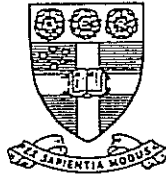


THE UNIVERSITY OF NEW ENGLAND

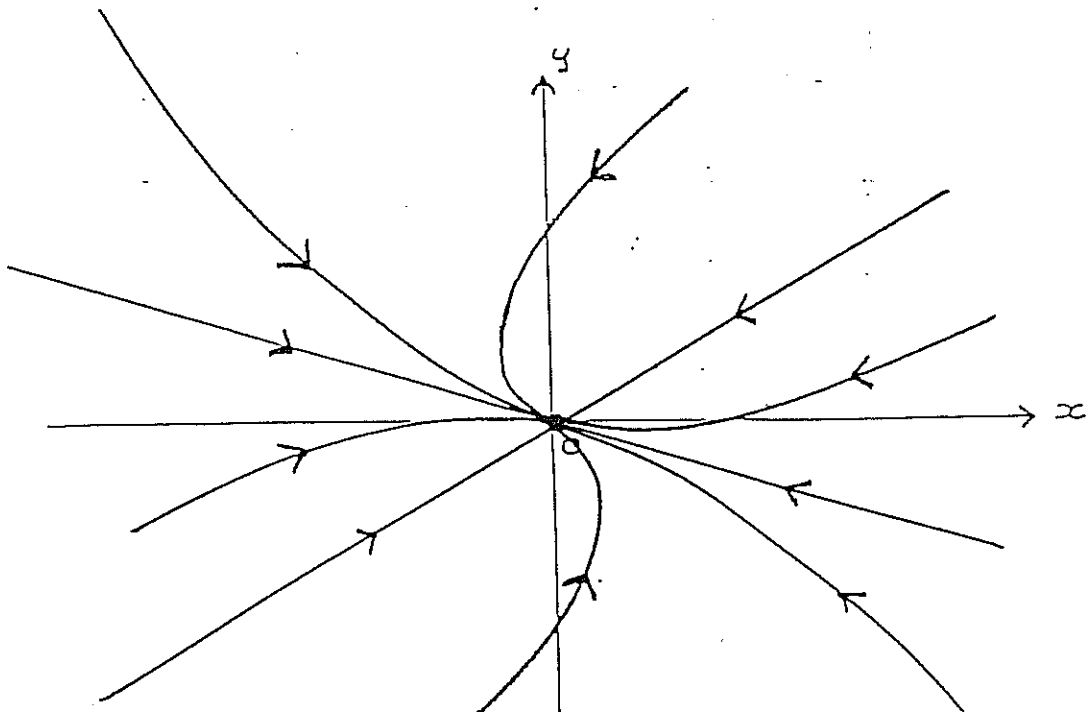


AN INTRODUCTORY  
COURSE OF  
LECTURES  
ON

The DIFFERENTIAL  
EQUATIONS  
of  
Biology

Brailey Sims

N. W. Taylor



THE UNIVERSITY OF NEW ENGLAND

AN INTRODUCTORY COURSE  
OF LECTURES ON  
THE DIFFERENTIAL EQUATIONS OF BIOLOGY

---

Brailey Sims  
N. W. Taylor

## PREFACE

These notes are a collection of mathematical results and techniques, from the theory of differential equations, which are of importance in the life sciences.

The division into 'lectures' is purely formal. Some of the 'lectures' would require more than one lecture period when presented to a class, while others would need less. On the average however they have been found fairly correct. Thus lectures 1-10, 13-15, 17-20 were the basis of an eighteen lecture, one term course on "the differential equations of biology" which has been given to second year students at the University of New England for several years now. Similarly, lectures 25 to 28, 11, 12 and 16 (qualitatively only) are the major part of a subsequent course on "Mathematical Methods in Biology", while lectures 21 to 24 formed part of an 18 lecture second year course on "Population Studies".

The lectures marked with an asterisk may be omitted or relocated without affecting the coherence of the course. Their location in the notes simply indicates their close connection with the preceding material.

The student working through these notes should gain a useful working knowledge of basic differential equation theory and its applications. However it is not the purpose of many of the lectures to provide such knowledge but rather to develop an awareness of potentially useful material. (To actually apply this material would often require a much more intensive study of it than these notes can hope to supply.)

I would like to draw the reader's attention to Dr. Taylor's section on "Delay Differential Equations" which renders an otherwise partial account more nearly complete.

I would like to thank Dr. E.W. Bowen and Dr. N.W. Taylor who carefully read parts of the draft manuscript and suggested many valuable improvements. I would also like to thank Mr. E.J. Hannah for his cooperation in the writing of Series VI, and my many colleagues in other departments who intentionally or otherwise helped supply the numerous examples and applications which occur throughout the notes. Lastly many thanks are due to Mrs. M. Christensen who expertly typed the manuscript with great patience and insight.

CONTENTS

Series I	- Single First Order Ordinary Differential Equations	B. Sims	
Lecture 1	What is a Differential Equation - its Meaning and Geometrical Interpretation (Slope Fields)?		1
Lecture 2	The Simplest First-Order Linear Equation and some Applications		5
Lecture 3	Another Useful First-Order Linear Equation		10
Lecture 4	Ageing Processes		16
Lecture 5	Non-Linear Differential Equations I, - the Logistic Equation		19
Lecture 6	Non-Linear Differential Equations II, - Allometry, Steven's Law		25
Lecture 7	The Numerical Solution of Ordinary First-Order Differential Equations - The Euler Tangent Method		29
Series II	- Systems of Differential Equations	B. Sims	
Lecture 8	Systems of Equations - Linear Systems		33
Lecture 9	Linear Systems from an Alternative View Point - Compartmental Analysis		40
Lecture 10	Non-Linear Systems - Lotka-Volterra Equations		44
*Lecture 11	Stability of Almost-Linear Systems		51
*Lecture 12	Periodic Solutions		58
Series III	- Single Second Order Ordinary Differential Equations	B. Sims	
Lecture 13	The General Equation and Solution of the Homogeneous Case with Constant Coefficients		62
Lecture 14	The Non-Homogeneous Equation (Laplace Transforms)		70
Lecture 15	Impulsive Inputs		81
*Lecture 16	Oscillatory Solutions - Sturm's Theory		86
Series IV	- Partial Differential Equations	B. Sims	
Lecture 17	Functions of More Than One Variable, Partial Differentiation		90
Lecture 18	Partial Differential Equations - The Diffusion Equation		95
Lecture 19	Propagation Equations		101
Lecture 20	A More General Second-Order Partial Differential Equation, Morphogenesis, Separation of Variables		107
Series V	- Delay Differential Equations	N.W. Taylor	
Lecture 21	The Exponential Solution of the Simple Population Growth Equation with Time Delay		114
Lecture 22	Oscillatory Solutions of the Simple Population Growth Equation with Time Delay		118
Lecture 23	Direct Methods of Solving the Delay Differential Equation		124
Lecture 24	The Method of Laplace Transforms		128

Series VI - Stochastic Models	B. Sims (and E.J. Hannah)	
Lecture 25	Yule's Pure Birth Process - I	132
Lecture 26	Yule's Pure Birth Process - II	139
Lecture 27	A Birth and Death Process	143
Lecture 28	Analysis of the Birth/Death Model	146

AN INTRODUCTORY COURSE OF LECTURES

ON THE

DIFFERENTIAL EQUATIONS OF BIOLOGY

SERIES I - Single First Order Ordinary Differential Equations

Lecture I    *What is a Differential Equation -*

*its meaning*

*and Geometrical Interpretation (Slope Fields)?*

§1.1 DEFINITION: Rather imprecisely, a *Differential Equation* is an equation relating an 'unknown' function and some of its derivatives. Usually such equations are coupled with some other subsidiary conditions, such as having the value of the unknown function and some of its derivatives specified at a point, in which case we say *initial conditions* have been prescribed. The problem is of course to find (or find out as much as we can about) functions satisfying the differential equation and the initial conditions.

The simplest instance of a differential equation is when the unknown function's derivative equals some given function  $f$ . i.e.

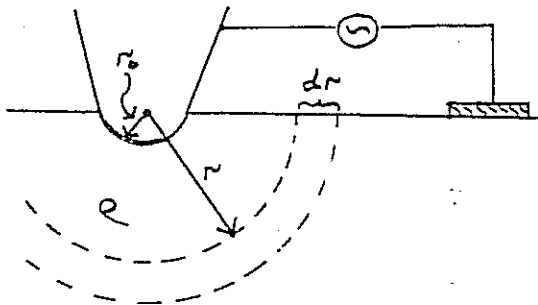
$$y' = f(x) \text{ or } \frac{dy}{dx} = f(x)^*$$

With the initial condition:  $y(0) = c$ , a given constant. This equation simply tells us that the unknown function  $y$  is an *antiderivative* of  $f$ , and so from the Fundamental Theorem of Calculus

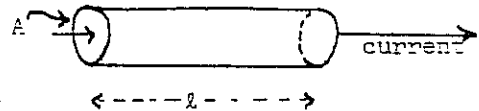
$$y(x) = \int_0^x f(t)dt + c$$

is easily seen to be the unique solution for this case.

ILLUSTRATIVE EXAMPLE: Consider a microelectrode, whose tip has radius  $r_0$ , pressed against the skin as shown in Figure 1. From elementary physics the



resistance of a bar of conducting material, with resistivity  $\rho$ , such as that shown below is  $R = \frac{\rho l}{A}$



So the resistance of the hemispherical shell indicated in figure 1 is

$$dR = \rho \frac{dr}{2\pi r^2} \quad \text{or} \quad \frac{dR}{dr} = \frac{\rho}{2\pi r^2}$$

Figure 1

Further

$R(r_0) = 0$  supplies the initial condition. From these it follows that

$$R(r) = \int_{r_0}^r \frac{\rho}{2\pi r^2} dr = \frac{\rho}{2\pi} \left( \frac{1}{r_0} - \frac{1}{r} \right)$$

\* We will use both notations,  $y'$  and  $\frac{dy}{dx}$ , interchangeably for the derivative.

For a point a long way away from the electrode,  $r$  is large and so  $\frac{1}{r}$  negligible. We conclude therefore that for points far removed from the electrode the resistance is inversely proportional to the diameter of the electrode's tip, i.e. the smaller the electrode's tip the larger the resistance, one argument against the use of sharply pointed electrodes.

Differential equations occur commonly in Ecology, Biophysics, Physiology and many other areas. Usually they involve both the unknown function  $y$  and  $y'$ , possibly even  $y''$ , and  $y'''$ , ... . We will see many examples of how such equations arise and how they may be treated. Frequently it proves impossible to solve explicitly these equations, in which case a number of questions arise.

- (1) Does the equation have a solution?
- (2) If there is a solution, is it unique?
- (3) What properties of the solutions can we infer directly from the differential equation; e.g. when is it zero, does it oscillate, will it tend to a limiting value for large values of  $x$ ?

We will assume that "yes" is the answer to (1) and also in most cases to (2). If a physical situation has been carefully modelled into a differential equation this usually proves to be the case. However, from the fact that the physical situation has a 'solution' we cannot conclude that our equations will. The proof that they do provides a first check on the plausibility of our model.

In later sections we will concentrate quite a lot on answering (3). The next section providing a first step in this direction.

§1.2 THE SLOPE FIELD OF A DIFFERENTIAL EQUATION: Let us consider differential equations which can be written in the form

$$y' = F(x, y)$$

where the right hand side is a known function of  $x$  and the unknown function  $y$ . The graph of a solution to this equation is a curve in the  $x$ - $y$  plane.

If the curve passes through a point  $(x_0, y_0)$  then its *slope* at that point is known. In fact it is

$$y'(x_0) = F(x_0, y_0).$$

If, at each of a large number of points in the  $x$ - $y$  plane, we hatch a short line with slope equal to that which a solution curve passing through would have, we obtain what is known as the slope field for the equation. From this it often proves possible to sketch in the specific solution curve passing through a given initial point and so gain some insight into the forms of the solution. To clarify this, consider the following.

ILLUSTRATIVE EXAMPLE. British Fisheries report that for a certain species of fish the average weight  $w$  of an individual at age  $t$  varies according to

$$\frac{dw}{dt} = 0.9w^{2/3} - 0.35w \quad (\text{weight in ounces})$$

(Can you see why this equation is reasonable? - Hint: apply dimensional considerations to the rate at which nutrients are assimilated and used in normal metabolic processes.)

In plotting the slope field for this equation, it is useful to note that  $w'$  has a constant value on any line of the form  $w = \text{constant}$ . The curves in the  $w$ - $t$  plane, in this instance straight lines parallel to the  $t$ -axis, along which  $w'$  has a constant value are *isoclines* (lines of constant slope). Straightforward numerical calculation gives

$w$	$w' = 0.9w^{2/3} - 0.35w$
0	0
.5	.39
1	.55
5	.88
10	.67
15	.22
20	-.368
30	-1.8

with  $w' = 0$  for  $w = 17$ .

This information is plotted on the  $w$ - $t$  plane in Figure 2, from which we see that the solution curve passing through the point  $(0, .5)$ , i.e. corresponding to the initial condition  $w(0) = 0.5$ , has slope .39 at this point. By the time it reaches the line  $w = 1$  it must have changed slope to 0.55 etc. Continuing in this way allows us to sketch in this solution curve. Similarly, a number of different solution curves which correspond to different initial values have been drawn in. From these it is apparent that independent of the initial conditions all solutions approach the line  $w = 17$  for large values of  $t$  and so we conclude that the expected mature weight of a fish is 17 ounces.



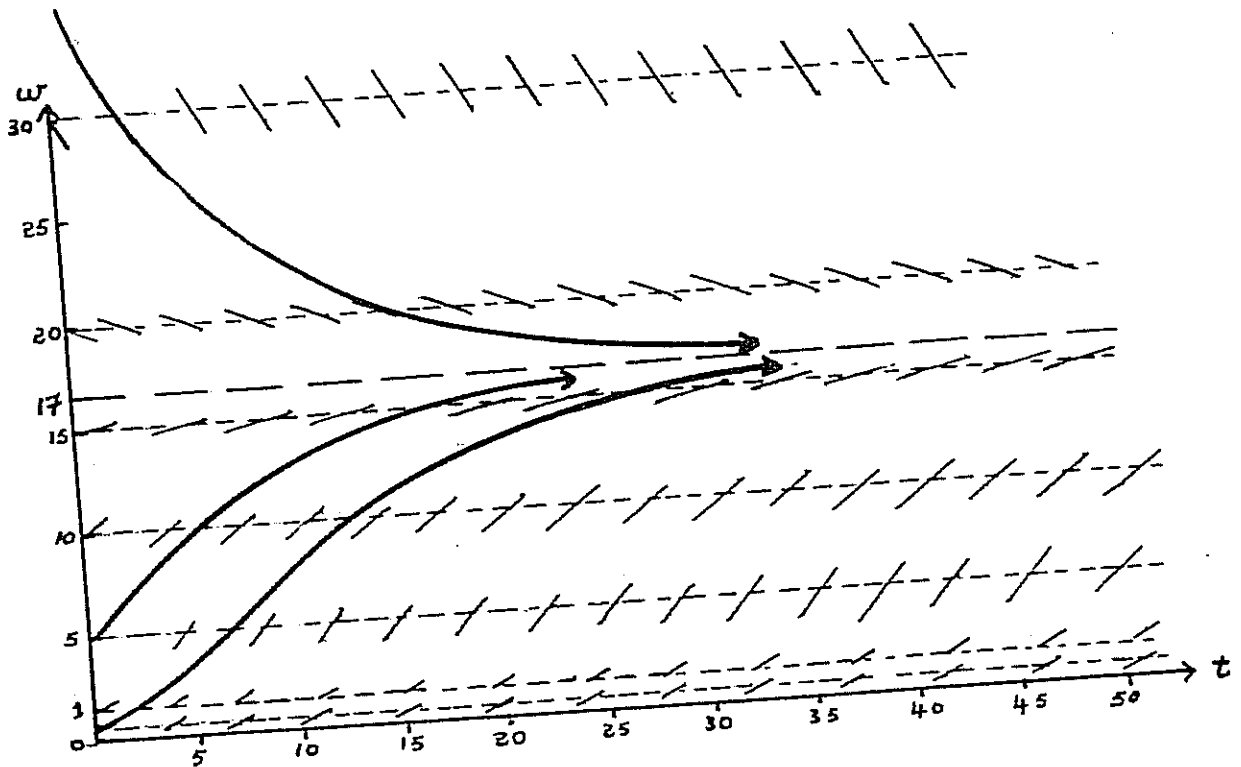


Figure 2. Slope Field for  $w' = 0.9w^{2/3} - 0.35w$

*Collateral Reading.*

A good exposition on the important question of existence and uniqueness of the solutions of differential equations may be found in Boyce and Di Prima "Elementary Differential Equations and Boundary Value Problems", Chapter 2, Sections 11 and 12.

EXERCISE.

1. The non-linear equation  $\frac{dM}{dt} = A - aM - bM^2$  may give a good approximation to the mass,  $M$ , of litter on a given area of forest floor. Using  $A = 1$ ,  $a = 0.5$ ,  $b = 0.5$  draw a slope field for this equation for  $0 \leq t < 10$ ,  $0 < M \leq 2$ . Use this to sketch the shape of some possible solutions.

Lecture 2    *The Simplest first order Linear Equation and some applications*

For motivation, let us consider the problem of modelling, in the simplest possible case, the growth of a population. Assume the only changes taking place in the population are due to births and deaths, then

$$\left\{ \begin{array}{l} \text{rate of change of} \\ \text{population with time} \end{array} \right\} = \left\{ \begin{array}{l} \text{number of births occur-} \\ \text{ring per unit of time} \end{array} \right\} - \left\{ \begin{array}{l} \text{number of deaths occur-} \\ \text{ring per unit of time} \end{array} \right\}$$

It seems reasonable to assume that the number of births occurring per unit of time will be proportional to the size of the population present at that time. Thus, letting  $N(t)$  be the population's size at time  $t$  we would have {number of births per unit of time} =  $\beta N(t)$ , where  $\beta$  is the proportionality constant known as the *specific birth rate* and equals the number of births occurring per unit of time per 'head' of population.

Similarly take

{number of deaths per unit of time} =  $\delta N(t)$ ,  $\delta$  being the specific death rate for the population (assumed constant).

Then, since  $\left\{ \begin{array}{l} \text{rate of change of population} \\ \text{with time} \end{array} \right\} = \frac{dN}{dt}$ , we obtain

$$\frac{dN}{dt} = \beta N - \delta N = (\beta - \delta)N$$

as the differential equation which describes the population's growth. Usually this would be combined with the initial condition  $N(0) = N_0$ ,  $N_0$  being the known initial size of the population. Except that we have used  $N$  in place of  $y$  and  $t$  instead of  $x$  this is a differential equation of the form;

$$y' = ky, \quad (k \text{ being constant})$$

and  $y(x_0) = y_0$

Draw a slope field for this equation (what are its isoclines?). Sketch in some typical solution curves.

Because of the simplicity of this particular equation, you might like to try guessing its solution before proceeding further, where the solution is obtained by a method applicable to a wide range of similar differential equations. Note that the equation is asking for a function  $y$  which when differentiated once becomes  $k$  times itself.

Take  $y_0 \neq 0$ . (If  $y_0 = 0$  then a solution is easy,  $y(x) = 0$  for all  $x$ .) Now assume  $y(x) \neq 0$  for any  $x$  (actually this can be deduced directly from the equation); we must check this assumption later, for if it proved false the following line of reasoning could be invalidated.

Rearrange the equation to

$$y^{-1} \frac{dy}{dx} = k$$

Integrating both sides from  $x_0$  to  $x$  we have

$$\int_{x_0}^x y^{-1} \frac{dy}{dx} dx = \int_{x_0}^x k dx = k(x - x_0).$$

From the integral change of variable theorem, the left hand side is

$$\ln y - \ln y_0 = \ln \frac{y}{y_0}$$

and so

$$\ln \frac{y}{y_0} = k(x - x_0)$$

or

$$y = y_0 \exp k (x - x_0) .$$

Clearly  $y \neq 0$  for any  $x$ , and so, as direct substitution will confirm, this is the required solution.

Returning to our population model, replacing  $x$  and  $y$  by the appropriate symbols, we have that the population size is

$$N(t) = N_0 \exp[(\beta - \delta)t].$$

So the population described by this simple model grows exponentially, increasing toward an infinite value if  $\beta > \delta$  or dying out if  $\beta < \delta$ .

The technique employed in solving  $y' = ky$ ,  $y(x_0) = y_0$  may be extended to any differential equation that can be arranged into the form

$$f(y)y' = g(x), \quad y(x_0) = y_0, \text{ provided}$$

both  $f$  and  $g$  are integrable functions. Such equations are termed *separable*, and for them the identity

$$\int_{y_0}^y f(y)dy = \int_{x_0}^x f(y)y'dx = \int_{x_0}^x g(x)dx$$

implicitly relates the solution  $y$  and the variable  $x$ . We will encounter many further examples of such equations.

Apart from the particular case of population already considered, any quantity whose rate of change is proportional to the quantity present is also described by the equation  $y' = ky$ . We now consider a number of examples in which this is the case.

### 2.1 The Growth of a single cell

For a cell of mass  $m$  its growth rate may be proportional to  $m$ , in which case

$$\frac{dm}{dt} = km \text{ and so } m = m_0 e^{kt}.$$

Of course growth of this form would not continue indefinitely (unless our cell were a truly exceptional one, which grew infinitely large), so usually

some restriction, like  $m < m_1$  would be operative, the cell undergoing division once mass  $m_1$  is reached rather than continuing to grow.

Possibly a more plausible model would be to assume that the growth rate of a cell is proportional to the rate at which it can absorb nutrient and so proportional to its surface area and hence to the two-third power of its mass. This leads to the equation

$$\frac{dm}{dt} = km^{\frac{2}{3}} \quad (\text{see exercises}).$$

### 2.2 The Weber-Fechner 'Law'

Physiological experiments suggest that the smallest change we can detect in a stimulus' magnitude (the 'just noticeable difference') varies in such a way that the fractional change is a constant (Weber's Law). That is, if  $S$  is the physical magnitude of the stimulus then we will just notice the change to  $S + \Delta S$  where  $\frac{\Delta S}{S} = k$ , a constant.

Fechner proposed that this constant represents a standard increase in the psychological magnitude of the stimulus,  $I$ . Thus  $\Delta I = \frac{\Delta S}{S}$  or  $\frac{\Delta S}{\Delta I} = S^*$ .

Treating  $\Delta I$ ,  $\Delta S$  as very small we then have as a first approximation, that  $S$  is related to  $I$  through the differential equation  $\frac{dS}{dI} = kS$ ,  $k$  being the constant of proportionality. So, from our previous work,

$$I = \left(\frac{1}{k}\right) \ln S + C \text{ where } C \text{ is a constant.}$$

In other words the psychological intensity of the stimulus is linearly related to the logarithm of its physical magnitude.

It is for this reason that the intensity of sounds (decibels), the brightness of stars (magnitudes) and many other similar quantities are measured on logarithmic scales.

### 2.3 Decay and Half-Lives

The rate at which a substance breaks down is frequently proportional to the amount of substance present. This is the case, for example in the decay of a radioactive substance, in the destruction of tissue due to harmful radiation, in the gradual break-down of a chemical compound, in the decay of compost on a forest floor.

\*

Writing this as  $\frac{S(I + \Delta I) - S(I)}{S(I)} = k\Delta I$  leads to the difference equation  $S(I + \Delta I) = (k\Delta I + 1) S(I)$  which may be

$$\text{solved to give } S(I_0 + n\Delta I) = S(I_0)(k\Delta I + 1)^n \quad n = 1, 2, 3, \dots$$

Here  $n$  is a measure of the psychological intensity (in units of  $\Delta I$ , above the base level  $I_0$ ) and taking logarithms gives  $n = \left(\log \frac{S}{S_0}\right) / \left(\log(k\Delta I + 1)\right)$ . So we have the same conclusion as derived from the differential equation.

In such cases, if  $M$  denotes the amount of undecayed material present we have the differential equation

$$\frac{dM}{dt} = -kM \text{ where } k > 0 \text{ is the proportionality constant,}$$

the minus sign indicates the decrease of  $M$  with  $t$ . So for  $M(0) = M_0$  we have  $M = M_0 e^{-kt}$ .  $M$  decays exponentially with  $t$ .

The question "how long before all the substance has decayed away?" is clearly meaningless, since there is no finite value of  $t$  at which  $M = 0$ . How then can we meaningfully measure the speed of decay? The size of the constant  $k$  is of course a measure of this but its interpretation is difficult. However, if we note that for any fixed constant  $\tau$  the ratio  $M(t + \tau)/M(t) = e^{-k\tau}$  is a constant independent of  $t$ , we have that the fraction of the material present at  $t$  remaining by  $t + \tau$  is a constant depending only on  $k$  and  $\tau$  but not on  $t$ . If we now choose  $\tau$  so that  $e^{-k\tau} = \frac{1}{2}$  (i.e.  $\tau = \frac{\ln 2}{k} = 0.69315/k$ ) we see that during any interval of length  $\tau$  the amount of undecayed material is reduced by a half. This value of  $\tau$  is thus called the *half-life* of the decay. It affords a convenient measure of the speed of decay. (Note, that from a knowledge of the half-life,  $k$  may be calculated and vice versa.)

#### EXERCISES:

1. In 1879 four hundred and thirty-five striped bass were planted in San Francisco bay from the Atlantic and allowed to breed for 20 years. Then in 1899 the net catch was 1,234,000 fish. Assuming that every tenth fish was caught find  $\lambda$  if the growth of the population is described by

$$\frac{dN}{dt} = \lambda N.$$

(Adapted from MacArthur & Connell)

2. For small doses the rate at which aspirin leaves the blood stream is proportional to the amount of aspirin in the blood. Typically, if the concentration of plasma aspirin is measured in mgms/litre and the time in hours, then the proportionality constant is 0.04. Derive a differential equation to describe this situation.

A 0.64 gm dose of aspirin is known to produce an initial 20 mgms/litre plasma aspirin concentration. Use this to find the plasma aspirin concentration, resulting from a 0.64 gm dose, as a function of time. After what time will the plasma aspirin concentration be reduced to half its initial value?

(Extracted from Waterson "Aspirin & Gastric Haemorrhage", Brit. Med. J. 1955-II pp. 1531-1533.)

3. When a beam of light, initially of intensity  $I$ , passes through a short distance  $\Delta x$  of liquid, its intensity is reduced by an amount  $\Delta I$  due to absorption.

3. (continued)

It is found that  $\frac{\Delta I}{I}/\Delta x$  is a constant.

e.g. For water  $\frac{\Delta I}{I}/\Delta x = \begin{cases} 2.9 \times 10^{-3} & \text{for red light (x in cms) and} \\ 4.6 \times 10^{-4} & \text{for blue light} \end{cases}$

Convert the relationship  $\frac{\Delta I}{I}/\Delta x = \lambda$  into a differential equation and hence find I as a function of  $I_0$ , the intensity at the surface, and x the depth of liquid which the light has traversed.

Assuming that the intensity of red and blue lights are equal at the surface, find their relative intensities at a depth of 10 metres. This constitutes an important environmental change for some marine creatures.

4. It has been found (Watt *et al*) that the growth of human population in developing countries is described by  $\frac{dN}{dt} = \lambda N^r$  for some constant  $r > 1$ . Find N as a function of t. (This same equation with  $r = \frac{2}{3}$  has been used to describe the mass of a growing cell.) If  $N(0) = 1$ ,  $\lambda = 1$  and  $r = 2$  plot the solution.
5. The number of worker bees of the species "Bombus humilis Ill" still alive after d-days from a hive initially containing 100 workers is  $N = 100 e^{-0.04d}$ . How many days must elapse before the number is halved? (Adapted from Brian, 1965)
6. Radioactive iodine  $^{131}\text{I}$ , used in diagnostic medicine, decays according to the law  $\frac{dN}{dt} = -0.886N$ , where t is time in days. N is the number of atoms present. What is the half-life of this isotope, after how many days will  $N = N_0/4$ ?

Lecture 3    *Another useful first order linear equation*

Consider the growth of a population due not only to births and deaths but also to immigration (emigration). In such a situation we have

$$\left\{ \begin{array}{l} \text{rate of change} \\ \text{of population} \end{array} \right\} = \left\{ \begin{array}{l} \text{number of births} \\ \text{per unit of time} \end{array} \right\} - \left\{ \begin{array}{l} \text{number of deaths} \\ \text{per unit of time} \end{array} \right\} + \left\{ \begin{array}{l} \text{number of immigrants} \\ \text{arriving per unit of} \\ \text{time} \end{array} \right\}$$

If we make the same assumptions concerning the occurrence of births and deaths as were used in the last lecture, this yields

$$\frac{dP}{dt} = \beta P - \delta P + f(t) \quad \text{where } P \text{ is the population size}$$

and  $f(t)$  is the rate of arrival of immigrants at time  $t$ , assumed known.

We are therefore led to the differential equation

$$y' = ky + f.$$

For a discussion of how this class of equations may be solved see Exercise 3. Here we will restrict attention to the case when  $f$  is a constant. In the terms of the example, this corresponds to a constant rate of immigration.

Hence let

$$y' = ky + a \quad (k, a \text{ being known constants})$$

and

$$y(x_0) = y_0.$$

Then under the change of variable  $z = y + \frac{a}{k}$  this becomes

$$z' = kz$$

$$z(x_0) = y_0 + \frac{a}{k}$$

and is of the form previously considered, for which we found

$$z = \left( y_0 + \frac{a}{k} \right) e^{k(x-x_0)}$$

and so we have for the solution to our present problem

$$y = \left( y_0 + \frac{a}{k} \right) e^{k(x-x_0)} - \frac{a}{k}.$$

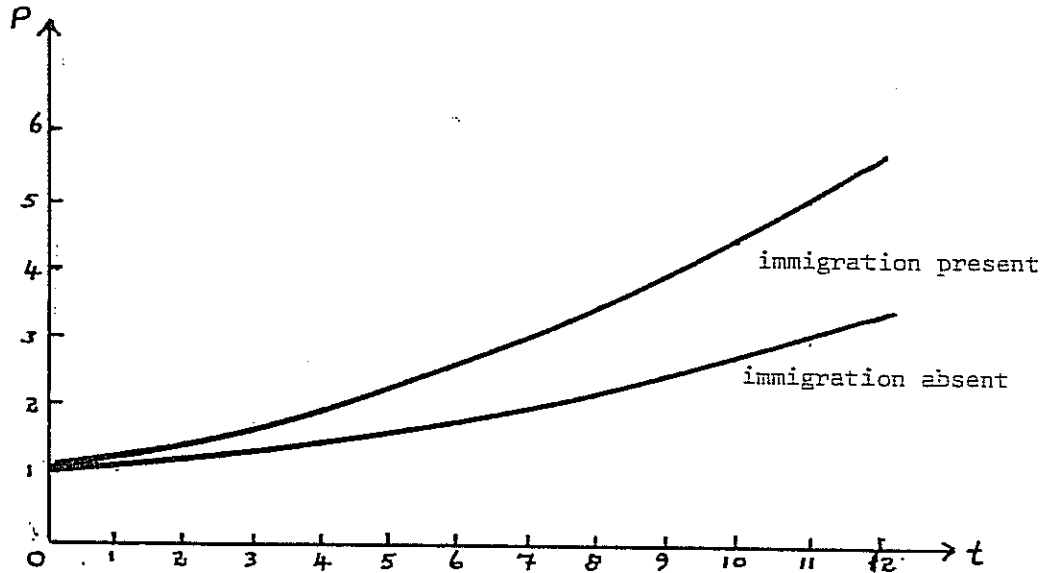
Interpreting this in the case of population growth in the presence of immigration leads to

$$P(\tau) = \left[ P_0 + \frac{r}{\beta - \delta} \right] e^{(\beta - \delta)\tau} - \frac{r}{\beta - \delta}$$

where  $r$  is the constant rate of immigration, or

$$P(t) = P_0 e^{(\beta - \delta)t} + \frac{r}{\beta - \delta} \left[ e^{(\beta - \delta)t} - 1 \right].$$

Here the last term represents the effect of immigration, the first term being the solution to the corresponding problem in the absence of any immigration. For a comparison see the graph below.



Following the previously established format we now look at some other instances in which the equation  $y' = ky + a$  appears.

3.1 *Development in the presence of growth-limiting factors.*

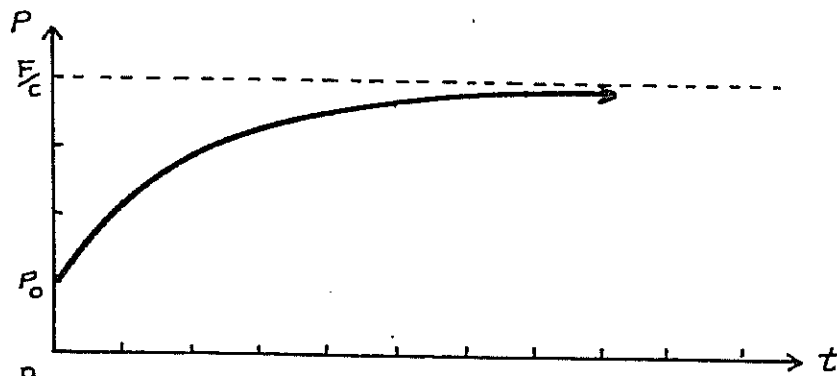
Here we consider the growth of a population whose rate of development is not proportional to the population's size,  $P$ , but rather to  $F - cP$ .  $F$  may represent the total space available to the population (or the total nutrient available) while  $c$  is the space occupied per 'head' of population. Thus  $F - cP$  measures the amount of unoccupied space (unutilized nutrient) into which the population can expand. In such a case the growth is described by

$$\frac{dP}{dt} = kF - kcP \quad (k \text{ being the constant of proportionality})$$

and reflects the fact that when  $F = cP$  the rate of growth is zero, further development being impossible since no more space (nutrient) is available. This is an equation of the form just considered, from the solution of which we have

$$P(t) = P_0 e^{-kct} + \frac{F}{c} (1 - e^{-kct}).$$

The graph of a typical function of this form appears below.



Note that the growth is 'cut-off' by the limiting size  $P_\infty = \frac{F}{c}$ .



### 3.2 Diffusion

For a single cell in a bath of metabolite held at a constant concentration  $c_e$ , it is found (or follows from the laws of statistical mechanics) that

$$\left\{ \begin{array}{l} \text{rate of diffusion of metabolite} \\ \text{across the cell's membrane} \end{array} \right\} = \left\{ \begin{array}{l} \text{rate of change of metabolite} \\ \text{concentration within the cell} \end{array} \right\}$$

$$= \left\{ \begin{array}{l} \text{difference in concentration of the metabolite} \\ \text{between the interior and exterior of the cell} \end{array} \right\}^*$$

or denoting by  $c$  the interior concentration

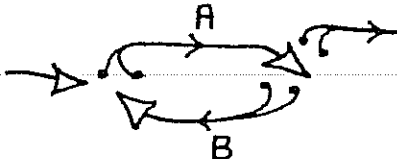
$$\frac{dc}{dt} = k(c_e - c).$$

From which it follows that  $c(t) = (c_0 - c_e)e^{-kt} + c_e$  where  $c_0 = c(0)$ , the initial concentration within the cell. Observe that regardless of the value of  $c_0$ ,  $c(t)$  eventually approaches  $c_e$  and so the interior concentration comes into equilibrium with that of the bath.

### 3.3 A model for the Central Nervous System

Here we give a very brief introduction to Rashevsky's model for neural pathways within the central nervous system.

Regard the continuous response to a stimulus of a neural pathway as due to the blending of a large number of on-off responses from individual axons and synapses. It will then be proportional to the number of excited neurons. Rashevsky assumes that these individual neurons are kept excited by means of 'neuron cycles' such as shown.



When A is triggered it excites B which in turn retriggers A.

So the number of excited neurons is proportional to the number of activated cycles, which we denote by  $\epsilon$ . Then

$$\left\{ \begin{array}{l} \text{rate of change} \\ \text{of } \epsilon \end{array} \right\} = \left\{ \begin{array}{l} \text{excitation of pathway} \\ \text{due to stimulus} \end{array} \right\} - \left\{ \begin{array}{l} \text{failure rate of synaptic mechanism} \\ \text{to continue cycle} \end{array} \right\}$$

Since failures of the synaptic mechanism occur at random, it is reasonable to take their failure rate proportional to the number of excited cycles present. We are therefore led to an equation of the form under consideration, viz.

$$\frac{d\epsilon}{dt} = AE - a\epsilon.$$

Here  $AE$  is a constant depending on the magnitude of the stimulus  $S$ , in fact

$$E = ah \log \frac{S/h}{1 + \frac{S}{h} e^{-Ea/ah}}$$

where  $a$ ,  $h$ ,  $Ea$  are appropriate constants.

---

\* Strictly, the concentration gradient across the cell membrane.

Such models have been elaborated to explain -

Reaction times -

Colour Vision -

Flicker -

and Gestalt.

### 3.4 *The Build-up of litter on a forest floor*

Denoting by  $M$  the mass of litter per unit of area of forest floor, it is plausible to assume

$$\frac{dM}{dt} = A - aM.$$

Here  $A$  represents the rate at which new litter accumulates and is regarded as a constant of the forest. While the last term accounts for the removal of litter through decay and other factors, taken to be proportional to  $M$ . More realistically, this last term could be replaced by  $-aM - bM^{\frac{2}{3}}$  the first component representing removal by decay and similar processes, while the second one accounts for the reduction in litter due to erosion etc., which is taken to be proportional to the exposed surface area of the litter (hence the two-thirds power). See Exercise 1 to lecture I.

#### *Collateral Reading*

The theory of neural pathways considered is developed in a much fuller form in

N. Rashevsky "Mathematical Biophysics, Physico-Mathematical Foundations of Biology." Vol. I and II, Dover, 1960.

as is the theory of cell development.

#### EXERCISES:-

1. Consider a population which has a fractional birth-rate of 0.2 and a fractional death-rate of 0.1 per year, also assume that the population is depleted by emigration at a constant rate of 1,000 per year. Derive a D.E. to describe the growth of the population. If the population were initially of size 10,000, find its size after 1, 2, 3, 4, 5 years.
2. Derive the differential equation describing a population which has a fractional birth-rate of 0.3 per year and a fractional death-rate of 0.1 per year and which is also supplemented by immigration at the constant rate of 500 per year. Solve the equation and so determine what the population will be in 10 years time if it starts off with a size of 10,000.
3. Consider a population, growing exponentially with time, which is host to a population of parasites of size  $y$  at time  $t$ . Then with assumptions similar to those used in lectures it is reasonable to assume that "rate of growth of parasite" is proportional to "amount of host currently unaffected by the parasite" which leads to an equation of the form  $y' = e^x - ky$  (where  $k$  is a positive constant).

3. (cont.)

Show that this can be written as  $(\mu y)' = e^x \mu$  where  $\mu$  is an appropriately chosen function of  $x$ , known as an integrating factor for the equation. Hence solve the equation. Use your solution to graph  $y$  as a function of  $x$  if  $y(0) = 1$ ,  $k = 1$ . Also show that the fraction of host infected by parasites tends to the value  $\frac{1}{1+k}$  for large values of time ( $x$ ). The use of an integrating factor, as illustrated by this problem, allows more general equations of the form  $y' + f(x)y = g(x)$  to be solved, (look up first order linear non-homogeneous equations and the use of integrating factors, in almost any book on differential equations).

*Modelling Project*

PASTURE GROWTH IN THE ABSENCE OF GRAZING

1. The General Model:

Consider an initially barren pasture containing  $N_0$  latent seeds per unit area. Let the number of seeds germinating between  $t$  and  $t+dt$  be  $n(t)dt$ , provided  $dt$  is sufficiently small.

Let the average weight of an individual plant, time  $t$  after germination, be  $w(t)$ . Then the total number of plants present at time  $t$  is;

$$N(t) = \int_0^t n(\tau) d\tau.$$

While the weight of pasture per unit area at time  $t$  is;

$$W(t) = \int_0^t \left\{ \begin{array}{l} n^0 \text{ of plants germinating} \\ \text{per unit time at time } T \end{array} \right\} \times \left\{ \begin{array}{l} \text{weight attained by a plant,} \\ \text{germinating at } T, \text{ by time } t \\ \text{is reached} \end{array} \right\} dT$$

so

$$W(t) = \int_0^t n(T) w(t-T) dT, \text{ the convolution of } n \text{ \& } w.$$

1. Assuming the individual plants have a constant specific growth rate of  $\lambda$  and initial weight  $w_0$  at germination, find an expression for  $w(\tau)$ .
2. If the probability of a seed, ungerminated at time  $t$ , germinating between  $t$  and  $t+dt$  is  $kdt$  ( $k$  a constant) show

$$n(\tau) = kN_0 e^{-k\tau} \text{ where } N_0 \text{ equals the}$$

initial number of seeds present.

3. Using the results of 1. and 2. and the above expressions, find  $N(\tau)$  and  $W(t)$ .

4. Using values  $k = 0.25$  ( $\text{days}^{-1}$ )  
 $\lambda = 0.1$  ( $\text{days}^{-1}$ )  
 $N_0 = 69$  ( $\text{seeds}/\text{dm}^2$ )  
and  $w_0 = 2.8$  ( $\text{mgms}$ )

Graphically compare the calculated values of  $N(t)$  and  $W(t)$  given by 3. with the following values obtained experimentally by R. C. Smith:

PLANT NUMBERS

$t(\text{days})$	$N(t)$ ( $1 \text{ dm}^2$ )
2	25
8	62
14	62
20	69

TOTAL PASTURE WEIGHT

$t(\text{days})$	$W(t)$ ( $\text{mgms}/\text{dm}^2$ )
2	74
8	281
14	543
20	1016

Lecture 4    *Aging Processes*

Under the assumption of an age-dependent mortality-rate, we examine the demise of a population initially consisting of  $P_0$  individuals all of the same age.

Denoting by  $t$  the elapsed time from the initial instant and by  $P(t)$  the corresponding size of the population we have

$$\left\{ \begin{array}{l} \text{rate of change} \\ \text{of population} \end{array} \right\} = - \left\{ \text{specific mortality rate} \right\} \times \left\{ \text{population size} \right\}$$

or  $\frac{dP}{dt} = -R_m(t)P$  where  $R_m(t)$  is the specific mortality-rate for individuals at time  $t$ .

Assuming  $P(t) \neq 0$  for any  $t$ ,  $0 \leq t \leq \tau$  we have upon separating variables and integrating

$$\ln P(\tau)/P_0 = - \int_0^\tau R_m(t) dt.$$

To progress further we need to determine plausible forms for  $R_m(t)$ .

Brody Failla Theory: Here some measure  $V$  of the individual's "vitality" is postulated, and the specific mortality-rate is assumed to be inversely proportion to  $V$ .

It is further assumed that vitality decreases with age at a rate proportional to its own amount. (This is plausible if we take the number of 'healthy' cells within the organism as a measure of the vitality. It being assumed that healthy cells are lost according to a simple decay process - see 2.3 - due to the effects of background radiation, random mutations etc.)

$$\text{Thus } \frac{dV}{dt} = -kV \text{ or } V = V_0 e^{-kt} \text{ and so}$$

$$R_m(t) = R_0 e^{kt}.$$

There is a considerable weight of experimental evidence vindicating this form for  $R_m(t)$ , see Figure 1. Although of course, many alternative rationales for this form have been offered.

For example, in the Simms-Jones model the concept of vitality is dispensed with,  $R_m$  simply being assumed proportional to the amount of decay present, while Strehler suggests a much more ambitious explanation based on "entropy" arguments.

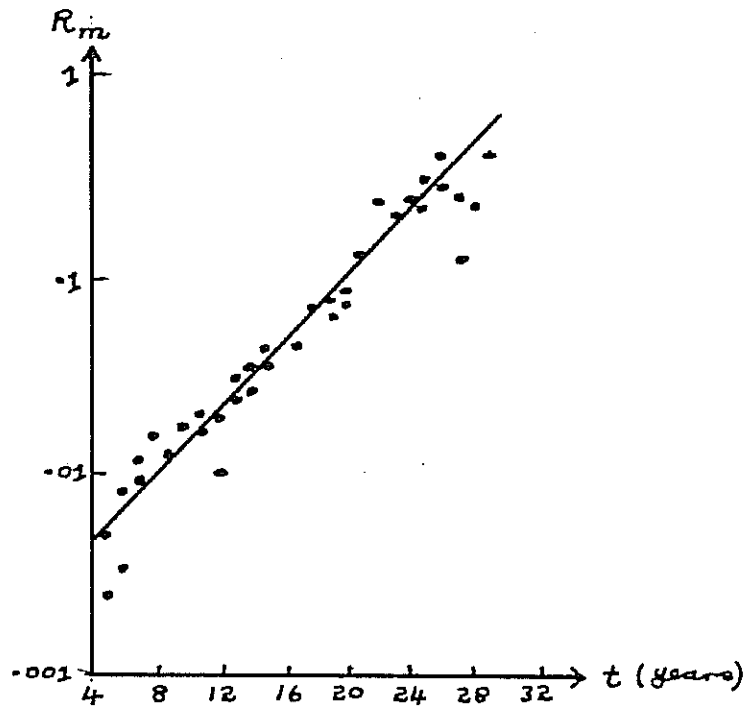


Figure 1. Specific mortality-rate of thoroughbred mares, plotted on a logarithmic scale, against age (adapted from Sacher).

Independently of these derivations, if we accept  $R_m(t) = R_0 e^{kt}$  (for white U.S. males it is  $0.12 e^{0.084t}$  where  $t$  is age in years) we then have

$$\frac{dP}{dt} = -R_0 e^{kt} P.$$

Separating variables and integrating gives

$$\int_{P_0}^P p^{-1} dp = -R_0 \int_0^t e^{kt} dt$$

So 
$$\ln \frac{P}{P_0} = \frac{-R_0}{k} [e^{kt} - 1]$$

or 
$$P(t) = \left[ P_0 e^{R_0/k} \right] e^{-\frac{R_0}{k} e^{kt}}.$$

This is known as Gompertz' Formula, the truth of which is well established from observation. It is of great importance in actuarial work. The graph for  $P$ , as given by Gompertz' formula, is illustrated in Figure 2.

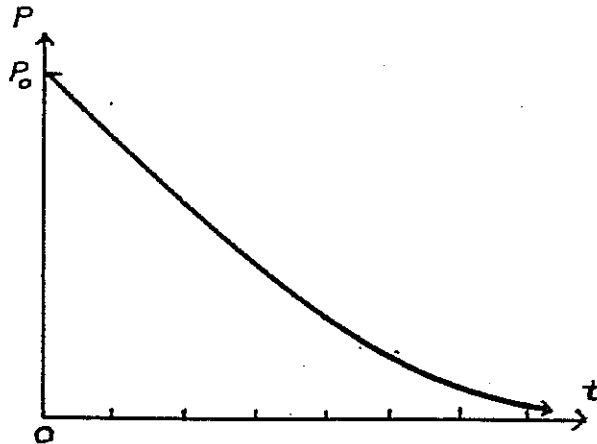


Figure 2. Gompertz' Curve

Makeham proposed that  $R_m$  be modified to  $R_m(t) = R_0 e^{kt} + A$ , the constant  $A$  allowing for age independent deaths due to factors like accidents. In some circumstances this Makeham variation leads to a more accurate model.

*Collateral Reading.*

A good account of the various theories of aging is given in Chapter IV (p. 86 on) of

B. L. Strehler, "Time, Cells and Aging".

Throughout the remainder of this book much of the experimental evidence and underlying biological theory may be found.

EXERCISE:

1. The Makeham variation of Gompertz' formula derives from the equation

$$\frac{dN}{dt} = -b(e^{kt} + c)N.$$

Solve this equation.

How does its solution differ from Gompertz' formula?

(Note the same equation might also be used to describe a population growing in the 'normal' way but whose death rate varies exponentially with time.)





$$= \frac{1}{r_1 - r_2} \left[ \ln(y - r_1) - \ln(y_0 - r_1) - \ln(y - r_2) + \ln(y_0 - r_2) \right].$$

So  $\ln \frac{y - r_1}{y - r_2} = (r_1 - r_2)a(x - x_0) + \ln \frac{y_0 - r_1}{y_0 - r_2}$

or taking exp of both sides

$$\frac{y - r_1}{y - r_2} = \left( \frac{y_0 - r_1}{y_0 - r_2} \right) e^{(r_1 - r_2)a(x - x_0)} = z, \text{ say.}$$

Happily, this implicit relationship between  $y$  and  $z$  (and hence  $x$ ) can be solved to yield  $y$  explicitly.

Thus  $y - r_1 = z(y - r_2)$

or  $y = \frac{r_1 - zr_2}{1 - z} = r_2 - \frac{r_2 - r_1}{1 - z}$

and so  $y = r_2 - \frac{r_2 - r_1}{1 - \left( \frac{y_0 - r_1}{y_0 - r_2} \right) e^{(r_1 - r_2)a(x - x_0)}}$

Further, provided  $y_0 \neq r_1$  or  $r_2$ ; we see that  $y = r_1$  or  $r_2$  implies that  $e^{(r_2 - r_1)a(x - x_0)} = 0$  or  $=$  which is not possible for any finite value of  $x$  provided  $r_1, r_2$  are real. So this is the required solution for all values of  $x$ .

Although we will simply refer back to this solution, its complexity is such that it is much better to learn the steps leading to it, rederiving the solution in any specific example, rather than try to remember its final form. To make sense of this solution when  $r_1$  and  $r_2$  are complex conjugates requires the Euler formula. However, this case arises extremely rarely in biological problems, and so will not be pursued further.

In the case of the population growing according to

$$\frac{dP}{dt} = kP(F - cP) \text{ we have } r_2 = 0, r_1 = \frac{F}{c} \text{ and } a = -kc \text{ so}$$

$$P(t) = \frac{F/c}{1 - Ae^{-kFt}} \quad \text{where } A = \frac{P_0 - F/c}{P_0}$$

As  $t \rightarrow \infty$ ,  $e^{-kFt} \rightarrow 0$  and so  $P \rightarrow F/c$ , this is the same limiting value as occurred in 3.1.

The *sigmoid* (S-like) graph of this function, known as the *logistic curve* is illustrated in figure 1. The corresponding curves for population growth according to our two earlier models are also indicated for the purpose of comparison.

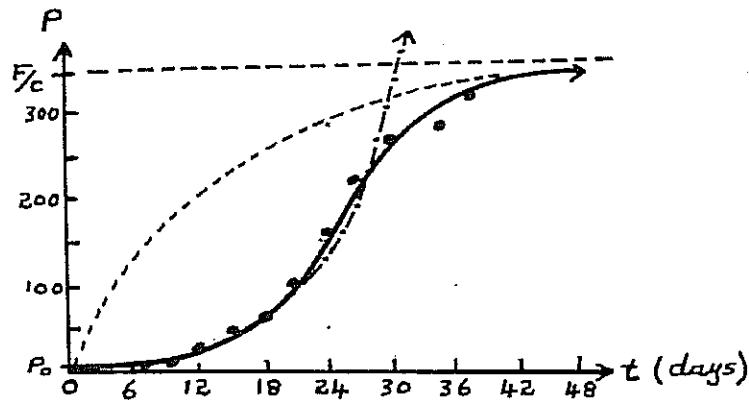


Figure 1. The Logistic Curve together with the corresponding curves (dotted) for our previous growth models. Dots represent experimental values of drosophila population.

The logistic curve is found to provide a good description in many instances of population development (particularly the growth of microbial populations). The particular curve illustrated in Figure 1 was chosen to describe the growth of a population of *Drosophila*. The dots represent the experimentally determined size of this population derived from data attributed to Pearl and Parker and adapted from Lotka.

Since a developing organism may be regarded as a growing colony of organised cells, the logistic curve is also useful in describing the growth of individual organisms. (Look up, and plot, some data on average height versus age for humans.)

*Epidemiology.* As we shall see, the logistic curve provides a useful description for the spread of a disease (or parasite).

Let  $n$  be the fixed size of a population composed of individuals all equally susceptible to infection by the disease. Take the disease to be such that once an individual is infected it remains so indefinitely. (These assumptions would be very nearly satisfied by many plant diseases; however, for epidemics in human populations more sophisticated treatments would be called for.)

Denoting by  $y(t)$  the number of individuals infected by time  $t$  and by  $x(t)$  the number susceptible to infection at time  $t$  we have  $y(t) + x(t) = n$  or  $x(t) = n - y(t)$ . If we now assume that the disease is transmitted through random contacts between infected and susceptible individuals we will have the rate of spread of the disease  $\left(\frac{dy}{dt}\right)$  proportional to the likely number of such contacts and so to both  $x(t)$  and  $y(t)$ .

Whence

$$\begin{aligned} \frac{dy}{dt} &= kyx \\ &= ky(n - y) \end{aligned}$$

and the epidemic progress according to a logistic curve.

*Collateral Reading.*

A careful discussion of growth according to the logistic curve and its relationship to other forms of growth is to be found in

Lotka "Elements of Mathematical Biology", Dover, 1956, Part II.

The application of logistic curves to the description of plant diseases is a dominant feature of

J.E. Van der Plank's "Plant Diseases", Academic, N.Y., 1963.

EXERCISES:

1. If the rate at which a mould grows on a piece of bread is assumed proportional to the amount of mould already present (i.e. to number of spores produced) and the amount of free bread surface left, develop a differential equation to describe the mould's growth.
2. Solve from first principles the differential equation

$$\frac{dy}{dt} = 0.05 y(10 - y), \text{ and check your answer.}$$

3. The logistic curve  $y = \frac{N}{1 + ke^{-\alpha t}}$  describes the number of infected individuals  $y$  at time  $t$  in a population of size  $N$  during an epidemic. Upon rearranging, this can be written as  $\ln \frac{y}{N - y} = \alpha t - \ln k$  i.e.  $\ln \frac{y}{N - y}$  is linearly related to  $t$ . (In fact in solving the differential equation  $y' = \alpha y - \frac{\alpha}{N} y^2$  we arrive at this relationship first, c.f. notes.) The simplest method of fitting a logistic curve to data is first to estimate  $N$  (as the value  $y$  tends to as  $t$  tends to infinity). Then for each data point plot  $\ln \frac{y}{N - y}$  against  $t$ . The slope and vertical intercept of the straight line of best fit through these points then give us the constants  $\alpha$  and  $\ln k$  respectively.

Arnold, 1971, in studying the epidemic progress of *Puccinia graminis tritici* on the heads of Lee variety Wheat obtained the following data:

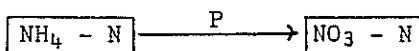
$y$ (Rust assessment)	$t$ (time in weeks)
0.25	10
2.3	17
8.75	22
9.5	28

By the above methods fit a logistic curve to this data. Plot both this theoretical curve and the actual data points on an accurate graph.

(Note: for calculation purposes it is better to write

$$y = \frac{N}{1 + ke^{-\alpha t}} \text{ as } y = \frac{N}{1 + e^{-\alpha t + \ln k}} .)$$

4. Part of the soil nitrogen system is illustrated diagrammatically by



where soil nitrogen, in the form of ammonium ions, is converted by a nitrifying population P to nitrates.

If N denotes the amount of nitrogen in the ammonium phase it is reasonable to take

$$\dot{N} = - aNP \text{ (where } a \text{ is a positive constant), why?}$$

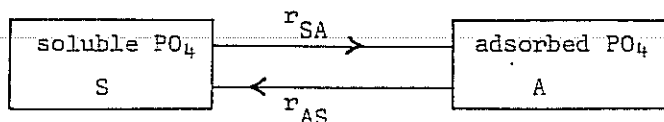
Further since P derives energy for its growth from this nitrogen we might assume that

$$\dot{P} = bNP \text{ (b is another positive constant).}$$

Show that this pair of equations implies that both N and P vary with time according to logistic curves.

PROJECT: A Simple Model for part of the Soil Phosphorus Cycle.

1. Introduction. No attempt is made to develop a complete model incorporating all the various states in which soil phosphorus occurs. Only the central transition between soluble phosphorus salts and adsorbed phosphorus is considered. Namely



Our basic assumptions concerning the transfer rates are

$$r_{SA} = S, (A_{sat} - A)$$

$$r_{AS} = A$$

Intuitively these are based on a simple capture - evaporation model.

*Rate of adsorption is proportional to the likelihood of a free PO<sub>4</sub> ion arriving at an adsorption site, and this in turn is proportional to the number of free PO<sub>4</sub> ions (measured by S) and to the number of adsorption sites (and so to A<sub>sat</sub> - A).*

*The rate of loss from the adsorbed state is proportional to the number of adsorbed ions with sufficient energy to escape, which, assuming a fixed energy distribution among the adsorbed ions, will be a fixed fraction of the adsorbed ions i.e. will be proportional to A.*

Thus  $r_{SA} = cS(A_{sat} - A)$

$$r_{AS} = kA$$

When the system is in equilibrium, i.e.  $r_{SA} = r_{AS}$ , show that  $S = \frac{k/c A}{A_{sat} - A}$  hence,

if  $A_{sat} = 6.4 \mu \text{ moles/cm}^3$  soil and the following values of S versus A are known

S( $\mu \text{ moles}$ )	A( $\mu \text{ moles/cm}^3$ soil)
5	4.4
10	5.3
20	5.9

determine an average value for  $k/c$ .

Graph the equilibrium relationship between S and A.

The Dynamic Model. In the closed system

$$-\frac{dA}{dt} = \frac{dS}{dt} = r_{AS} - r_{SA}$$

so 
$$\frac{dS}{dt} = kA - cS(A_{sat} - A)$$

$$= -\left[ cS^2 + (c(A_{sat} - P_t) + k)S - kP_t \right] \text{ where } P_t = A + S$$

a Ricatti equation, which may be integrated. Show that this leads to

$$S(t) = \frac{r_1 - r_2 C e^{ct} (r_2 - r_1)}{1 - C e^{ct} (r_2 - r_1)}$$

and so  $A(t) = P_t - S(t)$ , where  $r_2 < r_1$  are the roots of the quadratic

$$x^2 + (A_{sat} - P_t + \frac{k}{c})x - \frac{k}{c} P_t = 0.$$

Hence find  $r_1, r_2, C$  using  $S(0) = P_t = 10 \mu \text{ moles}$  and the value of  $\frac{k}{c}$  found above.

Graph this solution if  $c = 1$ ; what effect would changing the value of  $c$  have on the solution?

Lecture 6      *Non Linear Differential Equations II*  
 - allometry, Stevens' Law

Allometry is the study of relative growth, particularly of the distinct 'organs' of an individual. Thus, for example, it is readily observed that a child's limbs develop at a different rate from his torso.

Letting  $x_1(t)$  denote the size (length, volume, weight etc.) of an organ and  $x_2(t)$  that of another at time  $t$ , then it is reasonable to assume

$$\frac{dx_i}{dt} = f_i(x_i) \quad (i = 1, 2)$$

where  $f_1$  and  $f_2$  are known functions. (This is consistent with all of our previous growth models, some of which will be referred to soon.)

Our aim is to relate  $x_2$  to  $x_1$ . This would be worthwhile for instance, if the size  $x_1$  were easily measured while the direct measurement of  $x_2$  was difficult, perhaps requiring a fatal operation.

Using the function of a function theorem we have

$$\frac{dx_2}{dx_1} = \frac{dx_2/dt}{dx_1/dt} = \frac{f_2(x_2)}{f_1(x_1)}$$

This differential equation, relating  $x_2$  to  $x_1$ , may be separated and integrated to yield

$$\int_{x_2(0)}^{x_2} [f_2(x_2)]^{-1} dx_2 = \int_{x_1(0)}^{x_1} [f_1(x_1)]^{-1} dx_1$$

provided  $0 < x_i < z_i$  where  $z_i$  is the smallest positive zero of  $f_i$  ( $i = 1, 2$ ).

Turning now to specific cases, the simplest model of growth requires us to take  $f_i(x_i) = \gamma_i x_i$ ,  $\gamma_i$  being the specific growth rate of the  $i$ 'th organ assumed constant, and so

$$\frac{dx_2}{dx_1} = \left( \frac{\gamma_2}{\gamma_1} \right) \left( \frac{x_2}{x_1} \right) \quad \text{or}$$

$$\ln \frac{x_2}{x_2(0)} = \int_{x_2(0)}^{x_2} x_2^{-1} dx_2 = \left( \frac{\gamma_2}{\gamma_1} \right) \int_{x_1(0)}^{x_1} x_1^{-1} dx_1 = \left( \frac{\gamma_2}{\gamma_1} \right) \ln \frac{x_1}{x_1(0)}$$

All of which is justified since  $x_i(t) \neq 0$  for any  $t$ , and from which it follows that

$$x_2 = Ax_1^\gamma \quad \text{where } A = \frac{x_2(0)}{x_1(0)^\gamma} \quad \text{and } \gamma = \gamma_2/\gamma_1.$$

Thus  $x_2$  is related to  $x_1$  through a simple power function. This is the traditional form of the *Allometric Law* and is found to serve very well in practice.

For example, the weight of a mammal's brain  $x_2$  is found to be related to its forehead area  $x_1$  by

$$x_2(\text{grammes}) = 6x_1^{1.3} \quad (x_1 \text{ being measured in cm}^2)$$

to quite a high degree of accuracy.

We now examine the form of allometric law which results from one of our other models for growth. Hence, take

$$\frac{dx_i}{dt} = a_i - b_i x_i \quad (a_i, b_i > 0 \text{ being constants})$$

- see 3.1 .

Then following the argument outlined above, and observing that, for meaningful values of  $x_i$ ,  $\frac{dx_i}{dt} > 0$ , we obtain

$$x_2 = \frac{a_2}{b_2} - A(a_1 - b_1 x_1)^\gamma \text{ where } A = \frac{a_2 - b_2 x_2(0)}{b_2(a_1 - b_1 x_1(0))^\gamma}$$

$$\text{and } \gamma = b_2/b_1$$

This is a generalization of the traditional allometric law in the sense that for an appropriate choice of the constants it reduces to the traditional law ( $a_1 = a_2 = 0$  and  $b_1, b_2 < 0$ ), however such a choice of constants is quite inconsistent with the growth model assumed. We therefore have an alternative form of the allometric 'law' which is most conveniently written as

$$\frac{x_2(\infty) - x_2}{x_2(\infty) - x_2(0)} = \left\{ \frac{x_1(\infty) - x_1}{x_1(\infty) - x_1(0)} \right\}^\gamma$$

$x_i(\infty)$  denoting the maximum size of the fully developed  $i$ 'th organ i.e.  $x_i(\infty) = a_i/b_i$  ( $i = 1, 2$ ). A comparison of this and the traditional allometric law, arranged to agree at the initial and final organ sizes, is given in figure 1 below.

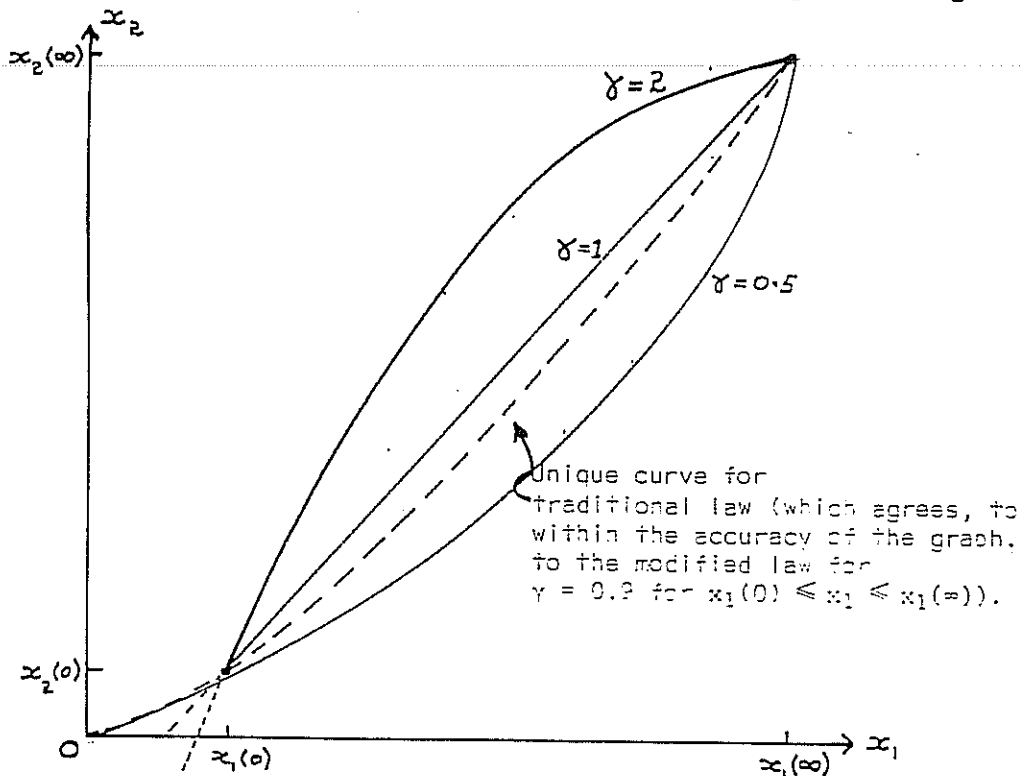


Figure 1

We could of course go on to derive the allometric laws corresponding to our other growth models, see Exercises 1 and 3.

The differential equation that produced the traditional allometric law,  $y' = ky/x$ , arises in connection with Stevens' 'Law' which appeared in the 1940's as a challenge to the long accepted Weber-Fechner Law, see 2.2 for background.

From a series of experiments Stevens concluded that, for *prothetic* stimuli at least\* (loudness, brightness, heaviness, taste, smell, temperature etc.), the constant fractional change in the physical magnitude of the stimulus expressed in Weber's Law corresponded to a constant fractional change in the psychological intensity (not an absolute change as postulated by Fechner).

This leads to

$$\frac{\Delta I}{I} = k \frac{\Delta S}{S} \text{ or } \frac{\Delta I}{\Delta S} = kI/S$$

and so for  $\Delta I$ ,  $\Delta S$  sufficiently small the relationship between  $I$  and  $S$  is expressed approximately by

$$\frac{dI}{dS} = kI/S.$$

Whence from our previous arguments  $I = AS^k$  and so a power relationship, rather than a logarithmic one, exists between  $I$  and  $S$ .

For example, brightness (in arbitrary psychological units) was found very nearly to vary according to  $0.01 S^{0.33}$  where  $S$  is the physical intensity in milliamberts.

In 1959 G. Ekman proposed that Weber's Law be modified to allow for the presence of a lower sensitivity threshold,  $t$ , below which a stimulus fails to register.

In which case Stevens' hypothesis produces

$$\frac{\Delta I}{I} = \frac{k\Delta S}{S+t} \text{ and so leads to the differential}$$

equation 
$$\frac{dI}{dS} = kI/(S+t).$$

Show that this may be solved to yield

$$I = A(S+t)^k \text{ - the Stevens' Ekman 'Law'.$$

In many cases  $|t|$  is sufficiently small that it may be safely neglected (Eg. for brightness  $t = -0.0372$ ) and so Stevens' Law applies. However, in other cases, such as heaviness, this is not so.

*Collateral Reading:*

A reasonably comprehensive account of allometry, from an essentially non mathematical stand point is given in

Alexander "Shape and Size".

For accounts of the work of Stevens and Ekman, reference is best made to their original publications, although the underlying mathematical simplicity of their assumptions is somewhat obscured.

---

\* There is some doubt as to which assumption is best in the case of metathetic stimuli such as pitch.



EXERCISES:

1. If organs grow according to differential equations of the form

$$\frac{dx}{dt} = \alpha x^r \text{ (some constants } \alpha, r > 0 \text{)}$$

Show that the size of two organs is related by the traditional allometric law

$$y = Ax^Y$$

if we assume  $y = 0$  when  $x = 0$ . (In fact this is the only common form of growth which leads to such an allometric law.)

2. It is suggested that the brain weight  $y$  (in gms) of a bird is related to its body weight  $x$  (in Kgms) by

$$y = k x^{0.6}$$

Assuming this, use the following data to estimate  $k$ .

Bird	$x$ (Kgms)	$y$ (gms)	$x^{0.6}$	$y/x^{0.6}$
Goldcrest	0.006	0.35		
House Sparrow	0.03	1		
Kestrel	0.2	4		
Common tern	0.25	3.2		
Buzzard	1	7.9		
Golden Eagle	5.6	22.4		

Average

Using this value of  $k$  graph the function  $y = k x^{0.6}$  and on the same graph indicate the above data points. Comment on the accuracy of the law.

(Adapted from: Alexander "Size and Shape".)

3. Derive an allometric law under the assumption that organs grow according to differential equations of the form

$$\frac{dx}{dt} = \alpha x(a - x).$$

Is the allometric law so deduced a "generalization" of the traditional law  $y = Ax^Y$ ?

Lecture 7 . . . . . *The numerical Solution of ordinary first order differential equations - The Euler Tangent Method.*

More complicated differential equations than those studied so far cannot, in very many cases, be solved explicitly. In such cases it is often necessary to fall back on a numerical solution. Here we examine the most simple method of obtaining such a 'solution' - The Euler tangent method\*.

Essentially the method is to convert the equation into a *forward difference equation* by replacing the derivative  $\frac{dy(t)}{dt}$  by the difference quotient  $\frac{y(t+h) - y(t)}{h}$  where h is an arbitrary (but fixed) increment.

We develop the method for differential equations of the form

$$\frac{dy}{dt} = f(y,t) \text{ with } y(0) = y_0$$

where, as we shall see, it affords a simple extension of the concept of a slope field.

If we choose to denote by  $y_n$  the *approximation* to the solution's  $y(t)$  value at  $t = nh$  we have

$$\frac{y_{n+1} - y_n}{h} \approx f(y_n, nh) \text{ or}$$

$$y_{n+1} \approx y_n + f(y_n, nh)h$$

From this relationship we are able to solve the problem iteratively as follows.

Knowing  $y_0$  we can find

$$y_1 = y_0 + f(y_0, 0)h$$

and so have an approximation to  $y$  at  $t = h$ . Then we may continue by finding

$$y_2 = y_1 + f(y_1, h)h$$

and so on till we have covered the range of  $t$  for which we are interested.

Graphically the methods proceed as indicated in figure 1.

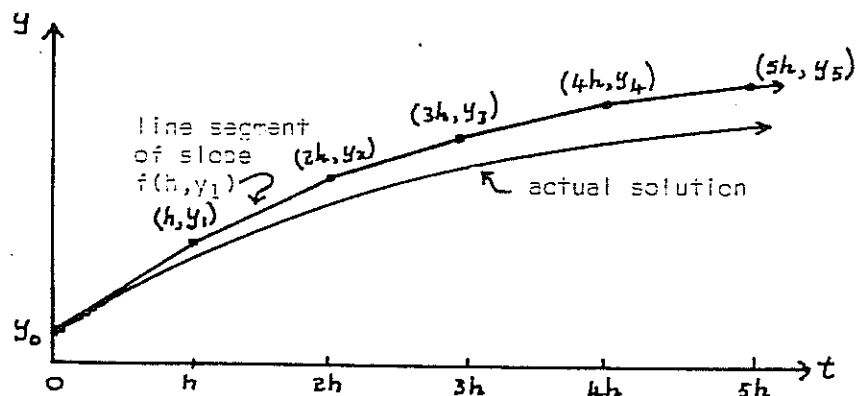


Figure 1

\* Because of the use Forrester made of this method, in connection with computer solutions to certain biological systems, it is, in some circles, known as Forrester's method. However one can hardly doubt the precedence of Leonhard Euler (1707 - 1783).

One disadvantage of the method is that it provides no means for estimating the error  $|y_n - y(nt)|$  at any stage. More advanced (and accurate) schemes for the numerical solution of such equations have been developed, some of which also provide error estimates. We will not consider these here, the interested reader being referred to the relevant books. One thing however should be clear, in most cases decreasing the value of  $h$  will improve the accuracy of the approximation

$$\frac{dy}{dt}(nh) \approx \frac{y_{n+1} - y_n}{h}$$

and so we would expect that taking smaller step sizes (smaller values for  $h$ ) would improve the accuracy of our approximations. Up to a certain point this is true, but another influence also comes into play. The smaller the value of  $h$  used the larger the number of steps needed to reach an approximation to the solution value at any  $t > 0$ . Since at each step inevitable round-off errors occur in the calculation, the accumulated effect of which increases as the number of steps does, we see that the error resulting from this factor will increase with decreasing step sizes.

Another disadvantage of the method is that an inadvertent error in the calculations at any stage will be propagated (and perhaps magnified) throughout the remaining calculations. Methods which avoid this last difficulty are known as *predictor-corrector* methods and are best employed when calculations are being done by hand and the possibility of slips is considerable.

Let us illustrate the Euler tangent method by applying it to the equation  $\frac{dy}{dt} = y(e^t - y)$  which might represent the growth of a parasite on an exponentially developing host (or any other situation where the "limiting factors" are increasing exponentially).

ILLUSTRATIVE EXAMPLE. For the differential equation

$$\frac{dy}{dt} = y(e^t - y) \text{ with } y(0) = y_0 = 0.5$$

we have for the approximate value of  $y$  at  $(n+1)h$

$$y_{n+1} = y_n + y_n(e^{nh} - y_n)h.$$

Taking  $h = 1$  we have

$$y_0 = 0.5 \text{ (given)}$$

$$y_1 = 0.5 + 0.5(e^0 - 0.5) \times 1 = 0.75, \text{ similarly}$$

$$y_2 = 2.23$$

$$y_3 = 13.7$$

While for  $h = 0.5$  we obtain

$$y_0 = 0.5$$

$$y_1 = 0.625$$

Note: here  $y_{2n}$  corresponds to  $y_n$  above.

$$y_2 = 0.95$$

$$y_3 = 1.78$$

$$y_4 = 4.18$$

$$y_5 = 10.89$$

$$y_6 = 17.9$$

$$y_7 = 37.3$$

These two approximations are compared graphically in Figure 2. If we proceed much further in either of these two instances the accumulation of errors becomes such that the value of  $y$  exceeds the corresponding value of  $e^t$  with the result that meaningless negative values are assigned to the solution. Thus except for values of  $t$  near 0 little credence can be placed on the approximations obtained. While the explicit solution to this equation cannot be given in terms of elementary functions the equation can nevertheless be solved (using the change of variable  $u = 1/y$  and an integrating factor - see Exercise 3.3) to give

$$y = e^{e^t} \left( e y_0^{-1} + \int_0^t e^{e^x} dx \right)^{-1}$$

From which we readily conclude that  $y > 0$  for all  $t$ . (Further an application of L'Hôspital's rule should quickly convince us that  $y \rightarrow \infty$  as  $t \rightarrow \infty$ . From analogy with 3. one might expect that  $y \rightarrow e^t$ , whereas in fact  $y \rightarrow e^t - 1$ .)

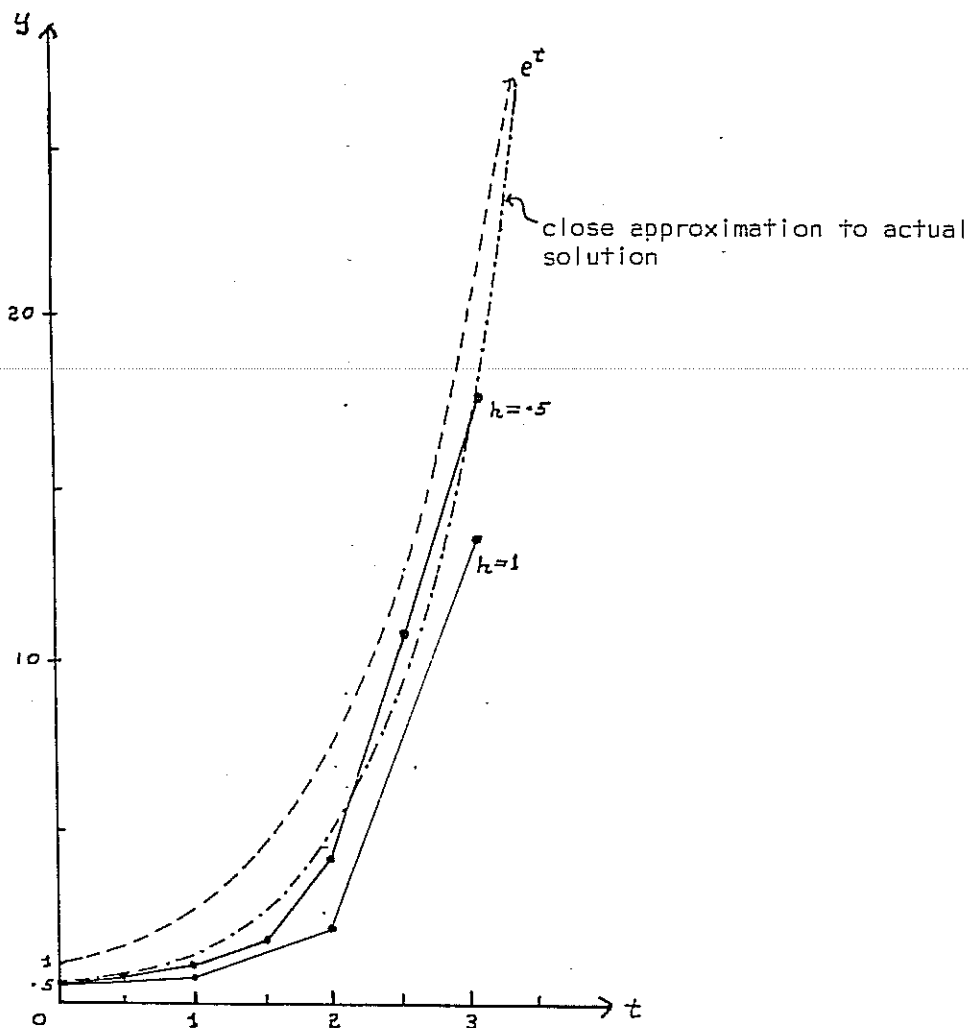


Figure 2

*Collateral Reading.*

Boyce and di Prima "Elementary Differential Equations and Boundary Value Problems", Wiley,

contains a useful account of the more advanced and accurate numerical methods available for solving differential equations, as do most books on Numerical Analysis.

EXERCISES:

1. Use the Euler tangent method to obtain approximate values for the solution of  $y' = 1 - 0.5y - 0.5y^{\frac{2}{3}}$  (see Exercise 1.1) at  $x = 0.5, 1, 1.5, \dots, 5$  if  $y(0) = 0$ .
2. The growth of a population (or the spread of an infection) is often described by an accumulated normal curve instead of a logistic one. Such curves are solutions of differential equations of the form  $\frac{dN}{dt} = Ae^{-(t-\beta)^2/\alpha}$ . Use the Euler tangent method with  $h = 0.1$  to plot the solution for the specific case when  $\frac{dN}{dt} = e^{-t^2}$  and  $N(0) = 1$ . (The study of curves such as these is the subject of Probit analysis.)
3. If you are familiar with a programming language, such as ALGOL or FORTRAN, you should develop a programme to numerically solve differential equations by the Euler-tangent method.

SERIES II - SYSTEMS OF DIFFERENTIAL EQUATIONS

Lecture 8 *Systems of Equations - Linear Systems*

"Traditionally" let us start by examining a population problem.

Consider a closed area containing two interacting species (one may provide; food, shelter, fertilization, etc. for the other, as in the instance of pollination of a plant species by bees, or alternatively it might inhibit the others growth through depletion of resources, etc.)

Denoting by  $x_1$  the size of the first population and by  $x_2$  that of the second we have

$$\frac{dx_1}{dt} = \left\{ \begin{array}{l} \text{rate of change of 1} \\ \text{in the absence of 2} \end{array} \right\} + \left\{ \begin{array}{l} \text{rate of change of 1 due to} \\ \text{the presence of 2} \end{array} \right\}$$

and

$$\frac{dx_2}{dt} = \left\{ \begin{array}{l} \text{rate of change of 2} \\ \text{in the absence of 1} \end{array} \right\} + \left\{ \begin{array}{l} \text{rate of change of 2 due to} \\ \text{the presence of 1} \end{array} \right\}$$

Making the simplest possible assumptions, that in the absence of interaction the rate of change of each species is proportional to the size of that species' population (c.f. 2) while the rate of change of one species due to the presence of the other interacting species is proportional to the size of the interacting species, the above equations become

$$\frac{dx_1}{dt} = ax_1 + bx_2$$

$$\frac{dx_2}{dt} = cx_1 + dx_2 \quad \text{where } a, b, c, d \text{ are constants.}$$

Thus our two unknown functions are related through a pair of simultaneous differential equations referred to as a *system of differential equations* (in this case a linear system, since each of the constituent equations is linear). Usually initial conditions, such as having  $x_1(0)$  and  $x_2(0)$  specified, are given along with such a system.

There is a great variety of means whereby systems such as this can be solved. The method we will follow (a matrix version of which is given in the appendix to this lecture) is a sort of guess and try procedure.

Noting that if  $b = 0$  the first equation of the system would yield  $x_1 = Ce^{at}$ , with a similar result for the second if  $c = 0$ , we are led to try for a solution of the form

$$x_1 = Ae^{\lambda t}$$

$$x_2 = Be^{\lambda t} \quad \text{with appropriate values of the constants } A, B \text{ and } \lambda.$$

Substituting these for  $x_1$  and  $x_2$  into the system and rearranging gives

$$(\lambda A - aA - bB)e^{\lambda t} = 0 \quad \text{and}$$

$(\lambda B - cA - dB)e^{\lambda t} = 0$  as a necessary and sufficient condition for them to be a solution. We will certainly have a solution, then, provided  $A, B$  and  $\lambda$  are chosen to satisfy

$$\lambda A - aA - bB = 0 \quad \text{and}$$

$$\lambda B - cA - dB = 0.$$

This is a pair of equations in three unknowns and so they do not uniquely determine all three of A, B and  $\lambda$ . However as we shall see they do uniquely determine  $\lambda$  and the ratio A/B.

*Determination of  $\lambda$ .* From the first of these two equations we have

$$\frac{B}{A} = \frac{\lambda - a}{b} \text{ while from the second } \frac{A}{B} = \frac{\lambda - d}{c}$$

whence 
$$\frac{A}{B} = \frac{\lambda - d}{c} = \frac{b}{\lambda - a} = \left(\frac{B}{A}\right)^{-1} \text{ or}$$

$$(\lambda - d)(\lambda - a) = bc.$$

This is a quadratic equation in  $\lambda$ ,  $\lambda^2 - (a + d)\lambda + (ad - bc) = 0$ , known as the *characteristic equation* of the system, from which the permissible values of  $\lambda$  may be obtained, viz.

$$\lambda = \frac{a + d \pm \sqrt{(a - d)^2 + 4bc}}{2}$$

A number of possible cases now arise.

CASE 1. There are two real and distinct values of  $\lambda$  (i.e.  $-4bc < (a - d)^2$ ).

Denote these values by  $\lambda_1$  and  $\lambda_2$ . Then for  $\lambda = \lambda_1$  we have

$$\frac{B}{A} = \frac{\lambda_1 - a}{b} \text{ (and } = \frac{c}{\lambda_1 - d} \text{)*}$$

Whence, leaving the value of A arbitrary, we have that

$$x_1 = Ae^{\lambda_1 t}$$

and

$$x_2 = \left(\frac{\lambda_1 - a}{b}\right)A e^{\lambda_1 t} \text{ as a solution of the system.}$$

Similarly, for  $\lambda = \lambda_2$  we have

$$x_1 = A'e^{\lambda_2 t}$$

and

$$x_2 = \left(\frac{\lambda_2 - a}{b}\right)A'e^{\lambda_2 t} \text{ as a second solution where}$$

in general A' need not equal A.

Since the system is linear a more general solution is obtained by adding these two together to give

$$x_1 = Ae^{\lambda_1 t} + A'e^{\lambda_2 t}$$

and

$$x_2 = \left(\frac{\lambda_1 - a}{b}\right)Ae^{\lambda_1 t} + \left(\frac{\lambda_2 - a}{b}\right)A'e^{\lambda_2 t}.$$

It can be shown that any solution of the system is in this case, of the above form. Values for the two arbitrary constants A and A' can now be chosen to

---

\* In some cases one of these two expressions for  $\frac{B}{A}$  may prove indeterminate, in which case the other should not be neglected. In the text above it is assumed that this is not the case.

satisfy the initial conditions, i.e. in the instance mentioned A and A' would be chosen as the solution to the pair of simultaneous equations

$$A + A' = x_1(0)$$

$$\text{and } \left(\frac{\lambda_1 - a}{b}\right)A + \left(\frac{\lambda_2 - a}{b}\right)A' = x_2(0).$$

ILLUSTRATIVE EXAMPLE. Let us consider two populations of size  $x_1$  and  $x_2$  interacting according to

$$\frac{dx_1}{dt} = 4x_1 + x_2$$

$$\frac{dx_2}{dt} = -2x_1 + x_2.$$

Thus the presence of  $x_2$  is advantageous to  $x_1$  while  $x_1$ 's existence is detrimental to the growth of  $x_2$ .

Substituting  $x_1 = Ae^{\lambda t}$ ,  $x_2 = Be^{\lambda t}$  yields

$$\lambda A - 4A - B = 0 \text{ and } \lambda B + 2A - B = 0 \text{ whence}$$

$$\frac{B}{A} = \frac{\lambda - 4}{1} = \left(\frac{A}{B}\right)^{-1} = \frac{-2}{\lambda - 1} \text{ so } (\lambda - 1)(\lambda - 4) = -2 \text{ or } \lambda^2 - 5\lambda + 6 = 0.$$

Factorizing this characteristic equation leads to

$$\lambda = 2 \text{ or } 3.$$

For  $\lambda = 2$  we have  $\frac{B}{A} = \frac{\lambda - 4}{1} = -2$  and so

$$x_1 = Ae^{2t}$$
$$x_2 = -2Ae^{2t}$$

while  $\lambda = 3$  gives

$$x_1 = A'e^{3t}$$
$$x_2 = -A'e^{3t}.$$

So the general solution is

$$x_1 = Ae^{2t} + A'e^{3t}$$
$$x_2 = -2Ae^{2t} - A'e^{3t}.$$

Thus if the initial sizes of the two populations were

$$x_1(0) = 25, \quad x_2(0) = 50, \text{ then}$$

$$A + A' = 25 \text{ and } -2A - A' = 50 \text{ from which we}$$

conclude  $A = -75$  and  $A' = 100$  so

$$x_1 = 100e^{3t} - 75e^{2t}$$

$$\text{and } x_2 = 150e^{2t} - 100e^{3t} \text{ (see Figure 1).}$$

Of course in the real situation once  $x_2$  was reduced to zero, at  $t = \ln 1.5 \approx 0.4$



it would remain so from then on, while  $x_1$  would continue to develop according to  $\frac{dx_1}{dt} = 4x_1$ . This situation is also illustrated in figure 1.

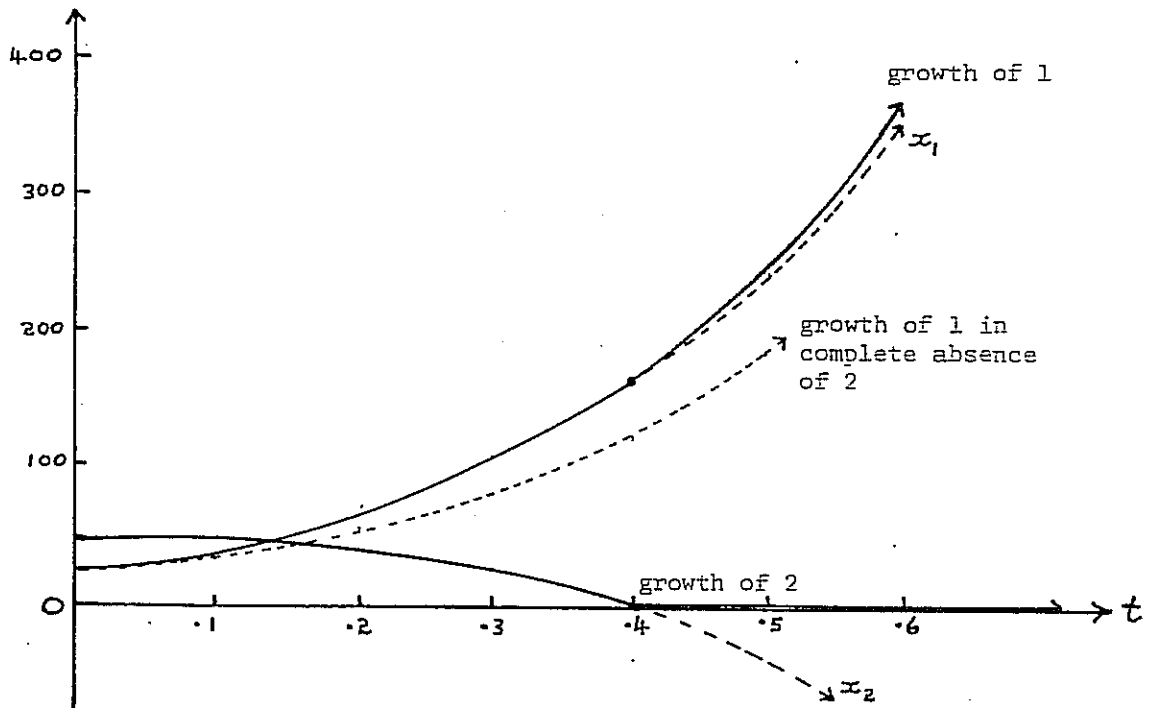


Figure 1

We will now consider the remaining two cases briefly.

CASE 2. The two values of  $\lambda$  are complex conjugates with non-zero imaginary parts (i.e.  $-4bc > (a - d)^2$ ). In this case denote the values of  $\lambda$  by  $\lambda_1 = \alpha + i\beta$  and  $\lambda_2 = \alpha - i\beta$  ( $\alpha, \beta$  real).

Here the calculations are precisely the same as in the previous case except that complex numbers are admitted throughout. Thus the solution is

$$x_1 = e^{\alpha t} (Ae^{i\beta t} + A'e^{-i\beta t})$$

$$x_2 = e^{\alpha t} \left( \frac{\alpha + i\beta - a}{b} Ae^{i\beta t} + \frac{\alpha - i\beta - a}{b} A'e^{-i\beta t} \right)$$

where  $A, A'$  are arbitrary complex numbers.

However it is readily seen that both the real and imaginary parts of this solution must also be solutions and so using the Euler formula we obtain real oscillatory solutions of the form

$$x_1 = e^{\alpha t} (B \sin \beta t + C \cos \beta t)$$

$$x_2 = e^{\alpha t} (D \sin \beta t + E \cos \beta t)$$

where  $B, C, D$  and  $E$  are appropriately inter-related constants whose values may be determined from the equation in conjunction with the initial conditions.

CASE 3. Here the two values of  $\lambda$  are real and equal (i.e.  $-4bc = (a - d)^2$ ). Since this case arises only infrequently in practice we will not pursue it further, except to note that the general solution is of the form

$$\begin{aligned} x_1 &= (At + B)e^{\lambda t} \\ x_2 &= (Ct + D)e^{\lambda t} \end{aligned} \quad \text{where } A, B, C \text{ and } D \text{ are appropriately}$$

interrelated constants.

*Collateral Reading.*

An account of systems of equations from a matrix viewpoint (see appendix) is to be found in

Boyce & DiPrima "Elementary Differential Equations and Boundary Value Problems.

A mathematically more advanced treatment is given in

H. Hochstadt "Differential Equations, a modern approach".

From the applications point of view a brief treatment is to be found in

S.I. Grossman and J. Turner "Mathematics for the Biological Sciences" and a slightly more extensive one in

E. Batschelet "Introduction to Mathematics for Life Scientists".

APPENDIX to Lecture 8.

MATRIX METHODS FOR FIRST ORDER LINEAR SYSTEMS OF DIFFERENTIAL EQUATIONS

(Most useful for systems of three or more equations, although we illustrate the method on systems of 2 equations.)

Consider the system

$$\frac{dx_1}{dt} = ax_1 + bx_2 \quad x_1 \equiv x_1(t)$$

where  $x_1$  and  $x_2$  are 2 unknown functions.

$$\frac{dx_2}{dt} = cx_1 + dx_2 \quad x_2 \equiv x_2(t)$$

If we form the vector  $\underline{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$  whose components are the two unknown functions and choose to define  $\frac{d\underline{x}}{dt}$  by

$$\frac{d\underline{x}}{dt} = \begin{pmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{pmatrix}, \text{ i.e. the vector whose}$$

components are the derivatives of the corresponding components of  $\underline{x}$ , then we may write the above system as the matrix equation

$$\frac{d\underline{x}}{dt} = A\underline{x} \text{ where } A \text{ is the } 2 \times 2 \text{ matrix } \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

From analogy with the single first order equation  $\frac{dx}{dt} = ax$  it is reasonable to try a solution of the form

$$\underline{x} = \underline{c}e^{\lambda t} = \begin{pmatrix} c_1 e^{\lambda t} \\ c_2 e^{\lambda t} \end{pmatrix} \text{ where } \underline{c} \text{ is a constant vector } \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

In which case, since  $\frac{dx}{dt} = \lambda x = \lambda c e^{\lambda t}$  (check that this follows trivially from the above definition), the above matrix equation becomes

$$\lambda c e^{\lambda t} = A c e^{\lambda t}$$

For this to be true for all values of  $t$ , as required if  $x = c e^{\lambda t}$  is to be a solution, we need  $\lambda$  and  $c$  to satisfy the matrix equation

$$\lambda c = A c$$

Rearranging this we therefore require  $c$  to be a non-trivial solution of

$$[A - \lambda I]c = 0$$

Now such an homogeneous system only has non-trivial solutions if the matrix  $[A - \lambda I]$  is singular and so the permissible values of  $\lambda$  are those for which the determinant  $|A - \lambda I| = 0^*$ , i.e. are roots of the "characteristic" equation of  $A$ .

Corresponding to each value of  $\lambda$  so determined,  $\lambda_i$  say, we can then find a non-zero vector  $c_i$  by solving  $[A - \lambda_i I]c_i = 0$ .

(Note: There is not one such  $c_i$  but infinitely many, for if  $c_i$  is a solution then so is  $k c_i$  for any arbitrary constant  $k \neq 0$ .)

For each distinct  $\lambda_i$  we are thus able to construct the solution  $x_i = c_i e^{\lambda_i t}$ .

If all the  $\lambda_i$  are distinct, the most general solution is a linear combination of these solutions, viz

$$\underline{x} = \sum_i k_i c_i e^{\lambda_i t}$$

**ILLUSTRATIVE SOLUTION:-** The system  $\frac{dx}{dt} = 2x + y$

$$\frac{dy}{dt} = 3x$$

may be written as the matrix equation  $\frac{dx}{dt} = Ax$  where  $x = \begin{pmatrix} x \\ y \end{pmatrix}$  and  $A = \begin{bmatrix} 2 & 1 \\ 3 & 0 \end{bmatrix}$  which has solutions  $x = c e^{\lambda t}$  where  $\lambda$  is determined from

$$|A - \lambda I| = \begin{vmatrix} 2-\lambda & 1 \\ 3 & -\lambda \end{vmatrix} = \lambda^2 - 2\lambda - 3 = 0$$

so  $\lambda = -1$  or  $3$  while  $c = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$  satisfies  $[A - \lambda I]c = 0$

so for  $\lambda = -1$   $\begin{bmatrix} 3 & 1 \\ 3 & 1 \end{bmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = 0$  or  $3c_1 + c_2 = 0$ . Taking  $c_1 = 1$  gives  $c_2 = -3$

and a suitable  $c$  is  $\begin{pmatrix} 1 \\ -3 \end{pmatrix}$  whence one solution is  $x_1 = \begin{pmatrix} 1 \\ -3 \end{pmatrix} e^{-t}$

similarly for  $\lambda = 3$   $\begin{bmatrix} -1 & 1 \\ 3 & -3 \end{bmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = 0$  so a suitable  $c$  is  $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$  and a second solution is  $x_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{3t}$

\* Such values of  $\lambda$  are called *eigenvalues* of  $A$  and the corresponding non-trivial solutions for  $c$  are then termed *eigenvectors*.

Thus the general solution is

$$\underline{x} = a \begin{pmatrix} 1 \\ -3 \end{pmatrix} e^{-t} + b \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{3t}$$

(where a, b are arbitrary constants whose values may be determined from the initial conditions) or in the language of our original problem

$$x = ae^{-t} + be^{3t}$$
$$y = -3ae^{-t} + be^{3t}.$$

EXERCISES:

1. The growth of two species of sizes  $N_1$  and  $N_2$  which react detrimentally is described by the system

$$\frac{dN_1}{dt} = 4N_1 - N_2$$

$$\frac{dN_2}{dt} = -4N_1 + 4N_2$$

Solve this system and find  $N_1, N_2$  if at time  $t = 0, N_1 = 400$  and  $N_2 = 1,000$ .

Draw rough graphs of their growth. On the same graphs also show the growth of each species in the absence of the other.

2. Two interacting populations have sizes of  $N_1 = A - x_1(t), N_2 = B - x_2(t)$  at time  $t$  where  $A, B$  are constants and where  $\frac{dx_1}{dt} = -x_1 + x_2$   
 $\frac{dx_2}{dt} = -x_1 - x_2$ .

If  $A = B = 1,000$  and at  $t = 0, N_1 = 1,000, N_2 = 800$ , show that this system has solutions of the form

$$x_1 = e^{-\mu t} (A_1 \sin \beta t + B_1 \cos \beta t)$$

$$x_2 = e^{-\mu t} (A_2 \sin \beta t + B_2 \cos \beta t)$$

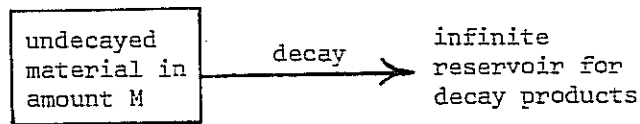
Hence find  $N_1$  and  $N_2$  as functions of  $t$ . Draw graphs of these. What can you say about  $N_1$  and  $N_2$  as  $t \rightarrow \infty$ ?

Lecture 9      *Linear Systems from an alternative view point*  
- *Compartmental Analysis*

A useful concept in developing biophysical models is that of a compartment, which Milhorn (in "Applications of control theory to physiological systems", 1966) defines as follows.

"If a substance is present in a biological system in several distinguishable forms or locations, and if it passes from one form or location to another form or location (at a measurable rate) then each form or location constitutes a separate *compartment* for the substance."

In Lecture 2.3 on decay processes, if we regard the undecayed material as forming one compartment the process could be schematically represented thus

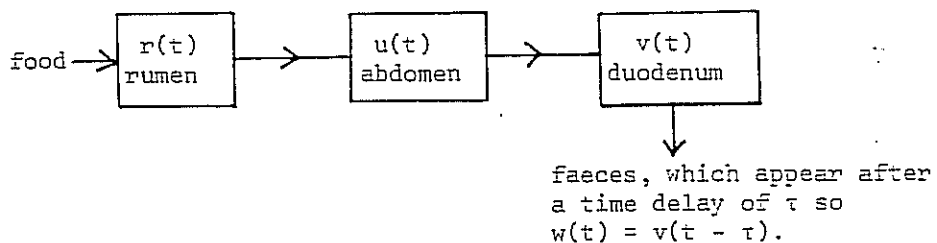


where the rate at which material is transferred from the compartment of undecayed material to an infinite reservoir for decay products was taken proportional to  $M$ . Some authors take this infinite reservoir to constitute a second compartment, but many do not. Quite arbitrarily we will side with the latter and so designate the above process as a one compartment model.

Dividing a complex biological process up into compartments with various transfers between different compartments often assists us in developing a system of equations which model the process.

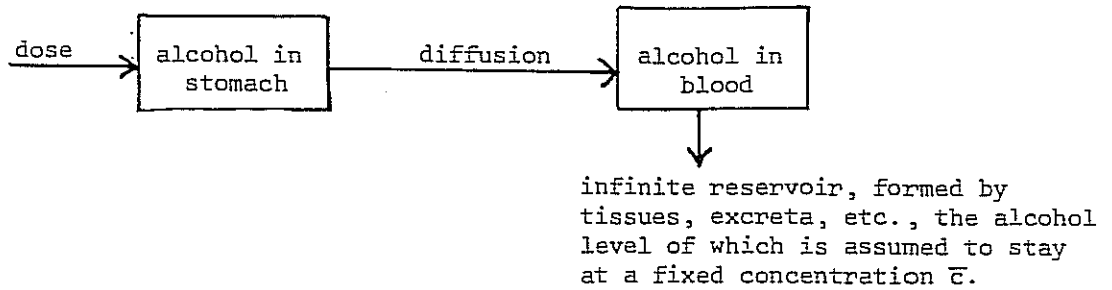
Other examples, in this case of two compartment models, were provided by the soil phosphorus model in the project to Lecture 5 and the soil nitrogen system of exercise 5.4. Many of the other situations so far considered could also be interpreted in terms of compartments (find a few).

Batschelet discusses the passage of food through the digestive tract of a ruminant in terms of a three compartment model, viz



Some researchers have developed more elaborate models involving nine and more compartments.

A simpler example, which we will consider in detail, is afforded by the *excretion of a drug*. For the sake of clarity we will take the case of the blood alcohol level resulting from a concentrated (at least in time) dose of alcohol. A suitable compartmental model for this situation is



Let  $c_1$  be the excess concentration of alcohol in the stomach at time  $t$  above the constant external concentration  $\bar{c}$ . Similarly let  $c_2$  be the excess blood alcohol concentration above  $\bar{c}$ . Initially  $c_2(0) = 0$  and  $c_1(0)$  is proportional to the size of the alcohol dose,  $D$ , administered i.e.  $c_1(0) = \alpha D$ .

Since the transfer mechanisms are essentially diffusion processes we have, from 3.2, that

$$\frac{dc_1}{dt} = \left\{ \begin{array}{l} \text{difference in alcohol concentration between the stomach and} \\ \text{the blood} \end{array} \right\}$$

and so

$$\frac{dc_1}{dt} = -a(c_1 - bc_2)$$

where  $a > 0$  is the proportionately constant and  $b > 0$  allows for  $c_1$  and  $c_2$  to be measured in different units. Also

$$\frac{dc_2}{dt} = \left\{ \begin{array}{l} \text{rate at which alcohol} \\ \text{enters the blood} \end{array} \right\} - \left\{ \begin{array}{l} \text{rate at which alcohol} \\ \text{leaves the blood} \end{array} \right\}$$

and so

$$\frac{dc_2}{dt} = -\frac{1}{b} \frac{dc_1}{dt} - dc_2 = \frac{a}{b} c_1 - (a + d)c_2 \text{ where } d > 0 \text{ is the coefficient of}$$

diffusion for alcohol leaving the blood stream.

We are thus led to the linear system of equations\*

$$\frac{dc_1}{dt} = -ac_1 + abc_2 \quad c_1(0) = \alpha D$$

$$\frac{dc_2}{dt} = \frac{a}{b} c_1 - (a + d)c_2 \quad c_2(0) = 0$$

with  $a, b$  and  $d$  positive constants, which may be solved by the methods discussed in the last lecture.

Thus if  $\lambda_1$  and  $\lambda_2$  are the two roots of

$$\lambda^2 + (2a + d)\lambda + ad = 0, \text{ which are real (since } (2a + d)^2 - 4ad = 4a^2 + d^2 > 0) \text{ and both negative;}$$

we have 
$$c_1 = Ae^{\lambda_1 t} + A'e^{\lambda_2 t}$$

$$c_2 = \frac{\lambda_1 + a}{ab} Ae^{\lambda_1 t} + \frac{\lambda_2 + a}{ab} A'e^{\lambda_2 t}$$

\* Not all compartmental models yield linear systems, viz the project to lecture 5, however with simple assumptions on the transfer rates many do.

So using  $c_2(0) = 0$  gives  $A' = -\frac{\lambda_1 + a}{\lambda_2 + a} A$  and then  $c_1(0) = \alpha D$  gives

$$A = \frac{\alpha}{1 + \frac{\lambda_1 + a}{\lambda_2 + a}} D$$

So finally we can conclude

$$c_2 = \frac{\alpha(\lambda_1 + a)(\lambda_2 + a)}{ab(\lambda_2 - \lambda_1)} D [e^{\lambda_1 t} - e^{\lambda_2 t}]$$

a result best written as

$$c_2 = kDe^{\lambda_1 t} [1 - e^{(\lambda_2 - \lambda_1)t}] \text{ where } k = \frac{\alpha(\lambda_1 + a)(\lambda_2 + a)}{ab(\lambda_2 - \lambda_1)}$$

and without loss of generality we may assume  $\lambda_2 \ll \lambda_1$ .

Drew, Colquhoun and Long in "Alcohol and a skill resembling driving" (H.M. Stationery Office, London, 1959) give a considerable amount of empirically determined values of  $c_2$  at various times and for various dose sizes. Fitting their findings to a curve of the form obtained above, by the important procedure known as *peeling* yields

$$c_2(\text{in mgms/100 mls of blood}) = 260De^{-0.01t} (1 - e^{-0.04t})$$

where  $D$  is measured in gms alcohol/Kgm of body weight and  $t$  is in minutes. This curve and their original data, for the case of  $D = 0.5$  gms/Kgm body weight, is plotted in Figure 1 from which the close agreement between the experimental results and the theoretically predicted curve should be noted.

As can be easily checked, the system

$$\frac{dc_1}{dt} = -0.02 (c_1 - bc_2)$$

$$\frac{dc_2}{dt} = \frac{0.02}{b} c_1 - 0.05c_2$$

yields approximately the correct values of  $\lambda_1$  and  $\lambda_2$  for any choice of  $b$  and so for appropriate choice of  $a$  and  $b$  ( $\frac{a}{b} \doteq 108$ ) has the above empirically determined function for its solution.

However, even disregarding the indeterminate nature of  $a$  and  $b$ , the values of  $a$  and  $d$  used here are not unique, as a few simple calculations will soon show. Thus from the completely determined nature of  $c_2$  we cannot assign a unique set of values to the underlying constants (parameters) of the model and so certainly cannot make any adequate predictions about the various transfer rates etc. from a knowledge of the systems out-put ( $c_2$ ) for various inputs ( $D$ ).

This is an extremely important point, the conclusions from it holding true of most systems. That is a *knowledge of the input-output response of a system does not uniquely determine the internal structure of the system*. A result well documented in engineering works on control theory, particularly in the paper of Kalman, who divides a system into observable (determinable) and unobservable

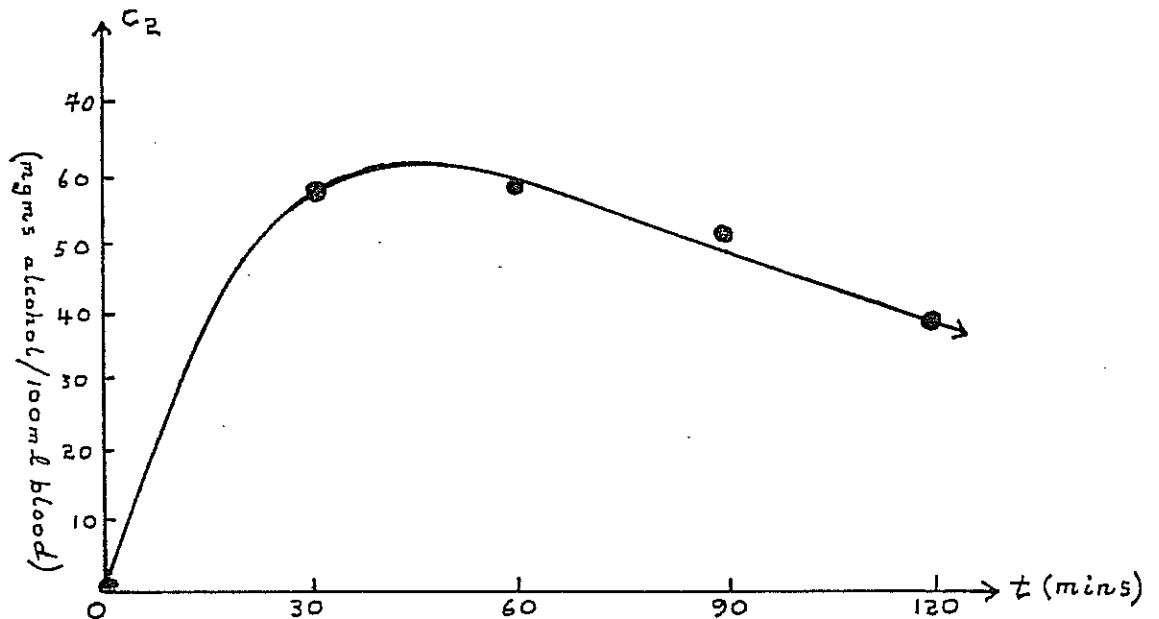


Figure 1. Dots represent average values as found by Drew et al, the curve is  $c_2 = 260e^{-0.01t}(1-e^{-0.04t})$ . (Here  $D = 0.5$  gms/Kgm body weight.)

(indeterminable) components, and shows that most systems have an unobservable part. Thus the somewhat common practice of adjusting the parameter values of a system till the known input - output response is obtained (often by means of an analogue computer) and then interpreting those parameter values as though they reflected the underlying physical situation is at the very least of doubtful value.

*Collateral Reading.*

Compartmental Models are discussed in sections 11.3(f) and 11.7(b) and (c) of

E. Batschalet "Introduction to Mathematics for Life Scientists".

The following book is devoted entirely to discussions of compartmental analysis and so serves as a detailed general reference:

M.A. Atkins, "Multicompartmental Models for Biological Systems", Methuen, 1969.

For the work of Kalman cited in the lecture, the most important paper is R.E. Kalman "Mathematical Description of Linear Dynamical Systems" J.S.I.A.M. Control, Vol. 1, No. 2 pp. 152-192.

EXERCISE:

1. Show that the peak (maximum) blood alcohol level occurs after a fixed number of minutes, independent of the initial dose size. Hence conclude that the peak blood alcohol level is directly proportional to the dose size (a conclusion found empirically by Drew et al).



Lecture 10    *Non-Linear Systems - Lotka-Volterra Equations.*

The particular non-linear system we will look at was first used by D'Ancona et al to examine the shark versus herbivorous fish populations of the Adriatic during the years 1910 to 1923. (When fishing resumed after the fishing fields had been 'rested' during World War I, it was expected that rich harvests would result. This did not turn out to be the case and explanations were sought by D'Ancona and others.) The system is however applicable to more general *Predator-Prey* and other situations, while the techniques illustrated in the analysis of the system can be used on any similar non-linear system of two equations.

If we assume that the environment is static and any genetic adaptation is slow enough as to be negligible (i.e. no external time-dependent influences are present) then a predator-prey situation will be described by

$$\left\{ \begin{array}{l} \text{rate of change of} \\ \text{prey} \end{array} \right\} = \left\{ \begin{array}{l} \text{natural rate of} \\ \text{increase of prey} \end{array} \right\} - \left\{ \begin{array}{l} \text{rate of destruction due} \\ \text{to predation} \end{array} \right\}$$

while

$$\left\{ \begin{array}{l} \text{rate of change of} \\ \text{predators} \end{array} \right\} = \left\{ \begin{array}{l} \text{rate of increase of} \\ \text{predator from} \\ \text{devouring prey} \end{array} \right\} - \left\{ \begin{array}{l} \text{natural death rate} \\ \text{of predators} \end{array} \right\}$$

Here we have further assumed that the predator's food supply is entirely composed of the prey and that once established, an individual predator always finds enough food and so its death is due to natural causes, not starvation. (Should any of these assumptions prove not to be the case, an allowance could be made, In some cases this would only lead to a change in parameter values, for example if starvation were admitted as a cause of predator deaths - why? - for others a considerable complexification of the model might result.)

Denoting by  $x$  the number of prey present and by  $y$  the number of predators, the simplest assumptions lead us to take

$$\{\text{natural rate of increase of prey}\} = ax$$

$$\text{and } \{\text{natural death rate of predators}\} = dy$$

Now it is reasonable to assume that the "destruction rate of prey by predator" is proportional to the likelihood of encounter between predator and prey and so to both  $x$  and  $y$ . It is then equally plausible to take

$$\left\{ \begin{array}{l} \text{rate of increase of} \\ \text{predators from devouring} \\ \text{prey} \end{array} \right\} = \left\{ \begin{array}{l} \text{rate of destruction} \\ \text{of prey by predators} \end{array} \right\}$$

and therefore to  $xy$ .

The presence of the cross-product terms " $xy$ " introduce non-linearity into the system which becomes

$$\frac{dx}{dt} = ax - bxy$$

$$\frac{dy}{dt} = cxy - dy \quad \text{where } a, b, c, d > 0 \text{ are constant.}$$

Single non-linear first order equations are often more difficult (if not impossible) to solve than linear ones. This applies even more so to non-linear systems. Thus for the relatively simple case derived above, it proves impossible to find a completely explicit solution. To find the explicit form of the solution in any particular case we would have to approximate it using numerical methods (see Exercise 3) probably requiring the use of a computer. However, even in the absence of an explicit solution, much insight into the form and behaviour of solutions can be gained.

A non-linear system of the form

$$\frac{dx}{dt} = f(x, y)$$

$$\frac{dy}{dt} = g(x, y)$$

where the right-hand sides do not depend explicitly on  $t$  is called an *autonomous* system, for example the system derived above is autonomous.

If  $x \equiv x(t)$ ,  $y \equiv y(t)$  is the solution of a given autonomous system, which satisfies the initial condition  $x(0) = x_0$ ,  $y(0) = y_0$ , we can regard it as the parametric representation of a 'curve',  $C$ , in the  $x$ - $y$  plane. In this context the  $x$ - $y$  plane is known as the *phase-plane* of the system.

This  $C = \{(x,y): x = x(t), y = y(t), t \in \mathcal{R}\}$  containing the point  $(x_0, y_0)$ , known as the *trajectory* of the solution passing through  $(x_0, y_0)$ , is the natural projection (shadow) of the solution in  $(x, y, t)$ -space onto the phase-plane. Although a knowledge of the trajectory passing through the point  $(x_0, y_0)$  does not give complete knowledge of the solution satisfying the initial condition  $x(0) = x_0$ ,  $y(0) = y_0$  it does however, as we shall see, yield much valuable information about that solution. It is therefore important that we be able to construct trajectories passing through various initial points. Under appropriate conditions on  $f$  and  $g$  the general theory of autonomous systems (and in particular the uniqueness of their solution) yields the following important result.

*If two trajectories have any points in common then they are identical, or less precisely, distinct trajectories cannot cross.*

Of particular importance are those trajectories which consist of a single point  $(a, b)$  say, such points being known as *critical points* of the system. A necessary and sufficient condition that  $(a, b)$  be a critical point is given by

$$\frac{da}{dt} = 0 = f(a, b)$$

$$\frac{db}{dt} = 0 = g(a, b)$$

i.e. that  $(a, b)$  be a solution of  $f(x, y) = 0$  and  $g(x, y) = 0$ .

Thus for our predator-prey system we have

$$ax - bxy = 0 \Rightarrow x = 0 \text{ or } y = \frac{a}{b}$$

$$cxy - dy = 0 \Rightarrow y = 0 \text{ or } x = \frac{d}{c} \quad \text{so the}$$

critical points are  $(0, 0)$  and  $(\frac{d}{c}, \frac{a}{b})$ .

---

\* Sufficient conditions would be that  $f$ ,  $g$ ,  $\frac{\partial f}{\partial x}$ ,  $\frac{\partial f}{\partial y}$ ,  $\frac{\partial g}{\partial x}$  and  $\frac{\partial g}{\partial y}$  are everywhere continuous (see later).

As a general rule, the other trajectories of a system either approach, recede from, or encircle the critical points which physically correspond to 'equilibrium states' for the system.\*

It should be apparent that  $C$  is a trajectory for the system if it is a solution curve of the first order equation

$$\frac{dy}{dx} \left( = \frac{dy/dx}{dt/dt} \right) = \frac{g(x,y)}{f(x,y)}$$

This provides one means of determining the system's trajectories.

So we seek trajectories for our predator-prey system as solutions of

$$\frac{dy}{dx} = \frac{cxy - d}{ax - bxy}$$

which may be separated to give

$$\int_{y_0}^y \frac{a - by}{y} dy = \int_{x_0}^x \frac{cx - d}{x} dx$$

or 
$$a \ln \frac{y}{y_0} - b(y - y_0) = c(x - x_0) - d \ln \frac{x}{x_0}$$

Unfortunately this implicit relationship cannot be solved to give  $y$  in terms of  $x$ . However for any fixed set of initial values  $(x_0, y_0)$  and any  $x$ -value we could first determine if the equation

$$a \ln y - by = c(x - x_0) - by_0 - d \ln \frac{x}{x_0} + a \ln y_0$$

(where the right-hand side is now a known constant) has a solution in  $y$  and if so find successive approximations to it by Newton's method of locating zeros, say. In this way a graph of  $y$  versus  $x$  could slowly be built up and so the trajectory through  $(x_0, y_0)$  determined. From such a procedure (or directly from the implicit relationship between  $y$  and  $x$ ) it would emerge that the trajectories form closed curves centred about the critical point  $\left(\frac{d}{c}, \frac{a}{b}\right)$  at least for  $(x_0, y_0)$  sufficiently near to this point.

A less strenuous approach is to construct a slope field for the equation

$$\frac{dy}{dx} = \frac{cxy - d}{ax - bxy}$$

Since from physical grounds we are only interested in  $x, y \geq 0$  we will concentrate on this quadrant. The slope field is best built up by first plotting isoclines. Thus observe that:

$$\begin{aligned} \frac{dy}{dx} = 0 & \text{ on the lines } y = 0 \text{ and } x = \frac{d}{c}; \\ \frac{dy}{dx} = \infty & \text{ on the lines } x = 0 \text{ and } y = \frac{a}{b}; \end{aligned}$$

---

\* The analysis of this situation is the subject of Stability Theory which will be taken up in later lectures.

$$\frac{dy}{dx} = 1 \text{ on the branches of the curve } y = \frac{ax}{(b+c)x - d}$$

$$\frac{dy}{dx} = -1 \text{ on the branches of the curve } y = \frac{ax}{(b-c)x + d}$$

These isoclines and the slopes along them are shown in figure 1 where a number of typical trajectories have also been sketched.

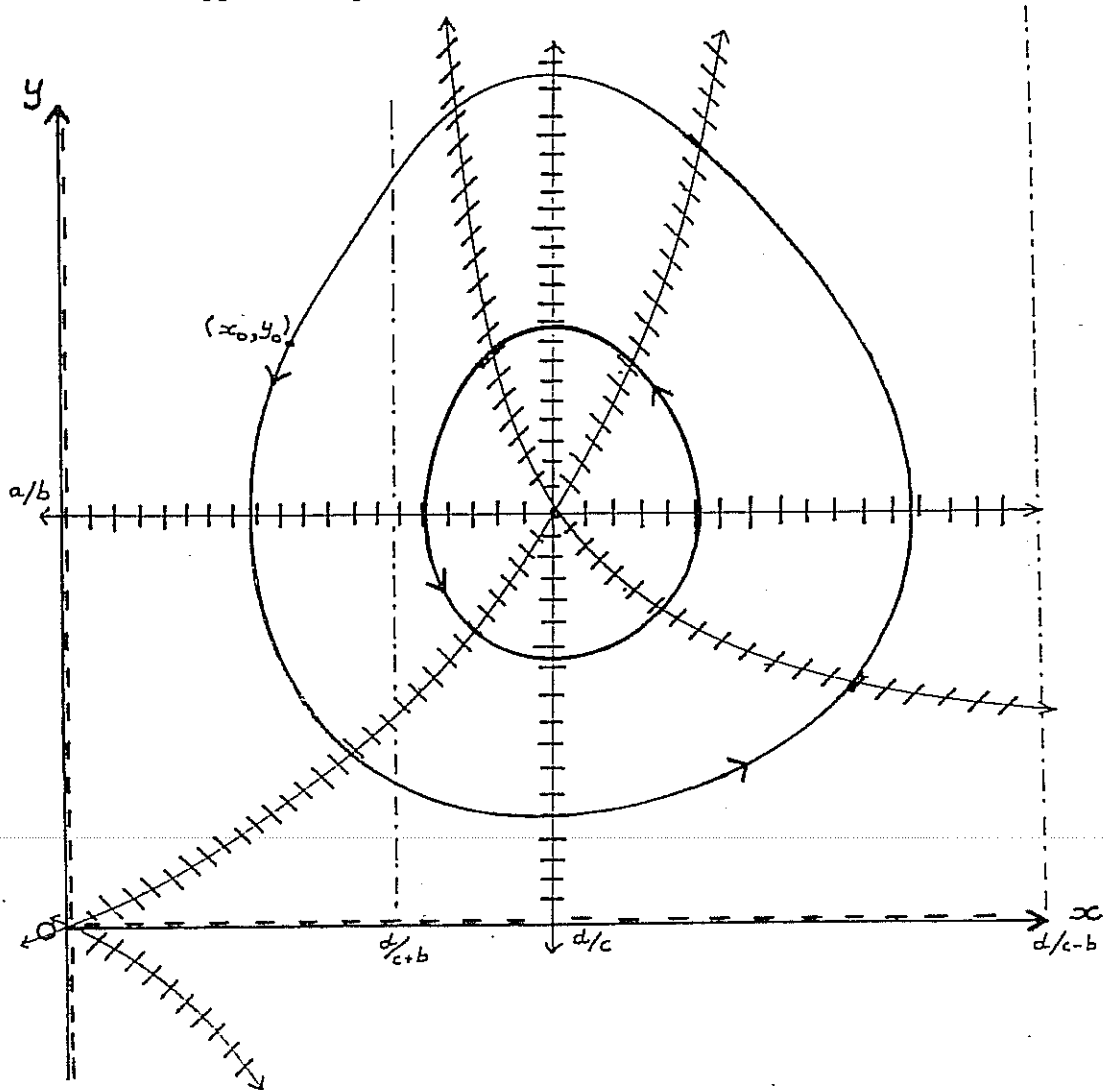


Figure 1. Phase-plane of the system  $\frac{dx}{dt} = ax - bxy$   
 $\frac{dy}{dt} = cxy - dy$  ( $c > b$ )

We now illustrate how the "phase-portrait" developed in figure 1 provides information about the solution of the system

$$\frac{dx}{dt} = ax - bxy$$

$$\frac{dy}{dt} = cxy - dy \quad \text{with } x(0) = x_0, y(0) = y_0 > 0.$$

As we have observed, points on the unique trajectory  $C$  passing through  $(x_0, y_0)$  are specified parametrically by  $x = x(t)$ ,  $y = y(t)$  where  $x(t)$ ,  $y(t)$  is the solution of the system. Since both  $x$  and  $y$  are continuous (in fact differentiable) functions of  $t$ , as the parameter value  $t$  increases the point  $(x(t), y(t))$  will move "smoothly" along  $C$ . For example, if  $(x_0, y_0)$  is such that  $C$  is one of the closed curves 'centred' about  $\left(\frac{d}{c}, \frac{a}{b}\right)$  and  $x_0 < \frac{d}{c}$ ,  $y_0 > \frac{a}{b}$ , then at  $(x_0, y_0)$  a simple calculation shows  $\frac{dx}{dt} < 0$  and  $\frac{dy}{dt} < 0$  and so as  $t$  increases the solution point moves anticlockwise around  $C$ . Thus by following the progress of a point around  $C$  and reading off the  $x$  and  $y$  values we can draw graphs of  $x$  and  $y$  versus  $t$  (see figure 2). However, since from what has been done so far, we have no indication of the rate at which a solution point moves along  $C$ , the  $t$ -units of such a graph must be completely arbitrary (we cannot even assert equal lengths along the axis correspond to equal increments in  $t$ ).

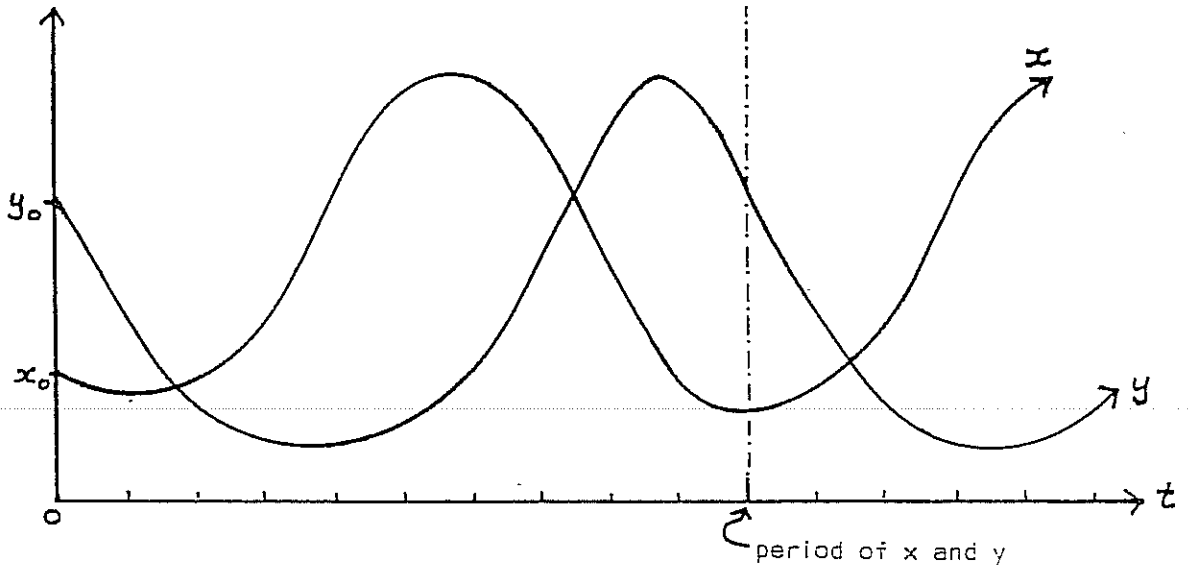


Figure 2

Nevertheless such a graph does give us a reasonable amount of information concerning the solution. Thus from figure 2 we see that for our particular predator-prey model if  $x_0 = \frac{d}{c}$  and  $y_0 = \frac{a}{b}$  the two populations would remain at these equilibrium values for all time. However, if for any reason either of the sizes was slightly different from this value, then the two population sizes would oscillate periodically about these values with  $y$  and  $x$  out of phase. When  $x$  has its maximum (minimum) value,  $y$  has the intermediate value  $\frac{a}{b}$  and similarly when  $y$  has its maximum,  $x$  has the intermediate value  $\frac{c}{d}$ .

From practical considerations this form of the solution is reasonable. When the number of prey increases, more food is available to the predator species and so after a time delay, its size will also increase. This gives rise to an increased predation rate and so reduces the number of prey available which will in turn lead to a reduction in the size of the predator species. But then as a result of reduced predation the prey species can again increase and so the whole cycle repeats itself.

This particular predator-prey model is appropriate for a great variety of situations besides animal versus animal or fish versus fish. One of the other "classical" situations is competition between grass types. For example in Clover/Rye-grass competition, the presence of clover leads to an increase in the available soil nitrogen which in turn enhances the growth of rye-grass. This latter grass type being hardier than clover then 'takes-over' to the detriment of the clover present. Thus supplementation of the soil nitrogen (and so, after some delay, also the rye grass) is reduced, allowing the clover to build up again.

*Collateral Reading:*

A brief account of non-linear system of the type considered here is given in

S.I. Grossman & J. Turner "Mathematics for the Biological Sciences".

A more detailed, though still brief, treatment may be found in

E. Batschelet "Introduction to Mathematics for Life Scientists".

A proof of the existence-uniqueness theorem for such systems is given in

D.A. Sánchez "Ordinary Differential Equations and Stability Theory".

The more detailed mathematical treatment of such systems is also contained in this book as well as in many other books on differential equations, such as Boyce and DiPrima "Elementary Differential Equations and Boundary Value Problems" where a discussion of the phase-plane and the concept of trajectories is also to be found.

EXERCISES:

1. Draw a phase plane for the system of equations

$$\frac{dx}{dt} = -x + xy$$

$$\frac{dy}{dt} = -y + xy$$

and hence discuss the form of its solutions.

2. By plotting "isoclines" corresponding to  $\frac{dy}{dx} = 0, \infty, 1$  and  $-1$  determine the "phase" plane for the non-linear system

$$\frac{dx}{dt} = x(1 - x - y/2)$$

$$\frac{dy}{dt} = y(x - y)$$

(Hint:  $ax^2 + by^2 + hxy + fx + gy + c = 0$  is a hyperbola if  $h^2 - 4ab > 0$ , a parabola if it is 0 and an ellipse if it is negative.)

3. (a) Using much the same reasoning as in the case of a single first order equation, develop an extended Euler-tangent method suitable for solving the first order system

$$\dot{x} = f_1(x, y, t)$$

$$\dot{y} = f_2(x, y, t)$$

- (b) Apply it to the system

$$\dot{x} = (x - y)(1 - x)$$

$$\dot{y} = y(x - y) \quad \text{with } x(0) = 0.1, y(0) = 0.01$$

and  $h = 1$ . Also graph the solution and on the same graph plot  $(y/x \times 100)$  for  $0 \leq t \leq 13$ .

(This system has been suggested as a crude model for the spread of rust ( $y$ ) in sunflowers, where  $x$  is the leaf area at time  $t$ . Can you see why such a model might work?)

- (c) Also analyse the system given in (b) by constructing a phase-portrait for it. Compare the form of the numerical solution obtained in (b) with the conclusions you might draw from consideration of the phase-portrait.

- (d) Find conditions under which the ratio  $y/x$  has a minimum value other than at the initial point. (This should have been the case with the specific initial values suggested in (b).)

4. In the course of the lecture it was shown that for the predator-prey model,  $y$  and  $x$  are related implicitly by

$$a \ln y - by = c(x - x_0) - by_0 - d \ln \frac{x}{x_0} + a \ln y_0 \quad (x_0 \text{ and } y_0 \text{ fixed}).$$

From this show that for each value of  $x > 0$  there corresponds at most two values of  $y$ . Hence conclude that the trajectories must form closed curves, such as those illustrated in figure 1, rather than spirals, a case which was not ruled out by the isoclines drawn in figure 1.

\*Lecture 11 *Stability of almost Linear Systems*

In this lecture the theory of phase-planes for autonomous systems is developed further. In particular we will be interested in the nature of the trajectories in a neighbourhood of a critical point. Since, as already observed a critical point corresponds to an 'equilibrium state' for the system, we can from such a study determine whether the equilibrium is "stable", i.e. whether the system, when slightly disturbed from equilibrium, tends to settle back toward equilibrium or not, and in general gain information about the behaviour of solutions for large values of time (i.e. about their *asymptotic* behaviour).

We begin by surveying the possible behaviour of solutions in the neighbourhood of a critical point for linear systems of the form

$$\frac{dx}{dt} = ax + by$$

$$\frac{dy}{dt} = cx + dy$$

and later extend our findings to certain non-linear systems.

We will further assume that (0,0) is the only critical point of the linear system, i.e.  $ax + by = 0$  and  $cx + dy = 0$  together imply  $x = y = 0$ \*. The general form of the solution for such a linear system was found in lecture 8 where it was seen to depend on the type of roots the characteristic equation,  $\lambda^2 - (a + d)\lambda + ad - bc = 0$ , possessed. In fact the nature of the two roots  $\lambda_1$  and  $\lambda_2$  characterises the behaviour of trajectories.

Thus, if  $\lambda_1 < \lambda_2 < 0$  the solution is

$$x = Ae^{\lambda_1 t} + A'e^{\lambda_2 t}$$

$$y = \frac{\lambda_1 - a}{b} Ae^{\lambda_1 t} + \frac{\lambda_2 - a}{b} A'e^{\lambda_2 t}$$

If we used this to plot a phase-portrait of the system, we would obtain a family of trajectories such as those illustrated in figure 1, in which four rather distinguished trajectories should be noted. The solution satisfying an initial condition of the form

$$x(0) = A, y(0) = \frac{\lambda_1 - a}{b} A$$

is

$$x = Ae^{\lambda_1 t}, \quad y = \frac{\lambda_1 - a}{b} Ae^{\lambda_1 t}$$

and so corresponds to the trajectory

$$y/x = \frac{\lambda_1 - a}{b} \quad (x \text{ between } A \text{ and } 0, y \text{ between } \frac{\lambda_1 - a}{b} \text{ and } 0)$$

a straight line segment with slope  $\frac{\lambda_1 - a}{b}$ . Similar situations arise in the three other cases:

\*

$$\text{equivalently } \det \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \neq 0$$



$$x(0) = A, y(0) = -\frac{\lambda_2 - a}{b} A; \quad x(0) = A', y(0) = \frac{\lambda_1 - a}{b} A';$$

and

$$x(0) = A', y(0) = -\frac{\lambda_2 - a}{b} A' .$$

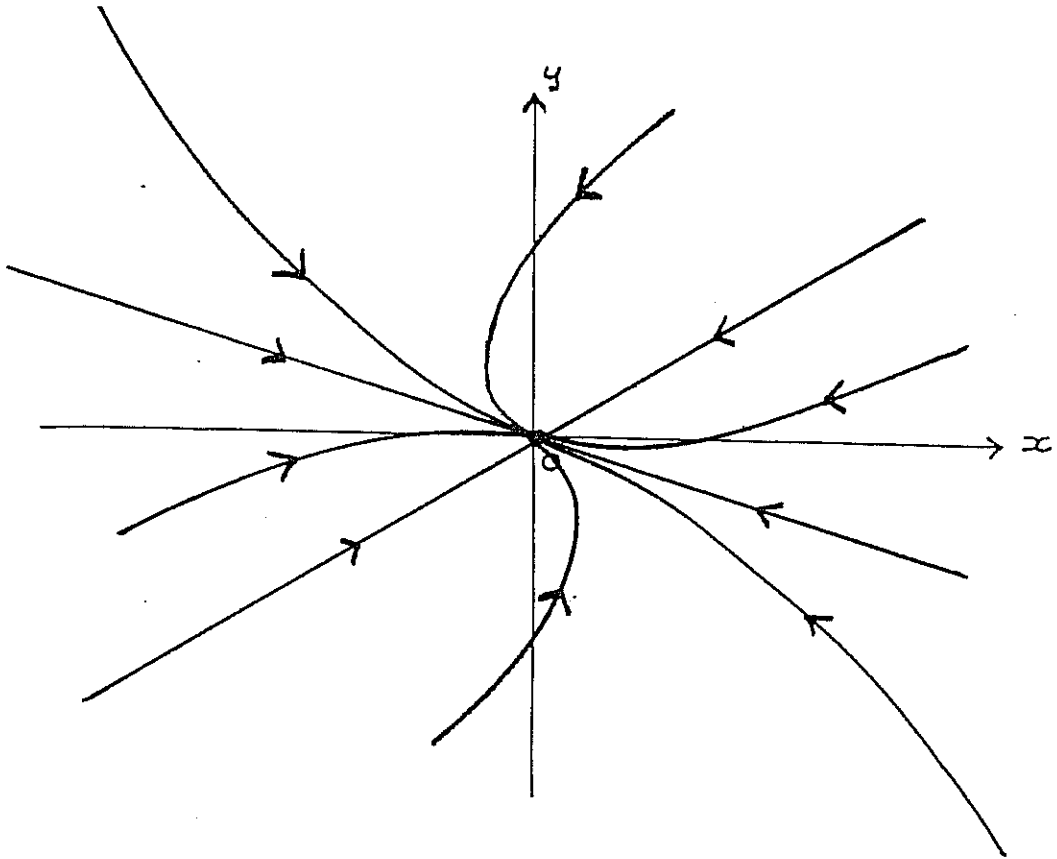


Figure 1

Since, as  $t \rightarrow \infty$ ,  $x$  and  $y \rightarrow 0$  (as  $\lambda_1 < \lambda_2 < 0$ ) we see that the point  $(x,y)$  in the phase-plane corresponding to the solution values  $x = x(t), y = y(t)$  at time  $t$ , will move along its trajectory with increasing values of  $t$  in such a way as to approach  $(0,0)$ . Hence the direction arrows on each of the trajectories shown in figure 1 indicate this fact.

Because of the way trajectories approach the critical point  $(0,0)$ , it is designated 'a node'. Further, since all trajectories, regardless of their initial point, remain arbitrarily close to the critical point  $(0,0)$  for all sufficiently large values of the parameter  $t$ , we say  $(0,0)$  is an *asymptotically stable critical point* of the system and of course corresponds to a 'stable' equilibrium state. It is worth observing that the word asymptotic is appropriate, since the solution does not reach  $(0,0)$  for any finite value of  $t$  unless it was initially  $(0,0)$  in which case it remains so for all  $t$ . This is generally true, for, as remarked

\* Precise mathematical definitions of various types of stability are available. However for us the intuitive idea should suffice.

in lecture 10, distinct trajectories cannot intersect and (0,0) is a trajectory in its own right.

A similar analysis could be carried out for each of the other possible forms of the two roots,  $\lambda_1, \lambda_2$  of the characteristic equation. The details of such an analysis is left to the reader, the results of it are summarised in Table 1, where the nature of the critical point (0,0) and the type of its "stability" is also given along with a typical phase-portrait for each of the cases. It is in terms of this phase-portrait that the various cases of "stability" should be interpreted.

An inspection of Table 1 leads to the following result.

**THEOREM:** *If (0,0) is the only critical point of the linear system*

$$\frac{dx}{dt} = ax + by$$

$$\frac{dy}{dt} = cx + dy,$$

*then it is asymptotically stable if and only if the real parts of the two roots of the characteristic equation  $\lambda^2 - (a + d)\lambda + ad - bc = 0$  are both strictly negative.*

We now see how these results extend to a certain class of non-linear systems where they are of considerably greater value, since explicit solutions for such systems may well be not available.

**DEFINITION:** The autonomous non-linear system

$$\frac{dx}{dt} = ax + by + f(x,y)$$

$$\frac{dy}{dt} = cx + dy + g(x,y)$$

is almost linear if

$$\text{Limit}_{r \rightarrow 0} \frac{f(x,y)}{r} = \text{Limit}_{r \rightarrow 0} \frac{g(x,y)}{r} = 0 \quad \text{where } r = \sqrt{x^2 + y^2}$$

The next result shows that many of the commonly occurring biophysical systems are almost linear.

**LEMMA:** *Any non-linear system of the form*

$$\frac{dx}{dt} = \sum_{m=1}^N \sum_{n=0}^m a_{mn} x^n y^{m-n}$$

$$\frac{dy}{dt} = \sum_{m=1}^M \sum_{n=0}^m b_{mn} x^n y^{m-n},$$

where the  $a_{mn}$  and  $b_{mn}$  are constants, is almost linear, i.e. any system for which the right-hand sides are polynomials in x and y.

**COROLLARY:** *Any quadratic system (c.f. lecture 10) is almost linear.*

TABLE 1

Nature of $\lambda_1, \lambda_2$	Form of Solution	Sketch of Trajectories	Nature of Critical Point	Type of Stability
$\lambda_1 < \lambda_2 < 0$	$\underline{x} = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t}$		(improper) NODE	Asymptotically STABLE
$\lambda_1 > \lambda_2 > 0$	$\underline{x} = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t}$		(improper) NODE	UNSTABLE
$\lambda_1 < 0 < \lambda_2$	$\underline{x} = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t}$		SADDLE POINT	UNSTABLE
$\lambda_1 = \lambda_2 < 0$	$\underline{x} = C_1 e^{\lambda t} + C_2 t e^{\lambda t}$ ( $\lambda < 0$ )		(proper or improper) NODE	Asymptotically STABLE
$\lambda_1 = \lambda_2 > 0$	$\underline{x} = C_1 e^{\lambda t} + C_2 t e^{\lambda t}$ ( $\lambda > 0$ )		(proper or improper) NODE	UNSTABLE
$\lambda_1 = \alpha + i\beta$ ( $\beta \neq 0$ ) $\lambda_2 = \overline{\lambda_1}$ $\alpha < 0$	$\underline{x} = e^{\alpha t} (A \sin \beta t + B \cos \beta t)$ ( $\alpha < 0$ )		SPIRAL POINT	Asymptotically STABLE
$\alpha > 0$	$\underline{x} = e^{\alpha t} (A \sin \beta t + B \cos \beta t)$ ( $\alpha > 0$ )		SPIRAL POINT	UNSTABLE
$\alpha = 0$	$\underline{x} = A \sin \beta t + B \cos \beta t$		CENTRE (periodic solutions)	"STABLE"

Proof. We will prove the lemma for the case when the right hand sides are quadratic i.e. when  $N = M = 2$ . The proof for right hand sides of higher degree should then be obvious.

Hence, consider the system

$$\frac{dx}{dt} = a_{11}x + a_{10}y + a_{22}x^2 + a_{20}y^2 + a_{21}xy$$

$$\frac{dy}{dt} = b_{11}x + b_{10}y + b_{22}x^2 + b_{20}y^2 + b_{21}xy$$

and let  $r = \sqrt{x^2 + y^2}$ , then  $x$  and  $y$  may be written as  $x = r \cos \theta$ ,  $y = r \sin \theta$  where  $\theta$  is chosen so that

$$\sin \theta = y/\sqrt{x^2 + y^2} \text{ and } \cos \theta = x/\sqrt{x^2 + y^2}$$

Substituting these for  $x$  and  $y$  gives

$$f(x,y) = a_{22}x^2 + a_{20}y^2 + a_{21}xy = r^2(a_{22}\cos^2\theta + a_{20}\sin^2\theta + a_{21}\sin\theta\cos\theta)$$

$$\text{and so } \lim_{r \rightarrow 0} \frac{f(x,y)}{r} = \lim_{r \rightarrow 0} r(a_{22}\cos^2\theta + a_{20}\sin^2\theta + a_{21}\sin\theta\cos\theta) = 0$$

similarly for  $g(x,y)$ .

The importance of almost-linear systems is due to the following theorem, a proof of which (although not difficult) will not be attempted.

**THEOREM:** *Let  $(0,0)$  be a critical point of the almost-linear system*

$$\frac{dx}{dt} = ax + by + f(x,y)$$

$$\frac{dy}{dt} = cx + dy + g(x,y)$$

*and the only critical point of the corresponding linear system*

$$\frac{dx}{dt} = ax + by$$

$$\frac{dy}{dt} = cx + dy,$$

*then:*

*if  $(0,0)$  is an asymptotically stable critical point of the linear system, it is also an asymptotically stable critical point of the almost-linear system;*

*if  $(0,0)$  is an unstable critical point of the linear system it is also an unstable critical point of the almost-linear system.*

*Further, if  $\lambda_1, \lambda_2$  are the two roots of the linear system's characteristic equation, then*

*$\lambda_1 \neq \lambda_2$  but both real, implies  $(0,0)$  is an improper node for the linear system and also for the almost-linear system;*

*$\lambda_1 < 0 < \lambda_2$  implies  $(0,0)$  is a saddle point for both the linear and almost-linear system;*

*$\lambda_1 = \alpha + i\beta$  with  $\alpha, \beta$  real and  $\alpha\beta \neq 0$  implies  $(0,0)$  is a spiral point of the linear system and also of the almost-linear system;*

*while, if  $\lambda_1 = \lambda_2, (0,0)$  is a node of the linear system and either a node or a spiral point of the almost linear system.*

Roughly speaking this somewhat lengthy theorem states that except when (0,0) is a centre for the linear system (a case about which almost nothing can be said in general) the trajectories near (0,0) for the almost-linear system are a facsimile to those of the linear system.

A non-linear system can of course have critical points other than (0,0), say (x<sub>0</sub>, y<sub>0</sub>). To analyse the behaviour near (x<sub>0</sub>, y<sub>0</sub>) we simply change origin to this point, through the transformation X = x - x<sub>0</sub> Y = y - y<sub>0</sub>, and then if appropriate apply the above theorem to the system rewritten in terms of X and Y.

ILLUSTRATIVE EXAMPLE. Consider the system

$$\frac{dx}{dt} = x(1 - x - \frac{y}{2})$$

$$\frac{dy}{dt} = y(x - y)$$

which has a critical point when  $\frac{dx}{dt} = \frac{dy}{dt} = 0$  i.e. when  $x(1 - x - \frac{y}{2}) = 0$  and  $y(x - y) = 0$ . Thus the critical points are (0,0),  $(\frac{2}{3}, \frac{2}{3})$  and (1,0).

We will examine the behaviour of the system near the critical point  $(\frac{2}{3}, \frac{2}{3})$ . To investigate the stability at this point we transfer the origin to it, viz. let  $X = x - \frac{2}{3}$ ,  $Y = y - \frac{2}{3}$  then the system becomes

$$\frac{dX}{dt} = (X + \frac{2}{3})(1 - X - \frac{2}{3} - \frac{Y}{2} - \frac{1}{3}) = -\frac{2}{3}(X - \frac{Y}{2}) - (X^2 + \frac{XY}{2})$$

$$\frac{dY}{dt} = (Y + \frac{2}{3})(X - Y) = \frac{2}{3}(X - Y) + (XY - Y^2).$$

First note the system is quadratic and so almost linear.

Then the linearized system is

$$\frac{dX}{dt} = -\frac{2}{3}X + \frac{1}{3}Y$$

$$\frac{dY}{dt} = \frac{2}{3}X - \frac{2}{3}Y$$

for which the characteristic equation is

$$\lambda^2 + \frac{4}{3}\lambda + \frac{2}{9} = 0$$

the roots of which are

$$\lambda_1 = \frac{-2 + \sqrt{2}}{3} \quad \lambda_2 = \frac{-2 - \sqrt{2}}{3}$$

both of which are real and negative. So (0,0) is an asymptotically stable node of this system. Whence  $(\frac{2}{3}, \frac{2}{3})$  is an asymptotically stable node for the original system. Compare this with the phase-plane you were asked to prepare of this system in Exercise 10.2.

Apart from the purely practical importance in determining the behaviour of non-linear systems, the theory developed here provides a sort of "meta-mathematical" justification for the suitability of differential equations in the modelling of physical situations. By assuring us that, in the appropriate cases, small changes in the initial conditions will produce only small (often decreasing) differences in the solution, the necessary uncertainty in the precis

initial values (due to inherent experimental errors etc.) does not present a major difficulty. Without the assurance however, the association of a differential equation's solution with the physical development of a system would be risky indeed. Of equal importance in this vein is the theorem (see Sánchez) which states that, under mild restrictions, the solution to a given system of equations varies continuously with respect to the parameter values of the system. So again a small amount of indeterminance in the parameter values (which are often determined experimentally) is not serious.

*Collateral Reading:*

Usually the mathematical discussion of systems is carried out in the language of matrix systems (see appendix to lecture 8). This is the case in the works cited below.

Along with many other valuable results a proof of the basic theorem on almost-linear systems may be found in D.A. Sánchez "Ordinary differential equations and stability theory", Freeman, San Francisco, 1968.

A widely applicable (though frequently very difficult actually to apply) method of determining the stability of non-linear systems, based on generalizations of the concept of 'energy', was developed by A.M. Liapounov. His method is described in

Boyce and DiPrima "Elementary Differential Equations and Boundary Value Problems", where the results described in this lecture are derived as a special case of Liapounov's method. A more extensive treatment is given in

H. Hochstadt "Differential Equations, a Modern Approach".

EXERCISES:

1. Examine the other two critical points of the example considered in the lecture. From this attempt to construct a phase-portrait for the system, compare it with your answer to exercise 10.2.
2. Consider the non-homogeneous linear system

$$\frac{dx}{dt} = ax + by + f(t)$$

$$\frac{dy}{dt} = cx + dy + g(t).$$

Let  $x_1 = x_1(t)$ ,  $y_1 = y_1(t)$  and  $x_2 = x_2(t)$ ,  $y_2 = y_2(t)$  be two solutions corresponding to distinct initial conditions.

Show that  $x_1 - x_2$  and  $y_1 - y_2$  both tend to zero as  $t$  goes to infinity if  $(0,0)$  is an asymptotically stable critical point of the linear system

$$\frac{dx}{dt} = ax + by$$

$$\frac{dy}{dt} = cx + dy$$

3. Determine the nature of the critical point  $\left(\frac{d}{c}, \frac{a}{b}\right)$  of the Lotka-Volterra system examined in lecture 10.

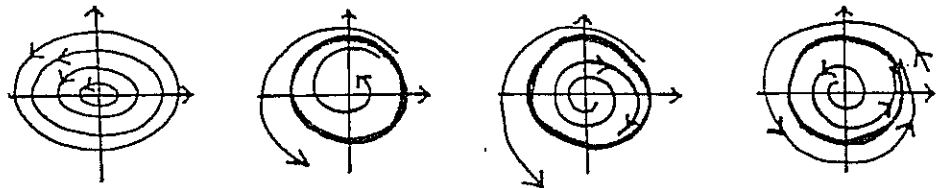
\*Lecture 12      *Periodic Solutions*

In lecture 8 we saw that linear systems could have periodic solutions (the case when the two roots of the characteristic equation are purely imaginary while in the last lecture we noted that this corresponded to  $(0,0)$  being a centre and the trajectories closed curves encircling it. Such a closed curve trajectory in the phase-plane of a system is termed a *cycle*. As may be easily seen the presence of a cycle is equivalent to the existence of a periodic solution for an autonomous system.

In lecture 10 we saw that the non-linear Lotka-Volterra system admitted cycles centred around the critical point  $(\frac{d}{c}, \frac{a}{b})$  and there commented on the corresponding periodic solutions. Our aim is to determine when a system has periodic solutions and how the corresponding cycles relate to other trajectories of the system. Only partial answers are known to these questions, the proofs of which require such a diversity of techniques that any attempt to give them would lead us too far afield. In consequence only a few of the more easily stated and useful results will be given but not proved. The first of these concerns the possible location of cycles.

(1) *A cycle must surround at least one critical point of the system.* (If it surrounds just one, and the system is almost linear at that critical point, then it cannot be a saddle-point - much more can be said about the possible configurations of critical points within a cycle, particularly through the use of a powerful concept known as the Poincaré index; see Hochstadt for a discussion of same.)

Just as with critical points, cycles may be classified according to the behaviour of nearby trajectories. Without entering into elaborate and precise definitions, this is best done through a diagram; accordingly the reader's attention is drawn to figure 1.



neutrally  
stable cycle  
(small changes  
remain small  
but do not  
decay to zero)

unstable  
cycle

semi-stable  
cycle  
(stable from  
within in  
the case  
illustrated)

stable limit  
cycle  
(orbital  
stability)

Figure 1

We now complete this lecture by presenting, with attendant explanations and examples, three general theorems regarding the possible existence of cycles (and hence periodic solutions).

(2) (Bendixson) Let 
$$\frac{dx}{dt} = F(x,y)$$

$$\frac{dy}{dt} = G(x,y)$$

be a non-linear system with the partial derivatives  $\frac{\partial F}{\partial x}$ ,  $\frac{\partial F}{\partial y}$ ,  $\frac{\partial G}{\partial x}$  and  $\frac{\partial G}{\partial y}$  defined and continuous in some simply connected region  $D$  (i.e.  $D$  is a region of the plane with no 'holes'), then if  $\frac{\partial F}{\partial x} + \frac{\partial G}{\partial y}$  has the same sign throughout  $D$  there is no cycle lying entirely in  $D$ .

EXAMPLE:- The Lotka-Volterra system

$$\frac{dx}{dt} = ax - bxy$$

$$\frac{dy}{dt} = cxy - dy$$

has  $\frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} = (a - d) + cx - by$  which has a constant sign on each side of the line

$$y = \frac{c}{b}x + \left(\frac{a}{b} - \frac{d}{b}\right)$$

so any cycle must cross this line. (Note  $(\frac{d}{c}, \frac{a}{b})$  lies on this line as expected.)

More general, though considerably more difficult to apply is the following:

(3) (Poincaré-Bendixson) Let 
$$\frac{dx}{dt} = F(x,y)$$

$$\frac{dy}{dt} = G(x,y)$$

be as in (2) i.e.  $F$  and  $G$  have defined and continuous partial derivatives throughout a simply connected region  $D$ . Let  $\mathcal{R}$  be a closed (i.e. includes its 'boundary') subregion of  $D$  which is not necessarily simply connected (i.e. may have 'holes') but which contains no critical points of the system, then, if there exists a trajectory starting in  $\mathcal{R}$  which stays within  $\mathcal{R}$  for all  $t$ ,  $\mathcal{R}$  contains a cycle.

The difficulty with this theorem is (a) to locate a suitable region  $\mathcal{R}$  and (b) to prove the existence of a suitable trajectory.

The following example is a standard one for illustrating this theorem.

EXAMPLE:-

$$\frac{dx}{dt} = x + y - x(x^2 + y^2)$$

$$\frac{dy}{dt} = y - x - y(x^2 + y^2)$$

Change variables to  $x = r \cos \theta$ ,  $y = r \sin \theta$  (where  $r = \sqrt{x^2 + y^2}$  and  $\theta$  is suitably chosen) to obtain

$$\frac{dr}{dt} = r(1 - r^2)$$

$$\frac{d\theta}{dt} = -1$$

---

\* If you are unfamiliar with the idea of partial differentiation, see



(This system can be integrated and  $r = 1$ , i.e. the unit circle  $x^2 + y^2 = 1$ , shown to be a cycle - however we will establish the existence of such a cycle through the use of (3).)

Now, take  $\mathcal{R} = \{(r, \theta) : \frac{1}{2} \leq r \leq 2\}$  it is easily seen that this closed annular region contains no critical points and is suitable for an application of (3). We will take for granted that the solution starting at  $\theta = 0, r = \frac{1}{2}$  exists, hence we need to show that it remains within  $\mathcal{R}$ . Assume it did not, then for some value of  $t$  either

$$r(t) = \frac{1}{2} \text{ and } r \text{ is decreasing } \underline{\text{i.e.}} \frac{dr}{dt} < 0 \text{ but at } r = \frac{1}{2}.$$

$$\frac{dr}{dt} = \frac{1}{2} \left(1 - \frac{1}{4}\right) > 0 \text{ a contradiction,}$$

or

$$r(t) = 2 \text{ and } r \text{ is increasing } \underline{\text{i.e.}} \frac{dr}{dt} > 0, \text{ but at } r = 2$$

$$\frac{dr}{dt} = 2(1 - 4) < 0 \text{ again a contradiction and so we conclude}$$

that the trajectory of this solution cannot leave  $\mathcal{R}$ . So by (3)  $\mathcal{R}$  contains a cycle and the system has periodic solutions.

The rather recent origin of the next, and last result, indicates the difficulty of research in this area.

(4) (Levinson-Smith, 1942) *If Liénard's System*

$$\frac{dx}{dt} = y$$

$$\frac{dy}{dt} = -g(x) - f(x)y$$

is such that:

(a)  $f$  is an even function and  $g$  is an odd function

(i.e.  $f(-x) = f(x)$  and  $g(-x) = -g(x)$  for all  $x$ );

(b)  $G(x) = \int_0^x g + \infty$  as  $x \rightarrow \infty$ ;

(c)  $F(x) = \int_0^x f + \infty$  as  $x \rightarrow \infty$ ;

(d) there exists  $x_0 > 0$  such that  $F(x) < 0$  for  $0 < x < x_0$  and  $F(x)$  is positive and monotonically increasing for  $x \geq x_0$ ,

then it has a (unique) cycle, which is a stable limit cycle.

EXAMPLE:- The Van der Pol system is

$$\frac{dx}{dt} = y$$

$$\frac{dy}{dt} = -x - \mu(x^2 - 1)y \text{ where } \mu > 0 \text{ is a constant,}$$

and for which  $g(x) = x$  is odd and  $G(x) = \frac{x^2}{2} + \infty$  as  $x \rightarrow \infty$ , while  $f(x) = \mu(x^2 - 1)$  is even,  $F(x) = \mu\left(\frac{x^3}{3} - x\right) + \infty$  as  $x \rightarrow \infty$ , for  $0 < x < \sqrt{3}(=x_0)$ ,  $F(x) < 0$ , and for  $x > \sqrt{3}$ ,  $F(x) \geq 0$ , with  $F'(x) = f(x) > 0$ .

So the conditions of (4) are satisfied and we conclude the system has a unique stable limit cycle and hence a unique periodic solution.

The phase-plane of Van der Pol's system with  $\mu = 5$  is illustrated in figure 2. (Note for smaller values of  $\mu$  the system is more nearly  $\frac{dx}{dt} = y$ ,  $\frac{dy}{dt} = -x$  and the limit cycle more nearly circular. Why might you expect this?)

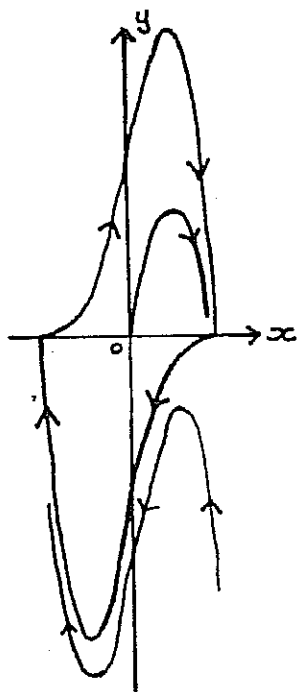


Figure 2

*Collateral Reading:*

Much of the material presented here is discussed in Boyce and diPrime "Elementary Differential Equations and Boundary Value Problems".

A more complete discussion is to be found in H. Hochstadt "Differential Equations, a Modern Approach".

An excellent review of the known results for quadratic systems is given in W.A. Coppel "A Survey of Quadratic Systems", J.D. Equations 2 (1966) pp. 293-304.

An interesting application of results similar to those considered here to biological problems may be found in J. Schnute & P. Van den Driessche, "Two Biological Applications of Dulac's Criterion", Applies Mathematics NOTES, Vol. 1, No. 2, Sept., '75, pp. 75-81.

EXERCISES:

1. Show that a sufficient condition for the linear autonomous system

$$\frac{dx}{dt} = ax + by$$

$$\frac{dy}{dt} = cx + dy$$

not to have a periodic solution other than  $x = 0$ ,  $y = 0$  is that  $a + d \neq 0$ .

2. Can the system

$$\frac{dx}{dt} = x + y + x^3 - y^2$$

$$\frac{dy}{dt} = -x + 2y + yx^2 + \frac{1}{3}y^3$$

have a periodic solution other than  $x = y = 0$ ?

3. Determine the location and nature of the critical points of Van der Pol's system.

4. Show the system

$$\frac{dx}{dt} = y$$

$$\frac{dy}{dt} = -x + y - x^5 - 2x^2y$$

has a periodic solution.

SERIES III - Single Second Order Ordinary Differential Equations.

Lecture 13 *The general equation and solution of the homogeneous case with constant coefficients.*

We will be concerned with equations of the form

$$\frac{d^2y}{dx^2} + a \frac{dy}{dx} + by = f(x).$$

The general form of such a linear second order non-homogeneous differential equation has both a and b functions of x, however we will initially restrict our attention to the case when a and b are constants. The general theory of such equations requires that two initial conditions say  $y(x_0) = y_0$  and  $y'(x_0) = y'_0$ , be specified to determine a solution uniquely. We will assume, at least implicitly, that this has been done.

To see how such an equation might in practice arise, let us consider the following ecological problem.

A lake system is supplied, via influent streams etc., with nutrient.

Let  $f(t)$  be the rate at which this nutrient is accumulating within the lake at any given moment  $t$ . Further let the lake support an aquatic population of size  $N(t)$  at time  $t$ . If we now assume this population expends nutrient at a rate proportional to its size and that the rate of change of the population is proportional to the amount of excess nutrient present (see section 3.1), then we have

$$\frac{dN}{dt} = \left\{ \begin{array}{l} \text{amount of excess} \\ \text{nutrient} \end{array} \right\} = \left\{ \begin{array}{l} \text{amount of nutrient} \\ \text{initially present, } f_0 \end{array} \right\} + \left\{ \begin{array}{l} \text{amount of nutrient which has} \\ \text{accumulated over a period of} \\ \text{duration } t \end{array} \right\} - \left\{ \begin{array}{l} \text{amount expended by the population} \\ \text{during the period of length } t \end{array} \right\}$$

$$\text{while } \left\{ \begin{array}{l} \text{amount of nutrient which has accumulated} \\ \text{over a period of duration } t \end{array} \right\} = \int_0^t f(\tau) d\tau$$

$$\text{and } \left\{ \begin{array}{l} \text{amount expended by the population} \\ \text{during the period of length } t \end{array} \right\} = \int_0^t N(\tau) d\tau.$$

$$\text{so } \frac{dN}{dt} = kF_0 + k \int_0^t f(\tau) d\tau - b \int_0^t N(\tau) d\tau$$

where  $k$  and  $b$  are the appropriate proportionality constants (both positive). This integro-differential equation is, upon differentiating throughout with respect to  $t$ , converted to the second order equation

$$\frac{d^2N}{dt^2} + bN = kF(t)$$

which is of the form

introduced above. The solution of such an equation will be taken up in some detail in subsequent lectures. For the present however, we will simplify the problem by postulating a constant rate of accumulation of nutrient within the lake, and see where this leads us.

Thus the differential equation now becomes

$$\frac{d^2N}{dt^2} + bN = r$$

where  $r$  is the appropriate rate constant.

This may be further simplified under the change of variable  $y = N - \frac{r}{b}$  to

$$\frac{d^2y}{dt^2} + by = 0.$$

We are therefore led to seek solutions of the second order linear homogeneous equation

$$\frac{d^2y}{dt^2} + a \frac{dy}{dt} + by = 0.$$

If for no better reason than it having worked in the past, we will try for solutions of the form

$$y(t) = Ae^{\lambda t}.$$

Substituting this into the equation we see that our requirement is that

$$(\lambda^2 + a\lambda + b)Ae^{\lambda t} = 0 \quad \text{for all } t.$$

A requirement which is met provided we choose  $\lambda$  to be a root of the *auxiliary equation*

$$\lambda^2 + a\lambda + b = 0, \text{ i.e. provided}$$

$$\lambda = \frac{-a \pm \sqrt{a^2 - 4b}}{2}.$$

Of course this leads to three possible cases.

*Case 1.* The two roots  $\lambda_1$  and  $\lambda_2$  are real and unequal.

In this case two distinct solutions have been furnished

$$y = Ae^{\lambda_1 t} \text{ and } y = A'e^{\lambda_2 t}$$

where the constants  $A, A'$  (not necessarily equal) are as yet arbitrary. Since the equation is linear and homogeneous it is readily checked that the sum of these two solutions is itself a solution. It is in fact the most general solution, viz

$$y = Ae^{\lambda_1 t} + A'e^{\lambda_2 t}$$

The values of  $A$  and  $A'$  can now be chosen to satisfy the initial conditions. Thus we may require that they be solutions of the two simultaneous equations

$$(e^{\lambda_1 t_0})A + (e^{\lambda_2 t_0})A' = y_0$$

$$(\lambda_1 e^{\lambda_1 t_0})A + (\lambda_2 e^{\lambda_2 t_0})A' = y'_0.$$

EXAMPLE.  $y'' + 2y' - 3y = 0$  has auxiliary equation  $\lambda^2 + 2\lambda - 3 = 0$  and so  $\lambda = -3$  or  $1$  and the general solution is

$$y(t) = Ae^t + A'e^{-3t}.$$

If  $y(0) = 1$  and  $y'(0) = 0$  we would then require  $A + A' = 1$

$$\text{and } A - 3A' = 0$$

So

$$y = \frac{3}{4}e^t + \frac{1}{4}e^{-3t}.$$

This curve is plotted in figure 1 along with several other distinctive forms for the solution curve where  $\lambda_1$  and  $\lambda_2$  are real and unequal.

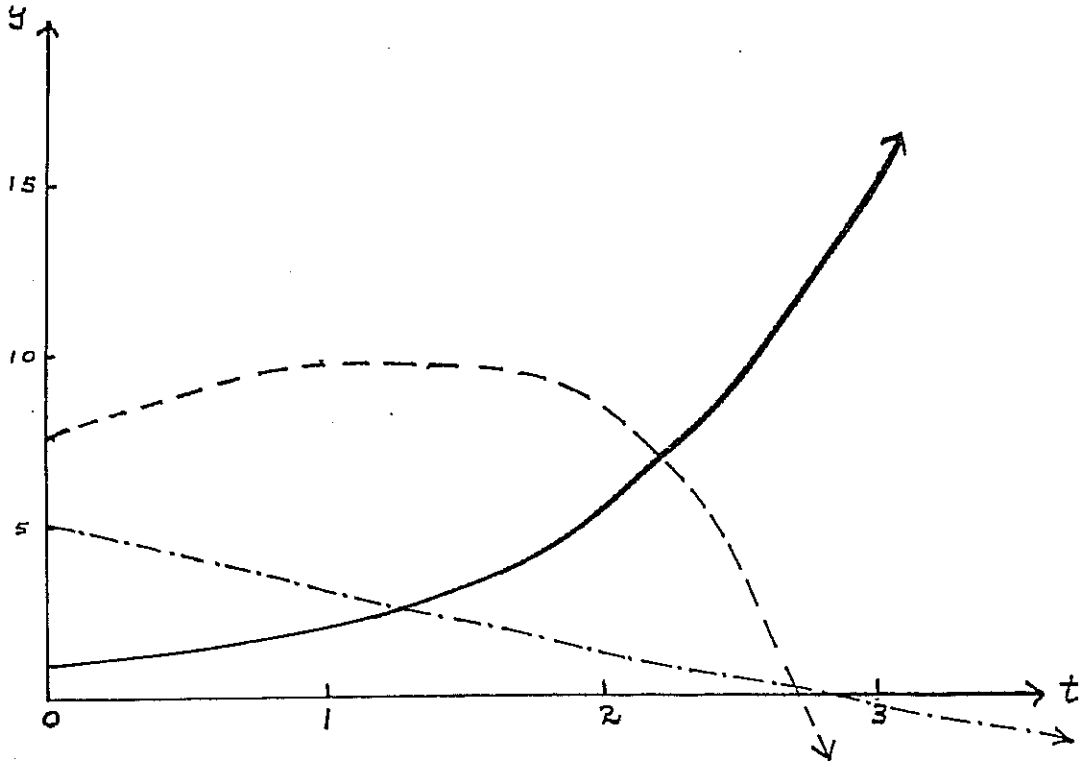


Figure 1

Case 2. The roots are complex conjugates,  $\alpha + i\beta$  and  $\alpha - i\beta$   
( $\alpha, \beta$  real and  $\beta \neq 0$ )

By reasoning similar to that above we find that the general form of the solution is

$$y(t) = e^{\alpha t} (Ae^{i\beta t} + A'e^{-i\beta t}).$$

Making use of the Euler formula and taking the two constants  $A, A'$  to be complex conjugates we arrive at the most general form of real solution

$$y(t) = e^{\alpha t} (A \cos \beta t + B \sin \beta t)$$

where  $A$  and  $B$  are arbitrary constants determined through the initial values. (That this is indeed a solution could profitably be checked by direct substitution into the equation.)

Our ecological problem produced the equation

$$\frac{d^2y}{dt^2} + by = 0$$

for which the auxiliary equation is

$$\lambda^2 + b = 0 \text{ or } \lambda = \pm i\sqrt{b} \text{ (recall } b > 0)$$

and so is an example of this case, having the general solution

$$N(t) = \frac{F}{b} + y(t) = \frac{F}{b} + A \cos \sqrt{b} t + B \sin \sqrt{b} t$$

It is clear from this that  $N(t) = \frac{r}{b}$  is a constant solution for the problem, all other solutions consisting of bounded, but undamped, periodic oscillations about this 'equilibrium' solution. A solution for which  $N(0) = \frac{r}{b}$  and  $\dot{N}(0) < 0$  is illustrated in figure 2(a). Note that if we had  $|\dot{N}(0)| \geq rb^{-\frac{1}{2}}$  extinction of the population would occur after a finite time had elapsed.

Figure 2(b) illustrates other possible forms of solution curves for Case 2.

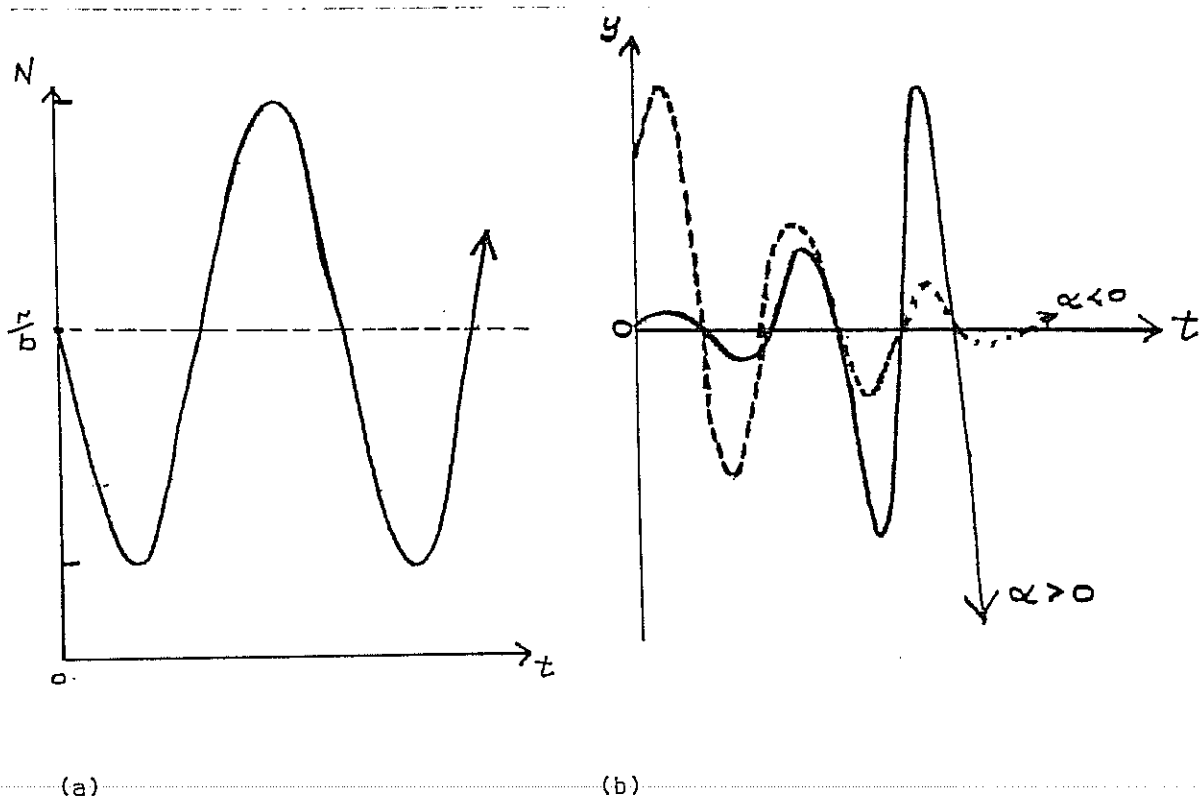


Figure 2

Case 3. The two roots are equal (and hence of necessity a real number  $\lambda$ , say). Here our trial solution yields only  $y = Ae^{\lambda t}$ , a solution containing only one arbitrary constant A and so unable in general to accommodate a pair of initial conditions. We might therefore be led to expect the existence of a second solution of a form other than exponential. This is indeed the situation, the general solution in this case is  $y(t) = (A + Bt)e^{\lambda t}$ , A and B arbitrary constants. Possible forms for such a solution are illustrated in figure 3.

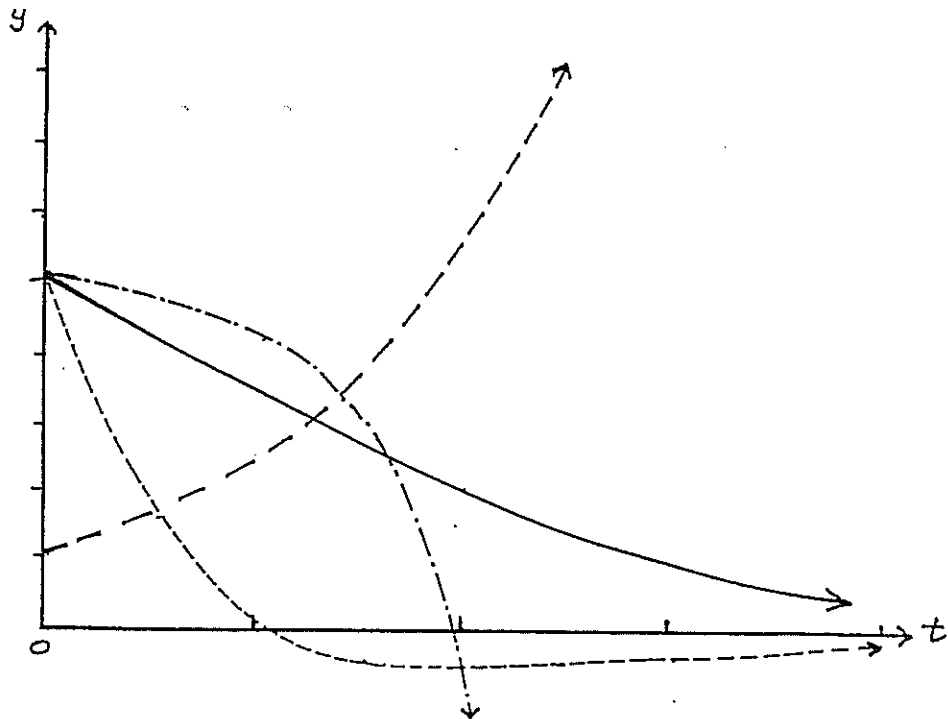


Figure 3.

EXAMPLE.  $y'' + 2y' + y = 0$  has  $\lambda = -1$  as the repeated root of its auxiliary equation, and so the general solution  $y(t) = (A + Bt)e^{-t}$ .

So if  $y(0) = 1, y'(0) = 0$  we have  
 $y(t) = (1 + t)e^{-t}$ .

As we will soon see, an ability to solve the homogeneous form of a second order linear equation is of use in solving the more complicated non-homogeneous case

$$\frac{d^2y}{dt^2} + a \frac{dy}{dt} + by = f(t).$$

Where, for reasons which should be apparent from the ecological example considered, the function  $f(t)$  is often referred to as the *input*, the solution  $y(t)$  then being regarded as the corresponding *output*.

More general methods for solving the homogeneous case are known, and also cover the case when  $a$  and  $b$  are functions of  $t$ . Two of these worthy of mention are 'reduction of order' whereby a second solution may be found if one is known (cf. Case 3, above) and 'the method of series solutions' (including that of Frobenius). Unfortunately any discussion of these would lead us too far away from our basic goal, the role of differential equations in biophysical theory.

*Collateral Reading.*

Second order linear equations are treated in most books containing a section on differential equations.

A good reference for much of the general theory is

Boyce and diPrima "Elementary Differential Equations and Boundary Problems", although proofs are often omitted.

J. Maynard-Smith in "Mathematical Ideas in Biology" gives a good account of such equations from a mechanics point of view. Applying them to the problem of limb motion and "deducing" that some form of neural motor reaction, and not just damping internal to the limb itself, is required to explain the rapid arresting of motion which is seen to be possible.

EXERCISES:

1. Solve the following second order D.E.'s:

$$(i) \quad \frac{d^2y}{dx^2} + 5 \frac{dy}{dx} + 6y = 0 \quad y(0) = 0, \frac{dy}{dx}(0) = 1$$

$$(ii) \quad \frac{d^2y}{dx^2} + 9y = 0 \quad y(0) = 2, \frac{dy}{dx}(0) = 0$$

$$(iii) \quad \frac{d^2y}{dx^2} - 2 \frac{dy}{dx} + 2y = 0 \quad y(0) = 0, \frac{dy}{dx}(0) = 1$$

$$(iv) \quad \frac{d^2y}{dx^2} + 8 \frac{dy}{dx} + 16y = 0 \quad y(0) = 2, \frac{dy}{dx}(0) = 0$$

and sketch graphs of their solutions.

2. In the lake ecology model considered in the lecture show that if, due to fishing, the lake is depopulated at a rate proportional to the population size, a term of the form  $ay'$  ( $a > 0$ ) must be included in the left-hand side of the equation. In the case of a constant rate of nutrient accumulation, what effect does this additional term have on the solution?

3. Show that the equation  $\frac{d^2y}{dt^2} + b \frac{dy}{dt} + cy = 0$  can be written as the system of first order equations

$$\frac{dy}{dt} = 0y + 1x \quad (\text{i.e. } = x)$$

$$\frac{dx}{dt} = -cy - bx$$

Thus second order equations can be regarded as special cases of systems of equations. This explains the obvious similarity in the form of their solutions, and also the statement that such equations require two initial conditions if the solution is to be uniquely determined.

(Can every linear system of two equations be written as a second-order linear equation?)

4. Referring back to that part of the soil nitrogen system described in problem 5.4, since the population of nitrifiens  $P$  is dependent only on the presence of  $NH_4 - N$  both for its supply of nutrient and energy and so has this as its only growth limiting factor, it is plausible to take

$$\frac{dP}{dt} = b(N - cP), \text{ instead of } \frac{dP}{dt} = bNP \text{ as in problem 5.4.}$$



4. (continued)

Since this ensures there is always sufficient N present to accommodate the needs of P it is also reasonable to replace the equation for N,  $\frac{dN}{dt} = -aNP$ , used in 5.4, by the more simple one  $\frac{dN}{dt} = -aP$ . We are thus led to an alternative model for the system, namely

$$\left. \begin{aligned} \frac{dN}{dt} &= -aP \\ \frac{dP}{dt} &= b(N - cP) \end{aligned} \right\}$$

Show that these equations may be combined to give the single second order equation for N,

$$\frac{d^2N}{dt^2} + bc \frac{dN}{dt} + ab N = 0.$$

By solving this, show that two possible cases can arise:

- (i) if  $bc^2 > 4a$  then both P and N can gradually reduce for an infinite period of time without either becoming zero;
- (ii) if  $bc^2 < 4a$  then of necessity both N and P must be exhausted after a finite time. (Find an expression for this time.)

Experimental results suggest values of  $a = 2.2$ ,  $b = 0.03$  and  $c = 2.3$ . Into which of the above two cases does this place the system? It is also known that  $P(0) = 1$  (positive tubes) and  $N(0) = 100$  (ppm). Using these values obtain a solution to the model, and illustrate the solution by plotting graphs of both N and P as functions of time.

5. For what values of  $\lambda$  will  $\frac{d^2y}{dx^2} + \lambda y = 0$  have a solution y such that  $y(0) = y(1) = 0$ ?

6. In the lake ecology problem of lectures, assume that the rate of nutrient accumulation is

$$f(t) = \alpha + \beta \sin 2\pi t \quad (\alpha, \beta > 0 \text{ are constants}),$$

thus allowing for possible seasonal variations (t in years). Show that under an appropriate change of variables this leads to the equation

$$\frac{d^2y}{dt^2} + by = k\beta \sin 2\pi t$$

If  $b = 1$  and  $k\beta = \frac{1}{2}$  show this has a solution of the form

$$y_p = C \sin 2\pi t + D \cos 2\pi t$$

where the parameters C and D may be determined by substitution into the differential equation. (That this is a reasonable guess to try follows since the only functions which differentiate to give trigonometric functions are trigonometric functions and the equation states that a multiple of y plus its second derivative equals a trigonometric function.) This method of solving non-homogeneous equations, which is tantamount to making an intelligent guess, is known as 'the method of undetermined coefficients'.

6. (continued)

Show that a more general solution, able to allow for varying initial conditions is

$$y = y_p + h$$

where  $h$  is any solution of the corresponding homogeneous equation

$$\frac{d^2y}{dt^2} + by = 0.$$

7. If  $y \equiv y_1(x)$  and  $y \equiv y_2(x)$  are two solutions of the linear homogeneous differential equation,  $y'' + a(x)y' + b(x)y = 0$ , show that  $y \equiv Ay_1(x) + By_2(x)$  is also a solution where  $A, B$  are constants.

Lecture 14      *The non-homogeneous equation*

Two approaches to the solution of the non-homogeneous equation

$$\frac{d^2y}{dx^2} + a(x) \frac{dy}{dx} + b(x)y = f(x),$$

with appropriate initial conditions, will be considered:

- (1) The use of Laplace Transforms, particularly when a and b are constants, is indicative of the numerous and powerful Transform methods available to the applied mathematician. This approach forms the material of the lecture;
- (2) An analytic approach, known as the method of Variation of Parameters, is detailed in an appendix to the lecture. This method provides an integral representation of the solution in terms of the general solution of the corresponding homogeneous equation

$$\frac{d^2y}{dx^2} + a(x) \frac{dy}{dx} + b(x)y = 0.$$

14.1 *The Laplace Transform.*

Since we are usually interested in the solution of the above second order equation only for  $x \geq 0$ , we will restrict our attention to such  $x$ .

Thus, let  $f$  be a function defined for  $\{x: 0 \leq x < \infty\}$  and such that the integral

$$\int_0^{\infty} e^{-sx} f(x) dx$$

exists, at least for all values of  $s$  greater than some  $s_0$ ,\* we will simply refer to any such  $f$ , extended to be 0 for  $x < 0$ , as a *suitable* function.

From any suitable function  $f$  we can form a new function  $F$  defined by

$$F(s) = \int_0^{\infty} e^{-sx} f(x) dx$$

known as the *Laplace Transform* of  $f$ .

Strictly speaking, we have defined a transformation, the Laplace transformation  $L$ , which associates with any suitable function  $f$  its Laplace Transform  $F$ , i.e.  $L(f) = F$ . It is important to note that both  $f$  and  $F$  are functions.  $f$ , by convention, a function of the variable  $x$  and  $F$  a function of the independent variable  $s$ .

\* A sufficient, though by no means necessary, condition for this to be so is that  $f$  has at most a finite number of discontinuities in any finite interval, i.e. be *piece-wise continuous*, and be of at most *exponential order*, i.e. there exists constants  $M$  and  $\lambda$  such that  $|f(t)| \leq Me^{\lambda t}$  for all  $t \geq t_0$ . For then the improper integral

$$\int_0^{\infty} e^{-sx} f(x) dx$$

is convergent for all  $s > \lambda$ .

EXAMPLE: If  $f(x) = e^{kx}$  ( $k$  a constant) we consider

$$\int_0^{\infty} e^{-sx} e^{kx} dx = \int_0^{\infty} e^{-(s-k)x} dx$$

which for  $s \leq k$  is divergent, but for  $s > k$  is

$$\left[ \frac{-e^{-(s-k)x}}{s-k} \right]_{x=0}^{\infty} = \frac{1}{s-k}.$$

So  $F(s) = \mathcal{L}\{e^{kx}\}(s) = \frac{1}{s-k}$  for  $s > k$ .

The graph of  $f$  and its Laplace transform  $F$  are given below in figure 1.

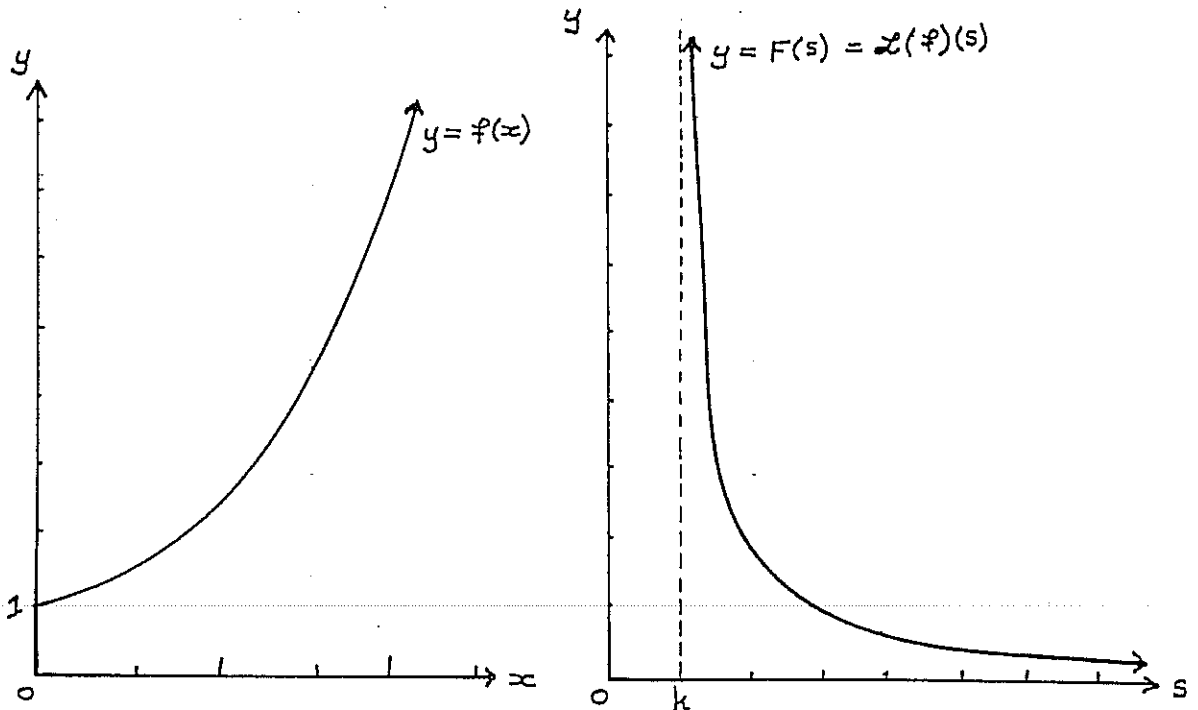


Figure 1

As a special case of this ( $k = 0$ ) we have  $\mathcal{L}(1) = \frac{1}{s}$ . A table of the Laplace transforms of some of the more common functions will be given later.

#### 14.2 Properties of Laplace Transforms.

We now develop those properties of Laplace transforms which are of use in our later work. The first two are direct consequences of linearity of the integral.

(I) For  $f$  a suitable function, and  $k$  a constant

$$\mathcal{L}(kf) = k\mathcal{L}(f)$$

Proof.

$$\begin{aligned} \mathcal{L}(kf)(s) &= \int_0^{\infty} e^{-sx} kf(x) dx \\ &= \int_0^{\infty} e^{-sx} f(x) dx \\ &= k\mathcal{L}(f)(s). \end{aligned}$$

(II) For  $f$  and  $g$  suitable functions, and  $s$  such that both  $F(s)$  and  $G(s)$  are defined,  $\mathcal{L}(f + g) = F + G = \mathcal{L}(f) + \mathcal{L}(g)$ .

Proof. 
$$\begin{aligned} \mathcal{L}(f + g)(s) &= \int_0^{\infty} e^{-sx} (f(x) + g(x)) dx \\ &= \int_0^{\infty} e^{-sx} f(x) dx + \int_0^{\infty} e^{-sx} g(x) dx \\ &= \mathcal{L}(f)(s) + \mathcal{L}(g)(s). \end{aligned}$$

The next string of results investigates the relationship between the operation of taking Laplace transforms and that of differentiation and provides the key for the use of Laplace transforms in solving differential equations.

(III) For  $f$  differentiable, with  $f'$  a suitable function:

$$\mathcal{L}(f')(s) = s\mathcal{L}(f)(s) - f(0)$$

Proof. Observe that

$$-f(0) = \left[ e^{-st} f(t) \right]_{t=0}^{\infty} \text{ provided } \lim_{t \rightarrow \infty} e^{-st} f(t) = 0,$$

a necessary condition if  $f$  is to be a suitable function.

Whence

$$\begin{aligned} -f(0) &= \int_0^{\infty} \frac{d}{dt} (e^{-st} f(t)) dt \quad (\text{provided } f(0) = \lim_{t \rightarrow 0^+} f(t)) \\ &= \int_0^{\infty} \left[ -se^{-st} f(t) + e^{-st} f'(t) \right] dt \\ &= -s \int_0^{\infty} e^{-st} f(t) dt + \int_0^{\infty} e^{-st} f'(t) dt \quad ** \\ &= -s\mathcal{L}(f)(s) + \mathcal{L}(f')(s) \end{aligned}$$

as required.

A corollary of this is

(IV) For  $f$  twice differentiable with  $f''$  a suitable function

$$\mathcal{L}(f'')(s) = s^2 \mathcal{L}(f)(s) - sf(0) - f'(0).$$

Proof. 
$$\begin{aligned} \mathcal{L}(f'')(s) &= \mathcal{L}((f')')(s) \\ &= s\mathcal{L}(f')(s) - f'(0) \quad \text{by (III)} \\ &= s(s\mathcal{L}(f)(s) - f(0)) - f'(0) \\ &= s^2\mathcal{L}(f)(s) - sf(0) - f'(0). \end{aligned}$$

\* (I) and (II) together assert that  $\mathcal{L}$  is a linear mapping from the vector space of suitable functions.

\*\* What we have tacitly used here is the important *integration by parts* formula

$$\int_a^b f(t)g'(t) dt = \left[ f(t)g(t) \right]_a^b - \int_a^b f'(t)g(t) dt.$$

Of course we could continue to obtain expressions for  $\mathcal{L}(f''')$  etc., however for our purposes it suffices to stop at the second derivative.

14.3 *Application of transform methods to  $y'' + ay' + by = f(x)$* , where  $f$  is a suitable function,  $a, b$  are constants and  $y(0) = y_0, y'(0) = y_0'$ .

Begin by taking Laplace transforms of both sides of the equation and, assuming the solution  $y$  is a suitable function with  $Y = \mathcal{L}(y)$ , use the properties listed in 14.2 to simplify the L.H.S. as far as possible.

Thus

$$\mathcal{L}(y'' + ay' + by) = \mathcal{L}(f) = F,$$

$$\mathcal{L}(y'') + a\mathcal{L}(y') + b\mathcal{L}(y) = F,$$

by (I) and (II). So, by (III) and (IV)

$$s^2Y - sy_0 - y_0' + asY - ay_0 + bY = F$$

$$\text{or } (s^2 + as + b)Y - (s + a)y_0 - y_0' = F.$$

We thus see that the differential equation for  $y$  is transformed into an algebraic equation for  $Y$ , the Laplace transform of  $y$ , from which we may determine  $Y$ .

$$Y(s) = \frac{F(s) + sy_0 + ay_0 + y_0'}{s^2 + as + b}$$

So, even though the solution  $y$  is still unknown, provided  $F = \mathcal{L}(f)$  and the initial values  $y_0$  and  $y_0'$  are known,  $Y$ , the Laplace transform of  $y$ , is readily determined. The problem is now one of inversion, for if from knowing the Laplace transform of a function we could determine the function, then knowing  $Y$  we would be able to find  $y$  and hence the sought-after solution.

EXAMPLE. Let us apply the above techniques to the equation

$$y'' + by = r,$$

already considered in Lecture 13, where

$$y(0) = \frac{r}{b} \text{ and } y'(0) = 0.$$

Taking Laplace transforms of both sides gives

$$\mathcal{L}(y'') + b\mathcal{L}(y) = \mathcal{L}(r) = r\mathcal{L}(1)$$

so

$$s^2Y - \frac{sr}{b} + bY = \frac{r}{s} \text{ (from before) or}$$

$$Y = \frac{\frac{r}{s} + \frac{r}{b}s}{s^2 + b} = \frac{r}{bs} \frac{(b + s^2)}{s^2 + b} = \left(\frac{r}{b}\right) \cdot \frac{1}{s}$$

So we are led to seek for the function  $y$  with Laplace transform  $\frac{r}{b} \cdot \frac{1}{s}$ .

In fact, for this case the information needed to accomplish this is already at hand.

If  $f$  is such that  $\mathcal{L}(f) = \frac{1}{s}$  then from (I)  $\mathcal{L}\left(\frac{r}{b}f\right) = \frac{r}{b} \cdot \frac{1}{s}$ , but as already noted  $\mathcal{L}(1) = \frac{1}{s}$  so  $y = \frac{r}{b} \cdot 1$  or  $y = \frac{r}{b}$ . That this constant function is the solution is easily verified upon direct substitution, indeed it was arrived at by other

methods in Lecture 13.

14.4 *Inverse Transforms.* In order that the method outlined in 14.3 may be used profitably to solve non-homogeneous second order differential equations, it is necessary to develop techniques for inverting the Laplace Transform, i.e. of being able to identify a function from a knowledge of its Laplace transform.

One of the first questions which must be settled is whether a function is uniquely identified by its Laplace transform, i.e. whether  $\mathcal{L}(f_1) = \mathcal{L}(f_2)$  implies  $f_1 = f_2$ ? Were this not the case, determining a  $y$  such that  $\mathcal{L}(y)$  equals the known  $Y$  would still leave us in doubt as to whether or not this  $y$  is the required solution of our differential equation.

Fortunately for us the answer to this question is the right one, as shown by the following corollary of Lerch's theorem (which we state without proof).

**THEOREM.** *If  $y_1$  and  $y_2$  are continuous functions (which in our context they must be, since, as solutions of a differential equation they must a priori be differentiable) and  $\mathcal{L}(y_1) = \mathcal{L}(y_2)$ , then  $y_1 = y_2$ .*

Exact inversion formulas allowing the determination of  $f$  from  $\mathcal{L}(f)$  are available, however the sophistication of the mathematics involved place them beyond the scope of these lectures.

A much used and sometimes effective, though rather crude, method of determining  $f$  from  $\mathcal{L}(f)$  is simply to use a table of Laplace transforms in reverse. Such a table giving  $\mathcal{L}(f)$  for some of the more common functions  $f$  is given below. Some of the transforms are rather easily established, others are quite difficult (you should certainly try to prove a few of them.)

TABLE 1.

COMMON LAPLACE TRANSFORMS\*

Function $f(t)$	Transform $F(s) = \mathcal{L}\{f(t)\}$
1	$1/s$
$e^{-at}$	$1/(s + a)$
$\frac{t^n}{n!}$	$1/s^{n+1}$
$\sin t$	$1/(s^2 + 1)$
$\cos t$	$s/(s^2 + 1)$
$\ln t$	$-(\gamma + \ln s)/s$ where $\gamma = \text{Euler's Constant} \doteq 0.577215$

Thus if  $f$  is known to have Laplace transform  $s/(s^2 + 1)$  we see from a consultation of table 1 that  $f(x) = \cos x$ .

If  $F = \mathcal{L}(f)$  we will write  $f = \mathcal{L}^{-1}(F)$ . For example

$$\mathcal{L}^{-1}\{s/(s^2 + 1)\} = \cos.$$

\* Many authorities reverse the table given above, placing  $\mathcal{L}(f)$  in the left-hand column and  $f$  in the right one. It is also common practice to use  $p$  instead of  $s$  (as the argument of a function's Laplace transform).

The power of such a table is greatly enhanced if the basic results of 14.2 on the properties of the laplace transform are borne in mind.

EXAMPLE: Find  $\mathcal{L}^{-1} \left\{ \frac{1}{s^2 - 1} \right\}$ .

Use partial fractions to write

$$\frac{1}{s^2 - 1} = \frac{1}{(s - 1)(s + 1)} = \frac{1}{2} \left( \frac{1}{s - 1} - \frac{1}{s + 1} \right).$$

Now from the table

$$\frac{1}{s + a} = \mathcal{L}\{e^{-at}\}(s),$$

so

$$\begin{aligned} \frac{1}{s^2 - 1} &= \frac{1}{2} (\mathcal{L}\{e^t\}(s) - \mathcal{L}\{e^{-t}\}(s)) \\ &= \mathcal{L}\left\{\frac{1}{2}(e^t - e^{-t})\right\}(s) \quad \text{by I and II of 14.2.} \end{aligned}$$

We therefore conclude that

$$\mathcal{L}^{-1}\left\{\frac{1}{s^2 - 1}\right\}(t) = \frac{1}{2} (e^t - e^{-t}).$$

The basic calculations of this example are a specific instance of a general theorem, which follows as an immediate consequence of (I) and (II) of 14.2 together with the uniqueness result stated in this section. It is

1)  $\mathcal{L}^{-1}(k_1 f_1 + k_2 f_2) = k_1 \mathcal{L}^{-1}(f_1) + k_2 \mathcal{L}^{-1}(f_2).$

The use of partial fractions in conjunction with this to expand a given F into components whose inverses can be found from the table (as in the above example) is a powerful technique in the use of Laplace transforms.

In general the basic skill needed for successful use of Laplace transform methods (a skill you must acquire for yourself - largely I suspect through numerous encounters with specific problems) is the ability to rearrange a given F into alternative forms, from which the desired inverse is more readily recognised. The following theorems, in conjunction with the one given above, are extremely useful in this regard.

2) For f a suitable function

$$\mathcal{L}\{f(kt)\}(s) = \frac{1}{k} F\left(\frac{s}{k}\right)$$

where  $F = \mathcal{L}(f)$ . (The proof follows easily from the definition of  $\mathcal{L}(f)$  and a simple change of variable.)

EXAMPLE: Find  $\mathcal{L}^{-1}\left\{\frac{1}{s^2 + 4}\right\}$ .

Now 
$$\frac{1}{s^2 + 4} = \frac{1}{2} \left( \frac{1}{\left(\frac{s}{2}\right)^2 + 1} \right).$$

The term in brackets is of the form  $\frac{1}{2} F\left(\frac{s}{2}\right)$  where  $F(s) = \frac{1}{s^2 + 1}$



and from Table I we see that  $F = \mathcal{L}(\sin)$ .

So 
$$\mathcal{L}^{-1} \left\{ \frac{1}{s^2 + 4} \right\} (t) = \frac{1}{2} \sin 2t.$$

3) For  $f$  a suitable function

$$\mathcal{L}\{e^{-at} f(t)\}(s) = \mathcal{L}(f)(s + a)$$

Proof.

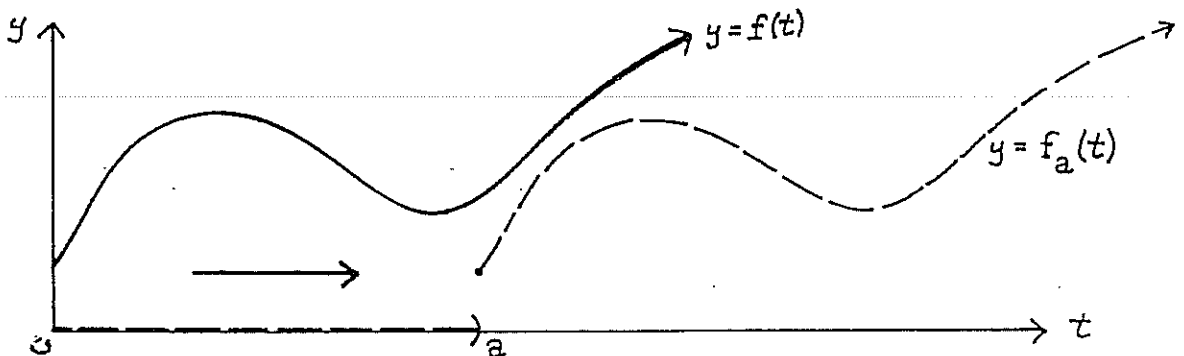
$$\begin{aligned} \mathcal{L}(f)(s + a) &= \int_0^{\infty} e^{-(s+a)t} f(t) dt \\ &= \int_0^{\infty} e^{-st} (e^{-at} f(t)) dt \\ &= \mathcal{L}\{e^{-at} f(t)\}(s). \end{aligned}$$

EXAMPLE:

$$\begin{aligned} \mathcal{L}^{-1} \left\{ \frac{6}{s^2 + 4s + 5} \right\} (t) &= 6 \mathcal{L}^{-1} \left\{ \frac{1}{(s + 2)^2 + 1} \right\} (t) = 6e^{-2t} \mathcal{L}^{-1} \left\{ \frac{1}{s^2 + 1} \right\} (t) \\ &= 6e^{-2t} \sin t. \end{aligned}$$

4) Given a suitable function  $f$  and  $a > 0$  define  $f_a$  by

$$f_a(t) = \begin{cases} 0 & \text{for } 0 \leq t < a \\ f(t - a) & \text{for } t \geq a, \end{cases}$$



then  $f_a$  is also a suitable function and

$$\mathcal{L}(f_a)(s) = e^{-as} \mathcal{L}(f)(s).$$

Proof.

$$\begin{aligned} \mathcal{L}(f_a)(s) &= \int_0^{\infty} e^{-st} f_a(t) dt \\ &= \int_a^{\infty} e^{-st} f(t - a) dt \end{aligned}$$

changing variables to  $x = t - a$  gives

$$\begin{aligned} \mathcal{L}(f_a)(s) &= \int_0^{\infty} e^{-s(x+a)} f(x) dx \\ &= e^{-sa} \int_0^{\infty} e^{-sx} f(x) dx \\ &= e^{-sa} \mathcal{L}(f)(s). \end{aligned}$$

EXAMPLE:  $\mathcal{L}^{-1}\{e^{-3s} s^{-3}\}(t) = f_3(t)$

where

$$\mathcal{L}(f)(s) = \frac{1}{s^3} \text{ and so } f(t) = \frac{t^2}{2!} .$$

Therefore

$$\mathcal{L}^{-1}\{e^{-3s} s^{-3}\}(t) = \begin{cases} 0 & \text{for } 0 \leq t < 3 \\ (t - 3)^2/2 & \text{for } t \geq 3. \end{cases}$$

Another useful result which we will not prove, although the proof is not difficult, is the convolution theorem.

5) If  $f, g$  are suitable functions with  $F = \mathcal{L}(f), G = \mathcal{L}(g)$  then

$$\mathcal{L}^{-1}(FG)(t) = \int_0^t f(x)g(t - x)dx$$

EXAMPLE: Let  $f = \mathcal{L}^{-1}(F)$  then, since  $\mathcal{L}^{-1}\left\{\frac{1}{s}\right\} = 1,$

we have

$$\mathcal{L}^{-1}\left\{\frac{F(s)}{s}\right\}(t) = \int_0^t f(x) \cdot 1 dx = \int_0^t f(x)dx$$

or equivalently  $\mathcal{L}\left\{\int_0^t f(x)dx\right\}(s) = \frac{\mathcal{L}(f)(s)}{s} = \frac{F(s)}{s} .$

To acquire suitable skill in the use of these techniques for the inversion of Laplace transforms is not easy, requiring a lot of practice and effort. Therefore, unless you specifically need to use them, it is probably sufficient that you develop an awareness of them and their potential use.

*Collateral Reading.*

For a good account of the Laplace transform and its applications the reader should consult the monograph by

Earl D. Rainville "The Laplace Transform, an introduction", Macmillan, 1964.

Boyce and DiPrima "Elementary Differential Equations and Boundary Value Problems" devotes a chapter to the Laplace Transform, which might serve as a useful reference.

For a discussion of the Laplace transform and its inversion from a theoretical view-point the reader could refer to Chapter XI of

H.F. Weinberger "A First Course in Partial Differential Equations".

Quite extensive tables of the Laplace transform may be found in many of the Mathematical Handbooks. For example in

R.S. Burington "Handbook of Mathematical Tables and Formulas", McGraw-Hill.

EXERCISES:

1. Use the method of Laplace transforms to solve

(i)  $\frac{d^2y}{dt^2} + y = 0, \quad y(0) = 0, \quad \frac{dy}{dt}(0) = 1$

(ii)  $\frac{dy}{dt} + 2y = 0, \quad y(0) = 1.$

Check your solutions by solving the equations a different way.

2. Using the information given in the table of common Laplace transforms, show that:

$$(i) \quad \mathcal{L}\{\sin nt\} = \frac{n}{s^2 + n^2}$$

$$(ii) \quad \mathcal{L}\{\cos nt\} = \frac{s}{s^2 + n^2}$$

$$(iii) \quad \frac{1}{(s + a)^2} \text{ is the Laplace transform of } te^{-at}.$$

3. Find  $\mathcal{L}\{e^{-t} \sin 2t\}$ .

4. Use Laplace transforms to solve:

$$(i) \quad \frac{d^2y}{dt^2} - y = e^{-t} \quad y(0) = 1, \quad \frac{dy}{dt}(0) = -1$$

$$(ii) \quad \frac{d^2y}{dt^2} + y = t \quad y(0) = \frac{dy}{dt}(0) = 0$$

$$(iii) \quad \frac{d^2y}{dt^2} + y = \sin 2t \quad y(0) = 0, \quad \frac{dy}{dt}(0) = 1.$$

(Hint: Use partial fractions to simplify the expression you get for  $\mathcal{L}\{y\}$ .)

Graph the solutions to (ii) and (iii), that of (iii) might describe the aquatic population of a lake into which nutrients are supplied at a rate varying sinusoidally throughout the year in accord with the seasons (see Exercise 13.)

5. See if you can find how the method of Laplace transforms might be applied to the problem of solving a linear system of equations, such as

$$\frac{dx_1}{dt} = ax_1 + bx_2 \quad x_1(0) = A$$

$$\frac{dx_2}{dt} = cx_1 + dx_2 \quad x_2(0) = B$$

The application of Laplace transforms to such systems, and more complicated non-homogeneous ones i.e.

$$\frac{dx_1}{dt} = ax_1 + bx_2 + f_1(t)$$

$$\frac{dx_2}{dt} = cx_1 + dx_2 + f_2(t)$$

is of importance in systems analysis.

\* APPENDIX to Lecture 14      *Variation of Parameters*

Let  $y(x) = Ay_1(x) + By_2(x)$  be the general solution of the second-order linear homogeneous equation

$$y'' + a(x)y' + b(x)y = 0$$

where  $y_1$  and  $y_2$  are themselves solutions and  $A, B$  are constants.

We seek for a solution of the non-homogeneous problem

$$y'' + a(x)y' + b(x)y = f(x)$$

of the form

$$y(x) = u(x)y_1(x) + v(x)y_2(x)$$

where the coefficient functions  $u$  and  $v$  are to be determined.

To determine these two functions  $u, v$  uniquely, we require them to satisfy two conditions. One such condition comes from the requirement that  $y$  be a solution of the non-homogeneous equation. The other may be selected arbitrarily. We will therefore choose it in such a way as to facilitate our calculations.

Now, we have

$$y'(x) = [u'(x)y_1(x) + v'(x)y_2(x)] + [u(x)y_1'(x) + v(x)y_2'(x)]$$

We will take as our first (arbitrary) relationship on  $u$  and  $v$  that the first term of this expression for  $y'$  vanish, i.e. we will require that

$$u'(x)y_1(x) + v'(x)y_2(x) = 0$$

and so  $y'$  has the simpler form

$$y'(x) = u(x)y_1'(x) + v(x)y_2'(x), \quad \text{whence}$$

$$y''(x) = u(x)y_1''(x) + v(x)y_2''(x) + u'(x)y_1'(x) + v'(x)y_2'(x)$$

Substituting these into the non-homogeneous equation and rearranging yields  $u(x)[y_1'' + ay_1' + by_1] + v(x)[y_2'' + ay_2' + by_2] + u'(x)y_1'(x) + v'(x)y_2'(x) = f(x)$

Now both the terms in brackets are zero, since  $y_1$  and  $y_2$  are solutions of  $y'' + ay' + by = 0$  and so as our second condition on  $u, v$  we have

$$u'(x)y_1'(x) + v'(x)y_2'(x) = f(x).$$

Solving these two conditions on  $u, v$  simultaneously we obtain

$$u'(x) = \frac{-y_2(x)f(x)}{y_1(x)y_2'(x) - y_1'(x)y_2(x)}$$

and

$$v'(x) = \frac{y_1(x)f(x)}{y_1(x)y_2'(x) - y_1'(x)y_2(x)}$$

Thus  $u$  and  $v$  are the solutions of the most trivial form of first-order equation (where the derivatives are equal to known functions of  $x$ ). From these,  $u$  and  $v$  may be determined by integration and so a solution  $y$  given by

$$y(x) = y_1(x) \int_0^x \frac{-y_2(t)f(t)}{y_1(t)y_2'(t) - y_1'(t)y_2(t)} dt + y_2(x) \int_0^x \frac{y_1(t)f(t)}{y_1(t)y_2'(t) - y_1'(t)y_2(t)} dt$$

provided the required integrals exist.

EXAMPLE. Solve  $y'' + 2y' + y = e^{-x}$ . The homogeneous equation  $y'' + 2y' + y = 0$  has solutions

$$y_1(x) = e^{-x}, \quad y_2(x) = xe^{-x}$$

(as may be shown by the methods of lecture 13).

So a solution of the non-homogeneous equation is

$$y(x) = e^{-x} \int_0^x \frac{(-te^{-t})e^{-t}}{e^{-2t}} dt + xe^{-x} \int_0^x \frac{(e^{-t})e^{-t}}{e^{-2t}} dt$$

since

$$y_1(t)y_2'(t) - y_1'(t)y_2(t) = e^{-2t}.$$

So

$$\begin{aligned} y(x) &= e^{-x} \int_0^x -t dt + xe^{-x} \int_0^x dt \\ &= \frac{1}{2}e^{-x} [-x^2] + xe^{-x} [x] \\ &= \frac{1}{2}x^2e^{-x}, \end{aligned}$$

a result easily verified by direct substitution. It is just as easily verified that

$$y(x) = \frac{1}{2}x^2e^{-x} + Ae^{-x} + Bxe^{-x}$$

is also a solution for arbitrary values of A and B, which may therefore be selected to satisfy initial conditions.

The method of variation of parameters not only provides a technique whereby certain non-homogeneous equations may be solved, it is also the basis for much of the known general theory of such equations. Of course while the method gives a solution of such equations in closed form it may not be easy (or indeed possible) to express the resulting integrals in terms of elementary functions. Nevertheless, even in such a case, this form of the solution allows numerical solutions to be obtained from the use of methods such as Simpson's\* to evaluate approximately the integrals which occur. Frequently it is more accurate and advantageous to determine the solution numerically in this way rather than directly by a method like that outlined in lecture 7.

---

\*  $\int_a^b f(x)dx \doteq \frac{(b-a)}{6n} (f_0 + 4f_1 + 2f_2 + 4f_3 + \dots + 4f_{2n-1} + f_{2n})$

where  $f_m = f\left(a + m \left(\frac{b-a}{2n}\right)\right)$

The size of the error decreases with increasing n, provided f satisfies certain smoothness conditions. (In fact the error is of order  $\frac{1}{n^4}$ .)

Lecture 15. *Impulsive Inputs.*

We examine the differential equation

$$\frac{d^2y}{dt^2} + a \frac{dy}{dt} + by = f(t)$$

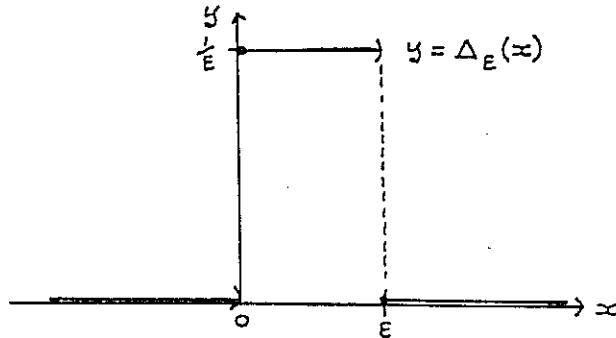
where the input function  $f(t)$  represents an impulse (such as would result from a sharp blow in the mechanical case). To a first approximation, such an impulse, occurring at 'time'  $t_0$ , could be represented by

$$f(t) = k\Delta_\epsilon(t - t_0)$$

where  $k$  is a constant and  $\Delta_\epsilon$  is the function defined by

$$\Delta_\epsilon(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1/\epsilon & \text{for } 0 \leq x < \epsilon \\ 0 & \text{for } x \geq \epsilon. \end{cases}$$

The approximation being better the smaller the value of  $\epsilon$  taken.



A number of properties possessed by the function  $\Delta_\epsilon$  are of importance.

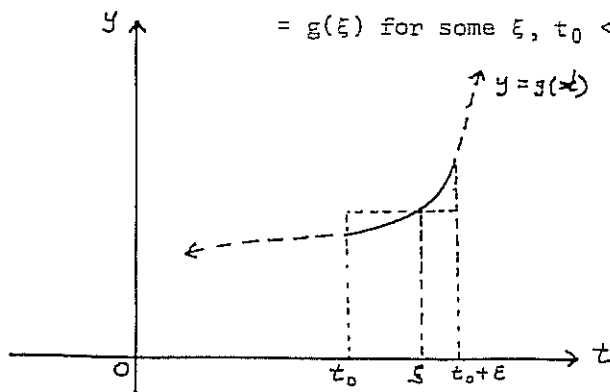
1.  $\Delta_\epsilon(t) = 0$  for  $t > \epsilon$ .

2.  $\int_0^\infty \Delta_\epsilon(t) dt = \int_0^\epsilon \frac{1}{\epsilon} dt = 1$

3. For  $g$  any continuous function, we have from the integral mean value theorem

$$\begin{aligned} \int_0^\infty g(t)\Delta_\epsilon(t - t_0)dt &= \frac{1}{\epsilon} \int_{t_0}^{t_0+\epsilon} g(t)dt \\ &= \frac{1}{\epsilon} g(\xi)(t_0 + \epsilon - t_0) \end{aligned}$$

$$= g(\xi) \text{ for some } \xi, t_0 < \xi < t_0 + \epsilon.$$



We idealise these results in the *Dirac delta 'function'*,  $\delta$ , which has the properties:

1.  $\delta(t) = 0 \quad t > 0$ ;
2.  $\int_0^{\infty} \delta(t) dt = 1$ ;
3. For any function  $g$ ,  $\int_0^{\infty} g(t) \delta(t - t_0) dt = g(t_0)$ ,

So  $\delta$  embodies the properties of  $\Delta_\epsilon$  which would be expected in the limit as  $\epsilon \rightarrow 0$ , and so represents an idealised impulsive input.

Strictly, no function having the properties assigned to  $\delta$  can exist. However, the concept of the Dirac delta function is, as we shall see, extremely useful and provided we regard  $\delta$  and all operations performed with it as purely formal\* no contradictions will arise.

Directly from the definition of  $\delta$  we have

$$\mathcal{L}\{\delta(t - t_0)\}(s) = \int_0^{\infty} e^{-st} \delta(t - t_0) dt = e^{-st_0}.$$

So

$$\frac{d^2y}{dt^2} + a \frac{dy}{dt} + by = \delta(t - t_0), \quad y(0) = y_0, \quad y'(0) = y'_0$$

can be solved using Laplace transform methods by the techniques outlined in lecture 14.

EXAMPLE. Solve  $\frac{d^2y}{dt^2} + by = \delta(t - t_0)$ ,  $y(0) = y'(0) = 0$ .

Letting  $Y = \mathcal{L}(y)$  and taking Laplace transforms of both sides gives

$$s^2Y(s) + bY(s) = e^{-st_0} \quad \text{or}$$

$$Y(s) = e^{-st_0}/(s^2 + b)$$

Now 
$$\frac{1}{s^2 + b} = \frac{1}{\sqrt{b}} \left( \frac{1}{\sqrt{b}} \frac{1}{\left(\frac{s}{\sqrt{b}}\right)^2 + 1} \right) = \mathcal{L}\left\{ \frac{1}{\sqrt{b}} \sin \sqrt{bt} \right\}(s)$$

by 2 of 14.4 and so by 4 of 14.4

$$y(t) = \mathcal{L}^{-1}\{e^{-st_0}/(s^2+b)\}(t) = 0 \text{ for } t < t_0$$

$$= \frac{1}{\sqrt{b}} \sin \sqrt{b}(t - t_0) \text{ for } t \geq t_0.$$

Apart from their immediate practical importance impulsive inputs, and hence the Dirac delta function, are of considerable theoretical importance because of the following result.

---

\* A justification of these formal procedures is provided by the theory of distributions recently developed by L. Schwartz.

THEOREM (Green's Convolution Theorem)

Let  $y = \phi(t)$  be a solution of

$$\frac{d^2y}{dt^2} + a \frac{dy}{dt} + by = \delta(t). \quad (a, b \text{ constants})$$

Then a solution of  $\frac{d^2y}{dt^2} + a \frac{dy}{dt} + by = f(t)$  is

$$y = \psi(t) = \int_0^{\infty} f(x) \phi(t-x) dx .$$

A given set of initial conditions can then be met by taking  $y = \psi + h$  where  $h$  is an appropriately chosen solution of the corresponding homogeneous equation  $y'' + ay' + by = 0$ .

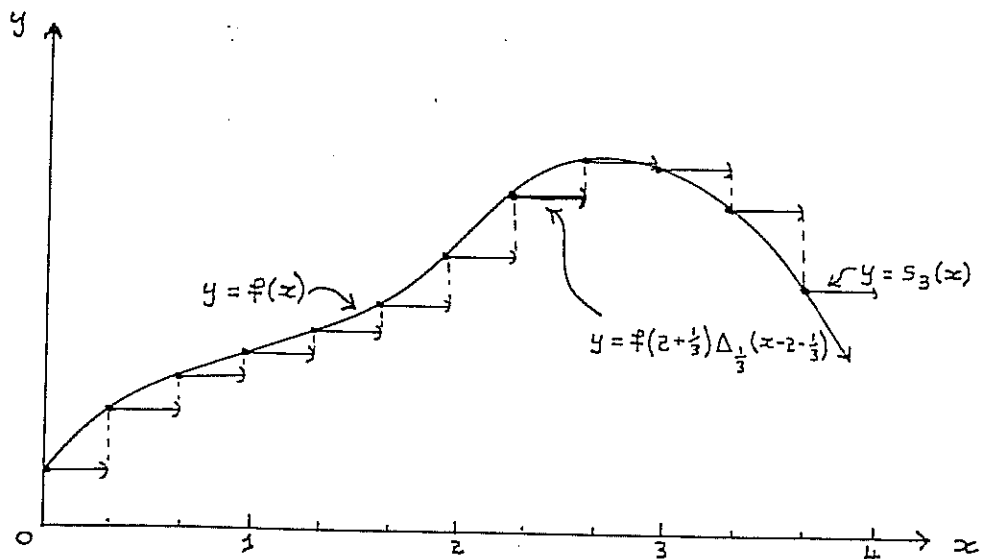
Thus, provided the equation can be solved for an impulsive input at  $t = 0$ , the solution resulting from any other input can be expressed in the form of the integral above. At worst this could be evaluated numerically (see comments on page 80) so yielding a numerical solution to our problem.

A rigorous proof of this result is beyond our scope; the following argument is intended to demonstrate its plausibility.

Proof (of Green's Convolution Theorem). We begin by making two observations.

1. If  $y = w(t)$  is the solution of  $y'' + ay' + by = g(t)$  then  $y = w(t-x)$  is the solution of  $y'' + ay' + by = g(t-x)$ , where  $x$  is any fixed constant. (Follows since  $a, b$  are constants.)
2. If  $y_i$  is a solution of  $y'' + ay' + by = g_i$  ( $i = 1, 2, \dots$ ), then  $y = \sum_i y_i$  is a solution of  $y'' + ay' + by = \sum_i g_i$ . (Follows from the linearity of the equation.)

Now any reasonable input function  $f$  can be approximated by a 'step function' each step of which has width  $\frac{1}{n}$  ( $n$  some natural number), the approximation improving the larger the value of  $n$  selected. If this is done as illustrated in figure 1





we have,

$$f(t) = s_n(t) = \sum_{m=0}^{\infty} \sum_{k=0}^{n-1} f(m + \frac{k}{n}) \Delta_{\frac{1}{n}}(t - m - \frac{k}{n}) \cdot \frac{1}{n}$$

since 
$$\Delta_{\frac{1}{n}}(t - m - \frac{k}{n}) \cdot \frac{1}{n} = \begin{cases} 1 & \text{for } m + \frac{k}{n} \leq t < m + \frac{k+1}{n} \\ 0 & \text{otherwise} \end{cases}$$

Let  $\phi_n(t)$  be a solution of

$$y'' + ay' + by = \Delta_{\frac{1}{n}}(t),$$

then from observations 1. and 2. it follows that a solution of

$$y'' + ay' + by = s_n(t) \quad \text{is}$$

$$y_n(t) = \sum_{m=0}^{\infty} \sum_{k=0}^{n-1} f(m + \frac{k}{n}) \phi_n(t - m - \frac{k}{n}) \cdot \frac{1}{n}$$

Now since  $s_n \rightarrow f$  and  $\Delta_{\frac{1}{n}}(t) \rightarrow \delta(t)$ , as  $n \rightarrow \infty$ , it is reasonable to expect

that  $\phi_n(t) \rightarrow \phi(t)$  and  $y_n(t) \rightarrow y(t)$ , where  $y$  is a solution of

$$y'' + ay' + by = f,$$

while the inner sum over  $k$  will convert to an integral between  $m$  and  $m + 1$ , so

$$\begin{aligned} y(t) &= \sum_{m=0}^{\infty} \int_m^{m+1} f(x) \phi(t-x) dx \\ &= \int_0^{\infty} f(x) \phi(t-x) dx \quad \text{as required.} \end{aligned}$$

EXAMPLE. Solve  $\frac{d^2y}{dt^2} + by = r$  (c.f. lectures 13 and 14)

From the earlier example of this lecture we have

$$\phi(t) = \frac{1}{\sqrt{b}} \sin \sqrt{bt} \quad \text{for } t \geq 0$$

is a solution of  $y'' + by = \delta(t)$ .

So a solution of the current problem is

$$\begin{aligned} y(t) &= \int_0^{\infty} r \phi(t-x) dx. \\ &= r \int_t^{\infty} \phi(t-x) dx + r \int_0^t \phi(t-x) dx. \end{aligned}$$

In the first of these two integrals  $x > t$  so  $t - x < 0$  and  $\phi(t-x) = 0$ , while in the second  $\phi(t-x) = \frac{1}{\sqrt{b}} \sin \sqrt{b}(t-x)$ .

Whence

$$y(t) = 0 + \frac{r}{\sqrt{b}} \int_0^t \sin \sqrt{b} (t - x) dx$$
$$= \left[ \frac{r}{b} \cos \sqrt{b}(t - x) \right]_{x=0}^{x=t}$$
$$= \frac{r}{b} - \frac{r}{b} \cos \sqrt{bt}.$$

*Collateral Reading:*

A fuller account of impulsive inputs from a physical point of view, together with further discussion and application of Green's Convolution Theorem may be found in

F.H. Raven "Mathematics of Engineering Systems"

OR

Papoulis "The Fourier Transform and its Applications", McGraw-Hill, 1962.

EXERCISES:

1. Solve  $y'' + y = \delta(t - 1)$  if  $y(0) = y'(0) = 0$ .
2. Using Green's Convolution Theorem find a solution of

$$y'' + y = \sin 2t.$$

Hence find the solution satisfying  $y(0) = 0, y'(0) = 1$ .

3. Show that it is reasonable to expect that, for any function  $f$ ,

$$f(x) = \int_0^{\infty} f(t) \delta(x - t) dt.$$

\* Lecture 16. *Oscillatory Solutions* - Sturm's Theory

This lecture should be read in conjunction with lecture 12 in which the question, "when does a given system have periodic solutions?" was discussed. It has been noted that the Lotka-Volterra system has periodic solutions as did the lake ecology model introduced in lecture 13. However it was also noted in lecture 13 that if the effect of 'fishing' was allowed for, an equation of the form

$$y'' + ay' + by = r$$

might result (exercise 13.2).

The solution of this equation is of the form

$$y(t) = \frac{r}{b} + Ae^{at} \sin(\beta t + \phi),$$

at least for small values of  $a > 0$ , and so is not periodic even though it continues to oscillate back and forward across the line  $y = \frac{r}{b}$  (see figure 13.2 (b)).

Similarly if  $(0,0)$  is a spiral point of any system for  $x$  and  $y$  we see from the phase-portrait (Table 12.1) the values of  $x$  and  $y$  continuously oscillate back and forth through  $x = 0$ . However neither  $x$  or  $y$  are periodic since their amplitude is either decreasing (asymptotically stable spiral point) or increasing (unstable spiral point).

In many contexts the existence of such oscillatory solutions is just as important as, and far more common than that of periodic solutions. Accordingly we examine conditions for the existence of oscillatory solutions to equations of the form

$$y'' + a(x)y' + b(x)y = 0.$$

Before proceeding further it is necessary to state more precisely what we shall mean by such an equation having an oscillatory solution.

The property of a function  $y = f(x)$  oscillating about the line  $y = 0$  is characterised by its repeated crossing of the line  $y = 0$ . We therefore offer the following definition.

DEFINITION. The function  $f$  is *oscillatory* (about  $y = 0$ ) if there exists a sequence of points  $x_1 < x_2 < x_3 < \dots < x_n < \dots$  with  $x_n \rightarrow \infty$  as  $n \rightarrow \infty$  such that  $f(x_i) = 0$   $i = 1, 2, \dots$

Since, whether or not  $y'' + ay' + by = 0$  has oscillatory solutions depends entirely on where the zeros of each solution are located, we could replace this equation by any more convenient one, provided its solution vanished at  $x = x_0$  if and only if the solution of  $y'' + ay' + by = 0$  also had a zero at  $x_0$ .

Thus let  $y(x) = u(x)v(x)$  where  $u$  is a function such that  $u(x) \neq 0$  for all  $x$ . Then  $a$  is a zero of  $v$  if and only if it is a zero of  $y$ , further  $v$  satisfies the differential equation obtained by substituting  $uv$  in place of  $y$  in  $y'' + ay' + by = 0$ , viz

$$v'' + \left(\frac{2u'}{u} + a\right)v' + \left(\frac{u''}{u} + a\frac{u'}{u} + b\right)v = 0$$

division by  $u$  being justified since  $u(x) \neq 0$ , for any  $x$ , by assumption.

It is particularly suitable to choose

$$u(x) = e^{-\frac{1}{2} \int_0^x a(t) dt}$$

for then  $u(x) \neq 0$ , any  $x$ , as required and  $\frac{2u'}{u} + a = 0$  so  $v$  satisfies the simpler equation

$$v'' + (b - \frac{1}{4} a^2 - \frac{1}{2} a')v = 0,$$

an equation of the form  $v'' + I(x)v = 0$ , known as the *normal form* of  $y'' + ay' + by = 0$ .

EXAMPLE. The normal form of  $y'' + 2y' - y = 0$  is  $v'' - 2v = 0$ .

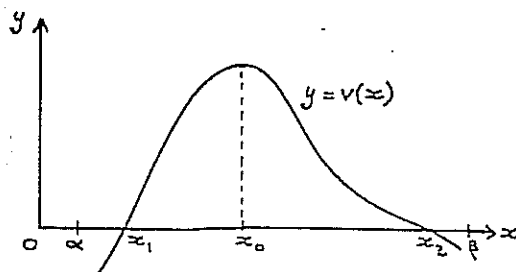
From the remarks made above concerning the existence of oscillatory solutions, it is sufficient to consider only the normal form of an equation. Henceforth, therefore, we will restrict our attention to equations of the form

$$v'' + I(x)v = 0.$$

THEOREM. (*Convexity*) If  $I(x) < 0$  for  $\alpha < x < \beta$ , then any non-trivial solution of  $v'' + I(x)v$  has at most one zero between  $\alpha$  and  $\beta$ .

Proof. Let  $v$  be a non-trivial solution and assume  $v$  has two consecutive zeros at  $x_1$  and  $x_2$  where  $\alpha < x_1 < x_2 < \beta$ . Then either  $v(x) > 0$  for  $x_1 < x < x_2$  or  $v(x) < 0$  for  $x_1 < x < x_2$ . (This follows since  $v$  is continuous and  $v'(x_1) \neq 0$ , otherwise the unique solution satisfying  $v(x_1) = v'(x_1) = 0$  would be  $v \equiv 0$ .)

In case  $v(x) > 0$  for  $x_1 < x < x_2$  we have  $v''(x) = -I(x)v(x) > 0$  for every  $x$  between  $x_1$  and  $x_2$  but, since  $v(x_1) = v(x_2) = 0$  there must exist a point  $x_0$  between  $x_1$  and  $x_2$  at which  $v$  attains a maximum so  $v'(x_0) = 0$  and



$v''(x_0) \leq 0$  a contradiction. We therefore conclude that in this case there does not exist consecutive zeros of  $v$  between  $\alpha$  and  $\beta$  and so  $v$  has at most one zero in the interval. A similar argument applies in the case of  $v(x) < 0$  for  $x_1 < x < x_2$ .

EXAMPLE. The equation  $y'' - x^3y = 0$  satisfies the conditions of the theorem for all  $x$  with  $\epsilon < x < \infty$  (where  $\epsilon$  is any strictly positive number). So, there is at most one zero greater than  $\epsilon$  and we may conclude that this equation admits no oscillatory solutions.

THEOREM (Sturm's Theorem). Let  $u$  and  $v$  be non-trivial solutions of  $v'' + I(x)v = 0$  and  $u'' + J(x)u = 0$  respectively where  $I(x) > J(x)$  for all  $x$  with  $\alpha < x < \beta$ . Then  $v$  has at least one zero between any two consecutive zeros of  $u$  provided they are both between  $\alpha$  and  $\beta$ .

Proof. Let  $x_1 < x_2$  be consecutive zeros of  $u$ , both lying between  $\alpha$  and  $\beta$ , and assume  $v(x) \neq 0$  for  $x_1 < x < x_2$ .

Then both  $u$  and  $v$  have a constant sign throughout the interval from  $x_1$  to  $x_2$ . Without loss of generality we take both to be strictly positive i.e.

$u(x) > 0$  and  $v(x) > 0$  for all  $x$  with  $x_1 < x < x_2$

(possible, since  $-u$  and  $-v$  are also solutions satisfying the required conditions).

Now construct

$$w(x) = v(x)u'(x) - u(x)v'(x), \text{ then}$$

$w(x_1) \geq 0$  (since  $u(x_1) = 0$  and  $u(x) > 0$  for  $x_1 < x < x_2$  and so  $u'(x_1) > 0$ ) similarly

$$w(x_2) \leq 0.$$

So (by the Mean Value Theorem) there is a point  $x_0$  between  $x_1$  and  $x_2$  for which  $w'(x_0) \leq 0$ .

But, 
$$w'(x) = v(x)u''(x) - u(x)v''(x)$$

$$= [I(x) - J(x)]v(x)u(x) > 0$$
 by the assumptions on  $I, J, v$  and  $u$  for  $x_1 < x < x_2$ , a contradiction, so  $v$  must have a zero between  $x_1$  and  $x_2$ .

While this theorem (which remains true under the weaker assumption  $I(x) \geq J(x)$ ,  $I \not\equiv J$  for  $\alpha < x < \beta$ ) is of importance in itself, of greater importance to us is the following.

**COROLLARY.** *If  $I$  is such that  $I(x) > m > 0$  for all  $x > \alpha \geq 0$ , then any non-trivial solution of  $v'' + I(x)v = 0$  is oscillatory.*

**Proof.** The equation  $u'' + mu = 0$  has a non-trivial solution  $u(x) = \sin \sqrt{m} x$  which has zeros at  $x = 0, \pi/\sqrt{m}, 2\pi/\sqrt{m}, 3\pi/\sqrt{m}, \dots$  and so, since Sturm's Theorem applies, there exists zeros  $x_1, x_2, x_3, \dots$  of any non-trivial solution of  $v'' + I(x)v = 0$  with

$$0 \leq \alpha \leq k\pi/\sqrt{m} < x_1 < (k+1)\pi/\sqrt{m} < x_2 < (k+2)\pi/\sqrt{m} < \dots,$$

for some positive integer  $k$ .

**EXAMPLE.** The normal form of Bessel's equation, which arises in biophysical models of cell chemistry, is

$$v'' + \left(1 + \frac{1 - 4v^2}{x^2}\right)v = 0$$

where  $v$  is a constant.

For this equation we have

$$I(x) = \left(1 + \frac{1 - 4v^2}{x^2}\right) > 1 \text{ for all } x \text{ if } |v| < \frac{1}{2}$$

and 
$$> \frac{1}{2} \text{ for } x > \sqrt{\frac{4v^2 - 1}{2}} \text{ any other } v.$$

So for any  $v$  any non-trivial solution of Bessel's equation is oscillatory.

When an equation satisfies the conditions of the above corollary, and so is known to have oscillatory solutions, more can be said about their spacing. We close by establishing one result in this direction.

**THEOREM (Spacing of Zeros)**

*Let  $I$  be such that  $0 < m < I(x) < M$  for  $\alpha < x < \beta$ , where  $m, M$  are constants, then if  $x_1$  and  $x_2$  are two consecutive zeros of  $v'' + I(x)v = 0$ , with  $\alpha < x_1 < x_2 < \beta$ , we have*

$$\frac{\pi}{\sqrt{M}} < x_2 - x_1 < \frac{\pi}{\sqrt{m}}.$$

Proof. Consider the comparison equation  $u'' + Mu = 0$  which has a non-trivial solution  $u(x) = \sin \sqrt{M}(x - x_1)$  with a zero at  $x_1$  and the next zero at  $x_1 + \pi/\sqrt{M}$ . Now by Sturm's Theorem there is a zero of  $u$  between  $x_1$  and  $x_2$  so we must have  $x_2 > x_1 + \pi/\sqrt{M}$  or

$$\pi/\sqrt{M} < x_2 - x_1.$$

The upper bound is established in the same manner.

EXAMPLE. For each  $x_0 > 0$  it is easily seen that we can select numbers  $m(x_0)$  and  $M(x_0)$  with

$$m(x_0) < \left(1 + \frac{1 - 4v^2}{x^2}\right) < M(x_0) \text{ for all } x > x_0$$

in such a way that  $m(x_0)$  and  $M(x_0)$  both tend to 1 as  $x_0 \rightarrow \infty$ . Thus for large values of  $x$ , the spacing between consecutive zeros of any non-trivial solution to Bessel's equation is very nearly  $\pi$ .

*Collateral Reading:*

Einar Hille's "Lectures on Ordinary Differential Equations" (Addison-Wesley, 1969) gives a full account of Sturm's Oscillation Theory.

EXERCISES:

1. A population grows according to the assumptions made in the lake ecology model introduced in lecture 1.3 with the rate of nutrient accumulation varying sinusoidally according to  $a + b \cos x$  where  $0 < b < a$ . If the per capita rate of expenditure of nutrient varies sympathetically with the rate of nutrient input, i.e. is proportional to it (a reasonable assumption since immediately after input the nutrient is in a readily available suspended state), show that upon an appropriate change of variable the resulting equation becomes Mathieu's equation,

$$v'' + k(a + b \cos x)v = 0 \text{ (where } k > 0 \text{)}.$$

Prove that Mathieu's equation has oscillatory solutions, and estimate bounds for the spacing between consecutive zeros.

2. Show that if  $y$  is a non-trivial solution of  $y'' + I(x)y = 0$  with  $I(x) > 0$  for  $\alpha < x < \beta$  where  $y'(\alpha) = y'(\beta) = 0$ , then  $y$  has a zero between  $\alpha$  and  $\beta$ .
3. Prove the normal form of  $y'' + \frac{1}{x}y' + \left(1 - \frac{v^2}{x^2}\right)y = 0$  is that of Bessel's equation as given in the lecture.
4. (The condition  $m > 0$  cannot be dropped from the corollary to Sturm's Theorem.) Show that Euler's equation  $y'' + \frac{1}{8}x^{-2}y = 0$  has solutions of the form  $y = Ax^{r_1} + Bx^{r_2}$  where  $r_1$  and  $r_2$  are appropriately chosen real numbers and so has at most one zero greater than 0, even though  $I(x) = \frac{1}{8}x^{-2} > 0$  for all  $x > 0$ .

SERIES IV - Partial Differential Equations

Lecture 17. *Functions of more than one Variable, Partial Differentiation*

We begin with a short 'review' of the concept of a function of two real variables. The extension of most of this work to functions of more than two variables is mechanical and should be obvious. For convenience we will denote the set of all ordered pairs of real numbers by  $\mathbb{R}^2$ .

DEFINITION. If  $f$  is a 'rule' which associates with each ordered pair of real numbers belonging to  $D \subseteq \mathbb{R}^2$  a unique real number (which usually varies from ordered pair to ordered pair), then we say  $f$  is a real valued *function of two variables*. The set  $D$  is called the *domain of definition*, domain for short, of  $f$ . Frequently we denote by  $f(x,y)$  the unique number which  $f$  associates with  $(x,y) \in D$ .

Functions of two variables have implicitly entered our work in many of the previous lectures. For example in discussing the spread of a disease through a community of size  $x$  we reasoned that the rate of spread of the disease might be  $r = ky(x-y)$  where  $y$  was the number of infected individuals and  $k > 0$  was a constant. Here  $x$  and  $y$  could assume values subject to the constraint  $0 < y \leq x$ . So  $r = f(x,y)$  where  $f$  is the function of two variables defined on the domain  $D = \{(x,y) : x > 0 \text{ and } y \leq x\}$  by

$$f(x,y) = ky(x - y)$$

Thus

$$f(5,2) = 6k \text{ for example.}$$

In general the right-hand sides of many of the first order equations, and most of the systems so far considered were functions of two variables.

It sometimes helps in visualising such functions if we realise that

$$\{(x,y,z) : z = f(x,y)\}$$

is a surface in three dimensional cartesian space which intersects any line perpendicular to the  $x$ - $y$  plane in at most one point; see Figure 1.

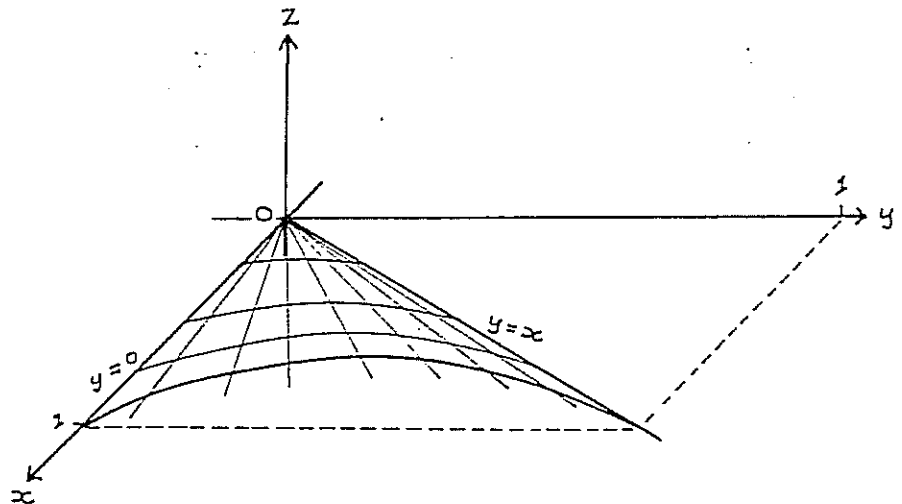


Figure 1

The surface  $\{(x,y,z) : z = y(x-y), 0 < y \leq x\}$

For any function  $f$  of two variables, defined on domain  $D$ , it should be clear that for any fixed value of  $y$ ,  $y_0$  say

$$z = \phi(x) = f(x, y_0)$$

defines a function of one variable on

$$\{x: (x, y_0) \in D\}.$$

For example if  $D = \{(x, y): 0 < y \leq x\}$ ,  $y_0 = 2$  and  $f(x, y) = y(x-y)$  we would have

$$z = \phi(x) = 2x - 4 \text{ for } x \geq 2.$$

Now if  $\phi$  is differentiable at  $x = x_0$ , i.e.

$$\phi'(x_0) = \lim_{h \rightarrow 0} \frac{\phi(x_0+h) - \phi(x_0)}{h} \text{ exists,}$$

then  $\phi'(x_0)$  measures the instantaneous rate of change of  $f$  with respect to  $x$  at  $x_0$  when  $y$  is held constant at  $y_0$ .

We will call this the *partial derivative of  $f$  with respect to  $x$  at  $(x_0, y_0)$*  and denote it by either  $\frac{\partial f}{\partial x}(x_0, y_0)$  or  $f_x(x_0, y_0)$ .

Thus 
$$f_x(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0) = \lim_{h \rightarrow 0} \frac{f(x_0+h, y_0) - f(x_0, y_0)}{h}$$

provided the limit exists.

For  $f$  as in the above example we would have

$$\frac{\partial f}{\partial x}(x_0, 2) = \phi'(x_0) = 2 \text{ for all } x_0 \geq 2.$$

Now let  $D_x$  be the set of ordered pairs  $(x_0, y_0) \in D$  at which the partial derivative of  $f$  with respect to  $x$  exists. Then we can construct a new function of two variables by associating with each  $(x_0, y_0) \in D_x$  the value of  $\frac{\partial f}{\partial x}(x_0, y_0)$ . We call this new function the *partial derivative of  $f$  with respect to  $x$*  and denote it simply by  $\frac{\partial f}{\partial x}$ ,  $f_x$  or better  $f_1$  (the subscript 1 denoting differentiation with respect to the first variable which of course need not be an  $x$ )

EXAMPLE: If for all  $(x, y) \in \mathbb{R}^2$ ,  $f(x, y) = ye^{xy}$  we have for any fixed value of  $y$

$$\phi(x) = ye^{xy}$$

and so

$$\phi'(x) = y^2e^{xy}$$

( $y$  being regarded as a constant for the purpose of the differentiation)

or 
$$\frac{\partial f}{\partial x} = y^2e^{xy} \text{ for all } (x, y) \in \mathbb{R}^2.$$

If  $D_x = D$ , as in the example, we say  $f$  is partially differentiable with respect to  $x$  on  $D$ . This will be the case for most functions encountered in the ensuing lectures and in such cases we will usually suppress all mention of domains, allowing the context to identify  $D$  implicitly

Of course the whole of the previous discussion could be repeated with the roles of  $x$  and  $y$  interchanged in which case we would arrive at the *partial derivative of  $f$  with respect to  $y$*  defined as



$$f_y(x,y) = \frac{\partial f}{\partial y}(x,y) = \lim_{h \rightarrow 0} \frac{f(x,y+h) - f(x,y)}{h}$$

provided this limit exists.

$\frac{\partial f}{\partial y}$  is determined in precisely the same way as  $\frac{\partial f}{\partial x}$ ,  $x$  being regarded as a constant for the purpose of differentiation with respect to  $y$ .

EXAMPLE. Let  $f$  be as above, i.e.  $f(x,y) = ye^{xy}$ , then  $\frac{\partial f}{\partial y} = e^{xy} + xye^{xy}$  (using the product rule to differentiate  $\phi(y) = ye^{xy}$ ,  $x$  being regarded as a constant).

Having obtained from  $f$  the new functions  $\frac{\partial f}{\partial x}$  and  $\frac{\partial f}{\partial y}$  we could now continue the process to obtain second and higher order partial derivatives by

$$\frac{\partial^2 f}{\partial x^2} = f_{xx} \equiv \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial x} \right), \quad \frac{\partial^2 f}{\partial y^2} = f_{yy} \equiv \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial y} \right)$$

$$\frac{\partial^2 f}{\partial x \partial y} = f_{yx} = \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial y} \right), \quad \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial x} \right) \quad \text{etc.}$$

The last two partial derivatives cited are termed "mixed" second-order partial derivatives.

EXAMPLE. For  $f(x,y) = ye^{xy}$  we have

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x} (y^2 e^{xy}) = y^3 e^{xy}$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x} (e^{xy} + xye^{xy}) = 2ye^{xy} + xy^2 e^{xy}.$$

What are the others?

An extensive calculus of partial differentiation has been developed, from which only the following results will be necessary for our purpose.

The first of these concerns the order in which mixed derivatives are calculated. While a proof of the result is not difficult it requires machinery outside our immediate area of interest and so will not be given.

THEOREM. If  $\frac{\partial^2 f}{\partial x \partial y}$  and  $\frac{\partial^2 f}{\partial y \partial x}$  both exist and are continuous, then they are equal.

Hence under suitable conditions the order in which we perform a mixed differentiation is immaterial, as the following example illustrates.

EXAMPLE. If  $f(x,y) = ye^{xy}$  we previously found  $\frac{\partial^2 f}{\partial x \partial y} = 2ye^{xy} + xy^2 e^{xy}$ .

Now

$$\begin{aligned} \frac{\partial^2 f}{\partial y \partial x} &= \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial x} \right) = \frac{\partial}{\partial y} (y^2 e^{xy}) \\ &= xy^2 e^{xy} + 2ye^{xy} \\ &= \frac{\partial^2 f}{\partial x \partial y} \end{aligned}$$

Lest we become too casual in our use of this theorem, an example such as

$$g(x,y) = \begin{cases} \frac{x^3 y - xy^3}{x^2 + y^2} & \text{for } (x,y) \neq (0,0) \\ 0 & \text{for } (x,y) = (0,0) \end{cases}$$

for which  $g_{xy}(0,0) = -1$  while  $g_{yx}(0,0) = 1$ , should be borne in mind. It frequently happens that both  $x$  and  $y$  are functions of a single variable,  $t$  say. i.e.  $x \equiv x(t)$  and  $y \equiv y(t)$ . (This would have been the case when  $x$  was population size and  $y$  the number of diseased individuals present  $t$  representing time.) In such a case, if  $f$  is a function of two variables, then we can define a new function of one variable,  $f^*$  by

$$f^*(t) = f(x(t), y(t))$$

(Many books fail to distinguish between the two functions  $f$  and  $f^*$  writing  $f$  for both of them, this can however lead to confusion.)

For example if  $x(t) = e^t$  and  $y(t) = 1 - e^{-t}$  while  $f(x,y) = y(x-y)$  we should have  $f^*(t) = f(x(t), y(t)) = e^t + 2e^{-t} - e^{-2t} - 2$ . In such a situation the following result is often useful.

**THEOREM (The Chain Rule).** *If  $f$  has continuous first-order partial derivatives in the domain  $D$ , and  $x, y$  are differentiable functions with  $(x(t), y(t)) \in D$  for all  $t$ , then, if  $f^*(t) = f(x(t), y(t))$ ,*

$$\frac{df^*}{dt} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt}$$

**EXAMPLE.** With  $f(x,y) = y(x - y)$ ,  $x(t) = e^t$  and  $y(t) = 1 - e^{-t}$  we have

$$\begin{aligned} \frac{df^*}{dt} &= (y(t))(e^t) + (x(t) - 2y(t))(e^{-t}) \\ &= e^t - 1 + 1 - 2e^{-t} + 2e^{-2t} = e^t - 2e^{-t} + 2e^{-2t}. \end{aligned}$$

A result easily checked from the expression for  $f^*(t)$  determined above.

The proof of this result, although more tedious, follows precisely the same lines as the proof of the chain rule for a function of one variable and so will not be given here.

Since partial differentiation amounts to ordinary differentiation with one variable being regarded as a constant, the above theorem has the following immediate corollary.

**COROLLARY:** *Let  $f \equiv f(x,y)$ ,  $x \equiv x(u,v)$  and  $y \equiv y(u,v)$ , then we can define the function of two variables,  $f^*$  by*

$$f^*(u,v) = f(x(u,v), y(u,v))$$

for which

$$\frac{\partial f^*}{\partial u} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial u}$$

and

$$\frac{\partial f^*}{\partial v} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial v}$$

provided the derivatives on the right-hand side exist and are continuous.

A particularly useful version of this corollary results when  $f$  is only a function of the single variable  $x$ . In which case we have:

**COROLLARY.** *If  $f \equiv f(x)$  and  $x \equiv x(u,v)$ , then we can define  $f^*(u,v) = f(x(u,v))$  for which*

$$\frac{\partial f^*}{\partial u} = \frac{df}{dx} \frac{\partial x}{\partial u}$$

and 
$$\frac{\partial f^*}{\partial v} = \frac{df}{dx} \frac{\partial x}{\partial v}$$

provided  $f$  is a differentiable function and  $x$  has continuous first order partial derivatives. (Note: Since  $f$  is a function of only one variable there is no distinction between  $\frac{\partial f}{\partial x}$  and  $\frac{df}{dx}$ .)

EXAMPLE: Let  $f$  be a differentiable function. We seek to find an expression for  $\frac{\partial}{\partial x}(f(x-y))$ . Now if  $t \equiv t(x,y) = x-y$  we have  $f^*(x,y) = f(t) = f(x-y)$

and 
$$\frac{\partial f^*}{\partial x} = \frac{df}{dt} \frac{\partial t}{\partial x} = \frac{df}{dt} \times 1 = f'(t) = f'(x-y).$$

So, for example

$$\frac{\partial}{\partial x}(\sin(x-y)) = \cos(x-y)$$

similarly 
$$\frac{\partial}{\partial y}(\sin(x-y)) = -\cos(x-y).$$

*Collateral Reading:*

There is a great number of good books on 'advanced' calculus, almost any of which would provide an adequate reference for the material of this lecture. For example:

Kaplan and Lewis "Calculus and Linear Algebra", Wiley.

Osserman "Two-Dimensional Calculus".

Lang "Analysis I".

EXERCISES.

1. Find all the first and second-order partial derivatives of the following functions:

- |                                |                               |
|--------------------------------|-------------------------------|
| (i) $\phi(x,y) = \sqrt{xy}$    | (ii) $f(x,y) = \cos(ax + by)$ |
| (iii) $e^{ax+by}$              | (iv) $e^{-ax} \sin(x + by)$   |
| (v) $\psi(x,t) = e^{-(x+t)^2}$ | (vi) $g(x,t) = e^{-x^2/t}$    |

2. If  $f(u,v) = u^2 + 2uv + v^2$  and  $u \equiv u(x,y) = x - y$ ,  $v \equiv v(x,y) = x + y$ , form  $f^*(x,y) = f(u(x,y), v(x,y))$  and find  $f_x^*$ .

3. (a) Let  $q(x,y) = f(x-y)$  where  $f$  is any differentiable function, show that

$$\frac{\partial q}{\partial x} + \frac{\partial q}{\partial y} = 0.$$

(b) Let  $q \equiv q(x,y)$  be any function of two variables such that  $\frac{\partial q}{\partial x} + \frac{\partial q}{\partial y} = 0$ . Set  $x = \frac{1}{2}(u+v)$ ,  $y = \frac{1}{2}(u-v)$ , [i.e.  $u = x+y$ ,  $v = x-y$ ] and form  $q^*(u,v) = q(x(u,v), y(u,v))$ . Show  $\frac{\partial q^*}{\partial u} = 0$  and conclude  $q^*$  is a function of  $v$  alone and so  $q(x,y) = q^*(v(x,y)) = q^*(x-y)$ . Hence the only functions satisfying  $\frac{\partial q}{\partial x} + \frac{\partial q}{\partial y} = 0$  are of the form given in part (a).

Lecture 18 *Partial Differential Equations - The diffusion equation*

For the diffusion of a substance (or the conduction of heat) through a body, experiments show that the rate of transfer across a surface (be it a real surface, such as a porous membrane, or simply an imaginary surface drawn through the body) is proportional to both the area of the surface and the concentration\* gradient across it.\*\*

We now examine in detail the diffusion of a substance along a tube such as that illustrated in figure 1.

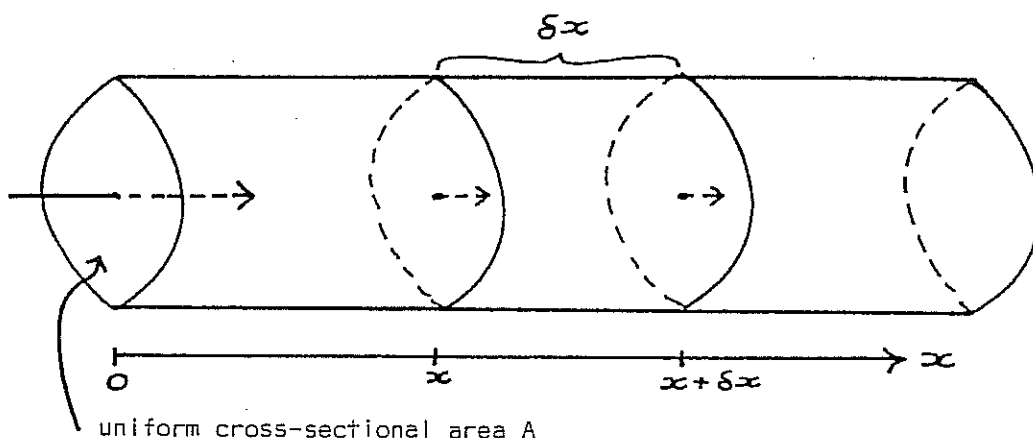


Figure 1

The concentration,  $c$ , of the diffusing substance, at a point distance  $x$  along the tube, at time  $t$  is a function of both these parameters i.e.  $c \equiv c(x,t)$ .

The rate at which substance is transferred into the element of tube between  $x$  and  $x + \delta x$  across the left-hand face at time  $t$  is, from the above remarks, proportional to both  $A$  and the concentration gradient at  $x$ ,  $\frac{\partial c}{\partial x}(x,t)$ , and so equals

$$- DA \frac{\partial c}{\partial x}(x,t),$$

where  $D > 0$  is a constant of proportionality, known as the coefficient of diffusion, and the minus sign indicates that transfer occurs from regions of higher concentration to those of lower concentration.

Similarly the rate at which substance is leaving the element through the right-hand face is

$$- DA \frac{\partial c}{\partial x}(x + \delta x, t)$$

\* In the case of heat flow, read temperature for concentration.

\*\* This conclusion is also confirmed by statistical mechanical arguments based on molecular models of the phenomenon.

The difference between these two rates

$$DA \left( \frac{\partial c}{\partial x} (x + \delta x, t) - \frac{\partial c}{\partial x} (x, t) \right)$$

is the rate at which substance is accumulating within the element and so equals  $\frac{\partial M}{\partial t}$  where M is the amount of substance in the element at time t. Now, provided  $\delta x$  is small enough so that we may take the concentration throughout the element to be approximately equal to  $c(x, t)$ , we have

$$M = c(x, t) \times \text{volume of element} = c(x, t)A\delta x$$

and so

$$\frac{\partial M}{\partial t} = A\delta x \frac{\partial c}{\partial t} (x, t).$$

Therefore

$$DA \left( \frac{\partial c}{\partial x} (x + \delta x, t) - \frac{\partial c}{\partial x} (x, t) \right) = A\delta x \frac{\partial c}{\partial t} (x, t)$$

or

$$D \left\{ \frac{\frac{\partial c}{\partial x} (x + \delta x, t) - \frac{\partial c}{\partial x} (x, t)}{\delta x} \right\} = \frac{\partial c}{\partial t} (x, t)$$

The approximation improving the smaller  $\delta x$ .

Taking the limit  $\delta x \rightarrow 0$ , the quantity in brackets on the left will be recognised as  $\frac{\partial}{\partial x} \left( \frac{\partial c}{\partial x} \right) (x, t)$ , from the definition of  $\frac{\partial}{\partial x}$ , and so we have

$$D \frac{\partial^2 c}{\partial x^2} = \frac{\partial c}{\partial t}.$$

An equation of this form, expressing as it does a relationship between an 'unknown' function,  $c$ , and some of its partial derivatives, is a *partial differential equation*, in this case the *diffusion (or heat) equation*.

Diffusion type processes occur in diverse areas of biology. As the heat equation it is of importance in physiology when the thermodynamics of organisms is under consideration. Since many chemical reactions involve interaction by diffusion, the equation finds frequent use in biochemistry and in particular in cell chemistry, where diffusion provides an efficient transport mechanism.\* Apart from these more obvious areas of application, the transfer of population from one region to another often satisfies the basic assumptions for a diffusion process and so the diffusion equation has application in certain ecological models. It has also been used by W. Feller and others in modelling population genetics and evolution.

Later on in these lectures we will outline a general technique for the solution of equations such as the diffusion equation. For the moment however, we will content ourselves by examining some properties of the solution which can be deduced directly from the partial differential equation itself and by identifying a few special solutions.

---

\* For larger organisms, diffusion is too slow and inefficient to provide an adequate transport mechanism, hence the need for circulatory systems. Indeed it is the absence of such circulatory systems and the consequent reliance on diffusion which limits the size of organisms such as the flat worm.

I. *Steady State Solutions*

For any partial differential equation involving time, a solution which does not change with time is termed a steady state solution. If  $\phi$  is any such solution, then by assumption  $\frac{\partial \phi}{\partial t} \equiv 0$  and so  $\frac{\partial^2 \phi}{\partial t \partial x}$  etc. are also all identically zero.

Thus in the case of the diffusion equation  $c$  is a steady state solution if  $\frac{\partial c}{\partial t} \equiv 0$  and so  $c$  is a function of  $x$  only in which case  $\frac{\partial c}{\partial x} = \frac{dc}{dx}$  and similarly  $\frac{\partial^2 c}{\partial x^2} = \frac{d^2c}{dx^2}$  thus the diffusion equation 'degenerates' into the ordinary differential equation

$$\frac{d^2c}{dx^2} = 0$$

Any solution of this is a steady state solution of the diffusion equation and conversely any steady state solution of the diffusion equation must satisfy it.

Now the general solution of  $\frac{d^2c}{dx^2} = 0$  is easily seen to be  $c(x) = ax + b$  ( $a, b$  constants). So we conclude that the concentration varies linearly along the tube when the system is in a steady state.

Consider the situation of a porous tube of length  $l$ , the two ends of which are connected to reservoirs held at constant concentrations of  $c_1$  and  $c_2$  respectively (see figure 2 (a)). When such a system is in a steady state, the concentration,  $c(x)$ , at a distance  $x$ , along the tube, from the end at concentration  $c_0$  will be  $c(x) = ax + b$ .

Further  $c(0) = c_0$  while  $c(l) = c_1$  and so we have

$$c(x) = \frac{c_1 - c_0}{l} x + c_0$$

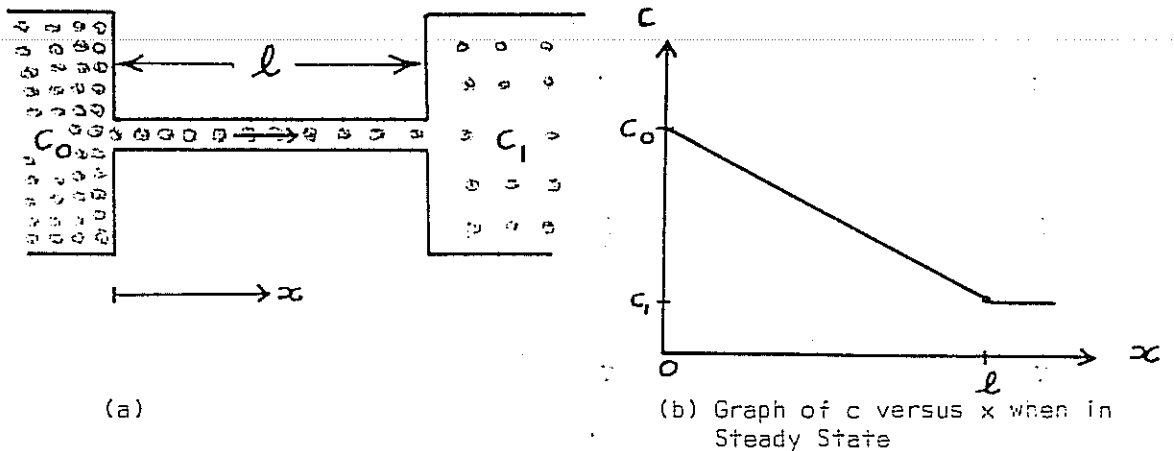


Figure 2

II. *Approach to the Steady State*

The steady state solutions of a partial differential equation are important as potential equilibrium states of the equation. (Playing much the same role as critical points do for autonomous systems of equations, clearly every equilibrium state is a steady state). That the linear steady state solutions of the diffusion equation found in I. do indeed represent asymptotically stable equilibrium states toward which a solution will tend

as  $t \rightarrow \infty$  is made plausible by the following considerations. For definiteness consider the situation described in I. where the two ends of the tube are maintained at constant concentrations,  $c_0$  and  $c_1$  and let  $c_s$ ,  $c_s(x) = \left(\frac{c_1 - c_0}{l}\right)x + c_0$  be the steady state solution.

If  $c = c(x,t)$  is any other solutions, for any value of time  $t_0$ , let  $x_1 < x_2$  be two consecutive points at which  $c$  cuts the steady state solution., i.e.  $c(x_i, t_0) = c_s(x_i)$  (for  $i = 1, 2$ ), then either  $c(x, t_0) > c_s(x)$  or  $c(x, t_0) < c_s(x)$  for all  $x$  between  $x_1$  and  $x_2$ . If  $c(x, t_0) < c_s(x)$  then the curve  $c = c(x, t_0)$  is convex downward (see figure 2) therefore  $\frac{\partial^2 c}{\partial x^2} > 0$ , in which case  $\frac{\partial c}{\partial t}(x, t_0) = D \frac{\partial^2 c}{\partial x^2}(x, t_0) > 0$  and, at  $t = t_0$ ,  $c(x, t)$  is increasing with time and so approaching the steady state solution. Similarly if  $c(x, t_0) > c_s(x)$  we have  $\frac{\partial^2 c}{\partial x^2} < 0$  and at  $t = t_0$   $c(x, t)$  is decreasing with time toward the steady state solution.

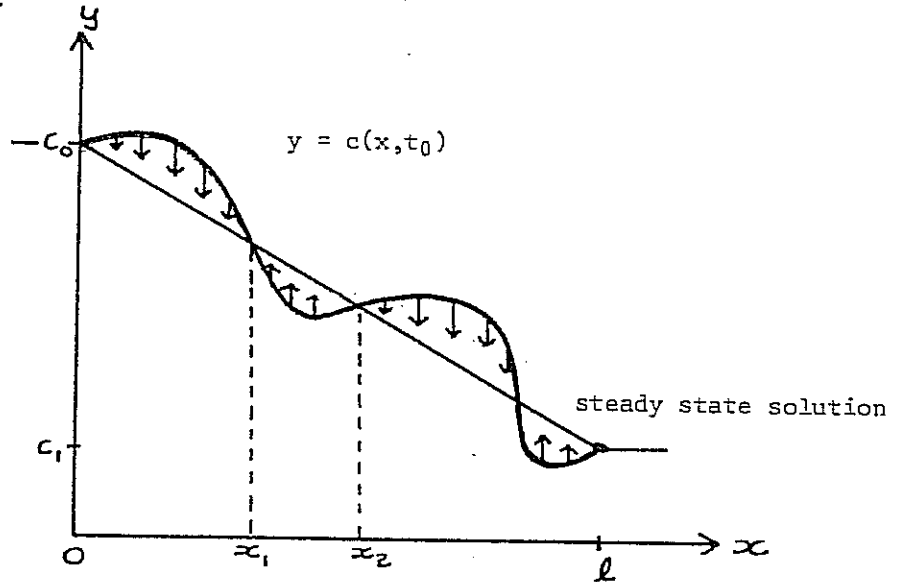


Figure 3

III. *Some simple solutions of the diffusion equation.*

Perhaps the simplest dynamic solution of the diffusion equation  $D \frac{\partial^2 c}{\partial x^2} = \frac{\partial c}{\partial t}$  is of the form  $c(x,t) = c_0 e^{ax + Da^2 t}$  where  $c_0$  and  $a$  are arbitrary constants. That this is indeed a solution is readily confirmed by direct substitution, hence if

$$c(x,t) = c_0 e^{ax + Da^2 t} \text{ we have}$$

$$\frac{\partial c}{\partial t} = Da^2 c_0 e^{ax + Da^2 t} \text{ while}$$

$$\frac{\partial^2 c}{\partial x^2} = \frac{\partial}{\partial x} \left[ a c_0 e^{ax + Da^2 t} \right] = a^2 c_0 e^{ax + Da^2 t}$$

and so 
$$\frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial x^2} \text{ as required.}$$

Another solution (which, from Euler's formula, arises as a special case:  $a = i\beta$ ,  $\beta$  real, of the above solution) is

$$c(x,t) = c_0 e^{-D\beta^2 t} \cos \beta x$$

Yet another solution is  $c(x,t) = c_0 \sqrt{t} e^{-x^2/4Dt}$ .

Graphs of  $c$  versus  $x$  at various values of  $t$  for these solutions are given in figure 4.

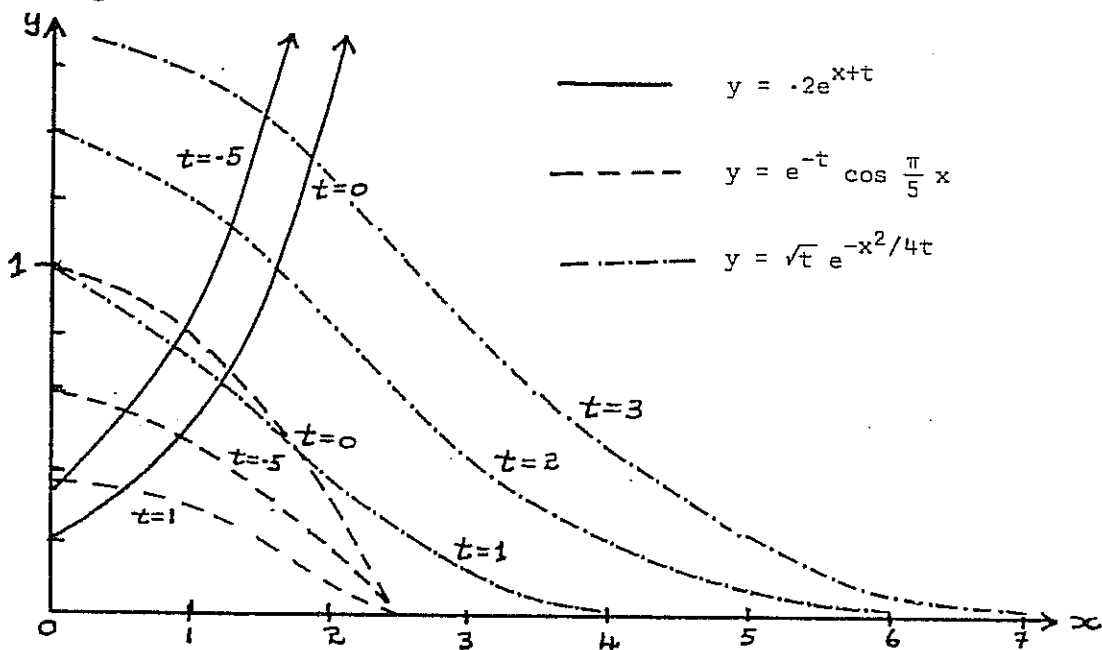


Figure 4  $y = c(x,t)$

*Collateral Reading.*

A discussion of partial differential equations at about this level, may be found in

Batschelet "Introduction to Mathematics for Life Scientists"

and

J. Maynard-Smith "Mathematical Ideals in Biology".

Steady state solutions are discussed in some detail in the second of these two books.

For a more mathematical discussion of the diffusion equation see

Boyce and DiPrima "Elementary Differential Equations and Boundary Value Problems".

EXERCISES.

1. If a narrow straight tube of length 10 cms connects two reservoirs maintained at fixed concentrations of 100 gms sugar/l and 4 gms sugar/l respectively, what is the steady state distribution of concentration along the tube?
2. Show that  $c(x,t) = c_0 e^{-DB^2t} \sin Bx$  and  $c(x,t) = t^{1/2} e^{-x^2/4Dt}$  are solutions of the diffusion equation  $D \frac{\partial^2 c}{\partial x^2} = \frac{\partial c}{\partial t}$ .
3. For diffusion radially outwards from the axis of a cylinder the diffusion equation is  $\frac{D}{r} \frac{\partial}{\partial r} (r \frac{\partial \phi}{\partial r}) = \frac{\partial \phi}{\partial t}$  where  $\phi$  is the concentration at time  $t$  a distance  $r$  radially out from the axis. However, for diffusion radially outwards from the centre of a sphere the diffusion equation becomes



3. (continued)

$$D \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial \phi}{\partial r} \right) = \frac{\partial \phi}{\partial t}$$

(can you think why the equation would be different in these cases?). Find the steady state solution in these cases. These could be used to find the distribution of concentration inside a cylindrical or a spherical cell at equilibrium.

4. If a constant amount of heat is absorbed at the skin's surface per unit time, then the temperature  $T$  at a distance  $x$  beneath the skin's surface at time  $t$  is described by the partial differential equation

$$\frac{\partial T}{\partial t} = a^2 \frac{\partial^2 T}{\partial x^2} \text{ (where } a^2 \text{ is a constant)}$$

If, however, heat is carried into the skin by penetrating radiation in such a way that the heat not absorbed at a depth  $x$  is  $H = H_0 e^{-rx}$  derive a partial differential equation to describe the variation of temperature  $T$  with  $x$  and  $t$ .

Note: For our purpose  $T$  = Heat/unit volume.

(Hint: The heat adsorbed in a layer of skin at depth  $x$  and of thickness  $\delta x$  will be  $\delta H \doteq \frac{dH}{dx} \delta x$ .)

Adapted from: Buettner, Konrad "Temperature changes of skin".

J. of Applied Physiology V3, No. 12, 1951, pp. 691-702.

Lecture 19 Propagation Equations

1. The Wave Equation

We investigate the propagation of a disturbance of constant 'spatial' profile, such as a sound wave travelling down the ear canal, a mechanical vibration along the basilar membrane or the passage of an impulse along a neural pathway.

Consider a 'wave' of constant profile, travelling from left to right as illustrated in figure 1, where the amplitude at time  $t$  a distance  $x$  from 0 is  $y \equiv y(x,t)$ .

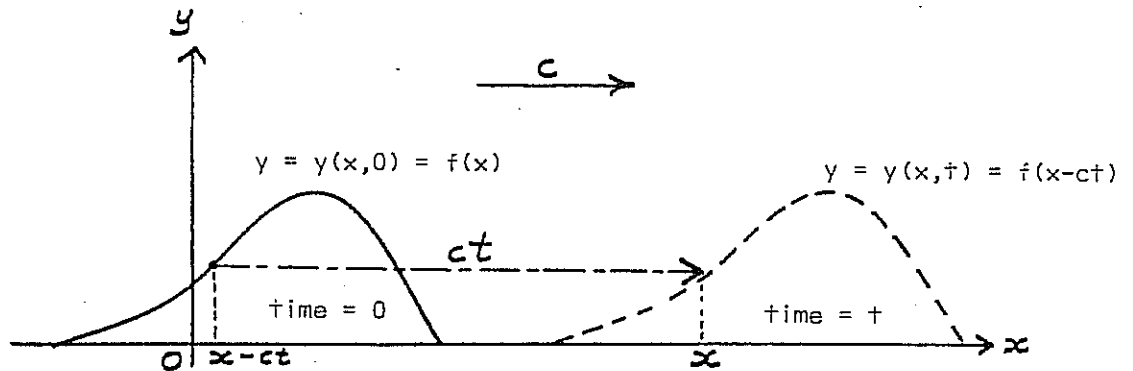


Figure 1

Let  $f(x) = y(x,0)$  specify the profile of the wave at time  $t = 0$ , then after a time  $t$  has elapsed the wave will have travelled a distance  $ct$  to the right, where  $c$  is the speed of propagation. So  $y(x,t)$  will be the same as the amplitude at time 0 for the point  $x - ct$  i.e.  $y(x,t) = f(x - ct)$ . This is the general form for a wave travelling from left to right. Similarly, a wave travelling from right to left would have amplitude

$$y(x,t) = g(x + ct)$$

where  $y = g(x)$  describes the profile of the wave at time 0.

The most general form of a propagating wave, results from the superposition of two such waves, one travelling left to right and the other right to left, i.e.

$$y(x,t) = f(x - ct) + g(x + ct)$$

where  $f$  and  $g$  are two arbitrary functions representing the initial profile of each of the component waves.

We seek a partial differential equation which is satisfied by such functions of two variables.

Thus, let  $u = x - ct$        $v = x + ct$ , and

$$y(u,v) = f(u) + g(v),$$

then, by the Chain Rule given on page 93,

$$\frac{\partial y}{\partial t} = \frac{\partial y}{\partial u} \frac{\partial u}{\partial t} + \frac{\partial y}{\partial v} \frac{\partial v}{\partial t} = -c \frac{df}{du} + c \frac{dg}{dv},$$

similarly

$$\frac{\partial^2 y}{\partial t^2} = c^2 \left( \frac{d^2 f}{du^2} + \frac{d^2 g}{dv^2} \right)$$

But 
$$\frac{\partial y}{\partial x} = \frac{\partial y^*}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial y^*}{\partial v} \frac{\partial v}{\partial x} = \frac{df}{du} + \frac{dg}{dv}$$

and so 
$$\frac{\partial^2 y}{\partial x^2} = \frac{\partial}{\partial u} \left( \frac{\partial y}{\partial x} \right)^* \frac{\partial u}{\partial x} + \frac{\partial}{\partial v} \left( \frac{\partial y}{\partial x} \right)^* \frac{\partial v}{\partial x} = \frac{d^2 f}{du^2} + \frac{d^2 g}{dv^2}$$

We see therefore that

$$\frac{\partial^2 y}{\partial x^2} = \frac{d^2 f}{du^2} + \frac{d^2 g}{dv^2} = \frac{1}{c^2} \frac{\partial^2 y}{\partial t^2}$$

and we are led to the *wave equation*

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 y}{\partial t^2}$$

which is satisfied by any function of the form

$$y(x,t) = f(x - ct) + g(x + ct)$$

i.e. by any function representing the propagation of a wave. We now show that this is the only form of function which satisfies the wave equation.

\* Thus, let  $y \equiv y(x,t)$  satisfy  $\frac{\partial^2 y}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 y}{\partial t^2}$  and as before set  $u = x + ct$ ,  $v = x - ct$  then  $y^*(u,v) \equiv y(x(u,v), t(u,v))$  where  $x = \frac{1}{2}(u + v)$ ,  $t = \frac{1}{2c}(v - u)$

Therefore 
$$\frac{\partial y^*}{\partial u} = \frac{\partial y}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial y}{\partial t} \frac{\partial t}{\partial u}$$

$$= \frac{1}{2} \frac{\partial y}{\partial x} - \frac{1}{2c} \frac{\partial y}{\partial t}$$

So 
$$\frac{\partial y^*}{\partial v \partial u} = \frac{\partial}{\partial v} \left( \frac{\partial y^*}{\partial u} \right) = \frac{\partial}{\partial x} \left( \frac{\partial y}{\partial u} \right) \frac{\partial x}{\partial v} + \frac{\partial}{\partial t} \left( \frac{\partial y}{\partial u} \right) \frac{\partial t}{\partial v}$$

$$= \frac{1}{2} \left( \frac{1}{2} \frac{\partial^2 y}{\partial x^2} - \frac{1}{2c} \frac{\partial^2 y}{\partial x \partial t} \right) + \frac{1}{2c} \left( \frac{\partial^2 y}{\partial t \partial x} - \frac{1}{2c} \frac{\partial^2 y}{\partial t^2} \right)$$

$$= \frac{1}{4} \left( \frac{\partial^2 y}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 y}{\partial t^2} \right) + \frac{1}{4c} \left( \frac{\partial^2 y}{\partial x \partial t} - \frac{\partial^2 y}{\partial t \partial x} \right)$$

$$= 0$$

The first term in brackets being zero by assumption and the second by the theorem on page 92, provided both  $\frac{\partial^2 y}{\partial x \partial t}$  and  $\frac{\partial^2 y}{\partial t \partial x}$  are continuous.

\* Thus, since  $\frac{\partial}{\partial v} \left( \frac{\partial y^*}{\partial u} \right) = 0$  we see that  $\frac{\partial y^*}{\partial u}$  is a function of  $u$  only i.e.  $\frac{\partial y}{\partial u} = f'(u)$  and so  $y^* = y(u,v)$  must be of the form  $y^* = f(u) + g(v)$  for some functions  $f$  and  $g$ , whence  $y = f(x - ct) + g(x + ct)$  as required.

It is important to note that the speed,  $c$ , of propagation of the wave appears explicitly in the wave equation.

For a uniform, tensioned string displaced slightly from its equilibrium position, the equations of classical mechanics lead to the equation of motion

$$\frac{\partial^2 y}{\partial x^2} = \frac{T}{\rho} \frac{\partial^2 y}{\partial t^2}$$

where  $y \equiv y(x,t)$  is the transverse displacement of the point at position  $x$  along the string at time  $t$ ,  $\rho$  is the mass per unit length of the string and  $T$  is the (longitudinal) tension in the string. (See appendix.)

From our arguments above we are therefore able to conclude that the motion of such a string consists of waves travelling up and down it. For this reason the wave equation plays a basic role in the theory of musical instruments and the physiology of hearing.

Various solutions to the wave equation are obtained by assigning  $f$  and  $g$  to be particular functions. Perhaps the best known example is

$$f(x) = A \sin(x) \text{ and } g \equiv 0 \text{ in which case}$$

$$y(x,t) = A \sin(x - ct) \text{ where } A > 0 \text{ is the (maximum)}$$

amplitude of the wave.

$$\text{Other solutions are: } y(x,t) = A \sin(x - ct) + B \sin(x + ct);$$

$$y(x,t) = e^{x} e^{ct}, \text{ etc.}$$

2. *The Telegraph Equation.* (So named because of its use to describe the transmission of signals by electric telegraph.)

In a mechanical situation the term  $\frac{\partial^2 y}{\partial t^2}$  in the wave equation represents the acceleration of the point with amplitude  $y(x,t)$  and so equals the force acting on a unit mass at that point and time. We see therefore that for such a mechanical system, the force applied by the system to a unit mass at the above point and time is  $c^2 \frac{\partial^2 y}{\partial x^2}$ . Frequently however an additional damping force acts on each of the particles in motion. It is hydrodynamically justifiable, and also usual, to take such a damping force proportional to the particle's velocity,  $\frac{\partial y}{\partial t}$ , but oppositely directed. In such a case we would have

$$\frac{\partial^2 y}{\partial t^2} = \{\text{force acting on unit mass at } (x,y(x,t))\}$$

$$= c^2 \frac{\partial^2 y}{\partial x^2} + \{\text{damping force}\}$$

$$= c^2 \frac{\partial^2 y}{\partial x^2} - k \frac{\partial y}{\partial t},$$

where  $k > 0$  is a proportionality constant, known as the coefficient of damping.

We are thus led to the telegraph equation

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{c^2} \left\{ \frac{\partial^2 y}{\partial t^2} + k \frac{\partial y}{\partial t} \right\}$$

Apart from describing the propagation of a mechanical disturbance in the presence of damping, this equation has also been used to describe the flow of fluid along a tube when viscosity is taken into account and so could be applied to arterial blood flow.

In the case of electrical signals the 'damping' term allows for the resistance of the conducting medium. In this context the telegraph equation played an important role in the Nobel prize winning theory of nerve impulse conduction along an axon, developed by Hodgkin and Huxley.

Previously, decay phenomena were found to take place according to a rule of the form  $e^{-\alpha t}$ . Since the presence of damping would be expected to lead to a decay in the amplitude of any disturbance propagating according to the telegraph equation, we will investigate the possibility of solutions with the form

$$y(x,t) = e^{-\alpha t} f(x - ct).$$

Upon substitution into the equation we require

$$\begin{aligned} e^{-\alpha t} \frac{\partial^2 f}{\partial x^2} &= \frac{1}{c^2} \left\{ \frac{\partial}{\partial t} \left( -\alpha e^{-\alpha t} f + e^{-\alpha t} \frac{\partial f}{\partial t} \right) + k \left( -\alpha e^{-\alpha t} f + e^{-\alpha t} \frac{\partial f}{\partial t} \right) \right\} \\ &= \frac{1}{c^2} \left\{ \alpha^2 e^{-\alpha t} f - 2\alpha e^{-\alpha t} \frac{\partial f}{\partial t} + e^{-\alpha t} \frac{\partial^2 f}{\partial t^2} - k\alpha e^{-\alpha t} f + k e^{-\alpha t} \frac{\partial f}{\partial t} \right\} \\ &= \frac{1}{c^2} \left\{ (\alpha^2 - k\alpha) e^{-\alpha t} f + (k - 2\alpha) e^{-\alpha t} \frac{\partial f}{\partial t} + e^{-\alpha t} \frac{\partial^2 f}{\partial t^2} \right\}, \end{aligned}$$

a relationship which is certainly true if

$$\frac{\partial^2 f}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 f}{\partial t^2} + \frac{1}{c^2} (k - 2\alpha) \frac{\partial f}{\partial t} + \frac{1}{c^2} (\alpha^2 - k\alpha) f$$

Now from the form of  $f$ ,  $f \equiv f(x - ct)$ , and part 1. of this lecture we have

$$\frac{\partial^2 f}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 f}{\partial t^2},$$

whence our requirement is that

$$(k - 2\alpha) \frac{\partial f}{\partial t} + (\alpha^2 - k\alpha) f = 0$$

where  $\alpha$  is as yet unspecified and so may be selected appropriately.

In general this relationship cannot be satisfied. However, if we choose  $\alpha$  to be  $\frac{k}{2}$  and assume the damping is so small that the term  $\frac{k^2}{4} f$  may be neglected, then the relationship is satisfied approximately. So, provided the coefficient of damping,  $k$ , is small,

$$y(x,t) = e^{-\frac{k}{2} t} f(x - ct),$$

for any bounded twice differentiable function  $f$ , is an 'approximate' solution of the telegraph equation

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{c^2} \left\{ \frac{\partial^2 y}{\partial t^2} + k \frac{\partial y}{\partial t} \right\}.$$

Thus one possible 'solution' would be

$$y(x,t) = e^{-\frac{k}{2} t} \sin(x - ct).$$

To determine how  $y(x,t)$  varies with  $x$  at any fixed time  $t$  we may proceed as follows.

Let  $y(x,t) = e^{-\frac{k}{2} t} f(x - ct)$ , then since  $f$  is an arbitrary function we can write it as

$$f(z) = e^{-kz/2c} g(z)$$

where  $g$  is again an arbitrary function. Substituting this into the expression for  $y(x,t)$  we then have

$$y(x,t) = e^{-\frac{k}{2c}x} g(x - ct)$$

a form for  $y$  identical to that previously given except that  $\frac{x}{c}$  replaces  $t$  in the exponential 'damping' factor.

So at any instant of time the solution of the telegraph equation is modulated by a factor decreasing exponentially with  $x$ .

A more generally applicable method, whereby exact solutions to the telegraph equation may be found, is presented in the next lecture.

*Collateral Reading.*

A more complete discussion of propagation phenomena from the mathematical point of view is to be found in the monograph

Coulson "Waves"

The application of the wave equation to the transmission of neural impulses and the development from this of a theory of delayed reflexes may be found in

Rashevsky "Mathematical Biophysics, Physico & Mathematical Foundations of Biology".

EXERCISES:

1. If a wave travelling from left to right at 20 cms/sec has initial shape  $y = e^{-x^2}$  find its shape after 10 secs.
2. Show that: (i)  $\phi = A \sin(x - ct) + B \cos(x - ct)$   
and (ii)  $\phi = f(x - ct) + g(x + ct)$  (for any  $f$  and  $g$ )  
are solutions of the wave equation  $\frac{\partial^2 \phi}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2}$ .
3. The energy of a wave is proportional to its amplitude squared. For a wave travelling according to the telegraph equation,

$$\frac{\partial^2 \phi}{\partial x^2} = \frac{1}{c^2} \left\{ \frac{\partial^2 \phi}{\partial t^2} + k \frac{\partial \phi}{\partial t} \right\},$$

after what time will its energy have been reduced by damping to a quarter of its initial value?

APPENDIX

*Equation of Motion of a Vibrating String*

Consider a horizontal, taut, uniform string clamped at each end and 'plucked' to assume some initial shape before being released. We seek to

determine the equation governing its subsequent motion, under the assumption that each point on the string moves only vertically up and down (i.e. the string is perfectly elastic), if we let the vertical displacement at time  $t$  of the point, whose horizontal distance from one end of the string is  $x$ , be  $\phi(x,t)$ , then the shape of the string at time  $t$  is  $y = \phi(x,t)$ .

Differential equation of Vibration: (D'Alembert), neglecting gravity, drag due to air resistance, etc.

By assumption, each point on the string has no horizontal motion, so the horizontal component  $H$  of the string's tension  $T$  is the same for each point, i.e.  $\frac{\partial H}{\partial x} = 0$  and so  $H \equiv H(t)$ .

Now, if  $\rho = \rho(x,t)$  is the mass per unit length of string, Newton's second law implies:

$$\Delta x \rho \frac{\partial^2 \phi}{\partial t^2} = \Delta V = \Delta(H \tan \theta)$$

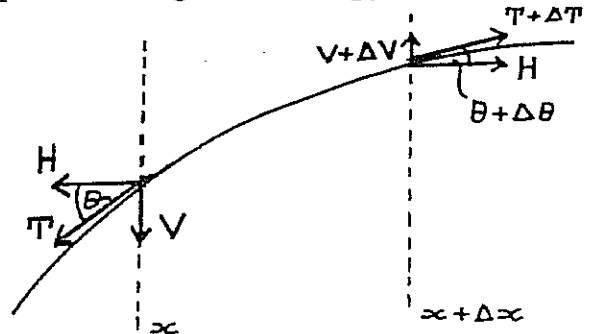
$$= H \Delta(\tan \theta)$$

but  $\tan \theta = \text{slope of string} = \frac{\partial \phi}{\partial x}$

so

$$\rho \frac{\partial^2 \phi}{\partial t^2} = H \frac{\Delta}{\Delta x} \frac{\partial \phi}{\partial x} \text{ and so taking the limit as } \Delta x \rightarrow 0 \text{ gives}$$

$$\frac{\partial^2 \phi}{\partial x^2} = \frac{\rho}{H} \frac{\partial^2 \phi}{\partial t^2}$$



Further, for small displacements, it can be proved that  $\sqrt{\frac{H}{\rho}} = c$ , a constant equal to the velocity of propagation of a wave along the string. (see Resnick and Haleday p. 402. Check for yourself that  $\text{dim} [H/\rho] = \frac{ML}{T^2} / \frac{M}{L} = \left(\frac{L}{T}\right)^2 = [\text{vel}]^2$ ) Therefore the equation of motion for the string is the wave equation

$$\frac{\partial^2 \phi}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2}$$

Lecture 20 *A more general second order partial differential equation, morphogenesis, separation of variables*

We begin with a discussion of the so-called *cable equation*\*

$$a^2 \frac{\partial^2 \phi}{\partial t^2} - \frac{\partial^2 \phi}{\partial x^2} = k\phi + D \frac{\partial \phi}{\partial t}$$

which represents a combination of all the processes discussed in the last two lectures and some other effects as well.

The equation can best be interpreted as follows.

Consider a substance infusing through some material by a variety of simultaneous processes. If  $\phi(x,t)$  denotes the concentration of the substance at position  $x$  and time  $t$ , then the equation represents the variation of  $\phi$  with respect to both  $x$  and  $t$  due to a number of combined causes:

- (i) Diffusion, accounted for by the presence of  $\frac{\partial^2 \phi}{\partial x^2}$  and  $\frac{\partial \phi}{\partial t}$ ;
- (ii) The passage of compression waves through the substance and/or surrounding material, allowed for by the terms  $\frac{\partial^2 \phi}{\partial x^2}$  and  $\frac{\partial^2 \phi}{\partial t^2}$ ;
- (iii) The decomposition of the substance into others through chemical reactions etc., this being assumed to take place at a rate proportional to the concentration and is represented by  $\frac{\partial \phi}{\partial t}$  and  $\phi$ .

Should there be two or more mutually interacting chemical substances present the above equation must be replaced by a system of two or more simultaneous partial differential equations. We will not consider this case, the interested reader being referred to the discussion in Maynard Smith's *Mathematical Ideas in Biology*. A complete solution of the cable equation will obviously be difficult. Before discussing techniques by which such a solution can be found, we will make some simple observations about the nature of the solution in certain special cases.

I. In case  $a = 0$  (i.e. (ii) above is absent) the equation becomes

$$\frac{\partial^2 \phi}{\partial x^2} + k\phi = -D \frac{\partial \phi}{\partial t}$$

If  $k = 0$  and  $D < 0$ , this is simply the diffusion equation. We will assume that  $D < 0$ ,  $k > 0$ , and consider the situation when  $y = \phi(x,0)$  has the form indicated in figure 1(a). For such a case  $\frac{\partial \phi}{\partial x}(x,0)$  has the form shown in figure 1(b) and so  $\frac{\partial^2 \phi}{\partial x^2}(x,0)$  is as illustrated in figure 1(c). But then  $\frac{\partial \phi}{\partial t} = -\frac{1}{D}(\frac{\partial^2 \phi}{\partial x^2} + k\phi)$  may appear as in figure 1(d) whence  $\phi$  is varying with  $t$  at the instant  $t = 0$  in the way shown in figure 1(e) and so at a slightly later moment,  $t = \tau$   $y = \phi(x,t)$  will have assumed the form illustrated in figure 1(f). Now, this same reasoning can be applied at each of the 'humps' in figure 1(f) to show that after a further passage of time  $y = \phi(x,t)$  will have developed into the form sketched in figure 1(g).

---

\* So named because of its original use to describe the transmission of signals via the first transatlantic cable, and its subsequent use in describing the operation of any co-axial cable.



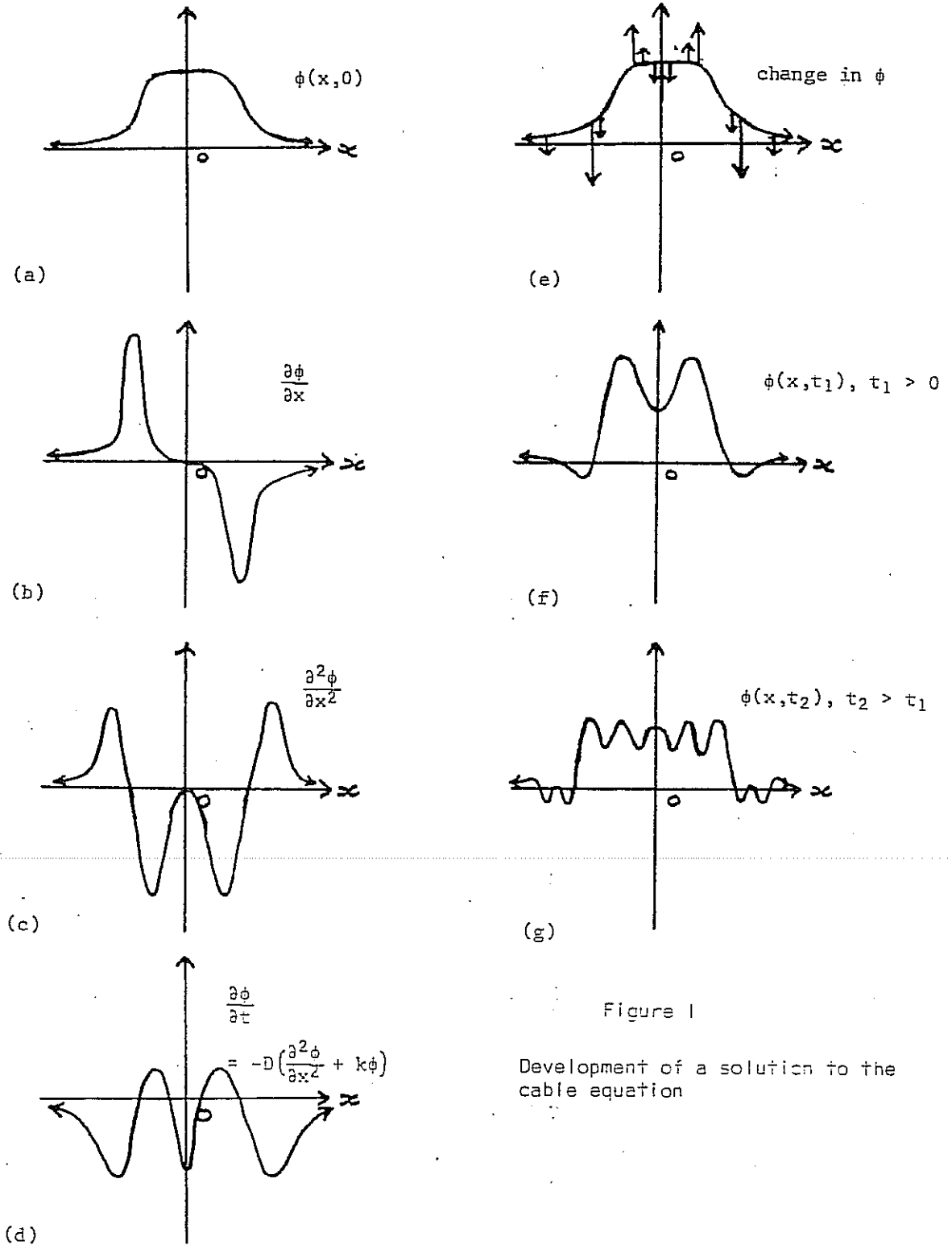


Figure 1

Development of a solution to the cable equation

Thus we are led to expect that one possible form of solution (resulting from the initial concentration distribution given in figure 1(a)) represents a wave-like disturbance which gradually spreads outward from  $x = 0$  in both directions.

To see this more clearly let us examine the steady state solution.

II. At the steady state  $\frac{\partial \phi}{\partial t} = 0$  and so we have  $\phi \equiv \phi(x)$  where

$$\frac{d^2 \phi}{dx^2} + k\phi = 0$$

which, provided  $k > 0$ , has the solution

$$\phi(x) = A \sin \sqrt{k} x + B \cos \sqrt{k} x$$

and so one possible steady state solution is that illustrated in figure 2.

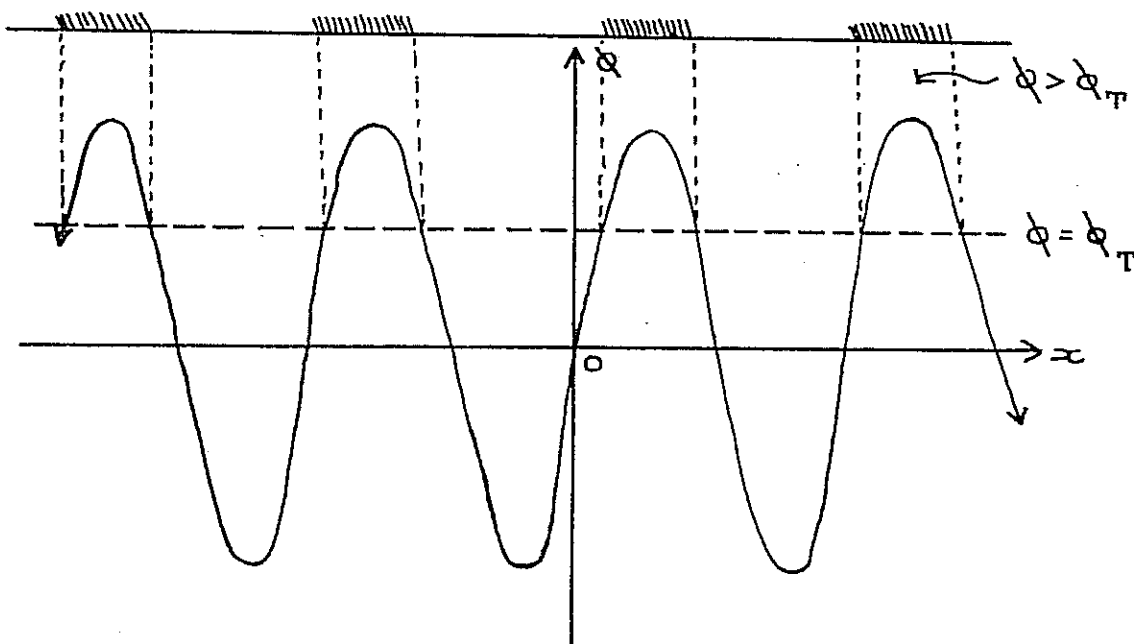


Figure 2

It is both important and perhaps somewhat surprising, that the equation considered in (I), which essentially represents a diffusion plus decay type process (the wave equation term  $\frac{\partial^2 \phi}{\partial x^2}$  being absent), should have wave-like solutions of this type.

It is also worth noting that the steady state solutions of the cable equation itself must satisfy

$$\frac{\partial^2 \phi}{\partial x^2} + k\phi = 0 \quad (\text{since } \frac{\partial^2 \phi}{\partial t^2} \equiv 0 \text{ at the steady state})$$

and so are the same as those given above.

The existence of regularly varying (periodic) steady state solutions to the cable equation, such as  $\phi(x) = A \cos \sqrt{k} x$ , has been put forward as a possible mechanism for *morphogenesis* (the development of regular patterns, such as the distribution of feather papillae, hair follicles, stomata in plants etc.).

In this case the equation is taken to describe the variation in concentration of some chemical across a strip of tissue. It being assumed that when the concentration exceeds some minimum threshold level,  $\phi_T$  say, it has the effect of inducing the cells to differentiate in some appropriate way. When operating in a steady state, this will result in the development of a regularly spaced structure as illustrated by figure 2.

With appropriate choices for the constants  $a$ ,  $k$  and  $D$  the cable equation was basic in the modified theory of electrically active cells developed by the late Prof. M. Lieberstein. This theory improves the original one of Hodgkin and Huxley (see page 103) by allowing for self induction which has the effect of 'stabilizing' the conduction of signals along the nerve axon.

Another particular form of the cable equation ( $a = 0$ ;  $D = i\hbar$ , an imaginary constant) is the now famous Schrödinger wave equation of modern quantum physics which is basic to much of molecular biochemistry.

### SEPARATION OF VARIABLES

#### *A General Method of solving some partial differential equations*

The method hinges on the observation that if  $\phi(x,t)$  can be "factorised" as  $\phi(x,t) = X(x)T(t)$  where  $X$  is a function of  $x$  only and  $T$  is a function of  $t$  only, then

$$\frac{\partial \phi}{\partial t} = X(x) \frac{dT}{dt}; \quad \frac{\partial \phi}{\partial x} = T \frac{dX}{dx}; \quad \frac{\partial^2 \phi}{\partial t^2} = X \frac{d^2 T}{dt^2} \quad \text{etc.}$$

[For example  $\phi(x,t) = e^{ax+bt}$  equals  $X(x)T(t)$  where  $X(x) = e^{ax}$  and  $T(t) = e^{bt}$ .]

For clarity we illustrate the method by applying it to the cable equation, (although it is in fact applicable to all the other partial differential equations we have considered, and also many others),

$$a^2 \frac{\partial^2 \phi}{\partial t^2} - \frac{\partial^2 \phi}{\partial x^2} = k\phi + D \frac{\partial \phi}{\partial t}$$

We first attempt to find solutions to our partial differential equation which can be factorised as  $X(x)T(t)$ .

- (i) Assuming  $\phi(x,t) = X(x)T(t)$  and using the above observations, rewrite the equation as

$$a^2 X \frac{d^2 T}{dt^2} - T \frac{d^2 X}{dx^2} = kXT + DX \frac{dT}{dt}$$

- (ii) Now attempt to "separate variables" by rearranging this so that the right-hand side contains only terms involving  $X$  and  $x$  and the left-hand side only terms involving  $T$  and  $t$ . In the case under consideration this is accomplished by dividing both sides by  $XT$  to obtain

$$a^2 \frac{1}{T} \frac{d^2 T}{dt^2} - D \frac{1}{T} \frac{dT}{dt} - k = \frac{1}{X} \frac{d^2 X}{dx^2}$$

- (iii) For any fixed value of  $x$  the right-hand side is a constant  $\lambda$ , to which the left-hand side is equal for all values of  $t$ . We thus conclude that the left-hand side and hence also the right-hand side is identically equal to this constant  $\lambda$ , known as the separation constant. i.e.

$$a^2 \frac{d^2 T}{dt^2} - D \frac{dT}{dt} = (\lambda + k)T \quad \text{and} \quad \frac{d^2 X}{dx^2} = \lambda X.$$

We are thus led to two ordinary second order linear homogeneous differentia

equations, which may be solved to give  $X(x)$  and  $T(t)$  for any particular value of  $\lambda$ . The product of these two solutions is then a solution of the original partial differential equation.

More general solutions may then be obtained by adding together solutions which arise from different values of the separation constant  $\lambda$ .

For example, if in the cable equation  $a^2 = 0$ ,  $D$  and  $k > 0$  we have

$$\frac{dT}{dt} = -\frac{1}{D}(\lambda + k)T \text{ and so } T(t) = Ae^{-\frac{1}{D}(\lambda+k)t} \text{ while } \frac{d^2X}{dx^2} = \lambda X \text{ and so, if we take}$$

the separation constant to be negative, i.e.  $\lambda = -\mu^2$  for some real number  $\mu$ ,

$$X(x) = B \sin \mu x.$$

Thus

$$\phi(x,t) = C e^{-\frac{1}{D}(k-\mu^2)t} \sin \mu x \text{ (where } C = AB \text{ is an arbitrary constant)}$$

is a solution for any value of  $\mu$ .

A more general solution (obtained by adding together two solutions corresponding to  $\mu = \sqrt{k}$  and  $\mu = \mu_0 < \sqrt{k}$ ) is

$$\phi(x,t) = C \sin \sqrt{k} x + C' e^{-\frac{1}{D}(k-\mu_0^2)t} \sin \sqrt{\mu_0} x$$

a form of solution which clearly indicates how  $\phi(x,t)$  approaches the steady state solution,  $C \sin \sqrt{k} x$ , as  $t \rightarrow \infty$ .

The need to construct solutions by this process which satisfy certain initial and boundary value conditions, such as, for example, a solution to the cable equation which has  $\phi(x,0)$  a prescribed function, has in part led to the development of the mathematical theory of *boundary value problems* and *Harmonic Analysis* (a particular case of which is Fourier Analysis). A study of these topics would take us too far away from our primary purpose. However a reader interested in partial differential equations and their use should certainly familiarise himself with this material.

*Collateral Reading.*

For a discussion of Morphogenesis in two dimensions the reader should see J. Maynard Smith "Mathematical Ideas in Biology".

For the applications of partial differential equations to electrically active cells and also to problems in blood flow see

M. Lieberstein "Mathematical Physiology, Blood Flow and Electrically Active Cells".

Material on the technique of separation of Variable, boundary value problems and Fourier Analysis is to be found in

Boyce and Di Prima "Elementary Differential Equations and Boundary Value Problems".

A good reference for partial differential equations in general, from a mathematical point of view is

Weinberger "Partial Differential Equations".

EXERCISE:

- Using the method of separation of variables, find a solution (depending on both  $x$  and  $t$ ) for the telegraph equation

$$\frac{\partial^2 \phi}{\partial x^2} = \frac{1}{c^2} \left\{ \frac{\partial^2 \phi}{\partial t^2} + k \frac{\partial \phi}{\partial t} \right\}.$$

How does it compare with the approximate form of solution obtained in lecture 19?

A PROJECT involving partial differential equations.

A DEMOGRAPHIC MODEL

Problem on Demise of an age structured population

- In the absence of births

Let  $p(x,y)$  denote the probability density that an individual is of age  $y$  at time  $x$ . i.e. The number of individuals of age between  $y$  and  $y + \Delta y$  at time  $x$  is  $p(x,y)\Delta y$ , for  $\Delta y$  sufficiently small. Assume the 'specific mortality rate' for individuals of age  $y$  is  $d(y) = ke^{ay}$  (in accordance with Gompertz' work).

Observe that: Number of individuals of age  $y + \Delta x$  at time  $x + \Delta x$  equals the number of age  $y$  at time  $x$  minus the number of age  $y$  dying during the time interval  $\Delta x$ .

Show that this leads to the "difference equation"

$$p(x + \Delta x, y + \Delta x) = (1 - d(y)\Delta x) p(x,y) \quad \dots (1)$$

Upon appropriate rearrangement show that in the limit as  $\Delta x \rightarrow 0$  this produces the partial differential equation

$$\frac{\partial p}{\partial x} + \frac{\partial p}{\partial y} = -d(y)p. \quad \dots (2)$$

If  $p(x,y) = q(x,y)e^{-\int^y d(t)dt}$  show that  $q$  satisfies the partial differential equation

$$\frac{\partial q}{\partial x} + \frac{\partial q}{\partial y} = 0 \quad \dots (3)$$

For any differentiable function  $f$ , show  $q(x,y) = f(x-y)$  is a solution of (3). (This is in fact the most general solution possible. See exercise 17.3.)

Hence deduce that if 
$$p(0,y) \begin{cases} = y e^{-\frac{k}{a} e^{ay}} & y \geq 0 \\ 0 & \text{otherwise,} \end{cases}$$

then

$$p(x,y) \begin{cases} = (y-x) e^{-\frac{k}{\alpha} e^{\alpha y}} & y \geq x \\ 0 & y < x. \end{cases}$$

Use this to sketch the age profile of the population at times  $x = 1$  and  $x = 2$  if  $k = \alpha = 0.1$ .

2. In the presence of births

Develop a model, similar to the above, which allows for the possibility of a non-zero, age dependent specific birth rate.

Hint: Let  $b(y)$  be the specific birth rate of individuals aged  $y$  and show that the change in  $p(x,y)$  during an interval of length  $\Delta x$  due to births is

$$\Delta x \int_0^{\infty} p(x-y,t)b(t)dt.$$

SERIES V - Delay Differential Equations.

Lecture 21 *The Exponential Solution of the Simple Population Growth Equation with Time Delay*

In Lecture 2 it is shown that a simple model of population growth is given by the first order linear differential equation

$$\frac{dN}{dt} = \beta N - \delta N$$

where  $N$  is the population number,  $\beta$  the constant birthrate per member of the population, and  $\delta$  the constant mortality rate per member. The theory to be given here will be based on the particular equation

$$\frac{dN}{dt} = \beta N,$$

which holds when there is no mortality, or perhaps in the early stages of some population histories, before members begin to die in appreciable numbers. Some extensions to more general cases will be indicated in the Exercises. The solution to this simplified equation is

$$N = N_0 e^{\beta t}$$

where  $N_0$  is the population number at time  $t = 0$ .

The above equation will be valid only for certain primitive organisms which begin reproducing almost from the moment they enter the population. However, if there is a significant time delay before a newly born member begins to reproduce, the effect of this age structure must be taken into account.

Suppose a typical member begins reproducing only after reaching the age of  $a$  time units. The number of members per unit time entering the adult (or reproducing) population is equal to the birthrate  $\beta$  multiplied by the number in the adult population at the time when they were born, i.e. at a time  $a$  units previously. Hence the adult population is given by

$$\frac{dN(t)}{dt} = \beta N(t - a)$$

where  $N(t)$  means the number at time  $t$  and  $N(t - a)$  the number at time  $t - a$ . An equation of this type is called a delay differential equation or an equation with retarded argument or a differential-difference equation.

Guided by the solution to the equation without delay, we test whether there is an exponential solution to this equation also. Suppose the solution is

$$N(t) = N_0 e^{\mu t}, \quad \mu \text{ constant.}$$

Substituting this into the equation gives

$$N_0 \mu e^{\mu t} = \beta N_0 e^{\mu(t-a)}$$

or

$$\mu = \beta e^{-\mu a}$$

This is an equation giving the required  $\mu$ .

In order to solve a transcendental equation of this type, approximation methods must be used, and it is necessary to have available tables of the exponential function, or a slide rule, or a calculator with an exponential function key. Perhaps the simplest way of reaching a solution is by trial and error, i.e., by guessing the value of  $\mu$  which makes the left hand side agree with the right hand side. This is illustrated by the following example.

EXAMPLE. In a certain insect population, members do not reproduce until they reach the age of 20 days, and then produce on the average 1 offspring per adult member per day. Find the adult population at any time, assuming it to be exponential.

Here,  $a = 20$  and  $\beta = 1$ , so that the equation to be solved is

$$\mu = e^{-20\mu}$$

$\mu$ (guessed)	$e^{-20\mu}$ (calculated)
0.5	$4.54 \times 10^{-5}$
0.1	0.135
0.15	0.050
0.12	0.091
0.11	0.111
0.111	0.109

Hence, approximately,  $\mu = 0.110$ .

So, the population of the adults is given by

$$N = N_0 \exp(0.110 t).$$

If all other conditions were the same, but with  $a = 40$ , the result would be  $\mu = 0.0674$ .

If time delay were neglected, so that  $a = 0$ , then  $\mu = \beta = 1$ .

Note the great reduction in the rate of increase when the time delay is introduced, but the relatively smaller change when the delay is doubled. A proposed method for reducing population growth is to increase the age at which members may reproduce. It can be seen from this present example that the reduction may not be as great as expected.

EXERCISES:-

1. In the Example given above, check the result for  $a = 40$ , and find the value of  $\mu$  when  $a = 50$ .
2. In the above Example, find the values of  $\mu$  for the given values of  $a$  when the reproduction rate is doubled, i.e., when  $\beta = 2$ .



3. A certain human adult population doubles in 30 years and the average generation time is 25 years. Making any necessary reasonable assumptions, find how long the population would take to double if the generation time were increased to 30 years, the reproduction constant  $\beta$  being the same in each case.
4. In a certain population, all members reach maturity at about the age  $a$  and reproduce at the rate  $\beta$  per member per unit time until they die at the age  $b$ . Show that an approximate simple model for the adult population is given by the equation

$$\frac{dN(t)}{dt} = \beta N(t - a) - \beta N(t - b).$$

Assuming that there is an exponential solution in the form

$$N = N_0 e^{\mu t},$$

find an equation giving  $\mu$ .

In an insect population, members reach maturity at age 20 days, produce on the average 1 offspring per member per day, and die at the age 30 days. Determine the population history, given that the number increases exponentially.

By comparing the result here with the Example worked out in the Lecture, it will be seen that the apparently drastic mortality produces an unexpected small effect. Thus the equation without the mortality term will work well enough in many cases.

5. In a certain population, in which the age of maturing is  $a$  units, the mortality of the adults is accidental and random so that a constant proportion  $\delta$  dies per unit time. Assuming that immature members suffer no appreciable loss, show that the adult population number is given by the equation

$$\frac{dN(t)}{dt} = \beta N(t - a) - \delta N(t).$$

Obtain the transcendental equation giving the exponential growth.

How is the equation modified if only a proportion  $k$  of immatures survives to adulthood?

6. Show that the equation quoted in Exercise 5 can be written in the form

$$\frac{dM(t)}{dt} = \gamma M(t - a)$$

where

$$M(t) = N(t) e^{t\delta}$$

and

$$\gamma = \beta e^{a\delta}.$$

Hence it can be treated like the simplified equation dealt with in the Lecture.

7. For the case treated in the Lecture, where the adult population is given by the equation

$$\frac{dN(t)}{dt} = \beta N(t - a),$$

7. (continued)

find the number  $I$  in the immature population by the method outlined below.

By taking the rate of increase of immatures to be equal to the rate at which they are produced by adults less the rate at which they are becoming adults, show that

$$\frac{dI(t)}{dt} = \beta N(t) - \frac{dN(t)}{dt}.$$

For a population increasing exponentially, show that the equation for  $I$  is

$$\frac{dI}{dt} = N_0 e^{\mu t} (\beta - \mu)$$

where  $N_0$  is the initial adult population, and  $\mu$  is given by

$$\mu = \beta e^{-\mu a}.$$

If initially  $I = 0$ , show that

$$I = N_0 \left( \frac{\beta}{\mu} - 1 \right) \left( e^{\mu t} - 1 \right).$$

Lecture 22 *Oscillatory Solutions of the Simple Population Growth Equation with Time Delay*

In Lecture 21 it was shown that a solution of the delay differential equation

$$\frac{dN(t)}{dt} = \beta N(t - a)$$

is

$$N = N_0 e^{\mu t}$$

where  $\mu$ , a solution of the transcendental equation

$$\mu = \beta e^{-\mu a},$$

is obtained by some numerical method.

The value of  $\mu$  may also be obtained graphically, as shown in Figure 1.

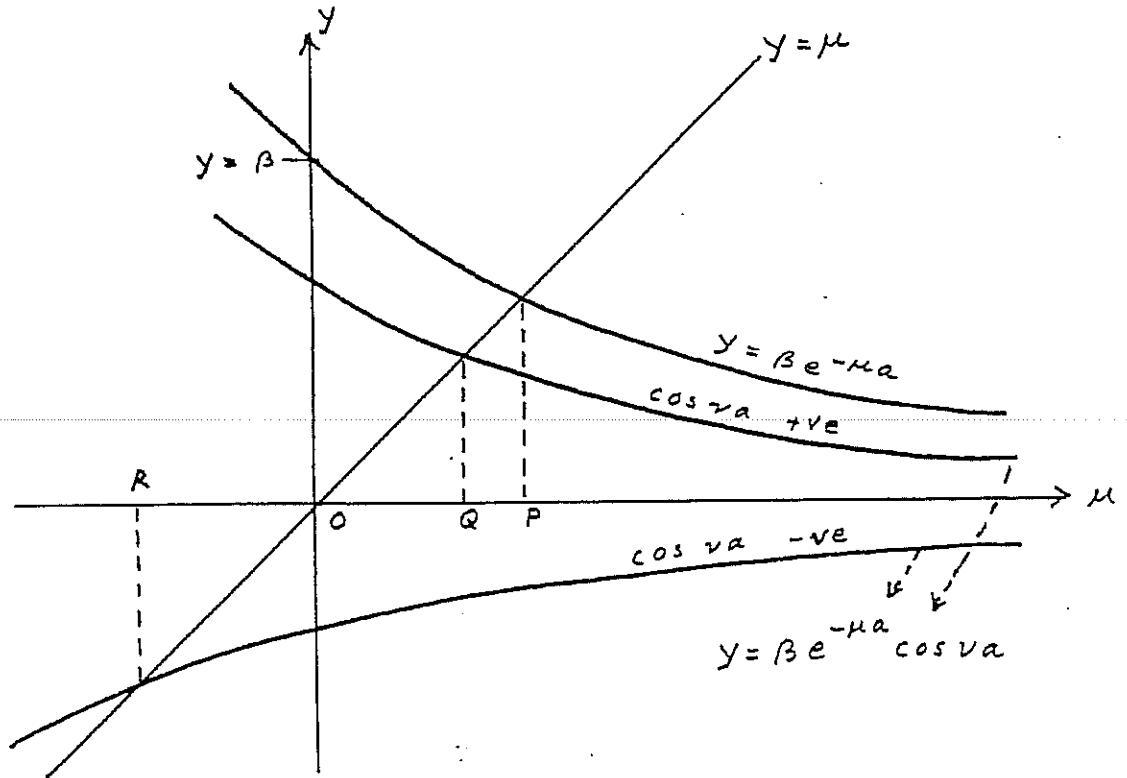


Figure 1. Graphical illustration of the equation  $\mu = \beta e^{-\mu a}$ , and of  $\mu = \beta e^{-\mu a} \cos va$  for positive and negative values of  $\cos va$ .

Draw the straight line  $y = \mu$  and the curve  $y = \beta e^{-\mu a}$  in the  $y, \mu$  plane. Then the solution of the transcendental equation is the value of  $\mu$  where these lines intersect, given by the point P in the diagram. Call this value  $\mu_0$ .

We now look for other solutions of the equation, and try

$$N = N_0 e^{\mu t} \sin(vt + \epsilon).$$

If this is substituted in the delay differential equation, the left hand side is

$$\frac{dN(t)}{dt} = N_0 e^{\mu t} (\mu \sin(vt + \epsilon) + v \cos(vt + \epsilon))$$

while the right hand side is

$$\begin{aligned} \beta N(t - a) &= \beta N_0 e^{\mu(t - a)} \sin(v(t - a) + \epsilon) \\ &= \beta N_0 e^{\mu t} e^{-\mu a} (\sin(vt + \epsilon) \cos va - \cos(vt + \epsilon) \sin va) \end{aligned}$$

If the two sides are to agree at all times, the coefficients of  $\sin(vt + \epsilon)$  on each side must be equal, and similarly for  $\cos(vt + \epsilon)$ . This will be so if

$$\mu = \beta e^{-\mu a} \cos va$$

$$v = -\beta e^{-\mu a} \sin va$$

These are two simultaneous equations giving values of  $\mu$  and  $v$ , and hence solutions of the differential equation.

[The reader can check that these two equations can be obtained also by assuming a solution in the form

$$N = N_0 \exp\{(\mu + iv)t\}.$$

This is an obvious generalisation of the trial solution used in the previous chapter and simply assumes a complex exponential solution instead of a real one. It shows that the trial solution containing the trigonometric function is not just a lucky guess.]

Consider the equation

$$\mu = \beta e^{-\mu a} \cos va$$

as an equation in  $\mu$  for a constant value of  $v$  ( $\neq 0$ ). From Figure I, we see that two kinds of solution may occur. If  $\cos va$  is positive, there will be a positive  $\mu$  given by the point Q. Since  $\cos va < 1$ , this will be less than the  $\mu_0$  occurring in the exponential solution

$$N = N_0 e^{\mu_0 t}.$$

Hence the present solution

$$N = N_0 e^{\mu t} \sin(vt + \epsilon)$$

will correspond to an oscillation which increases in amplitude, but at a slower rate than the exponential solution.

If  $\cos va$  is negative,  $\mu$  will be negative, given by the point R, and the oscillation will die away.

To obtain numerical solutions for  $\mu$  and  $v$ , even by trial and error, it is convenient to eliminate one of them, say  $\mu$ , to give an equation in  $v$  only. After finding  $v$ ,  $\mu$  can be determined. It is also convenient to express the  $\mu$ ,  $v$  and  $\beta$  in terms of  $a$  as time unit, and so we write

$$s = \mu a$$

$$x = va$$

$$\alpha = \beta a.$$

The equations then become

$$s = \alpha e^{-s} \cos x$$

$$x = -\alpha e^{-s} \sin x.$$

The terms  $\alpha e^{-s}$  can be eliminated from these to give

$$s = \frac{-x}{\tan x} . \quad \dots (i)$$

Substituting for  $s$  from this in the second equation gives an equation in  $x$  only,

$$\frac{x}{\sin x} = -\alpha \exp \left( \frac{x}{\tan x} \right). \quad \dots (ii)$$

The value of  $x$ , hence  $v$ , is given by solving (ii) numerically. Then without extra calculation,  $s$ , hence  $\mu$ , is given by (i).

When  $x = 0$ , the original two equations reduce to one,

$$s = \alpha e^{-s}.$$

This gives the first, the exponential solution

$$N = N_0 e^{\mu_0 t}.$$

The equation (ii) has an infinite sequence of solutions, giving  $v = v_1, v_2, v_3, \dots$ , say. Suppose the corresponding values of  $\mu$ , given by equation (i), are  $\mu_1, \mu_2, \mu_3, \dots$ , respectively. These solutions could be called the fundamental modes. Since the original delay differential equation contains only terms of the first degree in  $N$ , a more general solution is a sum of these,

$$N = A_0 e^{\mu_0 t} + A_1 e^{\mu_1 t} \sin (v_1 t + \epsilon_1) + A_2 e^{\mu_2 t} \sin (v_2 t + \epsilon_2) + \dots .$$

The constants  $A_i$  and  $\epsilon_i$  are determined by the initial conditions, and these are usually specified by the population history in an initial interval of length  $a$  units of time, or the age distribution of the population at a given instant. For the population to persist  $N$  must not become negative, and so cannot consist of the trigonometric terms only. The exponential mode  $A_0 e^{\mu_0 t}$  must be present. This will eventually dominate since  $\mu_0$  is greater than any other  $\mu$ .

In the Appendix to this Lecture it is shown that oscillations with increasing amplitude will occur only if  $\alpha > \frac{3\pi}{2}$  and that there will be only a finite number of these. Hence, unless the product of the birthrate  $\beta$  and time lag  $a$  is greater than  $\frac{3\pi}{2}$ , any oscillations will die away leaving little except the exponential increase.

EXERCISES:

1. If the solutions of  $\frac{dN(t)}{dt} = \beta N(t - a)$  are written  $N = Ae^{\mu t} \sin(\nu t + \epsilon)$ , show by direct substitution or otherwise that, for the case  $\alpha (= \beta a) = 1$ , some of the values of  $x (= \nu a)$  are  $x = 0, \frac{\pi}{2} (3 - 0.215), \frac{\pi}{2} (7 - 0.140)$ , approximately. (Note that  $x$  is in radians, and has values around  $(2M - \frac{1}{2})\pi$  where  $M$  is an integer.) Find the corresponding values of  $s (= \mu a)$ . Write down a solution for  $N(t)$  depending on these three modes. Find the next mode.
2. If  $\alpha = 12$ , show that some values of  $x$  are  $x = 0, \frac{\pi}{2} (3 + 0.114), \frac{\pi}{2} (7 + 0.005), \frac{\pi}{2} (11 - 0.013)$ , approximately. Find the corresponding values of  $s$  and state which modes are oscillations with increasing amplitude. Find the next mode.
3. Assuming that the equation

$$\frac{dN(t)}{dt} = \beta N(t - a) - \beta N(t - b), \quad b > a,$$

has solutions in the form

$$N = N_0 e^{\mu t} \sin(\nu t + \epsilon),$$

find the equations which determine  $\mu$  and  $\nu$ .

Examine similarly the equation

$$\frac{dN(t)}{dt} = \beta N(t - a) - \delta N(t).$$

4. For the equation

$$\frac{dN(t)}{dt} = \beta N(t - a)$$

show that if  $\alpha (= \beta a) = (2M - \frac{1}{2})\pi$  for  $M$  an integer, oscillations of constant amplitude can occur. That is, a solution exists for which  $\mu (= s/a) = 0$ .

The smallest value of  $\alpha$  for which this can happen is  $\frac{3\pi}{2}$ , the borderline value of  $\alpha$  for which oscillations of increasing amplitude can occur. (See Appendix to this Lecture.)

APPENDIX to Lecture 22 *The Number of Modes with Increasing Amplitude*

For oscillations with increasing amplitude,  $\mu$  must be positive. In discussing  $\mu$ , the sign of  $x (= va)$  can be chosen as positive, since a change in sign of  $x$  does not affect the equations containing  $s (= \mu a)$ . From equation (i) then,  $\tan x$  must be negative. From the equation  $x = -ce^{-s} \sin x$ ,  $\sin x$  also must be negative. Hence  $x$  must be of the form

$$x = 2M\pi - \xi$$

where  $0 < \xi < \frac{\pi}{2}$ , and  $M$  is an integer  $> 0$ . The equation (ii) can then be written

$$\frac{2M\pi - \xi}{\sin \xi} = \alpha \exp\{-(2M\pi - \xi)/\tan \xi\},$$

$$0 < \xi < \frac{\pi}{2}.$$

As  $\xi \rightarrow 0$ , L.H.S.  $\rightarrow \infty$  and R.H.S.  $\rightarrow 0$ . As  $\xi \rightarrow \frac{\pi}{2}$ , L.H.S.  $\rightarrow (2M - \frac{1}{2})\pi$  and R.H.S.  $\rightarrow \alpha$ . L.H.S. decreases as  $\xi$  goes from 0 to  $\frac{\pi}{2}$ . R.H.S. increases as  $\xi$  goes from 0 to  $\frac{\pi}{2}$ . Hence (see Figure 1), if  $\alpha > (2M - \frac{1}{2})\pi$ , there is a solution for  $\xi$  in the required range,  $0 < \xi < \frac{\pi}{2}$ .

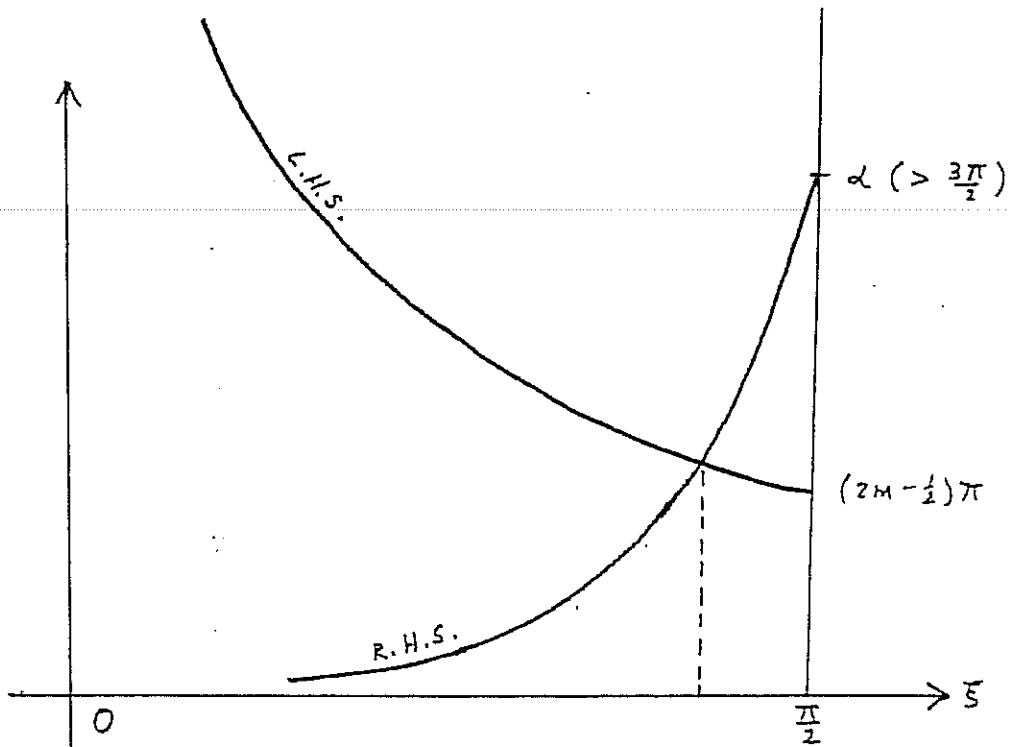


Figure 1. Graphical solution of the equation

$$(2M\pi - \xi)/\sin \xi = \alpha \exp\{-(2M\pi - \xi)/\tan \xi\},$$

for  $0 < \xi < \frac{\pi}{2}$ . The vertical scale is compressed.

If  $\alpha > (2M - \frac{1}{2})\pi$ , it is greater than  $(2[M - 1] - \frac{1}{2})\pi$ , and so on, down to  $\frac{3\pi}{2}$ . Thus, if  $\alpha > (2M - \frac{1}{2})\pi$ , but less than  $([2M + 1] - \frac{1}{2})\pi$ , there will be just  $M$  solutions for  $\xi$  in the range  $0 < \xi < \frac{\pi}{2}$ , and hence only  $M$  oscillatory modes with increasing amplitude. Such modes will occur only if  $\alpha (= a\beta)$  is greater than  $\frac{3\pi}{2}$ , corresponding to  $M = 1$ , the smallest value of  $M$ .

Since  $0 < \xi < \frac{\pi}{2}$ , then  $(2M - \frac{1}{2})\pi < x < 2M\pi$  and so  $\frac{1}{a}(2M - \frac{1}{2})\pi < v < \frac{2M\pi}{a}$ . This gives the limits of  $v$  for the various modes.

The reader can check that the results of Exercises 1 and 2 are in agreement with the results obtained in this Appendix.



Lecture 23 *Direct Methods of Solving the Delay Differential Equation*

The method of solving the equation

$$\frac{dN(t)}{dt} = \beta N(t - a)$$

given in the previous chapter is useful for such general theoretical matters as investigating the kinds of population growth that can occur. But because of the difficulty of determining the constants from the initial conditions, it is not easy to apply in practice to find out how a particular population will develop after starting from given initial conditions. This can be done better by integrating the equation directly for successive stages of length  $a$ , or by a numerical method similar to that described in Lecture 7.

1. Direct Integration

Integrate the differential equation from  $a$  to  $t$ :

$$\int_a^t \frac{dN(t)}{dt} dt = \int_a^t \beta N(t - a) dt$$

i.e.

$$N(t) = N(a) + \int_a^t \beta N(t - a) dt.$$

Now, the initial conditions give the function  $N(t)$  in the interval of range  $a$ ,  $0 \leq t < a$ . Hence the right hand side can be evaluated for any time between  $a$  and  $2a$ . Hence  $N(t)$  is now known in the interval  $a \leq t < 2a$ , which is the same as  $N(t - a)$  in the interval  $2a \leq t < 3a$ . Thus  $N(t)$  can be determined in the interval  $2a \leq t < 3a$ , and so on.

This method shows clearly that the initial conditions must be sufficient to allow the population to be specified over an initial interval of length  $a$  before it can be determined for all later times.

The method may be used for any equation, linear or non-linear, where  $N(t)$  can be expressed in terms which depend on values of  $N$  only up to time  $t - a$ , provided the integrals can be evaluated.

2. Numerical Method

If  $h$  is a (small) time increment, then we can take  $t = nh$ , where  $n$  is the number of increments in time  $t$ . Then, as in Lecture 7, we write  $N(t) = N_n$ . Suppose there are  $A$  increments of length  $h$  in the time  $a$ , so that  $a = Ah$ . Then  $N(t - a) = N_{n-A}$ . Hence, the simple delay differential equation can be written

$$\frac{N_{n+1} - N_n}{h} = \beta N_{n-A}$$

or

$$N_{n+1} = N_n + \beta N_{n-A} h.$$

The value of  $N$  at a given time now depends not only on the value one increment of time previously, but also on the value  $A + 1$  increments previously. As indicated in Chapter 7, this method is a very flexible one, and can be used to solve equations which are intractable by other means. As indicated also, there is the disadvantage of inaccuracies caused by the finite element size  $h$ . In the present case such errors will occur also if  $h$  is not small compared with  $a$ , i.e., the number  $A$  must be reasonably large.

The numerical method, particularly if programmed for a computer, can be used for much more complicated and life-like cases than those considered in the text and exercises of Lectures 21 and 22. For example, mortality due to both accident and natural death can be included in the programme. Systems of equations and non-linear equations with time delay can be treated. Such equations are of importance, not only for describing populations of interacting species in a restricted environment, but also for dealing with metabolic processes in an animal, where the transport of a substance from one compartment to another might take a finite time.

The methods described in this Lecture are useful mainly for dealing numerically with the particular problem under consideration, and provide little theoretical information on populations in general.

EXERCISES:

1. In laboratory experiments, populations are often started by choosing a number of newly matured adults. If there are no losses by accidents, this adult population will remain constant until their offspring begin to mature. Hence, an important case of the equation

$$\frac{dN(t)}{dt} = \beta N(t - a)$$

is that in which  $N(t)$  remains at the constant value  $N_0$  during the initial interval  $0 \leq t < a$ . By integrating the equation directly, show how this population develops over the next 3 time intervals of length  $a$ .

The expressions become tedious to write down after a few steps. Some simplification may be made by writing  $\int_{2a}^t N(t - a)dt$  where  $N(t - a)$  is given in the interval  $2a \leq t < 3a$  as  $\int_a^{t-a} N(t)dt$  where  $N(t)$  is given in the interval  $a \leq t < 2a$ . Similarly for  $\int_{3a}^t N(t - a)dt$  where  $N(t - a)$  is given in the interval  $3a \leq t < 4a$ , and so on.

A method leading to a neater, explicit answer will be given in the next Lecture.

2. Consider the example in Lecture 21, where  $\beta = 1$  and  $a = 20$  days. The equation is

$$\frac{dN(t)}{dt} = N(t - 20).$$

By taking time increments of length 4 days, show that this can be written approximately as

$$N_{n+1} = N_n + 4N_{n-5}.$$

Starting with the initial values  $N_{-5}$  to  $N_0$  equal to 1, 2, 3, 4, 6, 10, respectively, determine the population growth for the first 60 days after this, i.e., up to  $n = 15$ .

In Lecture 21 the exponential solution to the differential equation was found. With the same initial value  $N_0 = 10$  that has been chosen here, the solution is

$$N(t) = 10 \exp(0.110t).$$

All the initial values above for  $N_{-5}$  to  $N_0$  were chosen to agree approximately with this in the range  $t = -20$  to 0. Hence the numerical solution should agree with the exponential solution for later times. Check this by drawing graphs for the numerical and exponential solutions on the same diagram.

3. One method which has often been used is to expand  $N(t - a)$  as a series in  $a$ :

$$N(t - a) = N(t) - a \frac{dN(t)}{dt} + \frac{a^2}{2!} \frac{d^2N(t)}{dt^2} - \dots$$

and to express the time delayed equation as an ordinary differential equation with constant coefficients.

This method cannot be recommended. For even moderately large  $a$ , many terms in the series would need to be taken and so the equation would be of a high order. There would be difficulty also in matching the initial conditions for the delayed equation to the different kind of initial conditions for the ordinary differential equation.

Try to solve the equation

$$\frac{dN(t)}{dt} = 2N(t - 1)$$

by this method, taking terms only up to and including  $\frac{d^2N(t)}{dt^2}$ . With this solution, evaluate the first neglected term  $\frac{a^3}{3!} \frac{d^3N(t)}{dt^3}$  and compare it with the included terms. Compare the result also with the exponential solution of the delay differential equation. Discuss the specification of initial conditions for the two equations.

For further discussion, see A. Mazanov and K.P. Tognetti: "Taylor

3. (continued)

Series Expansion of Delay Differential Equations - A Warning".

J. Theor. Biol. 46 (1974), 271-282.

4. This is an exercise for the reader familiar with computer programming. Devise a programme to solve the equation

$$\frac{dN(t)}{dt} = \beta N(t - a) - \gamma N(t - b) - \delta N(t)$$

for an adult population. The programme should store no more information than is required to perform later steps. With programmes not much more complicated than this, the computer could easily run out of storage if too much information is kept.

As well as the adult population, arrange the programme to record the number of immatures in the population.

Lecture 24      *The Method of Laplace Transforms*

In this Lecture, the method of Laplace transforms described in Lecture 14 is applied to the solution of the delay differential equation

$$\frac{dN(t)}{dt} = \beta N(t - a),$$

given that the population is started with the number  $N_0$  at  $t = 0$ , and assumed to be zero before that time.

The main results from Lecture 14 needed here are the following:

The Laplace transform of the function  $f(x)$  is defined to be

$$F(s) = \int_0^{\infty} e^{-sx} f(x) dx.$$

This is written briefly as

$$F(s) = \mathcal{L}(f)(s)$$

or even

$$F = \mathcal{L}(f).$$

The inverse of the transform is written

$$f = \mathcal{L}^{-1}(F).$$

This means that  $f$  is the function whose Laplace transform is  $F(s)$ .

The Laplace transform of the derivative is

$$\mathcal{L}(f')(s) = s\mathcal{L}(f)(s) - f(0).$$

For the time delayed function defined by

$$f_a(t) = \begin{cases} f(t - a) & \text{for } t \geq a \\ 0 & \text{for } 0 \leq t < a, \end{cases}$$

the Laplace transform is

$$\mathcal{L}(f_a)(s) = e^{-sa} \mathcal{L}(f)(s).$$

This theory will now be applied to the population equation, which can be written

$$\dot{N} = \beta N_a,$$

where the dot denotes differentiation with respect to time  $t$ , and the subscript  $a$  denotes a time lag of  $a$  in a function which is zero before the initial time. Let the Laplace transform of  $N(t)$  be written as  $N^*$ . i.e.,

$$N^*(s) = \mathcal{L}(N)(s) = \int_0^{\infty} e^{-st} N(t) dt.$$

Then the Laplace transforms of the terms in the population equation are

$$\mathcal{L}(\dot{N})(s) = sN^*(s) - N_0$$

since  $N_0$  is the value of  $N(t)$  at  $t = 0$ , and

$$\mathcal{L}(N_a)(s) = e^{-sa} \mathcal{L}(N)(s) = e^{-sa} N^*(s).$$

Hence, we can take the Laplace transform of the population equation. In order to perform the inverse transformation later, some Laplace transforms involving the Heaviside step function,  $H(t)$ , will be needed.

Let  $H(t)$  be defined by

$$H(t) = \begin{cases} 1 & \text{for } t \geq 0 \\ 0 & \text{for } t < 0. \end{cases}$$

Then

$$\mathcal{L}(H) = \int_0^{\infty} e^{-st} H(t) dt = \int_0^{\infty} e^{-st} dt = \frac{1}{s}.$$

Write

$${}^n H_{na} \equiv (t - na)^n H(t - na)$$

That is

$${}^n H_{na} = \begin{cases} (t - na)^n & \text{for } t \geq na \\ 0 & \text{for } t < na. \end{cases}$$

Then

$$\begin{aligned} \mathcal{L}({}^n H_{na}) &= \int_0^{\infty} e^{-st} (t - na)^n H(t - na) dt \\ &= \int_{na}^{\infty} e^{-st} (t - na)^n dt \\ &= \int_0^{\infty} e^{-s(u+na)} u^n du \text{ where } u = t - na \\ &= e^{-sna} \int_0^{\infty} e^{-su} u^n du \\ &= e^{-sna} \int_0^{\infty} e^{-v} \left(\frac{v}{s}\right)^n \frac{dv}{s} \text{ where } v = su \\ &= \frac{e^{-sna}}{s^{n+1}} \int_0^{\infty} e^{-v} v^n dv. \end{aligned}$$

The number  $n$  in the integral can be reduced by successive integrations by parts, leading to the result  $n!$  for the integral when  $n$  is an integer. [In general, the integral defines the gamma function,  $\Gamma(n + 1)$ .]

Hence

$$\mathcal{L}({}^n H_{na}) = \frac{e^{-sna}}{s^{n+1}} n! \quad , \quad n \text{ an integer.}$$

With the convention that  $0! = 1$ , this also includes the result for  $\mathcal{L}(H)$ .

Now, take the Laplace transform of the equation

$$\dot{N} = \beta N_a$$

with the stated initial conditions. It is

$$sN^*(s) - N_0 = \beta e^{-sa} N^*(s).$$

Hence

$$\begin{aligned} N^*(s) &= \frac{N_0}{s - \beta e^{-sa}} \\ &= \frac{N_0}{s} \left( 1 - \frac{\beta}{s} e^{-sa} \right)^{-1} \\ &= \frac{N_0}{s} \sum_{n=0}^{\infty} \frac{\beta^n e^{-sna}}{s^n} \\ &= N_0 \sum_{n=0}^{\infty} \frac{\beta^n e^{-sna} n!}{n! s^{n+1}} \end{aligned}$$

So

$$\mathcal{L}(N)(s) = N_0 \sum_{n=0}^{\infty} \frac{\beta^n}{n!} \mathcal{L}({}^n H_{na})(s)$$

Thus, the inverse transform is

$$N = N_0 \sum_{n=0}^{\infty} \frac{\beta^n}{n!} {}^n H_{na},$$

or, in detail,

$$N(t) = N_0 \sum_{n=0}^{\infty} \frac{\beta^n}{n!} (t - na)^n \mathbb{H}(t - na).$$

This gives the adult population at any time  $t$ , having started with  $N_0$  at time  $t = 0$  and being zero before that time.

This method of solving a delay differential equation is used in the article by K.P. Tognetti and A. Mazanov: "A Two-Stage Population Model", *Mathematical Biosciences* 8 (1970), 371-378.

EXERCISES:

1. Show that the solution in terms of the Heaviside step function as obtained in this Lecture gives  $N(t) = 0$  for  $t < 0$  and  $N(t) = N_0$  at  $t = 0$ .

Show that it gives  $N(t) = N_0$  in the interval  $0 \leq t < a$ . (This is the initial condition for Exercise 23.1. Hence the solution given in this Lecture is the explicit form of the solution to Exercise 23.1.)

1. (continued)

Write down in ordinary algebraic terms, the forms of  $N(t)$  in the intervals  $a \leq t < 2a$  and  $2a \leq t < 3a$ . Check these results with those of Exercise 23.1.

2. Find, by the method of Laplace transforms, a solution to the equation

$$\frac{dN(t)}{dt} = \beta N(t - a) - \delta N(t)$$

with suitably chosen simple initial conditions.



SERIES VI - Stochastic Models

Lecture 25 *Yule's Pure Birth Process - I*

Throughout this final series of lectures it will be assumed that the reader has an understanding of the basic concepts of discrete probability theory. Such material may be found in most of the introductory books on either probability or statistics. By introducing probabilistic ideas we will show how to extend our models beyond the purely deterministic ones considered so far to a class of probability based ones known as Stochastic models. Our introduction to such models is neither comprehensive or adequate for most practical purposes, rather it is intended to demonstrate the possibility of such models and to place the conclusions drawn from our earlier deterministic work into better perspective.

We begin by looking at the simplest stochastic model for a growing population, namely the pure birth process examined by Yule in 1910.

Consider a population of individuals for which the only mechanism of change is through the birth of new individuals according to the following assumptions.

- (i) There is a finite, non-zero, 'conception/recovery' time for births and multiple births do not occur. That is, there exists a period  $h_0 > 0$  such that during an interval of duration less than  $h_0$  each individual can give birth to at most one new individual.
- (ii) The probability that any individual gives birth between  $t$  and  $t + h$  (where  $t$  is an arbitrary time and  $0 < h < h_0$ ) is  $\mu h$  where  $0 < \mu < h_0^{-1}$ .\*

REMARK. It is implicit in (ii) that the likelihood of an individual giving birth is independent of both the state of the remainder of the population and of the individual's age. In particular, immediately after birth a new individual is capable of reproducing.

Clearly these assumptions are only a crude approximation to the real situation for more advanced organisms\*\*.

---

\* More realistically, this probability might be  $\mu h + \mu_1 h^2 + \mu_2 h^3 + \dots$  (as would be the case, for example, if births were Poisson distributed over time). However, as our later analysis will show, we are only interested in the quotient of this probability with  $h$  in the limit as  $h$  tends to 0 and so only the first order term,  $\mu h$ , will prove to be of importance. Hence the simplified version given in (ii) leads to no real loss of generality.

\*\* This model is the stochastic analogue of a population growing with constant specific birth rate (of  $\mu$ ) and zero specific death rate, see lecture 2.

A much more complicated and finely structured model would be required to accurately portray such a population, however the assumptions are quite appropriate for many simple bacterial populations, at least during the early stages of their growth. (They may also be appropriate for the organised colony of cells forming a young plant.)

The first stage in analysing the development of a population satisfying the postulates of the pure birth process is the determination of  $P_n(t)$  the probability that the population is of size  $n$  at time  $t$ . We will assume that at time  $t = 0$  the population has the known initial size  $n_0$ , thus

$$P_n(0) = \begin{cases} 1 & \text{if } n = n_0 \\ 0 & \text{if } n \neq n_0. \end{cases}$$

Further, since from the assumptions the population size can only increase, we have

$$P_n(t) = 0 \text{ for all } t \text{ and } 0 \leq n < n_0$$

Thus it only remains to determine  $P_n(t)$  for  $t > 0$  and  $n \geq n_0$ ; to achieve this we proceed as follows.

From assumption (ii) it follows that the probability of a given individual not giving birth between  $t$  and  $t + h$  is  $1 - \mu h$  (provided of course  $0 < h < h_0$ , henceforth we will always understand this to be the case).

Now, the (conditional) probability that the population will have size  $n + m$  at time  $t + h$  given that it has size  $n$  at time  $t$  equals

$$\left\{ \begin{array}{l} \text{The number of ways } m \\ \text{individuals can be} \\ \text{selected from } n \end{array} \right\} \times \left\{ \begin{array}{l} \text{Probability that each} \\ \text{of the } m \text{ individuals} \\ \text{gives birth between } t \\ \text{and } t + h \end{array} \right\} \times \left\{ \begin{array}{l} \text{Probability that none of} \\ \text{the remaining } n - m \\ \text{individuals gives birth} \\ \text{between } t \text{ and } t + h \end{array} \right\}$$

and so recalling that the number of ways  $m$  individuals can be selected from among  $n$  is the binomial coefficient

$$\binom{n}{m} = \frac{n!}{(n-m)!m!} = \frac{n(n-1)(n-2)\dots(n-m+1)}{m(m-1)\dots 2.1}$$

we have for this probability

$$\binom{n}{m} (\mu h)^m (1 - \mu h)^{n-m}$$

Next we observe that if the population has size  $n$  at time  $t + h$  then one of the following (independent) events must have occurred.

- The population was of size  $n$  at time  $t$  and no birth occurred between  $t$  and  $t+h$
- The population was of size  $n-1$  at time  $t$  and 1 birth occurred between  $t$  and  $t+h$
- The population was of size  $n-2$  at time  $t$  and 2 births occurred between  $t$  and  $t+h$
- .....
- The population was of size  $\left[ \frac{n+1}{2} \right]$  at time  $t$  and  $\left[ \frac{n}{2} \right]$  births occurred between  $t$  and  $t+h$ .

Thus

$$\begin{aligned}
 P_n(t+h) &= \left\{ \begin{array}{l} \text{Probability that population has size } n \text{ at } t+h \\ \text{given it had size } n \text{ at time } t \end{array} \right\} \times P_n(t) \\
 &+ \left\{ \begin{array}{l} \text{Probability that population has size } n \text{ at } t+h \\ \text{given it had size } n-1 \text{ at time } t \end{array} \right\} \times P_{n-1}(t) \\
 &+ \left\{ \begin{array}{l} \text{Probability that population has size } n \text{ at } t+h \\ \text{given it had size } n-2 \text{ at time } t \end{array} \right\} \times P_{n-2}(t) \\
 &+ \dots \\
 &+ \left\{ \begin{array}{l} \text{Probability that population has size } n \text{ at } t+h \\ \text{given it had size } \left\lfloor \frac{n+1}{2} \right\rfloor \text{ at time } t \end{array} \right\} \times P_{\left\lfloor \frac{n+1}{2} \right\rfloor}(t)
 \end{aligned}$$

but from above

$$\left\{ \begin{array}{l} \text{Probability that population has size } n \text{ at } t+h \\ \text{given it had size } n-m \text{ at time } t \end{array} \right\} = \binom{n-m}{m} (\mu h)^m (1-\mu h)^{(n-2m)}$$

and so

$$\begin{aligned}
 P_n(t+h) &= (1-\mu h)^n P_n(t) \\
 &+ (n-1)(\mu h)(1-\mu h)^{n-2} P_{n-1}(t) \\
 &+ \binom{n-2}{2} (\mu h)^2 (1-\mu h)^{n-4} P_{n-2}(t) \\
 &+ \dots \\
 &= \sum_{m=0}^{\left\lfloor \frac{n}{2} \right\rfloor} \binom{n-m}{m} (\mu h)^m (1-\mu h)^{(n-2m)} P_{n-m}(t)
 \end{aligned}$$

Now, since  $0 < \mu h < 1$ , we may use the Binomial Theorem to expand  $(1-\mu h)^k$  as

$$1 - k\mu h + \binom{k}{2} (\mu h)^2 - \binom{k}{3} (\mu h)^3 + \dots$$

in which case the above expression for  $P_n(t+h)$  becomes

$$\begin{aligned}
 P_n(t+h) &= \left[ 1 - n\mu h + \binom{n}{2} (\mu h)^2 \dots \right] P_n(t) \\
 &+ (n-1)(\mu h - (n-2)(\mu h)^2 + \dots) P_{n-1}(t) \\
 &+ \binom{n-2}{2} ((\mu h)^2 \dots) P_{n-2}(t) \\
 &+ \dots \\
 &= P_n(t) - n\mu h P_n(t) + (n-1)\mu h P_{n-1}(t) + \text{terms involving } h^2, h^3, h^4, \dots
 \end{aligned}$$

Rearranging we obtain

$$\frac{P_n(t+h) - P_n(t)}{h} = (n-1)\mu P_{n-1}(t) - n\mu P_n(t) + \text{terms involving } h, h^2, h^3, \dots$$

Clearly the limit as  $h \rightarrow 0$  of the right-hand side exists and is

$$(n-1)\mu P_{n-1}(t) - n\mu P_n(t)$$

whence  $\lim_{h \rightarrow 0} \frac{P_n(t+h) - P_n(t)}{h}$  exists and is by definition  $\frac{dP_n(t)}{dt}$  or  $P'_n(t)$ ,

hence

$$P'_n(t) = (n-1)\mu P_{n-1}(t) - n\mu P_n(t).$$

This is a relatively simple differential equation from which  $P_n(t)$  may be determined provided we know  $P_{n-1}(t)$ . Thus if we first determine  $P_{n_0}(t)$  we may then successively determine  $P_{n_0+1}(t)$ ,  $P_{n_0+2}(t)$ , ...,  $P_{n_0+m}(t)$ , .....

Now, since  $P_{n_0-1}(t) = 0$ , the above differential becomes

$$P'_{n_0}(t) = -n_0\mu P_{n_0}(t),$$

and so

$$P_{n_0}(t) = Ce^{-n_0\mu t}$$

Further since  $P_{n_0}(0) = 1$  we see that  $C = 1$  and

$$P_{n_0}(t) = e^{-n_0\mu t}$$

showing that the probability of the population remaining at its initial size decreases exponentially with time due to the increasing likelihood of further births taking place (see figure 1).

Having determined  $P_{n_0}(t)$  we may now proceed to determine  $P_{n_0+1}(t)$  which satisfies

$$\begin{aligned} P'_{n_0+1}(t) &= n_0\mu P_{n_0}(t) - (n_0+1)\mu P_{n_0+1}(t) \\ &= n_0\mu e^{-n_0\mu t} - (n_0+1)\mu P_{n_0+1}(t) \end{aligned}$$

Setting  $y = P_{n_0+1}(t)$  and  $a = (n_0+1)$  this becomes an equation of the form

$$y' + ay = f(t), \text{ where, in this case, } f(t) = ke^{-kt} \text{ with } k = n_0\mu.$$

A method of solving such equations was outlined in Exercise 3 of lecture 3. First we note that upon multiplying throughout by the integrating factor  $e^{at}$  we have

$$e^{at}y' + ae^{at}y = e^{at}f(t),$$

which may be rewritten as

$$(e^{at}y)' = e^{at} f(t)$$

Upon integration this gives

$$e^{at}y = \int_0^t e^{at} f(t) dt + y(0)$$

or

$$y = e^{-at} \int_0^t e^{at} f(t) dt + y(0)e^{-at},$$

and so, substituting  $ke^{-kt}$  for  $f(t)$  we have

$$y = ke^{-at} \int_0^t e^{(a-k)t} dt + y(0)e^{-at}$$

or

$$y = \frac{k}{a-k} (e^{-kt} - e^{-at}) + y(0)e^{-at}$$

whence replacing  $a$  and  $k$  by their original values and noting  $y(0) = P_{n_0+1}(0) = 0$  we have

$$P_{n_0+1}(t) = n_0 e^{-n_0 \mu t} (1 - e^{-\mu t})$$

See figure 1 for a graph of this function.

The calculation of  $P_{n_0+2}(t)$  follows similarly, first we see that  $P_{n_0+2}(t)$  satisfies

$$\begin{aligned} P'_{n_0+2}(t) &= (n_0 + 1)\mu P_{n_0+1}(t) - (n_0 + 2)\mu P_{n_0+2}(t) \\ &= (n_0 + 1)n_0 \mu e^{-n_0 \mu t} (1 - e^{-\mu t}) - (n_0 + 2)\mu P_{n_0+2}(t) \end{aligned}$$

which is again of the form

$$y' + ay = f(t)$$

where

$$\begin{aligned} y &= P_{n_0+2}(t), \quad a = (n_0 + 2)\mu \text{ and} \\ f(t) &= be^{-kt}(1 - e^{-\mu t}) \text{ with } b = (n_0 + 1)n_0 \text{ and } k = n_0\mu. \end{aligned}$$

So, using the above expression for the solution of such an equation, we obtain

$$P_{n_0+2}(t) = \frac{(n_0 + 1)n_0}{2} e^{-n_0 \mu t} (1 - 2e^{-\mu t} + e^{-2\mu t})$$

The last term of this is of the form  $(1 - 2x + x^2)$  with  $x = e^{-\mu t}$  and so

$$P_{n_0+2}(t) = \frac{(n_0 + 1)n_0}{2} e^{-n_0 \mu t} (1 - e^{-\mu t})^2$$

We could now proceed similarly to obtain  $P_{n_0+3}(t)$ , however it is not hard to see that a likely result might be

$$P_{n_0+3}(t) = \frac{(n_0 + 2)(n_0 + 1)n_0}{3 \times 2} e^{-n_0 \mu t} (1 - e^{-\mu t})^3$$

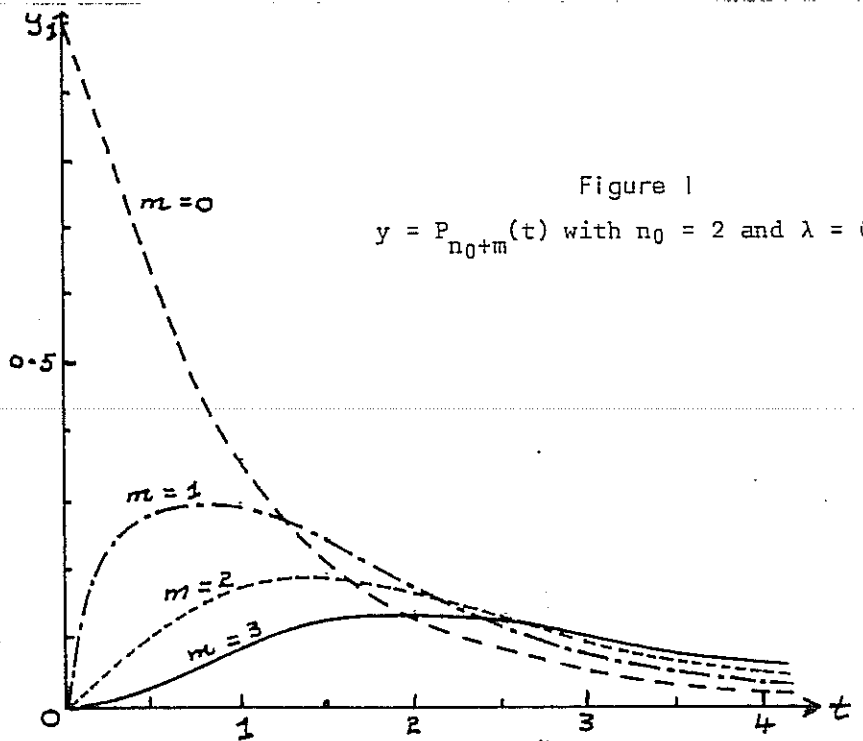
and that in general

$$P_{n_0+m}(t) = \binom{n_0+m-1}{m} e^{-n_0\mu t} (1 - e^{-\mu t})^m.$$

An application of the method of mathematical induction confirms that this is indeed the case, and so we have that the probability of a population, which satisfies the postulates of the pure birth process, having size  $n_0 + m$  at time  $t$  is

$$P_{n_0+m}(t) = \begin{cases} 0 & \text{if } m < 0 \\ \binom{n_0+m-1}{m} e^{-n_0\mu t} (1 - e^{-\mu t})^m & \text{for } m \geq 0. \end{cases}$$

Graphs of  $P_{n_0+m}(t)$  versus  $t$  for  $m = 0, 1, 2$  and  $3$  are illustrated in figure 1, from which we see that the probability of the population having size  $n_0 + m$  first increases to a maximum at  $t_m$  (except in the case of  $m = 0$  where  $t_m = 0$ ) and then decreases as the likelihood of further births increases.



*Collateral Reading.*

An account of basic probability theory may be found in any of the following works:

Alder and Roessler "An Introduction to Probability and Statistics",  
 Freeman, 1968.

Mosteller, Rourke, Thomas "Probability with Statistical Application",  
 Addison-Wesley, 1973.

W. Feller "An Introduction to Probability Theory and its applications",  
 Volume I, Wiley.

Chapter XVII of the last reference contains a good account of the Yule

pure birth process, however the basic assumptions and the derivation of the recursive system of differential equations for  $P_n(t)$  differ from those given here.

EXERCISES:

1. Complete the details of the calculation of  $P_{n_0+2}(t)$  and proceed to a determination of  $P_{n_0+3}(t)$ .
2. Using mathematical induction verify the form of  $P_{n_0+m}(t)$  suggested in the lecture.
3. For a fixed  $n$  show that  $P_{n_0+m}(t)$  has a unique maximum at  $t = t_m$ .  
Show that  $t_1 < t_2 < \dots < t_m < \dots$  as expected (?).  
Estimate  $P_{n_0+m}(t_m)$ .  
Draw a graph of  $n_0 + m$  versus  $t_m$ . What is the significance of such a graph?

Lecture 26 Yule's Pure Birth Process - II

In the previous lecture we established that for a population satisfying the postulates of the Pure Birth Process the probability of it having size  $n_0 + m$  at time  $t$  is

$$P_{n_0+m}(t) = \begin{cases} 0 & \text{if } m < 0 \\ \binom{n_0+m-1}{m} (e^{-\mu t})^{n_0} (1 - e^{-\mu t})^m & \text{for } m = 0, 1, 2, \dots \end{cases}$$

where  $n_0$  is the initial size of the population. Throughout that lecture we concentrated on this probability regarded as a function of time and in the exercises an investigation of this aspect was suggested.

Under some circumstances it is more natural to consider  $P_{n_0+m}(t)$  as a function of  $m$  for a fixed value of time. For this purpose let

$$q = e^{-\mu t}$$

where  $t$  is the arbitrary but for the moment fixed value of time under consideration and write

$$p_m = P_{n_0+m}(t) = \binom{n_0+m-1}{m} q^{n_0} (1-q)^m, \text{ for } m = 0, 1, 2, \dots$$

That is  $p_m$  is the probability that at time  $t$  the population has size  $n_0 + m$  and so

$$\sum_{m=0}^{\infty} p_m = p_0 + p_1 + \dots$$

represents the probability that the population is of size  $n_0$  or  $n_0 + 1$  or  $n_0 + 2$  or ... which should of course equal one. Thus, independently establishing

$$\sum_{m=0}^{\infty} p_m = 1,$$

will provide a check on the conclusions of the last lecture.

LEMMA. If  $S = \sum_{m=0}^{\infty} \binom{n+m-1}{m} q^n (1-q)^m$ , then  $S = 1$  for  $n, m > 0$  and  $0 < q < 1$

Proof.  $S = q^{n_0} \sum_{m=0}^{\infty} \binom{n_0+m-1}{m} (1-q)^m$

Now recalling the binomial theorem for negative exponents:

$$\frac{1}{(1-x)^k} = 1 + kx + \frac{k(k+1)}{2} x^2 + \dots + \binom{k+m-1}{m} x^m + \dots$$

provided  $-1 < x < 1$ , we can recognise that

$$\sum_{m=0}^{\infty} \binom{n+m-1}{m} (1-q)^m = \frac{1}{(1-(1-q))^{n_0}}$$



so 
$$\sum_{m=0}^{\infty} p_m = q^{n_0}/q^{n_0} = 1 \quad \text{as required.}$$

Thus the  $p_m$  do indeed define a discrete probability distribution since,  $q = e^{-\mu t}$  so  $0 < q < 1$ . This same distribution arises in other contexts as well and, for reasons which should be obvious, is referred to as the *negative binomial distribution*.

In order to obtain more tractable and convenient information from that already determined we proceed to calculate the *mean* population size

$$M = \sum_{m=0}^{\infty} (n_0 + m)p_m = n_0 + \sum_{m=0}^{\infty} m p_m, \text{ since } \sum_{m=0}^{\infty} p_m = 1.$$

Now

$$\sum_{m=0}^{\infty} m p_m = \sum_{m=1}^{\infty} m \binom{n_0+m-1}{m} q^{n_0} (1-q)^m$$

$$\begin{aligned} \text{where } m \binom{n_0+m-1}{m} &= m \frac{(n_0+m-1)(n_0+m-2)\dots(n_0+1)n_0}{m(m-1)\dots 2.1} \\ &= \frac{(n_0+m-1)(n_0+m-2)\dots(n_0+1)}{(m-1)(m-2)\dots 2.1} n_0 \\ &= \binom{n_0+m-1}{m-1} n_0 \end{aligned}$$

$$\text{and so } \sum_{m=1}^{\infty} m p_m = n_0 \sum_{m=1}^{\infty} \binom{n_0+m-1}{m-1} q^{n_0} (1-q)^m$$

which we choose to write as

$$\begin{aligned} &n_0 \sum_{m=1}^{\infty} \left[ \binom{(n_0+1)+(m-1)-1}{(m-1)} q^{n_0+1} (1-q)^{m-1} \right] q^{-1} (1-q) \\ &= n_0 q^{-1} (1-q) \sum_{m=1}^{\infty} \binom{(n_0+1)+(m-1)-1}{(m-1)} q^{n_0+1} (1-q)^{m-1} \end{aligned}$$

Now the sum is of precisely the form shown to add to one in the above lemma, whence

$$\sum_{m=0}^{\infty} m p_m = n_0 q^{-1} (1-q) = n_0 q^{-1} - n_0$$

and so

$$M = n_0 q^{-1} \quad *$$

Hence our stochastic model of population growth under the pure birth assumptions predicts a mean population size at time  $t$  of

$$M(t) = n_0 q^{-1} = n_0 e^{\mu t} \quad (q = e^{-\mu t}).$$

---

\* A generally more applicable and less tedious way to derive this result will be presented in the next lecture.

In other words the 'expected' population size increases exponentially. Furthermore its mean growth curve is precisely that predicted by the simple deterministic model of a population with constant specific birth rate  $\mu$  (zero specific death rate) starting from the same initial size (see p.132 and Lecture 2). Thus our deterministic model can now be seen to represent the mean behaviour of a stochastically prescribed situation - a conclusion which is generally valid.

To determine how good a predictor this mean growth is (and hence the deterministic model), we might look at the *variance* of the population size about the mean and so gain some measure of the likely departures from the mean in our stochastic model.

The variance at time  $t$  is

$$v(t) = \sum_{m=0}^{\infty} (m+n_0)^2 P_{m+n_0}(t) - n_0^2 e^{2\mu t}$$

and a similar, though slightly more tedious calculation, than that used to find  $M(t)$  shows that

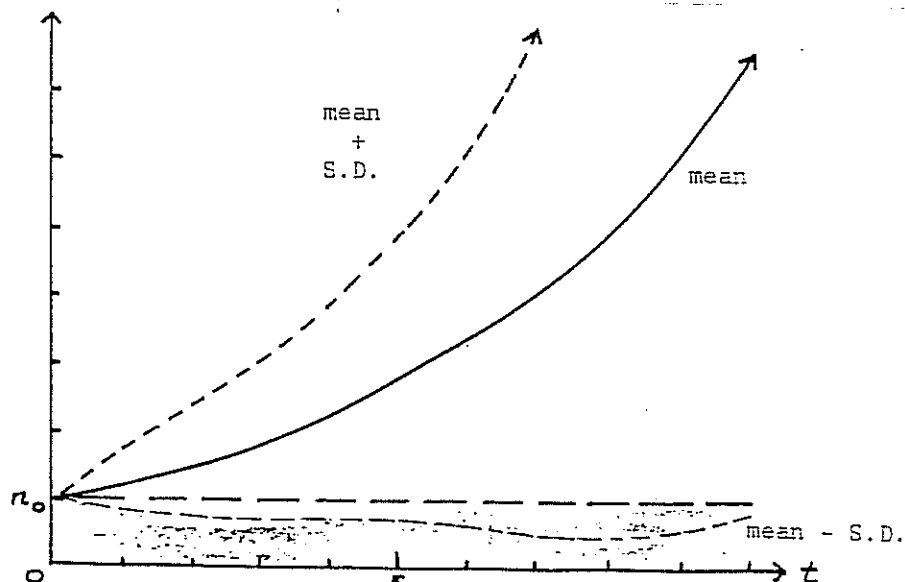
$$v(t) = n_0 e^{\mu t} (e^{\mu t} - 1)$$

a rapidly increasing function of  $t$ . In fact we see that

$$v(t) = M(t)^2/n_0 - M(t)$$

and so for large values of  $t$  the variance will increase like the square of the mean. Thus the *standard deviation* will be roughly proportional to the mean (see figure 1). Hence a particular population developing according to this, or similar models, is likely to deviate markedly from the expected population size. Thus any prediction of the size of a particular population, whether made from the stochastic model, or the corresponding deterministic model, is likely to be quite poor. However, for a large collection of populations satisfying the underlying assumptions the distribution of sizes at any time is reliably specified by the above stochastic solutions. Thus, you might say, that *if nothing more is known about a particular population's dynamics than its satisfying the assumptions of the pure birth process, then no useful long term model is possible.*

Figure 1



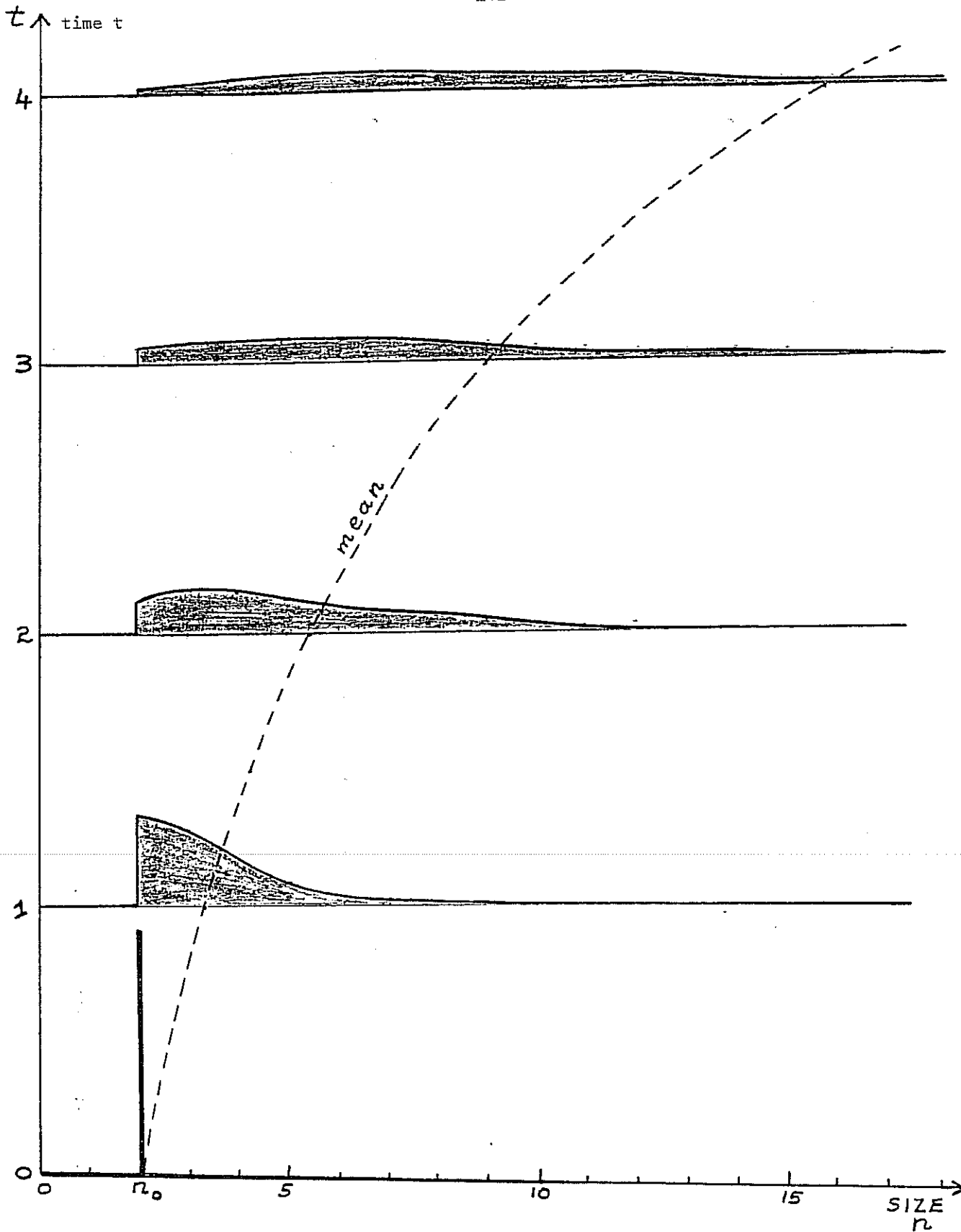


Figure 2  
Predicted distribution of population sizes at various times.  
( $\lambda = 0.5$ )

*Collateral Reading.*

Same as for the last lecture, particularly Feller.

EXERCISES:

1. For  $\mu = 0.1$  and  $n = 1$  plot graphs of  $p_m$  versus  $m$  for  $t = 0, 1, 2, 3$  and  $4$ , indicating the mean and standard deviation about the mean.
2. Derive the expression for  $v(t)$  given in the lecture.

Lecture 27. *A Birth and Death Process.*

In this lecture we refine the previous pure birth model by allowing for the occurrence of deaths.

The assumptions under which births take place will be the same as those previously used, namely:

There exists a period  $h_0 > 0$ , such that for  $0 < h < h_0$  the probability of any individual giving birth to one off-spring during an interval  $t$  to  $t+h$  is  $\mu h$  (where  $0 < \mu < h_0^{-1}$ ) and any one individual cannot give birth to more than one individual during such an interval.

We will further assume that, independent of other factors, any individual is likely to die between  $t$  and  $t+h$  with probability  $\lambda h$ , where without loss of generality  $0 < h < h_0$  and  $0 < \lambda < h_0^{-1}$ .\*

Let  $P_n(t)$  denote the probability of the population having size  $n \geq 0$  at time  $t$  and assume

$$P_n(0) = \begin{cases} 1 & \text{if } n = n_0 \\ 0 & \text{if } n \neq n_0 \end{cases},$$

that is the population is known to have initial size  $n_0$ .

By similar reasoning to that used previously we have that

$$\begin{aligned} P_n(t+h) &= \Delta(n-1,1)P_1(t) + \Delta(n-2,2)P_2(t) + \Delta(n-3,3)P_3(t) + \dots \\ &= \sum_{m=-\infty}^{n-1} \Delta(m,n-m)P_{n-m}(t) \end{aligned}$$

where  $\Delta(m,n-m)$  denotes the probability that the population size has changed by an amount  $m$  during the interval  $t$  to  $t+h$  given that it had size  $n-m$  at time  $t$ .

Such a change will occur if  $m+k$  births and  $k$  deaths occur between  $t$  and  $t+h$  where

$$\max\{0, -m\} \leq k \leq \min\{n-m, n-2m\}$$

(for values of  $k$  outside the range either the required number of births or the required number of deaths is impossible).

Thus

$$\Delta(m,n-m) = \sum_k \left\{ \begin{array}{l} \text{probability of } m+k \text{ births} \\ \text{out of population of size } n-m \end{array} \right\} \times \left\{ \begin{array}{l} \text{probability of } k \text{ deaths out of} \\ \text{population of size } n-m \end{array} \right\}$$

---

\* As in the previous model (p.132) we could, without effecting our final conclusions, replace  $\mu h$  by  $\mu h + \mu_1 h^2 + \mu_2 h^3 + \dots$  and  $\lambda h$  by  $\lambda h + \lambda_1 h^2 + \lambda_2 h^3 + \dots$

(Note: Here we have assumed that any individual can both give birth and then die during the interval  $t$  to  $t+h$  with probability  $\mu\lambda h^2$ , however since the probability of such occurrences is proportional to  $h^2$  their presence will not effect our final conclusions and so we could equally well have assumed births and deaths to be mutually exclusive events for any individual during an interval of duration  $h$ .)

Hence, by similar arguments to those used for the pure birth model, we have

$$\Delta(m, n-m) = \sum_k \binom{n-m}{m+k} (\mu h)^{m+k} (1-\mu h)^{n-2m-k} \times \binom{n-m}{k} (\lambda h)^k (1-\lambda h)^{n-m-k}$$

From this we see that except for  $m = -1, 0$  or  $1$ ,  $\Delta(m, n-m)$  contains a factor of  $h^q$  where  $q \geq 2$  and so for the purpose of our analysis it need not be explicitly determined.

Further

$$\Delta(-1, n+1) = (n+1)\lambda h + \text{terms involving } h^2, h^3 \text{ etc.}$$

$$\Delta(0, n) = 1 - n(\mu + \lambda)h + \text{terms involving } h^2, h^3 \text{ etc.}$$

$$\Delta(1, n-1) = (n-1)\mu h + \text{terms involving } h^2, h^3 \text{ etc.}$$

(unless  $n = 0$  in which case it does not exist).

Substituting these into the above expression for  $P_n(t+h)$  yields

$$P_n(t+h) = (n+1)\lambda h P_{n+1}(t) + P_n(t) - n(\mu + \lambda)h P_n(t) + (n-1)\mu h P_{n-1}(t) + \text{terms involving } h^2, h^3 \text{ etc.}$$

and so

$$\frac{P_n(t+h) - P_n(t)}{h} = (n+1)\lambda P_{n+1}(t) - n(\mu + \lambda)P_n(t) + (n-1)\mu P_{n-1}(t) + \text{terms involving } h, h^2 \text{ etc.}$$

Thus we may take the limit as  $h \rightarrow 0$  to obtain the system of differential equations

$$P_n'(t) = (n+1)\lambda P_{n+1}(t) + (n-1)\mu P_{n-1}(t) - n(\mu + \lambda)P_n(t)$$

(Note: In the case of  $n = 0$  the  $P_{n-1}(t)$  term is missing.)

This is not a simple recursive system of equations such as those produced by the pure birth model. To determine  $P_n(t)$  it is necessary not only to know its predecessor  $P_{n-1}(t)$  but also its successor  $P_{n+1}(t)$ . Naturally to determine the solution to a system such as this is an extremely difficult task - one which is far beyond the scope of these lectures. Nonetheless, solutions have been found. For example in the case  $n_0 = 1$  it can be shown that

$$P_0(t) = \lambda \left( \frac{1 - e^{(\mu-\lambda)t}}{\lambda - \mu e^{(\mu-\lambda)t}} \right)$$

$$P_n(t) = \left\{ 1 - \frac{\mu}{\lambda} P_0(t) \right\} \left\{ 1 - P_0(t) \right\} \left\{ \frac{\mu}{\lambda} P_0(t) \right\}^{n-1} \quad (n \geq 1)$$

(see exercise 1).

Instead of persisting in the quest for explicit solutions we will, in the next lecture, turn directly to the differential equations for information about the behaviour of  $P_n(t)$ .

EXERCISE:

1. Show that the form of  $P_n(t)$  given in the lecture is indeed a solution of the Birth/Death Process in case  $P_1(0) = 1$

Hint: Either substitute directly or assume  $P_n(t)$  has the form

$$P_n(t) = f(t) g^n(t) \text{ and try to determine } f \text{ and } g.$$

Lecture 28. *Analysis of the Birth Death Model.*

Here we are interested in analysing the development of a population for which the probability of it having size  $n$  at time  $t$  is  $P_n(t)$  where

$$P_n'(t) = (n+1)\lambda P_{n+1}(t) + (n-1)\mu P_{n-1}(t) - n(\mu+\lambda)P_n(t).$$

As in the case of the pure birth model, a very useful start would be the determination of the mean (expected) population size at time  $t$ ,

$$M(t) = \sum_{n=0}^{\infty} nP_n(t).$$

Let us accept that  $M(t)$  will be a differentiable function and further that

$$M'(t) = \sum_{n=0}^{\infty} nP_n'(t).$$

Then using the differential equation to replace  $P_n'(t)$  we obtain

$$M'(t) = \sum_{n=0}^{\infty} n \left[ (n+1)\lambda P_{n+1}(t) + (n-1)\mu P_{n-1}(t) - n(\mu+\lambda)P_n(t) \right]$$

which may be rewritten as

$$M'(t) = \sum_{n=0}^{\infty} \left[ (n+1)^2\lambda P_{n+1}(t) - (n+1)\lambda P_{n+1}(t) + (n-1)^2\mu P_{n-1}(t) + (n-1)\mu P_{n-1}(t) - n^2\mu P_n(t) - n^2\lambda P_n(t) \right]$$

Splitting this into individual sums and changing the variable of summation where appropriate, we obtain

$$M'(t) = \lambda \sum_{n=0}^{\infty} n^2 P_n(t) - \lambda \sum_{n=0}^{\infty} n P_n(t) + \mu \sum_{n=0}^{\infty} n^2 P_n(t) + \mu \sum_{n=0}^{\infty} n P_n(t) - \mu \sum_{n=0}^{\infty} n^2 P_n(t) - \lambda \sum_{n=0}^{\infty} n^2 P_n(t)$$

The terms involving  $\sum_{n=0}^{\infty} n^2 P_n(t)$  cancel and so we are left with

$$\begin{aligned} M'(t) &= (\mu - \lambda) \sum_{n=0}^{\infty} n P_n(t) \\ &= (\mu - \lambda)M(t). \end{aligned}$$

Thus

$$\begin{aligned} M(t) &= C e^{(\mu-\lambda)t} \\ &= n_0 e^{(\mu-\lambda)t} \quad \text{since } M(0) = \sum_{n=0}^{\infty} n P_n(0) \\ &= n_0. \end{aligned}$$

Again we see that the predicted mean growth corresponds to the solution of the corresponding deterministic model, in this case a population of initial size  $n_0$  growing with specific birth rate  $\mu$  and specific death rate  $\lambda$  (see lecture 2).

A similar argument allows us to determine that the variance

$$\begin{aligned} v(t) &= \sum_{n=0}^{\infty} n^2 P_n(t) - M^2(t) \\ &= e^{2(\mu-\lambda)t} \left\{ 1 - e^{(\lambda-\mu)t} \right\} \frac{\mu+\lambda}{\mu-\lambda}. \end{aligned}$$

We must distinguish between two cases -

- (1)  $\mu > \lambda$ , here the mean is an exponentially increasing function of time, while for large times the variance

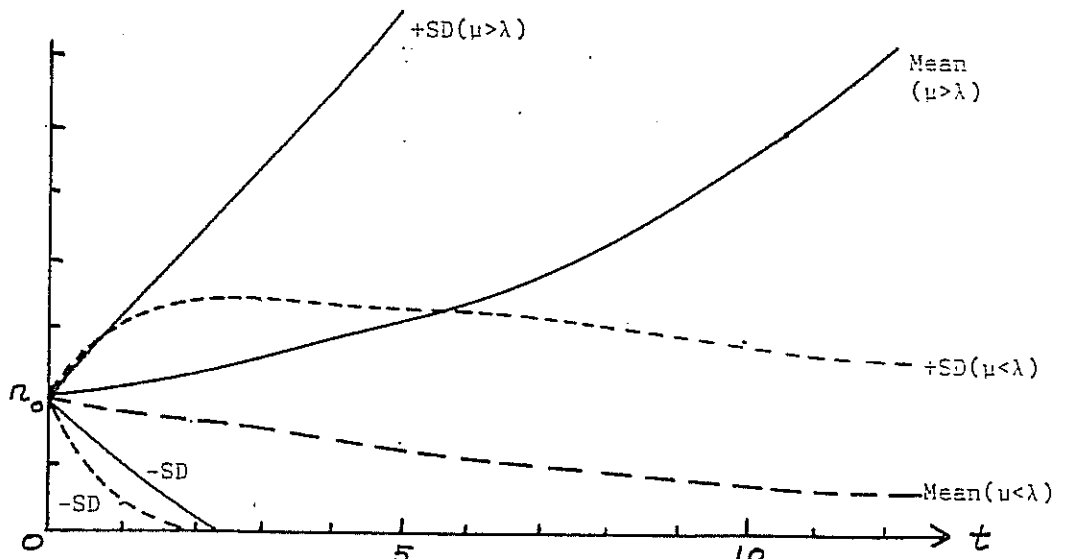
$$\begin{aligned} v(t) &\sim \frac{\mu+\lambda}{\mu-\lambda} e^{2(\mu-\lambda)t} \\ &= \frac{\mu + \lambda}{n_0^2(\mu-\lambda)} M^2(t). \end{aligned}$$

Thus the standard deviation is proportional to the mean and so again we must conclude that in the absence of any further information no reliable predictive model for such a population is possible.

- (2)  $\lambda > \mu$ , here the mean population size decays exponentially with time and the variance for large times is

$$\begin{aligned} v(t) &\sim \frac{\mu+\lambda}{\lambda-\mu} e^{(\mu-\lambda)t} \\ &= \frac{\mu + \lambda}{n_0(\lambda-\mu)} M(t). \end{aligned}$$

Thus the standard deviation decreases with time like the square root of the mean size. (See figure 1.)





At first this might appear a more healthy situation, however although in absolute terms the standard deviation does decrease the likely percentage error is proportional to

$$\frac{\text{standard deviation}}{\text{mean}} \propto M(t)^{-\frac{1}{2}}$$

and so increases with increasing time. An even worse situation than in case (1) where

$$\frac{\text{standard deviation}}{\text{mean}} \rightarrow \text{constant}$$

and so the percentage error is likely not to vary too greatly.

#### Ultimate Extinction

From the governing differential equations for the birth/death process, we have

$$P_0'(t) = \lambda P_1(t) \geq 0$$

and so  $P_0(t)$  is an increasing function, which may be interpreted as the probability that the population is extinct by time  $t$ . It is therefore natural to take

$$P_0 = \lim_{t \rightarrow \infty} P_0(t)$$

as the probability of *ultimate extinction*.

In the case of initial size  $n_0 = 1$  we have the explicit solution

$$P_0(t) = \lambda \left\{ \frac{1 - e^{(\mu-\lambda)t}}{\lambda - \mu e^{(\mu-\lambda)t}} \right\}.$$

If  $\mu < \lambda$ , we have  $e^{(\mu-\lambda)t} \rightarrow 0$  as  $t \rightarrow \infty$  and so

$$P_0 = \lim_{t \rightarrow \infty} P_0(t) = \frac{\lambda}{\lambda} = 1.$$

That is, ultimate extinction occurs with probability one.

Alternatively, when  $\mu > \lambda$  we have

$$P_0 = \lim_{t \rightarrow \infty} P_0(t) = \frac{\lambda}{\mu} (< 1).$$

as may be seen by writing

$$P_0(t) = \lambda \left\{ \frac{e^{(\lambda-\mu)t} - 1}{\lambda e^{(\lambda-\mu)t} - \mu} \right\}.$$

For  $n_0 > 1$ , call the descendants of any one of the initial  $n_0$  individuals an ancestral line. It has just been shown that

$$\text{Prob (an ancestral line dies out)} = \begin{cases} \frac{\lambda}{\mu} & \text{if } \lambda \leq \mu \\ 1 & \text{if } \lambda > \mu \end{cases}$$

and from the basic assumptions of the model it follows that this probability is independent of the fate of the other lines. Hence the probability that all ancestral lines die out is  $[\text{Prob (an ancestral line dies out)}]^{n_0}$ ,

i.e.  $\text{Prob (ultimate extinction of the population)} = \begin{cases} \left(\frac{\lambda}{\mu}\right)^{n_0} & \text{if } \lambda \leq \mu \\ 1 & \text{if } \lambda > \mu. \end{cases}$

Asymptotic behaviour

Here we investigate

$$p_n = \lim_{t \rightarrow \infty} P_n(t) \quad (n = 1, 2, 3, \dots)$$

which may be interpreted as the likelihood of finding the population having size  $n$  for large values of  $t$ .

It can be shown, and is certainly plausible, that

$$\lim_{t \rightarrow \infty} P'_n(t) = 0 \quad \text{for all } n,$$

and so from the differential equations for the  $P_n(t)$  we have

$$0 = \lim_{t \rightarrow \infty} P'_n(t) = \mu(n-1)p_{n-1} + \lambda(n+1)p_{n+1} - n(\lambda+\mu)p_n$$

or 
$$p_{n+1} = \frac{1}{(n+1)\lambda} [n(\lambda+\mu)p_n - \mu(n-1)p_{n-1}].$$

Thus  $p_1 = 0$  (the case when  $n = 0$ , since  $p_{-1} = 0$ ) and it then follows that  $p_2 = 0; p_3 = 0, \dots$

We can interpret these conclusions as follows.

With probability  $p_0$  (see previous section) the population reaches extinction, or else the population size increases beyond all bounds and this happens with probability  $1 - p_0$ . Hence *ultimately the population must either die out or reach arbitrarily large sizes.*

The style of analysis, and the type of conclusions reached, in this section may be carried through for a great many stochastic models besides the one considered here and so serves as an introduction to the general tools available for use by those working with stochastic models.

*Collateral reading:*

Apart from Feller (already referred to in previous lectures) there are a great many good books on stochastic processes and their applications, one such work is

Emanuel Parzen, "Stochastic Processes"

Holden-Day, San Francisco, 1962.

EXERCISES:

1. For the birth/death model, show

$$S = \sum_{n=0}^{\infty} P_n(t)$$

is a constant.

Hence conclude that for each value of  $t$  the distribution of population sizes is proper, that is

$$\sum_{n=0}^{\infty} P_n(t) = 1.$$

2. Derive the expression for the variance, of the birth/death process, given in the lecture.
3. Using the methods of this lecture re-establish the formulae for the mean and variance of the pure birth process (see Lecture 26).
4. In the case of  $n_0 = 1$  graph the explicit form of  $P_n(t)$  given in lectures if  $\mu = 0.3$  and  $\lambda = 0.2$ ;  
(a) regarding the  $P_n(t)$  as functions of  $t$  for  $n = 2,3,4$ .  
(b) regarding the  $P_n(t)$  as functions of  $n$  for  $t = 1,2,3,4$ .
5. Show that a population of initial size 11 growing according to the assumptions of the birth and death process, with  $\mu = 0.3$  and  $\lambda = 0.2$ , has about one chance in one-hundred of becoming extinct.
6. Tacitly we assumed in the course of this lecture that all solutions of the difference equation  $p(n) = p p(n+1) + q p(n-1)$  are of the form  $p(n) = A + B\left(\frac{q}{p}\right)^n$ . Prove this assumption.

PROJECT:                    *Simulation of the Birth/Death Process*

If you are familiar with a computer language, such as FØRTRAN, you should try the following exercise. (It would also be preferable if you had access to a computer, and so were able to obtain some actual results.)

Consider an "imaginary" population, initially of size  $N_0$  (for definiteness, take  $N_0 = 100$ ) such that during a unit interval of time the probability of any individual

- (a) giving birth to one new member is  $\mu$
- (b) dying is  $\lambda$
- (c) not giving birth or dying is  $1 - (\mu + \lambda)$ .

Thus, these are the only three possibilities.

By using a random number generator\* your task is to develop a computer program which will *simulate* the development of such a population for specified values

---

\* Usually the scientific subroutines library contains a random number generator. In the case of FØRTRAN each execution of the following sequence of instructions produces the number  $X$  which, while it may not be perfectly randomly distributed is quite adequate for our needs.

```
R = S + X
I = R
S = T
T = U
U = V
V = W
W = X
X = I
X = R - X
```

Of course you must avoid altering the values of any of these variables elsewhere in the program, and must initially assign "random values" (8 digit numbers between 0 and 1) to each of S, T, U, V, W and X.

of  $\lambda$  and  $\mu$ .<sup>\*\*</sup>

One way of doing this is to repeat the following *trial* as many times as there are individuals in the current population.

Generate  $X$ , a rectangularly distributed random number between 0 and 1. If  $0 \leq X \leq \mu$  then a birth has occurred, if  $\mu \leq X < \lambda + \mu$  then a death has occurred, while, if  $\lambda + \mu \leq X \leq 1$  there has been neither a birth nor a death. If at each trial the effect on the population's size of these births and deaths is recorded, then upon the conclusion of the trials we will have the population size one unit of time later. By repeating this whole procedure over and over again you can determine the population's size at the end of each successive time unit.

Using values of  $\lambda = 0.15$  and  $\mu = 0.2$  when running the program you develop, determine the population size at the end of 1,2,3,...,20 time units.

Note:- By running the program a large number of times you can build up a frequency table of population sizes for different times.

(k runs out of 2 gave a population size between 110 and 120 at time 6 units.)  
Such a frequency table should to some extent mirror the  $P_n(t)$  versus  $n$  curves for the birth/death model considered in the lecture and so provides one way of obtaining numerical information about the model's solution.

---

<sup>\*\*</sup>In this we have assumed that the age of an individual is unimportant, if this were not the case we would need to add in an age structure, assigning different probabilities of giving birth and death according to the individual's age. Thus rather than "pooling" all the individuals into one population we would have to record separately the evolution of each age group. You might try to elaborate on your program to include something like this.

\*APPENDIX TO LECTURE 28 - *Stochastic Differential Equations*

Over the last two decades a powerful new approach to Stochastic Modelling has been developed. This is the theory of Stochastic Differential Equations which is generally based on Itô's calculus of stochastic differentials. The newness of this approach has to some extent prevented its use widely in the biological area. However it is almost certain that its use will become more common in the future. Unfortunately, an appreciation of the theory and techniques of stochastic differential equations requires a sound understanding of Probability theory, in particular the concept of stochastic integrals. Any attempt to develop this background would take us far beyond the scope of these lectures and so we must content ourselves with a brief examination of a discrete model which may nonetheless serve to indicate the flavour of stochastic differential equations. (The continuous stochastic differential equation analogue of this model requires the extremely intricate theory of Feynman Paths and the existence of 'appropriate probability' measures on certain function spaces.)

Firstly consider the deterministic model of a population growing with constant specific birth rate  $\mu$  (and zero death rate). If  $N(t)$  denotes the size of the population at time  $t$  then

$$N'(t) = \mu N(t).$$

We may replace this by the discrete approximation

$$\frac{N(t+h) - N(t)}{h} = \mu N(t) \quad (\text{where } h > 0)$$

To find  $N(t)$  approximately using this we could divide the interval 0 to  $t$  into  $n$  equal pieces, set  $h = \frac{t}{n}$  and rewrite the above difference equation as

$$N_i = N_{i-1} + \mu N_{i-1} \frac{t}{n} \quad \text{where } N_i = N\left(\frac{it}{n}\right) \quad i = 0, 1, \dots, n.$$

or

$$N_i = N_{i-1} \left(1 + \frac{\mu t}{n}\right)$$

From this it is clear that

$$N_m = N_0 \left(1 + \frac{\mu t}{n}\right)^m \quad m = 0, 1, \dots, n \quad \text{where } N_0 = N(0)$$

whence

$$N(t) = N_n \approx N_0 \left(1 + \frac{\mu t}{n}\right)^n.$$

As is well known  $\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = e^x$ , and so the 'exact' result  $N(t) = N_0 e^{\mu t}$  follows upon taking the limit as  $n \rightarrow \infty$ .

If in this model  $\mu$  were not a constant but varied with time, then the appropriate difference equation would be

$$N_i = N_{i-1} \left(1 + \frac{a_i t}{n}\right) \quad \text{where } a_i = \mu\left(\frac{(i-1)t}{n}\right).$$

which is easily seen to have the solution

$$N(t) = N_0 \left(1 + \frac{a_1 t}{n}\right) \left(1 + \frac{a_2 t}{n}\right) \dots \left(1 + \frac{a_n t}{n}\right)$$

$$= N_0 \prod_{i=1}^n \left(1 + \frac{a_i t}{n}\right).$$

We now consider the analogous stochastic situation where the "instantaneous" specific birth rates  $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_n$  are independent random variables. For simplicity, we will assume that the  $\underline{a}_i$ 's are identically distributed, with mean  $\mu$  and variance  $\sigma^2$  (Note: it is not unreasonable to suppose that  $\sigma^2$  is a function of  $\frac{t}{n}$ ). This corresponds to the deterministic case with all the  $a_i$  equal.

By exactly the same reasoning as before we find that

$$\underline{N}(t) = N_0 \prod_{i=1}^n \left(1 + \frac{\underline{a}_i t}{n}\right)$$

is a random variable, equal to the population size at time  $t$ .

By a well known result in probability theory the mean of such a product is the product of the means of the factors and so we have

$$\text{Mean } (\underline{N}(t)) = N_0 \prod_{i=1}^n \text{Mean} \left(1 + \frac{\underline{a}_i t}{n}\right)$$

$$= N_0 \prod_{i=1}^n \left(1 + \frac{\mu t}{n}\right)$$

$$= N_0 \left(1 + \frac{\mu t}{n}\right)^n$$

Again we see that the mean behaviour of our process is that of the corresponding deterministic model.

The variance of a product of independent random variables can also be calculated and in this case shown to be

$$\text{Var } (\underline{N}(t)) = N_0^2 \left[ \left(1 + 2\mu \frac{t}{n} + (\sigma^2 + \mu^2) \frac{t^2}{n^2}\right)^n - \left(1 + \frac{\mu t}{n}\right)^{2n} \right]$$

One is tempted to give meaning to these results in the limit as  $n \rightarrow \infty$ , just as in the deterministic case, however this step can only be taken with great caution. The mean behaves as we would 'expect', approaching  $N_0 e^{\mu t}$ , on the other hand, unless  $\sigma^2 \sim \left(\frac{t}{n}\right)^{-1}$  the variance approaches 0 - a result which one would not accept as reasonable (?).

It is to the "limiting" case of models such as this that the theory of stochastic differential equations can apply.

---

\* Random variables will be indicated by underlining.

*Collateral reading.*

Over the last few years a great many books on stochastic differential equations have appeared. Some of these are listed below:

Arnold, Ludwig " Stochastic Differential Equations, theory and applications" Wiley, N.Y., 1974.

Gihman and Skorohod "Stochastic Differential Equations", Springer Verlag, 1972, Band 72.

Hersh, R. "Random Evolutions: A Survey of Results and Problems", Rocky Mountain J. of Maths, Vol. 4, No. 3, 1974, pp. 439-476.

McShane, E. "Stochastic Calculus and Stochastic Models" Academic Press, 1974.

Srinivasan and Basudevan "Introduction to Random Differential Equations and Their Applications", Elsevier, N.Y., 1971.