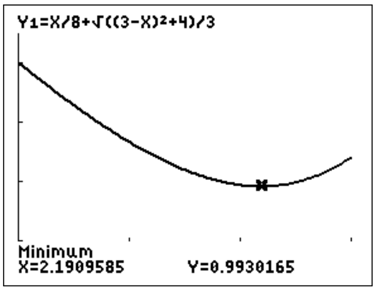
The time along the track is $x/8$ hours. The distance to W from where she leaves the track is $\sqrt{(3-x)^2+4}$, so the time across the desert is

$$T(x) = \frac{x}{8} + \frac{\sqrt{(3-x)^2+4}}{3} \text{ (hours).}$$

(c) Domain

A sensible domain is $0 \leqslant x \leqslant 3$. There is no point in walking away from the waterhole along the track from where she currently is, so $x \geqslant 0$. It is also not sensible to walk past the point directly opposite the waterhole, so $x \leqslant 3$.

(d) Plot $T(x)$ on $[0, 3]$.



window $[0, 3, 1] \times [0.9, 1.25, 0.1]$

(e) Find the global minimum graphically

The global minimum occurs at the local minimum with $x = 2.19$ (3SD) and $T = 0.993$ (3SD), using *minimum* on the TI-84/CE (claimed accuracy at least 5 significant digits).

She reaches the waterhole in a minimum time of about 0.993 hours or about 59.6 minutes if she leaves the track after about 2.19 km.

PTO

(f) Find the global minimum algebraically

T is a continuous function on a closed domain, so the global minimum exists and occurs either at an endpoint or a critical point.

Critical points

$$\begin{aligned}\frac{dT}{dx} &= \frac{1}{8}+\frac{1}{2}\cdot\frac{1}{3}((3-x)^2+4)^{-\frac{1}{2}}\cdot 2(3-x)\cdot -1\\ &= \frac{1}{8}-\frac{3-x}{3\sqrt{(3-x)^2+4}}.\end{aligned}$$

T is defined for all x, so any critical points occur when $T'(x)=0$.

$T'=0$ when

$$\begin{aligned}\frac{3-x}{3\sqrt{(3-x)^2+4}} &= \frac{1}{8}.\\ \therefore\ 8(3-x) &= 3\sqrt{(3-x)^2+4}.\\ \therefore 64(3-x)^2 &= 9\big((3-x)^2+4\big).\\ \therefore\ (3-x)^2 &= \frac{36}{55}.\\ \therefore\ (3-x) &= \pm\frac{6}{\sqrt{55}}.\\ \therefore\ x &= 3\pm\frac{6}{\sqrt{55}}.\end{aligned}$$

Only the critical point $x=3–6/\sqrt{55}\approx 2.19$ (to the nearest 10 m) lies in the domain.

$T(0)=1.20$ hour (3SD) endpoint

$T(2.19)=0.993$ hour (3SD) critical point

$T(3)=1.04$ hour (3SD) endpoint

Therefore, the global minimum lies at the critical point.

As found in (e), she reaches the waterhole in minimum time if she leaves the track after about 2.19 km.

(g) If she can take up to 5 minutes longer than the minimum time, the total time can be up to 1.08 hours (3SD). To find the range of possible x values, we have to solve $T(x)=1.08$ for x. We expect two solutions, one less than and one greater than the x value for minimum time.

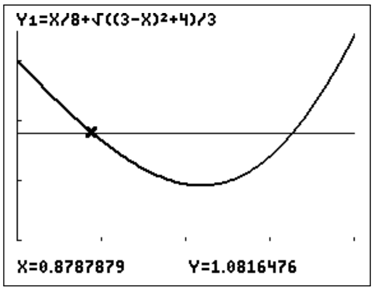
(i) *Graphically*, we plot $T(x)$ and trace along the curve to find the two x values at which $T(x)$ is as close as possible to 1.08.

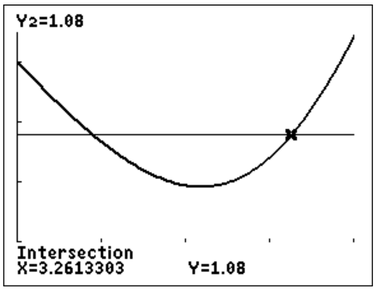
Numerically, we use one of the built-in calculator routines, for example *intersect*, to give approximations for the values of x at which $y=T(x)$ and $y=1.08$ intersect.

Algebraically, we solve $T(x)=1.08$ for x.

The graphical and numerical methods look easier here than the algebraic method.

(ii) Graphically, with the window in (d), the points on the screen for which $T(x)$ is closest to 1.08 give $0.88 < x < 3.26$.





window $[0, 4, 1] \times [0.9, 1.25, 0.1]$

Numerically, we get the same answers, but we now know the values are accurate to the number of digits given. Algebraically, we end up with a quadratic in x which gives the same answers (showing that the values obtained numerically were in fact accurate to 8 significant digits).

(iii) Clearly the decision on where to leave the track is not at all crucial. She can leave the track anywhere between 0.88 km and 3.26 km (actually past the waterhole on the track) from where she is and still only take up to 5 minutes or about 8% longer than the minimum time.

(h) The two most obvious assumptions that we make in the model are that she does indeed walk at a constant speed both on the track and across the sand, and that she can walk in a straight line across the sand. If she doesn't walk in a straight line, she will clearly take longer than the calculated minimum time, because she will cover a greater distance than assumed in the model. As far as the speed is concerned, provided her average speeds over the total distance along the track and the total distance across the sand are equal to the assumed values, the model will still be valid.

6. (a) Define variables and units

Let $T(x)$ be the time in hours for the sub to reach the ship. Let x as shown on the diagram be measured in km.

Formulate the equation

With the help of Pythagoras,

$$T(x) = \frac{\sqrt{200^2 + x^2}}{30} + \frac{300 - x}{60}.$$

Determine the domain

A sensible domain is $0 \leqslant x \leqslant 300$. Choose it to be closed.

PTO

Find the global minimum

T is a continuous function on a closed domain. The global minimum therefore exists and lies at a critical point or at an endpoint of the domain.

$$T'(x) = \frac{1}{2}\frac{2x}{30\sqrt{200^2+x^2}} - \frac{1}{60}.$$

T' is defined for all x, so any critical points occur when $T'(x)=0$.

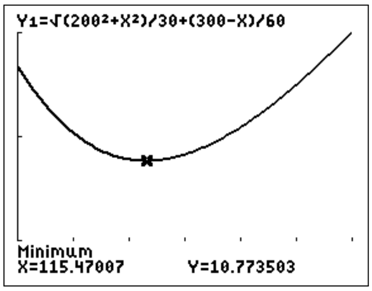
$$T'(x)=0 \;\Rightarrow\; \frac{x}{\sqrt{200^2+x^2}}=\frac{1}{2} \;\Rightarrow\; (2x)^2=200^2+x^2 \;\Rightarrow\; 3x^2=200^2.$$

Therefore $x^2=200^2/3$ or $x\approx\pm115.47$. $x\approx-115.47$ is not in the domain, so that $x\approx115.47$ is the only critical point.

$T(0)\approx11.7$ endpoint
$T(115.47)\approx10.8$ critical point
$T(300)\approx12.0$ endpoint

Therefore, the global minimum, $T(x)\approx10.8$ occurs at $x\approx115.47$.

Graphical check:

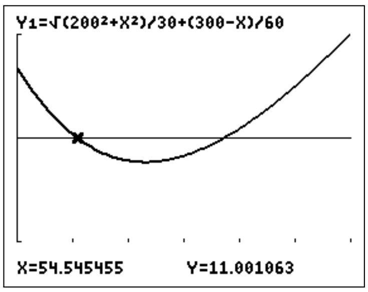


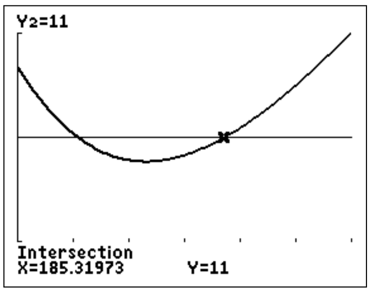
window $[0, 300, 50]\times[10, 12, 1]$

Answer the question

The value of x to minimise the time taken by the sub to reach the ship is about 115.5 km, giving a minimum time of about 10.8 hours.

(b) To answer this question, we need to solve $T(x)=11$. For an approximate solution, trace along the graph of T until $T(x)$ is as close to 11 as you can make it — we find approximately $54<x<185$.





window $[0, 300, 50]\times[10, 12, 1]$

For more accurate answers, find the intersections of the graph of $y=T(x)$ with $y=11$ using *intersect* to give $54.5<x<185$, these values accurate to 3 significant digits. Clearly the choice of x is not critical.

The equation $T(x)=11$ can also be solved algebraically — it reduces to a quadratic.

7. (a) Define variables and units

Let x km be the distance shown in the figure. Let D km be the total distance from the depot to the two camps.

Formulate the equation

By Pythagoras, the distance from S to Camp 1 is $\sqrt{x^2+1}$ and the distance to Camp 2 is $\sqrt{(4-x)^2+16}$.
Therefore, $D=\sqrt{x^2+1}+\sqrt{(4-x)^2+16}$.

The problem: find the global minimum of D with respect to the variable x.

Find the domain

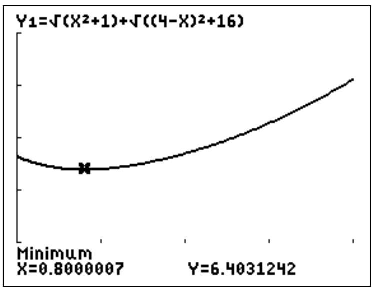
Whether the domain is open or closed is important in determining how we go about finding the global maximum. If we have the choice, we make the domain closed, even if we know the endpoints are not the optimum values, because that makes solving the problem mathematically more straightforward.
If $x=0$, S is directly opposite Camp 1; if $x=4$, S is directly opposite Camp 2. A *sensible* domain here is therefore $0 \leqslant x \leqslant 4$.

Find the global maximum

The function D is continuous and has a closed domain. The global minimum therefore exists and occurs at an endpoint of the domain or at a critical point.

Graphically

Plot $D(x)$ for $0<x<4$. The global minimum occurs at the local minimum $x=0.8$ (5SD), $D=6.40$ (3SD). The claimed calculator accuracy is at least 5 significant digits.



window $[0,4,1]\times[5,9,1]$

Algebraically

$$\frac{dD}{dx}=\frac{2x}{2\sqrt{x^2+1}}+\frac{-2(4-x)}{2\sqrt{(4-x)^2+16}}=\frac{x}{\sqrt{x^2+1}}-\frac{4-x}{\sqrt{(4-x)^2+16}}.$$

The derivative is defined for all x, so that it is defined for all x in the domain.

$$\frac{dD}{dx}=0 \;\Rightarrow\; \frac{x}{\sqrt{x^2+1}}=\frac{4-x}{\sqrt{(4-x)^2+16}}.$$

After squaring both sides and carrying out some algebra, we find $D'(x)=0$ when

$$15x^2+8x-16=0.$$

The quadratic formula gives $x=0.8$ or $x=-4/3$.

The value $x=-4/3$ does not lie in our domain; the only critical point is $x=0.8$.

$D(0)=6.66$ (3SD) endpoint

$D(0.8)=6.40$ critical point

$D(4)=8.12$ (3SD) endpoint

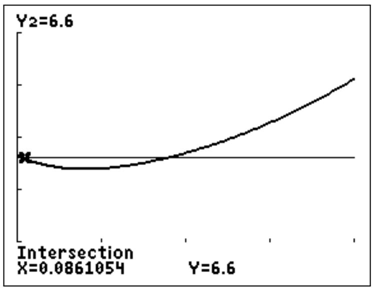
Therefore, the global minimum occurs at $x=0.8$. The corresponding total distance is $D=6.40$ km.

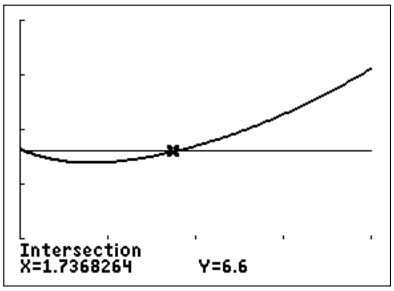
Answer the question

Locate the supply depot 0.8 km to the right (on the map) of the point on the river closest to Camp 1. The total distance to the camps is 6.40 km (3SD).

(b) A tolerance of 200 m = 0.2 km means the total distance can go up to 6.6 km. *Why can't it also go down to 6.2 km?* Find x for which $D(x) = 6.6$ graphically/numerically (easier) or algebraically (messy).

Plot $D(x)$ and the line $y=6.6$ and use *intersect*. Locating S anywhere between $x = 0.09$ km and $x = 1.7$ km will give a total distance within the tolerance. The graph of D versus x has a fairly flat minimum, so the placement of S is not critical.





window $[0, 4, 1] \times [5, 9, 1]$

8. Define variables and units

Let v be the speed of the boat in kilometres per hour.
Let C be the cost *per kilometre travelled.*
Let D be the cost *per hour.*

Formulate the equation

$D = D_f + 675$, where D_f is the fuel cost per hour. We are given that $D_f \propto v^3$, so that $D_f = kv^3$, where k is a constant. As $D_f = 100$ when $v = 10$, we find $k = 0.1$.

Therefore, $D(v) = 0.1v^3 + 675$.

We want to find the global minimum of C, which has units of dollars/km. D has units of dollars/hour, so that if we divide D by a quantity having units km/hour, i.e. by the velocity, we obtain C.

Therefore, $C(v) = \dfrac{D(v)}{v} = 0.1v^2 + \dfrac{675}{v}$.

The problem: Find the global maximum of C with respect to the variable v.

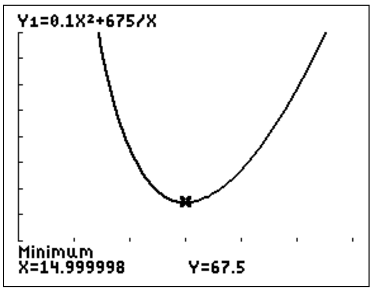
Determine the domain

The only restriction is $v > 0$, so the domain here is $(0, \infty)$.

Find the global minimum

The function C is continuous, but has an open domain. The global minimum, if it exists, will occur at a critical point.

Graphically: Plot $C(x)$ on $[0, 30]$ say. Using *minimum* on the calculator, the global minimum is at the local minimum $v = 15$, giving $C = 67.5$.



window $[0, 30, 5] \times [60, 100, 5]$

Algebraically

$$\frac{dC}{dv} = 0.2v - \frac{675}{v^2}.$$

The derivative is defined for all v in the domain.

$C'(v) = 0$ when $v^3 = 675/0.2$, giving $v = 15$.

The only critical point is $v = 15$, with $C(15) = 67.5$.

As $v \to 0$, $C(v) \to \infty$. As $v \to \infty$, $C(v) \to \infty$.

Therefore, the global minimum occurs at $v = 15$.

Answer the question

The boat should travel at 15 km/h to achieve a minimum fuel cost per km of \$67.50.

9. (a) Total cost $C(v)$ = cost of fuel + other costs.

Other costs: total time for trip is $\dfrac{1000}{v}\,\dfrac{\text{km}}{\text{km/hr}} = \dfrac{1000}{v}$ hr;

$$\text{total other costs} = 60\,\frac{\$}{\text{hr}} \times \frac{1000}{v}\ \text{h} = \$\,\frac{60\,000}{v}.$$

Fuel cost:

$$\begin{aligned}\text{fuel efficiency} &= 2-0.02(v-100)\ \text{km}/\ell\\ &= 2+2-0.02v\ \text{km}/\ell\\ &= 4-0.02v\ \text{km}/\ell;\end{aligned}$$

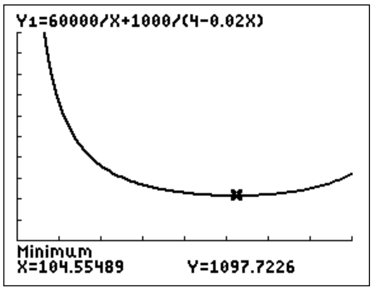
$$\text{total fuel required} = \frac{1000}{4-0.02v}\,\frac{\text{km}}{\text{km}/\ell} = \frac{1000}{4-0.02v}\,\ell;$$

$$\text{total fuel cost} = 1\,\frac{\$}{\ell} \times \frac{1000}{4-0.02v}\,\ell = \$\,\frac{1000}{4-0.02v}.$$

$$\therefore\ C(v) = \frac{60\,000}{v} + \frac{1000}{4-0.02v}.$$

(b) The domain is $v > 0$, but clearly the truck cannot exceed say 160 km/h, so a sensible domain is $(0, 160)$.

Graphically: Plot $C(v)$ vs v over its domain.



window $[0, 160, 20] \times [0, 5000, 500]$

From the graph, the global minimum on the domain lies at the local minimum $v = 104.6$ (4SD) (found using *minimum*,claimed accurate to at least 5 significant digits).

Algebraically

The function is defined on an open domain, so that the global minimum, if it exists, lies at a local minimnum, i.e. at a critical point.

Critical points

$$\begin{aligned}C'(v) &= -\frac{60\,000}{v^2} + \frac{0.02 \times 1000}{(4-0.02v)^2}\\ &= -\frac{60\,000}{v^2} + \frac{20}{(4-0.02v)^2}.\end{aligned}$$

C' is defined for all v in the domain $(0, 160)$.

$C'(v) = 0$ if

$$\begin{aligned} 60\,000(4-0.02v)^2 &= 20v^2. \\ \therefore\ 16 - 0.16v + 0.0004v^2 &= 0.000\dot{3}v^2. \\ \therefore\ 6.\dot{6}\times10^{-5}v^2 - 0.16v + 16 &= 0. \\ \therefore\ v &= \frac{0.16 \pm \sqrt{0.16^2 - 64 \times 6.\dot{6}\times10^{-5}}}{2\times6.\dot{6}\times10^{-5}} \\ &\approx \frac{0.16 \pm 0.146059}{1.\dot{3}\times10^{-4}} \\ &\approx 2215 \text{ or } 104.6 \text{ (4SD)}. \end{aligned}$$

Only $v \approx 104.6$ lies in the domain, so this is the only critical point.

$$C''(v) = \frac{120\,000}{v^3} + \frac{0.8}{(4-0.02v)^3},$$

which is positive for all v in the domain.

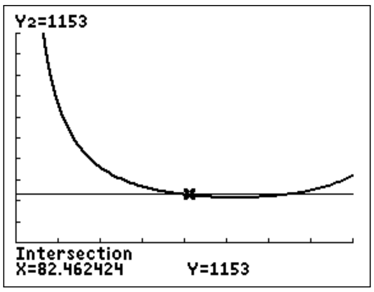
Therefore, C is concave up, the critical point is a local minimum, and because it is the only critical point, it is also the global minimum.

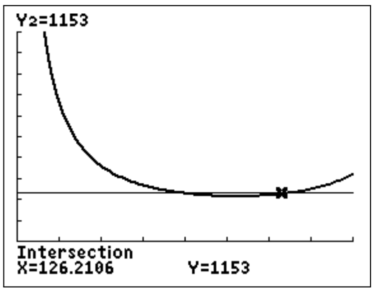
To the nearest km/h, the optimum speed is $v = 105$ km/h.

(c) The minimum total cost for the trip to the nearest dollar, either from using *minimum* on the graph or calculating $C(v_{\text{opt}})$, is \$1098.

(d) 5% more than the minimum total cost is $1.05 \times 1098 \approx \$1153$, so we must solve $C(v) = 1153$.

The simplest way to do this is numerically: graph $C(x)$ and use *intersect* to find where $C(v) = 1153$ or use *zero* to find where $C(v) - 1153 = 0$. To fall within budget, the driver can drive at a constant speed between 82 km/h and 126 km/h (both to the nearest km/h).





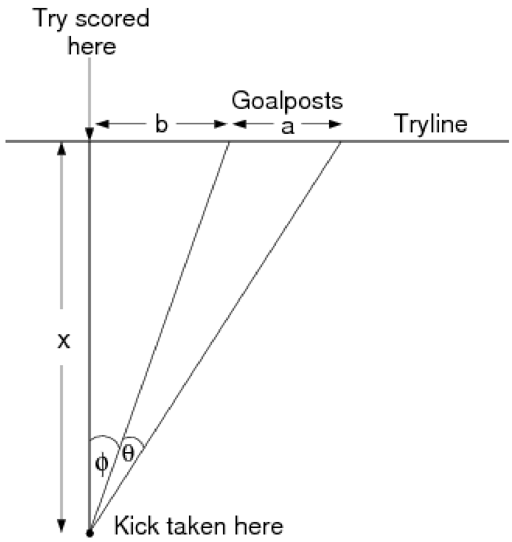
window $[0, 160, 20] \times [0, 5000, 500]$

Solving $C(v) = 1153$ algebraically is straightforward too — you end up finding the roots of a quadratic equation.

10. This problem allows students to develop a model (beyond just finding a maximum), to look at its implications and limitations in a context which should be familiar.

One approach: Take the ball back until the angular width of the goalposts as seen by the kicker is a maximum. Let a be the width of the goalposts, let the try be scored a distance b (measured along the tryline) from the lefthand goalpost and let x be the distance the ball is taken back from the tryline.

The steps to solving this problem and questions that arise subsequently in using the model are given below. The lab could be presented in this form for less mathematically sophisticated students.



(a) Show that the value of x that maximises the angular width of the goalposts is

$$x = \sqrt{b(b+a)}. \tag{1}$$

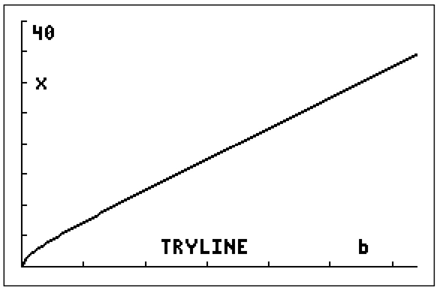
Solution: From the figure,

$$\theta = \arctan\left(\frac{b+a}{x}\right) - \arctan\left(\frac{b}{x}\right).$$

Setting $d\theta/dx = 0$ and solving the resulting quadratic gives $x = \pm\sqrt{b(b+a)}$. For our problem $x \geqslant 0$ (but see (i)), so we choose the + sign.

(b) Plot the optimum kicking distance x versus the distance b from the goalpost that the try was scored, given the width of the goalposts $a = 5.6$ m and the width of a rugby field is 70 m.

Solution: The distance from a goalpost to the touchline on that side is $(70-5.6)/2 = 32.2$ m. Plotting $x(b)$ with this domain gives



window $[0, 32, 5] \times [0, 40, 5]$

(c) However, suppose that the width a differs from ground to ground. Do we have to plot a separate graph for each ground?

The answer is no. We define the **non-dimensional** variables $X = x/a$ and $B = b/a$, and Equation (1) then becomes

$$X = \sqrt{B(B+1)}. \tag{2}$$

Now there are only two variables, so that a plot of X versus B will be valid for all grounds. Often using non-dimensional variables from the beginning makes the algebra easier too.

Derive Equation (2) using non-dimensional variables from the beginning.

Solution: With non-dimensional variables,

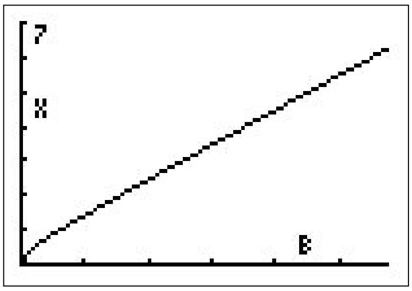
$$\theta = \arctan\left(\frac{B+1}{X}\right) - \arctan\left(\frac{B}{X}\right).$$

Setting $d\theta/dx = 0$ and solving the resulting quadratic gives $X = \pm\sqrt{B(B+1)}$.

(d) Plot X versus B. First work out the domain of the function.

Does the answer to the problem, as given by (2) and your graph (which can be regarded as a plan of the ground), make sense?

Solution: We found in (b) that the distance from a goalpost to the touchline on that side is $(70-5.6)/2 = 32.2$ m. The normalised distance is therefore $32.2/5.6 = 5.75$. With domain $0 \leqslant B \leqslant 5.75$, a graph of $X(B)$, valid for any ground is



(e) What happens as $B \to 0$? Does the answer for B just bigger than 0 predict what a kicker would do? Why not?

Solution: As $B \to 0$, $X \to 0$. For small B, the model predicts a small X, i.e. the kick would be taken close to the tryline. However, the model does not take into account the fact that the kick has also to pass over the crossbar. In practice, for small B the kicker would take the ball back further than predicted by the model.

(f) In reality, B can be as small as $-1/2$. What does $-1/2 < B < 0$ correspond to in practice? What does the model say when $-1/2 < B < 0$? What actually happens in practice when $-1/2 < B < 0$?

Solution: If $-1/2 < B < 0$, the try has been scored under the goalposts. The model gives no value for X in this case because the argument of the square-root function is negative. In practice, the kicker would take the ball far enough back that he or she was confident the kick would clear the crossbar.

(g) What happens as B gets big ($B >> 1$)? Is there a quick rule of thumb in this case? If the ground were arbitrarily wide, would the predictions of Equation (2) be useful for a kicker? Again what other (practical) considerations are there that are not included in the model?

Solution: As B gets big, X also gets big, with $X \approx B$. If the ground were arbitrarily wide, the kick would have to cover an arbitrarily large distance.

For example, with $B = 5.75$, corresponding to a try scored next to the touchline, according to the model the kicker should take the ball back a distance of nearly 35 m, giving a distance to the centre of the goalposts of nearly 50 m. The kicker may not be able to kick the ball this far and would therefore choose a smaller value of X to lessen the kicking distance. The apparent width of the goalposts would then no longer be a maximum.

(h) What happens if the try were scored on the other side of the goalposts to that in the diagram? What values of B correspond to this case? Is this handled by the model?

Solution: The goalposts occupy the interval $-1 \leqslant B \leqslant 0$. If the try were scored on the other side of the goalposts, we would have $B \leqslant -1$. The model handles this case, as you can readily see by allowing B to be negative in your graph.

(i) Why choose the positive square root in Equation (2)? What does the negative square root mean in practical terms? Define the domain and range of Equation (2) to correspond to an actual ground.

Solution: The positive square root in Equation (2) gives a ground running in the positive X direction. If we chose the negative square root, we would have the ground running in the negative X direction, i.e. on the other side of the goalposts. The ground is defined by $-6.75 \leqslant B \leqslant 5.75$ and $X \geqslant 0$. If the ground were 100 m from goalpost to goalpost, the maximum value of X would be $100/5.6 \approx 17.9$.

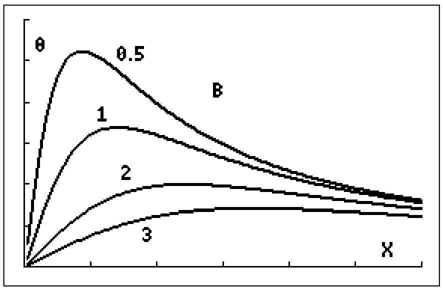
(j) Is the distance given by Equations (1) and (2) a global maximum? You should have checked this when answering the problem. There are several possibilities.

- first-derivative test: Does the first derivative change sign for values of X either side of that given by Equation (2)?
- second-derivative test (messy?)
- physical reasoning: Only one critical point in $0 \leqslant B$, and the angular width of the goalposts θ is a minimum ($= 0$) when $B = 0$ and $\theta \to 0$ as $B \to \infty$. Therefore the critical number given by Equation (2) must be a local and global maximum.

PTO

(k) How sharp is the maximum given by Equation (2)? Does it make a lot of difference if we choose X to be different to the 'best' value, i.e. if we choose to take the kick from a different distance back from the tryline?

Solution: Plot θ versus X for different values of B.



window $[0, 6, 1] \times [0, 0.6, 0.1]$

Clearly, if B is small (try scored close to the goalposts) it does matter: the angular width of the goalposts drops off quite sharply either side of the maximum. However, as we found above, the model is somewhat suspect for small B. For B greater than about 2, there is not much of a peak, and the position of the kick is not crucial. The distance that the ball will have to be kicked is probably just as important, if not more so.

11. $f(x) = x - 3\lambda x^{1/3}$ is a continuous function defined on a closed domain $[-1, 1]$, so that the global maximum and global minimum must exist for all values of λ. They must lie either at a critical point of f or at an endpoint of the domain.

$$f'(x) = 1 - \frac{\lambda}{x^{2/3}},$$

so that $f'(x) = 0$ when $x^{2/3} = \lambda$ or $x = \pm\lambda^{3/2}$. f' is undefined when $x = 0$.

The critical points are therefore $x = 0$ and $x = \pm\lambda^{3/2}$. The latter only lie in the domain if $\lambda \leqslant 1$.

$f(-1) = -1 + 3\lambda$ Endpoint 1

$f(-\lambda^{3/2}) = 2\lambda^{3/2}$ Critical Point 1

$f(0) = 0$ Critical Point 2

$f(\lambda^{3/2}) = -2\lambda^{3/2}$ Critical Point 3

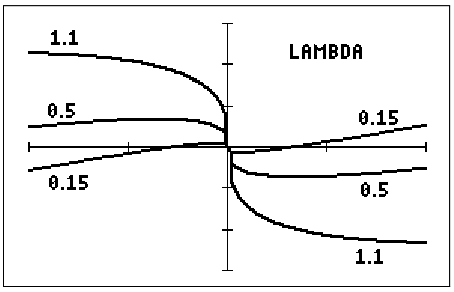
$f(1) = 1 - 3\lambda$ Endpoint 2

The biggest of these five numbers is the global maximum, the smallest the global minimum. Plot the functions $y = -1 + 3x$, $y = 2x^{3/2}$, $y = 0$, $y = -2x^{3/2}$ and $y = 1 - 3x$ for $x > 0$ to determine for which values of x (i.e. λ), which function is the biggest/smallest.

After some playing around with the window (try $[0, 1, 0.5] \times [-1.5, 1.5, 0.5]$) and *intersect*, we find the following.

- For $0 < \lambda < 1/4$, Endpoint 2 is the global maximum, Endpoint 1 the global minimum.
- At $\lambda = 1/4$, Endpoint 2 and Critical Point 1 are the global maxima, Endpoint 1 and Critical Point 3 the global minima.
- For $1/4 < \lambda < 1$, Critical Point 1 is the global maximum, Critical Point 3 the global minimum.
- At $\lambda = 1$, Critical Point 1 and Endpoint 1 are the same point $x = -1$; this is the global maximum. Critical Point 3 and Endpoint 2 are the same point $x = 1$; this is the global minimum.
- For $\lambda > 1$, the critical points no longer lie in $[-1, 1]$. Endpoint 1 is the global maximum, Endpoint 2 the global minimum.

$f(x) = x - 3\lambda x^{1/3}$



window $[-1, 1, 0.5] \times [-3, 3, 1]$

2.7 Bush Mathematics

The Problem

Your Maths class is out bush on a camp. One evening, a local farmer wanders into your camp and, hearing who you are, asks you to help him with a problem that has been keeping him awake at nights.

He has 4 km of fencing wire left over in his shed and, in the spirit of efficiency necessary these days for farms to make a profit, wants to make best use of it. After some discussions, you and he decide that 'best use' means that he wants to enclose maximum area with the wire. (For simplicity, we shall assume that the fence is single strand — if the fence had w strands, he would effectively have $4/w$ km of wire, and the problem would be the same.)

Following the steps below, try to find the best solution to the farmer's problem. Along the way, we may find out why most paddocks have four sides.

Part I

It is clear from looking around that rectangular paddocks are the norm. Your first task is to determine what shape rectangle encloses maximum area with the 4 km of wire.
Hint: Let the sides of the rectangle be x km and y km, and do the problem algebraically. Remember the scheme for tackling optimisation problems? See below.

- *Tell a lab instructor your solution to Part I. If it is correct, he or she will give you Part II.*
- *Write a progress report for the farmer: this report should contain a short summary in words of what you have done and discovered, a section detailing your calculations and a recommendation as to what he should do at this stage. The farmer has done Year-12 Mathematics.*

Procedure for Optimisation Problems

1. **Define variables with units.** Drawing a sketch is always a good idea too. At this stage, think about the problem: what should the answer be?
2. **Formulate the equation.** Write down an equation for the variable to be minimised/maximised as a function of the other variables. Use other equations to rewrite this equation in terms of one independent variable.
3. **Determine the domain of the function to be minimised/maximised.** Usually follows from the nature of the problem. Is the domain open or closed? If there is a choice, make it closed (endpoints included), because then there will always be a global minimum/maximum (if the function is continuous).
4. **Find the GLOBAL minimum/maximum.** Do this graphically or algebraically using the methods we have developed. If you find a **local** min/max, you have to argue in some way that it is also the **global** min/max.
5. **Answer the question.** Since the question was in words, your answer should be too. Remember units. Does your answer make sense?

Bush Mathematics
Part II

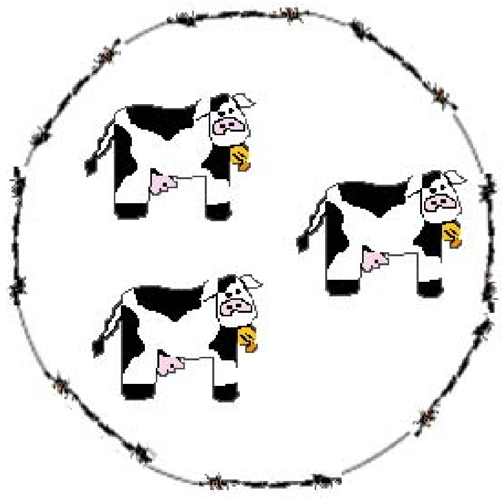
The farmer vaguely remembers something from school about circles and area. After some more discussion, you decide to work out what combination of a square and/or a circular paddock (a radical concept) will enclose maximum area with the 4 km of wire.

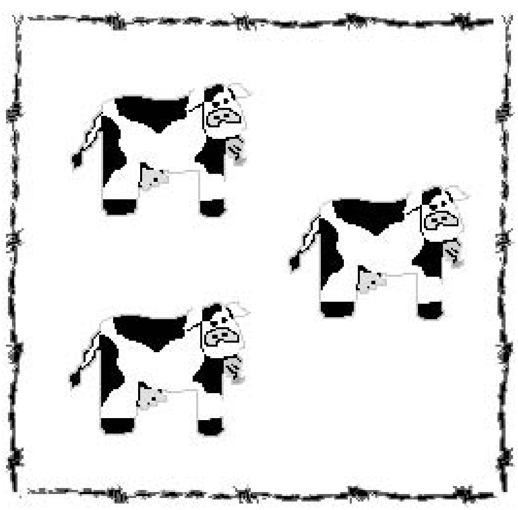
Hint: Have two paddocks, one a square with side x km, the other a circle of radius r km.

Again do the problem algebraically, guided by a graph. What is the maximum area you enclose?

Extension: The solution to this problem causes the farmer much excitement and he can't wait to tell his mates, who also have left-over wire, but of different lengths. You decide you had better solve the problem and find the maximum enclosed area for p km of wire. Then you can just apply the final formula to the length of wire that each farmer has.

- *Tell a lab instructor your solution to Part II. If it is correct, he or she will give you Part III.*
- *Don't forget a progress report for the farmer and a recommendation as to what he should do at this stage.*





Bush Mathematics
Part III

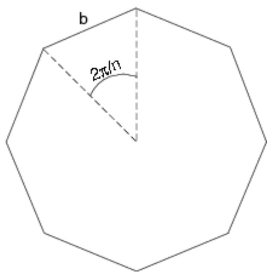
After the excitement over the discovery about circular paddocks has died down, you realise that to build a perfectly circular paddock, you will need a large number of posts — an infinite number in fact, if you assume that each post takes up negligible space. Realistically, what you will build is a polygonal paddock with n sides if you have n posts. The larger n, the closer the paddock to a circle. Archimedes went through a similar sort of paddock problem quite a while ago.

A polygonal paddock is no longer the optimal solution to your problem, so what sort of gain in area do you get with a paddock of n sides over a square one? An entry from a book of maths formulas, reproduced below, might help you to work out the area of a polygonal paddock of perimeter 4 km and n sides.[21]

Regular polygon of n sides, each of length b

$$\text{Area} = \frac{nb^2}{4\tan(\pi/n)}$$

$$\text{Perimeter} = nb$$

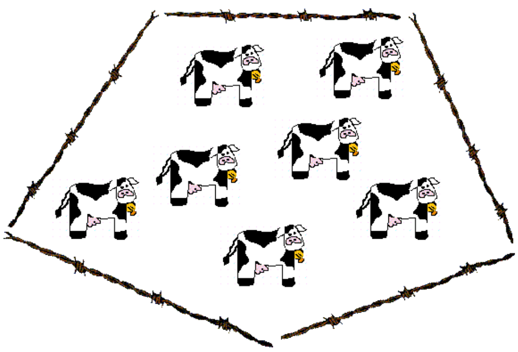


What's the best value of n for a perimeter of 4 km?

Hint: Graphics and/or a table of area values[22] are likely to be very handy here.[23]

Remember that here the independent variable n is discrete, not continuous, so we can't use differentiation.

- *Tell a lab instructor your solution to Part III. If it is correct, he or she will give you Part IV.*
- *Don't forget a progress report for the farmer and a recommendation as to what he should do at this stage.*



[21] You might even like to prove this result yourself. Look at the figure.

[22] Radian mode on your calculator is essential

[23] If you do the general case of p km of wire, you can plot/calculate values of Area/p^2, so that p does not enter into the determination of the optimum n. It will of course enter into the value for the maximum area.

Bush Mathematics
Part IV

So far we have not included the fact that posts cost money to buy and put in.

(a) Assume that there is a post at each vertex of the paddock. Let each of these n corner posts cost \$$c$ to buy and put in. As money is limited, a reasonable strategy now seems to be to maximise not area A enclosed by the 4 km of wire, but B, the area enclosed per dollar spent. What's the best n now? Does it depend on c?[24]

(b) We now have posts at each corner of the paddock, but fences also need smaller support posts to hold up the wire between the corner posts.[25] Suppose that these support posts cost \$$s$ to buy and put in ($s \leqslant c$).

Assume we need posts every 5 m, that is there is no length of fence between posts longer than 5 m. What value of n maximises area per dollar now for a perimeter of 4 km? *Hint*: What is the total number of posts? The total number of support posts? The total cost of the posts?

You should find that the optimum value of n depends on the ratio s/c (although the maximum value for area per dollar depends on c and s individually). Use $s/c = 0.005$ first, but then try a range of values $0 \leqslant s/c \leqslant 1$. Draw up a table of values of s/c and the corresponding optimum n.

Can your theory explain why most paddocks have $n = 4$?

(c) **Extension:** If the perimeter is p km and the maximum distance between posts is d km (*note*), what is the optimum n value?[26] You should find it now depends on two ratios, s/c and p/d, not on the four individual values.

Check your answer using $p/d = 4/0.005 = 800$, the values we used above.

What happens if $p/d = 400, 200, 100$? Is the optimum value of n sensitive to changes in the parameter p/d for the values used here?

- *Tell a lab instructor your solutions to Part IV. If they are correct, he or she will give you Part V.*
- *Don't forget a progress report for the farmer and a recommendation as to what he should do at this stage.*

[24] Graph/tablulate the function $c \times$ *area per dollar* $= cB$.

[25] This is still not the full story: a standard fence has support posts, typically star pickets, every 5 m. Roughly every tenth post is a larger wooden strainer post.

[26] The ceiling or least-integer function $C(x)$ might be useful here: $C(x)$ is the smallest integer greater than or equal to x: [math] NUM iPart(X + 0.9999) on your calculator.

Bush Mathematics

Part V

Another possible explanation for the choice of $n=4$ for paddocks is that farmers rarely have only one paddock. Square paddocks (and all other polygonal paddocks) can share sides if there is more than one paddock. Which of the three shapes of paddocks that have arisen so far, triangular, square and circular, comes out best when there are m paddocks sharing as many sides as possible? 'Best' here means that the m paddocks enclose maximum area for a total wire length of 4 km. Don't include the cost of fence posts just yet.

Hint: Circles are the easiest here. What is the total area of m identical circles for which the total length of wire used is 4 km? For triangles and squares, sketch how you would join together m paddocks ($m = 2, 3, 4, \dots$) most efficiently. Forming, as often as possible, hexagons with the triangles and an overall square with the squares, seems to be the best strategy. The whole question of 'best packing' arises here and leads off into many areas of chemistry and physics.

Now (for the very keen) include the cost of fence posts. What's the best n for m paddocks?

Write a short summary of your results in all parts of the lab. What advice are you going to give the farmer?

Congratulations on reaching and hopefully completing the last part of this lab.

Instructors' Guide

This is an open-ended lab, with the problems becoming successively harder. The idea is to see the steps in constructing a mathematical model, how to test it against the data (most paddocks are square or rectangular), then refine the model. The idea of handing out each part separately is that the group as a whole concentrates on one part at a time, rather than taking a part each. You can stop at any point that is appropriate for your class.

When you do stop, students should be asked to summarise all their results along the lines of ...

Write a full report for the farmer, summarising all your findings. Make a practical recommendation as to what he should do. What other considerations might have to be taken into account in finding the optimum solution to this problem? (Think about building a fence.)

Solutions and Discussion

Part I

This is the same problem as Question 2 of the *Optimisation* lab. The rectangle of maximum area is a square.

Part II

A standard problem in Calculus texts, usually involving string or wire. It is one of the few optimisation problems with an endpoint maximum, and therefore useful in pointing out this possibility to students. It was this problem and the attempt to put it in a more interesting context that led to this lab.

Define variables and units

Let $A\,\text{km}^2$ be the total area enclosed by the square and the circle. Let $x\,\text{km}$ be the side of the square and $r\,\text{km}$ be the radius of the circle.

Formulate the equation

$$\begin{aligned} A &= x^2 + \pi r^2 && \text{total area} \\ 4 &= 4x + 2\pi r && \text{total perimeter is } 4\,\text{km} \\ \therefore\ r &= \frac{4(1-x)}{2\pi} \\ \therefore\ A(x) &= x^2 + \frac{4}{\pi}(1-x)^2 && \text{area to be maximised} \end{aligned}$$

Find the domain

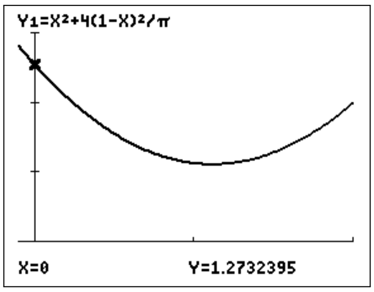
The domain here is $0 \leqslant x \leqslant 1$. $x=0$ means all the wire is in the circle, $x=1$ means all the wire is in the square. Choose the domain to be closed.

Find the global maximum

The function A is continuous and has a closed domain. The global minimum therefore exists and must occur at an endpoint of the domain or at a critical point.

Graphically

Plot $A(x)$ on $[0,1]$. The global maximum is clearly at the endpoint $x = 0$.



window $[-0.05, 1, 0.5] \times [0, 1.5, 0.5]$

Algebraically

$$\frac{dA}{dx} = 2x - \frac{8}{\pi}(1-x).$$

The derivative is defined for all x, so that it is defined for all x in the domain.

$A'(x) = 0$ when $x = 4/(4+\pi) \approx 0.56$: this is the only critical point.

$A(0) = 4/\pi \approx 1.27$ endpoint (all wire in circle)

$A\big(4/(4+\pi)\big) = 4/(4+\pi) \approx 0.56$ critical point

$A(1) = 1$ endpoint (all wire in square)

Therefore, the global maximum occurs at $x = 0$.

Answer the question

The maximum area is enclosed if all the wire goes into the circular paddock. The area is $4/\pi \approx 1.27\,\text{km}^2$, a 27% improvement over putting all the wire into a square paddock.

Extension

The equations for p km of wire are

$$\begin{aligned} A &= x^2 + \pi r^2 && \text{total area} \\ p &= 4x + 2\pi r && \text{total perimeter is 4 km} \\ \therefore\ r &= \frac{p - 4x}{2\pi} \\ \therefore\ A(x) &= x^2 + \frac{1}{4\pi}(p-4x)^2 && \text{area to be maximised} \end{aligned}$$

Find the domain

The domain here is $0 \leqslant x \leqslant p/4$. $x = 0$ means all the wire is in the circle, $x = p/4$ means all the wire is in the square. Choose the domain to be closed.

Find the global maximum

The function A is continuous and has a closed domain. The global minimum therefore exists and must occur at an endpoint of the domain or at a critical point.

$$\frac{dA}{dx} = 2x - \frac{1}{2\pi}\,(p - 4x).$$

The derivative is defined for all x, so that it is defined for all x in the domain.

$A'(x) = 0$ when $x = p/(4+\pi)$: this is the only critical point.

$A(0) = p^2/4\pi$ endpoint

$A\big(p/(4+\pi)\big) = p^2/4(4+\pi) < A(0)$ critical point

$A(p/4) = p^2/16 < A(0)$ endpoint

Therefore, the global maximum occurs at $x = 0$.

Answer the question

The maximum area is enclosed if all the wire goes into the circular paddock. The area is $p^2/4\pi$, still a 27% improvement over putting all the wire into a square paddock.

Part III

At this stage reality kicks in and makes the problem more interesting.

The area of a polygonal paddock of perimeter 4 km with n (equal) sides is

$$A = \frac{4}{n\tan(\pi/n)}\ \text{km}^2.$$

If we graph $A(x) = 4/x\tan(\pi/x)$ and only look at integer values or use the table feature of the calculator (probably easier), we find that A increases as n increases — there is no 'best' polygon. As $n \to \infty$, the area of the polygon approaches the area of the circle that we found in Part II.

In the table, X is n and Y1 is A.

```
\Y1=4/(Xtan(π/X))
```

X	Y1
3	0.7698
4	1
5	1.1011
6	1.1547
7	1.1866
8	1.2071
9	1.2211
10	1.2311
11	1.2384
12	1.244
13	1.2484

As the area of the square ($n = 4$) is 1, it is easy to see from a table of function values the improvement of successive polygons (with $n > 4$) over the square, but there is no obvious criterion for choosing the best polygon.

Part IV

Money had to come into the problem somewhere! The cost of the posts could also include the cost of stringing the wire from each post — students from a rural background may bring this up.

(a) The area enclosed by a polygon with n sides was found in Part III to be

$$A = \frac{4}{n\tan(\pi/n)}\ \text{km}^2.$$

The cost of n posts is $\$nc$, so the area enclosed per dollar is $B = A/nc$ or

$$B = \frac{4}{n^2 c\tan(\pi/n)}\ \text{km}^2 \text{ per dollar.}$$

To use a graph or table, we plot/tabulate $cB = 4/n^2\tan(\pi/n)$. The constant c does not affect the optimum n (if it exists), because the maximum of cB will occur at the same value of n as that for B. However, the value of c will determine the value of B for any given n.

In the table, X is n and Y1 is B.

```
Y1=4/(X²tan(π/X))
```

X	Y1
3	0.2566
4	0.25
5	0.2202
6	0.1925
7	0.1695
8	0.1509
9	0.1357
10	0.1231
11	0.1126
12	0.1037
13	0.096

B is only defined for $n \geqslant 3$ (for a viable paddock) and we find that it is a decreasing function of n: the paddock enclosing the maximum area per dollar spent is a triangular paddock.

(b) If there are posts every 5 m, there will be $4000/5 = 800$ posts altogether. As we have n corner posts, we will need $800-n$ support posts. The total cost of the posts is therefore $nc+s(800-n)$ dollars, and the area per dollar spent is

$$\begin{aligned} B &= \frac{4}{n\tan(\pi/n)} \cdot \frac{1}{nc+s(800-n)} \\ &= \frac{4}{nc\tan(\pi/n)} \cdot \frac{1}{n+\frac{s}{c}(800-n)}. \end{aligned}$$

The optimum value of n now depends on the ratio s/c.

Using a table of values of cB versus n, we find the following optimum values of n as a function of the ratio of post costs s/c. The relative values of area per dollar spent, cB, at optimum are also given.

s/c	**Optimum** n	cB	**Comments**
0	3	0.2566	corner posts only
0.005	4	0.1253	relative cost of support posts increasing
0.01	5	0.0850	↓
0.05	7	0.0254	
0.1	9	0.0139	
0.2	11	0.0073	
0.5	18	0.0031	
0.9	36	0.0018	
0.99	81	0.0016	
0.999	174	0.0016	
1	800	0.0016	maximum possible number of posts

As the price of support posts goes up, corner posts are relatively less expensive and so we can have more, enclosing greater area. While there is a value of s/c that gives an optimum value of $n = 4$, i.e. a square paddock, the optimum value is quite sensitive to changes in s/c. The graphs of cB versus n are also quite flat, particularly as s/c increases, so that choosing a value of n that is not optimum makes little difference to the value of B.

It therefore seems unlikely that this model explains why most paddocks have $n = 4$.

(c) The area enclosed by a polygon of perimeter p km with n sides is

$$A = \frac{p^2}{4n\tan(\pi/n)}\ \text{km}^2.$$

The total number of posts required so that the maximum spacing between posts is d km is $C(p/d)$, where $C(x)$ is the smallest integer greater than or equal to x, also called the ceiling function.[27] We therefore have n corner posts costing \$$c$ each and $C(p/d) - n$ support posts costing \$$s$ each. The total cost of posts in dollars is then

$$nc + s\big(C(p/d) - n\big) = nc\left(1 + \frac{s}{c}\left(\frac{1}{n}C(p/d) - 1\right)\right),$$

and the area enclosed per dollar spent as

$$\begin{aligned} B &= \frac{p^2}{4n\tan(\pi/n)} \cdot \frac{1}{nc\left(1 + \frac{s}{c}\left(C(p/d)/n - 1\right)\right)} \\ &= \frac{p^2}{4n^2c\tan(\pi/n)}\,\frac{1}{1 + \frac{s}{c}\left(C(p/d)/n - 1\right)}. \end{aligned}$$

[27] On a TI-84/CE, you can use iPart(X+0.9999) to generate the ceiling function.

To simplify the calculations, we assume that the perimeter is exactly divisible by the spacing between posts. Then p/d is an integer, giving

$$B = \frac{p^2}{4n^2 c \tan\left(\frac{\pi}{n}\right)} \frac{1}{1+\frac{s}{c}\left(\frac{p}{nd}-1\right)}.$$

The maximum of B will occur at the same value of n as that of

$$\frac{4c}{p^2} B = \frac{1}{n^2 \tan\left(\frac{\pi}{n}\right)} \frac{1}{1+\frac{s}{c}\left(\frac{p}{nd}-1\right)}.$$

and this is what we tabulate as a function of n for different values of s/c and p/d. For each tabulation, we then find the n to give maximum B.

The table below shows the optimum n for the same values of s/c as in (b) and with the number of posts $p/d = 800, 400, 200, 100$. The first of these values is what we used in (b), and corresponds to 5 m post spacing in a total perimeter of 4 km. The other values for p/d correspond to 10 m, 20 m and 40 m post spacings respectively in a total perimeter of 4 km.

	$\boldsymbol{p/d}$			
$\boldsymbol{s/c}$	**800**	**400**	**200**	**100**
0	3	3	3	3
0.05	7	6	5	4
0.1	9	7	6	5
0.2	11	9	7	6
0.5	18	14	11	9
0.9	36	29	23	18
1	800	400	200	100

As in (b), the graphs of $\left(\frac{4c}{p^2}\right)B$ versus n are quite flat, becoming flatter as s/c increases, and again the model does not seem to explain why $n=4$ in practice.

Part V

Joined paddocks, cost of posts not included

Circles

If there are m circular paddocks, each of radius r, with total perimeter p, we have total perimeter $p=2\pi rm$, so that $r=\dfrac{p}{2\pi m}$ and total area enclosed is

$$A = m\pi r^2 = \frac{p^2}{4\pi m}.$$

Squares and Triangles

For the number of paddocks $m = 1, 2, 3, \ldots$, count the number of units (length of one side of a paddock) required. The fixed perimeter then determines the length of one unit. You can then use the formula for the area of one paddock in terms of the length of one side to determine the total enclosed area. In all cases, we can take out a factor of p^2 and tabulate A/p^2. The aim in putting the paddocks together is to have as few sides as possible, so the more shared sides the better. However, the results below do not represent a full treatment of the problem, merely some experimentation on the part of the author. It is this sort of experimentation that we wish to encourage in our students.

The table below shows the values of A/p^2 as a function of m. The largest of the three values for each m (underlined) gives the best option of the three paddock shapes considered here.

	$\boldsymbol{A/p^2}$		
$\boldsymbol{m}$	**Circles**	**Squares**	**Triangles**
1	$\underline{0.0796}$	0.0625	0.0481
2	0.0398	$\underline{0.0408}$	0.0346
3	0.0265	$\underline{0.0300}$	0.0265
4	0.0199	$\underline{0.0278}$	0.0214
5	0.0159	$\underline{0.0222}$	0.0179
$\vdots$	$\vdots$	$\vdots$	$\vdots$

As $m \to \infty$, $A/p^2 \sim 1/4\pi m$ for circles, $A/p^2 \sim 1/4m$ for squares and $A/p^2 \sim 1/6.4m$ for triangles.

Except for $m = 1$, square paddocks are always best. They are better than circles because square paddocks can share sides and better than triangles presumably because they enclose more area for a given perimeter (Part III). Perhaps this is the reason why most paddocks are square (or rectangular).

Joined paddocks, cost of posts included

If we include the cost of posts, considering only corner posts, triangles come out better than squares for all m, and we suspect triangles are the the best of all shapes, based on our experience in Part IV (b). The overall strategy here seems to be the same as when the cost of the posts is not included, because the more shared sides, the fewer posts.

The table below shows B_m/p^2, where $B_m = A_m/nc$ is the area enclosed by m square or triangular paddocks divided by nc, the cost of n corner posts. B_m/p^2 is proportional to the area enclosed per dollar spent. The total number of posts required is also shown.

	Squares		**Triangles**	
m	B_m/p^2	**Posts**	B_m/p^2	**Posts**
1	0.0156	4	<u>0.0160</u>	3
2	0.0058	6	<u>0.0087</u>	4
3	0.0038	8	<u>0.0053</u>	5
4	0.0031	9	<u>0.0043</u>	6
5	0.0020	11	<u>0.0026</u>	7
6	0.0017	12	<u>0.0026</u>	7
7	0.0013	14	<u>0.0019</u>	8
8	0.0012	15	<u>0.0015</u>	9
9	0.00098	16	<u>0.0012</u>	10
10	0.00076	18	<u>0.0012</u>	10
11	0.00069	19	<u>0.00098</u>	11
12	0.00063	20	<u>0.00082</u>	12
⋮	⋮	⋮	⋮	⋮

Conclusions

For a single paddock, the best mathematical solution (maximum area enclosed for a given amount of wire) if the cost of posts and construction is not included is a circular paddock.

The best practical solution is a polygonal paddock with as many sides as possible, although the gain in area from adding an extra side decreases as the number of sides increases.

If the cost of posts and construction is included, the best solution (maximum area enclosed per dollar spent) is a triangular paddock.

If the cost of posts and construction is not included, the maximum area enclosed by m paddocks ($m > 1$) constructed using a fixed amount of wire is given by square paddocks joined optimally.

If the cost of posts and construction is included, the maximum area enclosed by m paddocks (constructed using a fixed amount of wire) per dollar spent is given by triangular paddocks joined optimally.

2.8 Rectangles, Area and the Definite Integral

The TI-84/CE program NINTGRPH/NINTGRCE[28] illustrates graphically how the area under a graph can be approximated by the areas of rectangles. As the number of rectangles covering the area increases, we obtain a better approximation to the area.

Here we calculate approximations to $\int_0^1 e^x \, dx$ by drawing the rectangles: the sum of their areas is called a *Riemann sum*.

- Put the function $f(x) = e^x$ in Y1.
- Set a [window] of $[0, 1, 0.2] \times [0, 3, 1]$.
- Run the program: press [prgm], press the number against its name and press [enter] to start the program.
- Set the integration limits A = 0 and B = 1.
- Set the number of rectangles $N = 5$.
- Choose the Left-Endpoint Rule (LER). The program will plot the function and draw in 5 rectangles, each rectangle touching the curve at its top-left corner. In this case, the area of the rectangles clearly underestimates the area under the graph.
- Press [enter] to see the area of the rectangles (LHSUM) as an approximation to the area under the graph. Put your answer in the table below. Round your answers to 3 decimal places.
- Press [enter], set $N = 5$ again and choose the Right-Endpoint Rule (RER). This time we obtain an overestimate of the area under the curve.
- Repeat the above two steps, doubling the number of rectangles each time.

  The mean of the two estimates is equivalent to the Trapezoidal-Rule approximation to the area, a more accurate approximation for a given N than either the Left- or Right-Endpoint Rules.

N	**LER**	**RER**	**Mean**
5			
10			
20			
40			
80			

Your best estimate:

$$\int_0^1 e^x \, dx \approx$$

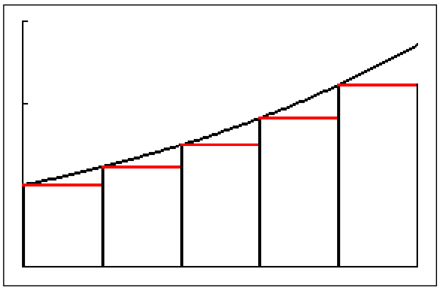
Compare this with the exact value if you know how to do the integral.

- When you've finished, select QUIT in the RULE menu.

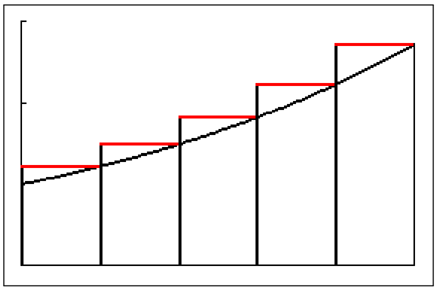
[28]NINTGRPH is for the TI-84; NINTGRCE for the TI-84CE.

Instructors' Guide

Answers *rounded to 3 decimal places.*



Left-Endpoint Rule



Right-Endpoint Rule

N	LER	RER	Mean
5	1.552	1.896	1.724
10	1.664	1.806	1.720
20	1.676	1.762	1.719
40	1.697	1.740	1.718
80	1.708	1.729	1.718

Best estimate is 1.718 from the Mean column. The fact that two successive values give 1.718 indicates that this is probably the exact answer rounded to 3 decimal places,

The exact answer is $e-1=1.718$ rounded to 3 decimal places.

2.9 Approximating Definite Integrals

Modified from an UNSW Canberra Maths Lab, which is itself based on a lab in *Resources for Calculus, Volume 1: Learning by Discovery*, Anita Solow, editor, Mathematical Association of America Note 26, 1993.

In this lab, we shall be comparing several numerical approximations to

$$\int_0^1 (5x^4 - 3x^2 + 1)\, dx$$

with the exact answer obtained by algebraic integration. This will give us a feel for some of the methods of numerical integration, which we can then use for any function, including those which cannot be integrated algebraically.

Question 1 *Algebraic integration — the exact answer*

What is the exact value of this integral? You may not realise it, but you are using the *Fundamental Theorem of Calculus* to do this definite integral exactly.

Question 2 *The Left-Endpoint Rule*

One approach to numerical integration is to approximate the definite integral of $y = f(x)$ with $a \leqslant x \leqslant b$ by the sum of the areas of a number of rectangles covering the region under the curve. If the top left-hand corner of each rectangle touches the curve, we have the *Left-Endpoint Rule*; if the top right-hand corner of each rectangle touches the curve, we have the *Right-Endpoint Rule.*

As the number of rectangles in the interval $[a, b]$ gets larger and larger (covering the integration range $a \leqslant x \leqslant b$ with more and more, thinner and thinner rectangles), both rules give numbers closer and closer to the definite integral (exact answer).

(a) On Figure 1 (at the end of this lab), draw and shade in the rectangles for the Left-Endpoint-Rule approximation to the definite integral $\int_a^b f(x)\, dx$ with N = 4 (4 rectangles).

(b) Using your sketch in (a), explain why the Left-Endpoint Rule with 4 rectangles approximates the area under the graph as

$$h\big(f(x_0) + f(x_1) + f(x_2) + f(x_3)\big),$$

where $x_0 = a$, $x_4 = b$ and the width of each rectangle is $h = (b - a)/4$.

(c) Use the NINTGRPH/NINTGRCE program (instructions below)[29] to estimate the definite integral $\int_0^1 (5x^4 - 3x^2 + 1)\, dx$ using the Left-Endpoint Rule with the number of rectangles N = 4. A suitable window is $[0, 1, 0.5] \times [0, 3, 1]$.

Note that the integrand here is positive, so that the definite integral corresponds to the area under the graph of f.

(d) Now use the program, doubling N until two successive answers from the Left-Endpoint Rule are the same when rounded to 2 decimal places. Write down the N value of the first of these two answers.

[29] NINTGRPH is for the TI-84; NINTGRCE for the TI-84CE.

Question 3 *The Trapezoidal Rule*

The Left-Endpoint and Right-Endpoint Rules approximate the area under a function by rectangles. In many cases, for example the function in Figure 1 with the rectangles you drew in, this is not a good approximation. We get a better approximation by using trapeziums: both top corners of each trapezium touch the curve.

(a) On Figure 2, draw and shade in the 4 trapeziums (N = 4), the total area of which approximates the definite integral $\int_a^b f(x)\,dx$.

The area of the trapezium in Figure 3 is $h(r+s)/2$. To see this result, split the trapezium into two regions — a triangle and a rectangle.

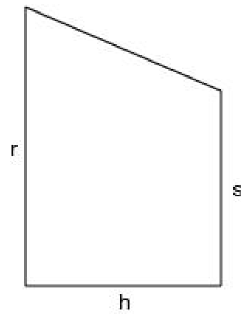


Figure 3

(b) Using your sketch in (a) and Figure 3, explain why the Trapezoidal Rule with 4 trapeziums approximates the area under the graph as

$$\frac{h}{2}\big(f(x_0) + 2f(x_1) + 2f(x_2) + 2f(x_3) + f(x_4)\big),$$

where $x_0 = a$, $x_4 = b$ and the width of each trapezium is $h = (b-a)/4$.

(c) Evaluate T_4, the Trapezoidal Rule with 4 trapeziums, as an estimate of the integral $\int_0^1 (5x^4 - 3x^2 + 1)\,dx$ using NINTGRPH/NINTGRCE.

How does this result compare with the Left-Endpoint result and the exact answer?

(d) Now use the program, doubling N, the number of trapeziums, until two successive answers are the same when rounded to 2 decimal places. Write down the N value of the first of these two answers. Compare it with the Left-Endpoint value.

Question 4 *Simpson's Rule*

A picture of Simpson's Rule for which N = 4 is given in Figure 4. We want to estimate the area under the solid curve. We do this by a fitting parabola to each set of 3 successive points on the graph and adding up the areas under the parabolas.

The dashed line in Figure 4 shows two parabolas: one through $\big(x_0, f(x_0)\big)$,$\big(x_1, f(x_1)\big)$ and $\big(x_2, f(x_2)\big)$; the other through $\big(x_2, f(x_2)\big)$, $\big(x_3, f(x_3)\big)$ and $\big(x_4, f(x_4)\big)$.

(a) On Figure 4, shade the area calculated by Simpson's Rule as an approximation to the definite integral $\int_a^b f(x)\,dx$.

(b) Evaluate S_4, Simpson's Rule with 4 sub-divisions of the integration interval, as an estimate of $\int_0^1 (5x^4 - 3x^2 + 1)\,dx$ using NINTGRPH/NINTGRCE.

The dotted lines are the two parabolas.

Compare your result with those from Questions 1 − 3.

(c) Now use the program, doubling N, the number of sub-divisions of the integration interval, until two successive answers are the same when rounded to 2 decimal places. Write down the N value of the first of these two answers.

Compare your result with those from Questions 2 and 3.

Question 5 *Comparing the methods*

Repeat your first (N = 4) and last calculations above for the three methods, this time keeping 5 decimal places. Put them in a summary table, together with the h value and the absolute value of the error $|E|$ for each entry (you know the exact answer).

The NUMINT/NUMINTCE program (no graphics) might be faster for this but note the N for Simpson's Rule (instructions below).

If $|E| = kh^m$, where k is a constant, find m for each method. *Hint*: Use the two sets of values of h and the corresponding errors to write down two equations for k and m; divide one equation by the other to obtain an equation for m. You'll need natural logs to isolate m. *Hint*: We are looking for integer values.

How many times do you have to double N in each method to improve the accuracy by a factor of 10?

What conclusions can you draw from your results regarding the different methods for estimating the definite integral? Which method would you choose to use? Why?

Calculator Programs

These programs calculate approximate values for $\int_{A}^{B} f(X)\, dX$.
The number N is an input to the program.

NINTGRPH/NINTGRCE approximates the integral using one of five different rules **with N sub-divisions** of the interval [A, B] (N+1 if N is odd, for Simpson's Rule), and draws the corresponding approximations to the function on each subinterval.

NUMINT/NUMINTCE approximates the integral using the *Left-Endpoint Rule* (**L**), the *Right-Endpoint Rule* (**R**), the *Trapezoidal Rule* (**T**) and the *Midpoint Rule* (**M**), **all with N sub-divisions**, and *Simpson's Rule* (**S**) **with 2N sub-divisions** to ensure an even number of sub-divisions.

Use: Type the function to be integrated into Y1.

- For NINTGRPH/NINTGRCE, first set a suitable window to display the function. Run the program and follow the prompts. Make sure B>A, otherwise things get mixed up. After the graph is plotted, press [enter] to see the numerical approximation to the integral.

  When you've finished, select QUIT in the RULE menu.

- For NUMINT/NUMINTCE, run the program and follow the prompts.

  [on] Quit stops the program.

Numerical Integration Lab Figures

The function here is not the function in Question 1.

Figure 1: Left-Endpoint Rule

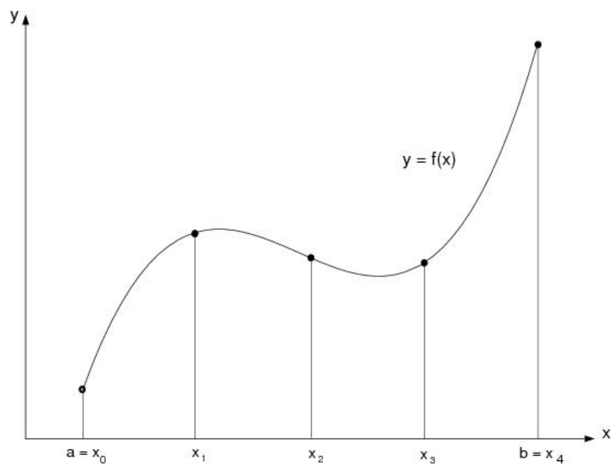


Figure 2: Trapezoidal Rule

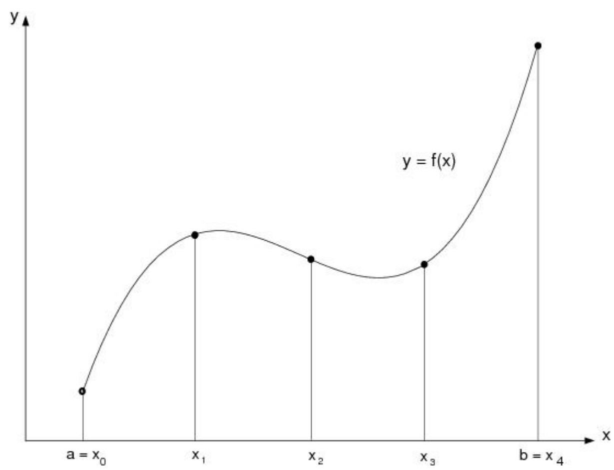
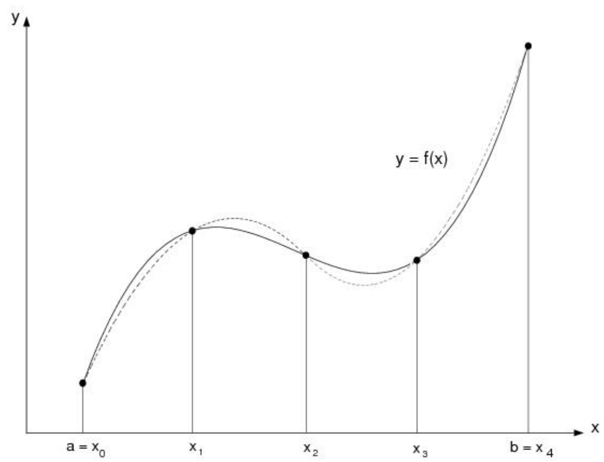


Figure 4: Simpson's Rule



Instructors' Guide

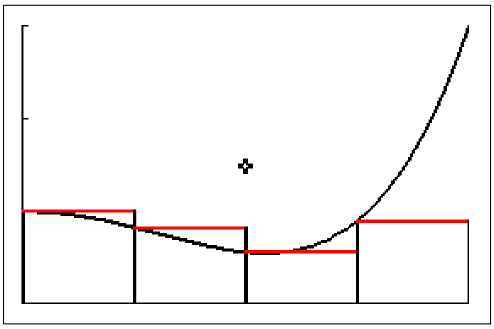
Question 1 *Algebraic integration — the exact answer*

$$\int_0^1 (5x^4 - 3x^2 + 1)\,dx \;=\; \Big[x^5 - x^3 + x\Big]_0^1 \;=\; 1.$$

Question 2 *The Left-Endpoint Rule*

(a) On Figure 1, draw and shade in the rectangles for the Left-Endpoint-Rule approximation to the definite integral $\int_a^b f(x)\,dx$ with $\mathrm{N} = 4$.

Done here with the function in Question 1 using the NINTGRCE program (no shading). The horizontal red lines show the approximation to the function f on each sub-interval.



window $[0, 1, 0.5] \times [0, 3, 1]$

(b) Using your sketch in (a), explain why the Left-Endpoint Rule with 4 rectangles approximates the area under the graph as

$$h\big(f(x_0) + f(x_1) + f(x_2) + f(x_3)\big),$$

where $x_0 = a$, $x_4 = b$ and the width of each rectangle is $h = (b - a)/4$.

The formula is just the sum of the areas of the 4 rectangles in (a) with h factored out.

(c) Use the NINTGRPH/NINTGRCE program to estimate $\int_0^1 (5x^4 - 3x^2 + 1)\,dx$ using the Left-Endpoint Rule with $\mathrm{N} = 4$. A suitable window is $[0, 1, 0.5] \times [0, 3, 1]$.

Note that the integrand here is positive, so that the definite integral corresponds to the area under the graph of f.

See the figure in (a).

The Left-Endpoint Rule with $\mathrm{N} = 4$ gives $\displaystyle\int_0^1 (5x^4 - 3x^2 + 1)\,dx \approx 0.822.$

(d) Now use the program, doubling N, the number of rectangles, until two successive answers from the Left-Endpoint Rule are the same when rounded to 2 decimal places. Write down the N value of the first of these two answers.

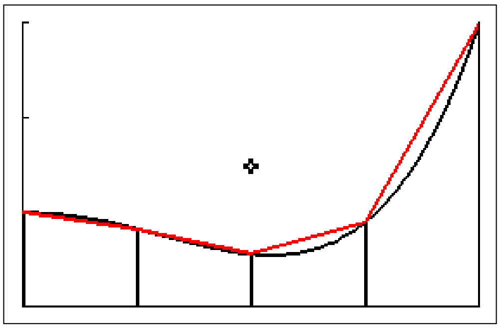
N	LER
4	0.82
8	0.89
16	0.94
32	0.97
64	0.98
128	0.99
256	1.00
512	1.00

The required N value is therefore 256. The last two values are the same as the exact answer rounded to 2 decimal places.

Question 3 *The Trapezoidal Rule*

(a) On Figure 2, draw and shade in the 4 trapeziums (N = 4), the total area of which approximates the definite integral $\int_a^b f(x)\,dx$.

Again done here with the function in Question 1 but without shading. The red straight lines show the approximation to the function f on each sub-interval.



window $[0, 1, 0.5] \times [0, 3, 1]$

(b) Using your sketch in (a) and Figure 3, explain why the Trapezoidal Rule with 4 trapeziums approximates the area under the graph as

$$\frac{h}{2}\big(f(x_0) + f(x_1) + 2f(x_2) + 2f(x_3) + f(x_4)\big),$$

where $x_0 = a$, $x_4 = b$ and the width of each trapezium is $h = (b-a)/4$.

The total area of the four trapeziums each of width $h = 0.25$ is, using the given formula,

$$h\,\frac{f(0)+f(x_1)}{2} + h\,\frac{f(x_1)+f(x_2)}{2} + h\,\frac{f(x_2)+f(x_3)}{2} + h\,\frac{f(x_3)+f(x_4)}{2}$$
$$= \frac{h}{2}\big(f(x_0)+2f(x_1)+2f(x_2)+2f(x_3)+f(x_4)\big).$$

(c) Evaluate T_4, the Trapezoidal Rule with 4 trapeziums, as an estimate of the definite integral $\int_0^1 (5x^4 - 3x^2 + 1)\,dx$ using NINTGRPH/NINTGRCE.
How does this result compare with the Left-Endpoint result and the exact answer?

$T_4 = 1.07$, considerably closer to the exact answer 1 than the Left-Endpoint Rule value with N = 4 of 0.82.

(d) Now use NINTGRPH/NINTGRCE, doubling N, the number of trapeziums, until two successive answers are the same when rounded to two decimal places. Write down the N value of the first of these two answers. Compare it with the Left-Endpoint N value.

N	TRAP
4	1.07
8	1.02
16	1.00
32	1.00

The required N value is therefore 16, compared with the much larger value of 256 for the Left-Endpoint Rule. The last two values are the same as the exact answer rounded to 2 decimal places.

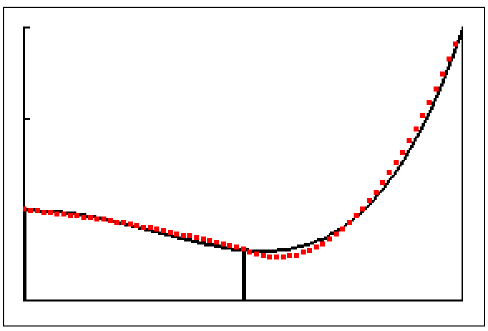
Question 4 *Simpson's Rule*

(a) On Figure 4, shade the area calculated by Simpson's Rule as an approximation to the definite integral $\int_a^b f(x)\,dx$.

Shade below the dotted curves in Figure 4.

(b) Evaluate S_4, Simpson's Rule with 4 sub-divisions of the integration interval (N = 4), as an estimate of $\int_0^1 (5x^4 - 3x^2 + 1)\,dx$ using NINTGRPH/NINTGRCE.

The red dotted lines show the two parabolic approximations to the function f on each two sub-intervals.



window $[0, 1, 0.5] \times [0, 3, 1]$

Compare your result with those from Questions 1 – 3.

$S_4 = 1.003$, considerably closer to the exact answer 1 than the other two values.

(c) Now use NINTGRPH/NINTGRCE, doubling the number N of sub-divisions of the integration interval each time, until two successive answers are the same when rounded to two decimal places. Write down the N value of the first of these two answers.

Compare it with the rectangle and trapezium values.

N	SIMP
4	1.00
8	1.00

The required N value is therefore 4, compared with 16 for the Trapezoidal Rule and the much larger value of 256 for the Left-Endpoint Rule. The two values here are the same as the exact answer rounded to 2 decimal places.

Question 5 *Comparing the methods*

If $|E| = kh^m$, where k is a constant, find m for each method.

Summary table $|E| = |1 - \text{Value}|$

Method	**N**	$\boldsymbol{h}$	**Value** (5DP)	$\boldsymbol{\lvert E\rvert}$ (5DP)
Left Endpoint	4	0.25	0.82227	0.17773
Left Endpoint	256	0.0039	0.99611	0.00389
Trapezoidal	4	0.25	1.07227	0.07227
Trapezoidal	16	0.0625	1.00455	0.00455
Simpson	4	0.25	1.00260	0.00260
Simpson	8	0.125	1.00016	0.00016

Assume error $|E| = kh^m$, where k is a constant. Therefore, if we have values E_1, h_1, E_2 and h_2 for a method,

$$\frac{E_1}{E_2} = \frac{h_1^m}{h_2^m} = \left(\frac{h_1}{h_2}\right)^m.$$

Taking natural logs of both sides and solving for m, we have

$$m = \frac{\ln\left(\frac{E_1}{E_2}\right)}{\ln\left(\frac{h_1}{h_2}\right)}.$$

Substituting in the two values for E and h for each method, we get $m \approx 1$ for the Left-Endpoint Rule, $m \approx 2.0$ for the Trapezoidal Rule and $m \approx 4.0$ for Simpson's Rule.

How many times do you have to double N in each method to improve the accuracy by a factor of 10?

If you double N, you halve h. Therefore, for the Left Endpoint Rule, you reduce the error by 0.5^1, i.e. by a factor of 2. For the Trapezoidal Rule, you reduce the error by 0.5^2, i.e. by a factor of 4. For Simpson's Rule, you reduce the error by 0.5^4, i.e. by a factor of 16.

To achieve an improvement in accuracy of 1 decimal place, you have to reduce the error by a factor of 10: therefore you have to double N (halve h) four times ($2^3 = 8 < 10$; $2^4 = 16 > 10$) using the Left Endpoint Rule, twice $\left((2^2)^2 = 16 > 10\right)$ for the Trapezoidal Rule and only once $\left((2^4)^1 = 16 > 10\right)$ for Simpson's Rule.

What conclusions can you draw from your results regarding the different methods for estimating the definite integral? Which method would you choose to use? Why?

For a given N or h, Simpson's Rule gives the most accurate approximation to the definite integral. To calculate an approximation to a given accuracy, it will therefore be the fastest of the three methods.

2.10 Projectile

Based on a problem suggested by John Rickert, Rose-Hulman Institute of Technology, Terre Haute, Indiana, USA.

Warm-up Mission

Send a projectile, with initial speed 25 m/s, as far as possible over level ground from a launch site. You might like to read the hints on the next page.

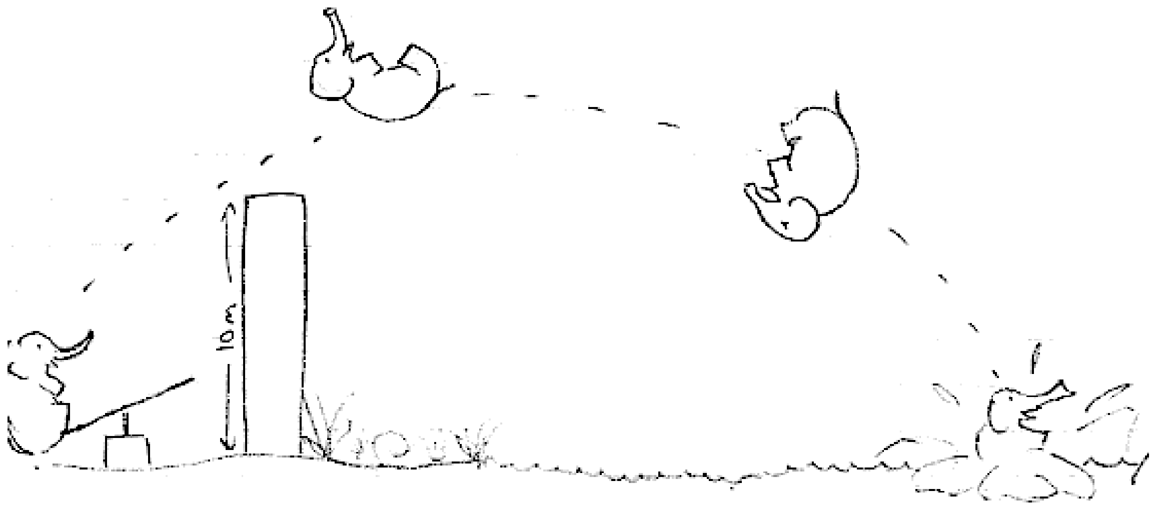
Main Mission

Send a projectile from ground level over a wall 10 m high, so that it lands as far as possible on the other side of the wall. The projectile can be launched at any distance from the wall. Its initial speed is 25 m/s. Write down all the assumptions you make. Draw a sketch of your final trajectory, labelling all relevant distances.

Note that the optimal launch angle for this mission is different to the one you obtained for the warm-up mission.

Scenario

In about half a page, say what the projectile is and why you want to send it over the wall (or other obstacle). Drawings/diagrams welcome.



Elephants negotiating the Great Wall of China by Caroline McKenna

Hints

- Draw a sketch for each mission on which you define the relevant variables. Let R be the range (the total horizontal distance covered by the projectile with or without the wall), let W be the horizontal distance from the launch site to the wall and let θ be the launch angle (to the horizontal).

  You need to derive expressions for $R(\theta)$ and $W(\theta)$, starting with the basic differential equations.

- First solve the differential equations[30] for *velocity* in the vertical (y) direction

$$\frac{dv_y}{dt} = -g$$

  and in the horizontal (x) direction

$$\frac{dv_x}{dt} = 0.$$

  Don't forget constants of integration. You need to supply the initial conditions.

- Then solve the differential equations for *position* in the vertical direction

$$\frac{dy}{dt} = v_y$$

  and in the horizontal direction

$$\frac{dx}{dt} = v_x.$$

  Note: Some of you know Newton's equations of motions. In this lab, you don't just write them down, you *derive* them from basics.

- In *both* missions, knowing a y value allows you to find the corresponding t, which allows you to find x.

- You should get for the range

$$R(\theta) = \frac{625\sin(2\theta)}{g}.$$

  Remember that $\sin(2\theta) = 2\sin(\theta)\cos(\theta)$. Take $g = 9.8\,\text{m/s}^2$.

- The PROJ/PROJCE program allows you to simulate the motion of the projectile. You can work out the answers using PROJ/PROJCE, but for the missions you need to find the expression for the distance over the wall (as a function of θ) algebraically.

[30] The parameter θ will occur in these equations.

Instructors' Guide

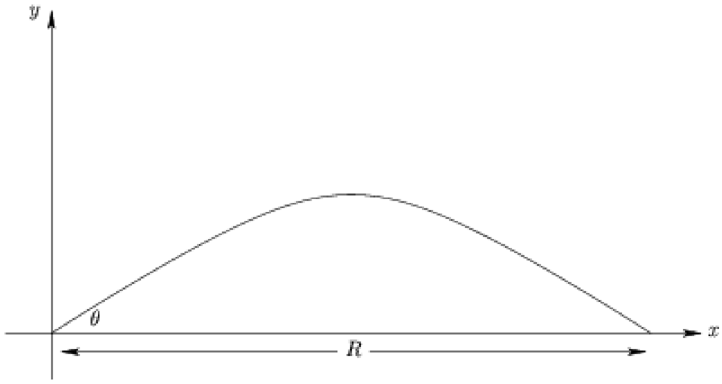
The PROJ/PROJCE program graphs the wall and the trajectory of the projectile, given a value for the launch angle θ. The graphics give the students a picture of what is happening, and the program also allows them to convince themselves (particularly students taking Physics) that the launch angle θ for maximum distance *beyond the wall* is not 45°. It may be a good idea to encourage students to use such a program first to obtain insight into the problem.

PROJ/PROJCE is available at *canberramaths.org.au* under *Resources*.

An extension to this lab for keen students is to include air resistance (more assumptions).

Solutions

Let the projectile be launched at the origin at an angle θ to the x axis.



Integrating the differential equation for v_y, we obtain

$$v_y(t) \;=\; -gt + C,$$

where C is an arbitrary constant. Now $v_y(0) = 25\sin(\theta)$, so that $C = 25\sin(\theta)$. Therefore,

$$v_y(t) \;=\; 25\sin(\theta) - gt.$$

Integrating $dy/dt = v_y$ using the expression just obtained for v_y and the initial condition $y(0)=0$ gives for the vertical position of the projectile,

$$y(t) \;=\; 25t\sin(\theta) - \frac{1}{2}gt^2.$$

Following similar steps, starting with the differential equation for v_x and using the initial conditions $v_x(0)=25\cos(\theta)$, $x(0)=0$,

$$\begin{aligned} v_x(t) &= 25\cos(\theta) \\ x(t) &= 25t\cos(\theta). \end{aligned}$$

The range R is the horizontal distance covered when the projectile lands, i.e. when $y(t)=0$. Solving $y(t) = 25t\sin(\theta) - \frac{1}{2}gt^2 = 0$ for time t gives $t=0$ (the launch) or

$$t_R \;=\; \frac{50\sin(\theta)}{g}.$$

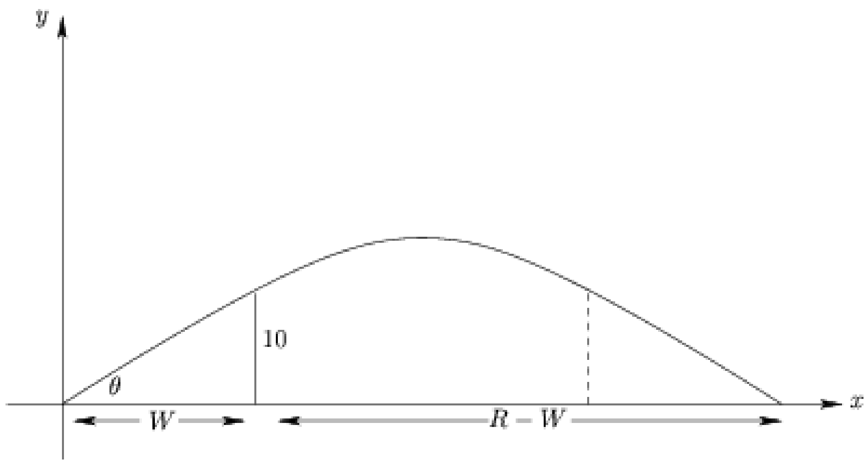
The range is $R = x(t_R)$, i.e.

$$R = \frac{1250\sin(\theta)\cos(\theta)}{g} = \frac{625\sin(2\theta)}{g}.$$

The maximum value $R = 625/g \approx 63.8\,\mathrm{m}$ is achieved when $\sin(2\theta) = 1$, i.e. when $\theta = \pi/4\,\mathrm{rad} = 45°$.

In the warm-up mission, we send the projectile a maximum horizontal distance of approximately 63.8 m by launching it at an angle of 45° to the horizontal.

For the main mission, we have a wall of height 10 m, whose base we take to be W m away from the launch position, i.e. at $x = W$. We then have to maximise the distance $R - W$.



We have

$$\begin{aligned} x(t) &= 25t\cos(\theta) \\ y(t) &= 25t\sin(\theta) - \frac{1}{2}gt^2 \\ R(\theta) &= \frac{625\sin(2\theta)}{g}. \end{aligned}$$

To find $W(\theta)$, we use the fact that the desired trajectory just touches the top of the wall, i.e. it passes through the point $(W, 10)$, so that the time the projectile takes to reach the wall is given by the solution of $y(t) = 10$. Once we know this time t_W, the distance to the wall is given by $W = x(t_W)$. Solving $y(t) = 10$, we have

$$\begin{aligned} 25t\sin(\theta) - \frac{1}{2}gt^2 &= 10. \\ \therefore\ gt^2 - 50t\sin(\theta) + 20 &= 0. \\ \therefore\ t &= \frac{50\sin(\theta) \pm \sqrt{(2500\sin^2(\theta) - 80g)}}{2g}. \end{aligned}$$

We expect to obtain two values for t, because the trajectory passes through $y = 10$ twice, once on the way up when it touches the wall and once on the way down (dashed line in the previous figure). To maximimize the distance beyond the wall, we clearly want the shorter of the two times, given by the minus sign. Therefore

$$t_W = \frac{50\sin(\theta) - \sqrt{(2500\sin^2(\theta) - 80g)}}{2g}.$$

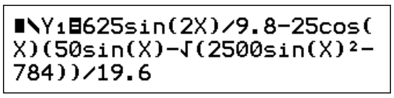
Therefore,

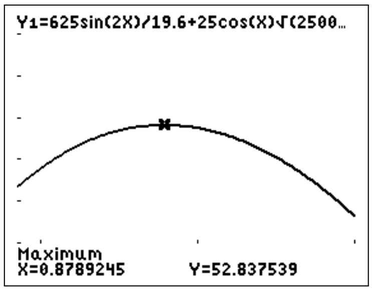
$$W = x(t_W) = \frac{25\cos(\theta)}{2g}\left(50\sin(\theta) - \sqrt{(2500\sin^2(\theta) - 80g)}\right),$$

and the distance beyond the wall as a function of launch angle θ is given by

$$R - W = \frac{625\sin(2\theta)}{g} - \frac{25\cos(\theta)}{2g}\left(50\sin(\theta) - \sqrt{(2500\sin^2(\theta) - 80g)}\right)$$

To find the maximum value of $R-W$, we plot $R-W$ as a function of θ (Radian mode) and use the calculator *maximum* operation (in the CALC menu).

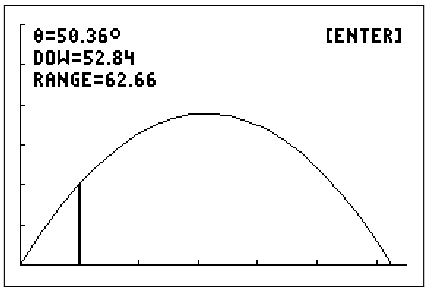




window $[\pi/4, 1, 0.1] \times [50, 55, 1]$

This gives the maximum value $R - W = 52.84$ m when $\theta \approx 0.8789$ rad or $\theta \approx 50.36°$.

In the main mission, we achieve a maximum distance beyond the wall of approximately 52.8 m when the launch angle is about 50.4°. This is confirmed by the PROJCE program, a screen of which is shown below.



window $[0, 65, 10] \times [0, 30, 5]$

Note that, at the optimum launch angle, the range is no longer a maximum, but the loss of range is more than compensated by being closer to the wall than if we launched the projectile at 45°, the angle for maximum range.

2.11 Parachuting

Based on the article *Minimal time of descent* by Jack Drucker, College Mathematics Journal 26: 232–235 (1995).

The Problem

You have graduated from the Academy, and in view of your excellent results in Mathematics Honours, you have been invited to join the Higher Mathematics Corps.

The Corps has been engaged to help out the elite Special Air Services Regiment, the SAS. In a big exercise coming up, the SAS plan to drop a spy behind enemy lines by parachute from an altitude of 400 m, the minimum height allowed by the local terrain. As the mission is crucial, it is important that the spy not be seen descending. She is much less visible during free fall than when floating down with her parachute open. Therefore, the SAS want to minimize the time the parachute will be open.

Specifically, they want to know

- what is the last possible moment the parachute can safely be opened?
- what is the minimum time required for the spy to descend safely?

The spy's mass m, including equipment, is 110 kg. Boots with special shock-absorbing insoles will let her withstand a maximum impact velocity of 10 m/s. The force of air resistance has been found experimentally to be proportional to velocity v, with proportionality constant $k \approx 19.96$ (so that $mg/k = 54$ m/s) during free fall and $k \approx 179.7$ (so that $mg/k = 6$ m/s) when the parachute is open. Take $g = 9.8\,\text{m/s}^2$.

A full report, understandable by officers of the SAS, is required. They like to see a few diagrams/graphs/pictures in reports too.

Some Preliminaries

Before we tackle the main problem, we need to revise a few facts about air resistance and falling bodies. The derivation and solution of the DEs below, although not required in this lab, are well within the capabilities of all students in the Maths 1 course.

- The total force acting on a falling body has two components, the gravitational force[31] $F_g = mg$, where m is the mass of the body, and the force due to air resistance F_r. The force due to air resistance is assumed to be proportional to velocity, i.e. $F_r = -kv$, where k is a positive constant. The coefficient of v in F_r is negative because air resistance acts in the opposite direction to velocity, i.e. in the $-v$ direction. Velocity v is a one-dimensional vector which can, in general, be positive or negative.
- From Newton's Second Law of Motion,

  $$m\,\frac{dv}{dt} \;=\; \text{sum of forces} \;=\; F_g + F_r \;=\; mg - kv.$$

  Dividing both sides by m ($\neq 0$) gives the DE

  $$\frac{dv}{dt} \;=\; g - \frac{kv}{m}.$$

[31] By choosing this force to be positive, we define the positive direction for velocity v and distance fallen x as downwards.

The equilibrium solution of this DE is the terminal velocity, which is the velocity of the body as $t \to \infty$ (a slope field will show you this). To obtain the equilibrium solution, we set the derivative equal to zero (the two forces are equal in magnitude, but opposite in direction), giving $g - kv/m = 0$. Solving for v gives the terminal velocity

$$v_T = \frac{mg}{k}.$$

The general solution to the DE

$$\frac{dv}{dt} = g - \frac{kv}{m},$$

is (separation of variables)

$$v(t) = v_T\big(1 - Ce^{-gt/v_T}\big),$$

where C is an arbitrary constant.

- To solve the parachute problem, you will also need a differential equation involving the velocity v and the distance fallen x. We have from the Chain Rule

$$\frac{dv}{dt} = \frac{dv}{dx}\frac{dx}{dt} = v\frac{dv}{dx},$$

so that, from our first DE,

$$v\frac{dv}{dx} = g - \frac{kv}{m}.$$

Dividing both sides by v ($\neq 0$) gives the DE

$$\frac{dv}{dx} = \frac{g}{v} - \frac{k}{m}.$$

The solution of this DE is also by separation of variables. It turns out that we can find the explicit general solution for x as a function of v, but not the other way around. The general solution is

$$x(v) = -\frac{v_T}{g}\left(v + v_T \ln\left|1 - \frac{v}{v_T}\right|\right) + D,$$

where D is an arbitrary constant.

Now solve on. Reading (and following) the hints is highly recommended. See if you can get the big picture from the graphs before starting on the details.

Hints for Solving the Problem

- Look at the sketch of velocity v against time t, covering the time between jump and landing (Figure 1 over the page).[32]
  - Put in what information you have, i.e. the points on the graph that you know. Let the chute open at time t_1, velocity v_1. Let the spy land at time t_2.
  - Write down the DEs and general solutions for each part of the graph.
  - Do you have enough information to find the particular solutions for each of the two parts of the graph? As you already have the general solution of the DE, it only remains to find the values for the constant C when the chute is closed and when it is open to give the particular solutions. You need to know one point on each curve to do this.
- Now look at the sketch of velocity v against distance fallen x (Figure 2).
  - Put in what information you have, i.e. the points on the graph that you know. Let the chute open at distance fallen x_1.
  - Write down the DEs and general solutions for each part of the graph.
  - Do you now have enough information to find the particular solutions, i.e. to find values for the constant D for each of the two parts of the graph?
- Having found particular solutions to each of the two parts of the v vs x graph, what are the unknowns now and how do you go about finding them?

Suggested Report Outline

- **Short introduction** — describe the problem.
- **Formulate the model** — write down the differential equations and solutions in the order you used them. Discuss any assumptions you make in the model. Include a filled-in copy of the sketches here.
- **Solutions of the problems**
  - the latest time to open the chute
  - the minimum time for descent
- **Conclusions and recommendations**

[32] Note that we only need to consider the *vertical* velocity and *vertical* distance fallen for this problem. For the sake of a simple picture, imagine she jumps from a hovering helicopter in conditions of no wind.

Parachute Lab

Please hand in one filled-in copy of these graphs with your group's report.

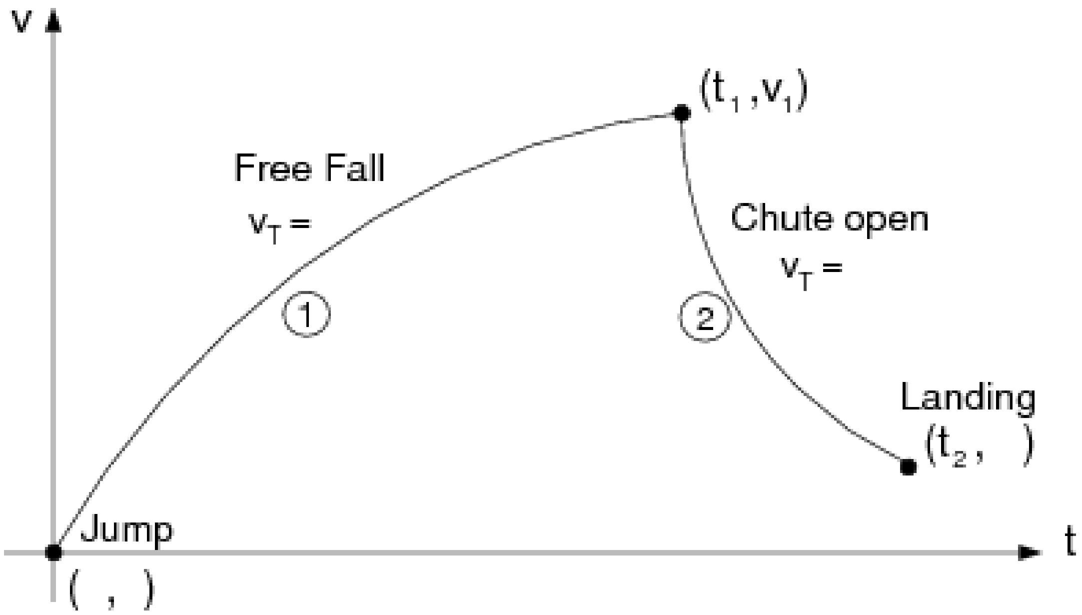


Figure 1: Velocity v as a function of time t.

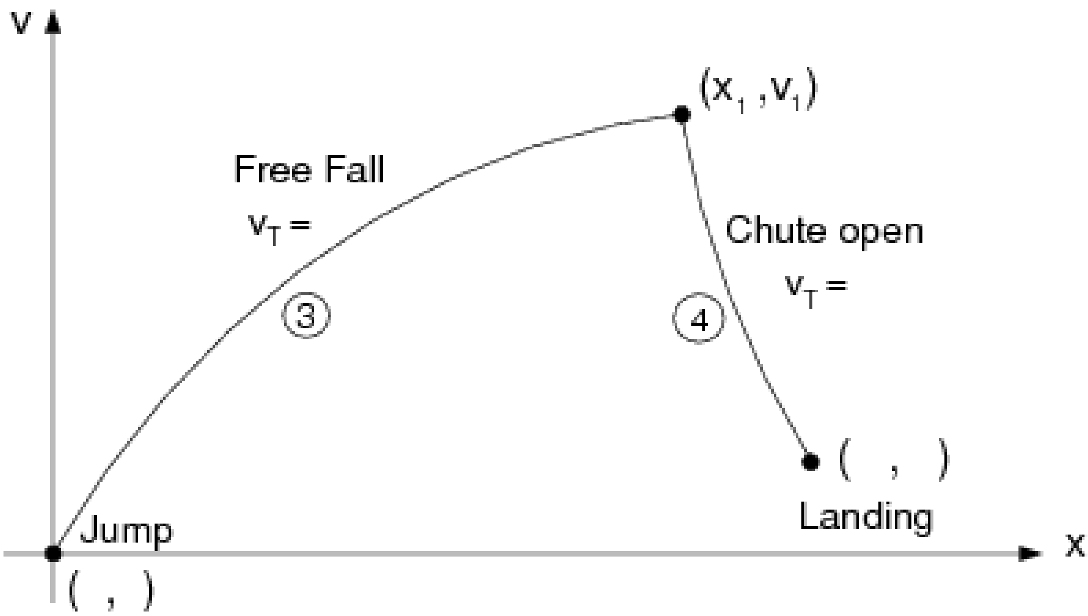


Figure 2: Velocity v as a function of distance fallen x.

Your calculator will plot distance fallen x as a function of velocity v. To make the graphs match, rotate the graph here 90° anti-clockwise and look at it from the reverse side.

Instructors' Guide

Solutions

Curve 1: The differential equation for this curve is

$$\frac{dv}{dt} = g - \frac{kv}{m},$$

with general solution

$$v(t) = v_T\big(1 - Ce^{-gt/v_T}\big),$$

where C is an arbitrary constant.

We have $v_T = 54$ m/s and initial condition $v(0) = 0$, giving $C = 1$ and the equation for Curve 1 is

$$v(t) = 54\big(1 - e^{-9.8t/54}\big).$$

Curve 2: The differential equation and general solution for this curve are the same as for Curve 1, but with $v_T = 6$ m/s. We don't know any points on Curve 2, so we can't find the arbitrary constant.

Therefore, we need to look for another method of solution, hence the need for the differential equation for velocity v as a function of distance fallen x.

Curve 3: The differential equation for this curve is

$$\frac{dv}{dx} = \frac{g}{v} - \frac{k}{m},$$

with general solution

$$x(v) = -\frac{v_T}{g}\left(v + v_T \ln\left|1 - \frac{v}{v_T}\right|\right) + D,$$

where D is an arbitrary constant.

We have $v_T = 54$ m/s and initial condition $v = 0$ when $x = 0$, giving $D = 0$ and the equation for Curve 3 as

$$x(v) = -\frac{54}{9.8}\left(v + 54 \ln\left|1 - \frac{v}{54}\right|\right).$$

Curve 4: The differential equation and general solution for this curve are the same as for Curve 3, but with $v_T = 6$,

$$x(v) = -\frac{6}{9.8}\left(v + 6 \ln\left|1 - \frac{v}{6}\right|\right) + D,$$

where D is an arbitrary constant.

We do know a point on Curve 4, the landing point, at which $x = 400$ and $v = 10$, the maximum velocity allowed for landing. Putting this into the general solution gives

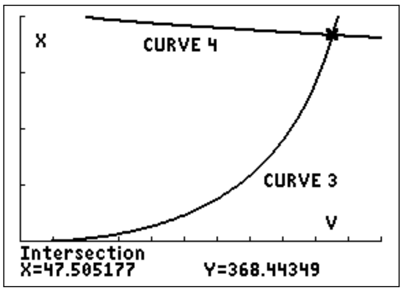
$$D = 400 + \frac{6}{9.8}\left(10 + 6 \ln\left|1 - \frac{10}{6}\right|\right) \approx 404.633.$$

Therefore, the equation for Curve 4 is

$$x(v) \approx 404.633 - \frac{6}{9.8}\left(v + 6\ln\left|1 - \frac{v}{6}\right|\right).$$

Our general strategy is now clear. The intersection of Curves 3 and 4 will give x_1 and v_1. Putting v_1 into the equation for Curve 1 will give t_1, the time at which the parachute should open. Putting t_1 and v_1 into the general solution for Curve 2 will give us the equation for Curve 2. We can then use the equation for Curve 2 to find the t_2, the time to landing, at which the velocity is 10 m/s.

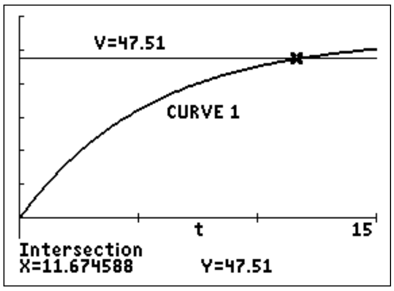
The intersection of Curves 3 and 4 is shown in the figure below. Students should make sure they understand this figure, perhaps following the graphs as time increases.



window $[0, 55, 5] \times [0, 400, 100]$

The point at which the parachute should open to give a landing velocity of 10 m/s is $x_1 \approx 368.44$ and $v_1 \approx 47.51$, that is she should fall about 368 m, at which position her velocity will be about 47.5 m/s.

Now we use the equation for Curve 1 to find the time t_1 at which $v = 47.51$. This can be done either algebraically or graphically to give $t_1 \approx 11.67$ s.



window $[0, 15, 5] \times [0, 60, 10]$

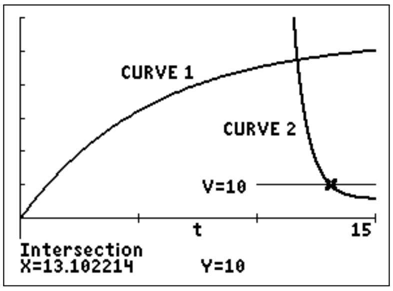
Now the point $(t_1, v_1) \approx (11.67, 47.51)$ lies on Curve 2, so we can find the arbitrary constant:

$$C \approx \left(1 - \frac{47.51}{6}\right) e^{9.8 \times 11.67/6} \approx -1.312 \times 10^9,$$

giving the equation for Curve 2 as

$$v(t) \approx 6\big(1 + 1.312 \times 10^9 e^{-9.8t/6}\big).$$

We use this equation to find the time to landing t_2, the time at which $v = 10$. Solving this algebraically or graphically gives $t_2 \approx 13.1$ s. The graphs of Curves 1 and 2 are shown below.



window $[0, 15, 5] \times [0, 60, 10]$

Therefore, the spy should open her chute after a time of 11.67 s, during which time she has fallen a distance 368 m and reached a velocity of 47.5 m/s. She should then land after 13.1 s at a velocity of 10 m/s.

The main assumption that would affect our results is that the parachute opens instantaneously. Clearly this is not the case, and clearly the margin for error in our problem is very small — she opens her chute 32 m above the ground! If we were to use this model, we would have to demand that the chute be fully open after 11.67 s, and work back from there to determine when she should open her chute. Alternatively, we could split the time between pulling the ripcord to open the chute and the chute being fully open into several time intervals; the value of k and hence the terminal velocity v_T would then decrease in several steps rather than instantaneously.

The other assumption which is not realistic is that the air resistance varies as the first power of velocity: a better assumption is that it varies as velocity squared, as discussed in L.N. Long and H. Weiss, *The velocity dependence of aerodynamic drag: a primer for mathematicians*, The American Mathematical Monthly, 106:127–135 (1999)[33]. Suitable values for the drag constant in this case are $k = 40$ kg/m with the parachute open and $k = 0.41$ kg/m with the parachute not open. These values can vary a bit, depending on the type of parachute and on the position of the parachutist during free fall.

The terminal velocity in this case is given by $v_T = \sqrt{\dfrac{mg}{k}}$, so that $v_T \approx 51.28$ with the parachute closed and $v_T \approx 5.191$ with the parachute open.

The integrals can still be done algebraically with a velocity-squared dependence of drag. We obtain the following equations for velocity as a function of time, and as a function of distance fallen.

$$v(t) = v_T \left(\frac{1 - e^{2kv_T t/m}}{1 + e^{2kv_T t/m}} \right) \quad \text{chute closed; } v(0)=0;\ k=0.41;\ v_T=51.28 > v.$$

$$v(t) = v_T \left(\frac{1+Fe^{-2kv_T t/m}}{1-Fe^{-2kv_T t/m}} \right) \quad \text{chute open; } F \text{ a constant; } k=40;\ v_T=5.191 < v.$$

$$v(x) = v_T\left(1-e^{-2kx/m}\right) \quad \text{chute closed; } v(0)=0;\ k=0.41;\ v_T=51.28 > v.$$

$$v^2(x) = v_T^2 + \left(100-v_T^2\right)e^{-2k(x-400)/m.} \quad \text{chute open; } v(400)=10;\ k=40;\ v_T=5.191 < v.$$

Note that we have to be careful how we express $v(x)$ for the last equation, otherwise we run into numerical problems with very large exponentials.

[33] available online

We solve the last two equations graphically to find the position and velocity at which the parachute opens (instantaneously). Substituting this velocity into the first equation tells us the time at which the parachute opened, and this allows us to determine F in the second equation. We are then able to answer the questions.

In this case, the spy should open her chute after a time of 11.84 s, during which time she has fallen a distance 395 m and reached a velocity of 50.2 m/s. She should then land after 12.9 s at a velocity of 10 m/s. A brave person!

2.12 Trial of the Session

Modified from an idea by Joel S. Foisy, SUNY Potsdam.

Aims

- To see a practical application of a mathematical model involving a differential equation that you are now capable of solving.
- To discover one way that unknown physical parameters in mathematical models may be determined experimentally.

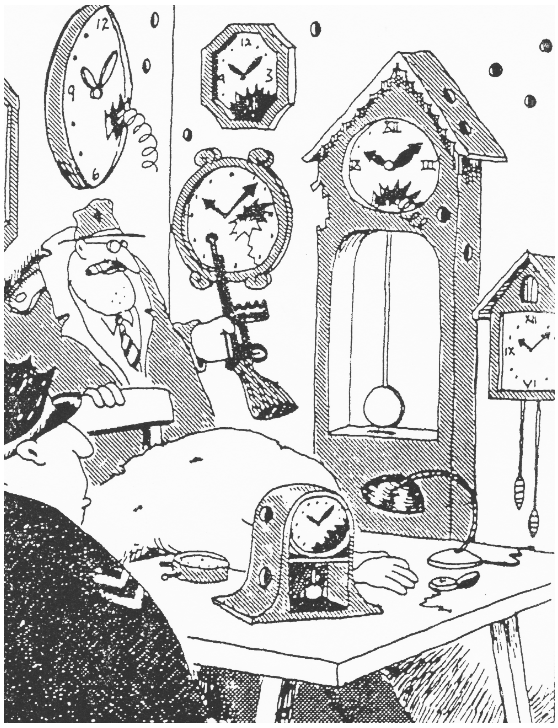
Newton's Law of Cooling
The rate of change in temperature of a cooling body is proportional to the difference between the temperature of the body and the surrounding temperature.

Scenario

Officer Cadet (OCDT) I.M. Innocent has been accused of the murder of Prof U.R. Knott.

The body of Prof Knott was found in the library at 0855 on the morning of Tuesday, 11 August. The cause of death was determined to be a bayonet stab wound to the heart. Death was adjudged to be instantaneous. The bayonet was found at the scene and had OCDT Innocent's fingerprints on it. As this bayonet was known to have been allocated to OCDT Innocent, this is not surprising. Needless to say, OCDT Innocent is the prime suspect. His alibi that he was in a tutorial can be verified, but only from 0800 to 0850.

The purpose of this trial is to determine OCDT Innocent's innocence or guilt using Newton's Law of Cooling. In other words, you will attempt to prove whether or not the crime was committed outside the time interval 0800 – 0850.



We've got the murder weapon and the motive . . . now if we can just establish time of death.
Gary Larson

Additional Facts Turned Up by Police Investigations

- The body temperature of the professor was 35.0°C at 0920. The body temperature was recorded again at 0928 and found to be 34.8°C. The temperature in the library was determined to be a constant 23.0°C.
- OCDT Innocent's tutorial was in a room 200 metres from the library.
- Three days before the murder, the professor visited a doctor, at which time the professor's body temperature was recorded to be exactly 38.7°C. Normal body temperature is 37.0°C.
- On the morning of the murder before coming to work, the professor wrote the following entry in his diary:

  > I have been feeling terrible ever since I had that run in with that horrible first-year student over his failing mid-year grade. What was his name? Innocent? Not likely. And then for the past few days I've had that terrible fever. I finally feel better today, not 100%, but nearly back to normal.

- All DNA evidence has been thrown out due to multiple lab errors.

Requirements for this Lab

- Write a brief for an expert witness explaining Newton's Law of Cooling, including assumptions that need to be made for its application and how the resulting equations may be solved. The body-temperature-versus-time curve for the professor's body should also be derived, with the estimated temperature at key times calculated. Think about the best way to present your results.

  Look (or at least *think*) before you leap!

  - What's a good time to choose for $t = 0$?
  - Your solution to the differential equation will initially have two unknown parameters in it. How many auxillary conditions will be needed in order to determine these two parameters? Where will you get them from?

- You are also required to produce a case for either the defence or the prosecution (the lab instructor will tell you which), based around Newton's Law of Cooling and the estimated time of death. Additional arguments consistent with the above evidence are also allowed.

  Any statements supporting (or otherwise) OCDT Innocent, in the event he is convicted, will be taken into account when passing sentence.

The Trial

For the purposes of the trial, one lab group will be asked to act as the prosecution team and another group as the defence team. Volunteers will also be needed for the accused and a bailiff (who will be provided with a script). The remaining students will be asked to act as the jury. One of the lab instructors will act as judge. Mini Mars Bars will be awarded to the team which comes up with either a successful defence or a successful prosecution.

Trial Procedure

- The prosecution team will present their case first. Their first witness will be an "expert witness" who will explain Newton's Law of Cooling; how it can be used and the resulting equations solved; any approximations used in solving the equations in this case; and the estimated time of death. The expert witness needs to be able to explain things so that a jury of non-mathematicians can understand what's going on. Other suitable witnesses may be concocted by the prosecution. The prosecuting attorney sums up the case for the prosecution.
- The defence proceeds with their case. They too will need an expert witness to explain how the mathematics and its assumptions support their case. Further arguments and/or witnesses can be concocted.
- The prosecution rebuts the defence's arguments (they will therefore need to have anticipated these) and makes a closing statement to the jury.
- The defence rebuts the prosecution's rebuttal and makes a closing statement to the jury.
- The jury is asked to consider the evidence and vote on a verdict.

Instructors' Guide

Mathematical working

Let $T(t)$ be the temperature of the body in degrees Centigrade at time t min after 0920.

Other choices for the time origin such as 0800 or 0850 make the analysis of the results easier; the choice here simplifies the working out a little by having one of the given temperatures be at $t=0$. Try both and see which suits your students better.

The differential equation for T is

$$\frac{dT}{dt} = -k(T-T_a),$$

where k is a constant and T_a is the ambient (room) temperature.

We are given

- T_a is constant at 23.0°C.
- $T(0) = 35.0$°C.
- $T(8) = 34.8$°C.
- $T_{\text{normal}} = 37.0$°C.
- $T_{\text{fever}} = 38.7$°C.
- OCDT Innocent has an alibi for 0800 − 0850, that is $-80 \leqslant t \leqslant -30$.
- The body was found at 0855.

The general solution to the differential equation — standard separation of variables or let $U = T-T_a$ and write down the exponential solution for U — is

$$T(t) = 23 + Ee^{-kt},$$

where E is a constant. We use the two given temperatures at different times to determine E and k.

$$T(0) = 35.0 \Rightarrow 35 = 23 + E \Rightarrow E = 12.$$

Therefore

$$T(t) = 23 + 12e^{-kt}.$$

$$T(8) = 34.8 \Rightarrow 34.8 = 23 + 12e^{-8k} \Rightarrow k = -\frac{1}{8}\ln\left(\frac{34.8-23}{12}\right) \approx 0.00210.$$

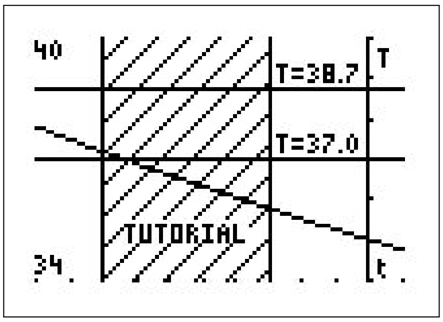
Therefore

$$T(t) \approx 23 + 12e^{-0.0021t}.$$

The question is could the murder have been committed before 0800, i.e. at a time for which OCDT Innocent does not have an alibi.

We can work out the time at which the temperature of the body was last normal, 37.0°C, the presumed time of death if the professor's temperature was indeed normal and not still above normal because of his fever. The time at which the temperature was last normal (found either graphically or algebraically) is $t \approx -73.4$ or at about 0806. At this time OCDT Innocent was in the tutorial and does therefore have an alibi.

If the professor's temperature was above normal, the picture is not so clear. The figure below shows a plot of body temperature against time, with the critical temperatures (normal and fever) and tutorial times marked (shaded region).



The figure shows that OCDT Innocent could have committed the murder if the professor's temperature was above normal. Tracing along the graph shows that if the professor's temperature was above about 37.3°C, OCDT Innocent could have committed the murder and still had between 3 and 4 minutes to get to his tutorial at 0800.

The rest is up to the prosecution, the defence and the jury, with plenty of scope for drama and invention.

If we consider that an error might have been made in measuring body temperature after the murder, the situation becomes more confused. For example, if the actual body temperature had been 34.7°C at 0928, an error of only 0.1°C, the time at which the body was last at normal temperature would be about 0831, right in the middle of the tutorial.

If the actual body temperature had been 34.9°C at 0928, an error of 0.1°C in the other direction, the time at which the body was last at normal temperature would be about 0653! Clearly the temperature measurements are crucial, particularly because the time between the measurements is so short. There is more to explore in this problem for the good student and plenty of scope for argument in the trial.

2.13 Some Like It Hot

In this lab, we solve some differential equations numerically using the calculator. Your job is to tell it what to calculate and to interpret its output.

Question 1 *Getting it hot*

A potato placed in an oven heats up at a rate proportional to the difference between its temperature and the oven temperature T_{oven} (Newton's Law of Heating).

(a) Explain how Newton's Law of Heating gives the differential equation for the temperature of the potato $T(t)$ at time t,

$$\frac{dT}{dt} = k(T - T_{\text{oven}}), \quad \text{where } k \text{ is a constant.}$$

(b) If $T_{\text{oven}} = 200°\text{C}$ (constant) and the potato is heating up at a rate of 2° per minute when its temperature is 120°C, show that $k = -0.025$ per minute.

Euler's Method and the *modified Euler's Method* are numerical methods to approximate solutions of a first-order differential equation.[34] In both methods, we have to specify an initial point (t_0, T_0) on the graph of temperature T versus time t and a step length H, which controls the accuracy of the result: the smaller H, the more accurate the result,[35] but the more calculations have to be performed, so the overall calculation takes longer.

Euler's Method is the basis of the EULER1 program, the modified Euler's Method the basis of the MODEULR1 program. Both programs graph an approximation to the solution of the differential equation and give the final values calculated. Instructions on how to use the programs are given over the page.

(c) The initial ($t_0 = 0$) temperature T_0 of the potato is 20°C. Use EULER1 with a step length H = 1 (minute) to estimate the temperature of the potato after 20 minutes. *Read the program instructions over the page before you do this. Answer*: 91.5°C.

(d) Now use EULER1 to find the temperature after 20 minutes, *accurate* to 3 significant digits. We can be reasonably certain that a result is accurate to n significant digits if, when we reduce H by a factor of 10, two successive values for T are the same when rounded to n digits. Document your procedure so someone else could repeat your calculations on their calculator. A table of values of H and the corresponding calculated temperatures is a good idea.

(e) Repeat (d) using MODEULR1. Which program would you prefer to use? Why?

(f) The exact general solution to the differential equation is $T(t) = 200 + Ee^{-0.025t}$, where E is an arbitrary constant.

(i) *Show that* this function $T(t)$ does satisfy the DE, i.e. show LHS of the DE = RHS *with this function.*

(ii) If the potato starts at a temperature of 20°C, show that the constant $E = -180$.

[34] The basis of both methods is approximating a curve by a tangent.

[35] With very small values of H, of the order of 10^{-10} or less, the accuracy of the result is limited by round-off error: the calculator cannot calculate the quantities in the calculations accurately enough and may therefore give meaningless answers. It will also take a very very long time to do the calculations.

(iii) Compare the exact values[36] of T with your results in (d) and (e). Were the numerical methods in (d) and (e) accurate to the number of significant digits claimed?

Question 2 *If it's not hot . . .*

The chef realises at 6:30pm that he has forgotten to preheat the oven to cook a potato to be served with the boss's dinner at 7:30. If it isn't done by 7:30, the chef will be fired. He puts the potato in the oven and turns the oven on. The oven temperature t minutes after it is turned on is

$$T_{\text{oven}}(t) = 200 - 175e^{-0.2t}\,°\text{C}.$$

The potato heats according to Newton's Law of Heating, that is the rate of change in the temperature $T(t)$ of the potato is given by

$$\frac{dT}{dt} = -0.025\big(T - T_{\text{oven}}\big).$$

The potato is done when it reaches a temperature of 160°C. The potato starts off at the same temperature as the oven, $T_{\text{oven}}(0)$, in a kitchen that's warmed up a bit since Question 1.

- What is the chef's fate? A full explanation is required, including estimates of the accuracy of your results.
- Would it have made any difference if the chef had remembered to preheat the oven to 200°?

Bonus question *A different scenario . . .*

How else could a heating or cooling problem like this arise? Write a scenario in which a similar problem could be set. Illustrations always welcome.

Solving first-order differential equations numerically

The EULER1 and MODEULR1 programs calculate and plot *approximate* solution curves to the first-order differential equation $\frac{dY}{dX} = f(X, Y)$.[37]

Using EULER1

- First write out your DE using X and Y as the variables. **The independent variable must be X and the dependent variable Y**.
- Type the derivative function f(X,Y) into Y1. *Hint*: There should be a Y in your expression in Y1 for the current problem.
- Set the [window] variables to appropriate values. **Xmin must be the initial X value.** If you have to estimate Y at some value of X, set Xmax to that X value.
- Run the program.

[36] actually decimal approximations generated by your calculator from the function

[37] See *Differential Equations* in Volume 2 of *Mathematics on a TI-84/CE* (at *www.XXX*) for more details.

- At the prompts, input a starting value of Y (corresponding to Xmin) and the step length H. You can move the cursor around the graph once it is plotted, but you can't trace the graph.
- Press enter to display a new menu. Press 3 to display the coordinates of the last point plotted, and enter to return to the menu to input new starting values. This allows you to plot a different solution curve, use a different step length or quit.
- The MODEULR1 program is a bit fancier, but works in essentially the same way as EULER1. Choose a TIME PLOT in the first menu.

Supplementary Question

Question 3 *Some like it hot — modelling from scratch*

You like your coffee white and hot. You have just poured a cup when the phone rings. Your friend needs to talk to you for 10 minutes.

(a) Should you add the milk before you go off or should you add it when you come back, so as to have your coffee as hot as possible?

- Use the DE for Newton's Law of Cooling

$$\frac{dT}{dt} = k(T - T_{\text{room}}),$$

where k is a constant and T_{room} is room temperature, assumed here to be constant over the 10 minutes. The general solution to this DE is $T(t) = T_{\text{room}} + Ee^{kt}$, where E is a constant.

- Make any reasonable assumptions you need to answer this question. For example, you will need to make up some numbers and you will have to decide what happens to the temperature of the coffee when you add milk. Don't forget to write all this in your report.

(b) Does it make much difference when you add the milk? Quantify your answer.

Instructors' Guide

This lab is a preliminary to learning about Euler's Method. Students are told, step by step, how to generate an Euler solution to an initial-value problem and are asked to relate the Euler results to the exact result. The modified Euler's Method is also introduced as a faster way of obtaining an approximate solution. Later in the course, we look at Euler's Method in detail to find out exactly what the calculator is doing.

The programs EULER1 and MODEULR1 are available at *canberramaths.org.au* under *Resources.*

Solutions

1. (a) The rate of heating is dT/dt, and the difference in temperature between the body and the oven is $T - T_{\text{oven}}$. Therefore,

$$\frac{dT}{dt} \propto T - T_{\text{oven}} \quad \Rightarrow \quad \frac{dT}{dt} = k(T - T_{\text{oven}}),$$

where k is a constant of proportionality.

(b) We have

$$2 = k(120-200) \ \Rightarrow \ k = \frac{2}{120-200} = -0.025.$$

(c) A table showing results for (c), (d) and (e) is given below. We use a window of $[0, 20, 5] \times [0, 200, 50]$, although any Y window will do as we are only interested in the final values. The approximations from the two methods are shown for different values of the step length H min. Temperatures are rounded to 3 significant digits.

H	**$T(20)$ EULER1**	**$T(20)$ MODEULR1**
1	91.5	90.8
0.1	90.9	90.8
0.01	90.8	–
0.001	90.8	–

The EULER1 estimate for $T(20)$, the temperature of the potato after 20 minutes, using a step length $\text{H} = 1$ minute is 91.5°C.

(d) The last two results for EULER1 in the table are the same to 3 significant digits, so that $T(20) = 90.8$°C, accurate to 3 significant digits. The last result in particular takes a long time.

(e) Using MODEULR1, we find the same result as in (d), but after only two runs. This is clearly the preferable method, because we achieve the same accuracy as EULER1 but with larger values of H. Consequently the time taken to achieve an accurate result is much less.

(f) (i) With the given function for T, we have

$$\begin{aligned}\text{LHS} &= \frac{dT}{dt} \\ &= -0.025Ee^{-0.025t}. \\ \text{RHS} &= -0.025(T-200) \\ &= -0.025(200+Ee^{-0.025t}-200) \\ &= -0.025Ee^{-0.025t} \\ &= \text{LHS}.\end{aligned}$$

Therefore, $T(t)=200+Ee^{-0.025t}$ is a solution to the differential equation.

(ii) $T(0)=20 \;\Rightarrow\; 20=200+Ee^0 \;\Rightarrow\; E=-180.$

(iii) The temperature after 20 minutes is $T(20)=200-180e^{-0.5}=90.8$, rounded to 3 significant digits. Our numerical approximations were indeed accurate to 3 significant digits.

2. The differential equation for the temperature of the potato is

$$\frac{dT}{dt} = -0.025(T-200+175e^{0.2t}),$$

with $T(0)=25$, the initial temperature of the oven. We need to find out if $T(60)\geqslant 160$. If so, the chef's job is safe. Using EULER1 or MODEULR1 (preferable), we obtain the following.

	***T*(60)**	
H	**EULER1**	**MODEULR1**
1	155.6	155.4
0.1	155.4	155.4

The temperature of the potato at 7.30pm is clearly less than 160°C, so the chef loses his job.

Had he remembered to preheat the oven to 200°C, the differential equation would have been the same as in Question 1, but with $T(0)=25$. In this case, the temperature of the potato would have reached 160°C, and the chef would still be working there.

PTO

3. We work with the general solution

$$T(t) = T_{\text{room}} + Ee^{kt},$$

where E is a constant that will be determined by the initial temperature. Assume that room temperature is 20°C, that the initial temperature of the coffee is 90°C and that the cup has been pre-heated to the same temperature to avoid a sudden initial cooling of the coffee.

(a) (i) *Milk added after 10 minutes*

In order to determine the constant k, we need to know/assume a temperature of the coffee at some time other than $t = 0$. In this theoretical treatment, we assume that, without milk added, the coffee cools from 90°C to 50°C in 10 minutes.

When we add the milk, the temperature of the coffee + milk will be different from that of the coffee alone. A reasonable assumption is that the new temperature will be the mean temperature of the coffee and the milk, weighted by volume. If we assume that there is 160 ml of coffee, 40 ml of milk and that the milk temperature is 5°C, the final temperature is

$$T_F = \frac{160 \times 50 + 40 \times 5}{200} = 41°\text{C}.$$

(ii) *Milk added immediately*

Here, we need to determine a value for k: this comes from the assumption that the coffee alone cools from 90°C to 50°C in 10 minutes. With $T(0) = 90$, we find $E = 70$ and

$$T(t) = 20 + 70e^{kt}.$$

Putting $T(10) = 50$ gives

$$\begin{aligned} 50 &= 20 + 70e^{10k}. \\ \therefore\ k &= \frac{1}{10} \ln\left(\frac{30}{70}\right) \approx -0.085. \end{aligned}$$

When we add the milk immediately, the initial temperature of the white coffee is again the weighted mean of the two temperatures,

$$T(0) = \frac{160 \times 90 + 40 \times 5}{200} = 73°\text{C}.$$

This is the initial temperature for the solution of the differential equation. Substituting $T = 73$ when $t = 0$ gives $E = 53$, so that

$$T(t) = 20 + 53e^{kt},$$

giving

$$T_F = T(10) = 20 + 53e^{10k} \approx 42.7°\text{C}.$$

With these parameters, adding the milk immediately is the better option, giving a coffee temperature after 10 minutes of 42.7°C, compared with a temperature of 41°C if we add the milk after 10 minutes.

Physically this makes sense, because a hotter body, the coffee without milk, loses heat faster than a cooler body, the coffee with milk.

It's an interesting modelling project to write the above equations in a more general form, investigate what happens when we vary the parameters, including the time, and to try to prove algebraically our assertion.

(b) *Does it make much difference when you add the milk?*

With the parameters assumed above, the difference is not great; an exploration of the problem as suggested above should show that this is mostly the case. One way is better than the other, but you probably wouldn't worry too much about which to choose unless you like large amounts of milk in your coffee.

2.14 Warm and Wealthy

The Problem

To use the lesser amount of energy, should you

A. turn off the heating in your house at night and have it come on again sometime early in the morning, or

B. leave the heating on all night?

How much more efficient is the more efficient method in percentage terms?

Some Assumptions

- The temperature in the house is 20°C at 2200 hours.
  - Using Scheme A, you turn off the heating at this time. The heating should then be switched on (by a suitably set time clock) so as to return the temperature to 20°C by 0700 hours the next morning.
- The heater is either on or off: its output cannot be varied. The heating is such that, *if there were no heat losses* (there are), it would raise the temperature of the house by 3°C per hour.[38]
- When the heating is on, it is controlled by a thermostat which switches on when the temperature drops to 18°C and off when the temperature reaches 20°C.

The Method

Use a mathematical model to help answer this question.[39]

- Sketch a graph of temperature versus time for each case. How are you going to find the equations of the curves? What points on the graph do you know?
- What further assumptions do you have to make to set up the model? Keep them as simple as possible for the first model. Think here about how you will write down a differential equation.
- For which constants do you need to know values?

Write these assumptions and constants down, then ask one of the lab staff to give you the values. If you can do this, you get BONUS marks. If you haven't worked this out after 20 minutes, ask for help.

[38] When building this condition into the DE that applies when the heater is on, think about units and the idea that the net rate of change of temperature is a balance between heat gains and heat losses.

[39] Write up your report so that a mathematically literate colleague not familiar with the problem could follow it. It should contain details of your calculations, together with any necessary explanations and conclusions; answers without explanations are not acceptable.

Further Assumptions

To give out to the groups after they have thought about the problem.

- Assume Newton's Law of Cooling: the *rate of change* of the house temperature is proportional to the difference between the house temperature and the outside temperature.
- The constant of proportionality $k = 0.125$°C per hour per degree of temperature difference. Think carefully about its sign.
- The outside temperature is a constant 5°C between 2200 and 0700.

Supplementary Questions

1. Would better insulation make a difference to your answer above? What sort of energy gains do you get with more insulation? A similar analysis to that in the original problem is expected here.

2. Would a bigger furnace change your answer? Do you use more or less energy overall by having a bigger furnace? Use the same k value as in the original question.

3. The outside temperature varies with time. Assume a sinusoidal daily outside temperature varying between a minimum of 0°C (at 0400) and a maximum of 10°C (at 1600). Write down the sine function that gives this outside-temperature variation. Then write down the differential equation that describes cooling in this case.[40] Use EULER1 or MODEULR1 to compare the solution from this DE to the cooling solution in the original problem from 2200 hours on. Explain the difference.

4. In some heating systems, the thermostat allows the inside temperature to drop by 5°C overnight (that is cycle between 13°C and 15°C), bringing the temperature back up to 20°C by the required time in the morning. Explore this scheme.

[40] The DE is not variables separable, but can be solved algebraically by another method.

Instructors' Guide

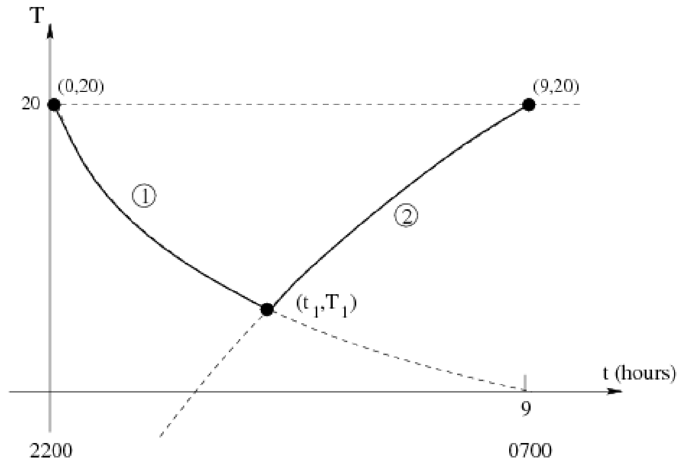
The EULER1 and MODEULR1 programs are available at *canberramaths.org.au* under *Resources*. Solving variables-separable differential equations is assumed knowledge.

Solutions

Let $T(t)$ be the temperature of the house in degrees Centigrade at time t hours after 2200.

Scheme A

The general scheme is shown in the figure below. The heating is turned off at 2200 ($t=0$) and switched back on at some time $t=t_1$ so that the temperature is back to 20°C by 0700 ($t=9$) the next morning. We have to find t_1: the heater will then be on for a total of $9-t_1$ hours.



The differential equation for cooling, Curve 1, is

$$\frac{dT}{dt} = -0.125(T-5),$$

using the given value for the constant k, a $-$ sign to give cooling and an outside temperature of 5°C. The general solution (separation of variables) is

$$T(t) = 5 + E_1 e^{-0.125t},$$

where E_1 is an arbitrary constant.

$T(0)=20 \Rightarrow E_1=15$, so that the equation for Curve 1 is

$$T(t) = 5 + 15e^{-0.125t}.$$

The differential equation for heating, Curve 2, is

$$\frac{dT}{dt} = 3 - 0.125(T-5) = -0.125(T-29).$$

The units of the left-hand side of this equation are °C per hour, showing that the heating term comes in simply as 3 on the right-hand side.

The general solution is

$$T(t) = 29 + E_2 e^{-0.125t},$$

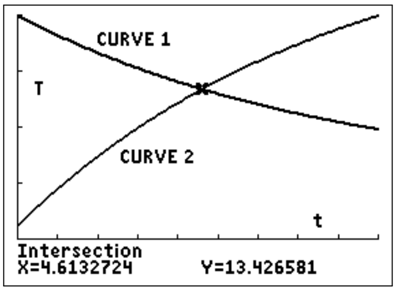
where E_2 is an arbitrary constant.

A point on Curve 2 is the final point $T = 20$ when $t = 9$ (0700 hours).

$T(9) = 20 \;\Rightarrow\; E_2 = -9e^{9/8}$, so that the equation for Curve 2 is

$$T(t) \;=\; 29 - 9e^{9/8}e^{-0.125t} \;=\; 29 - 9e^{-0.125(t-9)}.$$

The graphs of the two functions are shown below, together with their point of intersection, which determines t_1, the time at which the heater is switched on in the night.

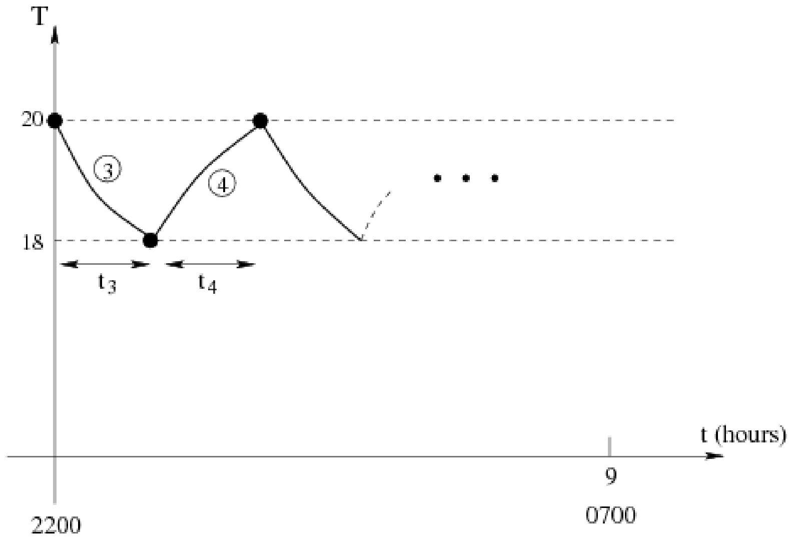


window $[0, 9, 1] \times [0, 20, 5]$

We find from the graph or algebraically that $t_1 = 4.61$, accurate to 3 significant digits (the heater is turned on at 0237 hours). For Scheme A, the heater is therefore on for $9 - 4.61 = 4.39$ hr.

Scheme B

Here the heater cycles on and off. One such cycle consisting of a cooling phase, Curve 3, and a heating phase, Curve 4, is shown below. We first need to find the total time, $t_3 + t_4$, for one cycle.



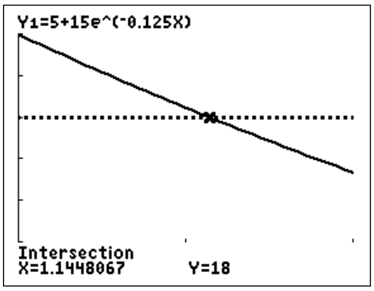
The equation for Curve 3 is the same as for Curve 1:

$$T(t) \;=\; 5 + 15e^{-0.125t}.$$

The time t_3 for the temperature to cool to 18°C is found by solving

$$18 = 5 + 15e^{-0.125t}$$

for t, giving (graphically or algebraically) $t_3 = 1.14$ hr, accurate to 3 significant digits.



window $[0, 2, 1] \times [15, 20, 1]$

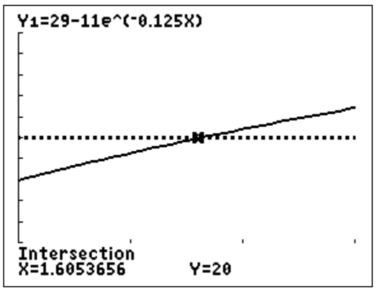
We could use the point $(t_3, 18)$ as the initial point on Curve 4, but, as we only need to find t_4, it is easier mathematically to start Curve 4 at $(0, 18)$. Following the same steps as for Curve 2, we find that the equation for Curve 4 starting at $t = 0$ is

$$T(t) = 29 - 11e^{-0.125t}.$$

The time t_4 taken for the temperature to reach 20°C is found by solving

$$20 = 29 - 11e^{-0.125t}$$

for t, giving (graphically or algebraically) $t_4 = 1.61$ hr, accurate to 3 significant digits.



window $[0, 3, 1] \times [15, 25, 1]$

The heating cycle in Scheme B is therefore 1.14 hours off, 1.61 hours on, for a total cycle time of 2.75 hours. Therefore, between 2200 and 0700, we go through 3 complete cycles (8.25 hr), plus $9 - 8.25 = 0.75$ hr. In this 0.75 hours, the heater is switched off, because the temperature has reached 20°C at the end of the last cycle and it takes 1.14 hours before the heater needs to switch on again.

Therefore, for Scheme B, the heating is on for 3 cycles, i.e. for $3 \times 1.61 = 4.83$ hours.

In Scheme A, the heating was on for 4.39 hours, so that Scheme A uses about 0.44 hours (26 minutes) less heating time than Scheme B or about 9% less energy. The reason for this is that the greater the house temperature, the greater the heat loss. Scheme A, with an overall average temperature less than that of Scheme B, results in less heat loss throughout the 9 hours.

If we were to make the comparison even fairer, we would observe that the temperature at 0700 in Scheme B is in fact less than 20°C, because the house has been cooling for 0.75 hours after reaching 20°C at the end of the second cycle. The temperature would then be $5 + 15e^{-0.125\times0.75} = 18.7$°C, accurate to 3 significant digits.

If we require from Scheme A that it only reach a temperature of 18.7°C at 0700, we find that the heater must switch on at $t_1 = 5.33$ hr, so that the heater is on for a total of 3.67 hr. This represents a saving of 24% over Scheme B.

Supplementary Questions

Some experimentation with different values of the appropriate parameter is required here, rather than general conclusions.

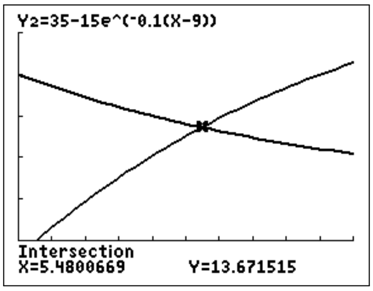
1. Greater insulation means a lesser rate of energy loss for a given temperature difference between inside and outside. Mathematically, this means that the constant k is reduced in magnitude. Let's take $k = 0.1$, rather than the value of 0.125 used above.

Scheme A

Following similar steps to those of the original problem, we find the equations for Curves 1 and 2 are

$$T(t) = 5 + 15e^{-0.1t} \qquad T(t) = 35 - 15e^{-0.1(t-9)},$$

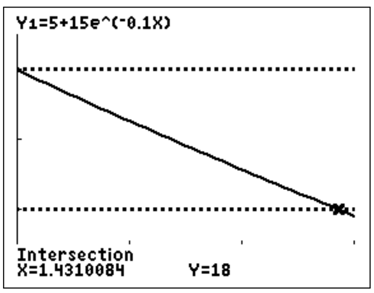
which intersect at $t_1 = 5.48$.

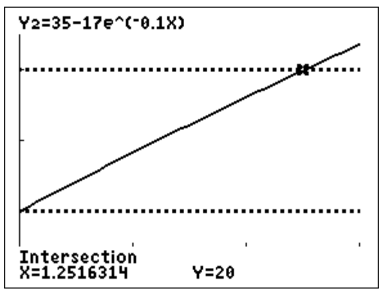


The heater is therefore on for $9 - 5.48 = 3.52$ hr. This is about 20% less than in the original question: here, a 20% decrease in k gives a 20% decrease in the amount of energy used in this case.

Scheme B

The equation for Curve 3 is $T(t) = 5 + 15e^{-0.1t}$, giving a time to cool from 20° to 18° as $t_3 = 1.43$. This is longer than the 1.14 hr of the original scheme, as we would expect with more insulation.





The equation for Curve 4 starting at $(0,18)$ is $T(t)=35-17e^{-0.1t}$, giving a time to heat from 18° to 20° as $t_4=1.25$.

The total cycle time, t_3+t_4, is 2.68 hours, giving 3.36 cycles in the 9 hours. In the last 0.36 of a cycle, the room is cooling, so that the heater is on for 3 cycles, i.e. for $3 \times 1.25 = 3.75$ hours.

Scheme A, with the heater on for 3.52 hours, is still better than Scheme B, with the heater on for 3.75 hours. Scheme A uses about 6% less energy than Scheme B, less than the value in the original question because of the greater insulation (smaller k). In the limiting case $k=0$ (perfect insulation), the two schemes would give the same answer — no heating required.

2. Assume a heater of say twice the capacity of the original heater and therefore capable of raising the temperature of the house by 6° per hour in the absence of heat losses. With the original value of k, $k=0.125$, the differential equation for the heating phase is then

$$\frac{dT}{dt} = 6 - 0.125(T-5) = -0.125(T-53),$$

with general solution

$$T(t) = 53 + E_2 e^{-0.125t}.$$

The differential equation for cooling is the same as in the original question. The equations for Curves 1 and 2 are

$$T(t) = 5 + 15e^{-0.125t} \qquad T(t) = 53 - 33e^{-0.125(t-9)},$$

with intersection point $t_1 = 7.10$, so the heater in Scheme A is on for $9-1.90 = 1.90$ hours.

Note that although the heater here is twice as powerful as the one in the original question, the total energy used is actually less: the heater in the original question was on for 4.39 hours, more than twice the time here.

The equation for Curve 3 is the same as for Curve 1, giving the time to cool from 20° to 18° as $t_3=1.14$. The equation for Curve 4 is

$$T(t) = 53 - 35e^{-0.125t},$$

giving the time to heat from 18° to 20° as $t_4 = 0.47$. The total cycle time is therefore 1.61 hours, so that there are 5.59 cycles betweeen 2200 hours and 0700 hours. Again the house is cooling in the 0.59 hours, so the heater in Scheme B is on for $5 \times 0.47 = 2.35$ hours. Scheme A is about 19% better than Scheme B in this case.

3. Here we start with a general sine curve for the outside temperature,

$$T_{\text{out}} = A \sin\big(B(t-C)\big) + D,$$

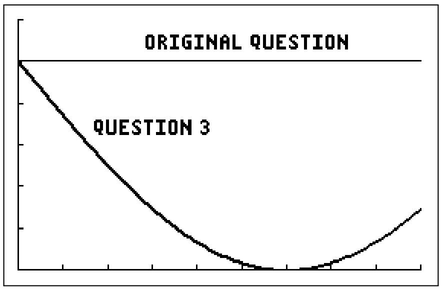
where we have to choose values for A, B, C and D so that the curve has a minimum of 0 at $t=6$ (0400 hours) and a maximum of 10 at $t=18$ (1600 hours).

For the curve to oscillate between 0 and 10, it must have a mean value $D=5$ and an amplitude $A=5$. The period $2\pi/B$ is 24 hours, so $B=2\pi/24=\pi/12$. The curve attains its mean value of 5 halfway between the minimum at $t=6$ and the maximum at $t=18$, i.e. at $t=12$. $t-C=0$ at this value, so that $C=12$.

The outside temperature is therefore given by

$$T_{\text{out}} = 5\sin\left(\frac{\pi(t-12)}{12}\right)+5.$$

The outside temperature for the original question (a constant 5°) and for the current question is shown below for $0<t<9$.



window $[0,9,1]\times[0,6,1]$

To find the temperature inside the house at any given time, we now have to solve

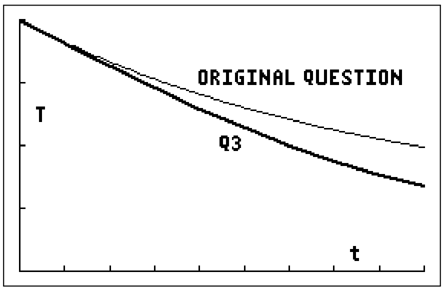
$$\frac{dT}{dt} = -0.125(T-T_{\text{out}}),$$

with the initial condition $T(0)=20$. This cannot be done using the method of separation of variables (it can be solved algebraically using an integrating factor), so we turn to a numerical method, either Euler's method or the modified Euler's method.

To set this up for the EULER1 or MODEULR1 programs:

- put $\text{Y}_1=-0.125\Big(\text{Y}-5\sin\big(\pi(\text{X}-12)/12\big)-5\Big)$;
- set a window of $[0,9,1]\times[-5,20,5]$;
- set Radian mode;
- run the program with an initial Y of 20 and a step length of 1. A step length of 0.1 gives more or less the same curve, showing it is reasonably accurate.

The figure below shows the house temperature assuming a constant outside temperature of 5° (the original question) and the present case with a variable outside temperature. As the outside temperature is always lower in the second case, so is the house temperature.



4. This problem requires exploration. What happens depends on the values used for the parameters. For example, using the values in the original problem, the house temperature does not drop to 13° before the heater switches on to bring the temperature back to 20° by 0700 hours. There is therefore no cycling between 13° and 15°, and the result is the same as the original Scheme A. With a sufficiently large value for k, cycling will occur. The problem then is to determine when to leave the heater on to bring the temperature back to 20° by 0700 hours.

In the cases where the house temperature would drop below 13° in Scheme A, Scheme A will use less energy than the present scheme, which will in turn use less energy than Scheme B. This is because the average house temperature will be lowest in Scheme A and highest in Scheme B, with the current scheme in between. The lower the average house temperature, the less the heat loss that has to be made up by the heater.

2.15 Mousecapades

A very irate Ray, the owner of *Ray's Rest and Recuperation Restaurant and Retreat*, charged out through the kitchen's swinging door because the new chef had forgotten to preheat the oven for the potatoes (Section 2.13). At the same time as Ray left the kitchen, an opportunistic mouse entered the kitchen in search of food scraps on the floor. The swinging door (see the figure below) has adjustment screws that control the amount of friction in the hinges. Its motion is governed by the initial-value problem

$$I\,\frac{d^2\theta}{dt^2}+b\,\frac{d\theta}{dt}+k\theta=0 \qquad \theta(0)=\theta_0 \qquad \frac{d\theta}{dt}(0)=v_0,$$

where θ is the angle that the door is open in radians, t is the time in seconds, I is the moment of inertia of the door about the hinges, $b>0$ is the damping constant that varies with the amount of friction on the door, $k>0$ is the spring constant associated with the swinging door, θ_0 is the initial angle that the door is opened and v_0 is the initial angular velocity imparted to the door.

The width of the door in question is 0.75 m, $I=2.5\,\mathrm{kg\,m^2}$, $b=2.5\,\mathrm{kg\,m^2/s}$, $k=10.625\,\mathrm{kg\,m^2/s^2}$ and, in storming out of the kitchen, Ray gave the door an initial angular velocity of $v_0 = 3\pi\,\mathrm{rad/s}$, letting it go when $\theta_0=\pi/3\,\mathrm{rad}$.

What is the maximum amount of time that the mouse can spend in the kitchen and still escape through the swinging door?

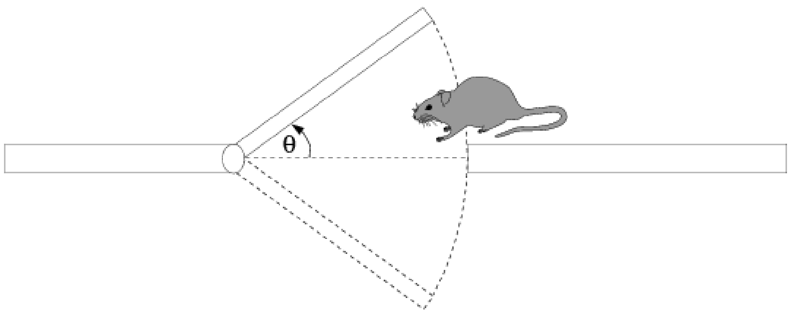
Note: The only way in and out of the kitchen is through this door. The well-fed mouse's size and speed is such that as long as the gap in the doorway is at least 0.05 m, it will be able to get through. You may assume that for small θ, the arc length traced out by the swing of the door is approximately equal to the the gap in the doorway.

A Useful Result

The general solution to the second-order differential equation $\theta''+\theta'+4.25\theta = 0$ is

$$\theta(t) \;=\; e^{-0.5t}\Big(A\cos(2t)+B\sin(2t)\Big),$$

where A and B are arbitrary constants. *You should prove this in your report.*



PTO for the sequel

Son of Mousecapades

The mouse starring in *Mousecapades* stayed in the kitchen for 10 s and was caught. However, one of its offspring, Rudolf (a very smart mouse), worked out that it had about 8.6 s in which to forage in the kitchen, then escape. For a while, Rudolf and friends ate well.

Ray saw the mice a number of times and decided that if he tightened the friction screws in the door sufficiently, the door wouldn't oscillate when he came charging out. Ray tightened the screws so that $b = 22.5\,\mathrm{kg\,m^2/s}$ and reckoned he'd then be able to trap the mice provided they stayed in the kitchen longer than 6 s. Rudolf did some calculations on the back of a cheese wrapper, and the mice stayed in the kitchen for 7 s. If Ray continued to charge out the door in his usual fashion, did he catch the mice?

If you have time, think about the best thing for Ray to do.

Another Useful Result

The general solution to the second-order differential equation $\theta'' + 9\theta' + 4.25\theta = 0$ is

$$\theta(t) = Ae^{-0.5t} + Be^{-8.5t},$$

where A and B are arbitrary constants. *You should prove this too in your report.*

Scenario Writers: Not keen on mice? Make up another script using similar mathematics to *Mousecapades/Son of Mousecapades*. Alternatively, design a better mousetrap.

Instructors' Guide

Mousecapades Solutions

Putting in the values for I, b and k gives a differential equation

$$\frac{d^2\theta}{dt^2} + \frac{d\theta}{dt} + 4.25\theta \;=\; 0,$$

with solution (from *A Useful Result*)

$$\theta(t) \;=\; e^{-0.5t}\Big(A\cos(2t) + B\sin(2t)\Big),$$

where A and B are arbitrary constants. To prove this is a solution, either show that it satisfies the differential equation (a little messy) or use the method of the characteristic equation to solve the differential equation (straightforward).

Initial condition: $\theta(0)=\pi/3 \;\Rightarrow\; A=\pi/3$.

Therefore,

$$\theta(t) \;=\; e^{-0.5t}\left(\frac{\pi}{3}\cos(2t) + B\sin(2t)\right),$$

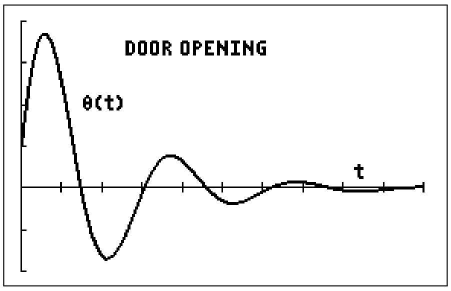
giving

$$\theta'(t) \;=\; e^{-0.5t}\left(-0.5\left(\frac{\pi}{3}\cos(2t) + B\sin(2t)\right) - \frac{2\pi}{3}\sin(2t) + 2B\cos(2t)\right).$$

Then, $\theta'(0) = -\pi/6+2B = 3\pi$ from the given initial condition, so that $B=19\pi/12$ and

$$\theta(t) \;=\; e^{-0.5t}\left(\frac{\pi}{3}\cos(2t) + \frac{19\pi}{12}\sin(2t)\right).$$

A graph of $\theta(t)$ for $0<t<10$ and $-2<\theta<4$ is shown below.

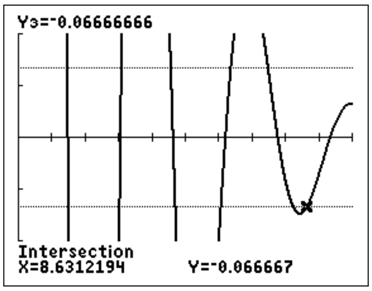


window $[0, 10, 1]\times[-2, 4, 1]$

For small θ, the gap in the door is approximately 0.75θ m, and for the mouse to escape this has to be at least 0.05 m. Therefore,

$$\theta_{\min} \approx \frac{0.05}{0.75} = 0.0\dot{6}.$$

As the door can be open in either direction (θ positive or negative), the latest time that the mouse can leave the kitchen is therefore when $\theta(t)$ is last equal to $\pm 0.0\dot{6}$. We solve this graphically by plotting the lines $y = \pm 0.0\dot{6}$ over the graph of $\theta(t)$ and looking for the largest value of t at which the θ graph intersects either of these lines. This is made easier if we replot the graphs with $-0.1 < \theta < 0.1$.



$[0, 10, 1] \times [-0.1, 0.1, 0.05]$

Using *intersect* on the calculator, we find the latest time at which the door is open the minimum amount is $t = 8.63\,\text{s}$ (intersection with $y = -0.0\dot{6}$).

Therefore, the mouse can spend up to about 8.6 s in the kitchen and still escape.

Son-of-Mousecapades Solutions

Putting in the new value for b gives the differential equation

$$\frac{d^2\theta}{dt^2} + 9\frac{d\theta}{dt} + 4.25\theta = 0,$$

with solution (from *Another Useful Result*)

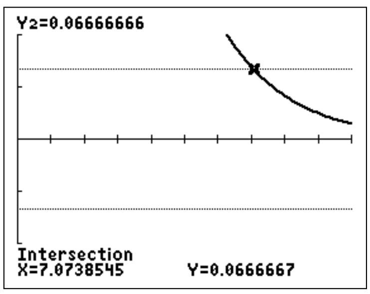
$$\theta(t) = Ae^{-0.5t} + Be^{-8.5t},$$

where A and B are arbitrary constants. As before, prove this is a solution either by showing that it satisfies the differential equation or by using the method of the characteristic equation to solve the differential equation.

Putting in the initial conditions gives

$$\theta(t) = \frac{35\pi}{48}e^{-0.5t} - \frac{19\pi}{48}e^{-8.5t}.$$

Plot as before and find the latest (and only) intersection is with $y = 0.0\dot{6}$ at time $t = 7.07$s. Rudolf was right and the mice escaped.



One thing Ray could do is to open the door less violently, reducing the initial opening θ_0 and the initial velocity v_0 as much as possible, so the door is open for a shorter time.

If Ray is unable to modify his door-opening habits, he could adjust the hinges to obtain critical damping, which gives the fastest approach to equilibrium without oscillations. The analysis is most easily done algebraically, using the method of the characteristic equation.

Assume that $\theta(t) = e^{rt}$, where r is a parameter whose value we have to find so that e^{rt} satisfies the differential equation. Then, $\theta'(t) = re^{rt}$ and $\theta''(t) = r^2 e^{rt}$. Substituting these into the differential equation

$$I\,\frac{d^2\theta}{dt^2} + b\,\frac{d\theta}{dt} + k\theta \;=\; 0$$

and cancelling out a (non-zero) factor of e^{rt} gives the characteristic equation

$$Ir^2 + br + k \;=\; 0,$$

with solutions

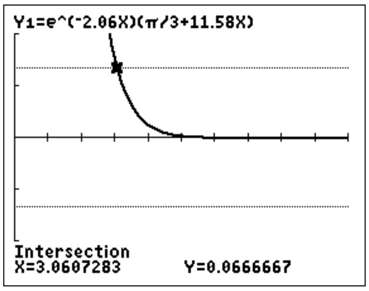
$$r = \frac{-b \pm \sqrt{b^2 - 4Ik}}{2I}.$$

Critical damping is obtained when $\sqrt{b^2 - 4Ik} = 0$ or $b = 2\sqrt{Ik} \approx 10.31$ for the values of I and k given here. Then $r = -b/(2I) \approx -2.06$, and the general solution to the differential equation is

$$\theta(t) \;\approx\; e^{-2.06t}\big(A + Bt\big),$$

where A and B are arbitrary constants. The initial conditions $\theta(0) = \pi/3$ and $\theta'(0) = 3\pi$ give the particular solution

$$\theta(t) \;\approx\; e^{-2.06t}\left(\frac{\pi}{3} + 11.58t\right).$$



The graph of $y = \theta(t)$ only intersects the graph of $y = 0.0\dot{6}$, with the only intersection at time $t = 3.06$ s.

If Ray adjusts the hinges so that $b = 10.31$, the door closes enough to trap the mice in just over 3 s.

2.16 Higher Mathematics Inc

Clients of your firm, Higher Mathematics Inc, have developed a Super Cool Night Vision Detection System for the military. In order to give a digital read-out in appropriate units, the System requires the calculation of natural logs. The System does contain a microprocessor, but, as you know, a microprocessor can only carry out additions and multiplications, that is evaluate polynomials.

The clients require an approximation to the function $y = \ln(1+x)$ in the region of $x = 0$, where x can take on all values > -1.

You, the Head of the Polynomials Section at Higher Mathematics Inc, decide that Taylor polynomials should be used, and now have to write a report for your clients, persuading them that your approach is correct and giving them details of how to carry out the calculations.

The clients have only limited (Year 12) mathematical knowledge but want to be able to understand most of the report (remember they pay the bill). You will therefore have to begin the report with a short explanation of the ideas behind Taylor polynomials and include explanations of all your calculations.

The clients have mathematically more knowledgeable employees who will be checking the details. They have TI-84CE calculators, and so will need to be able to duplicate all your results from the information you give in the report.

Following are some of the details your clients would like to know — any additional useful information you give is likely to make the clients happier about the service you provide and more likely to hire you next time.

1. What do the first five Taylor polynomials look like compared to the function?

2. What is the expression for the nth-degree polynomial $P_n(x)$ (in case they need to extend your calculations later)?

3. Are there x values for which the polynomials do not provide a good approximation, even when the degree n is increased?

4. The clients' favourite number is 5, and on that basis, they would like to use a fifth-degree polynomial in their microprocessor. Over what range of x values is the fifth-degree polynomial accurate to within 0.0001?[41] The clients need to know the endpoints of this range of x values accurate (not just rounded) to 3 decimal places.

Write the report

Suitably explained graphical, numerical or algebraic methods are all acceptable to the clients.

[41] That is, find all x such that $|f(x) - P_5(x)| \leqslant 0.0001$. A graphical method might be useful here.

Calculator Hints

- When programming the polynomials into your calculator, make use of the fact that the polynomial of degree n contains the polynomial of degree $n-1$. Thus, for example, if your polynomials are

  $$P_1(x) = x$$
  $$P_2(x) = x + x^2$$
  $$P_3(x) = x + x^2 + x^3$$
  $$P_4(x) = x + x^2 + x^3 + x^4$$

  you can program these as (assuming you have the function f in Y1)

  Y2 = X

  Y3 = Y2 + X²

  Y4 = Y3 + X³

  Y5 = Y4 + X⁴

  and so on. This scheme is especially useful when the coefficients are complicated.

- A graphical intersection or root-finding routine might be useful for the last part of the problem, but how accurate is it?

The SCNVDS under field trial. The microprocessor screen is blank because the client is awaiting your report.

Instructors' Guide

Students are expected to come up with a reasonably comprehensive report here. Below we just give details of the calculations.

Solutions

1. The Taylor polynomial of degree 5 about $x=0$ for a function $f(x)$ is

$$P_5(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \frac{f^{(4)}(0)}{4!}x^4 + \frac{f^{(5)}(0)}{5!}x^5.$$

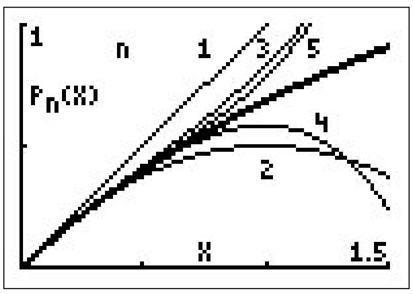
Here,

$$\begin{array}{ll} f(x) = \ln(x+1) & f(0) = 0 \\ f'(x) = \dfrac{1}{x+1} & f'(0) = 1 \\ f''(x) = \dfrac{-1}{(x+1)^2} & f''(0) = -1 \\ f'''(x) = \dfrac{2}{(x+1)^3} & f'''(0) = 2 = 2! \\ f^{(4)}(x) = \dfrac{-3!}{(x+1)^4} & f^{(4)}(0) = -3! \\ f^{(5)}(x) = \dfrac{4!}{(x+1)^5} & f^{(5)}(0) = 4! \end{array}$$

Therefore,

$$\begin{aligned} P_5(x) &= 0 + x - \frac{1}{2!}x^2 + \frac{2}{3!}x^3 - \frac{3!}{4!}x^4 + \frac{4!}{5!}x^5 \\ &= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} \\ &= \sum_{i=1}^{5} (-1)^{i+1}\,\frac{x^i}{i}. \end{aligned}$$

Each polynomial P_n, $n = 1, 2, 3, 4$, is obtained by truncating P_5 at x^n. Graphs of the first five Taylor polynomials and the function $f(x) = \ln(x+1)$ (bold curve) are shown below for $0<x<1.5$, $0<y<1$.



2. Continuing the above process, we see that

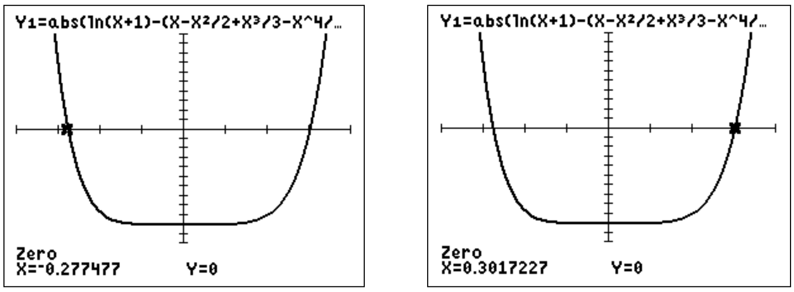
$$P_n(x) = \sum_{i=1}^{n} (-1)^{i+1} \frac{x^i}{i}.$$

3. Graphing the higher-degree polynomials in Question 1 above (or using the Ratio Test if you have covered it) shows that the Taylor series for $f(x) = \ln(x+1)$ converges for $|x| < 1$ The series actually converges at $x = 1$, but not at $x = -1$, where f is undefined. Therefore, the Taylor polynomials will not be useful for $x > 1$.

4. Finding the zeros of $\left|\ln(x+1) - P_5(x)\right| - 0.0001$ using *zero* on the calculator (accurate to 3 decimal places according to the manual) shows that P_5 is accurate to 0.0001 for

$$-0.277 < x < 0.302,$$

both numbers accurate to 3 decimal places.



window $[-0.4, 0.4, 0.1] \times [-1.2\times10^{-4}, 1\times10^{-4}, 1\times10^{-5}]$

2.17 Exploring the Ideas of Linear Equations

Aims

- To see examples of problems leading to linear equations.
- To practise formulating and solving problems.
- To see extensions of the basic linear-equations problem.
- To relax a little over some challenging problems!

Puzzles

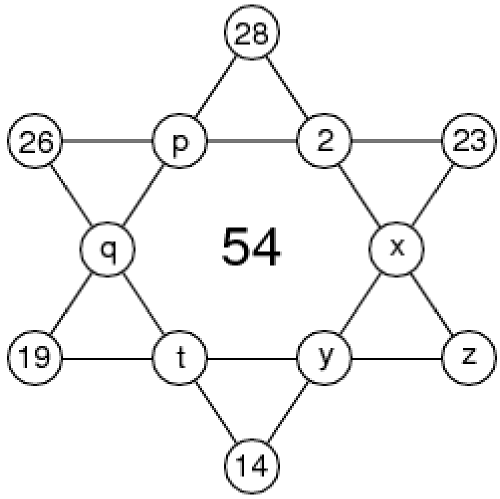
G.H. Hardy (1877 – 1947) was Professor of Mathematics at Cambridge University and one of the finest mathematicians of his time. He wrote a little book, *A Mathematician's Apology*, in which he gave his views about Mathematics, the nature of the subject and its importance. In the book, which is now very famous, Hardy claimed that Mathematics was much more popular than most people realised, and he referred to the puzzles that are given in so many newspapers and magazines. Many puzzles involve sorting out information and using linear equations. Questions 1 and 2 provide examples and Question 5 shows how puzzles lead us to develop some general results for linear equations.

Question 1

I have labelled the unknowns in this popular "Magic Star" puzzle as x, y, z, p, q, t. Find p first, then q and then t. Now explain by writing out the linear equations how your method is like "back substitution". Now solve for x, y, z.

Magic Star

Find the missing numbers. The four numbers along every line add up to the magic number in the center.



The solution set for linear equations may be restricted if we want only positive integers, as naturally happens in some puzzles. Here is an old puzzle written with an American flavour.

Question 2

Find all the ways in which 100 coins — pennies (one cent), dimes (10c) and quarters (25c) — can be worth exactly \$4.99.

Hints: How many unknowns are there?
How many equations are there?
Does your solution involve a parameter?
How does this particular problem affect the range of parameters to use?

You might wonder when linear equations will or will not have integer solutions. We return to this in Question 5.

Question 3 *Beyond equality*

Here is a problem of the type that commonly occurs in all sorts of industrial and other planning problems. It leads to an area of mathematics called *Linear Programming.*

The Fruit Basket

The produce manager of a grocery is making up fruit baskets to sell as gifts. They are to sell for no more than \$5, and contain only apples and oranges. She wants to get 24c per orange, 12c per apple, and 68c for the basket. No more than 26 pieces of fruit will fit in the basket. Suppose she uses x oranges and y apples. How can and should x and y be chosen?

(a) Show that we are looking for those x and y values that satisfy the inequalities

$$\begin{aligned} x &\geq 0 \qquad & x + y &\leq 26 \\ y &\geq 0 \qquad & 2x + y &\leq 36. \end{aligned}$$

(b) When the equality signs are used, we have the equations for 4 lines. Draw a diagram showing those 4 lines and then find an area of the xy plane in which useful x and y values must be found. Shade in that area. *Remember to explain what you are doing.*

(c) Which of these (x, y) values could the manager use?

$$(5, 10) \quad (10, 5) \quad (5, 25) \quad (20, 5) \quad (-2, 10)$$

(d) If she makes a profit of 3 cents on every orange sold and 2 cents for every apple sold, what is the equation for the total profit p?

Draw in the $p = 30$ and $p = 42$ lines on your diagram.

Can you see how to get the maximum profit? *Explain your reasoning.*

What is the maximum profit?

Question 4 *Beyond linear equations*

Even if we have sets of equations that are no longer linear, we may still be able to use our basic ideas to find solutions:

manipulate equations to get *equivalent equations*,

solve using *back substitution*

Here is an example that comes from using the Global Positioning System (GPS).

GPS is based on a constellation of 24 high-altitude satellites — officially known as Navstar GPS (short for Navigation System with Timing and Radar). A simple example of using the system works as follows (although of course more signal analysis and more-elaborate algorithms will be used in practice).

Suppose you know there are three satellites, with position coordinates (a_1, b_1, c_1), (a_2, b_2, c_2) and (a_3, b_3, c_3). By bouncing a signal off each one, you measure their distances r_1, r_2 and r_3 from yourself. Now your (unknown) coordinates (x, y, z) must satisfy

$$\begin{aligned}(x-a_1)^2+(y-b_1)^2+(z-c_1)^2 &= r_1^2\\(x-a_2)^2+(y-b_2)^2+(z-c_2)^2 &= r_2^2\\(x-a_3)^2+(y-b_3)^2+(z-c_3)^2 &= r_3^2.\end{aligned}$$

We have a set of nonlinear equations, but perhaps we can use our ideas developed in Linear Algebra to find a way to get a solution. Let's try with a simple example.

Consider the set of equations

$$x^2+y^2+z^2 = 6 \quad (1)$$

$$(x-1)^2+y^2+(z+1)^2 = 10 \quad (2)$$

$$(x+1)^2+(y+1)^2+(z-1)^2 = 9. \quad (3)$$

(a) *Find solutions for x, y and z.*

Here are some suggestions to help you.

- Can you use Equation (1) to convert Equations (2) and (3) into a pair of *linear* equations?
- Can you solve those linear equations?
- Can you use the idea of back substitution to re-use Equation (1) now? What sort of equation do you have to solve?
- Now can you give the full solution?

(b) *How many solutions did you get?*

Do you think that will always be the case for any set of equations like the general ones we began with?

Why? Can a geometric argument help? What shape is represented by the equations?

Question 5 *Finding some general mathematical results*

Puzzles very often have integer answers, and in fact in ancient times many people could only handle integers. About 2000 years ago, the Greek Diophantus wrote the first real algebra books.

Equations that involve integers and count only integers as possible solutions are called Diophantine equations.

The linear Diophantine equation with two unknowns x and y is

$$my = nx + k \qquad \text{where } m, n \text{ and } k \text{ are given integers.}$$

$$\text{e.g. } 3y = 2x + 7.$$

This example clearly has $x = 7$, $y = 7$ as a (Diophantine) solution.

(a) *All linear Diophantine equations with 2 unknowns have solutions.*

Is that statement true or false?
Notice that examples can never prove it is true in all cases, but we only need one example (a "counter-example") to prove the statement "false". *Explain that in your report.*
Does the example $2y = 6x+5$ help?

Can you give other examples? Perhaps at the end of the lab, if you have spare time, you can come back and try to give a general condition for when

$$my = nx+k$$

has Diophantine solutions.

(b) Our example $3y = 2x+7$ does have a solution and here are some more:

$\boldsymbol{x}$	1	4	7	10	13
$\boldsymbol{y}$	3	5	7	9	11

Can you see a pattern? Write down a formula involving a parameter t for your guess for all the Diophantine solutions. Check that your formula satisfies the equation.

(c) Find a general formula for the solutions of $my = nx+k$ as follows.

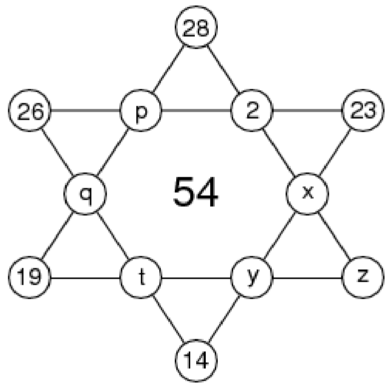
- Assume we have found one solution. Call it $x = x_1$, $y = y_1$.
  So we must have $my_1 = nx_1+k$.
- Now use that to check that $x = x_1+mt$, $y = y_1+nt$ is also a solution.
  Specify the possible t values (for this Diophantine equation, remember).
- In summary, *what can you say about the form and possible number of solutions for linear Diophantine equations for two unknowns?*

Instructors' Guide

Introductory Problem

Magic Star

Find the missing numbers. The four numbers along every line add up to the magic number in the center.



I have labelled the unknowns in this popular "Magic Star" puzzle as x, y, z, p, q and t. Find p first, then q and then t. Now explain by writing out the linear equations how your method is like "back substitution". Now solve for x, y, z.

The equations are:

$$\begin{aligned} x+z &= 54-(28+2) = 24 && (4)\\ x+y &= 54-(14+23) = 17 && (5)\\ y+z+t &= 54-19 = 35 && (6)\\ t+q &= 54-(26+14) = 14 && (7)\\ p+q &= 54-(19+28) = 7 && (8)\\ p &= 54-(26+2+23) = 3 && (9) \end{aligned}$$

The last equation gives $p = 3$.

Substituting this into Eq. (8) gives $q = 7-p = 4$ (back substitution).

Substituting this into Eq. (7) gives $t = 14-q = 10$ (back substitution).

Substituting this into Eq. (6) gives

$$y + z = 25 \qquad (10)$$

Eq. (4) − Eq. (10) gives

$$x - y = -1 \qquad (11)$$

Eq. (5) + Eq. (11) gives $2x = 16$, so that $x = 8$.

Eq. (11) then gives $y = 9$, and Eq. (10) gives $z = 16$ (back substitution).

Therefore, $x = 8$, $y = 9$, $z = 16$, $p = 3$, $q = 4$ and $t = 10$.

Check by substituting back into the original equations.

1. Suppose there are p pennies, d dimes and q quarters.

Then the problem translates into

$$\begin{aligned} p+d+q &= 100 \\ p+10d+25q &= 499 \end{aligned}$$

The solution is $q = t$, $d = (133-8t)/3$, $p = (167+5t)/3$, with $t \in \mathbb{R}$.

We want p, d, q to be non-negative integers

As $q = t$, we must have $t \geqslant 0$. For $t > 16$, d becomes negative, so $0 \leqslant t \leqslant 16$.

On the calculator, if X is the number of quarters, $Y_1 = (133-8X)/3$ is the number of dimes and $Y_2 = (167+5X)/3$ is the number of pennies. Set these equations up using table and look for integer solutions with $0 \leqslant X \leqslant 16$.

X	Y_1	Y_2
0	44.333	55.667
1	41.667	57.333
2	39	59
3	36.333	60.667
4	33.667	62.333
5	31	64
6	28.333	65.667
7	25.667	67.333
8	23	69
9	20.333	70.667
10	17.667	72.333

X	Y_1	Y_2
11	15	74
12	12.333	75.667
13	9.6667	77.333
14	7	79
15	4.3333	80.667
16	1.6667	82.333
17	-1	84
18	-3.667	85.667
19	-6.333	87.333
20	-9	89
21	-11.67	90.667

PTO

2. **(a)** The numbers of oranges and apples are zero or positive, so that

$$x \geq 0 \qquad y \geq 0.$$

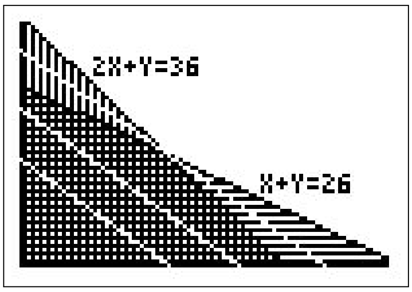
There are to be no more than 26 pieces of fruit, so that $x + y \leq 26$.
The total cost of fruit plus basket is to be no more than \$5, so that

$$\begin{aligned} 24x + 12y + 68 &\leq 500 \\ \text{or} \quad 2x + y &\leq 36. \end{aligned}$$

(b) $x = 0$ and $y = 0$ give the axes. The shaded area bounded by the four lines contains the useful x and y values.

(c) $(5, 10)$ and $(10, 5)$ are in the shaded area, so those x, y values could be used. The other x, y values are not in the shaded area and could not be used.

(d) The profit is given by $p = 3x + 2y$.
The lines for $p = 30$ and 45 are shown below.
All profit lines will be parallel to those. The largest value of p is for the profit line farthest to the right and still intersecting the shaded region, i.e. the line through the corner point $(10, 16)$, giving $p = 62$. The maximum profit is therefore 62 cents, obtained with 10 oranges and 16 apples in a basket.



window $[0, 25, 5] \times [0, 35, 5]$

The white lines are the constant-profit lines for $p = 30$ (bottom left), 45 and 62 (top right).

3. **(a)** Subtracting Eq. (1) from Eqs. (2) and (3) respectively leads to the linear equations

$$\begin{aligned} -x + z &= 1 \\ x + y - z &= 0. \end{aligned}$$

These can be solved to give $x = z - 1$, $y = 1$.

Substituting this into Eq. (1) leads to the quadratic equation

$$\begin{aligned} z^2 - z - 2 &= 0 \\ \text{or} \quad (z-2)(z+1) &= 0. \end{aligned}$$

Thus, we deduce the solutions $z = 2 \quad y = 1 \quad x = 1$
and $z = -1 \quad y = 1 \quad x = -2.$

(b) We get two solutions.

Equations like these always have two solutions, or do they?

Each equation represents a sphere, and solving them simultaneously finds the points in which three spheres intersect, if they do intersect. Two spheres intersect, in general, in a circle. This circle of intersection then intersects the third sphere at two points (two solutions), at one point (circle touches third sphere — one solution) or not at all (no solution).

4. (a) The statement is false.

One counter-example: $2y = 6x+5$ can be written as $2y-6x = 5$, but if x and y are integers, $2y-6x$ must be an even number and hence can never equal 5.

The general equation has a solution if and only if k is a multiple of the greatest common divisor of m and n.

(b) $x = 7+3t$, $y = 7+2t$, where $t = 0, \pm1, \pm2, \pm3, \ldots$, i.e. t is any integer.

(c) Check that $x = x_1+mt$, $y = y_1+nt$, where (x_1, y_1) is a solution of $my = nx+k$.

If $x=x_1$, $y=y_1$ is a solution, we have

$$my_1 = nx_1+k. \tag{12}$$

Then,

$$\begin{aligned} my &= m(y_1+nt) \\ &= my_1 + mnt \\ &= nx_1 + k + mnt \qquad \text{using Equation (12)} \\ &= n(x-mt) + k + mnt \qquad \text{as } x = x_1+mt \\ &= nx + k. \end{aligned}$$

Therefore, $x=x_1+mt$, $y=y_1+nt$, with t any integer, is a solution. It turns out that all solutions are of this form.

Linear Diophantine equations in two unknowns therefore have no solution or an infinite number of solutions.

2.18 Population Models: Matrices and Eigenvalues

Aims

- To see problems that can be modelled mathematically using matrices.
- To explore matrix problems using the TI-84/CE.
- To develop a bit more algebra concerning eigenvalues.

Preliminary work *Program POP/POPCE*[a]

This will allow you to multiply repeatedly a column vector $\boldsymbol{v}$ by a transition matrix $\boldsymbol{T}$ and to see the sum of the components of $\boldsymbol{v}$. $\boldsymbol{T}$ will be a 3×3 matrix stored in [**A**], $\boldsymbol{v}$ will be a 3×1 column matrix stored in [**B**].

Prgm POP/POPCE

[A][B] → [B]	gives $\boldsymbol{Tv}$ and puts the answer into [B] ready for the next multiplication.
[B](1,1) + [B](2,1) + [B](3,1)	gives the sum of the components of $\boldsymbol{v}$.
Disp [B], Ans	displays $\boldsymbol{v}$ and the sum of its components.

You can test this program in Question 1(b) below.
Select this program and press enter to run it. Each time you press enter, it will run again. Do as many enters as you need repeated $\boldsymbol{T}$ multiplications.

[a] available at *www.YYY*

A reminder before you begin

(a) For a matrix $\boldsymbol{T}$, there are special vectors $\boldsymbol{v}$ called **eigenvectors**: multiplying $\boldsymbol{v}$ by $\boldsymbol{T}$ gives a scaled version of $\boldsymbol{v}$, i.e.

$$\boldsymbol{Tv} = \lambda \boldsymbol{v},$$

where the scaling constant λ is the **eigenvalue**. If $\boldsymbol{v}_1$ is an eigenvector with $\lambda_1 = 1$, multiplying by $\boldsymbol{T}$ gives back exactly $\boldsymbol{v}_1$, i.e. $\boldsymbol{Tv}_1 = \boldsymbol{v}_1$.

(b) If $\boldsymbol{u}$ is **not** an eigenvector of $\boldsymbol{T}$, $\boldsymbol{Tu} \neq$ a constant times $\boldsymbol{u}$.

(c) Eigenvalues are the roots of the characteristic equation

$$\left|\boldsymbol{T} - \lambda \boldsymbol{I}\right| = 0,$$

and the eigenvector $\boldsymbol{v}_i$ corresponding to eigenvalue λ_i is found by solving

$$\left(\boldsymbol{T} - \lambda_i \boldsymbol{I}\right)\boldsymbol{v}_i = \boldsymbol{0}.$$

More details on the use of matrices on a TI-84/CE and further examples can be found in *Matrix and Vector Operations* in Volume 3 of *Mathematics on a TI-84/CE*.[42]

[42] available at *www.XXX*

Question 1 *Leslie matrices and beetles*

A population is divided into a number of classes — we shall assume three classes referring to three age groups: *young* y, *adults* a and *seniors* s. The population is then described by the vector

$$\boldsymbol{v} = \begin{bmatrix} y \\ a \\ s \end{bmatrix}.$$

There will be a 3×3 matrix $\boldsymbol{T}$ that tells us how the population evolves. For example, if the population is

$$\begin{bmatrix} y_1 \\ a_1 \\ s_1 \end{bmatrix}$$

to start with, after one cycle it is

$$\begin{bmatrix} y_2 \\ a_2 \\ s_2 \end{bmatrix} = \boldsymbol{T} \begin{bmatrix} y_1 \\ a_1 \\ s_1 \end{bmatrix}.$$

In problems leading to a Leslie matrix:

- in each year or cycle, members of the other classes produce a certain number of new young in Class 1;
- a certain fraction of each class survives to move into the next class and the rest die;
- all members of the top class die.

This leads to a transition matrix that is zero everywhere except possibly

- along the top row after the first element and
- in the elements along the diagonal parallel to and just below the main diagonal.

$$\begin{bmatrix} 0 & * & * & * & * & \cdots & * & * \\ * & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & * & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & * & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & * & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & * & 0 \end{bmatrix}$$

Leslie discovered these matrices in the 1940s when he pioneered this way of exploring how populations can develop. He had TB and taught himself matrix algebra while he was in hospital.

(a) For our beetle population: during each cycle, each adult produces on average 2.75 young and each senior produces on average 2.5 young; one quarter of the young survive to become adults and one half of the adults survive to become seniors. In a Leslie-matrix problem, all the seniors die.

Find the Leslie transition matrix $\boldsymbol{T}$.

Good strategy: Write out the linear equations for y_2, a_2, s_2 in terms of y_1, a_1, s_1 and convert to matrix form.

$$\begin{aligned} y_1 &= ?y_0 + ?a_0 + ?s_0 \\ a_1 &= ?y_0 + ?a_0 + ?s_0 \\ s_1 &= ?y_0 + ?a_0 + ?s_0 \end{aligned}$$

(b) If we start with 40 young and no adults or seniors, show that after one cycle

$$\boldsymbol{v} = \begin{bmatrix} y \\ a \\ s \end{bmatrix} = \begin{bmatrix} 0 \\ 10 \\ 0 \end{bmatrix}.$$

If you do not get this, check your $\boldsymbol{T}$ with a lab instructor. You should have entered $\boldsymbol{T}$ into [A] and the initial $\boldsymbol{v}$ into [B].

(c) Multiply repeatedly by $\boldsymbol{T}$ (using POP/POPCE), and record the population vector $\boldsymbol{v}$ and the population total $P = y + a + s$ after 11, 12 and 13 cycles.

(d) Looking at the values of P, what is your guess for an eigenvalue λ_1 of $\boldsymbol{T}$?
What do you guess for the eigenvector $\boldsymbol{v}_1$?
Check the accuracy of your guess by working out $\boldsymbol{T}\boldsymbol{v}_1$ and seeing whether the result is equal to $\lambda_1 \boldsymbol{v}_1$.

Question 2 *A mathematical check*

Make sure you have done 1(d) **before** *you do Question 2.*

(a) For the Leslie matrix $\boldsymbol{T}$ in Question 1, find the characteristic equation $\left|\boldsymbol{T} - \lambda \boldsymbol{I}\right| = 0$ for the eigenvalues λ.

(b) Check that $\lambda_1 = 1$ satisfies that characteristic equation.

(c) Now check that the eigenvector $\boldsymbol{v}_1$ satisfies $\left(\boldsymbol{T} - \lambda \boldsymbol{I}\right)\boldsymbol{v} = \boldsymbol{0}$ for this case: as $\lambda_1 = 1$, you should expect $\left(\boldsymbol{T} - \boldsymbol{I}\right)\boldsymbol{v}_1 = \boldsymbol{0}$.

Let $\boldsymbol{v}_1 = \begin{bmatrix} 8 \\ 2 \\ 1 \end{bmatrix}$ and check that this homogeneous equation is really satisfied.

Would any other $\boldsymbol{v}_1$ satisfy the equation?
You should now be able to confirm your guess in Question 1(d).

Question 3 *You are the consultant*

An orchardist has beetles in her trees and knows the beetle population evolves as described in Question 1, with a cycle taking one year. At the start of last year she had employed a pest-control company to reduce the beetle numbers and they did that (they claim) by using a spray that killed off the "senior" beetles. This year the orchardist did a survey over 50 trees and found 800 young beetles, 200 adults and 100 seniors. She now says the pest control company failed and she wants her money back.

Discuss her claim and how you would support or not support her case in court.

Question 4 *Populations and oscillations*

Workers other than Leslie had independently used matrix algebra in population models. The first was Harro Bernardelli, who had published a paper in 1941 in the Journal of the Burma Research Society with the title *Population Waves*. Bernardelli's paper was unusual in focussing not on the eventual stability of the population structure, but on intrinsic oscillations in the population structure. He had observed oscillations in the age structure of the Burmese population between 1901 and 1931. As an abstract model for such oscillations, he proposed a matrix model for the evolution of the population with

$$\boldsymbol{T} = \begin{pmatrix} 0 & 0 & 8 \\ 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \end{pmatrix}$$

and showed by numerical calculations that this gave rise to apparently permanent oscillations in the age structure.

(a) Set the vector $\boldsymbol{v}$ initially to $\begin{bmatrix} 1 \\ 0.01 \\ 0.01 \end{bmatrix}$ (population in millions).

Then use POP/POPCE to run Bernardelli's matrix $\boldsymbol{T}$ for 12 cycles, recording the total population at each cycle. Plot P versus cycle number, joining up the points with straight lines.

(b) Discuss your findings.

Could you have predicted something like that just by thinking how the population classes develop according to Bernardelli's model?

Check the case where initially $y=1$, $a=0$ and $s=0$. *You do not even need a calculator to get results in this very simple case!*

(c) Repeat (a) using $\boldsymbol{T} = \begin{bmatrix} 0 & 0 & 5 \\ 0.7 & 0 & 0 \\ 0 & 0.5 & 0 \end{bmatrix}$. Describe your results in words.

Question 5 *Review*

You have now seen examples of systems specified by a vector and with a matrix producing transitions to show how that system's vector evolves. What types of behaviour have you observed? Summarise what you see as the various possibilities and speculate on other types of results you think could be obtained in these linear problems.

Supplementary Question

Question 6 *Managing a car hire business*

You are the area manager for a car rental company with three hiring locations P, Q and R, and 600 cars under your control. You must decide where to base your cars. Luckily you have done Maths 1 so you know how to build a mathematical model to tackle this problem.

A little monitoring of car hire and returns gives you the following data or Weekly Distribution History:

location P: 60% of cars starting at P remain there at the end of the week, 10% go to Q, 30% go to R

location Q: 80% of cars starting at Q remain there at the end of the week, 10% go to P, 10% go to R

location R: 70% of cars starting at R remain there at the end of the week, 10% go to P, 20% go to Q.

(a) Let p_n, q_n and r_n be the number of cars at locations P, Q and R respectively at the start of week n. They can be the components of a column vector. What is the transition matrix $\boldsymbol{T}$ which tells you the distribution p_{n+1}, q_{n+1} and r_{n+1} at the start of Week $n+1$?
Answers are available from the lab supervisors if you want to check at any stage. Record in your report whenever you check at a particular stage.

(b) If Week 1 starts with 200 cars at each location, how are they distributed at the start of Week 2?
Work this out by hand, then use program POP/POPCE to check. Another check: has the total number of cars remained constant?

(c) How do you assign your staff (clerks, cleaners, mechanics) to the three locations?
Hint: You really need staff numbers at each location proportional to the number of cars there. Do you want to have to keep moving them around?

Explain how an eigenvector of $\boldsymbol{T}$ helps and why the eigenvalue should be 1.

(d) Substitute $\lambda = 1$ into the characteristic equation $|\boldsymbol{T} - \lambda \boldsymbol{I}| = 0$ to check that 1 is indeed an eigenvalue of $\boldsymbol{T}$.

(e) Now find the corresponding eigenvector $\boldsymbol{v}_1$. Make it suitable for a total of 600 cars. Check $\boldsymbol{T}\boldsymbol{v}_1$ really does give you back the same distribution and therefore convince yourself that this will keep cars located in a steady-state way.

(f) You employ a total of 20 car cleaners. How many should you, smart manager, base at each location P, Q and R?

Bonus Parts

(g) Check the other solutions to the eigenvalue problem for $\boldsymbol{T}$:

$$\lambda_2 = 0.6 \qquad \boldsymbol{v}_2 = t\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \qquad \lambda_3 = 0.5 \qquad \boldsymbol{v}_3 = t\begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}.$$

(h) You might wonder: is there some test we can do to detect a mistake in the eigenvalues of a matrix $\boldsymbol{A}$. Suppose $\boldsymbol{A}$ is $n \times n$ and let $p(\lambda)$ be the characteristic polynomial. We know

- $p(\lambda) = \left|\boldsymbol{A} - \lambda \boldsymbol{I}\right|$ and the characteristic equation is $p(\lambda) = 0$;
- if the eigenvalues are $\lambda_1, \lambda_2, \cdots, \lambda_n$, it must be possible to write

$$p(\lambda) \;=\; (-1)^n(\lambda - \lambda_1)(\lambda - \lambda_2)(\lambda - \lambda_3)\ldots(\lambda - \lambda_n).$$

Now have a think about $p(0)$. Work it out using each of the above two results and hence obtain your test for the eigenvalues.

(i) Check that the three eigenvalues satisfy your test for the matrix $\boldsymbol{T}$.

Instructors' Guide

Solutions

1. **(a)** $\boldsymbol{T} = \begin{bmatrix} 0 & 2.75 & 2.5 \\ 0.25 & 0 & 0 \\ 0 & 0.5 & 0 \end{bmatrix}.$

(b) $\boldsymbol{T}\begin{bmatrix} 40 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 10 \\ 0 \end{bmatrix}.$

(c)

After cycle	$\boldsymbol{v}$	$\boldsymbol{P}$
1	$\begin{bmatrix} 0 \\ 10 \\ 0 \end{bmatrix}$	10
2	$\begin{bmatrix} 27.5 \\ 0 \\ 5 \end{bmatrix}$	32.5
3	$\begin{bmatrix} 12.5 \\ 6.875 \\ 0 \end{bmatrix}$	19.375
⋮	⋮	⋮
11	$\begin{bmatrix} 17.320 \\ 4.305 \\ 2.184 \end{bmatrix}$	23.809
12	$\begin{bmatrix} 17.299 \\ 4.330 \\ 2.153 \end{bmatrix}$	23.781
13	$\begin{bmatrix} 17.289 \\ 4.325 \\ 2.165 \end{bmatrix}$	23.779

(d) It appears that P is becoming close to a constant value, so we guess $\lambda_1 = 1$ (and that this is the eigenvalue with the largest absolute value).

The eigenvector is close to $\begin{bmatrix} 17.289 \\ 4.325 \\ 2.165 \end{bmatrix} = 2.165 \begin{bmatrix} 7.99 \\ 2.00 \\ 1 \end{bmatrix}.$

In fact the exact answer is $\lambda_1 = 1$, $\boldsymbol{v}_1 = t\begin{bmatrix} 8 \\ 2 \\ 1 \end{bmatrix}$, t any real number.

2. **(a)** The characteristic equation is $\lambda^3 - 0.6875\lambda - 0.3125 = 0$.

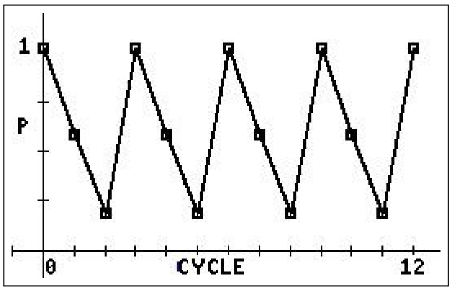
(c) Any multiple of $\boldsymbol{v}_1$ will satisfy the equation, but all other vectors will not.

3. From Question 2, we see that the survey numbers give an eigenvector ($100\boldsymbol{v}_1$ in Question 2) of $\boldsymbol{T}$, with eigenvalue 1. Therefore, the numbers of beetles the year before (after the spraying) must have been the same as now — the pest-control company appears to have killed no beetles at all.

4. **(a)** After 1 cycle: $\begin{bmatrix} 0 & 0 & 8 \\ 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0.01 \\ 0.01 \end{bmatrix} = \begin{bmatrix} 0.08 \\ 0.5 \\ 0.0025 \end{bmatrix}$, so $P = 0.5825$.

From successive cycles, we build up the table

cycle	0	1	2	3	4	5	6	···
$\boldsymbol{P}$	1.02	0.5825	0.185	1.02	0.5825	0.185	1.02	···



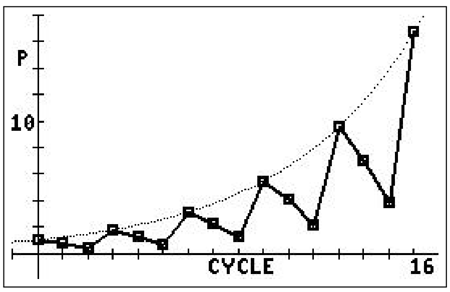
window $[0, 12, 1] \times [0, 1.2, 0.25]$

(b) The population is oscillating or going in waves, with no overall growth or decline. A group of young first becomes adults with a decline of 50%; the group then has to become seniors before producing young, and the process repeats itself.

If you begin with $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, it is even easier to see that the cycle is

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 0 \\ 0.5 \\ 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 0 \\ 0 \\ 0.125 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \longrightarrow \cdots\cdots$$

(c) The populations are now 1.02, 0.755, 0.41, 1.785, 1.321, 0.718, 3.124, 2.312, 1.256, 5.467, 4.046, 2.197, 9.566, 7.081, 3.846, 16,741, ... The population is oscillating, but growing overall.



window $[0, 16, 1] \times [0, 18, 2]$

6. (a) We have $\boldsymbol{T} = \begin{bmatrix} 0.6 & 0.1 & 0.1 \\ 0.1 & 0.8 & 0.2 \\ 0.3 & 0.1 & 0.7 \end{bmatrix}$, so that $\begin{bmatrix} p_{n+1} \\ q_{n+1} \\ r_{n+1} \end{bmatrix} = \begin{bmatrix} 0.6 & 0.1 & 0.1 \\ 0.1 & 0.8 & 0.2 \\ 0.3 & 0.1 & 0.7 \end{bmatrix} \begin{bmatrix} p_n \\ q_n \\ r_n \end{bmatrix}$.

(b) $\begin{bmatrix} 200 \\ 200 \\ 200 \end{bmatrix} \longrightarrow \begin{bmatrix} 160 \\ 220 \\ 220 \end{bmatrix}$ after 1 cycle.

(c) The staff should be distributed in proportion to the number of cars at each location. A stable distribution would be sensible, so the staff do not have to be shifted around.

Mathematically we need to see if $\boldsymbol{T}$ has an eigenvalue $\lambda = 1$ and use the eigenvector for the steady-state distribution. ($\lambda \neq 1$ means the total number of cars goes up or down — not what we want.)

(d) With $\lambda = 1$, $\left|\boldsymbol{T} - \lambda \boldsymbol{I}\right| = \left|\boldsymbol{T} - \boldsymbol{I}\right| = \begin{vmatrix} -0.4 & 0.1 & 0.1 \\ 0.1 & -0.2 & 0.2 \\ 0.3 & 0.1 & -0.3 \end{vmatrix} = 0$ (check).

Therefore, $\lambda = 1$ is an eigenvalue.

(e) Solve $(\boldsymbol{T} - \boldsymbol{I})\boldsymbol{v} = \boldsymbol{0}$. Using Gaussian elimination,

$$\begin{bmatrix} -4 & 1 & 1 & 0 \\ 1 & -2 & 2 & 0 \\ 3 & 1 & -3 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & -2 & 2 & 0 \\ -4 & 1 & 1 & 0 \\ 3 & 1 & -3 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & -2 & 2 & 0 \\ 0 & -7 & 9 & 0 \\ 0 & 7 & -9 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & -2 & 2 & 0 \\ 0 & -7 & 9 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Back-substitute to get $\boldsymbol{v} = \begin{bmatrix} 4t/7 \\ 9t/7 \\ t \end{bmatrix}$. For 600 cars, $20t/7 = 600$, so that $t = 210$

and $\boldsymbol{v} = \begin{bmatrix} 120 \\ 270 \\ 210 \end{bmatrix}$. Check that $\boldsymbol{T} \begin{bmatrix} 120 \\ 270 \\ 210 \end{bmatrix} = \begin{bmatrix} 120 \\ 270 \\ 210 \end{bmatrix}$.

(f) Distribute the staff in proportion to number of cars. Therefore, at P, we have $20 \times (120/600) = 4$ staff; similarly 9 staff at Q and 7 staff at R.

(g) Check these are solutions to $\boldsymbol{T}\boldsymbol{v} = \lambda \boldsymbol{v}$.

(h) Notice that $p(0) = \left|\boldsymbol{A}\right|$ and $p(0) = \lambda_1 \lambda_2 \lambda_3 \ldots \lambda_n$, i.e.

product of eigenvalues = determinant of matrix.

(i) $\left|\boldsymbol{T}\right| = 0.3$ and $\lambda_1 \lambda_2 \lambda_3 = 1 \times 0.6 \times 0.5 = 0.3$.

2.19 Gambling Returns

Question 1 *Means and variances for a roulette wheel*

In Australia, casinos have adopted the European roulette wheel. This wheel has 37 slots numbered 0 and 1 to 36. (It gives better odds than the Las Vegas wheels which have 38 slots numbered 00, 0 and 1 to 36.) Half the slots 1 to 36 are coloured red and half are black. The 0 slot is neither red nor black. The casinos pay even money for *red/black* and *odd/even* bets, that is you win as much as you bet.[43]

To see what sort of margin the casino has over the player, we shall examine two of the many possible bets — *straight-up* bets, where a player bets on one particular number, and *red/black* bets, where the player bets on either red or black.

A *straight-up* \$1 bet returns \$35 for a win (plus the \$1 you bet), otherwise the \$1 is lost. The probability of a *straight-up* win is 1/37, while the probability of a loss is 36/37.

To calculate the margin operating in favour of the casino, we calculate the expected return to the gambler for a \$1 bet. The player's return is +\$35 for a win and −\$1 for a loss. Let r_i be the ith return ($i = 1, 2$) and $p_R(r_i)$ be the probability of return r_i. Then $r_1 = 35$, with $p_R(r_1) = 1/37$, and $r_2 = -1$, with $p_R(r_2) = 36/37$. The expected or mean return is

$$\begin{aligned} E[R] &= \sum_{i=1}^{2} r_i\, p_R(r_i) \\ &= r_1\, p_R(r_1) + r_2\, p_R(r_2) \\ &= 35 \times \frac{1}{37} - 1 \times \frac{36}{37} \\ &= -\$\frac{1}{37} \\ &\approx -\$0.027 \\ &= -2.7\% \text{ of the initial bet.} \end{aligned}$$

We say that the casino's margin for the *straight-up* bet is 2.7%, i.e. for every \$100 bet on the game, it will take, *on average*, \$2.70 for itself.

Now let us examine the outcome for a *red/black* bet. A \$1 wager on red, say, returns \$1 for a win, otherwise the \$1 bet is lost. The probability of a win on red is 18/37, while the probability of a loss is 19/37.

(a) Calculate the expected return for a player placing a *red/black* bet.

You should have found that the casino margin was again 2.7%. In fact, by the same argument, you can show that *the margin is the same for every bet on the European wheel.* Hence, whatever bet you make, your expected loss is still 2.7%. What this means is that, *on average*, a player will lose \$2.70 for every \$100 wagered, no matter which way he or she bets.

[43] If you win at a gambling game, you get back what you win (the *return*) plus what you bet.

However, while the two betting strategies yield the same expected return, the spread or variability of the return differs greatly in each case.

(b) **(i)** Suppose a player bets \$1 one hundred times in succession on *red* each time. What is the most she could win? What is the most she could lose?

(ii) Suppose a player bets \$1 one hundred times in succession on a single number. What is the most he could win? What is the most he could lose?

To complete the picture we need to look at the *variance* of the return: the greater the variance, the more volatile the fortune of the player.

To illustrate this, we examine again the cases of a *straight-up* bet and a *red/black* bet. For the *straight-up* \$1 bet, we have

$$\begin{aligned} E[R^2] &= \sum_{i=1}^{2} r_i^2 p_R(r_i) \\ &= (35)^2 \times \frac{1}{37} + (-1)^2 \times \frac{36}{37} \\ &\approx 34.08. \end{aligned}$$

Hence the variance

$$\begin{aligned} \mathrm{Var}(R) &= E[R^2] - (E[R])^2 \\ &= 34.08 - (-\frac{1}{37})^2 \\ &\approx 34.08. \end{aligned}$$

(c) **(i)** Calculate the variance of the return for a \$1 bet on *red/black*.

(ii) Compare the variances for *the straight-up* and *red/black* bets, and explain what this difference in variability means in practice. Note that the greater the variance of R, the more variation R exhibits about the mean return of -2.7%.

If we did not know the formulas to calculate these quantities, we could *estimate* these quantities by performing a large number of simulated spins of the wheel. In this lab, we shall simulate 999 spins on a TI-84/CE.[44] We calculate the *sample mean return*

$$\overline{R} = \frac{R_1 + R_2 + \ldots + R_{999}}{999} = \frac{1}{999}\sum_{i=1}^{999} R_i$$

to estimate $E[R]$, and to estimate $\mathrm{Var}(R)$ we calculate the *sample variance*

$$S^2 = \frac{(R_1 - \overline{R})^2 + (R_2 - \overline{R})^2 + \ldots + (R_{999} - \overline{R})^2}{999} = \frac{1}{999}\sum_{i=1}^{999}(R_i - \overline{R})^2.$$

[44] Lists on the TI-84/CE only take up to 999 values. We need the values in a list to carry out our statistical analysis.

(d) Use the program WINBET/WINBETCE (see next page) to simulate returns from each of 5 runs of 999 *straight-up* bets and from each of 5 runs of 999 *red/black* bets. Share this around — do two or three runs each. For each simulation write down the *return*, the *mean return per bet* $\overline{R}$ (= $R/999$) and the *variance* S^2 (the last two to 4SD); the program calculates these using the stat operations on the TI-84/CE (see next page).

(e) Write down a formula for the return as a function of the number of wins in 999 bets. *Hint*: What return would you get from 1 win? from 2 wins? from n wins? What is the minimum number of wins for a profit (positive return)?

(f) We would now like to combine, if possible, the statistics from our 5 runs to calculate the corresponding statistics for a single run of $5\times 999 = 4995$ spins. Show mathematically, using the formula for $\overline{R}$ on page 153, that the mean of $5\overline{R}$ values is equal to the mean of a single run of the same 4995 spins, i.e. of all spins done in one go. Is this true for the variance? What is your value for $\overline{R}$ from all 5 runs of the *straight-up* bets combined? From all 5 runs of the *red/black* bets combined?

(g) Compare your results in (d) and (f) with the theoretical means and variances obtained above.[45]

(h) In a paragraph, summarize your findings in (a) – (g) as an explanation to a novice gambler.

Question 2 *The Law of Averages or The Law of Large Numbers*

Many people misunderstand the Law of Averages.

A sales book said something along the lines of

> If 1 in 2 people are interested in buying your product and you get a knockback at one house, the **Law of Averages** says you **will** get a sale at the very next house.

It makes you wonder whether this guy had ever tried it himself.

It's like saying that if red came up on the roulette wheel last time, black *must* come up next. If that were the case, you could correctly predict the next umpteen trillion spins of the wheel and retire tomorrow.

(a) Let the probability of an event be p. Let the number of times this event occurs in n trials be x. To what value would you expect the ratio x/n to approach as $n\to\infty$?[46]

This is the Law of Averages, which concerns the *long-term behaviour* of a process. In the statistical literature, this is referred to as the Law of Large Numbers.

The Law of Averages does *not* say that one event or a series of events causes another event to occur next, in order to even things out. In fact the opposite is true: the Law of Averages depends for its validity on each event being independent, i.e. unaffected by earlier events.

(b) Use the program COINTOSS/CNTOSSCE (see next page) to simulate 999 tosses of a fair coin. This program plots H/N against N, where H is the number of heads thrown in N tosses. To what value does H/N converge?

[45] These are the values we would obtain in the limit as the number of simulated spins tends to ∞.

[46] If you can't work this out, let the event be *throwing a 6* on a die. What would you expect in this case?

(c) From your simulation in (b), find how many tosses it takes before H/N *stays within* 10% of the limiting value you found in (b). This number is given on the graph.

Do this for two further runs each of 500 tosses.

(d) Summarize your findings in (a) – (c) with regard to the Law of Averages.

Statistics Operations on the TI-84/CE

The statistics operations are in the [stat] menu, and operate on data contained in the lists. List names $L_1 - L_6$ are on the keys [1]–[6], and are selected by pressing the [2nd] key first. [stat] Edit gets you into the list editor to look at the lists, [2nd] [quit] out of it.

To calculate means and variances, you need to use one-variable statistics on list L_1: you can do this manually by selecting *1-Var Stats* from the [stat] CALC menu, followed by [2nd] [1] (L_1) [enter], or you can let the WINBET/WINBETCE program do exactly this for you.

The mean is $\overline{x}$ and the variance we want in Question 1 is $(\sigma \text{x})^2$.

Using the Programs

These programs are available at *www.YYY*.

WINBET/WINBETCE

This program simulates a game such as roulette using, as do the other simulation programs, the *rand* function on the TI-84/CE ([math] PRB menu): this generates a random number between 0 and 1.

Enter the amount you win on a $1 bet, the probability of a win (you can enter a fraction here, e.g. 1/37) and the number of trials (spins). For example, for a *straight-up* bet you would win $35 for a $1 bet and the probability of a win would be 1/37. Press [enter] to start the simulation. The counter tells you the number of the current trial.

After the simulation is over, the program tells you the return, that is how much you won (+) or lost (−). The results of each trial are stored in L_1. Press [enter] to calculate the mean and variance.

Note: Storing values in the calculator uses up LOTS of memory, about 1K per 100 values. You may need to delete something ([mem]) on a TI-84 so you have enough memory; the CE has heaps of memory. Remember to delete the lists when you have finished.

COINTOSS/CNTOSSCE

This program simulates tossing a coin (fair or biassed). Enter the probability of *heads* and the number of trials (up to 999). The program plots the ratio H/N versus N, where H is the total number of *heads* in N tosses. The program also displays the value of N beyond which H/N remains within 10% of the probability of *heads*.

Values of H/N are stored in list L_1; you should clear this when you have finished.

Instructors' Guide

Solutions

1. **(a)** $E[R] = \$1\times(18/37) - \$1\times(19/37) = -\$1/37 \approx -\$0.027$, 2.7% of the initial bet.

(b) **(i)** The most she could win is \$100 — win \$1 each time.
The most she could lose is \$100 — lose \$1 each time.

(ii) The most he could win is \$3500 — win \$35 each time.
The most he could lose is \$100 — lose \$1 each time.

(c) **(i)** $E[R^2] = (1)^2\times(18/37) + (-1)^2\times(19/37) = 1$.
Hence the variance $\mathrm{Var}(R) = E[R^2]-(E[R])^2 = 1-(-1/37)^2 \approx 0.9993$.

(ii) The variance for the return on a straight-up bet is much higher than the variance for the return on a red-black bet. This is not surprising, because we would expect the winnings or losses of a *red-black* player to remain much more steady than a *straight-up* player, who could easily have a long losing streak before winning or, conversely, may have big wins very quickly. The fortunes of a *straight-up* player are more volatile; he or she will win 'big' or 'go bust' much faster than the *red-black* player.

(d) Examples of using the WINBET/WINBETCE program are given below. Obviously the numbers will vary from trial to trial. The values for sample mean and sample variance have been rounded to 4 significant digits.

Straight-up Bet: \$1 bet; return on win \$35; win probability 1/37; 999 trials.

Simulation Number	Total Return	Sample Mean $\overline{R}$	Sample Variance S^2
1	−63	−0.0631	32.85
2	225	0.2252	42.61
3	261	0.2613	43.81
4	−99	−0.0991	31.62
5	−207	−0.2072	27.91

Red-Black Bet: \$1 bet; return on win \$1; win probability 18/37; 999 trials.

Simulation Number	Total Return	Sample Mean $\overline{R}$	Sample Variance S^2
1	5	0.0050	1.000
2	−55	−0.0551	0.9970
3	−47	−0.0470	0.9978
4	−47	−0.0470	0.9978
5	−9	−0.0090	0.9999

(e) For Simulation 1 in the straight-up bets, the return is −\$63. If this is a result of n wins, then $-63 = 35n + (999-n)(-1)$, so that $36n=936$ or $n=26$.

For a profit, return $35n+(999-n)(-1) > 0$ or $n>999/36=27.75$. Hence, you need to win at least 28 times out of 999 to make a profit.

(f) We have 5 runs of 999 bets with means $\overline{R_1}$, $\overline{R_2}$, ...,$\overline{R_5}$. Then, for the first run,

$$999\overline{R_1} = \sum_{i=1}^{999} R_i,$$

and similarly for each run. If we add up these 5 equations, we have

$$999(R_1 + R_2 + R_3 + R_4 + R_5) = \sum_{i=1}^{4995} R_i,$$

i.e. the right-hand-side is the sum of the returns from all 5 runs. Dividing both sides by 4995, we have

$$\frac{R_1 + R_2 + R_3 + R_4 + R_5}{5} = \frac{1}{4995}\sum_{i=1}^{4995} R_i,$$

that is, the mean of the means is the overall mean, as required.

The same does not work for the variance, because the mean in the overall variance is the overall mean, whereas the mean in each individual variance is the individual mean. The sum of the variances does not turn into the variance of the overall sum because of the square in each term.

The mean from all 5 runs of the straight-up bets in (d) combined is 0.023, while the mean from all 5 runs of the red/black bets combined is −0.031.

(g) The theoretical mean in both cases is −0.027. We see from the straight-up results, both individual and combined, that there is considerable variation from this value, as we would expect given the large variances. We would need to carry out a lot more bets before the experimental values will approach the theoretical value.

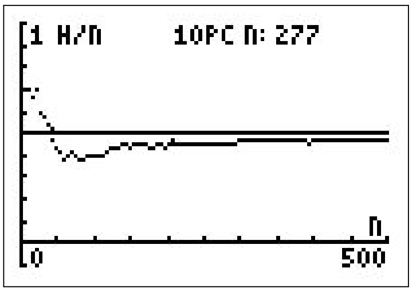
In the red/black case, the individual results, and especially the combined result, are reasonably close to the theoretical value. Again, we would expect this, given the small variances in this case.

2. **(a)** $P(\textit{throwing a 6 on a die}) = 1/6$.

For a large number of throws, we'd expect *no of 6s / no of throws* $\to 1/6$.

Similarly, the Law of Averages states:
If the probability of an event is p and the event occurs x times in n trials, the ratio $x/n \to p$ as $n \to \infty$.

(b) A sample run of COINTOSS is shown below.



(c) From four runs of the program, it took 439, 124, 277 (the above run) and 57 tosses respectively before the value of H/N remained within 0.05 (10%) of the limiting value of 0.5.

2.20 Gambling Tactics

Question 1 *Bold play vs cautious play*

The scenario is that you have \$9, but this by itself is of no use to you. Payday is tomorrow, but tonight you need \$20 for a movie/meal/drink or whatever.

As you are passing the casino, you decide to bet the \$9 to try to get it to \$20. If you lose the \$9 you walk home, but if you get your bankroll to \$20 you are going to enjoy a night out. Suppose also that you have decided to bet on red/black at roulette and that the minimum table bet is \$1.

Is cautious play or bold play preferable?

(a) Bold Play

Use the program LABRLTTE/LBRLTECE. This program simulates a European roulette wheel, with 37 numbers (0 − 36). Your starting bank is \$9.

- Bet the maximum you have to in order to reach your target. That is, on your first bet place \$9. Bet on red or black. If you win, then your next bet is \$1, as that is all that is required to get you to \$20. In general, bet as much as you can or as much as you need to reach your target in one bet, whichever is smaller.
- Run this program 20 times and record the number of times you reached your target of \$20.

(b) Cautious Play

You can also simulate cautious play with LABRLTTE/LBRLTECE, betting \$1 each time, but even with a starting bank of \$9, it could take a lot of bets to reach \$20 or \$0, and therefore a long time. It's quicker to use the WALK/WALKCE program.

In the WALK/WALKCE program, A (you) wins 1 (from the bank B) with probability p or loses 1 (to B). The program continues until either A reaches 0 (you lose) or B reaches 0 (and therefore A reaches 20 and you win). The final I value tells you just how many bets it took for this to occur.

- Run WALK/WALKCE, setting A's capital to 9 and B's to 11.
- Set the probability of a win to 18/37.
- Run the program 20 times and record the number of times A reached the target of 20.

(c) When you have completed (a) and (b), put your results up on the board, so we can get some pooled data. Comment on what you would **expect** the results to be, based on the theory. Compare the theory with what your group obtained and with the overall class results.

Question 2 *Systems betting*

The claim is that the Martingale and Labouchere systems are dangerous in that they can produce a long string of small wins which don't prepare you for the massive loss which is around the corner. Now you will run some simulations so you can check this for yourself, and also see how Martingale and Labouchere systems differ from their reverse systems.

(a) Use the program GAMBLSYS/GBLSYSCE to compare the outcomes of five different betting strategies:

(i) betting the same amount each time;

(ii) Martingale;

(iii) reverse Martingale;

(iv) Labouchere (sequence 1,2,3,4); and

(v) reverse Labouchere (sequence 1,2,3,4).

For each system, set the starting bank to \$100, the probability of a win to 18/37 and the maximum number of bets to 100.

- Run each system 10 times.
- For each simulation, record the minimum, maximum and final bank of the player (press [enter] after the graph is plotted to see these numbers, and [enter] again to return to the main menu). Note that you are allowed to go negative in these simulations. In reality, you would have lost your money.
- When you have completed all the simulations, put your results on the board. Comment on the differences apparent between the five systems, both from your results and the pooled class results.

(b) Complete one copy of the handout tree diagram below for each of the Martingale and Labouchere systems. For Labouchere, use the sequence $\{1, 2, 3, 4\}$, so the initial bet is 5 units. For Martingale, start with a bet of 1 unit.

(i) Fill in the amounts bet, won and lost at each stage, and the return for each branch. Write the returns under the *Return* column. As a check, the returns should add up to 0.

(ii) Let p denote the probability of a win, so that $q = 1-p$ is the probability of a loss. Write down the probability of each return on your diagram in terms of p and q. Write this under the *Probability* column.

(iii) Multiply each term in the *Return* column by the corresponding term in the *Probability* column and total these. This gives you the expected return $E[R]$.

PTO

(iv) In a similar way, the expected amount bet $E[B]$ can be calculated.
For the Martingale system, it is

$$E[B] \;=\; 4p^3 + 15p^2q + 22pq^2 + 15q^3.$$

For the Labouchere system it is

$$E[B] \;=\; 20p^3 + 61p^2q + 76pq^2 + 26q^3.$$

Don't do it now, but after the lab you should try to verify these results.

(v) Evaluate $E[R]$ and $E[B]$ for the Martingale system when $p = 18/37$ (roulette).
What percentage of $E[B]$ is $E[R]$?

(vi) Evaluate $E[R]$ and $E[B]$ for the Labouchere system when $p = 18/37$.
What percentage of $E[B]$ is $E[R]$?

(vii) Comment on the results in (v) and (vi).

SYSTEM:

Return ↓ **Probability** ↓

$E[R] =$

$E[B] =$

$E[R]$ is ______ % of $E[B]$.

Instructors' Guide

This lab was used in a course in which gambling, gambling strategies and gambling systems were used as a vehicle to teach probability to first-year university students. The lab relies on some knowledge of several well-known gambling systems, which are discussed in the Appendix at the end of this section, as is some of the background to bold vs cautious play. The TI-84/CE programs LABRLTTE/LBRLTECE, WALK/WALKCE and GAMBLSYS/GBLSYSCE are available at *www.YYY*.

Solutions

1. **(a)** Clearly the numbers will vary here, but the general result should be clear, especially if class data are pooled. You should allow time for the pooled results to be discussed by the class and conclusions drawn.

In 20 trials of bold play, the \$20 was reached 8 times.

(b) In 20 trials of cautious play, the \$20 was reached 6 times.

According to the theory (see the Appendix), with the odds of winning (18/37) in the house's favour, bold play should be the better strategy.

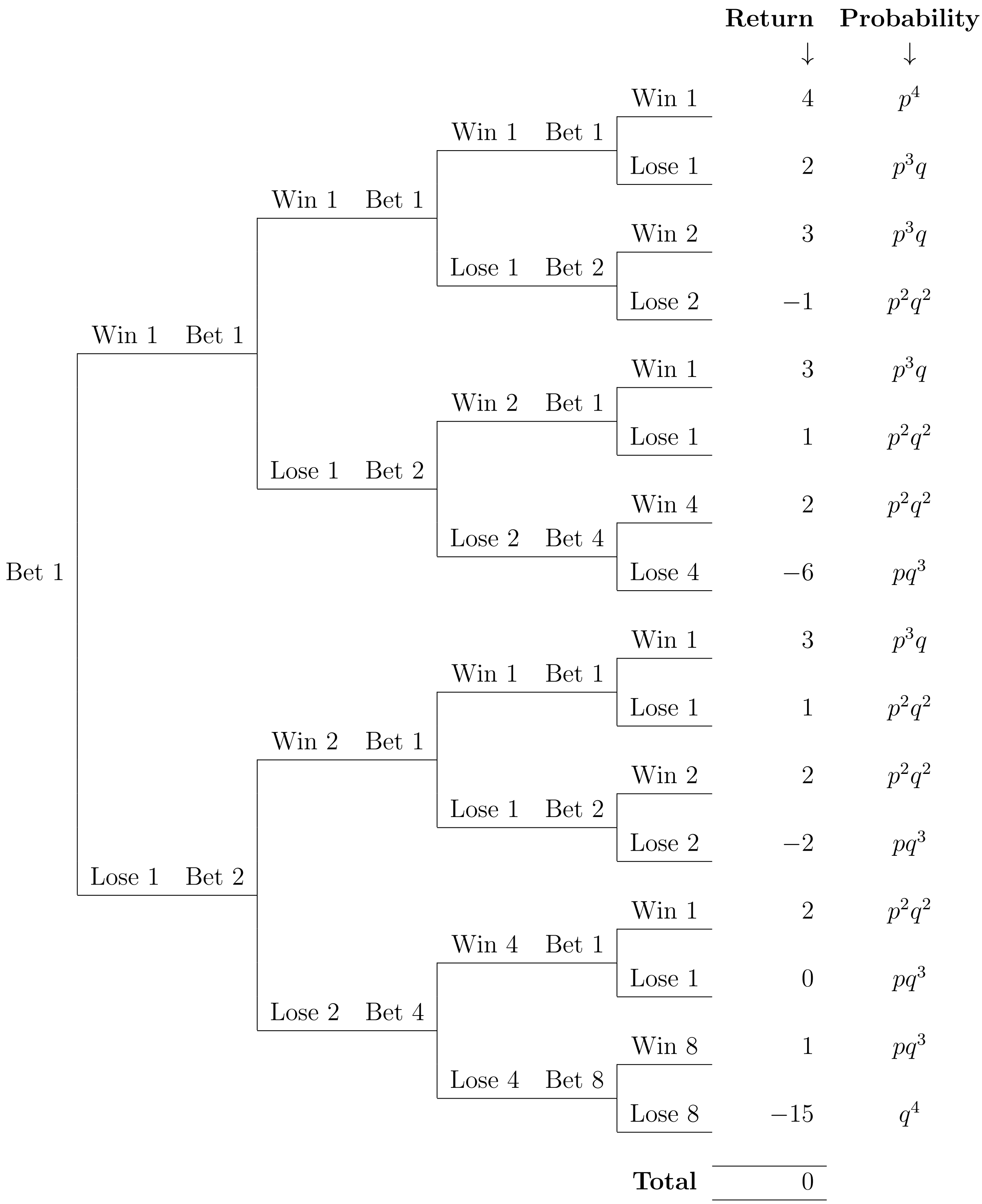
2. **(a)** Again results will vary, but the pooled class data should allow some conclusions to be drawn. Again, you should allow time for the pooled results to be discussed by the class and conclusions drawn.

A table of some results is shown below. Starting bet \$100.

Sim	**Same Amount**			**Martingale**			**Rev Martingale**		
Number	**Min**	**Max**	**Final**	**Min**	**Max**	**Final**	**Min**	**Max**	**Final**
1	88	104	92	90	149	149	50	108	50
2	84	100	88	−119	136	−119	45	131	46
3	82	104	88	79	152	152	46	127	46
4	84	102	86	−21	106	−21	43	119	43
5	96	104	98	11	150	143	52	211	52

Sim	**Labouchere**			**Rev Labouchere**		
Number	**Min**	**Max**	**Final**	**Min**	**Max**	**Final**
1	67	227	182	0	127	0
2	−19	181	−19	0	100	0
3	−2	126	−2	0	113	0
4	−1	185	−1	30	406	92
5	69	210	151	0	221	0

SYSTEM: MARTINGALE



$$E[R] = 4p^4 + 11p^3q + 7p^2q^2 - 7pq^3 - 15q^4 \approx -0.19289 \text{ if } p = 18/37.$$

$$E[B] = 4p^3 + 15p^2q + 22pq^2 + 15q^3 \approx 7.13697 \text{ if } p = 18/37.$$

$E[R]$ is -2.7% of $E[B]$, the usual margin for roulette.

SYSTEM: LABOUCHERE

Outcome	Return ↓	Probability ↓
Win 5	20	p^4
Lose 5	10	p^3q
Win 6	11	p^3q
Lose 6	-1	p^2q^2
Win 3	10	p^3q
Lose 3	4	p^2q^2
Win 9	2	p^2q^2
Lose 9	-16	pq^3
Win 3	10	p^3q
Lose 3	4	p^2q^2
Win 8	3	p^2q^2
Lose 8	-13	pq^3
Win 7	3	p^2q^2
Lose 7	-11	pq^3
Win 8	-10	pq^3
Lose 8	-26	q^4
Total	0	

$$E[R] = 20p^4 + 41p^3q + 15p^2q^2 - 50pq^3 - 26q^4 \approx -0.62126 \text{ if } p = 18/37.$$
$$E[B] = 20p^3 + 61p^2q + 76pq^2 + 26q^3 \approx 22.9866 \text{ if } p = 18/37.$$

$E[R]$ is -2.7% of $E[B]$.

Appendix

Bold Play vs Cautious Play

In their 1965 book *How to gamble if you must*, re-published in 1976 under the really catchy title *Inequalities for Stochastic Processes*, Dubins and Savage proved the following revealing theorem.

> Suppose a gambler has a fixed target to reach — say to achieve a capital of \$20,000 from an initial \$9,000 bankroll. In this scenario it is all or nothing — the player will keep playing until either he reaches the target or goes bust. What strategy maximizes the probability that he reach his goal? (We call this an optimal strategy).

Dubins and Savage proved that

- when the house has an edge, BOLD play is optimal,
- when the player has an edge, CAUTIOUS play is optimal.

By BOLD play, they mean bet your **entire bankroll** on each bet, or as much as necessary to reach your goal.

By CAUTIOUS play, they mean bet the minimum amount allowable on each bet.

For example, if you wish to get your capital to \$20 from an inital \$9, and you are betting on red-black in roulette ($P(win) = 18/37$), then adopt BOLD play (minimum bet \$1).

- Bet \$9 on your first bet. If you lose, that's it — the fat lady has sung.
- If you win, you now have \$18. Bet \$1, as this is all you need to reach your target. If you win, you pack up and go home rich. The same is true if you win at any of the following stages.
- If you lose the \$1, now bet \$2.
- If you lose the \$2, now bet \$3.
- If you lose the \$3, now bet \$4.
- If you lose the \$4, you now have \$8. Bet the \$8.
- If you win, you're there. If not, you have nothing.

Caution: Bold play doesn't guarantee a profit. Nor does it shift the edge in your favour. It only maximizes the probability you will reach your goal before going broke.

With the aid of a tree diagram, it is possible to work out the probability of losing all your bankroll. If p denotes the probability of winning a single bet, the probability of losing all the bankroll is

$$\begin{aligned} P(capital\ reaches\ zero) &= 1-p+p(1+p)(1-p)^4 \times \\ &\quad \left(1+p^2(1-p)^2+\left(p^2(1-p)^2\right)^2+\left(p^2(1-p)^2\right)^3+\dots\right) \\ &= 1-p+\frac{p(1+p)(1-p)^4}{1-p^2(1-p)^2} \qquad \left(\dots\right) \text{ is a GP with } r=p^2(1-p)^2. \end{aligned}$$

So using $p = 18/37$ for a red-black bet in roulette, we get $P(\textit{capital reaches zero}) = 0.567$, or conversely $P(\textit{capital reaches \$20}) = 0.433$.

We can show that with CAUTIOUS play, the corresponding probability of reaching the $20 target by betting $1 each time is only 0.321 (the Gambler's Ruin problem).

Gambling Systems

Gambling systems are sets of betting strategies which are often purported to offer you some edge against the house and increase your chances of winning. The truth of the matter is that you **cannot alter the house margin** in roulette, two-up or craps by using a system.[47] In fact far from helping a gambler to better profits, gambling systems have the built-in danger that they can lead people to **bet money they can't afford to lose**.

Let's now look at four such systems.

The Martingale or Double-Up System

Probably the oldest, easily the most popular, and definitely the most dangerous of all gambling systems.

Darwin Ortiz, *On Casino Gambling*, page 173

This system is considered by casino staff as the mark of a true amateur.

The method: Start with a 1-unit bet — say $1.

Double the size of the bet every time you lose.

Whenever you win, your next bet is $1 (or another option is to stop betting as soon as you score a win).

This applies to all even-money bets such as on red/black in roulette, on heads/tails in two-up and line bets in craps.

The problem with this is that you end up chasing minute winnings while risking a huge loss. Consider a table on the Las Vegas strip with a minimum bet of $2 and a maximum of $500. If you stop as soon as you win, the possible outcomes are:

win $2
lose $2, win $4, (net gain $2)
lose $2, lose $4, win $8 (net gain $2)
lose $2, lose $4, lose $8, win $16 (net gain $2)
lose $2, ... lose $16, win $32 (net gain $2)
lose $2, ... lose $32, win $64 (net gain $2)
lose $2, ... lose $64, win $128 (net gain $2)
lose $2, ... lose $128, win $256 (net gain $2)
lose $2, ... lose $128, lose $256: can't bet $512 as this exceeds the table limit (net **loss** $510)

That is, with 8 consecutive losses, you have lost $510, chasing a $2 profit! With this system, when you win you win small; when you lose, it's big-time!

[47] I'm excluding Blackjack here; a 'basic' Blackjack strategy can improve your return and a card-counting strategy can give you an edge over the house — but the latter is a very demanding, complicated strategy based on the run of cards already in play. See for example *Blackjack for Profit* by Jady Davis or, the vastly more complicated, *Professional Blackjack* by Stanford Wong. The system strategies I'm referring to in this section are not specific to any game.

The Labouchere or Cancellation System

Popularized by Henry Labouchere (1831 – 1912), an English world traveller, member of Parliament and gambler. Invented by French mathematician the Marquis de Condorcet (1743 – 1794).

The method: Write down a sequence of four numbers — e.g. 1, 2, 3, 4.

Bet the sum of the first and last numbers in the row — 5 the first time.

Each time you lose a bet, you append the amount of that bet to the end of the sequence; each time you win, cross out the first and last numbers in the sequence. Continue betting the sum of the first and last numbers in the sequence.

When you succeed in crossing out all the numbers, you have a net profit equalling the total of your original sequence of numbers.

To illustrate this system, let's consider some possible outcomes, starting with the sequence 1, 2, 3, 4.

Outcome	**Number Sequence**	**Bet**	**Result**	**Accummulated Gain/Loss**
1	1234	5	Win 5	+ 5
	23	5	Win 5	+ 10
2	1234	5	Lose 5	− 5
	12345	6	Lose 6	− 11
	123456	7	Win 7	− 4
	2345	7	Win 7	+ 3
	34	7	Win 7	+ 10

In both cases, we do in fact end up winning $10.

The problem with the system, however, is that, like the Martingale system, it can end up winning a small amount of money for the player very often, but occasionally losing a huge amount which outweighs the winnings.

A computer simulation of 160,000 spins of the roulette wheel found the following:

Player wins the target of $30 twenty out of twenty-one times played.

Player loses an average of $1012 once in twenty-one times played.

This gives a net loss of $412.

The system works 95.2% of the time, BUT STILL LOSES MONEY.

The Reverse Martingale System

This is a mirror image of the Martingale System.

The method: Start with a 1-unit bet — say $1.

Double the size of the bet every time you WIN.

Whenever you lose, your next bet is $1.

The comparison between the Martingale system and the Reverse Martingale can be best expressed as follows:

The Martingale system is an example of a multiplicative system — where the wagers rise rapidly with increasing losses; it offers a high probability of a small win and a small probability of a large loss.

The Reverse Martingale system has the wagers rising rapidly with increasing winnings; it offers a high probability of a small loss and a small probability of a large win.

The Reverse Labouchere System

This proceeds as for the Labouchere, except that every **win** results in appending that bet to the sequence and every **loss** results in cancelling the first and last numbers in the sequence. Like the Reverse Martingale, it offers a high probability of a small loss and a small probability of a large win. Two possible outcomes follow.

Outcome	Number Sequence	Bet	Result	Accummulated Gain/Loss
1	1234	5	−5	−5
	23	5	−5	−10
2	1234	5	−5	−5
	23	5	+5	0
	235	7	−7	−7
	3	3	−3	−10

Final Notes

Systems thrive despite their drawbacks for one reason: because people who try them find they are winning most of the time (e.g. 95% of the time!), they declare the system works. Either they haven't yet hit the wall with a bad session or they have only assessed their performance in terms of percentage wins instead of overall gain/loss.

The Reverse Labouchere is the system reputedly used by Norman Leigh's syndicate — *Thirteen against the bank* — with great success in the European casinos. Without having read the book, I expect the syndicate must have been consistent 'small' losers before their big wins.

2.21 Kangaroo Management

Adapted from *Stimulating Mathematical Interest with Dynamical Systems* by M.B. Durkin, The Maths Teacher 89, 242–24 (1996).

Aim

The main purpose of this lab is to help you appreciate the way in which discrete non-linear dynamics can be applied to problems in the 'real world'.

Introduction

You are hired by the State Forestry Department, with your main task to assist in the management of the kangaroo population in a remote forest called Hamt Reserve. The possibility of culling of kangaroos in the reserve is under consideration.

The kangaroo population in the reserve is given by the Discrete Logistic Model, a difference equation,

$$P_{n+1} = 1.8P_n - 0.8(P_n)^2, \tag{1}$$

where P_n is the number of kangaroos in the reserve at the end of year n in tens of thousands, i.e. *one unit of P equals 10, 000 kangaroos.* At the end of 2005, there were 8000 kangaroos in the reserve ($P_0 = 0.8$).

The first task

As a training exercise, management asks you to model and report on a scenario containing several events that would affect the kangaroo population.

Write a short report on the outcome of the following scenario. The report should include a mathematical analysis with calculations, tables and/or graphs to substantiate your conclusions.

The scenario

- If there were no natural disasters in 2006, what would the kangaroo population be at the end of 2006? Do this and the following calculations manually (without a program) using Eq. (1).[48]
- Unfortunately, at the end of 2006, there was a short but fatal outbreak of the dreaded rootoxis which kills around 4000 kangaroos. What would the population of kangaroos be at the end of 2007? When would the kangaroo population recover to more than 9000 kangaroos if there were no more natural disasters?
- Following the rootoxis epidemic, on Christmas Day 2008 there was a forest fire in a nearby forest which resulted in 2000 kangaroos from that forest migrating into Hamt Reserve. What would the population of kangaroos in Hamt Reserve be at the end of 2009?
- After these two events, there were no more natural disasters. What would the kangaroo population be after a long time? The number here is the limiting capacity or maximum sustainable population of the reserve.

[48] *Calculator hint*: Store the initial population in memory P and repeatedly execute the calculation $1.8\text{P} - 0.8\text{P}^2 \to \text{P}$ by pressing [enter] the required number of times. Make sure you understand why this works.

Effect of culling

Impressed by your previous report, management has now put you in charge of undertaking a feasibility study into whether culling of kangaroos is necessary/desirable in Hamt Reserve. Your analysis will be a crucial factor in the decision-making process.

Write a report addressing the following questions. Again, a mathematical analysis including calculations, tables and/or graphs is required to substantiate your conclusions. Add an executive summary for your boss, summarising your findings and making suitable recommendations.

1. What is the modified form of Eq. (1) if H kangaroo units are culled each year?

   We assume here, for simplicity, that all the kangaroos are killed close to the end of the year, otherwise the killing of the female kangaroos in particular would affect the number of births and deaths, and consequently the growth rate.

2. What would happen if 720 kangaroos were culled each year ($H = 0.072$), a value used in a nearby reserve? Assume the initial population is that given above for the year 2005, $P_0 = 0.8$. What is the long-term population?

   What if the initial population were $P_0 = 0.3$? $P_0 = 0.095$?[49]

3. What would happen if 2400 kangaroos were culled each year ($H = 0.24$)? Assume again that $P_0 = 0.8$. What is the long-term population?

   What if the initial population were $P_0 = 1$? $P_0 = 1.5$?

4. What about $H = 0.2$? It turns out[50] that this is the largest number of kangaroos which could be culled annually without the kangaroos dying out in Hamt Reserve. Note that the initial population must be larger than 0.5. What is the long-term population in this case?

LOGISTIC/LGSTCE

Use: Run the program. Select KANGAROOS. Input the appropriate parameters at the prompts ($0 < \text{u}(0) < 1$ for the bacteria). The program plots population versus cycle number (time). Use the arrow keys to trace the graph or press enter to return to the main menu.

When the program has finished, choose QUIT from the main menu. Here you can either keep the equations for manual plotting (e.g. with a different window; Option 1) or you can have the equations and other settings deleted (Option 2). If you choose Option 1, when you have finished rerun the program, QUIT and select Option 2 to tidy up.

[49]The LOGISTIC/LGSTCE program (available at *www.YYY*) might help here.
[50]Experiment and see

Instructors' Guide

The calculation of the population P_n can be done manually on a calculator[51] or by using the built-in sequence grapher, depending on the sophistication of your students and the type of calculator they have. To do the time plots, it is desirable to use the sequence grapher.

The LOGISTIC/LGSTCE program (available at *www.YYY*) sets up the sequence grapher for the problem here.

Solutions

We have the logistic difference equation for the kangaroo population[52]

$$P_{n+1} = 1.8P_n - 0.8(P_n)^2,$$

with $P_0 = 0.8$ corresponding to the (end of) year 2005.

Using this and incorporating the rootoxis outbreak in 2006 by subtracting 0.4 (4000 kangaroos) from the 2006 population, we have the following number of kangaroos in subsequent years.

Year	n	P_n	**Number of kangaroos**
2005	0	0.8	8000
2006	1	0.928−0.4=0.528	5280
2007	2	0.7274	7274
2008	3	0.8860	8860
2009	4	0.9668	9668

The number of kangaroos has recovered to 9668 by the end of the year 2009.

If we include the migration of 2000 kangaroos at the end of 2008, we have the following numbers.

Year	n	P_n	**Number of kangaroos**
2008	3	0.8860+0.2=1.0860	10, 860
2009	4	1.0113	10, 113
2010	5	1.0022	10, 022
2011	6	1.0004	10, 004
2012	7	1.0001	10, 001
2013	8	1.0000	10, 000

The population in the reserve at the end of 2009 would be 10, 113. In subsequent years, the population declines to the carrying capacity or maximum sustainable population of 10, 000, the population after a long time.

[51] Faster if you store the initial population in memory P and repeatedly execute the calculation $1.8\text{P} - 0.8\text{P}^2 \rightarrow \text{P}$ by pressing [enter] the required number of times.

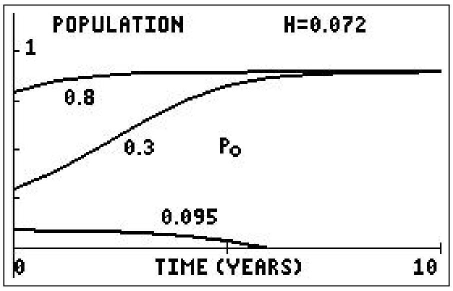
[52] The sequence grapher on the TI-84Plus writes P_n in terms of P_{n-1}, so it is necessary to rewrite the difference equation as $P_n = 1.8P_{n-1} - 0.8(P_{n-1})^2$ if you use this method. The CE does both.

Effect of culling

1. If H kangaroo units are killed each year, this number is subtracted from the value for P_{n+1} that we calculated above, giving the difference equation

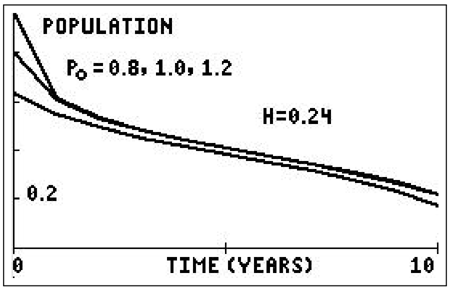
$$P_{n+1} = 1.8P_n - 0.8(P_n)^2 - H.$$

2. With $H = 0.072$ and an initial population of 0.8 units, the long-term population will be 0.9 units or 9000 kangaroos.

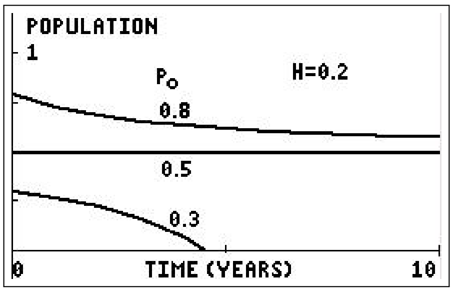


We find that[53] if the initial population is greater than 0.1 kangaroo units, the population will tend toward a stable value of 0.9. If the initial population is less than 0.1 kangaroo units, the population will tend to 0.

3. With $H = 0.24$, the population will die out, no matter what the initial population.



4. With $H = 0.2$, the long-term population will be 0.5 units or 5000 kangaroos, the maximum sustainable population with this level of hunting, provided that the initial population is greater than 5000. If the initial population is less than 5000, the population will die out.



If this level of hunting were chosen, any natural disaster that killed more than a few kangaroos after the population had levelled off at 5000 would bring the population to below 5000, and it would therefore die out; there is no margin for error. In practice, a smaller value than 2000 would be chosen for the number of kangaroos killed annually, thereby leaving a margin to allow for natural disasters.

[53]Theory helps a lot here, but you can reach the same conclusions by experimenting with numbers on your calculator. Using the LOGISTIC/LGSTCE program should help with this.

2.22 Epidemic

The population on an idyllic Pacific island is stable at $N = 200$. The daily birth rate per capita $b = 0.00001$ is balanced by an equal death rate. The island's post office employs five people[54] who are all clustered round when a Christmas hamper is opened for customs inspection. Unfortunately, the white powder in the hamper is not artificial snow. Within a short time, all five workers are infected with a mysterious bacterium, and the resulting severe disease starts to spread to others.

You are the only person on the island with some knowledge of epidemiology, gained from Maths in Year 12. The island's Chief Medical Officer needs to know how many people might need to be treated for the virus and whether to call for emergency hospital facilities. The island's hospital can cope with at most 50 patients at any one time. The CMO asks for your help in predicting the course of the epidemic.

Based on the post-office experience, you assume that the incubation period is less than 1 day. If the mean infectious period is $1/\alpha$, where α has to be guessed/estimated, a possible set of difference equations governing the spread of such a disease is

$$S_{n+1} = S_n - \beta S_n I_n + bN - bS_n \quad (1)$$

$$I_{n+1} = I_n + \beta S_n I_n - bI_n - \alpha I_n \quad (2)$$

$$R_{n+1} = R_n - bR_n + \alpha I_n, \quad (3)$$

where S_n, I_n and R_n are, respectively, the number of susceptible persons, number of infected persons and number of recovered persons after n time intervals and β is a constant to be determined. The time interval is 1 day.

Question 1 *The initial model*

(a) Show that $S+I+R$ is a constant.
Hint Use mathematical induction: show that $S_{n+1}+I_{n+1}+R_{n+1} = S_n+I_n+R_n$.

If this is true, we only need to calculate S_n and I_n; R_n is given by $R_n = N - S_n - I_n$.

(b) Assuming that the terms involving births and deaths can be neglected because of the relatively short duration of the epidemic, write down the two simplified equations for S and I (the model).

(c) For your initial modelling, you need to estimate α and β. Based on other similar diseases, you take the mean infectious period $1/\alpha$ to be 5 days, so that $\alpha = 0.2$, and $\beta = 0.0025$. These values will have to be reviewed after a few days, once you have some data on the course of the disease on the island.

- How will the disease run its course according to your model? Plot S_n and I_n against n to find out, with $S_0 = 195$ and $I_0 = 5$. Details on how to do this are given below.
- How many days will it take before the disease dies out ($I < 1$)?
- When will the peak of infection occur?
- How many people avoid catching the disease?

[54]The island's economy is based around issuing stamps, banking and processing asylum seekers.

- From the simplified equation for I_{n+1} above, the maximum value of I occurs ($I_{n+1} = I_n$) when $S = \alpha/\beta$. Compare this value with the value you obtained from your graph. Why might they not be exactly the same?

Question 2 *The improved model*

After a week, you have some actual data on your epidemic. You find that the mean infectious period is actually 4 days, not 5, and that after a week there are 30 infected persons, that is $I_8 = 30$.

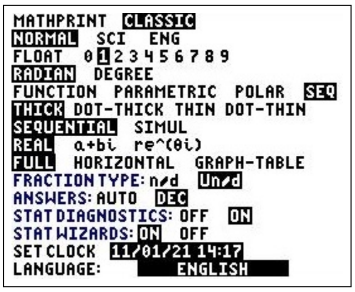
(a) Experiment with different values of β to find its true value for the island epidemic — for each β value, use either a graph and trace or table (easier) to find I_8 until you match the actual value (or as close to it as possible).

(b) When you have found the best β value, sketch the graphs of S_n and I_n against n and use them to predict, hopefully more accurately than from the initial model, how many days it will take before the disease dies out ($I < 1$) and when the peak of infection will occur. Will the hospital be able to cope? How many people will avoid catching the disease?

Sequence Graphing

Population Modelling 2: Logistic and Epidemic Models[55] has an introduction to discrete models, in particular the Discrete Logistic Model.

The TI-84/CE calculator has three built-in sequences u, v and w (on the 7, 8 and 9 keys). To access them, press mode; select SEQ (fifth line on the CE, fourth on the 84) with the cursor and press enter. Select 1 decimal place as well (third/second line).



Now press y= and you will see where to define the sequence functions.

TI-84Plus: The sequences $u(n)$ and $v(n)$ are defined in terms of $u(n-1)$, $v(n-1)$, etc. The simplified equations for S and I written in this form are

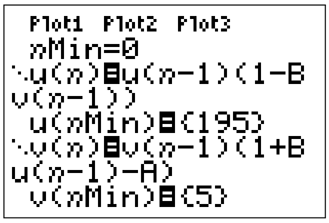
$$S_n = S_{n-1} - \beta S_{n-1} I_{n-1} = S_{n-1}\big(1 - \beta I_{n-1}\big).$$
$$I_n = I_{n-1} + \beta S_{n-1} I_{n-1} - \alpha I_{n-1} = I_{n-1}\big(1 + \beta S_{n-1} - \alpha\big).$$

With $S_n \to u(n)$, $I_n \to v(n)$, $\alpha \to A$ and $\beta \to B$, the equations in calculator variables are

$$u(n) = u(n-1)\big(1 - Bv(n-1)\big).$$
$$v(n) = v(n-1)\big(1 + Bu(n-1) - A\big).$$

[55] in Volume 2 of *Mathematics on a TI-84/CE*, available at *www.XXX*

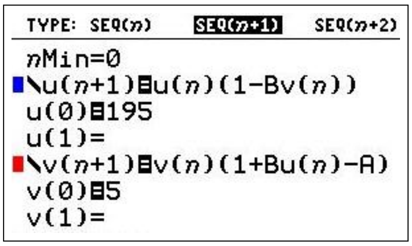
Enter these equations into the calculator as shown below. Set nMin=0, that is we start with S_0 and I_0. Then S_1 is the S value after 1 day, etc. The initial conditions are contained in u(nMin) and v(nMin); note the curly brackets here.



In SEQ mode, the X,T,θ,n key now gives n, the independent variable for the sequence functions. You can't use the letter N.

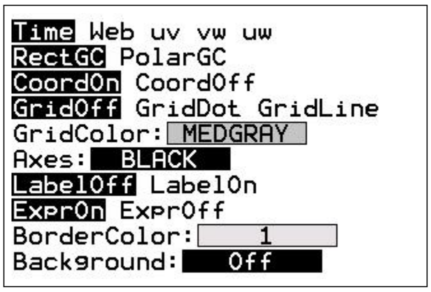
TI-84CE: The sequences can be input in the original simplified form (note the top line in the figure). Set nMin=0, that is we start with S_0 and I_0. Then S_1 is the S value after 1 day, etc. The initial conditions are contained in u(0) and v(0).

$$
\begin{aligned}
u(n+1) &= u(n)\big(1 - Bv(n)\big).\\
v(n+1) &= v(n)\big(1 + Bu(n) - A\big).
\end{aligned}
$$



Time plots

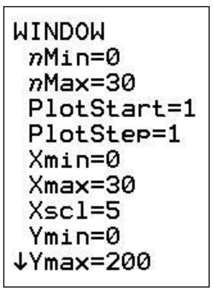
Press 2nd format and select *Time* as shown below.[56] The X axis is n (time in days), the Y axis S and I.



The final step before plotting is to choose a window. As well as the usual window settings for the X and Y axes, we have to specify nMax, the maximum n value we want.

Start by plotting 30 points, corresponding to running the system through 30 days. Note that we must also set Xmax to 30.

As $N = 200$ is the maximum value for both S and I, setting Ymax to 200 seems like a good starting point. Yscl = 50.



[56] uv would give a phase plot

Store the values of α and β in memories A and B respectively. Press [graph] to plot graphs of S and I versus n.

Adjust the [window], if necessary, until the graphs more or less fill the screen.

Use [trace] to explore the values. You can go directly to the point with a particular n value by just typing in the n value and pressing [enter].

Adjust nMax and Xmax (both corresponding to Tmax) in [window] so that the disease runs its full course ($I < 1$).

You can also see the values of S and I in a table by pressing [table] ([2nd] [graph]).
If the n values in the table don't start at 1 and/or don't increment in steps of 1, fix this in [tblset] ([2nd] [window]).

Instructors' Guide

The calculations for this lab can be done by hand (using an ordinary calculator), but the availability of sequence plotting makes life a lot easier and more visual.

1. (a) We are given $S_0 = 195$, $I_0 = 5$ and, as the recovery period is of the order of days, $R_0 = 0$. Therefore $S_0 + I_0 + R_0 = N = 200$.

Assume that $S_n + I_n + R_n = N$. Adding the three given equations gives

$$\begin{aligned} S_{n+1} + I_{n+1} + R_{n+1} &= S_n + I_n + R_n + b\big(N - (S_n + I_n + R_n)\big) \\ &= N \qquad \text{by our assumption.} \end{aligned}$$

Therefore, by induction, $S + I + R = N$ and is constant.

(b) The two simplified equations are obtained by omitting any terms containing b.

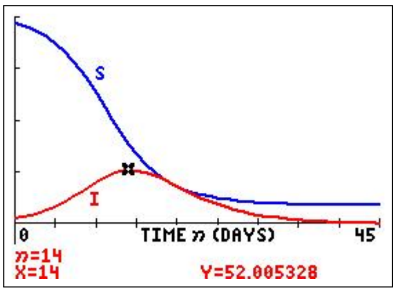
$$\begin{aligned} S_{n+1} &= S_n - \beta S_n I_n \\ I_{n+1} &= I_n + \beta S_n I_n - \alpha I_n \;=\; (1-\alpha) I_n + \beta S_n I_n \end{aligned}$$

To write these for the TI-84, replace n by $n - 1$.

(c) $\alpha = 0.2$, $\beta = 0.0025$. A suitable window is (Yscl = 50)

```
WINDOW
 nMin=0
 nMax=45
 PlotStart=1
 PlotStep=1
 Xmin=0
 Xmax=45
 Xscl=5
 Ymin=0
↓Ymax=200
```

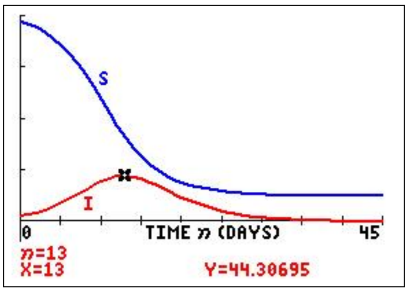
with the resulting graph of I and S vs time in days:



For these values of α and β (using trace on the calculator graph),

- the disease runs for 44 days ($I_{44} < 1$)
- the maximum I, the peak of infection, of 52 occurs at $n = 14$, i.e. after 14 days.
- the value of S at I_{max} is 76.9, close to the theoretical value $\alpha/\beta = 80$.

2. A mean infectious period $1/\alpha = 4$ gives $\alpha = 0.25$. We are given $I_8 = 30$. The value of $\beta = 0.00281$ gives $I_8 = 30$.



For these values of α and β (again using trace):

- the disease runs for 37 days ($I_{37} < 1$);
- the maximum I of 44 occurs at $n = 13$, i.e. after 13 days; the corresponding value of S, $S = 81.5$, is again close to the theoretical value $\alpha/\beta = 89$ (after 12 days, $S = 93$);
- the hospital therefore copes;
- 25 persons (S_{37}) avoid catching the disease.

3 Labs for which a TI-84/CE is useful

3.1 Siting a School — An Exercise in Mathematical Thinking

Aims

- To promote the appreciation and use of mathematical thinking and its applications.
- To illustrate themes in introductory Linear Algebra
  - the algebra-geometry link
  - how maths responds to practical problems
  - how the methods and problems of maths are extended.
- To practise a little using vectors.
- To learn about a famous problem.

Procedure

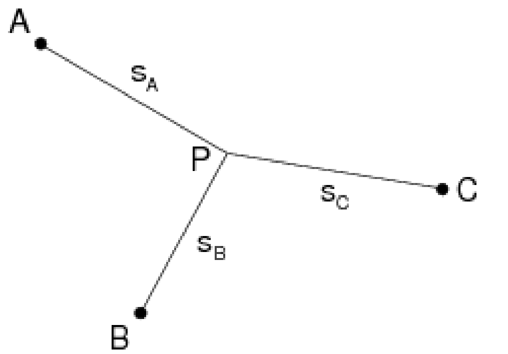
This lab will lead you through a series of questions and problems. At each stage you should puzzle over the material and discuss it as a group. Then all help the scribe write up your proposals, arguments and solutions before you move on to the next stage.

The general problem

Three villages A, B and C each have 30 children. They decide to build a school at P for all the children to attend.

Where should the school be placed so that the shortest total length of road is built?

Mathematically: given points A, B, C find the point P so that the total distance $s = s_A + s_B + s_C$ is a minimum.



Question 1 *Preliminary thoughts*

Discuss how you might go about solving or exploring the problem, without going into the details. Strategies?

What things might the point P depend on? e.g. size — if we magnify up the triangle ABC, does P move?

Would a rule for finding P depend on the shape of the triangle ABC?

What makes this problem hard? easy?

Stop! Write up your thoughts before going on.

This problem has a long history. P is sometimes called the Fermat point because we know that Fermat (1601 – 1 665) posed the problem like this:

> Given a triangle ABC, find the point P so that s, the sum of the distances from P to the corners, is a minimum,

and Torricelli (1608 – 1647) solved it (as well as inventing the barometer).

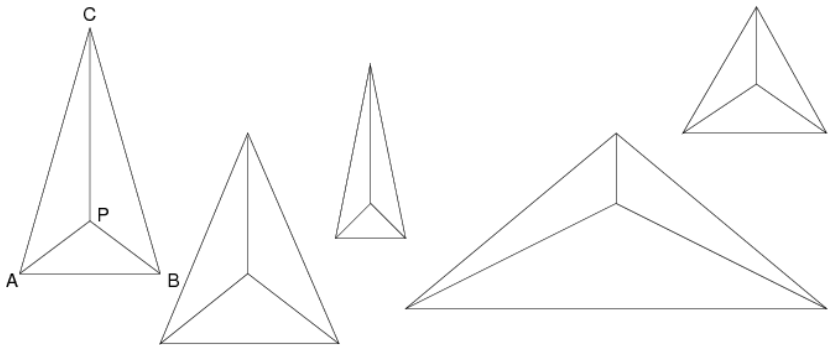
I think it is a hard problem because the point can move around in two dimensions — in modern terms, we have to find two coordinates.

Question 2 *Finding a conjecture*

Mathematicians often explore a difficult problem in different ways to come up with a guess at the answer, or conjecture, which they then try to prove correct.

Idea 1 *Look at some simple special cases*

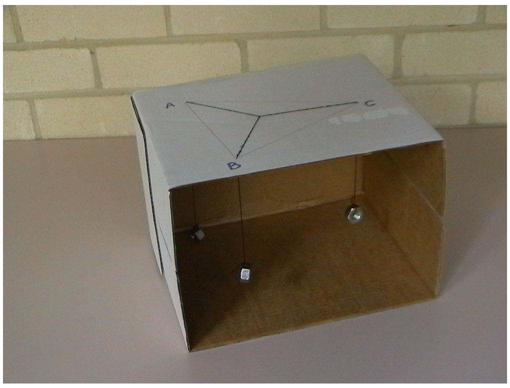
If we make ABC an **isosceles** triangle with $AC = BC$, symmetry suggests (correctly) that P is on the line joining C to the mid-point of AB. Finding P is now a straightforward one-variable Calculus problem. Here are some examples for you to observe.



Idea 2 *Use Archimedes' approach*

In his book *The Method*, Archimedes described how he used mechanical problems and devices to discover mathematical results (the beginning of the analogue-computer idea). For our problem we can use the Varignon Frame, invented by Pierre Varignon (1654–1722), which works like this:

On a smooth horizontal surface, holes are made at the triangle corners A, B and C. A string goes through each hole. On top of the surface, the three strings are knotted together at K. Below the surface, the strings hang down with equal weights tied at the end of each one.



If the arrangement is disturbed (tapped gently), the device will always come back to an equilibrium; **the position of K then gives the Fermat point P**.

We come to the reason later, but for now, look at the models and observe what happens.

Action: Use your observations from Ideas 1 and 2 to make a guess or conjecture about the General Problem (as on page 179) where no special triangle shape is assumed. What characterises the point P?

Stop! Write up your ideas before going on.

I hope you found that the angles θ_A, θ_B and θ_C made at P were always equal to $120°$.

Perhaps you wondered what happens if an angle in the triangle becomes bigger than $120°$. A very clever or lucky guess would have given this (correct) conjecture.

For any triangle:

- if no angle in the triangle is $\geqslant 120°$, the required point P is the one inside the triangle making $\theta_A = \theta_B = \theta_C = 120°$;
- if an angle is $\geqslant 120°$, point P is located at the corresponding corner.

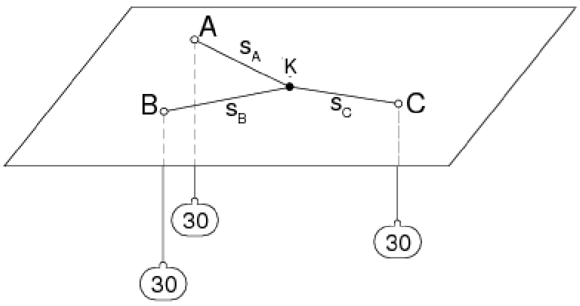
Remarkably this result does not depend on the size or shape of the triangle.

Our next job is to PROVE that the conjecture is correct. We will concentrate on the interesting case where no angle is greater than or equal to $120°$.

We could look for a proof of the mathematical problem as stated on page 179. However, it is interesting to consider the Varignon Frame some more: why does it work and why does it give $120°$ angles?

Question 3

Imagine a smooth horizontal surface with holes drilled in it at the points A, B and C of the triangle you are interested in: an ideal Varignon Frame. Through each hole there are strings of lengths L_A, L_B and L_C, tied together in a common point K.



On the end of each string, under the surface, is tied a weight of say 30 grams (as there are 30 children in each of the original villages).

Use the following argument to prove that the point K moves to the required minimizing point P. Given

- the potential energy of each weight is $-30 \times g \times$ *distance below the surface*, where g is the gravitational constant (**note the minus**);
- this mechanical system comes to a resting or equilibrium state when the **total** potential energy E is a **MINIMUM**.

Method: Write down a formula for E in terms of the string lengths L_A, L_B and L_C, and the lengths s_A, s_B and s_C on the surface. Remember L_A, L_B and L_C are fixed, but s_A, s_B and s_C can vary. Relate $s = s_A + s_B + s_C$ to the minimizing of E, and so explain why the mechanical method works.

Stop! Work as a group! Write up your ideas.

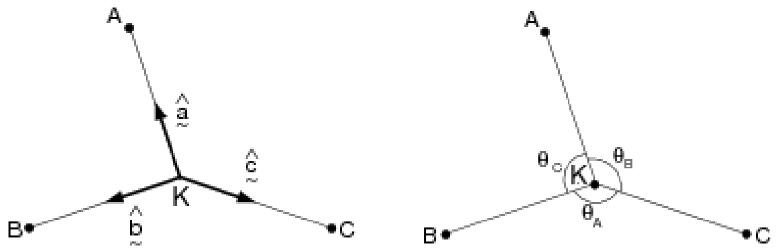
The Varignon Frame DOES give the correct angles — now to a general proof.

Question 4

Think of the knot where the strings meet at K as a little particle. The tension due to the equal weights pulls on the particle with forces of equal magnitude T along each string. When the system is in equilibrium there is no nett force acting on the particle.

That means the TOTAL force acting in any direction is zero.

As in the diagram below, define unit vectors $\hat{\boldsymbol{a}}$, $\hat{\boldsymbol{b}}$, $\hat{\boldsymbol{c}}$ pointing from K to A, B and C respectively.



Write the total force vector acting at K in terms of T and those unit vectors.

Now get an equation involving the unit vectors by saying that total force = $\mathbf{0}$.

That equation tells us about the angles. There are two options.

Algebraic: Take the dot product of the equation with $\hat{\boldsymbol{a}}$, then with $\hat{\boldsymbol{b}}$, then with $\hat{\boldsymbol{c}}$. You should now have three equations involving $\cos(\theta_A)$, $\cos(\theta_B)$ and $\cos(\theta_C)$. Use them to deduce each angle must be 120°,
or
Geometric: Draw a diagram to show how the vectors add to $\mathbf{0}$. What sort of a triangle do you get? Does that give information about angles? So what must θ_A, θ_B and θ_C be?

Careful write-up now done?

Question 5

Write a few comments about this lab and what things you think it might have illustrated for you or taught you.

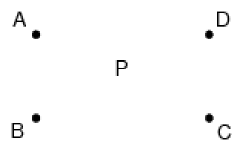
What mathematical questions remain? How does the work extend?

If you have breezed through the lab, you could consider one of these extensions.

1. How would you change the problem if the 90 children were distributed as 10, 30 and 50 over the three villages. Minimize something else? The total distances travelled by all the children? But how to do it?

2. Suppose there were four villages forming a rectangle.

   How to place the school and the roads?
   Could your result for the triangle case help?



This lab is a tiny introduction to an enormous and important topic in Applied Mathematics dealing with optimised-network planning.

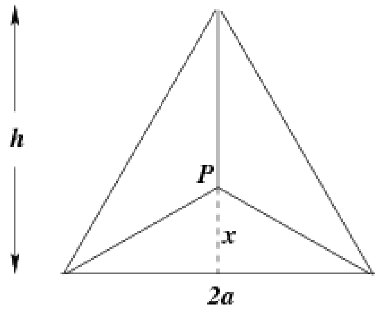
Instructors' Guide

The Varignon frame, made from a cardboard box as shown in the photograph, was taken to the lab for students to experiment with while pondering Question 2.

Solutions

2. Although not required for this lab (it could be), it is useful to show that, for an isosceles triangle at least, s is minimised when the angles between the strings are 120°.

Consider the isosceles triangle shown below.



The total length s is given by

$$s(x) = h - x + 2\sqrt{a^2+x^2}, \quad \text{using symmetry and Pythagoras,}$$

with $0 \leqslant x \leqslant h$. This is a continuous function on a closed domain, so the global minimum of s must exist and lies either at a critical point of s or at an endpoint of the domain.

$$\frac{ds}{dx} = -1 + \frac{2x}{\sqrt{a^2+x^2}},$$

which is defined for all x. Any critical points must therefore occur where the derivative is zero.

$$\begin{aligned} \frac{ds}{dx} = 0 \ &\Rightarrow \ \frac{2x}{\sqrt{a^2+x^2}} = 1 \\ &\Rightarrow \ 2x = \sqrt{a^2+x^2} \\ &\Rightarrow \ 4x^2 = a^2 + x^2 \\ &\Rightarrow \ 3x^2 = a^2 \\ &\Rightarrow \ x_{cr} = \frac{a}{\sqrt{3}} \qquad x \geqslant 0. \end{aligned}$$

There is only one critical point, $x_{cr} = a/\sqrt{3}$.

Now $s'' = 2a^2/(a^2+x^2)^2$ is always positive, so the function s is everywhere concave up. As s is continuous and has only one critical point: that point must be a global minimum.

$\tan(\theta/2) = a/x_{cr} = a/(a/\sqrt{3}) = \sqrt{3}$. Therefore, $\theta/2 = 60°$, so that $\theta = 120°$.

Therefore, at this global minimum, the angles between the three lines at P are each 120°.

3. The potential energy of the system is given by

$$\begin{aligned} E &= -30g(L_A - s_A) - 30g(L_B - s_B) - 30g(L_C - s_C) \\ &= -30g(L_A + L_B + L_C) + 30g(s_A + s_B + s_C) \\ &= -30g(L_A + L_B + L_C) + 30gs. \end{aligned}$$

Therefore E is a minimum (and the system comes to its equilibrium state) when s is a minimum, as we require in our problem.

4. The total force acting at K is zero, so that

$$T\hat{\boldsymbol{a}} + T\hat{\boldsymbol{b}} + T\hat{\boldsymbol{c}} = \mathbf{0},$$

giving, as $T \neq 0$,

$$\hat{\boldsymbol{a}} + \hat{\boldsymbol{b}} + \hat{\boldsymbol{c}} = \mathbf{0}.$$

Taking the dot product of this equation with $\hat{\boldsymbol{a}}$, $\hat{\boldsymbol{b}}$ and $\hat{\boldsymbol{c}}$ respectively, and using the results $\hat{\boldsymbol{a}} \cdot \hat{\boldsymbol{b}} = \cos(\theta_C)$, $\hat{\boldsymbol{a}} \cdot \hat{\boldsymbol{c}} = \cos(\theta_B)$ and $\hat{\boldsymbol{c}} \cdot \hat{\boldsymbol{b}} = \cos(\theta_A)$ gives us the 3 equations

$$\begin{aligned} 1 + \cos(\theta_C) + \cos(\theta_B) &= 0 \\ \cos(\theta_C) + 1 + \cos(\theta_A) &= 0 \\ \cos(\theta_B) + \cos(\theta_A) + 1 &= 0. \end{aligned}$$

Solving these equations gives

$$\cos(\theta_A) = \cos(\theta_B) = \cos(\theta_C) = -\frac{1}{2},$$

that is, all angles are 120°.

Geometrically, because the force vectors add to zero, the vector diagram gives a triangle. The triangle is an equilateral triangle because all sides are the same length, T. The internal angles of the triangle are then each 60°, with the angles θ, the supplements of the internal angles, each 120°.

3.2 Messing about (with Vectors) in Boats

Aims

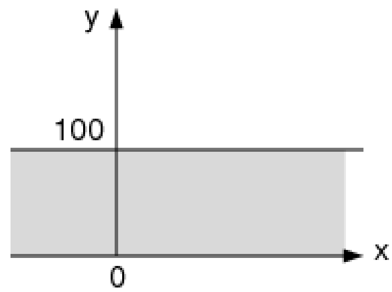
- To practise using vectors.
- To see how vectors are applied physically.
- To learn how measurements lead to "inverse problems".
- To see the link between "the number of unknowns" and the "pieces of information available".

> **Procedure**
> In this lab you will be led through a series of exercises and questions. At each stage you should record your work on the exercises and your answers to the questions. If you feel uncertain, check the answer with a lab instructor before you go on — but do have a good ponder over things before you rush to the answers. Ask for help if you cannot get answers.

This lab is about boats and their speed and the speed of water; the ideas are exactly the same when applied to planes, air speeds and wind speeds, but I thought you might visualise the boats better.

Question 1 *The situation*

A river is 100 m wide and the boat begins at the origin using these axes ...
Assume the river has the same speed w everywhere, so the water velocity is $\boldsymbol{w} = w\boldsymbol{i}$.



Relative to the **water**, the boat has velocity $\boldsymbol{b}$ and speed $b = ||\boldsymbol{b}||$.
Relative to the **ground** the boat has velocity $\boldsymbol{v}$, speed $v = ||\boldsymbol{v}||$.

Let's begin by checking the basics. The boat starts from the origin 0.

(a) If $b = 40\,\text{m/min}$, the river is still ($w = 0$) and the boat goes directly across the river, what is **(i)** the vector $\boldsymbol{b}$; **(ii)** the vector $\boldsymbol{v}$; **(iii)** the time taken to cross the river? Write the vectors in terms of $\boldsymbol{i}$ and $\boldsymbol{j}$.

(b) If $b = 0$ and $w = 30\,\text{m/min}$, what is $\boldsymbol{v}$ and where is the boat after 2 minutes? Don't forget to write down the explanation for your answer.

Now we go on to the general case where $\boldsymbol{b}$ and $\boldsymbol{w}$ are both non-zero. The velocity $\boldsymbol{v}$ relative to the ground depends on both $\boldsymbol{b}$ and $\boldsymbol{w}$.

> If you find trouble with this, try a little experiment: let the table top be the "ground" and use a piece of paper as the "river"; then draw a line on the paper as the boat moves with velocity $\boldsymbol{b}$ and someone can move the paper to simulate the river velocity $\boldsymbol{w}$. Try $\boldsymbol{w} = \boldsymbol{0}$, then non-zero.

Question 2

(a) What is the general formula for $\boldsymbol{v}$ in terms of $\boldsymbol{b}$ and $\boldsymbol{w}$?

(b) As an example, suppose the boat is headed across the river with $\boldsymbol{b} = 40\boldsymbol{j}$ and the river flows with speed 30 m/min.

What are the velocity $\boldsymbol{v}$ and speed v?

How long does the boat take to cross the river?

Where does the boat reach the other side?

How far has it travelled?

The above problems are typical "direct problems" and they are the ones we most commonly solve. You were given velocities and asked what happened to the boat. Inverse problems start with what happened and then ask what was going on to get there. This is actually very common in science and engineering.

An example

Direct problem: Given the thermal expansion properties of mercury and the temperature, how high does a column of mercury rise up a tube.

Inverse problem: Given the height of mercury in this tube (thermometer), what is the temperature? (*Of course we solve that one by calibration.*)

Inverse problems can be very tricky — do we have **enough** information to figure out what is going on? how accurate must our information be? and so on. We now explore those things using our river and boat example.

Question 3 *An inverse problem*

(a) Case 1

If $b = 20$ m/min, $w = 10$ m/min and the boat is steered so that it actually travels straight across the river (so that $\boldsymbol{v} = v\boldsymbol{j}$), what is the angle α between $\boldsymbol{b}$ and $\boldsymbol{v}$?
Draw a vector diagram.

(b) Case 2

The boat leaves the origin and reaches the other bank at $x = 75$ m, $y = 100$ m after 5/3 min.

How far did the boat travel?

What was the speed v?

Find $\boldsymbol{v}$ in terms of $\boldsymbol{i}$ and $\boldsymbol{j}$.

Hint: Write $\boldsymbol{v} = v\cos(\theta)\boldsymbol{i} + v\sin(\theta)\boldsymbol{j}$ and look at the boat path to get $\cos(\theta)$ and $\sin(\theta)$. Be careful to label the correct angle as θ.

(c) Let's continue with Case 2. You have found $\boldsymbol{v}$, and I now tell you that the river speed is $w = 34$ m/min.

What is the boat velocity $\boldsymbol{b}$?

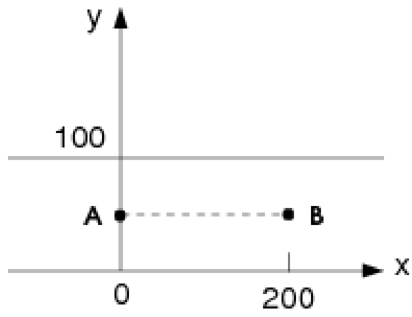
What is the boat speed b?

Draw a picture to show what is happening, but you will probably find it easier to calculate $\boldsymbol{b}$ by letting $\boldsymbol{b}=b_1\boldsymbol{i}+b_2\boldsymbol{j}$ and use the algebraic equation linking $\boldsymbol{b}$, $\boldsymbol{w}$ and $\boldsymbol{v}$ that you found in 2(a).

Question 4 *Measuring boat speeds*

Suppose we want to measure our boat's speed b, perhaps to check its speedometer. In general **we will not know the water speed** w, so we shall need to eliminate it from our final equations.

(a) Suppose we head directly downstream 200 m from A to B with $\boldsymbol{b}_1 = b\boldsymbol{i}$ and it takes 1 min. What is $\boldsymbol{v}_1$?



Now we make the return journey in 2 min. (The boat **water speed** is still b of course.) What are $\boldsymbol{b}_2$ and $\boldsymbol{v}_2$?

Write down two vector equations, eliminate $\boldsymbol{w}$ and find b. How many unknowns were there? How many pieces of information were measured?

(b) In my next try at measuring a boat's speed, I time the boat as it goes **directly** across the river and find it takes $t_1=5/6$ min. Remember $\boldsymbol{w}=w\boldsymbol{i}$, but I do not know w, i.e. I know the direction of the river's flow but not its speed.

What is $\boldsymbol{v}_1$? (Remember: the boat goes directly across.)

Can you find the boat speed b? If not, why not?

Explain using: **(i)** equations (it is useful to put $\boldsymbol{b}=b_1\boldsymbol{i}+b_2\boldsymbol{j}$, so $b=\sqrt{b_1^2+b_2^2}$); and **(ii)** diagrams.

Would it help if I timed the return journey straight across the river?

You should have found that you can find the boat's speed in (a) but not in (b). **Make sure you have explained clearly why.**

Question 5

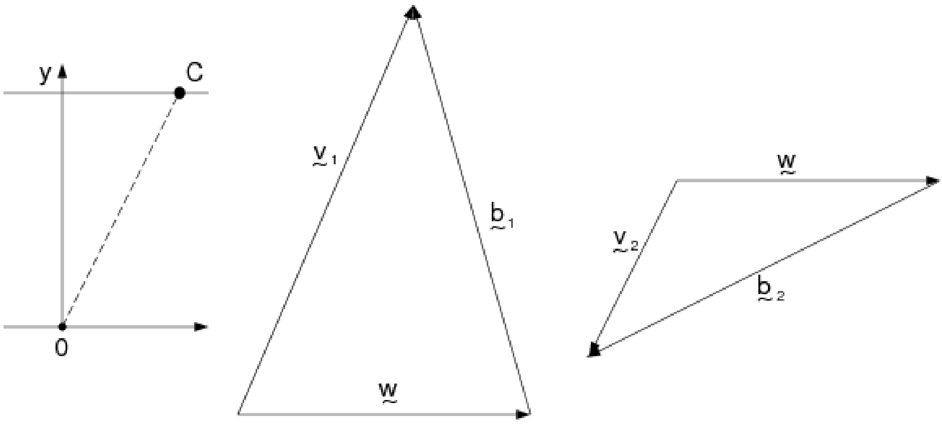
We have seen that steering the boat directly down the river and back enables us to measure its speed, but going directly across does not work. Now let's move to the next level of generality: suppose the boat goes **diagonally** across the river and back.

Do you think we will now be able to deduce the value of the speed b? Why?

Hopefully you said yes! We can check this out nicely using a geometric approach.

First, I want you to follow through the following direct-problem working, then you will be asked to modify it to solve the inverse problem.

Suppose we go across the river from O to C and back and measure times t_1 and t_2. Then $\boldsymbol{v}_1$ will be in the direction OC and have magnitude OC/t_1; $\boldsymbol{v}_2$ will be in the opposite direction, CO and have magnitude OC/t_2. We have two vector diagrams for velocities.

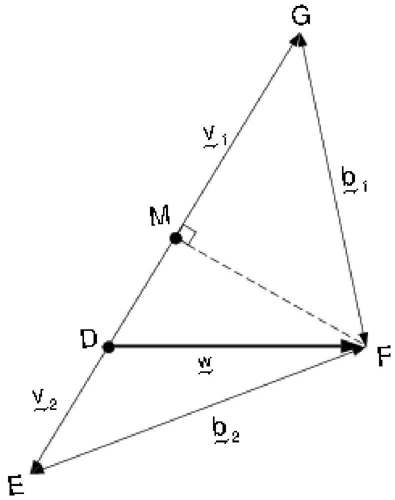


Remember $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ both have length b, the speed we want, and the direction of $\boldsymbol{w}$ is known to be $\boldsymbol{i}$.

Now I want you to join $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ together at the point D and form a triangle EFG as below, i.e. join together the above two triangles using their common side $\boldsymbol{w}$.

Now join the midpoint M of EG to F.

(a) Actually, instead of doing that I could have said, find the midpoint M of EG, draw the line through M and perpendicular to EG, and you will find it goes to F. *Why is that so?*



Now we are ready to tell someone how to deduce the boat speed b, after they have timed diagonal journeys across and back.

(b) Write out instructions for finding the boat speed b. Use your own words and remember you are telling someone exactly what they must do — think of it as writing a manual. Show how their diagram should look at each stage.

One possible outline would be:

Calculate the two vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ and join them together at D, so that EDG is a straight line with $\overrightarrow{DE} = \boldsymbol{v}_2$ and $\overrightarrow{DG} = \boldsymbol{v}_1$.

Draw a horizontal line (i.e. in the $\boldsymbol{i}$ direction) through . . . .

Next, find M, the . . . .

Through M, draw a line . . . .

Denote by F the point given by . . . .

Now you can measure the length of . . . . to get the boat speed b.

Question 6 *The grand finale!*

If I am in open water, I will not even know the **direction** of $\boldsymbol{w}$. (In a plane, we will be flying from O to C and back without knowing details of the cross-wind.)

Before you read on: do you think the scheme in 5(b), where we go from a point O to a point C and back, still works? Think about it. Discuss it. Write down your comments.

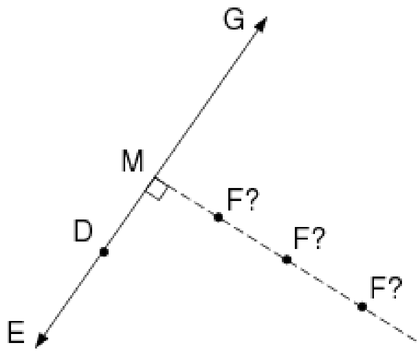
(a) Very generally: where is the difficulty? How many unknowns are there and how many pieces of information are given by the measurements?

Let's see if we can identify the difficulty very clearly using our geometric approach.

We can still measure $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$, form the line EDG and draw in the perpendicular line through midpoint M.

We still know that there is a point F **somewhere** on that last line so that $FE = FG =$ required speed b.

(b) What does the $\overrightarrow{DF}$ represent?
So why does $FE = FG = b$ give the speed still?
But do we know the length or direction of DF?
So can we **actually** draw in DF and then find b?
What do you suggest we need to do? (Be polite!)

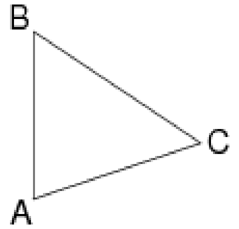


So how DOES the general case work?

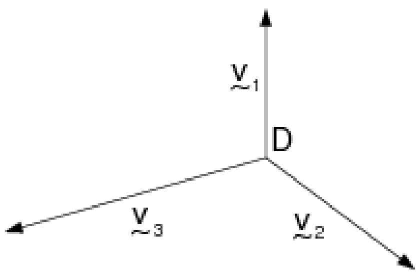
So that you are not driven crazy wondering about it, this is how the general case works. This solution was given by Von Mises about eighty years ago as a way of measuring aircraft speed when the wind speed and direction are unknown. I'll write it out for a boat as before: water velocity $\boldsymbol{w}$, as well as boat speed b, unknown.

Travel a triangular course ABC measuring the three times taken. This will give $\boldsymbol{v}_1$ in direction AB, magnitude AB/t_1; $\boldsymbol{v}_2$ in direction BC, magnitude BC/t_2; $\boldsymbol{v}_3$ in direction CA, magnitude CA/t_3.

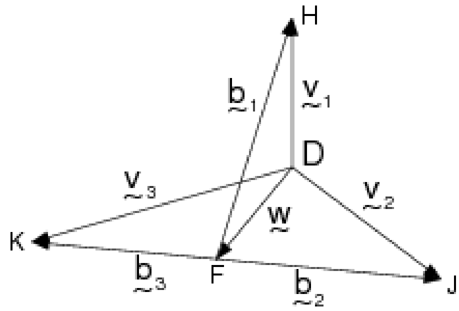
As before we join the $\boldsymbol{v}$ vectors together at a point D (see below). Now we know that there **is** a vector diagram using $\boldsymbol{v} = \boldsymbol{b} + \boldsymbol{w}$ for each case and, as before, we **could** put those all together so that F is the end of the water flow vector $\boldsymbol{w}$ (see below), but to actually do that we need to know $\boldsymbol{w}$! And we do **not** know it.



The course



The measured velocities



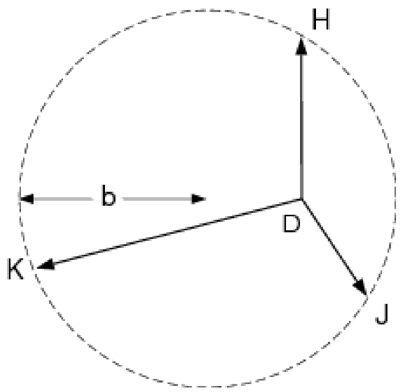
The $\boldsymbol{v} = \boldsymbol{b} + \boldsymbol{w}$ vector diagrams

If only we knew F, we could then measure FH, FJ or FK, and they would all give us the wanted boat speed b.

But wait, FH, FJ and FK all have the same length b, so points H, J and K lie on a circle centre F. Got it!

Draw the measured vectors $\overrightarrow{DH} = v_1$, $\overrightarrow{DJ} = v_2$ and $\overrightarrow{DK} = v_3$.

Now draw a circle through the points H, J and K: its radius will be the required boat speed b
.



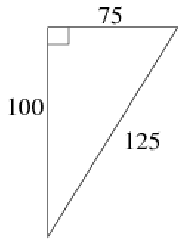
Just a nagging doubt: is there only **one** *circle you can draw through three given points A, B and C? i.e. is our answer unique?*

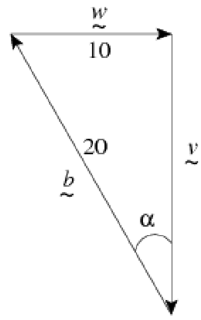
Instructors' Guide

Solutions

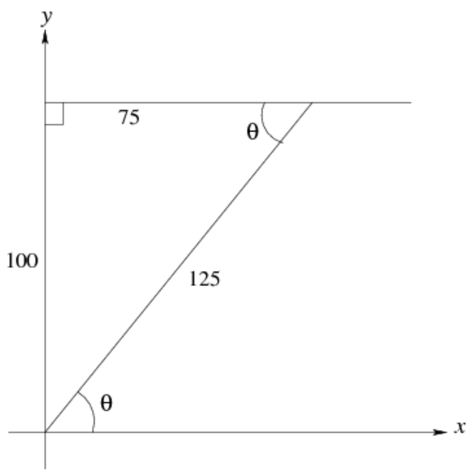
1. **(a)** **(i)** $\boldsymbol{b} = 40\boldsymbol{j}$; **(ii)** $\boldsymbol{v} = 40\boldsymbol{j}$; **(iii)** to cross takes $100/40 = 2.5$ minutes.
   (b) The boat drifts down the river with speed 30 m/min, so that after 2 minutes it is at $x{=}60$, $y{=}0$.

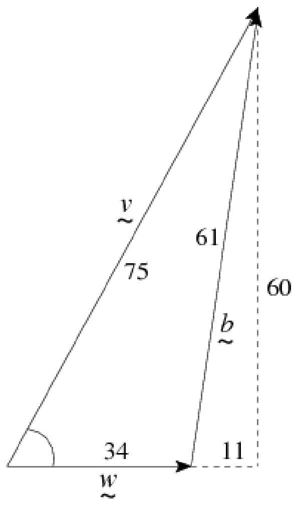
2. **(a)** $\boldsymbol{v} = \boldsymbol{b} + \boldsymbol{w}$.
   (b) *Left-hand figure below.* Velocity $\boldsymbol{v} = 30\boldsymbol{i} + 40\boldsymbol{j}$; speed $= ||\boldsymbol{v}|| = 50$ m/min.
   The boat reaches the other side after 2.5 minutes at $x{=}2.5{\times}30{=}75$m.
   It has travelled 125m.





3. **(a)** *Right-hand figure above.* Because we have a right-angled triangle, the sine of the angle between $\boldsymbol{b}$ and $\boldsymbol{v}$ is $10/20 = 0.5$. Therefore the angle α is $\pi/6$ radians or 30°.
   (b) *Left-hand figure below.* The boat travels 125m. The speed $||\boldsymbol{v}||$ is $125/(5/3) = 75$ m/min. Since $\cos(\theta){=}3/5$, $\sin(\theta){=}4/5$ and $v{=}75$, we get $\boldsymbol{v} = 45\boldsymbol{i} + 60\boldsymbol{j}$.





   (c) *Right-hand figure above.* We now have $\boldsymbol{w} = 34\boldsymbol{i}$, and since $\boldsymbol{v} = \boldsymbol{b} + \boldsymbol{w}$, we have $\boldsymbol{b} = \boldsymbol{v} - \boldsymbol{w} = 11\boldsymbol{i} + 60\boldsymbol{j}$. Boat speed $b = ||\boldsymbol{b}|| = \sqrt{11^2 + 60^2} = 61$ m/min.

4. **(a)** As the speed in travelling A to B is 200 m/min, $\boldsymbol{v}_1 = 200\boldsymbol{i}$, with $\boldsymbol{b}_1 = b\boldsymbol{i}$.
As the speed in travelling B to A is $200/2 = 100$ m/min, $\boldsymbol{v}_2 = -100\boldsymbol{i}$, with $\boldsymbol{b}_1 = -b\boldsymbol{i}$.
The two equations are

$$\begin{aligned} \boldsymbol{v}_1 &= \boldsymbol{b}_1 + \boldsymbol{w} \quad \text{or} \quad 200\boldsymbol{i} = b\boldsymbol{i} + w\boldsymbol{i} \\ \text{and } \boldsymbol{v}_2 &= \boldsymbol{b}_2 + \boldsymbol{w} \quad \text{or} \quad -100\boldsymbol{i} = -b\boldsymbol{i} + w\boldsymbol{i} \end{aligned}.$$

Solving these gives $b = 150$ m/min.

There are two unknowns, b and w. Two pieces of information were measured, the times of the journeys.

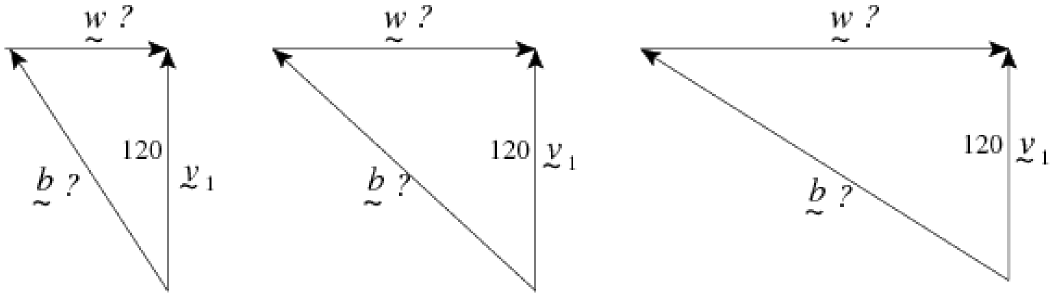
(b) As the boat goes directly across, the direction of $\boldsymbol{v}_1$ is $\boldsymbol{j}$.
The speed is $100/(5/6)$, so $||\boldsymbol{v}_1|| = 120$ m/min. Therefore $\boldsymbol{v}_1 = 120\boldsymbol{j}$.

(i) To find b we need $\boldsymbol{b}$, as $b = ||\boldsymbol{b}||$.
Now our basic vector equation gives $\boldsymbol{v} = \boldsymbol{b} + \boldsymbol{w}$, so that $\boldsymbol{b} = \boldsymbol{v}_1 - \boldsymbol{w}$.
If we let $\boldsymbol{b} = b_1\boldsymbol{i} + b_2\boldsymbol{j}$, this equation gives us

$$b_1\boldsymbol{i} + b_2\boldsymbol{j} = 120\boldsymbol{j} - w\boldsymbol{i}.$$

So we know $b_2 = 120$. But we do not know b_1, as $b_1 = w$, and we are not given the water speed w.
Because $b = ||\boldsymbol{b}|| = \sqrt{b_1^2 + b_2^2}$, this means we cannot find b.

(ii) We know what the vector diagram looks like but we do not know how long to make $\boldsymbol{w}$, so we cannot draw in $\boldsymbol{b}$ and calculate its length.



No, we gain no new information by timing the return journey.

5. **(a)** Because EF and FG both have length b, we have an isosceles triangle. Therefore, because FM bisects EG, it is perpendicular to it.

(b) Calculate the two vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$, and join them together at D, so that EDG is a straight line, with $\overrightarrow{DE} = \boldsymbol{v}_2$ and $\overrightarrow{DG} = \boldsymbol{v}_1$.
Draw the horizontal line (i.e. in the $\boldsymbol{i}$ direction) through D.
Next find the midpoint of the line EG.
Through M, draw a line perpendicular to EG.
Denote by F the point of intersection of the lines drawn above through M and D.
Now you can measure the length of FE or FG to get the boat speed b.

6. **(a)** There are really three unknowns — the water speed and direction and the boat speed b. However, we only have two pieces of information — the times of the journeys.

(b) $\overrightarrow{DF}$ represents the water velocity $\boldsymbol{w}$.

$FE = FG = b$ still gives the water speed, because the $\boldsymbol{v} = \boldsymbol{b} + \boldsymbol{w}$ vector diagrams can be drawn in as before. $\overrightarrow{FG}$ still gives $\boldsymbol{b}_1$ and $\overrightarrow{FE}$ still gives $\boldsymbol{b}_2$.

We do **not** know the length or direction of $\overrightarrow{DF}$ because we are not given $\boldsymbol{w}$.

No we cannot actually draw in DF to then find b.

We need more information. A triangular course will do the job, as explained on the last page of the lab.

3.3 Epidemics, Airline Routes and Assignment Problems

Aims

- To become more familiar with matrix notation.
- To become more familiar with matrix multiplication.
- To see how matrices organise data.
- To see how simple matrix manipulations aid our thinking and analysis of various problems.

Do Questions 1 and 2 first, then ask the lab instructor for Question 3. The maximum mark you can get for 1 and 2 only is 7/10.

Each of these problems should really take no more than 30 minutes. Get to work — use the "group power" — ask for help if needed.

Question 1 *Epidemics: direct and indirect contact with a contagious disease*

In this example we show how data about diseases can be stored in matrices and how matrix multiplication can be used to model the spread of a contagious disease. Suppose that 3 individuals have contracted such a disease. This group has contacts with 5 people in a second group. We can represent these contacts, called *direct contacts*, by a 3×5 matrix. An example of such a matrix is given below.

Direct-Contact Matrix: first and second groups

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

We set $a_{ij} = 1$ if the ith person in the first group has made contact with the jth person in the second group. For example, the 1 in the 2,4 position, i.e. $a_{24} = 1$, means that the second person in the first (infected) group has been in contact with the fourth person in the second group. So row numbers refer to people in the first group and column numbers refer to people in the second group.

(a) Which members of Group 1 was Member 5 of Group 2 in contact with?

(b) What does the sum of all the elements in column 1 represent? in column k?

(c) What does the sum of all the elements in row 1 represent? in row p?

Now suppose that a third group of 4 people has had a variety of direct contacts with individuals of the second group. We can also represent this by a matrix.

Direct-Contact Matrix: second and third groups

$$\boldsymbol{B} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$b_{ij} = 1$ if member i of the second group has made contact with member j of the third group. So that $b_{21} = 0$, which means that the second person in the second group has had no contact with the first person in the third group.

We now analyse how the disease might spread from the infected Group 1 to Group 3.

(d) What does $a_{21}b_{13} + a_{22}b_{23} + a_{23}b_{33} + a_{24}b_{43} + a_{25}b_{53}$ represent in terms of Member 2 in Group 1 passing the disease to Member 3 in Group 3?

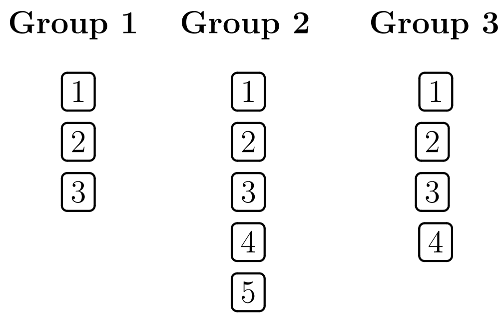
(e) So what does the matrix $\boldsymbol{C} = \boldsymbol{A}\,\boldsymbol{B}$ represent and tell us? Explain why it is called the **indirect- or second-order-contact matrix**.

(f) Calculate $\boldsymbol{C}$ for the $\boldsymbol{A}$ and $\boldsymbol{B}$ given above (by hand).

(g) Which member of the third group is **most likely** to get the disease? (Perhaps the answer to a problem like (b) is relevant here.)

Who would you expect **not** to get the disease? Explain your reasoning.

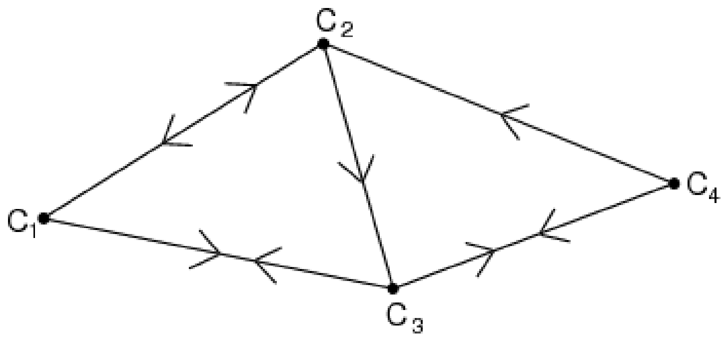
Perhaps it would help to construct a diagram of the groups. Draw a line to represent a contact.



Question 2 *Airline routes: graphs and matrices*

In the last question we finished with a diagram. We often start with a network or **graph** (vertices + joining edges) and analyse it using matrices.

Here is an airline route map. Arrows indicate how planes fly between the four cities.



The matrix $\boldsymbol{M}$ summarises the route data.

$m_{ij}=1$ means a plane flies directly from City i to City j
$m_{ij}=0$ means there is no service directly from i to j.

$$\boldsymbol{M} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{bmatrix}$$

(a) Complete rows 3 and 4 of $\boldsymbol{M}$.

(b) How can we interpret the matrix

$$\boldsymbol{M}^2 = \begin{bmatrix} 2 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 2 & 2 & 0 \\ 2 & 0 & 1 & 1 \end{bmatrix}?$$

Explain in terms of the route map.

Hint: If $\boldsymbol{S} = \boldsymbol{M}^2$, then

$$s_{ij} = m_{i1}m_{1j} + m_{i2}m_{2j} + m_{i3}m_{3j} + m_{i4}m_{4j}.$$

What does it mean if one of these product terms is one? or zero?

(c) Calculate $\boldsymbol{M} + \boldsymbol{M}^2$.
Are there any zeros in your result? What does this tell you about this airline network? Is there anything special about column 3? How would that help the airline with its boast "we can always get you to City 3"?

Question 3 *Assignment problems*

These problems turn up in many forms: assigning workers to jobs, equipment to worksites, players to team positions, even grooms to brides (marriage brokers or matchmakers) and so on. Here is an example, with a little introduction to the systematic treatment of the Assignment Problem.

The situation

A college intends to install air-conditioning in three of its buildings during a one-week spring break. It invites three contractors to submit separate bids for the work involved in each of the three buildings. The bids it receives (in 1000-dollar units) are listed in the following table.

Bids

	Bldg 1	Bldg 2	Bldg 3
Contractor 1	57	96	37
Contractor 2	47	87	41
Contractor 3	60	80	36

Each contractor can install the air-conditioning in only one building during the one-week period, so the college must assign a different contractor to each building.

The problem

Decide which building each contractor should be assigned to in order to minimize the total cost (sum of the corresponding bids).

Definitions

The **cost matrix** for this problem is the 3×3 matrix $\begin{bmatrix} 57 & 96 & 37 \\ 47 & 87 & 41 \\ 60 & 80 & 36 \end{bmatrix}$.

For a 3×3 cost matrix, an **assignment** is a set of 3 matrix elements, no two of which lie in the same row or column.

The **cost of an assignment** is the sum of the 3 entries.

An **optimal assignment** has the smallest cost.

PTO

Your tasks

(a) Explain why that definition of assignment makes sense for our example. What does it correspond to in our case? What would it mean if there was more than one element chosen in a row? or a column?

(b) How many possible assignments are there?

Here is one.
We can indicate assignments with circles.
It uses the 1,2, the 2,1 and the 3,3 elements and its cost is 96+47+36=179.

$$\begin{bmatrix} 57 & \textcircled{96} & 37 \\ \textcircled{47} & 87 & 41 \\ 60 & 80 & \textcircled{36} \end{bmatrix}$$

(c) Find the costs and the optimum for the possible assignments in our example.

Don't grumble — it's simple arithmetic and a little each if you work as a group.

Go back to the original application: which contractors should do which job?

(d) *Obviously such a process starts to get even more tedious and time consuming (even for a computer) when the size of the problem increases beyond* 3×3. *Time for some mathematical thinking! Here is a relevant result. (Perhaps if it seems very obvious or you have time at the end you might comment on it and on why it is true.)*

> **Theorem:** If a number is added to or subtracted from all of the entries of any one row or column of a cost matrix, an optimal assignment for the resulting cost matrix is also an optimal assignment for the original cost matrix.

How might that help? Well, a cost matrix with non-zero elements but with lots of zeros in it is easy to analyse — if one assignment has a cost of assignment equal to zero, it is certainly the best. Try it.

Take the cost matrix in this example and

first, take the smallest element in each row away from all elements in that row,

second, take the smallest element in each new column away from all elements in that column.

What do you get? Can you spot a zero-cost assignment? Is it the same as your original assignment?

Remarks

- The idea of generating "equivalent problems" is one we used for linear equations.
- The method you just used is called *The Hungarian Method*. It may involve more steps than the ones you had to make.

Instructors' Guide

Solutions

1. **(a)** Member 5 of Group 2 was in contact with Members 1 and 3 of Group 1 (look at the 5th column).

   (b) The sum of all the elements in column k is the total number of contacts Member k of Group 2 had with members of Group 1.

   (c) The sum of all the elements in row p is the total number of contacts Member p of Group 1 had with members of Group 2.

   (d) $a_{2j}b_{j3} = 1$ if Member 2 of Group 1 had contact with Member j of Group 2 **and** Member j of Group 2 had contact with Member 3 of Group 3. Otherwise $a_{2j}b_{j3}$ is zero. $j = 1, 2, 3, 4, 5$.

   So the sum gives the number of times Member 2 of Group 1 and Member 3 of Group 3 interact through intermediate contacts with members of Group 2.

   (e) $\boldsymbol{C}$ displays all the indirect contacts.

   c_{ij} is the number of indirect contacts between Member i of Group 1 and Member j of Group 3.

   (f) $$\boldsymbol{C} = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 2 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}.$$

   (g) Adding up the columns gives the total number of indirect contacts between members of Group 3 and Group 1:

   Member 1 has 2 indirect contacts
   Member 2 has 0 indirect contacts
   Member 3 has 1 indirect contact
   Member 4 has 3 indirect contacts.

   Member 4 is most likely to catch the disease, but Member 2 should not.

Group 1	**Group 2**	**Group 3**
1	1	1
2	2	2
3	3	3
	4	4
	5	

The lines (to be put in) show the paths or links from Group 1 to Group 3.
There is no path from Group 1 to Member 2 of Group 3. But there are paths for the others.

2. **(a)** $\boldsymbol{M} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}.$

(b) Put $\boldsymbol{S} = \boldsymbol{M}^2$. Then $\boldsymbol{S} = [s_{ij}]$, with $s_{ij} = m_{i1}m_{1j} + m_{i2}m_{2j} + m_{i3}m_{3j} + m_{i4}m_{4j}$.

Each term tells us about the possibility of going from City i to City j via an intermediate city, e.g. $m_{i2}m_{2j}$ tells us about going from City i to City 2, then from City 2 to City j. If both of these are possible routes, we get $m_{i2}m_{2j} = 1 \times 1 = 1$. If either is not a possible route, we get 0.

So the elements of $\boldsymbol{M}^2$ tell us how many routes there are between the cities when we travel through an intermediate city.

(c) $\boldsymbol{M} + \boldsymbol{M}^2 = \begin{bmatrix} 2 & 1 & 2 & 1 \\ 2 & 1 & 2 & 1 \\ 1 & 2 & 2 & 1 \\ 2 & 1 & 2 & 1 \end{bmatrix}.$

All cities are linked using either a direct flight or with one change of flight at an intermediate city. Some cities can be reached from others in more than one way. City 3 can always be reached in 2 ways (column 3), so if there is a problem on one route, City 3 can still be reached by another route.

3. **(a)** Two entries in a row would mean a contractor was doing two jobs, and that is not possible.

Two entries in a column would mean that a building is being air-conditioned by two contractors, which is not what we want.

This definition of assignment gives us a different contractor for each building, as we require.

(b) We can let Contractor 1 do any one of the 3 buildings.

Then Contractor 2 can be sent to either of the remaining 2 buildings.

Contractor 3 does the building not taken care of so far.

So the total number of possibilities is $3 \times 2 \times 1 = 6$.

PTO

(c) $\begin{bmatrix} \boxed{57} & 96 & 37 \\ 47 & \boxed{87} & 41 \\ 60 & 80 & \boxed{36} \end{bmatrix}$ $\begin{bmatrix} \boxed{57} & 96 & 37 \\ 47 & 87 & \boxed{41} \\ 60 & \boxed{80} & 36 \end{bmatrix}$ $\begin{bmatrix} 57 & \boxed{96} & 37 \\ \boxed{47} & 87 & 41 \\ 60 & 80 & \boxed{36} \end{bmatrix}$

cost = 180 cost = 178 cost = 179

$\begin{bmatrix} 57 & \boxed{96} & 37 \\ 47 & 87 & \boxed{41} \\ \boxed{60} & 80 & 36 \end{bmatrix}$ $\begin{bmatrix} 57 & 96 & \boxed{37} \\ 47 & \boxed{87} & 41 \\ \boxed{60} & 80 & 36 \end{bmatrix}$ $\begin{bmatrix} 57 & 96 & \boxed{37} \\ \boxed{47} & 87 & 41 \\ 60 & \boxed{80} & 36 \end{bmatrix}$

cost = 197 cost = 184 cost = 164

The optimum assignment is the last one, using the $1,3$, the $2,1$ and the $3,2$ elements.

So we should get Contractor 1 to work on Building 3, Contractor 2 to work on Building 1 and Contractor 3 to work on Building 2.

(d) $$\begin{bmatrix} 57 & 96 & 37 \\ 47 & 87 & 41 \\ 60 & 80 & 36 \end{bmatrix} \rightarrow \begin{bmatrix} 20 & 59 & 0 \\ 6 & 46 & 0 \\ 24 & 44 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 14 & 15 & 0 \\ 0 & 2 & 0 \\ 18 & 0 & 0 \end{bmatrix}.$$

$\begin{bmatrix} 14 & 15 & \boxed{0} \\ \boxed{0} & 2 & 0 \\ 18 & \boxed{0} & 0 \end{bmatrix}$ is optimum.

3.4 Transformations and Matrices

Aims

- To practise matrix manipulations.
- To explore the use of matrices for describing geometric operations.
- To learn about symmetry.
- To see how a different area of algebra can be developed.

Preamble

Suppose that a transformation T changes a position vector $\boldsymbol{r}_1$ into $\boldsymbol{r}_2$.

We can describe this algebraically using column vectors $\boldsymbol{r}_1 = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$ and $\boldsymbol{r}_2 = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$

The transformation T is **represented** by a matrix $\boldsymbol{T}$ such that $\boldsymbol{r}_2 = \boldsymbol{T}\boldsymbol{r}_1$.

If it exists, $\boldsymbol{T}^{-1}$ will transform $\boldsymbol{r}_2$ back into $\boldsymbol{r}_1$, i.e. $\boldsymbol{r}_1 = \boldsymbol{T}^{-1}\boldsymbol{r}_2$.

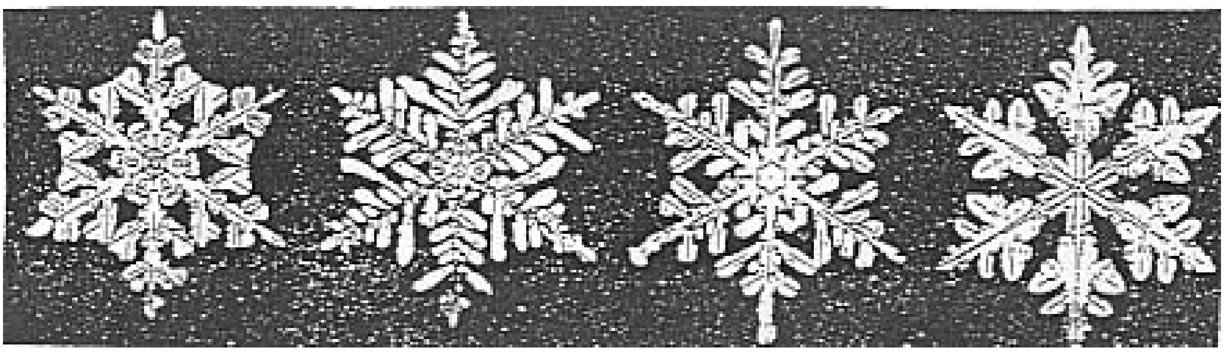
If it takes two transformations to get from $\boldsymbol{r}_1$ to $\boldsymbol{r}_2$, and if the one represented by $\boldsymbol{T}_1$ is done first and the one represented by $\boldsymbol{T}_2$ is done next, then $\boldsymbol{r}_2 = \boldsymbol{T}_2\boldsymbol{T}_1\boldsymbol{r}_1$.

A summary sheet is on page 207 — have a quick look before you begin.

Don't forget to explain with words and pictures how it all works.

Question 1 *Setting the scene*

It is often said that no two snowflakes are the same and that is clear in the little sample shown below. But there are some similarities. Describe what it is that the snowflakes below appear to have in common.



Question 2 *Using the algebra–geometry link*

Two rotation matrices can be multiplied in any order to get the same result, i.e. they "commute". How about a mixture of rotation and reflection matrices?

(a) Draw a diagram to see whether rotating a vector by 45° and then reflecting it in the y axis gives the same vector as doing the reflection first, then the rotation. Just pick one or two vectors and see what happens to them.

(b) So, do you expect

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} ?$$

Explain your reasoning by identifying the geometric operations represented by these matrices and referring back to your answer in (a). Then check it out by doing the multiplications.

Question 3 *Manipulating some matrices*

We now work with the following set of four matrices:

$$\boldsymbol{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \boldsymbol{A} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \boldsymbol{B} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \quad \boldsymbol{C} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

(a) Work out the multiplication table for $\boldsymbol{M}_1$, $\boldsymbol{M}_2$, where $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$ can be any of the four matrices.

Calculator? No! Easy by hand. Think: what must the product be where $\boldsymbol{I}$ *is involved?*

$\boldsymbol{M}_1$ / $\boldsymbol{M}_2$	$\boldsymbol{I}$	$\boldsymbol{A}$	$\boldsymbol{B}$	$\boldsymbol{C}$	
$\boldsymbol{I}$					
$\boldsymbol{A}$					
$\boldsymbol{B}$				$\boldsymbol{A}$	← e.g. $\boldsymbol{CB}=\boldsymbol{A}$
$\boldsymbol{C}$					

(b) **(i)** Does every product give $\boldsymbol{I}$, $\boldsymbol{A}$, $\boldsymbol{B}$ or $\boldsymbol{C}$?

(ii) Explain how you can use your table to find inverses for $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$.
What are $\boldsymbol{A}^{-1}$, $\boldsymbol{B}^{-1}$, $\boldsymbol{C}^{-1}$?
Is it true that the set contains the inverses of all matrices in it?

(iii) Is it true that $(\boldsymbol{M}_1\boldsymbol{M}_2)\boldsymbol{M}_3 = \boldsymbol{M}_1(\boldsymbol{M}_2\boldsymbol{M}_3)$, no matter how you choose $\boldsymbol{M}_1$, $\boldsymbol{M}_2$ and $\boldsymbol{M}_3$? *Hint*: do you need to check all cases or is there a particular property of matrix multiplication you can refer to?

(c) What is $\boldsymbol{ACBC}$?

Question 4 *Broadening your algebraic horizons a little*

Now we define a GROUP. Then we (i.e. you) will find an example, and after that we will find out why groups are important in Mathematics and Science.

Definition A **group** is a set of elements I, A, B, C, D, ... with a rule for combining any pair of them — in our case we shall call it multiplication — for which the following four properties are satisfied:

(i) there must be an identity element I so that $IQ = Q$ for any element Q in the set;

(ii) the inverse of every element must also be in the set;

(iii) multiplication is associative, i.e. $Q(RS) = (QR)S$ for any elements Q, R and S in the set; and most important of all

(iv) multiplying together any two elements in the set must always give an element in the set.

(a) If the set of elements is the set of matrices in Question 3 and "multiplication" is the usual matrix multiplication, do they form a group? *Explain how you check all four properties, referring back to Question 3 and the table as necessary.*

(b) If the set is reduced to just $\{\boldsymbol{I}, \boldsymbol{B}\}$, do you still have a group? *Give reasons.*

(c) If the set is reduced to $\{\boldsymbol{I}, \boldsymbol{A}, \boldsymbol{B}\}$, do you still have a group? *Give reasons.*

Question 5 *Appreciating the importance of groups*

So who cares about groups? Well, people interested in algebra and mathematical structures do, but people involved in chemistry, physics and other sciences also need to use them. Let's see if we can understand why.

(a) Identify the matrices $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ given in Question 3 with particular rotation matrices $\boldsymbol{R}(\alpha)$.

(b) By thinking about what happens to any vector when repeatedly acted upon by $\boldsymbol{I}$, $\boldsymbol{A}$, $\boldsymbol{B}$ or $\boldsymbol{C}$ (and in particular four times by $\boldsymbol{A}$), why do you think this group is called "the cyclic group of order 4", denoted C_4? *Explain. Drawings needed*?

(c) Now look at the swastika shown below in Figure 1. Choose any point on it and draw in the position vector. Show what happens when you multiply that vector by each of the elements of our group C_4.

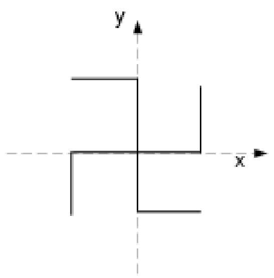


Fig.1

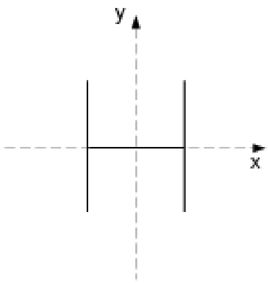


Fig.2

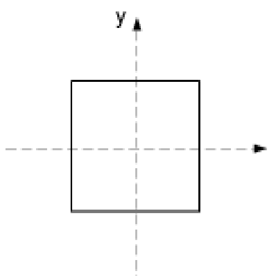


Fig.3

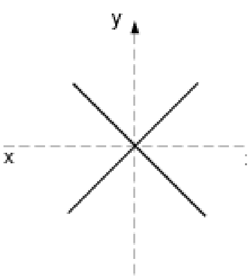


Fig.4

(d) What happens if you act on the **whole** swastika with any of those transformations (rotations). What changes? Or is it left looking unchanged? *Explain/report on your reasoning.*

(e) Can you do any reflections (e.g. through the x or y axes, or the lines $y=x$ or $y=-x$) to leave the swastika looking the same? Give an example of what happens.

(f) The swastika is obviously symmetric in certain ways and now we can state this in a precise mathematical way: "the symmetry group of the swastika is C_4". What do you think that means? Perhaps you can answer by writing down how you would explain it to a fellow student.

> Keep going! Do Question 6 and at least read through Question 7. It will be treated as a bonus question if I find this lab is a little long. But do leave a little time for Question 8.

Question 6 *Another symmetry group*

The symmetry group for the letter H, as in Figure 2, has elements

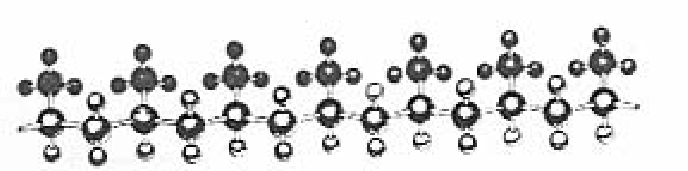
$$\boldsymbol{I}=\begin{bmatrix}1&0\\0&1\end{bmatrix}\qquad \boldsymbol{R}_2=\begin{bmatrix}-1&0\\0&-1\end{bmatrix}\qquad \boldsymbol{M}_x=\begin{bmatrix}1&0\\0&-1\end{bmatrix}\qquad \boldsymbol{M}_y=\begin{bmatrix}-1&0\\0&1\end{bmatrix}.$$

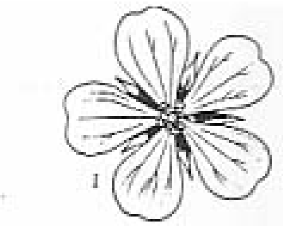
Explain what that means! If you have spare time, you could construct the multiplication table for this group.

Question 7 *Lots of symmetry*

(a) The square (Figure 3) and letter X (Figure 4) are obviously more symmetrical than the swastika. What extra "symmetry operations" can you do to them which will leave them looking the same?

(b) *A more symmetric figure will be described by an extended group.* What does that mean?

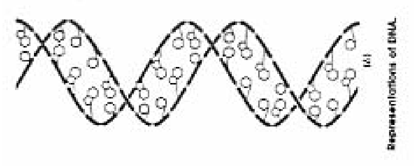


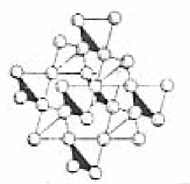


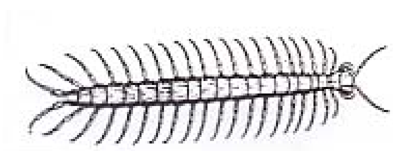
* *

The different types of atoms, molecules, crystals and other structures found in nature have a variety of symmetry properties; groups are the mathematical things we use to describe and classify them, and to explore their properties.

* *







Question 8 *Summarizing*

Think about what you have been doing in this lab. What do you think the key points have been? What have you discovered? How would you **now** answer Question 1?

Supplementary Question

Question 9 *Extending to three dimensions*

Suppose we now go to three dimensions and position vectors $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$, which are represented by column vectors $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$.

You may need to draw some pictures and do some "experiments" using pens, rulers, etc as vectors and axes.

(a) What geometric operations do these three matrices represent algebraically?

$$\boldsymbol{M} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \boldsymbol{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \boldsymbol{R} = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(b) What does $\boldsymbol{P}$ do to all the vectors $\begin{bmatrix} 1 \\ 2 \\ t \end{bmatrix}$, where t can be any number?

Can you use that fact to decide whether $\boldsymbol{P}$ has an inverse?

Perhaps you need to re-read the Preamble and think what $\boldsymbol{P}^{-1}$ *is suppposed to do.*

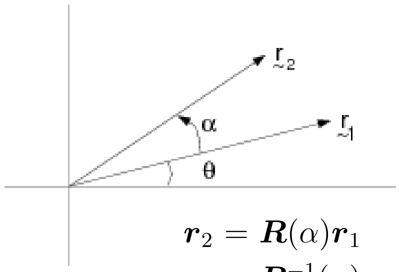
Give an algebraic argument to support your answer about the existence of $\boldsymbol{P}^{-1}$.
Hint: Perhaps you could think about the bottom row produced when $\boldsymbol{P}$ multiplies any 3×3 matrix and recall which matrix must be produced when $\boldsymbol{P}$ multiplies its inverse.

Matrices Representing Geometric Operations on Vectors

ROTATIONS

$$\boldsymbol{R}(\alpha) = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}$$

$$\boldsymbol{R}(\alpha) \begin{bmatrix} r\cos(\theta) \\ r\sin(\theta) \end{bmatrix} = \begin{bmatrix} r\cos(\theta+\alpha) \\ r\sin(\theta+\alpha) \end{bmatrix}$$



$$\boldsymbol{r}_2 = \boldsymbol{R}(\alpha)\boldsymbol{r}_1$$
$$\boldsymbol{r}_1 = \boldsymbol{R}^{-1}(\alpha)\boldsymbol{r}_2$$

Inverse: $\boldsymbol{R}^{-1}(\alpha) = \boldsymbol{R}(-\alpha)$.

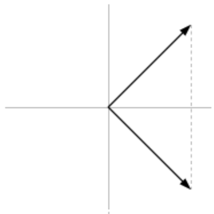
Products: If $\boldsymbol{r}_2 = \boldsymbol{R}(\alpha)\boldsymbol{r}_1$ and $\boldsymbol{r}_3 = \boldsymbol{R}(\beta)\boldsymbol{r}_2$, then $\boldsymbol{r}_3 = \boldsymbol{R}(\beta)\boldsymbol{R}(\alpha)\boldsymbol{r}_1 = \boldsymbol{R}(\alpha+\beta)\boldsymbol{r}_1$.

Products commute: $\boldsymbol{R}(\beta)\boldsymbol{R}(\alpha) = \boldsymbol{R}(\alpha)\boldsymbol{R}(\beta)$.

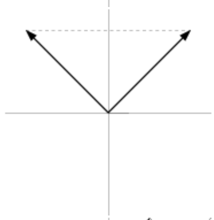
Examples: $\boldsymbol{R}\left(\frac{\pi}{4}\right) = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$ $\quad \boldsymbol{R}\left(\frac{\pi}{3}\right) = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$.

REFLECTIONS

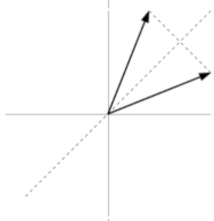
Reflection in $\boldsymbol{x}$ axis: $\quad \boldsymbol{M}_x \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ -y \end{bmatrix}$



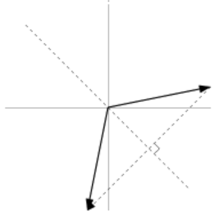
Reflection in $\boldsymbol{y}$ axis: $\quad \boldsymbol{M}_y \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ y \end{bmatrix}$



Reflection in line $\boldsymbol{x} = \boldsymbol{y}$: $\quad \boldsymbol{M}_{x-y} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix}$



Reflection in line $\boldsymbol{x} = -\boldsymbol{y}$: $\quad \boldsymbol{M}_{x+y} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -y \\ -x \end{bmatrix}$

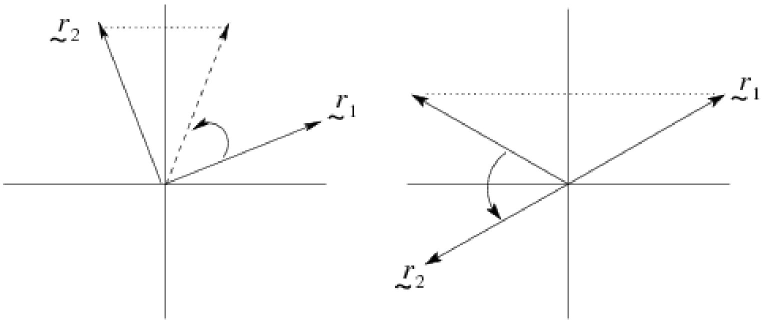


Each reflection matrix is its own inverse.

Instructors' Guide

Solutions

2. **(a)**



A different order of operations generates a different result.

(b) These are the matrices representing the reflection and rotation in (a). The answer in (a) suggests order of multiplication is important.

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

3. **(a)** Work out the multiplication table for $\boldsymbol{M}_1\boldsymbol{M}_2$.

$\boldsymbol{M}_1$ / $\boldsymbol{M}_2$	$\boldsymbol{I}$	$\boldsymbol{A}$	$\boldsymbol{B}$	$\boldsymbol{C}$
$\boldsymbol{I}$	$\boldsymbol{I}$	$\boldsymbol{A}$	$\boldsymbol{B}$	$\boldsymbol{C}$
$\boldsymbol{A}$	$\boldsymbol{A}$	$\boldsymbol{B}$	$\boldsymbol{C}$	$\boldsymbol{I}$
$\boldsymbol{B}$	$\boldsymbol{B}$	$\boldsymbol{C}$	$\boldsymbol{I}$	$\boldsymbol{A}$
$\boldsymbol{C}$	$\boldsymbol{C}$	$\boldsymbol{I}$	$\boldsymbol{A}$	$\boldsymbol{B}$

(b) **(i)** Yes.

(ii) For each matrix, look up the matrix giving $\boldsymbol{I}$ as the product in the table. $\boldsymbol{A}^{-1}=\boldsymbol{C}$, $\boldsymbol{B}^{-1}=\boldsymbol{B}$, $\boldsymbol{C}^{-1}=\boldsymbol{A}$. All inverses are in the set.

(iii) This is the associative property for matrix multiplication.

(c) $\boldsymbol{A}\,\boldsymbol{C}\,(\boldsymbol{B}\,\boldsymbol{C})=\boldsymbol{A}\,\boldsymbol{C}\,\boldsymbol{A}=\boldsymbol{A}\,(\boldsymbol{C}\,\boldsymbol{A})=\boldsymbol{A}\,\boldsymbol{I}=\boldsymbol{A}$.

4. **(a)** Yes, all 4 properties check out. The solutions to Question 3 show this.

(b) We only have to check properties (ii) and (iv), as the other two properties follow from (a): $\boldsymbol{I}\,\boldsymbol{B}=\boldsymbol{B}$, $\boldsymbol{B}\,\boldsymbol{B}=\boldsymbol{I}$, $\boldsymbol{B}^{-1}=\boldsymbol{B}$.

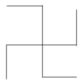
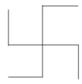
(c) No, because $\boldsymbol{A}\,\boldsymbol{B}=\boldsymbol{C}$, and $\boldsymbol{C}$ has been removed from the set.

5. **(a)** $\boldsymbol{A}{=}\boldsymbol{R}(90°)$, $\boldsymbol{B}{=}\boldsymbol{R}(180°)$, $\boldsymbol{C}{=}\boldsymbol{R}(270°){=}\boldsymbol{R}(-90°)$.

(b) Operating with any of these matrices 4 times gets back to where we started, for example $\boldsymbol{A}^4{=}\boldsymbol{I}$.

(c) We get another point on the swastika.

(d) Nothing changes. We always end up with the original swastika.

(e) No, it looks different, e.g. reflect  in the x axis to get 

(f) The operations described by C_4 are those that leave the swastika unchanged and so they describe its symmetry.

6. You can do the operations represented by those matrices — rotation by 180°, reflection in the x or y axes — on the letter H and it is not changed.

That group describes the symmetry of the letter H.

7. Reflections in the x and y axes and in the lines $y = x$ and $y = -x$ leave $\square$ and $\times$ unchanged.

9. **(a)** Matrix $\boldsymbol{M}$ represents reflection in the xy plane — it changes the sign of the z component of a vector.

Matrix $\boldsymbol{P}$ projects a vector onto the xy plane — it makes the z component of a vector zero.

Matrix $\boldsymbol{R}$ represents a rotation by an angle α about the z axis.

(b) $\boldsymbol{P}\begin{bmatrix}1\\2\\t\end{bmatrix} = \begin{bmatrix}1\\2\\0\end{bmatrix}$.

$\boldsymbol{P}$ projects any vector of the form $\begin{bmatrix}1\\2\\t\end{bmatrix}$ onto the vector $\begin{bmatrix}1\\2\\0\end{bmatrix}$ in the xy plane.

Clearly this operation cannot have an inverse, because an infinite number of vectors are transformed to $\begin{bmatrix}1\\2\\0\end{bmatrix}$ by $\boldsymbol{P}$. If $\boldsymbol{P}^{-1}$ existed, it would transform $\begin{bmatrix}1\\2\\0\end{bmatrix}$ back to only one of these.

If $\boldsymbol{P}$ had an inverse, $\boldsymbol{P}\boldsymbol{P}^{-1}{=}\boldsymbol{I}$. However, the bottom row of $\boldsymbol{P}$ is all zeros and so cannot multiply any matrix to produce a 1 in the bottom row of the product, as is required to produce $\boldsymbol{I}$.

3.5 Complex Numbers, Populations, Matrices and Eigenvalues

Aims

- To practise manipulating complex numbers.
- To link up with some work on matrices.

Preamble

Remember that we have three representations of a complex number

$a+ib$ $\qquad r(\cos\theta+i\sin\theta)$ $\qquad$ a point in the complex plane

We can move between these representations to solve problems in different ways and to visualise our working.

For matrices,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax+by \\ cx+dy \end{bmatrix}.$$

If $\boldsymbol{w}$ is an eigenvector of $\boldsymbol{M}$ with eigenvalue λ, then $\boldsymbol{M}\boldsymbol{w}=\lambda\boldsymbol{w}$.

Details on using a TI-84/CE with complex numbers can be found in *Complex Numbers* in Volume 3 of *Mathematics on a TI-84/CE*.[57]

Question 1

Earlier in the course, we studied Linear Algebra and solved equations using real numbers. That part of Mathematics extends very naturally to equations involving complex numbers, and this problem gives you a little insight into that.

(a) Let $\boldsymbol{u}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\boldsymbol{u}_{n+1} = \boldsymbol{M}\boldsymbol{u}_n$, with $\boldsymbol{M} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$.

Thinking of vectors as $\begin{bmatrix} x \\ y \end{bmatrix}$, plot in the xy plane the vectors $\boldsymbol{u}_0$, $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, $\boldsymbol{u}_3$, $\boldsymbol{u}_4$.

Describe in words what seems to be happening to the vectors.

We could explain this result by going back to our work on matrices as operators, but let's see if we can use a different approach, exploiting some of our other results in order to develop a formula for $\boldsymbol{u}_n$.

(b) **Verify**, i.e. check by seeing that they do actually satisfy $\boldsymbol{M}\boldsymbol{w}=\lambda\boldsymbol{w}$, that

$$\boldsymbol{w}_1 = \begin{bmatrix} 1 \\ -i \end{bmatrix} \quad \text{and} \quad \boldsymbol{w}_2 = \begin{bmatrix} 1 \\ i \end{bmatrix}$$

are eigenvectors of $\boldsymbol{M}$ with eigenvalues $\lambda_1=1+i$ and $\lambda_2=1-i$.

[57] available at *www.XXX*

(c) Write the eigenvalues in the form $\lambda_1 = r(\cos\theta + i\sin\theta)$ and $\lambda_2 = r(\cos\theta - i\sin\theta) = r\big(\cos(-\theta)+i\sin(-\theta)\big)$, so r is the modulus and the arguments are θ and $-\theta$ (in radians of course).

We knew the eigenvalues had to come out as a pair like that didn't we? Of course! The characteristic equation giving the eigenvalues is a polynomial and complex zeros of polynomials always come in conjugate pairs. Previously our examples had real eigenvalues.

(d) Now find numbers a and b so you can write $\boldsymbol{u}_0 = a\boldsymbol{w}_1 + b\boldsymbol{w}_2$. Then we will be able to use the laws of matrix algebra to write $\boldsymbol{M}\boldsymbol{u}_0 = a\boldsymbol{M}\boldsymbol{w}_1 + b\boldsymbol{M}\boldsymbol{w}_2$.

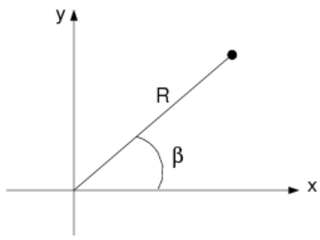
(e) Now the big step! Remembering that if $\boldsymbol{w}$ is an eigenvector of $\boldsymbol{M}$ with eigenvalue λ, then $\boldsymbol{M}^n\boldsymbol{w} = \lambda^n\boldsymbol{w}$, prove that $\boldsymbol{u}_n = \boldsymbol{M}^n\boldsymbol{u}_0 = r^n \begin{bmatrix} \cos(n\theta) \\ \sin(n\theta) \end{bmatrix}$, where r and θ are the modulus and angle you found in (c).

Remember the easy way to raise a complex number to the power n?

(f) Using polar coordinates R, β, so that

$$\boldsymbol{u}_n = \begin{bmatrix} x_n \\ y_n \end{bmatrix} = \begin{bmatrix} R_n\cos(\beta_n) \\ R_n\sin(\beta_n) \end{bmatrix},$$

explain how your answer in (e) confirms your observations in (a).



Notice that we began with a real (as opposed to complex) problem in (a), and in (e) we have derived a general real answer. However, we made use of complex numbers to get to that answer. This is another example of the strange value of complex numbers that emerged 500 years ago when people were learning to solve cubic equations.

Question 2

We can now move to two topics: an understanding of some of our previous results in population modelling and then a look at classifying behaviour in linear problems.

Finding eigenvalues

$\det\big(\boldsymbol{M} - \lambda\boldsymbol{I}\big) = 0$ gives the eigenvalues λ_i of $\boldsymbol{M}$.

And determinants are found using the rules

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$$

or an expansion using other rows or columns.

(a) A population of young and adults was seen to evolve from year n to year $n+1$ according to

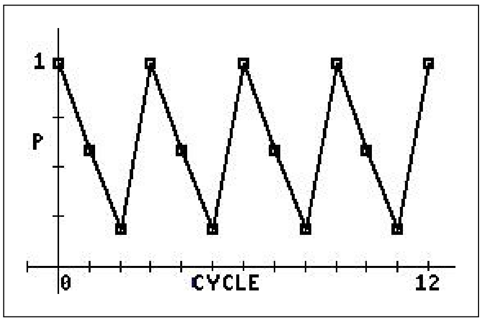
$$\begin{bmatrix} y_{n+1} \\ a_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & b \\ m & s \end{bmatrix} \begin{bmatrix} y_n \\ a_n \end{bmatrix},$$

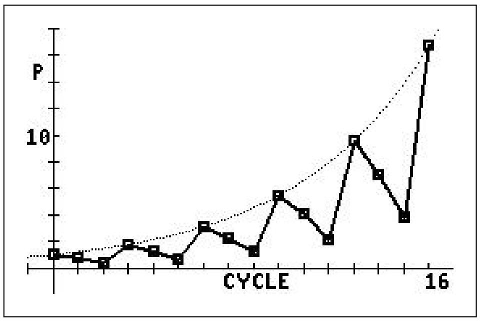
where b is birthrate, m is maturing rate and s is survival rate for adults. Previously we saw how populations grow or decline, depending on whether the dominant (real) eigenvalue is > 1 or < 1. Show that such an ultimate growth or decline must occur for the above population model, as b, m and s are all positive numbers and so there can be no complex eigenvalues.
Hint: Find a formula for λ and show it always gives real answers.

(b) In the lab *Population Models: Matrices and Eigenvalues*, you explored how populations evolved when they had age structure (e.g. young, adults and seniors), and the transitions were described by Leslie matrices. For the case with

$$\boldsymbol{T} = \begin{bmatrix} 0 & 0 & 8 \\ 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \end{bmatrix},$$

you found oscillations in the population as it evolved. Here are two examples of plots of the total population.





What do you think this means about the eigenvalues of $\boldsymbol{T}$? Check your guess by finding them.

(c) *Because the characteristic equation for eigenvalues is a polynomial and we now know what roots of polynomials can be, we can actually classify all behaviours of linear matrix problems. To finish this lab, let's do a little of that for 2×2 matrices.*

We saw that real eigenvalues can be associated with scaling vectors — stretching or shrinking them. Our example in Question 1 above shows that complex eigenvalues can also introduce vector rotations.

Sketch out what sort of different diagrams you would have got in (a) if it had turned out that the modulus of the complex eigenvalues satisfied

$$r>1 \qquad \text{or} \qquad r=1 \qquad \text{or} \qquad r<1$$

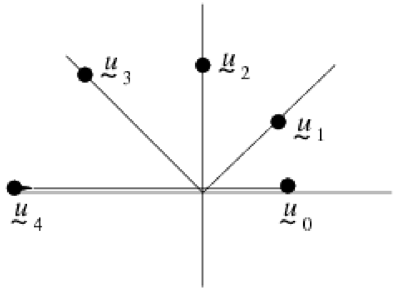
(Just put a spot for the ends of the vectors.)

Indicate also on other sketches the effects of finding larger and smaller θ values.

Instructors' Guide

Solutions

1. (a)



Each multiplication by $\boldsymbol{M}$ rotates the vector by $\pi/4$ and increases its length by $\sqrt{2}$.

(b) $\boldsymbol{M}\boldsymbol{w}_1 = \boldsymbol{M}\begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} 1+i \\ 1-i \end{bmatrix} = (1+i)\begin{bmatrix} 1 \\ -i \end{bmatrix} = \lambda_1 \boldsymbol{w}_1.$

$\boldsymbol{M}\boldsymbol{w}_2 = \boldsymbol{M}\begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} 1-i \\ 1+i \end{bmatrix} = (1-i)\begin{bmatrix} 1 \\ i \end{bmatrix} = \lambda_2 \boldsymbol{w}_2.$

(c) The modulus of λ_1 and λ_2 is $\sqrt{2}$. The argument of λ_1 is $\arctan(1)=\pi/4$ and that of λ_2 is $\arctan(-1)=-\pi/4$. Therefore,

$$\begin{aligned}
\lambda_1 &= \sqrt{2}\left(\cos\left(\frac{\pi}{4}\right)+i\sin\left(\frac{\pi}{4}\right)\right) \\
\lambda_2 &= \sqrt{2}\left(\cos\left(\frac{-\pi}{4}\right)+i\sin\left(\frac{-\pi}{4}\right)\right) \\
&= \sqrt{2}\left(\cos\left(\frac{\pi}{4}\right)-i\sin\left(\frac{\pi}{4}\right)\right).
\end{aligned}$$

(d) Setting $\boldsymbol{u}_0 = a\boldsymbol{w}_1 + b\boldsymbol{w}_2$ and solving the two simultaneous equations for a and b gives $\boldsymbol{u}_0 = (\boldsymbol{w}_1 + \boldsymbol{w}_2)/2$.

(e) We have, with $r=\sqrt{2}$, $\theta=\pi/4$ and using de Moivre's theorem,

$$\begin{aligned}
\boldsymbol{u}_n &= \boldsymbol{M}^n \boldsymbol{u}_0 \\
&= \frac{1}{2}\left(\boldsymbol{M}^n\boldsymbol{w}_1 + \boldsymbol{M}^n\boldsymbol{w}_2\right) \\
&= \frac{1}{2}\lambda_1^n\boldsymbol{w}_1 + \frac{1}{2}\lambda_2^n\boldsymbol{w}_2 \\
&= \frac{1}{2}\Big(r\big(\cos\theta+i\sin\theta\big)\Big)^n\boldsymbol{w}_1 + \frac{1}{2}\Big(r\big(\cos(-\theta)+i\sin(-\theta)\big)\Big)^n\boldsymbol{w}_2 \\
&= \frac{1}{2}r^n\big(\cos(n\theta)+i\sin(n\theta)\big)\boldsymbol{w}_1 + \frac{1}{2}r^n\big(\cos(-n\theta)+i\sin(-n\theta)\big)\boldsymbol{w}_2 \\
&= \frac{1}{2}r^n\big(\cos(n\theta)+i\sin(n\theta)\big)\boldsymbol{w}_1 + \frac{1}{2}r^n\big(\cos(n\theta)-i\sin(n\theta)\big)\boldsymbol{w}_2
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2}r^n\Big(\big(\cos(n\theta)+i\sin(n\theta)\big)\begin{bmatrix}1\\-i\end{bmatrix}+\big(\cos(n\theta)-i\sin(n\theta)\big)\begin{bmatrix}1\\i\end{bmatrix}\Big)\\
&= \frac{1}{2}r^n\begin{bmatrix}2\cos(n\theta)\\2\sin(n\theta)\end{bmatrix}\\
&= r^n\begin{bmatrix}\cos(n\theta)\\\sin(n\theta)\end{bmatrix}.
\end{aligned}
$$

(f) We have $R_n = r^n$, so that R is multiplied by $r=\sqrt{2}$, and $\beta_n = n\theta$, so that β increases by $\theta=\pi/4$, each time n increases by 1. Increasing n by 1 corresponds to multiplying $\boldsymbol{u}_n$ by $\boldsymbol{M}$, as we did in (a).

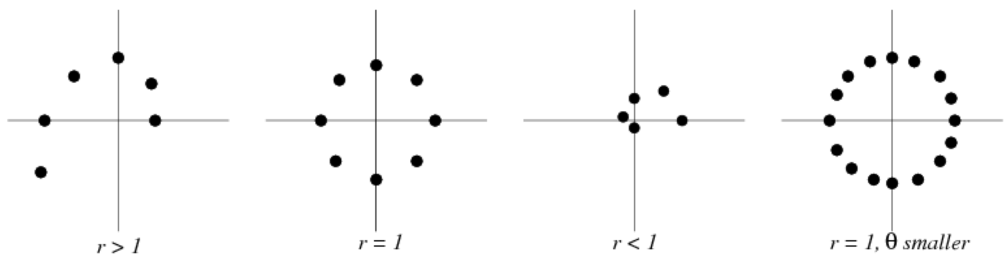
2. **(a)** We find that $\lambda = \big(s\pm\sqrt{s^2+4bm}\big)/2$. As $s^2+4bm>0$, we always have real eigenvalues.

(b) The total population is oscillating, indicating that the eigenvalues of $\boldsymbol{T}$ are complex. In the first case, the modulus of the dominant eigenvalue is 1 because the amplitude is constant; in the second case, the modulus is larger than 1 because the amplitude is increasing.

$\big|\boldsymbol{T} - \lambda\boldsymbol{I}\big| = 0$ leads to $\lambda^3-1=0$. Hence, we obtain a pair of complex eigenvalues, each with modulus 1:

$$
\begin{aligned}
\lambda &= 1,\quad \operatorname{cis}(2\pi/3),\quad \operatorname{cis}(4\pi/3)\\
&= 1,\quad -\frac{1}{2}+\frac{\sqrt{3}}{2}i,\quad -\frac{1}{2}-\frac{\sqrt{3}}{2}i.
\end{aligned}
$$

(c) Diagrams like



When θ is smaller, each rotation is less, so the dots are closer together.

3.6 Using Probability to Model Championship Tennis

Adapted from *Championship Tennis as a Probabilistic Modelling Context* by Peter Galbraith, Teaching Mathematics and its Applications 15 (4): 161–166 (1996).

Introduction

We are familiar with the common practice of radio and television commentators relaying statistics in the course of a match.

For example, 'Graf is hitting 73% of first serves into play' or 'Seles is only winning 50% of her second serves'.

This upsurge in statistical interest has led to the publication of all kinds of analytical information regarding match play. *The Australian* newspaper published the data in the table below relating to the women's final in the 1995 US Open, in which Steffi Graf defeated Monica Seles 7–6 (8–6) 0–6 6–3.

Note that it's NOT necessary for you to understand all the details in the table.

Final Statistics: US Open Women's Final 1995

Statistic	**Graf**	**Seles**
First serves in	73%	72%
Aces	7	8
Double faults	4	4
Wins from first serve, given first serve is in	72%	69%
Wins from second serve	73%	50%
Winners (including aces)	24	38
Unforced errors	32	42
Break points converted	17%	50%
Total points won	96	95

In this lab you will:

- use some of the data in the table to estimate the probability that Graf/Seles wins a service *point*;
- use these service-point results to estimate the probability that Graf/Seles wins a service *game*;
- use these service-game results to estimate the probability that Graf wins a set 6–3 when serving first, as in the final set of the 1995 US Open; and
- write a brief interpretation of some of your calculations.

You might like to review *Useful Results* on page 219 before starting the questions.

Question 1 *Probability of winning a service point*

Let p be the probability of a player *winning a service point.*

Let r be the probability that a player's *first serve is 'in'.*

Let f be the probability that a player *wins a service point if her first serve is 'in'.*

Let s be the probability that a player *wins a service point on her second serve*, i.e. the probability that a player wins a service point, given her first serve is 'out'.

(a) Use the data in the table above to estimate r, f and s for Seles and for Graf.

(b) Use the above definitions and the Law of Total Probability to show that the probability of winning a service point is given by

$$p = rf + (1-r)s = rf + s - rs.$$

Draw a tree diagram.

(c) Now use your results in (a) and (b) to obtain estimates for Seles' and Graf's probabilities of winning their service points.

(d) Hence find q, the probability of losing a service point, for Seles and Graf.

(e) Clearly both players would like to increase the probability of their winning a service point. Investigate (and explain) what happens to p for each player if she concentrates on one particular aspect of her serve:

(i) getting more first serves in, that is increasing r, say by 0.05;

(ii) improving her second serve, that is increasing s, say by 0.05.

What would you recommend to each player?

Some Assumptions

In order to model outcomes of games, we assume that elite players such as Graf and Seles exhibit a level of consistency, i.e.

- the probability of winning/losing a service point remains constant from point to point;
- each point may be considered to be independent of the preceding point.

You might like to think about whether these assumptions are valid.

Question 2 *Probability of winning a service game*

A player can win a service game 'to 0' (called winning 'to love') or 'to 15' or 'to 30' or 'to 40'. That is, if a player wins her service game, her opponent could have scored '0' or '15' or '30' or '40'. (If both players reach '40', it is called 'deuce'.)

We can calculate the probability of winning a service game 'to 15' using the negative binomial distribution (page 219) as follows.

$$\begin{aligned} P(\textit{wins service to 15}) &= P(\textit{server wins 3 of 1st 4 points } \underline{\textit{and}} \textit{ wins next point}) \\ &= P(\textit{server wins 3 of 1st 4 points}) \times P(\textit{server wins point}) \\ &= \binom{4}{3} p^3 q \times p \\ &= 4p^4 q, \end{aligned}$$

as $P(\textit{wins service point}) = p$, $P(\textit{loses service point}) = q$, and the number of wins on 4 serves is distributed binomially with 4 trials and probability of success p.

It can be shown that (see the article cited on page 215)

$$P(\textit{wins service to 40}) = \frac{20p^5q^3}{1-2pq}.$$

(a) Using the derivation of $P(\textit{wins service to 15})$ as a guide, calculate the probabilities of winning a service game 'to 0' and 'to 30'.

(b) Use the above results and Useful Result 2 to find an expression for the probability of winning a service game in terms of p and q. Check your expression using the values $p = 0, \frac{1}{2}, 1$.

(c) Hence estimate Seles' and Graf's probabilities of winning their service games.
Answers (to 2 significant digits): 0.81; 0.93.

PTO

We can use our results in Questions 1 and 2 to estimate probabilities of the outcomes of various sets in the 1995 US Open Women's Final. Here is one example.

Question 3 *Probability that Graf wins a set 6–3, serving first*

Use your result in 2(c) to calculate the probability that Graf wins the deciding set 6–3, given that she served in the first game of the set. *Answer*: 0.31 (to 2 significant digits).

Some Hints

- Let p_G and p_S be the probabilities that Graf and Seles win their respective service games.
- We have, using Useful Result 1,

$$\begin{aligned} P(\textit{Graf wins 6–3}) &= P(\textit{Graf leads 5–3 and Graf wins the next game}) \\ &= P(\textit{Graf leads 5–3}) \times P(\textit{Graf wins the next game}). \end{aligned}$$

  Note that the score after the previous (eighth) game must have been 5–3. If it had been 6–2, Graf would have already won the set.

- How can Graf be leading 5–3?
  - In these eight games, each player has four service games.
  - There are *four* ways in which Graf can be leading 5–3. Note the use of the binomial distribution (page 219).
  - *One* way is for Graf to win 4 service games, win one Seles service game and lose 3 Seles service games.
    * That is, Graf wins all 4 of her service games (with probability p_G^4) and Seles wins 3 of her 4 service games $\left(\text{with probability } \binom{4}{3} p_S^3\, q_S\right)$.
    * Thus $P(\textit{Graf wins 4 service games, wins one Seles service game and loses 3 Seles service games}) = \binom{4}{3} p_S^3\, q_S\, p_G^4$.
  - *Another* way is for Graf to win 3 out of 4 service games, win 2 Seles service games and lose 2 Seles service games.
    * That is, Graf wins 3 of her 4 service games $\left(\text{with probability } \binom{4}{3} p_G^3\, q_G\right)$ and Seles wins 2 of her 4 service games $\left(\text{with probability } \binom{4}{2} p_S^2\, q_S^2\right)$.
    * Thus $P(\textit{Graf wins 3 service games, wins 2 Seles service games and loses 2 Seles service games}) = \binom{4}{3} p_G^3\, q_G \times \binom{4}{2} p_S^2\, q_S^2$.
  - Now you work out the probabilities for the other *two* ways, and hence (using Useful Result 2) $P(\textit{Graf leads 5–3})$.

Scenario

Monica Seles and Steffi Graf agree to participate in a special *New Millennium* charity tennis match here. Your lab group is chosen to be the statistical advisors to the **SPORT** Television commentary team. By considering ANY of your calculations in this lab, write a couple of sentences to be used by the **SPORT** commentators in their preliminary discussion of the match.

Useful Results

1. $P(A \cap B) = P(A)\,P(B)$ if A and B are independent.

2. $P(\cup A_i) = \sum P(A_i)$ when all the A_i are mutually exclusive.

3. *The Law of Total Probability*: If events E_1, E_2, ... form a partition of sample space S, then for any event A,

$$P(A) \;=\; \sum P(A \cap E_i) \;=\; \sum P(E_i)\,P(A|E_i).$$

4. *The binomial distribution*: If X is the number of successes in n independent Bernoulli trials (only two outcomes: success or failure) with probability of success p on each trial,

$$P(X\!=\!k) \;=\; \binom{n}{k} p^k q^{n-k} \quad k = 0, 1, \ldots, n.$$

5. *The negative binomial distribution*: If N is the number of attempts needed to achieve k successes in a sequence of Bernoulli trials, then

$$\begin{aligned} P(N = j) \;&=\; P(k-1 \text{ successes in the first } j-1 \text{ trials})\,P(\text{success on the } j\text{th}) \\ &=\; \binom{j-1}{k-1} p^{k-1} q^{(j-1)-(k-1)} \cdot p \\ &=\; \binom{j-1}{k-1} p^k q^{j-k} \quad j = k,\, k+1,\, \ldots. \end{aligned}$$

Instructors' Guide

Solutions

1. *Probability of winning a service point*

(a) **Seles:** $r=0.72 \quad f=0.69 \quad s=0.5$. **Graf:** $r=0.73 \quad f=0.72 \quad s=0.73$.

(b) $P(\textit{wins service point})$

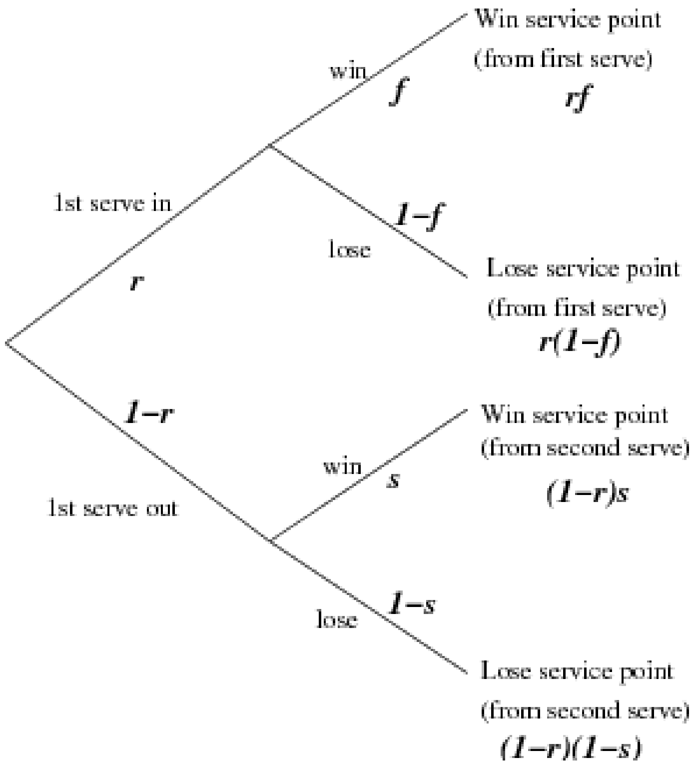
$= P(\textit{wins service point on 1st serve}) + P(\textit{wins service point on 2nd serve})$

$= P(\textit{wins service point/1st serve in}) \times P(\textit{1st serve in}) +$ Law of

$\quad P(\textit{wins service point/1st serve out}) \times P(\textit{1st serve out})$ Total Probability

$= f \times r + s \times (1-r)$.

Therefore, $p = rf + (1-r)s = rf + s - rs$, as required.



(c) Using (a),

$$P(\textit{Seles wins service point}) = 0.72 \times 0.69 + (1-0.72) \times 0.50 = 0.6368.$$

$$P(\textit{Graf wins service point}) = 0.73 \times 0.72 + (1-0.73) \times 0.73 = 0.7227.$$

(d) $P(\textit{loses service point}) = 1 - P(\textit{win service point})$, so that

$P(\textit{Seles loses service point}) = 1-0.6368 = 0.3632$.

$P(\textit{Graf loses service point}) = 1-0.7227 = 0.2773$.

(e) We have $p = rf + s - rs$.

Seles: $r{=}0.72,\ f{=}0.69,\ s{=}0.5$.

Fix $f = 0.69$ and $r = 0.72$	$s = 0.5$	$\Rightarrow$	$p = 0.6368$
	$s = 0.55$	$\Rightarrow$	$p = 0.6508$
Fix $f = 0.69$ and $s = 0.5$	$r = 0.72$	$\Rightarrow$	$p = 0.6368$
	$r = 0.77$	$\Rightarrow$	$p = 0.6463$

Therefore, increasing s (the probability of winning a second serve) by 0.05 increases p more than by increasing r (the probability that the first serve is 'in') by 0.05.

Graf: $r{=}0.73,\ f{=}0.72,\ s{=}0.73$.

Fix $f = 0.72$ and $r = 0.73$	$s = 0.73$	$\Rightarrow$	$p = 0.7227$
	$s = 0.78$	$\Rightarrow$	$p = 0.7362$
Fix $f = 0.72$ and $s = 0.73$	$r = 0.73$	$\Rightarrow$	$p = 0.7227$
	$r = 0.78$	$\Rightarrow$	$p = 0.7222$

Therefore, increasing s by 0.05 increases p more than by increasing r by 0.05. In fact, for Graf, increasing r by 0.05 actually reduces p.

If students have covered partial differentiation, they can derive the following general results.

$$\Delta p \approx \frac{\partial p}{\partial s}\Delta s = (1-r)\Delta s \qquad \Delta p \approx \frac{\partial p}{\partial r}\Delta r = (f-s)\Delta r$$

Putting in the respective numbers for the two players gives, *for Seles*,

$$\Delta p \approx 0.28\Delta s \qquad \Delta p \approx 0.19\Delta r,$$

showing that, for an equal increment of r and s, p increases more when s is increased.

For Graf, $\Delta p \approx 0.27\Delta s$ and $\Delta p \approx -0.01\Delta r$, again confirming our findings above.

Therefore, Seles and Graf should concentrate on increasing s, that is improving their **second** serves.

We could also look at the effect of increasing f, the probability of winning a first serve given that it is 'in'. Clearly r and f are not independent. Interesting discussion?

2. *Probability of winning a service game*

(a) Following the derivation of $P(\textit{win serve to 15})$ in the lab sheet,

$$\begin{aligned} P(\textit{wins serve to 0}) &= P(\textit{server wins all 4 service points}) \\ &= p^4 \qquad \text{(special case of negative binomial distribution).} \end{aligned}$$

$$\begin{aligned} P(\textit{wins serve to 30}) &= P(\textit{server wins 3 of 1st 5 points and wins next point}) \\ &= P(\textit{server wins 3 of 1st 5 points}) \times P(\textit{server wins point}) \\ &= \binom{5}{3} p^3 q^2 \times p \qquad \text{(negative binomial distribution)} \\ &= \binom{5}{3} p^4 q^2 \\ &= 10p^4q^2. \end{aligned}$$

(b) The probability of winning a service game is the sum of the probabilities of the different ways that a player can win a service game (Useful Result 2).

$$\begin{aligned} P(\textit{wins service game}) &= P(\textit{wins service game to 0 or 15 or 30 or deuce}) \\ &= p^4 + 4p^4q + 10p^4q^2 + \frac{20p^5q^3}{1-2pq}. \end{aligned}$$

Note: $P(\textit{wins service game}) = \begin{cases} 0 & \text{when } p = 0 \\ 0.5 & p = 0.5 \\ 1 & p = 1 \end{cases}.$

(c) For Seles, $p = 0.6368$ from 1(d), so that $q = 1 - p = 0.3632$. Putting these values in the above expression gives

$$p_S = P(\textit{Seles wins service game}) \approx 0.81.$$

For Graf, $p = 0.7227$ from 1(d), so that $q = 0.2773$ and

$$p_G = P(\textit{Graf wins service game}) \approx 0.93.$$

The simplest and least-error-prone method of calculating the above numbers is to store the values of p and q in memories P and Q of the calculator, then type out the expression for $P(\textit{wins service game})$ in terms of P and Q and evaluate it. Store the other set of values for p and q and use the recall facility on the calculator to evaluate $P(\textit{wins service game})$ for the other player.

3. *Probability that Graf wins a set 6–3, serving first*

To find $P(\textit{Graf leads 5–3})$, we use two results: Useful Result 1 and the binomial distribution.

$P(\textit{Graf wins 4 service games and Seles wins 3 out of 4 service games})$

$$= \binom{4}{3} p_S^3\, q_S\, p_G^4.$$

$P(\textit{Graf wins 3 out of 4 service games and Seles wins 2 out of 4 service games})$

$$= \binom{4}{3} p_G^3\, q_G \times \binom{4}{2} p_S^2\, q_S^2.$$

$P(\textit{Graf wins 2 out of 4 service games and Seles wins 1 out of 4 service games})$

$$= \binom{4}{2} p_G^2\, q_G^2 \times \binom{4}{1} p_S\, q_S^3.$$

$P(\textit{Graf wins 1 out of 4 service games and Seles wins 0 out of 4 service games})$

$$= \binom{4}{1} p_G\, q_G^3 \times q_S^4.$$

Therefore, using Useful Result 2, we have

$$P(\textit{Graf leads 5–3}) = \binom{4}{3} p_S^3\, q_S\, p_G^4 + \binom{4}{3}\binom{4}{2} p_G^3\, q_G\, p_S^2\, q_S^2 + \binom{4}{2}\binom{4}{1} p_G^2\, q_G^2\, p_S\, q_S^3 + \binom{4}{1} p_G\, q_G^3\, q_S^4.$$

By substituting $p_S = 0.81,\ q_S = 0.19,\ p_G = 0.93,\ q_G = 0.07$, we obtain

$$P(\textit{Graf leads 5–3}) \approx 0.335.$$

Therefore, using Useful Result 1, we have

$$\begin{aligned} \therefore\ P(\textit{Graf wins 6–3}) &= P(\textit{Graf leads 5–3}) \times P(\textit{Graf wins next game}) \\ &\qquad\qquad\qquad\qquad\qquad \text{(her service game)} \\ &\approx 0.335 \times p_G \\ &\approx 0.335 \times 0.93 \\ &\approx 0.312. \end{aligned}$$

4 Lab Manual

Aims of labs

To give students the opportunity to

1. **come to a greater depth of understanding of the course material through: the problems chosen; the time spent on these problems; being able to get help from both their peers and staff.**
2. **be exposed to and raise their ability/confidence to tackle, in groups, problems of greater complexity and difficulty than we could expect them to tackle individually.**
3. **practise oral and written *mathematics-related* communication skills.**

4.1 Lab organisation

- Lab work will be in groups, usually of four people. Each group will be expected to work together on a set of questions handed out at the beginning of the lab.
- For the first lab, you will be in a randomly assigned group. For the remaining labs, you will then be asked to form your own groups, which will remain the same for the rest of the semester.
- For the labs you will need a calculator each, pens, pencils, rulers, paper to write on and *at least one copy of the text and/or lecture notes per group.*
- Each group will be required to produce a lab report *by the end of each lab session.* We expect the group to work together and *one* person to scribe. **Questions scribed by someone other than the original scribe will not be marked.** The job of scribing should rotate week by week among the members of the group. We keep a record of who scribes each lab.
- *Each member of the group must sign his or her name at the end of the report.*
- Lab reports will be collected at the end of the lab, marked and returned with comments, usually by the next lab period. Each member of the group will receive the group mark for the lab. Each lab will be worth 2% of your final mark. You will also receive a mark for how you contributed to the group (see Formal Groupwork Evaluation).

4.2 Lab reports

- The lab report should be a thoughtful, well-written and neatly organised document that summarises both your experiences in the lab and what you learned as a result of that experience. The questions in the lab sheets will help to guide you as to what to do (and should be answered), but you should feel free to make any observations or comments that you think are appropriate. We will be giving you some guidance in writing lab reports.
- We expect reports to be written in proper English, with correct spelling and appropriate punctuation. Any mathematics should read as English sentences when the symbols are read out. *Lab reports not written in this manner are likely to be viewed less than favourably by the marker.* There is no fixed format for lab reports, but some of the points below may help you in your writing up.
- **Introduction:** This summarises briefly what the lab is about and how you are going to tackle it in the report. Sometimes it is easier to write the Introduction at the end, though it should of course come first in your report.
- **Graphs:** The calculators take all the hard work out of these. In the questions, the words *plot* or *graph* means graph it on your calculator. The words *sketch* or *draw* mean sketch it in your report. We don't expect a work of art or accuracy down to the micron level: a reasonable sketch on ordinary paper showing *labelled axes with scales* and the *salient features of the graph* is fine. Graphs should also have a short caption explaining what they are about.
- **Calculations:** If you do some calculations, present them succinctly. Explain briefly what you did, perhaps by writing out one typical calculation in full.
- **Data:** Summarise any data you collect in a succinct, easy-to-grasp form such as a table or a diagram with labels. Keep in mind that we are interested in your answers, thoughts and analysis, rather than in lots of numbers.
- **Conclusions:** We'd like to see these and will often prompt for them in the questions. They should be the inferences that you draw from your data and calculations. Here is your chance to show that you understood the purpose of the lab, saw patterns in the data and gained significant insights.

4.3 What the mark on your lab report means

Marks are given for completeness and correctness of answers, and for presentation. By presentation, we mean things like

- appropriate use of figures, tables, graphs, equations and explanatory remarks
- answers in clear and complete English or mathematical sentences
- descriptive and meaningful sentences to guide a reader through the report
- graphs with scale marks on the axes and labelled axes and curves
- data succinctly summarised in tables
- legible writing.

GRADE	**MARK**	**STANDARD REQUIRED**
High Distinction	$\geqslant 8.5$	Excellent presentation. Virtually every question done, answered well, clearly and correctly. *Outstanding.*
Distinction	7.5 – 8	Presentation very good. Clear answers. Most questions done with few errors. *Very Good.*
Credit	6.5 – 7	Good presentation. Many questions done, with most done fairly well. *Pretty reasonable to fairly good.*
Pass	5 – 6	Adequate presentation. Reasonable number of questions attempted. More right than wrong. *Barely acceptable to just OK.*
Fail	< 5	Poor presentation. Many errors. Many questions not answered at all. *Awful.*

4.4 Sample lab reports — good vs poor

Exercise: For the following lab, two reports were handed in.

1. Critique each report for its

- **Readability:** Can it be read as a stand-alone document? i.e. without either continually referring to the lab questions or knowledge of the complete worked solution.
- **Clarity:** Can you understand (relatively easily) what's been said and done?
- **Completeness:** Has the report completely answered all questions?
- **Correctness:** Are the explanations and calculations correct?
- **Presentation:** Is the material well organised? Is it easy to find things?

Basically, a good report would be clear and informative to a *student* who missed the lab for some reason.

2. According to the above criteria, identify at least six things that make the better of the two reports better.

Lab X: Snow White and the Seven Dwarves revisited — was the prince a hero or just lucky?[58]

The problem

According to the tale of *Snow White and the Seven Dwarves*, Snow White collapsed comatose to the ground immediately after eating the poisoned apple given to her by her wicked stepmother, the queen. The dwarves found her one hour later. After three days, Snow White's condition remained unchanged. Fearing her dead, the dwarves carried her to the centre of the forest where they left her in a glass coffin. Seven days after she had eaten the apple, a prince happened by and Snow White awoke to his kiss. Was the kiss magical or was the prince just lucky?

Additional information

Assume that

- the concentration C of the drug in Snow White's bloodstream as a function of time t can be modelled by the equation $C(t) = ate^{-bt}$ for some positive constants a and b.
- the drug reached the minimum effective concentration in the minimum time allowable physiologically, 30 seconds according to the medicos.

[58] Based on an after-dinner Maths-conference speech by Ansie Meiring.

Report 1

Lab X

Given $C(t) = ate^{-bt}$

$$C(1) = C(73)$$

$$ae^{-b} = 73ae^{-73b}$$

$$73e^{-72b} = 1$$

$$-72b = \ln(1/73)$$

$$b = \frac{\ln(73)}{72} = 0.05959$$

$$C(t_r) = C_{min} = C(1/120)$$

$$at_r e^{-0.05959t_r} = a\frac{1}{120}e^{-0.05959/120}$$

$$t_r e^{-0.05959t_r} = 0.008329$$

Plotting on calculator and using TRACE, $t_r = 166 = 6.92$ days.

Snow White was better before the prince arrived so there was nothing magical about his kiss. He was just lucky!!

Report 2

Lab X: Snow White and the Seven Dwarves revisited — was the prince a hero or just lucky?

Introduction

The concentration C of the drug in Snow White's bloodstream was given by

$$C(t) = ate^{-bt} \qquad (1)$$

for some constants a and b, so we plotted this function for some sample values of a and b, and found it always had the basic shape shown in Fig. 1 below.

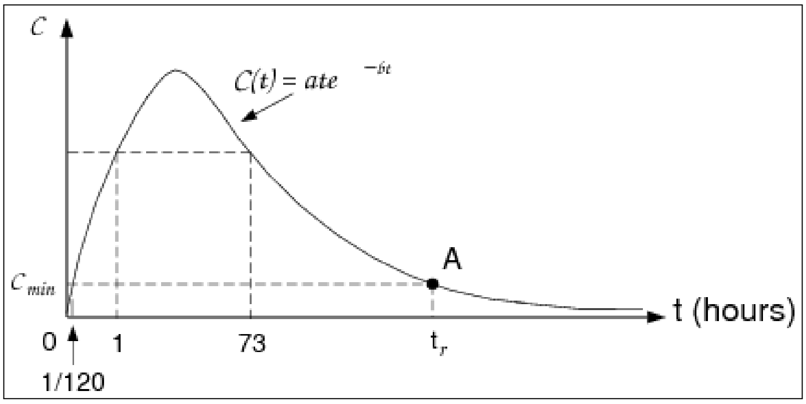


Figure 1: Concentration of drug in Snow White's blood as a function of time. Note that the time axis is not to scale so that the relevant points may be seen clearly.

After discussing the problem for a while, we came to the conclusion that the prince would have been a hero if the concentration of the drug was still above its minimum effective value when he found Snow White. However, if it had fallen *below* its minimum effective value when he found her, then Snow White would have been just resting peacefully at the time and so his kiss would have just woken her from normal sleep. The problem thus reduces to: "Is the time shown at the point A in Fig. 1 before or after 7 days have elapsed?"

Assumptions

In order to solve the re-posed problem, we need to determine the exact shape of the curve, i.e. we need to determine the constants a and b in Eq. (1). To do this, we need some points on the curve $C(t)$.

From the given information, we will assume that:

1. $C_{\min} = C(1/120)$, where $C_{\min}$ is the minimum effective concentration of the drug, and this is reached when $t = 1/120$ hour (i.e. 30 seconds). Note that we have set $t = 0$ to be the time when Snow White ate the apple and that the unit of time has been chosen to be hours.

2. From the fact that Snow White's condition was unchanged 3 days after the dwarves found her, we will assume that $C(1) = C(73)$ (3 days = 72 hours and the dwarves found her 1 hour after she ate the apple).

Calculations

From Eq. (1), Assumption 2 implies that

$$\begin{aligned} ae^{-b} &= 73ae^{-73b} \\ &\Rightarrow 73e^{-72b} = 1 \\ &\Rightarrow -72b = \ln(1/73) \\ &\Rightarrow b = \frac{\ln(73)}{72} \approx 0.05959. \qquad (2) \end{aligned}$$

Now let the recovery time t_r be the time when C has fallen back to $C_{\min}$, its minimum effective value. Therefore, from Assumption 1 above,

$$\begin{aligned} C(t_r) &= C_{\min} = C(1/120) \\ &\Rightarrow at_re^{-bt_r} = a\frac{1}{120}e^{-b/120} \\ &\Rightarrow t_re^{-bt_r} \approx 0.008329, \qquad (3) \end{aligned}$$

using b from Eq. (2) above.

Equation (3) cannot be solved algebraically, so we used our TI-84 calculators to obtain a numerical solution. Plotting $\mathrm{Y}_1 = \mathrm{X}e^{-0.05959\mathrm{X}}$ and $\mathrm{Y}_2 = 0.008329$ on our calculators with an appropriate window and zooming in, we found that the two curves intersect near $t = 166$ hours. Using the *intersect* function in the CALC menu of our calculator, we found the two curves to intersect at $t_r \approx 166.15$ hours or approximately 6.92 days. Thus after seven days, C will have dropped below C_{min}, and the drug will no longer be effective.

Conclusions

Our investigations have shown that some plausible assumptions lead to the conclusion that after seven days, the concentration of drug in Snow White's bloodstream would have dropped below its minimum effective amount. Therefore, when the prince found her, Snow White would have just been sleeping naturally. His kiss therefore, was not magical. He was just lucky!!

4.5 Groupwork

Objectives of learning in groups

- To promote learning by communication and interaction between the group members.
- To stimulate critical thinking.
- To promote learning by increasing the motivation of group members.
- To give group members the opportunity of expressing and clarifying their own ideas.
- To allow group members to pool and pass on information.
- To enhance oral communication skills.
- To allow group members to take more responsibility for their own learning.

Some benefits of doing mathematics in groups

Students find that they

- learn by explaining mathematics to their group in their own terms and consequently often develop a deeper understanding of the mathematical concepts through having to talk about them
- persist longer on a tough problem when they are working together (i.e. they don't give up as quickly)
- are better motivated and more productive because of the presence of others
- realise that a group discussion often generates a variety of ways of solving a problem and that a group solution may be better than that which any individual could have produced
- discover that their peers often have similar difficulties in mathematics
- become more willing to test their ideas and more willing to explore new ways of solving old problems
- become more willing to attempt a solution to a problem that they don't immediately know how to solve.

Over and above the value working in a group has on your learning in Maths, many of the above objectives lead to skills which are highly valued in the workforce. Many jobs in the workplace are either too big or too complex to be handled effectively by any one person. Hence the need to be able to tackle problems cooperatively in a group and to communicate solutions effectively to others. Employer groups are continually complaining that new graduates lack oral and written communication skills, and lack the ability to work effectively in teams.

Formal groupwork evaluation

Rationale

One characteristic of the most effective groups is that they periodically review how they are functioning and performing as a group, with the aim of improving their performance. To encourage you to do the same, we will give you help in improving your performance as group members, and once each semester there will be a formal evaluation of your groupwork performance. These evaluations will count for 2% of your total mark.

Procedure

You will be asked to rate your performance and the performance of the other members of your group, using the criteria and evaluation scheme shown on the next page. Your Groupwork Evaluation mark will be the mean of the total mark given to you by yourself and by each of the other members of your group. *Individual evaluations will be completely confidential.*

Is this a fair and valid assessment procedure?

Some concerns have been expressed as to the fairness and validity of this assessment scheme. We think it is fair and valid; the points below say why.

- Assessment of groupwork performance is better done by students than by staff, because it is the students in a group who know best how well each member is contributing to the performance of the group. Educational research has shown that, for some sorts of assessment tasks, provided students are given clear guidelines on what and how to mark, *and their marking is monitored*, they can assess each other just as well as a member of staff can, and in this case probably better. The information on the following page gives clear guidelines on what and how to mark. We check the marking, *and will not accept unreasonably high or low marks*, so the procedure is monitored.
- To allow for bias or differing standards of marking, we take the mean of marks given to a person by themselves and every other member of the group. If the marks vary wildly from person to person, we investigate the discrepancy and re-evaluate where necessary.
- From our perspective (with a few exceptions which we dealt with), this scheme has worked well in the past, and we believe the procedure to be reasonable and fair. Student surveys have also shown that a significant number of students find the feedback they have received from these evaluations to be quite helpful.

Groupwork Evaluation Sheet

Use the following five-point scale to evaluate the performance of yourself and the other members of your group. Use only integers 1 – 5.

5 = excellent (done regularly / is effective / stands out in a positive way)

3 = okay (done sometimes / reasonable, but room for improvement / doesn't stand out)

1 = poor (done rarely, if at all / needs lots of improvement / stands out in a negative way)

Lab participation

1. Contributes regularly to group discussions.

2. Involved in all aspects of completing the lab.

Oral communication skills

3. Explains ideas / concepts well.
(Explains things logically, clearly, simply.)

4. Is prepared to listen to others.
(Interrupts only for clarification; listens well: uses ideas of others as a springboard for further ideas.)

Leadership

5. Helps keep the group "on task" and "on track".
(Refocusses the group after some socialising; leads and structures discussions; suggests "plans of attack".)

6. Encourages all group members to contribute.
(Doesn't let one member of the group dominate; doesn't tolerate non-contributors.)

7. Monitors the work of the scribe.
(Checks clarity of expression; suggests and checks appropriate use of tables/graphs/figures/....)

Supportive of group members

8. Values the learning of all group members.
(Doesn't leave anyone behind; helps others to understand.)

9. Maintains positive working relationships.
(Provides positive feedback to others; supports innovation; criticises the ideas of others constructively, not personally.)

Your name	**The names of the other members of your group**		

4.6 Instructors' guide

First lab

1. As students walk in, they are given a number and a Lab Manual.
2. Students find their seat in the room according to the number they are given.
3. Introduce lab instructors.
4. Explain what labs are about (Lab Manual).
5. Explain that they are in random groups in this lab, but for the remaining labs they should choose their own groups (in which they will remain for the rest of the semester).
6. Hand out the lab and 'set them loose'.

Second lab

1. Form their own groups if not already done.
2. Briefly explain why we think working in groups is a good idea (more details in the Lab Manual for them to read at their leisure) and that we'll be stopping with 15 minutes to go so that we can discuss what makes an effective group effective.
3. Set them loose on the lab.
4. *Reminder*: Finishing up in 5 minutes.
5. *What makes an effective group effective?*

Discussion

- **To students:** You've now had experience with working in a group in a lab and undoubtedly most of you have had other experiences of working in groups. Some groups are more effective than others. What we'd like you to do now is, for the next 5 minutes, discuss with the other members of your group, what you think makes an effective group effective. Jot down your conclusions and at the end of the 5 minutes, we'll pool your ideas.
- **Five minutes later:** Ask a variety of groups for one idea that they came up with. Write on the OHP or whiteboard. Keep asking until no new ideas are forth coming.
- Point out the key themes and tell students: "We'll distil these ideas and remind you about them during the semester". Tell them they'll be asked to evaluate each other on how effectively they work together as a group later in the semester. More details to be given about this in the next lab.

Third lab

1. Distribute notes on conclusions from the *What makes an effective group effective?* discussion.

2. Discuss the groupwork-evaluation questions.

Later lab

1. Students to come out to the front to fill in a Groupwork Evaluation Sheet. They are not to discuss with each other how they fill this in. Scores which are wildly inconsistent or totally unrealistic (from our perspective) will be investigated by a lab instructor.

2. Results to be returned next lab.

End of semester

Review how the scheme worked and whether any changes need to be made.

5 Programs

5.1 Copying programs

From the website

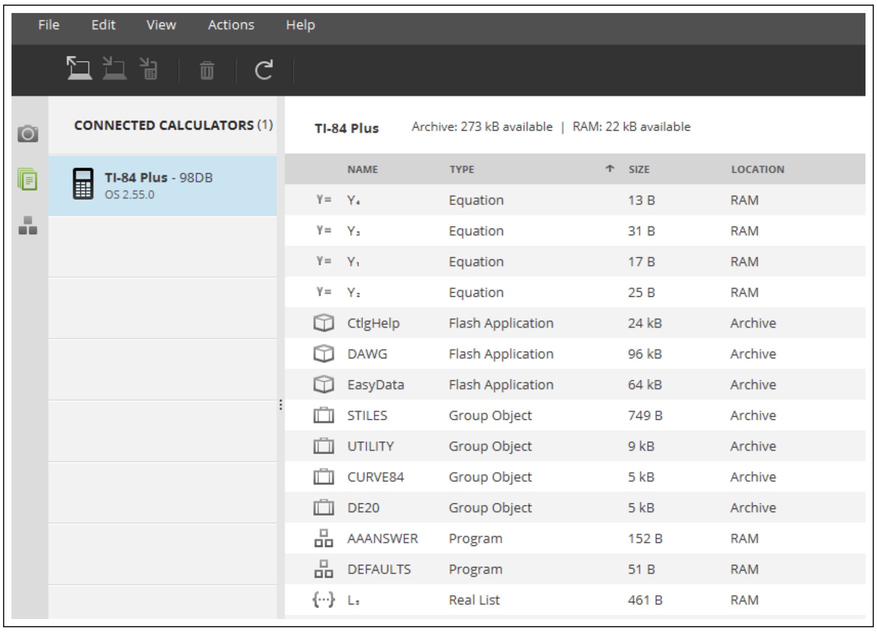
You will need either *TI Connect* or *TI Connect CE* for this. If your TI-84 has a mini-USB port,[59] use *TI Connect CE*, otherwise use *TI Connect*. Both are available for free at *education.ti.com*. Click on *Downloads* and follow through.

To copy a program from the website, open up/unzip the program folder on the website and drag the program you want (see Section 5.3 *Program Protection* below) onto your desktop or into another folder.

Connect your calculator to your computer with the appropriate cord and start *TI Connect* or *TI Connect CE*.

TI Connect CE

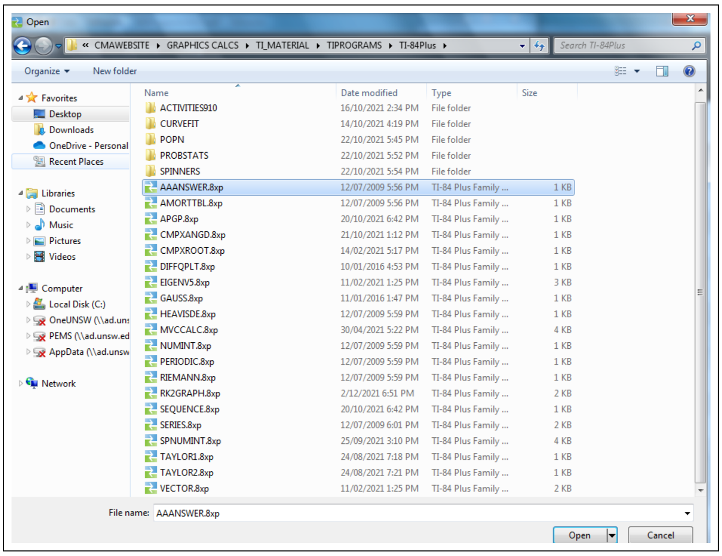
TI Connect CE should recognise your calculator. The screenshot below[60] shows *TI Connect CE*, set up for program transfers, with a TI-84 Plus calculator connected.



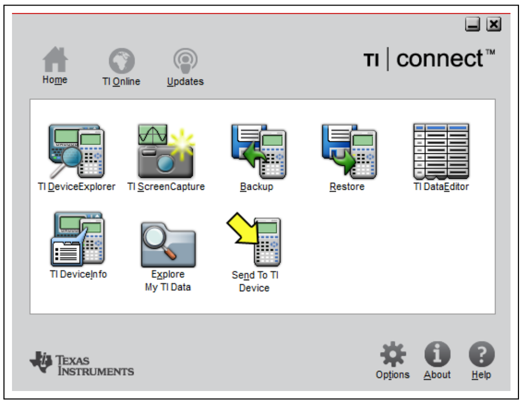
Click on the central icon of the three icons in the left-hand column; this is for transfers between computer and calculator.[61] As in the screenshot above, the contents of your calculator should be shown.

On the left of the second row of icons at the top of the screen is an icon for copying programs from your computer; click on this, select the desktop or folder where you have copied the program, then click on the program (screenshot over the page). Click on *Open* (bottom right of the screen) and follow the prompts to complete the transfer.

[59] all CE and 84Plus calculators except for some early model 84Plus
[60] taken from a PC running Windows 7
[61] The top icon is for downloading calculator screen shots, the bottom icon for editing programs.

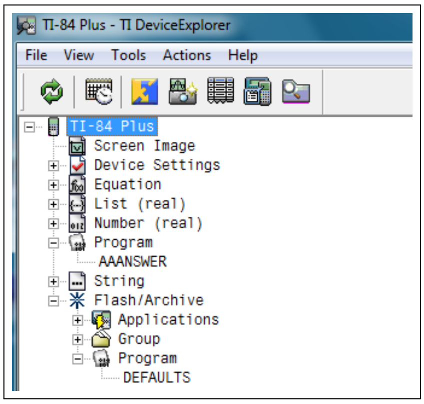


TI Connect



Click on *TI Device Explorer*. A window for your calculator should open, showing you what's on it (figure on the next page).

Click on the + sign next to *Program* to show the unarchived programs on the calculator. To see the archived programs (see Section 5.2), click on the + sign next to *Program* under *Flash/Archive*.



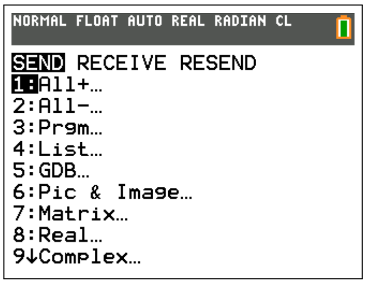
Drag the program from your desktop or folder where you have copied the program onto the calculator window, and the transfer should happen.

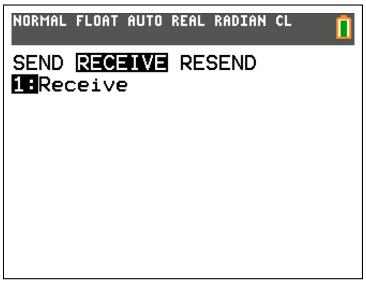
TI Connect will also work with the mini-USB cable and both calculator types but is slower and lacks the program editor of TI Connect CE.

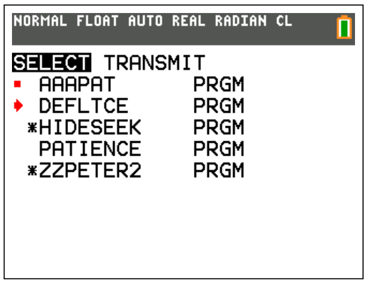
From another calculator

Do this using the [link] key ([2nd] [X,T,θ,n]). This gives a menu like that below left (TI-84CE).

The receiver moves the cursor to RECEIVE with the right arrow (below centre) and presses [enter]. This needs to be done before the sender transmits.







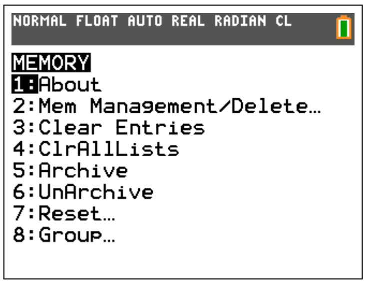
The sender presses [3] (Prgm...), selects the program(s) to be sent with the cursor and [enter] (above right), moves the cursor to TRANSMIT with the right arrow and presses [enter] to start transmission.

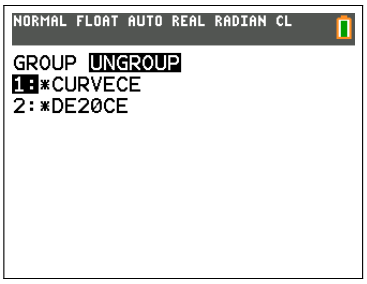
If a program is already on the receiving calculator, you will get a menu of options.

Press [quit] on both calculators to return to the Home screen.

Group programs

Sometimes, several programs are grouped together in a group file, for example when a program has several subroutines. This ensures that, when you copy the group file to your calculator, you have all the necessary programs.

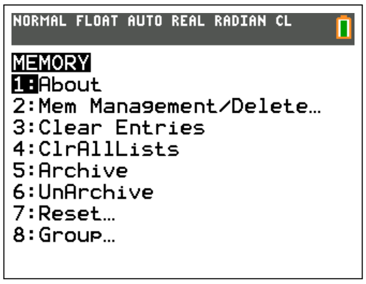


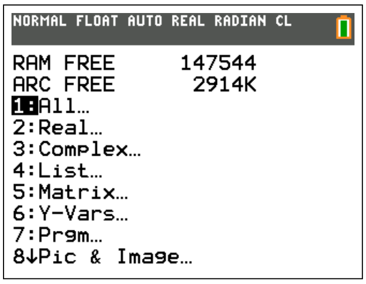


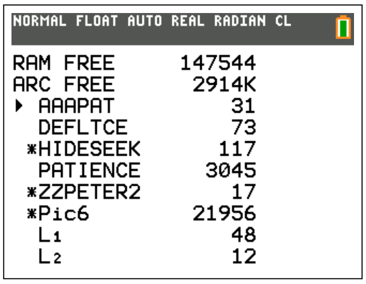
To recover the individual programs, press [mem] ([2nd] [+]) (above left), scroll down to *Group...* and press [enter]. Here you can create a new group file or ungroup an existing one (above centre). The latter is what we need here. Move the cursor to UNGROUP and press [enter]. The group file you have just copied to your calculator should be listed here (above right). Press the number against its name and all should happen. If you already have some of the programs on your calculator, a menu will appear giving you options.

5.2 Archiving and unarchiving

Programs on your calculator can be either in main memory (RAM), and available to run and edit, or archived (ARC) and not available to run (except on the latest CE) or edit. Archived programs are indicated by an asterisk in the program list.







To unarchive (or archive) a program, press mem (2nd +) 2 (Memory Management) (above left) and 1 (All...) (above centre). You will see a list there of everything on your calculator, starting with the programs (above right). To unarchive a program, move the cursor onto it and press enter; the asterisk should disappear. To archive it, press enter, which just toggles between the two states.

Press quit to return to the Home screen.

5.3 Program protection

Programs in the PROTECTED folders have been locked or protected. This means they can be run and deleted but not altered in any way. Programs in the other, corresponding, folders are not protected.

Alteration usually happens when the user stops a program using on but then (inadvertently) selects GoTo instead of Quit. Any subsequent key presses will alter the program, meaning it will not work as it should or at all. This is the greatest source of frustration in a class. The only cure is to copy the program again from someone else.

Using the protected version of the program means that this cannot happen. Recommended for use in class.

The downside of protecting a program is that the program cannot be read and perhaps modified on purpose. That is why both protected and unprotected programs are provided. The only way of knowing whether a program on your calculator is protected is to try to edit it; the calculator won't let you if it is.

Protecting and unprotecting a program is done in the editor of TI-Connect CE (but not TI-Connect). The editor is the bottom symbol in the column of three symbols on the left-hand side of the TI-Connect CE screen. Once a program is displayed, there is a box to click on to protect/unprotect it.