# THE INTERNATIONAL ENCYCLOPEDIA OF PHYSICAL CHEMISTRY AND CHEMICAL PHYSICS

THE INTERNATIONAL
ENCYCLOPEDIA
OF PHYSICAL CHEMISTRY
AND CHEMICAL PHYSICS

Topic 1. MATHEMATICAL TECHNIQUES

EDITOR: PROFESSOR H. JONES

Volume 1

# ORDINARY DIFFERENTIAL EQUATIONS

BY

E. R. LAPWOOD, M.A., PH.D.

**THE INTERNATIONAL ENCYCLOPEDIA**
**OF PHYSICAL CHEMISTRY AND CHEMICAL PHYSICS**

# THE INTERNATIONAL ENCYCLOPEDIA OF PHYSICAL CHEMISTRY AND CHEMICAL PHYSICS



*List of Topics and Editors*

| | |
|---|---|
| 1. Mathematical Techniques | H. JONES, *London* |
| 2. Classical and Quantum Mechanics | R. MCWEENY, *Keele* |
| 3. Electronic Structure of Atoms | C. A. HUTCHISON, JR., *Chicago* |
| 4. Molecular Binding | J. W. LINNETT, *Cambridge* |
| 5. Molecular Properties | |
| (*a*) Electronic | J. W. LINNETT, *Cambridge* |
| (*b*) Non-electronic | N. SHEPPARD, *East Anglia* |
| 6. Kinetic Theory of Gases | E. A. GUGGENHEIM, *Reading* |
| 7. Classical Thermodynamics | D. H. EVERETT, *Bristol* |
| 8. Statistical Mechanics | J. E. MAYER, *La Jolla* |
| 9. Transport Phenomena | J. C. MCCOUBREY, *Birmingham* |
| 10. The Fluid State | J. S. ROWLINSON, *London* |
| 11. The Ideal Crystalline State | M. BLACKMAN, *London* |
| 12. Imperfections in Solids | Editor to be appointed |
| 13. Mixtures, Solutions, Chemical and Phase Equilibria | M. L. MCGLASHAN, *Exeter* |
| 14. Properties of Interfaces | D. H. EVERETT, *Bristol* |
| 15. Equilibrium Properties of Electrolyte Solutions | R. A. ROBINSON, *Washington, D.C.* |
| 16. Transport Properties of Electrolytes | R. H. STOKES, *Armidale* |
| 17. Macromolecules | C. E. H. BAWN, *Liverpool* |
| 18. Dielectric and Magnetic Properties | J. W. STOUT, *Chicago* |
| 19. Gas Kinetics | A.F.TROTMAN-DICKENSON, *Aberystwyth* |
| 20. Solution Kinetics | R. M. NOYES, *Eugene* |
| 21. Solid and Surface Kinetics | F. C. TOMPKINS, *London* |
| 22. Radiation Chemistry | R. S. LIVINGSTON, *Minneapolis* |

# ORDINARY DIFFERENTIAL EQUATIONS

BY

E. R. LAPWOOD, M.A., PH.D.

DEPARTMENT OF APPLIED MATHEMATICS AND THEORETICAL PHYSICS
UNIVERSITY OF CAMBRIDGE

PERGAMON PRESS

OXFORD · LONDON · EDINBURGH · NEW YORK
TORONTO · SYDNEY · PARIS · BRAUNSCHWEIG

Pergamon Press Ltd., Headington Hill Hall, Oxford
4 & 5 Fitzroy Square, London W.1
Pergamon Press (Scotland) Ltd., 2 & 3 Teviot Place, Edinburgh 1
Pergamon Press Inc., 44–01 21st Street, Long Island City, New York 11101
Pergamon of Canada Ltd., 207 Queen's Quay West, Toronto 1
Pergamon Press (Aust.) Pty. Ltd., Rushcutters Bay, Sydney, N.S.W.
Pergamon Press S.A.R.L., 24 rue des Écoles, Paris 5e
Vieweg & Sohn GmbH, Burgplatz 1, Braunschweig



First edition 1968

Library of Congress Catalog Card No. 68-21278

PRINTED IN GREAT BRITAIN BY BELL AND BAIN LTD., GLASGOW

08 003409 8

# INTRODUCTION

THE International Encyclopedia of Physical Chemistry and Chemical Physics is a comprehensive and modern account of all aspects of the domain of science between chemistry and physics, and is written primarily for the graduate and research worker. The Editors-in-Chief, Professor D. D. ELEY, Professor J. E. MAYER and Professor F. C. TOMPKINS, have grouped the subject matter in some twenty groups (General Topics), each having its own editor. The complete work consists of about one hundred volumes, each volume being restricted to around two hundred pages and having a large measure of independence. Particular importance has been given to the exposition of the fundamental bases of each topic and to the development of the theoretical aspects; experimental details of an essentially practical nature are not emphasized although the theoretical background of techniques and procedures is fully developed.

The Encyclopedia is written throughout in English and the recommendations of the International Union of Pure and Applied Chemistry on notation and cognate matters in physical chemistry are adopted. Abbreviations for names of journals are in accordance with *The World List of Scientific Periodicals.*

# CONTENTS

# PREFACE

THIS book has been written to explain mathematical techniques useful in setting up and solving the *ordinary differential equations* met by physical chemists. Such equations arise most commonly either from studies of reaction rates in chemical kinetics or from separation of variables in partial differential equations of diffusion theory or quantum theory.

Methods of solution are introduced only if they apply to the equations of physical chemistry; methods which are included in elementary texts mainly because they sometimes produce exact solutions in terms of elementary functions have been omitted.

The first four chapters deal with differential equations which describe chemical reactions. In them I have explained the method of setting up the equations in more detail than is usual. Chapters 5 and 6 deal with the ordinary differential equations which are obtained from partial differential equations of physical chemistry by the process of separating variables. I have paid special attention to approximate analytical methods, but numerical methods are not within the scope of this book. Full references are provided wherever further theory or techniques may be required.

I am glad to acknowledge valuable help from Miss V. G. Bozman and Mr. J. A. Hulse, from my colleagues Professor P. G. Ashmore, Dr. B. A. Thrush and Dr. A. J. Stone, and especially from Dr. J. D. P. Meldrum, who checked mathematical detail and read proofs with great care.

E. R. L.

CHAPTER 1

# THE SETTING UP OF AN ORDINARY DIFFERENTIAL EQUATION

## 1.1. Translation of working hypothesis into differential equation

The *scientific method* may be analysed into successive steps as follows.

(1) Collection and systematic arrangement of data, by experiment where possible.
(2) Search for fundamental relationships hidden in the data.
(3) Setting up of working hypotheses.
(4) Examination of the logical consequences of the hypotheses, both to explain known phenomena and to propose new tests.
(5) Making of the experiments proposed in (4), as tests of the working hypotheses (3).
(6) Modification of the hypotheses in the light of the experiments (5).

Step (4) will be best achieved if the hypotheses are exact enough to enable mathematical techniques to be brought into action. The rigorous deductive methods of mathematics make it possible to say that from certain hypotheses certain conclusions follow inexorably. Moreover it will usually be possible to forecast quantitative results which can be subjected to experimental proof or disproof and so form a crucial test of the working hypotheses.

Most of the working hypotheses of physics and chemistry are concerned with *changing* quantities. We therefore need a branch of mathematics that deals with relations between changing quantities: this is the calculus. A relation set up between certain quantities and their rates of change with regard to a single independent variable will constitute an *ordinary differential equation*. In most of the problems that we meet, this single parameter will be the *time* that has elapsed since a certain origin of time. We begin with elementary examples of the setting up of a differential equation that relates a single dependent

variable to a single independent variable. Methods of solution of the resulting equation will be treated in later chapters.

(a) *The law of natural decay*

Let us suppose that it is found that the amount of a certain material is decreasing, and that the rate of decrease at a certain moment is proportional to the amount of the material that exists at that moment (as, for instance, in radioactive decay).

We first select our variables. Here we take the time $t$, measured from a certain instant, to be the independent variable. We take $x$ to represent the quantity of the material (measured in convenient units), so that $x$ varies with $t$, and $x(t)$ denotes the value of $x$ at time $t$. Then the rate of change of $x$ is $\mathrm{d}x/\mathrm{d}t$. [We have assumed that $\mathrm{d}x/\mathrm{d}t$ exists and is unique. That is, we have assumed that $x$ is a *differentiable* function of $t$. We add a comment on this assumption later.]

Now we formulate our hypothesis in exact terms: we assume that $\mathrm{d}x/\mathrm{d}t$ is strictly proportional to $x$, and that this relation completely determines the way in which $x$ changes as long as $t$ lies within a certain interval of time. Thus we write

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -kx \quad \text{for} \quad t_0 \leqq t \leqq t_1, \tag{1.1.1}$$

where $k$ is a positive constant. (We take $k$ positive since $x$ is naturally positive and we are told that the amount of material is decreasing, which means that $\mathrm{d}x/\mathrm{d}t$ is negative.)

We may next proceed, by methods which are described later, to find $x$ as a function of $t$. Since integration is involved, there will be introduced a parameter which cannot be determined from the data given so far. The result, which is easily verified by substitution into (1), is

$$x(t) = A\mathrm{e}^{-kt}, \tag{1.1.2}$$

where $\mathrm{e}^{-kt}$ is the exponential function of $(-kt)$ and $A$ is an arbitrary constant.

If we now introduce further information—namely that at $t_0 = t$ the quantity of material is $x_0$—we can select $A$ to satisfy this requirement. Thus

$$x_0 = A\mathrm{e}^{-kt_0}$$

and

$$x(t) = x_0\mathrm{e}^{-k(t-t_0)}. \tag{1.1.3}$$

This expression (3) satisfies both the differential equation (1), which describes the process during the interval $t_0 \leqq t \leqq t_1$, and the *initial condition* that $x = x_0$ when $t = t_0$.

Consider now the graphical interpretation of our procedure. We take $t$ as abscissa and $x$ as ordinate in a rectangular cartesian coordinate system, as shown in Fig. 1.1. Then our hypothesis states that if $x$ can be represented as a function of $t$ by a curve $\Gamma$ between $t = t_0$ and $t = t_1$, the slope of $\Gamma$ at any point $P$ with abscissa $t$ is $-k$ times the ordinate $x$ at $P$.

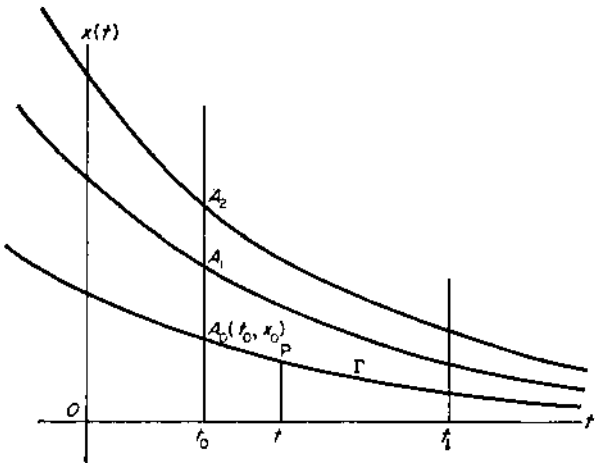


FIG. 1.1. Curves of the family $\{\Gamma\}$ representing natural decay.

There is not a unique curve with this property. In fact the relation (2) shows that $\Gamma$ is one of a *family*, which we denote by $\{\Gamma\}$, of curves obtained by giving all possible real values to $A$ in (2). By specifying that $x = x_0$ when $t = t_0$ we select from the family the curve that passes through the point $(t_0, x_0)$; the appropriate value of $A$ is then determined as $x_0 e^{kt_0}$. We may say that each point on the ordinate at $t_0$ lies on (or determines) one member of the family $\{\Gamma\}$. Given that point (i.e. the value of $x$ when $t = t_0$) and the differential equation, the subsequent history of the system during the period $t_0 \leqq t \leqq t_1$ is completely determined.

It is sometimes useful to obtain from the differential equation general information about the behaviour of the system. Here, for instance, we observe from (1) that, $k$ being positive, positive values of $x$

imply negative values of $dx/dt$—that is, all curves $\Gamma$ above the axis of $t$ have negative slope, so that $x$ gets smaller as $t$ increases. Moreover, the absolute value of the slope is smaller for smaller $x$. Furthermore, if we differentiate the differential equation we have

$$\frac{d}{dt}\left(\frac{dx}{dt}\right) = -k\frac{dx}{dt} = k^2x, \tag{1.1.4}$$

which is positive where $x$ is positive, showing that the slope increases as $t$ increases. In other words, for $x > 0$ the curves $\Gamma$ are all convex towards the axis of $t$. Thus we obtain, without solving the differential equation, a general idea of the solution. Figure 1.1 shows some members of the family $\{\Gamma\}$ obtained by graphical representation of (2) for various values of $A$. The form of these curves bears out the rough description of the family as deduced from the differential equation itself.

Note that although (2) may be used to define curves over the interval $-\infty < t < \infty$, our differential equation has been set up only for the range $t_0 \leqq t \leqq t_1$, and is not to be expected to describe the behaviour of the system outside that range.

(b) *The law of natural growth*

If a quantity is *increasing* at a rate proportional to the amount of that quantity, as in population growth, we have in place of (1) the equation of *natural growth*

$$\frac{dx}{dt} = kx, \quad k > 0, \quad t_0 \leqq t \leqq t_1,$$

with solution

$$x = Ae^{kt}.$$

The properties of this equation and its solution may be discussed in exactly the same way as those of the law of natural decay. The reader may show that the family $\{\Gamma'\}$ of solutions has the form shown in Fig. 1.2.

[*Note on continuity and differentiability.* When we deal with distances comparable with inter-molecular distances and with time intervals comparable with the period of vibration of a molecule, it is necessary to take into account the essentially discontinuous nature of materials and processes. But if we are dealing (as is usually the case in chemistry) with an amount of matter of linear dimensions large compared with

inter-molecular distances, we regard it as continuous—that is, for purposes of mathematical analysis we consider instead of the real material an ideal material which is a continuum. The error thus introduced is negligible as long as we are dealing with lengths great compared with $10^{-8}$ cm. Similarly changes may be considered as continuous if we are dealing with intervals of time long compared with $10^{-13}$ sec.

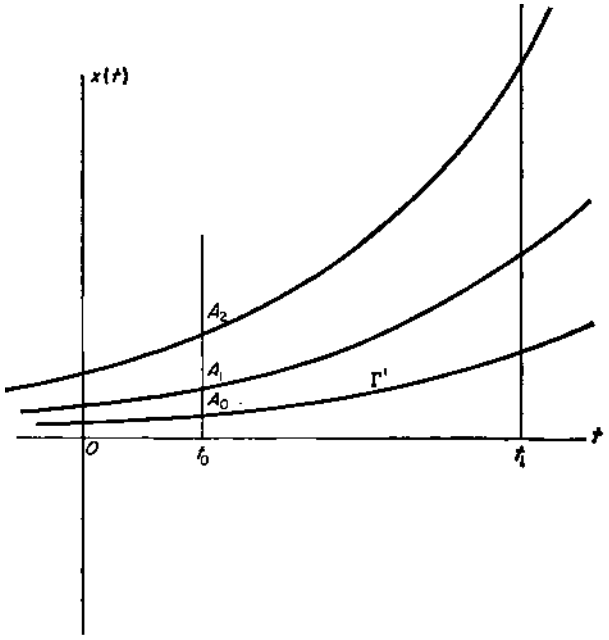


FIG. 1.2. Curves of family $\{\Gamma'\}$ representing natural growth.

The mathematical definitions of continuity and differentiability are as follows:

$x(t)$ is a *continuous* function of $t$ at $t = t'$ if, given any small positive number $\varepsilon$ (however small),

$$| x(t) - x(t') |$$

can be made less than $\varepsilon$ by taking $t$ near enough to $t'$. $x(t)$ is *differentiable* at $t = t'$ if in addition to being continuous there it has a unique value of

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \lim_{t \to t'} \frac{x(t) - x(t')}{t - t'}.$$

If we illustrate a continuous function $x(t)$ by a continuous curve, then in Fig. 1.3 curve (a) is continuous but not differentiable at $t'$ while curve (b) is both continuous and differentiable there.]

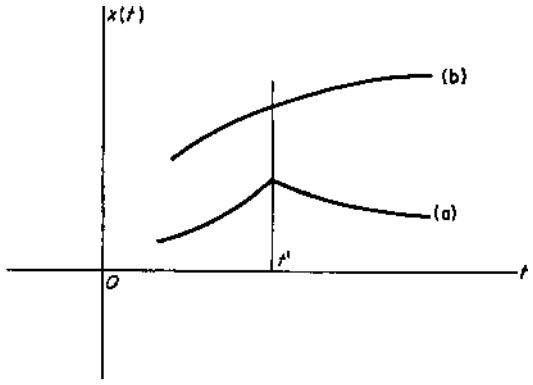


FIG. 1.3. (b) is differentiable at $t'$ but (a) is not.

(c) *Linear motion about a centre of attraction or repulsion*

Let us now suppose that the acceleration of a point which moves on a straight line is proportional to its distance from a given origin on the line, and that this relation is true for all time.

We use $\xi$ to denote the distance of the moving point P from the fixed origin $O$. Then if $t$ denotes time, the acceleration of P is $\mathrm{d}^2\xi/\mathrm{d}t^2$. We formulate our hypothesis in the differential equation

$$\frac{\mathrm{d}^2\xi}{\mathrm{d}t^2} = \lambda\xi, \quad -\infty < t < \infty, \tag{1.1.5}$$

where $\lambda$ is a constant. It has not been stated whether $\lambda$ is positive or negative.

Since a curve in the $(t, \xi)$ plane corresponding to a solution $\xi = \xi(t)$ of the differential equation (5) will be convex towards the $t$-axis for regions where both $\xi > 0$ and $\lambda > 0$, and concave towards the $t$-axis for $\xi > 0$ and $\lambda < 0$, the sign of $\lambda$ must play a decisive part in determining the nature of the solutions. In fact $\lambda > 0$ for motion about a centre of repulsion and $\lambda < 0$ for motion about a centre of attraction, and we will see later (Section 3.2) that the form of solution differs for the two cases.

Here, since the differential equation contains a second derivative, the solution will involve two integrations and therefore the introduction of two arbitrary parameters. Consequently the solution may be made to satisfy two initial conditions or two other conditions. Complete specification of the problem, so as to lead to a solution free from arbitrariness, must include the statement of these conditions.

(d) *Damped vibrations*

We consider the motion of a particle of mass $m$ which moves on a straight line. This particle is subject to an attractive force directed towards a fixed origin on the line and proportional to the distance from the origin (as by an elastic bond). It is subject also to a resistance which is proportional to the speed. We set up the differential equation of motion of the particle as follows.

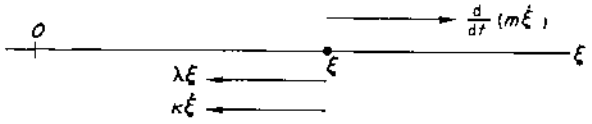


FIG. 1.4. Damped vibrations.

We denote displacement from the fixed origin by $\xi$ (Fig. 1.4). Using the Newtonian fluxional notation, in which a superior dot means a time derivative ($\dot{\xi} \equiv \mathrm{d}\xi/\mathrm{d}t$), we set down the rate of change of momentum as

$$\frac{\mathrm{d}}{\mathrm{d}t}(m\dot{\xi}). \tag{1.1.6}$$

We will take $m$ to be constant, so that (6) becomes $m\ddot{\xi}$. [It should be noted that in problems where the mass varies the form (6) must be retained.] This is in the direction of increasing $\xi$. The elastic restoring force is $\lambda\xi$ in the opposite direction, and the resistance is $\kappa\dot{\xi}$ opposing the motion, where $\lambda$ and $\kappa$ are positive constants which describe the intrinsic properties of the system. Applying Newton's Second Law, we write down the equation of motion

$$m\ddot{\xi} = -\kappa\dot{\xi} - \lambda\xi. \tag{1.1.7}$$

The range of $t$ for which equation (7) is valid has not yet been stated. If the motion was started at time $t_0$ and is allowed to proceed without

interference for an indefinitely long time, we may take the range of applicability of (7) as $t_0 \leqq t < \infty$.

## 1.2. Classification of differential equations

The equations (1), (5) and (7) of Section 1.1 are elementary examples, containing only one independent and one dependent variable and being linear in the dependent variable and its derivatives. We formulate definitions to cover more complicated equations. An *ordinary differential equation* is one in which all derivatives have been formed with respect to the same single independent variable. A *partial differential equation* is one in which occur derivatives with respect to more than one independent variable. In this book we confine attention to ordinary differential equations.

The *order* of a derivative is the number of times that the operation of differentiation has been performed in obtaining it. Thus

$$x, \frac{\mathrm{d}x}{\mathrm{d}t}, \frac{\mathrm{d}^2x}{\mathrm{d}t^2}, \ldots, \frac{\mathrm{d}^nx}{\mathrm{d}t^n}$$

are the derivatives of zeroth, first, second, . . . , $n$th order. The order of the highest derivative in a differential equation is the *order of the equation*. The power to which the highest derivative is raised is the *degree* of the equation. Thus

$$\left(\frac{\mathrm{d}^3x}{\mathrm{d}t^3}\right)^2 + \left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^3 + x^4 = f(t), \quad t_0 \leqq t \leqq t_1$$

is an ordinary differential equation of order 3 and degree 2. A differential equation in which each derivative (including the zeroth) occurs in the first degree is called a *linear* differential equation. Linear differential equations assume great importance because a first approximation to the solution of a physical problem may sometimes be found by reducing the differential equation to linear form by neglect of higher powers, and also because linear equations have certain properties which simplify the mathematical theory (see Sections 5.2, 5.3).

We frequently meet sets of differential equations in which there is one independent variable (commonly the *time*) but there are several dependent variables. These will be called sets of *simultaneous ordinary differential equations*. There will be provided by the physical or chemical laws as many simultaneous equations as there are dependent variables. We give two examples.

(a) *Plane motion of a particle subject to gyroscopic and restoring forces*

A particle P of mass $m$ moves in the $(x, y)$ plane under the control of two forces $F$ and $G$. $F$ acts towards the origin and is proportional to $OP$. $G$ acts at right angles to the velocity $V$ of P (in the sense shown in Fig. 1.5) and is proportional to $V$. The force $G$ is of the kind that appears to act in gyroscopic motion, and is called a *gyroscopic force* (even where no spinning body is concerned). The force experienced by an electron moving in a magnetic field is a gyroscopic force. Let $OP$ have length $r$ and make an angle $\theta$ with $Ox$, and let $V$ make an angle $\psi$

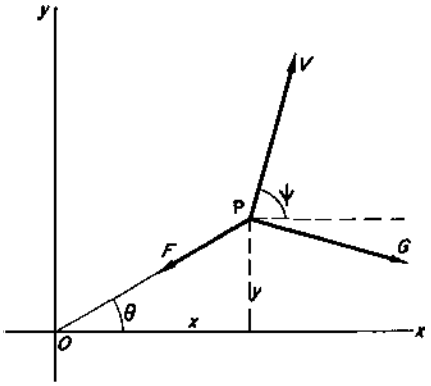


FIG. 1.5. Motion under elastic restoring force F and gyroscopic force G.

with $Ox$. Then the equations of motion of P are (by taking components in the directions $Ox$ and $Oy$):

$$\left.\begin{aligned} m\ddot{x} &= -F\cos\theta + G\sin\psi, \\ m\ddot{y} &= -F\sin\theta - G\cos\psi. \end{aligned}\right\} \tag{1.2.1}$$

We take $F = m\lambda r$, $G = 2m\gamma V$, where $\lambda$ and $\gamma$ are constants, so that

$$F\cos\theta = m\lambda r\frac{x}{r} = m\lambda x, \qquad F\sin\theta = m\lambda y,$$

$$G\cos\psi = 2m\gamma V\cos\psi = 2m\gamma\dot{x}, \qquad G\sin\psi = 2m\gamma\dot{y}.$$

Then the equations of motion (1) become

$$\left.\begin{aligned} \ddot{x} - 2\gamma\dot{y} + \lambda x &= 0, \\ \ddot{y} + 2\gamma\dot{x} + \lambda y &= 0. \end{aligned}\right\} \tag{1.2.2}$$

[Readers familiar with vectors will see that the equations (2) could have been derived directly from the vectorial equation of motion

$$\ddot{\mathbf{r}} = -\lambda\mathbf{r} + 2\gamma\dot{\mathbf{r}} \times \mathbf{k},$$

where $\mathbf{r}$ is the position vector $\overrightarrow{OP}$ and $\mathbf{k}$ is the unit vector in the direction $Oz$ which forms a right-handed orthogonal triad with $Ox$, $Oy$.]

The equations (2) form a pair of simultaneous linear differential equations of the second order with constant coefficients. They are set up for a range of $t$ which must be specified, and if they are to determine $x(t)$ and $y(t)$ uniquely we must add data sufficient to fix the four arbitrary parameters which arise in integration. Methods of integration are described in Chapters 3 and 4.

(b) *Series of first-order chemical reactions*

Consider the series of reactions in which a substance $\mathrm{X}_1$ breaks down according to the law of natural decay to give the substance $\mathrm{X}_2$, and $\mathrm{X}_2$ in turn decays to produce the substance $\mathrm{X}_3$ which is the end-product. Let $x_1$, $x_2$, $x_3$ denote the concentrations of $\mathrm{X}_1$, $\mathrm{X}_2$, $\mathrm{X}_3$ at time $t$ after the start of the reaction, and let the rate constants be $k_1$, $k_2$. Then $x_1$ is decreasing at the rate $k_1x_1$. $x_2$ increases because of production from $\mathrm{X}_1$, and decreases by its own decay. $x_3$ increases as it is produced from $\mathrm{X}_2$. Thus we can set down three equations representing these three relations respectively:

$$\left.\begin{aligned} \dot{x}_1 &= -k_1x_1, \\ \dot{x}_2 &= +k_1x_1 - k_2x_2, \\ \dot{x}_3 &= +k_2x_2, \end{aligned}\right\} \quad t \geqq t_0. \tag{1.2.3}$$

Here we have three simultaneous linear differential equations to determine $x_1$, $x_2$, $x_3$. We note that the first is independent of the other two and can be solved to give $x_1(t)$ if we know the original concentration $x_1(t_0)$. If we substitute the resulting function $x_1(t)$ into the second equation we can solve that for $x_2(t)$ provided we know $x_2(t_0)$. Finally this $x_2(t)$ substituted into the third equation gives $x_3(t)$ provided $x_3(t_0)$ is known. Thus the solution is unique if the data consist of the equations (3) together with the starting concentrations $x_1(t_0)$, $x_2(t_0)$, $x_3(t_0)$.

## 1.3. Initial conditions in Marching Problems

In each of the problems discussed in Sections 1.1 and 1.2 the differential equations have described the changes in the system as time proceeded, and the data in addition to the differential equations have been supplied by *initial conditions*, that is, by information about quantities and rates of change at the start of the range throughout which the differential equation or equations hold good. We now examine further the part played by the initial conditions, starting with examples.

*Example* (a). In problem (a) of Section 1.1 the equation $\mathrm{d}x/\mathrm{d}t = -kx$ was of the first order: its solution involved one integration and we introduced one arbitrary parameter. The specification of the problem was completed when one initial condition was stated, enabling this parameter to be uniquely determined. We achieved this by means of the statement that $x(t_0) = x_0$. If we had given a second condition also, such as $\dot{x}(t_0) = u_0$, it would in general have been impossible to satisfy both conditions—the problem would have been *over-determined*.

*Example* (b). In problem (c) of Section 1.1 the equation

$$\frac{\mathrm{d}^2\xi}{\mathrm{d}t^2} = \lambda\xi \tag{1.3.1}$$

was of second order. Solution of this equation would involve two integrations, and to fix the two arbitrary constants thus introduced we would need two initial conditions. These would naturally be supplied by stating the initial position and velocity

$$\xi(t_0) = \xi_0, \qquad \dot{\xi}(t_0) = \eta_0.$$

In general, the $n$th order equation

$$F(t, x, \dot{x}, \ddot{x}, \ldots, \overset{(n)}{x}) = 0, \quad t_0 \leqq t \leqq t_1, \tag{1.3.2}$$

where $\dot{x}, \ddot{x}, \ldots, \overset{(n)}{x}$ represent the first, second, . . ., $n$th derivatives of $x$ with regard to $t$, and $F$ is any continuous function of its arguments, must be integrated $n$ times to rid it of derivatives. Thus $n$ constants of integration will be introduced, and $n$ conditions are needed to determine their values. A natural method of specifying the initial state is to give

the values of the function and its first $(n-1)$ derivatives at time $t_0$. This was done in examples (a) and (b) above. But other initial conditions than these might be given. To take a simple illustration, in example (b) the initial values of $\xi+\dot{\xi}$ and $\xi-\dot{\xi}$ might be given. This would be equivalent to the specification of $\xi$ and $\dot{\xi}$ and would be equally valid. But note that if we were to give the initial values of $\xi+2\dot{\xi}$ and $2\xi+4\dot{\xi}$ the specification would be inadequate. For if the value given for the second of these expressions was just twice that for the first, the two conditions would be consistent but equivalent to one condition, while if the value of the second expression was not twice that of the first, the two would be inconsistent. Thus the problem would be either *under-determined* or *not self-consistent*. In the general case, any $n$ *independent* and *consistent* initial conditions will determine the problem correctly and completely. If the conditions are consistent but not independent, the problem is under-determined and a search must be made for one or more other conditions to complete the determination. If the conditions are inconsistent the problem has been wrongly posed.

Returning to example (b), we note that when the initial value of $\xi$ is given the differential equation fixes the initial value of $\ddot{\xi}$. Differentiating the differential equation we obtain

$$\frac{\mathrm{d}^3\xi}{\mathrm{d}t^3} = \lambda \frac{\mathrm{d}\xi}{\mathrm{d}t},$$

and if $\mathrm{d}\xi/\mathrm{d}t$ is known so is $\mathrm{d}^3\xi/\mathrm{d}t^3$. Used in this way, the differential equation determines every derivative at $t = t_0$ in terms of the initial values of $\xi$ and $\dot{\xi}$, and every coefficient is known in the expression

$$\xi(t_0)+(t-t_0)\dot{\xi}(t_0)+\tfrac{1}{2}(t-t_0)^2\ddot{\xi}(t_0)+\ldots$$

$$\ldots+\frac{1}{n!}(t-t_0)^n\overset{(n)}{\xi}(t_0)+\ldots \qquad (1.3.3)$$

But this is the Taylor expansion of $\xi(t)$ in the neighbourhood of $t = t_0$. Thus the solution is uniquely given by its Taylor expansion in the region where the series (3) converges.

Problems for which the solution is uniquely determined by the differential equation and the initial conditions are called *marching problems*. The solution marches forward from the point $t_0$, its behaviour for $t > t_0$ being determined from the conditions at the start.

## 1.4. More general boundary conditions; fully determined problems

So far we have used $t$ as independent variable, and have very naturally considered problems in which the data additional to the differential equations have been supplied by *initial conditions*. But we shall meet problems in which the independent variable is not time, but a length or an angle or some other parameter. Then it will be seen that the appropriate additional conditions may not all apply at one value of the independent variable. To emphasize the increased generality thus envisaged, we here follow the general usage and employ $x$ as independent variable and $y$ as dependent variable. We discuss the $n$th order equation

$$F(x, y, y', y'', \ldots, y^{(n)}) = 0, \quad x_0 \leqq x \leqq x_1, \tag{1.4.1}$$

where $y'$, $y''$, . . ., $y^{(n)}$ are the first, second, . . ., $n$th derivatives of $y$ with respect to $x$, and $F$ may be any continuous function of its arguments.

A *solution* or *primitive* of the equation (1) is any relation

$$y = f(x), \quad x_0 \leqq x \leqq x_1, \tag{1.4.2}$$

free from derivatives of $y$. Since such a relation must in effect be the result of integrating the equation (1) $n$ times we see that, if we have used no other information than that given in the differential equation, our solution will contain $n$ constants of integration, that is, $n$ parameters which may be fixed to satisfy $n$ further conditions. Let us write the *general solution*, in which these arbitrary parameters appear, as

$$y = f(x, a_1, a_2, \ldots, a_n). \tag{1.4.3}$$

This solution is said to have $n$ *degrees of freedom*. If we want to obtain $y$ as an exact function of $x$ free from arbitrariness, we must specify enough further conditions to determine $a_1, a_2, \ldots, a_n$. In general $n$ conditions will be needed. They may be any conditions which engender relationships between the parameters $a_1, a_2, \ldots, a_n$, but with few exceptions they refer to one or both of the end-points $x_0, x_1$ of the region of validity of the differential equation, and not to any intermediate point. They are consequently called *boundary conditions*. Initial conditions are now seen to form one restricted type of boundary conditions in which all data refer to the beginning $x_0$ of the range.

With our previous examples in mind, we must stipulate that the conditions provided for the fixing of the parameters $a_1, a_2, \ldots, a_n$ must be *independent* and *consistent*.

[The mathematician is concerned first with the question "Under what conditions does the differential equation (1) possess any solution?" and much of the pure mathematicians' study of differential equations is concerned with *existence theorems* which establish with rigour the conditions which must be imposed on $F$ in order that the equation may be soluble. The applied mathematician is usually prepared to take it for granted that a problem arising from physics or chemistry and correctly set up will possess one and only one solution: he is mainly concerned with methods of obtaining that solution. The conditions which must be obeyed by the function $F$, if there is to exist a unique solution of the equation (1) with appropriate boundary conditions, are discussed by J. C. Burkill[1] and E. L. Ince.[2]]

Problems in which the boundary conditions are imposed at both ends of the range are called *jury problems*. Since the conditions at the first end-point do not determine the problem completely, solutions of the differential equation which satisfy those conditions will retain one or more degrees of freedom. The boundary conditions at the second end-point will select from all those solutions the one that satisfies all the data: hence the name.

We now give examples of jury problems.

(a) *Equilibrium of a cylindrical film of small curvature.* Consider the surface S of separation between two liquids that do not mix. It is found by experiment that equilibrium can exist when there is a difference of pressure on the two sides of the surface S, provided the surface itself is curved. This we explain in terms of a physical hypothesis, namely that the surface of separation acts like a skin or membrane under tension. We assume further that this state of tension is such that if we describe a curve $\Gamma$ on the surface S, then the molecules of the surface on one side of $\Gamma$ exert on those on the other side of $\Gamma$ forces whose resultant across an element $\delta s$ of $\Gamma$ is a force $T\delta s$, where $T$ is a constant. The direction of the force $T\delta s$ is normal to $\delta s$ and in the tangent plane to S containing $\delta s$. This hypothesis is illustrated in Fig. 1.6(a).

For simplicity we consider a cylindrical surface of separation, the pressure on the side (1) exceeding that on the side (2) by $p$. Conditions are assumed to be identical in every plane section perpendicular to the generators of the cylinder S, and the problem is reduced from one in three dimensions to one in two dimensions. Figure 1.6(b) shows one plane section: C is the curve of intersection of S with the plane. We

take axes $Ox$, $Oy$ in this plane, and an axis $Oz$ forming with them a right-handed triad. $Oz$ is then parallel to the generators of S. Let P, with coordinates $(x, y, 0)$, be a point on C, and Q, with coordinates $(x+\delta x, y+\delta y, 0)$, a nearby point on C. Let the tangent to C at P make an angle $\psi$ with $Ox$, as shown in Fig. 1.6(b).

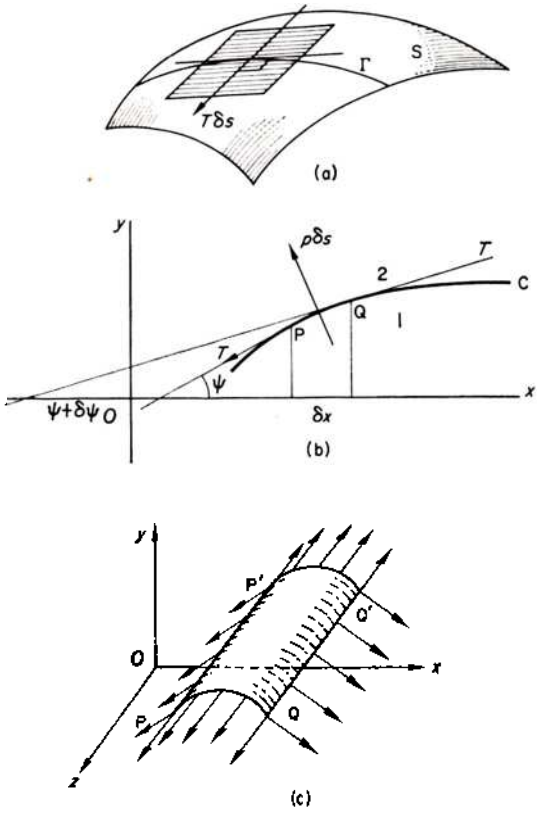


FIG. 1.6. Surface tension on a surface of separation.

We examine the equilibrium of a strip of S which stands on the element PQ as base and extends for unit distance along generators of S to the element P'Q' (parallel and congruent to PQ). Figure 1.6(c) shows this strip. For equilibrium it is necessary that the forces acting on the strip have zero resultant: thus the components of their resultant in the directions $Ox$, $Oy$, $Oz$ must vanish. The force across PQ due to surface tension is $T\delta s$ parallel to $Oz$ (where $\delta s$ is the length of the arc PQ). This

is balanced by the force $T\delta s$ in the opposite direction across P′Q′. There are no other forces parallel to $Oz$. We have therefore to set down the condition that the forces of surface tension across PP′ and QQ′ are in equilibrium with the force due to pressure-difference. By symmetry the resultants of these three will act in the plane of mid-section of the strip, where the geometry is identical with that in the plane $z = 0$.

We now take PQ so small that the pressure-difference $p$ may to a sufficient approximation be regarded as constant over PQ. Then by elementary hydrostatics the pressure-difference across PQ gives forces $-p\delta y$ and $p\delta x$ parallel to $Ox$ and $Oy$ respectively. The surface tension across PP′ (which is of unit length) gives a force $T$ with components $(-T\cos\psi, -T\sin\psi)$. Since QQ′ has abscissa $x+\delta x$, the corresponding components of surface tension across QQ′ are approximately

$$\left\{T\cos\psi + \frac{\mathrm{d}}{\mathrm{d}x}(T\cos\psi)\delta x,\ T\sin\psi + \frac{\mathrm{d}}{\mathrm{d}x}(T\sin\psi)\delta x\right\}.$$

[We have here in each case the first two terms of the Taylor expansion

$$f(x+\delta x) = f(x) + \left\{\frac{\mathrm{d}}{\mathrm{d}x}f(x)\right\}\delta x + \tfrac{1}{2}\left\{\frac{\mathrm{d}^2}{\mathrm{d}x^2}f(x)\right\}(\delta x)^2 + \ldots$$

Since we are going to take the limit as $\delta x \to 0$, it is enough here to retain only terms up to the first order in $\delta x$.]

Writing $\delta y = (\mathrm{d}y/\mathrm{d}x)\delta x = \tan\psi\ \delta x$, and equating to zero the sum of components parallel to $Ox$, we have

$$-T\cos\psi + T\cos\psi + \frac{\mathrm{d}}{\mathrm{d}x}(T\cos\psi)\delta x - p\tan\psi\ \delta x = 0,$$

i.e.

$$\frac{\mathrm{d}}{\mathrm{d}x}(T\cos\psi) - p\tan\psi = 0. \tag{1.4.4}$$

Similarly for components parallel to $Oy$,

$$\frac{\mathrm{d}}{\mathrm{d}x}(T\sin\psi) + p = 0. \tag{1.4.5}$$

Since we have assumed that $T$ is uniform these both reduce to

$$\cos\psi\,\frac{\mathrm{d}\psi}{\mathrm{d}x} + \frac{p}{T} = 0. \tag{1.4.6}$$

Now

$$\tan\psi = \frac{\mathrm{d}y}{\mathrm{d}x},$$

so that

$$\sec^2\psi \frac{\mathrm{d}\psi}{\mathrm{d}x} = \frac{\mathrm{d}^2y}{\mathrm{d}x^2},$$

and

$$\cos\psi \frac{\mathrm{d}\psi}{\mathrm{d}x} = \frac{\mathrm{d}^2y}{\mathrm{d}x^2}\cos^3\psi.$$

If we now assume that $\psi$ is everywhere small, $\sin\psi$ will be approximately equal to $\psi$, and $\cos^3\psi$, being $(1-\sin^2\psi)^{\frac{3}{2}} \doteqdot 1-\frac{3}{2}\sin^2\psi$, is approximately equal to unity. Then the equation (6) becomes

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2}+\frac{p}{T} = 0. \tag{1.4.7}$$

[The reason both components give the same equation (6) is simple. The resultant force due to surface tension must balance the resultant force due to pressure. This is an equality between two vectors. It will lead to the same equation whether we take components of the vectors parallel to $Ox$ or parallel to $Oy$. If we had resolved along the normal at the midpoint of PQ we would have obtained the same relation in the simple (intrinsic) equation $p = T\mathrm{d}\psi/\mathrm{d}s$.]

If $p$ is uniform, the equation (7) is of very simple type and integration is immediate, but if $p$ is a function of $x$ and $y$ integration is in general not straightforward and the differential equation must be subjected to detailed study. Since the equation is in any case of second order, two boundary conditions will complete the specification. We might for instance have given end-points (0, 0) and ($l$, 0). But we might have given slope, e.g. $\mathrm{d}y/\mathrm{d}x = \tan\alpha$ at $x = 0$: this would happen if we knew the angle of contact of the surface of separation at a wall, but the point of contact was to be calculated.

Note that the differential equation (7) was obtained after we had made two assumptions, one a physical hypothesis about the mechanism of surface tension, and the other a mathematical hypothesis aimed at simplification of the analysis. The differential equation therefore represents the physical situation *provided* (a) the surface of separation does in fact act as a membrane with uniform tension $T$, and (b) the slope is everywhere small enough for the square of the slope to be neglected.

(b) *Steady laminar viscous flow under a uniform pressure-gradient.* An incompressible viscous liquid flows steadily between parallel plane boundaries under a uniform pressure-gradient parallel to the boundaries. We set up the differential equation of flow. Let us take an origin on one of the bounding planes, and axes such that these planes are $y = 0$, $y = h$, and the pressure-gradient is parallel to $Ox$. $Oz$ completes the right-handed orthogonal triad. Since there is no stress to cause flow in the direction $Oz$, $\partial/\partial z \equiv 0$ and the problem is two-dimensional. We therefore consider flow of the liquid between the planes $z = 0$ and $z = 1$. Since no point on the $x$-axis differs from any other with respect to the flow-pattern, there is no variation of pattern with $x$ and $\partial/\partial x \equiv 0$. Then the velocity $(u, v, 0)$ at the point $(x, y, 0)$ must depend on $y$ only.

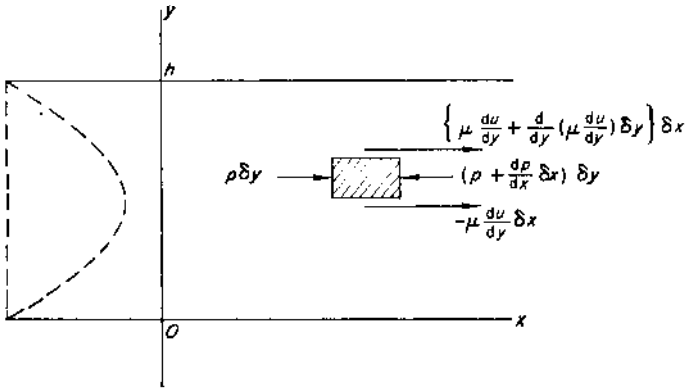


FIG. 1.7.

Two principles (or natural laws) are immediately available, without any further hypothesis about viscous flow—the law of conservation of mass and Newton's second law of motion. The first shows that $v$ must be zero. For the net flow into any cylinder parallel to $Oy$ must be zero, and therefore the component $v$ of velocity must be the same for all $y$. But $v = 0$ at $y = 0$, $y = h$ and therefore $v = 0$ for all $y$. Secondly, since the motion is steady the net force on any element of the liquid is zero. This force is partly due to the pressure-gradient and partly due to viscosity. We must therefore introduce at this point our hypotheses about the nature of viscous flow. This hypothesis is an exact formulation of the intuitive concept of viscous drag. Consider a small plane element parallel to $y = 0$ and hence parallel to the lines of flow. Then we assume that the liquid on the side of this element away from the origin exerts

on the liquid on the side nearer to the origin a tangential stress (force per unit area),

$$\mu \frac{\mathrm{d}u}{\mathrm{d}y},$$

where $\mu$ is a constant, the *viscosity*.

Now we are able to set down the condition that the net force on an element of the liquid is zero. Let the element be rectangular in cross-section, with sides $\delta x$, $\delta y$ as shown in Fig. 1.7: then the components of force on the faces of the element are as shown, and their sum vanishes. Thus

$$\left\{\mu \frac{\mathrm{d}u}{\mathrm{d}y}+\frac{\mathrm{d}}{\mathrm{d}y}\left(\mu \frac{\mathrm{d}u}{\mathrm{d}y}\right)\delta y\right\}\delta x-\mu \frac{\mathrm{d}u}{\mathrm{d}y}\,\delta x+p\delta y-\left(p+\frac{\mathrm{d}p}{\mathrm{d}x}\,\delta x\right)\delta y = 0.$$

Here we have retained only the first-order terms in the Taylor expansions of $\mu \mathrm{d}u/\mathrm{d}y$ and $p$. Hence when $\delta x \to 0$, $\delta y \to 0$ we find, in the limit

$$\frac{\mathrm{d}}{\mathrm{d}y}\left(\mu \frac{\mathrm{d}u}{\mathrm{d}y}\right)-\frac{\mathrm{d}p}{\mathrm{d}x} = 0.$$

We now take the pressure-gradient to be uniform, and (if pressure decreases as $x$ increases) we put

$$\frac{\mathrm{d}p}{\mathrm{d}x} = -G,$$

where $G$ is a positive constant. So, $\mu$ being constant,

$$\frac{\mathrm{d}^2u}{\mathrm{d}y^2}+\frac{G}{\mu} = 0.$$

This is the required differential equation. We have next to consider boundary conditions. Here we introduce a second assumption about the nature of viscous flow, namely that the tangential velocity of the liquid at a solid boundary is zero. The liquid "sticks to the bounding surface". In our problem,

$$u = 0 \quad \text{when } y = 0,\ y = h.$$

Thus we have two boundary conditions and the problem is stated in complete form.

The differential equation is easily integrated. If we incorporate the boundary conditions we obtain

$$u = \tfrac{1}{2}\frac{G}{\mu}y(h-y),$$

so that the velocity distribution is parabolic. Its form is shown by the broken line in Fig. 1.7.

*Simultaneous ordinary differential equations*

If a problem leads to simultaneous ordinary differential equations (as in examples (a) and (b) of Section 1.2) the number of boundary conditions needed for complete specification may again be discovered by counting up the number of integrations needed. For example, the set

$$\ddot{x} = my$$

$$\ddot{y} = nx$$

needs four boundary conditions. Note that we must *not* argue as follows: "By substitution,

$$\ddot{\ddot{x}} = mnx,$$

and the general solution for $x(t)$ contains four arbitrary constants. Similarly

$$\ddot{\ddot{y}} = mny,$$

and $y(t)$ contains four arbitrary constants, making eight in all." In fact, if we first solve for $x(t)$, then $y(t) = (1/m)\ddot{x}(t)$ gives $y(t)$ in terms of the same four constants as $x(t)$. No more may be introduced.

## 1.5. Choice of key variables and relations; simplification of problems

The mathematical formulation of a problem inevitably involves simplification. It will be necessary to modify the hypotheses until the problem comes within the scope of available techniques of analysis, or at least within the range of attack by numerical computation. This will mean that the mathematical theory or the numerical result does not apply strictly to the real problem, but only to the ideal problem which corresponds to the hypotheses adopted. The mathematician usually prefers to start by considering a problem in its simplest form. He then

proceeds to extend the range of application of his work by successive modification of the hypotheses to attain greater generality.

Consider for example the steady diffusion of one fluid into another along a uniform pipe which connects one reservoir with another. We first formulate the problem in the simplest way, noting carefully what assumptions are made. We suppose

(a) that the material in which diffusion occurs is *isotropic* (i.e. having the same physical properties in all directions from a point) and *homogeneous* (i.e. having the same physical constants at all points),

(b) that no chemical changes occur,

(c) that the diffusion is *isothermal* (i.e. taking place at a constant uniform temperature),

(d) that diffusion takes place parallel to the axis of the pipe in such a way that the concentration is always uniform across any cross-section.

We note that (a) and (b) simply restrict the type of materials to which our theory will apply, but (c) and (d) are hypotheses, about the actual process of diffusion, which are not immediately obvious and will only be justified by success. That is, we will only know whether the changes are isothermal by comparing results obtained on this assumption with experiment. Assumption (d) is very unlikely to be exactly true and depends on the boundary conditions and the scale of the problem. It may for instance be true for slow diffusion through a wide pipe but not for rapid diffusion through a narrow pipe.

In such problems we commonly need a physical hypothesis concerning the mechanism governing flow. Here we make it in the form

(e) that there is a direct proportionality between the local gradient of concentration and the rate of transfer.

Finally, we make an assumption concerning the continuity and conservation of the quantity under discussion. In this case it is

(f) that the total net rate of flow of the diffusing material into a given fixed volume of space is equal to the total net rate of change of the amount of that material within the volume.

If we were to consider diffusion varying with time, we would have two independent variables (time, and distance along the pipe) leading to the introduction of partial derivatives and a partial differential equation.

So for the present, in order to set up the simplest differential equation, we restrict consideration to a state of *steady flow*, in which there is no variation with time ($\partial/\partial t \equiv 0$).

We measure $x$ from a fixed point in the pipe, and consider the volume bounded by cross-sections at $x$ and $x+\delta x$. Denote the area of cross-section of the pipe by $A$, the local concentration of the diffusing fluid by $c(x)$, and the rate of transfer of diffusing fluid across unit area of cross-section by $F(x)$. Assumption (e) states

$$F(x) = -D\frac{\mathrm{d}c}{\mathrm{d}x}, \tag{1.5.1}$$

where $D$ is the *diffusion constant*. Assumption (f), applied to steady flow, gives

$$F(x)A = F(x+\delta x)A, \tag{1.5.2}$$

and, in the limit when $\delta x \to 0$,

$$\frac{\mathrm{d}F}{\mathrm{d}x} = 0.$$

Hence

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(D\frac{\mathrm{d}c}{\mathrm{d}x}\right) = 0,$$

and by hypothesis (a) this is

$$\mathrm{d}^2c/\mathrm{d}x^2 = 0. \tag{1.5.3}$$

This simple equation has the general solution

$$c = Bx+E, \tag{1.5.4}$$

where $B, E$ are constants which will be fixed by the boundary conditions. For instance, if the pipe connects a large reservoir of concentration $c_0$ at $x_0$ with another of concentration $c_1$ at $x_1$, we obtain

$$c = c_0+(c_1-c_0)\frac{x-x_0}{x_1-x_0}, \tag{1.5.5}$$

where we have assumed the reservoir so large that $c_0$ and $c_1$ are effectively constant. Here the problem is extremely elementary and the reader could probably have written down equation (5) immediately by inspection. But nevertheless the argument which has been set down at some length has a double value. First, it brings out explicitly the various

assumptions that have been made. Secondly, and following from this, it opens the way for an attack on harder problems. For instance, if we keep our assumptions except for (d), which makes the problem one-dimensional, then we need a form of (e) (relation between rate of transfer and concentration gradient) which is suitable for two- and three-dimensional problems. This is, in cartesian coordinates,

$$\left.\begin{aligned} F_x &= -D\frac{\partial c}{\partial x}, \\ F_y &= -D\frac{\partial c}{\partial y}, \\ F_z &= -D\frac{\partial c}{\partial z}, \end{aligned}\right\} \qquad (1.5.6)$$

where $F_x$, $F_y$, $F_z$ are the components of the rate of transfer $\mathbf{F}$. (Readers familiar with vectors will recognize the equations (6) as the expression in cartesian coordinates of the general vector relation

$$\mathbf{F} = -D\,\mathbf{grad}\,c.)$$

Just as the one-dimensional equation (1) has been generalized to give the three-dimensional form (6), so we must generalize the one-dimensional equation of continuity (2) to its three-dimensional equivalent. We must state in an equation the principle that in the steady state the net transfer into a rectangular box of edges $\delta x$, $\delta y$, $\delta z$ is zero. Thus

$$\begin{aligned} &F_x\delta y\delta z - \left\{F_x\delta y\delta z + \frac{\partial}{\partial x}(F_x\delta y\delta z)\delta x\right\} \\ &\quad + F_y\delta z\delta x - \left\{F_y\delta z\delta x + \frac{\partial}{\partial y}(F_y\delta z\delta x)\delta y\right\} \\ &\quad + F_z\delta x\delta y - \left\{F_z\delta x\delta y + \frac{\partial}{\partial z}(F_z\delta x\delta y)\delta z\right\} \\ &\quad + \text{higher order terms} = 0. \end{aligned}$$

Dividing by $\delta x\delta y\delta z$ and proceeding to the limit as $\delta x \to 0$, $\delta y \to 0$, $\delta z \to 0$ we obtain

$$\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} = 0. \qquad (1.5.7)$$

(This is the cartesian expression of the general vector equation

$$\operatorname{div} \mathbf{F} = 0.)$$

Combining (6) and (7) we get

$$\frac{\partial}{\partial x}\left(D\frac{\partial c}{\partial x}\right)+\frac{\partial}{\partial y}\left(D\frac{\partial c}{\partial y}\right)+\frac{\partial}{\partial z}\left(D\frac{\partial c}{\partial z}\right) = 0, \tag{1.5.8}$$

a partial differential equation which is far from trivial. [The corresponding more general vector equation is

$$\operatorname{div} (D \operatorname{\mathbf{grad}} c) = 0.]$$

When $D$ is constant it can be cancelled from (8). But if the material is isotropic but not homogeneous (so that assumption (a) is modified) we must regard $D$ as a function of position and retain it in (8).

A further generalization of the problem may be obtained by abandoning the assumption of isotropy. In a crystal, for instance, the value of $D$ will depend on the direction of flow relative to the axes of the crystal. Whereas in an isotropic body the direction of transfer is the same as that of the gradient of concentration, in an anisotropic body these directions differ in a way characteristic of the particular type of anisotropy. We write

$$\left.\begin{aligned} -F_x &= D_{11}\frac{\partial c}{\partial x}+D_{12}\frac{\partial c}{\partial y}+D_{13}\frac{\partial c}{\partial z}, \\ -F_y &= D_{21}\frac{\partial c}{\partial x}+D_{22}\frac{\partial c}{\partial y}+D_{23}\frac{\partial c}{\partial z}, \\ -F_z &= D_{31}\frac{\partial c}{\partial x}+D_{32}\frac{\partial c}{\partial y}+D_{33}\frac{\partial c}{\partial z}. \end{aligned}\right\} \tag{1.5.9}$$

Here the array $\{D_{ij}\}$, $i = 1, 2, 3, j = 1, 2, 3$ characterizes the material. When the components of $\mathbf{F}$ from (9) are substituted into (7) we obtain the differential equation governing steady flow.

The relations (1), (6) and (9) are all linear—that is, components of $\mathbf{F}$ and $\operatorname{\mathbf{grad}} c$ occur in the first degree only. This is an approximation to the true physical relation, holding only for a limited range of the variables concerned. If we wish to include in our theory conditions outside that range we must abandon the assumption of linearity.

Further generalizations of the theory could be obtained by allowing for non-steady flow, chemical reactions and thermodynamic effects, thus modifying assumptions (b) and (c).

In the above example we have shown how a particular problem may be formulated first in a very simple form, and how by successive modifications of the hypotheses more complicated and realistic problems may be treated. The solution at each stage advances from the previous one, and may often be tested by checking whether it reduces to the previous one when restrictions are re-introduced. The method is general, and enables the scientist to introduce gradually mathematical techniques of increasing difficulty. It also ensures that the mathematical formulation is closely related to the physical hypotheses at each stage: the scientist's intuition is then able to act as a check on mathematical solutions and to help to provide detailed interpretations of the mathematical equations and resulting formulae.

### 1.6. General survey of methods of solution

Let us suppose that we have been able to set up an ordinary differential equation describing the behaviour of a certain material system; simplifications and approximations will have been introduced, and if not justified at least brought into the open. We suppose further that the boundary conditions of the system have been correctly specified, so that the mathematical problem is uniquely determined. We now embark on the search for the solution. Various possibilities arise: we list methods in order, starting from those which are mathematically most complete and finishing with those purely numerical methods which cannot afford much analytical insight but which may nevertheless provide the quantitative estimates that practical applications demand.

(a) There may exist an exact solution in closed form (that is, given by a formula which does not involve an infinite series or an infinite process), or it may be possible to transform the equation into one with such a solution. Chapters 2, 3 and 4 deal with some of the equations for which this is true.

(b) It may be possible to use an integral transform technique, reducing the problem to a search through tables of transforms or to the evaluation of an integral. Chapter 4 introduces briefly the Laplace transform, which is the transform of most use in solving ordinary differential equations.

(c) There may be no solution in closed form in terms of known functions (such as hyperbolic functions, Bessel functions, error functions), but the equation may be taken (with its boundary conditions) to define a new function, the properties of which can be explored and

catalogued by discussion of the differential equation. Many of the functions which, since they have been tabulated and much studied, are now regarded as "known functions" were first introduced as solutions of differential equations.

(d) If no solution exists in terms of known functions, it may be possible to derive an infinite series (either convergent or asymptotic) to represent the solution (see Sections 5.3, 6.1).

(e) Alternatively, it may be possible to find an iterative method such that the function derived at each iteration is a better approximation than that obtained at the previous step (see Section 6.4).

(f) If all of the above methods fail, and no hope is entertained of finding a closed expression or an infinite series to represent the true solution, recourse may be had to variational methods. The solving of the equation is replaced by the search for the function that minimizes a certain integral. Instead of the correct function we may set out to find that combination of elementary functions satisfying the boundary conditions which makes the integral smallest. The combination thus found will represent the true solution as well as can be done when choice is restricted to that set of elementary functions (see Section 6.7).

(g) The method of (f) may be inapplicable, and direct recourse to numerical methods may be needed. Then all of the parameters that occur in the problem must be given numerical values and either the differential equation replaced by difference equations or the process of integration replaced by summation (see Section 6.9).

(h) It is sometimes possible to find a second chemical or physical system which leads to the same differential equation and boundary conditions. The second system (for example a system of electronic circuits) may be used as an analogue if the dependent variable, that is, the unknown function in the first system, can be found by measurement or exhibited directly (as on an oscilloscope) in the second system.

### 1.7. Discussion of ordinary differential equations by classes

It is to be expected that differential equations of low order and low degree will have been found most amenable to solution and will have been most studied. We here list these classes, with references to the sections of this book or of other books where methods of solution are given. Equations of little interest in physical chemistry will not be treated in this book.

(1) First order, first degree
  (a) variables separable (Chapter 2)
  (b) variables separable after substitution
    (Examples at end of Chapter 2)
    Ince.[2] Chapter 2
  (c) Solution by iteration (Section 6.1)
  (d) Graphical method (Section 6.1)
  (e) Numerical methods (Section 6.4)
  (f) The linear equation

$$\frac{dx}{dt}+p(t)x = f(t) \quad \text{(Section 5.2).}$$

(2) First order, higher degree
    McLachlan.[3]

(3) Second order
  (a) Linear equations with constant coefficients

$$\frac{d^2x}{dt^2}+a\frac{dx}{dt}+bx = f(t) \quad \text{(Chapters 3, 4)}$$

  (b) Linear equations with variable coefficients

$$\frac{d^2x}{dt^2}+p(t)\frac{dx}{dt}+q(t)x = f(t) \quad \text{(Chapter 5)}$$

  (c) Non-linear equations
    McLachlan.[3]

(4) Higher order
  (a) Linear equations with constant coefficients (Chapter 4)
  (b) Linear equations with variable coefficients
    Ince.[2] Chapter 5.

Two books are recommended as dictionaries of differential equations; in these books will be found extensive lists of equations of which the solutions are available.

Kamke, E.:[4] *Differentialgleichungen, Lösungsmethoden und Lösungen*, Vol. 1: *Gewohnliche Differentialgleichungen.* (In German.)

Murphy, G. M.:[5] *Ordinary Differential Equations and their Solutions.*

Kamke catalogues over a thousand differential equations with their solutions and references to books or articles where the solutions are

derived. Murphy's book lists 2315 equations, with indexes and cross-references. A reader who on meeting a particular ordinary differential equation wishes first to see whether a standard solution exists should consult one of these two books.

## REFERENCES

1. BURKILL, J. C., *Theory of Ordinary Differential Equations*, 1956, Ch. 1.
2. INCE, E. L., *Ordinary Differential Equations*, 1927, Ch. 3.
3. MCLACHLAN, N. W., *Ordinary Non-linear Differential Equations in Engineering and Physical Sciences*, 2nd edn., 1956.
4. KAMKE, E., *Differentialgleichungen, Lösungsmethoden und Lösungen*, Leipzig, 1944.
5. MURPHY, G. M., *Ordinary Differential Equations and their Solutions*. New York, 1960.

## Exercises

**1.1.** A test tube contains liquid and, above the liquid, a mixture of vapour and air. The vapour diffuses upward, from saturation vapour pressure at the surface of the liquid to zero vapour pressure at the top of the tube where a current of air blows past. Set up the differential equation for vapour pressure as a function of position in the tube, with the boundary conditions. State explicitly the assumptions on which the equation is derived.

**1.2.** Viscous liquid flows in a cylindrical pipe of radius $R$, under a pressure gradient $G$. Taking distance from the axis of the pipe as $r$, obtain the differential equation of flow in the form:

$$\frac{\mathrm{d}^2v}{\mathrm{d}r^2}+\frac{1}{r}\frac{\mathrm{d}v}{\mathrm{d}r}+\frac{G}{\mu}=0, \quad 0 \leqq r < R.$$

[Note that here it is necessary to consider an element bounded by radii which include an angle $\delta\theta$ and by circular segments of radii $r$ and $r+\delta r$. The force across the inner arc being

$$\mu\frac{\partial v}{\partial r}r\delta\theta,$$

that across the outer arc is

$$\mu\frac{\partial v}{\partial r}r\delta\theta+\frac{\partial}{\partial r}\left(\mu\frac{\partial v}{\partial r}r\delta\theta\right)\delta r.]$$

**1.3.** A simple circuit contains condenser of capacity $C$, resistance $R$, coil of self-inductance $L$ and battery of electromotive force $E(t)$. Obtain the differential equation satisfied by the current in the circuit, and initial conditions when it is switched on. Formulate explicitly the assumptions that have been made.

**1.4.** It is assumed that when an elastic beam is slightly bent, so that a point on the central axis of the beam at distance $x$ from one end is displaced through distance $y$, the bending moment at $x$ is $M = B(\mathrm{d}^2y/\mathrm{d}x^2)$, where $B$ is a constant (the *flexural rigidity*). Prove that if the beam is subjected to a non-uniform load $w(x)$ the equation of equilibrium is $\mathrm{d}^2M/\mathrm{d}x^2 = w(x)$, and hence that the shape of the distorted axis is given by the solution of

$$B\frac{\mathrm{d}^4y}{\mathrm{d}x^4}=w(x),$$

with appropriate boundary conditions. Find the boundary conditions at (a) a supported unclamped end, (b) a clamped end, (c) a free end.

**1.5.** A unimolecular reaction is followed by a bimolecular reaction. In the first material A decays according to the law of natural decay to form B. B reacts with C at a rate proportional to the product of the concentrations of B and C. With a suitable choice of variables and constants obtain the differential equations of the process:

$$\frac{dx_1}{dt} = -k_1x_1,$$

$$\frac{dx_2}{dt} = k_1x_1 - k_2x_2x_3,$$

$$\frac{dx_3}{dt} = k_2x_2x_3.$$

State clearly the assumptions made, and formulate suitable initial conditions.

**1.6.** A charged particle of mass $m$ and charge $q$ moves in an electromagnetic field where the electric intensity at any point is $\mathbf{E}$ and the magnetic intensity $\mathbf{H}$. Assuming that the force exerted on the particle due to the magnetic field is equal to $(q/c)\ \mathbf{v}\times\mathbf{H}$, where $\mathbf{v}$ is the velocity of the particle and $\times$ denotes the vector product, show that the equation of motion of the particle is

$$\frac{d}{dt}m\mathbf{v} = q\mathbf{E} + \frac{q}{c}\mathbf{v}\times\mathbf{H}.$$

Taking rectangular cartesian coordinate axes, write out the three components of this equation.

CHAPTER 2

# ORDINARY DIFFERENTIAL EQUATIONS OF THE FIRST ORDER WHICH ARE SEPARABLE OR EXACT

## 2.1. Separable equations

The differential equations which represent simple chemical reactions are very often of the kind called *separable*. A separable equation is one that can be rearranged to bring the independent variable and its differential to one side of the equation and the dependent variable with its differential to the other. Thus the equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x)g(y), \tag{2.1.1}$$

where $f(x)$ is a given function of $x$ only, and $g(y)$ of $y$ only, can be written as

$$\frac{\mathrm{d}y}{g(y)} = f(x)\mathrm{d}x,$$

and if $y = y_0$ when $x = x_0$ we have

$$\int_{y_0}^{y} \frac{\mathrm{d}y}{g(y)} = \int_{x_0}^{x} f(x)\mathrm{d}x. \tag{2.1.2}$$

When an equation like (2) has been obtained, the problem is said to have been "reduced to quadrature", i.e. reduced to straightforward integration. The differential equation is then regarded as solved, since methods of exact or approximate evaluation of integrals are available. Methods of exact evaluation of those indefinite integrals which can be evaluated in closed form (i.e. in a finite number of terms) are given in standard texts on calculus such as

Gillespie, R. P., *Integration*,[1]
Stewart, C. A., *Advanced Calculus*.[2]

A very useful list of formulae for evaluation of common integrals is given in

Peirce, B. O., *A Short Table of Integrals.*[3]

The following book also contains a list of useful integrals:

*Handbook of Chemistry and Physics.*[4]

Numerical methods of integration are explained in

Hartree, D. R., *Numerical Analysis,*[5]

Buckingham, R. A., *Numerical Methods.*[6]

*Equations separable after transformation*

In some equations the variables are not separable as they stand, but a change of variable will produce a separable equation. The most important equation that can be made separable is the homogeneous equation, which can be written (after suitable rearrangement)

$$\frac{dy}{dx} = F\left(\frac{y}{x}\right), \tag{2.1.3}$$

where $F$ is a known function, and $x$ and $y$ occur in $F$ only in the combination $y/x$.

To separate the variables, we introduce a new dependent variable

$$v = y/x,$$

so that

$$y = vx$$

and

$$\frac{dy}{dx} = v + x\frac{dv}{dx}.$$

The differential equation thus becomes

$$v + x\frac{dv}{dx} = F(v),$$

or

$$\frac{dx}{x} = \frac{dv}{F(v) - v}, \tag{2.1.4}$$

and the problem has been reduced to quadrature.

*Example:* A point moves on a curve in the plane $Oxy$ in such a way that the angle made by the tangent with $Ox$ is three times the angle between the radius vector and $Ox$. Find the cartesian equation of the family of curves satisfying this condition.

With the notation of Fig. 2.1(a), $\psi = 3\theta$, so that

$$\begin{aligned}\frac{dy}{dx} &= \tan\psi \\ &= \tan 3\theta \\ &= \tan\theta\,\frac{3-\tan^2\theta}{1-3\tan^2\theta} \\ &= \frac{y}{x}\,\frac{3x^2-y^2}{x^2-3y^2}.\end{aligned} \tag{2.1.5}$$

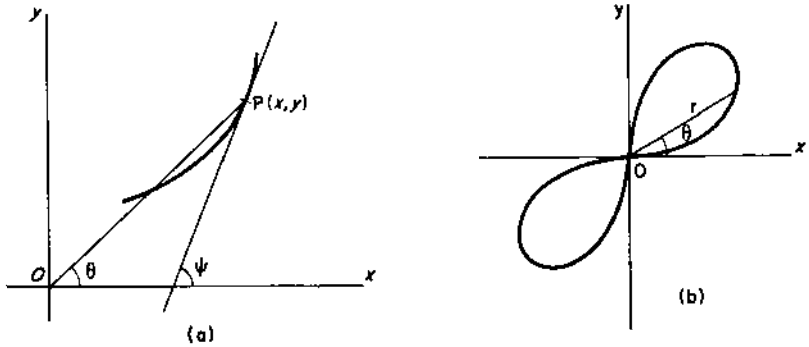


FIG. 2.1.

Putting $y = vx$ we obtain

$$v + x\frac{dv}{dx} = v\,\frac{3-v^2}{1-3v^2},$$

and

$$x\frac{dv}{dx} = \frac{2v(1+v^2)}{1-3v^2}.$$

Then

$$2\frac{dx}{x} = \frac{1-3v^2}{v(1+v^2)}\,dv. \tag{2.1.6}$$

The right-hand side is integrated by the use of partial fractions. Put

$$\frac{1-3v^2}{v(1+v^2)} = \frac{A}{v} + \frac{B+Cv}{1+v^2}.$$

Then $1-3v^2 = A(1+v^2)+(B+Cv)v$, for a range of $v$. Hence, equating coefficients, we have

$$A = 1, \quad B = 0, \quad C = -4,$$

and

$$2\frac{\mathrm{d}x}{x} = \frac{\mathrm{d}v}{v} - \frac{4v\mathrm{d}v}{1+v^2}. \tag{2.1.7}$$

Integrating, we have

$$\ln x^2 = \ln v - 2\ln(1+v^2) + \ln \lambda,$$

where $\lambda$ is an arbitrary constant. (Here ln is used to denote the logarithm to the base e: log will be used to denote a logarithm to the base 10, i.e. $\ln N \equiv \log_e N$ and $\log N \equiv \log_{10} N$.) Thus the solution of (5) is

$$x^2 = \frac{\lambda v}{(1+v^2)^2} = \frac{\lambda x^3 y}{(x^2+y^2)^2}$$

and the required family of curves is

$$(x^2+y^2)^2 = \lambda xy. \tag{2.1.8}$$

The form of these curves is best seen by writing (8) in polar coordinates, so that we have

$$r^2 = \tfrac{1}{2}\lambda \sin 2\theta.$$

A typical member of the family is shown in Fig. 2.1(b). The curves are Bernoulli's *lemniscates*.

[Note that this is one of a wide range of problems specified by $\tan\psi = f(\tan\theta)$. Some of those problems might be solved by *ad hoc* geometrical methods (e.g. that in which $\psi = 2\theta$). But the solution by way of the differential equation is to be preferred as *more general* and *more powerful*.]

## 2.2. Rates of chemical reactions in simple kinetic systems

Let $A_1$, $A_2$, ... be reacting molecules in a single stage chemical reaction of which the products are $A_1'$, $A_2'$, ..., the stoichiometric relation being

$$n_1A_1 + n_2A_2 + \ldots \rightarrow n_1'A_1' + n_2'A_2' + \ldots. \tag{2.2.1}$$

Let $c_1, c_2, \ldots, c_1', c_2', \ldots$ be the concentrations† of $A_1, A_2, \ldots, A_1', A_2', \ldots$ at time $t$ after the start of the reaction. We expect that as the reaction proceeds the concentration $c_1$ will decrease steadily; the rate of decrease will be $(-dc_1/dt)$. Since $n_1$ molecules of $A_1$ combine with $n_2$ molecules of $A_2, \ldots$, to produce $n_1'$ molecules of $A_1'$, $n_2'$ molecules of $A_2', \ldots$, we see that

$$\begin{aligned} -\frac{1}{n_1}\frac{dc_1}{dt} &= -\frac{1}{n_2}\frac{dc_2}{dt} = \ldots \\ &= \frac{1}{n_1'}\frac{dc_1'}{dt} = \frac{1}{n_2'}\frac{dc_2'}{dt} = \ldots \\ &= R. \end{aligned} \tag{2.2.2}$$

This quantity $R$ we will call the *rate of advancement of the reaction.*

[Chemists commonly use the phrase "rate of reaction" to describe $dc_1/dt, dc_2/dt, \ldots$] The above argument is stated in terms of reacting molecules, and (1) and (2) must be regarded as describing the overall result of molecular processes. We might state the argument in terms of concentrations measured in moles per litre; then $n_1, n_2, \ldots$ would denote numbers of moles, and (1) and (2) would be unchanged.

It is found by experiment that within certain restrictions there exists for a given single stage reaction a relation between $(-dc_1/dt)$ and the concentrations of the various molecules present in the reaction. We first consider reactions where this relation is one between concentrations of the reacting substances $A_1, A_2, \ldots$ only. From the experiments we are led to seek relations of the form

$$\frac{dc_1}{dt} = -k^{(1)}c_1^{\alpha_1}c_2^{\alpha_2}\ldots, \tag{2.2.3}$$

where $\alpha_1, \alpha_2, \ldots$ are small integers, and $k^{(1)}$ is the *rate constant* for the substance $A_1$. We shall see that the equation (3) leads to a separable differential equation. The reaction which gives rise to equation (3) is said to be *of order* $\alpha_1$ with regard to $A_1$, of order $\alpha_2$ with regard to $A_2, \ldots$ and to be overall of order $\alpha_1 + \alpha_2 + \ldots$. Normally each $\alpha$ takes

† The *concentration* of a substance is defined as the number of molecules/cm³, and its rate of change is measured in molecules/cm³ per sec. Alternatively, concentration may be given in moles/litre. When reactions take place in solution the concentrations are commonly given in moles/litre. The concentration of a substance A is commonly denoted by [A], but this notation will not be used much here, as it is rather inconvenient in mathematical manipulations.

one of the values 0, 1, 2, 3, and the number of reacting substances is small.

(a) *Zero order* reactions are occasionally observed: for such the differential equation (3) is (using $c$ for $c_1$)

$$\frac{dc}{dt} = -k_0, \quad c = c_0 \text{ when } t = 0, \tag{2.2.4}$$

with an immediate integral $c = c_0 - k_0 t$, where $c_0$ is the initial concentration.

(b) The *first order* reaction for which

$$\frac{dc}{dt} = -k_1 c, \text{ and } c = c_0 \text{ when } t = 0, \tag{2.2.5}$$

gives the law of natural decay, which was discussed in Section 1. Separating the variables in (5) we have

$$\frac{dc}{c} = -k_1 dt,$$

giving

$$\ln c = -k_1 t + A,$$

where

$$\ln c_0 = A,$$

by the initial condition.

Hence

$$\ln c - \ln c_0 = -k_1 t \tag{2.2.6}$$

or

$$c = c_0 e^{-k_1 t}. \tag{2.2.7}$$

The form (6) is convenient for graphical work; $-k_1$ is the slope of the curve of $\ln c$ graphed against $t$.

The *half-life* $\tau_{\frac{1}{2}}$ is defined as the time taken for $c$ to fall to half the initial value $c_0$.

Thus

$$\ln (\tfrac{1}{2} c_0) - \ln c_0 = -k_1 \tau_{\frac{1}{2}},$$

giving

$$\tau_{\frac{1}{2}} = \frac{1}{k_1} \ln 2.$$

(c) The *second order* reaction in which one substance only is involved gives

$$\frac{dc}{dt} = -k_2 c^2, \quad c = c_0 \text{ when } t = 0, \tag{2.2.8}$$

which may be rewritten

$$-\frac{dc}{c^2} = k_2 dt, \tag{2.2.9}$$

so that

$$\frac{1}{c} - \frac{1}{c_0} = k_2 t. \tag{2.2.10}$$

(d) The *second order* reaction in which two substances are involved leads to

$$\frac{dc_1}{dt} = -k_2 c_1 c_2, \text{ where at } t = 0,\ c_1 = c_{10},\ c_2 = c_{20}. \tag{2.2.11}$$

Suppose that, in time $t$, $c_1$ decreases by $x$ (molecules/cm$^3$). Then $c_2$ will decrease by $(n_2/n_1)x$.

Thus

$$c_1 = c_{10} - x,$$

$$c_2 = c_{20} - (n_2/n_1)x,$$

and

$$\frac{dx}{dt} = k_2(c_{10} - x)\left(c_{20} - \frac{n_2}{n_1}x\right). \tag{2.2.12}$$

Let us for convenience rewrite this equation, by setting

$$c_{10} = a, \quad c_{20} = \frac{n_2}{n_1}b, \quad k_2 = \frac{n_1}{n_2}k,$$

as

$$\frac{dx}{dt} = k(a - x)(b - x). \tag{2.2.13}$$

$a$ may be greater than or less than $b$: let us assume without loss of generality that $a > b$. Separating the variables in (13), we have

$$k dt = \frac{dx}{(a - x)(b - x)}. \tag{2.2.14}$$

The right-hand side may be integrated by use of partial fractions. We seek $A$ and $B$ such that

$$\frac{1}{(a-x)(b-x)} = \frac{A}{a-x} + \frac{B}{b-x}.$$

This gives

$$1 = A(b-x) + B(a-x)$$

for all $x$. Putting $x = a$ we have $A = -1/(a-b)$, and putting $x = b$ we have $B = 1/(a-b)$. Hence (14) becomes

$$k\mathrm{d}t = -\frac{1}{a-b}\frac{\mathrm{d}x}{a-x} + \frac{1}{a-b}\frac{\mathrm{d}x}{b-x}.$$

When $t = 0$, $x = 0$; incorporating this condition as we integrate we get

$$\begin{aligned} kt &= \frac{1}{a-b}\{\ln(a-x) - \ln a\} \\ &\quad - \frac{1}{a-b}\{\ln(b-x) - \ln b\} \\ &= \frac{1}{a-b}\ln\frac{b(a-x)}{a(b-x)}, \end{aligned}$$

or

$$\frac{b(a-x)}{a(b-x)} = \exp(a-b)kt.$$

Putting back the original constants, we obtain

$$\ln\frac{1-x/c_{10}}{1-n_2x/n_1c_{20}} = k_2\left(\frac{n_2}{n_1}c_{10} - c_{20}\right)t.$$

If we write this as

$$\ln\frac{c_{10}-x}{c_{20}-\frac{n_2}{n_1}x} - \ln\frac{c_{10}}{c_{20}} = \frac{n_2c_{10}-n_1c_{20}}{n_1}k_2t,$$

the left-hand side is the difference between the values at times $t$ and zero of the logarithms of the ratio of the concentrations of $A_1$ and $A_2$. When

the logarithm of the ratio of concentrations is plotted against time, the slope of the straight line thus obtained is

$$\frac{n_2c_{10}-n_1c_{20}}{n_1}k_2,$$

so that $k_2$ may be found if the stoichiometric equation and the initial concentrations are known.

## 2.3. Opposing or concurrent chemical reactions

If the products of a reaction themselves react to form the original substances we have *opposing reactions*. If the reacting substances may combine in more than one way to form products we have *concurrent reactions*.

(a) *Opposing reactions of the first order*

Let A break down with rate constant $k_1$ to form A′ while A′ breaks down with rate constant $k_1'$ to form A again. Let the concentrations of A, A′ at $t = 0$ be $c_0$, $c_0'$, and at a subsequent time $t$ be $c$, $c'$. Then if both reactions are of the first order we can write the differential equation of the opposing reactions as

$$\frac{dc}{dt} = -k_1c+k_1'c'. \tag{2.3.1}$$

Here $k_1c$ is the rate of decrease of the concentration $c$ of A by intrinsic breakdown, while $k_1'c'$ is the rate of increase of $c$ by first order production from A′. By the conservation of matter

$$c+c' = c_0+c_0'. \tag{2.3.2}$$

Substituting $c'$ from (2) into (1) we have

$$\frac{dc}{dt} = k_1'(c_0+c_0')-(k_1+k_1')c$$

or

$$dt = \frac{dc}{k_1'(c_0+c_0')-(k_1+k_1')c}, \tag{2.3.3}$$

with $c = c_0$ when $t = 0$.

Hence

$$t = -\frac{1}{k_1+k_1'}\begin{bmatrix}\ln\{k_1'(c_0+c_0')-(k_1+k_1')c\} \\ -\ln\{k_1'(c_0+c_0')-(k_1+k_1')c_0\}\end{bmatrix}$$

or

$$-(k_1+k_1')t = \ln\frac{k_1'c'-k_1c}{k_1'c_0'-k_1c_0}. \tag{2.3.4}$$

This may also be expressed as

$$k_1'c'-k_1c = (k_1'c_0'-k_1c_0)\exp\{-(k_1+k_1')t\}. \tag{2.3.5}$$

From (5) we see that as $t \to \infty$, $k_1'c'-k_1c \to 0$. Thus as time advances the ratio of the concentrations tends to a constant value—the *equilibrium constant* $k_1/k_1'$, i.e.

$$\frac{c'}{c} \to \frac{k_1}{k_1'} \text{ as } t \to \infty.$$

(b) *Opposing reactions of the second order*

Let $A_1$ and $A_2$ react to produce $A_1'$ and $A_2'$, which in turn react to produce $A_1$ and $A_2$. The opposing reactions are represented by

$$A_1+A_2 \rightleftharpoons A_1'+A_2'.$$

In this process the concentrations of $A_1$ and $A_2$ each fall by $x$ while those of $A_1'$ and $A_2'$ each rise by $x$. Thus after time $t$ the concentrations have changed from the initial values $c_{10}$, $c_{20}$, $c_{10}'$, $c_{20}'$ to $c_{10}-x$, $c_{20}-x$, $c_{10}'+x$, $c_{20}'+x$. If each reaction is a simple second order one, both reagents appearing in the first order, the differential equation for $c_1$ is

$$\frac{dc_1}{dt} = -k_2c_1c_2+k_2'c_1'c_2'. \tag{2.3.6}$$

Substituting $c_1 = c_{10}-x$, etc., we obtain

$$\begin{aligned}\frac{dx}{dt} &= k_2(c_{10}-x)(c_{20}-x)-k_2'(c_{10}'+x)(c_{20}'+x) \\ &= px^2+qx+r,\end{aligned} \tag{2.3.7}$$

where

$$\left.\begin{aligned} p &= k_2-k_2', \\ q &= -k_2(c_{10}+c_{20})-k_2'(c_{10}'+c_{20}'), \\ r &= k_2c_{10}c_{20}-k_2'c_{10}'c_{20}'. \end{aligned}\right\} \tag{2.3.8}$$

Separating the variables in (7) we have

$$\mathrm{d}t = \frac{\mathrm{d}x}{px^2+qx+r}. \tag{2.3.9}$$

The form of the right-hand side of (9) after integration depends on whether the denominator can be factorized into real factors or not. There will be real factors if

$$px^2+qx+r = 0 \tag{2.3.10}$$

has real roots. Now the roots are

$$x_1, x_2 = \frac{-q \pm \sqrt{(q^2-4pr)}}{2p}, \tag{2.3.11}$$

and these are real if $q^2-4pr \geqq 0$. Substituting from (8) we obtain

$$\begin{aligned} q^2-4pr &= k_2^2(c_{10}+c_{20})^2+2k_2k_2'(c_{10}+c_{20})(c_{10}'+c_{20}')+k_2'^2(c_{10}'+c_{20}')^2 \\ &\quad -4\{k_2^2c_{10}c_{20}-k_2k_2'(c_{10}c_{20}+c_{10}'c_{20}')+k_2'^2c_{10}'c_{20}'\} \\ &= k_2^2(c_{10}-c_{20})^2 \\ &\quad +2k_2k_2'\{(c_{10}+c_{20})(c_{10}'+c_{20}')+2(c_{10}c_{20}+c_{10}'c_{20}')\} \\ &\quad +k_2'^2(c_{10}'-c_{20}')^2 \\ &> 0, \end{aligned} \tag{2.3.12}$$

since all terms are positive.

We therefore write $px^2+qx+r = p(x_1-x)(x_2-x)$, and integrate by use of partial fractions. (We write the factors in this way rather than as $(x-x_1)(x-x_2)$ because we know that the process begins at $x = 0$ and we wish to avoid writing logarithms of negative quantities.) Then

$$p\mathrm{d}t = \frac{1}{x_2-x_1}\frac{\mathrm{d}x}{x_1-x}+\frac{1}{x_1-x_2}\frac{\mathrm{d}x}{x_2-x}.$$

Take $x_1$ to be the greater root. Then since $x = 0$ when $t = 0$,

$$\begin{aligned} (x_1-x_2)pt &= \ln(x_1-x)-\ln x_1-\{\ln(x_2-x)-\ln x_2\} \\ &= \ln\frac{(x_1-x)x_2}{(x_2-x)x_1}, \end{aligned}$$

or

$$\frac{x_1-x}{x_2-x} = \frac{x_1}{x_2}\exp(x_1-x_2)pt, \tag{2.3.13}$$

where

$$x_1 = \frac{1}{2p}\{-q+\sqrt{(q^2-4pr)}\},$$

$$x_2 = \frac{1}{2p}\{-q-\sqrt{(q^2-4pr)}\},$$

so that

$$\frac{2px+q-\sqrt{(q^2-4pr)}}{2px+q+\sqrt{(q^2-4pr)}} = \frac{q-\sqrt{(q^2-4pr)}}{q+\sqrt{(q^2-4pr)}} \exp t\sqrt{(q^2-4pr)}. \tag{2.3.14}$$

(c) *Concurrent first and second order reactions in decomposition of a substance into the same final product*

We have to solve

$$\frac{dc}{dt} = -k_1c-k_2c^2, \text{ with } c = c_0 \text{ when } t = 0. \tag{2.3.15}$$

Separating the variables, and writing $k_2/k_1 = \mu$,

$$-k_1dt = \frac{dc}{c(1+\mu c)}$$

$$= dc\left(\frac{1}{c}-\frac{\mu}{1+\mu c}\right).$$

Integrating, and incorporating the initial condition,

$$-k_1t = (\ln c-\ln c_0)-\{\ln(1+\mu c)-\ln(1+\mu c_0)\}$$

$$= \ln\left(\frac{c}{1+\mu c}\frac{1+\mu c_0}{c_0}\right),$$

so that

$$\frac{c}{1+\mu c} = \frac{c_0}{1+\mu c_0}e^{-k_1t}. \tag{2.3.16}$$

This relation between $c$ and $t$ may be rearranged to express either $c/c_0$ or $(c_0-c)/c_0$ as a function of $t$. For

$$c(1+\mu c_0) = c_0(1+\mu c)e^{-k_1t},$$

giving

$$\frac{c}{c_0} = \frac{e^{-k_1 t}}{1+\mu c_0(1-e^{-k_1 t})} \tag{2.3.17}$$

and

$$\frac{c_0-c}{c_0} = \frac{1+\mu c_0(1-e^{-k_1 t})-e^{-k_1 t}}{1+\mu c_0(1-e^{-k_1 t})}$$

$$= \frac{1-e^{-k_1 t}}{1-\nu e^{-k_1 t}}, \tag{2.3.18}$$

where

$$\nu = \frac{\mu c_0}{1+\mu c_0} = \frac{k_2 c_0}{k_1+k_2 c_0}.$$

Other examples of opposing and concurrent reactions will be found in the Exercises at the end of this chapter.

## 2.4. Exact equations

If the equation

$$\frac{dy}{dx} = F(x, y), \quad y = y_0 \text{ when } x = x_0, \tag{2.4.1}$$

is not separable as it stands, nor separable because $F$ can be expressed as a function of $y/x$, it may still be integrable. We rewrite the equation:

$$P(x, y)dx + Q(x, y)dy = 0. \tag{2.4.2}$$

*If* we can identify the left-hand side of (2) as the differential of some function of $x$ and $y$, say $U(x, y)$, then we can write (2) as

$$dU(x, y) = 0, \tag{2.4.3}$$

and we can integrate to obtain

$$U(x, y) = C \text{ (constant)}. \tag{2.4.4}$$

When this process can be carried out, $P(x, y)dx + Q(x, y)dy$ is said to be an *exact differential*, and the equation (2) is said to be *exact*.

We will obtain a *necessary and sufficient* condition for (2) to be exact. First we derive a *necessary* condition. If the equation (4) is the integral of (2), then since (4) implies (on differentiating)

$$\frac{\partial U}{\partial x}dx + \frac{\partial U}{\partial y}dy = 0, \tag{2.4.5}$$

this equation (5) must be identified with (2) for arbitrary $dx$, $dy$. This is possible only if

$$P = \lambda \frac{\partial U}{\partial x}, \qquad Q = \lambda \frac{\partial U}{\partial y},$$

where $\lambda$ is constant, and hence

$$\frac{\partial P}{\partial y} = \lambda \frac{\partial^2 U}{\partial y \partial x} = \lambda \frac{\partial^2 U}{\partial x \partial y} = \frac{\partial Q}{\partial x}. \tag{2.4.6}$$

The condition

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

is therefore *necessary* for (2) to be integrable: we next show that it is also sufficient, and in showing that we develop a method by which the unknown $U(x, y)$ may be discovered.

Suppose it is verified that in a certain case $\partial P/\partial y = \partial Q/\partial x$. We have to show that the function $U(x, y)$ will then exist. We first define a function $V(x, y)$ by

$$V(x, y) = \int_{x_0}^{x} P(x, y)dx + R(y), \tag{2.4.7}$$

where we have taken the lower limit of integration at the initial value $x_0$ (though any other value within the range where (2) holds could be used), and $R(y)$ is a function of $y$ only, which is subsequently to be determined. Then

$$V(x, y) = \text{constant} \tag{2.4.8}$$

is a solution of (2) if

$$\frac{\partial V}{\partial x} dx + \frac{\partial V}{\partial y} dy = 0$$

is identical with (2). But our method of construction of $V$ ensures that $\partial V/\partial x = P(x, y)$. $\partial V/\partial y$ must therefore equal $Q(x, y)$. Thus

$$Q(x, y) = \int_{x_0}^{x} \frac{\partial P}{\partial y} dx + \frac{\partial R}{\partial y},$$

from (7),

$$= \int_{x_0}^{x} \frac{\partial Q}{\partial x} dx + \frac{dR}{dy},$$

by (6) and since $R(y)$ is a function of $y$ only, giving

$$\frac{\mathrm{d}R}{\mathrm{d}y} = Q(x_0, y).$$

Hence, integrating from $y_0$,

$$R(y) - R(y_0) = \int_{y_0}^{y} Q(x_0, y)\mathrm{d}y, \tag{2.4.9}$$

and (8) becomes

$$\int_{x_0}^{x} P(x, y)\mathrm{d}x + \int_{y_0}^{y} Q(x_0, y)\mathrm{d}y = \text{constant}, \tag{2.4.10}$$

where $R(y_0)$ has been absorbed into the constant on the right-hand side. (10) is the general solution of the differential equation, $y$ being given implicitly as a function of $x$. When we insert the condition $y = y_0$ when $x = x_0$ we get the value zero for the arbitrary constant. Solving the equation

$$\int_{x_0}^{x} P(x, y)\mathrm{d}x + \int_{y_0}^{y} Q(x_0, y)\mathrm{d}y = 0$$

for $y$ in terms of $x$ we get the particular solution $y(x)$ that satisfies the boundary condition $y = y_0$ when $x = x_0$.

We have thus shown how to integrate (2) when (6) is given. In other words, the condition (6)

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

is *sufficient* for the integrability of the equation (2). It is therefore the *necessary and sufficient* condition for

$$P(x, y)\mathrm{d}x + Q(x, y)\mathrm{d}y = 0$$

to be exact.

*Example* (a). To show that $(x+ky)\mathrm{d}x + (y-kx)\mathrm{d}y = 0$ is not exact, but that it becomes exact if multiplied by $(x^2+y^2)^{-1}$; and to find the solution for which $y = 0$ when $x = 1$. Since

$$\frac{\partial}{\partial y}(x+ky) = k, \qquad \frac{\partial}{\partial x}(y-kx) = -k,$$

the equation is not exact as it stands. But

$$\frac{\partial}{\partial y}\{(x+ky)(x^2+y^2)^{-1}\} = k(x^2+y^2)^{-1}-(x+ky)2y(x^2+y^2)^{-2}$$
$$= (x^2+y^2)^{-2}(kx^2-2xy-ky^2),$$

$$\frac{\partial}{\partial x}\{(y-kx)(x^2+y^2)^{-1}\} = -k(x^2+y^2)^{-1}-(y-kx)2x(x^2+y^2)^{-2}$$
$$= (x^2+y^2)^{-2}(kx^2-2xy-ky^2),$$

so that the modified equation is exact.

By (10) the solution is

$$\int_1^x \frac{x+ky}{x^2+y^2}\,dx + \int_0^y \left(\frac{y-kx}{x^2+y^2}\right)_{x=1} dy = 0,$$

i.e.

$$\left[\tfrac{1}{2}\ln(x^2+y^2)\right]_1^x + \left[k\tan^{-1}\frac{x}{y}\right]_1^x + \left[\tfrac{1}{2}\ln(1+y^2)\right]_0^y - \left[k\tan^{-1}y\right]_0^y = 0,$$

i.e.

$$\tfrac{1}{2}\ln\left\{\frac{x^2+y^2}{1+y^2}(1+y^2)\right\} + k\left(\tan^{-1}\frac{x}{y} - \tan^{-1}\frac{1}{y} - \tan^{-1}y\right) = 0.$$

But

$$\tan^{-1}\frac{x}{y} = \frac{\pi}{2} - \tan^{-1}\frac{y}{x}, \qquad \tan^{-1}y + \tan^{-1}\frac{1}{y} = \frac{\pi}{2},$$

and so

$$\tfrac{1}{2}\ln(x^2+y^2) - k\tan^{-1}\frac{y}{x} = 0$$

is the required solution.

*Example* (b). If the equation (2) is separable, giving

$$P(x)dx + Q(y)dy = 0,$$

$\partial P/\partial y = 0 = \partial Q/\partial x$, and it is exact. Thus the separable equation is a special case of the exact equation.

*Example* (c). Let the $(x, y)$ plane be the field of a force which has at the point $(x, y)$ the components $X(x, y)$, $Y(x, y)$. Then the work done by the field when the force acts through a displacement $(dx, dy)$ is, to the first order,

$$X dx + Y dy.$$

Thus the work done by the field when the point of application of the force is moved by the path $\gamma$ from $P_1$ to $P_2$ (Fig. 2.2) is

$$W_1^2(\gamma) = \int_{1(\gamma)}^{2} (X dx + Y dy), \tag{2.4.11}$$

the notation indicating that in general the value of the integral (and the

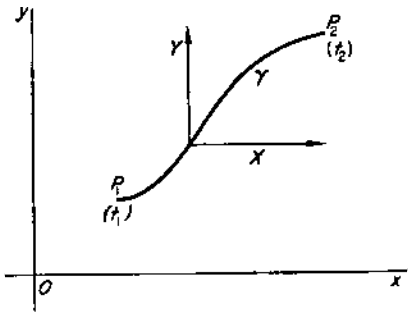


FIG. 2.2. Work done along the path $\gamma$.

work done) depends on the path $\gamma$. (The value of the integral is computed as follows : let the equation of the curve $\gamma$ be $x = x(t)$, $y = y(t)$, where $t$ is a parameter fixing position on $\gamma$. Then

$$W_1^2(\gamma) = \int_{t_1}^{t_2} \left[ X\{x(t), y(t)\} \frac{dx}{dt} + Y\{x(t), y(t)\} \frac{dy}{dt} \right] dt, \tag{2.4.12}$$

and this is the integral of a function of $t$ alone.)

It may happen that the force-field $(X, Y)$ is *conservative*, i.e. such that $W_1^2$ is independent of the path $\gamma$. Then $W_1^2$ must depend only on the end-points $P_1$ and $P_2$. In other words there must exist a scalar function of position (the work function) $W(x, y)$ such that

$$W_1^2 = W(x_2, y_2) - W(x_1, y_1).$$

Thus

$$[W(x, y)]_1^2 = \int_1^2 X\mathrm{d}x + Y\mathrm{d}y,$$

and consequently there is a scalar function $W(x, y)$ such that

$$\mathrm{d}W = X\mathrm{d}x + Y\mathrm{d}y.$$

Thus in a conservative field $X\mathrm{d}x + Y\mathrm{d}y$ is the differential of a function $W(x, y)$ and is therefore an exact differential. The condition for this is, as shown above,

$$\frac{\partial X}{\partial y} - \frac{\partial Y}{\partial x} = 0.$$

[Readers familiar with vector analysis will see that the preceding discussion could be more elegantly displayed as follows:

$$W_1^2(\gamma) = \int_{1(\gamma)}^2 \mathbf{F}.\mathbf{dr}.$$

If $W(\mathbf{r})$ exists,

$$\mathrm{d}W = \mathbf{F}.\mathbf{dr}, \quad \mathbf{F} = \mathbf{grad}\ W,$$

and

$$\mathbf{curl}\ \mathbf{F} = \mathbf{curl}\ \mathbf{grad}\ W = 0.]$$

## REFERENCES


1. Gillespie, R. P., *Integration*, Oliver & Boyd, 1955.
2. Stewart, C. A., *Advanced Calculus*, Methuen, 1946.
3. Peirce, B. O., *Short Table of Integrals*, Ginn, 1929.
4. *Handbook of Chemistry and Physics*, Chemical Publishing Co.
5. Hartree, D. R., *Numerical Analysis*, 2nd edn., Oxford, 1958.
6. Buckingham, R. A., *Numerical Methods*, Pitman, 1957.
7. Hinshelwood, C., *Kinetics of Chemical Change in Gaseous Systems*, Oxford, 1936.


## Exercises

**2.1.** A substance A decomposes according to the rate law

$$-\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{A}] = k[\mathrm{A}].$$

Show that the value of [A] at time $t$, denoted by $[\mathrm{A}]_t$, is a linear function of $[\mathrm{A}]_{t+\alpha}$, where $\alpha$ is a fixed interval.

**2.2.** A decomposition is known not to be of first order, and is suspected of being of fractional order. Taking the order to be $\alpha$ ($\neq 1$), we assume that the rate equation is of the form

$$\frac{dc}{dt} = -kc^{\alpha}, \quad c = c_0 \text{ when } t = 0.$$

Solve this equation and suggest a way of finding $\alpha$ by means of a graph of the solution.

**2.3.** Show that a general expression for the half-life $\tau_{\frac{1}{2}}$ of a decomposition of order $n$ (rate law $-d[A]/dt = k_n[A]^n$, with $[A] = [A]_0$ at $t = 0$), $n$ being not less than 2, is

$$\tau_{\frac{1}{2}} = \frac{1}{k_n(n-1)}\left(\frac{2^{n-1}-1}{[A]_0^{n-1}}\right).$$

**2.4.** The rate of formation of iodine by the reaction

$$H_2O_2 + 2H^+ + 2I^- \rightarrow I_2 + 2H_2O$$

is given by the rate law

$$\frac{d}{dt}[I_2] = k_1[H_2O_2][I^-] + k_2[H_2O_2][I^-][H^+].$$

Set up differential equations for (a) stoichiometric mixtures and (b) non-stoichiometric mixtures, and show how to solve the two equations.

**2.5.** Show, both from the differential equations and from the corresponding solutions, how opposed first order reactions may be regarded to a good approximation as simple if one reaction is slow compared with the other.

**2.6.** In the muta-rotation of glucose two opposed first order reactions occur:

$$\alpha G \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \beta G.$$

Show that starting from pure $\alpha G$ the process is described by

$$\frac{\alpha_t - \alpha_e}{\alpha_0 - \alpha_e} = \exp\{-(k_1 + k_2)t\},$$

where $\alpha_0$, $\alpha_t$, $\alpha_e$ are the concentrations of $\alpha G$ at time zero, time $t$, equilibrium, respectively.

**2.7.** *Thermal decomposition of hydrogen iodide.* Taking $x$ to represent the fraction of hydrogen iodide that has decomposed at time $t$, and allowing for the reversible nature of the reaction, show that the differential equation of the opposed reactions is

$$\frac{dx}{dt} = k(1-x)^2 - k'\left(\frac{x}{2}\right)^2.$$

Solve this equation. Show how to find $k$ when the amount decomposed at time $t$ and the amount decomposed at equilibrium are known. (Hinshelwood, *Kinetics of Chemical Change in Gaseous Systems*,[7] Chapter 2.)

**2.8.** The rate of hydrolysis of methyl acetate in a solution of fixed pH follows the law

$$\frac{dx}{dt} = k_1(a-x) - k_2x^2,$$

where $(a-x)$ is the concentration of the ester and $x$ is the concentration of each product. Show that

$$\ln\frac{ax_e + x(a-x_e)}{a(x_e-x)} = k_1\left(\frac{2a-x_e}{x_e}\right)t,$$

where $x_e$ is the equilibrium concentration of each product.

**2.9.** The rate of decomposition of ammonia over a platinum surface

$$(NH_3 \rightarrow \tfrac{1}{2}N_2 + \tfrac{3}{2}H_2)$$

is given by the equation

$$-\frac{d}{dt}p_{NH_3} = k\frac{p_{NH_3}}{p_{H_2}},$$

where $p_{NH_3}$ and $p_{H_2}$ are the partial pressures of ammonium and hydrogen. If $p_0$ is the initial pressure of pure ammonium and $p$ the total pressure at time $t$, show that

$$\ln\frac{2p_0-p}{p_0} = 1 - \frac{p}{p_0} - \frac{2k}{3p_0}t.$$

**2.10.** An ester A is catalysed by hydrogen ions so that the reaction is autocatalytic as the acid B is produced. If the rate of hydrolysis is proportional to [A] and to [B], and the initial concentrations of A and B are $[A]_0$ and $[B]_0$, set up and solve the differential equation of the reaction.

**2.11.** Show that if $P(x, y)dx + Q(x, y)dy$ is not an exact differential, it becomes exact when multiplied by $R(x, y)$ provided $R(x, y)$ satisfies the condition:

$$P\frac{\partial R}{\partial y} - Q\frac{\partial R}{\partial x} + \left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}\right)R = 0.$$

[In general this procedure is of no help towards the integration of

$$Pdx + Qdy = 0,$$

since the partial differential equation for $R(x, y)$ is harder to solve than the original differential equation. In special cases, however, progress can be made.]

Show that the equation

$$dy + \{p(x)y - f(x)\}dx = 0$$

has an integrating factor

$$\exp\{\int^x p(x)dx\}.$$

CHAPTER 3

# ORDINARY DIFFERENTIAL EQUATIONS THAT ARE LINEAR WITH CONSTANT COEFFICIENTS: SYSTEMS OF SUCH EQUATIONS

## 3.1. The linear equation of the first order with a constant coefficient

The laws of natural decay and growth were stated in Section 1.1, and the differential equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \mp kx, \quad k > 0, t > 0, \tag{3.1.1}$$

was derived. In Section 2.2(b) this equation was solved by separation of variables, the solution that takes the value $x_0$ at $t = 0$ being

$$x = x_0 \exp(\mp kt). \tag{3.1.2}$$

Another method of attack, worth describing because of its value in harder problems, goes as follows. The equation

$$\mathrm{d}x + kx\mathrm{d}t = 0 \tag{3.1.3}$$

is not exact, since

$$\frac{\partial}{\partial t}(1) = 0, \quad \frac{\partial}{\partial x}(kx) = k \quad \text{(Section 2.4)}.$$

Let us look for an *integrating factor*, that is, a function $g(t)$ such that when the equation (3) is multiplied through by $g(t)$ it becomes exact.

$$g(t)\mathrm{d}x + g(t)kx\mathrm{d}t = 0 \tag{3.1.4}$$

is exact if

$$\frac{\partial}{\partial t}\{g(t)\} = \frac{\partial}{\partial x}\{g(t)kx\},$$

i.e.

$$g'(t)/g(t) = k.$$

Hence (omitting an irrelevant factor)

$$\ln g(t) = kt,$$

and

$$g(t) = \exp kt.$$

(4) now becomes

$$\mathrm{d}(x\mathrm{e}^{kt}) = 0.$$

We may integrate this equation from $t = 0$ to $t = t$ and obtain

$$x\mathrm{e}^{kt} - x_0 = 0,$$

and

$$x = x_0\mathrm{e}^{-kt}$$

as before.

*The inhomogeneous equation*

Let us now consider the process of natural decay modified by an external disturbing contribution, such that the quantity measured by $x$ is being increased at a rate $F(t)$ whilst it decays naturally. We now have

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -kx + F(t), \quad t > 0, \quad x(0) = x_0,$$

or

$$\frac{\mathrm{d}x}{\mathrm{d}t} + kx = F(t). \tag{3.1.5}$$

Our equation (1), which is obtained from (5) when $F(t) = 0$, contains only a linear combination of $x$ and its derivatives, and is called a *homogeneous* equation. Equation (5), on account of the presence of the term $F(t)$, is called an *inhomogeneous* equation. The homogeneous equation (1) is said to *correspond* to the inhomogeneous equation (5), since it differs only in the term $F(t)$.

We cannot separate the variables in (5), but we can make the left-hand side exact. Following our second method of solving the corresponding homogeneous equation (1), we multiply (5) by the integrating factor $\exp kt$. It then becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}(x\mathrm{e}^{kt}) = F(t)\mathrm{e}^{kt}, \tag{3.1.6}$$

and we can integrate to obtain

$$xe^{kt} - x_0 = \int_0^t F(u)e^{ku}du,$$

or

$$x = x_0e^{-kt} + e^{-kt}\int_0^t F(u)e^{ku}du. \tag{3.1.7}$$

[The reader will note that the variable of integration has been written as $u$ rather than $t$, for we are now using $t$ to denote the upper end-point of the range of integration. Since the integral is definite we are at liberty to use any symbol for the variable of integration, and we choose $u$ to avoid confusion with $t$.]

In (7) the problem has been reduced to quadrature. Although it appears that an exponentially increasing factor has been introduced as a multiplier of $F$, a factor $\exp(-kt)$ multiplies the integral, and if we write

$$\int_0^t F(u)e^{-k(t-u)}\,du \tag{3.1.8}$$

we see that $F$ is multiplied by a factor less than or equal to unity before integration.

The first term of the expression for $x$ in (7) is what would be obtained if $F(t) = 0$; it measures the intrinsic behaviour of the system. The second term measures the response to the external disturbing term $F(t)$; naturally it contains also $k$ which affects that response.

*Example.* $F(t)$ might be the rate of increase of $x$ due to the breakdown of some other substance which is itself decaying naturally. Thus if the other substance is measured by

$$y = y_0 \exp(-mt),$$

then

$$-\frac{dy}{dt} = my_0 \exp(-mt),$$

and this is the value of $F(t)$. Substituting into (7) we get

$$x = x_0e^{-kt} + e^{-kt}\int_0^t my_0e^{(k-m)u}\,du,$$

where we have assumed that $m$ is less than $k$, though this is not necessary to our argument but makes for more definite operations. Then

$$x = x_0 e^{-kt} + e^{-kt} \frac{my_0}{k-m} \{e^{(k-m)t} - 1\}$$

$$= x_0 e^{-kt} + \frac{m}{k-m} y_0 (e^{-mt} - e^{-kt}). \tag{3.1.9}$$

This expression for $x$ represents the sum of that fraction of the original substance which remains after time $t$ (first term) together with that fraction of the continually added substance which has not decayed again at time $t$ (second term). These two fractions are shown in Fig. 3.1; their relative size depends on the ratio of $y_0$ to $x_0$ as well as of $m$ to $k$.

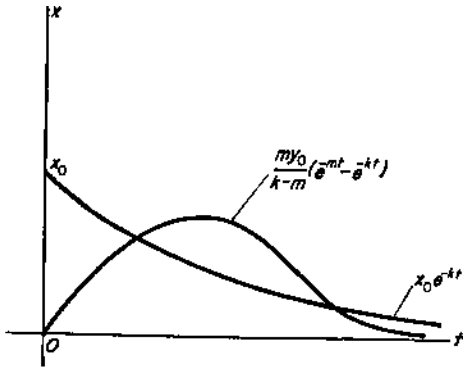


FIG. 3.1. Contributions to $x(t)$.

It will be seen that the contribution from the second substance rises to a maximum at the moment when the rate of production from dwindling $y$ just equals the rate of decay from increasing $x$; it thereafter sinks gradually to zero.

## 3.2. The linear equation of the second order with constant coefficients

Suppose that a particle of mass $m$ moves along the straight line $Ox$ under two forces—(1) a central force in which when the particle is at distance $x$ from $O$ the force towards $O$ is $\lambda x$, where $\lambda$ is a positive constant, and (2) a resistance $k\dot{x}$ proportional to the velocity $\dot{x}$ ($k$ being

a positive constant) and opposing the motion. Then the equation of motion is

$$m\ddot{x}+k\dot{x}+\lambda x = 0, \quad t > 0,$$

i.e.

$$\ddot{x}+2b\dot{x}+cx = 0, \quad t > 0, \tag{3.2.1}$$

where $2b = k/m$ and $c = \lambda/m$. We must also specify two boundary conditions; let them be the initial conditions

$$x = x_0, \quad \dot{x} = u_0, \quad \text{when } t = 0.$$

We may solve equation (1) by various methods, of which two are given here and one in Chapter 4.

*First method*

Since there is a linear relation, true for all $t > 0$, between $x$, its first derivative, and its second derivative, these three quantities must all be functions of $t$ of essentially the same type. But the simplest function satisfying this condition is the exponential function. So let us seek a solution $x = A \exp \alpha t$, where $A$ is a constant and $\alpha$ a parameter which will be selected so that this $x$ satisfies the differential equation (1).

Substituting $x = Ae^{\alpha t}$, $\dot{x} = A\alpha e^{\alpha t}$, $\ddot{x} = A\alpha^2 e^{\alpha t}$ into (1) we have

$$(\alpha^2+2b\alpha+c)Ae^{\alpha t} = 0, \quad t > 0.$$

This can be true for all $t$ only if

$$\alpha^2+2b\alpha+c = 0, \tag{3.2.2}$$

i.e. if $\alpha$ is chosen as a root of (2).

Let us first suppose that $b^2-c \neq 0$, so that (2) has two distinct roots, which may be both real or a pair of conjugate complex numbers. Let these roots be called $\alpha_1$ and $\alpha_2$. Then $x_1 = A_1 \exp \alpha_1 t$, where $A_1$ is any constant, is a solution of (1), and $x_2 = A_2 \exp \alpha_2 t$, is a second solution. Now it is easily verified that if $x_1(t)$ and $x_2(t)$ are solutions of (1), then $x_1(t)+x_2(t)$ is also a solution. Thus a solution which contains two arbitrary constants is

$$x = A_1 e^{\alpha_1 t}+A_2 e^{\alpha_2 t}. \tag{3.2.3}$$

We may now determine the constants $A_1$ and $A_2$ so as to satisfy the initial conditions. Thus

$$x_0 = A_1+A_2,$$
$$u_0 = \alpha_1 A_1+\alpha_2 A_2,$$

whence

$$A_1 = \frac{u_0 - \alpha_2 x_0}{\alpha_1 - \alpha_2}, \qquad A_2 = \frac{-u_0 + \alpha_1 x_0}{\alpha_1 - \alpha_2},$$

giving the solution

$$x(t) = \frac{1}{\alpha_1 - \alpha_2}\{u_0(e^{\alpha_1 t} - e^{\alpha_2 t}) - x_0(\alpha_2 e^{\alpha_1 t} - \alpha_1 e^{\alpha_2 t})\}. \qquad (3.2.4)$$

If $\alpha_1$ and $\alpha_2$ are real, this form is satisfactory. But if $\alpha_1$ and $\alpha_2$ are complex conjugates it contains complex expressions which can be avoided as follows. Let $\alpha_1 = \gamma + i\delta$, $\alpha_2 = \gamma - i\delta$, where $\gamma$, $\delta$ are real. Then

$$\begin{aligned} \alpha_1 - \alpha_2 &= 2i\delta, \\ e^{\alpha_1 t} - e^{\alpha_2 t} &= e^{\gamma t}(e^{i\delta t} - e^{-i\delta t}) = 2i \sin \delta t\, e^{\gamma t}, \\ \alpha_2 e^{\alpha_1 t} - \alpha_1 e^{\alpha_2 t} &= e^{\gamma t}\{(\gamma - i\delta)e^{i\delta t} - (\gamma + i\delta)e^{-i\delta t}\} \\ &= e^{\gamma t}(2i\gamma \sin \delta t - 2i\delta \cos \delta t), \end{aligned}$$

so that

$$x = \frac{e^{\gamma t}}{\delta}\{u_0 \sin \delta t - x_0(\gamma \sin \delta t - \delta \cos \delta t)\}. \qquad (3.2.5)$$

We have still to deal with the special case when $b^2 - c = 0$ and (2) has coincident roots. Then $\alpha_2 = \alpha_1$ and $A_2 \exp \alpha_2 t$ is not essentially different from $A_1 \exp \alpha_1 t$. In this case (3) does not represent a general solution: it is of the form $(A_1 + A_2) \exp \alpha_1 t$, and in reality contains only one arbitrary constant. We have to find a new and different solution of (1) in order to construct a general solution. This we achieve by means of a manoeuvre frequently used in this subject.

Let us form a function $e^{\alpha_1 t}v(t)$, $\alpha_1$ being the repeated root of (2), and try to determine $v(t)$ so that $x(t) = e^{\alpha_1 t}v(t)$ satisfies (1). Substituting

$$\begin{aligned} \dot{x} &= \alpha_1 e^{\alpha_1 t}v + e^{\alpha_1 t}\dot{v}, \\ \ddot{x} &= \alpha_1{}^2 e^{\alpha_1 t}v + 2\alpha_1 e^{\alpha_1 t}\dot{v} + e^{\alpha_1 t}\ddot{v}, \end{aligned}$$

we get

$$e^{\alpha_1 t}\{(\alpha_1{}^2 v + 2\alpha_1 \dot{v} + \ddot{v}) + 2b(\alpha_1 v + \dot{v}) + cv\} = 0,$$

i.e.

$$e^{\alpha_1 t}\{(\alpha_1{}^2 + 2b\alpha_1 + c)v + 2(\alpha_1 + b)\dot{v} + \ddot{v}\} = 0,$$

But since in this case (2) has equal roots $(\alpha_1, \alpha_1)$

$$\alpha_1^2 + 2b\alpha_1 + c = 0,$$

$$\alpha_1 + b = 0,$$

and $e^{\alpha_1 t}$ cannot vanish. Hence

$$\ddot{v} = 0,$$

and

$$v(t) = Bt + C, \tag{3.2.6}$$

where $B$ and $C$ are constants. Thus a second solution is $e^{\alpha_1 t}(Bt + C)$; but $e^{\alpha_1 t}C$ is merely a multiple of the first solution, so that the second solution can be taken as $Bte^{\alpha_1 t}$. The general solution is then

$$(A + Bt)e^{\alpha_1 t}. \tag{3.2.7}$$

This satisfies the initial conditions if

$$x_0 = A,$$

$$u_0 = \alpha_1 A + B,$$

and hence

$$x(t) = \{x_0 + (u_0 - \alpha_1 x_0)t\}e^{\alpha_1 t}. \tag{3.2.8}$$

*Second method*

Let us denote the operation of differentiating with respect to $t$ by $D$, so that $\dot{x} = Dx$ and

$$\ddot{x} = D(Dx) = D^2 x.$$

The operator $D$ commutes with any constant, i.e.

$$D(ax) = aDx,$$

but does not commute with any function of $t$, since

$$D\{f(t)x\} = f'(t)x + f(t)Dx$$

$$\neq f(t)Dx$$

unless $f'(t) = 0$ (and so $f(t)$ is a constant). Then the differential equation

$$\ddot{x} + 2b\dot{x} + cx = 0 \tag{3.2.1}$$

can be written

$$(D^2 + 2bD + c)x = 0. \tag{3.2.9}$$

Now the operator $D^2+2bD+C$ is identical in form with

$$\alpha^2+2b\alpha+c = (\alpha-\alpha_1)(\alpha-\alpha_2),$$

where $\alpha_1$ and $\alpha_2$ have the same meaning as before (but here we make no condition that they should be distinct). We can therefore factorize the operator as

$$(D-\alpha_1)(D-\alpha_2),$$

since $D$ commutes with the constants $\alpha_1$ and $\alpha_2$ to give

$$\begin{aligned}(D-\alpha_1)(D-\alpha_2) &= D(D-\alpha_2)-\alpha_1(D-\alpha_2)\\ &= D^2-\alpha_2D-\alpha_1D+\alpha_1\alpha_2\\ &= D^2+2bD+c,\end{aligned}$$

where we have used the relations between roots and coefficients:

$$\alpha_1+\alpha_2 = -2b, \quad \alpha_1\alpha_2 = c.$$

Thus we have to solve

$$(D-\alpha_1)(D-\alpha_2)x = 0, \quad t > 0, \tag{3.2.10}$$

with $x = x_0$, $\dot{x} = u_0$ at $t = 0$. We accomplish the work in two steps, making use of an auxiliary variable. Let us write

$$(D-\alpha_2)x = y, \tag{3.2.11}$$

so that when $t = 0$,

$$y = y_0 = u_0-\alpha_2x_0. \tag{3.2.12}$$

Then substituting from (11) into (10) we have

$$(D-\alpha_1)y = 0, \ y = y_0 \text{ when } t = 0.$$

This is a linear equation of the first order, with the solution

$$y = y_0e^{\alpha_1t}. \tag{3.2.13}$$

Substituting this value of $y(t)$ into (11) we have

$$(D-\alpha_2)x = y_0e^{\alpha_1t}, \ x = x_0 \text{ when } t = 0. \tag{3.2.14}$$

But this is another first order linear equation, this time inhomogeneous. We may quote the solution using (1.7) or we may use the method of Section 3.1 as follows. The integrating factor is $\exp(-\alpha_2t)$, giving

$$D(xe^{-\alpha_2t}) = y_0e^{(\alpha_1-\alpha_2)t}, \ x = x_0 \text{ when } t = 0. \tag{3.2.15}$$

Integrating from $t = 0$,

$$xe^{-\alpha_2 t} - x_0 = \frac{y_0}{\alpha_1 - \alpha_2}\{e^{(\alpha_1-\alpha_2)t} - 1\},$$

so that

$$\begin{aligned} x &= x_0 e^{\alpha_2 t} + \frac{u_0 - \alpha_2 x_0}{\alpha_1 - \alpha_2}(e^{\alpha_1 t} - e^{\alpha_2 t}) \\ &= \frac{1}{\alpha_1 - \alpha_2}\{u_0(e^{\alpha_1 t} - e^{\alpha_2 t}) - x_0(\alpha_2 e^{\alpha_1 t} - \alpha_1 e^{\alpha_2 t})\} \end{aligned} \tag{3.2.4}$$

as before (4).

This method is available (and very elegant) if the roots $\alpha_1$ and $\alpha_2$ coincide. For then the above working stands as far as (14), with $\alpha_2 = \alpha_1$. (15) becomes

$$D(xe^{-\alpha_1 t}) = y_0$$

which integrates to

$$xe^{-\alpha_1 t} - x_0 = y_0 t,$$

so that, as before,

$$x = \{x_0 - (u_0 - \alpha_1 x_0)t\}e^{\alpha_1 t}. \tag{3.2.8}$$

The first method is more immediate where $\alpha_1$ and $\alpha_2$ are distinct, or where the general solution only is required, initial conditions not being specified. When the first method is used for the inhomogeneous equation various special tricks or methods of guesswork are introduced. They will not be dealt with in this book, since our second method is available in all cases. The second method has the advantages that it deals with equal roots without modification, that it extends immediately to the inhomogeneous equation, and that it incorporates the initial conditions in the course of the work, without the introduction of arbitrary constants.

*Example.* In a nitric oxide–chlorine system, chlorine atoms are generated at constant rate $n$ and are removed at a rate proportional to the number of atoms present (rate constant $k$). The chlorine atoms diffuse through the system with diffusion coefficient $D$ and are removed when they reach the walls of the container. We consider steady diffusion between parallel plane walls.

In the steady state, if $c$ is the concentration of chlorine atoms at a distance $x$ from one wall and $F$ is the rate of flow of chlorine atoms

in the direction of $x$, the balance of chlorine atoms in a cylinder of unit cross-section and length $\delta x$ gives

$$n\delta x - kc\delta x - \frac{\mathrm{d}F}{\mathrm{d}x}\delta x = 0,$$

while Fick's law gives

$$F = -D\frac{\mathrm{d}c}{\mathrm{d}x}.$$

Thus we derive

$$D\frac{\mathrm{d}^2c}{\mathrm{d}x^2} - kc = -n, \tag{3.2.16}$$

with $c = 0$ at $x = 0$ and $x = l$.

An inhomogeneous equation such as (16), which contains on the right-hand side a simple constant, can be made homogeneous by a small change of variable. The equation is

$$D\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(c - \frac{n}{k}\right) - k\left(c - \frac{n}{k}\right) = 0.$$

We therefore write

$$c - \frac{n}{k} = z,$$

and obtain

$$\frac{\mathrm{d}^2z}{\mathrm{d}x^2} - m^2z = 0, \tag{3.2.17}$$

where $m^2 = k/D$, with $z = -n/k$ when $x = 0, l$.

Since this is a jury problem, with one boundary condition at each end, it is convenient first to find the general solution and then determine constants. The equation (17) has general solution

$$z = Ae^{mx} + Be^{-mx}, \tag{3.2.18}$$

as is found immediately by either of our methods. The end conditions are

$$A + B = -N, \text{ where } N = n/k,$$

and

$$Ae^{ml} + Be^{-ml} = -N.$$

Then

$$A = -N\frac{1-e^{-ml}}{e^{ml}-e^{-ml}}, \qquad B = N\frac{1-e^{ml}}{e^{ml}-e^{-ml}},$$

so that

$$c = \frac{n}{k}\left\{1-\frac{(1-e^{-ml})e^{mx}-(1-e^{ml})e^{-mx}}{e^{ml}-e^{-ml}}\right\}$$

$$= \frac{n}{k}\left\{1-\frac{\sinh mx+\sinh m(l-x)}{\sinh ml}\right\}. \tag{3.2.19}$$

This answer clearly makes $c = 0$ at $x = 0$ and $x = l$ and is, as might be expected, symmetrical in $x$ and $l-x$. The graph of concentration against $x$ is shown in Fig. 3.2 (the curve is an inverted catenary).

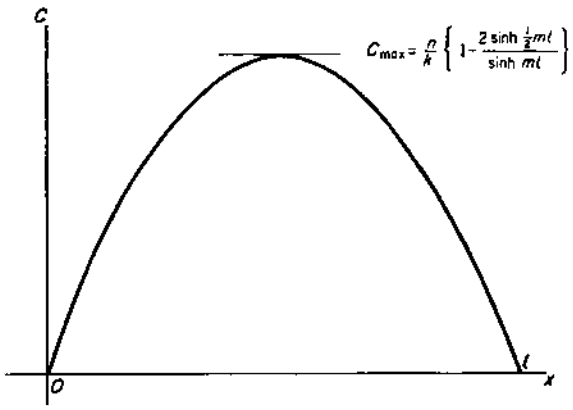


FIG. 3.2. Graph of concentration against $x$,
$c = \frac{n}{k}\left\{1-\frac{\sinh mx+\sinh m(l-x)}{\sinh ml}\right\}$

*The inhomogeneous equation*

Returning to the physical problem of linear motion with which this section began, we now make the additional assumption that the particle is subject to a further (time-dependent) force $mf(t)$, so that instead of (1) we have

$$\ddot{x}+2b\dot{x}+cx = f(t). \tag{3.2.20}$$

(20) is the inhomogeneous equation corresponding to the homogeneous equation (1). We solve it by our second method. (20) is

$$(D-\alpha_1)(D-\alpha_2)x = f(t),$$

with $x = x_0$, $\dot{x} = u_0$ at $t = 0$. Take $(D-\alpha_2)x = y$, so that $y = y_0 = u_0 - \alpha_2 x_0$ at $t = 0$. Then

$$(D-\alpha_1)y = f(t). \tag{3.2.21}$$

This is a linear equation with integrating factor $\exp(-\alpha_1 t)$, so that

$$ye^{-\alpha_1 t} - y_0 = \int_0^t f(u)e^{-\alpha_1 u}du,$$

$$y(t) = y_0 e^{\alpha_1 t} + \int_0^t f(u)e^{\alpha_1(t-u)}\,du. \tag{3.2.22}$$

We substitute this expression into the equation $(D-\alpha_2)x = y$, and solve the resulting linear equation in the form

$$x(t) = x_0 e^{\alpha_2 t} + \int_0^t y(v)e^{\alpha_2(t-v)}dv, \tag{3.2.23}$$

where $y(v)$ is found from (22). Thus

$$\begin{aligned} x(t) &= x_0 e^{\alpha_2 t} + \int_0^t y_0 e^{\alpha_1 v} e^{\alpha_2(t-v)}dv \\ &\qquad + \int_0^t e^{\alpha_2(t-v)} \int_0^v f(u)e^{\alpha_1(v-u)}du\,dv \\ &= x_0 e^{\alpha_2 t} + (u_0 - \alpha_2 x_0)e^{\alpha_2 t}\,\frac{e^{(\alpha_1-\alpha_2)t}-1}{\alpha_1-\alpha_2} \\ &\qquad + \int_0^t e^{\alpha_2 t + (\alpha_1-\alpha_2)v} \int_0^v f(u)e^{-\alpha_1 u}du\,dv. \end{aligned} \tag{3.2.24}$$

(3.2.24) is a formal answer to the problem. But the double integral may be reduced to a single integral by means of an interchange of order of integration. [It is proved in textbooks on advanced calculus that for the change of order in a double integral to be valid it is *sufficient* that the ranges of integration be finite and the integrand be continuous in each variable.] We have to take care that when the order is changed the limits of integration are correctly adjusted.

We rewrite the double integral as

$$\int_0^t dv \int_0^v f(u)e^{\alpha_2 t+(\alpha_1-\alpha_2)v-\alpha_1 u}du, \tag{3.2.25}$$

and observe that in the $(u, v)$ plane the integration extends over the triangle OAB (Fig. 3.3), the double integral being regarded as the limit of a process of summing first along strips PQ parallel to the axis of $u$ and then over all strips between O and AB.

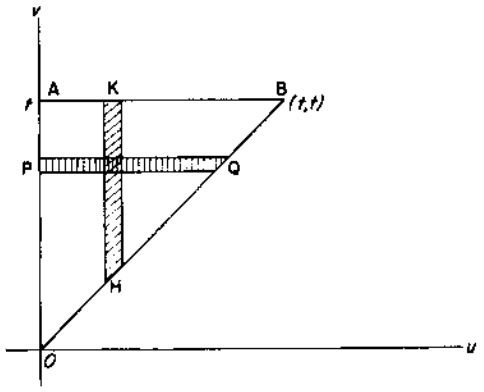


FIG. 3.3. Change of order of integration.

When we change the order of integration, we regard the double integral as the limit of a process of summing first along strips HK parallel to the axis of $v$ and then over all such strips between OA and B. Thus the new inner integral must have as end-points $v = u$ and $v = t$, and the outer must have $u = 0$ and $u = t$. Then (25) becomes

$$\int_0^t du \int_u^t f(u)e^{\alpha_2 t+(\alpha_1-\alpha_2)v-\alpha_1 u}dv$$

$$= \int_0^t du f(u)e^{\alpha_2 t-\alpha_1 u}\frac{1}{\alpha_1-\alpha_2}\{e^{(\alpha_1-\alpha_2)t}-e^{(\alpha_1-\alpha_2)u}\}$$

$$= \int_0^t f(u)\frac{e^{\alpha_1(t-u)}-e^{\alpha_2(t-u)}}{\alpha_1-\alpha_2}du. \tag{3.2.26}$$

This manoeuvre succeeds because $f$ is a function of $u$ only, so that when the order of integration is inverted $f(u)$ can be taken outside the inner integral, leaving an easily evaluated function. Substituting (26) into (24) we find the solution

$$x(t) = \frac{1}{\alpha_1 - \alpha_2}\left[u_0(e^{\alpha_1 t} - e^{\alpha_2 t}) - x_0(\alpha_2 e^{\alpha_1 t} - \alpha_1 e^{\alpha_2 t})\right.$$

$$\left. + \int_0^t f(u)\{e^{\alpha_1(t-u)} - e^{\alpha_2(t-u)}\}du\right]. \qquad (3.2.27)$$

The first two terms within the square bracket in (27) are seen to be those obtained in (4): they represent the intrinsic response of the system to the disturbance given by the initial displacement and velocity. The third term is the response to the "forcing term" $f(t)$. These responses may be added to give the full solution because the differential equation is linear.

A substitution of $t-w$ for $u$ in the integrand gives the integral in the equivalent form

$$\int_0^t f(t-w)(e^{\alpha_1 w} - e^{\alpha_2 w})dw, \qquad (3.2.28)$$

which is to be preferred for some forms of $f(t)$.

The method given above is general and powerful. It is valid for all $f(t)$ for which the integrals converge and for any $\alpha_1$ and $\alpha_2$, real or complex. If $\alpha_1 = \alpha_2$, the first two terms are modified as in (8), and the double integral in (24) simplifies (since $(\alpha_1 - \alpha_2)v = 0$) to give, in place of (26),

$$\int_0^t f(u)e^{\alpha_1(t-u)}(t-u)du. \qquad (3.2.29)$$

*Example.* A common special case of the linear motion considered above is that of a system which if left to itself would perform simple harmonic oscillations of period $2\pi/n$. Then the governing equation (20) becomes

$$\ddot{x} + n^2 x = f(t), \qquad (3.2.30)$$

so that $\alpha_1 = in$, $\alpha_2 = -in$, and (27) reduces to

$$x(t) = \frac{1}{2in}\left[u_0(e^{int} - e^{-int}) - x_0(-ine^{int} - ine^{-int}) + \int_0^t f(u)\{e^{in(t-u)} - e^{-in(t-u)}\}du\right]$$

$$= \frac{1}{n}\left[u_0 \sin nt + nx_0 \cos nt + \int_0^t f(u) \sin n(t-u)du\right]. \tag{3.2.31}$$

### 3.3. The general linear equation with constant coefficients

The work on second order equations in the preceding section can be generalized to give the solution of a linear equation of order $n$ with constant coefficients. Let $D \equiv d/dt$ and let $L(D)$ denote the linear operator

$$D^n + a_1 D^{n-1} + a_2 D^{n-2} + \ldots + a_n. \tag{3.3.1}$$

Then

$$\{L(D)\}x = f(t), \quad t > 0, \tag{3.3.2}$$

is an inhomogeneous linear equation of degree $n$, and

$$\{L(D)\}x = 0, \quad t > 0, \tag{3.3.3}$$

is the corresponding homogeneous equation.

We first introduce the concepts of *linear dependence* and *linear independence* of a set of functions. If $\phi_1(t), \phi_2(t), \ldots, \phi_n(t)$ are functions given for a certain range of $t$, and if a set of constant multipliers $\alpha_1, \alpha_2, \ldots, \alpha_n$ can be found such that

$$\alpha_1\phi_1(t) + \alpha_2\phi_2(t) + \ldots + \alpha_n\phi_n(t) = 0$$

for all $t$ within the given range, then $\phi_1(t), \phi_2(t), \ldots, \phi_n(t)$ are said to be linearly dependent. If no such set of constants exists, the functions are said to be linearly independent. No one of them can then be expressed as a linear combination of the remainder.

*First method*

We now begin with an attack on the homogeneous equation (3). We observe that if $x = \phi_1(t)$ and $x = \phi_2(t)$ are linearly independent solutions of (3) and $A_1$, $A_2$ are any constants, then

$$\{L(D)\}(A_1\phi_1 + A_2\phi_2) = A_1\{L(D)\}\phi_1 + A_2\{L(D)\}\phi_2 = 0,$$

so that $A_1\phi_1 + A_2\phi_2$ is a solution of (3) containing two arbitrary constants $A_1$ and $A_2$. Since $\phi_1$ and $\phi_2$ are linearly independent, $\phi_2$ cannot be replaced by a multiple of $\phi_1$, thus the expression cannot be equated to an expression containing only one arbitrary constant. Now since (3) is of order $n$, and its solution will involve $n$ integrations, the most general solution must contain $n$ arbitrary constants. But a generalization of the argument just given shows directly that if

$$\phi_1(t), \phi_2(t), \ldots, \phi_n(t)$$

are $n$ linearly independent solutions of (3), then

$$A_1\phi_1 + A_2\phi_2 + \ldots + A_n\phi_n \tag{3.3.4}$$

is a solution containing $n$ arbitrary constants $A_1, A_2, \ldots, A_n$ which cannot be reduced in number, and so is the general solution. This property is a consequence of the *linearity* of (3).

We have now to find $n$ linearly independent solutions of (3) : our first method follows from the first method of Section 3.2. Since for all $t > 0$ the sum of multiples of the derivatives of $x$ is zero, all the derivatives must be essentially similar to $x(t)$ itself. This suggests trying as a solution $x = \exp \lambda t$, where $\lambda$ is a parameter to be determined. Substituting into (3) we find

$$e^{\lambda t}(\lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \ldots + a_n) = 0, \tag{3.3.5}$$

so that $\lambda$ must satisfy the $n$th order equation

$$\lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \ldots + a_n = 0. \tag{3.3.6}$$

This equation has $n$ roots (real and complex) which may not all be distinct. Let them be denoted by

$$\lambda_1, \lambda_2, \ldots, \lambda_n.$$

We first suppose that these roots are all different. Then

$$\exp \lambda_1 t, \exp \lambda_2 t, \ldots, \exp \lambda_n t$$

are $n$ linearly independent solutions of (3), so that the general solution will be

$$A_1e^{\lambda_1 t}+A_2e^{\lambda_2 t}+\ldots+A_ne^{\lambda_n t}$$

or

$$\sum_{r=1}^{n} A_re^{\lambda_r t}. \tag{3.3.7}$$

This form of solution is available whether the roots are real or complex, but it is sometimes desirable to deal with a pair of conjugate complex roots together, introducing real trigonometric functions as follows. Without loss of generality we can take $\lambda_1$ and $\lambda_2$ to be the complex conjugates, and write

$$\lambda_1 = \gamma+i\delta, \qquad \lambda_2 = \gamma-i\delta, \tag{3.3.8}$$

where $\gamma$, $\delta$ are real. Then

$$\begin{aligned} A_1e^{\lambda_1 t}+A_2e^{\lambda_2 t} &= A_1e^{\gamma t}(\cos\delta t+i\sin\delta t) \\ &\quad +A_2e^{\gamma t}(\cos\delta t-i\sin\delta t) \\ &= e^{\gamma t}(B\cos\delta t+C\sin\delta t), \end{aligned} \tag{3.3.9}$$

where

$$B = A_1+A_2, \qquad C = i(A_1-A_2).$$

[This does not mean that $C$ must be imaginary; for instance if $A_1$ and $A_2$ are complex conjugates both $B$ and $C$ will be real.] If the equation has several pairs of complex roots, then corresponding to each conjugate pair there will be a pair of terms, as in the right-hand side of (9), containing trigonometric functions of $t$.

If the equation (6) for $\lambda$ has repeated roots, then the sum (7) of exponentials cannot represent the general solution. For if, for example, $\lambda_1$ is a double root, (7) contains as its first two terms

$$A_1e^{\lambda_1 t}+A_2e^{\lambda_1 t}$$

and therefore does not contain $n$ separate adjustable constants. We may guess from our work on the second order equation that corresponding to a double root $\lambda_1$ of (6) there should be included instead of $A_1e^{\lambda_1 t}+A_2e^{\lambda_2 t}$ a pair of terms

$$(A_1+B_1t)e^{\lambda_1 t},$$

where $A_1$ and $B_1$ are arbitrary constants. This can be established as follows. We examine a process in which two roots $\lambda_1$ and $\lambda_2$, originally

distinct, come to coincidence as $\lambda_2 \to \lambda_1$. Let $\lambda_2 = \lambda_1 + \varepsilon$. Then $\exp \lambda_1 t$ and $\exp (\lambda_1 + \varepsilon)t$ are linearly independent solutions when $\varepsilon \neq 0$, but the second becomes identical with the first when $\varepsilon = 0$. If, however, we form the expression

$$\{\exp (\lambda_1 + \varepsilon)t - \exp \lambda_1 t\}/\varepsilon,$$

this is linearly independent of $\exp \lambda_1 t$, since it cannot be expressed in terms of $\exp \lambda_1 t$ alone, and moreover its limit as $\varepsilon \to 0$ is a function of $t$ linearly independent of $\exp \lambda_1 t$. This limit is

$$e^{\lambda_1 t} \lim_{\varepsilon \to 0} \frac{e^{\varepsilon t} - 1}{\varepsilon} = te^{\lambda_1 t}.$$

We have thus constructed a second linearly independent solution, so that a general combination of two linearly independent solutions is $(A_1 + B_1 t) \exp \lambda_1 t$.

This method can be modified to deal with double complex roots, and can be extended to roots of multiplicity greater than two. But we do not pursue this first method further in those directions, since other methods given in this chapter and Chapter 4 prove to be more powerful and of more general application.

Let us assume that we have found $n$ linearly independent integrals of the homogeneous equation,

$$\phi_1, \phi_2, \ldots, \phi_n,$$

so that we have the general solution

$$\phi_h = \sum_{r=1}^{n} A_r \phi_r. \tag{3.3.10}$$

We now show how this can be used in obtaining the general solution of the *inhomogeneous* equation

$$\{L(D)\}x = f(t), \quad t > 0. \tag{3.3.2}$$

Suppose that by some means we have found any particular integral $x = X(t)$ of (2). Then

$$\begin{aligned} \{L(D)\}(\phi_h + X) &= \{L(D)\}\phi_h + \{L(D)\}X \\ &= 0 + f(t), \end{aligned} \tag{3.3.11}$$

since $\phi_h$ and $X$ are solutions of the homogeneous and inhomogeneous equations respectively. Thus $\phi_h + X$ is a solution of the inhomogeneous

These initial values of $y_1, y_2, \ldots, y_{n-1}$ can be introduced at each stage of integration until (16) is reached. Then $x(0) = b_0$ is used in the integration, and the solution for $x(t)$ is obtained with all initial conditions incorporated.

It is clear that this method of solution, though systematic, might become extremely cumbersome, especially if the successive integrals could not be evaluated in simple closed form. One is also led to wonder whether it is necessary to integrate the function $f(t)$ (with various factors) $n$ times, seeing that when $n = 2$ the double integral reduces to a single integral. In fact it is not necessary; the Laplace transform method (Chapter 4) leads to a single integral.

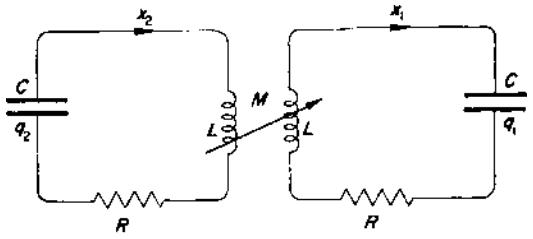


FIG. 3.4. Coupled circuits.

*Example.* Two electric circuits are coupled (Fig. 3.4). Each has resistance $R$, capacitance $C$, and self-inductance $L$. The mutual inductance is $M$. Initially the first circuit has charge $Q$ in the condenser, and no current flowing, while the second has no charge or current. We wish to find the subsequent variations in charge on the first condenser.

Let the charge in the first condenser at time $t$ be $q_1(t)$, so that the current in the first circuit is $x_1 = \dot{q}_1(t)$. Similarly let $q_2(t)$ and $x_2 = \dot{q}_2(t)$ refer to the second circuit. Then the differential equations of the two circuits are

$$\left.\begin{aligned} L\dot{x}_1 + M\dot{x}_2 + Rx_1 + q_1/C &= 0, \\ M\dot{x}_1 + L\dot{x}_2 + Rx_2 + q_2/C &= 0, \end{aligned}\right\} \quad t > 0, \tag{3.3.17}$$

with

$$q_1(0) = Q, \quad \dot{q}_1(0) = 0, \quad q_2(0) = 0, \quad \dot{q}_2(0) = 0. \tag{3.3.18}$$

Thus

$$\begin{aligned} (LD^2 + RD + 1/C)q_1 + MD^2q_2 &= 0, \\ MD^2q_1 + (LD^2 + RD + 1/C)q_2 &= 0. \end{aligned} \tag{3.3.19}$$

equation. But it contains $n$ adjustable constants and is therefore the general solution.

Thus the finding of the general solution of the inhomogeneous equation is broken down into two steps, (1) the finding of the general integral of the corresponding homogeneous equation, and (2) the finding of any particular integral of the inhomogeneous equation. Because of this, the homogeneous equation is sometimes called the *auxiliary equation.* The general solution (10) of the auxiliary equation is called the *complementary function.*

*Example.* Given the *inhomogeneous* equation

$$\ddot{x} - 4\dot{x} + 3x = 6t - 8,$$

the corresponding *homogeneous* (or auxiliary) equation is

$$\ddot{x} - 4\dot{x} + 3x = 0.$$

The general solution of this (i.e. the *complementary function*) is

$$x = Ae^{t} + Be^{3t},$$

and a particular integral of the inhomogeneous equation is seen to be

$$x = 2t.$$

Hence the general solution of the inhomogeneous equation is

$$x = Ae^{t} + Be^{3t} + 2t.$$

The search for the needed particular integral may be carried out by a combination of experience with trial and error (see exercises at end of chapter), or sometimes by manipulations with the operator $L(D)$. The latter method requires the use of processes of uncertain validity, and the memorizing of various formulae. It is not introduced here or advocated, for the method now to be described is preferred. Better still in many problems is the use of the Laplace transform (Chapter 4).

*Second method*

Our second method of Section 3.2 can be established theoretically for the linear equation of any order with constant coefficients. Its convenience in practice will depend on the ease with which the operator $L(D)$ can be factorized.

Since the form of $L(D) = D^n + a_1 D^{n-1} + \ldots + a_n$ is identical with that of the polynomial in the equation (6) for $\lambda$, the roots of which are $\lambda_1, \lambda_2, \ldots, \lambda_n$ we can write

$$L(D) = (D - \lambda_1)(D - \lambda_2) \ldots (D - \lambda_n). \tag{3.3.12}$$

This factorization of the operator is always possible, whether the roots are real or complex, single or multiple: it rests only on the fact that

the operator $D$ commutes with constants (see Section 3.2). Extending the method previously used for the second order equation, we now introduce extra variables $y_1, y_2, \ldots, y_{n-1}$, defined by the equations

$$\begin{aligned}(D-\lambda_n)x &= y_1,\\ (D-\lambda_{n-1})y_1 &= y_2,\\ &\cdots\\ (D-\lambda_2)y_{n-2} &= y_{n-1}.\end{aligned} \tag{3.3.13}$$

These are all uniquely defined in terms of $x$ and its derivatives. Then the differential equation is

$$(D-\lambda_1)(D-\lambda_2)\ldots(D-\lambda_n)x = f(t), \quad t > 0, \tag{3.3.14}$$

or

$$(D-\lambda_1)y_{n-1} = f(t), \quad t > 0. \tag{3.3.15}$$

(15) is a first order linear equation for $y_{n-1}$. Having solved it, we may insert $y_{n-1}(t)$ into the last equation of the set (13) and solve for $y_{n-2}(t)$. Proceeding thus up the set of equations (13) we eventually reach the equation

$$(D-\lambda_n)x = y_1(t), \tag{3.3.16}$$

where the right-hand side has been found as a function of $t$. The solution of (16) is the required solution of (14). Each step, containing an integration, will introduce an arbitrary constant, which remains in the solution if not determined by an appropriate boundary condition. We could then obtain the general solution of (14) containing $n$ arbitrary constants. But if, as is more usual, the problem is completely specified by the statement of initial conditions as well as the differential equation, these conditions can be introduced during the process of solution, as follows. Let the conditions at $t = 0$ be

$$\begin{aligned}x(0) &= b_0,\\ \dot{x}(0) &= b_1,\\ &\cdots\\ x^{(n-1)}(0) &= b_{n-1}.\end{aligned}$$

Then

$$\begin{aligned}y_1(0) &= y_{10} = b_1-\lambda_n b_0,\\ y_2(0) &= y_{20} = b_2-(\lambda_n+\lambda_{n-1})b_1+\lambda_n\lambda_{n-1}b_0,\\ &\cdots\\ y_{n-1}(0) &= y_{(n-1)0} = b_{n-1}\ldots+(-1)^{n-1}\lambda_n\lambda_{n-1}\ldots\lambda_2 b_0.\end{aligned}$$

Eliminating $q_2$ we have

$$\{(LD^2+RD+1/C)^2-(MD^2)^2\}q_1 = 0,$$

i.e.

$$\{(L-M)D^2+RD+1/C\}\{(L+M)D^2+RD+1/C\}q_1 = 0.$$

To simplify the algebra, we consider the case when $R\dot{q}_1$ can be neglected in comparison with $(L \pm M)\ddot{q}_1$ and $q_1/C$. Then writing

$$\frac{1}{C(L-M)} = m^2, \qquad \frac{1}{C(L+M)} = n^2 \tag{3.3.20}$$

(we have assumed $L \neq M$) we obtain

$$(D^2+m^2)(D^2+n^2)q_1 = 0, \quad t > 0. \tag{3.3.21}$$

Substituting

$$(D^2+n^2)q_1 = y \tag{3.3.22}$$

we have

$$(D^2+m^2)y = 0, \quad t > 0, \tag{3.3.23}$$

and we need the initial values of $y$ and $\dot{y}$. Substituting the initial values into the differential equations (19) we obtain

$$L\ddot{q}_1(0)+M\ddot{q}_2(0) = -Q/C,$$

$$M\ddot{q}_1(0)+L\ddot{q}_2(0) = 0,$$

so that

$$\ddot{q}_1(0) = -\frac{LQ}{C(L^2-M^2)}.$$

If we differentiate the equations (19) and substitute the initial conditions, we find (taking $R = 0$)

$$L\dddot{q}_1(0)+M\dddot{q}_2(0) = 0,$$

$$M\dddot{q}_1(0)+L\dddot{q}_2(0) = 0,$$

so that, since $L \neq M$, $\dddot{q}_1(0) = 0$.

Then

$$y(0) = \ddot{q}_1(0)+n^2Q = -\frac{MQ}{C(L^2-M^2)}, \tag{3.3.24}$$

$$\dot{y}(0) = \dddot{q}_1(0)+n^2\dot{q}_1(0) = 0. \tag{3.3.25}$$

For the solution of the equation (23) for $y$, with the initial conditions (24) and (25), we may quote (2.31), and find

$$y(t) = y_0 \cos mt.$$

Then (22) is

$$(D^2+n^2)q_1 = y_0 \cos mt,$$

with $q_1 = Q$, $\dot{q}_1 = 0$ when $t = 0$. Again using (2.31) we obtain

$$q_1(t) = Q \cos nt + \frac{1}{n}\int_0^t y_0 \cos mu \sin n(t-u)\mathrm{d}u.$$

Evaluating the integral and simplifying we reach the result

$$q_1(t) = \tfrac{1}{2}Q(\cos mt + \cos nt), \tag{3.3.26}$$

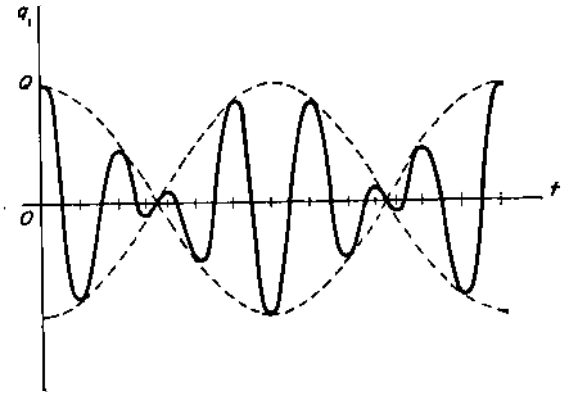


FIG. 3.5. Motion compounded of two simple harmonic motions

$$q_1(t) = \tfrac{1}{2}Q(\cos mt + \cos nt)$$
$$= Q\left(\cos\frac{m-n}{2}t \cos\frac{m+n}{2}t\right).$$

Here $m+n = 6(m-n)$.

so that $q_1$ oscillates with a motion compounded of two modes of periods $2\pi/m$ and $2\pi/n$. If we write this as

$$q_1(t) = Q \cos\frac{m-n}{2}t \cos\frac{m+n}{2}t, \tag{3.3.27}$$

we can regard (27) as a harmonic oscillation of period $4\pi/(m+n)$ and of varying amplitude $Q\cos\frac{1}{2}(m-n)t$. If $m$ and $n$ differ only slightly, the amplitude will vary slowly, as shown in Fig. 3.5. The motion is then an oscillation with beats.

## 3.4. Systems of linear equations with constant coefficients: particular examples

*Consecutive chemical reactions of the first order*

Consider a reaction in which a substance $A_1$ breaks down with rate constant $k_1$ to form the substance $A_2$ which in turn breaks down at rate $k_2$ to form $A_3$. If $x_1$, $x_2$, $x_3$ denote the concentrations of $A_1$, $A_2$, $A_3$ at time $t$, so that initially $x_1 = x_{10}, x_2 = 0, x_3 = 0$, the equations describing the process are

$$\left.\begin{aligned} \dot{x}_1 &= -k_1 x_1, \\ \dot{x}_2 &= k_1 x_1 - k_2 x_2, \\ \dot{x}_3 &= k_2 x_2, \end{aligned}\right\} \quad t > 0, \tag{3.4.1}$$

with $x_1(0) = x_{10}$, $x_2(0) = 0$, $x_3(0) = 0$.

The first equation (1) is solved as in Section 3.1 to give

$$x_1 = x_{10} e^{-k_1 t}. \tag{3.4.2}$$

Substituting this value into the second equation, we have the first order linear equation

$$\dot{x}_2 + k_2 x_2 = k_1 x_{10} e^{-k_1 t}. \tag{3.4.3}$$

The solution of this equation has been found in (1.7) and (1.8). It is

$$x_2(t) = \frac{k_1 x_{10}}{k_2 - k_1} (e^{-k_1 t} - e^{-k_2 t}). \tag{3.4.4}$$

We could substitute this value (4) into the third equation of (1) and integrate to find $x_3(t)$. But if we observe that by adding the three differential equations we get

$$\frac{d}{dt}(x_1 + x_2 + x_3) = 0,$$

and hence

$$x_1 + x_2 + x_3 = x_{10},$$

we may deduce

$$\begin{aligned} x_3(t) &= x_{10} - x_1 - x_2 \\ &= x_{10}\left\{1 - e^{-k_1 t} - \frac{k_1}{k_2 - k_1}(e^{-k_1 t} - e^{-k_2 t})\right\}. \end{aligned} \tag{3.4.5}$$

It would be easy if desired to modify these results for more general initial conditions $x_1(0) = x_{10}$, $x_2(0) = x_{20}$, $x_3(0) = x_{30}$.

The sequence of events in radioactive breakdown is of this kind, but may involve many steps. The method just given may be extended to cover the whole sequence to the final product. A slightly different method may be preferable if the boundary conditions of the problem are not straightforward initial conditions. Suppose that the process is described by the differential equations

$$\begin{aligned} \dot{x}_1 &= -k_1 x_1, \\ \dot{x}_2 &= k_1 x_1 - k_2 x_2, \\ &\cdots \\ \dot{x}_r &= k_{r-1} x_{r-1} - k_r x_r, \\ &\cdots \\ \dot{x}_n &= k_{n-1} x_{n-1}, \end{aligned} \tag{3.4.6}$$

where all the $k$'s are different.

We rewrite these equations as

$$\begin{aligned} (D+k_1)x_1 &= 0, \\ (D+k_2)x_2 &= k_1 x_1, \\ &\cdots \\ (D+k_r)x_r &= k_{r-1} x_{r-1}, \\ &\cdots \\ Dx_n &= k_{n-1} x_{n-1}. \end{aligned} \tag{3.4.7}$$

Operating on the second equation of (7) by $D+k_1$ we obtain

$$(D+k_1)(D+k_2)x_2 = 0.$$

Proceeding in this way we obtain

$$(D+k_1)(D+k_2)\ldots(D+k_r)x_r = 0, \tag{3.4.8}$$

and

$$(D+k_1)(D+k_2)\ldots(D+k_{n-1})Dx_n = 0. \tag{3.4.9}$$

The general solution of (9) is

$$x_n = a_1 e^{-k_1 t} + a_2 e^{-k_2 t} + \ldots + a_{n-1} e^{-k_{n-1} t} + a_n, \tag{3.4.10}$$

where $a_1, a_2, \ldots, a_n$ are arbitrary constants. We can choose them to fit the boundary conditions.

We find

$$x_{n-1} = \frac{1}{k_{n-1}} Dx_n,$$

$$x_{n-2} = \frac{1}{k_{n-2}} (D + k_{n-1})x_{n-1},$$

. . . . . .

and thus all of $x_n, x_{n-1}, \ldots, x_1$ are found in terms of the expressions occurring in (10). Inserting, for instance, the initial values of $x_1, x_2, \ldots, x_n$ we have $n$ equations from which the parameters $a_1, a_2, \ldots, a_n$ are uniquely determined.

*General system of first order reactions*

Rather than treat separately various problems that have arisen or might arise and lead to systems of first order differential equations, we now examine a method available for all such problems. The problem is set up in a general and complete manner.

If we have a general system of $n$ components $A_1, A_2, \ldots, A_n$ in which any two components may be related by simple opposing reactions at different rates, i.e.

$$A_r \to A_s, \text{ rate constant } k_{rs},$$

$$A_s \to A_r, \text{ rate constant } k_{sr},$$

the differential equations describing the whole system will be, if $x_r$ is the concentration of $A_r$, and $t > 0$,

$$\begin{aligned}
\dot{x}_1 &= -k_{12}x_1 - k_{13}x_1 \ldots - k_{1n}x_1 + k_{21}x_2 + k_{31}x_3 + \ldots + k_{n1}x_n \\
&\quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \\
\dot{x}_r &= -k_{r1}x_r - k_{r2}x_r \ldots - k_{r,r-1}x_r - k_{r,r+1}x_r \ldots - k_{rn}x_r \\
&\quad + k_{1r}x_1 + k_{2r}x_2 \ldots + k_{r-1,r}x_{r-1} + k_{r+1,r}x_{r+1} \ldots + k_{nr}x_n \\
&\quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots
\end{aligned}$$

These may be expressed more neatly if we introduce an auxiliary constant $k_{rr}$ defined by

$$k_{rr} = -k_{r1} - k_{r2} \ldots - k_{r,r-1} - k_{r,r+1} \ldots - k_{rn}, \qquad (3.4.11)$$

so that

$$\dot{x}_r = \sum_{s=1}^{n} k_{sr}x_s. \qquad (3.4.12)$$

We now introduce the *summation convention*, by which if a suffix occurs twice in any term the summation of such terms for values 1, 2, . . ., $n$ of the suffix is implied: thus

$$a_{rr} \equiv \sum_{r=1}^{n} a_{rr} = a_{11} + a_{22} + \ldots + a_{nn},$$

$$a_{rs}b_s \equiv \sum_{s=1}^{n} a_{rs}b_s = a_{r1}b_1 + a_{r2}b_2 + \ldots + a_{rn}b_n. \tag{3.4.13}$$

We are thus enabled to omit the sign

$$\sum_{r=1}^{n},$$

but we must always remember that the doubled suffix implies summation and that one term represents $n$ terms. Then (12) is

$$\dot{x}_r = k_{sr}x_s. \tag{3.4.14}$$

This simple equation represents a set of $n$ equations ($r = 1, 2, \ldots, n$) in each of which the left-hand side contains one term and the right-hand side $n$ terms.

Guided by previous experience, we seek a solution in terms of exponential functions: we try $x_r = a_r \exp(-\lambda t)$, and see whether there is any value of $\lambda$ which enables the set of values $x_r$ to satisfy the differential equations (at present we leave aside the boundary conditions; the satisfying of the boundary conditions is a further step in the method). Then

$$-\lambda a_r \exp(-\lambda t) = k_{sr}a_s \exp(-\lambda t), \quad t > 0. \tag{3.4.15}$$

The work that follows can be most elegantly formulated if at this stage we introduce the *Kronecker symbol* $\delta_{rs}$, which is defined by

$$\delta_{rs} = 1 \text{ if } r = s,$$

$$\delta_{rs} = 0 \text{ if } r \neq s. \tag{3.4.16}$$

Then

$$\delta_{rs}a_s = a_r, \tag{3.4.17}$$

since the only term of the sum on the left-hand side that is not zero is the one for which $s = r$. On account of this property, $\delta_{rs}$ is often

called the *substitution operator*. Using (17) in (15) we obtain, for all $t > 0$, since $\delta_{sr} = \delta_{rs}$,

$$(k_{sr} + \lambda\delta_{sr})a_s \exp(-\lambda t) = 0.$$

This can be true for all $t$ only if for each $r$

$$(k_{sr} + \lambda\delta_{sr})a_s = 0. \tag{3.4.18}$$

Here we have a set of $n$ homogeneous linear algebraic equations, each containing $n$ terms (since a sum over $s$ is implied). They will be consistent only if the determinant of the coefficients is zero, i.e. if

$$\det(k_{sr} + \lambda\delta_{sr}) = 0. \tag{3.4.19}$$

This is an equation of degree $n$ in $\lambda$, as is clearly understood if we write it out in full as

$$\begin{vmatrix} k_{11}+\lambda & k_{21} & k_{31} & . & k_{n1} \\ k_{12} & k_{22}+\lambda & k_{32} & . & k_{n2} \\ . & . & . & . & . \\ k_{1n} & k_{2n} & k_{3n} & . & k_{nn}+\lambda \end{vmatrix} = 0. \tag{3.4.20}$$

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the roots of (19). If we substitute $\lambda_1$ back into (18) we obtain a consistent set of equations from which the corresponding *ratios* $a_1 : a_2 : \ldots : a_n$ can be determined. Thus each $a_s$ will be determined except for an arbitrary factor common for all $a_s$. Let the solution corresponding to $\lambda_1$ be

$$a_s = B_1 a_s^{(1)},$$

where $B_1$ is arbitrary but $a_s^{(1)}$ have been found in terms of the rate constants $k_{sr}$ and the root $\lambda_1$ (which is also known in terms of $k_{sr}$). Then a solution of the system of differential equations (12) is the set

$$x_r^{(1)} = B_1 a_r^{(1)} e^{-\lambda_1 t}. \tag{3.4.21}$$

Similarly a second solution is

$$x_r^{(2)} = B_2 a_r^{(2)} e^{-\lambda_2 t}.$$

In all, if there are $n$ distinct roots $\lambda_p$, $p = 1, 2, \ldots, n$, there will be $n$ linearly independent solutions

$$x_r^{(p)} = B_p a_r^{(p)} e^{-\lambda_p t}, \tag{3.4.22}$$

from which we can build the general solution

$$x_r = \sum_{p=1}^{n} B_p a_r^{(p)} e^{-\lambda_p t}. \tag{3.4.23}$$

We observe that one of the roots of (19) is always $\lambda = 0$. For if we form a new first row of the determinant (20) by adding all rows, then every element of that row is $\lambda$ (by use of the definition (11) of $k_{rr}$). $\lambda$ is thus a factor of the determinant, and $\lambda = 0$ is a root. The corresponding solution (22) is a set of constant values $B_0 a_r^{(0)}$. We may take $\lambda_n$ to be the root which is zero, so that $B_0 a_r^{(0)}$ replaces $B_n a_r^{(n)} e^{-\lambda_n t}$ in the sum (23).

So far we have not introduced boundary conditions. Let us use initial conditions

$$x_r = x_{r0} \text{ at } t = 0.$$

These $n$ conditions will be sufficient to determine the $n$ arbitrary constants $B_p$ according to the system of equations

$$x_{r0} = \sum_{p=1}^{n} B_p a_r^{(p)}. \tag{3.4.24}$$

(23) then gives the final solution.

It will be found that none of the roots $\lambda_p$ is real and negative. This corresponds to the physical fact that we are dealing with natural decay. Negative $\lambda_p$, corresponding to increase without limit as $t \to \infty$, could not represent a real situation.

*Coupled circuits*

We can now give a more powerful method of solving problems of coupled electric circuits. Let us take the same arrangement of two circuits as in Section 3 (p. 70), but introduce into the first circuit an E.M.F. of given time-dependence $E(t)$, which starts to operate at $t = 0$ when charges and currents are zero. We wish to find the subsequent current in the second circuit. With the same notation as before, our equations replacing (3.16) are

$$\left.\begin{aligned} L\dot{x}_1 + M\dot{x}_2 + Rx_1 + q_1/C &= E(t), \\ M\dot{x}_1 + L\dot{x}_2 + Rx_2 + q_2/C &= 0. \end{aligned}\right\} \quad t > 0. \tag{3.4.25}$$

The initial conditions are now $q_1 = 0$, $x_1 = 0$, $q_2 = 0$, $x_2 = 0$ at $t = 0$. Then equations (3.18) are replaced by

$$\left.\begin{aligned} (LD^2 + RD + 1/C)q_1 + MD^2 q_2 &= E(t), \\ (LD^2 + RD + 1/C)q_2 + MD^2 q_1 &= 0. \end{aligned}\right\} \tag{3.4.26}$$

The second equation has been rearranged so that the next step seems a natural one. If we now add the two equations (26), writing $q_1+q_2 = u$, and then subtract them, writing $q_1-q_2 = v$ we find

$$(LD^2+RD+1/C)u+MD^2u = E,$$
$$(LD^2+RD+1/C)v-MD^2v = E.$$

We have thus separated the two dependent variables. As before, we take $R = 0$ for simpler algebra and introduce $m$ and $n$ defined by

$$m^{-2} = C(L+M), \qquad n^{-2} = C(L-M),$$

to get

$$\ddot{u}+m^2u = Cm^2E,$$
$$\ddot{v}+n^2v = Cn^2E. \tag{3.4.27}$$

By this manoeuvre we have obtained two second-order equations instead of the one fourth-order equation we found before. The initial conditions give $u = 0$, $\dot{u} = 0$, $v = 0$, $\dot{v} = 0$ at $t = 0$. So using the solution (2.31) we have

$$u = Cm \int_0^t E(\tau) \sin m(t-\tau)\mathrm{d}\tau,$$

$$v = Cn \int_0^t E(\tau) \sin n(t-\tau)\mathrm{d}\tau,$$

and the answer to the original question is provided by

$$q_2 = \tfrac{1}{2}(u-v) = \tfrac{1}{2}C \int_0^t E(\tau)\{m \sin m(t-\tau)-n \sin n(t-\tau)\}\mathrm{d}\tau,$$

$$x_2 = \frac{\mathrm{d}}{\mathrm{d}t} q_2 = \tfrac{1}{2}C \int_0^t E(\tau)\{m^2 \cos m(t-\tau)-n^2 \cos n(t-\tau)\}\mathrm{d}\tau. \tag{3.4.28}$$

We observe from the equations (27) that in the absence of external disturbing E.M.F. $E(t)$, the system is capable of free oscillations with periods $2\pi/m$ and $2\pi/n$. These two harmonic oscillations are not coupled: the system can oscillate in one mode alone if suitably started. Such a mode is called a *normal mode* of the system, and a variable describing the motion in a certain normal mode is called the corresponding *normal coordinate*. Normal modes and normal coordinates describe the intrinsic behaviour of the system when oscillating freely.

Our method of changing the independent variables so as to produce uncoupled differential equations is in fact a way of finding the normal coordinates of the system. If the normal coordinates cannot be seen by inspection we can find them as follows. Let us examine a system in which the circuits have self-inductances $L_1$, $L_2$, capacitances $C_1$, $C_2$. Then with the same notation as before, we have to solve

$$L_1\ddot{q}_1 + M\ddot{q}_2 + q_1/C_1 = E,$$
$$M\ddot{q}_1 + L_2\ddot{q}_2 + q_2/C_2 = 0. \tag{3.4.29}$$

We seek a linear combination of $q_1$ and $q_2$ which obeys a single second-order differential equation. Let us multiply the first equation by $\beta$ and add it to the second equation. Then

$$D^2\{(\beta L_1 + M)q_1 + (\beta M + L_2)q_2\} + \{\beta q_1/C_1 + q_2/C_2\} = \beta E, \tag{3.4.30}$$

and the two quantities in curved brackets must be multiples of the same combination of $q_1$ and $q_2$. Hence the ratio of coefficients in the first bracket must equal the ratio of coefficients in the second, i.e.

$$\frac{\beta L_1 + M}{\beta M + L_2} = \frac{\beta C_2}{C_1},$$

or

$$\beta^2 MC_2 + \beta(L_2C_2 - L_1C_1) - MC_1 = 0. \tag{3.4.31}$$

This is a quadratic for $\beta$, with roots $\beta_1$, $\beta_2$. Let us use the value $\beta_1$ and insert it into the differential equation (30). Then since

$$\beta_1 L_1 + M = \frac{\beta_1 C_2}{C_1}(\beta_1 M + L_2),$$

the equation is

$$(\beta_1 M + L_2)D^2\left\{\frac{\beta_1 C_2}{C_1}q_1 + q_2\right\} + \frac{1}{C_2}\left\{\frac{\beta_1 C_2}{C_1}q_1 + q_2\right\} = \beta_1 E,$$

i.e.

$$\ddot{u} + m^2 u = \frac{\beta_1}{\beta_1 M + L_2}E, \tag{3.4.32}$$

where

$$u = \frac{\beta_1 C_2}{C_1}q_1 + q_2 \quad \text{and} \quad m^2 = \frac{1}{C_2(\beta_1 M + L_2)}.$$

$u$ is then one normal coordinate; the other, found from $\beta_2$, is

$$v = \frac{\beta_2 C_2}{C_1} q_1 + q_2, \tag{3.4.33}$$

satisfying

$$\ddot{v} + n^2 v = \frac{\beta_2}{\beta_2 M + L_2}, \quad \text{with} \quad n^2 = \frac{1}{C_2(\beta_2 M + L_2)}.$$

The initial conditions (or other boundary conditions) can now be expressed in terms of $u$ and $v$, and the problem solved by straightforward integration of equations (32) and (33). The variables $u$ and $v$ are now the normal coordinates, and the corresponding normal modes of oscillation have periods $2\pi/m$ and $2\pi/n$. As stated above, the system can be started so as to oscillate with simple harmonic motion in one of the normal modes (the corresponding normal coordinate alone varying). But for general conditions of starting both modes will be excited and the variation of $q_1$ or $q_2$ will be obtained by a superposition of normal modes with appropriate amplitudes.

*Orbital motion with gyroscopic disturbing force*

Another application of the method just introduced is to the motion of a particle in a plane when elastic and gyroscopic terms occur in the differential equations. These are (from (1.2.2))

$$\begin{aligned} \ddot{x} - 2\gamma\dot{y} + \lambda x &= 0, \\ \ddot{y} + 2\gamma\dot{x} + \lambda y &= 0. \end{aligned} \tag{3.4.34}$$

The form of (34) is such that we can combine the two equations into one equation for $z = x + iy$, $(i = \sqrt{(-1)})$, by multiplying the second equation by $i$ and adding to the first. Then

$$\ddot{z} + 2i\gamma\dot{z} + \lambda z = 0. \tag{3.4.35}$$

This is a second order linear equation with constant coefficients. We therefore seek a solution

$$z \propto \exp i\eta t,$$

where we have introduced a factor $i$ into the exponent for algebraic convenience. Substituting, we get

$$\eta^2 + 2\gamma\eta - \lambda = 0,$$

so that $\eta = -\gamma \pm \sqrt{(\gamma^2+\lambda)}$, and the general solution of (35) is

$$z = e^{-i\gamma t}(Ae^{i\mu t} + Be^{-i\mu t}), \tag{3.4.36}$$

where $\mu = \sqrt{(\gamma^2+\lambda)}$, and $A$ and $B$ are constants which may be complex. $A$ and $B$ will be determined by four independent conditions. For example, let us take

$$x(0) = a, \qquad \dot{x}(0) = 0,$$
$$y(0) = 0, \qquad \dot{y}(0) = V.$$

Then $z(0) = a$, $\dot{z}(0) = iV$. Substituting for $z$ from (36) we have

$$a = A + B,$$
$$V = (\mu-\gamma)A - (\mu+\gamma)B,$$

so that

$$A = \tfrac{1}{2}a + \tfrac{1}{2}\frac{\gamma a + V}{\mu}, \quad B = \tfrac{1}{2}a - \tfrac{1}{2}\frac{\gamma a + V}{\mu}.$$

Hence

$$z = e^{-i\gamma t}\left[a\cos\mu t + i\frac{\gamma a + V}{\mu}\sin\mu t\right]. \tag{3.4.37}$$

The motion represented by (37) is best envisaged as follows.

Write

$$\zeta = \xi + i\eta = a\cos\mu t + i\frac{\gamma a + V}{\mu}\sin\mu t. \tag{3.4.38}$$

Then

$$\frac{\xi^2}{a^2} + \frac{\eta^2}{(\gamma a/\mu + V/\mu)^2} = 1,$$

so that $\zeta$ moves on an ellipse. But $z = e^{-i\gamma t}\zeta$, so that at any time $t$, $z$ is obtained from $\zeta$ by rotating $\zeta$ clockwise about the origin through the angle $\gamma t$. Hence the point $(x, y)$ has a motion which is to be envisaged as periodic motion on an ellipse which rotates with angular velocity $\gamma$ about its centre.

The motion of the tip of a Foucault pendulum has this character, the gyroscopic force being the coriolis force due to the Earth's rotation. The motion of an orbiting electron when a transverse magnetic field is applied provides a further example (see Exercise 3.9 on the Zeeman effect).

## 3.5. Systems of inhomogeneous linear first order equations with constant coefficients: general theory

The general system of linear first order differential equations with constant coefficients when there are $n$ dependent variables has the form

$$\begin{aligned}
&a_{11}\dot{x}_1 + a_{12}\dot{x}_2 + \ldots + a_{1n}\dot{x}_n + b_{11}x_1 + b_{12}x_2 + \ldots + b_{1n}x_n = f_1(t),\ t > 0,\\
&a_{21}\dot{x}_1 + a_{22}\dot{x}_2 + \ldots + a_{2n}\dot{x}_n + b_{21}x_1 + b_{22}x_2 + \ldots + b_{2n}x_n = f_2(t),\\
&\qquad\qquad\qquad\qquad \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\\
&a_{n1}\dot{x}_1 + a_{n2}\dot{x}_2 + \ldots + a_{nn}\dot{x}_n + b_{n1}x_1 + b_{n2}x_2 + \ldots + b_{nn}x_n = f_n(t).
\end{aligned} \tag{3.5.1}$$

Using the summation convention we write this as

$$a_{rs}\dot{x}_s + b_{rs}x_s = f_r(t), \quad r, s = 1, 2, \ldots, n, \quad t > 0, \tag{3.5.2}$$

or

$$(a_{rs}D + b_{rs})x_s = f_r, \qquad t > 0.$$

The problem will be completely specified if the $n$ initial values are given:

$$x_s(0) = x_{s0}.$$

Let us denote the operator $a_{rs}D + b_{rs}$ by $c_{rs}$, and introduce the determinant of operators

$$C = \det c_{rs} = \begin{vmatrix} c_{11} & c_{12} & . & . & c_{1n} \\ c_{21} & c_{22} & . & . & c_{2n} \\ . & . & . & . & . \\ c_{n1} & c_{n2} & . & . & c_{nn} \end{vmatrix}. \tag{3.5.3}$$

Let $C_{rs}$ denote the cofactor of $c_{rs}$ in this determinant (i.e. $(-1)^{r+s}$ times the determinant obtained by striking out the $r$th row and $s$th column). If each element of a certain column of $C$ is multiplied by its cofactor, and the products are summed, the sum is $C$. But if each element of a column is multiplied by the cofactor of the corresponding element of a different column, the sum of those products, being a determinant with two equal columns, is zero. Using these properties, we now solve the system of equations (2), i.e.

$$c_{rs}x_s = f_r, \tag{3.5.4}$$

for $x_s$. Multiply the $r$th equation by $C_{rs}$ and sum over $r$. All terms on the left-hand side will vanish except that containing $x_s$. We obtain

$$Cx_s = C_{rs}f_r. \tag{3.5.5}$$

$C$ when expanded is seen to be a polynomial in $D$ of degree $n$, and $C_{rs}$ a polynomial of degree $n-1$. Thus the right-hand side of (5) is a known function of $t$, and the left-hand side can in principle be expressed as the product of $n$ linear factors in $D$. Let

$$C = C_0(D-\lambda_1)(D-\lambda_2)\ldots(D-\lambda_n), \tag{3.5.6}$$

where $C_0 = \det a_{rs}$. Then the general solution of (5) is

$$x_s = A_{sq}e^{\lambda_q t} + X_s(t), \tag{3.5.7}$$

where $A_{sq}$ are arbitrary constants and it is understood that the first term on the right is a sum from $q = 1$ to $q = n$, and where $X_s(t)$ is any particular integral of (5).

If this process is performed for each $x_s$, we obtain the same set $\lambda_q$ every time, but the constants $A_{sq}$ and the particular integrals $X_s(t)$ will be different for different $s$. We thus have $n$ variables expressed in terms of $n^2$ constants. We shall show that these $n^2$ constants can be expressed in terms of $n$ constants. Substituting from (7) into (2), we obtain

$$\begin{aligned}\sum_{q=1}^{n}(a_{rs}\lambda_q + b_{rs})A_{sq}e^{\lambda_q t} &= f_r(t) - (a_{rs}D + b_{rs})X_s \\ &= 0,\end{aligned} \tag{3.5.8}$$

since $X_s$ is a particular integral. Now (8) must hold for all $t > 0$, and so the coefficients on the left must vanish. Thus for each $q$, say $q_p \equiv q'$,

$$(a_{rs}\lambda_{q'} + b_{rs})A_{sq'} = 0, \quad r = 1, 2, \ldots, n,$$

(not summed over $q'$). Here we have $n$ equations which determine the ratios of

$$A_{1q'}, A_{2q'}, \ldots, A_{nq'}, \tag{3.5.9}$$

and therefore determine each constant of the set (9) in terms of one of them, say $A_{1q'}$.

Similarly for every $q$, the values of $A_{2q}, A_{3q}, \ldots, A_{nq}$ may be determined in terms of $A_{1q}$. We are thus left with $n$ undetermined constants $A_{11}, A_{12}, \ldots, A_{1n}$ and the differential equations are satisfied. If now we put in the $n$ conditions $x_s(0) = x_{s0}$ we can determine the values of $A_{1s}$. Thus the solution of the problem is obtained. The reason for the apparent superfluity of constants at one stage is that equation (5) was obtained by introducing $n-1$ further differentiations

—in the original set only the first derivative of $x_s$ had been involved. Thus when the equations (7) were obtained they referred not to the original equations but to these derived equations. Re-imposing the conditions that the original set of equations must be satisfied we were able to get rid of the extra constants and leave only enough arbitrary parameters for the satisfying of the original boundary conditions.

We have given the method for the general case in which the values are distinct. If some roots are repeated the solution must be modified in a way which is clear enough in a particular problem: it is unnecessary to write out the various modifications in the general case.

An example of the use of this method will be found in Exercise 3.9. The reader must understand that the value of the method lies in its systematic attack on problems of considerable complexity: to use it on a simple problem of two dependent variables is to use a sledgehammer to crack a nut. But the application to a simple problem does enable the student to obtain a closer and more definite view of the steps in the process of solution of the equations.

## Exercises

**3.1.** A circuit of negligible capacitance contains resistance $R$, self-inductance $L$ and electromotive force $E\cos\omega t$. At time $t = 0$ it is switched on. Set up the differential equation for the current $I$ in the circuit and solve with appropriate initial conditions. Show that after the transient has died down the current oscillates with the period of the electromotive force but out of phase with it by $\tan^{-1}(\omega L/R)$: find the amplitude of this current.

**3.2.** Show how to derive (3.2.8) from (3.2.4) by a limiting process in which we put $\alpha_2 = \alpha_1 + \delta$ and let $\delta \to 0$.

**3.3.** (a) From the result (3.2.19) for the concentration of chlorine atoms, show that the average concentration is

$$\frac{n}{k}\left(1 - \frac{2}{ml}\tanh\frac{ml}{2}\right).$$

(b) Given that the condition $c = 0$ at $x = 0$, $l$ is replaced by $F = \mp\gamma$ at $x = 0, l$, where $\gamma$ is a constant (this condition specifies removal of chlorine atoms at a constant rate at the walls), show that

$$c = \frac{n}{k} - \frac{\gamma}{mD}\left\{\frac{\cosh mx + \cosh m(l-x)}{\cosh ml}\right\}.$$

**3.4.** The particle whose motion is described by the differential equation (3.2.20) starts from rest at the origin, and the force $mf(t)$ is described by

$$\begin{aligned} f(t) &= 0, \quad t < 0, \\ &= P/\varepsilon, 0 < t < \varepsilon, \\ &= 0, \quad \varepsilon < t, \end{aligned}$$

where $\varepsilon$ is constant. Show that the position of the particle after time $t$ is given by

$$x(t) = \frac{1}{\alpha_1 - \alpha_2} \int_0^{\varepsilon} \frac{P}{\varepsilon} \{e^{\alpha_1(t-u)} - e^{\alpha_2(t-u)}\} du.$$

Show that as $\varepsilon \to 0$, $x(t)$ tends to the same expression as is found when the particle starts with initial velocity $P$ and no force acts subsequently. Why is this?

**3.5.** *Some particular integrals.*
Given the equation

$$\ddot{x} - 3\dot{x} + 2x = f(t),$$

find a particular integral when

(a) $f(t) = \exp 3t$ (try $x = C \exp 3t$ and determine $C$),

(b) $f(t) = \sin 3t$ (try $x = C \cos 3t + D \sin 3t$ and determine $C$ and $D$),

(c) $f(t) = t^2$ (try $x = Ct^2 + Dt + E$ and determine $C$, $D$ and $E$),

(d) $f(t) = e^{-t} \cos 3t$ (write $e^{-t} \cos 3t = \mathrm{Re}\{e^{(-1+3i)t}\}$ and try $x = C \exp\{(-1+3i)t\}$),

(e) show that such methods break down if $f(t) = e^t$ or $f(t) = e^{2t}$, and in general if $f(t)$ is a solution of the auxiliary equation. (In such cases use the second method given in Section 3.2.)

**3.6.** Using the *second method* of Section 3.3, solve

$$(D^4 - n^4)x = \cos nt,$$

given that when $t = 0$, $x = 1$, $Dx = 0$, $D^2x = 0$, $D^3x = 0$.

(Hint: write

$$(D^2 + n^2)x = y, \ (D^2 - n^2)y = \cos nt,$$

and prove

$$x = \tfrac{1}{2}\left(1 - \frac{1}{2n^4}\right) \cos nt - \frac{1}{4n^3} t \sin nt - \tfrac{1}{2}\left(1 + \frac{1}{2n^4}\right) \cosh nt.)$$

**3.7.** Discuss the problem of a system of first order reactions (Section 3.4) for the special case $n = 2$, showing that the values of $\lambda$ are then zero and $(k_{12} + k_{21})$.
Discuss the similar problem for $n = 3$, showing that the cubic for $\lambda$ is

$$\lambda\{\lambda^2 - \lambda(k_{23} + k_{32} + k_{31} + k_{13} + k_{12} + k_{21}) + (k_{21} + k_{12} + k_{23})(k_{31} + k_{13} + k_{21}) - (k_{23} - k_{13})(k_{32} - k_{12})\} = 0,$$

and hence that the non-zero roots have $\mathrm{Re}(\lambda) > 0$.

**3.8.** A particle of mass $m$ moves in a central field of force, the force being $-m\lambda\mathbf{r}$ (where $\lambda$ is constant) when the particle is at $\mathbf{r}$. Prove that the particle moves in one plane and that its orbit is an ellipse with centre at the origin.
A magnetic field $\mathbf{H}$ is then imposed, and it is given that the particle carries charge $e$ and experiences a force $e\dot{\mathbf{r}} \times \mathbf{H}/c$ due to the magnetic field ($c$ being the speed of light). If $\mathbf{H}$ is perpendicular to the plane of the original ellipse, prove that its effect is to make the orbit rotate with angular velocity $eH/2mc$ (Classical Zeeman effect).

**3.9.** Carry through the analysis of Section 3.5 for the pair of equations

$$(LD + R)x_1 + MDx_2 = Ee^{-\mu t},$$
$$MDx_1 + (LD + R)x_2 = 0$$

(where $L$, $M$, $R$, $E$, $\mu$ are constants) which describe coupled circuits. Obtain the general solution (valid provided $L \neq M$ and $(L \pm M)\mu \neq R$)

$$x_1 = A_{11}e^{\lambda_1 t} + A_{12}e^{\lambda_2 t} + \frac{R-\mu L}{L^2-M^2}\,\frac{Ee^{-\mu t}}{(\mu+\lambda_1)(\mu+\lambda_2)},$$

$$x_2 = -\frac{R+\lambda_1 L}{\lambda_1 M}A_{11}e^{\lambda_1 t} - \frac{R+\lambda_2 L}{\lambda_2 M}e^{\lambda_2 t} + \frac{\mu M}{L^2-M^2}\,\frac{Ee^{-\mu t}}{(\mu+\lambda_1)(\mu+\lambda_2)},$$

where $$\lambda_1 = -\frac{R}{L-M}, \qquad \lambda_2 = -\frac{R}{L+M}.$$

CHAPTER 4

# LAPLACE TRANSFORM METHOD OF SOLUTION OF ORDINARY LINEAR DIFFERENTIAL EQUATIONS WITH CONSTANT COEFFICIENTS

## 4.1. Introduction

An unsatisfactory feature of the methods so far described for solving ordinary linear differential equations with constant coefficients is the difficulty of finding particular integrals. This difficulty is highlighted in Section 3.3, where our second method gives the particular integral for an $n$th order equation as an $n$-ple integral: the first method is even more vulnerable at this point, since it depends on experienced guesswork or on manipulations of doubtful validity and inadequate scope (not treated in this book).

A second difficulty—if the problem is solved by our second method—is the cumbersome form taken by the initial conditions that are to be inserted at each integration. Here again, as far as marching problems are concerned, the first method makes still heavier weather. The solution is first written in general form, with a particular integral (assumed known) and $n$ arbitrary constants multiplying linearly independent solutions of the auxiliary equation. Then the initial conditions are imposed and the adjustable constants determined by solving $n$ simultaneous linear equations.

Both of these difficulties can be met, and reduced as much as possible, by a method of solution essentially due to Oliver Heaviside but now developed in terms of Laplace transforms. The crux of the method is the reduction of a differential equation for $x(t)$ to an algebraic equation for its *transform* $\bar{x}(p)$, which is derived from $x(t)$ as shown in (4.2.2) below. The algebraic equation incorporates the information supplied by the initial conditions. Its solution for $\bar{x}(p)$ is straightforward, and there remains the derivation of $x(t)$ from its transform $\bar{x}(p)$. In Section 4.2 the method is expounded in its simplest application—to the linear equation of the first order with constant coefficients. Later sections

extend the application to equations of any order and to sets of simultaneous equations. For problems of this kind the Laplace transform method is so much more economical and effective than other methods that it is almost always used in research work, for instance in electric circuit analysis.

The introduction given here is a brief and utilitarian account of the Laplace transform method: it is adequate for the ordinary differential equations that arise in physical chemistry. A more complete account will be found in Churchill[1] and Carslaw and Jaeger.[2]

## 4.2. The linear equation with constant coefficients

Given

$$a\frac{dx}{dt}+bx = f(t), \quad t > 0, \tag{4.2.1}$$

with $x(0) = x_0$, let us multiply each term by $\exp(-pt)$, where $p$ is positive, and integrate from $t = 0$ to $t = \infty$. We assume that the given function $f(t)$ is such that

$$\int_0^\infty f(t)e^{-pt}dt \tag{4.2.2}$$

exists for all $p$ greater than a certain $p_0$, and that the integrals containing the unknown functions $x(t)$ and $dx(t)/dt$ have the same property.

The expression (4.2.2) is then a function of $p$ and not of $t$. It is called the *Laplace transform* of $f(t)$ and is denoted by

$$\mathscr{L}\{f(t)\},$$

or by $\bar{f}(p)$. Similarly the Laplace transform of $x(t)$ will be denoted by $\bar{x}(p)$.

The usefulness of this transformation of the equation (1) appears when we examine the transform of $dx/dt$. For

$$\begin{aligned}\mathscr{L}\left\{\frac{dx}{dt}\right\} &= \int_0^\infty \frac{dx}{dt}e^{-pt}dt \\ &= \int_0^\infty e^{-pt}dx \\ &= [xe^{-pt}]_0^\infty - \int_0^\infty x(-pe^{-pt})dt \\ &= p\bar{x} - x_0 \end{aligned} \tag{4.2.3}$$

provided $xe^{-pt}$ vanishes at $t = \infty$. Thus the differential equation transforms into the algebraic equation

$$a(p\bar{x} - x_0) + b\bar{x} = \bar{f}(p), \tag{4.2.4}$$

and this we may solve for $\bar{x}$ to get

$$\bar{x} = \frac{\bar{f}(p) + ax_0}{ap + b}. \tag{4.2.5}$$

Thus we have found the Laplace transform of $x(t)$: moreover the initial condition $x(0) = x_0$ has been incorporated into our solution, showing its presence by the term $ax_0$ in the numerator in (5). If we could deduce $x(t)$ from a knowledge of $\bar{x}(p)$ our problem would be solved. It would of course be necessary to check that the assumptions on which our solution has been derived are in fact obeyed, namely

$$x(t)e^{-pt} \to 0 \text{ as } t \to \infty$$

and

$$\left.\begin{array}{l} \displaystyle\int_0^\infty x(t)e^{-pt}dt \\ \\ \displaystyle\int_0^\infty \frac{dx}{dt}e^{-pt}dt \end{array}\right\} \text{exist for } p > \text{some } p_0. \tag{4.2.6}$$

Our situation is like one in which we have "reduced a problem to quadrature". In that case it remains to evaluate an integral, i.e. to construct the function whose derivative is the integrand: this is done either by use of tables or (ultimately) by reference to experience of the results of differentiating. Here we can proceed from $\bar{x}(p)$ to $x(t)$ either by use of tables of transforms or by reference to accumulated experience of transforming elementary functions. We choose the second method and so construct our own table, but the reader will find lists of transforms in the references given on p. 114.

We first observe that it follows directly from the definition that if $k$ is a constant

$$\mathscr{L}\{kx(t)\} = k\mathscr{L}\{x(t)\}, \tag{4.2.7}$$

and that

$$\mathscr{L}\{x_1(t) + x_2(t)\} = \mathscr{L}\{x_1(t)\} + \mathscr{L}\{x_2(t)\}, \tag{4.2.8}$$

so that in general

$$\mathscr{L}\{\sum_r A_r f_r(t)\} = \sum_r A_r \mathscr{L}\{f_r(t)\}.$$

This character of the transformation is described by saying that it is *linear*.

We now proceed to obtain a few simple transforms for our dictionary. If $\alpha$ is real

$$\mathscr{L}\{e^{\alpha t}\} = \int_0^\infty e^{-(p-\alpha)t}dt = \frac{1}{p-\alpha}, \quad \text{provided } p > \alpha. \tag{4.2.9}$$

If $\alpha$ is complex it is easily proved that the same formula holds provided $p > \mathrm{Re}(\alpha)$. Writing $\alpha = \beta + i\gamma$ we have

$$\mathscr{L}\{e^{(\beta+i\gamma)t}\} = \frac{1}{p-\beta-i\gamma},$$

and taking real and imaginary parts we have

$$\mathscr{L}\{e^{\beta t}\cos \gamma t\} = \frac{p-\beta}{(p-\beta)^2+\gamma^2}, \tag{4.2.10}$$

$$\mathscr{L}\{e^{\beta t}\sin \gamma t\} = \frac{\gamma}{(p-\beta)^2+\gamma^2}. \tag{4.2.11}$$

Putting $\beta = 0$, we obtain the important special cases

$$\mathscr{L}\{\cos \gamma t\} = \frac{p}{p^2+\gamma^2}, \tag{4.2.12}$$

$$\mathscr{L}\{\sin \gamma t\} = \frac{\gamma}{p^2+\gamma^2}, \tag{4.2.13}$$

The equation (9) obeys the conditions under which we can differentiate with regard to the parameter $\alpha$. Differentiating $r$ times we get

$$\mathscr{L}\{t^r e^{\alpha t}\} = \frac{r!}{(p-\alpha)^{r+1}}. \tag{4.2.14}$$

Proceeding to the limit as $\alpha \to 0$ we obtain

$$\mathscr{L}\{t^r\} = \frac{r!}{p^{r+1}}. \tag{4.2.15}$$

which could have been obtained directly by integrating by parts. We have already obtained a list of transforms long enough to enable us to solve many elementary problems.

*Example.* It is required to solve the equation

$$\frac{dx}{dt}+2x = \cos t, \tag{4.2.16}$$

with the initial condition $x(0) = 1$.

Since

$$\mathscr{L}\{\cos t\} = \frac{p}{p^2+1},$$

we obtain, by substituting into (5),

$$\begin{aligned}\bar{x}(p) &= \frac{p/(p^2+1)+1}{p+2} \\ &= \frac{p^2+p+1}{(p^2+1)(p+2)}.\end{aligned} \tag{4.2.17}$$

To interpret (17) we express it in partial fractions as

$$\tfrac{1}{5}\left(\frac{2p+1}{p^2+1}+\frac{3}{p+2}\right),$$

and by reference to (9), (12) and (13) we find

$$x(t) = \tfrac{1}{5}(2\cos t+\sin t+3e^{-2t}). \tag{4.2.18}$$

We observe that (18) satisfies the conditions (6) provided we have worked with $p > 2$. A check by substitution shows that it is the correct solution of (16).

## 4.3. Methods for the inversion of a transform

If $\bar{f}(p) = \mathscr{L}\{f(t)\}$, we may extend our notation to write

$$f(t) = \mathscr{L}^{-1}\{\bar{f}(p)\}, \tag{4.3.1}$$

which we read as

"$f(t)$ is the *inverse Laplace transform* of $\bar{f}(p)$".

The process of finding $f(t)$ from $\bar{f}(p)$ is called the *inversion* of the transform.

Our definition of $\mathscr{L}\{f(t)\}$ ensures that the transform of a function is unique whenever it exists. We have assumed without mention that the converse is true, namely that given $\bar{f}(p)$ there is one and only one $f(t)$ of which it is the transform. The proof of this uniqueness is given

in *Lerch's theorem*, where conditions for its validity are stated. Lerch's theorem is discussed in the following texts:

R. V. Churchill, *Modern Operational Mathematics*[1]

H. S. Carslaw and J. C. Jaeger, *Operational Methods*[2]

The reader who has understood the previous sections can proceed to solve problems by the use of the Laplace transform, obtaining the operational solution and then inverting entirely by reference to tables of transforms. But if he wishes to have a deeper understanding of the step-by-step correspondence between the problem and its mathematical analysis he will wish to follow the process of inversion more closely rather than by rule of thumb. He will then find the ensuing discussion useful.

Our power to invert a transform is greatly increased by the following theorem, which enables us to construct the inverse of the product of two transforms whose separate inverses are known.

*The convolution theorem*

If

$$\mathscr{L}^{-1}\{\bar{f}(p)\} = f(t),$$

and

$$\mathscr{L}^{-1}\{\bar{g}(p)\} = g(t),$$

then

$$\mathscr{L}^{-1}\{\bar{f}(p)\bar{g}(p)\} = \int_0^t f(\tau)g(t-\tau)\mathrm{d}\tau. \tag{4.3.2}$$

The proof, in outline, goes as follows: we have

$$\bar{f}(p) = \int_0^\infty \mathrm{e}^{-pt}f(t)\mathrm{d}t,$$

$$\bar{g}(p) = \int_0^\infty \mathrm{e}^{-pu}g(u)\mathrm{d}u,$$

so that

$$\begin{aligned}\bar{f}(p)\bar{g}(p) &= \int_0^\infty \mathrm{e}^{-pt}f(t)\mathrm{d}t \int_0^\infty \mathrm{e}^{-pu}g(u)\mathrm{d}u \\ &= \int_0^\infty \int_0^\infty \mathrm{e}^{-p(\tau+u)}f(\tau)g(u)\mathrm{d}\tau\mathrm{d}u,\end{aligned}$$

where $\tau$ has been written in place of $t$.

We now make the change of variables

$$\tau = \tau$$
$$u = t - \tau.$$

[The reader who is unfamiliar with the technique of changing variables in a multiple integral may refer to Courant: *Differential and Integral Calculus*[3] (Vol. II, Ch. 4).]

The Jacobian of this transformation is

$$\frac{\partial(\tau, u)}{\partial(\tau, t)} = 1,$$

and the quadrant of the $(\tau, u)$ plane (see Fig. 4.1) is described by allowing $\tau$ to vary from 0 to $t$, and $t$ from 0 to $\infty$. The double integral is then

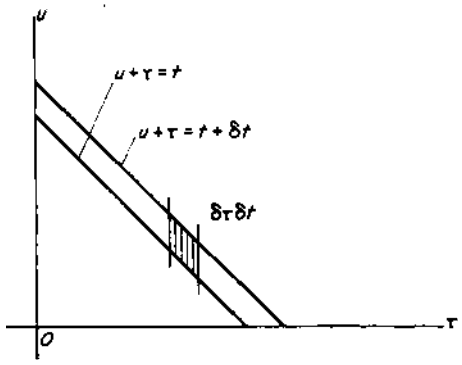


FIG. 4.1. Change of variable in double integral.

regarded as the limit of the sum of strips lying at right angles to the line which bisects the angle between the axes of $\tau$ and $u$. Thus

$$\bar{f}(p)\bar{g}(p) = \int_0^\infty \mathrm{d}t e^{-pt}\{\int_0^t f(\tau)g(t-\tau)\mathrm{d}\tau\}$$
$$= \mathscr{L}\{\int_0^t f(\tau)g(t-\tau)\mathrm{d}\tau\}, \qquad (4.3.3)$$

which is the required result. A simple change of variable shows that the right-hand side of (3) is equal to

$$\mathscr{L}\{\int_0^t f(t-\tau)g(\tau)\mathrm{d}\tau\}.$$

It should be noted that the above formal derivation of the convolution theorem is incomplete, since it has not been shown that the integrands all obey the conditions required for the validity of the change of variable. A proof which includes such considerations may be found in

R. V. Churchill, *Modern Operational Mathematics.*[1]

It is sufficient that $f(t)$ and $g(t)$ be sectionally continuous and that as $t \to \infty$ they do not diverge more severely than $\exp \alpha t$, where $\alpha$ is some constant. These conditions are usually satisfied in the problems with which we deal.

The convolution integral is sometimes indicated by the notation

$$f(t)*g(t) = \int_0^t f(\tau)g(t-\tau)\mathrm{d}\tau. \tag{4.3.4}$$

*Example* (a). Since

$$\frac{1}{p-\alpha} = \mathscr{L}\{\mathrm{e}^{\alpha t}\},$$

$$\frac{1}{p-\beta} = \mathscr{L}\{\mathrm{e}^{\beta t}\},$$

by the convolution theorem

$$\begin{aligned}\frac{1}{(p-\alpha)(p-\beta)} &= \mathscr{L}\Big\{\int_0^t \mathrm{e}^{\alpha\tau}\mathrm{e}^{\beta(t-\tau)}\mathrm{d}\tau\Big\} \\ &= \mathscr{L}\left\{\mathrm{e}^{\beta t}\left(\frac{1}{\alpha-\beta}\,\mathrm{e}^{(\alpha-\beta)\tau}\right)_0^t\right\} \\ &= \mathscr{L}\left\{\frac{1}{\alpha-\beta}\,(\mathrm{e}^{\alpha t}-\mathrm{e}^{\beta t})\right\}. \end{aligned} \tag{4.3.5}$$

This result could also have been obtained by expressing $1/(p-\alpha)(p-\beta)$ in partial fractions.

*Example* (b). Since

$$\frac{1}{p^{n+1}} = \mathscr{L}\left\{\frac{t^n}{n!}\right\},$$

$$\frac{1}{p^2+\gamma^2} = \mathscr{L}\left\{\frac{1}{\gamma}\sin \gamma t\right\}$$

then

$$\frac{1}{p^{n+1}(p^2+\gamma^2)} = \mathscr{L}\left\{\int_0^t \frac{\tau^n}{n!}.\frac{1}{\gamma}\sin\,\gamma(t-\tau)\mathrm{d}\tau\right\}. \tag{4.3.6}$$

The integral on the right may be evaluated by integration by parts.

*Example* (c). *Def.* The *Heaviside step-function* $H(t-t_0)$ is defined by

$$\begin{aligned} H(t-t_0) &= 0, \quad t < t_0, \\ &= 1, \quad t \geqq t_0. \end{aligned} \tag{4.3.7}$$

It is represented graphically by the simple unit step shown in Fig. 4.2.

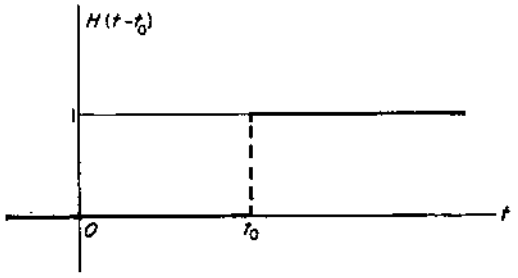


FIG. 4.2. The Heaviside step-function $H(t-t_0)$.

We find

$$\begin{aligned} \mathscr{L}\{H(t-t_0)\} &= \int_0^\infty H(t-t_0)\mathrm{e}^{-pt}\mathrm{d}t \\ &= \int_{t_0}^\infty \mathrm{e}^{-pt}\mathrm{d}t \\ &= \left[-\frac{1}{p}\mathrm{e}^{-pt}\right]_{t_0}^\infty \\ &= \frac{1}{p}\mathrm{e}^{-pt_0}. \end{aligned} \tag{4.3.8}$$

An important special case is obtained by taking $t_0 = 0$: then

$$\mathscr{L}\{H(t)\} = \frac{1}{p}. \tag{4.3.9}$$

Note : If we refer back to (2.9), putting $\alpha = 0$ there, we see that

$$\mathscr{L}\{1\} = \frac{1}{p}.$$

How is this to be reconciled with the formula (9)? We must remember that the Laplace transform utilizes only that part of the transformed function which lies in the range $0 < t < \infty$. Values taken by the function in the range $t < 0$ are irrelevant. Thus the transform of 1 and the transform of $H(t)$ are identical because the values taken by these functions in the range $0 < t < \infty$ are identical. From the standpoint of physics, we are concerned with the behaviour of a system subsequent to a certain origin of time. We are given initial conditions at that origin of time : what happened before $t = 0$ is irrelevant—it does not affect our solution except through the initial conditions.

We can now write

$$\frac{1}{p}\bar{f}(p) = \mathscr{L}\{\int_0^t f(\tau)H(t-\tau)\mathrm{d}\tau\}.$$

But
$$H(t-\tau) = 1 \quad \text{when} \quad \tau \leqq t$$
$$= 0 \quad \text{when} \quad \tau > t,$$

so that

$$\frac{1}{p}\bar{f}(p) = \mathscr{L}\{\int_0^t f(\tau)\mathrm{d}\tau\}, \tag{4.3.10}$$

and we see that multiplication of a transform by $1/p$ corresponds to integration of the original function from $t = 0$ to $t = t$.

*Example* (d). Let us define, by superposition of two step-functions,

$$\Delta(t_0, \varepsilon) = \frac{1}{2\varepsilon}[H\{t-(t_0-\varepsilon)\} - H\{t-(t_0+\varepsilon)\}]. \tag{4.3.11}$$

The graph of this function is shown in Fig. 4.3. The area of the rectangle standing on the axis of $t$ is unity. We can rewrite (11) as

$$\Delta(t_0, \varepsilon) = 0, \qquad t < t_0-\varepsilon,$$
$$= \frac{1}{2\varepsilon}, \qquad t_0-\varepsilon \leqq t < t_0+\varepsilon,$$
$$= 0, \qquad t_0+\varepsilon \leqq t.$$

Then

$$\mathscr{L}\{\Delta(t_0, \varepsilon)\} = \frac{1}{2\varepsilon}\int_{t_0-\varepsilon}^{t_0+\varepsilon} e^{-pt}dt$$

$$= -\frac{1}{2\varepsilon p}\{e^{-p(t_0+\varepsilon)} - e^{-p(t_0-\varepsilon)}\}$$

$$= e^{-pt_0}\frac{\sinh p\varepsilon}{p\varepsilon}. \tag{4.3.12}$$

Now let $\varepsilon \to 0$, so that the rectangle, of width $2\varepsilon$, height $1/2\varepsilon$ and unit area, tends to a limiting configuration of zero width, infinite

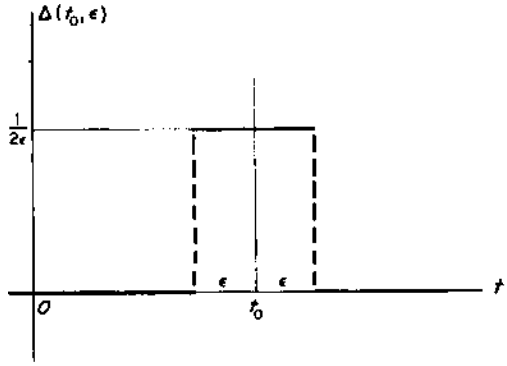


FIG. 4.3. The function $\Delta(t_0, \varepsilon)$.

height and unit area. This limit, the "infinite spike" is the *Dirac delta-function* $\delta(t-t_0)$. We write

$$\delta(t-t_0) = \lim_{\varepsilon\to 0}\Delta(t_0, \varepsilon). \tag{4.3.13}$$

If we assume that the order of the operations

$$\lim_{\varepsilon\to 0} \quad \text{and} \quad \int_0^\infty \ldots dt$$

can be interchanged, we obtain

$$\mathscr{L}\{\delta(t-t_0)\} = \lim_{\varepsilon\to 0} e^{-pt_0}\frac{\sinh p\varepsilon}{p\varepsilon}$$

$$= e^{-pt_0} \tag{4.3.14}$$

(since $\sinh p\varepsilon = p\varepsilon +$ higher powers of $p\varepsilon$). Thus $e^{-pt_0}$ is the Laplace transform of $\delta(t-t_0)$.

There is another approach to this question that is illuminating. Let us examine the expression

$$\int_0^\infty \delta(t-t_0)\ f(t)\mathrm{d}t.$$

This is

$$\int_0^\infty \lim_{\varepsilon\to 0} \Delta(t_0,\ \varepsilon)\, f(t)\mathrm{d}t.$$

Now

$$\begin{aligned}
\int_0^\infty \Delta(t_0,\ \varepsilon)\, f(t)\mathrm{d}t &= \frac{1}{2\varepsilon}\int_{t_0-\varepsilon}^{t_0+\varepsilon} f(t)\mathrm{d}t \\
&= \frac{1}{2\varepsilon}\int_{-\varepsilon}^{\varepsilon} f(t_0+\xi)\mathrm{d}\xi \\
&= \frac{1}{2\varepsilon}\int_{-\varepsilon}^{\varepsilon} \{f(t_0)+\xi f'(t_0)+\tfrac{1}{2}\xi^2 f''(t_0)+\ \ldots\}\mathrm{d}\xi \\
&= \frac{1}{2\varepsilon}\{f(t_0).2\varepsilon + f''(t_0).\tfrac{1}{3}.2\varepsilon^3+\ \ldots\} \\
&\to f(t_0) \quad \text{as} \quad \varepsilon\to 0,
\end{aligned}$$

so that *formally*† (i.e. presuming that the various manipulations of the working can be justified)

$$\int_0^\infty \delta(t-t_0)\, f(t)\mathrm{d}t = f(t_0). \tag{4.3.15}$$

Thus the delta-function, occurring as part of an integrand, picks out the value of the remaining factor of the integrand at the point where the

† The properties of the delta-function and the step-function are treated with rigour in the theory of distributions. See Lighthill, M. J.: *Fourier Analysis and Generalized Functions*.[4]

Note that in (15) it is quite unnecessary that the lower and upper end-points of integration should be 0 and $\infty$. All that is necessary is that $t_0$ should lie between them. Thus

$$\begin{aligned}
\int_a^b \delta(t-t_0)f(t)\mathrm{d}t &= 0 \quad \text{if } t_0 \text{ lies outside } (a,\ b) \\
&= f(t_0) \quad \text{if } t_0 \text{ lies in } (a,\ b).
\end{aligned}$$

delta-function is discontinuous. Applying (15) to the Laplace integral we get

$$\mathscr{L}\{\delta(t-t_0)\} = \int_0^\infty \delta(t-t_0)e^{-pt}dt$$

$$= e^{-pt_0},$$

as in (14).

Using (14) with the convolution theorem we have

$$e^{-p\alpha}\bar{f}(p) = \mathscr{L}\{\int_0^t \delta(\tau-\alpha)\,f(t-\tau)d\tau\}.$$

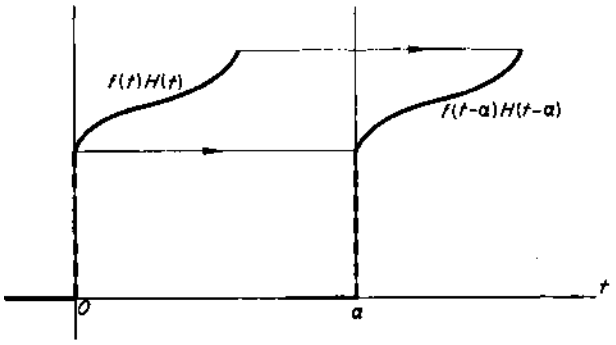


FIG. 4.4. The shift operation.

Now if $\alpha$ lies within $(0, t)$ the integral becomes $f(t-\alpha)$, while if $\alpha$ lies outside $(0, t)$ the integral is zero. Thus the integral can be evaluated as $f(t-\alpha)H(t-\alpha)$. Our result can be exhibited as

$$\bar{f}(p) = \mathscr{L}\{f(t)H(t)\},$$

$$e^{-p\alpha}\bar{f}(p) = \mathscr{L}\{f(t-\alpha)H(t-\alpha)\}. \qquad (4.3.16)$$

Now the graph of $f(t-\alpha)H(t-\alpha)$ is obtained from that of $f(t)H(t)$ by a shift through a distance $\alpha$ in the direction of $t$ increasing, as shown in Fig. 4.4; $\exp(-p\alpha)$ is therefore called the *shift operator*. If a transform is multiplied by $\exp(-p\alpha)$ the corresponding function is shifted through a time $\alpha$, the ordinate being zero on the left of $t = \alpha$.

## 4.4. The second order equation with constant coefficients

The Laplace transform method extends immediately to differential equations of higher order. Reverting to (3.2.6), which is

$$\ddot{x}+2b\dot{x}+cx = f(t), \quad x(0) = x_0, \ \dot{x}(0) = u_0, \tag{4.4.1}$$

and applying the Laplace transformation, we note that

$$\begin{aligned}\mathscr{L}\{\ddot{x}(t)\} &= \int_0^\infty \mathrm{e}^{-pt}\mathrm{d}\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)\\ &= \left[\frac{\mathrm{d}x}{\mathrm{d}t}\,\mathrm{e}^{-pt}\right]_0^\infty + p\int_0^\infty \frac{\mathrm{d}x}{\mathrm{d}t}\,\mathrm{e}^{-pt}\mathrm{d}t\\ &= p(p\bar{x}-x_0)-\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)_{t=0}\\ &= p^2\bar{x}-px_0-u_0\end{aligned} \tag{4.4.2}$$

provided we assume that $x$, $\mathrm{d}x/\mathrm{d}t$ and $\mathrm{d}^2x/\mathrm{d}t^2$ obey suitable restrictions concerning integrability and convergence. The differential equation transforms into

$$p^2\bar{x}-px_0-u_0+2b(p\bar{x}-x_0)+c\bar{x} = \bar{f}(p),$$

i.e.

$$\begin{aligned}\bar{x} &= \frac{\bar{f}+(p+2b)x_0+u_0}{p^2+2bp+c}\\ &= \frac{\bar{f}+(p+2b)x_0+u_0}{(p-\alpha_1)(p-\alpha_2)},\end{aligned} \tag{4.4.3}$$

where $\alpha_1$, $\alpha_2$ have the same meaning as before (Section 3.2). Then if $f(t)$ is known $\bar{f}(p)$ can be found, and the solution $x(t)$ can be discovered by looking up in a table of transforms the expressions that make up the solution $\bar{x}(p)$ in (3).

We observe that

$$\begin{aligned}\frac{(p+2b)x_0+u_0}{(p-\alpha_1)(p-\alpha_2)} &= \frac{(p-\alpha_1-\alpha_2)x_0+u_0}{(p-\alpha_1)(p-\alpha_2)}\\ &= \mathscr{L}\left[\frac{1}{\alpha_1-\alpha_2}\{u_0(\mathrm{e}^{\alpha_1 t}-\mathrm{e}^{\alpha_2 t})-x_0(\alpha_2\mathrm{e}^{\alpha_1 t}-\alpha_1\mathrm{e}^{\alpha_2 t})\}\right].\end{aligned} \tag{4.4.4}$$

As for the term containing $\bar{f}$, we see that

$$\frac{\bar{f}}{(p-\alpha_1)(p-\alpha_2)} = \frac{1}{\alpha_1-\alpha_2}\left(\frac{\bar{f}}{p-\alpha_1} - \frac{\bar{f}}{p-\alpha_2}\right)$$
$$= \mathscr{L}\left[\frac{1}{\alpha_1-\alpha_2}\left\{\int_0^t f(\tau)e^{\alpha_1(t-\tau)}d\tau - \int_0^t f(\tau)e^{\alpha_2(t-\tau)}d\tau\right\}\right] \tag{4.4.5}$$

by the convolution theorem. Thus we have the same result as in (3.2.11).

## 4.5. The linear equation of order $n$ with constant coefficients

Given the differential equation (3.3.2)

$$L(D)x \equiv D^n x + a_1 D^{n-1}x + a_2 D^{n-2}x + \ldots + a_r D^{n-r}x + \ldots + a_n x = f(t), \quad t > 0, \tag{4.5.1}$$

with

$$x(0) = x_0, \quad \dot{x}(0) = x_1, \quad \ldots, \quad (D^{n-1}x)_{t=0} = x_{n-1},$$

we wish to solve for $x(t)$. Since

$$\mathscr{L}\{D^r(x)\} = \int_0^\infty e^{-pt}dD^{r-1}(x)$$
$$= [e^{-pt}D^{r-1}(x)]_0^\infty + p\int_0^\infty e^{-pt}D^{r-1}(x)dt$$
$$= p\mathscr{L}\{D^{r-1}(x)\} - x_{r-1}$$

(provided $e^{-pt}D^{r-1}(x) \to 0$ as $t \to \infty$)

$$= p[p\mathscr{L}\{D^{r-2}(x)\} - x_{r-2}] - x_{r-1} \tag{4.5.2}$$

and so on, we transform (1) to get

$$\begin{aligned} &p^n\bar{x} - p^{n-1}x_0 - p^{n-2}x_1 \ldots - px_{n-2} - x_{n-1} \\ &+a_1(p^{n-1}\bar{x} - p^{n-2}x_0 - p^{n-3}x_1 \ldots - x_{n-2}) \\ &+a_2(p^{n-2}\bar{x} - p^{n-3}x_0 \ldots) \\ &+ \ldots \\ &+a_n\bar{x} \qquad\qquad = \bar{f}(p), \end{aligned}$$

i.e.

$$\begin{aligned}(p^n+a_1p^{n-1}+\ldots+a_rp^{n-r}+\ldots+a_n)\bar{x}(p)&\\ =\bar{f}(p)+(p^{n-1}+a_1p^{n-2}+\ldots+a_{n-1})x_0&\\ +(p^{n-2}+a_1p^{n-3}+\ldots+a_{n-2})x_1&\\ +\ldots&\\ +(p+a_1)x_{n-2}&\\ +x_{n-1}.&\end{aligned}\tag{4.5.3}$$

Thus

$$\bar{x}(p)=\frac{\bar{f}(p)+R(p)}{L(p)},$$

where $R(p)$ is the polynomial of degree $n-1$ which appears on the right-hand side of (3), and $L(p)$ is the polynomial of degree $n$ obtained by substituting $p$ for $D$ in the operator $L(D)$ in (1).

*Heaviside's formula*

We seek a formula for the inversion of the transform

$$R(p)/L(p),\tag{4.5.4}$$

where $L(p)$ is a polynomial in $p$ of degree $n$ and $R(p)$ is a polynomial of degree less than $n$. We first write

$$L(p)=\prod_{r=1}^{n}(p-\alpha_r),\tag{4.5.5}$$

where the $n$ roots $\alpha_r$ of $L(p)=0$ are assumed to be all different, but may be complex. We use partial fractions, and determine constants $A_r$ such that

$$\frac{R(p)}{L(p)}=\sum_{1}^{n}\frac{A_r}{p-\alpha_r}.\tag{4.5.6}$$

Then

$$R(p)=\sum_{1}^{n}A_r\Pi_r^*(p-\alpha),\tag{4.5.7}$$

where $\Pi_r^*(p-\alpha)$ is the product of all the factors of $L(p)$ except $(p-\alpha_r)$. Putting $p=\alpha_r$ in (7) we see that all terms on the right vanish, except that containing $A_r$, and

$$R(\alpha_r)=A_r\Pi_r^*(\alpha_r-\alpha).$$

If we differentiate $L(p)$ we obtain

$$L'(p) = \Pi_1^*(p-\alpha)+\Pi_2^*(p-\alpha)+\ldots+\Pi_n^*(p-\alpha),$$

so that

$$L'(\alpha_r) = \Pi_r^*(\alpha_r-\alpha),$$

all the other terms on the right vanishing owing to the presence of the factor $(p-\alpha_r)$. Hence

$$R(\alpha_r) = A_r L'(\alpha_r)$$

and

$$\frac{R(p)}{L(p)} = \sum_1^n \frac{R(\alpha_r)}{L'(\alpha_r)} \frac{1}{p-\alpha_r}, \tag{4.5.8}$$

so that

$$\mathscr{L}^{-1}\left\{\frac{R(p)}{L(p)}\right\} = \sum_1^n \frac{R(\alpha_r)}{L'(\alpha_r)} e^{\alpha_r t}. \tag{4.5.9}$$

This is Heaviside's formula, available when the $\alpha_r$ are all different. If the $\alpha_r$ are all *real*, (9) is the final form. If there are complex roots $\alpha_r$, they must occur in conjugate pairs. Let $\beta \pm i\gamma$ be a typical pair. This pair gives rise, in (9), to the pair of terms

$$\frac{R(\beta+i\gamma)}{L'(\beta+i\gamma)} e^{(\beta+i\gamma)t} + \frac{R(\beta-i\gamma)}{L'(\beta-i\gamma)} e^{(\beta-i\gamma)t}$$

$$= e^{\beta t}.2\,\mathrm{Re}\left\{\frac{R(\beta+i\gamma)}{L'(\beta+i\gamma)} e^{i\gamma t}\right\} \tag{4.5.10}$$

$$= e^{\beta t} B\cos(\gamma t+\delta), \tag{4.5.11}$$

where $B$ and $\delta$ are constants depending on $\beta$ and $\gamma$, found by evaluating (10).

If we wish to invert $R(p)S(p)/L(p)$, where $R(p)$ and $L(p)$ have the same meanings as above, and $S(p)$ is some other function of $p$, we first seek the function of which $S(p)$ is the transform. Then we cast $R(p)/L(p)$ into the form (8) and invert by means of convolution.

*Example.* We solve the equation

$$(D^4-m^4)x = f(t),\ t > 0, \tag{4.5.12}$$

with $x = x_0$, $Dx = 0$, $D^2x = 0$, $D^3(x) = 0$ at $t = 0$.

Transforming (12) according to the rule (3) we get

$$p^4\bar{x} - p^3x_0 - m^4\bar{x} = \bar{f}(p),$$

so that

$$\bar{x} = \frac{p^3}{p^4 - m^4}\,x_0 + \frac{1}{p^4 - m^4}\,\bar{f}(p). \tag{4.5.13}$$

In the term containing $x_0$ in (13) we have, with the notation of (4)

$$\begin{aligned} R(p) &= p^3x_0, \\ L(p) &= p^4 - m^4 = (p-m)(p+m)(p-im)(p+im), \\ L'(p) &= 4p^3. \end{aligned}$$

Thus $R(p)/L'(p) = \frac{1}{4}x_0$, and the inverse of the first term of (13) is

$$\begin{aligned} &\tfrac{1}{4}(e^{mt} + e^{-mt} + e^{imt} + e^{-imt})x_0 \\ &\qquad = \tfrac{1}{2}(\cosh mt + \cos mt)x_0. \end{aligned} \tag{4.5.14}$$

The second term of (13) must be inverted by means of a convolution.

$$\begin{aligned} \frac{1}{p^4 - m^4} &= \frac{1}{4m^3}\,\frac{1}{p-m} + \frac{1}{4(-m)^3}\,\frac{1}{p+m} \\ &\qquad + \frac{1}{4(im)^3}\,\frac{1}{p-im} + \frac{1}{4(-im)^3}\,\frac{1}{p+im} \\ &= \frac{1}{4m^3}\left(\frac{1}{p-m} - \frac{1}{p+m} + \frac{i}{p-im} - \frac{i}{p+im}\right) \end{aligned}$$

so that

$$\mathscr{L}^{-1}\left\{\frac{1}{p^4 - m^4}\right\} = \frac{1}{2m^3}\,(\sinh mt - \sin mt)$$

and

$$\mathscr{L}^{-1}\left\{\frac{\bar{f}(p)}{p^4 - m^4}\right\} = \frac{1}{2m^3}\int_0^t f(\tau)\{\sinh m(t-\tau) - \sin m(t-\tau)\}\,d\tau. \tag{4.5.15}$$

The solution is the sum of (14) and (15). (14) represents the response of the system to the particular conditions of starting, while (15) gives the response to the disturbing force.

*Repeated roots of $L(p)$*

Let $L(p)$ contain the factor $(p-\alpha)^2$. We write

$$\frac{R(p)}{L(p)} = \frac{1}{(p-\alpha)^2}Q(p)$$

$$= \frac{A_1}{p-\alpha}+\frac{A_2}{(p-\alpha)^2}+Q^*(p), \tag{4.5.16}$$

where $A_1$ and $A_2$ are to be found, and $Q^*(p)$ contains no factor $(p-\alpha)$ in the denominator.

Multiplying by $(p-\alpha)^2$ we have

$$Q(p) = A_1(p-\alpha)+A_2+Q^*(p)(p-\alpha)^2. \tag{4.5.17}$$

Let $p \to \alpha$, then

$$A_2 = Q(\alpha).$$

Differentiate (17) with regard to $p$ and again let $p \to \alpha$, and we obtain

$$A_1 = Q'(\alpha).$$

Then the terms in

$$\mathscr{L}^{-1}\left\{\frac{R(p)}{L(p)}\right\}$$

that correspond to the factor $(p-\alpha)^2$ in $L(p)$ are

$$Q'(\alpha)e^{\alpha t}+Q(\alpha)te^{\alpha t}. \tag{4.5.18}$$

The extension of this formula to the case where $L(p)$ contains a factor $(p-\alpha)^m$ is given in Exercise 4.7.

## 4.6. Sets of simultaneous equations

(a) *The set of simultaneous first order linear equations with constant coefficients*

Many of the sets of equations which describe chains of chemical reactions are special cases of the general type (3.5.2), namely

$$(a_{rs}D+b_{rs})x_s = f_r(t), \quad t > 0, \tag{4.6.1}$$

where $D$ represents $\mathrm{d}/\mathrm{d}t$ and $r, s$ take values $1, 2, \ldots, n$. These are to be solved with the initial conditions

$$x_s = x_{s0} \quad \text{when} \quad t = 0. \tag{4.6.2}$$

(We use the summation convention as in Sections 3.4 and 3.5.)

Transforming with regard to $t$ we obtain

$$a_{rs}(p\bar{x}_s - x_{s0}) + b_{rs}\bar{x}_s = \bar{f}_r(p),$$

i.e.

$$(a_{rs}p + b_{rs})\bar{x}_s = \bar{f}_r(p) + a_{rs}x_{s0}. \tag{4.6.3}$$

We denote $a_{rs}p + b_{rs}$ by $c_{rs}$, $\det c_{rs}$ by $C$, and the cofactor of $c_{rs}$ in $C$ by $C_{rs}$. Then the solution of the set of equations (3), rewritten

$$c_{rs}\bar{x}_s = \bar{f}_r(p) + a_{rs}x_{s0},$$

is obtained by multiplying the $r$th equation by $C_{rq}$ and summing over $r$. Using the fact that

$$c_{rj}C_{rk} = \delta_{jk}C$$

($\delta_{jk} = 0$ if $j \neq k$, $\delta_{jj} = 1$), we see that all terms on the left vanish except $C\bar{x}_q$, so that

$$C\bar{x}_q = C_{rq}\{\bar{f}_r + a_{rs}x_{s0}\}.$$

This solution

$$\bar{x}_q = \frac{C_{rq}}{C}\{\bar{f}_r + a_{rs}x_{s0}\} \tag{4.6.4}$$

is the sum of $n^2$ terms, since both $r$ and $s$ are to be summed over all values. For a more detailed account of this procedure for solution of a set of simultaneous linear equations, see

Margenau, H. and Murphy, G. M., *Mathematics of Physics and Chemistry*[5] (Section 10.9).

$C$ is a polynomial in $p$ of degree $n$, and $C_{rq}$ a polynomial in $p$ of degree $n-1$. Thus the right-hand side is a function of $p$ and may be inverted (e.g. by Heaviside's formula) to give the solution $x_q(t)$.

Comparing this treatment of the equations (1) with that of Section 3.5, we see that the incorporation of the initial conditions into our solution at the early stage (equation (3)) makes for very much simpler algebra. The conciseness and economy of our notation is deceptive, and particular problems, such as the following Example, may sometimes be solved more easily by *ad hoc* methods.

*Example.* In Section 3.4 we obtained, for consecutive chemical reactions of the first order, the equations

$$\left.\begin{aligned} \dot{x}_1 &= -k_1x_1 \\ \dot{x}_2 &= k_1x_1 - k_2x_2 \\ \dot{x}_3 &= k_2x_2 \end{aligned}\right\} \quad t > 0, \tag{4.6.5}$$

with $x_1(0) = x_{10}$, $x_2(0) = 0$, $x_3(0) = 0$.

Transforming, we have

$$\begin{aligned} p\bar{x}_1 - x_{10} &= -k_1\bar{x}_1, \\ p\bar{x}_2 &= k_1\bar{x}_1 - k_2\bar{x}_2, \\ p\bar{x}_3 &= k_2\bar{x}_2, \end{aligned}$$

so that

$$\bar{x}_1 = \frac{x_{10}}{p+k_1}, \quad \bar{x}_2 = \frac{k_1\bar{x}_1}{p+k_2} = \frac{k_1 x_{10}}{(p+k_1)(p+k_2)}.$$

If we want to find only $x_2$ we invert $\bar{x}_2$ and obtain

$$\bar{x}_2 = \frac{k_1 x_{10}}{k_2 - k_1}\left(\frac{1}{p+k_1} - \frac{1}{p+k_2}\right),$$

$$x_2 = \frac{k_1}{k_2 - k_1} x_{10}(e^{-k_1 t} - e^{-k_2 t}),$$

as in (3.4.4). One of the advantages of this method is that we evaluate only that part of the solution which is desired.

In Section 3.4 we considered the general problem of solving the equations of radioactive breakdown. There the boundary conditions were left unspecified. If they are given, as is common, in terms of the initial concentrations of the various products of the series, the Laplace transform method becomes by far the most economical.

Thus the problem is specified by the differential equations (3.4.6) and boundary conditions:

$$\left.\begin{aligned} \dot{x}_1 &= -k_1 x_1 & x_1(0) &= x_{10} \\ \dot{x}_2 &= k_1 x_1 - k_2 x_2 & x_2(0) &= x_{20} \\ &\cdots & &\cdots \\ \dot{x}_r &= k_{r-1}x_{r-1} - k_r x_r & x_r(0) &= x_{r0} \\ &\cdots & &\cdots \\ \dot{x}_{n-1} &= k_{n-2}x_{n-2} - k_{n-1}x_{n-1} & x_{n-1}(0) &= x_{(n-1)0} \\ \dot{x}_n &= k_{n-1}x_{n-1} & x_n(0) &= x_{n0} \end{aligned}\right\} \tag{4.6.6}$$

Transforming these equations, we obtain

$$\left.\begin{aligned} (p+k_1)\bar{x}_1 &= x_{10}, \\ -k_1\bar{x}_1 + (p+k_2)\bar{x}_2 &= x_{20}, \\ \cdots & \\ -k_{r-1}\bar{x}_{r-1} + (p+k_r)\bar{x}_r &= x_{r0}, \\ \cdots & \\ -k_{n-2}\bar{x}_{n-2} + (p+k_{n-1})\bar{x}_{n-1} &= x_{(n-1)0}, \\ -k_{n-1}\bar{x}_{n-1} + p\bar{x}_n &= x_{n0}. \end{aligned}\right\} \tag{4.6.7}$$

These equations could be solved by the use of determinants as in the general theory. A more elementary method, however, is quicker, and we observe that we may write

$$\bar{x}_1 = \frac{x_{10}}{p+k_1},$$

$$\bar{x}_2 = \frac{x_{20}}{p+k_2} + \frac{k_1}{p+k_2}\bar{x}_1$$

$$= \frac{x_{20}}{p+k_2} + \frac{k_1}{(p+k_2)(p+k_1)}x_{10},$$

$$\bar{x}_3 = \frac{x_{30}}{p+k_3} + \frac{k_2}{(p+k_3)(p+k_2)}x_{20} + \frac{k_2k_1}{(p+k_3)(p+k_2)(p+k_1)}x_{10},$$

and so on.

In the simplest case, where all $x_{r0}$ are zero except $x_{10}$ (a start from a pure substance), we have, assuming $k_s$ all different and using Heaviside's method,

$$\bar{x}_r = \frac{k_{r-1}k_{r-2}\ldots k_1}{(p+k_r)(p+k_{r-1})\ldots(p+k_1)}x_{10}$$

$$= \sum_{s=1}^{r} \frac{k_{r-1}k_{r-2}\ldots k_1 x_{10}}{(k_r-k_s)\ldots(k_{s+1}-k_s)(k_{s-1}-k_s)\ldots(k_1-k_s)}\frac{1}{p+k_s},$$

and

$$x_r(t) = \sum_{s=1}^{r} \frac{k_{r-1}k_{r-2}\ldots k_1 x_{10}}{(k_r-k_s)\ldots(k_{s+1}-k_s)(k_{s-1}-k_s)\ldots(k_1-k_s)}e^{-k_s t}. \tag{4.6.8}$$

This formula is valid for $r = 1$ to $r = n-1$, but not for $r = n$ owing to the deviation of the final equation of (7) from the pattern of the others for which $r > 1$. The expression for $x_n(p)$ is easily found and can be cast into partial fractions for inversion; but the simplest method of obtaining $x_n(t)$ is from the fact that

$$x_1 + x_2 + \ldots + x_n = x_{10}.$$

Thus if we have found $x_r(t)$ for $r \leqq n-1$ we subtract their sum from $x_{10}$ to obtain $x_n(t)$.

*(b) Sets of simultaneous equations of higher order*

Given a set of linear equations, of order higher than the first, with constant coefficients, we can proceed in the same manner : transform the equations, solve for the transforms of the dependent variables, then invert. We illustrate the method by a worked problem.

*Example.* Returning to the problem of Sections 1.2 and 3.4, where the equations of orbital motion under restoring and gyroscopic forces were

$$\begin{aligned} \ddot{x} - 2\gamma\dot{y} + \lambda x &= 0, \\ \ddot{y} + 2\gamma\dot{x} + \lambda y &= 0, \quad t > 0, \end{aligned} \tag{4.6.9}$$

we find the motion following the initial conditions

$$\begin{aligned} x(0) &= a, \quad \dot{x}(0) = 0, \\ y(0) &= 0, \quad \dot{y}(0) = V. \end{aligned}$$

The transformed equations are

$$\begin{aligned} (p^2 + \lambda)\bar{x} - 2\gamma p\bar{y} &= pa, \\ 2\gamma p\bar{x} + (p^2 + \lambda)\bar{y} &= 2\gamma a + V, \end{aligned} \tag{4.6.10}$$

with solution

$$\bar{x} = \frac{1}{f(p)}\{ap^3 + (a\lambda + 4a\gamma^2 + 2\gamma V)p\}, \tag{4.6.11}$$

$$\bar{y} = \frac{1}{f(p)}\{Vp^2 + (2\gamma a\lambda + V\lambda)\},$$

where

$$\begin{aligned} f(p) &= (p^2 + \lambda)^2 + 4\gamma^2 p^2 \\ &= (p^2 + \alpha^2)(p^2 + \beta^2), \end{aligned} \tag{4.6.12}$$

in which we have introduced $\mu = \sqrt{(\lambda + \gamma^2)}$ and written

$$\begin{aligned} \alpha &= \mu + \gamma, \\ \beta &= \mu - \gamma. \end{aligned} \tag{4.6.13}$$

To invert the operational solution (11) we need the partial fractions

$$\frac{1}{f(p)} = -\frac{1}{\alpha^2 - \beta^2}\left(\frac{1}{p^2 + \alpha^2} - \frac{1}{p^2 + \beta^2}\right),$$

$$\frac{p^2}{f(p)} = \frac{1}{\alpha^2 - \beta^2}\left(\frac{\alpha^2}{p^2 + \alpha^2} - \frac{\beta^2}{p^2 + \beta^2}\right),$$

which lead immediately to the formulae

$$\mathscr{L}^{-1}\left\{\frac{1}{f(p)}\right\} = -\frac{1}{4\gamma\mu}\left(\frac{1}{\alpha}\sin\alpha t - \frac{1}{\beta}\sin\beta t\right),$$

$$\mathscr{L}^{-1}\left\{\frac{p}{f(p)}\right\} = -\frac{1}{4\gamma\mu}(\cos\alpha t - \cos\beta t),$$

$$\mathscr{L}^{-1}\left\{\frac{p^2}{f(p)}\right\} = \frac{1}{4\gamma\mu}(\alpha\sin\alpha t - \beta\sin\beta t),$$

$$\mathscr{L}^{-1}\left\{\frac{p^3}{f(p)}\right\} = \frac{1}{4\gamma\mu}(\alpha^2\cos\alpha t - \beta^2\cos\beta t).$$

Inserting these formulae into (11) and simplifying we obtain the solution, giving parametric equations of the orbit:

$$\begin{aligned} 2\mu x &= a(\beta\cos\alpha t + \alpha\cos\beta t) + V(-\cos\alpha t + \cos\beta t), \\ 2\mu y &= a(-\beta\sin\alpha t + \alpha\sin\beta t) + V(\sin\alpha t + \sin\beta t). \end{aligned} \tag{4.6.14}$$

This is best interpreted by use of the complex variable $z = x + iy$

$$\begin{aligned} 2\mu z &= a(\beta e^{-i\alpha t} + \alpha e^{i\beta t}) + V(-e^{-i\alpha t} + e^{i\beta t}) \\ &= e^{-i\gamma t}\{(a\alpha + V)e^{i\mu t} + (a\beta - V)e^{-i\mu t}\} \\ &= 2\mu\zeta e^{-i\gamma t} \end{aligned} \tag{4.6.15}$$

where

$$\begin{aligned} 2\mu\zeta &= a(\alpha+\beta)\cos\mu t + i\{a(\alpha-\beta)+2V\}\sin\mu t \\ &= 2\mu a\cos\mu t + i(2\gamma a + 2V)\sin\mu t. \end{aligned} \tag{4.6.16}$$

Thus if $\zeta = \xi + i\eta$, $\xi = a\cos\mu t$, $\eta = \{(\gamma a + V)/\mu\}\sin\mu t$, and $(\xi, \eta)$ moves on an ellipse of semi-axes $a$, $(\gamma a + V)/\mu$. Meanwhile, since $z$ is derived from $\zeta$ by multiplication by $\exp(-i\gamma t)$, $z$ moves on that ellipse as it rotates with angular velocity $\gamma$. Our result thus tallies with that of Section 3.4, (16) being identical with (3.4.38).

## 4.7. Transforms of periodic functions

*Def.* The function $f(t)$ is *periodic* with period $T$ if, for all $t$ and all integers $k$,

$$f(t + kT) = f(t).$$

For such a function

$$\mathscr{L}\{f(t)\} = \int_0^T f(t)e^{-pt}dt + \int_T^{2T} f(t)e^{-pt}dt + \ldots + \int_{kT}^{(k+1)T} f(t)e^{-pt}dt + \ldots.$$

In the general term, we write $t = kT + \tau$, and using $f(kT+\tau) = f(\tau)$ we obtain

$$e^{-pkT} \int_0^T f(\tau) e^{-p\tau}\, d\tau,$$

so that

$$\mathscr{L}\{f(t)\} = \sum_{k=0}^{\infty} e^{-pkT} \int_0^T f(t) e^{-pt} dt$$

$$= \frac{1}{1-e^{-pT}} \int_0^T f(t) e^{-pt} dt.$$

*Example.* If

$$f(t) = 1 \quad \text{when} \quad 0 < t \leqq \tfrac{1}{2}T$$

$$= -1 \quad \text{when} \quad \tfrac{1}{2}T < t \leqq T$$

and is periodic with period $T$, then

$$\int_0^T f(t) e^{-pt} dt = \int_0^{\frac{1}{2}T} e^{-pt} dt - \int_{\frac{1}{2}T}^T e^{-pt} dt$$

$$= \int_0^{\frac{1}{2}T} e^{-pt}(1 - e^{-\frac{1}{2}pT}) dt$$

$$= \frac{1}{p}(1 - e^{-\frac{1}{2}pT})^2,$$

so that

$$\mathscr{L}\{f(t)\} = \frac{1}{1-e^{-pT}} \frac{1}{p}(1 - e^{-\frac{1}{2}pT})^2$$

$$= \frac{1}{p} \frac{1 - e^{-\frac{1}{2}pT}}{1 + e^{-\frac{1}{2}pT}}$$

$$= \frac{1}{p} \tanh \tfrac{1}{4}pT.$$

This "square wave" or "meander function" is shown in Fig. 4.5. Transforms of other periodic functions are given in the Exercises at the end of this chapter. Such functions are useful in problems where a periodic disturbing force is acting, for instance problems concerning electric circuits.

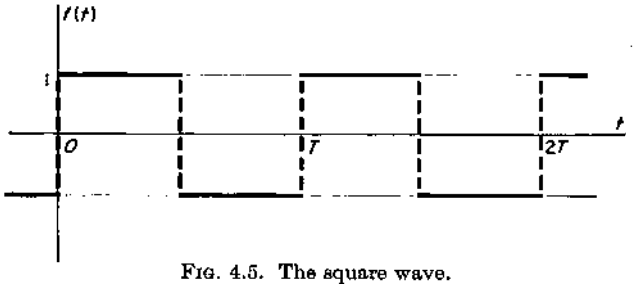


FIG. 4.5. The square wave.

## 4.8. More advanced methods

The reader may ask whether there is not some more general and powerful method of inverting a transform than those so far exhibited. Can we not solve the equation

$$\bar{f}(p) = \int_0^\infty e^{-pt} f(t) dt \tag{4.8.1}$$

to obtain an explicit expression for $f(t)$ in terms of $\bar{f}(p)$, as for instance by an integral containing $\bar{f}(p)$?

The answer is that such an expression does exist, but since it involves integration along a contour in the complex plane we merely state the inversion theorem here, referring the reader to other texts for proof and applications.

*Inversion theorem*

If $\bar{f}(p)$ is a function of $p$ (regarded as complex variable) that is analytic in some half-plane $\mathrm{Re}(p) \geqq \gamma$, and for which $| p\bar{f}(p) |$ is bounded as $| p | \to \infty$ in the same half-plane, then

$$f(t) = \frac{1}{2\pi i} \int_L e^{pt} \bar{f}(p) dp, \tag{4.8.2}$$

where $L$ is the *Bromwich contour*, which runs parallel to the imaginary axis from $c - i\infty$ to $c + i\infty$, where $c \geqq \gamma$.

Various methods are available for the evaluation of the inversion integral which occurs in (2). It may be converted into a real integral by change of variable with distortion of path, it may be evaluated by residue theory, or it may give an asymptotic series for $f(t)$ by the method of steepest descents. These methods are dealt with in

Churchill, R. V., *Modern Operational Methods*.[1]

Jeffreys, H. and Jeffreys, B. S., *Methods of Mathematical Physics*.[6]

In applications of the Laplace transform method, it is desirable to make use of a Table of Transforms. The most extensive table that is generally available is

Erdelyi, Magnus, Oberhettinger and Tricomi, *Tables of Integral Transforms*,[7] Vol. 1.

A short table, easily accessible to chemists and physicists, will be found in

*Handbook of Chemistry and Physics*.[8]

## REFERENCES


1. Churchill, R. V., *Modern Operational Mathematics*, 2nd edn, McGraw-Hill, 1958.
2. Carslaw, H. S. and Jaeger, J. C., *Operational Methods*, Oxford, 1948.
3. Courant, R., *Differential and Integral Calculus*, Blackie, 1937.
4. Lighthill, M. J., *Fourier Analysis and Generalized Functions*, Cambridge, 1958.
5. Margenau, H. and Murphy, G. M., *Mathematics of Physics and Chemistry*, Van Nostrand, 1956.
6. Jeffreys, H. and Jeffreys, B. S., *Methods of Mathematical Physics*, 3rd edn, Cambridge, 1956.
7. Erdelyi, A., Magnus, W., Oberhettinger, F. and Tricomi, F. G., *Tables of Integral Transforms*, McGraw-Hill, 1954.
8. *Handbook of Chemistry and Physics*, 44th edn, Chemical Publishing Co.


## Exercises

[All of the following exercises should be attacked by the Laplace transform method.]

**4.1.** (a) Given $dx/dt = -kx$, $t > 0$, with $x(0) = x_0$, solve for $x(t)$. (Compare Section 3.1.)

(b) Given $dx/dt + kx = F(t)$, $t > 0$, with $x(0) = x_0$, solve for $x(t)$. Find $x(t)$ if $F(t) = my_0e^{-mt}$. (Compare (3.1.9).)

**4.2.** Given $\ddot{x} + n^2x = f(t)$, $t > 0$, $x(0) = x_0$, $\dot{x}(0) = u_0$, solve for $x(t)$. (Compare (3.2.31).)

**4.3.** Given $\{L(D)\}x = f(t)$, $t > 0$, as in (4.5.1), discuss (a) the case where $L(p)$ has equal roots, (b) the case where $f(t) = \exp \alpha t$, $\alpha$ being a root of $L(p) = 0$.

**4.4.** Given

$$\left.\begin{aligned} L\ddot{q}_1 + M\ddot{q}_2 + q_1/C &= 0, \\ M\ddot{q}_1 + L\ddot{q}_2 + q_2/C &= 0, \end{aligned}\right\} \quad t > 0,$$

with

$$q_1(0) = Q, \quad \dot{q}_1(0) = 0,$$
$$q_2(0) = 0, \quad \dot{q}_2(0) = 0,$$

prove

$$q_1(t) = \tfrac{1}{2}Q(\cos mt + \cos nt),$$

where

$$m^2 = \{C(L-M)\}^{-1}, \qquad n^2 = \{C(L+M)\}^{-1}$$

(as in (3.3.26)).

**4.5.** Given

$$\left.\begin{aligned} L\ddot{q}_1 + M\ddot{q}_2 + q_1/C &= E(t), \\ M\ddot{q}_1 + L\ddot{q}_2 + q_2/C &= 0, \end{aligned}\right\} \quad t > 0,$$

with $q_1(0) = 0$, $\dot{q}_1(0) = 0$, $q_2(0) = 0$, $\dot{q}_2(0) = 0$, find $q_2(t)$. (Compare (3.4.28).)

**4.6.** Given $(D^4 - n^4)x = \cos nt$, $t > 0$, with $x(0) = 1$, $\dot{x}(0) = 0$, $\ddot{x}(0) = 0$, $\dddot{x}(0) = 0$, prove

$$x(t) = \tfrac{1}{2}\left(1 - \frac{1}{2n^4}\right)\cos nt - \frac{1}{4n^3}\, t \sin nt - \tfrac{1}{2}\left(1 + \frac{1}{2n^4}\right)\cosh nt.$$

**4.7.** (Heaviside's formula.) If in (4.5.4) $L(p)$ contains the factor $(p-\alpha)^m$, show by the method of Section 4.5 that the inverse transform of

$$\frac{R(p)}{L(p)} \equiv \frac{1}{(p-\alpha)^m}\, Q(p)$$

contains terms

$$e^{\alpha t} \sum_{s=1}^{m} \frac{Q^{(m-s)}(\alpha)}{(m-s)!}\, \frac{t^{s-1}}{(s-1)!}$$

corresponding to partial fractions arising from the factor $(p-\alpha)^m$.

**4.8.** If $f(t) = 1$ when $0 < t \leqq \frac{1}{2}T$,
$= 0$ when $\frac{1}{2}T < t \leqq T$,

and $f(t+T) = f(t)$,

prove $\bar{f}(p) = \dfrac{1}{2p}\, e^{\frac{1}{2}pT} \operatorname{sech} \tfrac{1}{4}pT.$

**4.9.** (The rectified sine wave.) Show that the Laplace transform of the function $\left|\sin(2\pi t/T)\right|$ is

$$\frac{2\pi/T}{\{p^2 + (2\pi/T)^2\}} \coth \tfrac{1}{2}pT.$$

CHAPTER 5

# ORDINARY LINEAR DIFFERENTIAL EQUATIONS WITH VARIABLE COEFFICIENTS

## 5.1. Introduction

In Chapters 3 and 4 we have considered only such linear equations as have constant coefficients. But it is essential to develop methods of solution which will apply to equations with variable coefficients—that is, with coefficients that are functions of the independent variable. As an introduction to new ideas, and to provide a necessary tool, we treat in Section 5.2 the first order equation with variable coefficients, called *Euler's equation*.

But the main effort of this chapter will be aimed at second order equations, because of the tremendous importance of such equations in mathematical physics. They arise in the following way: very commonly physical laws when translated into mathematical form will give partial differential equations of the second order, such as Laplace's equation, Poisson's equation, the Wave equation and Schrödinger's equation. The most powerful method of attack on such equations is by *separation of the variables*: this method leads to ordinary differential equations of the second order. The types of ordinary equation that appear are determined by the choice of coordinates, but all have the common property that they are linear, usually with variable coefficients.

In Chapters 5 and 6 we will be concerned with methods of analysis applicable to such equations. We will not discuss physical or chemical entities or relationships, but will fashion tools without which the mathematics describing those relationships cannot be constructed.

Although second order equations are by far the most important, linear equations of higher order are also met in mathematical physics. Some of the methods of Chapters 5 and 6 can be extended to deal with equations of order higher than the second. The more general theory is given by

Ince, E. L., *Ordinary Differential Equations*,[1] Chapters 5, 9, 11.

## 5.2. The linear equation of the first order (Euler's equation)

In Section 3.1 we considered the equation of natural decay disturbed by a given function of the time :

$$\frac{dx}{dt}+kx = f(t), \quad t > t_0, \tag{5.2.1}$$

$k$ being a positive constant. The equation was solved by the use of an integrating factor.

Let us now move on to the corresponding equation in which $k$ is a given function of the time, $k(t)$. This describes decay in which the intrinsic rate of decrease of $x$ is at any instant proportional to the amount of $x$ at that instant, the parameter of proportionality varying with time. There is also, as before, an externally imposed contribution $f(t)$ to the rate of change of $x$. Thus we start from the first order linear equation

$$\frac{dx}{dt}+k(t)x = f(t), \quad t > t_0. \tag{5.2.2}$$

The method that uses an integrating factor, introduced in Section 3.1, is again available here. We multiply equation (2) by a function $g(t)$ that is to be such that the left-hand side is the derivative of $x(t)g(t)$. This entails

$$g(t)k(t) = g'(t),$$

so that

$$g(t) = \exp \int^t k(\tau)d\tau, \tag{5.2.3}$$

where the lower end-point of integration can be selected according to convenience, since a change in it merely introduces a constant multiplier into $g(t)$. (2) is then

$$\frac{d}{dt}\{x(t)g(t)\} = f(t)g(t), \tag{5.2.4}$$

and so

$$x(t)g(t)-x(t_0)g(t_0) = \int_{t_0}^t f(u)g(u)du. \tag{5.2.5}$$

Thus our problem has been reduced to one of quadrature and is regarded as solved. The factor (3) is the *integrating factor* and is of such

importance in this subject that the reader will do well to commit it to memory. Equation (2) was first solved by Euler, and bears his name.

*Example* 1. To solve

$$\frac{\mathrm{d}x}{\mathrm{d}t} - \tan t\, x = \sin t, \text{ with } x(t_0) = x_0.$$

An integrating factor is

$$\begin{aligned} &\exp\Big(-\int_c^t \tan\tau\, \mathrm{d}\tau\Big) \\ &\quad = \exp(\ln\cos t - \ln\cos c) \\ &\quad = \cos t \end{aligned}$$

if we choose $c = 0$. Hence the differential equation is

$$\frac{\mathrm{d}}{\mathrm{d}t}(x\cos t) = \sin t\cos t,$$

giving

$$[x\cos t]_{t_0}^{t} = \int_{t_0}^{t} \sin u\cos u\, \mathrm{d}u$$

and

$$x\cos t - x_0\cos t_0 = \tfrac{1}{2}(\sin^2 t - \sin^2 t_0).$$

*Example* 2. To solve

$$\frac{\mathrm{d}x}{\mathrm{d}t} + (a+bt)x = f(t), \text{ with } x(0) = x_0.$$

An integrating factor is

$$\exp\int_c^t (a+bt)\mathrm{d}t = \exp\,(at + \tfrac{1}{2}bt^2)$$

if we take $c = 0$. Hence

$$\frac{\mathrm{d}}{\mathrm{d}t}(x\mathrm{e}^{at+\frac{1}{2}bt^2}) = f(t)\mathrm{e}^{at+\frac{1}{2}bt^2},$$

giving

$$x\mathrm{e}^{at+\frac{1}{2}bt^2} - x_0 = \int_0^t f(u)\mathrm{e}^{au+\frac{1}{2}bu^2}\mathrm{d}u.$$

[An equation like this would arise if we were using $a+bt$ as an approximation to $k(t)$ over a certain range : this approximation is a replacement of part of the graph of $k(t)$ by a straight line fitted by suitable choice of $a$ and $b$.]

If the linear equation has its most general form

$$p_0(t)\frac{\mathrm{d}x}{\mathrm{d}t}+p_1(t)x = f(t) \tag{5.2.6}$$

we first divide through by $p_0(t)$, and then proceed as above : the net result is that equation (6) will have been multiplied through by

$$\frac{1}{p_0}\exp\int^t \frac{p_1(\tau)}{p_0(\tau)}\,\mathrm{d}\tau.$$

This is valid as long as the range of $t$ that concerns us does not contain a zero of $p_0(t)$. The vanishing of $p_0(t)$ within or at an end-point of the range will correspond to some physical feature of the problem, such as resonance. If the vanishing of $p_0(t)$ does not cause the integrals that contain it to diverge our solution will stand, but if the integrals diverge the "solution" will be mathematically and physically unacceptable, and we will be forced to reconsider the applicability of our differential equation in the neighbourhood of the value of $t$ where $p_0(t) = 0$.

We now proceed to a more detailed examination of the meaning of the solution (5). Let us rewrite it using the notation

$$\int^t k(\tau)\mathrm{d}\tau = K(t), \tag{5.2.7}$$

so that

$$g(t) = \exp K(t).$$

Then (5) becomes

$$x(t)\mathrm{e}^{K(t)} - x(t_0)\mathrm{e}^{K(t_0)} = \int_{t_0}^{t} f(u)\mathrm{e}^{K(u)}\mathrm{d}u$$

or

$$x(t) = x_0\mathrm{e}^{-\{K(t)-K(t_0)\}} + \int_{t_0}^{t} f(u)\mathrm{e}^{-\{K(t)-K(u)\}}\mathrm{d}u, \tag{5.2.8}$$

where $x_0 = x(t_0)$.

Here we have exhibited the solution as the sum of two terms. The initial value $x_0$ appears only in the first, and the disturbing force $f(t)$

only in the second. $K(t)$, which appears in both terms, is derived from $k(t)$ and depends on the intrinsic character of the system.

As in Section 3.1, we will call equation (2) the *inhomogeneous equation*, and the equation

$$\frac{\mathrm{d}x}{\mathrm{d}t}+k(t)x = 0, \quad t > t_0, \tag{5.2.9}$$

obtained from (2) by putting $f(t) \equiv 0$, the *corresponding homogeneous equation*. We now see that the first term of (8) is a solution of the homogeneous equation : it represents the response of the system if left free from external force and started with $x(t_0) = x_0$.

The second term of (8), on the other hand, may be isolated by writing $x_0 = 0$. Let us examine it further. If in

$$\int_{t_0}^{t} f(u)\mathrm{e}^{-\{K(t)-K(u)\}}\mathrm{d}u \tag{5.2.10}$$

we put $f(t) = \delta(t-T)$, it reduces to

$$\begin{cases} \mathrm{e}^{-\{K(t)-K(T)\}} & \text{if } T \text{ lies within } (t_0, t), \\ 0 & \text{if } T \text{ lies outside } (t_0, t), \end{cases}$$

and since we are interested only in times later than $t_0$ we may combine these two expressions into

$$\mathrm{e}^{-\{K(t)-K(T)\}}H(t-T). \tag{5.2.11}$$

This is the response to a delta-function disturbance at $t = T$. Superposing such solutions, we see that the response to $f(u)\delta(t-u)$ is

$$f(u)\mathrm{e}^{-\{K(t)-K(u)\}}H(t-u),$$

and since an integral is the limit of a sum, the response to

$$\int_{t_0}^{t_1} f(u)\delta(t-u)\mathrm{d}u \tag{5.2.12}$$

is

$$\int_{t_0}^{t_1} f(u)\mathrm{e}^{-\{K(t)-K(u)\}}H(t-u)\mathrm{d}u. \tag{5.2.13}$$

Let us take $t_1 > t$, so that the delta-function falls at $u = t$ within $(t_0, t_1)$. Then, by Section 4.3, (12) becomes $f(t)$. In (13), $H(t-u)$ is

unity as long as $t_0 < u < t$, and zero when $t < u < t_1$, so that the integral reduces to

$$\int_{t_0}^{t} f(u)\mathrm{e}^{-\{K(t)-K(u)\}}\mathrm{d}u \tag{5.2.14}$$

(the upper end-point now lying at $t$). Thus the statement containing (12) and (13) now becomes :

$$\text{the response to } f(t) \text{ is } \int_{t_0}^{t} f(u)\mathrm{e}^{-\{K(t)-K(u)\}}\mathrm{d}u. \tag{5.2.15}$$

Formula (15), giving the response to an arbitrary $f(t)$, is built up from $f(u)$ and

$$\mathrm{e}^{-\{K(t)-K(u)\}}H(t-u), \tag{5.2.16}$$

which is the response to the delta-function. (16) is called the *Green's function* for this system. Returning to (8) we see that the integral there represents the response to $f(t)$ which is built up of the responses to delta-functions.

## 5.3. The linear equation of the second order

The second order linear equation usually arises from the separation of space variables in a second order partial differential equation, and so the single independent variable that enters the ordinary differential equation will usually be not time but length or angle or some other parameter of position in space. We will therefore use not $t$ but $s$ to denote the independent variable. Further, the boundary conditions will commonly apply at the two extremes of $s$ (as for instance when temperature is specified at the two bounding faces of a slab) so that the problems will be *jury problems* (see Section 1.4).

We are thus led to examine the inhomogeneous equation (in which primes denote derivatives with regard to $s$)

$$p_0(s)x'' + p_1(s)x' + p_2(s)x = f(s), \quad s_0 < s < s_1, \tag{5.3.1}$$

and the corresponding homogeneous equation

$$p_0(s)x'' + p_1(s)x' + p_2(s)x = 0, \quad s_0 < s < s_1. \tag{5.3.2}$$

We will assume that $p_0(s)$, $p_1(s)$ and $p_2(s)$ are continuous functions of $s$ in $(s_0, s_1)$. Usually they will be simple polynomials. The appropriate boundary conditions will be introduced later.

Certain general theorems have been established regarding the solutions of (1) and (2). We state them here, referring the reader who wishes to follow the proofs to

Burkill, J. C., *The Theory of Ordinary Differential Equations.*[2] Ch. 2.
Ince, E. L., *Ordinary Differential Equations.*[1] Ch. 5.

*Theorem I.* There cannot exist more than two linearly independent solutions of the homogeneous equation (2).

(Linear independence is defined in Section 3.3.)
Let us denote any two solutions of (2) by $X_1(s)$ and $X_2(s)$.

*Theorem II.* A necessary and sufficient condition for $X_1$ and $X_2$ to be linearly independent solutions is

$$W(s) = X_1X_2' - X_1'X_2 \neq 0. \tag{5.3.3}$$

This function $W(s)$ is called the *Wronskian* of $X_1$ and $X_2$.

*Theorem III.* If $X_1$ and $X_2$ are linearly independent solutions of (2), and $X^*(s)$ is any solution of (1), then the general solution of (1) is

$$x(s) = a_1X_1 + a_2X_2 + X^*, \tag{5.3.4}$$

where $a_1$ and $a_2$ are arbitrary constants.

*Example.* We find the general solution of

$$sx'' - x' - 4s^3x = s^3.$$

The corresponding homogeneous equation is

$$sx'' - x' - 4s^3x = 0,$$

and it can be verified by direct substitution or by use of the change of variable $s = u^{\frac{1}{2}}$ that the general solution of this is

$$x = Ae^{s^2} + Be^{-s^2},$$

where $A$ and $B$ are constants. (It is easily seen that $e^{s^2}$ and $e^{-s^2}$ have Wronskian $-4s$, which is not in general zero, and are therefore linearly independent, by Theorem II.

By inspection, a particular integral of the inhomogeneous equation is $x = -\frac{1}{4}$. Hence, by Theorem III the general solution of the inhomogeneous equation is

$$x(s) = Ae^{s^2} + Be^{-s^2} - \tfrac{1}{4}.$$

*Discovery of a second solution of the homogeneous equation when one solution is known*

Let $x = Y(s)$ be a solution of (2). We seek a second solution by trying

$$x = Y(s)Z(s),$$

where $Z(s)$ is to be found. Substituting into (2), we have

$$p_0(Y''Z + 2Y'Z' + YZ'') + p_1(Y'Z + YZ') + p_2YZ = 0.$$

Since $Y$ is a solution of (2) the terms containing $Z$ add to zero. We are left with

$$p_0YZ'' + (2p_0Y' + p_1Y)Z' = 0. \tag{5.3.5}$$

This is a linear equation for $Z'$ and can be solved by Euler's method (Section 5.2). Hence in principle we can find $Z(s)$ and the second solution. We note that $p_0(s)$ and $p_1(s)$ must be such that the integrals that arise do exist.

*Example.* Given that $x = e^{s^2}$ is a solution of

$$x'' - 2sx' - 2x = 0,$$

to find a second solution.

We write

$$x = e^{s^2}Z(s),$$

so that

$$x' = 2se^{s^2}Z + e^{s^2}Z'$$

$$x'' = (2e^{s^2} + 4s^2e^{s^2})Z + 4se^{s^2}Z' + e^{s^2}Z'',$$

and

$$\begin{aligned} x'' - 2sx' - 2x &= e^{s^2}\{(2 + 4s^2)Z + 4sZ' + Z'' - 4s^2Z - 2sZ' - 2Z\} \\ &= e^{s^2}(Z'' + 2sZ'). \end{aligned}$$

Writing $v$ for $Z'$, we see that we have to solve

$$v' + 2sv = 0.$$

A solution is $v = \exp(-s^2)$ (by separation of variables), so that

$$Z = \int^s e^{-\sigma^2}d\sigma$$

and the required second solution is

$$x = e^{s^2}\int^s e^{-\sigma^2}d\sigma.$$

[The lower end-point of the integral may be fixed at our convenience, for a change in it merely adds to our second solution a multiple of the first. Again, insertion of an arbitrary parameter in $v$ means only the multiplication of our second solution by a constant.]

*Reduction of the second order linear equation to standard form*

Is it possible by some change of variables to simplify the form of the general second order equation (1)? Such simplification may lead to the discovery of properties of the solutions that are obscured by the complexity of the relation between $x$, its two derivatives, and the forcing term $f(s)$.

We will show that it is possible to reduce the number of terms in the equation, and to derive a form that for some purposes is much simpler. Let us again seek a factor $q(s)$ such that after multiplication by this factor the first two terms become an exact derivative. Thus

$$qp_0x'' + qp_1x'$$

is to be an exact derivative. The first term is one part of

$$\frac{\mathrm{d}}{\mathrm{d}s}(qp_0x').$$

Equating the second term to the other part we have

$$qp_1 = \frac{\mathrm{d}}{\mathrm{d}s}(qp_0) = q'p_0 + qp_0',$$

so that

$$\frac{q'}{q} = \frac{p_1}{p_0} - \frac{p_0'}{p_0}$$

and

$$\ln q = \int^s \frac{p_1}{p_0}\,\mathrm{d}s - \ln p_0,$$

giving

$$q = \frac{1}{p_0}\exp\int^s \frac{p_1}{p_0}\,\mathrm{d}s.$$

The differential equation now becomes

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(qp_0\frac{\mathrm{d}x}{\mathrm{d}s}\right) + qp_2x = qf. \tag{5.3.6}$$

If we now introduce a new variable $\eta$, defined by

$$qp_0 \frac{dx}{ds} = \frac{dx}{d\eta}, \tag{5.3.7}$$

the equation (6) becomes

$$\frac{d^2x}{d\eta^2} + q^2 p_0 p_2 x = q^2 p_0 f. \tag{5.3.8}$$

This will be an equation for $x$ as a function of $\eta$ if $p_0$, $p_2$, $f$ and $q$ are expressed in terms of $\eta$. This requires an equation giving $s$ in terms of $\eta$, which is provided by (7) if we write

$$d\eta = \frac{ds}{q(s)p_0(s)}$$

and integrate. Substituting $s = s(\eta)$ into (8) we obtain

$$\frac{d^2x}{d\eta^2} + r(\eta)x = g(\eta), \tag{5.3.9}$$

which is theoretically the simplest form into which the equation (1) may be transformed. Equation (9) is called the *standard form*. It is clear that in practice the difficulties of deriving (9) may be formidable. We will make use of (9) in Section 5.4 where we investigate oscillating solutions, and again in Chapter 6 when we seek approximate solutions. But for other purposes the original form (1) may be more convenient, especially if $s$ can be given a clear physical meaning whereas $\eta$ cannot, or if $p_0$, $p_1$ and $p_2$ are simple polynomials and we seek a solution in series, as in the later part of this section.

We note that the discovery of $q(s)$ and the change of variable to $\eta$ may fail if $p_0$ has a zero within or at an end-point of the range $(s_0, s_1)$.

*Solution of the inhomogeneous equation when solutions of the homogeneous equation are known*

If the solution of the (simpler) homogeneous equation (2) is known, can we use it, as with equations having constant coefficients, to find the solution of the inhomogeneous equation (1)? We can, by the method called *variation of parameters* and due to Lagrange.

Let $X_1(s)$ and $X_2(s)$ be known linearly independent solutions of the homogeneous equation (2). Then by (3)

$$W(s) = X_1X_2' - X_1'X_2 \neq 0. \tag{5.3.10}$$

We now make the assumption that a solution of (1) may be found in the form

$$x(s) = u_1(s)X_1(s) + u_2(s)X_2(s),$$

where $u_1(s)$ and $u_2(s)$ are functions which we have to discover. Differentiating, we obtain

$$x'(s) = u_1X_1' + u_2X_2' + u_1'X_1 + u_2'X_2. \tag{5.3.11}$$

Let us demand that the unknown functions satisfy

$$u_1'X_1 + u_2'X_2 = 0, \tag{5.3.12}$$

so that in (11) only the first two terms remain. Differentiating again we find

$$x''(s) = u_1X_1'' + u_2X_2'' + u_1'X_1' + u_2'X_2'.$$

Substituting for $x'$ and $x''$ into (1), we have

$$p_0(u_1X_1'' + u_2X_2'' + u_1'X_1' + u_2'X_2') + p_1(u_1X_1' + u_2X_2') + p_2(u_1X_1 + u_2X_2) = f.$$

In this equation the coefficient of $u_1$ is $p_0X_1'' + p_1X_1' + p_2X_1$, which vanishes since $X_1$ is a solution of (2). Similarly the coefficient of $u_2$ vanishes, and we are left with

$$u_1'p_0X_1' + u_2'p_0X_2' = f. \tag{5.3.13}$$

The condition (10) ensures that the two equations (12) and (13) can be solved for $u_1'$ and $u_2'$. They yield

$$u_1' = \frac{-X_2}{X_1X_2' - X_1'X_2}\frac{f}{p_0}, \quad u_2' = \frac{X_1}{X_1X_2' - X_1'X_2}\frac{f}{p_0}, \tag{5.3.14}$$

and these can be integrated to give $u_1$ and $u_2$, and hence

$$x(s) = X_1(s)\int^s \frac{-X_2(\sigma)}{W(\sigma)}\frac{f(\sigma)}{p_0(\sigma)}\,d\sigma + X_2(s)\int^s \frac{X_1(\sigma)}{W(\sigma)}\frac{f(\sigma)}{p_0(\sigma)}\,d\sigma \tag{5.3.15}$$

where $W$ is the Wronskian $X_1X_2' - X_1'X_2$. The lower end-point of each integral is at our disposal, and in each case a change in the end-point at most adds a constant multiple of $X_1(s)$ or $X_2(s)$ to the solution $x(s)$. Thus we have in (15) a solution of (1). To obtain a solution

which can be fitted to two boundary conditions, we add to (15) the terms $AX_1+BX_2$ which form the general solution of the corresponding homogeneous equation (2).

*Example.* It is required to find the solution of

$$x''+n^2x = f(s),\ n \text{ real and constant},\ 0 < s < l, \qquad (5.3.16)$$

which satisfies the boundary conditions $x(0) = 0$, $x(l) = 0$. Here we take as our independent solutions of the auxiliary equation $\cos ns$ and $\sin ns$. Then

$$W = \cos ns(n\cos ns)-(-n\sin ns)\sin ns = n,$$

so that formula (15) gives

$$x(s) = \cos ns\int_0^s \frac{(-\sin n\sigma)}{n} f(\sigma)\mathrm{d}\sigma+\sin ns\int_0^s \frac{\cos n\sigma}{n} f(\sigma)\mathrm{d}\sigma,$$

where we have arbitrarily taken the lower end-point as zero in each integral. Hence a particular solution of (16) is

$$x(s) = \frac{1}{n}\int_0^s \sin n(s-\sigma) f(\sigma)\mathrm{d}\sigma. \qquad (5.3.17)$$

Adding the general solution of the corresponding homogeneous equation, we have as complete solution of (16)

$$x(s) = C\cos ns+D\sin ns+\frac{1}{n}\int_0^s \sin n(s-\sigma) f(\sigma)\mathrm{d}\sigma.$$

The boundary conditions can now be satisfied by suitable choice of $C$ and $D$. $x(0) = 0$ and $x(l) = 0$ give

$$C = 0$$

$$D\sin nl+\frac{1}{n}\int_0^l \sin n(l-\sigma) f(\sigma)\mathrm{d}\sigma = 0,$$

and the required solution is

$$x(s) = -\frac{\sin ns}{n \sin nl}\int_0^l \sin n(l-\sigma)\, f(\sigma)\mathrm{d}\sigma + \frac{1}{n}\int_0^s \sin n(s-\sigma)\, f(\sigma)\mathrm{d}\sigma. \tag{5.3.18}$$

If we split off from the first integral an integral from 0 to $s$, which we then combine with the second integral, we obtain

$$x(s) = -\int_0^s \frac{\sin n\sigma \sin n(l-s)}{n \sin nl} f(\sigma)\mathrm{d}\sigma - \int_s^l \frac{\sin ns \sin n(l-\sigma)}{n \sin nl} f(\sigma)\mathrm{d}\sigma$$

$$= \int_0^l G(s, \sigma) f(\sigma)\mathrm{d}\sigma, \tag{5.3.19}$$

where
$$G(s, \sigma) = \begin{cases} -\dfrac{\sin n\sigma \sin n(l-s)}{n \sin nl} & \text{in} \quad 0 < \sigma < s \\ -\dfrac{\sin ns \sin n(l-\sigma)}{n \sin nl} & \text{in} \quad s < \sigma < l. \end{cases}$$

We observe that if we take $f(s)$ to be the delta-function at $s = \xi$, namely $\delta(s-\xi)$, then the solution (19) becomes

$$x(s) = \begin{cases} -\dfrac{\sin n\xi \sin n(l-s)}{n \sin nl} & \text{if} \quad 0 < \xi < s < l, \\ -\dfrac{\sin ns \sin n(l-\xi)}{n \sin nl} & \text{if} \quad 0 < s < \xi < l, \end{cases}$$

$$= G(s, \xi).$$

Thus $G(s, \sigma)$ is the response to a delta-function disturbance at $s = \sigma$, and if we know it we can immediately build up the response (19) to $f(s)$. $G(s, \sigma)$ is the *Green's function* for this problem.

## 5.4. Solution in series

Only a few very special types of second order homogeneous linear equations possess solutions expressible in terms of a finite number of elementary functions. The example worked above is one of these.

But in general we must expect that $X_1(s)$ and $X_2(s)$ will be new functions, which will in fact be defined by the differential equation; for instance the equation

$$x'' - sx = 0, \quad -\infty < s < \infty, \tag{5.4.1}$$

which is very simple in form, defines the Airy functions $Ai(s)$ and $Bi(s)$ (see Section 5.5). Some of the properties of the new functions thus defined may be deduced directly from their differential equations (see for instance the discussion of eigenvalues with the help of Bessel's equation in Burkill, § 46). But for other properties we must seek to represent the functions by infinite series. We therefore introduce here the method of *solution in series*.

As an example we solve the *Airy equation* (1). Let us substitute

$$x = a_0 + a_1 s + a_2 s^2 + \ldots + a_r s^r + \ldots \tag{5.4.2}$$

so that we obtain

$$\begin{aligned} &2a_2 + 3.2.a_3 s + 4.3.a_4 s^2 + \ldots + (r+3)(r+2)a_{r+3}s^{r+1} + \ldots \\ &-( \quad a_0 s + \quad a_1 s^2 + \ldots + \quad a_r s^{r+1} + \ldots) = 0. \end{aligned}$$

This must hold for all $s$, and therefore the coefficient of each power of $s$ must vanish. We have

$$\begin{aligned} s^0 : \quad & a_2 = 0, \\ s^1 : \quad & 3.2.a_3 - a_0 = 0, \\ s^2 : \quad & 4.3.a_4 - a_1 = 0, \\ & \cdots \\ s^{r+1} : \quad & (r+3)(r+2)a_{r+3} - a_r = 0, \\ & \cdots \end{aligned}$$

These equations do not determine $a_0$ or $a_1$; but if we take assumed values for these two coefficients all others will be determined in terms of them. Thus

$$a_3 = \frac{1}{3.2}a_0, \qquad a_4 = \frac{1}{4.3}a_1,$$

$$a_6 = \frac{1}{6.5}a_3 = \frac{1}{6.5.3.2}a_0, \quad \text{etc.},$$

while all terms of the sequence $a_{2+3k}$ ($k$ a positive integer) will vanish since $a_2$ vanishes.

Our solution is therefore

$$x = a_0\left(1+\frac{1}{3.2}s^3+\frac{1}{6.5.3.2}s^6+\ldots\right)$$
$$+a_1\left(s+\frac{1}{4.3}s^4+\frac{1}{7.6.4.3}s^7+\ldots\right). \qquad (5.4.3)$$

Each of the bracketed series converges for all values of $s$, and since $a_0$ and $a_1$ are at our disposal and the series are linearly independent (the

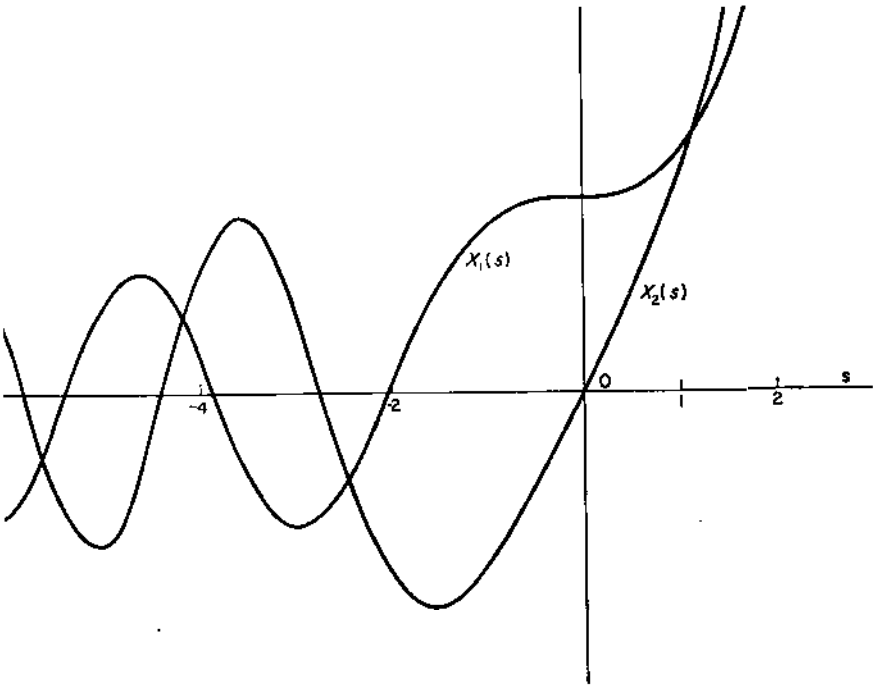


FIG. 5.1. $X_1(s)$ and $X_2(s)$, solutions of Airy's equation.

second not being a multiple of the first) we have the general solution for $x(s)$. The functions represented by the two series are shown in Fig. 5.1.

From our success in this example, it might be supposed that whatever differential equation we meet we can assume that there is a solution in series of the form (2). But that is correct only when the unknown solution does in fact possess a Taylor expansion in powers of $s$ in the neighbourhood of $s = 0$. For many of the most interesting equations this is not true.

A point at which no expansion of a function as a Taylor series exists is called a *singularity* (or singular point) of that function. For instance, $(1-s)^{-1}$ has a singularity at $s = 1$.

For let

$$f(s) = \frac{1}{1-s};$$

then near $s = s_0$, some point in the interval $(0, 1)$,

$$f(s) = \frac{1}{1-s_0-(s-s_0)}$$

$$= \frac{1}{1-s_0}\left\{1+\frac{s-s_0}{1-s_0}+\frac{(s-s_0)^2}{(1-s_0)^2}+\ldots\right\},$$

provided $|s-s_0| < 1-s_0$, which is a Taylor expansion in powers of $s-s_0$. But the region of validity shrinks as $s_0$ approaches 1, and we cannot expand $1/(1-s)$ in a series of positive powers of $(s-1)$. At the point $s_0 = 1$ the Taylor series fails to exist. In fact the function and all its derivatives are infinite at $s = 1$, which is a singularity. Points at which a Taylor expansion exists are called *regular points.*

An example in which the solution cannot be expressed as a Taylor series in powers of $s$, with a suggestion as to how to proceed, is provided by the simple first order equation

$$x' - \frac{\alpha}{s}x = 0, \text{ where } \alpha \text{ is a constant.} \tag{5.4.4}$$

This equation is immediately soluble by separation of variables, giving

$$x = As^{\alpha}, \quad A \text{ being a constant.} \tag{5.4.5}$$

If $\alpha$ is a positive integer this solution is itself a vestigial Taylor expansion at the origin, consisting of a single term. But if $\alpha$ is not a positive integer no such Taylor expansion exists. If $\alpha$ is negative, $x$ and all its derivatives become infinite at $s = 0$, while if $\alpha$ is positive but not an integer all the derivatives of $x$ beyond a certain derivative become infinite at $s = 0$. In either case the Taylor expansion cannot be constructed.

We observe that in (4) the coefficient of $x$ has a singularity at $s = 0$: it is this that introduces the danger of breakdown of the Taylor series at $s = 0$. Taking this hint, we write the second order equation (3.2) in the form

$$x'' + (p_1/p_0)x' + (p_2/p_0)x = 0.$$

A singularity of either of the coefficients $p_1/p_0$, $p_2/p_0$ will be called a *singularity of the equation*. If there is such a singularity at the origin, we will seek a solution of the form

$$x = s^\gamma(a_0 + a_1 s + a_2 s^2 + \ldots + a_r s^r + \ldots), \tag{5.4.6}$$

determining $\gamma$ and the coefficients $a_0, a_1, \ldots, a_r, \ldots$ so that the equation is satisfied for all $s$ within a certain region.

With these principles in mind, let us look for solutions of *Bessel's equation*

$$s^2x'' + sx' + (s^2 - \nu^2)x = 0, \tag{5.4.7}$$

where for the time being we impose on $\nu$ the condition that it shall not be an integer. If we write the equation as

$$x'' + \frac{1}{s}x' + \left(1 - \frac{\nu^2}{s^2}\right)x = 0, \tag{5.4.8}$$

we see that the origin is a singular point. We therefore seek a solution in the form (6), which we take to mean that $a_0$ is the first *non-zero* coefficient. Substituting this into (7), we have

$$\begin{aligned}
&\gamma(\gamma-1)a_0 s^\gamma + (\gamma+1)\gamma a_1 s^{\gamma+1} + \ldots + (\gamma+r)(\gamma+r-1)a_r s^{\gamma+r} + \ldots \\
+\;&\gamma a_0 s^\gamma + (\gamma+1)a_1 s^{\gamma+1} + \ldots + (\gamma+r)a_r s^{\gamma+r} + \ldots \\
+\;&a_0 s^{\gamma+2} + \ldots + a_{r-2}s^{\gamma+r} + \ldots \\
-\;&\nu^2 a_0 s^\gamma - \nu^2 a_1 s^{\gamma+1} - \ldots - \nu^2 a_r s^{\gamma+r} - \ldots = 0.
\end{aligned}$$

Annulling all coefficients we obtain the set of equations:

$$\begin{aligned}
s^\gamma: &\quad (\gamma^2 - \nu^2)a_0 = 0, \\
s^{\gamma+1}: &\quad \{(\gamma+1)^2 - \nu^2\}a_1 = 0, \\
s^{\gamma+2}: &\quad \{(\gamma+2)^2 - \nu^2\}a_2 + a_0 = 0, \\
&\quad \cdots \\
s^{\gamma+r}: &\quad \{(\gamma+r)^2 - \nu^2\}a_r + a_{r-2} = 0, \\
&\quad \cdots
\end{aligned} \tag{5.4.9}$$

The first equation of (9) will be satisfied if $\gamma = \pm\nu$ ($a_0 = 0$ has been excluded by our hypothesis). Then, by the second equation, $a_1 = 0$. The remaining equations show that all $a_r$ vanish when $r$ is odd, and that

$$a_r = -\frac{1}{(\gamma+r)^2 - \nu^2}a_{r-2}$$

when $r$ is even. Thus taking $\gamma = +\nu$ we obtain

$$a_r = -\frac{1}{r(2\nu+r)}\, a_{r-2}, \tag{5.4.10}$$

and a solution

$$X_1(s) = a_0 s^\nu \left\{1 - \frac{1}{2(2\nu+2)}\, s^2 + \frac{1}{4(2\nu+4)2(2\nu+2)}\, s^4 - \ldots\right\}, \tag{5.4.11}$$

while taking $\gamma = -\nu$ we obtain

$$a_r = \frac{1}{r(2\nu-r)}\, a_{r-2}, \tag{5.4.12}$$

and a solution

$$X_2(s) = a_0' s^{-\nu} \left\{1 + \frac{1}{2(2\nu-2)}\, s^2 + \frac{1}{4(2\nu-4)2(2\nu-2)}\, s^4 + \ldots\right\}. \tag{5.4.13}$$

The general solution is a linear combination of these two. As would have been expected, (13) may be derived from (11) by writing $(-\nu)$ for $\nu$. $X_1(s)$ and $X_2(s)$ are shown in Fig. 5.2 for $\nu = 3/2$, $a_0' = -3a_0$.

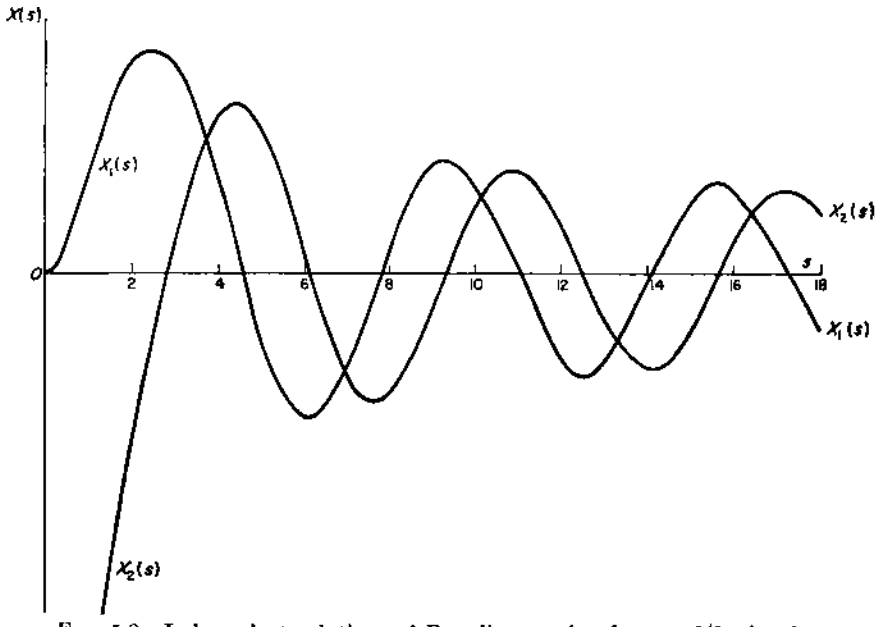


FIG. 5.2. Independent solutions of Bessel's equation for $\nu = 3/2$, showing interlacing (p. 150).

We see why we imposed the condition that $\nu$ should not be an integer. If $\nu$ is a positive integer, (12) will give an infinite ratio $a_r/a_{r-2}$ at some stage and the series (13) cannot be employed : the series (11), however, still obtains. Similarly if $\nu$ is a negative integer we have the single solution (13). Thus when $\nu$ is an integer we have determined one solution only and the second is still to be found, but if $\nu$ is not an integer we have two independent solutions and can construct a general solution.

As a second example of a differential equation for which the origin is a singular point, we consider the *Laguerre equation*

$$sx'' + (1-s)x' + \nu x = 0, \tag{5.4.14}$$

where $\nu$ is a constant. Substituting as before the series (6) for $x$ we obtain

$$\begin{aligned}
&\gamma(\gamma-1)a_0 s^{\gamma-1} + (\gamma+1)\gamma a_1 s^{\gamma} + \ldots + (\gamma+r)(\gamma+r-1)a_r s^{\gamma+r-1} + \ldots \\
+\quad &\gamma a_0 s^{\gamma-1} + (\gamma+1)a_1 s^{\gamma} + \ldots + (\gamma+r)a_r s^{\gamma+r-1} + \ldots \\
- \quad &\gamma a_0 s^{\gamma} - \ldots - (\gamma+r-1)a_{r-1} s^{\gamma+r-1} + \ldots \\
+ \quad &\nu a_0 s^{\gamma} + \ldots + \nu a_{r-1} s^{\gamma+r-1} + \ldots = 0,
\end{aligned}$$

leading to the set of equations

$$\begin{aligned}
s^{\gamma-1}: &\quad \gamma^2 a_0 = 0, \\
s^{\gamma}: &\quad (\gamma+1)^2 a_1 - (\gamma-\nu)a_0 = 0, \\
&\quad \ldots \\
s^{\gamma+r-1}: &\quad (\gamma+r)^2 a_r - (\gamma+r-1-\nu)a_{r-1} = 0.
\end{aligned} \tag{5.4.15}$$

The first equation, seen as an equation for $\gamma$, has two equal roots $\gamma = 0$. Then

$$a_r = \frac{r-1-\nu}{r^2} a_{r-1},$$

so that as long as $\nu$ is not a positive integer we obtain the single series solution

$$X(s) = a_0 \left\{ 1 - \nu s + \frac{\nu(\nu-1)}{2^2} s^2 - \ldots \right\}. \tag{5.4.16}$$

If $\nu$ is an integer this series ends at the term containing $s^{\nu}$, and the solution reduces to a polynomial. Such polynomials occur in the solution of Schrödinger's equation for the harmonic oscillator.

We now turn to the general second order linear equation and consider the steps in solution by series. If the equation has no singular point we may proceed immediately to an ordinary power series solution as we did for the Airy equation (1). If the equation has singular points, we can transfer the origin to each one in turn. We suppose this to be done, so that henceforth we examine only solutions in the neighbourhood of a singularity at the origin. If the singularity is such that $sp_1/p_0$ and $s^2p_2/p_0$ are regular at the origin, the equation is said to have a *regular singularity* there. All of the equations with which we will be concerned are of this kind.

If the origin is a regular singularity, and we expand regular functions in Taylor series there, we may write

$$\begin{aligned} sp_1/p_0 &= b_0+b_1s+b_2s^2+ \ldots +b_rs^r+ \ldots, \\ s^2p_2/p_0 &= c_0+c_1s+c_2s^2+ \ldots +c_rs^r+ \ldots, \end{aligned} \tag{5.4.17}$$

so that the differential equation is

$$x''+\frac{1}{s}(b_0+b_1s+ \ldots)x'+\frac{1}{s^2}(c_0+c_1s+ \ldots)x = 0.$$

Substituting $x = s^\gamma(a_0+a_1s+ \ldots +a_rs^r+ \ldots)$ we obtain, after multiplying through by $s^2$,

$$\begin{aligned} &\gamma(\gamma-1)a_0s^\gamma+(\gamma+1)\gamma a_1s^{\gamma+1}+ \ldots +(\gamma+r)(\gamma+r-1)a_rs^{\gamma+r}+ \ldots \\ &+(b_0+b_1s+ \ldots)\{\gamma a_0s^\gamma+(\gamma+1)a_1s^{\gamma+1}+ \ldots +(\gamma+r)a_rs^{\gamma+r}+ \ldots\} \\ &+(c_0+c_1s+ \ldots)\{a_0s^\gamma+a_1s^{\gamma+1}+ \ldots +a_rs^{\gamma+r}+ \ldots\} = 0, \end{aligned}$$

and the vanishing of all coefficients gives

$$\begin{aligned} &\{\gamma(\gamma-1)+b_0\gamma+c_0\}a_0 = 0, \\ &\{(\gamma+1)\gamma+b_0(\gamma+1)+c_0\}a_1+(b_1\gamma+c_1)a_0 = 0, \\ &\qquad \ldots \\ &\{(\gamma+r)(\gamma+r-1)+b_0(\gamma+r)+c_0\}a_r+\sum_{n=1}^{r-1}\{b_n(\gamma+r-n)+c_n\}a_{r-n} = 0, \\ &\qquad \ldots \end{aligned} \tag{5.4.18}$$

Since we take $a_0$ to be non-zero the first of the equations (18) gives two values of $\gamma$. It is called the *indicial equation.* The remaining equations determine successive $a_r$ in terms of coefficients already found. Thus the problem is solved provided the values of $\gamma$ are unequal and the expression

$$(\gamma+r)(\gamma+r-1)+b_0(\gamma+r)+c_0 \tag{5.4.19}$$

never vanishes. But this does vanish if $\gamma+r$ satisfies the indicial equation

$$\gamma(\gamma-1)+b_0\gamma+c_0 = 0,$$

i.e. if the second root differs from the first by $r$, which is an integer. We have therefore to consider further the cases where the indicial equation has equal roots or roots separated by an integer.

Let us write $\Gamma(\gamma)$ for (19), and *without putting* $\Gamma(\gamma) = 0$ let us solve the remaining equations for the coefficients in terms of $a_0$, obtaining after substitution

$$X(s, \gamma) = s^\gamma\{a_0+a_1(a_0, \gamma)s+a_2(a_0, \gamma)s^2+\ldots\}.$$

If this expression is substituted into the differential equation, we obtain

$$\left(p_0\frac{\mathrm{d}^2}{\mathrm{d}s^2}+p_1\frac{\mathrm{d}}{\mathrm{d}s}+p_2\right)X(s, \gamma) = p_0\Gamma(\gamma)s^\gamma$$
$$= p_0(\gamma-\gamma_1)(\gamma-\gamma_2)s^\gamma, \qquad (5.4.20)$$

$\gamma_1$, $\gamma_2$ being the roots of the indicial equation.

If we now set $\gamma = \gamma_1$ or $\gamma = \gamma_2$ the right-hand side vanishes and $X(s, \gamma_1)$, $X(s, \gamma_2)$ are shown to be solutions. If $\gamma_1 = \gamma_2$ these solutions are identical : but we observe that if we then differentiate the equation (20) with respect to $\gamma$ we get

$$\left(p_0\frac{\mathrm{d}^2}{\mathrm{d}s^2}+p_1\frac{\mathrm{d}}{\mathrm{d}s}+p_2\right)\frac{\mathrm{d}}{\mathrm{d}\gamma}X(s, \gamma) = 2p_0(\gamma-\gamma_1)s+p_0(\gamma-\gamma_1)^2\frac{\mathrm{d}}{\mathrm{d}\gamma}(s^\gamma), \qquad (5.4.21)$$

and now when we put $\gamma = \gamma_1$, the right-hand side vanishes and we get

$$\left\{\frac{\mathrm{d}}{\mathrm{d}\gamma}X(s, \gamma)\right\}_{\gamma=\gamma_1}$$

as a second solution.

If the solution corresponding to $\gamma_2$ fails because (19) vanishes for some $r$, this means that all the coefficients beyond $a_{r-1}(a_0, \gamma)$ have in the denominator the factor $(\gamma-\gamma_2)$. In this case we can discover a second independent solution by writing

$$Y(s, \gamma) = (\gamma-\gamma_2)X(s, \gamma),$$

so that

$$\left(p_0\frac{\mathrm{d}^2}{\mathrm{d}s^2}+p_1\frac{\mathrm{d}}{\mathrm{d}s}+p_2\right)Y(s, \gamma) = p_0(\gamma-\gamma_1)(\gamma-\gamma_2)^2s^\gamma.$$

If we now differentiate with respect to $\gamma$, we find

$$\left(p_0\frac{\mathrm{d}^2}{\mathrm{d}s^2}+p_1\frac{\mathrm{d}}{\mathrm{d}s}+p_2\right)\frac{\mathrm{d}}{\mathrm{d}\gamma}\,Y(s,\gamma)$$

$$=p_0\left[(\gamma-\gamma_2)^2\frac{\mathrm{d}}{\mathrm{d}\gamma}\{(\gamma-\gamma_1)s^\gamma\}+2(\gamma-\gamma_2)(\gamma-\gamma_1)s^\gamma\right]$$

and again the right side vanishes if $\gamma=\gamma_2$, showing that

$$\left\{\frac{\mathrm{d}}{\mathrm{d}\gamma}Y(s,\gamma)\right\}_{\gamma=\gamma_2} \tag{5.4.22}$$

is a solution.

We now seem to have solutions $X(s,\gamma_1)$, $Y(s,\gamma_1)$, $Y(s,\gamma_2)$ and

$$\left\{\frac{\mathrm{d}}{\mathrm{d}\gamma}\,Y(s,\gamma)\right\}_{\gamma=\gamma_2},$$

but in fact the second and third of these are merely multiples of the first, and the fourth is the required new solution. This method is due to *Frobenius*.

We have seen that successive coefficients $a_r$ can be calculated from the set of equations (18). If each equation contains only two coefficients we can find an explicit form for $a_r$ in terms of $r$. But if each equation contains three or more coefficients it is in general not possible to find such an explicit form. Three-term relationships are discussed in

Jeffreys and Jeffreys, *Methods of Mathematical Physics*,[5] C.U.P.

When the relationship is two-term, the labour of finding the series can be cut down by a method described by W. G. Bickley.[3] We will explain the method by way of a particular equation—Airy's equation

$$x''-sx=0. \tag{5.4.1}$$

We first note that if for $x$ in (1) we were to substitute one power of $s$, say the expression $a_rs^r$, the first term would become $r(r-1)a_rs^{r-2}$ and the second would become $a_rs^{r+1}$. The difference between the indices in these expressions is 3. This we will call the *index interval*. The recurrence relation by which the coefficient of $s^r$ is annulled will link $a_{r-1}$ and $a_{r+2}$, i.e. constants whose suffixes differ by the index interval. Let us make use of this information and seek a series solution of (1) in the form

$$x=s^\gamma(a_0+a_3s^3+\ldots+a_{3n}s^{3n}+\ldots).$$

Substituting, we have

$$\gamma(\gamma-1)a_0s^{\gamma-2}+(\gamma+3)(\gamma+2)a_3s^{\gamma+1}+\ldots$$
$$+(\gamma+3n)(\gamma+3n-1)a_{3n}s^{\gamma+3n-2}+\ldots$$
$$-a_0s^{\gamma+1}-a_3s^{\gamma+4}\ldots-a_{3n}s^{\gamma+3n+1}+\ldots=0.$$

The indicial equation is $\gamma(\gamma-1)=0$ and the successive coefficients are determined by

$$(\gamma+3)(\gamma+2)a_3-a_0=0,$$
$$\ldots$$
$$(\gamma+3n)(\gamma+3n-1)a_{3n}-a_{3n-3}=0,$$
$$\ldots$$

giving

$$a_{3n}=\frac{1}{(\gamma+3n)(\gamma+3n-1)}a_{3n-3}.$$

Taking $\gamma=0$, we obtain $a_3=\frac{1}{3.2}a_0$ etc., and the solution

$$x=a_0\left(1+\frac{1}{3.2}s^3+\ldots\right),\quad a_0 \text{ arbitrary.}$$

Taking $\gamma=1$, we obtain $a_3=\frac{1}{4.3}a_0$ etc., and the solution

$$x=a_0\left(s+\frac{1}{4.3}s^4+\ldots\right),\quad a_0 \text{ arbitrary.}$$

These are the required independent solutions. They are shown graphically in Fig. 5.1.

Again, substituting $s^r$ for $x$ in Bessel's equation

$$s^2x''+sx'+(s^2-\nu^2)x=0, \tag{5.4.7}$$

we obtain terms with indices

$$r, r, r+2, r,$$

so that the index interval is 2. We therefore employ the series

$$x=s^\gamma(a_0+a_2s^2+\ldots+a_{2m}s^{2m}+\ldots), \tag{5.4.23}$$

and obtain after substitution

$$\gamma(\gamma-1)a_0s^\gamma+(\gamma+2)(\gamma+1)a_2s^{\gamma+2}+\ldots$$
$$+(\gamma+2m)(\gamma+2m-1)a_{2m}s^{\gamma+2m}+\ldots$$
$$+\quad \gamma a_0s^\gamma+\quad (\gamma+2)a_2s^{\gamma+2}+\ldots+(\gamma+2m)a_{2m}s^{\gamma+2m}+\ldots$$
$$+\quad a_0s^{\gamma+2}+\ldots+\quad a_{2m-2}s^{\gamma+2m}+\ldots$$
$$-\quad \nu^2a_0s^\gamma-\quad \nu^2a_2s^{\gamma+2}-\ldots-\quad \nu^2a_{2m}s^{\gamma+2m}-\ldots=0,$$

so that the indicial equation is

$$\gamma(\gamma-1)+\gamma-\nu^2 = 0$$

and the coefficients are related by

$$\{(\gamma+2m)(\gamma+2m-1)+(\gamma+2m)-\nu^2\}a_{2m}+a_{2m-2} = 0.$$

Hence $\gamma = \pm\nu$ and

$$a_{2m} = -\frac{1}{(\gamma+2m)^2-\nu^2}\,a_{2m-2},$$

leading to two distinct solutions if $\nu$ is not an integer.

In general if the index interval is $j$, we have the relation between successive coefficients

$$a_{mj}F\{\gamma+mj\}+a_{(m-1)j}G\{\gamma+(m-1)j\} = 0 \qquad (5.4.24)$$

for $m = 1, 2, 3 \ldots$, where $F$ and $G$ are known functions. The indicial equation is

$$F(\gamma) = 0. \qquad (5.4.25)$$

Let the use of the relations (24) in the series lead to the expression

$$x = X(s, \gamma).$$

Then if the roots $\gamma_1$, $\gamma_2$ of the indicial equation do not differ by an integer the two solutions sought are $X(s, \gamma_1)$ and $X(s, \gamma_2)$. If the roots are equal, the solutions are

$$X(s, \gamma_1) \quad \text{and} \quad \left[\frac{\mathrm{d}}{\mathrm{d}\gamma}X(s, \gamma)\right]_{\gamma=\gamma_1}.$$

If the roots are $\gamma_1$ and $\gamma_2 = \gamma_1+rj$, then $F(\gamma_1+rj) = 0$, and a zero appears in the denominator of $a_{rj}$ and all succeeding coefficients, by (24). Then the required solutions are

$$X(s, \gamma_1) \quad \text{and} \quad \left[\frac{\partial}{\partial\gamma}\{(\gamma-\gamma_1-rj)X(s, \gamma)\}\right]_{\gamma=\gamma_1+rj}.$$

Applications of Frobenius' method will be found in Section 5.5 and in the Exercises at the end of this chapter.

## 5.5. Sturm–Liouville theory

In (5.3.6) we obtained, by change of variable, a standard form for the general linear differential equation of the second degree; this we now rewrite, with $s$ as independent variable, as

$$x''+r(s)x = 0.$$

In this section we derive properties of the solutions of this equation for very general forms of $r(s)$ and the boundary conditions. These properties will belong to many of the special functions, such as Airy functions, Bessel functions, and Legendre functions, which satisfy differential equations that can be transformed into the above form. The general tenor of our analysis is set by the observation that if $r(s)$ were a negative constant the solution of this equation would be a hyperbolic function $A \sinh \{(-r)^{\frac{1}{2}}s + \alpha\}$, whereas if $r(s)$ were a positive constant the solution would be a trigonometric function $B \sin (r^{\frac{1}{2}}s + \beta)$, $A, B, \alpha, \beta$ being adjustable parameters. The graph of the hyperbolic function cuts the $s$-axis once only, while the graph of the trigonometric

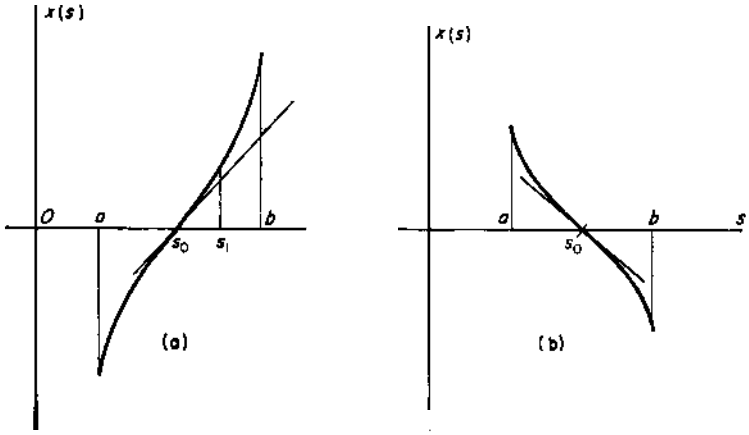


FIG. 5.3. Graph of $x(s)$ for $r(s) < 0$.
(a) where $x'(s_0) > 0$.
(b) where $x'(s_0) < 0$.

function is an oscillatory curve that cuts the $s$-axis in infinitely many equidistant points. We may expect that if $r(s)$ deviates slightly from constancy these features may be somewhat modified but not completely changed.

Consider a range $(a, b)$ of $s$, and let $r(s)$ be negative in this range. Then we can prove that there is not more than one zero of $x(s)$ in $(a, b)$. For let $s_0$ be a zero of $x(s)$, so that $x(s_0) = 0$. First assume $x'(s_0) > 0$, i.e. that the graph of $x(s)$ passes through $s_0$ with positive slope (see Fig. 5.3). Then near $s_0$ and on the right of it there is a region $(s_0, s_1)$ where $x(s) > 0$ and consequently

$$x''(s) = -r(s)x(s) > 0,$$

so that the slope $x'(s)$ is increasing in that region. [Note that we have assumed that the solution can be represented by a "well-behaved" curve, which is free from sudden jumps or changes in direction in the near neighbourhood of $s_0$. For the solution to have these properties there must be some restriction on $r(s)$. It can be shown that it is sufficient that $r(s)$ is continuous.] Since $x'(s)$ is increasing the curve is concave *up* in the range $(s_0, s_1)$. Thus $x(s_1) > 0$, $x'(s_1) > 0$, so that we can find a further region $(s_1, s_2)$, such that the curve is concave *up* in the region. Proceeding in this way we show that the curve $x(s)$ cannot meet the $s$-axis between $s = s_0$ and $s = b$.

On the left of $s_0$, on the other hand, there must be a region $(s_{-1}, s_0)$ where $x(s) < 0$ and consequently $x''(s) = -r(s)x(s) < 0$, so that the curve is concave down and $x(s_{-1}) < 0$. Again we extend step by step the range in which the curve is concave down, until we reach $s = a$. Thus we conclude that the graph of $x(s)$ is concave down in $(a, s_0)$ and concave up in $(s_0, b)$, having a point of inflexion at $s_0$ since $x''(s_0) = -r(s_0)x(s_0) = 0$. The slope is everywhere positive in $(a, b)$.

By similar arguments we show that if $x'(s_0) < 0$ the slope is everywhere negative in $(a, b)$, the graph being concave up in $(a, s_0)$ and concave down in $(s_0, b)$ (Fig. 5.3b). In either case there is only one zero of $x(s)$ in $(a, b)$. [The reader may enquire whether the graph might touch the $s$-axis at $s_0$. It is easily shown that if at $s_0$ both $x(s_0) = 0$ and $x'(s_0) = 0$ then every derivative is zero there and the function $x(s)$ is identically zero. We do not admit this as a solution of the differential equation.]

Next consider the corresponding problem in which it is given that $r(s)$ is *positive* in $(a, b)$. Again let $x(s_0) = 0$ and $x'(s_0) > 0$ for some $s_0$ in $(a, b)$. Then by arguments like that given above $x(s)$ is concave down on the right of $s_0$, bending down to meet the $s$-axis again, and concave up on the left of $s_0$, bending up towards the $s$-axis. At $s_0$ a point of inflexion divides the regions of opposite concavity. This is true in the neighbourhood of every zero of $x(s)$, so that $x(s)$ may be oscillatory (see Fig. 5.4). We now consider further this oscillation, relating its rapidity to the magnitude of $r(s)$.

Let us compare solutions of the two equations

$$x'' + r(s)x = 0, \quad a < s < b, \tag{5.5.1}$$

$$x'' + R(s)x = 0, \quad a < s < b, \tag{5.5.2}$$

that start with the same initial conditions, and let us postulate that $R(s) > r(s)$ for all $s$ in $(a, b)$. We will denote the solution of (1) that

takes the value $h\ (>0)$ at $s=a$ and has slope $k\ (>0)$ at $s=a$ by $y(s)$, and the solution of (2) that obeys the same conditions at $s=a$ by $Y(s)$. Thus the graphs of $y(s)$ and $Y(s)$ pass through the point $(a, h)$ with the same slope $k$ (see Fig. 5.5). Since $y$ and $Y$ are solutions of (1) and (2) respectively,

$$y''+ry = 0, \tag{5.5.3}$$

$$Y''+RY = 0, \tag{5.5.4}$$

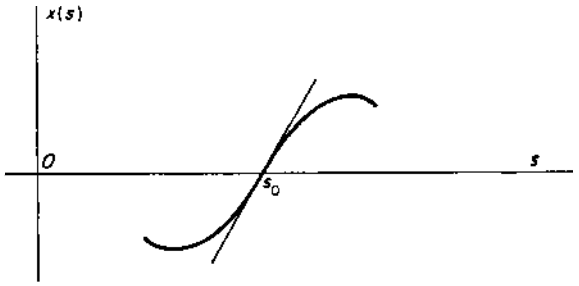


FIG. 5.4. Graph of $x(s)$ near $s_0$ when $r(s) > 0$.

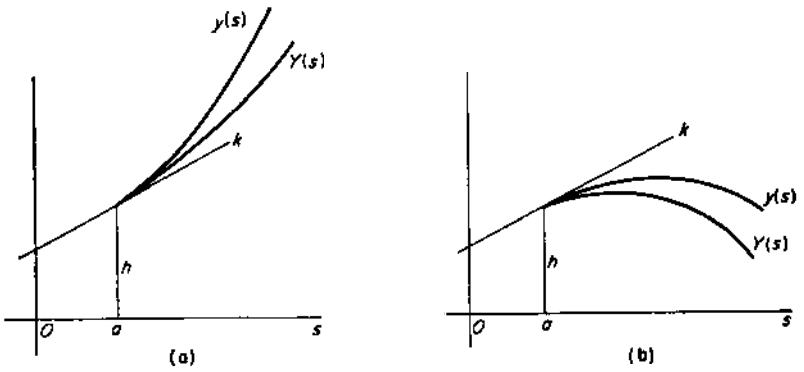


FIG. 5.5. Graphs of $y(s)$ and $Y(s)$.
(a) for $r(s) < R(s) < 0$,
(b) for $0 < r(s) < R(s)$.

for all $s$ in $(a, b)$. We multiply (3) by $Y$ and (4) by $-y$ and add to get

$$y''Y-yY''+(r-R)yY = 0.$$

Integrating from $a$ to $s$ we have

$$[y'Y-yY']_a^s+\int_a^s (r-R)yY\,ds = 0,$$

and since $y'(a)Y(a) - y(a)Y'(a) = 0$ we get

$$y'(s)Y(s) - y(s)Y'(s) = \int_a^s (R-r)yY\,\mathrm{d}s. \tag{5.5.5}$$

But in (5) the left side is

$$Y^2 \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{y}{Y}\right),$$

and the right side is positive as long as $y$ and $Y$ are positive. Hence, as long as $y$ and $Y$ are both positive on the right of $a$, $y/Y$ is increasing; but it is unity at $s = a$. Hence on the right of $a$, as long as $y$ and $Y$ are positive, $y$ is greater than $Y$ and the ratio $y/Y$ is increasing. Thus as $s$ increases from $a$, $Y$ falls away below $y$. This is true whether $r$ and $R$ are positive or negative, provided $R > r$. Figure 5.5 shows two possible relations of the graphs.

If $R(s) > r(s) > 0$, both graphs are concave down (Fig. 5.5b), and

$$\frac{y'}{Y'} > \frac{y}{Y} > 1$$

in the region on the right of $s = a$ where both $y$ and $Y$ are positive. As the graphs swerve towards the axis the slope of $y$ is greater than that of $Y$, and where both slopes are negative $Y$ is steeper. $Y$ therefore hits the axis before $y$, i.e. the next root of $Y$ lies nearer to $s = a$ than does that of $y$. Thus increase of $r(s)$ increases the rapidity of oscillation of $y(s)$.

Now suppose that within $(a, b)$

$$n^2 < r(s) < N^2, \tag{5.5.6}$$

where $n$ and $N$ are constants. Then comparing the solution $y$ of

$$x'' + rx = 0$$

with the solutions of

$$x'' + n^2x = 0, \tag{5.5.7}$$

$$x'' + N^2x = 0, \tag{5.5.8}$$

which have angular frequencies $n$ and $N$ respectively, we see that the angular frequency of oscillations of $y(s)$ lies between $n$ and $N$.

*Eigenvalues and eigenfunctions*

A vibrating string, fixed at the two ends, has a set of *normal modes* of vibration. The periods of the fundamental, first harmonic, second

harmonic, and so on, form a sequence of discrete numbers that characterize the intrinsic properties of the system. Many other systems with two-point boundary conditions show the same phenomenon : they contain a parameter, often related to frequency of oscillations, that can take only certain values if the boundary conditions are to be satisfied : these values are the *eigenvalues* (characteristic values). We will analyse a typical eigenvalue problem that arises naturally in Sturm–Liouville theory.

Suppose we want to solve the equation

$$x'' + \lambda r(s)x = f(s) \tag{5.5.9}$$

with the boundary conditions $x(a) = 0$, $x(b) = 0$. Here $\lambda$ is a parameter whose value is at our disposal. We can solve the inhomogeneous equation (9) if we know two independent solutions of the homogeneous equation

$$x'' + \lambda r(s)x = 0, \tag{5.5.10}$$

using the method of variation of parameters as in (5.3.15). We therefore discuss solutions of (10). Let $X_1(s, \lambda)$, $X_2(s, \lambda)$ be any two independent solutions of (10) : the parameter $\lambda$ has been inserted as a reminder that the solutions depend upon it. Then a general solution of (10) is

$$X(s, \lambda) = d_1X_1(s, \lambda) + d_2X_2(s, \lambda), \tag{5.5.11}$$

where $d_1$ and $d_2$ are constants. The boundary conditions demand

$$\begin{aligned} d_1X_1(a, \lambda) + d_2X_2(a, \lambda) &= 0, \\ d_1X_1(b, \lambda) + d_2X_2(b, \lambda) &= 0, \end{aligned} \tag{5.5.12}$$

and in general these have no solution for $d_1, d_2$, other than $d_1 = 0$, $d_2 = 0$, which annihilate $X(s, \lambda)$ in (11). There will be a non-zero solution only if the two equations (12) are *consistent*, that is, if

$$\begin{vmatrix} X_1(a, \lambda) & X_2(a, \lambda) \\ X_1(b, \lambda) & X_2(b, \lambda) \end{vmatrix} = 0. \tag{5.5.13}$$

(13) may be regarded as an equation for $\lambda$. The determinant will vanish for a set (which will be shown to be infinite) of discrete values $\lambda_1, \lambda_2, \ldots$ of $\lambda$. These roots, for which the differential equation and the boundary conditions have a non-zero solution, are the *eigenvalues*. If $\lambda_n$ is the $n$th eigenvalue, we can solve either of the equations (12) (with $\lambda_n$ inserted for $\lambda$) to get the corresponding ratio $d_1 : d_2$. $X(s, \lambda_n)$ is then given by

(11) except for an arbitrary constant multiplier. This $X(s, \lambda_n)$ is the *eigenfunction* corresponding to the eigenvalue $\lambda_n$.

*Example.* Let us take $r(s) = 1$ in (10), so that we have a familiar equation whose linearly independent solutions may be taken to be

$$X_1 = \sin \lambda^{\frac{1}{2}} s, \quad X_2 = \cos \lambda^{\frac{1}{2}} s$$

The consistency condition (13) becomes

$$\begin{vmatrix} \sin \lambda^{\frac{1}{2}} a & \cos \lambda^{\frac{1}{2}} a \\ \sin \lambda^{\frac{1}{2}} b & \cos \lambda^{\frac{1}{2}} b \end{vmatrix} = 0,$$

$$\text{i.e.} \quad \sin \lambda^{\frac{1}{2}} (b-a) = 0.$$

Thus consistency is ensured by taking $\lambda^{\frac{1}{2}}(b-a) = n\pi$, where $n$ is an integer. This has solutions

$$\lambda_n = \frac{n^2\pi^2}{(b-a)^2}, \quad n = 0, 1, 2, \ldots$$

$\lambda_0 = 0$ gives the trivial solution that is identically zero. Corresponding to a certain $\lambda_n$ we can determine the ratio $d_1 : d_2$ from either of the (consistent) boundary conditions, say

$$d_1 \sin \lambda_n^{\frac{1}{2}} a + d_2 \cos \lambda_n^{\frac{1}{2}} a = 0,$$

and the $n$th eigenfunction is

$$d_1 \left( \sin \lambda_n^{\frac{1}{2}} s - \frac{\sin \lambda_n^{\frac{1}{2}} a}{\cos \lambda_n^{\frac{1}{2}} a} \cos \lambda_n^{\frac{1}{2}} s \right) = \frac{d_1}{\cos \dfrac{n\pi a}{b-a}} \sin n\pi \frac{s-a}{b-a}.$$

We return now to the equation (10) with its boundary conditions $x(a) = 0$, $x(b) = 0$. Let us assume $r(s) > 0$ in $(a, b)$. Then if $\lambda < 0$ there cannot be any zero to the right of $s = a$ (compare Fig. 5.3a). We consider $\lambda \geqq 0$ and enquire whether from all the solutions that satisfy $X(a, \lambda) = 0$ we can select some that satisfy also $X(b, \lambda) = 0$. We concentrate attention on solutions $X_a(s, \lambda)$ with a zero at $s = a$ and an arbitrarily fixed slope $k$ there (see Fig. 5.6); we examine the effect of increasing $\lambda$. We have shown that since $r(s) > 0$ an increase in $\lambda$ will have the effect of increasing the frequency of oscillation of solutions and crowding the zeros. It can be proved (see Burkill,[2] Ch. 3) that any particular zero of $X_a(s, \lambda)$ (greater than $a$) moves *continuously* to the left along the $s$-axis as $\lambda$ is increased continuously. When $\lambda = 0$, $X_a(s, \lambda)$, being a solution of $x'' = 0$, is represented by a straight line of

slope $k$ and has no zero in $s > a$, but as soon as $\lambda$ becomes positive the solution becomes oscillating. For small $\lambda$ the zeros will be widely spaced, and at first the next zero $Z_1$ above $s = a$ will lie in the region $s > b$ (Fig. 5.6). As $\lambda$ is increased continuously this zero moves in towards $s = a$, and at some value $\lambda_1$ of $\lambda$ it coincides with $s = b$. The corresponding solution $X_a(s, \lambda_1)$ satisfies the differential equation and *both* boundary conditions. Thus $\lambda_1$ is the first eigenvalue and $X_a(s, \lambda_1)$ the corresponding eigenfunction. $X_a(s, \lambda_1)$ has no zero between $a$ and $b$; it is the fundamental mode.

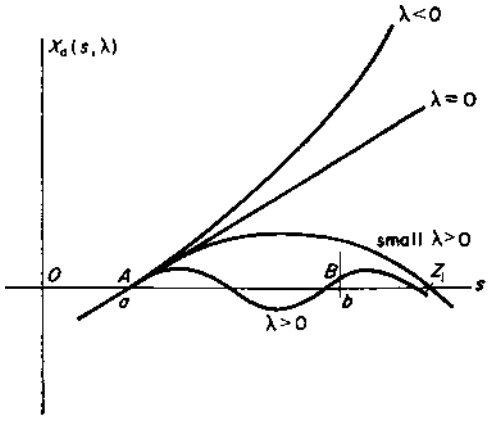


FIG. 5.6. Solutions for various $\lambda$.

Let $\lambda$ increase further. Then at some $\lambda = \lambda_2$ a second zero reaches $B$. $\lambda_2$ is the second eigenvalue, and the corresponding eigenfunction $X_a(s, \lambda_2)$ has one zero between $A$ and $B$. It is the second mode. So we proceed, increasing $\lambda$ to successive eigenvalues and obtaining successive eigenfunctions, each having one more zero in $(a, b)$ than its predecessor. If $\lambda$ can tend to infinity there will be an infinite set of discrete eigenvalues and correspondingly an infinite set of eigenfunctions, each oscillating more rapidly than its predecessor.

Now suppose that $\lambda_m$ and $\lambda_n$ are two *different* eigenvalues, and that $X_m$ and $X_n$ are the corresponding eigenfunctions. Then

$$X_m'' + \lambda_m r X_m = 0, \tag{5.5.14}$$

$$X_n'' + \lambda_n r X_n = 0, \tag{5.5.15}$$

and

$$X_m(a) = 0, \qquad X_m(b) = 0, \tag{5.5.16}$$

$$X_n(a) = 0, \qquad X_n(b) = 0. \tag{5.5.17}$$

Multiply (14) by $X_n$ and (15) by $-X_m$ and add. Then

$$X_m''X_n - X_mX_n'' + (\lambda_m - \lambda_n)rX_mX_n = 0.$$

Integrating from $s = a$ to $s = b$ we obtain

$$[X_m'X_n - X_mX_n']_a^b + (\lambda_m - \lambda_n)\int_a^b rX_mX_n ds = 0. \tag{5.5.18}$$

The expression within square brackets vanishes at both limits, by (16) and (17), and we have

$$\int_a^b rX_mX_n ds = 0 \tag{5.5.19}$$

as long as $m \neq n$. We must not conclude that

$$\int_a^b rX_m^2 ds$$

vanishes: in fact it equals a positive constant, since $r$ and $X_m^2$ are positive within $(a, b)$. If we denote this constant by $M_m^2$, we see that since the eigenfunctions have thus far been arbitrary to the extent of a constant multiplier we may divide each $X_m$ by the corresponding $M_m$ and obtain eigenfunctions which have the property

$$\int_a^b rX_m^2 ds = 1. \tag{5.5.20}$$

The eigenfunctions are then said to be *normalized.*

The property (19) is described by saying that different eigenfunctions are *orthogonal.* [In three-dimensional coordinate geometry two straight lines with direction cosines $(f_1, f_2, f_3)$, $(g_1, g_2, g_3)$ are orthogonal (at right angles) if

$$\sum_{i=1}^{3} f_i g_i = 0.$$

By analogy, extending the concept, we say that the two vectors $(f_1, f_2, \ldots f_n)$, $(g_1, g_2, \ldots g_n)$ in $n$-dimensional space are orthogonal if

$$\sum_{i=1}^{n} f_i g_i = 0.$$

From this a further step leads to the definition of orthogonality over the range $(a, b)$ of two functions $f(s)$ and $g(s)$ by

$$\int_a^b f(s)g(s) ds = 0.$$

Finally we may introduce a "weighting function" or *kernel* $r(s)$ and define orthogonality with respect to the kernel $r(s)$ and the range $(a, b)$ by

$$\int_a^b f(s)g(s)r(s)\mathrm{d}s = 0.]$$

Referring back to our theory of simultaneous differential equations of the second order with constant coefficients (Section 3.4) we see that the eigenvalues of the present section are a generalization of the angular frequencies of the normal modes, and the eigenfunctions correspond to the normal modes themselves. The orthogonality of the eigenfunctions corresponds to the orthogonality of the normal modes, being the mathematical expression of their mutual independence. Further, just as any motion of a system with normal modes can be analysed as a sum of constituent motions in the modes, so any solution of our Sturm–Liouville equation with the boundary conditions can be expressed as a sum of eigenfunctions, as follows.

Let $X_m$ be normalized eigenfunctions, and let us try to approximate to a given function $F(s)$, which vanishes at $A$ and $B$, by a sum of multiples of eigenfunctions, namely

$$\sum_{k=1}^{n} A_k X_k(s) \tag{5.5.21}$$

where $A_k$ are constants. We need some criterion by which to judge the accuracy with which (21) fits $F(s)$ in $(a, b)$ : we will use Gauss' criterion of "least sum of squares of deviations", which here becomes the demand that $A_1, A_2, \ldots, A_n$ are to be chosen so as to make the "error"

$$E = \int_a^b r(s)\{F(s) - \sum_{k=1}^{n} A_k X_k(s)\}^2 \mathrm{d}s \tag{5.5.22}$$

as small as possible. $E$ is a quadratic function of the $A_k$ and cannot be negative. It will be least when

$$\frac{\partial E}{\partial A_k} = 0, \quad k = 1, 2, \ldots, n.$$

This gives

$$\int_a^b r\{F - \sum_{1}^{n} A_l X_l\} X_k \mathrm{d}s = 0, \quad k = 1, 2, \ldots, n. \tag{5.5.23}$$

But, by the orthogonality and normalization of the eigenfunctions,

$$\int_a^b rX_l X_k \mathrm{d}s = \delta_{lk}, \tag{5.5.24}$$

and substituting this into (23) we have

$$\int_a^b rFX_k \mathrm{d}s = \sum_1^n A_l \delta_{lk} = A_k,$$

giving $A_k$ uniquely.

Using these values of $A_k$, we evaluate $E$ as

$$\begin{aligned} E &= \int_a^b r(F^2 - 2F\sum_1^n A_l X_l + \sum_1^n \sum_1^n A_l X_l A_k X_k)\mathrm{d}s \\ &= \int_a^b rF^2 \mathrm{d}s - 2\sum_1^n A_l A_l + \sum_1^n \sum_1^n A_l A_k \delta_{lk} \\ &= \int_a^b rF^2 \mathrm{d}s - \sum_{k=1}^n A_k{}^2. \end{aligned} \tag{5.5.25}$$

Each term added to (21) therefore reduces $E$, and owing to the orthogonality of the eigenfunctions the addition of a new term $A_{n+1}X_{n+1}$ does not affect the constants $A_1, A_2, \ldots, A_n$ already determined. If $E \to 0$ as $n \to \infty$ for any well behaved $F(s)$ the set of eigenfunctions is capable of representing any such function to any required accuracy of approximation, and the set is called *complete*.

Examples of analysis in eigenfunctions are given in Exercises 5.

*Interlacing of zeros of a pair of independent solutions*

We now leave the eigenvalue problem and return to equation (1): we show that any two linearly independent solutions have interlacing roots. Let two such solutions be $X_1$ and $X_2$. Then the Wronskian does not vanish, i.e.

$$X_1'X_2 - X_1X_2' \neq 0.$$

But since

$$X_1'' + rX_1 = 0,$$

$$X_2'' + rX_2 = 0,$$

it follows that

$$X_1''X_2 - X_1X_2'' = 0. \tag{5.5.26}$$

L

Let $s = a$ and $s = b$ be consecutive zeros of $X_1$, so that $X_1(a) = 0$, $X_1(b) = 0$, and $X_1'(a)$, $X_1'(b)$ have opposite signs. Integrate (26) from $s = a$ to $s = b$, to get

$$[X_1'X_2 - X_1X_2']_a^b = 0.$$

The term $X_1X_2'$ vanishes at both limits, so that

$$X_1'(b)X_2(b) - X_1'(a)X_2(a) = 0,$$

or

$$\frac{X_2(b)}{X_2(a)} = \frac{X_1'(a)}{X_1'(b)} < 0.$$

Hence $X_2$ has opposite signs at $s = a$ and $s = b$ and must have at least one root between. Similarly we show that $X_1$ has at least one root between successive roots of $X_2$. Thus the zeros of $X_1$ and $X_2$ must in fact occur alternately: the graphs interlace. This is shown in the graphs of solutions of the Airy and Bessel equations in Figs. 5.1 and 5.2.

An introduction to Sturm–Liouville theory is given by Burkill[2] (Ch. 3) and Ince[1] (Ch. 10).

## 5.6. Particular second order equations

In this section we list second order linear differential equations which arise from problems of mathematical physics, giving references in which further information may be found. The equations are

(a) $x'' - sx = 0$, (Airy)

(b) $s^2x'' + sx' + (s^2 - \nu^2)x = 0$, (Bessel)

(c) $(1 - s^2)x'' - 2sx' + \nu(\nu + 1) = 0$, (Legendre)

(d) $x'' - 2sx' + 2\nu x = 0$, (Hermite)

(e) $sx'' + (1 - s)x' + \nu x = 0$, (Laguerre)

(f) $(1 - s^2)x'' - sx' + \nu^2 x = 0$, (Tschebycheff)

(g) $4s(1 - s)x'' + 2(1 - 2s)x' + (a - 16b + 32bs)x = 0$. (Mathieu)

Here $a$, $b$ and $\nu$ are constants, and $\nu$ is frequently an integer.

*(a) Airy's equation*

The equation

$$x'' - sx = 0 \tag{5.6.1}$$

has no singularity at any finite point. We can find two independent solutions at the origin; these were shown in Section 5.4 (p. 130) to be

$$\left.\begin{aligned} X_1(s) &= 1+\frac{1}{2.3}s^3+\frac{1}{2.3.5.6}s^6+\ldots \\ X_2(s) &= s+\frac{1}{3.4}s^4+\frac{1}{3.4.6.7}s^7+\ldots \end{aligned}\right\} \qquad (5.6.2)$$

Any two linearly independent linear combinations of $X_1$ and $X_2$ could be taken as the fundamental pair of solutions. We define the *Airy functions* as

$$\left.\begin{aligned} Ai(s) &= \alpha X_1-\beta X_2 \\ Bi(s) &= \sqrt{3}(\alpha X_1+\beta X_2) \end{aligned}\right\} \quad \text{where} \quad \alpha=\frac{1}{3^{\frac{1}{3}}\Gamma(\frac{2}{3})},$$

$$\beta=\frac{1}{3^{\frac{1}{3}}\Gamma(\frac{1}{3})}, \qquad (5.6.3)$$

$\Gamma(\nu)$ denoting the *Gamma function*

$$\int_0^\infty \mathrm{e}^{-t}t^{\nu-1}\mathrm{d}t,$$

which reduces to $(\nu-1)!$ when $\nu$ is a positive integer (see Margenau and Murphy, *Mathematics of Physics and Chemistry*,[4] Section 3.2). The particular combination $\alpha X_1-\beta X_2$ was selected as the one which remains finite as $s\to\infty$. In fact it can be shown that when $s$ is large and positive

$$Ai(s)\sim\frac{1}{2\sqrt{\pi}}s^{-\frac{1}{4}}\exp(-\tfrac{2}{3}s^{\frac{3}{2}}),$$

$$Bi(s)\sim\frac{1}{\sqrt{\pi}}s^{\frac{1}{4}}\exp(+\tfrac{2}{3}s^{\frac{3}{2}}), \qquad (5.6.4)$$

where we use the notation $f(s)\sim\phi(s)$ to indicate that $f(s)/\phi(s)\to 1$ as $s\to\infty$. Thus any other combination of $X_1$ and $X_2$ behaves like $Bi(s)$ for large positive $s$. $Ai(s)$ and $Bi(s)$ are shown in Fig. 5.7.

The choice of coefficients in $Bi(s)=\sqrt{3}(\alpha X_1+\beta X_2)$ was made to ensure that $Ai(s)$ and $Bi(s)$, which interlace for $s<0$ according to Sturm–Liouville theory, shall differ in phase by $\frac{1}{2}\pi$ when $s\to-\infty$.

For $s$ large and negative, we write $s = -\sigma$, $\sigma$ being large and positive, and

$$Ai(-\sigma) \sim \frac{1}{\sqrt{\pi}} \sigma^{-\frac{1}{4}} \sin\left(\tfrac{2}{3}\sigma^{\frac{3}{2}} + \tfrac{1}{4}\pi\right),$$

$$Bi(-\sigma) \sim \frac{1}{\sqrt{\pi}} \sigma^{-\frac{1}{4}} \cos\left(\tfrac{2}{3}\sigma^{\frac{3}{2}} + \tfrac{1}{4}\pi\right). \tag{5.6.5}$$

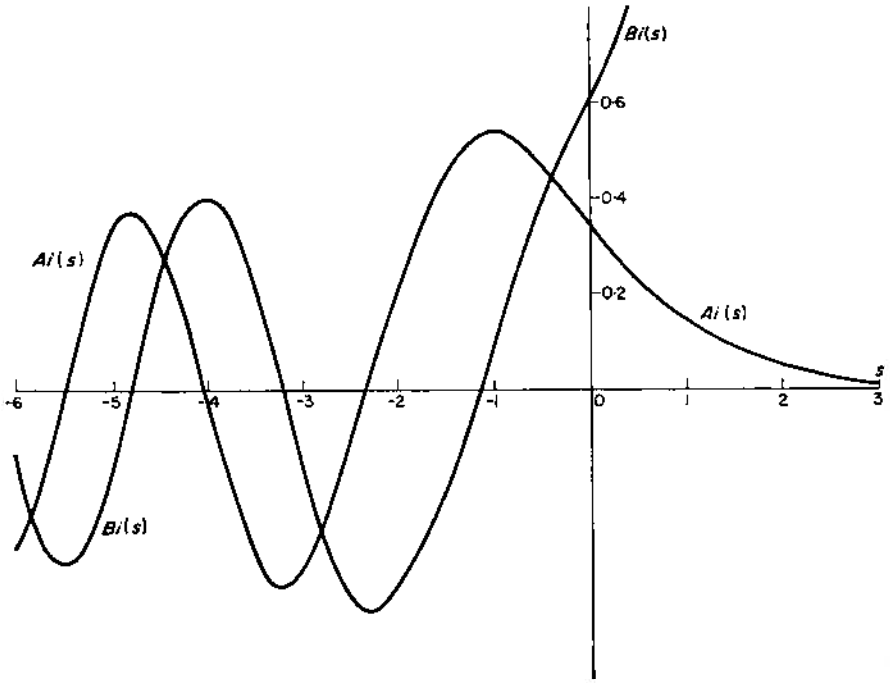


FIG. 5.7. $Ai(s)$ and $Bi(s)$, standard solutions of Airy's equation.

These choices determine the *ratio* of coefficients in $Ai$ and $Bi$: the actual values of $\alpha$ and $\beta$ were chosen to make

$$Ai(s) = \frac{1}{\pi} \int_0^\infty \cos\left(su + \tfrac{1}{3}u^3\right) \mathrm{d}u. \tag{5.6.6}$$

This integral, which can easily be shown to be a solution of the differential equation (1), is the form in which Airy first introduced the function in his discussion in 1838 of intensity of illumination in the neighbourhood of a caustic.

An account of Airy functions may be found in

Jeffreys, H. and Jeffreys, B. S., *Methods of Mathematical Physics*,[5] Ch. 17,

and tables of $Ai$ and $Bi$ in

Miller, J. C. P., *British Association Mathematical Tables*.[6]

An important application of Airy functions arises in the discussion of approximate solutions of second order differential equations (see Chapter 6).

(*b*) *Bessel's equation*

The equation

$$s^2x'' + sx' + (s^2 - \nu^2)x = 0 \tag{5.6.7}$$

has a singularity at the origin. If $\nu$ is not an integer, the solutions (found on p.133)

$$\begin{aligned} X_1(s) &= s^{\nu}\left\{1 - \frac{1}{\nu+1}\left(\frac{s}{2}\right)^2 + \frac{1}{1.2.(\nu+1)(\nu+2)}\left(\frac{s}{2}\right)^4 - \ldots\right\}, \\ X_2(s) &= s^{-\nu}\left\{1 - \frac{1}{1-\nu}\left(\frac{s}{2}\right)^2 + \frac{1}{1.2.(1-\nu)(2-\nu)}\left(\frac{s}{2}\right)^4 - \ldots\right\}, \end{aligned} \tag{5.6.8}$$

are linearly independent. Introducing a factor $(2^{\nu}\nu!)^{-1}$, where the factorial notation has been extended to apply to non-integers according to the identity $\nu! \equiv \Gamma(\nu+1)$, we define *Bessel functions of the first kind*

$$\begin{aligned} J_{\nu}(s) &= \sum_{r=0}^{\infty} (-1)^r \frac{(\frac{1}{2}s)^{\nu+2r}}{r!(\nu+r)!}, \\ J_{-\nu}(s) &= \sum_{r=0}^{\infty} (-1)^r \frac{(\frac{1}{2}s)^{-\nu+2r}}{r!(-\nu+r)!}. \end{aligned} \tag{5.6.9}$$

The multiplying factor has been chosen so that when $\nu$ is an integer, which we denote by $n$,

$$J_n(s) = \frac{1}{\pi}\int_0^{\pi} \cos(s\sin\theta - n\theta)\mathrm{d}\theta, \tag{5.6.10}$$

the form originally studied by Bessel in his discussion of the orbital motion of a planet.

It can be shown that, when $s$ is large,

$$J_\nu(s) \sim \left(\frac{2}{\pi s}\right)^{\frac{1}{2}} \cos\left(s - \tfrac{1}{2}\nu\pi - \tfrac{1}{4}\pi\right). \tag{5.6.11}$$

When $\nu$ is an integer $n$, $J_{-n}$ is simply $(-1)^n J_n$, and we must look further for a second independent solution, which we will call $Y_n(s)$. We select that solution which for large positive $s$ differs in phase from $J_\nu$ by $\frac{1}{2}\pi$. This is

$$Y_\nu(s) \sim \left(\frac{2}{\pi s}\right)^{\frac{1}{2}} \sin\left(s - \tfrac{1}{2}\nu\pi - \tfrac{1}{4}\pi\right). \tag{5.6.12}$$

A function for which this asymptotic approximation holds is defined by

$$Y_\nu(s) = \frac{J_\nu(s)\cos\nu\pi - J_{-\nu}(s)}{\sin\nu\pi}. \tag{5.6.13}$$

$Y_n(s)$ is the limit of this function as $\nu$ approaches the integer $n$. This can be shown to be

$$Y_n(s) = \frac{1}{\pi}\left\{\frac{\partial}{\partial n} J_n(s) - (-1)^n \frac{\partial}{\partial n} J_{-n}(s)\right\}. \tag{5.6.14}$$

$Y_n$ is the *Bessel function of the second kind*, introduced by Neumann. The zeros of $J_n$ and $Y_n$ interlace.

For some purposes it is more convenient to use a different pair as our fundamental solutions. We define the *Hankel functions of the first and second kind* by

$$\begin{aligned} Hs_\nu(s) &= J_\nu(s) + iY_\nu(s), \\ Hi_\nu(s) &= J_\nu(s) - iY_\nu(s) \end{aligned} \tag{5.6.15}$$

then from (11) and (12) we see that for large $s$

$$\begin{aligned} Hs_\nu(s) &\sim \left(\frac{2}{\pi s}\right)^{\frac{1}{2}} \exp\{i(s - \tfrac{1}{2}\nu\pi - \tfrac{1}{4}\pi)\}, \\ Hi_\nu(s) &\sim \left(\frac{2}{\pi s}\right)^{\frac{1}{2}} \exp\{-i(s - \tfrac{1}{2}\nu\pi - \tfrac{1}{4}\pi)\}, \end{aligned} \tag{5.6.16}$$

so that Hankel and Bessel functions are related in the same way as exponential and trigonometric functions.

Bessel functions and their applications are dealt with in

Jeffreys and Jeffreys, *Methods of Mathematical Physics*,[5] Chs. 21, 22.

Sneddon, I. N., *Special Functions of Mathematical Physics and Chemistry*,[7] Ch. 4.

Watson, G. N., *Bessel Functions*.[8]

(c) *Legendre's equation*

The equation

$$(1-s^2)x''-2sx'+\nu(\nu+1)x = 0 \tag{5.6.17}$$

has singular points at $s = \pm 1$ but is well-behaved at the origin. We note that the index interval (p. 137) is 2, so we seek a series solution

$$x(s) = s^\gamma\{a_0+a_2s^2+a_4s^4+ \ldots +a_{2r}s^{2r}+ \ldots\}.$$

We get

$$\begin{aligned}(1-s^2)\{&\gamma(\gamma-1)a_0s^{\gamma-2}+(\gamma+2)(\gamma+1)a_2s^\gamma+ \ldots \\ &+(\gamma+2r)(\gamma+2r-1)a_{2r}s^{\gamma+2r-2}+ \ldots\} \\ -2s\{&\gamma a_0s^{\gamma-1}+(\gamma+2)a_2s^{\gamma+1}+ \ldots +(\gamma+2r)a_{2r}s^{\gamma+2r-1}+ \ldots\} \\ +\nu(\nu+1)\{&a_0s^\gamma+a_2s^{\gamma+2}+ \ldots +a_{2r}s^{\gamma+2r}+ \ldots\} = 0.\end{aligned}$$

The indicial equation is

$$\gamma(\gamma-1) = 0, \tag{5.6.18}$$

and the general relation between coefficients is

$$\begin{aligned}(\gamma+2r+2)(\gamma+2r+1)a_{2r+2}&-(\gamma+2r)(\gamma+2r-1)a_{2r} \\ &-2(\gamma+2r)a_{2r}+\nu(\nu+1)a_{2r} = 0,\end{aligned}$$

so that

$$a_{2r+2} = \frac{(\gamma+2r+1)(\gamma+2r)-\nu(\nu+1)}{(\gamma+2r+2)(\gamma+2r+1)}a_{2r}. \tag{5.6.19}$$

Putting $\gamma = 0$ we get

$$a_{2r+2} = \frac{(2r+\nu+1)(2r-\nu)}{(2r+1)(2r+2)}a_{2r}, \tag{5.6.20}$$

so that the corresponding solution is

$$x(s) = a_0\left\{1-\frac{\nu(\nu+1)}{2!}s^2+\frac{\nu(\nu+1)(\nu+3)(\nu-2)}{4!}s^4-\ldots\right\}. \tag{5.6.21}$$

If $\nu$ is an even integer, say $2m$, the coefficients beginning from $a_{2m+2}$ all vanish and the solution becomes a polynomial of degree $2m$. It is convenient then to reverse the order of terms, starting the series with

$a_{2m}s^{2m}$ and obtaining other coefficients, by the inverted form of (20) with $\nu = 2m$,

$$a_{2r-2} = \frac{(2r-1)2r}{(2r-1+2m)(2r-2-2m)}\, a_{2r},$$

to get the polynomial solution

$$x(s) = a_{2m}\left\{s^{2m} - \frac{2m(2m-1)}{2(4m-1)}\, s^{2m-2} + \ldots\right\}. \tag{5.6.22}$$

For the second solution we put $\gamma = 1$, so that (19) becomes

$$a_{2r+2} = \frac{(2r+\nu+2)(2r+1-\nu)}{(2r+2)(2r+3)}\, a_{2r}, \tag{5.6.23}$$

with corresponding solution

$$x(s) = a_0\left\{s - \frac{(\nu-1)(\nu+2)}{3!}\, s^3 + \frac{(\nu-1)(\nu+2)(\nu-3)(\nu+4)}{5!}\, s^5 + \ldots\right\}. \tag{5.6.24}$$

This is an infinite series that converges as long as $|s| < 1$. When $\nu = 2m$ it is the infinite series

$$x(s) = a_0\left\{s - \frac{(2m-1)(2m+2)}{3!}\, s^3 + \ldots\right\} \tag{5.6.25}$$

and this with (22) gives the complete solution for $|s| < 1$ and $\nu = 2m$, an even integer.

If $\nu$ is an odd integer, let $\nu = 2m+1$; then the series (21) does not terminate, but (24) does terminate since $a_{2m+2} = 0$. Then using (23) in the form

$$a_{2r-2} = -\frac{2r(2r+1)}{(2m+2r+1)(2m-2r+2)}\, a_{2r}, \tag{5.6.26}$$

we get the polynomial solution

$$x(s) = a_{2m}\left\{s^{2m+1} - \frac{(2m+1)2m}{2(4m+1)}\, s^{2m-1} + \ldots\right\}. \tag{5.6.27}$$

(27) is essentially of the same form as (22). Choosing the multiplying constant according to accepted usage, we can combine (22) and (27)

in the following series for $P_n(s)$, valid when $n$ is either an even or an odd integer :

$$P_n(s) = \frac{(2n)!}{2^n(n!)^2}\left\{s^n - \frac{n(n-1)}{2(2n-1)}s^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2.4(2n-1)(2n-3)}s^{n-4} - \ldots\right\}. \tag{5.6.28}$$

This is the *Legendre polynomial* of order $n$. A full discussion of the properties of Legendre polynomials and other solutions of Legendre's equation may be found in

MacRobert, T. M., *Spherical Harmonics.*[9]

Sneddon, I. N., *Special Functions of Mathematical Physics and Chemistry*[7] (Ch. 3).

(*d*) *Hermite's equation*

Independent solutions of

$$x'' - 2sx' + 2\nu x = 0, \tag{5.6.29}$$

which has no finite singularity and has index interval 2, are found to be

$$x_1(s) = a_1\left\{1 - \frac{2\nu}{2!}s^2 + \frac{2^2.\nu(\nu-2)}{4!}s^4 - \ldots\right\},$$

$$x_2(s) = a_2\left\{s - \frac{2(\nu-1)}{3!}s^3 + \frac{2^2(\nu-1)(\nu-3)}{5!}s^5 - \ldots\right\},$$

where $a_1$ and $a_2$ are constants at our disposal. If $\nu$ is an even integer the first series terminates, if odd the second. By suitable choice of $a_1$ and $a_2$ the polynomials can be expressed, by the same method as in (*c*) (Exercise 5.11) as

$$H_n(s) = (2s)^n - \frac{n(n-1)}{1!}(2s)^{n-2} + \frac{n(n-1)(n-3)(n-4)}{2!}(2s)^{n-4} - \ldots, \tag{5.6.30}$$

where $\nu$ has been given the integral value $n$. This is the *Hermite polynomial* of order $n$, which occurs in the wave mechanics of the harmonic oscillator.

Schrödinger's equation for the harmonic oscillator can be reduced to

$$\psi'' + (2n + 1 - s^2)\psi = 0, \tag{5.6.31}$$

where $n$ is an integer.

Write

$$\psi = e^{-\frac{1}{2}s^2}\chi,$$

so that

$$\psi' = e^{-\frac{1}{2}s^2}\chi' - se^{-\frac{1}{2}s^2}\chi,$$

$$\psi'' = e^{-\frac{1}{2}s^2}\chi'' - 2se^{-\frac{1}{2}s^2}\chi' + s^2e^{-\frac{1}{2}s^2}\chi - e^{-\frac{1}{2}s^2}\chi,$$

and

$$\psi'' + (2n+1-s^2\psi) = e^{-\frac{1}{2}s^2}\{\chi'' - 2s\chi' + (s^2-1)\chi + (2n+1-s^2)\chi\}$$

$$= e^{-\frac{1}{2}s^2}\{\chi'' - 2s\chi' + 2n\chi\},$$

whence $\chi$ satisfies Hermite's equation, so that

$$e^{-\frac{1}{2}s^2}H_n(s)$$

is a solution of (31).

Hermite polynomials and the associated *Hermite functions* are treated in

Sneddon, I. N., *Special Functions of Mathematical Physics and Chemistry*[7] (Ch. 5).

(e) *Laguerre's equation*

The equation

$$sx'' + (1-s)x' + vx = 0 \tag{5.6.32}$$

has a regular singularity at the origin. Its index interval is 1. In (5.4.16) we found a solution

$$x(s) = a_0\left\{1 - vs + \frac{v(v-1)}{2^2}s^2 - \frac{v(v-1)(v-2)}{(3!)^2}s^3 + \ldots\right\}. \tag{5.6.33}$$

A second solution can be found by Frobenius' method as

$$\left\{\frac{d}{d\gamma}x(s,\gamma)\right\}_{\gamma=0}.$$

But we are interested in the solution that reduces to a polynomial when $v$ is a positive integer $n$. This is (33). If we put $a_0 = (-1)^n n!$ and invert the order of terms we obtain the *Laguerre polynomial*:

$$L_n(s) = (-1)^n\left\{s^n - \frac{n^2}{1!}s^{n-1} + \frac{n^2(n-1)^2}{2!}s^{n-2} - \ldots + (-1)^n n!\right\}. \tag{5.6.34}$$

Taking $\nu = n$ in (32) and differentiating the equation $m$ times with respect to $s$ we have

$$sD^{m+2}x + (m+1-s)D^{m+1}x + (n-m)D^m x = 0$$

where $D$ denotes d/d$s$. Thus the differential equation

$$sx'' + (m+1-s)x' + (n-m)x = 0 \tag{5.6.35}$$

is satisfied by $L_n{}^m(s) = D^m L_n(s)$. This is the *Associated Laguerre polynomial*. The function defined as

$$R_{nl}(s) = \mathrm{e}^{-\frac{1}{2}s} s^l L_{n+l}{}^{2l+1}(s) \tag{5.6.36}$$

is the *Laguerre function*, used in discussion of the wave function of the hydrogen atom.

(*f*) *Tschebycheff's equation*

The equation

$$(1-s^2)x'' - sx' + \nu^2 x = 0 \tag{5.6.37}$$

has singularities at $s = \pm 1$ but is regular at the origin. The index interval is 2. Solutions in the neighbourhood of the origin are (Exercise 5.12)

$$\begin{aligned} x_1(s) &= a_1\left\{1 - \frac{\nu^2}{2!}s^2 + \frac{\nu^2(\nu^2-2^2)}{4!}s^4 - \ldots\right\}, \\ x_2(s) &= a_2\left\{s - \frac{\nu^2-1}{3!}s^3 + \frac{(\nu^2-1)(\nu^2-3^2)}{5!}s^5 - \ldots\right\}. \end{aligned} \tag{5.6.38}$$

According as $\nu$ is an even or odd integer the first or second terminates, becoming the *Tschebycheff polynomial* for that integer.

(*g*) *Mathieu's equation*

Mathieu's equation is met when the wave equation is expressed in elliptic cylindrical coordinates. It is the simplest second order linear equation in which a coefficient is a trigonometric function of the independent variable :

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + (a + 16b\cos 2t)x = 0. \tag{5.6.39}$$

If we substitute $s = \cos^2 t$ this is easily transformed to

$$4s(1-s)x'' + 2(1-2s)x' + (a-16b+32bs)x = 0, \qquad (5.6.40)$$

which has regular singularities at $s = 0$ and $s = 1$. The attempt to solve for $x(s)$ as an infinite series leads to the indicial equation $\gamma(2\gamma-1) = 0$. But here there is no single index interval, and the general relation between coefficients contains three successive coefficients instead of two as in previous problems (see Jeffreys and Jeffreys[5]). Solutions in series are useful in astronomical applications, but in problems where there is an elliptic boundary, e.g. oscillations of an elliptic membrane, we are more interested in periodic solutions, which must be found by other methods.

Mathieu's equation is treated by

Whittaker, E. T. and Watson, G. N., *Modern Analysis*,[10] Ch. 19,
Ince, E. L., *Ordinary Differential Equations*,[1] Ch. 7.

The particular equations listed in this section are all, with the exception of Mathieu's equation, special cases of the *Hypergeometric equation*

$$s(s-1)x'' + \{(1+\alpha+\beta)s - \gamma\}x' + \alpha\beta x = 0, \qquad (5.6.41)$$

where $\alpha$, $\beta$, $\gamma$ are constants (see Exercise 5.10). There is an extensive literature of this equation and its various degenerate forms. Reference may be made to

Jeffreys and Jeffreys, *Methods of Mathematical Physics*,[5] Ch. 23
Whittaker and Watson, *Modern Analysis*,[10] Ch. 14
Ince, E. L., *Ordinary Differential Equations*,[1] Ch. 7
Sneddon, I. N., *Special Functions of Mathematical Physics and Chemistry*,[7] Ch. 2.

## REFERENCES


1. Ince, E. L., *Ordinary Differential Equations*, Longman, 1926.
2. Burkill, J. C., *Theory of Ordinary Differential Equations*, Oliver & Boyd, 1962.
3. Bickley, W. G., *Mathematical Gazette*, 1963, **360**, pp. 120–3.
4. Margenau, H. and Murphy, G. M., *Mathematics of Physics and Chemistry*, Van Nostrand, 1956.
5. Jeffreys, H. and Jeffreys, B. S., *Methods of Mathematical Physics*, Cambridge, 1956.
6. Miller, J. C. P., *British Association Mathematical Tables*, Vol. 2, 1932.
7. Sneddon, I. N., *Special Functions of Mathematical Physics and Chemistry*, Oliver & Boyd, 1956.
8. Watson, G. N., *Bessel Functions*, Cambridge, 1941.
9. MacRobert, T. M., *Spherical Harmonics*, Methuen, 1947.
10. Whittaker, E. T. and Watson, G. N., *Modern Analysis*, Cambridge, 1927.

## Exercises

**5.1.** Solve by the use of an integrating factor:

(a) $\sin s\,\dfrac{dx}{ds}+3\cos s\,x = \operatorname{cosec} s$, with $x\left(\dfrac{\pi}{2}\right) = 1$.

(b) $s(s-1)\dfrac{dx}{ds}+x = 2s^3$, with $x(2) = 0$.

**5.2.** By use of (2.5) find in the form of an integral the solution of

$$\frac{dx}{dt}-\frac{1}{1+t}x = f(t), \text{ with } x(0) = 0,$$

and evaluate this solution when $f(t) = 1-t^2$.

**5.3.** Given that the differential equation

$$x''+\tfrac{1}{2}s^{-1}x = 0$$

has a solution

$$X_1(s) = (s^{\frac{1}{2}}+i)\exp is^{\frac{1}{2}},$$

write down a second independent solution $X_2(s)$. (Hint: consider the complex conjugate.)

Prove that the Wronskian of $X_1$ and $X_2$ is constant.

Hence using (3.15) find the general solution of

$$x''+\tfrac{1}{2}s^{-1}x = f(s).$$

**5.4.** Show that a solution of

$$sx''+x'+sx = 0$$

is

$$x = 1-\frac{s^2}{2^2}+\frac{s^4}{2^2.4^2}-\ldots+\frac{(-1)^n s^{2n}}{2^{2n}(n!)^2}+\ldots$$

This is $J_0(s)$, the zero order Bessel function. Show that $x = vJ_0$ is a second solution if

$$v = \int^s \frac{d\sigma}{\sigma J_0^2(\sigma)},$$

and expand $v$ in the form $a+\ln s+bs^2+O(s^4)$, for small $s$, where $b$ is to be found.

**5.5.** Reduce to standard form the second order differential equation

$$\frac{d^2x}{dt^2}-\frac{1}{t}\frac{dx}{dt}-4t^4(t^2-1)x = 0,$$

and show that it may be reduced further to

$$x''-s(s-1)x = 0.$$

**5.6.** Prove that the Wronskian of two independent solutions of

$$p_0(s)x''+p_1(s)x'+p_2(s)x = 0$$

is

$$A\exp\left\{-\int^s \frac{p_1(\sigma)d\sigma}{p_0(\sigma)}\right\},$$

where $A$ is constant. Hence show that the Wronskian of the standard form is constant.

**5.7.** Given the differential equation $x'' = \delta(s-\xi)$, show that

$$x' = H(s-\xi)+A$$

and

$$x = (s-\xi)H(s-\xi)+As+B,$$

where $\delta(\ )$ and $H(\ )$ are the Dirac and Heaviside functions and $\xi$, $A$ and $B$ are constants.

Apply the boundary conditions $x(0) = 0$ and $x(1) = 0$ to obtain the Green's function

$$\begin{aligned} x &= -s(1-\xi), \quad 0 < s < \xi, \\ &= -(1-s)\xi, \quad \xi < s < 1. \end{aligned}$$

**5.8.** Solve in series (by use of the index interval)

$$2s(1-s^2)x''+x'+12sx = 0,$$

and determine the range of $s$ for which the solution is valid.

**5.9.** Following the method of Section 5.3, construct the solution, valid in $0 < s < l$, of

$$x''-n^2x = \delta(s-\xi), \text{ with } x(0) = 0,\ x(l) = 0.$$

**5.10.** Show that the *Hypergeometric equation*

$$s(1-s)x''+\{\gamma-(\alpha+\beta+1)s\}x'-\alpha\beta x = 0,$$

where $\gamma$ is not an integer, has as a solution the series

$$1+\frac{\alpha.\beta}{1.\gamma}s+\frac{\alpha(\alpha+1)\beta(\beta+1)}{1.2.\gamma(\gamma+1)}s^2+\frac{\alpha(\alpha+1)(\alpha+2)\beta(\beta+1)(\beta+2)}{1.2.3.\gamma(\gamma+1)(\gamma+2)}s^3+\ \ldots$$

What is the region of convergence of this series?

Denoting this solution by $F(\alpha, \beta, \gamma; s)$, show that the general solution is $AF(\alpha,\beta,\gamma; s)+Bs^{1-\gamma}F(\alpha-\gamma+1,\ \beta-\gamma+1,\ 2-\gamma; s)$, where $A$ and $B$ are arbitrary constants.

Show that the Legendre polynomial $P_n(s)$ is given by

$$F\left(n+1,\ -n,\ 1; \frac{1-s}{2}\right),$$

and that

$$F\left(n,\ -n,\ \tfrac{1}{2}, \frac{1-s}{2}\right)$$

is a solution of Tschebycheff's equation.

**5.11.** Prove the formulae given in $(d)$ of Section 5.6 for series solutions of Hermite's equation

$$x''-2sx'+2\nu x = 0.$$

**5.12.** Prove the formulae (6.38) for solutions of Tschebycheff's equation (6.37).

CHAPTER 6

# METHODS OF APPROXIMATE SOLUTION

## 6.1. Introduction

The previous chapters have dealt with methods of finding exact solutions of ordinary differential equations : some solutions are given in closed form and some as infinite series. But there are very few types of equation for which such methods succeed. Infinite series, for instance, can be derived only for linear equations where the recurrence relation contains not more than two coefficients. Unless the differential equation is of simple form it will in general be necessary to resort to a method of approximate solution.

Sometimes the approximate solution will be in analytic form and will give insight into the whole family of solutions ; but other methods of approximation will involve the insertion of numbers at some stage, and will give only particular solutions corresponding to designated values of the parameters. If the solution is numerical we must discover ways of testing the accuracy and refining the approximation; if the approximate solution is analytic we must specify with care the region over which the approximation is good. Numerical methods are not developed in this text : they may be found in the references cited in Section 6.9. The procedures described here are, however, closely related to those of numerical solution.

*Reduction of an nth order differential equation to a system of first order equations*

We first show how the complexity of the general problem may be reduced. The equation of order $n$ :

$$F\{x^{(n)}, x^{(n-1)}, \ldots, x', x, s\} = 0, \tag{6.1.1}$$

with given values of $x, x', \ldots x^{(n-1)}$ at $s = 0$, can in principle be solved for $x^{(n)}$ to give

$$x^{(n)} = f_n\{x^{(n-1)}, \ldots, x', x, s\}, \tag{6.1.2}$$

and if we now introduce new variables $x_0, x_1, \ldots, x_{n-1}$, defined by

$$\begin{aligned} x_0 &= x, \\ x_1 &= x', \\ x_2 &= x'', \\ &\cdots \\ x_{n-1} &= x^{(n-1)}, \end{aligned}$$

then

$$x'_{n-1} = f_n(s, x_0, x_1, \ldots, x_{n-1}), \tag{6.1.3}$$

and this is to be solved with

$$\begin{aligned} x'_{n-2} &= x_{n-1}, \\ &\cdots \\ x_1' &= x_2, \\ x_0' &= x_1, \end{aligned} \tag{6.1.4}$$

and initial conditions

$$x_0(0) = x_{00}, x_1(0) = x_{10}, \ldots, x_{n-1}(0) = x_{n-1,0}. \tag{6.1.5}$$

The equations (3) and (4) are a special case of the more general set of simultaneous first order equations

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}s} x_0 &= f_1(s, x_0, x_1, \ldots, x_{n-1}), \\ \frac{\mathrm{d}}{\mathrm{d}s} x_1 &= f_2(s, x_0, x_1, \ldots, x_{n-1}), \\ &\cdots \\ \frac{\mathrm{d}}{\mathrm{d}s} x_{n-1} &= f_{n-1}(s, x_0, x_1, \ldots, x_{n-1}). \end{aligned} \tag{6.1.6}$$

Since first order equations are much easier to handle than higher order, we will develop methods of solution of (6).

The most simple problem of this type is that of the solution of

$$\frac{\mathrm{d}}{\mathrm{d}s} x = f(s, x), \tag{6.1.7}$$

with $x(0) = x_0$, and we start from this in Section 6.2.

It will soon become clear that there is an enormous increase in difficulty when we pass from linear to non-linear equations. A linear equation of order $n$ has the property that if $\phi_1(s)$ and $\phi_2(s)$ are two independent solutions then $k_1\phi_1 + k_2\phi_2$ is also a solution, for any pair of constants $k_1, k_2$. Consequently solutions may be superposed and suitable combinations found to fit given boundary conditions. This property does not hold for non-linear equations, and solutions cannot be superposed. Again if a solution is sought in the form of a series with coefficients to be determined, the relations between the unknown coefficients for a linear equation are linear. But for a non-linear equation the series and its derivatives may occur in higher powers and the relations between coefficients are extremely complex. The reader may test this assertion by looking for a solution *as a power series in* $s$ of the simple differential equation

$$x'' + x^2 = 0.$$

The difference between marching problems and jury problems was noted in Chapter 1. In general it is much easier to find an approximate solution of a marching problem than a jury problem. This applies especially to numerical solutions, for which the only way to solve a jury problem may be to construct a set of solutions for arbitrarily chosen initial conditions, find which come closest to satisfying the final conditions, and proceed by interpolation and iteration.

When appropriate analytical tools are lacking, graphical methods may give a very useful insight into the nature of the solutions of a differential equation—see Section 6.2. But it must be realized that graphical methods suffer from the grave disadvantages that (a) the accuracy is limited by the draughtsman's technique, (b) the judgement is subjective, and (c) the error is difficult to compute.

Several of the methods listed below have been developed to deal with various special forms of Schrödinger's equation or other second order differential equations arising from physics and chemistry (6.4), (6.5), (6.6), (6.7).

Others have been invented to attack non-linear problems met in orbital theory or electric circuit theory (see Van der Pol's equation, etc., 6.5). Non-linear problems are so difficult that much attention has been paid to those few equations which possess an exact solution, even where that involves an advanced field of analysis such as that of elliptic functions. From a full analysis of such a solution insight can be gained as to the type of singularities to be expected.

In this chapter the methods at our disposal are described under the following heads:

Graphical methods (6.2)
Taylor series expansions (6.3)
Iteration (6.4)
Perturbation (6.5)
Variational formulation (6.7)
Construction of integral equations (6.8)
Approximation by difference equations (6.9)
Use of analogue machines (6.9)

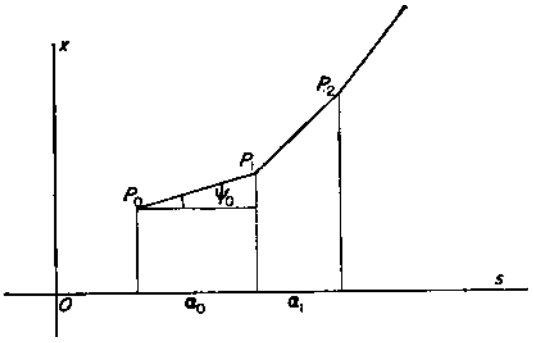


FIG. 6.1.

## 6.2. Graphical methods

Consider the differential equation

$$x' = f(s, x), \tag{6.2.1}$$

with $x(s_0) = x_0$. The most naïve method of solution would be to calculate $x'$ at $s_0$ from (1) and draw a short segment from the point $P_0(s_0, x_0)$, with slope $x'(s_0) = \tan \psi_0$, whose projection on the $s$-axis is of length $\alpha_0$. The end-point will be at $P_1(s_0 + \alpha_0, x_0 + \alpha_0 \tan \psi_0)$. Using (1) again we can calculate a slope at $P_1$ and draw a second segment $P_1P_2$ with this slope.

Proceeding in this way we draw a line (Fig. 6.1) which we regard as a first approximation to the solution. It is a very poor approximation since the slope always refers to the one end-point, so that the line continually veers across the family of true solutions. It is clear that we should attempt to find points such as $P_1'$, $P_2'$ all on the member of the family that passes through $P_0$. This suggests that instead of the

slope at $P_0$ we should take the slope at say $s_0^* = s_0 + \frac{1}{2}\alpha_0$, $x_0^* = x_0 + \frac{1}{2}\alpha_0 \tan\psi_0$ (where $x_0^*$ is calculated from the previous approximation) (see Fig. 6.2). This would bring the new point $P_1^*$ closer to the true solution $P_1'$, and we ought then to seek an iterative procedure to improve the approximation further. This idea is the basis of several numerical methods, and goes back to Euler.

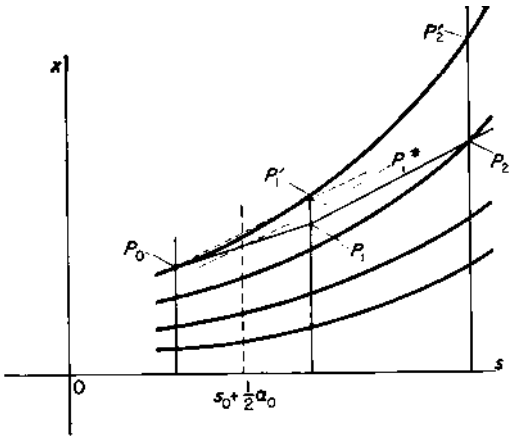


FIG. 6.2.

*Method of isoclines*

If in the equation $x' = f(s, x)$ we give a fixed value to $x'$, say $m_1$, we obtain the curve $\Gamma_1$,

$$f(s, x) = m_1, \tag{6.2.2}$$

on which lie all points where solutions have slope $m_1$. In particular, $f(s, x) = 0$ passes through all points where solutions have tangents parallel to the $s$-axis. Again

$$x'' = \frac{\partial f}{\partial s} + f\frac{\partial f}{\partial x} = 0$$

passes through all points of inflexion of solutions. We can use these ideas as illustrated in the following example :

*Example.* To solve

$$\frac{\mathrm{d}x}{\mathrm{d}s} = sx - 1. \tag{6.2.3}$$

Loci of equal slope are

$$sx = 1 + c;$$

these are the rectangular hyperbolae shown by broken lines for $c = -5, -4, -3, -2, -1, 0, 1, 2, 3$. Since

$$\frac{\mathrm{d}^2x}{\mathrm{d}s^2} = x + s\frac{\mathrm{d}x}{\mathrm{d}s} = x + s(sx - 1),$$

the locus of inflexions is

$$x + s(sx - 1) = 0.$$

i.e.
$$x = \frac{s}{1 + s^2};$$

this is shown by the heavy broken line in Fig. 6.3.

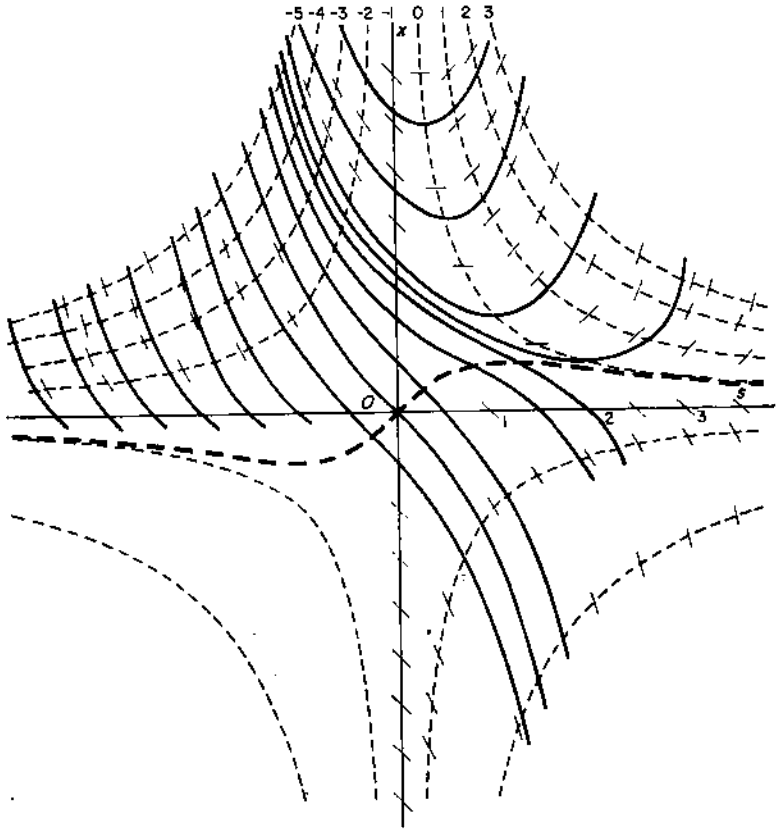


FIG. 6.3. Solutions of $\frac{\mathrm{d}x}{\mathrm{d}s} = sx - 1$ by the method of isoclines (continuous lines).

Part of the family of solutions is shown. It has been constructed so that each curve of the family has the correct slope at each point of intersection with the hyperbolae $c = -5, -4, \ldots, 2, 3$, and has a point of inflexion on the broken line, if it meets that line.

This method may be extended to other types of equation.

If

$$x'' = f(x', x), \tag{6.2.4}$$

$s$ being absent, the substitution

$$x' = v,\ x'' = \frac{\mathrm{d}}{\mathrm{d}s}v = v\frac{\mathrm{d}v}{\mathrm{d}x},$$

gives

$$\frac{\mathrm{d}v}{\mathrm{d}x} = \frac{1}{v}f(v, x).$$

The form of the family of solutions of this equation can be found by the isocline method, giving $v = v(x, c)$. The value of $c$ will be fixed by the initial conditions. Then by integration of

$$\frac{\mathrm{d}x}{\mathrm{d}s} = v$$

we obtain

$$s = \int \frac{\mathrm{d}x}{v(x, c)}, \tag{6.2.5}$$

again fixing the constant of integration by the initial conditions.

*Example.* Integrate

$$2x'' - x'^2 - x = 0, \text{ with } x(0) = 0,\ x'(0) = 1.$$

We put $x' = v$, so that

$$2v\frac{\mathrm{d}v}{\mathrm{d}x} - v^2 - x = 0.$$

If we now write $v^2 = w$ we get the linear equation

$$\frac{\mathrm{d}w}{\mathrm{d}x} - w = x, \qquad w = 1 \text{ when } x = 0.$$

This may be solved by use of the integrating factor $\mathrm{e}^{-x}$ to give

$$w - 1 = \mathrm{e}^x \int_0^x \xi \mathrm{e}^{-\xi}\mathrm{d}\xi$$

$$= \mathrm{e}^x - 1 - x.$$

Hence

$$\frac{dx}{ds} = \pm\sqrt{(e^x - x)}$$

and

$$s = \pm\int_0^x \frac{du}{\sqrt{(e^u - u)}},$$

which must be evaluated numerically. Fig. 6.4 shows how we may derive the form of $x(s)$ graphically.

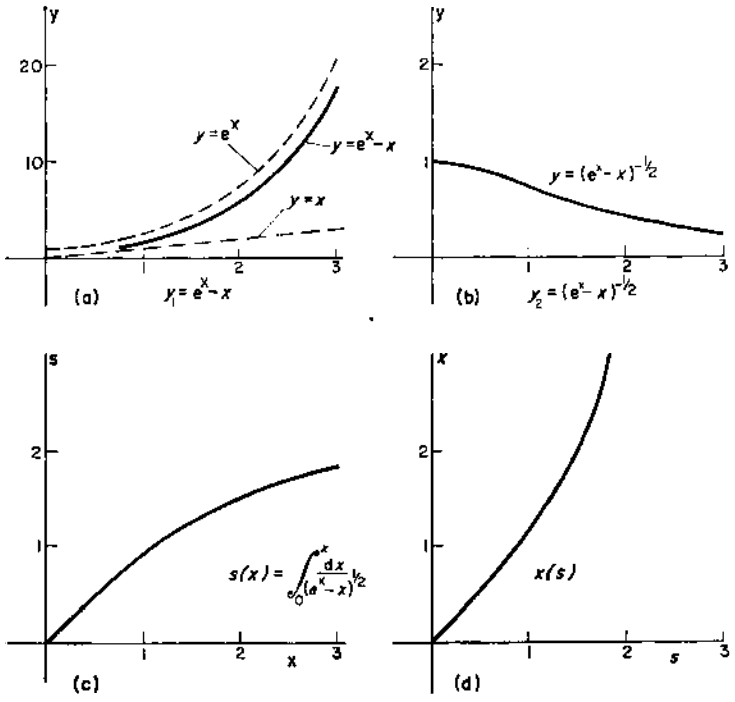


FIG. 6.4. Graphical illustration of the process of finding $x(s)$ from

$$\frac{dx}{ds} = (e^x - x)^{\frac{1}{2}}.$$

(a) $y_1 = e^x - x$ is obtained from values of $e^x$ and $x$.
(b) $y_2 = (e^x - x)^{-\frac{1}{2}}$ is next computed.
(c) $s(x)$ is now obtained by numerical integration.
(d) $x(s)$ is the function inverse to $s(x)$ and is found by exchanging axes.

*Method of curvatures*

Consider the equation, in which $x$ is independent variable,†

$$y'' = f(x, y, y'), \quad y(x_0) = y_0, y'(x_0) = u_0. \tag{6.2.6}$$

Put $y' = \tan\psi$, where $\psi$ is angle between tangent and $x$-axis. Then

$$y'' = \frac{\mathrm{d}}{\mathrm{d}x}\tan\psi = \sec^2\psi\,\frac{\mathrm{d}\psi}{\mathrm{d}s}.\frac{\mathrm{d}s}{\mathrm{d}x}.$$

But $\mathrm{d}s/\mathrm{d}\psi = \rho$, radius of curvature, and $\mathrm{d}x/\mathrm{d}s = \cos\psi$, so that

$$\sec^3\psi\,.\frac{1}{\rho} = f(x, y, \tan\psi),$$

or

$$\rho = \frac{\sec^3\psi}{f(x, y_1 \tan\psi)}. \tag{6.2.7}$$

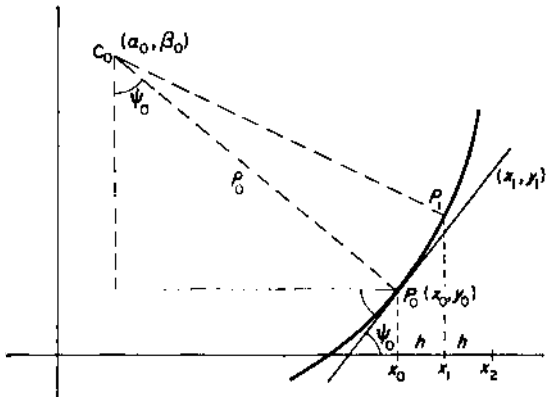


FIG. 6.5. Use of osculating circle in constructing solution of $y'' = f(x, y, y')$.

Since the initial values of $y$ and $y' = \tan\psi$ are known, we can use formula (7) to compute the initial radius of curvature $\rho_0$ at $P_0$, and draw the *osculating circle* there. Note that if $\rho = \mathrm{d}s/\mathrm{d}\psi$ is positive the tangent is turning counterclockwise and the centre of curvature is on the left of the curve as described with increasing $s$. Figure 6.5 shows the centre $C_0$ of the osculating circle, and an arc of the circle passing

†We now use $x$, $y$ as independent and dependent variables so as to obtain familiar formulae for curvature, etc.

through $P_0$. $C_0$ is the point $(\alpha_0, \beta_0)$, where $\alpha_0 = x_0 - \rho_0 \sin\psi_0$, $\beta_0 = y_0 + \rho_0 \cos\psi_0$, and the circle has as its equation

$$(x-\alpha_0)^2 + (y-\beta_0)^2 = \rho_0^2.$$

Taking an interval $h$ of the abscissa $x$ in computing the integral curve, we next find the point $P_1$, $(x_1, y_1)$ on the osculating circle, where $x_1 = x_0 + h$ and

$$y_1 = \beta_0 \pm \{\rho_0^2 - (x_1-\alpha_0)^2\}^{\frac{1}{2}},$$

the sign being chosen according to the direction of turning of the tangent—negative if $y''$ is positive and positive if $y''$ is negative. In effect, we have used the arc of the osculating circle between $x_0$ and $x_1$ in place of the true arc of the integral curve. We now have a second point $P_1$ and a corresponding slope $\tan\psi_1$ given by the direction of the tangent to the osculating circle,

$$\tan\psi_1 = -\frac{x_1-\alpha_0}{y_1-\beta_0}.$$

The routine of calculation can now be repeated to give a third point $P_2$, and so on.

*Example.* To find the solution, between $x = 0$ and $x = 1$, of

$$y'' = x^2 + y, \tag{6.2.8}$$

with initial conditions $y(0) = 1$, $y'(0) = 0$.

Let us take an interval 0·2 of $x$. Using the given initial conditions we have

$$y''(0) = 1, \quad \rho_0 = 1, \quad \alpha_0 = 0, \quad \beta_0 = 2,$$

and so when

$$x = x_1 = 0{\cdot}2, \quad y_1 = 2 - \{1^2 - (0{\cdot}2)^2\}^{\frac{1}{2}} \doteqdot 1{\cdot}020$$

and

$$\tan\psi_1 = -(x_1-0)/(y_1-2) \doteqdot 0{\cdot}204.$$

Computing as explained above we obtain the values of $y_c$ (to three decimal places) shown in the following table. Since this equation is in fact capable of exact solution by the methods of Chapter 3, giving

$$y = 3\cosh x - x^2 - 2,$$

we list for comparison the corresponding accurate values $y_a$. The table gives an idea of the accuracy of the method of curvatures in a

particular example, and Fig. 6.6 shows how the error increases as the curve steepens and the arc-length corresponding to $h$ increases. A discussion of the magnitude of the errors in this method is given by H. T. Davis in *Introduction to Non-linear Differential and Integral Equations*,[1] Ch. 9.

| $x$ | $y_c$ | $y_a$ |
|---|---|---|
| 0·0 | 1·000 | 1·000 |
| 0·2 | 1·020 | 1·020 |
| 0·4 | 1·084 | 1·083 |
| 0·6 | 1·201 | 1·195 |
| 0·8 | 1·384 | 1.371 |
| 1·0 | 1·667 | 1·629 |

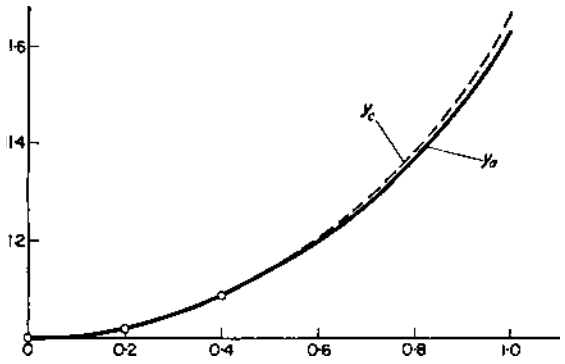


FIG. 6.6. Solutions of $y' = x^2+y$.
$y_a$: accurate solution
$y_c$: computed solution by method of curvatures.

*The phase plane*

Given the pair of equations

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

$$\frac{dy}{dt} = Y(x, y), \tag{6.2.9}$$

with $x = x_0$, $y = y_0$ when $t = 0$, we can seek the solution of

$$\frac{dy}{dx} = \frac{Y(x, y)}{X(x, y)} \tag{6.2.10}$$

which passes through $(x_0, y_0)$ : if this is $y = y^*(x)$, then integrating the first of (9) we get

$$t = \int_{x_0}^{x} \frac{\mathrm{d}x}{X(x, y^*(x))}. \tag{6.2.11}$$

When integrated this gives $t$ as a function of $x$, and in principle we may invert to get $x = x(t)$. Again writing the solution of (10) as $x = x^*(y)$ we get from the second of (9)

$$t = \int_{y_0}^{y} \frac{\mathrm{d}y}{Y(x^*(y), y)}.$$

Integrating and inverting we obtain $y = y(t)$.

The parametric representation $\{x(t), y(t)\}$ refers to the curve which satisfies (10) and passes through $(x_0, y_0)$ where $t = 0$. The $(x, y)$ plane is called the *phase plane* and the curve a *phase trajectory*. If this curve is closed and $x(T) = x_0$, $y(T) = y_0$ then the solution is periodic with period $T$.

If the curve is not closed it may approach a limiting curve as $t \to \infty$ : the solution is then oscillating but not periodic. Detailed examination of the family of phase trajectories will lead to qualitative understanding of the behaviour of solutions. For discussion of the method of phase trajectories see

Davis, H. T., *Introduction to Non-linear Differential and Integral Equations*,[1] Chs. 10, 11.

*Example.* To solve $\dot{x} = x(1-y)$, $\dot{y} = y(x-1)$, with $x(0) = 0$, $y(0) = 0$.

Here the phase trajectories are given by

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{y(x-1)}{x(y-1)}. \tag{6.2.12}$$

There are two points where the slope is not defined, namely (0, 0) and (1, 1). Such points are called *singular points*. We examine the trajectories in the neighbourhoods of these points.

(*a*) *The neighbourhood of* (0, 0)

Near (0, 0)

$$\frac{dy}{dx} = -\frac{y}{x}$$

approximately, i.e. $xy$ = constant, approximately. The trajectories are almost branches of rectangular hyperbolas, and motion started near (0, 0) moves away and does not return. Thus the motion near (0, 0) is unstable.

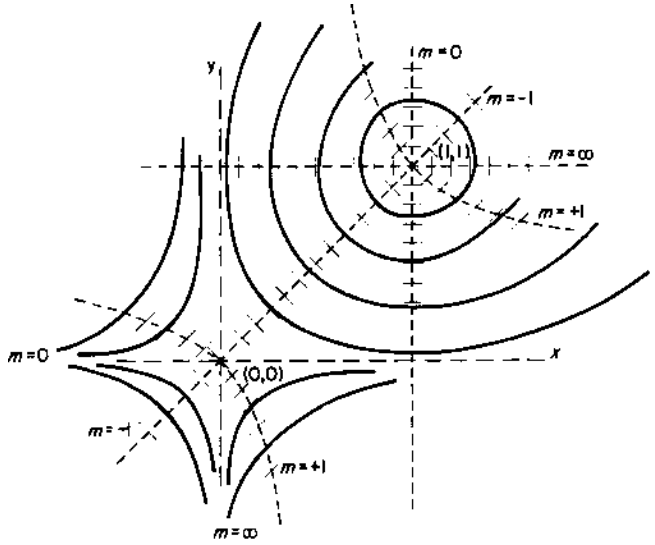


FIG. 6.7. Phase trajectories of $\dot{x} = -x(y-1)$, $\dot{y} = y(x-1)$. $m$ denotes slope used in isocline method. Near (1, 1) motion is stable, near (0, 0) unstable.

(*b*) *The neighbourhood of* (1, 1)

Near (1, 1) we put $x = 1+\xi$, $y = 1+\eta$, where $\xi$, $\eta$ are small. Then

$$\frac{d\eta}{d\xi} = -\frac{\xi}{\eta}$$

approximately, and $\xi^2+\eta^2$ = constant, approximately. Thus the trajectories are approximately circles and the motion is stable if started near (1, 1).

It may be objected that although a motion with a circle as trajectory is not only stable but periodic, we have not proved that the exact motion governed by (12) is so. However, if we draw the trajectories by the method of isoclines (see Fig. 6.7) we note that there is exact

symmetry about the axis $y = x$. Thus any half-loop on one side of the axis has a matching half-loop on the other side, and the loops close exactly. Moreover $\dot{x}$ and $\dot{y}$ are non-zero except at the singular points, so that a loop around (1, 1) is described in a finite time. The motion then repeats itself and so is periodic.

This problem is made easy by the symmetry, and in fact the differential equation of the phase trajectories can be integrated to give

$$xy = Ae^{(x+y)},$$

where $A$ is a constant. In general such simplification does not occur, and the question of periodicity must be investigated by methods beyond the scope of this book. (See Davis, H. T., *loc. cit.*[1] Ch. 11.)

### 6.3. Taylor series expansions

We showed in Section 6.1 that the general problem may be reduced to integration of a system of equations of the type

$$\frac{\mathrm{d}}{\mathrm{d}s} x_i = f_{i+1}(s, x_0, x_1, \ldots, x_{n-1}), \tag{6.3.1}$$

where values of $x_0, x_1, \ldots, x_{n-1}$ are given at $s = 0$. Consider first the simplest equation of this type

$$x' = f(s, x), \quad \text{with } x(0) \text{ given.} \tag{6.3.2}$$

We see that $x'(0)$ is given by (2).

Differentiating (2) we have

$$x'' = f_s(s, x) + f_x(s, x)x',$$

so that we can find $x''(0)$. Proceeding in this way we can find the values of the first $n$ derivatives at $s = 0$ provided the function is such that these exist. Then by Taylor's theorem we can express $x(s)$ as a finite power series near $s = 0$,

$$x(s) = x(0) + x'(0).s + \tfrac{1}{2}x''(0).s^2 + \frac{1}{3!}x'''(0).s^3$$

$$+ \ldots + \frac{1}{n!}x^{(n)}(0)s^n + \frac{1}{(n+1)!}x^{(n+1)}(\theta s)s^{n+1} \tag{6.3.3}$$

where $0 < \theta < 1$ (see Jeffreys and Jeffreys[5] (1.133)).

Under suitable conditions there is some constant $S$ such that successive terms of this series will decrease rapidly as long as $|s| < S$,

and $x(s)$ will be given approximately by a small number of terms for $s$ in this range.

*Example.* Consider again

$$x' = sx - 1, \text{ with } x(0) = 0. \tag{6.3.4}$$

Then

$$\begin{aligned} x'' &= x + sx', \\ x''' &= 2x' + sx'', \\ x^{\text{iv}} &= 3x'' + sx''', \\ &\cdots \\ x^{(n)} &= (n-1)x^{(n-2)} + sx^{(n-1)}. \end{aligned}$$

When $s = 0$,

$$x' = -1,\ x'' = 0,\ x''' = -2,\ x^{\text{iv}} = 0,\ x^{\text{v}} = -8, \ldots$$

and

$$x = -s - \frac{2}{3!}s^3 - \frac{8}{5!}s^5$$

is an approximation valid for small $s$.

Similarly in the neighbourhood of some other point $s_0$

$$x(s) = x(s_0) + x'(s_0)(s - s_0) + \tfrac{1}{2}x''(s_0)(s - s_0)^2 + \ldots, \tag{6.3.5}$$

and under suitable conditions we can use the first few terms of this series to obtain an approximate solution of the differential equation valid near $s = s_0$.

This method breaks down if $x$ or any derivative becomes infinite at $s = s_0$. In fact the Taylor series converges only for $|\, s - s_0 \,|$ less than $|\, s_1 - s_0 \,|$, where $s_1$ is the nearest singularity of the function defined by the differential equation and initial condition.

*Example.* Given

$$x' = -x^2 \quad \text{with} \quad x = -\tfrac{1}{2} \quad \text{when} \quad s = 0. \tag{6.3.6}$$

Here

$$\begin{aligned} x'' &= -2xx' = 2x^3, & x''(0) &= -2!(\tfrac{1}{2})^3, \\ x''' &= 6x^2x' = -6x^4, & x'''(0) &= -3!(\tfrac{1}{2})^4, \\ x^{\text{iv}} &= -24x^3x' = 24x^5, & x^{\text{iv}}(0) &= -4!(\tfrac{1}{2})^5, \\ &\cdots & &\cdots \\ x^{(n)} &= (-)^n n!\, x^{n+1}, & x^{(n)}(0) &= -n!(\tfrac{1}{2})^{n+1}, \end{aligned}$$

so that near $s = 0$

$$x = -\frac{1}{2} - \frac{1}{2^2}.s - 2.\left(\frac{1}{2}\right)^3 \frac{s^2}{2!} - 3!\left(\frac{1}{2}\right)^4 \frac{s^3}{3!} - 4!\left(\frac{1}{2}\right)^5 .\frac{s^4}{4!} - \ldots - n!\left(\frac{1}{2}\right)^{n+1} \frac{s^n}{n!}$$

$$= -\frac{1}{2}\left(1 + \frac{s}{2} + \frac{s^2}{2^2} + \frac{s^3}{2^3} + \frac{s^4}{2^4} + \ldots + \frac{s^n}{2^n} + \ldots\right)$$

$$= -\frac{1}{2}.\frac{1}{1-(s/2)}, \text{ provided } \left|\frac{s}{2}\right| < 1,$$

$$= \frac{1}{s-2},$$

as could have been found by separation of variables. Thus the Taylor series converges to the exact solution within the range $|s| < 2$. We see that the function $1/(s-2)$ has an infinity at $s = 2$, and it is this singularity that bounds the region of application of the Taylor series.

We can modify this method to give a solution valid for large values of $s$.

In

$$x' = f(s, x)$$

we write $s = 1/\xi$ so that

$$x' = \frac{dx}{d\xi}\frac{d\xi}{ds} = \frac{dx}{d\xi}(-\xi^2).$$

Then

$$-\xi^2 \frac{dx}{d\xi} = f\left(\frac{1}{\xi}, x\right),$$

$$\frac{dx}{d\xi} = -\frac{1}{\xi^2} f\left(\frac{1}{\xi}, x\right).$$

We now seek a Taylor approximation for $x(\xi)$ valid near $\xi = 0$ : this can be transformed back into an approximation $x(s)$ valid near $s = \infty$.

*Simultaneous equations*

If we have

$$\left.\begin{aligned} x' &= X(s, x, y), \\ y' &= Y(s, x, y), \end{aligned}\right\} \tag{6.3.7}$$

then

$$\begin{aligned} x'' &= X_s + X_x x' + X_y y', \\ y'' &= Y_s + Y_x x' + Y_y y', \end{aligned}$$

and if at $s = s_0$ $x$ and $y$ are given we can evaluate $x'(s_0)$, $y'(s_0)$, $x''(s_0)$, $y''(s_0)$ and so on. Thus the Taylor series

$$\left.\begin{aligned} x &= x(s_0)+(s-s_0)x'(s_0)+\frac{1}{2!}(s-s_0)^2x''(s_0)+\ldots, \\ y &= y(s_0)+(s-s_0)y'(s_0)+\frac{1}{2!}(s-s_0)^2y''(s_0)+\ldots, \end{aligned}\right\} \tag{6.3.8}$$

can be constructed and an approximate solution obtained, provided all the derivatives of $x$ and $y$ with respect to $s$ exist at $s_0$. Similarly we can extend our method to the general set of simultaneous first order equations and hence to the general differential equation of order $n$. The conditions of application will correspond to those stated in the above examples, and the bounds of the region of validity will be determined by the singularities of the functions sought.

The usefulness of this method depends on the ease with which successive derivatives can be evaluated, and on the rapidity of convergence of the series derived.

The method has been extended by H. T. Davis to become a useful tool for numerical solution, under the name "continuous analytic continuation" (*loc. cit.*[1] Ch. 9). Other numerical methods, for instance the Adams–Bashford method, are based on the Taylor series.

## 6.4. Iteration

The method of iteration is one of approximation by successive application of a given operation. We first demonstrate in an example :

$$x' = sx-1, \text{ with } x(0) = 0. \tag{6.4.1}$$

The equation is equivalent to

$$x = \int_0^s (sx-1)\mathrm{d}s. \tag{6.4.2}$$

1*st approx.* Put $x = 0$ in the integral. (This is the value of $x$ at the end-point $s = 0$.) We get

$$x_1 = -s.$$

2*nd approx.* Put $x = x_1 = -s$ in the integral.

$$x_2 = \int_0^s (-s^2-1)\mathrm{d}s = -\tfrac{1}{3}s^3-s.$$

*3rd approx.* Put $x = x_2 = -s - \frac{1}{3}s^3$ in the integral.

$$x_3 = \int_0^s (-s^2 - \tfrac{1}{3}s^4 - 1)\mathrm{d}s = -s - \tfrac{1}{3}s^3 - \tfrac{1}{15}s^5.$$

This approximation is the same as that obtained by Taylor's series in Section 6.3.

This is a simple example of *Picard's method*, which can be applied very generally. Consider now

$$\frac{\mathrm{d}x}{\mathrm{d}s} = f(s, x), \quad \text{with} \quad x(s_0) = x_0. \tag{6.4.3}$$

The equation is equivalent to

$$x - x_0 = \int_{s_0}^{s} f(s, x)\mathrm{d}s. \tag{6.4.4}$$

No great advance has been made, for the integral contains the unknown function $x(s)$. We have exchanged the differential equation for an integral equation. The latter does however incorporate the initial condition, and this makes it superior to the differential equation for some purposes.

Our first approximation is obtained by using the initial value $x_0$ for $x(s)$ in the integral. Then

$$x_1(s) = x_0 + \int_{s_0}^{s} f(s, x_0)\mathrm{d}s.$$

The second approximation uses $x_1(s)$ within the integral:

$$x_2(s) = x_0 + \int_{s_0}^{s} f(s, x_1)\mathrm{d}s.$$

Proceeding thus, we obtain the $n$th approximation

$$x_n(s) = x_0 + \int_{s_0}^{s} f(s, x_{n-1})\mathrm{d}s.$$

Have we any guarantee that each step improves the approximation? *Picard's theorem* states that convergence of the sequence $\{x_n(s)\}$ to a unique solution of the differential equation with boundary conditions

is ensured provided the function $f(s, x)$ is continuous and bounded in the appropriate region of the $(s, x)$ plane and provided

$$| f(s, x_a) - f(s, x_b) | < K | x_a - x_b |$$

where $(s, x_a)$, $(s, x_b)$ are any two points on the same ordinate in that region, and $K$ is a positive constant.

The method of iteration can be used to find approximate solutions of systems of differential equations. For instance, the equations :

$$\left.\begin{aligned} \frac{dx}{ds} &= F(x, y, s) \qquad \text{with} \quad x(0) = x_0, \\ \frac{dy}{ds} &= G(x, y, s) \qquad\qquad y(0) = y_0, \end{aligned}\right\} \tag{6.4.5}$$

are equivalent to

$$\left.\begin{aligned} x_{r+1} &= x_0 + \int_0^s F(x_r, y_r, s)ds, \\ y_{r+1} &= y_0 + \int_0^s G(x_r, y_r, s)ds. \end{aligned}\right\} \tag{6.4.6}$$

*Example.* To solve

$$\frac{dx}{ds} = x^2 + y^2, \qquad x(0) = 1,$$

$$\frac{dy}{ds} = x + y, \qquad y(0) = 0.$$

Then

$$x_{r+1} = 1 + \int_0^s (x_r^2 + y_r^2)ds,$$

$$y_{r+1} = \int_0^s (x_r + y_r)ds.$$

We obtain

$$\begin{aligned} x_0 &= 1, & y_0 &= 0, \\ x_1 &= 1 + s, & y_1 &= s, \\ x_2 &= 1 + (s + s^2 + \tfrac{2}{3}s^3), & y_2 &= s + s^2, \\ x_3 &= \ldots, & y_3 &= \ldots. \end{aligned}$$

For small values of $s$ the convergence will be fast.

## 6.5. Perturbation

If we know the exact solution to a certain problem, we can use this to find an approximate solution to another problem in which either the differential equation or the boundary conditions have been slightly altered—i.e. perturbed. The method of perturbations is illustrated in the following examples :

*Example* 1. To solve

$$x''+\{n^2+\varepsilon f(s)\}x = 0, \quad x(0) = x_0, x'(0) = u_0, \tag{6.5.1}$$

where $n$ is constant, $f(s)$ is a given function, and $\varepsilon$ is a small constant. We can write this as

$$x''+n^2x = -\varepsilon f(s)x(s).$$

The integral is found as in Section 3.2 (p. 63), but retaining the unknown function $x(s)$. Thus

$$\begin{aligned}(D^2+n^2)x &= -\varepsilon fx,\\ (D+in)(D-in)x &= -\varepsilon fx.\end{aligned}$$

Put

$$(D-in)x = y,$$

so that

$$\begin{aligned}y_0 = y(0) &= u_0-inx_0,\\ (D+in)y &= -\varepsilon fx.\end{aligned}$$

Hence (as in Section 3.2)

$$ye^{ins}-y_0 = -\varepsilon\int_0^s f(\sigma)x(\sigma)e^{in\sigma}d\sigma.$$

Then

$$(D-in)x = e^{-ins}\{u_0-inx_0-\varepsilon\int_{s_0}^s f(\sigma)x(\sigma)e^{in\sigma}d\sigma\}.$$

Similarly

$$(D+in)x = e^{ins}\{u_0+inx_0-\varepsilon\int_{s_0}^s f(\sigma)x(\sigma)e^{-in\sigma}d\sigma\},$$

and by subtraction

$$x(s) = \frac{1}{n}\{u_0\sin ns+nx_0\cos ns - \varepsilon\int_{s_0}^s f(\sigma)x(\sigma)\sin n(s-\sigma)d\sigma\}. \tag{6.5.2}$$

The first approximation $x_1(s)$ is found by substituting $x_0$ for $x(\sigma)$ in the integral. Then $x_1(\sigma)$ is substituted for $x(\sigma)$ in the integral to give $x_2(s)$, and so on. We obtain a power series in $\varepsilon$.

*Example* 2. Second method.

If we omit the term $\varepsilon f(s)$ we find the solution

$$X(s) = x_0 \cos ns + \frac{1}{n} u_0 \sin ns. \tag{6.5.3}$$

Let us seek a solution $x = X + \varepsilon h(s)$ to the modified equation. Substituting this expression into the differential equation we get

$$X'' + \varepsilon h'' + (n^2 + \varepsilon f)(X + \varepsilon h) = 0, \text{ with } h(0) = 0, h'(0) = 0,$$

i.e.

$$\varepsilon h'' + \varepsilon f X + n^2 \varepsilon h + \varepsilon^2 f h = 0,$$

i.e.

$$h'' + n^2 h = -fX - \varepsilon f h.$$

We omit $\varepsilon fh$ and get an approximation

$$h(x) = -\frac{1}{n}\int_0^s f(\sigma)X(\sigma) \sin n(s-\sigma)\mathrm{d}\sigma. \tag{6.5.4}$$

The approximation $X + \varepsilon h$ differs by a term in $\varepsilon^2$ from our previous approximate solution (2). We could now proceed to find a further approximation

$$X + \varepsilon h(s) + \varepsilon^2 k(s).$$

*Example* 3. Bessel's equation of order zero: to find approximate solutions for large $s$. Given Bessel's equation with $\nu = 0$, we can show that it is equivalent to

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(s\frac{\mathrm{d}x}{\mathrm{d}s}\right) + sx = 0. \tag{6.5.5}$$

We have first to transform the equation (5) into such a form that the effect of taking $s$ large can be more easily examined. Let us start by reducing the first term to a simple second derivative.

We write

$$x = s^{-\frac{1}{2}}y,$$

so that

$$s\frac{\mathrm{d}x}{\mathrm{d}s} = s^{\frac{1}{2}}y' - \tfrac{1}{2}s^{-\frac{1}{2}}y$$

and

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(s\frac{\mathrm{d}x}{\mathrm{d}s}\right) = s^{\frac{1}{2}}(y'' + \tfrac{1}{4}s^{-2}y).$$

Thus the differential equation is transformed to

$$y'' + \left(1 + \frac{1}{4s^2}\right)y = 0, \tag{6.5.6}$$

and our first approximation, obtained by neglecting the small term $(1/4s^2)y$, is

$$y = A\cos s + B\sin s, \tag{6.5.7}$$

where $A$ and $B$ are constants.

Second approximation : let $y = \left(A + \frac{A'}{s}\right)\cos s + \left(B + \frac{B'}{s}\right)\sin s$,

where $A'$ and $B'$ are constants, giving

$$y' = -\left(A + \frac{A'}{s}\right)\sin s + \left(B + \frac{B'}{s}\right)\cos s$$
$$-\frac{A'}{s^2}\cos s - \frac{B'}{s^2}\sin s,$$

and

$$y'' = -\left(A + \frac{A'}{s}\right)\cos s - \left(B + \frac{B'}{s}\right)\sin s$$
$$+\frac{2A'}{s^2}\sin s - \frac{2B'}{s^2}\cos s$$
$$+\frac{2A'}{s^3}\cos s + \frac{2B'}{s^3}\sin s.$$

Thus

$$\cos s\left\{-\left(A + \frac{A'}{s}\right) - \frac{2B'}{s^2} + \frac{2A'}{s^3} + \left(1 + \frac{1}{4s^2}\right)\left(A + \frac{A'}{s}\right)\right\}$$
$$+\sin s\left\{-\left(B + \frac{B'}{s}\right) + \frac{2A'}{s^2} + \frac{2B'}{s^3} + \left(1 + \frac{1}{4s^2}\right)\left(B + \frac{B'}{s}\right)\right\} = 0$$

for all $s$ in a certain range. The terms without $s$ and the coefficients of $1/s$ vanish. Equating to zero the coefficients of $1/s^2$ we have

$$-2B' + \tfrac{1}{4}A = 0,$$
$$2A' + \tfrac{1}{4}B = 0,$$

and this equation is satisfied to order $1/s^2$.

Hence the second approximation for $y$ is

$$A\left(\cos s+\frac{1}{8s}\sin s\right)+B\left(\sin s-\frac{1}{8s}\cos s\right), \tag{6.5.8}$$

and $x$ is $s^{-\frac{1}{2}}$ times this.

*Example* 4. *Van der Pol's equation.* Van der Pol's equation describes the self-oscillation of grid potential in a thermionic valve circuit. After simplification the non-linear differential equation to be solved for $x(t)$ is

$$\ddot{x}-\varepsilon(1-x^2)\dot{x}+ax = 0, \tag{6.5.9}$$

where $\varepsilon$ is small and positive and $a$ is positive. $a$ is to be chosen so that the solution is periodic. The term $-\varepsilon(1-x^2)\dot{x}$ arises from the expression of anode current as a cubic function of the grid potential.

We take as initial conditions $x(0) = 0$, $\dot{x}(0) = u_0$, and seek a solution in series of ascending powers of $\varepsilon$. If we put $\varepsilon = 0$ we get

$$\ddot{x}+ax = 0, \tag{6.5.10}$$

with solution

$$x_0(t) = \frac{u_0}{\sqrt{a}}\sin \sqrt{(a)}t. \tag{6.5.11}$$

This is the first approximation to our solution, and we wish to refine it. Since we are interested in sustained self-oscillation we seek a *periodic* solution. We take

$$\left.\begin{aligned} x(t) &= x_0(t)+\varepsilon x_1(t)+\varepsilon^2 x_2(t)+\ldots, \\ a &= a_0+\varepsilon a_1+\varepsilon^2 a_2+\ldots, \end{aligned}\right\} \tag{6.5.12}$$

and substitute these series into the differential equation (9), obtaining

$$\ddot{x}_0+\varepsilon\ddot{x}_1+\varepsilon^2\ddot{x}_2+\ldots-\varepsilon(\dot{x}_0+\varepsilon\dot{x}_1+\varepsilon^2\dot{x}_2+\ldots)\{1-(x_0+\varepsilon x_1+\ldots)^2\}$$
$$+(a_0+\varepsilon a_1+\varepsilon^2 a_2+\ldots)(x_0+\varepsilon x_1+\varepsilon^2 x_2+\ldots) = 0. \tag{6.5.13}$$

This equation will be satisfied, to order $\varepsilon^2$, for all $t$, if we arrange that the coefficients of $\varepsilon^0$, $\varepsilon^1$ and $\varepsilon^2$ vanish. We obtain

$$\ddot{x}_0+a_0x_0 = 0, \tag{6.5.14}$$

$$\ddot{x}_1+a_0x_1 = -a_1x_0+\dot{x}_0(1-x_0^2), \tag{6.5.15}$$

$$\ddot{x}_2+a_0x_2 = -a_2x_0-a_1x_1-2x_0\dot{x}_0x_1+(1-x_0^2)\dot{x}_1, \tag{6.5.16}$$

with the initial conditions

$$x_0(0) = 0, \dot{x}_0(0) = u_0, x_1(0) = 0, \dot{x}_1(0) = 0, x_2(0) = 0, \dot{x}_2(0) = 0. \tag{6.5.17}$$

The solution of (14) with the initial conditions on $x_0(t)$ is $X_0 \sin\psi$, where $\psi = n_0 t$, $n_0 = \sqrt{a_0}$, and $X_0 = u_0/n_0$. This is our approximation (11) again, except that $a_0$ replaces $a$. Substituting it into (15) we have

$$\begin{aligned} \ddot{x}_1 + a_0 x_1 &= -a_1 X_0 \sin\psi + n_0 X_0 \cos\psi(1 - X_0^2 \sin^2\psi) \\ &= -a_1 X_0 \sin\psi + n_0 X_0(1 - \tfrac{1}{4}X_0^2)\cos\psi + \tfrac{1}{4} n_0 X_0^3 \cos 3\psi. \end{aligned} \tag{6.5.18}$$

Now since $\sin\psi$ and $\cos\psi$ are solutions of the homogeneous equation $\ddot{x}_1 + a_0 x_1 = 0$, the contributions to the particular integral of (18) from the terms on the right-hand side of (18) containing $\sin\psi$ and $\cos\psi$ as factors will include terms of type $t \sin a_0^{\frac{1}{2}} t$ and $t \cos a_0^{\frac{1}{2}} t$ which are not periodic. In order to exclude these we equate to zero the coefficients of $\sin\psi$ and $\cos\psi$ on the right-hand side of (18), obtaining $a_1 = 0$ and $X_0 = 2$. Then $n_0 = \frac{1}{2}u_0$ and $a_0 = n_0^2 = \frac{1}{4}u_0^2$. The equation (18) then becomes

$$\ddot{x}_1 + a_0 x_1 = 2n_0 \cos 3\psi, \quad x_1(0) = 0, \dot{x}_1(0) = 0 \tag{6.5.19}$$

with solution

$$x_1 = \frac{1}{4n_0}(\cos\psi - \cos 3\psi). \tag{6.5.20}$$

In order to achieve this periodic solution we have chosen $a = a_0 = \frac{1}{4}u_0^2$.

We next proceed to solve the equation (16) for $x_2(t)$, using $x_0(t)$ and $x_1(t)$ already found. The work is left to the reader (see Exercise 6.6).

## 6.6. The WKBJ method

In Section 5.3 the equation

$$x'' + p_1(t)x' + p_2(t)x = 0 \tag{6.6.1}$$

was reduced to normal form

$$\frac{\mathrm{d}^2 x}{\mathrm{d}\xi^2} + r(\xi)x = 0. \tag{6.6.2}$$

We now show how to find useful approximate solutions of (2). If $r(\xi)$

were a constant, the equation would be the familiar equation of simple harmonic or exponential motion with the solution

$$x = A \cos \sqrt{r}.\xi + B \sin \sqrt{r}.\xi \quad \text{if} \quad r > 0, \tag{6.6.3}$$

$$x = A \exp \sqrt{(-r)}\xi + B \exp \{-\sqrt{(-r)}\xi\} \quad \text{if} \quad r < 0, \tag{6.6.4}$$

where $A$ and $B$ denote arbitrary constants. If $r$ is not constant we may still expect that in any region of $\xi$ for which $r(\xi)$ varies slowly the solutions (3) and (4) will in some sense describe the motion, showing it to be essentially oscillatory where $r(\xi) > 0$ and exponential where $r(\xi) < 0$. But the parameter $\sqrt{r}$ or $\sqrt{(-r)}$ will vary with $\xi$ and so will the multipliers $A$ and $B$ which determine amplitude.

We now try to refine this rough qualitative statement. Let us write

$$x = \exp \eta(\xi). \tag{6.6.5}$$

Then

$$x_\xi = \eta_\xi \exp \eta,$$
$$x_{\xi\xi} = (\eta_{\xi\xi} + \eta_\xi^2) \exp \eta,$$

and our equation becomes

$$\eta_{\xi\xi} + \eta_\xi^2 + r(\xi) = 0. \tag{6.6.6}$$

Writing $\eta_\xi = \zeta$, we obtain

$$\zeta' + \zeta^2 + r(\xi) = 0, \tag{6.6.7}$$

which is of first order but non-linear. We can get a good idea of its solutions, for given $r(\xi)$, by the method of isoclines. But let us proceed by an iterative method, finding successive solutions of

$$\zeta^2 = -r(\xi) - \zeta'.$$

We first take $r(\xi) < 0$. Neglecting $\zeta'$ we obtain $\zeta = \pm \sqrt{(-r)}$. Choosing the positive sign we get for the derivative

$$\zeta' = \frac{-r'}{2\sqrt{(-r)}}.$$

Hence, taking the positive sign again since we are proceeding by iteration, we find the second approximation

$$\begin{aligned} \zeta &= \{-r + \tfrac{1}{2}r'(-r)^{-\frac{1}{2}}\}^{\frac{1}{2}} \\ &= (-r)^{\frac{1}{2}}\{1 + \tfrac{1}{2}r'(-r)^{-1}\}^{\frac{1}{2}} \\ &\doteq (-r)^{\frac{1}{2}}\{1 + \tfrac{1}{4}r'(-r)^{-1}\} \\ &= (-r)^{\frac{1}{2}} + \tfrac{1}{4}\frac{r'}{(-r)}. \end{aligned} \tag{6.6.8}$$

This is $d\eta/d\xi$. Hence an integral of (6) is

$$\eta = -\tfrac{1}{4}\ln(-r) + \int^{\xi} (-r)^{\frac{1}{2}} d\xi,$$

giving

$$x = (-r)^{-\frac{1}{4}} \exp\{\int^{\xi} (-r)^{\frac{1}{2}} d\xi\}. \tag{6.6.9}$$

If we had carried through the same analysis choosing the negative sign for $\zeta$ we would have obtained the integral

$$x = (-r)^{-\frac{1}{4}} \exp\{-\int^{\xi} (-r)^{\frac{1}{2}} d\xi\}, \tag{6.6.10}$$

and hence an approximate general solution of (2) is, for $r(\xi) < 0$,

$$x = (-r)^{-\frac{1}{4}}[A \exp\{\int^{\xi} (-r)^{\frac{1}{2}} d\xi\} + B \exp\{-\int^{\xi} (-r)^{\frac{1}{2}} d\xi\}], \tag{6.6.11}$$

where $A$ and $B$ are arbitrary constants.

If $r(\xi) > 0$ and we follow through a similar argument we get the general solution

$$x = (r)^{-\frac{1}{4}}[C \cos\{\int^{\xi} r^{\frac{1}{2}} d\xi\} + D \sin\{\int^{\xi} r^{\frac{1}{2}} d\xi\}], \tag{6.6.12}$$

where $C$ and $D$ are arbitrary constants. The lower limits of the integrals can obviously be chosen at our convenience.

This method was first explored by Jeffreys, and then used in problems arising from quantum mechanics by Wenzel, Kramers and Brillouin. Hence the string of initials by which the method is known.

*Example.* To find WKBJ approximations for the solution of Bessel's equation of order $\nu$,

$$x'' + \frac{1}{s}x' + \left(1 - \frac{\nu^2}{s^2}\right)x = 0, \tag{6.6.13}$$

which are valid for large $s/\nu$.

We first reduce the differential equation (13) to standard form (Section 5.3) by means of the substitution $x = s^{-\frac{1}{2}}y$. We obtain

$$y'' + \left\{1 - \frac{1}{s^2}(\nu^2 - \tfrac{1}{4})\right\}y = 0. \tag{6.6.14}$$

This is of the form (2), with

$$r(s) = 1 - \nu_1^2/s^2,$$

where $\nu_1^2 = \nu^2 - \frac{1}{4}$. $r(s)$ is positive when $s > \nu_1$.

Here

$$\int^{s} r^{\frac{1}{2}}\mathrm{d}s = \int^{s} \sqrt{(1-\nu_1^2/s^2)}\mathrm{d}s$$
$$= \nu_1\left(\sqrt{(s^2-\nu_1^2)}/\nu_1 - \cos^{-1}\frac{\nu_1}{s}\right)$$

and

$$y = (1-\nu_1^2/s^2)^{-\frac{1}{4}} \begin{matrix}\sin\\ \cos\end{matrix}\left\{\sqrt{(s^2-\nu_1^2)} - \nu_1\cos^{-1}\frac{\nu_1}{s}\right\},$$

so that

$$x = (s^2-\nu_1^2)^{-\frac{1}{4}} \begin{matrix}\sin\\ \cos\end{matrix}\left\{\sqrt{(s^2-\nu_1^2)} - \nu_1\cos^{-1}\frac{\nu_1}{s}\right\}. \qquad (6.6.15)$$

The particular combination of sine and cosine that we choose depends on whether we wish to represent $J_\nu(s)$, $Y_\nu(s)$, $Hs_\nu(s)$ or $Hi_\nu(s)$.

The approximations (11) and (12) can be satisfactory as long as $r(\xi)$ is not near zero, but they are certain to fail in the neighbourhood of a zero of $r(\xi)$. Then we proceed as follows. Let the zero be at $\xi = \xi_0$. Then $r(\xi_0) = 0$ and near $\xi_0$ the Taylor expansion is

$$r(\xi) = (\xi-\xi_0)r'(\xi_0) + \tfrac{1}{2}(\xi-\xi_0)^2 r''(\xi_0) + \ldots = (\xi-\xi_0)r^*(\xi_0).$$

The differential equation is now replaced by the approximation

$$x_{\xi\xi} + (\xi-\xi_0)r'(\xi_0)x = 0.$$

Put $(\xi-\xi_0) = \alpha t$, $\alpha$ being independent of $x$ and $t$, so that

$$\frac{\mathrm{d}}{\mathrm{d}\xi} = \frac{1}{\alpha}\frac{\mathrm{d}}{\mathrm{d}t}.$$

Then the differential equation is

$$\alpha^{-2}x_{tt} + \alpha t r'(\xi_0)x = 0.$$

We choose $\alpha$ so that $\alpha^3 r'(\xi_0) = -1$ and the differential equation reduces to Airy's equation :

$$x_{tt} - tx = 0.$$

Its solutions are $Ai(t)$ and $Bi(t)$ [see Section 5.6(a)] where

$$t = \frac{1}{\alpha}(\xi-\xi_0) = \{-r'(\xi_0)\}^{\frac{1}{3}}(\xi-\xi_0).$$

These Airy functions, which give valid solutions throughout the region connecting $r > 0$ and $r < 0$, provide the means of following a solution through the zero of $r(\xi)$. The problem, which arises in solutions of Schrödinger's equation, is fully discussed in

Jeffreys, H., *Asymptotic Approximations.*[2]
Heading, J., *Phase Integral Methods.*[3]

## 6.7. Variational methods

If we are faced with a new differential equation of the second order, and it deviates too far from known differential equations to make a perturbation useful, we may be able to get an approximate solution by an entirely new method. This method arises naturally from physical problems. If we wish to solve a problem of equilibrium of a rigid body we may either formulate equations of equilibrium or seek the configuration in which the potential energy is a minimum. If we seek the solution of a problem of equilibrium of an elastic body we may either set down differential equations of equilibrium with boundary conditions or find the configuration in which the sum of potential energy in the force field and potential energy of strain is minimized. Again, Hamilton's principle replaces the solution of Lagrangian equations of motion of a dynamical system by minimization of the action integral. Let us look for a method of constructing a function whose stationary values occur when a given differential equation is satisfied.

Here, in order to link up with the usual notation of the calculus of variations, we take $x$ as independent variable and $y$ as dependent variable. Let us deal with the second order differential equation (which is the one which commonly arises, as from the Schrödinger or diffusion equation):

$$\frac{\mathrm{d}}{\mathrm{d}x}\left\{p(x)\frac{\mathrm{d}y}{\mathrm{d}x}\right\} - q(x)y - r(x) = 0 \quad \text{in} \quad (x_0, x_1), \tag{6.7.1}$$

with the jury conditions

$$y(x_0) = y_0, \quad y(x_1) = y_1. \tag{6.7.2}$$

Let $y^*(x)$ be the function, represented by the curve $\Gamma$, that satisfies the differential equation (1) and the boundary conditions (2). Let $y = y^*(x) + \delta y(x)$ be another function that *satisfies the same end conditions* but not the differential equation, and differs from $y^*$ by the

arbitrary *variation* $\delta y(x)$. $\delta y(x)$ may be any small continuous function of $x$ for which

$$\delta y(x_0) = 0, \quad \delta y(x_1) = 0. \tag{6.7.3}$$

Then

$$\int_{x_0}^{x_1} \left[ \frac{\mathrm{d}}{\mathrm{d}x} \left\{ p(x) \frac{\mathrm{d}}{\mathrm{d}x} y^* \right\} - q(x) y^* - r(x) \right] \delta y \mathrm{d}x = 0$$

for arbitrary $\delta y$, since $y^*$ is a solution of (1) throughout the range of integration. Thus, integrating by parts,

$$\left[ \left\{ p(x) \frac{\mathrm{d}}{\mathrm{d}x} y^* \right\} \delta y \right]_{x_0}^{x_1} - \int_{x_0}^{x_1} \left\{ p(x) \frac{\mathrm{d}}{\mathrm{d}x} y^* \right\} \frac{\mathrm{d}}{\mathrm{d}x} \delta y \mathrm{d}x - \int_{x_0}^{x_1} \{ q(x) y^* + r(x) \} \delta y \mathrm{d}x = 0. \tag{6.7.4}$$

The integrated part vanishes, by (3). Next consider $(\mathrm{d}/\mathrm{d}x)\delta y$. This is the slope of the variation $\delta y$ and is

$$\lim_{\Delta x \to 0} \frac{1}{\Delta x} \{ \delta y(x + \Delta x) - \delta y(x) \}$$

The quantity $\delta(\mathrm{d}y/\mathrm{d}x)$, on the other hand, is the variation in the slope, and is

$$\begin{aligned} &\frac{\mathrm{d}y}{\mathrm{d}x} - \frac{\mathrm{d}}{\mathrm{d}x} y^* \\ &= \lim_{\Delta x \to 0} \left[ \frac{1}{\Delta x} \{ y(x + \Delta x) - y(x) \} - \frac{1}{\Delta x} \{ y^*(x + \Delta x) - y^*(x) \} \right] \\ &= \lim_{\Delta x \to 0} \left[ \frac{1}{\Delta x} \{ \delta y(x + \Delta x) - \delta y(x) \} \right]. \end{aligned}$$

These two quantities are therefore equal as long as all these limits exist and are finite. We may write (4) as

$$\int_{x_0}^{x_1} \left[ \left\{ p(x) \frac{\mathrm{d}}{\mathrm{d}x} y^* \right\} \frac{\mathrm{d}}{\mathrm{d}x} \delta y + \{ q(x) y^* + r(x) \} \delta y \right] \mathrm{d}x = 0. \tag{6.7.5}$$

Now

$$\delta(y^2) = (y^* + \delta y)^2 - y^{*2} = 2y^*\delta y + (\delta y)^2$$

and

$$\delta(y'^2) = \left(y^{*\prime} + \frac{\mathrm{d}}{\mathrm{d}x}\delta y\right)^2 - (y^{*\prime})^2 = 2y^{*\prime}\frac{\mathrm{d}}{\mathrm{d}x}\delta y + \left(\frac{\mathrm{d}}{\mathrm{d}x}\delta y\right)^2,$$

so that (5), on being multiplied by 2, becomes

$$\int_{x_0}^{x}\left[p(x)\left\{\delta(y'^2) - \left(\frac{\mathrm{d}}{\mathrm{d}x}\delta y\right)^2\right\} + qx\{\delta(y^2) - (\delta y)^2\} + 2r(x)\delta y\right]\mathrm{d}x = 0,$$

and since neither $x$ nor $p(x)$, $q(x)$, $r(x)$ are varied

$$\delta\int_{x_0}^{x_1}\{p(x)y'^2 + q(x)y^2 + 2r(x)y\}\mathrm{d}x$$

$$= \int_{x_0}^{x_1}\left\{p(x)\left(\frac{\mathrm{d}}{\mathrm{d}x}\delta y\right)^2 + q(x)(\delta y)^2\right\}\mathrm{d}x$$

$$= \int_{x_0}^{x_1}\left\{p(x)\left(\delta\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2 + q(x)(\delta y)^2\right\}\mathrm{d}x. \qquad (6.7.6)$$

Thus if

$$I = \int_{x_0}^{x_1}\{p(x)y'^2 + q(x)y^2 + 2r(x)y\}\mathrm{d}x, \qquad (6.7.7)$$

$\delta I$ is a function containing integrals of $(\delta y)^2$ and $(\delta(\mathrm{d}y/\mathrm{d}x))^2$ : in other words $I$ changes a quantity of second order when a first order change is made from $y^*$. $I(y^*)$ is therefore stationary with regard to such variations. [$f(x)$ is *stationary* at $x_0$ if $f(x_0 + \Delta x)$ differs from $f(x_0)$ by a quantity of second order in $\Delta x$. At such a point the Taylor expansion is $f(x_0) + \frac{1}{2}(\Delta x)^2 f''(x_0) + \ldots$ and $f(x)$ has a horizontal tangent.] This analysis has been built on the assumption that $y^*$ is a solution of the differential equation. The satisfying of the differential equation by $y^*$ is thus a sufficient condition for the integral $I$ to have a stationary

value. The argument is reversible, and we can prove that the satisfying of the differential equation is necessary for a stationary value of $I$. This equivalence means that if we find $y^*$ that makes $I$ stationary, the end points not being varied, this $y^*$ is a solution of the differential equations with its end conditions. We see that if $p(x)$ and $q(x)$ are both positive throughout the range $(x_0, x_1)$ the right-hand side of (6) is positive. Thus the value of $I$ when $y^*$ is inserted is not only stationary but is in fact a minimum.

So far we have worked in terms of $y^*$, which is the unknown solution of the problem. But the choice of $y^*$ to make $I$ stationary is no easier and may be harder than the choice of $y^*$ to satisfy the differential equation. However, the finding of an approximation to $y^*$ is easier. For let $\phi_i$, $i = 1, 2, \ldots, n$, be a set of specific functions satisfying the end conditions (but not the differential equation). We form

$$y = \sum A_i \phi_i, \tag{6.7.8}$$

where $A_i$ are adjustable parameters. Inserting this $y$ into $I$ we obtain an integral in which the parameters $A_i$ occur in second degree, coefficients being integrals of given functions (products of $\phi_i$ and their derivatives). In principle the integrals can be evaluated and we are left with $I$ expressed as a quadratic function of the $A_i$. It is a straightforward matter to find the sets of values $A_{i0}$ of $A_i$ that make $I$ stationary. These are the solutions of

$$\frac{\partial}{\partial A_i} I = 0.$$

There are $n$ simultaneous linear equations in $A_{i0}$. Solving and substituting in (7) we obtain the function which is the best linear combination of the $\phi_i$. This solution may not give a true minimum of $I$, but it gives the smallest $I$ that can be obtained using the set $\phi_i$. In other words $\sum A_{i0}\phi_i$ will not in general be a solution of the differential equation, but it will be as near as we can get using the functions $\phi_i$.

*Example.* Legendre's equation.

$$\frac{\mathrm{d}}{\mathrm{d}x}\{(1-x^2)y'\} + l(l+1)y = 0, \quad y(-1) = 1, y(1) = 1.$$

To illustrate the method, we take $l = 2$ and find the best fitting quadratic

$$y = a + bx + cx^2.$$

This has to satisfy boundary conditions

$$a+b+c = 1,$$

$$a-b+c = 1,$$

so that $b = 0$, $a = 1-c$, and $y = 1-c+cx^2$.

Thus (7) becomes, with $p = 1-x^2$, $q = -6$, $r = 0$,

$$I = \int_{-1}^{1} [(1-x^2)(2cx)^2 - 6\{(1-c)+cx^2\}^2]\mathrm{d}x$$

and we select $c$ so that this is stationary.

$$I = 4c^2(\tfrac{2}{3}-\tfrac{2}{5}) - 6\{(1-2c+c^2).2 + 2c(1-c).\tfrac{2}{3} + c^2.\tfrac{2}{5}\}$$

$$= \tfrac{16}{15}c^2 - 6(2-\tfrac{8}{3}c+\tfrac{16}{15}c^2)$$

$$= -\tfrac{16}{3}c^2 + 16c - 12.$$

Then

$$\frac{\partial I}{\partial c} = -\tfrac{32}{3}c + 16 = 0 \quad \text{if} \quad c = \tfrac{3}{2}, \quad a = -\tfrac{1}{2}.$$

The required function is $\frac{1}{2}(3x^2-1)$. [This is the exact answer, since the required solution is in fact a quadratic function of $x$. For $l = 4$ we can find exactly the quartic in $x$ that satisfies the equation, and so on for even $l$. But in most cases the method will give an approximate solution.]

## 6.8. Integral equations

We have already seen that the equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x, y),$$

with the initial condition $y(x_0) = y_0$, is equivalent to the equation

$$y - y_0 = \int_{x_0}^{x} f(x, y)\mathrm{d}x,$$

where $y$ is an unknown function of $x$. This is an *integral equation*. In some cases the integral equation can be solved exactly; in others i

can be solved approximately by iterative methods, of which Picard's is an example. Again

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = g(x, y)$$

can be written

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + n^2y = n^2\left(y + \frac{1}{n^2}g\right),$$

with corresponding integral equation :

$$y = A\cos nx + B\sin nx + n\int_0^x \left[y(\xi) + \frac{1}{n^2}g\{\xi, y(\xi)\}\right]\sin n(\xi - x)\mathrm{d}\xi.$$

We can seek an approximate solution of this equation by iterative methods, and $A$ and $B$ can be chosen to fit the boundary conditions.

See Levy, H. and Baggott, E. A., *Numerical Solutions of Differential Equations*,[4] Section 17.2.

## 6.9. Other methods

Numerical solution of ordinary differential equations is dealt with in

Jeffreys, H. and Jeffreys, B. S., *Methods of Mathematical Physics*.[5]

Levy, H. and Baggott, E. A., *Numerical Solutions of Differential Equations*.[4]

Noble, B., *Numerical Methods*,[6] Vol. 2.

The usual procedures are developments of the methods of approximate solution of Picard and Euler. A different attack, depending on replacement of the differential equation by a set of difference equations which are then solved by *successive relaxation*, is given by

Southwell, R., *Relaxation Methods*,[7] Sect. 209.

Allen, D. N. de G., *Relaxation Methods in Engineering and Science*,[8] Ch. 4.

The reader should not overlook the possibility of finding an *analogue* of his problem. For instance, a non-linear equation arising in his work may be such that a suitably designed system of electric or electronic circuits will be described by the same equation. By constructing this system he may be able to produce on an oscilloscope a graph of the

solution he is seeking, and then to explore the effect of varying different parameters.

This relationship has been exploited in the design and use of *analogue computers* on which differential equations of considerable complexity can be set up and solutions displayed or digitized and printed out. Information about such computers can be found in specialist journals, such as the publications of the Institute of Electrical and Electronics Engineers. Resort to an analogue or digital computer is almost essential when a non-linear differential equation is to be solved.

## REFERENCES


1. Davis, H. T., *Introduction to Non-linear Differential and Integral Equations*, Dover, 1962.
2. Jeffreys, H., *Asymptotic Approximations*, Oxford, 1962.
3. Heading, J., *Phase Integral Methods*, Methuen, 1962.
4. Levy, H. and Baggott, E. A., *Numerical Solutions of Differential Equations*, Dover, 1950.
5. Jeffreys, H. and Jeffreys, B. S., *Methods of Mathematical Physics*, Cambridge, 1956.
6. Noble, B., *Numerical Methods*, Oliver & Boyd, 1964.
7. Southwell, R., *Relaxation Methods*, Oxford, 1940.
8. Allen, D. N. de G., *Relaxation Methods in Engineering and Science*, McGraw-Hill, 1954.


## Exercises

**6.1.** Use the method of isoclines to find the general form of the integral curves of $\mathrm{d}x/\mathrm{d}s = s^2 - x^2$.

**6.2.** Given $x'' = 1 + x^2$, $x(0) = 1$, $x'(0) = 0$, compute $x(1)$ by the method of curvatures. (Take steps of $0{\cdot}2$ in $s$ in computation.)

**6.3.** Solve Exercise 6.2 by Picard's method of iteration.

**6.4.** Writing Van der Pol's equation as

$$\dot{x} = y,$$
$$\dot{y} = \varepsilon(1 - x^2)y - ax,$$

show that in the phase plane there is only one finite singular point, namely the origin.

[In Section 6.5 we found that there can exist a periodic solution corresponding to a particular relation between the initial conditions. In the phase plane this gives a unique closed trajectory to which all other trajectories are asymptotic (see Davis, *loc. cit.*[1] Ch. 12).]

**6.5.** Given $x'' - (h^2\psi_0 + h\psi_1 + \psi_2)x = 0$, where $h$ is a large parameter, and $\psi_0$, $\psi_1$, $\psi_2$ are given functions of $s$, assume that there is a solution of the form

$$x = \phi \mathrm{e}^{hw}\left(1 + \frac{1}{h}f_1 + \frac{1}{h^2}f_2 \cdots\right)$$

where $\phi$, $w$, $f_1$, $f_2$ are functions of $s$, and by the method of perturbation obtain the equations

$$w'^2 = \psi_0,$$
$$\phi'/\phi = (\psi_1 - w'')/2w',$$
$$2w'f_1' = \psi_2 - \phi''/\phi.$$

Hence show that

$$w = \int^s \psi_0^{\frac{1}{2}} ds,$$
$$\phi = \psi_0^{-\frac{1}{4}} \exp \int^s \tfrac{1}{2}\psi_1\psi_0^{-\frac{1}{2}} ds$$

and

$$f_1 = \int^s \tfrac{1}{2}\psi_0^{-\frac{1}{2}}(\psi_2 - \phi''/\phi) ds.$$

**6.6.** Complete the solution of Van der Pol's equation to the second order in $\varepsilon$ (Section 6.5), obtaining

$$x = (2 - 29\varepsilon^2/96n_0^2)\sin n_0 t + (\varepsilon/4n_0)(\cos n_0 t - \cos 3\, n_0 t) + (\varepsilon^2/16n_0^2)(3\sin 3n_0 t - \tfrac{5}{3}\sin 5n_0 t)$$

with

$$n_0^2 = a - \tfrac{1}{8}\varepsilon^2.$$

**6.7.** *Duffing's equation* is the non-linear differential equation

$$\ddot{x} + ax + bx^3 = f\cos \omega t,$$

where the term $bx^3$ is always small. Let $x = A\cos \omega t$ be the steady solution when $b = 0$. Putting $x = A\cos \omega t + bx_1(t)$ into the equation, show that an approximate solution is

$$x = A\left(1 + \tfrac{3}{4}\frac{A^2 b}{a - \omega^2}\right)\cos \omega t + \tfrac{1}{4}\frac{A^3 b}{a - 9\omega^2}\cos 3\omega t,$$

where

$$A(a - \omega^2) + \tfrac{3}{4}A^3 b = f.$$

**6.8.** Show that the WKBJ method gives the exact solution when

$$r(\xi) = (A\xi + B)^{-4},$$

where $A$ and $B$ are any constants.

**6.9.** Show that approximate solutions of Airy's equation

$$x'' - sx = 0$$

valid for large positive $s$ are $s^{-\frac{1}{4}} \exp(\pm \tfrac{2}{3}s^{\frac{3}{2}})$, and approximate solutions valid for large negative $s$ are

$$(-s)^{-\frac{1}{4}} \begin{matrix}\sin \\ \cos\end{matrix} \{\pm \tfrac{2}{3}(-s)^{\frac{3}{2}}\}.$$

(Compare (5.6.4) and (5.6.5).)

**6.10.** Show that approximate solutions of Schrödinger's equation for the harmonic oscillator

$$x'' + (2n + 1 - s^2)x = 0,$$

valid when $2n + 1 - s^2$ is large and positive, are

$$x = (2n + 1 - s^2)^{-\frac{1}{4}} \begin{matrix}\sin \\ \cos\end{matrix} \left\{\tfrac{1}{2}s\sqrt{(2n + 1 - s^2)} + \tfrac{1}{2}(2n + 1)\sin^{-1}\frac{s}{\sqrt{(2n + 1)}}\right\}.$$

o

# HINTS AND ANSWERS TO EXERCISES

**Ch. 1, p. 28**

**1.1.** Let $v$ be concentration of vapour. Diffusion upwards takes place at rate proportional to local concentration gradient. Measuring $x$ upwards from liquid surface, $\mathrm{d}^2v/\mathrm{d}x^2 = 0$. At liquid level $v = v_s$ (saturation). At top $\mathrm{d}v/\mathrm{d}x = 0$.

**1.3.** See C. A. Coulson, *Electricity* (Oliver & Boyd), Sect. 95.

Assuming (a) negligible resistance in wire and battery,

(b) $C$, $R$ and $L$ constants,

(c) Ohm's law,

(d) Law of electromagnetic induction,

we obtain, for current $I$,

$$L\,\mathrm{d}^2I/\mathrm{d}t^2 + R\,\mathrm{d}I/\mathrm{d}t + I/C = \mathrm{d}E/\mathrm{d}t.$$

**1.4.** See H. Lamb, *Statics* (Cambridge University Press), Sect. 146.

(a) $y = 0$

(b) $y' = 0$ (for end clamped horizontally)

(c) $y'' = 0$ (zero torque).

**1.5.** Take $x_1$, $x_2$, $x_3$ to be amounts of A, B, C at time $t$, $k_1$ decay constant for A→B, $k_2$ reaction rate for B and C. We may take $x_1 = X_1$, $x_2 = X_2$, $x_3 = X_3$ at $t = 0$.

**1.6.** See Coulson (*loc. cit.*), Sect. 62.

$$\mathrm{d}(mv_1)/\mathrm{d}t = qE_1 + q(v_2H_3 - v_3H_2)/c$$

and two similar equations.

**Ch. 2, p. 47**

**2.1.** We obtain $[\mathrm{A}]_t = [\mathrm{A}]_{t\,t^{\alpha}}e^{k\alpha}$, which is the required linear relation, since $k$ and $\alpha$ are fixed.

**2.2.**
$$(c_0{}^{1-\alpha} - c^{1-\alpha})/(1-\alpha) = \mathrm{k}t.$$

Measure $c$ as a function of $t$ and estimate $m = -\mathrm{d}c/\mathrm{d}t$. Since $m = \mathrm{k}c^{\alpha}$, $\log m = \alpha \log c + \log \mathrm{k}$, so that $\alpha$ is the slope of the graph of $\log m$ against $\log c$.

**2.3.** Separate variables and obtain solution

$$[\mathrm{A}]_0{}^{(1-n)}/(n-1) - [\mathrm{A}]^{(1-n)}/(n-1) = -k_nt.$$

Put $[\mathrm{A}] = \frac{1}{2}[\mathrm{A}]_0$ when $t = \tau_{\frac{1}{2}}$ to obtain given half-life.

**2.4.** See P. G. Ashmore, *Catalysis and Inhibition of Chemical Reactions* (Butterworth), p. 62.

$$\text{(a)} \quad dx/dt = 2k_1(a-x)^2 + 4k_2(a-x)^3,$$

with solution

$$2k_1t = \frac{1}{a-x} - \lambda \ln\left(\frac{1}{a-x} + \lambda\right), \quad \lambda = 2k_2/k_1.$$

$$\text{(b)} \quad dx/dt = 2k_1(a-x)(b-x) + 4k_2(a-x)(b-x)^2,$$

with solution

$$2k_1t = \frac{1}{c(1-\lambda c)} \ln(a-x) - \frac{1}{c} \ln(b-x) - \frac{\lambda}{1-\lambda c} \ln\{1 + \lambda(b-x)\},$$

where $c = a-b$.

This reduces to the solution of (a) when $c = 0$.

**2.5.** If $k_1 \gg k_1'$, (3.1) reduces to $dc/dt = -k_1c$ and (3.4) to $-k_1t = \ln(c/c_0)$ approximately.

**2.6.** See Ashmore (*loc. cit.*), p. 34.

**2.10.** At time $t$ let $[A] = [A]_0 - x$, $[B] = [B]_0 + x$. Then $dx/dt = k([A]_0 - x)([B]_0 + x)$, with solution

$$[A]_0([B]_0 + x)/[B]_0([A]_0 - x) = \exp([A]_0 + [B]_0)kt.$$

**Ch. 3, p. 85**

**3.1.** $L dI/dt + RI = E \cos \omega t$; $I = 0$ when $t = 0$.

Solution:

$$I = \{R \cos \omega t + \omega L \sin \omega t - R \exp(-Rt/L)\}E/(R^2 + \omega^2L^2).$$

Amplitude: $E(R^2 + \omega^2L^2)^{-\frac{1}{2}}$.

**3.4.** Substitution into equation (2.27) gives the answer immediately. The impulse of $mf(t)$ is $mP$, and the particle has initial velocity $P$.

**3.8.** From the equation of motion $\ddot{\mathbf{r}} + \lambda\mathbf{r} = \mathbf{0}$ we can prove $\dfrac{d}{dt}(\dot{\mathbf{r}} \times \mathbf{r}) = \mathbf{0}$, so that the normal to the plane of $\mathbf{r}$, $\dot{\mathbf{r}}$ is fixed. Taking coordinates, we find $x = A \cos nt + B \sin nt$, $y = C \cos nt + D \sin nt$, $A$, $B$, $C$, $D$ constants, $\lambda = n^2$. The locus is an ellipse with centre at the origin.

When $\mathbf{H}$ is imposed, the equation of motion is

$$\ddot{\mathbf{r}} - 2\gamma\dot{\mathbf{r}} \times \mathbf{k} + \lambda\mathbf{r} = \mathbf{0},$$

where $\mathbf{k}$ is a unit vector perpendicular to the plane of the orbit and $2\gamma = eH/mc$. Taking coordinates we obtain equations (4.34), and the subsequent analysis follows that for the Foucault pendulum.

**Ch. 4, p. 114**

**4.3.** (a) See R. V. Churchill, *Operational Mathematics*, New York, Sect. 16. If $L(p) \equiv (p-\beta)^r M(p)$, write

$$R(p)/L(p) = N(p)/(p-\beta)^r = \sum_{q=1}^{r} A_q/(p-\beta)^q + H(p),$$

and prove $$A_q = N^{(r-q)}(\beta)/(r-q)!$$

(b) $f(p)/L(p)$ will have repeated linear factor $(p-\alpha)$ in denominator, and procedure follows (a).

**Ch. 5, p. 161**

**5.1.** (a) Multiply by $v(s)/\sin s$. The left-hand side is an exact derivative if $3v \cot s = \mathrm{d}v/\mathrm{d}s$, giving $v = \sin^3 s$, $x = (1-\cos s)/\sin^3 s$.

(b) Divide through by $s(s-1)$. Integrating factor is $(s-1)/s$.
$x = s(s^2-4)/(s-1)$.

**5.2.** Integrating factor: $\exp \int \{-\mathrm{d}t/(1+t)\} = 1/(1+t)$,
$x = (1+t) \int_0^t f(\tau)\mathrm{d}\tau/(1+\tau)$. When $f(t) = 1-t^2$, $x = (1+t)(t-\frac{1}{2}t^2)$.

**5.3.** Since the d.e. is real, there must be a second solution

$$X_2(s) = (s^{\frac{1}{2}} - i)\exp(-is^{\frac{1}{2}}),$$

the complex conjugate of $X_1(s)$.

$$W = -\tfrac{2}{3}i,$$

$$x(s) = \int^s 3f(\sigma)\mathrm{d}\sigma\{(s^{\frac{1}{2}}\sigma^{\frac{1}{2}} + 1)\sin(s^{\frac{1}{2}} - \sigma^{\frac{1}{2}}) - (s^{\frac{1}{2}} - \sigma^{\frac{1}{2}})\cos(s^{\frac{1}{2}} - \sigma^{\frac{1}{2}})\}.$$

**5.4.** Put $x = vJ_0$ and show $v'' + v'(2sJ_0' + J_0)/sJ_0 = 0$. Integrating factor: $sJ_0^2$. $v' = 1/(sJ_0^2)$. $b = \frac{1}{4}$.

**5.5.** Multiplying by $1/t$ we obtain

$$\frac{1}{t}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{t}\frac{\mathrm{d}x}{\mathrm{d}t}\right) - 4t^2(t^2-1)x = 0.$$

Put $t^2 = s$.

**5.8.** Solutions are

$$x_1(s) = a_0\left(1 - 2s^2 + \frac{4}{7}s^4 \ldots\right), \quad \frac{u_{2n}}{u_{2n-2}} = \frac{4n-10}{4n-1}s^2,$$

$$x_2(s) = a_0 s^{\frac{1}{2}}\left(1 - \frac{5}{4}s^2 + \frac{5}{32}s^4 \ldots\right), \quad \frac{u_{2n}}{u_{2n-2}} = \frac{4n-9}{4n}s^2.$$

These series converge for $-1 < s < 1$.

**5.9.** cosh $ns$ and sinh $ns$ replace cos $ns$ and sin $ns$ throughout.

**5.10.** See Ref. 7, Ch. 2.

**Ch. 6, p. 196**

**6.1.** Isoclines are rectangular hyperbolae with asymptotes $x = \pm s$. On the asymptotes the slopes are zero. Inflexions lie on the three-branched curve $s^2x - s - x^3 = 0$.

**6.3.** For small $s$, successive approximations are

$$x_0 = 1,$$
$$x_1 = 1 + s^2,$$
$$x_2 = 1 + s^2 + \frac{1}{6}s^4 + \frac{1}{30}s^6.$$

**6.4.** Phase trajectories $dy/dx = \{\varepsilon(1-x^2)y - ax\}/y$, with only singular point at $x = 0, y = 0$. Near this point, for small $\varepsilon$, $y^2 + ax^2 = \text{const}$.

**6.8.** In WKBJ formula $x = r^{-\frac{1}{4}} \exp \int ir^{\frac{1}{2}} d\xi$ write $r^{\frac{1}{2}} = V(\xi)$. Show that $V$ satisfies $VV'' = 2V'^2$. Integrate twice to obtain $V = (A\xi + B)^{-1}$. With this $V$, $x$ is an exact solution of $x'' + r(\xi)x = 0$.

**6.9.** Apply WKBJ method directly.

**6.10.** Apply WKBJ method with $r(s) = 2n + 1 - s^2$.
In the integral substitute $s = (2n+1)^{\frac{1}{2}} \sin\theta$.

# INDEX