PHYS220 – SCIENTIFIC MODELLING
2001
Weeks 7-12

Department of Physics, Macquarie University
Lecturer: M.S. Wheatland (E7A306)

REFERENCES

In preparing this part of the course I have used the following references:

- Banks, R.B. (1998) *Towing Icebergs, Falling Dominoes, and Other Adventures in Applied Mathematics*, Princeton University Press: Princeton
- Bridgman, P.W. (1931) *Dimensional Analysis*, Yale University Press: New Haven
- May, R.M. (1976) *Theoretical Ecology – Principles and Applications*, Blackwell Scientific: Oxford
- Mesterton-Gibbons, M. (1989) *A Concrete Approach to Mathematical Modelling*, Addison-Wesley: Redwood City
- Prigogine, I. and Herman, R. (1971) *Kinetic Theory of Vehicular Traffic*, American Elsevier: New York
- Shepard, R. (1997) *Amateur Physics for the Amateur Pool Player* (available from: http://www.playpool.com/apapp/)
- Sivia, D.S. (1996) *Data Analysis – A Bayesian Tutorial*, Clarendon Press: Oxford

MATLAB will be used for all numerical work in this part of the course. Codes will be made available at: http://www.physics.mq.edu.au/units/phys220/lectures/msw/. For questions concerning numerical procedures, I recommend:

- Press, W.H.,Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. (1992) *Numerical Recipes in C – The Art of Scientific Computing*, Cambridge University Press: Cambridge

  The codes in different editions of this book are in different programming languages. However, the real value of the book is in its explanation of numerical methods, which can be appreciated without knowledge of the specific language.

CONTENTS:

1. THE MODELLING PROCESS :

Scientific modelling consists of the application of scientific principles to the solution of problems. An important part of this process is the construction of mathematical models, i.e. mathematical representations of the behaviour of a system. There are no absolute principles in modelling, but there are recommended procedures, & in this chapter we consider first the steps that scientists usually take in modelling. Emphasis is placed on the idea of modelling as a process, involving the refinement of mathematical models & the continual testing of assumptions. Two procedures are introduced that often yield useful results without detailed modelling: dimensional analysis & order of magnitude estimation. Finally we consider the process of model validation.

1.1 Steps in modelling:

The process of modelling usally proceeds thru a number of steps:

Step 1. Identifying a problem: A phenomenon of interest is identified that requires description or explanation. Problems are often identified from data analysis.

SRR 2. Formulating a model :

- May require hypothesizing a mechanism to explain the phenomenon
- Dimensional analysis / o.o.m. estimates may be helpful at this stage (described below)
- Quantitative description may require a mathematical model. Formulation may involve physics, biology, chemistry, etc., or the model may be more abstract (e.g. game of life)
- Start with the simplest possible model (minimum number of parameters, simplest assumptions)
- Model may have "free parameters", i.e. quantities with values not prescribed by the model alone. Models without free parameters are preferred: "occam's razor"

SRR 3. Solving / analysing the model :

- Involves analytic / numerical work. Analytic solutions are more informative than numerical ones, so (if possible) the model should be simplified until an analytic solution is found

Step 4. Testing the model :

- Model should produce results which can be tested against observations. Ideally, also lead to new predictions which, if confirmed, lend further support to the model.

- All models are idealisations, involving assumptions & approximations.

- If the model does not reproduce the observations go back to Step 2., or examine assumptions / simplifications made in Step 3.

- Even if the model reproduces the observations, it may

    - make an incorrect prediction
    - be flawed in some way

  & it may be necessary to return to Step 2.

- A new observation may turn out to be in conflict with an 'accepted' model. If so, ...

The process of testing (often called 'validation') is non trivial & requires comment.

- First, check if the model qualitatively reproduces the data (i.e. has the correct general behaviour)

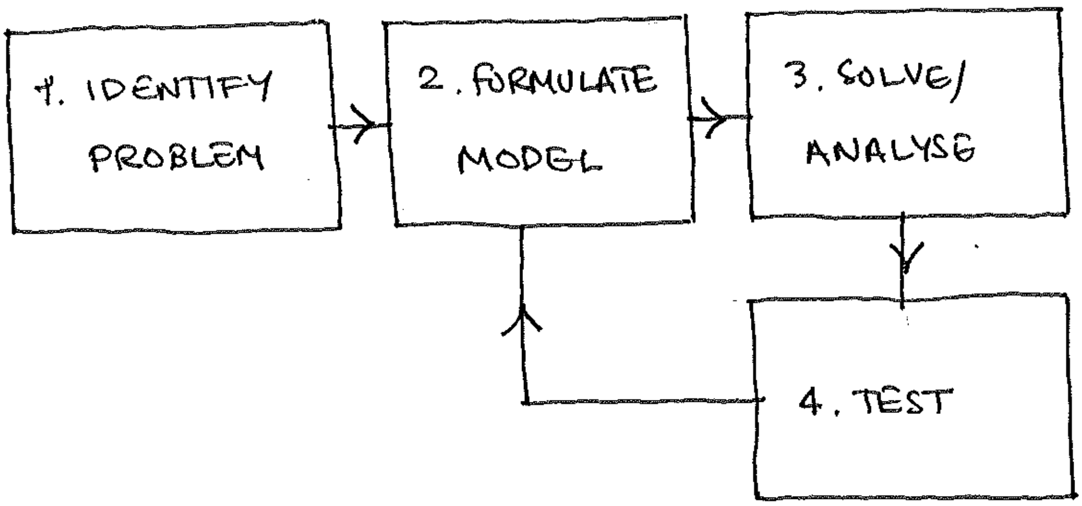
- Two quantitative problems:

  - PARAMETER ESTIMATION (determining the free parameters of the model, given the observations)

  - HYPOTHESIS TESTING (deciding whether the model is correct, given the observations).

We will consider a new, generic approach to these problems, based on Bayes' theorem, as well as considering the classical method.

In summary we can sketch a 'flow chart' of the modelling process:



Two points should be stressed:

- modelling is iterative
- modelling is motivated by problems

Next we consider two techniques that may provide the solution to a problem without detailed modelling.

1.2 Dimensional analysis

- Science deals with physical quantities (variables & constants), which have associated <u>units</u> based on the measurements required to determine the quantities.

  e.g. - distance is measured in m, cm, etc.

  - velocity is constructed from $\frac{\text{distance}}{\text{time}}$, & so has ~~dimens~~ units ms^{-1}, kmh^{-1}, etc.

- The numerical value of a quantity depends on the units

  e.g. the distance $\ell = 1m = 100cm$

- There are a certain minimum number of units needed to describe a given set of quantities. For most problems we choose basic units in mass, length & time.

- The dimension of a quantity is the combination of basic units associated with the quantity. We write M, L, T for the dimensions of masses, lengths & times.. Then e.g. the dimensions of velocity are LT^{-1}

  We write $[\text{velocity}] = LT^{-1}$

Exercise : What are the dimensions of energy?

A : $[\text{energy}] = [\text{force} \times \text{distance}]$

$= [\text{mass} \times \text{acc'n} \times \text{distance}]$

$= M(LT^{-2})L = ML^2T^{-2}$

- Dimensions impose constraints on the functional forms of equations relating physical quantities. For example, consider

$$A = \ell^2 \qquad (1)$$

relating the area A of a square to the length ℓ of a side. We have $[A] = [\ell^2] = L^2$, i.e. area has dimensions of length squared.

The general formula for the area of a pentagon cannot be

$$A = 5\ell^3 \qquad \times$$

because the dimensions of the two sides of the equation do not match.

- Principle of dimensional homogeneity:

  all terms in an equation
  relating physical quantities
  should have the same dimension

- This principle requires some equations to include dimensional constants, described in more detail below.

This principle is useful for finding errors in equations arrived at after lengthy algebra: if a term has the wrong dimensions it is incorrect.

~~It follows from the principle of dim. homog.~~
~~A related feature of (1) is~~ that if we
A related feature of (1)
change the units of the quantities, the same equation works

e.g. $\ell = 10\,\text{cm}$: area $A = \ell^2 = (10)^2\,\text{cm}^2$
$= 100\,\text{cm}^2$

$\ell = 0.1\,\text{m}$: area $A = \ell^2 = (0.1)^2\,\text{m}^2$
$= 0.01\,\text{m}^2$

& these are equivalent.

Bridgman refers to this property of equations relating physical quantities as "completeness". All equations ~~representing~~ relating physical quantities can be put in complete form.

Next note that (1) can be rewritten

$$\frac{A}{\ell^2} - 1 = 0$$

which is a non-dimensional form. Introducing the dimensionless variable

$$\pi = A/\ell^2 \tag{2}$$

we can write

$$F(\pi) = \pi - 1 = 0. \tag{3}$$

Notice that our original formula involving two dimensional variables (A & ℓ) is now replaced by a relationship involving only one dimensionless variable (π), namely the statement that it is a constant (unity).

This may seem trivial, but now we present the generalisation of these ideas, known as the Pi theorem (Buckingham, 1914).

Suppose we have n quantities $v_1, v_2, \ldots, v_n$ involving a total of m dimensions. Suppose further that we have a complete physical equation among the ~~vari~~ quantities:

$$f(v_1, v_2, \ldots, v_n) = 0$$

$$\cancel{v_1 = f(v_2, \ldots, v_n).} \qquad (4)$$

This relationship is equivalent to a dimensionless relationship among $n-m$ dimensionless variables $\pi_1, \pi_2, \ldots, \pi_{n-m}$:

$$F(\pi_1, \ldots, \pi_{n-m}) = 0 \qquad (5)$$

some function

where the π_i are formed by taking products of powers of the original quantities.

A proof of the Pi theorem is provided on a sheet.

Bridgman's proof of the Pi theorem

Reference: Bridgman, P.W., *Dimensional Analysis*, Yale University Press, 1931

Assume that there are n quantities $\alpha, \beta, ..., \xi, \zeta$ involving the m units $m_1, m_2, ..., m_m$, such that the dimensions of the n quantities are

$$\begin{aligned} [\alpha] &= m_1^{\alpha_1} m_2^{\alpha_2} ... m_m^{\alpha_m}, \\ [\beta] &= m_1^{\beta_1} m_2^{\beta_2} ... m_m^{\beta_m}, \quad \text{etc.} \end{aligned} \tag{1}$$

Consider the effect of decreasing the size of the units $m_1, ..., m_m$ by factors $x_1, ..., x_m$ respectively. The new numerical values of the quantities $\alpha, \beta, ...$ will be $\alpha', \beta', ...$, where

$$\begin{aligned} \alpha' &= x_1^{\alpha_1} x_2^{\alpha_2} ... x_m^{\alpha_m} \alpha, \\ \beta' &= x_1^{\beta_1} x_2^{\beta_2} ... x_m^{\beta_m} \beta, \quad \text{etc.} \end{aligned} \tag{2}$$

Suppose that we have a "complete equation" among the variables, i.e. a relation

$$\phi(\alpha, \beta, ...) = 0 \tag{3}$$

that retains its functional form for any choice of the size of the units of the quantities. Bridgman argues that any equation that describes a physical relationship can be put into a complete form. By definition we must have

$$\phi(\alpha', \beta', ...) = 0, \tag{4}$$

or

$$\phi\left[(x_1^{\alpha_1} x_2^{\alpha_2} ... x_m^{\alpha_m})\alpha, (x_1^{\beta_1} x_2^{\beta_2} ... x_m^{\beta_m})\beta, ...\right] = 0. \tag{5}$$

Differentiating (5) partially with respect to x_1 gives

$$(\alpha_1 x_1^{\alpha_1 - 1} x_2^{\alpha_2} ...)\alpha\phi_1\left[(x_1^{\alpha_1} ...)\alpha, ...\right] + (\beta_1 x_1^{\beta_1 - 1} x_2^{\beta_2} ...)\beta\phi_2\left[(x_1^{\alpha_1} ...)\alpha, ...\right] + ... = 0, \tag{6}$$

where the subscripts $1, 2, ...$ to ϕ denote differentiation with respect to that argument. Setting all of the x_i equal to unity gives

$$\alpha_1 \alpha \frac{\partial \phi}{\partial \alpha} + \beta_1 \beta \frac{\partial \phi}{\partial \beta} + ... = 0. \tag{7}$$

Next introduce the new variables

$$\alpha'' = \alpha^{1/\alpha_1}, \quad \beta'' = \beta^{1/\beta_1}, ... \tag{8}$$

which evidently make α'' etc. of first degree in m_1. Under this change of variable the derivatives in (7) transform according to

$$\frac{\partial}{\partial \alpha} = \frac{d\alpha''}{d\alpha} \frac{\partial}{\partial \alpha} = \frac{\alpha''}{\alpha_1 \alpha} \frac{\partial}{\partial \alpha''}, \quad \text{etc.} \tag{9}$$

and hence (7) becomes

$$\alpha'' \frac{\partial \phi}{\partial \alpha''} + \beta'' \frac{\partial \phi}{\partial \beta''} + ... = 0. \tag{10}$$

Next introduce the new variables

$$z_1 = \frac{\alpha''}{\zeta''}, \quad z_2 = \frac{\beta''}{\zeta''}, ... \quad z_{n-1} = \frac{\xi''}{\zeta''}, \tag{11}$$

where $\zeta'' = \zeta^{1/\zeta_1}$, and ζ is the nth of our original variables. The new variables are dimensionless in m_1, by construction. In terms of the new variables

$$\phi(\alpha'', \beta'', ..., \zeta'') = \phi(z_1\zeta'', z_2\zeta'', ..., z_{n-1}\zeta'', \zeta''). \tag{12}$$

The function on the RHS of (12) can be shown to be independent of ζ'':

$$\begin{aligned} \frac{\partial \phi}{\partial \zeta''} &= z_1\phi_1 + z_2\phi_2 + ... + z_{n-1}\phi_{n-1} + \phi_n \\ &= (\alpha''\phi_1 + \beta''\phi_2 + ... + \zeta''\phi_n)/\zeta'' \\ &= 0, \end{aligned} \tag{13}$$

using (10). Hence $\phi(\alpha'', \beta'', ..., \zeta'')$ can be written as a function of the $n-1$ variables $z_1, z_2, ..., z_{n-1}$, say

$$\phi(\alpha'', \beta'', ..., \zeta'') = \Psi(z_1, z_2, ..., z_{n-1}), \tag{14}$$

where the arguments z_i are dimensionless in the unit m_1.

This argument can be repeated for the unit m_2, taking as the starting point the function $\Psi(z_1, z_2, ..., z_n - 1)$, and then repeated for m_3, etc. The result is the "Pi theorem," stated below.

Pi theorem: If the equation $\phi(\alpha, \beta, ..., \zeta) = 0$ describes a complete equation among n variables $\alpha, \beta, ..., \zeta$ involving m units, then it is equivalent to a relationship

$$F(\Pi_1, \Pi_2, ..., \Pi_{n-m}) = 0 \tag{15}$$

among $n - m$ dimensionless variables $\Pi_1, \Pi_2, ..., \Pi_{n-m}$ that are formed by taking products of the original variables.

comments :

- The ~~basic~~ important point is that the number of basic variables is reduced by the number of dimensions in the problem. This has important consequences: it tells us how many variables are needed to model a system.
- ~~The Π_i are products of powers of the v_i~~
- The Π_i are not unique : in general there are many possible choices, but there will always be $n-m$ of them.
- Eq. (5) may be rewritten

$$\Pi_1 = G(\Pi_2, \ldots, \Pi_{n-m}), \qquad (6)$$

i.e. we can solve for one non-dim. variable in terms of the others

- If there are dimensional constants in the problem,

e.g. $G = 6.67 \times 10^{-11}\ m^3 s^{-2} kg^{-1}$, (the Universal gravitational constant),

then these must be treated as a variable, i.e. included in the v_i

How does the Pi theorem apply to our simple example?

variables :	A	l
dimensions :	L^2	L

So we have $n=2$ variables in $m=1$ dimensions, & the Pi theorem tells us that there is only $n-m=1$ non-dimensional variable. The choice $\Pi = A/\ell^2$ will do (but so would $\Pi' = \ell^2/A$, $\Pi'' = (\ell^2/A)^{\frac{1}{2}}$, etc.) & then there is a relationship that this variable satisfies. For the special case of one non-dimensional variable, (5) or (6) implies

$$\Pi = \text{constant}$$

& in fact $\Pi = 1$, as we have seen.

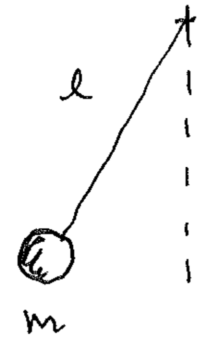
This is an example of an important special case:

- if there is one more variable than the number of units involved, then there is only one non-dimensional variable, <u>which must be constant</u>

$$\boxed{n-m=1 \quad : \quad \Pi = \text{constant}} \qquad (7)$$

Next we consider some examples of the application of the Pi theorem.

<u>Ex. 1: Simple pendulum</u>



What is the formula for the period τ of a simple pendulum?

Suppose you don't know any physics, but you suspect the period depends on the length l of the pendulum, the acc'n $g = 9.8\,ms^{-2}$ of the bob due to gravity, & the mass m of the bob, i.e. you suspect a relationship

$$\tau = f(l, g, m).$$

(Note that g, a dimensional constant, is included as a variable.)

quantities:	τ	l	g	m
dimensions:	T	L	LT^{-2}	M

?

We have $n = 4$ variables in $m = 3$ dimensions, so the Pi theorem tells us there is only one non-dimensional variable which is constant,

$$\Pi = \text{const}.$$

Since Π is a product of powers of the variables, it follows that

$$\tau = l^{\alpha} g^{\beta} m^{\gamma}.\,\text{const}$$

for some α, β, γ. Both sides of this relation must have the same dimensions, so writing

$$T = L^{\alpha}(LT^{-2})^{\beta} M^{\gamma}$$

& equating powers of like dimensions gives

$$\left.\begin{aligned} L:&\quad 0 = \alpha + \beta \\ T:&\quad 1 = -2\beta \\ M:&\quad 0 = \gamma \end{aligned}\right\} \Rightarrow \gamma = 0,\ \beta = -\tfrac{1}{2},\ \alpha = \tfrac{1}{2}$$

$$\text{so } \boxed{\tau = \text{const.}\left(\frac{l}{g}\right)^{\frac{1}{2}}} \quad (8)$$

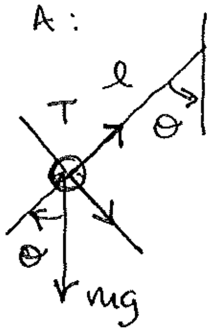
In fact the correct answer, at least for small amplitude swings, is

$$\tau = 2\pi \left(\frac{\ell}{g}\right)^{\frac{1}{2}}. \qquad (9)$$

Notice how the answer does not depend on m. Even though this parameter was included at the beginning, the method tells us that it is not in the answer.

<u>Exercise</u>: Demonstrate (9) from Newton's laws.

A:



Resolve forces $\perp$ to string:

$$mg\sin\theta = ma_{\perp} = m\ell\ddot{\theta}$$

so $\ddot{\theta} = \frac{g}{\ell}\sin\theta \simeq \frac{g}{\ell}\theta$ for small θ

This has solution $\theta = \theta_0 \sin\omega t$ where $\omega = (g/\ell)^{\frac{1}{2}}$, so $\tau = 2\pi/\omega = 2\pi(\ell/g)^{\frac{1}{2}}$.

The method fails to give us the value of the constant, & this is always the case. However, we have obtained the correct functional relationship. It would only take one measurement of a pendulum's period to work out that const $= 2\pi$. Hence DA provides a valuable head start in a phenomenological determination of the formula for τ.

The procedure followed above is standard, & it takes only a few practices to get the hang of it.

Ex. 2: Simple pendulum revisited:

What if we thought the period depended on the amplitude θ_0 of the swing (in fact it does for large swings)?

We have

$$\begin{array}{ccccc} \tau & \ell & g & m & \theta_0 \\ T & L & LT^{-2} & M & 1 \text{ (non-dim.)} \end{array}$$

i.e. 5 variables in three units. The Pi theorem tells us there are two non-dim. variables. We can take

$$\Pi_1 = \theta_0$$

$$\& \quad \Pi_2 = \frac{1}{\tau}\left(\frac{\ell}{g}\right)^{\frac{1}{2}},$$

using the results of the previous exercise. Note that the possibility of dependence on m is excluded.
Then we have

$$\Pi_2 = f(\Pi_1)$$

leading to

$$\tau = \left(\frac{\ell}{g}\right)^{\frac{1}{2}} \frac{1}{f(\theta_0)} = \left(\frac{\ell}{g}\right)^{\frac{1}{2}} g(\theta_0) \qquad (10)$$

where $g(\theta_0)$ is an arbitrary function of θ_0.

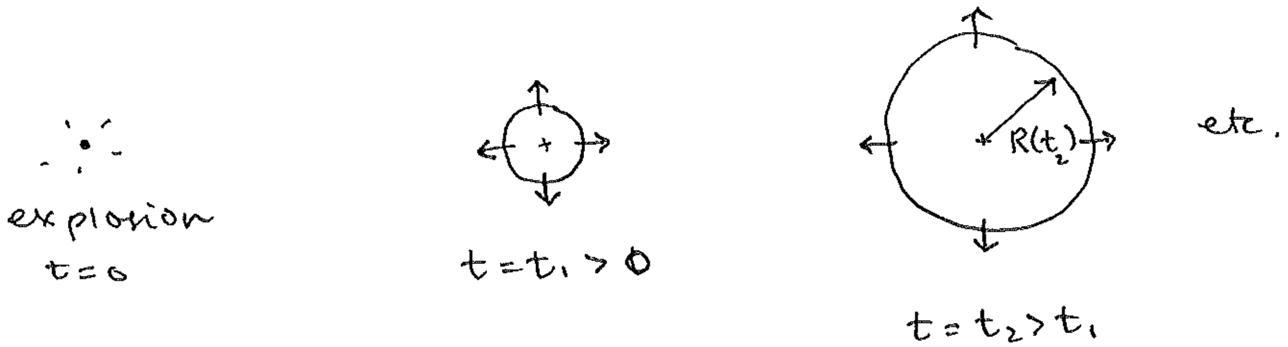
For $\theta_0 \lesssim \frac{\pi}{2}$, observations indicated that $g(\theta_0) \approx 2\pi$.[+]

+ Actually $\tau = 2\pi\left(\frac{\ell}{g}\right)^{\frac{1}{2}}\left(1 + \frac{\theta_0^2}{16} + \ldots\right)$

Next we consider a famous historical example, due to Taylor (1950).

Ex. 3 : Blast waves in air

An explosion occurs in air. The explosion produces a spherical expanding shock wave (a blast wave). How does the radius of the blast wave vary with time?



This is a non-trivial problem in compressible fluid dynamics. Interestingly DA gives the right answer with very little effort.

What are the relevant variables?

E - energy of explosion

ρ - density of air

t - time

R - radius of blast wave

And we expect a relationship $R=f(t,\rho,E)$

Dimensions:	E	ρ	t	R
	ML^2T^{-2}	ML^{-3}	T	L

So we have 4 variables in 3 dimensions, & the Pi theorem tells us there is only one non-dimensional variable

$$\pi = \text{const.}$$

$$\text{or} \quad R = \text{const.}\ E^{\alpha} \rho^{\beta} t^{\gamma}$$

Replacing quantities by their dimensions gives

$$L = (ML^2T^{-2})^{\alpha}(ML^{-3})^{\beta}\, T^{\gamma}$$

& equating powers of like dimensions:

$$L: \quad 1 = 2\alpha - 3\beta \qquad (a)$$

$$M: \quad 0 = \alpha + \beta \qquad (b)$$

$$T: \quad 0 = -2\alpha + \gamma \qquad (c)$$

(b) $\Rightarrow \beta = -\alpha$ & subst. in (a): $2\alpha + 3\alpha = 1 \Rightarrow \alpha = \frac{1}{5}$ so $\beta = -\frac{1}{5}$, & (c) $\Rightarrow \gamma = 2\alpha = \frac{2}{5}$. So:

$$\boxed{R = \text{const.}\ E^{\frac{1}{5}} \rho^{-\frac{1}{5}} t^{\frac{2}{5}}} \qquad (11)$$

which is known as the Taylor-Sedov solution. The velocity of expansion is

$$v = \frac{dR}{dt} = \frac{2}{5}.\ \text{const.}\ E^{\frac{1}{5}} \rho^{-\frac{1}{5}} t^{-3/5}. \qquad (12)$$

These solutions may be confirmed by solution of the compressible fluid equations in spherical geometry.

Taylor examined a movie of the first atomic bomb explosion at New Mexico in 1945. He plotted $\log_{10} R^{5/2} = \frac{5}{2}\log_{10} R$ versus $\log_{10} t$. According to (12) these quantities should follow a straight line with slope unity. The results are shown in Figure 1.1 : the expected dependence was observed. (What about early behaviour?)

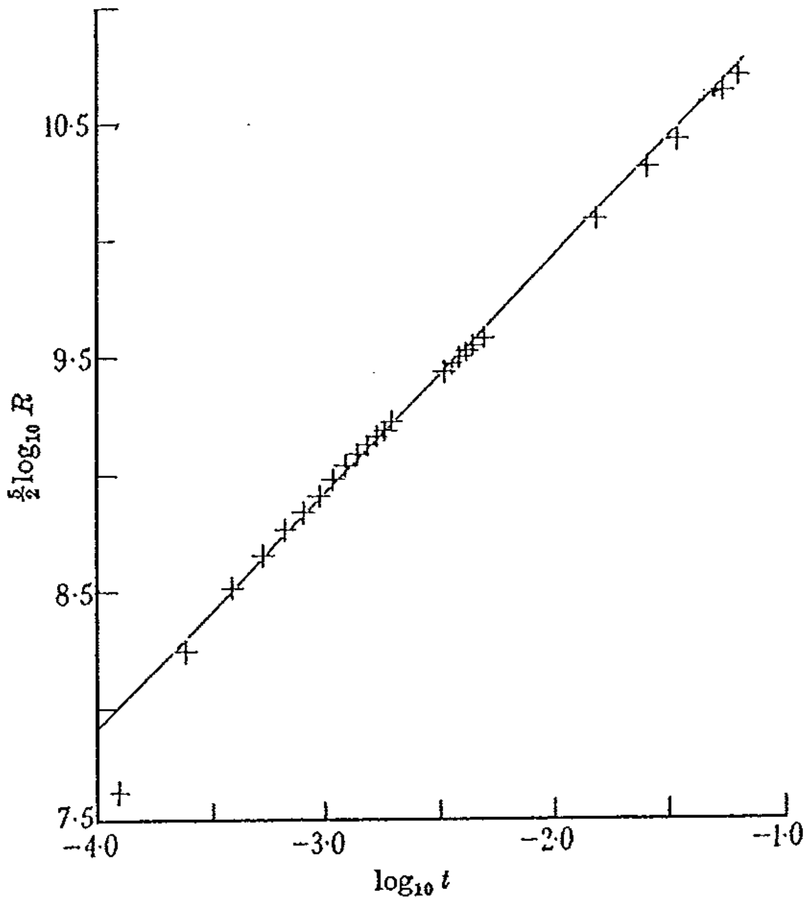
Ex.4: Kolmogorov spectrum : (maybe omit)

Another well-known example of DA relates to the spectrum of turbulence in fluids. So far we have considered examples where there is an exact analytic theory that confirms the results of DA. In the case of turbulence in fluids there is no complete theory, yet DA gives an important result which is confirmed by many observations (due to Kolmogorov, 1941) for the shape of the spectrum of fully-developed turbulence.

The idea is that in turbulence there is a cascade of energy from large scales to small scales. Energy is put into the flow on large scales & is dissipated on small scales (e.g. by viscosity), so there must be a continuous transfer of energy to smaller scales. The intermediate scales where the transfer occurs define the "inertial region".

Figure 1.1

Figure 6.4
Logarithmic plot showing that $r^{5/2}$ was proportional to t for the atomic explosion in New Mexico in 1945. From Taylor (1950a). (©The Royal Society. Reproduced with permission from *Proceedings of the Royal Society*.)



From "The Physics of Fluids & Plasmas", Choudhuri, CUP (1998)

Exercise: Estimate the energy of the explosion based on Figure 1.1

In the inertial region there are flows with length scales ℓ, for a range of ℓ. It is customary to describe these in terms of the reciprocal wavenumber $k = 1/\ell$. A flow with wavenumber k has a characteristic velocity v_k. Kolmogorov reasoned that v_k can depend only on k & the power per unit mass in the dissipation process, ε. So we have

$$v_k = v_k(k, \varepsilon)$$

where

$$[v_k] = LT^{-1}$$

$$[k] = L^{-1}$$

$$[\varepsilon] = \frac{ML^2T^{-2}}{MT} = L^2T^{-3}.$$

There are 3 variables in two dimensions, so there is only one non-dimensional variable which is constant. We can write

$$v_k = \text{const. } k^{\alpha} \varepsilon^{\beta}$$

& the usual procedure gives $\alpha = -\frac{1}{3}$, $\beta = \frac{1}{3}$:

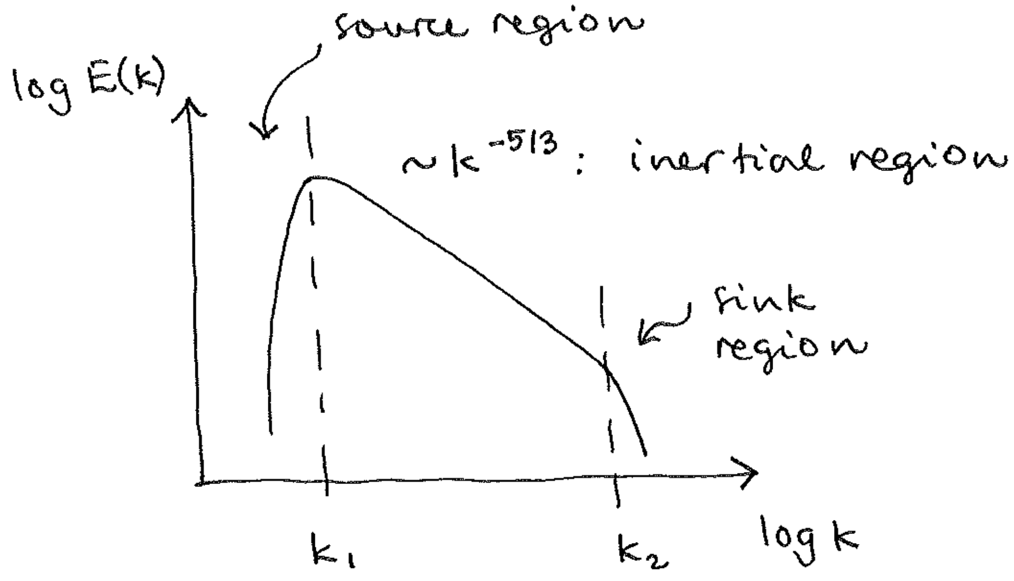
$$v_k = \text{const. } (\varepsilon/k)^{1/3}.$$

The energy per unit mass & per unit wavenumber defines the spectrum $E(k)$, so we have $E(k) \propto v_k^2/k$

$$\Rightarrow \quad E(k) \propto \varepsilon^{\frac{2}{3}} k^{-5/3} \qquad (13)$$

which is the Kolmogorov spectrum.

This result is confirmed by observations: typically turbulent spectra look like



At first sight DA may seem magical: results can be obtained without knowledge of the underlying processes. However, the choice of the correct variables in a problem may require considerable physical insight. A word of caution: if you choose the wrong variables, then the method can go astray.

You may be interested in trying the following additional examples.

Supplementary problems: (More comprehensive list is Latexed)

1. The velocity v of a wave in shallow water is hypothesized to depend on the depth d of the water, the density ρ of the water, & the acceleration due to gravity, g. Derive the functional form of the velocity using DA.

2. The difference in pressure Δp between the surface of a body of water & the pressure at depth z depends on the density ρ of water, the acceleration g due to gravity & the depth z. Use dimensional analysis to establish that

$$\Delta p = \rho g z .$$

1.3 Order of magnitude (OOM) estimation :

Sometimes we can avoid detailed modelling because an educated guess is good enough. An order of magnitude is a factor of 10. Guessing something to within a factor of 10 is often easy. We will consider a ~~few~~ couple of examples.

Ex1: What is the volume of a person?

People are mostly water, so they are only a bit denser than water :

$$\rho_{water} = 1\,g\,cm^{-3} = 10^{-3}\,kg/(10^{-6}\,m^3)$$
$$= 10^{3}\,kg\,m^{-3}$$

A typical mass of a person is, say, $m = 70\,kg$, so

$$V = \frac{m}{\rho} = \frac{70\,kg}{10^{3}\,kgm^{-3}} = 0{\cdot}07\,m^3$$
$$0{\cdot}01 - $$
$$\sim 0{\cdot}1\,m^3$$

where we use '$\sim$' to denote an OOM.

PHYS220 Scientific Modelling 2001

Exercises in dimensional analysis

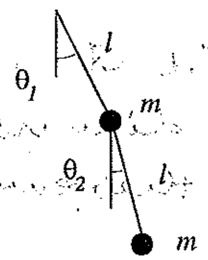
1. Assume that the velocity v of a surface wave in shallow water depends on the depth d of the water, the density ρ of water and the acceleration g due to gravity.
   (a) Derive the functional form of the velocity.
   (b) Use this result to explain why waves break as they approach the shore.

2. The difference in pressure Δp between a point at a depth z and a point at the surface of a body of water depends on the density ρ of water and the acceleration g due to gravity.
   (a) Show that $\Delta p \doteq \rho g z$.
   (b) How deep do you have to be to experience an increase in pressure of one atmosphere?

3. A medical researcher studying spermatozoa determines that the speed of propagation depends on the length l of the organism, the viscosity μ of the fluid, and the rate of expenditure of energy per unit time and per unit volume $\dot{\epsilon}$ by the organism. (The speed does not depend on the density because on the small scales of relevance, viscous forces dominate over inertial forces.)
   (a) Determine how the velocity depends on μ, l and $\dot{\epsilon}$. [Note: μ has dimensions $\mathrm{ML^{-1}T^{-1}}$.]
   (b) If a spermatazoan begins with a total energy per unit volume ϵ_0, obtain an approximate expression for how far the organism can swim.

4. At a critical angular speed of rotation Ω_c a ball of incompressible fluid that is self gravitating becomes unstable. Obtain an expression for Ω_c and show that it is independent of the radius of the ball and depends on the square root of density.

5. A double pendulum (shown) involves two equal masses m and strings of equal length l. The system can be set into motion so that the two masses exhibit steady oscillations with angular amplitudes θ_1 and θ_2, as shown.



   (a) Show that the period of this motion does not depend on m, and

(b) obtain an expression for the period. [The exact answer is that there are two such modes with periods $T_{\pm} = 2\pi(1 \pm \theta_1/\theta_2)^{1/2}(l/g)^{1/2}$, where $\theta_1/\theta_2 = 1/\sqrt{2}$.]

6. In 1908 Andrew Stephenson, a mathematics lecturer at Manchester University, showed that a simple rigid pendulum can be maintained stably in the upside-down position by oscillating the pivot up and down at a high frequency. For pivot vibrations with a small amplitude a, the critical frequency ω_c for stability depends only on a, the length l of the pendulum and the acceleration g due to gravity. Determine an expression for ω_c. [The exact answer is $\omega_c = (2gl)^{1/2}/a$.]

Comments that should be added to notes:

1. Pi theorem assumes only one rel'n $\phi(\alpha, \beta, \dots) = 0$. If the variables are related by other rel'ns, assumptions of proof not met. Then need to reduce to 1 relation.

2. The Pi theorem gives the minimum number of non-dim. variables required to ~~relat~~ represent a relation, in general. For specific relations fewer variables may suffice.

e.g. $a = b.c/d$ ⊛, where a, b, c, d have the same ~~relad~~ dimension. We have 4 var's in 1 dim, so Pi theorem says 3 non-dim var's,

e.g. $\Pi_1 = \frac{a}{b}, \quad \Pi_2 = \frac{c}{b}, \quad \Pi_3 = \frac{d}{b}$

& then ⊛ is $\Pi_1 = \Pi_2/\Pi_3 = f(\Pi_2, \Pi_3)$.

However, knowing ⊛ we could choose $\Pi_1 = a/b$, $\Pi_2 = c/d$, & then ⊛ is

$$\Pi_1 = \Pi_2$$

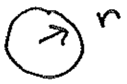
which involves only 2 non-dim. variables. So for a given relation, try non-dimensionalising to find out how many variables are needed.

Ex 2. What is the surface area of a person?

Surface area is a bit harder – we cannot appeal to mass. However, consider a cube & sphere:

cube: 

$$A = 6\ell^2, \quad V = \ell^3 \quad \Rightarrow \quad \frac{A}{V} = \frac{6}{\ell}$$

sphere: 

$$A = 4\pi r^2, \quad V = \frac{4}{3}\pi r^3 \quad \Rightarrow \quad \frac{A}{V} = \frac{3}{r}$$

We see that the ratio of area to volume for these solids is a few times the reciprocal of the characteristic dimension. If we 'model' the human body as a sphere, we can determine an effective radius from our estimate for the volume:

$$r = \left(\frac{3V}{4\pi}\right)^{\frac{1}{3}} = \left(\frac{3 \cdot 0.07}{4\pi}\right)^{\frac{1}{3}} \text{ m}$$

$$\approx 0.25\text{m} \qquad (0.27 \text{ for } 80\text{kg})$$

& then

$$A = \frac{3}{r}V = \frac{3}{0.25} \cdot 0.07 \text{ m}^2$$

$$\approx 0.84 \text{ m}^2 \qquad (0.9)$$

$$\sim 1 \text{ m}^2$$

& so a person's skin has a surface area of order 1m^2. [On the web I found 15-20 sq. ft which is $1.4\text{m}^2 - 1.9\text{m}^2$]

[Adopting the cube model, $\ell = (V)^{\frac{1}{3}} \approx 0.41\text{m}$

so $A = \frac{6}{\ell}V = \frac{6}{0.41} \cdot 0.07\text{m}^3 \approx 1\text{m}^2$.]

~~Rather than doing a lot of these,~~ in class, ~~I have produced an example sheet.~~

Ex 3. How many hairs are there on your head?

If you're Andre Agassi, easy: none. Otherwise, lets assume that at the roots the hairs are about $\ell = 1\,mm$ apart. The area of your head is $A = L^2$ where L is a characteristic linear size. Let's take $L = 30\,cm$. Then the number of hairs is

$$N = \left(\frac{L}{\ell}\right)^2 = \left(\frac{30\,cm}{0.1\,cm}\right)^2 \sim 100000$$

[On the web I found (100-150)000.]

There are a lot of questions along these lines. Rather than doing a lot in class, I have produced a sheet of exercises. The last few are examples of simple geometric modelling, & are more accurate than OOM estimates.

Next we consider the process of testing models, at first from a Bayesian perspective.

PHYS220 Scientific Modelling 2001

Exercises in order of magnitude estimation and simple modelling

1. How many rice grains would fit into a wine bottle? How many sand grains?
2. How many times will your heart beat in your lifetime?
3. Estimate the total volume of water on Earth.
4. How many cows would a meat-eater consume in their lifetime?
5. Estimate the number of piano tuners working in Sydney. [This question has been used in job interviews by Management Consultancy companies.]
6. DNA can be crudely modelled as a cylinder with radius $7.9 \times 10^{-4}\,\mu$m. A typical cell nucleus has radius $1.7\,\mu$m. If a DNA molecule is 1.5m long, what fraction of a cell nucleus is DNA?
7. Suppose you are standing on a hill 100 m above sea level looking out to sea. How far away is the horizon?
8. Around 230 BC Erastothenes estimated the radius of the Earth, based on the following observation. At noon during the northern Summer solstice (i.e. when the Sun follows its most northerly path in the sky), the Sun is directly overhead in Syene, whereas at the same time the Sun's rays are inclined to the vertical by 7.2 degrees in Alexandria, which is 787 km due north of Syene. Reproduce Erasthothenes' estimate. [The actual radius of the Earth is about 6380 km. What is the percentage error in Erasthothenes' result?]
9. Aristarchus was a contemporary of Erastothenes, who estimated the distance to the Moon in the following way. The duration of a lunar eclipse (when the moon passes through the Earth's shadow) is about 3 hours. Use this value to estimate the ratio of the radius of the Moon's orbit to the radius of the Earth. Hence use Erastothenes' value for the radius of the Earth to obtain the distance to the Moon. [The average value for the orbital distance is 384000 km. What is the percentage error in Aristarchus' result?]

1.4 Testing models

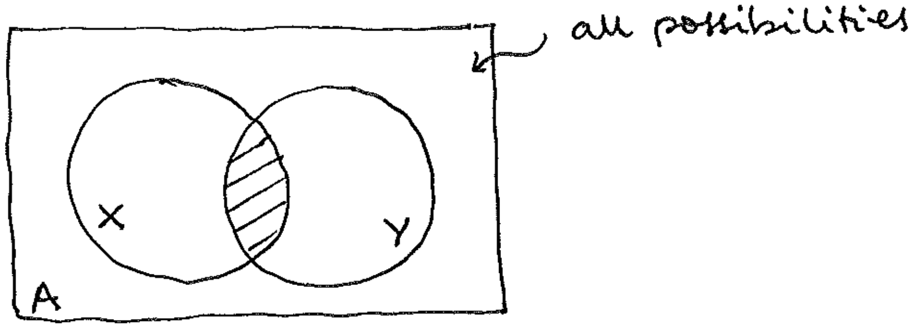
1.4.1 Bayesian inference :

We begin with some results from conditional probability. If prob (X,Y) denotes the probability that two propositions are both true (the comma is read as "and"), then

$$\text{prob}(X,Y) = \text{prob}(X|Y)\,\text{prob}(Y) \qquad (14)$$

where prob (X|Y) denotes the probability that X is true, given that Y is true. The quantity prob(X|Y) is a conditional probability.

A justification of (14) follows from a diagram:



The areas correspond to probabilities, e.g. the area of the circle Y divided by the area of the rectangle A is prob(Y). Clearly:

$$\text{shaded area} = \text{fraction of Y shaded} \times \text{area of Y}$$

or

$$\text{shaded area}/\text{area A} = \text{fraction of Y shaded} \times \frac{\text{area of Y}}{\text{area of A}}$$

The quantity on the LHS is prob(X,Y) & the fraction of Y that is shaded is prob(X|Y), so we have established (14).

Similarly we can write

$$\text{prob}(X,Y) = \text{prob}(Y|X)\cdot\text{prob}(X) \qquad (15)$$

& combining (14) & (15) we have

$$\text{prob}(X|Y) = \frac{\text{prob}(Y|X)\,\text{prob}(X)}{\text{prob}(Y)}, \qquad (16)$$

which is one version of Bayes' theorem, due to Reverend Thomas Bayes (1763). Bayes pointed out how to use this identity for scientific inference, as follows.

Make the identifications

$X \rightarrow$ hypothesis or model, H

$Y \rightarrow$ available data, D

& then we have

$$\boxed{\text{prob}(H|D) \propto \text{prob}(D|H)\cdot\text{prob}(H).} \qquad (17)$$

It is usual to neglect the term $1/\text{prob}(D)$ leaving a proportionality rather than an equality. The quantity on the LHS is a probability, so if it is summed/integrated over all possible hypotheses the result is unity, & this procedure is generally used to determine the constant of proportionality.

The terms in (14) are given names:

- prob(H) is the "prior probability": this is the state of knowledge before we consider the data
- prob(H|D) is the "posterior probability": this is the state of knowledge in the light of the data.
- prob(D|H) is the "likelihood function": this is the probability of getting ~~that~~ the data that we did, given that the model is true.

Equation (14) may be used directly for parameter estimation - we will consider an example shortly. Regarding hypothesis testing, consider the situation ~~where there are~~ of two competing hypotheses, H_1 & H_2. Taking ratios of Bayes' theorem applied in each case gives

$$\frac{\text{prob}(H_1|D)}{\text{prob}(H_2|D)} = \frac{\text{prob}(D|H_1)}{\text{prob}(D|H_2)} \cdot \frac{\text{prob}(H_1)}{\text{prob}(H_2)} . \quad (18)$$

Note that in this case the term prob(D) is common & has cancelled: so we have equality again.

Eq.(18) tells us that the relative probability of the models is the ratio of the likelihoods, modulated by the ratio of the prior probabilities. If the two models are considered to be a priori equally likely, then it may be

reasonable to assume $prob(H_1) = prob(H_2)$.
Eq. (18) provides a way to do hypothesis testing, provided we can specify the competing hypotheses.

Now we consider in details a problem in parameter estimation, using Bayes' theorem.

Example: Is a given coin fair?

Suppose you have a coin that you acquired at Star City Casino. In 10 tosses you observe 3 heads. Is the coin fair?

Let H denote the probability of obtaining a head with one toss. We will call this the bias of the coin. If $H = \frac{1}{2}$ the coin is fair.

Bayes' theorem tells us

$$prob(H|D) \propto prob(D|H)\, prob(H)$$

where now we interpret $prob(H|D)$ as a differential distribution, i.e. $prob(H|D)\,dH$ is the probability that the bias is in the range $(H, H+dH)$.

We need to decide on $prob(H)$, the prior. Since the coin is from a ~~[illegible]~~ casino it might be prudent to allow all possibilities. Hence we ~~will~~ consider a "uniform prior",

$$prob(H) = \begin{cases} 1 & 0 \le H \le 1 \\ 0 & \text{otherwise.} \end{cases} \tag{19}$$

The probability of obtaining r heads in n tosses is given by the Binomial theorem:

$$\text{prob}(r|n) = \frac{n!}{r!(n-r)!} H^r (1-H)^{n-r}.$$

This expression is easy to understand: $H^r(1-H)^{n-r}$ is the probability of obtaining r heads & $(n-r)$ tails [since H is the probability of obtaining a head, $1-H$ is the probability of a tail in one toss] & the factor out the front is the number of ways in which this outcome (r heads, n tails) can be achieved (without regard to the actual order of heads, tails).

We take r heads in n tosses to be the available data, so

$$\text{prob}(D|H) \propto H^r (1-H)^{n-r} \tag{20}$$

where the factor out the front can be neglected since it does not depend on H.

Putting the pieces [(19) & (20)] into Bayes' theorem [(17)] gives the posterior probability:

(21)

$$\text{prob}(H|D) \propto \begin{cases} H^r(1-H)^{n-r} & 0 \le H \le 1 \\ 0 & \text{otherwise} \end{cases} \tag{21}$$

To turn this into an equality we just need to enforce $\int_0^1 \text{prob}(H|D)\, dH = 1$. Hence Eq. (21) gives us a way to assign a probability to the hypothesis that a coin has a given bias, based on the result of n tosses.

Let's see how this works in practise. A code (biased-coin.m : a Matlab script) has been written to simulate tosses of a biased coin. The bias is $H = 0.25$ (one quarter of the time the coin comes down heads). A certain number of tosses of the coin are simulated, & at each toss the posterior probability is evaluated. The prior probability is taken to be uniform. Figure 1.2 shows the posterior probability dist'n function (pdf) after 10 tosses & after 100 tosses. After 10 tosses the posterior pdf is broad but after 100 tosses the pdf is a narrow peak around $H = 0.25$.

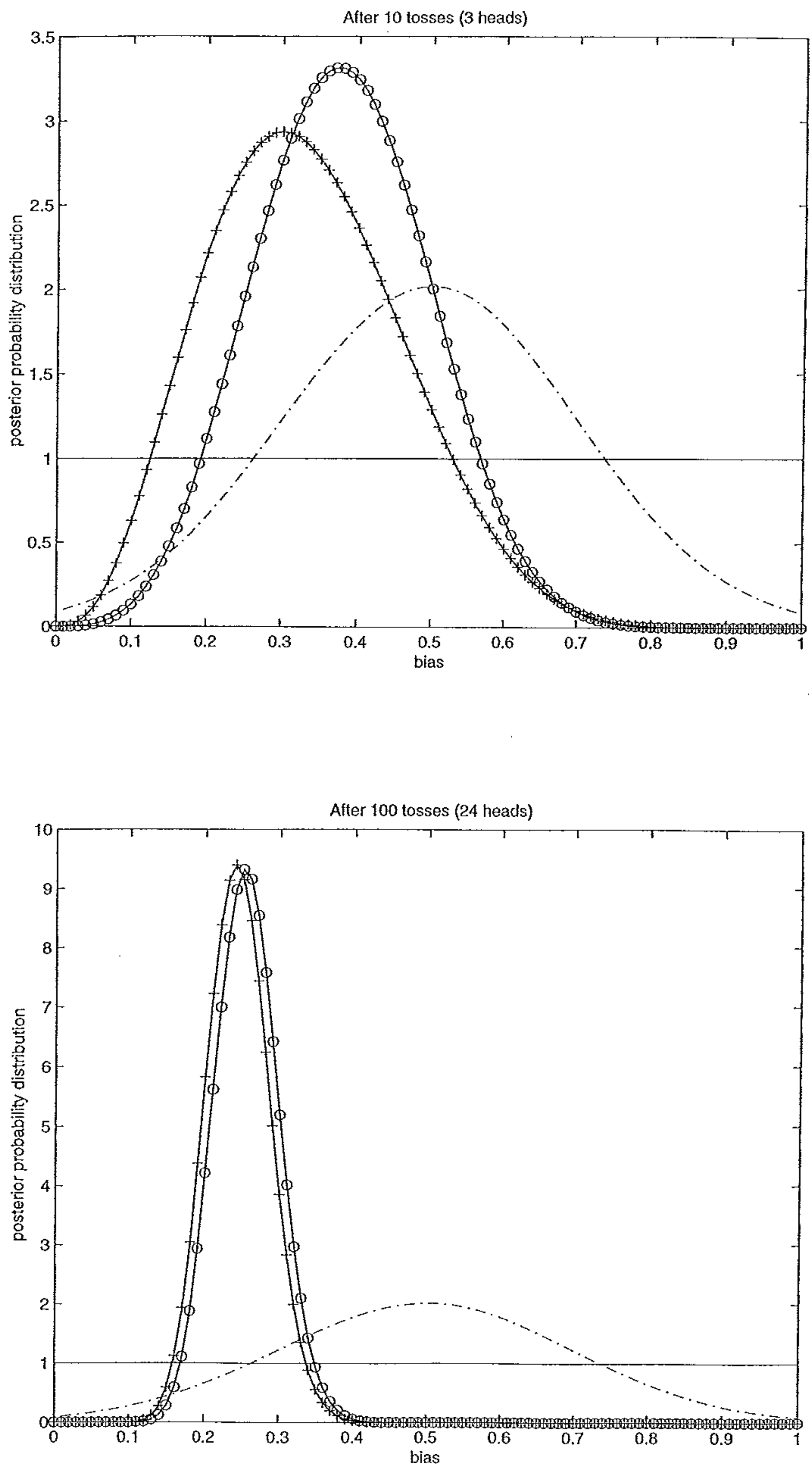
Obviously the numerical value of prob(H|D) depends on the choice of prior: however, for large numbers of tosses the same results are obtained independent of the choice of prior. To see this the simulation also works out the posterior pdf for a Gaussian prior (the dot-dashed curve), which asserts that the coin is most likely fair, & unlikely to have extreme values of H. After 10 tosses the posterior pdf is different from that found for the uniform prior, but after 100 tosses they are similar.

- If we wanted to we could make a "best estimate" of the value of H after a certain number of tosses: this is the maximum of the posterior pdf. We could also assign an uncertainty, based on the width of the pdf.

- To address the original question, based on observing 3 heads in 10 tosses it is not really possible to make a definitive statement about whether the coin is biased (e.g. in the top panel of Figure 1.2, the value of $H = \frac{1}{2}$ still has a large associated value of the posterior pdf). However, if we tossed the coin another 90 times, we could make fairly confident statements.

You are welcome to run this code yourself, & try modifying it: e.g. try different biases & different priors. The results are ~~also~~ animated, so you watch the evolution of the posterior pdf. The address of the code is given on the first sheet I handed out, or else you can get to it from the page for this course.

Figure 1.2


After 10 tosses (3 heads)
posterior probability distribution
bias
After 100 tosses (24 heads)
posterior probability distribution
bias

```
% biased_coin.m
%
% Bayesian treatment of a biased coin, following  D.S. Sivia "Data Analysis
% (A Bayesian Tutorial)," Clarendon Press, Oxford, 1996. A certain number of
% tosses of a biased coin are simulated. Based on the results, the posterior
% probability of the coin having a given bias is calculated, for a uniform
% prior and for a Gaussian prior.
%
% M.S. Wheatland, 20 April 2001


NTOSS=100; % number of  tosses
BIAS=0.25; % bias in coin
STEP=0.01; % plotting step
SIGMA=0.2; % width of Gaussian prior

% generate uniform deviates

x=rand(1,NTOSS);

bvals=[0:STEP:1]; % bias values for plot
nn=size(bvals);
nn=nn(2);

prob_unif=0*bvals;
prob_gauss=0*bvals;
prior_unif=ones(nn);
prior_gauss=exp(-(bvals-0.5).*(bvals-0.5)/(2*SIGMA^2));
prior_gauss=prior_gauss/(STEP*sum(prior_gauss)); % normalise

r=0; % counter for number of heads
for i=1:NTOSS, % at each toss evaluate posterior probabilities...

  if(x(i)<BIAS) % if toss is heads
    r=r+1;
    n=i;
  else % if toss is tails
    n=i;
  end
  for j=1:nn, % work out posterior probabilities
    prob_unif(j)=prior_unif(j)*(bvals(j)^r)*(1-bvals(j))^(n-r);
    prob_gauss(j)=prior_gauss(j)*(bvals(j)^r)*(1-bvals(j))^(n-r);
  end

  nprob_unif=prob_unif/(STEP*sum(prob_unif)); % normalise
  nprob_gauss=prob_gauss/(STEP*sum(prob_gauss)); % normalise

  % plot
  plot(bvals,nprob_unif,'+',bvals,nprob_unif,bvals,nprob_gauss,'o',bvals,...
    nprob_gauss,bvals,prior_unif,'-',bvals,prior_gauss,'-.') %
  title(['After ',num2str(i),' tosses (',num2str(r),' heads)'])
  xlabel('bias')
  ylabel('posterior probability distribution')
  pause(0.001) % pause to allow animation effect

end

% reset random number generator, so different results are obtained

rand('state',sum(100*clock))
```

1.4.2 Maximum likelihood & least squares:

Two common procedures ~~used~~ for parameter estimation are "maximum likelihood" & "least squares". How are these methods related to the Bayesian approach?

Suppose that we have a model involving M free parameters $X_1, X_2, \ldots, X_M$, which we will write briefly as $\underset{\sim}{X}$. Also we have N measured data $D_1, D_2, \ldots, D_N$, or $\underset{\sim}{D}$. Bayes' theorem tells us

$$\text{prob}(\underset{\sim}{X}|\underset{\sim}{D}) \propto \text{prob}(\underset{\sim}{D}|\underset{\sim}{X})\,\text{prob}(\underset{\sim}{X})$$

If we are ignorant about the relative merit of different models (different choices of $\underset{\sim}{X}$) then we might assume (ASSUMPTION 1) a uniform prior:

$$\text{prob}(\underset{\sim}{X}) = \text{const.}$$

compare, for example, the coin problem.
← In this case we have

$$\text{prob}(\underset{\sim}{X}|\underset{\sim}{D}) \propto \text{prob}(\underset{\sim}{D}|\underset{\sim}{X}). \qquad (22)$$

The best estimate of $\underset{\sim}{X}$ is the value $\underset{\sim}{X}_0$ for which the posterior probability $\text{prob}(\underset{\sim}{X}|\underset{\sim}{D})$ is a maximum. Clearly according to (22) this is also the value of $\underset{\sim}{X}$ which makes the likelihood $\text{prob}(\underset{\sim}{D}|\underset{\sim}{X})$ a maximum, & for this reason, in this context $\underset{\sim}{X}_0$ is called the "maximum likelihood estimate."

Assuming the data are independent (ASSUMPTION 2) we can write

$$\text{prob}(\underset{\sim}{D}|\underset{\sim}{X}) = \text{prob}(D_1|\underset{\sim}{X})\,\text{prob}(D_2|\underset{\sim}{X}) \ldots \times \text{prob}(D_N|\underset{\sim}{X}) \quad (23)$$

or $$\text{prob}(\underset{\sim}{D}|\underset{\sim}{X}) = \prod_{i=1}^{N} \text{prob}(D_i|\underset{\sim}{X}), \quad (23')$$

where Π denotes a product, in the same way Σ denotes a sum.

Next we introduce the data $\underset{\sim}{F}$ that the model would produce given the parameters $\underset{\sim}{X}$, & in the absence of noise:

$$F_i := f(\underset{\sim}{X}, i). \quad (24)$$

The noise associated with the i^{th} datapoint is described by a standard deviation σ_i.

If the noise can be described as Gaussian (ASSUMPTION 3) then the probability of an individual datapoint is

$$\text{prob}(D_i|\underset{\sim}{X}) = \frac{1}{(2\pi)^{\frac{1}{2}}\sigma_i}\, e^{-(F_i-D_i)^2/2\sigma_i^2}. \quad (25)$$

(This just says that the observed datapoint is distributed like a Gaussian around the true value.).

Putting (23') & (25) together we have

$$\text{prob}(\underset{\sim}{D}|\underset{\sim}{X}) \propto \prod_{i=1}^{N} e^{-(F_i-D_i)^2/2\sigma_i^2}$$

$$\text{or} \quad \text{prob}(\underset{\sim}{D}|\underset{\sim}{X}) \propto \exp\left[-\sum_{i=1}^{N} \frac{(F_i - D_i)^2}{2\sigma_i^2}\right]$$

$(e^a e^b \cdots = e^{a+b+\cdots})$

$$\text{i.e.} \quad \text{prob}(\underset{\sim}{D}|\underset{\sim}{X}) \propto \exp(-\chi^2/2) \qquad (26)$$

where

$$\chi^2 \equiv \sum_{i=1}^{N} \frac{(F_i - D_i)^2}{\sigma_i^2} \qquad (27)$$

is "chi-squared".

Using (26), (27) & (22), the natural logarithm of the posterior probability is then

$$L = \ln[\text{prob}(\underset{\sim}{X}|\underset{\sim}{D})] = \text{const} - \frac{\chi^2}{2}, \qquad (28)$$

where the constant is the logarithm of the constant of proportionality in (22). The maximum of L occurs when χ^2 is a minimum: the optimal solution $\underset{\sim}{X}_0$ is called the "least squares estimate" because it is obtained by minimizing χ^2, which involves a sum of squares.

<u>Least squares Procedure</u>: minimise χ^2 to determine $\underset{\sim}{X}_0$.

Note that the least squares estimate is the maximum likelihood estimate subject to the additional assumptions of Gaussian errors & independent datapoints. The maximum likelihood estimate is also the Bayesian solution (i.e. provides a maximum posterior probability), provided a uniform prior is accepted.

These comments show how maximum likelihood & least squares estimates are approximations to the Bayesian solution. For this reason the Bayesian solution is preferred.

We have not considered how to minimise χ^2 to obtain $\underset{\sim}{X}_0$. If the function (24) is linear in $\underset{\sim}{X}$ then exact analytic solutions are possible: otherwise numerical methods are needed: consult Numerical Recipes, for example.

1.4.3 Classical hypothesis testing

We have seen how to do hypothesis testing in the Bayesian scheme, following Eq. (18). Note that it is necessary to specify the competing hypotheses, H_1 & H_2.

There is also a classical approach to hypothesis testing, that we will briefly discuss. In this approach only one hypothesis, H, is considered, & the method is meant to compare H & $\bar{H}$ ("not H").

In the discussion in the previous section we assumed that the model is correct for some choice of $\underset{\sim}{X}$. If the minimum value of χ^2 is large (in some sense) then perhaps the ~~the~~ model is wrong, or maybe the errors have been underestimated, or the data are just plain wrong. In classical hypothesis testing the latter possibilities are neglected, & the value of χ^2 is used to test the hypothesis that the model is correct. (other statistics than χ^2 are also commonly used, but the approach is the same.)

PROCEDURE (The "χ^2 test")

- Determine $\underset{\sim}{X}_0$ & hence $\chi_0^2 \equiv \chi^2(\underset{\sim}{X}_0)$, the minimum value of χ^2.

- Calculate the "significance", $P_{\bar{N}}(\chi^2 > \chi_0^2)$. This is the probability of obtaining a larger value of χ^2 than χ_0^2, assuming the model is correct (with parameters $\underset{\sim}{X}_0$). The quantity $\bar{N} = N - M$ is called the "number of degrees of freedom." Consult a reference (e.g. NR) for ways to calculate the significance.

- If $P_{\bar{N}}(\chi^2 > \chi_0^2)$ is sufficiently small, the model is "rejected." For example, if $P_{\bar{N}}(\chi^2 > \chi_0^2) = 0.009$ then "the model is rejected at the 1% significance level."

A few comments on this procedure are in order.

1. In the context of the procedure it is not possible to prove the hypothesis, only reject it. A small value of χ_0^2 ($P_{\bar{N}}(\chi^2 > \chi_0^2) \simeq 1$) does not mean the model is correct.

2. Bayesians reject this procedure altogether. They argue that

   - χ^2 is a measure of prob(D|H) rather than prob(H|D), which is what you want, &
   - even if the model is "rejected," no alternative hypotheses have been given, & so the procedure is not helpful. (For example, an alternative hypothesis is that the data are faulty. What then?)

I urge you to consider these arguments. In my opinion the classical tests serve some purpose, although it is important to be aware of their shortcomings.

Finally, some comments on the numerical content of this part of the course. I will use MATLAB for numerical work, as needed. Terry is providing an introduction to MATLAB, as well as spending some time looking at a number of numerical methods. My lectures concentrate on modelling & I won't address the details of numerical methods. When I use MATLAB it will be as a "black box," i.e. I assume the routines work as advertised. However, you should be sceptical of numerical results. If they look wrong, test them (either against available analytic results or by repeating calculations using other packages).

The following sections of the course present examples of modelling, illustrating the application of ideas given in this section.

2. FLUID MODELS FOR TRAFFIC

Problem: What is the maximum amount of traffic a particular highway can carry?

2.1 The fluid analogy:

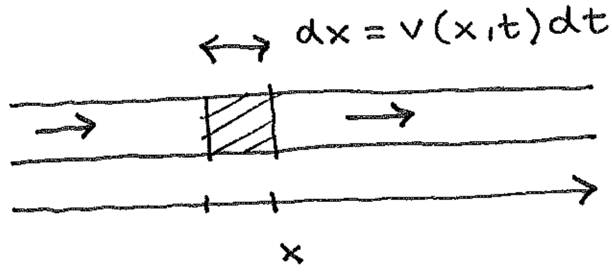
To address this problem we consider simple models in which the flow of traffic along a single lane highway is treated by analogy to the flow of a compressible† fluid in a pipe. A row of discrete moving vehicles is replaced by a continuous stream of fluid, an approach which dates back to Lighthill & Whitham (1955).

Consider a single lane highway with cars moving in the $+x$ direction. A density $\rho = \rho(x,t)$ specifies the number of cars per unit length at position x at time t. In practice this is determined by averaging the number of cars over some length around x: for example if there are N cars in a length L around x then the density at x is $\rho = \frac{N}{L}$. The average distance between cars at x is $\frac{L}{N} = \rho^{-1}$.

We can also assign an average velocity $v = v(x,t)$ to cars passing the point x at time t.

† So the fluid can be a gas but not a liquid.

How many cars pass the position x per unit time, at time t? Consider the time interval $(t, t+dt)$. In this time all cars within a distance $dx = v(x,t)\,dt$ to the left of x will pass the point x:



The number of cars in dx is $\rho(x,t)\,dx$, & hence the number passing x per unit time is

(29):
$$\mathcal{F}(x,t) = \frac{\rho(x,t)\,dx}{dt} = \rho(x,t)\,v(x,t)$$

We will call this quantity the flux (some books use "flow"), by analogy with fluid dynamics.

2.2 A basic equilibrium model:

Following the principle of considering the simplest possible model first we begin with the uniform equilibrium situation, where the density & velocity are independent of x & t. This is appropriate for a uniform stretch of highway with a constant volume of traffic.

We have two basic variables: ρ & v. It is reasonable to assume they are related, as follows. If the density of cars is low, then all cars should be able to move at the speed limit, v_{lim}, i.e.

$$\rho \to 0 \quad \Rightarrow \quad v \to v_{lim}. \qquad (30)$$

There is a physical limit to the density, namely that there can only be one car per car length ℓ. The density in this situation is $\rho = 1/\ell$, & the velocity must be zero for this density. In fact the velocity will be zero for a density less than $1/\ell$, which we will label ρ_{max}: this is the "traffic jam density":

$$\rho \to \rho_{max} \quad \Rightarrow \quad v \to 0. \qquad (31)$$

In between $\rho = 0$ & $\rho = \rho_{max}$ we expect that the equilibrium velocity will decrease with density, because drivers are forced to be more cautious when the distance to the next car is smaller.

To summarise, we have a model with four variables (ρ, v, ρ_{max}, v_{lim}), & a relationship between them:

$$v = v(\rho, \rho_{max}, v_{lim}).$$

At this point we can try dimensional analysis. The dimensions of the variables are:

$$\begin{array}{cccc} \rho & v & \rho_{max} & v_{lim} \\ L^{-1} & LT^{-1} & L^{-1} & LT^{-1} \end{array}$$

so we have 4 variables in 3 dimensions, & the Pi theorem tells us there are two non-dimensional variables. Obviously we can choose

$$\Pi_1 = \frac{\rho}{\rho_{max}}, \qquad \Pi_2 = \frac{v}{v_{lim}}$$

& we have

$$\frac{v}{v_{lim}} = F\left(\frac{\rho}{\rho_{max}}\right) \qquad (32)$$

for some function F. This is the "velocity-density relationship."

We can expand the RHS of (32) as a Taylor series:

$$\frac{v}{v_{lim}} = F(0) + \frac{\rho}{\rho_{max}} F'(0) + \frac{1}{2}\left(\frac{\rho}{\rho_{max}}\right)^2 F''(0) + \ldots \qquad (33)$$

Once again we appeal to the principle of simplicity. The quantity v/v_{lim} must depend on ρ/ρ_{max}, so the simplest choice is to keep just the first two terms on the RHS, i.e. consider a linear model

$$\frac{v}{v_{lim}} = F(0) + \frac{\rho}{\rho_{max}} F'(0).$$

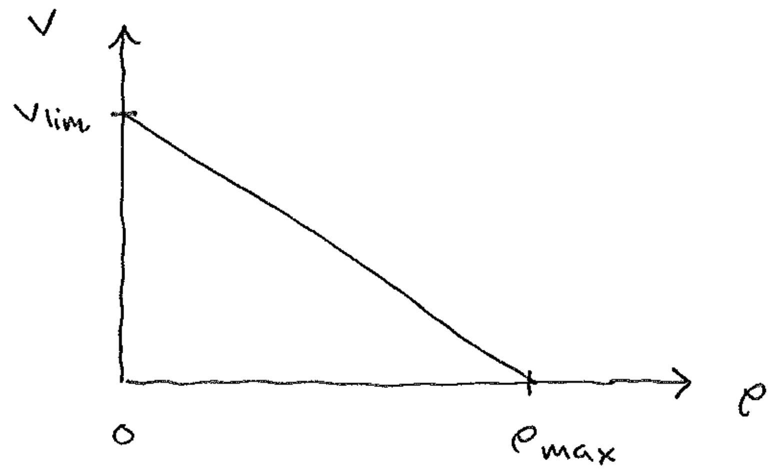
Imposing the constraints (30) & (31) implies $F(0) = 1$ & $F'(0) = -1$, & hence

$$v = v_{lim}(1 - \rho/\rho_{max}).$$

This is appropriate for $\rho \leq \rho_{max}$: more generally we write

$$(34): \quad v = \begin{cases} v_{lim}(1 - \rho/\rho_{max}) & \rho \leq \rho_{max} \\ 0 & \text{otherwise} \end{cases}.$$

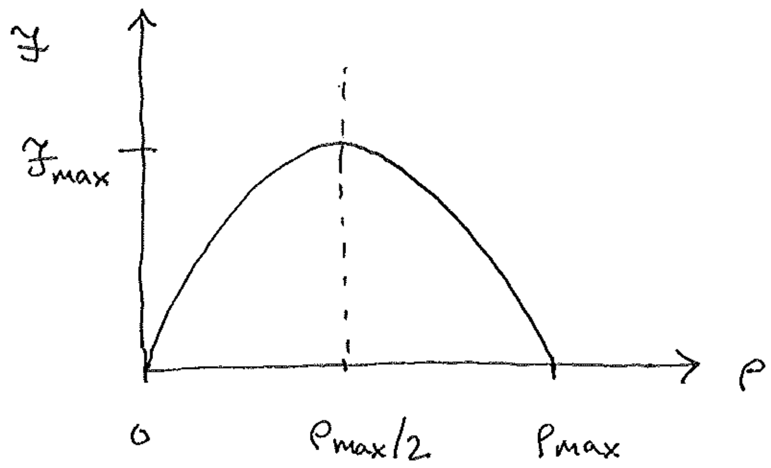
This constitutes our "basic" model:



The flux in this model is

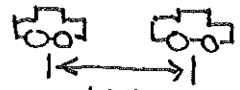
$$\mathcal{F} = \begin{cases} v_{lim}\,\rho\,(1 - \rho/\rho_{max}) & \rho \leq \rho_{max} \\ 0 & \text{otherwise} \end{cases} \quad (35)$$

which is a parabola (for $0 \leq \rho \leq \rho_{max}$) :



This graph is sometimes called the "fundamental diagram". We see that $\mathcal{F}$ has a maximum at $\rho_* = \frac{1}{2}\rho_{max}$, & the maximum value is

$$\mathcal{F}_{max} = \frac{\rho_{max} V_{lim}}{4} . \qquad (36)$$

Observations suggest that the maximum density ρ_{max} is about 100 cars per km, i.e. $\rho_{max} = 0.1\, m^{-1}$ (this is one car every 10m: since the length of a car is about 4m, this corresponds to about a car length between cars  10m). In that case we have, for $V_{lim} = 60\, km\, h^{-1}$

$$\mathcal{F}_{max} = \frac{1}{4} \cdot 0.1\, m^{-1} . \frac{60 \times 10^3}{3600}\, ms^{-1}$$

$$\approx 0.41 s^{-1} = 0.41 \times 3600\, hr^{-1}$$

$$\approx 1500\, hr^{-1}$$

& so we have a first answer to our motivating problem.

What is the corresponding density?

$$\rho_* = \frac{1}{2} \rho_{max} = \frac{1}{2} \cdot 0.1 m^{-1} = 0.05 m^{-1},$$

i.e. the cars have a spacing $\rho_*^{-1} = 20m$.

The speed of the cars is

$$V_* = \frac{\mathcal{F}_{max}}{\rho_*} = \frac{0.41 s^{-1}}{0.05\, m^{-1}} \qquad \left(= \frac{1}{2} V_{lim}\right)$$

$$\approx 8.2 ms^{-1}$$

$$= \overset{30}{\cancel{29.5}}\, km\, h^{-1}.$$

Are these results reasonable? Prigogine & Herman's book shows observational data for traffic in two tunnels in New York City, the Holland

& Lincoln tunnels. Figure 2.1 shows the flux-density relationships for the tunnels.

- Notice first that the shape of the datapoints is qualitatively consistent with the basic model: the flux rises & then decays with increasing density. (The curves in the figure are not the basic model & should be ignored.)

- The maximum flux in the Holland tunnel is a bit larger than 1200 cars per hour, & this is qualitatively similar to our results. However, based on other data the maximum speed (speed limit) in the tunnel is around $80 \, km h^{-1}$ (55 mph), in which case the basic model predicts $\frac{80}{60} \cdot 1500 \, hr^{-1} \approx 2000 \, hr^{-1}$. Hence the basic model overestimates the maximum flux, by comparison with the observations.

- The flux-density points are clearly asymmetrical: the maximum is skewed towards lower densities. Our basic model gives a symmetrical curve, a parabola.

What then is wrong with the basic model? How can it be improved to the point of

Figure 2.1

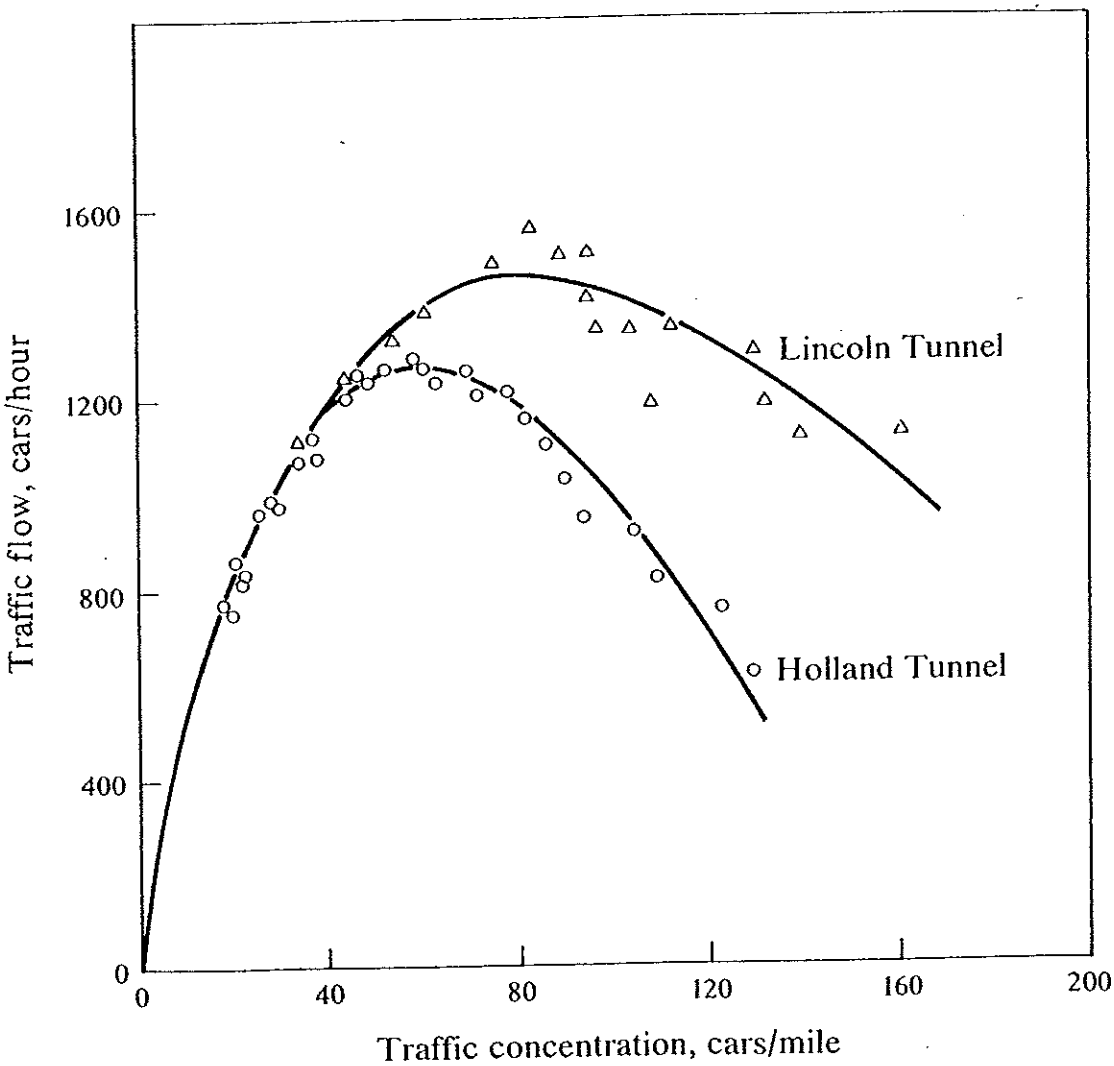


FIG. 1.3. Flow (cars per hour) versus vehicle concentration (cars per mile) for the Lincoln and Holland Tunnels. The Lincoln Tunnel data are those of Greenberg (1959). The solid curves correspond to least squares fits of the reciprocal-spacing car-following model to the data.

From Prigogine & Herman

being able to quantitatively reproduce the observations? Figure 2.2 shows the velocity - density relationship for the Holland tunnel data. We see that the points do not follow a linear decrease, i.e. do not obey (34). This is not surprising: (34) was chosen for simplicity, & has no firm justification. The data tells us that it is wrong, & hence ~~it is~~ this aspect of the model ~~that~~ needs revision. Will this fix the problems of the basic model highlighted by Figure 2.1? From Figure 2.2 we see that velocity decreases initially more rapidly with density than expected from simple linear decrease. This will tend to skew the maximum of $\mathcal{F} = \rho V$ to smaller densities, & will decrease the maximum value of the flux, as required.

How can the behaviour represented by Figure 2.2 be understood? Clearly the figure describes how drivers in the tunnel adapt to increasing density, & must arise from the way in which drivers follow one another. Hence we consider a revised model which attempts to describe this process.

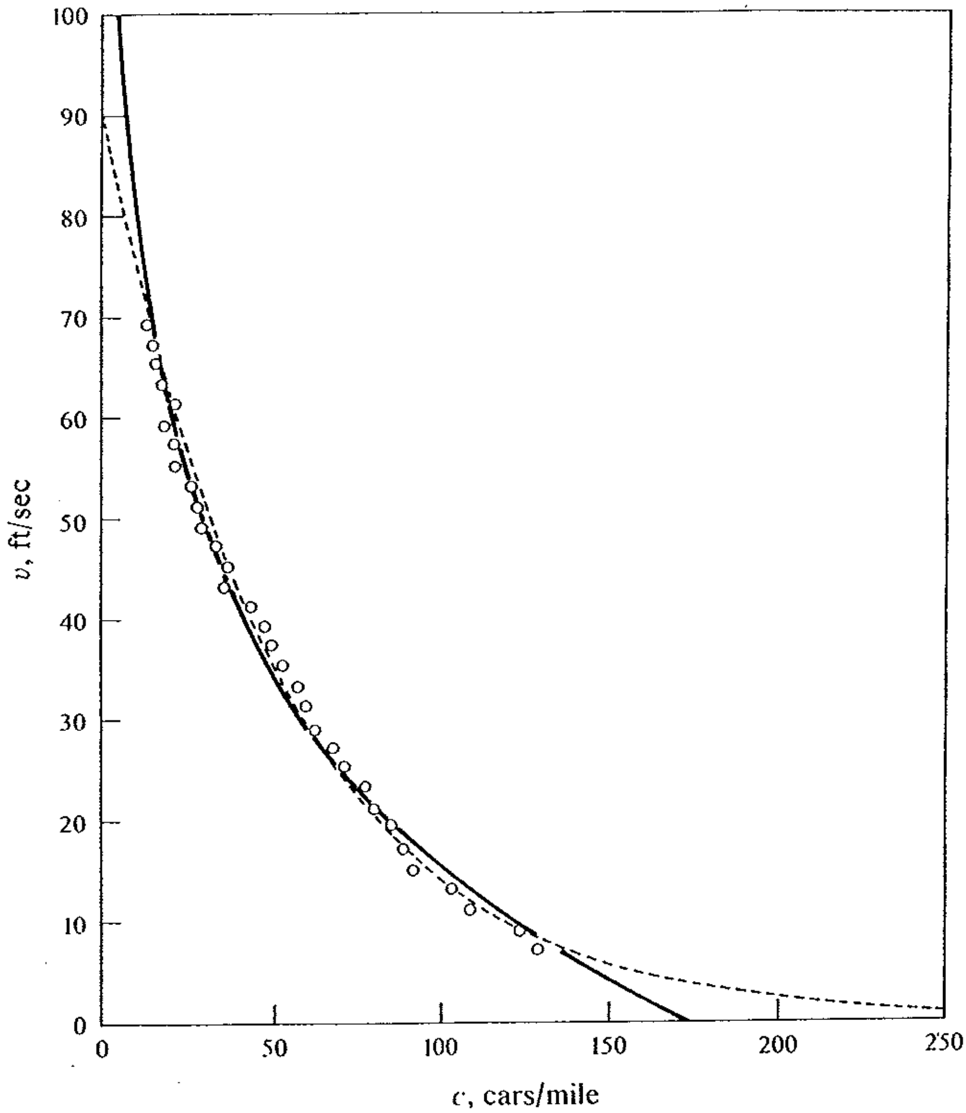


FIG. 1.6. Data on speed (feet per second) versus vehicle concentration (cars per mile) plotted together with two speed–concentration relations derived from car-following models. The data were obtained in the Holland Tunnel. The solid curve is based on the reciprocal-spacing car-following model and the dotted curve is from Edie's model (see Section 1.2).

From Prigogine & Herman

Figure 2.2

2.3 Follow-the-leader equilibrium model :

On a busy road drivers do not crash into one another because they hit the brakes when they see the gap between themselves & the car ahead closing. We label the vehicle positions x_i $(i=1,2,\dots)$ & the gaps ahead $\Delta x_i = x_{i+1} - x_i$ $(i=1,2,\dots)$. These variables are functions of time, i.e. $x_i = x_i(t)$, $\Delta x_i = \Delta x_i(t)$.

We assume :

- drivers brake when Δx_i decreases with time ($\frac{d\Delta x_i}{dt}$ negative),
- braking is more severe when Δx_i is small.

A simple model incorporating these assumptions as well as a response time T for drivers is that

$$\frac{d^2 x_i(t+T)}{dt^2} = \lambda \frac{d(\Delta x_i(t))}{dt} \Big/ \Delta x_i(t) \qquad (37)$$

i.e. the acceleration at time $t+T$ is proportional to the rate of change of the gap ahead & inversely proportional to the gap size. The factor λ is the constant of proportionality : it is a "free parameter" in our model, but has dimensions of velocity, from (37).

Eq. (37) is not the only possible model of acc'n we could choose to embody the basic stated assumptions. However, it is a choice which leads to a simple analytic model, as follows.

We have

$$\frac{d}{dt}\left(\frac{dx_i}{dt}\right)^{(t+T)} = \lambda \frac{\frac{d(\Delta x_i)}{dt}}{\Delta x_i} = \lambda \frac{d}{dt} \ln(\Delta x_i)$$

which is directly integrable:

$$\frac{dx_i}{dt}^{(t+T)} = \lambda \ln(\Delta x_i) + \alpha_i , \qquad (38)$$

where α_i is the constant of integration.

Now consider the equilibrium situation, in which all cars move with a constant speed v with the same distance between them. Then $v = dx_i/dt$, $\rho = \frac{1}{\Delta x_i}$, & $\alpha_i = \alpha$, which does not depend on i:

$$v = \lambda \ln \frac{1}{\rho} + \alpha .$$

We also require $v(\rho_{max}) = 0$

$$0 = \lambda \ln \frac{1}{\rho_{max}} + \alpha$$

$$\alpha = -\lambda \ln \frac{1}{\rho_{max}} = \lambda \ln \rho_{max}$$

so: $$v = \lambda \ln\left(\frac{\rho_{max}}{\rho}\right) \qquad (39)$$

This equation implies that $v \to \infty$ as $\rho \to 0$. Clearly we require $v = v_{lim}$ for $\rho \leq \rho_c$, where

$$v_{lim} = \lambda \ln\left(\frac{\rho_{max}}{\rho_c}\right)$$

$$\text{or} \quad \rho_c = \rho_{max}\, e^{-v_{lim}/\lambda} . \qquad (40)$$

Hence the velocity-density relationship is

$$v = \begin{cases} v_{lim} & \rho \leq \rho_c \\ \lambda \ln\left(\frac{\rho_{max}}{\rho}\right) & \rho_c \leq \rho \leq \rho_{max} \\ 0 & \rho \geq \rho_{max} \end{cases} \qquad (41)$$

Note that although λ is a free parameter we expect that cars will only move at the speed limit for relatively low densities, & (40) then implies that v_{lim}/λ will be relatively large, or we expect λ sig. less than v_{lim}.

The flux corresponding to (41) is

$$\mathcal{Y} = \rho v = \begin{cases} \rho v_{lim} & \rho \leq \rho_c \\ \lambda \rho \ln\left(\frac{\rho_{max}}{\rho}\right) & \rho_c \leq \rho \leq \rho_{max} \\ 0 & \rho \geq \rho_{max} \end{cases} \qquad (42)$$

What does the equilibrium follow-the-leader solution [(41) & (42)] look like? To facilitate comparison with the basic model it is useful to non-dimensionalise the variables in the models. Specifically we introduce

$$u \equiv \frac{v}{v_{lim}} \quad \& \quad \xi \equiv \frac{\rho}{\rho_{max}} \qquad (43)$$

& a non-dimensional flux,

$$F = \frac{\mathcal{Y}}{\rho_{max} v_{lim}} . \qquad (44)$$

Then the <u>basic model</u> becomes

$$u = \begin{cases} 1-\xi & \xi \le 1 \\ 0 & \xi \geq 1 \end{cases} \qquad F = \begin{cases} \xi(1-\xi) & \xi \le 1 \\ 0 & \xi \geq 1 \end{cases} \tag{45}$$

For the <u>follow-the leader model</u> we need another non-dimensional variable, say

$$a = \frac{\lambda}{V_{lim}} \tag{46}$$

& then we have

$$u = \begin{cases} 1 & \xi \le \xi_c \\ a \ln \xi^{-1} & \xi_c \le \xi \le 1 \\ 0 & \xi \geq 1 \end{cases} \qquad F = \begin{cases} \xi & \xi \le \xi_c \\ a\,\xi \ln \xi^{-1} & \xi_c \le \xi \le 1 \\ 0 & \xi \geq 1, \end{cases} \tag{47}$$

where $\xi_c = \dfrac{\rho_c}{\rho_{max}} = e^{-1/a}$.

Figure 2.3 compares the two models, for $a = 1/3$.

<u>Exercise</u>: Show that for the follow-the-leader model the maximum flux is

$$\mathcal{F}_{max} = \frac{\lambda \rho_{max}}{e} = \frac{a}{e} \rho_{max} V_{lim}, \tag{48}$$

where $e \approx 2.718$ is the base of the natural logarithm, & this flux occurs when the density & velocity are

$$\rho_* = \rho_{max}/e \quad \& \quad v_* = \lambda . \tag{49}$$

Note that from this exercise λ has a simple interpretation: it is the speed at maximum flux.

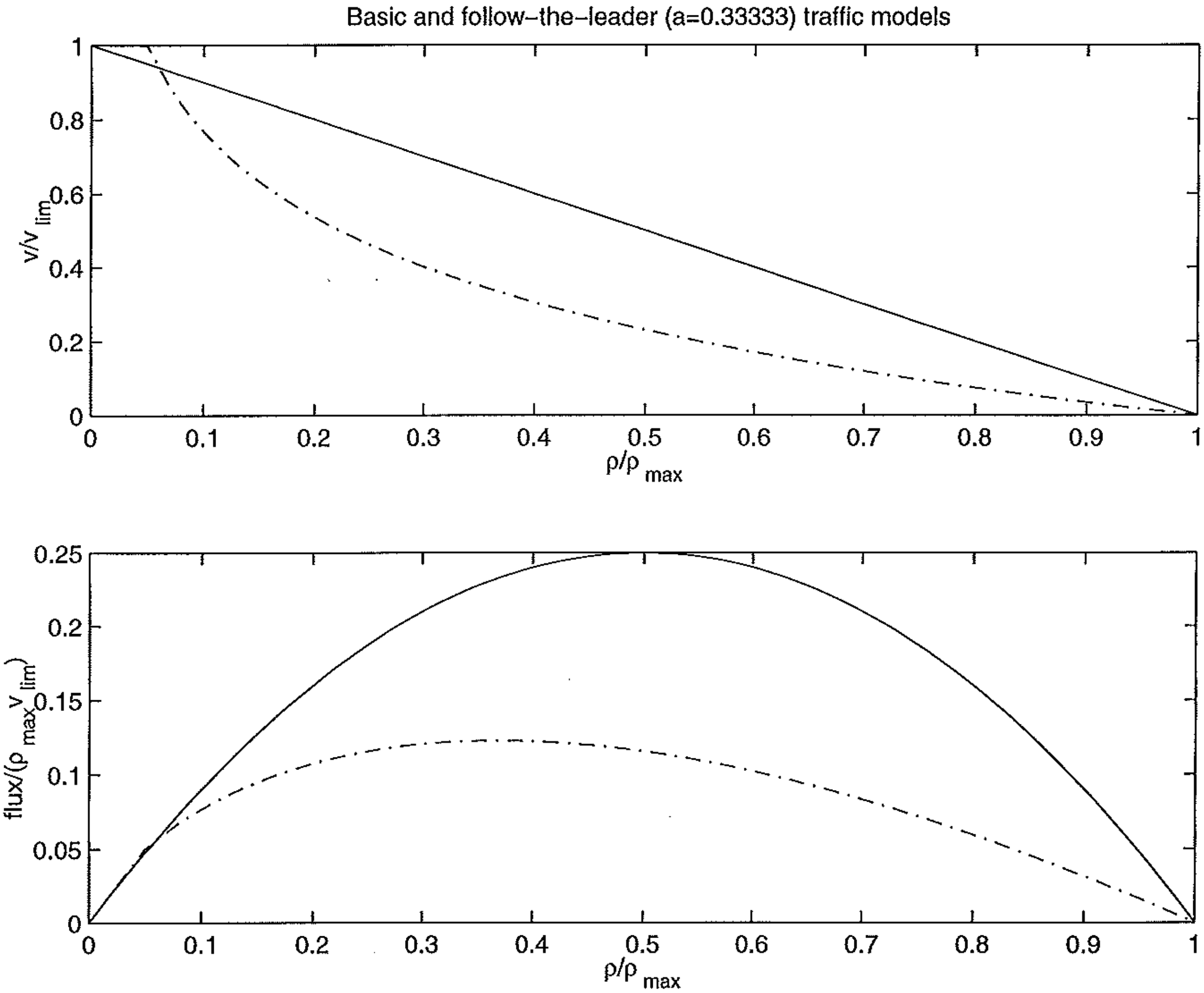

Basic and follow-the-leader (a=0.33333) traffic models
v/v lim
ρ/ρ max
flux/(ρ max v lim)
ρ/ρ max

Figure 2.3

Returning to Figure 2.2, the solid curve is a least squares fit to the follow-the-leader model, which Prigogine & Herman refer to as ~~the~~ Edie's model. ~~An acceptable~~ The best fit is obtained for

$$\rho_c = 10 \text{ cars/mile}, \quad \rho_{max} = 175 \text{ cars/mile},$$

in the units of the figure. Hence

$$a = -\left[\ln\left(\frac{\rho_c}{\rho_{max}}\right)\right]^{-1} \approx 0.35,$$

corresponding to $\lambda = a\,v_{lim} = 0.35 \cdot 55$ mph ≈ 19 mph.

Note that the figure fails to show the correct behaviour for $\rho < \rho_c$ (the solid curve should be flat at $v \approx 80$ ft/sec which corresponds to $v_{lim} = 55$ mph), but this is not important to the fit because there are no observations for densities this low.

The density at maximum flux according to the model is [Eq. (49)]

$$\rho_* = \frac{1}{e}\rho_{max} \approx 64 \text{ cars/mile}$$

& the maximum flux [Eq. (48)] is

$$\mathcal{F}_{max} = \frac{a}{e}\rho_{max} v_{lim} = \frac{0.35 \cdot 175 \text{ cars/mile} \cdot 55 \text{ mph}}{e}$$

$$\approx 1240 \text{ cars/h},$$

which agrees ~~well~~ better with Figure 2.1. We have not discussed the uncertainties in the data. Prigogine & Herman state that this model is consistent with the data, although the data is not sufficient to discriminate between this model & other similar models.

Obviously it would be possible to develop this model further. If we looked hard enough at better data, there would doubtless be discrepancies. Also, the follow-the-leader derivation is reasonable but not compelling. For example, as a car ahead approaches, the driver brakes more severely, which seems OK. However, Eq.(37) also implies that for a receding car ahead, the further that car is away, the less the driver accelerates, which seems curious.

Instead of refining the specific model, we will look more generally at the behaviour of fluid models for traffic.

2.4 Density waves in traffic :

The sign of a good model is that it not only accounts for the data which motivated the model, but also predicts new phenomena, or explains other phenomena not countenanced in the construction of the model. We will see that this is true for our fluid models of traffic.

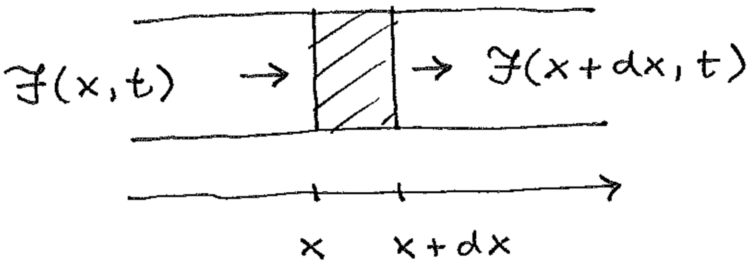
So far we have mostly considered equilibrium situations. Now we consider the more general case. Consider an element of road of length dx. The number of cars in this element at time t is

swap arguments

$$N(t,x) = \rho(t,x)\,dx$$

& the number at time $t+dt$ is

$$N(t+dt,x) = \rho(t+dt,x)\,dx.$$



In time dt, $\mathcal{F}(x,t)\,dt$ cars move into the element & $\mathcal{F}(x+dx,t)\,dt$ move out. Since cars are conserved, we require

$$\begin{array}{c}\text{change in}\\ \text{number in}\\ dx\end{array} = \text{number in} - \text{number out}$$

i.e.

$$N(x,t+dt) - N(x,t) = \mathcal{F}(x,t)\,dt - \mathcal{F}(x+dx,t)\,dt$$

$$\frac{\rho(x,t+dt)-\rho(x,t)}{dt} = -\,\frac{\mathcal{F}(x+dx,t)-\mathcal{F}(x,t)}{dx}$$

or in the limit as $dx, dt \to 0$

$$\frac{\partial \rho}{\partial t} = -\frac{\partial \mathcal{F}}{\partial x},$$

or $$\boxed{\frac{\partial \rho}{\partial t} + \frac{\partial \mathcal{F}}{\partial x} = 0.} \qquad (50)$$

This is the "continuity equation" for traffic. Since $\mathcal{F} = \rho v$ it can also be written

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho v) = 0,$$

which is the usual continuity equation for fluids. The continuity equation is a partial differential equation (PDE). Most of the important equations in physics are PDEs: you have already met the heat equation.

Previously we argued that, in equilibrium, $v = v(\rho)$. Now we will assume that this is also true out of equilibrium, & see where it leads us. If we have $\mathcal{F} = \mathcal{F}(\rho)$, then

$$\frac{\partial \mathcal{F}}{\partial x} = \frac{d\mathcal{F}}{d\rho} \cdot \frac{\partial \rho}{\partial x}$$

& the continuity equation becomes

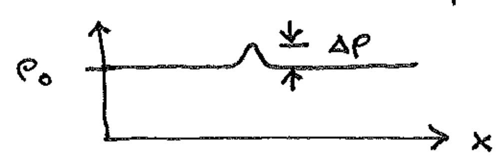
$$\frac{\partial \rho}{\partial t} + w \frac{\partial \rho}{\partial x} = 0 \qquad (51)$$

where

$$w = w(\rho) = \frac{d\mathcal{F}}{d\rho}. \qquad (52)$$

Now consider a uniform density ρ_0 of traffic with a localized enhancement $\Delta\rho$:

$$\rho = \rho_0 + \Delta\rho$$



where $\Delta\rho \ll \rho_0$, & $\frac{\partial \rho_0}{\partial t} = \frac{\partial \rho_0}{\partial x} = 0.$

Then we have $w(\rho) \approx w(\rho_0) \equiv w_0$, &
Equation (51) becomes

$$\frac{\partial}{\partial t}(\Delta\rho) + w_0 \frac{\partial}{\partial x}(\Delta\rho) = 0. \qquad (53)$$

Exercise: Establish (53) more formally.

The solution to (53) is

$$\Delta\rho(x,t) = f(x - w_0 t), \qquad (54)$$

where f is a completely arbitrary function!
To see this, note that differentiating (54):

$$\frac{\partial}{\partial t}(\Delta\rho) = -w_0 f'(x - w_0 t),$$

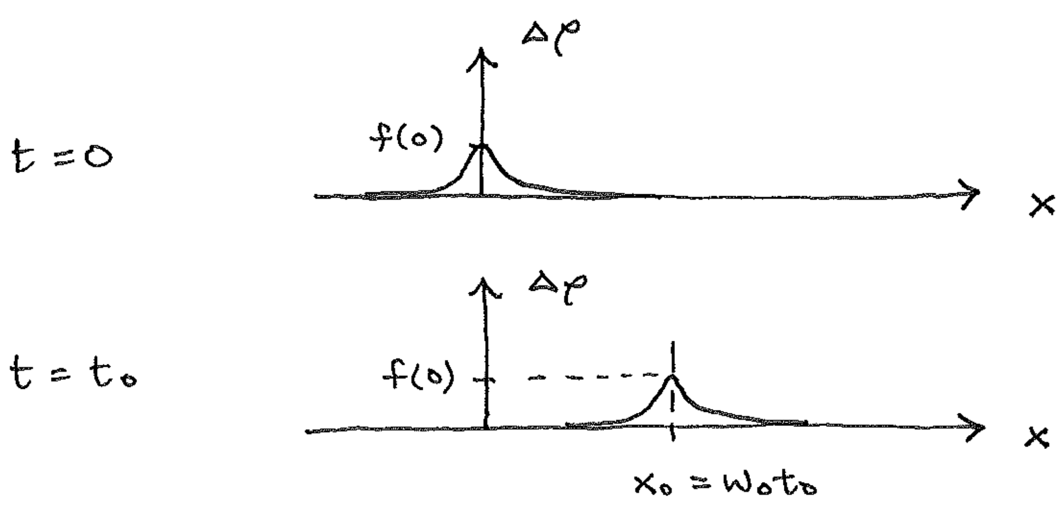
where the prime denotes differentiation w.r.t. the argument. Also

$$\frac{\partial}{\partial x}(\Delta\rho) = f'(x - w_0 t),$$

so

$$\frac{\partial}{\partial t}(\Delta\rho) + w_0 \frac{\partial}{\partial x}(\Delta\rho) = -w_0 f'(\;) + w_0 f'(\;)$$
$$= 0.$$

What does (54) look like?



The point $x_0 = w_0 t_0$ has the same value of $\Delta\rho$ at time t_0 as was at the origin at $t=0$. So the peak moves along with speed w_0. The solution (54) represents density waves propagating with speed w_0. Density bumps in traffic propagate as waves. These waves are called "kinematic waves", because Eq. (53) has not been derived from Newton's laws, but relies on the continuity equation & a local relationship between speed & density.

<u>Comment:</u> The usual wave equation is

$$\frac{\partial^2 y}{\partial t^2} - c^2 \frac{\partial^2 y}{\partial x^2} = 0 \qquad (55)$$

where the constant c is the wave speed. The d'Alembert solution to the wave equation is

$$y = F(x \pm ct). \qquad (56)$$

The relation between (55) & (56) & density waves in traffic is made clear by factorising (55):

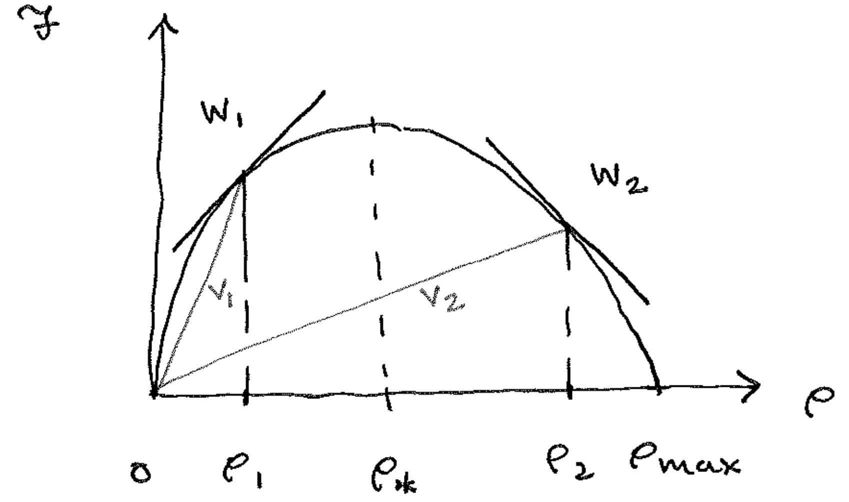
$$\left(\frac{\partial}{\partial t} - c\frac{\partial}{\partial x}\right)\left(\frac{\partial}{\partial t} + c\frac{\partial}{\partial x}\right)y = 0. \qquad (57)$$

The regular wave equation is the product of two PDEs like (53), with differing signs in front of the wavespeed. By comparison with (57), (53) permits wave propagation in only one direction, as we shall see.

The wavespeed is $W(\rho_0) = \left.\dfrac{d\mathcal{F}}{d\rho}\right|_{\rho=\rho_0}$.

To evaluate this for different ρ_0's we need a flux-density relationship:

e.g.



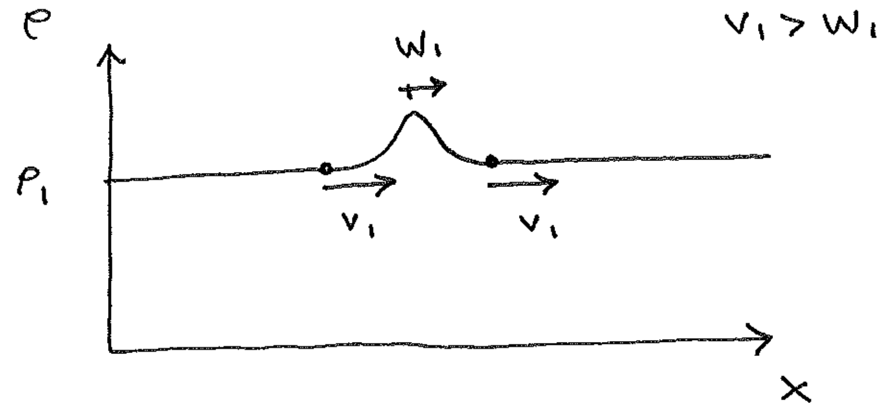
- The kinetic wavespeed is the gradient of the flux-density curve at the relevant density, as illustrated for two density values, ρ_1 & ρ_2.

- The velocity of the vehicles is the slope of the straight line from the origin to the relevant point on the curve, since $v = \mathcal{F}/\rho$. The values v_1 & v_2 are illustrated in the figure.

We can identify two types of behaviour, corresponding to ρ_1 & ρ_2:

1. <u>If $\rho < \rho_*$</u> (e.g. ρ_1): The kinetic wavespeed is positive, so the wave moves in the same direction as the traffic $(+x)$.

- However, $w < v$, so whilst the wave moves downstream with respect to the highway, with respect to the cars it is moving upstream, i.e. backwards

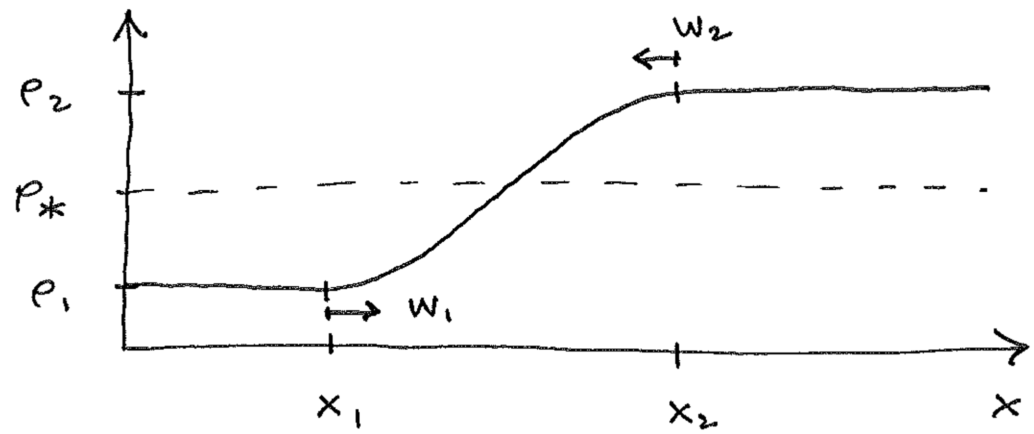


cars at the rear of the bump move into the bump, & cars at the front move out of the bump. Because the density in the bump is larger, the cars must slow down in moving through the bump.

2. <u>If $\rho > \rho*$</u> (e.g. ρ_2) the kinematic wave speed is negative, so the wave moves in the $-x$ direction, & is going backwards both w.r.t. the highway & the cars.

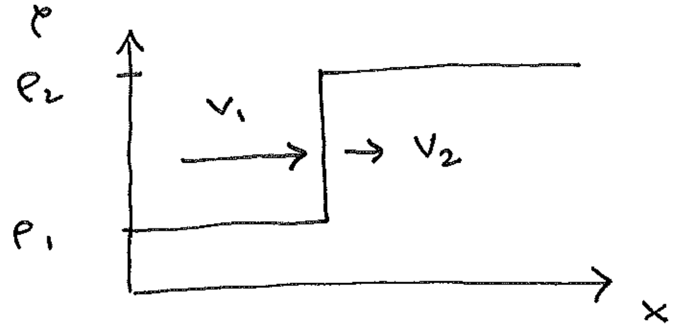
We have all experienced these density waves in traffic. Sometimes the car ahead slows down & you are forced to slow down too, even though there is no obvious cause for the slow down. The density in the vicinity of your car increases. Then you speed up again. Behind you the cars have been forced to slow down in turn. You have experienced a backwards propagating density wave! Hence we see that the fluid model explains a traffic phenomenon we had not anticipated.

So far we have considered how small perturbations in density propagate. What about large density variations? Consider the following situation:



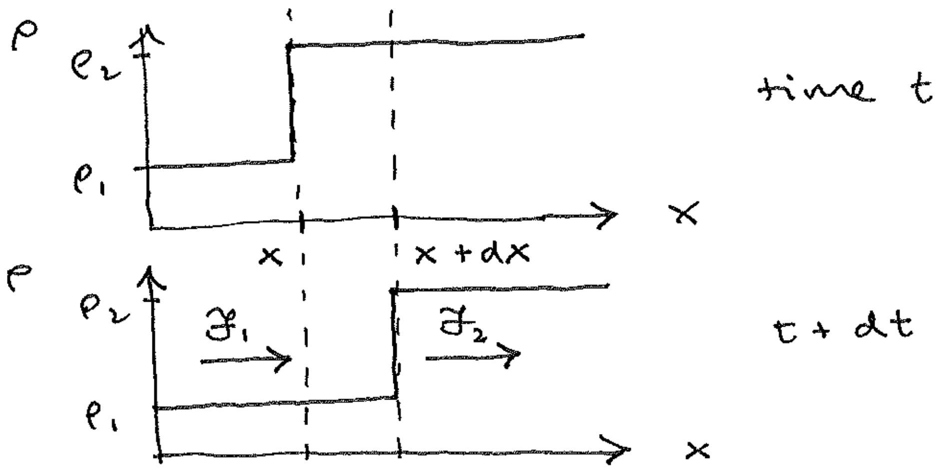
i.e. low density traffic with high density traffic ahead. Consider the points x_1 & x_2, before & after the ramp. Ripples in density at x_2 have velocity $w_2 < 0$ & hence move backwards. Ripples in density at x_1 have velocity $w_2 > 0$ & move forwards. In each case the ripples move density ~~towards~~ into the ramp. The effect is that the ramp steepens with time:

Eventually there is a discontinuity in density: a shock wave. This is analogous to shockwaves in air due to an explosion, or due to the passage of a supersonic aircraft, etc.* Ahead of the "shock front" the cars are still moving (provided $\rho_2 < \rho_{max}$), so the cars move through the shock front, experiencing a sudden slow down.



If $\rho_2 \geqslant \rho_{max}$ we have $v_2 = 0$ & this is a <u>traffic jam</u>. No doubt you have also experienced sudden traffic slowdowns & jams.

The shockfront has a range of densities. What speed does it propagate with? In time dt, assume the front moves a distance dx as shown



* The steepening of the wave is analogous to the steepening of water waves before they break at the shore.

consider the number of cars in the region x to $x+dx$:

at time t, $\rho_2\,dx$

at time $t+dt$, $\rho_1\,dx$

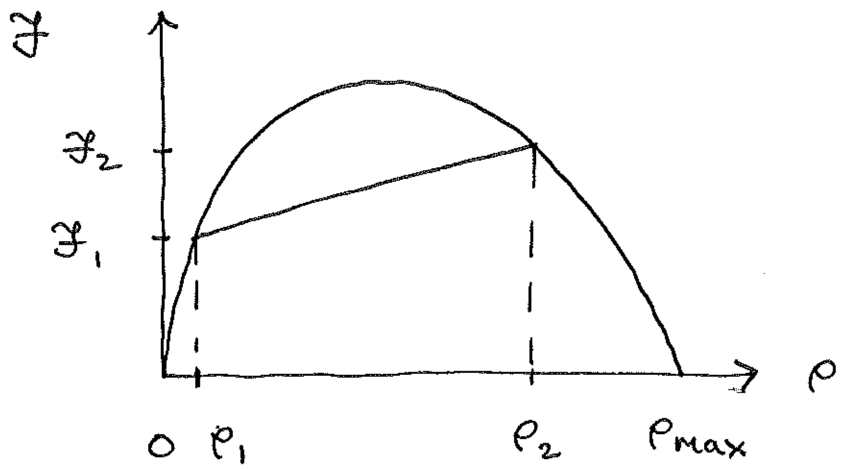
In time dt $\mathcal{F}_1\,dt$ cars crossed x & $\mathcal{F}_2\,dt$ crossed $x+dx$. We require:

$$\begin{array}{c} \text{change in number} \\ \text{in } (x, x+dx) \end{array} = \begin{array}{c} \text{number} \\ \text{in} \end{array} - \begin{array}{c} \text{number} \\ \text{out} \end{array}$$

i.e. $$\rho_1\,dx - \rho_2\,dx = \mathcal{F}_1\,dt - \mathcal{F}_2\,dt$$

or $$\boxed{\frac{dx}{dt} = \frac{\mathcal{F}_2 - \mathcal{F}_1}{\rho_2 - \rho_1},} \qquad (58)$$

which is the speed of propagation of the ~~wa~~ shock front.



From the diagram we see that this is the slope of the line joining the relevant points on the flux-density curve.

More generally, whenever kinematic waves coalesce or collide, shock waves are formed.

There are many more details of this model that can be explored: this is just a brief introduction. However, we have seen that the fluid model for traffic provides a qualitative & quantitative description of traffic flow, & also provides considerable insight into the perplexing problem of traffic jams.

Next we consider models for growth, in particular population models, which are a mainstay of courses such as this one.

PHYS220 Scientific Modelling 2001
Assignment 2 – due Monday May 21

1. For a certain 'fractal' curve, it is found that the measured length l of the curve is proportional to $\delta^{-\alpha}$, where δ is the length of the ruler used in measuring the length, and $\alpha \approx 0.3$. Given that l has dimensions L and $\delta^{-\alpha}$ has dimensions $L^{-\alpha}$, how can this be?

2. Consider a sphere moving through air.

   (a) For small spheres at low speeds, viscous forces dominate over inertial forces. In this case the drag force F_D on the sphere can be assumed to depend on the viscosity μ, the radius r of the sphere and the velocity v of the sphere, and *does not* depend on the air density. Using dimensional analysis, derive an expression for F_D.

   (b) For larger spheres at higher speeds, the viscosity is not important, and the force F_D depends on the air density ρ, the radius r of the sphere and the velocity v of the sphere. Using dimensional analysis, derive an expression for F_D.

   The expression derived in (a) is appropriate to describe the force which keeps the water droplets in clouds suspended in the air, and the expression derived in (b) describes air resistance on macroscopic falling objects, e.g. sky divers.

3. Give an order of magnitude estimate of the total mass of air in the Earth's atmosphere.

4. The NSW RTA recommends that drivers maintain a time interval of at least $\Delta t = 3\,\mathrm{s}$ between themselves and the car ahead. We take this to mean that there is at least three seconds between when the rear of the car ahead of a driver passes a point on the road and when the front of the driver's own car passes the same point.

   Consider an equilibrium situation involving 'RTA drivers' (i.e. drivers who maintain a gap of exactly three seconds whenever possible) all moving with the same speed.

   (a) Derive a velocity-density relationship for these drivers, taking into account the existence of a speed limit on the road. Your expression for the velocity v should depend only on the speed limit v_{lim}, the density ρ, the interval Δt and the length l of the cars.

   (b) At what density is the velocity zero? Is this reasonable?

   (c) Write down the flux-density relationship for the RTA drivers.

   (d) Sketch the velocity v and flux $\mathcal{F}$ as functions of ρ.

   (e) For $v_{\mathrm{lim}} = 60\,\mathrm{km\,h^{-1}}$ and $l = 4\,\mathrm{m}$, how does this model compare with the 'follow-the-leader' model presented in the lectures? Specifically, is the maximum flux better or worse?

3. MODELLING GROWTH

Problem: How do populations grow with time?

Population modelling is a common topic in modelling courses. Here I will attempt a broad discussion, including single populations, discrete modelling, coupled populations, & the evolution of age structure within a population. The application of the models to other growth & spreading processes will also be discussed.

3.1 Single populations:

3.1.1 Exponential growth:

In 1798 Thomas Malthus published a series of essays in which he argued that populations grow naturally according to the law of geometric progression, i.e. they double in a fixed time τ_2. If a population starts at one million & becomes two million after 10 years then in another 10 years it will be four million. We can write

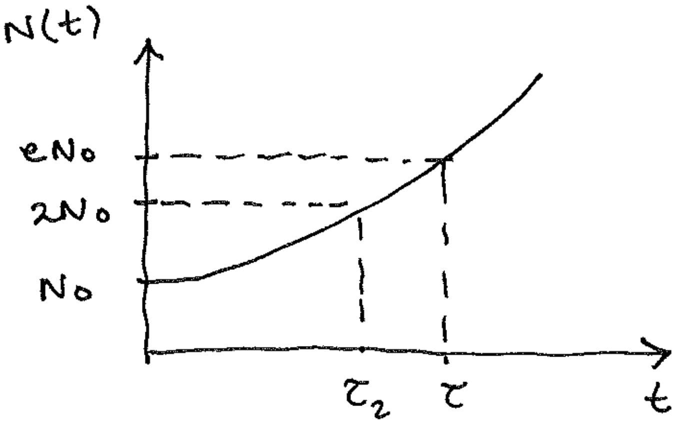
$$N(t) = N_0 \, 2^{t/\tau_2} \qquad (59)$$

where $N(t)$ is the population at time t & $N_0 \equiv N(0)$.

Using $a^b = e^{b \ln a}$ we can write

$$N(t) = N_0 e^{\frac{t}{\tau_2}\ln 2} = N_0 e^{t/\tau}, \qquad (60)$$

where $\tau \equiv \tau_2/\ln 2$ is the time for the population to increase by a factor of e. Hence Malthusian growth is exponential growth.



Differentiating (60) wrt time gives

$$\frac{dN}{dt} = \frac{N_0}{\tau} e^{t/\tau} = \frac{1}{\tau} N$$

or $$\boxed{\frac{dN}{dt} = \mu N} \qquad (61)$$

where $\mu \equiv 1/\tau$. This is the ordinary differential equation (ODE) describing exponential growth.

What is the basis for Malthusian growth? Consider a population $N(t)$. In time Δt assume the population increases by ΔN, which must be equal to the number of births minus the number of deaths:

$$\Delta N = B(t, \Delta t) - D(t, \Delta t).$$

It is reasonable to assume that the number of births is proportional to Δt (for small

Δt) & is also proportional to the population:

$$B(t, \Delta t) = \alpha . N . \Delta t .$$

Similarly it is reasonable to assume the number of deaths is proportional to Δt & N:

$$D(t, \Delta t) = \beta . N . \Delta t .$$

Combining these relations we have

$$\Delta N = (\alpha - \beta) N \Delta t$$

& in the limit $\Delta t \to 0$

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

Provided α & β are constant with time we have (61), with

$$\boxed{\mu \equiv \alpha - \beta .} \qquad (62)$$

This quantity is the net growth rate of the population, per individual (often called the "specific growth rate"). The quantities α & β are the birth & death rates per individual.†

This derivation makes Malthusian growth plausible, but it is by no means inescapable. For example, consider an animal population with no females of childbearing age: in this case B is independent of N. Also, the neglect of time dependence in α & β is questionable. In fact, Malthusian growth

† Note that if the death rate exceeds the birth rate we have exp. decay

is always incorrect, eventually.

To see why, consider the life cycle of a prolific insect, the cockroach. Female cockroaches produce 320 eggs in their one-year lifetime. A naive application of the ideas developed so far would suggest

$$\alpha = 160 \text{ year}^{-1} \;\&\; \beta = 1 \text{ year}^{-1}$$

provided there are equal numbers of male & female cockroaches. Hence $\mu = \alpha - \beta = 159$ year^{-1}, & so in 1 year, starting with two cockroaches you would have $2 \cdot e^{159} \approx 10^{69}$ cockroaches! This is ridiculous. Suppose a cockroach weighs a fraction of a gram (0·1g, say): then all of these cockroaches would weigh 10^{65} kg. The mass of the known universe is $\sim 10^{53}$ kg.

There are many reasons why this calculation is incorrect (e.g. female cockroaches produces eggs in discrete bunches, ~~also~~ there is a maturation period of a few months, not all cockroaches survive...). However, it highlights the fact that Malthusian growth produces ridiculous numbers given sufficient time.

For short times Malthusian growth provides a reasonable description of many populations. For longer times it is necessary to consider restraints or limits to growth, imposed for example by the environment, or by competition among members of the population.

Before we proceed note also that Eq. (61) describes the growth of quantities besides populations. For example:

- growth in a country's economic debt
- increase in the number of people infected with a disease
- growth of plants
- spread of a joke

may all be described by Eq. (61), at least for short times.

3.1.2 Hyperbolic growth:

Although exponential growth is fast, it is by no means the fastest imaginable type of growth. For example, consider

$$\boxed{\frac{dN}{dt} = \mu N^2.} \qquad (63)$$

This equation is directly integrable:

$$\int \frac{dN}{N^2} = \mu \int dt$$

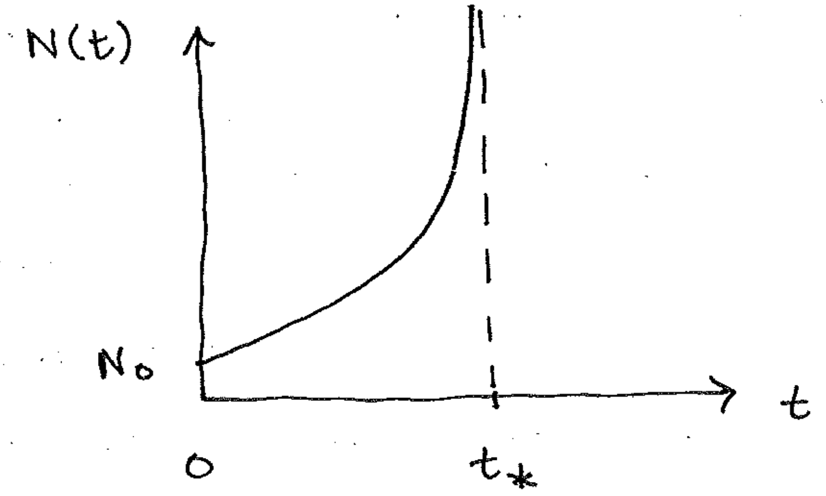
$$\Rightarrow \quad -\frac{1}{N} = \mu t + C$$

Imposing $N = N_0$ at $t = 0$ $\Rightarrow$ $C = -\frac{1}{N_0}$

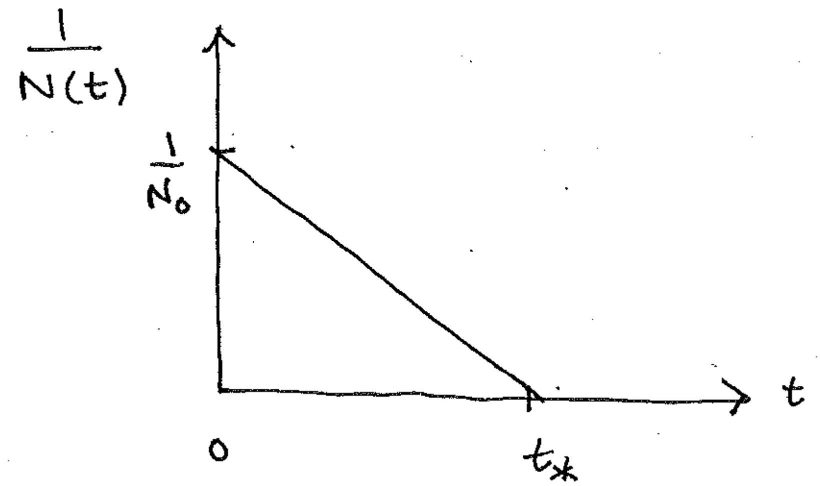
$$\frac{1}{N} = \frac{1}{N_0} - \mu t$$

$$\boxed{N = \frac{N_0}{1 - \mu N_0 t}} \qquad (64)$$

Note that $N \to \infty$ when $t \to t_* = \frac{1}{\mu N_0}$.
The population becomes infinite in a finite time (cf. exponential growth).



This is "hyperbolic" growth. If we plot $\frac{1}{N}$ versus t we get a straight line which intercepts the t axis at t_*



Interestingly there is a population which has exhibited hyperbolic growth for some time: the human world population. Contrary to popular belief the world population has not grown like an exponential but has grown much more quickly. This aspect of world population growth was first pointed out by von Foerster et al. in 1960 (Science <u>132</u>, 1291, 1960)

Figure 3.1 plots world population estimates for 0 AD – present (the numbers are from Kramer 1993, Quarterly J. Economics <u>108</u> 681–716)

- Top panels show that the growth is not exponential, except possibly for recent decades

- Bottom panels illustrate a remarkable adherance to hyperbolic growth, in particular since around 1600. The dashed line in the bottom right panel is the phenomenological law

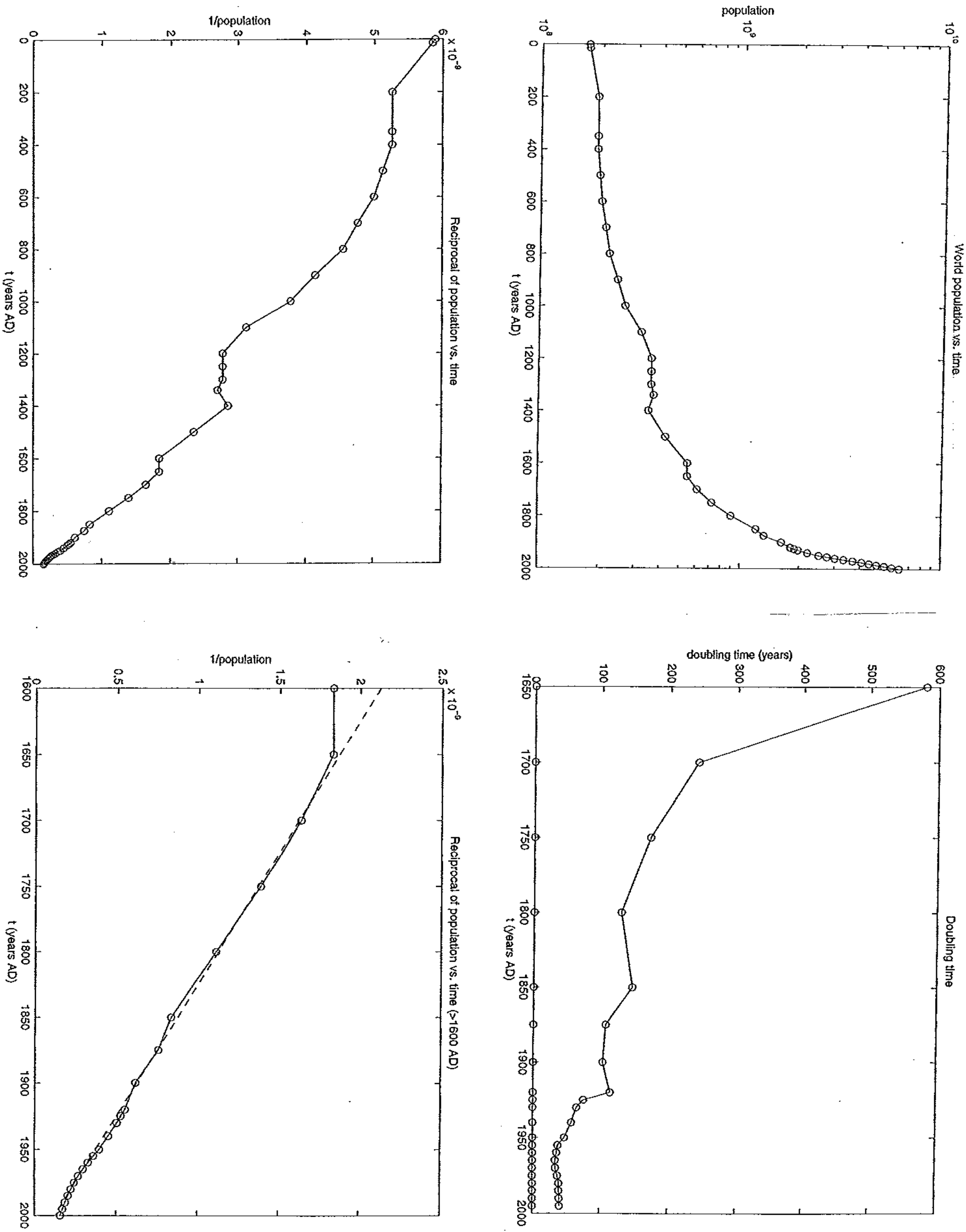
t is in years. $\rightsquigarrow$

$$N = \frac{200 \times 10^9}{2025 - t} \qquad (65)$$

(i.e. $\frac{1}{\mu} = 200 \times 10^9$, $t_* = 2025$ in Eq. (64))
suggested by von Hoerner (J. British Interplanetary Society <u>28</u> 691, 1975). In fact the world population estimates are somewhat uncertain; ~~some~~ other estimates suggest an adherance to this law since 0 AD.

- Note that $t_* = 2025$: according to this law, the world population will be infinite in 2025 AD. Of course this is ridiculous. Extrapolating the law into the distant past also gives an absurd result, viz. that there were 10 people around 20×10^9 years ago, i.e.

Figure 3.1


World population vs. time
population
t (years AD)
Doubling time
doubling time (years)
t (years AD)
Reciprocal of population vs. time
1/population
t (years AD)
Reciprocal of population vs. time (>1600 AD)
1/population
t (years AD)

around the time of the big bang. Regarding ~~the~~ "doomsday 2025", it is already possible to see a departure from the hyperbolic trend.

Why does the human population grow in this way? As far as I can tell nobody knows. The human population is the only biological population exhibiting hyperbolic growth. Presumably the organisation & co-operation within human society, together with economic, social & technological changes have lead to an increase in the specific birth rate with population,

$$\alpha = k . N$$

where k is roughly constant. The human capacity to multiply (which alone produces exponential growth) is only part of the mechanism.

An interesting question in this area is: how many people have ever lived? If the world population is $N(t)$ then the answer is approximately

$$N_{tot} = \frac{1}{\tau} \int N(t)\, dt \qquad (66)$$

where the integral runs over the history of our species & where τ is the average human lifespan for that time.

If we adopt the model (65) & assume it has held for 0 – 2000 AD & adopt $\tau \approx 45$ years then we obtain a total of 20×10^9 people in that time.

Exercise : Confirm this estimate.

However, although early human populations were low, the timescales are large, & it turns out that far more humans lived prior to 0 AD than after 0 AD. Typical estimates for N_{tot} are in the range 80 to 100 billion.

3.1.3 Logistic growth :

For many populations if there are too many individuals there is competition for finite resources, which is sometimes called "crowding." We can model this crudely by assuming that the death rate for individuals increases with population. A simple choice is

$$\beta = \beta_0 + \beta_1 N . \qquad (67)$$

In this case we have

$$\frac{dN}{dt} = (\alpha - \beta) N = (\alpha - \beta_0 - \beta_1 N) N$$
$$= (a - \beta_1 N) N ,$$

where $a \equiv \alpha - \beta_0$, or

$$\frac{dN}{dt} = aN\left(1 - \frac{N}{a/\beta_1}\right).$$

So $\frac{dN}{dt} = 0$ for $N = N_* = a/\beta_1$, which is called the "carrying capacity." The usual form of the "logistic" equation is

$$\boxed{\frac{dN}{dt} = aN\left(1 - \frac{N}{N_*}\right).} \qquad (68)$$

There are other ways to justify (68)†, but it should be seen as a simple correction to Malthusian growth. The logistic equation was first suggested by Pierre Francois Verhulst in 1838, based on studies of populations in Europe.

For a & N_* constant Eq. (68) is easy to solve as follows:

$$\int \frac{dN}{N(1-N/N_*)} = \int a\,dt$$

& $$\frac{1}{N\left(1-\frac{N}{N_*}\right)} = \frac{A}{N} + \frac{B}{1-N/N_*}$$

$$\Rightarrow \quad 1 = A\left(1 - \frac{N}{N_*}\right) + BN$$

Setting $N = 0 \Rightarrow A = 1$
" $N = N_* \Rightarrow B = 1/N_*$

So we have

$$\int\left(\frac{1}{N} + \frac{1}{N_* - N}\right)dN = at + C$$

† e.g. it may arise from a depression in birth rates instead

$$\ln N - \ln(N_* - N) = at + c$$

$$\frac{N}{N_* - N} = Ae^{at} \qquad (A = e^c)$$

$$N(0) = N_0 \Rightarrow \frac{N_0}{N_* - N_0} = A$$

so
$$N = \frac{N_* - N}{N_* - N_0} N_0 e^{at}$$

$$N\left(1 + \frac{N_0}{N_* - N_0} e^{at}\right) = \frac{N_* N_0}{N_* - N_0} e^{at}$$

$$N = \frac{N_* N_0 e^{at}}{N_* - N_0 + N_0 e^{at}}$$

$$= \frac{N_*}{1 + \frac{N_* - N_0}{N_0} e^{-at}}$$

or
$$\boxed{N = \frac{N_*}{1 + (N_*/N_0 - 1)e^{-at}}} \qquad (69)$$

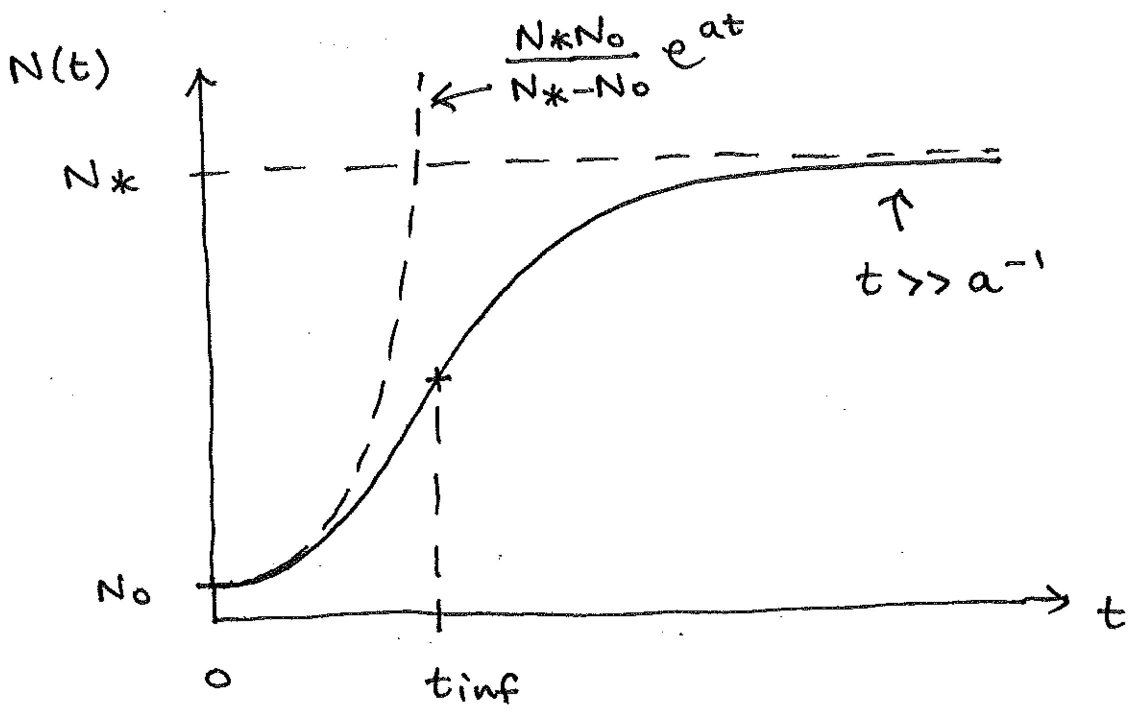
- Note that $N \to N_*$ as $t \to \infty$

- For small t $e^{-at} \approx 1$. Assuming the final population is large $\frac{N_*}{N_0} - 1 \gg 1$, * the 2nd term on the denominator is much larger than unity, so we have

$$N \approx \frac{N_* N_0}{N_* - N_0} e^{at} \qquad (at \ll 1,\ N_* \gg N_0)$$

i.e. the growth is initially exponential.

so it looks like the following:



a characteristic S shaped curve.

Exercise: Show that, provided $N_* > N_0$, the logistic curve has an inflection point at

$$t_{inf} = \frac{1}{a} \ln\left(\frac{N_*}{N_0} - 1\right)$$

Growth consistent with the logistic equation has been observed in certain laboratory controlled growth experiments. Examples include bacteria, yeast, & the height of plants. Figure 3.2 shows two examples. Outside of the laboratory logistic growth is likely to be a cruder approximation, for example because of variations in the environment.

Exercise: Sketch N(t) for $N_0 > N_*$

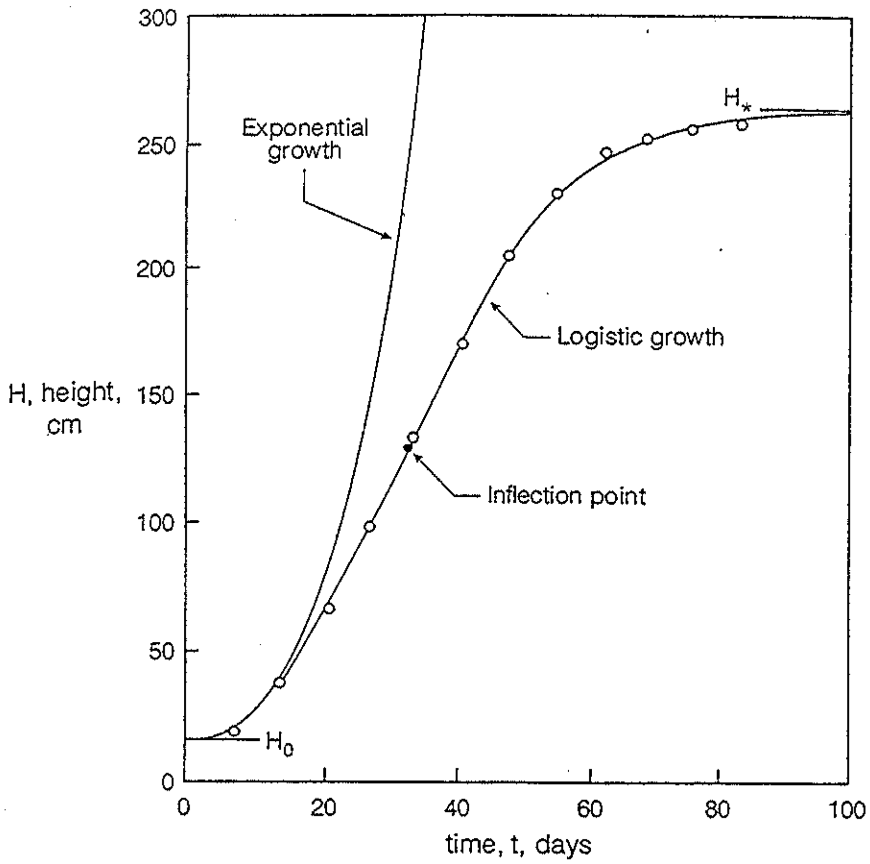


FIG. 3.1

Growth of sunflower plants. Data of Reed and Holland. (From Lotka 1956.)

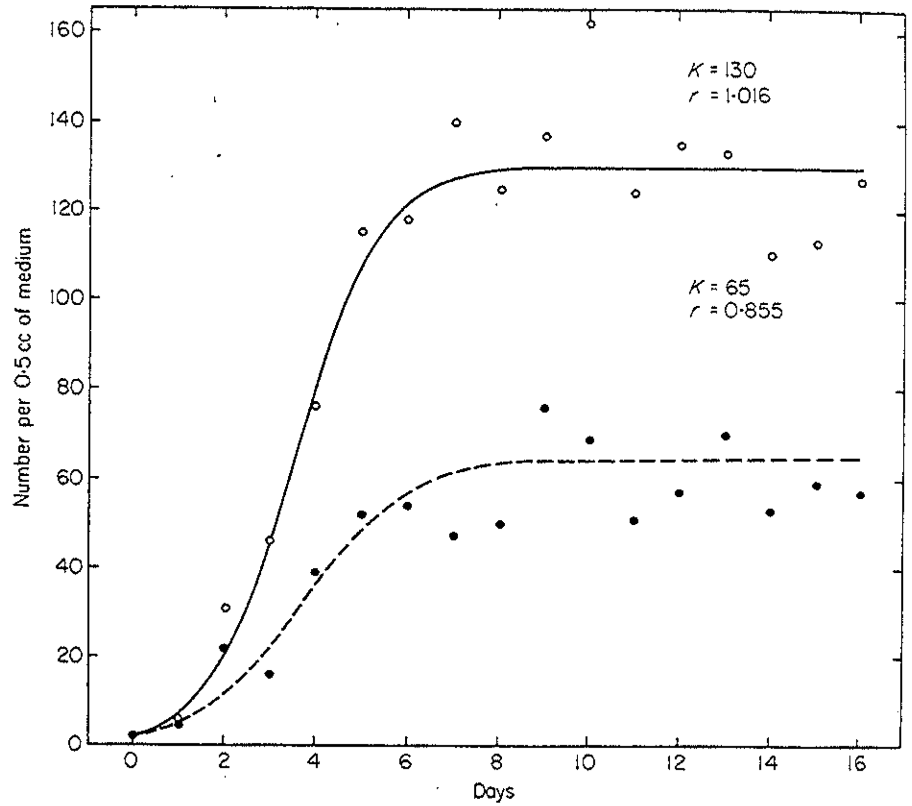


Fig. 6.1. Growth of *Paramecium* populations grazing a vegetation renewed at a constant rate.

From Banks, p. 27

From May, p. 98

$r = a, \quad N_* = K$

Figure 3.2

Comment:

· You do not get chaos (limit cycles, bifurcation, etc.) out of (68): you get these things out of the ~~log~~ discrete version, the logistic map. You might say, 'but I discretise in numerically solving ...' - yes, but you need a small time step Δt to solve the equation accurately, & this ensures you are not in the parameter range needed for chaos.

To see this:

$$\frac{dN}{dt} = a\left(1-\frac{N}{N_*}\right)N$$

Discretise:

$$\frac{N_{i+1}-N_i}{\Delta t} = a\left(1-\frac{N_i}{N_*}\right)N_i$$

$$N_{i+1} = N_i + a\Delta t\left(1-\frac{N_i}{N_*}\right)N_i$$

$$= (1+a\Delta t)N_i\left[1-\frac{N_i}{N_*\left(\frac{1}{a\Delta t}+1\right)}\right]$$

set $x_i = \dfrac{N_i}{N_*\left(\frac{1}{a\Delta t}+1\right)}$, $r = 1+a\Delta t$

get: $\quad x_{i+1} = r x_i(1-x_i)$

the logistic map. But: $r=(1+a\Delta t)\approx 1$ for an accurate solution. You need $r > 3.57\ldots$ for chaos. There is no chaos in (68)!

3.1.4 Population crashing :

There are other ways in which growth can be limited. For example, early attempts to demonstrate logistic growth in bacterial populations in the laboratory often showed a rise followed by a decline to extinction, rather than the expected rise to the carrying capacity. This phenomenon was called "crashing." In the 1930's Volterra & Kostitzin explained this behaviour as being due to the contamination of the growth environment by lethal products produced by the bacteria themselves (self-poisoning). They proposed the form for the death rate per bacterium:

$$\beta = \beta_0 + \beta_1 N + \beta_2 \int_0^t N \, dt \qquad (70)$$

where the experiment begins at $t=0$. Equation (70) involves the logistic death rate plus a term representing an increase in β proportional to all the bacteria that have lived. This is explicable if each bacterium produces a certain amount of toxin, which remains, & the death rate is then proportional to the accumulated amount of the toxin.

The growth equation becomes

$$\frac{1}{N}\frac{dN}{dt} = \alpha - \beta = \alpha - \beta_0 - \beta_1 N - \beta_2 \int_0^t N\,dt$$

$$= a - \beta_1 N - \beta_2 \int_0^t N\,dt \quad (71)$$

where $a = \alpha - \beta_0$.

Equation (71) is an example of an "integro-differential equation". In general these are hard to solve, but (71) is not too difficult. For simplicity (& to make the role of the integral term clearer) we will ignore the crowding term, i.e. we will set $\beta_1 = 0$. Also we will write $\beta_2 = b$, so we wish to solve

$$\frac{1}{N}\frac{dN}{dt} = a - b\int_0^t N\,dt . \quad (72)$$

Writing

$$P = \int_0^t N\,dt \quad (73)$$

clearly

$$\frac{dP}{dt} = N . \quad (74)$$

Also notice that we can write

$$\frac{dN}{dt} = \frac{dN}{dP}\frac{dP}{dt} = \frac{dN}{dP}\cdot N , \quad (75)$$

in light of (74). Using (73) & (75) Eq. (72) becomes

$$\frac{dN}{dP} = a - bP$$

which is integrable:

$$N = aP - \tfrac{1}{2}bP^2 + \text{const.}$$

when $t=0$ $P=0$ & $N=N_0$. Hence

$$N = N_0 + aP - \tfrac{1}{2}bP^2. \qquad (76)$$

Even without proceeding any further, we can guess the outcome from (76). The function $P(t)$ is strictly increasing. Eventually the term $\frac{1}{2}bP^2$ will be larger than $aP+N_0$, & so the population N will begin to decline with time: it will crash. (The same conclusions can be reached directly from (71) or (72).)

To continue, from (74) & (76) we have

$$\frac{dP}{dt} = N_0 + aP - \tfrac{1}{2}bP^2. \qquad (77)$$

Exercise: Show that the solution to (77) is

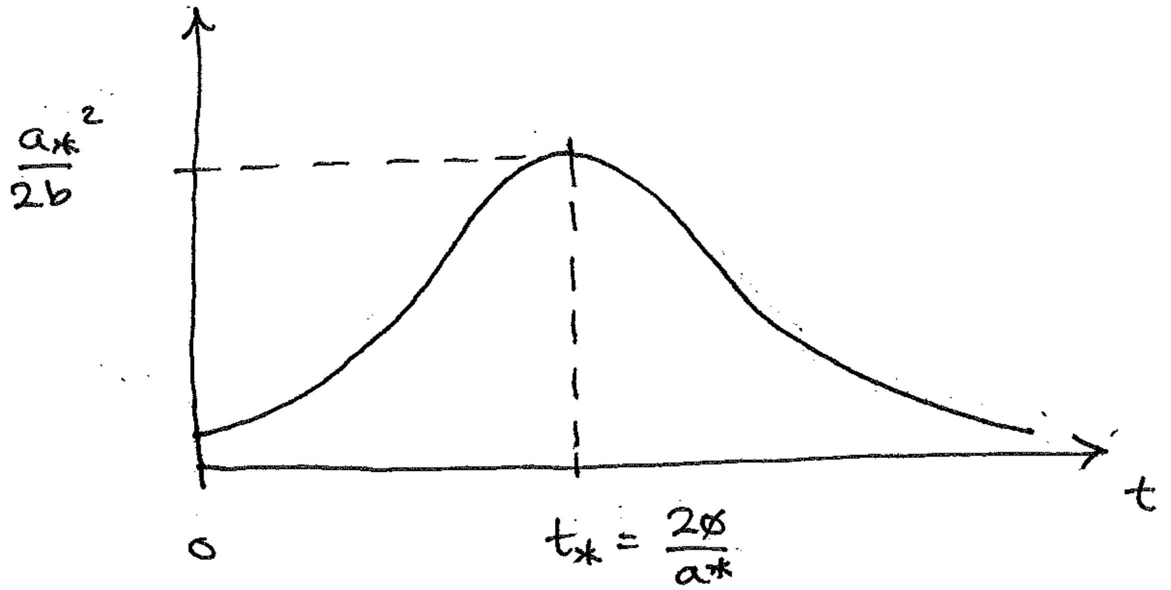
$$P = \frac{a}{b}\left[1 + \frac{a_*}{a}\tanh\left(\tfrac{1}{2}a_*t - \phi\right)\right] \qquad (78)$$

where $\phi = \tanh^{-1}\dfrac{a}{a_*}$, $a_* \equiv (a^2 + 2bN_0)^{\frac{1}{2}}$.

Differentiating (78) gives N:

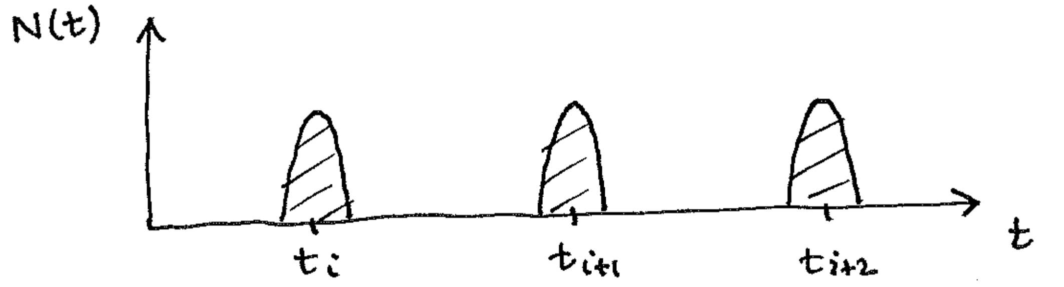
$$N = \frac{dP}{dt} = \frac{a_*^2}{2b}\operatorname{sech}^2\left(\tfrac{1}{2}a_*t - \phi\right). \qquad (79)$$

This form of the solution makes the behaviour of the population obvious: the sech^2 function is a symmetrical peak, & the argument in (79) means that it is 'shifted to the right' by $t_* = 2\phi/a_*$:



3.1.5 Discrete modelling:

Many populations are essentially made up of a single generation, with no overlap between successive generations:



Examples include annual flowers, certain insects (e.g. periodical cicadas). Obviously in these cases members of a generation must disperse seeds, lay eggs, etc. to contribute to the next generation.

In these cases the appropriate model is a difference equation relating the population N_{i+1} of generation $i+1$ to the population N_i of generation i:

$$\boxed{N_{i+1} = F(N_i)} \qquad (80)$$

The simplest example is the linear equation

$$N_{i+1} = \lambda N_i \qquad (81)$$

which Barry introduced. This is the discrete analogue of (61) & also produces exponential growth.

Exercise: Show that the solution to (81) is $N_i = N_0 e^{ri}$, where $r = \ln\lambda$.

The discrete analogue of logistic growth is

$$N_{i+1} = N_i\left[1 + \lambda\left(1 - \frac{N_i}{N_*}\right)\right] \qquad (82)$$

which is ~~the~~ equivalent to the "logistic map" discussed by Barry. This nonlinear difference equation exhibits stable points, stable cycles & chaos, as discussed by Barry. We will not go into these details, but a few comments are in order.

- The continuous version [Eq. (68)] is nonlinear but does not exhibit chaos: with continuous equations (ODEs) a second order equation [involving d^2N/dt^2] is needed for chaos (among other things).

- Although Barry discussed the application of (82) to bacteria, in that case the generations overlap & so (82) is invalid: Eq. (68) is needed. Hence there will be no limit cycles, chaos, etc. – sorry.
- Eq. (82) is of limited use in biology because the population can become negative. This equation is preferred by mathematicians because it is the simplest nonlinear difference equation.

To conclude this discussion of single populations, we note that there are many other possibilities that we have neglected. You may wish to consider the following modifications of the models presented here:

- Time-dependent birth & death rates, $\alpha = \alpha(t)$, $\beta = \beta(t)$.
- A time-delay to represent a response time in the environment or within the species. For example, if an animal

eats vegetation which takes a time T to grow back, then we can write

$$\frac{dN}{dt} = aN\left[1 - \frac{N(t-T)}{N*}\right]$$

as a crude model. The "crowding" effect depends on how much eating was done at an earlier time.

3.2 Coupled populations :

Next we consider simple ecosystems consisting of two populations which interact. The procedure is to consider simple models & then to refine them, following the process outlined in §1.

3.2.1 Predator-prey models :

Consider a population $N(t)$ of herbivores & a population $P(t)$† of predators who eat the herbivores.

- $\mu_1 = \frac{1}{N}\frac{dN}{dt}$ is the net rate of increase, per individual, of the herbivores. If there are no predators we expect growth $(\mu_1 > 0)$: if there are a lot of predators μ_1 should become negative. the "simplest model" is a

† UNRELATED to P in §3.1.4

linear dependence on P:

$$\mu_1 = a_1 - b_1 P . \tag{83}$$

- $\mu_2 = \frac{1}{P}\frac{dP}{dt}$ is the specific rate of increase of the predator population. We expect growth ($\mu_2 > 0$) if there are lots of prey, but there should be decay ($\mu_2 < 0$) in the absence of prey. The simplest model is again the linear choice

$$\mu_2 = -a_2 + b_2 N . \tag{84}$$

Putting the pieces together we have

$$\boxed{\begin{aligned} \frac{dN}{dt} &= (a_1 - b_1 P) N \\ \frac{dP}{dt} &= (-a_2 + b_2 N) P \end{aligned}} \tag{85}$$

which are coupled first order ODEs. These are the Lotka-Volterra equations, studied around 1920.

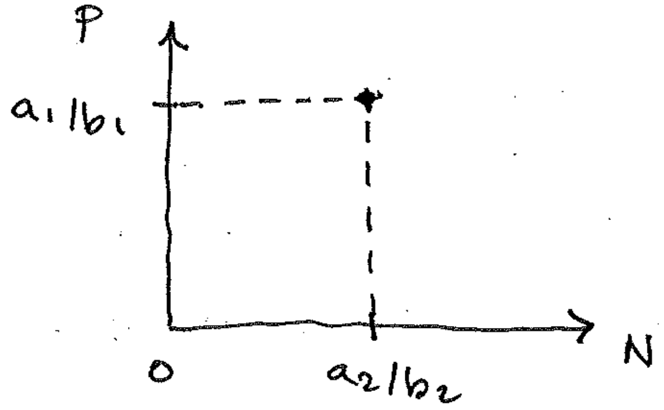
First note that we will have an equilibrium if $\frac{dN}{dt} = \frac{dP}{dt} = 0$ (the populations do not change with time). Setting the RHS of (85) to zero, this implies

$$N = P = 0$$

or

$$\boxed{N = N_{eq} \equiv a_2/b_2 , \quad P = P_{eq} \equiv a_1/b_1 .} \tag{86}$$

We ignore the first possibility, because we want there to be populations to study. Hence we have a single static equilibrium point at $(a_2/b_2, a_1/b_1)$ in the P-N plane:



We call this plane the "phase space," & use it to describe our system at different times.

If we have an initial population with $N = N_{eq}$ & $P = P_{eq}$ & perturb it a little (e.g. add a few individuals to each population by migration), will the system return to the equilibrium situation? Mathematically we write

$$N = \frac{a_2}{b_2} + \varepsilon_2$$
$$P = \frac{a_1}{b_1} + \varepsilon_1 \qquad (87)$$

where ε_1 & ε_2 are small increases in each population (or decreases).

Substituting the first of (87) into the first of (85) gives (check):

$$\frac{dN}{dt} = \frac{d\varepsilon_2}{dt} = (a_1 - a_1 - b_1\varepsilon_1)\left(\frac{a_2}{b_2} + \varepsilon_2\right) = -b_1\varepsilon_1\left(\frac{a_2}{b_2} + \varepsilon_2\right)$$

Since ε_1 & ε_2 are small, we can neglect the $\varepsilon_1\varepsilon_2$ term on the RHS by comparison with the other term, & we have:

$$\frac{d\varepsilon_2}{dt} \approx -\frac{a_2 b_1}{b_2}\varepsilon_1 . \qquad (88)$$

Similarly substituting the second of (17) into the second of (85) leads to

$$\frac{d\varepsilon_1}{dt} \approx \frac{b_2 a_1}{b_1}\varepsilon_2 . \qquad (89)$$

Differentiating (89) gives

$$\frac{d^2\varepsilon_1}{dt^2} = \frac{b_2 a_1}{b_1}\frac{d\varepsilon_2}{dt} = -a_1 a_2 \varepsilon_1 , \qquad (90)$$

using (88). The solution to (90) is

$$\varepsilon_1 = A\sin(\omega t + \phi) \qquad (91)$$

where

$$\omega = (a_1 a_2)^{\frac{1}{2}} . \qquad (92)$$

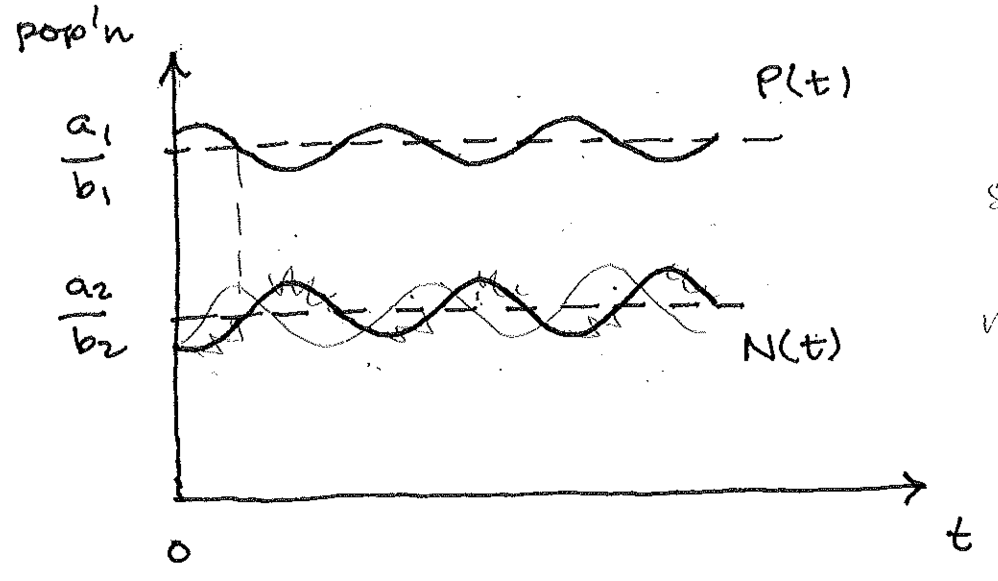
Using (89) gives the solution for ε_2:

$$\varepsilon_2 = B\cos(\omega t + \phi) \qquad (93)$$

where $B = \frac{b_1}{b_2}\left(\frac{a_2}{a_1}\right)^{\frac{1}{2}} A$. Equations (91) & (93) indicate that the populations oscillate around the equilibrium positions when perturbed. The period of the oscillation is

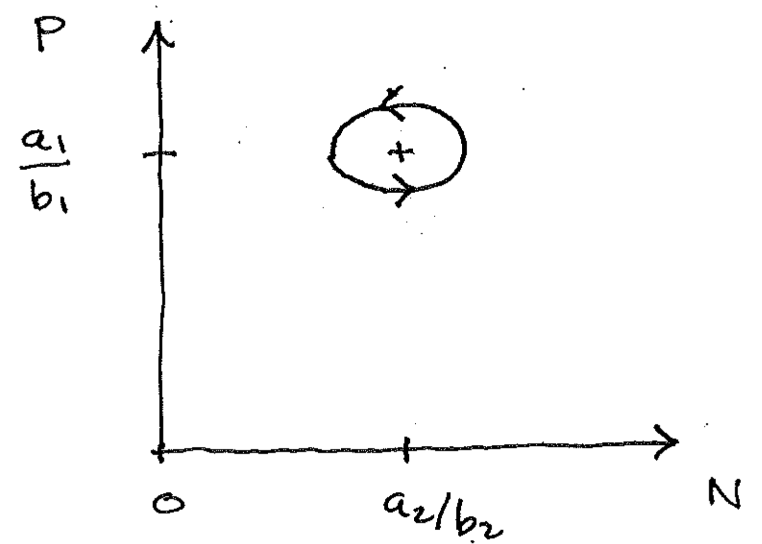
$$\boxed{T = \frac{2\pi}{\omega} = \frac{2\pi}{(a_1 a_2)^{\frac{1}{2}}} ,} \qquad (94)$$

& the oscillations in P & N are 90° out of phase.



should be 90°, not 180°!

In the phase space diagram we have a small ellipse:



This procedure (find equilibrium, perturb, determine behaviour) is called a "stability analysis." In throwing away the $\varepsilon_1 \varepsilon_2$ terms we have linearised the equations, & so the solution [(91) & (93)] is only valid for small perturbations. What about large excursions? There is no exact analytic solution to (85), so we ~~will~~ need to solve the ODEs numerically.

Before you solve any equation numerically you should non-dimensionalise. This is necessary because there may be fewer free parameters in the problem (and hence fewer things you need to vary in numerical solutions) than there appear to be. In non-dimensionalising a given set of equations it is not generally helpful to apply the Pi theorem (§1.2) because the results of the Pi theorem describe a general (unknown) relationship between variables. For a specific, known relationship fewer non-dimensional variables may be needed. The procedure is to construct non-dimensional variables for a given problem by trial & error.

First, introduce yet-to-be identified scale factors for the problem. For Eqs (85) we introduce t_s, N_s & P_s as scaling factors for time & the populations, & define the non-dimensional variables

$$\bar{t} = t/t_s \qquad \bar{N} = N/N_s \qquad \bar{P} = P/P_s . \tag{95}$$

Next rewrite (85) in terms of these variables:

$$\begin{aligned} \frac{d\bar{N}}{d\bar{t}} &= (a_1 t_s - b_1 P_s t_s \bar{P})\bar{N} \\ \frac{d\bar{P}}{d\bar{t}} &= (-a_2 t_s + b_2 N_s t_s \bar{N})\bar{P} \end{aligned} \tag{96}$$

Notice that all of the terms in (96) have no dimensions. Now we need to make choices for the scaling factors. If we choose

$$t_s = \frac{1}{a_1} \quad \& \quad P_s = \frac{a_1}{b_1} \left(= P_{eq}\right) \qquad (97)$$

& then the first of (96) becomes

$$\frac{d\bar{N}}{dt} = (1-\bar{P})\bar{N} \qquad (98)$$

which involves no constants. Making the analogous choice for Ns:

$$N_s = \frac{a_2}{b_2}\left(= N_{eq}\right) \qquad (99)$$

the second of (96) becomes

$$\frac{d\bar{P}}{dt} = \left(-\frac{a_2}{a_1} + \frac{a_2}{a_1}\bar{N}\right)\bar{P}$$

& we see that if we introduce

$$\bar{a} \equiv \frac{a_2}{a_1} \qquad (100)$$

we have the simpler form

$$\frac{d\bar{P}}{d\bar{t}} = -\bar{a}\,(1-\bar{N})\,\bar{P}. \qquad (101)$$

To summarise, our non-dimensional ODEs are

$$\boxed{\begin{aligned} \frac{d\bar{N}}{d\bar{t}} &= (1-\bar{P})\bar{N} \\ \& \quad \frac{d\bar{P}}{d\bar{t}} &= -\bar{a}(1-\bar{N})\bar{P} \end{aligned}} \qquad (102)$$

which now involve only one free parameter $\bar{a}$! We started with four (a_1, b_1, a_2, b_2) but in fact we see that we only need one. Hence we see the value of non-dimensionalising. Let me restate:

$\boxed{\text{NEVER SOLVE DIMENSIONAL EQUATIONS NUMERICALLY! NON-DIMENSIONALISE!}}$

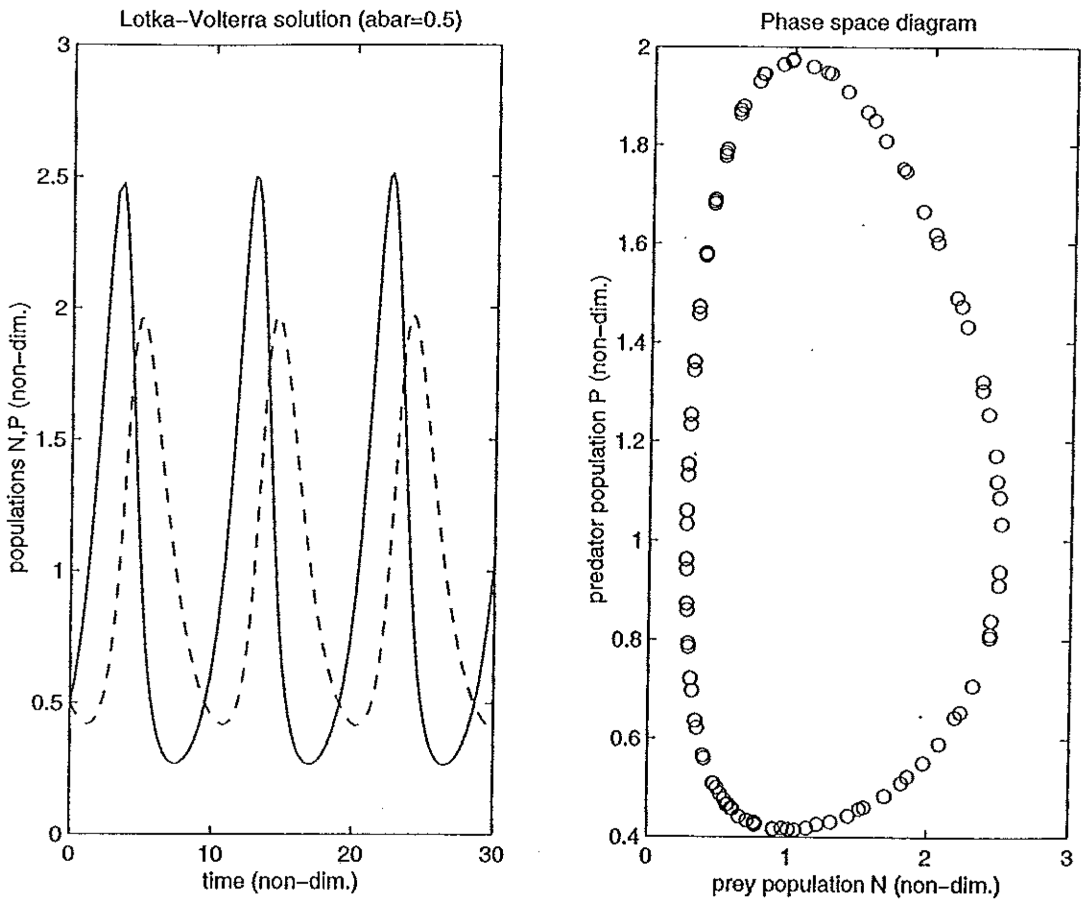
Anyway we need to solve (102), subject to specified $\bar{N}(0), \bar{P}(0)$, & for a given $\bar{a}$.

An example of a numerical solution (for $\bar{N}(0) = \frac{1}{2}$, $\bar{P}(0) = \frac{1}{2}$ & $\bar{a} = \frac{1}{2}$) is shown in Figure 3.3. The ODEs were solved in MATLAB using the ode45 routine, a built-in ODE solver. The script used to produce the solution (predator-prey1.m) is shown, together with the function (lotka-volterra.m) used to define the RHS of the ODEs.†

The numerical solution exhibits steady oscillations, ~~after~~ just as the linearised solution did, although in this case the oscillations are not sinusoidal, & $\bar{N}$ & $\bar{P}$ are not exactly 90° out of phase. The top left panel shows $\bar{N}(t)$ & $\bar{P}(t)$, & the top right panel shows the phase space diagram.

---

† These scripts are on the web




```
% predator_prey1.m
%
% Solve non-dimensional Lotka-Volterra equations
%
% M.S. Wheatland, 9 May 2001

global a % makes this variable global so it is visible to fn defining
         % RHS of ODEs

a=0.5;
N0=0.5;
P0=0.5;
TMAX=30; % maximum value of non-dimensional time

options = odeset('RelTol',1e-4,'AbsTol',[1e-4 1e-4]); % specify accuracy

% solve for t=0 to 10, using built-in procedure ode45
% RHS of ODEs is in lotka_volterra.m

[t,y]=ode45('lotka_volterra',[0 TMAX],[N0 P0]);

% plot results

subplot(1,2,1)
plot(t,y(:,1),'-',t,y(:,2),'--')
xlabel('time (non-dim.)')
ylabel('populations N,P (non-dim.)')
title(['Lotka-Volterra solution (abar=',num2str(a),')'])
subplot(1,2,2)
plot(y(:,1),y(:,2),'o')
xlabel('prey population N (non-dim.)')
ylabel('predator population P (non-dim.)')
title('Phase space diagram')
```

```
function dy=lotka_volterra(t,y)

% RHS of Lotka-Volterra ODEs

global a % recognise these variables from main program

dy=zeros(2,1); % define column vector

dy(1)=(1-y(2))*y(1);
dy(2)=-a*(1-y(1))*y(2);
```

Figure 3.3

Why do the populations oscillate?

- If the prey population increases then there is more for the predators to eat, & so the predator population increases. However, the increased predation reduces the number of prey. Then there is less for the predators to eat, & their population decreases. which allows the prey numbers to increase, & so on.

Equation (94) gives a rough estimate of the oscillation period (it is only accurate for small oscillations). We can rewrite it as a non-dimensional period,

$$\bar{T} = T/\left(\frac{1}{a_1}\right) = 2\pi\left(\frac{a_1}{a_2}\right)^{\frac{1}{2}} = \frac{2\pi}{\bar{a}^{1/2}} \qquad (103)$$

so for $\bar{a} = \frac{1}{2}$ we get $\bar{T} = 2\sqrt{2}\pi \simeq 8.9$. From the figure the period appears to be around 10, so this is not too bad, even though the oscillations are large.

Are such oscillations observed in nature? There is a famous example, that is still the subject of some debate: the lynx & snowshoe hare of America's northern coniferous forests. Trading records from the Hudson Bay Company show that the number of pelts obtained by trappers exhibit a roughly 10-year oscillation in numbers, with the number of lynx pelts lagging behind the number of hare pelts. If the numbers reflect the populations, then this may be an example of an oscillation in a predator-prey ecosystem (the lynx eats the hare, by the way). Figure 3.4 shows the trading records, over 90 years. There is a relatively convincing oscillation in both time series, with the lynx numbers lagging the hare numbers, most of the time.

Does the Lotka-Volterra result (e.g. figure 3.3) explain the oscillation implied by Figure 3.4? This is a difficult question. Some insight is gained by the observation that the majority of predator-prey ecosystems observed in nature do not exhibit cycles, e.g. the caribou-wolf populations in Alaska do not oscillate[†]. However, the Lotka-Volterra ODEs always produce oscillations, unless the populations are at the single equilibrium point $(N_{eq}, P_{eq}) = (a_2/b_2, a_1/b_1)$, or in non dimensional terms $(\bar{N}_{eq}, \bar{P}_{eq}) = (1,1)$. For real populations it is expected that there are always perturbations, & so the system would not remain at the equilibrium point.

† See e.g. May

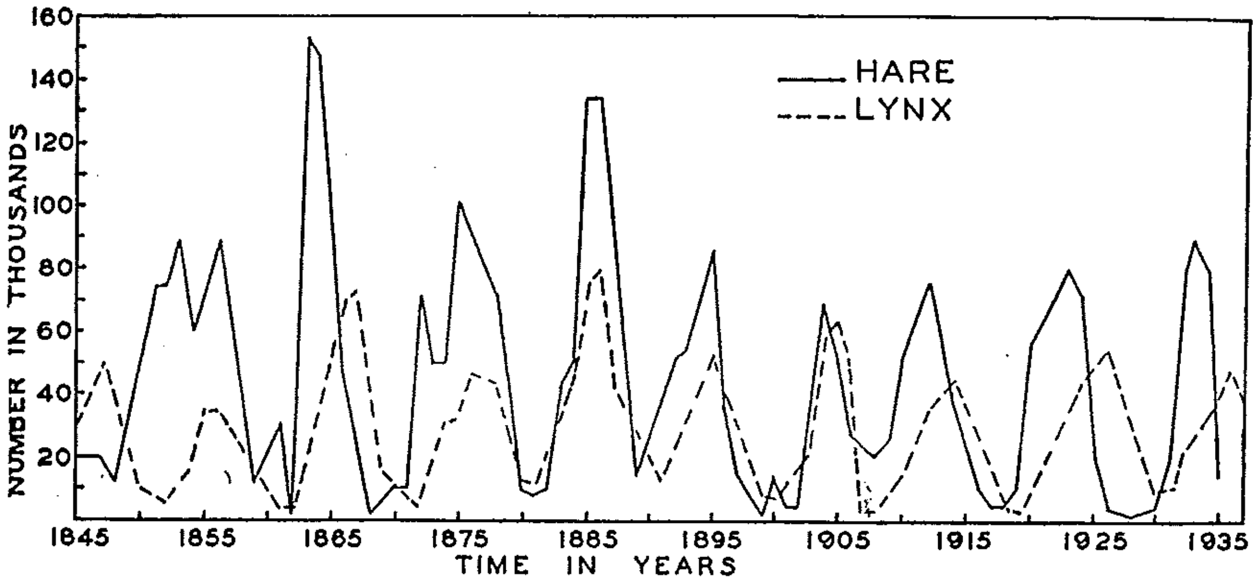
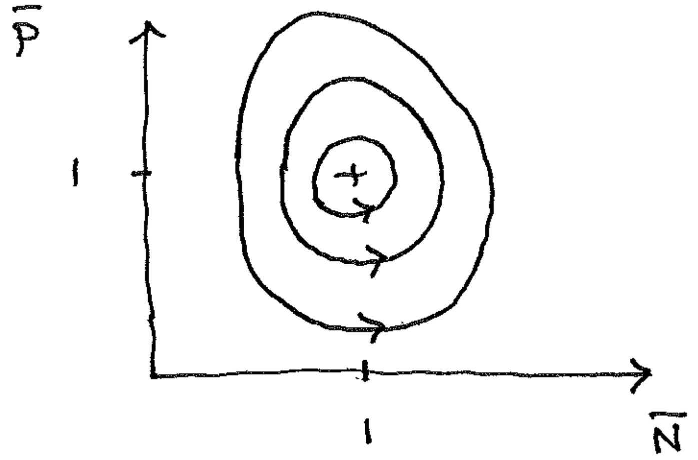


Figure 54. Changes in the abundance of the lynx and the snowshoe hare, as indicated by the number of pelts received by the Hudson Bay Company. This is a classic case of cyclic oscillation in population density. (Redrawn from MacLulich, 1937.)

Figure 3.4

From: Odum (1959), Fundamentals of Ecology, W.B. Saunders Company, Philadelphia

Away from the equilibrium point there are only oscillations, with an amplitude determined by $\bar{N}(0)$ & $\bar{P}(0)$, & a shape determined by $\bar{a}$:



Fixed $\bar{a}$, different $(\bar{N}(0), \bar{P}(0))$

Hence we conclude that the Lotka-Volterra model is unable to account for the majority of predator-prey ecosystems, which do not oscillate.

A second problem with the L-V model becomes apparent when we try to improve the model a bit. A criticism of the model is that, in the absence of predators, the numbers of prey grow exponentially. We know that Malthusian growth is silly for large times, so it seems reasonable to adopt our solution from §3.1.3 & assume logistic growth, in the absence of predators:

$$\frac{1}{N}\frac{dN}{dt}\bigg|_{P=0} = a_1\left(1-\frac{N}{N_*}\right)$$

where N_* is the carrying capacity of the prey (sans predators).

Including predators it then seems reasonable to write

$$\frac{1}{N}\frac{dN}{dt} = a_1\left(1-\frac{N}{N_*}\right) - b_1 P \qquad (104)$$

& for prey we still have

$$\frac{1}{P}\frac{dP}{dt} = -a_2 + b_2 N . \qquad (105)$$

Since N_* is expected to be large, the term N/N_* should be small. We have added a small term to the RHS, & hence it seems reasonable to assume that the results should not change much. Right?

Exercise: Show that the equilibrium points for (104) & (105) are

$$N = P = 0$$

or $\quad P = 0, \quad N = N_*$

or $\quad N = a_2/b_2, \quad P = \frac{a_1}{b_1}\left(1-\frac{a_2}{b_2 N_*}\right). \qquad (106)$

Exercise: Try a linear stability analysis about the non-trivial equilibrium, as follows. Look for a solution of the form

$$N = \frac{a_2}{b_2} + \varepsilon_2 , \quad P = \frac{a_1}{b_1}\left(1-\frac{a_2}{b_2 N_*}\right) + \varepsilon_1 ,$$

& show that ε_2 satisfies

$$\frac{d^2\varepsilon_2}{dt^2} + a_1\delta\frac{d\varepsilon_2}{dt} + a_1 a_2(1-\delta)\varepsilon_2 = 0 \qquad (107)$$

where $\boxed{\delta \equiv \frac{a_2/b_2}{N_*} = \frac{N_{eq}}{N_*}}$ (108) is a parameter that can be assumed to be small. By looking for a solution to (107) of the form $\varepsilon_2 = A e^{\lambda_2 t}$ show that

$$\lambda_2 \simeq -\tfrac{1}{2} a_1 \delta \pm i\,(a_1 a_2)^{\frac{1}{2}}\left(1 - \tfrac{1}{2}\delta\right).$$

The imaginary part of λ_2 corresponds to oscillation & the real part (which is negative) corresponds to a decay of the oscillation in time.

The exercise above suggests that the modified Lotka-Volterra equations (104) & (105) do not exhibit sustained oscillations, but that the oscillations damp out with time. To verify this small-amplitude result we can numerically solve the equations. Before we do that we need to non-dimensionalise.

Once again we can use

$$\bar{t} = a_1 t$$

$$\bar{P} = P/(a_1/b_1)$$

$$\bar{N} = N/(a_2/b_2)$$

& then it is straightforward to show that the non-dimensional version of (104),(105) is

$$\frac{d\bar{N}}{d\bar{t}} = \left[(1 - \delta \bar{N}) - \bar{P}\right]\bar{N} \qquad (109)$$

$$\frac{d\bar{P}}{d\bar{t}} = -\bar{a}\,(1 - \bar{N})\,\bar{P},$$

where δ is given by (108) & where (as before) $\bar{a} \equiv a_2/a_1$.

Exercise: confirm (109).

Figure 3.5 shows a numerical solution (for $N(0) = P(0) = \bar{a} = \frac{1}{2}$, $\delta = 0.2$) which confirms the result of the stability analysis: the cycles die away.† Although we have only shown one choice of $N(0)$, $P(0)$, $\bar{a}$ & δ, this is true in general. So we have made a small change to the RHS of the Lotka-Volterra equations & we have significantly changed their behaviour. Whereas previously they apparently accounted for population oscillations & could not explain steady state predator-prey ecosystems, now it seems the situation is reversed. This is not very satisfactory, & makes us doubt the ability of the equations to explain predator-prey systems.

How can we fix this situation? Clearly we need to consider further modification of our equations:

$$\frac{dN}{dt} = \left[a_1\left(1 - \frac{N}{N_*}\right) - b_1 P\right] N \qquad (104)$$

$$\frac{dP}{dt} = (-a_2 + b_2 N)P. \qquad (105)$$

† The scripts are on the web: predator-prey2.m & lotka-volterra_mod.m

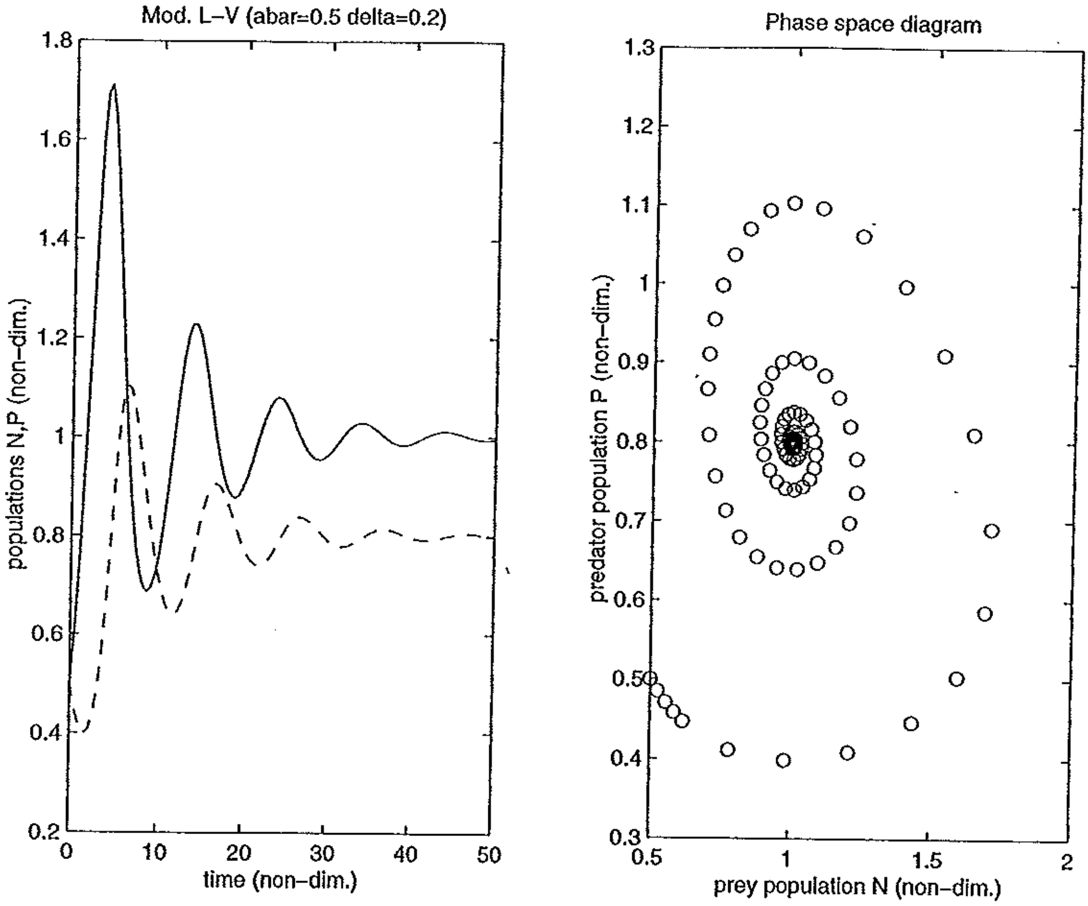

Mod. L–V (abar=0.5 delta=0.2)
populations N,P (non–dim.)
time (non–dim.)
0
10
20
30
40
50
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
Phase space diagram
predator population P (non–dim.)
prey population N (non–dim.)
0.5
1
1.5
2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
1.1
1.2
1.3

Figure 3.5

Note that (105) implies that the predator population decays exponentially if $N < a_2/b_2$, i.e. if the prey population is below its equilibrium value. This seems too drastic. A milder assumption is that the predator population obeys logistic growth with a capacity proportional to the prey population:

$$\frac{dP}{dt} = a_2 P\left(1 - \frac{P}{P_*}\right), \qquad (106)$$

where $P_* = b_2 N$. In this case when the prey population is low the predator population does not head towards zero, but heads towards a small value commensurate with the prey numbers.

Next consider the prey equation, (104). The final term $-b_1 PN$ on the RHS represents the rate of removal of prey by predators. Hence $b_1 N$ is the rate of removal of prey by each predator. Obviously there is a limit to how much prey an individual predator can remove in a given time, & so this term is unrealistic for large N.

- Better to assume each predator removes $\phi(N)$ prey per unit time, where

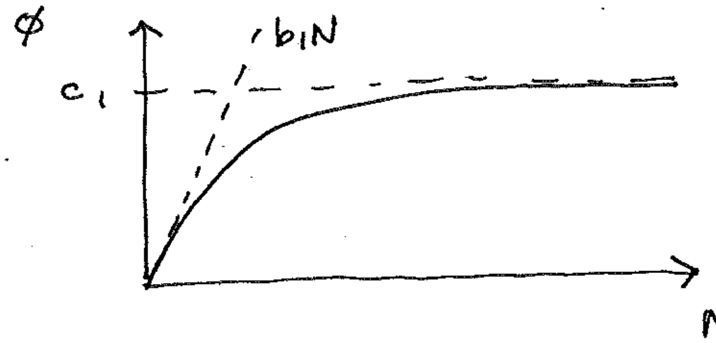
$$\phi(N) \approx b_1 N \quad \text{for small } N$$

$$\phi(N) \to c_1, \text{ a constant}, \quad \text{as } N \to \infty.$$

The parameter c_1 is the maximum rate of removal of prey by an individual predator.

There are many possible choices for $\phi(N)$. Following Mesterton-Gibbons we consider the simple choice

$$\phi(N) = \frac{c_1 N}{N + c_1/b_1}, \qquad (107)$$



which has a history in the literature (May). With this choice the prey equation becomes

$$\frac{dN}{dt} = a_1\left(1 - \frac{N}{N_*}\right)N - \frac{c_1 N P}{N + c_1/b_1}, \qquad (108)$$

& to restate the new predator equation,

$$\frac{dP}{dt} = a_2 P\left(1 - \frac{P}{b_2 N}\right). \qquad (106)$$

We need to solve these equations. We could attempt a stability analysis, but instead will just numerically solve them. Once again we begin by non-dimensionalising.

<u>Exercise</u>: Introducing $\bar{N} = \frac{N}{N_*}$, $\bar{P} = \frac{P}{b_2 N_*}$, $\bar{t} = a_2 t$

show that the non-dimensional versions of (106) & (108) are

$$\frac{d\bar{N}}{d\bar{t}} = \left[a(1-\bar{N}) - \frac{b\bar{P}}{\bar{N}+c}\right]\bar{N}$$
$$\frac{d\bar{P}}{d\bar{t}} = \bar{P}(1-\bar{P}/\bar{N}) \tag{109}$$

where $a \equiv \dfrac{a_1}{a_2}$ $\quad b \equiv \dfrac{c_1 b_2}{a_2}$ $\quad c \equiv \dfrac{c_1}{b_1 N_*}$. (110)

Now we have three parameters (a,b,c), which have interpretations as follows.

- From (108) the parameter a_1 is the maximum specific growth rate of the prey in the absence of predators & ~~a~~ from (106) a_2 is the max. specific growth rate of the predators. Hence $a \equiv a_1/a_2$ is the ratio of growth rates under ideal conditions.

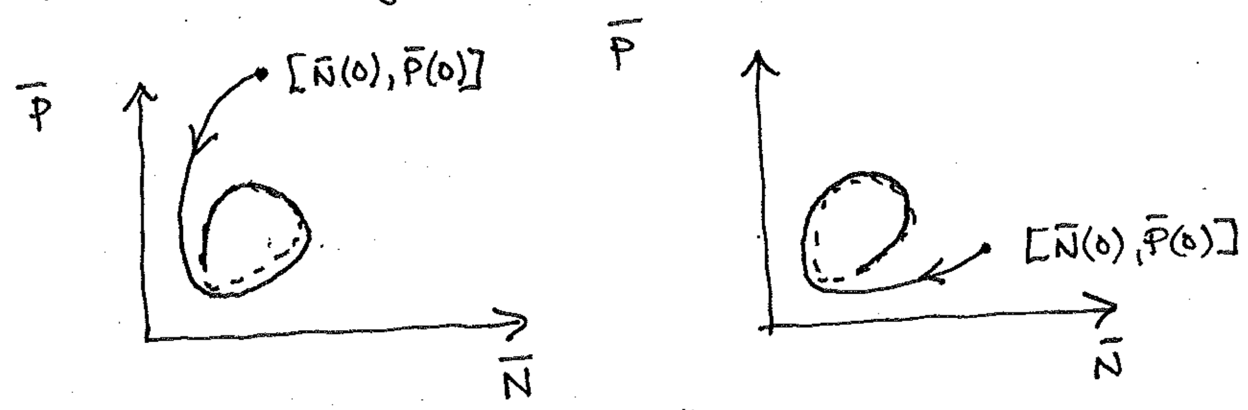
- From (107) the quantity c_1/b_1 is the prey population at which the removal rate is half of its maximum value, & from (108) N_* is the prey capacity in the absence of predation. Hence $c \equiv c_1/b_1 N_*$ is a measure of the effectiveness of the predators at removing prey: if c is small the predators' rate of removal of prey saturates at a fraction of the prey capacity.

- Finally $b \equiv \dfrac{c_1 b_2}{a_2}$: this is a measure of the predator's maximum prey removal rate (c_1) relative to its ideal growth rate (a_2), modulated by its ability to cope with small prey populations (b_2).

Figure 3.6 illustrates two numerical solutions (for $\bar{N}(0) = \bar{P}(0) = \frac{1}{2}$, $a = 5, b = 10, c = 0.15$, & for $\bar{N}(0) = \bar{P}(0) = \frac{1}{2}$, $a = 5$, $b = 10$, $c = 0.5$). If c is small (e.g. the top panels), i.e. the predators are relatively ineffective at removing prey, you get cycles. If c is large, i.e. the predator is good at removing prey, the system goes to an equilibrium point (e.g. the bottom panels).

Exercise: Determine the equilibrium points for (109)

- Note that the cycle in the top panel is not a rigid oscillation (cf. the Lotka-Volterra solution, Figure 3.3) with amplitude dictated by the initial conditions. Instead we have a cycle that is approached from $(\bar{N}(0), \bar{P}(0))$. If you start at a different $(\bar{N}(0), \bar{P}(0))$ you end up on the same cycle[†]:



This is a "limit cycle", & represents a more robust behaviour than rigid oscillation, & is better suited to modelling biological populations. (See e.g. May)

† You can try this: the scripts predatorprey3.m & lotka-volterra_Mod2.m

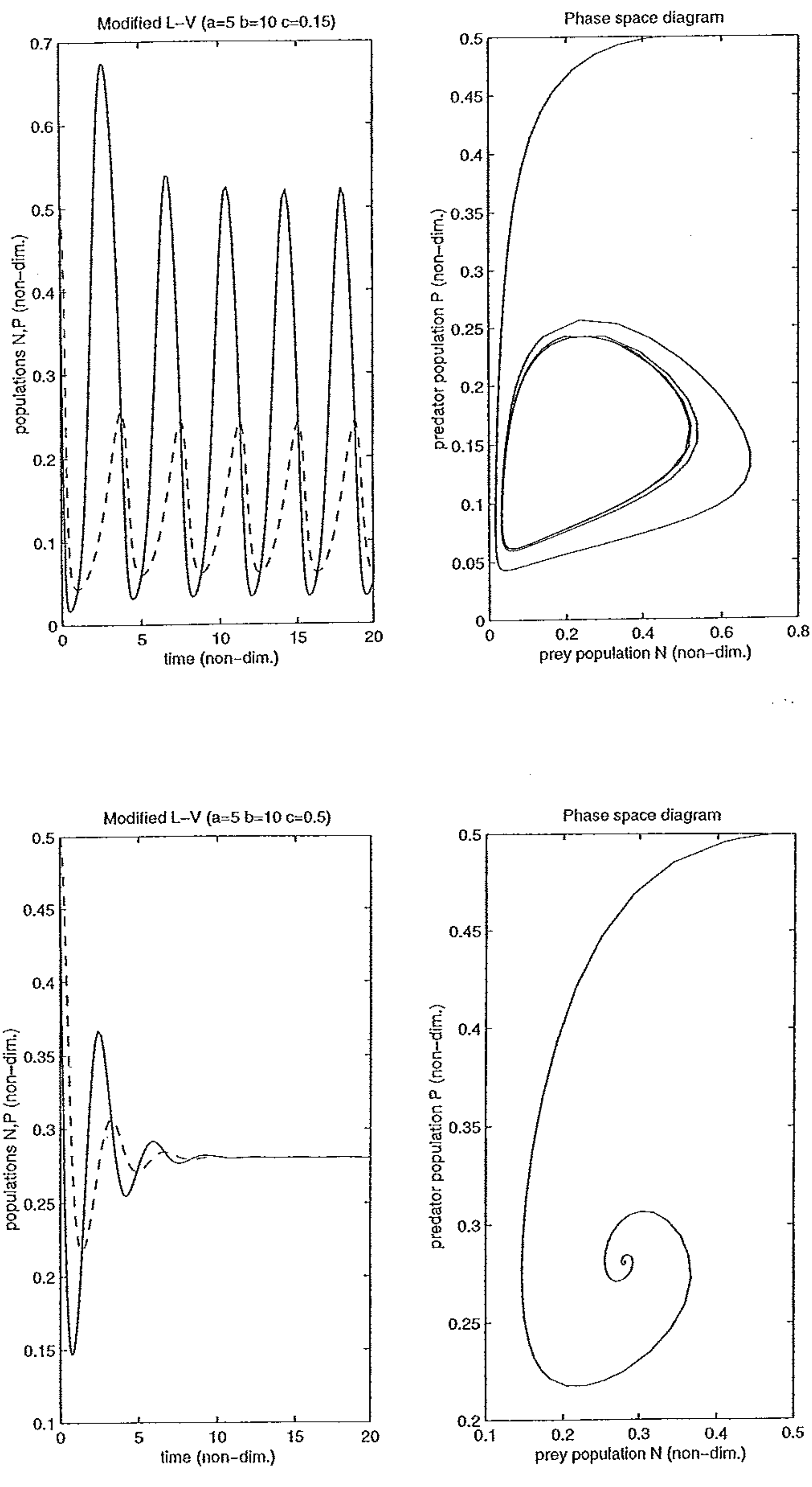

Modified L–V (a=5 b=10 c=0.15)
populations N,P (non–dim.)
time (non–dim.)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
5
10
15
20
Phase space diagram
predator population P (non–dim.)
prey population N (non–dim.)
0.05
0.15
0.25
0.35
0.45
0.8
Modified L–V (a=5 b=10 c=0.5)
populations N,P (non–dim.)
time (non–dim.)
Phase space diagram
predator population P (non–dim.)
prey population N (non–dim.)

Figure 3.6

• The new model can account for both equilibria & cycles in predator-prey systems, & also predicts when each type of behaviour is expected. This - together with the stability properties already mentioned - make it a better model for real predator-prey systems, & it may explain oscillations such as those depicted (or suggested) by Figure 3.4.

3.2.2 Competing populations :

We will briefly consider a second simple model of an ecosystem consisting of two species in competition for the same food supply (e.g. herbivores competing for grass). Sticking with the labels $N(t)$ & $P(t)$ for the two populations we have

$$\frac{1}{N}\frac{dN}{dt} = \mu_1 , \quad \frac{1}{P}\frac{dP}{dt} = \mu_2 \qquad (111)$$

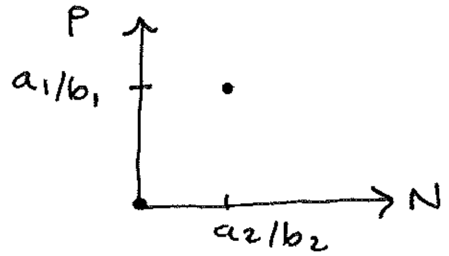
where μ_1 & μ_2 are the specific growth rates of each population. Based on the reasoning leading to the Lotka-Volterra equations (85) we can try

$$\mu_1 = a_1 - b_1 P , \qquad \mu_2 = a_2 - b_2 N . \qquad (112)$$

Exercise: Show that the equilibrium points for (111) with (112) are:

$$N = P = 0$$

$$\text{or} \qquad N = \frac{a_2}{b_2}, \; P = \frac{a_1}{b_1}$$



Exercise: Consider a perturbation around the non-trivial equilibrium:

$$N = \frac{a_2}{b_2} + \varepsilon_2 \qquad P = \frac{a_1}{b_1} + \varepsilon_1 .$$

Show that ε_1 satisfies

$$\frac{d^2\varepsilon_1}{dt^2} = a_1 a_2 \varepsilon_1 ,$$

& that this has solution $\varepsilon_1 = Ae^{\lambda t} + Be^{-\lambda t}$ where $\lambda = (a_1 a_2)^{\frac{1}{2}}$.

This exercise shows that there is a growing exponential component to any perturbation from the equilibrium, & hence the system moves away from equilibrium when disturbed. The ~~so~~ equilibrium is "unstable".

There is no exact analytic solution to (111) – (112). We could solve the system numerically, but instead we will examine the global behaviour of solutions using the phase space diagram.

Exercise: Show that the non-dimensional version of (111) – (112) is

$$\frac{d\bar{N}}{d\bar{t}} = (1-\bar{P})\bar{N} \qquad (113)$$

$$\frac{d\bar{P}}{d\bar{t}} = \bar{a}(1-\bar{N})\bar{P}$$

where

$$\bar{N} = N/(a_2/b_2), \quad \bar{P} = P/(a_1/b_1), \quad \bar{t} = t/(1/a_1)$$

& where $\bar{a} = a_2/a_1$.

In a short time $\Delta\bar{t}$ the populations change by $\Delta\bar{N}$ & $\Delta\bar{P}$, where

$$\Delta\bar{N} \approx \frac{d\bar{N}}{d\bar{t}}\,\Delta\bar{t} = (1-\bar{P})\bar{N}\,\Delta\bar{t}$$

$$\& \quad \Delta\bar{P} \approx \frac{d\bar{P}}{d\bar{t}}\,\Delta\bar{t} = \bar{a}(1-\bar{N})\bar{P}\,\Delta\bar{t}\,.$$

If we plot the vector $\underset{\sim}{V} = (\Delta\bar{N}, \Delta\bar{P})$ in the $\bar{N}$-$\bar{P}$ plane at the point $(\bar{N},\bar{P})$, then this vector describes how the populations change in time $\Delta\bar{t}$. In particular the direction of the vector indicates the direction of evolution of the populations, i.e. the instantaneous trajectory of the system. Since we are only interested in the direction, we can introduce the vector

$$\underset{\sim}{v} = \underset{\sim}{V}/|\underset{\sim}{V}| = \frac{[(1-\bar{P})\bar{N},\ \bar{a}(1-\bar{N})\bar{P}]}{\{(1-\bar{P})^2\bar{N}^2 + \bar{a}^2(1-\bar{N})^2\bar{P}^2\}^{\frac{1}{2}}} \tag{114}$$

which does not depend on $\Delta\bar{t}$.

Instead of solving (113), we can just plot the vector $\underset{\sim}{v}$ throughout the $\bar{N}$-$\bar{P}$ plane, & by following the vectors, guess how the system evolves.

Figure 3.7 shows the result (for $\bar{a} = 0.8$). It is easy to infer from this diagram that the trajectories are as follows:

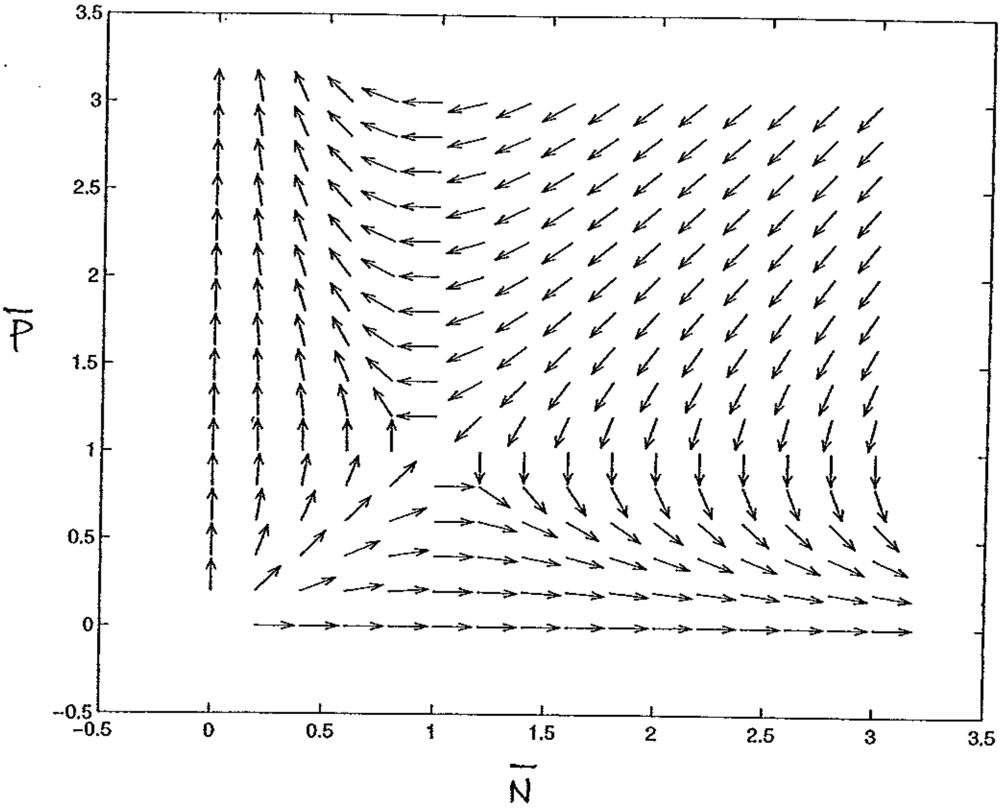
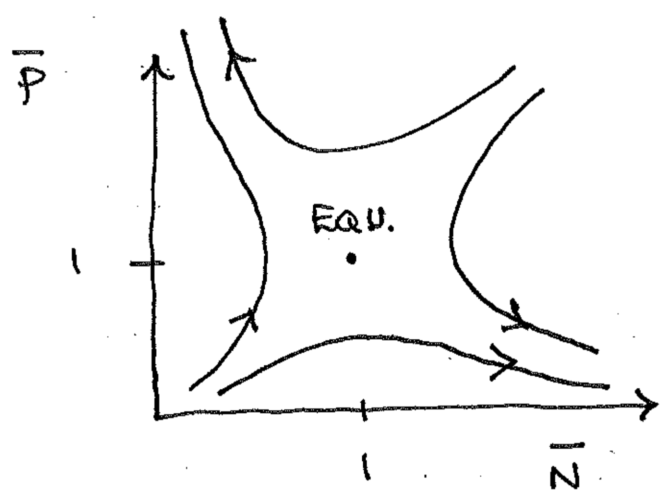


Figure 3.7



- Notice that one population always heads to zero, irrespective of the starting point (you could start at the equilibrium, but as we have seen this is not stable, so you would end up following one of the trajectories shown).

- This is true irrespective of $\bar{a}$, although we have only tried one value.

In fact there is evidence in nature in favour of the proposition that, if an ecosystem consists of two species competing for the same food supply, only one will ultimately survive. This rule is often called "Gause's ~~exclusionary~~ principle of competitive exclusion" & Equations (111) - (112) are called Gause's equations.

But how firm is our theoretical basis for such an exclusion principle? With the Lotka-Volterra equations we found that introducing a small change to the RHS of the equations produced drastically different behaviour. Will the same thing happen here, making us doubt the conclusions of the model?

Each species grows exponentially in the absence of the other. As we have seen this is unrealistic, & the usual solution is to assume logistic growth. This suggests the revised equations

$$\frac{1}{N}\frac{dN}{dt} = a_1\left(1-\frac{N}{N_*}\right) - b_1 P \qquad (115)$$

$$\frac{1}{P}\frac{dP}{dt} = a_2\left(1-\frac{P}{P_*}\right) - b_2 N.$$

For simplicity of notation we will write $c_1 = a_1/N_*$, $c_2 = a_2/P_*$, & hence

$$\boxed{\begin{aligned}\frac{dN}{dt} &= (a_1 - c_1 N - b_1 P)N \\ \frac{dP}{dt} &= (a_2 - c_2 P - b_2 N)P\end{aligned}} \qquad (116)$$

First consider the equilibria, i.e. N & P satisfying $\frac{dN}{dt} = \frac{dP}{dt} = 0$. Clearly we have

$$N = 0 \quad \text{or} \quad a_1 - c_1 N - b_1 P = 0 \qquad (117)$$

$$\& \quad P = 0 \quad \text{or} \quad a_2 - c_2 P - b_2 N = 0.$$

We exclude the possibility $N = P = 0$ since we want there to be populations to study.

If $N=0$ & $P \neq 0$ then $P = a_2/c_2$. Hence one equilibrium is

EQ1: $$\boxed{N=0, \quad P = a_2/c_2.} \tag{118}$$

Alternatively if $P=0$ & $N \neq 0$ then $N = a_1/c_1$, & we have a second equilibrium:

EQ2: $$\boxed{N = a_1/c_1, \quad P=0.} \tag{119}$$

Finally if both $N \neq 0$ & $P \neq 0$ then there may be an equilibrium defined by the intersection of two lines in the (N,P) plane:

$$\boxed{\begin{array}{ll} L_1: & a_1 - c_1 N - b_1 P = 0 \\ L_2: & a_2 - c_2 P - b_2 N = 0. \end{array}} \tag{120}$$

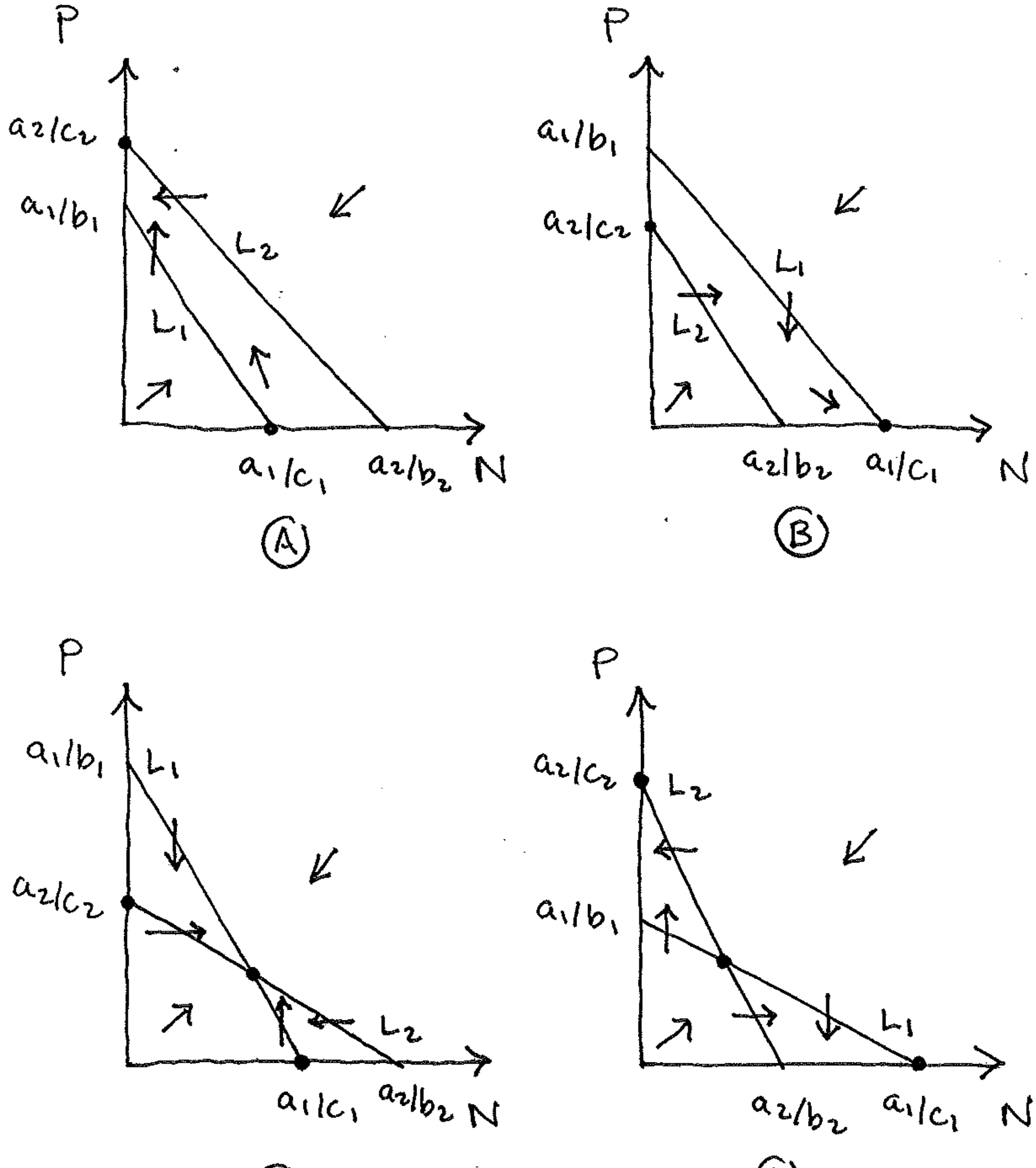
These lines may not intersect (in the region $N>0$, $P>0$).

Exercise: Show that the solution to (120) is

EQ3: $$N = \frac{a_1 c_2 - a_2 b_1}{c_1 c_2 - b_1 b_2}, \quad P = \frac{a_1 b_2 - a_2 c_1}{b_1 b_2 - c_2 c_1}. \tag{121}$$

The possible equilibria are shown graphically in Figure 3.8. There are four possible situations, depending on the relative sizes of a_1/c_1 & a_2/b_2 (the N intercepts of L_1 & L_2) and the relative sizes of a_1/b_1 & a_2/c_2 (the P intercepts of L_1 & L_2). The arrows in the figure indicate the direction of trajectories, as specified by the vector $(dN/dt, dP/dt)$, at various points in the plane.

Figure 3.8


P
a_2/c_2
a_1/b_1
L_2
L_1
a_1/c_1
a_2/b_2
N
A
P
a_1/b_1
a_2/c_2
L_1
L_2
a_2/b_2
a_1/c_1
N
B
P
a_1/b_1
L_1
a_2/c_2
L_2
a_1/c_1
a_2/b_2
N
C
P
a_2/c_2
L_2
a_1/b_1
L_1
a_2/b_2
a_1/c_1
N
D

It is easy to work out these directions by noting that:

$$\frac{dN}{dt} > 0 \quad \text{below } L_1$$

$$\frac{dN}{dt} = 0 \quad \text{on } L_1$$

$$\frac{dN}{dt} < 0 \quad \text{above } L_1$$

$$\frac{dP}{dt} > 0 \quad \text{below } L_2$$

$$\frac{dP}{dt} = 0 \quad \text{on } L_2$$

$$\frac{dP}{dt} < 0 \quad \text{above } L_2 ,$$

based on the RHS's of (116).

Based on these representative directions of trajectories it is easy to see that in Ⓐ the attracting equilibrium is $(0, a_2/c_2)$, i.e. EQ.1. In Ⓑ the attracting equilibrium is $(\frac{a_1}{c_1}, 0)$, i.e. EQ.2. In Ⓒ the attracting equilibrium is EQ3, & in Ⓓ both EQ1 & EQ2 are locally attracting.

Hence there are three possible stable equilibria. The different scenarios (Ⓐ, Ⓑ, Ⓒ or Ⓓ) are determined by the relative sizes of $\frac{a_1}{c_1}, \frac{a_2}{b_2}$ & $\frac{a_1}{b_1}, \frac{a_2}{c_2}$. In only one case is EQ3 apparently a stable equilibrium, namely Ⓒ, or

$$\boxed{\frac{a_2}{c_2} < \frac{a_1}{b_1} \quad \& \quad \frac{a_1}{c_1} < \frac{a_2}{b_2}.} \qquad (122)$$

This is an interesting case because EQ. 3 then represents a stable equilibrium with the two species coexisting in harmony. This contradicts Gause's principle of competitive exclusion, & is at odds with our original equations [(111) & (112)]. What is going on?

To understand this we need to understand the meaning of the inequalities (122). Note that, according to (116),

$\frac{a_1}{c_1}$ describes the ratio of intrinsic growth of N to competition of N with itself (crowding)

$\frac{a_2}{b_2}$ describes the ratio of intrinsic growth of P to competition of P with N

Hence if $\frac{a_1}{c_1} < \frac{a_2}{b_2}$ then competition of N with itself is more important in limiting the growth of N than competition of P with N is in limiting the growth of P. Similar reasoning shows us that the other inequality in (122) states that competition of P with itself is a more important factor in limiting the growth of P than competition of N with P is in limiting the growth of N.

In simpler terms, if (122) hold then both species compete more with themselves than they do with the other species.

Hence we see that the two species can coexist, but it requires that they are only in weak competition with one another. The other two equilibria (EQ.1, EQ.2) represent exclusion of one species by the other, & apply when there is strong competition. Hence we see that the exclusion principle still applies, albeit in a weaker form.

In conclusion, (116) is probably a better model than (111)-(112), & we have learnt something new in its formulation. Unfortunately biological data is pretty equivocal, & it ~~has~~ is not ~~been~~ possible to rigorously test models of interacting populations. The principle of competitive exclusion has some evidence in its favour, & the model (116) provides a theoretical basis. You may be interested to read further on this topic - see e.g. May.

Next we consider age structure within a population, beginning with a phenomenological law of mortality for human populations.

3.3 Age structure :

So far we have not considered the age structure of a population. Obviously the age structure is important (for many populations) in determining the birth & death rates.

PHYS220 Scientific Modelling 2001
Assignment 3 – due Monday June 4

1. The logistic equation introduced in lectures was

$$\frac{1}{N}\frac{dN}{dt} = a(1 - N/N_*). \tag{1}$$

By making the replacements

$$\frac{dN}{dt} \to \frac{N_{i+1} - N_i}{\Delta t}, \quad N \to N_i$$

discretise the logistic equation and put it into the form of the logistic map,

$$x_{i+1} = r(1 - x_i)x_i, \tag{2}$$

where $x_i = N_i/A$ for some factor A.

   (a) Identify the factor A and the coefficient r (they depend on N_*, a and Δt).

   (b) Use the result of (a) to explain why an accurate numerical solution of Equation (1) using the discrete version (2) will not produce limit cycles or chaos. (You may need some results concerning the logistic map from Barry's half of the course.)

2. State whether the following populations are best described by a differential equation or by a (finite-step) difference equation:

   (a) a small population of mammals,

   (b) Pacific salmon that spawn once a year, with the adults dying shortly after spawning,

   (c) bacteria in a Petri dish, with a very large growth rate.

   Also, which of the above could (in principle) produce chaotic behaviour?

3. Consider the variation on logistic growth,

$$\frac{1}{N}\frac{dN}{dt} = a\left[1 - (N/N_*)^2\right] \tag{3}$$

(the "crowding" depends on the square of the population).

   (a) Write down the ODE satisfied by $u = N^2$.

   (b) Using the result of (a) or otherwise, solve Equation (3) subject to $N = N_0$ at $t = 0$.

   (c) Sketch the solution found in (b) for $N_* \gg N_0$, and on the same graph show the solution to the logistic equation (1) (for the same values of a, N_* and N_0).

4. The Lotka-Volterra equations may be written

$$\begin{aligned} \frac{dN}{dt} &= (a_1 - b_1 P)N \\ \frac{dP}{dt} &= (-a_2 + b_2 N)P. \end{aligned} \tag{4}$$

Show that trajectories (in the phase space) of solutions to Equation (4) satisfy

$$a_2 \ln N + a_1 \ln P - b_2 N - b_1 P + C = 0,$$

where C is a constant which labels a solution for a given $N(0)$ and $P(0)$.

How can we describe this dependence?

We begin with a morbid question: how do individuals in a population die?

3.3.1 Modelling mortality:

Demographers introduce a special kind of population called a 'birth cohort', which consists of a group of individuals all born at the same time. This population only decreases with time, as members die: we ignore the numbers of offspring produced by the individuals in a birth cohort. If we assume all members are born at time $t=0$ then we can replace the time variable by an age variable z, & then $N(z)$ is the number of members of the cohort who survive to age z. The growth equation for this population is a decay equation (since there are only deaths):

$$\boxed{\frac{1}{N}\frac{dN}{dz} = -\mu(z)} \qquad (123)$$

where $\mu(z)$ is the specific death rate, as a function of age.† We include age dependence in the death rate to model how members of the cohort die.

What is the functional form of $\mu(z)$? Benjamin Gompertz (1825) was the first person to address this question for the human

† Demographers call this the "mortality," hence μ.

population. He wrote of an individual that:

"at the end of equal infinitely small intervals of time, he lost equal portions of his remaining power to avoid destruction."

If we denote the "power to avoid destruction" by P, then we can interpret this statement mathematically as

$$\frac{\Delta P}{P} = -C\,\Delta z, \qquad (C \text{ const.})$$

i.e. the fractional rate of decrease of P is proportional to the increment in age Δz, for small increments. In the limit $\Delta z \to 0$ we have

$$\frac{dP}{dz} = -CP$$

which has solution

$$P = P_0 e^{-Cz}. \tag{124}$$

If we interpret the "power to avoid destruction" to be inversely proportional to the probability of death per unit age,

$$P = \frac{a}{\mu}$$

for some a, then (124) becomes

$$\frac{P_0}{P} = \frac{\mu}{\mu_0} = e^{Cz}$$

i.e. $\boxed{\mu = \mu_0 e^{Cz},}$ (125)

which is the usual form of the "Gompertz law" of mortality. The probability of death per unit time (age) increases with age, at an increasing rate.

Exercise: The Gompertz form of $\mu(z)$ is not normalisable, i.e. we require $\int_0^\infty \mu(z)\,dz = 1$ (why?) but this is not possible for the functional form (125). Resolve this problem.

The Gompertz law provides a reasonable description of the way real human populations die. To see this I have used a U.S. "life table" constructed by the US National Centre for Health Statistics in 1996. This table (which is constructed using observed death rates) lists the probability of dying within a year as a function of age. This probability (given as a proportion of an initial birth cohort dying) is plotted in Figure 3.9. The graph is log-linear, so an exponential shows up as a straight line. We see that for older individuals there is a clear straight line trend. There is also a peak around zero (infant mortality may be different), & a small bump around age 20.

It is reasonable to assume that in addition to the natural deteriotation underlying the Gompertz law, individuals are subject to a ~~constant~~ risk of death by accident throughout their life. To a first approximation this risk may be assumed constant. In that case an improved version of (125) is

$$\mu = \mu_0 + \mu_1 e^{cz} \qquad (126)$$

where μ_0 is a constant. This form gives an

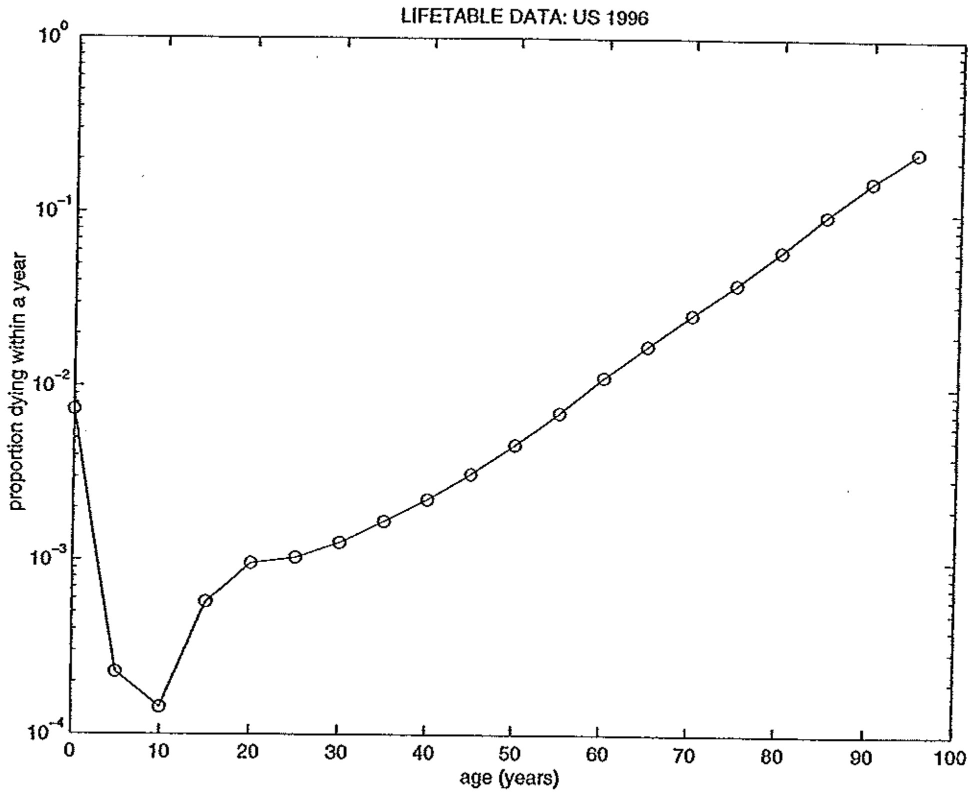

LIFETABLE DATA: US 1996
proportion dying within a year
age (years)
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
0
10
20
30
40
50
60
70
80
90
100

Figure 3.9

improved fit to data & is due to Malcoham (1860). However, it still fails to account for infant mortality & the bump in mortality around age 20. Presumably infant mortality is different, & young individuals take more risks. Incidentally, other populations do not necessarily follow the Gompertz law.

Returning to (123), the general solution is obtained by integrating:

$$\int \frac{dN}{N} = -\int \mu \, dz$$

or

$$\ln N = -\int \mu \, dz + \text{const}$$

which gives

$$\boxed{N = N_0 \exp\left[-\int_0^z \mu(z')dz'\right]} \qquad (127)$$

where N_0 is the initial number in the birth cohort.

The quantity $N(z)/N_0$ is the fraction of individuals left at age z: demographers call this the "survival function" & label it $\ell(z)$:

$$\boxed{\ell(z) = \exp\left[-\int_0^z \mu(z')dz'\right].} \qquad (128)$$

What is the probability distribution $p(z)$ for dying at a given age z? Formally $p(z)dz$ is the probability of dying between the ages of z & $z+dz$. Clearly

$$\begin{aligned} p(z)dz &= \text{prob(getting to } z) \times \text{prob(dying in } z, z+dz) \\ &= \ell(z).\mu(z)dz \end{aligned}$$

i.e. $p(z) = \mu(z)\ell(z)$ (129)

or $$\boxed{p(z) = \mu(z)\exp\left[-\int_0^z \mu(z')dz'\right].} \quad (130)$$

Exercise: Work out $N(z)$, $\ell(z)$ & $p(z)$ for a population obeying the Gompertz law.

If the probability of death per unit time is a constant, $\mu = \mu_0$, then $N = N_0 e^{-\mu_0 z}$, $\ell = e^{-\mu_0 z}$ & $p = \mu_0 e^{-\mu_0 z}$. This is an example of a 'Poisson process'. The classic example in physics of a Poisson process is the decay of radioactive atoms.

3.3.2 Modelling age structure:

We can introduce a differential distribution $f(z,t)$ to describe the age structure of a pop'n at time t. Formally $f(z,t)\,dz$ is the number of individuals aged between z & $z+dz$ at time t, so that, for example,

$$\int_a^b f(z,t)\,dz$$

is the number of individuals aged between a & b at time t. Clearly if $N(t)$ is the total population at time t, then

$$\boxed{N(t) = \int_0^\omega f(z,t)\,dz} \quad (131)$$

where ω is the oldest an individual can live to ($\omega \approx 115$ years for humans?). Note that we require

$$\boxed{f(\omega, t) = 0} \tag{132}$$

or else there are individuals who will have an age greater than ω after time t.

Also note that $f(0,t)\,dz$ is the number of people aged between 0 & dz at time t. Since these people must have been born in a time $dt = dz$ before time t, the instantaneous rate of births is $f(0,t)\,dz/dt$ where $dt = dz$, i.e. the rate is $f(0,t)$. In terms of the specific birth rate α introduced in §3.1.1, we must have

$$\boxed{f(0,t) = \alpha(t)\, N(t).} \tag{133}$$

It is also useful to introduce the distribution describing fractions of the population,

$$\boxed{\pi(z,t) = \frac{f(z,t)}{N(t)},} \tag{134}$$

so that $\pi(z,t)\,dz$ is the fraction of individuals in the population in the age range $(z, z+dz)$. This distribution satisfies

$$\int_0^{\omega} \pi(z,t)\,dz = 1, \tag{135}$$

so it behaves like a pdf.

Next let the number of deaths per unit of time & per unit age range be denoted $d(z,t)$, i.e. $d(z,t)\,dz$ is the ~~number~~ rate of death of individuals aged z to $z+dz$ at time t. Since:

$$\text{rate of death} = \begin{array}{c}\text{probability of}\\ \text{death per unit}\\ \text{time, per individual}\end{array} \times \begin{array}{c}\text{number}\\ \text{of}\\ \text{individuals}\end{array}$$

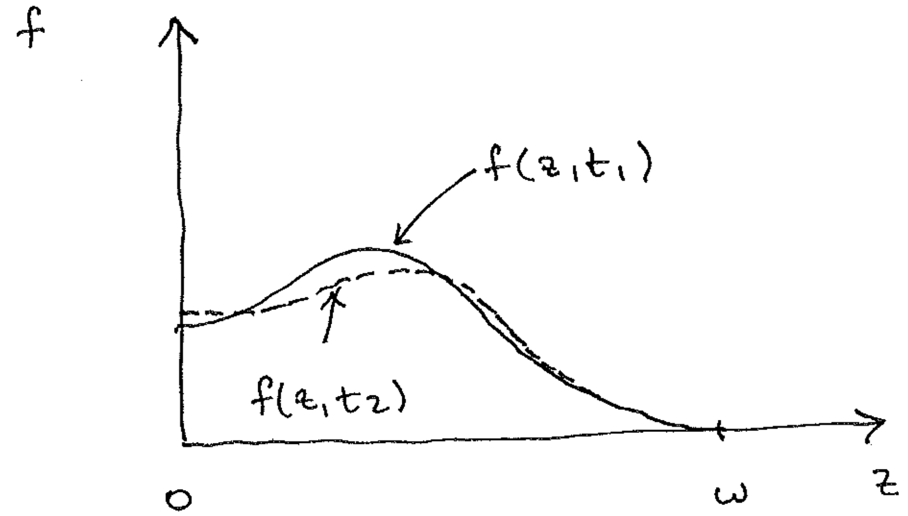
we have:

$$d(z,t)\,dz = \mu(z,t) \times f(z,t)\,dz$$

where $\mu(z,t)$ is the mortality introduced earlier, although here it is allowed to vary with time as well as age. Hence:

$$\boxed{d(z,t) = \mu(z,t)\, f(z,t).} \qquad (136)$$

Next we need to work out the equation governing $f(z,t)$. This will tell us how $f(z,t)$, defined on $0 \leq z \leq \omega$, evolves with time.



two possible distributions f for times t_1 & t_2.

Consider the number of individuals in the age range z to $z+dz$:

number at time $t = f(z,t)\,dz$,

& number at time $t+dt = f(z,t+dt)\,dz$.

In time dt the population ages by $dz = dt$. Hence, neglecting deaths, we can write

$$f(z, t+dt)\,dz = f(z-dt, t)\,dz$$

i.e. individuals aged z to $z+dz$ at time $t+dt$ were aged $z-dt$ to $z-dt+dz$ $(=z)$ at time t. Expanding the RHS as a Taylor series about age z, we have:

$$f(z, t+dt)\,dz \simeq \left[f(z,t) - dt.\frac{\partial f}{\partial z}(z,t) \right] dz.$$

Hence the change in the number of individuals aged z to $z+dz$ due to aging is

$$\Delta N_{\text{aging}} = f(z,t+dt)\,dz - f(z,t)\,dz$$

$$= -dt\,\frac{\partial f}{\partial z}(z,t).\,dz.$$

The change in the number due to deaths follows from the definition of $d(z,t)\,dz$:

$$\Delta N_{\text{deaths}} = -\underbrace{d(z,t)\,dz}_{\text{deaths/unit time}} \times dt.$$

Hence the total change is

$$\Delta N_{\text{tot}} = \Delta N_{\text{aging}} + \Delta N_{\text{deaths}}$$

i.e.

$$f(z,t+dt)dz - f(z,t)dz = -dt\,\frac{\partial f}{\partial z}(z,t)\,dz - d(z,t)\,dz\,dt$$

or $$\frac{f(z,t+dt) - f(z,t)}{dt} = -\frac{\partial f}{\partial z}(z,t) - d(z,t)$$

or in the limit $dt \to 0$,

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial z} + d = 0$$

& then using (136) we have

$$\boxed{\frac{\partial f}{\partial t} + \frac{\partial f}{\partial z} + \mu(z,t) f = 0} \qquad (137)$$

as the governing equation for $f = f(z,t)$. This is really a continuity equation for individuals, & bears close resemblance to the continuity equation obtained for traffic, Equation (51):

$$\frac{\partial \rho}{\partial t} + w \frac{\partial \rho}{\partial x} = 0. \qquad (51)$$

If we rewrite (137) as

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial z} = -\mu f \qquad (138)$$

we see that the LHS takes the form of a kinetic wave equation, with wave speed unity: individuals age at the rate of one year per year.

By comparison with (51) this is an "inhomogenous" wave equation: there is a sink term on the RHS. Individuals disappear when they die, whereas cars are strictly conserved. Note also that ρ & f are essentially the same thing: ρ is the number of cars per unit length, & f is the number of individuals per unit age. In fact (138) is considerably simpler than (51) in that the "wave speed" is constant, which makes the problem linear. There is no analog of shockwaves for populations.

How is (137) related to our earlier equations of growth for the whole population? If we integrate (137) over age from zero to ω, we get

$$\int_0^\omega \frac{\partial f}{\partial t}\,dz + \int_0^\omega \frac{\partial f}{\partial z}\,dz + \int_0^\omega \mu f\,dz = 0$$

$$\text{i.e.}\quad \frac{d}{dt}\left(\int_0^\omega f\,dz\right) + f(\omega,t) - f(0,t) + \int_0^\omega \mu f\,dz = 0$$

$$\text{i.e.}\quad \frac{dN}{dt} + 0 - \alpha(t)N(t) + \int_0^\omega d(z,t)\,dz = 0 \qquad (139)$$

using (131), (132), (133) & (136). The final term is the total rate of death at time t for the population, which must be equal to $\beta(t)N(t)$ where $\beta(t)$ is the specific death rate for the pop'n, i.e. the death rate per individual:

$$\boxed{\beta(t)N(t) = \int_0^\omega d(z,t)\,dz.} \qquad (140)$$

Hence (139) becomes

$$\boxed{\frac{dN}{dt} = [\alpha(t) - \beta(t)]\,N(t),} \qquad (141)$$

which is a time-dependent form of (61), (62). Hence we have arrived at a more rigorous justification of our original growth equation for populations. Note that we have also clarified when exponential growth is appropriate: it will occur whenever $f(0,t)/N(t) = \pi(0,t) = \alpha(t)$ & $\int_0^\omega (\mu f)\,dz / N(t) = \int_0^\omega \mu\pi\,dz = \beta(t)$ are independent of time for a population.

In the following we consider the case that $\mu = \mu(z)$, i.e. the mortality does not depend on time. This assumption is not appropriate for human populations over more than a few decades, since improvements in health care & safety have seen continual reductions in the mortality rate with time. However, the assumption is reasonable for shorter time periods. For many other populations the assumption is quite reasonable. Subject to this assumption our equation is

$$\boxed{\frac{\partial f}{\partial t} + \frac{\partial f}{\partial z} + \mu(z) f = 0} \qquad (142)$$

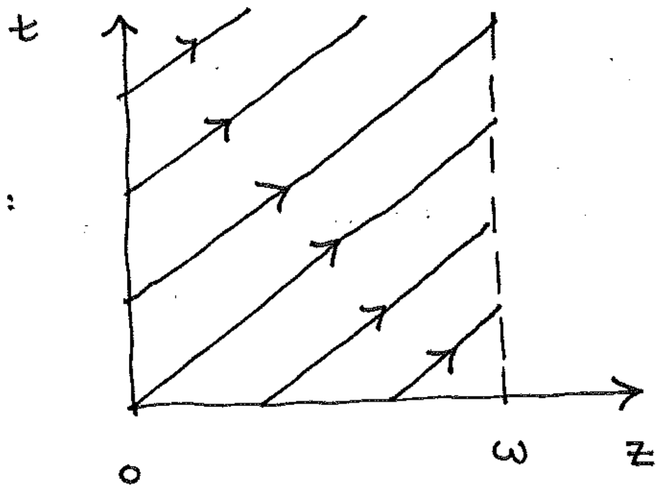
where $f = f(z,t)$.

Before proceeding, note that the lines

$$t - z = \text{const.}$$

in the z-t plane describe the "lifelines" of individuals in the population. If an individual is born on a line with a particular value of const., they stay on this line as they age.

"lifelines":



This observation suggests that we look for a solution to (142) of the form

$$f(z,t) = u(t-z)\, g(z) \tag{143}$$

where u & g are as-yet unspecified functions. The reasoning is that if $t-z=$ const we are dealing with a birth cohort (a group of individuals all born at the same time), & in this case we expect the population of the cohort to vary only with age. Setting $t-z$ constant in (143) implies f varies only with z.

Exercise: Adopting the form (143), show that $g(z)$ satisfies

$$\frac{dg}{dz} = -\mu(z)\, g \tag{144}$$

Exercise: Using (144), show that a formal solution to

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial z} + \mu(z) f = 0$$

for the boundary conditions

$$\begin{aligned} f(0,t) &= B(t) \qquad 0 \leq t \leq \infty \\ f(z,0) &= C(z) \qquad 0 \leq z \leq \omega \end{aligned} \tag{145}$$

(where B & C are specified functions) is

$$f(z,t) = \begin{cases} B(t-z) \exp\left[-\int_0^z \mu(z')dz'\right] & \text{if } t \geq z \\ C(z-t) \exp\left[-\int_{z-t}^z \mu(z')dz'\right] & \text{if } t < z. \end{cases} \tag{146}$$

Note that we require $B(0) = C(0)$, if f is to be continuous at $z = t$.

The formal solution written down in (146) is not all that much use because it requires $B(t)$ to be specified ahead of time. In general $B(t) = f(0,t) = \alpha(t) N(t)$ depends itself on the evolution of the population, e.g. for a human population it depends on how many women there are of child-bearing age, among other things. To develop the theory further we will concentrate on a human population. We see the importance of considering separately the time evolution of the female population.

To this end, let $f(z,t)$ denote the age distribution of the female population, & let $m(z,t)$ denote the rate at which a female of age z produces female children at time t, i.e. $m(z,t)f(z,t)dz$ is the number of girls born to females aged z to $z+dz$ per unit time, at time t. Then

$$B(t) = \int_0^\omega m(z,t)f(z,t)dz \qquad (146)$$

is the total rate of production of baby girls at time t. The fn $m(z,t)$ is the "fertility function".

The male population is treated by introducing the "sex ratio" $S(z,t)$, such that $S(z,t)f(z,t)dz$ is the number of men aged z to $z+dz$. The total population at time t is then

$$N(t) = \int_0^\omega [1+S(z,t)]f(z,t)dz. \qquad (147)$$

If we know (or can estimate $S(z,t)$), then knowledge of the female population implies knowledge of the entire population. It is often assumed that $S = S(z)$, e.g. because there is equality of numbers of men & women at birth, but women live longer on average.

It remains to solve

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial z} + \mu(z) f = 0$$

on $0 \le z \le \omega$ for $t \geq 0$ subject to

$$f(0,t) = B(t) = \int_0^\omega m(z,t)f(z,t)dz$$

& for a specified initial age structure

$$f(z,0) = C(z).$$

Fertility changes with time, due e.g. to sociological reasons. However, we have restricted ourselves to a short-term description by assuming $\mu = \mu(z)$, so it is reasonable to also assume $m = m(z)$. (Note that in any case the age dependence of m will be more important than the time dependence.)

We assume women can only bear children within a narrow range of ages, say $a \leq z \leq b$:

$$m(z) = 0 \quad \text{if } z > b \text{ or } z < a.$$

Then we need to solve (142) subject to

$$f(0,t) = \int_a^b m(z) f(z,t)\,dz \quad (t>0) \qquad (148)$$

$$\& \quad f(z,0) = c(z). \qquad (0 \leq z \leq \omega) \qquad (149)$$

A general solution to this problem is outside our scope. Instead we will look for a solution of the form

$$\boxed{f(z,t) = c(z)e^{rt}} \qquad (150)$$

where r is a constant. Note that this solution automatically satisfies (149). Why would we be interested in such a solution? Note that in this case, assuming equality of the sexes,

$$N(t) = 2\int_0^\omega f(z,t)\,dz = e^{rt}\, 2\int_0^\omega c(z)\,dz$$

$$= e^{rt}\, 2\int_0^\omega f(z,0)\,dz$$

$$= N(0)\,e^{rt}, \qquad (151)$$

i.e. this case corresponds to Malthusian growth. (Of course we also know that exponential growth must be wrong for large times, but our restrictions $\mu = \mu(z)$ & $m = m(z)$ limit our description to short times anyway.)

Note that for the case (150) we have

$$\pi(z,t) = \frac{2f(z,t)}{N(t)} = \frac{2C(z)e^{rt}}{N(0)e^{rt}} = \frac{2C(z)}{N(0)}, \qquad (152)$$

so $\pi = \pi(z)$. Hence the fraction of the population in a given age range does not change with time (although the population in any age range does, it grows exponentially). Demographers say that a population with this property has a "stable age structure."

For (150) we have

$$\frac{\partial f}{\partial t} = rC(z)e^{rt}, \qquad \frac{\partial f}{\partial z} = C'(z)e^{rt}$$

& hence the continuity equation (142) becomes

$$rCe^{rt} + C'e^{rt} + \mu Ce^{rt} = 0$$

$$\text{or} \quad \frac{dC}{dz} + [\mu(z)+r]C = 0. \qquad (153)$$

The B.C. (148) becomes

$$f(0,t) = C(0)e^{rt} = \left(\int_a^b m(z)C(z)\,dz\right)e^{rt}$$

$$\text{i.e.} \quad C(0) = \int_a^b m(z)C(z)\,dz. \qquad (154)$$

The solution to (153) is obtained by integrating:

$$\frac{1}{C}\frac{dC}{dz} = -[\mu(z)+r]$$

$$\ln C(z) = -\int_0^z [\mu(z')+r]\,dz' + \text{const.}$$

$$= -rz - \int_0^z \mu(z')\,dz' + \text{const.}$$

$$\text{or} \quad \boxed{C(z) = C(0)\,e^{-rz}\,e^{-\int_0^z \mu(z')\,dz'}} \qquad (155)$$

or

$$c(z) = c(0)\,\ell(z)\,e^{-rz}, \qquad (156)$$

in terms of the survival function $\ell(z) = \exp\left[-\int_0^z \mu(z')dz'\right]$.

Our solution is then

$$\boxed{f(z,t) = c(0)\,\ell(z)\,e^{r(t-z)}.} \qquad (157)$$

- For any fixed age group, the population grows exponentially.
- The age distribution has a roughly exponential decay with z (a little faster than exponential, if the Gompertz law applies)

Finally, substituting (156) into the BC (154) gives

$$\boxed{\int_a^b m(z)\,\ell(z)\,e^{-rz}dz = 1.} \qquad (158)$$

If a, b, $m(z)$ & $\ell(z)$ are specified, this equation may be solved for r, the net specific growth rate of the population. In this way the model 'predicts' the growth rate of a population with a stable age structure, based on information about fertility & mortality.

This seems a good place to conclude our discussion of modelling populations. We have come full circle, returning to the original model of Malthus for natural growth of a population. However, we now understand the conditions needed for Malthusian growth, & can predict the growth rate based on more fundamental considerations.

PHYS220 Scientific Modelling 2001
Assignment 2 – due Monday May 21

3 1. For a certain 'fractal' curve, it is found that the measured length l of the curve is proportional to $\delta^{-\alpha}$, where δ is the length of the ruler used in measuring the length, and $\alpha \approx 0.3$. Given that l has dimensions L and $\delta^{-\alpha}$ has dimensions $L^{-\alpha}$, how can this be?

2. Consider a sphere moving through air.

   5 (a) For small spheres at low speeds, viscous forces dominate over inertial forces. In this case the drag force F_D on the sphere can be assumed to depend on the viscosity μ, the radius r of the sphere and the velocity v of the sphere, and *does not* depend on the air density. Using dimensional analysis, derive an expression for F_D.

   5 (b) For larger spheres at higher speeds, the viscosity is not important, and the force F_D depends on the air density ρ, the radius r of the sphere and the velocity v of the sphere. Using dimensional analysis, derive an expression for F_D.

   The expression derived in (a) is appropriate to describe the force which keeps the water droplets in clouds suspended in the air, and the expression derived in (b) describes air resistance on macroscopic falling objects, e.g. sky divers.

4 3. Give an order of magnitude estimate of the total mass of air in the Earth's atmosphere.

4. The NSW RTA recommends that drivers maintain a time interval of at least $\Delta t = 3\,\mathrm{s}$ between themselves and the car ahead. We take this to mean that there is at least three seconds between when the rear of the car ahead of a driver passes a point on the road and when the front of the driver's own car passes the same point.

   Consider an equilibrium situation involving 'RTA drivers' (i.e. drivers who maintain a gap of exactly three seconds whenever possible) all moving with the same speed.

   4 (a) Derive a velocity-density relationship for these drivers, taking into account the existence of a speed limit on the road. Your expression for the velocity v should depend only on the speed limit v_{lim}, the density ρ, the interval Δt and the length l of the cars.

   2 (b) At what density is the velocity zero? Is this reasonable?

   2 (c) Write down the flux-density relationship for the RTA drivers.

   4 (d) Sketch the velocity v and flux $\mathcal{F}$ as functions of ρ.

   1 (e) For $v_{\text{lim}} = 60\,\mathrm{km\,h^{-1}}$ and $l = 4\,\mathrm{m}$, how does this model compare with the 'follow-the-leader' model presented in the lectures? Specifically, is the maximum flux better or worse?

SOLUTION TO PHYS220 ASSIGNMENT 2 :

1. Dimensional analysis can be applied to equations, & the statement that ℓ is proportional to $\delta^{-\alpha}$ is not an equation. To form an equation we write $\ell = k\delta^{-\alpha}$ where k is the constant of proportionality. Both sides of this equation must have the same dimension, which implies $[k] = L^{1+\alpha}$, i.e. the constant of proportionality has dimensions of length to the power $1+\alpha$.

2. (a). We have the following quantities & dimensions:

F_D	μ	r	v
MLT^{-2}	$ML^{-1}T^{-1}$	L	LT^{-1}

so there are four variables & three dimensions & hence the Pi theorem tells us there is only one non-dimensional variable, which must be constant. Hence we can write

$$F_D = \text{const.}\, \mu^{\alpha} r^{\beta} v^{\gamma}$$

which implies the relation among dimensions

$$MLT^{-2} = (ML^{-1}T^{-1})^{\alpha} L^{\beta} (LT^{-1})^{\gamma}$$

& because the dimensions on each side must match we have

$$M: \quad 1 = \alpha$$

$$L: \quad 1 = -\alpha + \beta + \gamma$$

$$T: \quad -2 = -\alpha - \gamma.$$

The solution to these equations is $\alpha = \beta = \gamma = 1$, & hence

$$\boxed{F_D = \text{const.}\, \mu r v.}$$

The exact relationship is $F_D = 6\pi\mu r v$, which is known as Stokes' law.

(b). In this case we have

$$\begin{array}{cccc} F_D & \rho & r & v \\ MLT^{-2} & ML^{-3} & L & LT^{-1} \end{array}$$

which is again 4 variables in 3 dimensions. Hence there is only one non-dimensional variable, which must be constant, so we can write

$$F_D = \text{const.}\, \rho^{\alpha} r^{\beta} v^{\gamma}$$

or $$MLT^{-2} = (ML^{-3})^{\alpha} L^{\beta} (LT^{-1})^{\gamma}$$

in terms of dimensions. Requiring the same dimensions on each side we have

$$M: \quad 1 = \alpha$$
$$L: \quad 1 = -3\alpha + \gamma + \beta$$
$$T: \quad -2 = -\gamma$$

which have solution $\alpha = 1$, $\beta = 2$, $\gamma = 2$, & hence

$$\boxed{F_D = \text{const.}\, \rho r^2 v^2.}$$

The exact r'ship is usually written $F_D = \frac{1}{2} C_D \pi r^2 \rho v^2$ where C_D is the drag coefficient (for air $C_D \approx 0.45$).

3. The atmosphere is a thin shell of thickness Δr around the Earth, which has radius r_E. Hence the mass of air is approximately

$$M = 4\pi r_E^2 \Delta r.\, \rho_0$$

where ρ_0 is the air density at ground level.

choosing $\Delta r = 10\text{km}$, the other numbers are $r_E = 6400\text{km}$, $\rho_0 \approx 1.29\ \text{kg m}^{-3}$, & hence

$$M = 4\pi\ (6.4\times10^6)^2 . 10^4 . 1.29\ \text{kg}$$
$$\approx 7\times10^{18}\ \text{kg}.$$

[This is ok provided our estimate for Δr is OK. Recall the atmospheric model from the last assignment:

$$\frac{dp}{dz} = -\rho g$$

Assuming an ideal gas $p = \frac{\rho}{\bar{m}} k_B T$ where $\bar{m}$ is the mean molecular mass. For a constant temperature atmosphere we have

$$\frac{d\rho}{dz} = \frac{-\rho}{z_0}$$

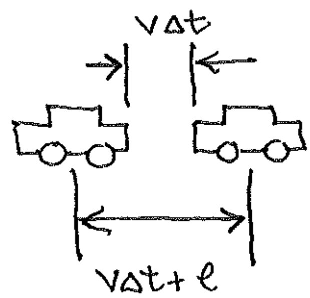
where $z_0 = k_B T/(\bar{m} g)$. The distance z_0 is the density scale height, i.e. the thickness of the atmosphere, i.e. our Δr. For Nitrogen we have $\bar{m} = 28 . 1.67\times10^{-27}\ \text{kg}$ & taking $T = 300\text{K}$ we estimate

$$\Delta r = \frac{1.38\times10^{-23} . 300}{28 . 1.67\times10^{-27} . 9.8}\ \text{m}$$
$$\approx 9000\text{m}$$

which is close enough to 10km for me.]

4. (a). The distance between bumpers is $v\Delta t$.

The distance between the centres of adjacent cars is $v\Delta t + l$, where l is the car length (see diagram). The density is the reciprocal of the distance between car centres, hence



$$\rho = \frac{1}{v\Delta t + l}.$$

Rearranging this we have

$$v = \frac{1}{\Delta t}\left(\frac{1}{\rho} - \ell\right).$$

Note that $v \to \infty$ as $\rho \to 0$. Clearly we need to impose the constraint that cars travel at the speed limit for low densities. Note that $v = v_{lim}$ when $\rho = \rho_c = \dfrac{1}{v_{lim}\Delta t + \ell}$; we require that cars travel at the speed limit for $\rho \le \rho_c$, & hence the velocity-density relationship is

$$\boxed{v = \begin{cases} v_{lim} & \rho \le \rho_c \\ \dfrac{1}{\Delta t}\left(\dfrac{1}{\rho} - \ell\right) & \rho_c \le \rho \le \dfrac{1}{\ell} \end{cases}}$$

where $\rho_c = (v_{lim}\Delta t + \ell)^{-1}$.

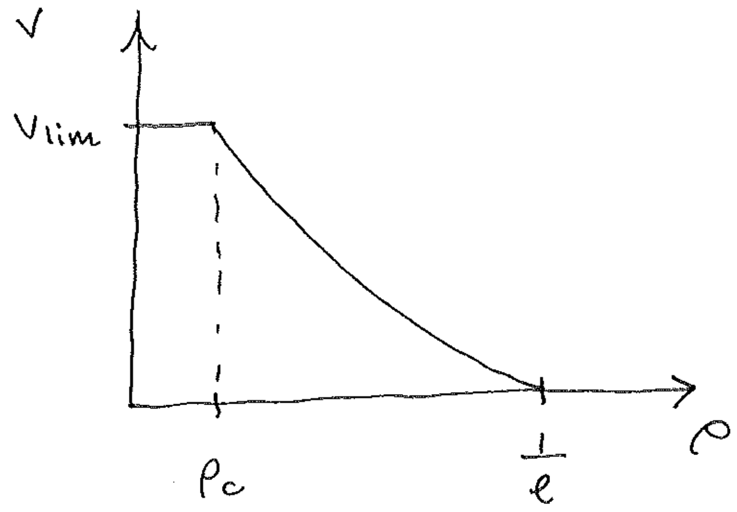
(b). We have $v = 0$ when $\rho = \frac{1}{\ell}$. This corresponds to the cars touching bumpers. As discussed in lectures studies of roads indicate $v = 0$ for $\rho = \rho_{max}$ where $\rho_{max} < \frac{1}{\ell}$. Hence it is not reasonable : in reality the road ceases to move for densities less than the absolute maximum density $1/\ell$.

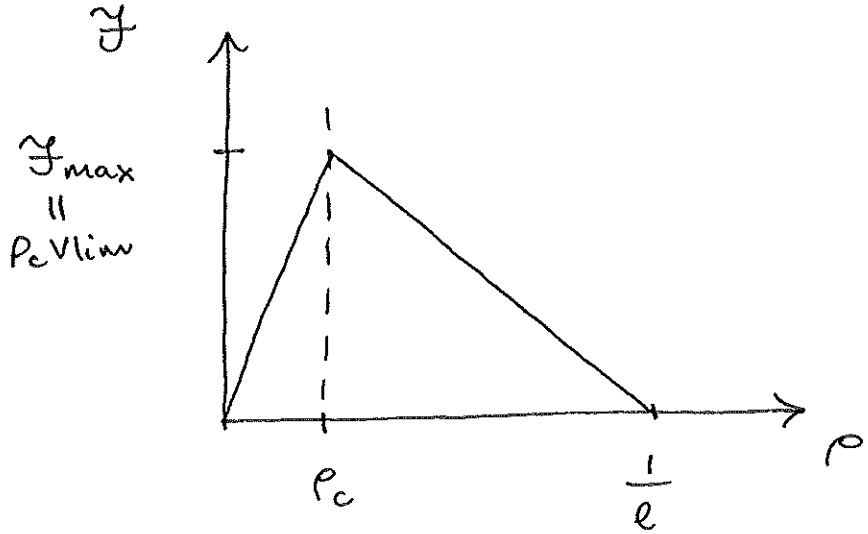
(c). The flux $\mathcal{F}$ is the product of density & velocity, & hence

$$\boxed{\mathcal{F} = \begin{cases} \rho v_{lim} & \rho \le \rho_c \\ \dfrac{1}{\Delta t}(1 - \rho\ell) & \rho_c \le \rho \le \dfrac{1}{\ell}, \end{cases}}$$

which represents a linear increase for $\rho \le \rho_c$ followed by a linear decrease for $\rho_c \le \rho \le \frac{1}{\ell}$.

(d).





(e). The maximum flux is achieved for $\rho = \rho_c$, hence

$$\mathcal{F}_{max} = \rho_c V_{lim} = \frac{V_{lim}}{V_{lim}\Delta t + \ell}.$$

For the given values we have

$$\mathcal{F}_{max} = \frac{6\times 10^4 \text{ m / hr}}{\left(\frac{6\times10^4}{3600}\cdot 3 + 4\right)\text{m}}$$

$$\simeq 1100 \text{ hour}^{-1}$$

It is actually hard to compare this result with the follow-the-leader model, because that model contains a free parameter (a point which should have been stressed more in lectures).

The free parameter is λ (or equivalently a), as described in lectures. If we take the Holland tunnel data & keep the same a & ρ_{max} ($a = 0.35$, $\rho_{max} = 175$ cars/mile ≈ 109 cars/km) then the predicted maximum flux is

$$\mathcal{F}_{max} = \frac{a}{e} \cdot \rho_{max} v_{lim}$$

$$= \frac{0.35}{2.71} \cdot 109 \text{ cars/km} \cdot 60 \text{ km/h}$$

$$\approx 845 \text{ cars/hour}$$

which suggests the RTA model is better. Of course, the comparison is limited by the validity of taking the a value from the Holland tunnel data & applying it with a different v_{lim}: perhaps a depends on v_{lim}.

This model gives a large flux because the velocity at low densities is large, compared with the follow-the-leader model. The RTA rule appears to move cars quite efficiently, when everyone obeys it.

PHYS220 Scientific Modelling 2001
Assignment 3 – due Monday June 4

1. The logistic equation introduced in lectures was

$$\frac{1}{N}\frac{dN}{dt} = a(1 - N/N_*). \tag{1}$$

By making the replacements

$$\frac{dN}{dt} \to \frac{N_{i+1} - N_i}{\Delta t}, \quad N \to N_i$$

discretise the logistic equation and put it into the form of the logistic map,

$$x_{i+1} = r(1 - x_i)x_i, \tag{2}$$

where $x_i = N_i/A$ for some factor A.

4 (a) Identify the factor A and the coefficient r (they depend on N_*, a and Δt).

2 (b) Use the result of (a) to explain why an accurate numerical solution of Equation (1) using the discrete version (2) will not produce limit cycles or chaos. (You may need some results concerning the logistic map from Barry's half of the course.)

2. State whether the following populations are best described by a differential equation or by a (finite-step) difference equation:

2 (a) a small population of mammals,

2 (b) Pacific salmon that spawn once a year, with the adults dying shortly after spawning,

2 (c) bacteria in a Petri dish, with a very large growth rate.

Also, which of the above could (in principle) produce chaotic behaviour?

3. Consider the variation on logistic growth,

$$\frac{1}{N}\frac{dN}{dt} = a\left[1 - (N/N_*)^2\right] \tag{3}$$

(the "crowding" depends on the square of the population).

2 (a) Write down the ODE satisfied by $u = N^2$.

2 (b) Using the result of (a) or otherwise, solve Equation (3) subject to $N = N_0$ at $t = 0$.

2 (c) Sketch the solution found in (b) for $N_* \gg N_0$, and on the same graph show the solution to the logistic equation (1) (for the same values of a, N_* and N_0).

4. The Lotka-Volterra equations may be written

$$\begin{aligned} \frac{dN}{dt} &= (a_1 - b_1 P)N \\ \frac{dP}{dt} &= (-a_2 + b_2 N)P. \end{aligned} \tag{4}$$

Show that trajectories (in the phase space) of solutions to Equation (4) satisfy

2

$$a_2 \ln N + a_1 \ln P - b_2 N - b_1 P + C = 0,$$

where C is a constant which labels a solution for a given $N(0)$ and $P(0)$.

20

2. (a). In this case generations overlap, & so a differential equation provides a better description. The fact that the numbers are small is irrelevant (for small populations, the differential equation may be interpreted as describing the evolution of the expected number in the population). In this case there can be no chaos.

(b). In this case successive generations are essentially distinct, & so the finite-step difference equation provides a better description. In principle chaos is possible.

(c). In this case generations overlap, so the differential equation is appropriate. Even for a large growth rate there will be no chaos.

3. (a). Setting $u = N^2$, note that

$$\frac{du}{dt} = 2N\frac{dN}{dt} = 2a\left[1-\left(\frac{N}{N_*}\right)^2\right]N^2,$$

using (3). Rewriting the RHS in terms of u,

$$\boxed{\frac{du}{dt} = 2a\left(1-\frac{u}{u_*}\right)u} \qquad (1)$$

where $u_* \equiv N_*^2$. Hence we see that u obeys the logistic equation, with a maximum specific growth rate $2a$ & a carrying capacity u_*.

(b). Based on the solution to the logistic equation given in lectures the solution to (1) must be

$$u = \frac{u_*}{1+\left(\frac{u_*}{u_0}-1\right)e^{-2at}}, \qquad (2)$$

SOLUTION TO PHYS220 ASSIGNMENT 3:

1. (a). Making the suggested replacements gives

$$\frac{N_{i+1}-N_i}{\Delta t} = a\left(1-\frac{N_i}{N_*}\right)N_i$$

or $$N_{i+1} = N_i\left[1 + a\Delta t - \frac{N_i}{N_*/(a\Delta t)}\right]$$

$$= (1+a\Delta t)N_i\left[1-\frac{N_i}{N_*\left(1+\frac{1}{a\Delta t}\right)}\right].$$

Clearly we can make the choices

$$\boxed{r \equiv 1+a\Delta t, \qquad A \equiv N_*\left(1+\frac{1}{a\Delta t}\right)}$$

& then we have

$$x_{i+1} = r(1-x_i)x_i$$

where $x_i \equiv N_i/A$.

(b). For an accurate numerical solution the timestep must be small: hence we expect $a\Delta t \ll 1$, & the quantity r will be just a bit larger than unity. However, limit cycles and chaos are obtained for large values of r in the logistic map, as Barry showed. In particular, $r > 3$ is needed for limit cycles and $r \gtrsim 3.57$ for chaos. Hence there will be no limit cycles / chaos. [Note that the discretisation used in this example, so-called Euler timestepping, should not be used in practise in any circumstance: there are better ways to discretise ODEs.]

where $u_0 = u(t=0)$. Rewriting this in terms of N we have

$$\boxed{N^2 = \frac{N_*^2}{1 + \left[\left(\frac{N_*}{N_0}\right)^2 - 1\right] e^{-2at}}.}$$

(c). Setting $x = N/N_*$, $\tau = at$ & $\varepsilon = N_0/N_*$, we need to plot

$$x_1 = \frac{\varepsilon}{\varepsilon + (1-\varepsilon)e^{-\tau}}$$

(the logistic map) &

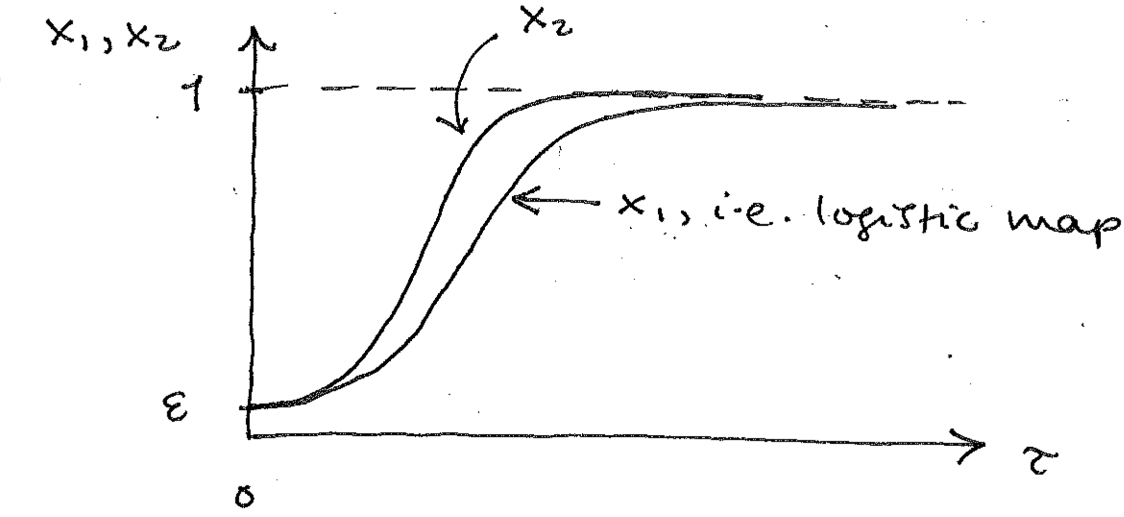
$$x_2 = \frac{\varepsilon}{[\varepsilon^2 + (1-\varepsilon^2)e^{-2\tau}]^{\frac{1}{2}}}.$$

It is easy to show that for small τ both x_1 & x_2 increase like εe^{τ}. For large τ note that

$$x_1 = \frac{\varepsilon}{\varepsilon\left[1 + \left(\frac{1}{\varepsilon} - 1\right)e^{-\tau}\right]} \simeq 1 - \left(\frac{1}{\varepsilon} - 1\right)e^{-\tau}$$

& $$x_2 = \frac{\varepsilon}{\varepsilon\left[1 + \left(\frac{1}{\varepsilon^2} - 1\right)e^{-2\tau}\right]^{\frac{1}{2}}} \simeq 1 - \frac{1}{2}\left(\frac{1}{\varepsilon^2} - 1\right)e^{-2\tau}.$$

Irrespective of the size of ε, for sufficiently large τ, $e^{-2\tau} \ll e^{-\tau}$. Hence for large τ x_2 will be closer to unity, i.e. larger than x_1. Considering an expansion in ε for small τ, it is also possible to show that $x_2 > x_1$ for small τ. Hence the graphs are as shown (otherwise you could just plot them using the computer):



4. Taking the ratio of the equations (4) we have:

$$\frac{dN}{dP} = \frac{(a_1 - b_1 P)N}{(-a_2 + b_2 N)P}$$

which is separable:

$$\int \left(-\frac{a_2}{N} + b_2\right) dN = \int \left(\frac{a_1}{P} - b_1\right) dP .$$

Integrating gives

$$-a_2 \ln N + b_2 N = a_1 \ln P - b_1 P + \text{const}$$

or $\quad a_2 \ln N + a_1 \ln P - b_2 N - b_1 P + C = 0, \quad (3)$

where C is the constant of integration. This constant is determined by the requirement

$$a_2 \ln N(0) + a_1 \ln P(0) - b_2 N(0) - b_1 P(0) + C = 0,$$

& hence (3) describes trajectories in the phase space, with different C's for different trajectories (defined by the initial conditions).

PART B:
Solve any two of the following three problems

4. (a) (5 marks)
A medical researcher studying spermatozoa determines that the speed of propagation depends on the length l of the organism, the viscosity μ of the fluid, and the rate of expenditure of energy per unit time and per unit volume $\dot{\epsilon}$ by the organism. (The speed does not depend on the density because on the small scales of relevance, viscous forces dominate over inertial forces.)
/3 Use dimensional analysis to determine how the velocity depends on μ, l and $\dot{\epsilon}$. If a spermatazoan begins with a total energy per unit volume ϵ_0, obtain an approximate
/2 expression for how far the organism can swim.

/4 (b) (5 marks)
Suppose that an alien is attempting to determine the formula describing the gravitational force F between two small point masses m_1 and m_2. The alien has worked out that the force depends on the masses, a dimensional constant she calls G (with dimensions $\mathrm{M}^{-1}\mathrm{L}^3\mathrm{T}^{-2}$), and the distance r between the masses. The alien applies dimensional analysis to determine how F depends on the other quantities. What will she discover? Describe briefly how the alien could perform experiments involving varying only one of the quantities to guess the exact form of the relationship. only π : 1/5 π_1, π_2 : 2/5 all correct : 4/5

/1

(c) (5 marks)
Estimate the number of cells in your body (a typical cell has diameter 50μm).

(d) (5 marks)
3 Recall the Bayesian treatment of a biased coin given in lectures. It was shown that the posterior probability of the bias of a coin (the bias is the probability of a head in a single toss) being H given r heads in n tosses is

$$P(H|r \text{ heads in } n \text{ tosses}) \propto H^r(1-H)^{n-r}f(H).$$

where $f(H)$ describes the prior probability assigned to a bias H. Show that for a uniform prior the most likely value of the bias is $H_* = r/n$. What about for an arbitrary $f(H)$?

2

5 2) Consider the 'follow the leader' equilibrium traffic model introduced in lectures, involving the free parameters v_{lim}, ρ_{max} and λ. The following questions concern this model.

(a) (10 marks)
Work out the time Δt between when the rear of a driver's car passes a point on the road and when the front of the following driver's car passes the same point, *as a function of speed v*. (Your answer should depend on v, ρ_{max}, λ and l, the length of a car.) Explain the behaviour of Δt as $v \to 0$. Plot Δt versus v/v_{lim} using the values of the free parameters given in the lectures for Holland tunnel traffic and assuming $l = 4\,\text{m}$ (you may use a computer for this step if you wish). How does the time Δt compare with that recommended by the RTA? What can you infer about the way drivers drive?

5 2 2 1

(b) (5 marks)
Assuming that the follow-the-leader flux vs. density relationship also holds in time-dependent situations, work out the speed of small amplitude density disturbances in traffic. Sketch this speed, as a function of ρ. Show that this speed is zero when the flux is a maximum.

2 2 1

(c) (5 marks)
Consider traffic waiting at a set of traffic lights, which can be modelled by the initial density profile

$$\rho(x, t = 0) = \begin{cases} \rho_{\text{max}} & \text{if } x \leq 0 \\ 0 & \text{if } x > 0. \end{cases}$$

The lights change to green at time $t = 0$. Show that the information that the lights have changed propagates backwards through the queue of traffic with speed λ, and the same information propagates forwards into the open road with speed v_{lim}.

3 1 2

3. Consider the following modification of the Lotka-Volterra equations:

$$\begin{aligned}\frac{dN}{dt} &= (a_1 - b_1 P)N + r \\ \frac{dP}{dt} &= (-a_2 + b_2 N)P, \end{aligned} \qquad (1)$$

where the symbols have the same meaning as in lectures. The extra term r is a positive constant which describes a rate of migration of prey into the ecosystem.

4/5 if they don't say $N \le 0$ unphys.

one error: 4/5

(a) (5 marks)
Determine the equilibrium point(s) for Equations (1). incomplete: 2/5

(b) (5 marks)
Non-dimensionalise the ODEs (1). A suitable non-dimensional version of r is $\delta = rb_2/(a_1 a_2)$.

(c) (5 marks)
4 → Numerically solve the non-dimensional ODEs obtained in (b) for a choice of initial conditions and for $\delta = 0.2$. Briefly explain the method used and show the results. (A suitable procedure would be to modify modify one of the MATLAB scripts provided on the web.) Briefly describe the behaviour of the system, and comment on what it says about the suitability of the Lotka-Volterra equations to describe real predator-prey ecosystems. 1

(d) (5 marks)
Confirm the result found in (c) using a stability analysis about an equilibrium point, assuming $\delta \ll 1$.

Must have correct procedure else 0/5

1.

(a) Quantities: $\quad v \quad \mu \quad \ell \quad \dot{\varepsilon}$

dimensions: $\quad LT^{-1} \quad ML^{-1}T^{-1} \quad L \quad \dfrac{MLT^{-2}L}{TL^3} = ML^{-1}T^{-3}$

so 4 quantities in 3 dimensions $\Rightarrow$ only one non-dim variable, & it is constant:

$$\pi = \text{const.}$$

$$v = \mu^{\alpha} \ell^{\beta} (\dot{\varepsilon})^{\gamma} \cdot \text{const}$$

dimensions:

$$LT^{-1} = (ML^{-1}T^{-1})^{\alpha} L^{\beta} (ML^{-1}T^{-3})^{\gamma}$$

$M: \quad 0 = \alpha + \gamma \quad \Rightarrow \quad \alpha = -\gamma$

$L: \quad 1 = -\alpha + \beta - \gamma \Rightarrow \beta = 1 + \alpha + \gamma = 1$

$T: \quad -1 = -\alpha - 3\gamma \quad \Rightarrow \quad 2\gamma = 1 \Rightarrow \gamma = \frac{1}{2}$

$\Rightarrow \alpha = -\frac{1}{2}$

so $\quad \boxed{v = \text{const.} \dfrac{\ell (\dot{\varepsilon})^{\frac{1}{2}}}{\mu^{\frac{1}{2}}}}$

total energy per unit volume ε_0

suppose the organism loses energy at a constant rate

$$(\dot{\varepsilon})_{loss} = -k = -(\dot{\varepsilon})_{\text{expended}}$$

$$\varepsilon = \varepsilon_0 - kt = \varepsilon_0 - \dot{\varepsilon}t$$

so it takes time $T = \dfrac{\varepsilon_0}{\dot{\varepsilon}}$ to lose all energy.

$$v = \frac{ds}{dt} = \text{const.}\ \frac{\ell(k)^{\frac{1}{2}}}{\mu^{\frac{1}{2}}} \qquad \cancel{v = \text{const.}\ \ell(\dot{\varepsilon})}$$

so $$s = \text{const.}\ \frac{\ell(k)^{\frac{1}{2}}}{\mu^{\frac{1}{2}}} \cdot T$$

is total distance covered

$$s = \text{const.}\ \frac{\ell(k)^{\frac{1}{2}}}{\mu^{\frac{1}{2}}} \cdot \frac{\varepsilon_0}{k} = \frac{\ell \cdot \dot{\varepsilon}^{\frac{1}{2}} \varepsilon_0}{\mu^{\frac{1}{2}}\, \dot{\varepsilon}}$$

$$= \text{const.}\ \frac{\ell\, \varepsilon_0}{\mu^{\frac{1}{2}} (\dot{\varepsilon})^{\frac{1}{2}}} \checkmark$$

$$(\dot{\varepsilon} = \text{const.})$$

suppose it loses energy at a rate prop. to tere energy:

$$\cancel{\langle \dot{\varepsilon} \rangle = \int_0^\infty \dot{\varepsilon} . \varepsilon_0 e^{-k}}$$

$$\dot{\varepsilon} = -k\varepsilon$$

$$\varepsilon = \varepsilon_0 e^{-kt}$$

$$\dot{\varepsilon} = \varepsilon_0(-k)\, e^{-kt}$$

~~so~~

$$\frac{ds}{dt} = \text{const.}\ \frac{\ell\, \varepsilon_0^{\frac{1}{2}} k^{\frac{1}{2}}\, e^{-kt/2}}{\mu^{\frac{1}{2}}}$$

$$s = \text{const.}\ \frac{\ell\, \varepsilon_0^{\frac{1}{2}} k^{\frac{1}{2}}}{\mu^{\frac{1}{2}}} \int_0^\infty e^{-kt/2}\, dt$$

$$= \text{const.}\ \frac{\ell\, \varepsilon_0^{\frac{1}{2}} k^{\frac{1}{2}}}{\mu^{\frac{1}{2}}} \left(-\frac{2}{k}\right) \left[e^{-kt/2}\right]_0^\infty$$

$$= 2 \cdot \text{const.}\ \frac{\ell\, \varepsilon_0^{\frac{1}{2}}}{\mu^{\frac{1}{2}} k^{\frac{1}{2}}} = \text{const.}\ \frac{\ell\, \varepsilon_0^{\frac{1}{2}}}{\mu^{\frac{1}{2}}}\, \tau^{\frac{1}{2}}$$ where τ is timescale

1(b)

G	m_1	m_2	r	F
$L^3T^{-2}M^{-1}$	M	M	L	MLT^{-2}

5 in 3 $\Rightarrow$ two non-dim. variables

$$\left[\frac{F}{G}\right] = M^2L^{-2}$$

$$f\left(\frac{2m_1}{m_2}\right) = 2f\left(\frac{m_1}{m_2}\right)$$

$$f(2x) = 2f(x)$$

$$f\left(\frac{m_1}{m_2}\right)$$

$$\cancel{f'(2x)\cdot 2 = 2f'(x)}$$

$$\cancel{f'(x) = f'(2x)}$$

$$\pi_1 = \frac{m_1}{m_2}$$

$$\pi_2 = \frac{F}{G}\cdot\frac{r^2}{m_1^2}$$

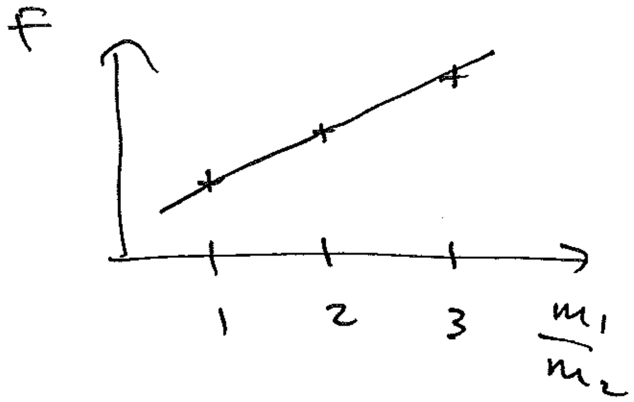
$$\pi_2 = f(\pi_1)$$

$$\frac{Fr^2}{Gm_1^2} = f\left(\frac{m_1}{m_2}\right) \Rightarrow F = \frac{Gf\left(\frac{m_1}{m_2}\right)m_1^2}{r^2}$$

$G \quad m_1 \quad m_2 \quad r$

$x = \frac{m_1}{m_2}$

double m_2



(c).

cell volume $\delta V = \frac{4}{3}\pi\left(\frac{d}{2}\right)^3$

volume of body: $V = \frac{m}{\rho}$

so number $N = \frac{V}{\delta V} = \frac{m \cdot 24}{4\pi d^3 \rho}$

$\approx \frac{\cancel{4}\cdot 80\,\text{kg}\cdot 6}{\frac{\cancel{4}}{3}\pi\cdot(50\times10^{-6})^3\cdot 1}$

$\approx 3\times10^{10}$ ~~4×10^{12}~~

~~30 billion!~~

cell volume $\delta V = \frac{4}{3}\pi\left(\frac{d}{2}\right)^3$

or $\delta V = d^3$ will do

volume of body $V = \frac{m}{\rho}$

So $N = \frac{V}{\delta V} = \frac{m}{\rho}\cdot\frac{1}{d^3}$

$= \frac{80}{10^3(50\times10^{-6})^3}$

$\approx 6.4\times10^{11} \approx 10^{12}$

or with spheres:

$N = \frac{m}{\rho}\cdot\frac{1}{\frac{4}{3}\pi\left(\frac{d}{2}\right)^3} = \frac{6.m}{\pi\rho d^3} = \frac{6}{\pi}\times$ above

so $\approx 1.2\times10^{12}$

(d)

biased dice:

toss 100 times, observe 25 6's.
Give an estimate for the probability of getting a 6.

$$\text{prob}(H|D) \propto H^r (1-H)^{n-r}$$

$$H = \text{prob. of } 6$$

plot posterior prob

$$\text{prob}(\underset{\text{bias } H}{\cancel{25\ 6's}} \mid 25 \text{ heads in } 100)$$

$$= H^{25}(1-H)^{75} = P$$

$$P = CH^r (1-H)^{n-r}$$

$$y = x^a$$

$$\frac{dP}{dH} = rH^{r-1}(1-H)^{n-r}$$

$$\frac{dy}{dx} = a x^{a-1}$$

$$+ H^r (\ )^{n-r-1} (n-r)\cdot(-1) = 0$$

$$rH^{r-1}(1-H)^{n-r} - H^r(1-H)^{n-r-1}(n-r) = 0$$

$$rH^{r-1}(1-H) - (n-r)H^r = 0$$

$$r \cdot \frac{1-H}{H} - (n-r) = 0$$

$$\frac{1}{H} - 1 = \frac{n-r}{r}$$

$$\frac{1}{H} = \frac{n-r}{r} + 1$$

$$= \frac{n-r+r}{r} = \frac{n}{r}$$

$$H = \frac{r}{n}$$

If not uniform,

$$P = c\, H^r (1-H)^{n-r} \cdot f(H)$$

$$P' = c\, r H^{r-1} (1-H)^{n-r} f$$

$$+ c H^r (n-r)(1-H)^{n-r-1} \cdot (-1) f$$

$$+ c H^r (1-H)^{n-r} f'$$

$P' = 0 \Rightarrow \;/H^{r-1}$

$$c\, r (1-H)^{n-r} f - cH(n-r)(1-H)^{n-r-1} f$$

$$+ cH(1-H)^{n-r} f' = 0$$

$/(1-H)^{n-r-1}$ solution to:

$$\cancel{c}\, r(1-H) f - \cancel{c} H(n-r) f + \cancel{c} H(1-H) f' = 0$$

Solution to this equation gives max. posterior prob. in general case.

1. (a). The follow-the-leader model has the speed-density r'ship: 1.

$$v = \begin{cases} v_{lim}, & \rho \le \rho_c \\ \lambda \ln \rho_{max}/\rho, & \rho_c \le \rho \le \rho_{max} \\ 0 & \rho > \rho_{max} \end{cases} \quad (*)$$

where $\rho_c \equiv \rho_{max} e^{-v_{lim}/\lambda}$.

The time Δt between cars is defined by

$\frac{1}{\rho} = v\Delta t + l$ where v is the speed & l is the length of a car.

so $\Delta t = \frac{1}{v}\left(\frac{1}{\rho} - l\right)$

We can rewrite (*) for $\rho_c < \rho < \rho_{max}$:

$$\frac{\rho}{\rho_{max}} = e^{-v/\lambda}$$

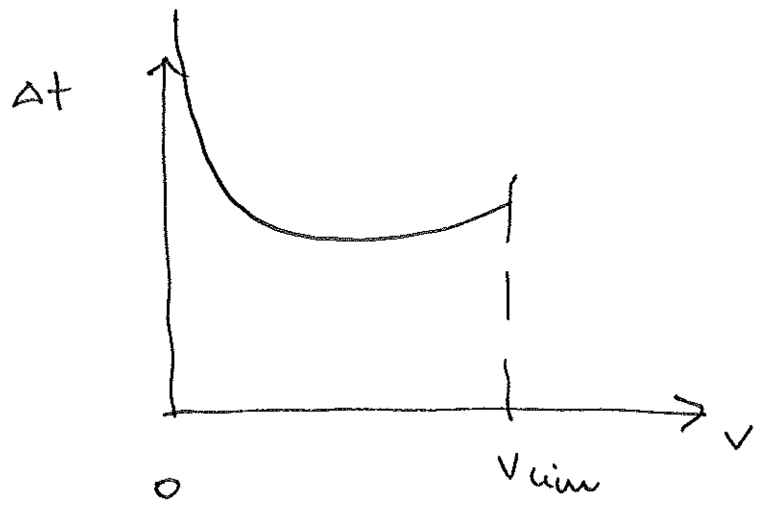
$$\rho = \rho_{max} e^{-v/\lambda}$$

so $$\boxed{\Delta t = \frac{1}{v}\left(\frac{1}{\rho_{max}} e^{+v/\lambda} - l\right)}$$

is the time between cars, as a function of v. This holds for $0 \le v \le v_{lim}$

~~When~~ As $v \to 0$, $\Delta t \to \infty$

As $v \to v_{lim}$, $\Delta t \to \frac{1}{v_{lim}}\left(\frac{1}{\rho_{max}} e^{+v_{lim}/\lambda} - l\right)$



For $v_{lim} = 55\text{mph} \approx 88\text{kph} \approx 24.4\,\text{ms}^{-1}$

$\rho_{max} = 175$ cars/mile ≈ 109 cars/km $\approx 0.11\,\text{m}^{-1}$

(Holland tunnel data)

$\lambda = 0.35\,v_{lim}$ ⇒ $a = \dfrac{\lambda}{v_{lim}} = 0.35$

set $u = v/v_{lim}$ so graph runs from 0 to 1

so: $$\Delta t = \frac{1}{v_{lim}} \frac{1}{u}\left(\frac{1}{\rho_{max}} e^{u v_{lim}/\lambda} - l \right)$$

set $l = 4\text{m}$

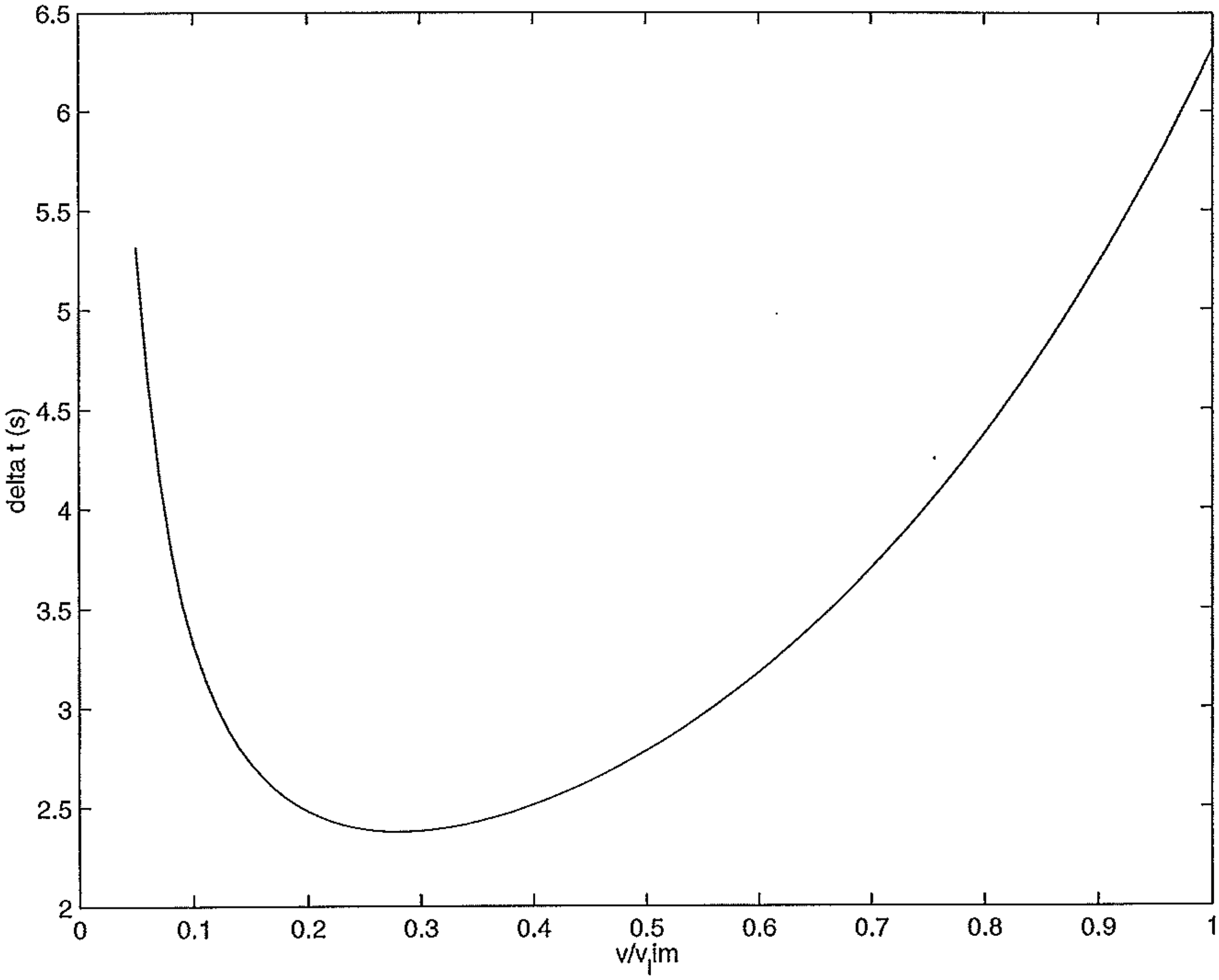
$$\Delta t = \frac{1}{v_{lim}} \frac{1}{u}\left(\frac{1}{\rho_{max}} e^{u/a} - l \right)$$

The figure over shows this graph, for the parameters above.

We see that $\Delta t \to \infty$ as $v \to 0$. This is because there is a finite ~~distance~~ separation between cars when $v = 0$ in this model: $\frac{1}{\rho} = v\Delta t + l$ implies that if Δt is finite for $v = 0$ then $\frac{1}{\rho} = l$, i.e. there is no separation.

3.

We see that $\Delta t < 3s$ for a range of speeds: hence these drivers are following more closely than recommended by the RTA, & their driving may be considered risky.



(b). For this model

$$\mathcal{F} = \rho . V = \begin{cases} V_{lim}\, \rho & \rho \leq \rho_c \\ \lambda \rho \ln \dfrac{\rho_{max}}{\rho} & \rho_c \leq \rho \leq \rho_{max} \\ 0 & \rho \geqslant \rho_{max} \end{cases}$$

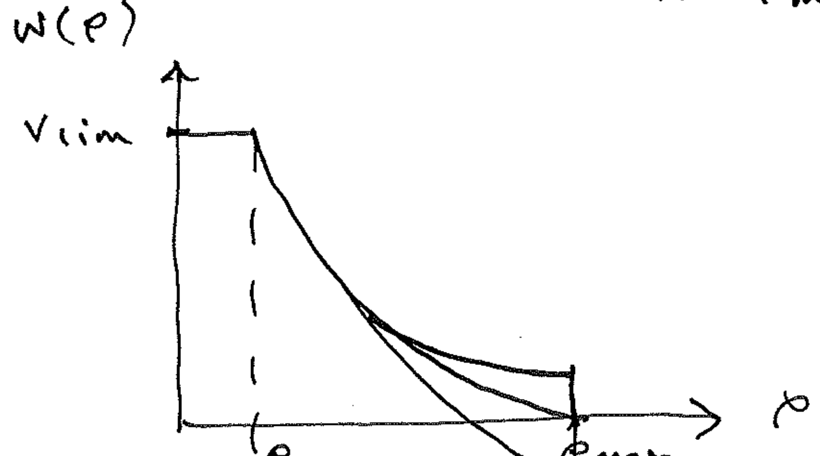
Small amplitude disturbances in density have the speed $W(\rho) = \dfrac{d\mathcal{F}}{d\rho}$, as shown in lectures

$$\text{So: } W = \begin{cases} V_{lim} & \rho \leq \rho_c \\ \lambda \ln \dfrac{\rho_{max}}{\rho} + \lambda \cancel{\rho} . \dfrac{\cancel{\rho_{max}}}{\cancel{\rho_{max}}} . \left(- \dfrac{\cancel{\rho_{max}}}{\cancel{\rho^2}} \right) & \rho_c \leq \rho \leq \rho_{max} \\ 0 & \rho \geqslant \rho_{max} \end{cases}$$

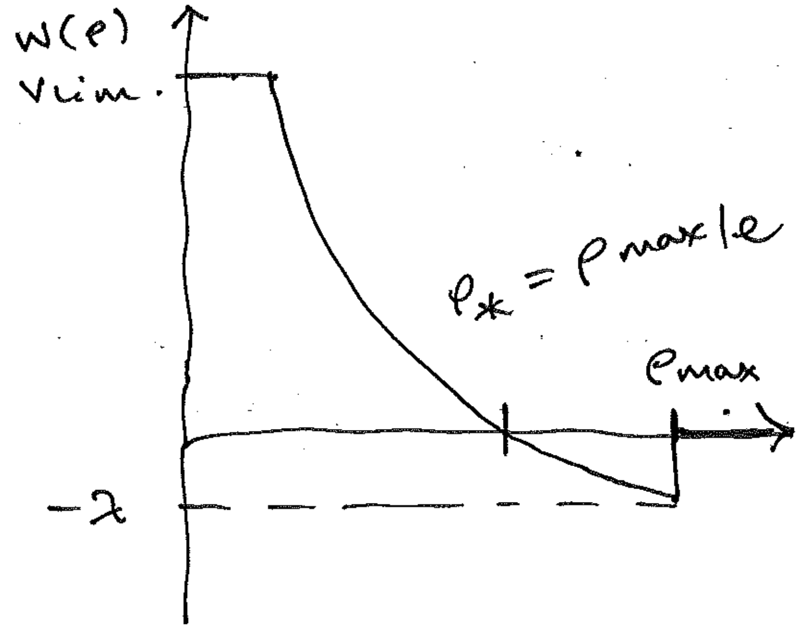
$$\text{i.e. } W = \begin{cases} V_{lim} & \rho \leq \rho_c \\ \lambda \left(\ln \dfrac{\rho_{max}}{\rho} - 1 \right) & \rho_c \leq \rho \leq \rho_{max} \\ 0 & \rho \geqslant \rho_{max} \end{cases}$$

Note that $\quad \lambda \ln \dfrac{\rho_{max}}{\rho} - 1 = \lambda \ln \rho_{max} - \lambda \ln \rho - 1$

$$= \lambda \ln \rho_{max} - 1 - \lambda \ln \rho$$



Woops:



$w = 0$ when $\ln \frac{\rho_{max}}{\rho} = 1$

$$\frac{\rho_{max}}{\rho} = e$$

$$\rho = \frac{\rho_{max}}{e}$$

From lectures this is ρ_*, the ~~flux~~ density at which the flux is maximum.

(c). The information propagates with the speed of ripples in density, i.e. at the speed $w(\rho)$. For $x<0$ $\rho = \rho_{max}$, so the speed is $-\lambda$, as shown by the diagram above (Note that this requires the density is actually $\rho_{max} - \varepsilon$ where ε is a small number, but this is a detail).

The info propagates to the right at $w(\rho=0) = v_{lim}$.

3. Migration:

$$\frac{dN}{dt} = (a_1 - b_1 P)N + r$$ rate of migration (const.) $(r > 0)$

$$\frac{dP}{dt} = (-a_2 + b_2 N)P$$

(a) Equilibria :

$$(a_1 - b_1 P)N + r = 0 \quad ①$$

$$(-a_2 + b_2 N)P = 0 \quad ②$$

② $\Rightarrow$ $P = 0$ or $N = \frac{a_2}{b_2}$

$\downarrow$ in ① :

$N = -\frac{r}{a_1} < 0$

so not possible if $r > 0$

$\downarrow$

$$\frac{a_1 a_2}{b_2} + r = \frac{b_1 a_2}{b_2} P$$

$$P = \frac{b_2 r + a_1 a_2}{b_1 a_2}$$

so there is a new equilibrium point,

$$P = \frac{b_2 r + a_1 a_2}{b_1 a_2}$$

$$= \frac{b_2}{b_1 a_2} r + \frac{a_1}{b_1} > \frac{a_1}{b_1}$$

so prey pop'n goes up a bit.

$$P = \frac{1}{b_1}\left(\frac{b_2}{a_2} r + a_1\right)$$

Numerically:

(b) $\bar{N} = N/(a_2/b_2)$

$\bar{P} = P/(a_1/b_1)$

$\bar{t} = a_1 t$

$\bar{a} = \frac{a_2}{a_1}$

$$\frac{a_2 a_1}{b_2} \frac{d\bar{N}}{d\bar{t}} = \left(a_1 - \cancel{b_1} \frac{a_1}{\cancel{b_1}} \bar{P}\right) \frac{a_2}{b_2} \bar{N} + r$$

$$\frac{d\bar{N}}{d\bar{t}} = (1 - \bar{P}) \bar{N} + r.\frac{b_2}{a_1 a_2}$$

$$\boxed{\frac{d\bar{N}}{d\bar{t}} = (1 - \bar{P}) \bar{N} + \delta}$$

$$\delta = \frac{r b_2}{a_1 a_2}$$

$$\frac{\cancel{a_1}}{\cancel{b_1}} a_1 \frac{d\bar{P}}{d\bar{t}} = \left(-a_2 + \cancel{b_2} \cdot \frac{a_2}{\cancel{b_2}} \bar{N}\right) \frac{\cancel{a_1}}{\cancel{b_1}} \bar{P}$$

$$\boxed{\frac{d\bar{P}}{d\bar{t}} = -\bar{a}(1 - \bar{N}) \bar{P}}$$

(c).. Modified Scripts:

```
% predator_prey2.m
%
% Solve non-dimensional modified Lotka-Volterra equations, with the prey
% following logistic growth (in the absence of predators).
%
% M.S. Wheatland, 9 May 2001

global a DELTA % makes these variables global so they are visible to
               % the function defining the RHS of ODEs

a=0.5;
DELTA=0.2;
N0=0.5;
P0=0.5;
TMAX=50; % maximum value of non-dimensional time

options = odeset('RelTol',1e-4,'AbsTol',[1e-4 1e-4]); % specify accuracy

% solve for t=0 to TMAX, using built-in procedure ode45
% RHS of ODEs is in lotka_volterra_mod.m

[t,y]=ode45('lotka_volterra_mod_e',[0 TMAX],[N0 P0]);

% plot results

subplot(1,2,1)
plot(t,y(:,1),'-',t,y(:,2),'--')
xlabel('time (non-dim.)')
ylabel('populations N,P (non-dim.)')
title(['Mod. L-V (a=',num2str(a),' delta=',num2str(DELTA),')'])
subplot(1,2,2)
plot(y(:,1),y(:,2))
xlabel('prey population N (non-dim.)')
ylabel('predator population P (non-dim.)')
title('Phase space diagram')
```

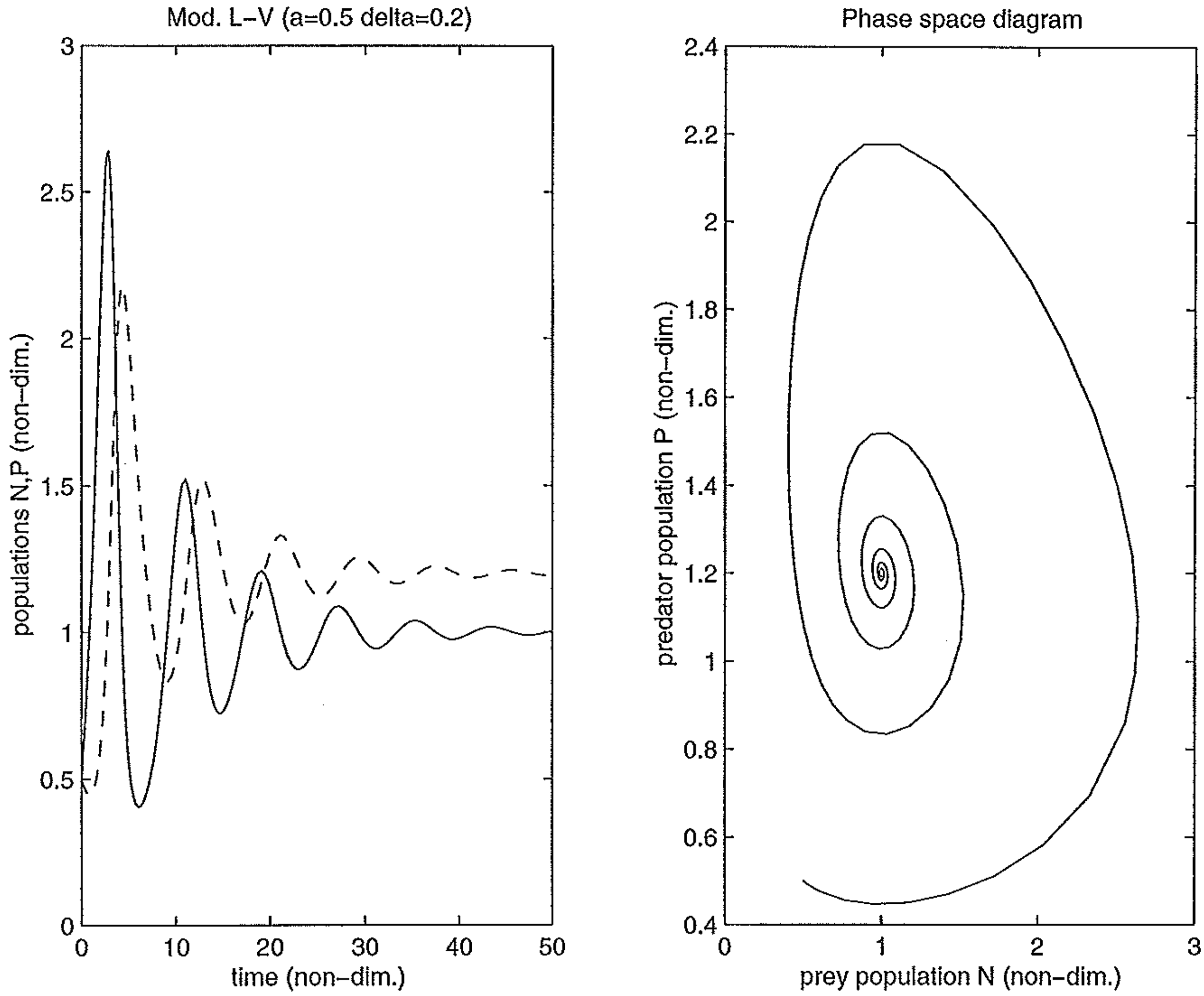
```
function dy=lotka_volterra_mod_e(t,y)

% RHS of modified Lotka-Volterra ODEs

global a DELTA % recognise these variables from main program

dy=zeros(2,1); % define column vector

dy(1)=(1-y(2))*y(1)+DELTA;
dy(2)=-a*(1-y(1))*y(2);
```



Mod. L–V (a=0.5 delta=0.2)
populations N,P (non–dim.)
time (non–dim.)
0
0.5
1
1.5
2
2.5
3
0
10
20
30
40
50
Phase space diagram
predator population P (non–dim.)
prey population N (non–dim.)
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
2.2
2.4
0
1
2
3

Stability analysis:

$$N = \frac{a_2}{b_2} + \varepsilon_2$$

$$P = \frac{b_2}{b_1 a_2} r + \frac{a_1}{b_1} + \varepsilon_1$$

In L-V (mod):

$$\frac{d\varepsilon_2}{dt} = \left[\cancel{a_1} - \frac{b_2}{a_2} r - \cancel{a_1} - b_1 \varepsilon_1\right]\left(\frac{a_2}{b_2} + \varepsilon_2\right) + r$$

$$\frac{d\varepsilon_2}{dt} = -\cancel{r} - \frac{b_2}{a_2} r \varepsilon_2 - \frac{a_2 b_1}{b_2} \varepsilon_1 + \cancel{r} \quad (3)$$

other:

$$\frac{d\varepsilon_1}{dt} = \left(-\cancel{a_2} + \frac{\cancel{b_2 a_2}}{\cancel{b_2}} + b_2 \varepsilon_2\right)\left(\frac{b_2}{b_1 a_2} r + \frac{a_1}{b_1} + \varepsilon_1\right)$$

$$= \frac{b_2^2}{b_1 a_2} r \varepsilon_2 + \frac{b_2 a_1 \varepsilon_2}{b_1} \quad (4)$$

$$= \left(\frac{b_2^2}{b_1 a_2} r + \frac{b_2 a_1}{b_1}\right) \varepsilon_2$$

$$\frac{d^2\varepsilon_2}{dt^2} = -\frac{b_2}{a_2} r \frac{d\varepsilon_2}{dt} - \frac{a_2 b_1}{b_2} \frac{d\varepsilon_1}{dt}$$

$$\frac{d^2\varepsilon_2}{dt^2} = -\frac{b_2 r}{a_2} \frac{d\varepsilon_2}{dt} - \frac{a_2 \cancel{b_1}}{\cancel{b_2}}\left(\frac{b_2^{\cancel{2}}}{\cancel{b_1} a_2} r + \frac{\cancel{b_2} a_1}{\cancel{b_1}}\right)\varepsilon_2$$

$$= 0$$

$$\frac{d^2\varepsilon_2}{dt^2} + \frac{b_2 r}{a_2}\frac{d\varepsilon_2}{dt} + a_2\left(\frac{b_2 r}{a_2} + a_1\right)\varepsilon_2 = 0$$

$$\frac{d^2\varepsilon_2}{dt^2} + \frac{b_2 r}{a_2}\frac{d\varepsilon_2}{dt} + (b_2 a_2 r + a_1 a_2)\varepsilon_2 = 0$$

so if $r = 0$,

$\frac{d^2\varepsilon_1}{dt^2} = -\omega^2\varepsilon_2$

$$\frac{d^2\varepsilon_2}{dt^2} + a_1 a_2 \varepsilon_2 = 0 \quad \checkmark$$

Now: $\varepsilon_2 = Ae^{\lambda t}$

$$\lambda^2 + \frac{b_2 r}{a_2}\lambda + (b_2 a_2 r + a_1 a_2) = 0$$

$$\lambda = \frac{-\frac{b_2 r}{a_2} \pm \sqrt{\frac{b_2^2 r^2}{a_2^2} - 4(b_2 a_2 r + a_1 a_2)}}{2}$$

r is small cf. other term

$[a_i] = \frac{1}{T}$ $\qquad \frac{r}{a_1} \ll \frac{a_2}{b_2}$

$$\frac{r b_2}{a_1 a_2} = \delta$$

or $\quad r = \frac{\delta a_1 a_2}{b_2}$

$$\lambda = \frac{-\delta a_1 \pm \sqrt{\delta^2 a_1^2 - 4\left(\cancel{b}_2 a_2 \frac{\delta a_1 a_2}{b/2} + a_1 a_2\right)}}{2}$$

$$= \frac{-\delta a_1 \pm \sqrt{\delta^2 a_1^2 - 4 a_1 a_2 - 4\delta a_1 a_2^2}}{2}$$

$$\lambda = \frac{-\delta a_1 \pm i\sqrt{4 a_1 a_2 + 4\delta a_1 a_2^2 - \delta^2 a_1^2}}{2}$$

$$= -\frac{\delta a_1}{2} \pm i (a_1 a_2)^{\frac{1}{2}} \sqrt{1 + \delta \frac{a_2}{a_1} + \cdots}$$

$$\approx -\frac{\delta a_1}{2} \pm i (a_1 a_2)^{\frac{1}{2}} \left[1 + \frac{1}{2}\delta \frac{a_2}{a_1}\right]$$

so oscillation & collapse.

Non-dim version:
Equilibrium points:

$$N = \frac{a_2}{b_2} \Rightarrow \bar{N} = 1$$

$$P = \frac{a_1}{b_1} + \frac{b_2}{b_1 a_2} r$$

$$\bar{P} = P/(a_1/b_1) = 1 + \frac{\cancel{b_1}}{a_1} \cdot \frac{b_2}{\cancel{b_1} a_2} r$$

$$= 1 + \frac{b_2 r}{a_1 a_2}$$

$$= 1 + \delta$$

$$\frac{d\bar{N}}{d\bar{t}} = (1 - \bar{P})\bar{N} + \delta$$

$$\frac{d\bar{P}}{d\bar{t}} = -\bar{a}(1 - \bar{N})\bar{P}$$

Perturbation:

$$\bar{N} = 1 + \cancel{\varepsilon_k}\ \varepsilon$$

$$\bar{P} = 1 + \delta + \cancel{\varepsilon}\ \Delta$$

$$\frac{d\varepsilon}{dt} = (\cancel{1} - \cancel{1} - \overset{\delta - \Delta}{\cancel{\varepsilon}})(1 + \varepsilon) + \delta$$

$$= -\cancel{\delta} - \delta\varepsilon - \Delta - \Delta.\varepsilon + \cancel{\delta}$$

$$= -(\delta + \Delta)\varepsilon - \Delta$$

$$\frac{d\varepsilon}{dt} = -\delta\varepsilon \cancel{- \Delta\varepsilon} - \Delta \qquad (1)$$

&

$$\frac{d\Delta}{dt} = -\bar{a}\,(\cancel{1} - \cancel{1} - \varepsilon)(1+\delta+\Delta)$$

$$= \bar{a}\,\varepsilon(1+\delta+\Delta)$$

$$= \bar{a}\varepsilon + \bar{a}\varepsilon\delta \cancel{+ \bar{a}\varepsilon\Delta} \qquad (2)$$

diff. (1):

$$\frac{d^2\varepsilon}{dt^2} = -\delta\frac{d\varepsilon}{dt} - \frac{d\Delta}{dt}$$

$$= -\delta\frac{d\varepsilon}{dt} - \bar{a}\varepsilon - \bar{a}\varepsilon\delta$$

$$\frac{d^2\varepsilon}{dt^2} + \delta\frac{d\varepsilon}{dt} + \bar{a}(1+\delta)\varepsilon = 0$$

$$\varepsilon = \frac{-\delta \pm \sqrt{\delta^2 - 4\bar{a}(1+\delta)}}{2}$$

$$= \frac{-\delta \pm 2(\bar{a})^{\frac{1}{2}} i\sqrt{1+\delta-\frac{\delta^2}{4\bar{a}}}}{2}$$

$$\approx -\frac{\delta}{2} \pm (\bar{a})^{\frac{1}{2}} i\left(1+\tfrac{1}{2}\delta\right)$$

Linear second order ODEs

Reference: Zwillinger, D. *Handbook of Differential Equations*, Academic Press Inc., Boston, 1989

The treatment of linear stability analysis in the lectures involves solution of second order linear ODEs. The general form is:

$$\frac{d^2y}{dx^2} + a\frac{dy}{dx} + by = 0, \tag{1}$$

where a and b are constants. Substituting the trial solution $y = Ce^{\lambda x}$ leads to the *characteristic equation*

$$\lambda^2 + a\lambda + b = 0, \tag{2}$$

which is a quadratic with solutions

$$\lambda_{\pm} = \frac{-a \pm \sqrt{a^2 - 4b}}{2}. \tag{3}$$

The general solution to (1) is then

$$y = C_+ e^{\lambda_+ x} + C_- e^{\lambda_- x}, \tag{4}$$

where $C_{\pm}$ are constants.

An important special case is when $a = 0$, i.e. when we are solving

$$\frac{d^2y}{dx^2} + by = 0. \tag{5}$$

In this case the solutions to the characteristic equation are $\lambda_{\pm} = \pm(-b)^{1/2}$. There are two possibilities that you should learn to recognise, as follows.

1. If $b > 0$, then writing $b = k^2$ where k is a positive constant, we are solving

   $$\frac{d^2y}{dx^2} + k^2 y = 0.$$

   The solutions to the characteristic equation are $\lambda_{\pm} = \pm(-k^2)^{1/2} = \pm ki$, and the general solution to the ODE is $y = C_+ e^{ikx} + C_- e^{-ikx}$. Noting that $e^{i\theta} = \cos\theta + i\sin\theta$ it follows that the solution can be written in the form

   $$y = A\cos(kx) + B\sin(kx) = A'\cos(kx + \phi),$$

   where A, b, and A' are constants. For our purposes the solutions are always real, so the constants will be real.

2. If $b < 0$, then writing $b = -k^2$ where k is a positive constant, we are solving

   $$\frac{d^2y}{dx^2} - k^2 y = 0.$$

   The solutions to the characteristic equation are $\lambda_{\pm} = \pm(k^2)^{1/2} = \pm k$, and the general solution to the ODE is $y = C_+ e^{kx} + C_- e^{-kx}$.

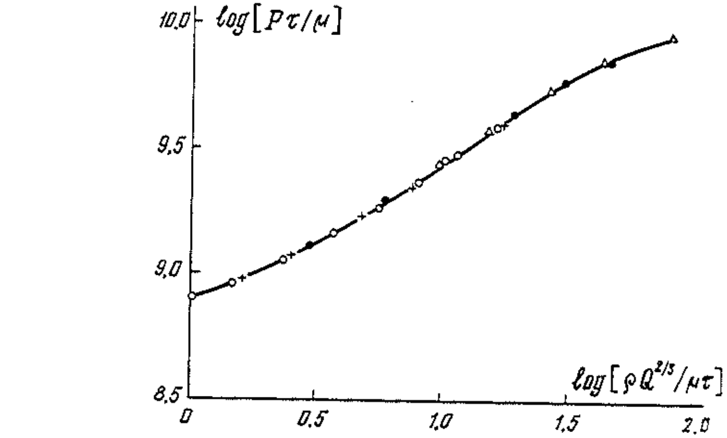


Figure 1.3. Results of the experiments of E. Bose, D. Rauert, and M. Bose in the form in which Th. von Kármán presented them, using dimensional analysis. All the experimental points lie on a single curve.

This formula shows that if by some means one measures the radius of the shock wave at various instants of time, then in logarithmic coordinates $\frac{5}{2}\log r_f$, $\log t$ the experimental points must lie on the straight line

$$\frac{5}{2}\log r_f = \frac{5}{2}\log CE^{1/5}\rho_0^{-1/5} + \log t,$$

having slope equal to one. G. I. Taylor confirmed this extremely well, making use of a movie film of the spread of the fireball, taken by J. Mack at the time of an American nuclear test (Fig. 1.4). As a more detailed calculation shows (cf. the next chapter), the factor C is close to one. Knowing this, from the experimental dependence of the radius of the front on time one can determine the

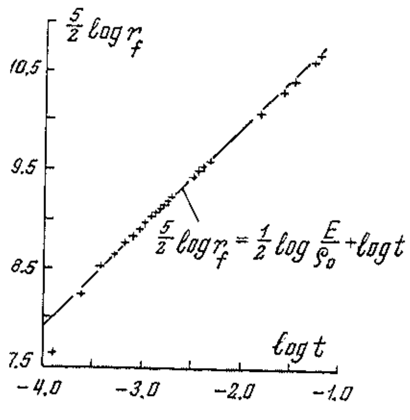


Figure 1.4. Propagation of the shock wave of an atomic explosion. The experimental points determined from J. Mack's motion picture lie, in the coordinates $\frac{5}{2}\log r_f$, $\log t$, on a straight line with slope equal to unity over a large time interval. Analysis of Mack's film enabled G. I. Taylor to determine the energy of the explosion.

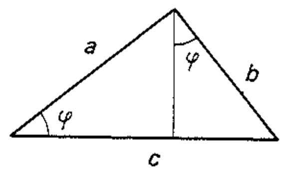


Figure 1.5. Proof of the Pythagorean theorem with the help of dimensional analysis.

energy of the explosion. Taylor's publication of this quantity (which turned out to be equal to $\sim 10^{21}$ ergs) evoked in its time, in his words, considerable confusion in American governmental circles, since this number was considered entirely secret, although Mack's film was not secret.

We give another rather more amusing example of the application of dimensional analysis, using it to prove the Pythagorean theorem [cf. also Migdal (1977)]. The area S of a right triangle is determined by the size of its hypotenuse c and, for definiteness, the lesser of the acute angles φ: $S = f(c, \varphi)$. Obviously, dimensional analysis gives $S = c^2\Phi(\varphi)$. The altitude perpendicular to the hypotenuse (Fig. 1.5) divides the basic triangle into two right triangles that are similar to it, and whose hypotenuses are the legs a and b of the basic triangle. Thus, their areas are equal to $S_1 = a^2\Phi(\varphi)$, $S_2 = b^2\Phi(\varphi)$, where $\Phi(\varphi)$ is just the same as in the case of the basic triangle. But the sum of the areas S_1 and S_2 is equal to the area of the basic triangle S: $S = S_1 + S_2$, whence $c^2\Phi(\varphi) = a^2\Phi(\varphi) + b^2\Phi(\varphi)$, so that $c^2 = a^2 + b^2$, which was to be proved. It is evident that the theorem is based essentially on Euclidean geometry: in Riemannian and Lobachevskian geometry there is an intrinsic parameter λ having the dimensions of length, and the proof does not go through: to the number of arguments of the function Φ one must append the ratio of the hypotenuse to λ.

The examples just considered show that apparently trivial considerations of dimensional analysis can yield quite significant results. The most important element here is a proper definition of the set of governing parameters. Finding the set of governing parameters is simple if one has a mathematical formulation of the problem—it is the set of independent variables and parameters of the problem appearing in the equations, boundary conditions, initial conditions, etc., that determine a solution of the problem in a unique way. The proper choice of governing parameters in a problem that does not have an explicit mathematical formulation is dependent, first of all, on the intuition of the investigator—success here depends on a proper understanding of which parameters are really important and which can be neglected. This question will be considered in detail below.

3. Similarity

In most cases, before some large and expensive structure is built, such as a ship or airplane, in order to obtain its best characteristics under future working