STURM-LIOUVILLE THEORY

CONTENTS

STURM-LIOUVILLE THEORY

Introduction

The mathematical expression of many problems from the natural sciences: Physics, Chemistry, Geology and the life sciences, leads to differential equations of the form

$$a(x)y'' + b(x)y' + c(x)y = f(x).$$

Some of the more commonly occurring of these are listed in the following table together with a few of the situations in which they arise.

$y'' + by' + cy = f(x)$ (b, c constants)	Forced vibrations of mechanical and electrical systems. Simple ecological models.
Euler (or Cauchy) Equation $x^2y'' + bxy' + cy = f(x)$ (b, c constants)	Some potential problems with circular symmetry.
Bessel Equation (solutions of order ν) $x^2y'' + xy' + (x^2 - \nu^2)y = 0$	Vibrational, gravitational and electromagnetic potential problems with cylindrical symmetry. Diffraction problems (astronomy) resolving power of optical instruments. Chemistry, Biochemistry.
Airy Equation $y'' - xy = 0$ (Essentially a Bessel Equation)	Scattering problems (atomic collisions, rainbows). Quantum description of a particle in a uniform field.
Legendre Equation $(1 - x^2)y'' - 2xy' + n(n + 1)y = 0$	Potential problems (as for Bessel's) with spherical symmetry. Including quantum models of the Hydrogen atom.
Laguerre Equation $xy'' + (1 - x)y' + ny = 0$	Radial structure of the quantum mechanical hydrogen atom.
Hermite Equation $y'' + 2xy' + 2ny = 0$	Quantum mechanical Harmonic Oscillator.

Tchebyshev Equation $(1 - x^2)y'' - xy' + n^2y = 0$	Theory of Filters (Telecommunications).
Hypergeometric Equation $x(1 - x)y'' + (c - (a + b + 1)x)y' - aby = 0$	Unifies a broad class of special functions. Statistics.
Mathieu Equation $y'' + (a + b \sin x)y = 0$	Vibration of elliptical membranes.

The above applications, and most of the later examples, are drawn from physics, not because analogous problems do not occur in all the other branchs of science, but simply because it is the area of which I am least ignorant.
No such table would be even partially complete without adding the one-dimensional Schrödinger Equation

$$\frac{\hbar^2}{2m}\psi'' + [E - P(x)]\psi = 0$$

determining the probability density function, $|\psi(x)|^2$, for the position of a particle, mass m, of total energy E and potential energy P(x) confined to 'motion' along a straight line.

Except in special cases, none of the above equations (with the possible exception of the first two) have solutions which can be expressed as finite combinations of the elementary functions* ($x \mapsto x$, sin, cos, exp, and their inverses). Their solutions belong to the class of *special functions* (Bessel's functions, gamma functions etc.) whose individual properties were intensively researched during the 19th century, and are still of great interest to the applied mathematician and scientist of today.

In the absence of such elementary closed form solutions it becomes important to establish techniques whereby the behaviour of solutions can be studied directly from the equations. Even in cases where a closed form solution is available, it frequently proves less tedious to work from the equation itself rather than employ the solution.

In this course we develop some general theories which enable us to probe (at least qualitatively) certain aspects of solutions to such equations.

The thorny question of whether or not the equations under consideration do indeed possess solutions will not be tackled (see PMI notes and Boyce and diPrime §2.11), we will assume that solutions exist and when necessary that they are unique.

---

* In many cases however, solutions can be expressed as infinite series, products, etc. of these elementary functions.

Throughout it is well to bear in mind the trivial, yet powerful, observation that since a solution to a differential equation must be a priori differentiable, it is therefore continuous.

Examples

We now examine specific cases which may help motivate the type of questions considered in the ensuing work and how they can arise. The underlying physics is included for interest only and is not an examinable part of the course.

1. VIBRATING MEMBRANES

We begin by deriving the equation of motion for small amplitude transverse vibrations of an 'infinite' plane membrane of density ρ per unit of area, under a uniform tension τ (per unit of length). For our later work it will be convenient if we work in polar coordinates (r, ϕ).

An 'infinitesimal' element of the membrane surface such as illustrated in figure 1(a) is subject to the forces indicated on the diagram.

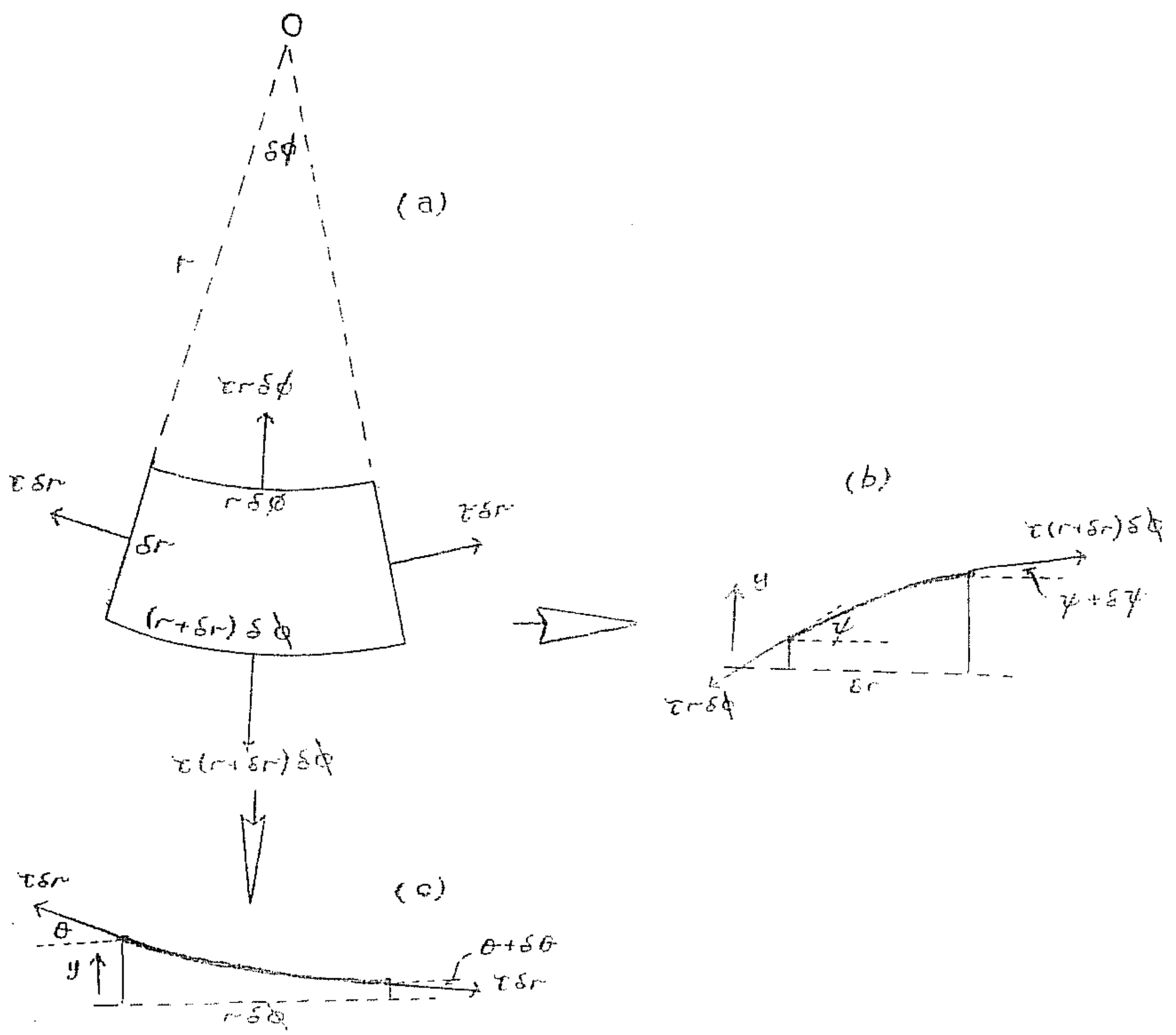


Figure 1

If, due to slight vertical distortions of the membrane, the cross-sectional views of the element are as shown in figures 1(b) and (c), we have

total vertical force on element

$$= \tau\delta r(\sin(\theta + \delta\theta) - \sin\theta) + \tau((r + \delta r)\delta\phi \sin(\psi + \delta\psi) - r\delta\phi \sin\psi)$$

$$\doteq \tau\delta r\delta\theta + \tau r\delta\phi\delta\psi + \tau\delta r\delta\phi\psi$$

(since θ and ψ are small and $\delta\theta$, $\delta\psi$ even smaller).

So, if y denotes the vertical displacement of the element, we have, from Newton's second law

$$\rho r\delta\phi\delta r \frac{\partial^2 y}{\partial t^2} \simeq \tau(r\delta\phi\delta\psi + \delta r\delta\phi\psi + \delta r\delta\theta)$$

The approximation becoming more nearly exact the smaller the values of δr and $\delta\phi$.

Now, $\psi \simeq \tan\psi = \frac{\partial y}{\partial r}$ and $\theta \simeq \tan\theta = \frac{\partial y}{r\partial\phi}$

whence

$$\rho\frac{\partial^2 y}{\partial t^2} \simeq \tau\left(\frac{\delta\left(\frac{\partial y}{\partial r}\right)}{\delta r} + \frac{1}{r}\frac{\partial y}{\partial r} + \frac{\delta\left(\frac{\partial y}{r\partial\phi}\right)}{r\delta\phi}\right)$$

and so, passing to the limit as δr, $\delta\phi \to 0$ we recognise this as

$$\rho\frac{\partial^2 y}{\partial t^2} \simeq \tau\left(\frac{\partial^2 y}{\partial r^2} + \frac{1}{r}\frac{\partial y}{\partial r} + \frac{1}{r^2}\frac{\partial^2 y}{\partial\phi^2}\right)$$

or

$$\nabla^2 y = \frac{\partial^2 y}{\partial r^2} + \frac{1}{r}\frac{\partial y}{\partial r} + \frac{1}{r^2}\frac{\partial^2 y}{\partial\phi^2} = \frac{\rho}{\tau}\frac{\partial^2 y}{\partial t^2}$$

where ∇^2 (= Δ) is the two-dimensional Laplacian operator, here expressed for polar coordinates. This partial differential equation is the two dimensional <u>wave equation</u> in polar coordinates. (What is the form of the equation in rectangular coordinates?)

Solutions to the wave equation, of the form $y(r,\phi,t) = R(r)\Phi(\phi)T(t)$ may be obtained by separation of variables, <u>i.e.</u> observing that $y = R\Phi T$ implies

$$\frac{\partial y}{\partial r} = \Phi T\frac{dR}{dr}, \frac{\partial^2 y}{\partial r^2} = \Phi T\frac{d^2R}{dr^2} \quad \text{etc.}$$

we obtain

$$\Phi T\frac{d^2R}{dr^2} + \frac{1}{r}\Phi T\frac{dR}{dr} + \frac{1}{r^2}RT\frac{d^2\Phi}{d\phi^2} = \frac{\rho}{\tau}R\Phi\frac{d^2T}{dt^2}$$

or dividing throughout by $y = R\Phi T$,

$$\frac{\frac{d^2R}{dr^2} + \frac{1}{r}\frac{dR}{dr}}{R} + \frac{\frac{1}{r^2}\frac{d^2\Phi}{d\phi^2}}{\Phi} = \frac{\rho}{\tau}\frac{\frac{d^2T}{dt^2}}{T}$$

which is of the form: a function of r and ϕ only equals a function of t only, and so we conclude that both sides must equal a common constant μ. (Prove)

Thus T is determined by

$$\frac{d^2T}{dt^2} = \frac{\mu\tau}{\rho} T$$

while

$$\frac{\frac{d^2R}{dr^2} + \frac{1}{r}\frac{dR}{dr}}{R} + \frac{\frac{1}{r^2}\frac{d^2\Phi}{d\phi^2}}{\Phi} = \mu$$

or

$$r^2\left(\frac{\frac{d^2R}{dr^2} + \frac{1}{r}\frac{dR}{dr}}{R} - \mu\right) = \frac{-\frac{d^2\Phi}{d\phi^2}}{\Phi} = \text{a constant}$$

λ, say (for similar reasons to above).

Consequently Φ is given by

$$\frac{d^2\Phi}{d\phi^2} = -\lambda\Phi \text{ (solutions } \Phi(\phi) = {\sin \atop \cos} \sqrt{\lambda}\,\phi)$$

and clearly, unless our membrane is schizophrenic we require $\Phi(\phi + 2\pi) = \Phi(\phi)$ for all ϕ, which in general necessitates $\sqrt{\lambda} = n$, an integer.

R is then a solution of

$$r^2 \frac{d^2R}{dr^2} + r\frac{dR}{dr} + (-\mu r^2 - n^2)R = 0$$

which transforms to the Bessel equation

$$x^2\frac{d^2R}{dx^2} + x\frac{dR}{dx} + (x^2 - n^2)R = 0$$

where $x = \sqrt{-\mu}\, r$. (Note: for wave-like solutions we need $T(t)$ to vary periodically, or $\mu < 0$, so $\frac{x}{r} = k > 0$.)

Circular Nodes. If the circle $r = a$ is to be a permanent node (curve of zero amplitude "vibration") we require

$$y(a,\phi,t) = R(a)\,\Phi(\phi)\,T(t) = 0 \text{ for all } \phi \text{ and } t$$

and so (unless $\Phi(\phi)$ or $T(t) \equiv 0$, i.e. the membrane is static)

$$R(a) = 0$$

or $x = ka$ is a zero of the solution to the above Bessel Equation.

Vibration of a circular drum skin. The clamped outer edge of the skin, at $r = \frac{D}{2}$ (where D is the diameter of the drum head), must be a node. Thus if $0 \leqslant z_1 < z_2 < z_3 < \ldots$ are successive zeros of the solution to the Bessel equation

$$x^2y'' + xy' + (x^2 - n^2)y = 0$$

we require k to be such that

$$\frac{D}{2} = kz_j \quad (j = 1, 2, \ldots)$$

Thus the possible values of $\mu = -k^2$ are

$$\mu = -\frac{D^2}{4z_j^2} \quad (j = 1, 2, 3, \ldots)$$

and via the equation for T, these values determine the frequency of vibration of the skin.

The mode of vibration corresponding to each of these frequencies is known as an "harmonic".

We therefore see that it is desirable to determine the existence and location of zero's to solutions of equations such as the one considered here.

Often, by touching the centre of the skin, a drummer introduces a node at the centre.

For a fixed $n \in \{1,2,\ldots\}$, the problem then becomes;

find values of μ such that

$$r^2\frac{d^2R}{dr^2} + r\frac{dR}{dr} - n^2R = \mu r^2R$$

has a non-trivial solution (i.e. other than $R \equiv 0$) on the domain $0 \leqslant r \leqslant \frac{D}{2}$ which vanishes at the two boundary points of the domain, i.e. is such that $R(0) = 0$, $R(\frac{D}{2}) = 0$. (Again the values of μ determine the harmonics of the drum.) For obvious reasons such a problem is usually referred to as a *Boundary value problem*, in contrast to the *Initial value problem:* find solutions satisfying initial conditions,

$$R(r_0) = R_0, \; R'(r_0) = R'_0,$$

specified at one point (the initial point) $r = r_0$.

(A simpler example, and one which the student might well investigate himself, is afforded by the vibration of a taut, uniform string clamped at each end, which, unlike the above illustration, is completely tractable in terms of the elementary functions.)

2. HYDROGENIC ION

Schrödinger's equation for the 'spacial' probability density function of a single electron (mass m, charge -e) bound to a considerably more massive neucleus of charge Ze at the origin of a set of spherical polar coordinates (r,θ,ϕ), is

$$\frac{\hbar^2}{2m}\nabla^2\psi + \left(\frac{Ze^2}{r} - E\right)\psi = 0 \qquad (\text{where } E > 0).$$

For spherically symmetric solutions, ψ is a function of r only, in which case, the form of the Laplacian operator for spherical polar coordinates, leads to the equation

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\psi}{dr}\right) + \left(\frac{2mZe^2}{\hbar^2 r} - \frac{2mE}{\hbar^2}\right)\psi = 0$$

For the change of variables $\psi(r) = e^{-\beta x}y(x)$, $x = \alpha r$ we have

$$\frac{d\psi}{dr} = \frac{dx}{dr}\frac{d}{dx}\left(e^{-\beta x}y(x)\right) = \alpha e^{-\beta x}\,y' - \alpha\beta e^{-\beta x}\,y$$

and similarly

$$\frac{d^2\psi}{dr^2} = \alpha^2 e^{-\beta x}\,y'' - 2\alpha^2\beta e^{-\beta x}\,y' + \alpha^2\beta^2 e^{-\beta x}\,y,$$

which, upon substitution, leads to the transformed equation

$$xy'' + (2 - 2\beta x)y' + \left[\left(\frac{2mZe^2}{\alpha\hbar^2} - 2\beta\right) + \left(\beta^2 - \frac{2mE}{\hbar^2\alpha^2}\right)x\right]y = 0$$

and so choosing $\beta = \tfrac{1}{2}$, $\alpha = \sqrt{\frac{8mE}{\hbar^2}}$

we have the "associated Laguerre Equation"

$$xy'' + (2 - x)y' + \lambda y = 0, \text{ where } \lambda = \frac{Ze^2}{\hbar}\sqrt{\frac{m}{2E}}\ .$$

We are thus lead to seek solutions of this equation for the domain $0 \leqslant x < \infty$ which satisfy the "natural" boundary conditions

$$y(0) \text{ and } \lim_{x\to\infty} y(x) \text{ are both finite.}$$

The values of E corresponding to those values of λ $\left(E = \frac{Z^2me^4}{2h^2\lambda^2}\right)$ for which the above <u>Boundary value problem</u> has a non-trivial solution, represent the permissible energy levels for the electron.
The <u>zero's</u> (if any) of the solution for a permissible energy level correspond to "forbidden" regions for the electron.
As we have tried to indicate, many questions of applied mathematics result in

boundary value problems, a theory of Boundary value problems is therefore of considerable practical importance.

STURM'S OSCILLATION THEORY.

In this section we investigate the existence and location of zeros for solutions to equations of the form

$$y'' + a(x)y' + b(x)y = 0.$$

RECALL x_0 is a *zero* of the function y if $y(x_0) = 0$.

The Liouville Normal Form

If y and v are functions related by $y(x) = u(x)v(x)$ where $u(x) > 0$, for all x, then

(a) x_0 is a zero of y if and only if x_0 is a zero of v

and (b) if y satisfies

$$y'' + ay' + by = 0$$

then upon substitution of uv for y, we find v satisfies

$$uv'' + (2u' + au)v' + (u'' + au' + bu)v = 0.$$

Thus by choosing u appropriately (subject to the constraint $u > 0$) we may obtain an equation for v, of a "simpler" form than the original equation for y, whose solutions have precisely the same zeros as solutions to the original equation.

It is particularly suitable, to choose u such that $2u' + au = 0$

i.e. $$\frac{u'}{u} = -\tfrac{1}{2}a \text{ or } u(x) = e^{-\frac{1}{2}\int^{x} a(t)dt},$$

in which case

$$u(x) > 0 \text{ provided } \left|\int^{x} a(t)dt\right| < \infty$$

and the equation for v becomes

$$v'' + \left(\frac{u''}{u} + a\frac{u'}{u} + b\right)v = 0$$

or $$v'' + (b - \tfrac{1}{2}a' - \tfrac{1}{4}a^2)v = 0$$

since, $\frac{u'}{u} = -\tfrac{1}{2}a$ and $u'' = -\tfrac{1}{2}(au)' = -\tfrac{1}{2}(a'u - \tfrac{1}{2}a^2u)$ (provided a is differentiable)

so $$\frac{u''}{u} = -\tfrac{1}{2}a' + \tfrac{1}{4}a^2.$$

The equation

$$v'' + (b(x) - \tfrac{1}{2}a'(x) - \tfrac{1}{4}a^2(x))v = 0$$

is the *normal form* of

$$y'' + a(x)y' + b(x)y = 0$$

It is clear from the above derivation that the presence and location of zero's can be studied from the normal form.

Henceforth, we will therefore only consider equations of the form

$$y'' + I(x)y = 0.$$

EXAMPLES. Airy's, Mathieu's and the 1-dimensional Schrödinger equation are already in normal form.

Writing Bessel's equation as

$$y'' + \frac{1}{x}y' + \left(1 - \frac{\nu^2}{x^2}\right)y = 0 \qquad (x > 0)$$

we see that

$$a(x) = \frac{1}{x} \text{ (and } \left|\int_1^x \frac{1}{t}\,dt\right| = |\ell n\, x| < \infty \text{ for all finite } x)$$

while $\quad b(x) = \left(1 - \frac{\nu^2}{x^2}\right),$

and so the normal form of Bessel's equation is

$$v'' + \left(1 - \frac{\nu^2}{x^2} - \frac{1}{4x^2} + \frac{1}{2x^2}\right)v = 0$$

or

$$v'' + \left(1 + \frac{1 - 4\nu^2}{4x^2}\right)v = 0 \qquad (\text{for } x > 0)$$

EXERCISES.

Noting any necessary restrictions, obtain the normal form for each of the equations listed in the previous table.

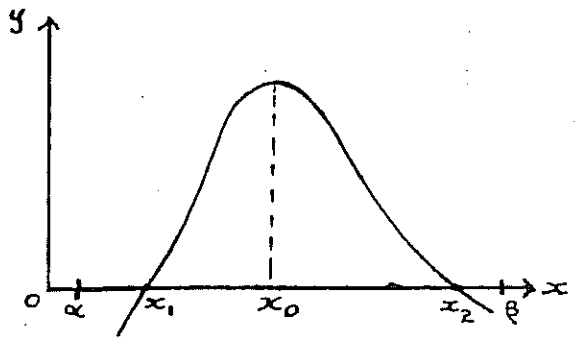
Note: Rather than remembering the precise form of the v coefficient in the normal form (i.e. $b - \frac{1}{2}a' - \frac{1}{4}a^2$) it is nearly as easy to work from first principles. You should try at least one of the above cases in this way.

Basic Theory

THEOREM. (*Convexity*) *If* $I(x) < 0$ *for* $\alpha < x < \beta$, *then any non-trivial solution of* $v'' + I(x)v = 0$ *has at most one zero between* α *and* β.

Proof. Let v be a non-trivial solution and assume v has two consecutive zeros at x_1 and x_2 where $\alpha < x_1 < x_2 < \beta$. Then either $v(x) > 0$ for $x_1 < x < x_2$ or $v(x) < 0$ for $x_1 < x < x_2$. (This follows since v is continuous and $v'(x_1) \neq 0$, otherwise the unique solution satisfying $v(x_1) = v'(x_1) = 0$ would be $v \equiv 0$.)

In case $v(x) > 0$ for $x_1 < x < x_2$ we have $v''(x) = -I(x)v(x) > 0$ for every x between x_1 and x_2 but, since $v(x_1) = v(x_2) = 0$ there must exist a point x_0 between x_1 and x_2 at which v attains a maximum; then $v'(x_0) = 0$ and $v''(x_0) \leqslant 0$, contradicting $v''(x_0) > 0$. We therefore conclude that in this case there do not exist consecutive zeros of v between α and β and so v has at most one zero in the interval. A similar argument applies in the case



$$v(x) < 0 \text{ for } x_1 < x < x_2.$$ ■

EXAMPLE. The equation $y'' - x^3y = 0$ satisfies the conditions of the theorem for all x with $\varepsilon < x < \infty$ (where ε is any strictly positive number). So, there is at most one zero greater than ε.

THEOREM (<u>*Sturm's Theorem*</u>). *Let* u *and* v *be non-trivial solutions of* $v'' + I(x)v = 0$ *and* $u'' + J(x)u = 0$ respectively *where* $I(x) > J(x)$ *for all* x *with* $\alpha < x < \beta$. *Then* v *has at least one zero between any two consecutive zeros of* u, *provided these are both between* α *and* β.

Proof. Let $x_1 < x_2$ be consecutive zeros of u, both lying between α and β, and assume $v(x) \neq 0$ for $x_1 < x < x_2$.

Then both u and v have a constant sign throughout the interval from α to β. Without loss of generality we take both to be strictly positive <u>i.e.</u>

$$u(x) > 0 \text{ and } v(x) > 0 \text{ for all } x \text{ with } x_1 < x < x_2$$

(possible, since $-u$ and $-v$ are also solutions satisfying the required conditions).

Now construct

$w(x) = v(x)u'(x) - u(x)v'(x)$, then

$w(x_1) \geqslant 0$ (since $u(x_1) = 0$ and $u(x) > 0$ for $x_1 < x < x_2$ and so $u'(x_1) > 0$) similarly

$w(x_2) \leqslant 0$.

So (by the Mean Value Theorem) there is a point x_0 between x_1 and x_2 for which $w'(x_0) \leqslant 0$.

But, $w'(x) = v(x)u''(x) - u(x)v''(x)$

$= [I(x) - J(x)]v(x)u(x) > 0$ by the assumptions on I, J, v and u for $x_1 < x < x_2$, contradicting $w'(x_0) \leqslant 0$. So v must have a zero between x_1 and x_2. ■

NOTE: This theorem remains true under the weaker assumption $I(x) \geqslant J(x)$, $I \not\equiv J$ for $\alpha \leqslant x \leqslant \beta$.

COROLLARY (Spacing of Zeros Theorem)

Let I *be such that* $0 < m < I(x) < M$ *for* $\alpha < x < \beta$, *where* m, M *are constants, then if* x_1 *and* x_2 *are two consecutive zeros of* $v'' + I(x)v = 0$, *with* $\alpha < x_1 < x_2 < \beta$, *we have*

$$\pi/\sqrt{M} < x_2 - x_1 < \pi/\sqrt{m} \ .$$

Proof. Consider the comparison equation $u'' + Mu = 0$ which has a non-trivial solution $u(x) = \sin \sqrt{M}(x - x_1)$ with zero at x_1 and the next zero at $x_1 + \pi/\sqrt{M}$. Now by Sturm's Theorem there is a zero of u between x_1 and x_2 so we must have $x_2 > x_1 + \pi/\sqrt{M}$ or

$$\pi/\sqrt{M} < x_2 - x_1 \ .$$

The upper bound is established in the same manner. ■

EXAMPLE.

From the normal form of Bessel's Equation

$$v'' + \left[1 + \frac{1 - 4\nu^2}{4x^2}\right]v = 0$$

we have

$$I(x) = 1 + \frac{1 - 4\nu^2}{4x^2} > \begin{cases} 1 \text{ for all } x > 0, \text{ if } |\nu| < \frac{1}{2} \\ \frac{1}{2} \text{ for } x > \sqrt{\frac{4\nu^2 - 1}{2}} \quad \text{all other } \nu \end{cases}$$

Thus, taking $J(x) = \frac{1}{2}$, we see that for sufficiently large x any solution of Bessel's equation has a zero between $n\sqrt{2}\pi$ and $(n + 1)\sqrt{2}\pi$, consecutive zeros of a solution of $u'' + \frac{1}{2}u = 0$.

Further for each $x_0 > 0$ it is easily seen that we can select numbers $m(x_0)$ and $M(x_0)$ with

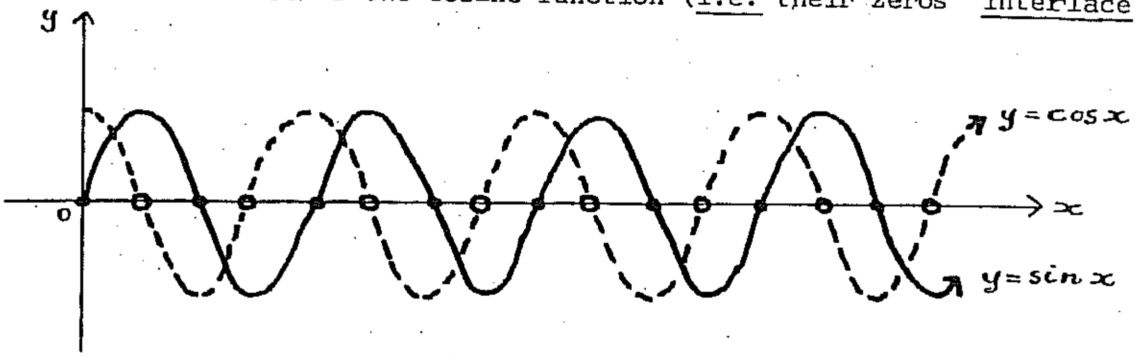
$$m(x_0) < \left(1 + \frac{1 - 4\nu^2}{x^2}\right) < M(x_0) \text{ for all } x > x_0$$

in such a way that $m(x_0)$ and $M(x_0)$ both tend to 1 as $x_0 \to \infty$. Thus for large values of x, the spacing between consecutive zeros of any non-trivial solution to Bessel's equation is very nearly π.

Our next result extends the observation:

$y = \sin x$ and $y = \cos x$ are linearly independent[(1)]

solutions of $y'' + y = 0$ and between each pair of consecutive zeros of the sine function there is a zero of the cosine function (i.e. their zeros interlace).



THEOREM (Interlacing of Zeros). *Let* v_1, v_2 *be two linearly independent, non-trivial, solutions of* $v'' + I(x)v = 0$, *then between each pair of consecutive zeros of* v_1 *there is a zero of* v_2.

Proof. Since v_1 and v_2 are linearly independent, twice differentiable functions their Wronskian $W(v_1, v_2) = v_1'v_2 - v_1v_2'$ is never zero,[(2)] and so has constant sign. Without loss of generality take $W(v_1, v_2) > 0$ (otherwise replace W by -W in the following argument).

Let $x_1 < x_2$ be two consecutive zeros of v_1 and assume $v_2(x) \neq 0$ for all x between x_1 and x_2. So we may form the quotient $\frac{v_1}{v_2}$ for all x between x_1 and x_2.

---

(1) Recall: Two functions f and g are linearly independent if $a\cdot f(x) + b\cdot g(x)=0$ for all x implies the constants a and b are both zero.

(2) If $W(v_1, v_2)(x_0) = 0$ for some x_0 we show there exist non-zero constants a, b with $av_1 + bv_2 = 0$ and so v_1 and v_2 are not linearly independent.

Now, since v_1 is non-trivial, there exists a point x_1 with $v_1(x_1) \neq 0$ and so by the continuity of v_1 there is an open interval J containing x_1 on which v_1 does not vanish. So for $x \in J$ we can form the quotient $\frac{v_2}{v_1}$ for which

$$\left(\frac{v_2}{v_1}\right)' = \frac{W(v_1, v_2)}{v_1^2} = 0 \qquad \left\{ \begin{aligned} \text{as } W' &= v_1''v_2 - v_1v_2'' \\ &= -Iv_1v_2 + Iv_1v_2 \\ &= 0 \\ \text{So } W &= \text{constant} \\ &= 0 \text{ as } W(x_0) = 0 \end{aligned} \right.$$

So for $x \in J$ $\qquad \frac{v_2}{v_1} = k$, a constant

or $\qquad v_2 = kv_1$

We now show this holds true for all x.

From $y = v_2 - kv_1$, then $y(x) = 0$ for all $x \in J$ and so in particular $y(x_0) = 0$ and further since x_0 is an interior point of J $y'(x_0) = 0$. Further y, being a linear combination of the solutions v_1 and v_2 is itself a solution of the second order equation $v'' + J(x)v = 0$, which as we have seen satisfies the initial conditions $y(x_0) = y'(x_0) = 0$. Clearly another such solution would be the zero function. However the solution to such a problem is unique and we therefore conclude that $y = 0$ or $v_2 = kv_1$ and so v_1 and v_2 are not linearly independent.

Now $$\int_{x_1}^{x_2} \left(\frac{v_1}{v_2}\right)' = \left[\frac{v_1}{v_2}\right]_{x_1}^{x_2} = 0 \text{ (as } v_1(x_1) = v_1(x_2) = 0)$$

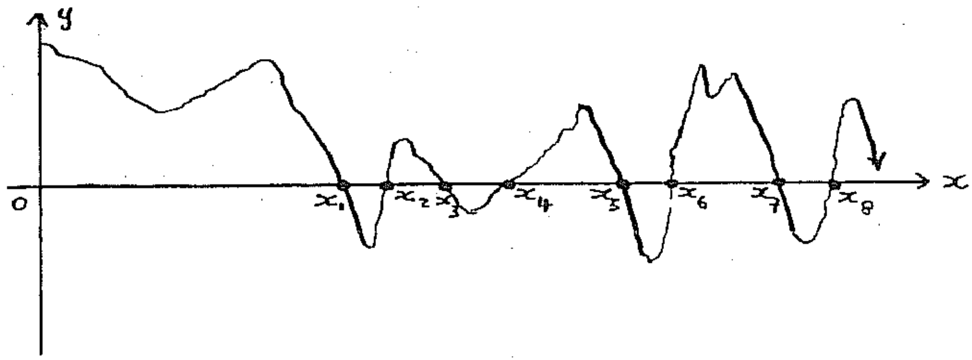
But $$\left(\frac{v_1}{v_2}\right)' = \frac{v_1'v_2 - v_1v_2'}{v_2^2} = \frac{W(v_1, v_2)}{v_2^2} > 0$$

so $$\int_{x_1}^{x_2} \left(\frac{v_1}{v_2}\right)' = \int_{x_1}^{x_2} \frac{W(v_1, v_2)}{v_2^2} > 0 \quad \text{a contradiction.}$$

Hence we cannot have $v_2(x) \neq 0$ for all x between x_1 and x_2 whence there exists $x_0(x_1 < x_0 < x_2)$ with $v_2(x_0) = 0$ i.e. v_2 has a zero between x_1 and x_2. ∎

Oscillatory Solutions

The presence (or absence) of oscillatory behaviour is of importance in many physical and biological situations. Clearly the property of a function, $y = f(x)$, oscillating about the line $y = 0$ is characterized by its repeated crossing of the x-axis.*



We therefore offer the following definition.

DEFINITION. The function f *is oscillatory* (about $y = 0$) if there exists a sequence of points $x_1 < x_2 < x_3 < \ldots < x_n < \ldots$ with $x_n \to \infty$ as $n \to \infty$ such that $f(x_i) = 0$, $i = 1,2,\ldots$ and $f(x) \neq 0$ if $x \neq x_n$ any n.

THEOREM. *If* I *is such that* $I(x) > m > 0$ *for all* $x > \alpha \geqslant 0$, *then any non-trivial solution of* $v'' + I(x)v = 0$ *is oscillatory.*

Proof. The equation $u'' + mu = 0$ has a non-trivial solution $u(x) = \sin \sqrt{m}\, x$ which has zeros at $x = 0$, $\pi/\sqrt{m}$, $2\pi/\sqrt{m}$, $3\pi/\sqrt{m}$, ... and so, since Sturm's Theorem applies, there exists zeros $x_1, x_2, x_3, \ldots$ of any non-trivial solution of $v'' + I(x)v = 0$ with

$$0 \leqslant \alpha \leqslant k\pi/\sqrt{m} < x_1 < (k+1)\pi/\sqrt{m} < x_2 < (k+2)\pi/\sqrt{m} < \ldots .$$

∎

* Oscillatory behaviour is not to be confused with periodicity, which is a special case. Thus $e^{-t} \sin t$ is oscillatory but not periodic.

As a consequence of the convexity theorem we also have

THEOREM. *If* I *is such that* $I(x) < 0$ *for all* $x > \alpha \geqslant 0$, *then* $v'' + I(x)v = 0$ *has no oscillatory solutions.*

Frequently an application of the spacing of zeros theorem allows something to be said on the "frequency" of the oscillations.

EXAMPLES.

For Bessel's equation in normal form

$$I(x) = 1 + \frac{1 - 4\nu^2}{4x^2} > \tfrac{1}{2} \text{ for all } x > \sqrt{\frac{|4\nu^2 - 1|}{2}}$$

and so solutions to Bessel's equation are oscillatory.

For $x > 0$ the Airy equation $y'' - xy = 0$ has $I(x) = -x$ negative and so solutions are not oscillatory. Reversing the direction of x by the substitution $t = -x$ we obtain

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

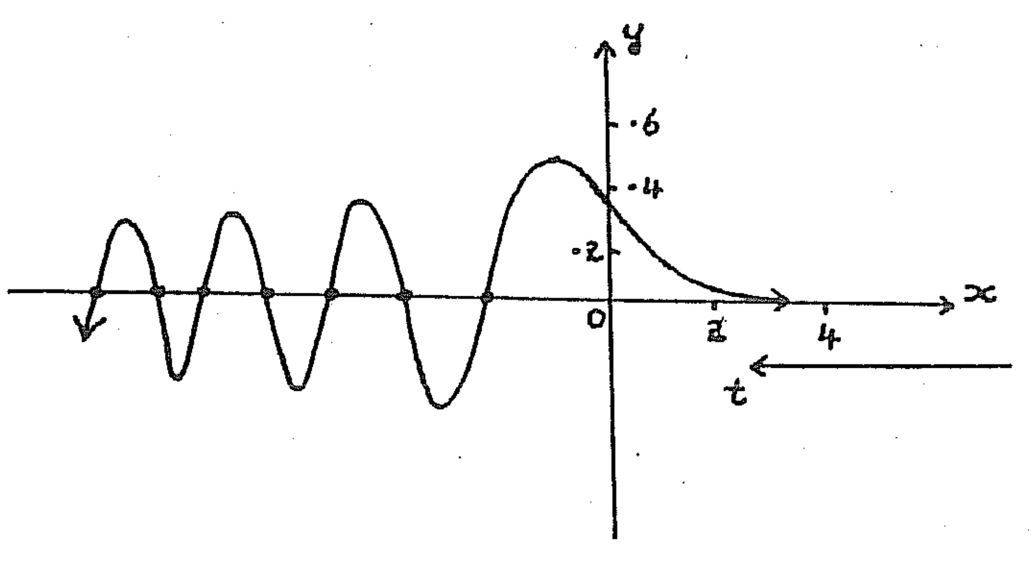
which is oscillatory for $t > 0$ (i.e. $x < 0$).

Further, from the spacing of zeros theorem, we conclude that

$$t_2 \leqslant t_1 + \pi/\sqrt{t_1} \text{ where } t_2 > t_1$$

are consecutive zeros, i.e. the maximum spacing between zeros decreases as $t \to \infty$.

Thus, the zeros must be distributed as illustrated below, where the form of a solution curve has also been indicated.



EXERCISES

1. By obtaining explicit solutions for the equation $y'' + ky = 0$ (k a constant) show that the theorems obtained so far are 'natural' generalizations of what this equation might lead you to suspect.
2. Under what conditions does

$$y'' + ay' + by = 0 \quad (a,\ b \text{ constants})$$

have

(a) oscillatory solutions

(b) periodic solutions?

3. Discuss the existence of oscillatory solutions to Mathieu's equation

$$v'' + (a + b \sin x)v = 0 \ (a,\ b \text{ constants}).$$

4. (The condition $m > 0$ cannot be dropped from the corollary to Sturm's Theorem.) Show that the normal form of Euler's equation, $y'' + \frac{1}{8}x^{-2}y = 0$, has solutions of the form $y = Ax^{r_1} + Bx^{r_2}$ where r_1 and r_2 are appropriately chosen real numbers and so has at most one zero greater than 0, even though

$$I(x) = \frac{1}{8}x^{-2} > 0 \text{ for all } x > 0.$$

5. Show that if v is a non-trivial solution of $v'' + I(x)v = 0$ with $I(x) > 0$ for $\alpha < x < \beta$ where $v'(\alpha) = v'(\beta) = 0$, then v has a zero between α and β.

*6. Using the result obtained for Bessel's equation, investigate the higher harmonics of a circular drum.

Boundary Value Problems

Finite-dimensional Analogues. Our treatment of boundary-value problems will involve generalizations of finite-dimensional vector space concepts, some of which were discussed in Sections VII,VIII of the Linear Algebra Notes, e.g. *eigenvalue* has essentially the same meaning, *eigenvector* becomes *eigenfunction*. Another relevant finite-dimensional concept is the following:

On the vector space X of complex n-tuples we can define a scalar (colloquially, "dot") product by

$$\underset{\sim}{x}\cdot\underset{\sim}{y} = (x_1, x_2, \ldots, x_n)\cdot(y_1, y_2, \ldots, y_n)$$

$$= x_1\overline{y_1} + x_2\overline{y_2} + \ldots + x_n\overline{y_n} = \sum_{j=1}^{n} x_j\overline{y_j}$$

which has the readily established properties

1. $\underset{\sim}{x}\cdot\underset{\sim}{x} = 0$ if and only if $\underset{\sim}{x} = \underset{\sim}{0}$.
2. $(\lambda\underset{\sim}{x})\cdot\underset{\sim}{y} = \lambda(\underset{\sim}{x}\cdot\underset{\sim}{y})$
3. $(\underset{\sim}{x} + \underset{\sim}{y})\cdot\underset{\sim}{z} = \underset{\sim}{x}\cdot\underset{\sim}{z} + \underset{\sim}{y}\cdot\underset{\sim}{z}$
4. $\underset{\sim}{y}\cdot\underset{\sim}{x} = \overline{\underset{\sim}{x}\cdot\underset{\sim}{y}}$

and consequently

$$\underset{\sim}{x}\cdot(\lambda \underset{\sim}{y}) = \bar{\lambda}\underset{\sim}{x}\cdot\underset{\sim}{y},\ \underset{\sim}{x}\cdot(\underset{\sim}{y} + \underset{\sim}{z}) = \underset{\sim}{x}\cdot\underset{\sim}{y} + \underset{\sim}{x}\cdot\underset{\sim}{z}.$$

A linear operator $T:X \to X$ is such that

$$T(\underset{\sim}{x} + \lambda \underset{\sim}{y}) = T(\underset{\sim}{x}) + \lambda T(\underset{\sim}{y})$$

T is *self-adjoint* or *hermitian* (= *symmetric* in the real case) if its matrix $[t_{ij}]$ (w.r.t. some basis) is hermitian, i.e. $t_{ji} = \overline{t_{ij}}$, or equivalently,

$$T(\underset{\sim}{x})\cdot\underset{\sim}{y} = \underset{\sim}{x}\cdot T(\underset{\sim}{y}) \qquad \text{(for all } \underset{\sim}{x}, \underset{\sim}{y} \in X).$$

The eigenvalues of a self-adjoint operator (i.e. values of λ for which there exists $\underset{\sim}{x} \neq 0$ with $T(\underset{\sim}{x}) = \lambda\underset{\sim}{x}$) are real.

If λ_1, λ_2 are two distinct eigenvalues of the self-adjoint operator T and $\underset{\sim}{e}_1$, $\underset{\sim}{e}_2$ corresponding eigenvectors (i.e. $T(\underset{\sim}{e}_j) = \lambda_j\underset{\sim}{e}_j$, $j = 1,2$), then $\underset{\sim}{e}_1$ and $\underset{\sim}{e}_2$ are orthogonal i.e. $\underset{\sim}{e}_1\cdot\underset{\sim}{e}_2 = 0$.

Definition of the Problem (see earlier examples)

We now investigate Boundary value problems of the form:

Find those values of λ (eigenvalues) for which there are non-trivial solutions (corresponding eigen*functions*) to the second order equation

$$a(x)y'' + b(x)y' + c(x)y = -\lambda r(x)y$$

satisfying prescribed boundary conditions on the closed interval

$$[a, b] = \{x \in R: a \leqslant x \leqslant b\}.^{*}$$

We will only consider boundary conditions of the following three types

1. Separated Linear Homogeneous Boundary Conditions.
   $\alpha_1 y(0) + \beta_1 y'(0) = 0$
   $\alpha_2 y(1) + \beta_2 y'(1) = 0 \qquad (\alpha_1, \alpha_2, \beta_1, \beta_2$ constants)
   which include the simplest boundary conditions
   $y(0) = y(1) = 0.$

2. General Linear Homogeneous Boundary Conditions.
   $\alpha_1 y(0) + \beta_1 y'(0) + \gamma_1 y(1) + \delta_1 y'(1) = 0$
   $\alpha_2 y(0) + \beta_2 y'(0) + \gamma_2 y(1) + \delta_2 y'(1) = 0$
   (Note 1 is a special case of 2, in which the conditions at 0 are separated from the conditions at 1).

---

* Except when otherwise stated we will take a = 0 and b = 1, indeed unless either $a = -\infty$ or $b = \infty$ under the change of variable $x = \dfrac{t - a}{b - a}$ the interval [a, b] becomes [0, 1], and so it is sufficient to consider problems defined on [0, 1].

3. General Homogeneous Boundary Conditions.

Any conditions on y, y' at 0 and 1 which ensure that for some suitable function p(x),

$$\Big[p(x)(y_1'(x)y_2(x) - y_1(x)y_2'(x))\Big]_{x=0}^{x=1}$$

$$= p(1)(y_1'(1)y_2(1) - y_1(1)y_2'(1)) - p(0)(y_1'(0)y_2(0) - y_1(0)y_2'(0)) = 0$$

where y_1, y_2 are eigenfunctions of the equation, possibly for different values of λ.

Note. If the boundary conditions (1) hold we have

$$\frac{y_1(0)}{y_1'(0)} = \frac{\beta_1}{\alpha_1} = \frac{y_2(0)}{y_2'(0)}$$

or
$$y_1(0)y_2'(0) - y_1'(0)y_2(0) = 0$$

and similarly

$$y_1(1)y_2'(1) - y_1'(1)y_2(1) = 0$$

for any two solutions y_1, y_2, and so

(1) implies a special case (3).

The other case of interest, where (3) applies is when either p(0) or p(1) = 0 in which case a sufficient condition at 0 (or 1) is that y and y' assume finite values (i.e. are bounded).

Although the character of the boundary value problem is largely dependent on the particular boundary conditions prescribed and in the absence of any would be ill-posed, as we shall now see, the boundary conditions can be placed in the background (at least for the purpose of general theory).

Operator Formulation

Since, a priori the eigenfunctions of a given boundary value problem must be twice differentiable and must satisfy the boundary conditions we need never look beyond the *set H of twice differentiable functions which satisfy the boundary conditions.*

Note. (1) H will vary from problem to problem as we change the boundary conditions, and

(2) There is no assumption that the elements of H satisfy the differential equation. H is determined entirely from the boundary conditions.

For convenience we take the elements of H to be complex valued functions i.e. functions of the form

$$f(x) = u(x) + iv(x).$$

LEMMA. *The set* H *of twice differentiable functions satisfying boundary conditions of the type (1), (2) or (3) is a vector space under point-wise defined addition and scalar multiplication.*

Proof. Recall: By f + g we mean the function defined by $(f + g)(x) = f(x) + g(x)$, and by λf the function for which $(\lambda f)(x) = \lambda f(x)$. Since under these operations the set of all functions with domain [0, 1] is a vector space (see linear algebra) it suffices to show H is a subspace, i.e. $f, g \in H$ implies $f + \lambda g \in H$.
From elementary calculus $f + \lambda g$ is twice differentiable if f and g are. Indeed $(f + \lambda g)'' = f'' + \lambda g''$.
Now for boundary conditions of the type (1) and $f, g \in H$ we have

$$\begin{aligned}
&\alpha_1(f(0) + \lambda g(0)) + \beta_1(f'(0) + \lambda g'(0)) \\
&= (\alpha_1 f(0) + \beta_1 f'(0)) + \lambda(\alpha_1 g(0) + \beta_1 g'(0)) \\
&= \qquad 0 \qquad + \qquad 0
\end{aligned}$$

as both f and g satisfy the boundary conditions.
A similar argument applies at x = 1 and so we conclude that $f + \lambda g \in H$.
For boundary conditions of type (2) and (3) analogous arguments (give them) establish the result.

■

It was largely to ensure the truth of the above lemma that we restricted our attention to the three particular types of boundary conditions. Had we for example allowed nonhomogeneous boundary conditions,
e.g. $y(0) = a \neq 0$ and $y(1) = b$,
then, for f and g satisfying the boundary conditions we have

$$(f + g)(0) = f(0) + g(0) = 2a \neq a$$

and so their sum f + g does not satisfy the boundary conditions.

We now define an operator L on H as follows. L maps $f \in H$ to the new function $af'' + bf' + cf$ (a, b and c are the given functions appearing in the differential equation)
i.e. $L(f)(x) = a(x)f''(x) + b(x)f'(x) + c(x)f(x)$ (all $x \in [0, 1]$).

EXAMPLE. For the boundary value problem

$$(1 - x^2)y'' - 2xy' = -\lambda y$$

with $y(0) = 0$, $y(1)$ finite
we have associated the Legendre operator L such that

$$L(f)(x) = (1 - x^2)f''(x) - 2xf'(x).$$

So, if $f(x) = x^3$ (clearly a member of H in this case) we have

$$\begin{aligned} L(f)(x) &= (1 - x^2)6x - 6x^2 \\ &= 6x - 6x^2 - 6x^3. \end{aligned}$$

Similarly $$L(\exp)(x) = (1 - 2x - x^2)e^x.$$

LEMMA. L, *as defined above, is a linear operator on* H.

Proof. This follows trivially, since differentiation is a linear operator. Thus, for $f, g \in H$ we have

$$\begin{aligned} L(f + \lambda g) &= a(f + \lambda g)'' + b(f + \lambda g)' + c(f + \lambda g) \\ &= a(f'' + \lambda g'') + b(f' + \lambda g') + c(f + \lambda g) \\ &= af'' + bf' + cf + \lambda(ag'' + bg' + cg) \\ &= L(f) + \lambda L(g) \end{aligned}$$

(Because of this the differential equation

$L(y) = ay'' + by' + cy = 0$ is often referred to as 'linear'.)

In terms of the vector space H and linear operator L, our boundary value problem may be restated as -

Find values of λ for which there is a non-trivial element y of H such that

$$L(y) = -\lambda r(x)y.$$

Except for the factor $-r(x)$ on the R.H.S. this should be highly suggestive of the eigenvalue/eigenvector problem of linear algebra, hence the parallel terminology (eigenvalue, eigenfunctions).

Since symmetric operators on finite dimensional vector spaces have the richest eigenvalue-theory, we attempt to develop an analogous theory for a second order linear differential operator L on H.

An inner-product for H

Since H is in general an "infinite dimensional" vector space, it is not possible to represent L by a matrix, as linear transformations of finite dimensional vector spaces can be. Thus if we are going to try and specify what a self-adjoint (symmetric) operator on H might be, we must look for a definition other than the usual matrix one of finite dimensional linear algebra.

A suggestion comes from the scalar product characterization of self-adjoint linear transformations viz.

The linear transformation T is self-adjoint if and only if $\underset{\sim}{x}\cdot T(\underset{\sim}{y}) = T(\underset{\sim}{x})\cdot\underset{\sim}{y}$ for all $\underset{\sim}{x}$ and $\underset{\sim}{y}$.

In order to use this as a definition in our case, we must first however, define a scalar product on H, and what could one mean by the "dot" product of two functions?

A clue is to be found by examining the definition of dot product in the finite case.

For $\underset{\sim}{x} = (x_1, x_2, x_3, \ldots, x_n)$

$\underset{\sim}{y} = (y_1, y_2, y_3, \ldots, y_n)$ (x_j, y_j complex numbers)

$$\underset{\sim}{x} \cdot \underset{\sim}{y} \overset{\text{defn}}{=\!=} x_1\overline{y_1} + x_2\overline{y_2} + x_3\overline{y_3} + \ldots + x_n\overline{y_n}$$

$$= \sum_{j=1}^{n} x_j\overline{y_j}$$

Now, the vector $\underset{\sim}{x}$ may be regarded as the "baby" function, whose domain $\{1,2,3,\ldots,n\}$ is mapped into the complex numbers according to

$$\begin{aligned} 1 &\longmapsto x_1 \\ 2 &\longmapsto x_2 \\ 3 &\longmapsto x_3 \\ &\vdots \\ n &\longmapsto x_n \end{aligned}$$

which we summarize by writing $\underset{\sim}{x}: j \mapsto x_j$. Similarly $\underset{\sim}{y}$ is equivalent to the function $\underset{\sim}{y}: j \mapsto y_j$.

From this point of view the individual terms, $x_1\overline{y_1}$, $x_2\overline{y_2}$, ..., of the scalar product $\underset{\sim}{x} \cdot \underset{\sim}{y}$ represent the value of $\underset{\sim}{x}$ at each domain point j (x_j) multiplied by the conjugate of the value of $\underset{\sim}{y}$ at the same domain point. The scalar product is obtained by summing these products over all possible domain points.

In the case of interest to us, our vectors are again complex valued functions defined on the domain [0, 1]. For f, g $\in$ H the corresponding point value product is $f(x)\,\overline{g(x)}$ and the "sum" over all such values corresponds to

$$\int_0^1 f(x)\overline{g(x)}dx$$

We are therefore led to define a <u>scalar product</u>* in H by

$$f \cdot g = \int_0^1 f(x)\overline{g(x)}dx \quad \text{for all } f, g \in H.$$

---

* In more advanced works, scalar products are usually termed inner-products and denoted by (f, g) instead of f·g.

It is readily verified that "." satisfies all the axioms of a scalar product.

1. $f \cdot f = 0$ if and only if $f = 0$
2. $(\lambda f) \cdot g = \lambda(f \cdot g)$
3. $(f + g) \cdot h = f \cdot h + g \cdot h$
4. $g \cdot f = \overline{f \cdot g}$

For example: $f \cdot f = 0 \Leftrightarrow \int_0^1 |f(x)|^2 dx = 0$

$\Leftrightarrow |f(x)|^2 = 0$ for all x (f is twice differentiable and so continuous)

$\Leftrightarrow f = 0$

$$g \cdot f = \int_0^1 g(x)\overline{f(x)}dx = \int_0^1 \overline{\overline{g(x)}f(x)}dx$$

$$= \overline{\int_0^1 f(x)\overline{g(x)}dx} = \overline{f \cdot g} \ .$$

(Proofs of the remaining two are left to you.)

From these we also have the useful identities:

$$f \cdot (g + h) = \overline{(g + h) \cdot f} = \overline{g \cdot f + h \cdot f}$$

$$= \overline{g \cdot f} + \overline{h \cdot f} = f \cdot g + f \cdot h$$

and

$$f \cdot (\lambda g) = \overline{(\lambda g) \cdot f} = \overline{\lambda (g \cdot f)}$$

$$= \overline{\lambda}(\overline{g \cdot f}) = \overline{\lambda}(f \cdot g) \ .$$

EXAMPLE: If H contains $f(x) = x + ix^2$ and $g(x) = x^3 + ix^5$ we have

$$f \cdot g = \int_0^1 (x + ix^2)\overline{(x^3 + ix^5)}dx$$

$$= \int_0^1 (x + ix^2)(x^3 - ix^5)dx$$

$$= \int_0^1 (x^4 + x^7) + i(x^5 - x^6)dx$$

$$= \int_0^1 (x^4 + x^7)dx + i \int_0^1 (x^5 - x^6)dx$$

$$= \frac{1}{5} + \frac{1}{8} + i\left(\frac{1}{6} - \frac{1}{7}\right)$$

$$= \frac{13}{40} + {}^{i}/42.$$

Self-adjoint (Hermitian) operators

Having established a scalar product in H we now define a (linear) operator L on H to be *self-adjoint* if

$$(Lf)\cdot g = f\cdot(Lg) \text{ for all } f, g \in H.$$

(i.e. if $\int_0^1 L(f)(x)\,\overline{g(x)}\,dx = \int_0^1 f(x)\,\overline{L(g)(x)}\,dx$)

EXAMPLES. (1) The simplest type of operator we define on H is multiplication by a fixed function.

i.e. $$M_{r(x)}(f)(x) = r(x)f(x)$$

E.g. $$M_{e^{-x}}(\sin)(x) = e^{-x}\sin x.$$

Provided $r(x)$ is a real valued function, the operator $M_{r(x)}$ is self-adjoint on H. $\overline{r(x)} = r(x)$, so

$$(M_{r(x)}f)\cdot g = \int_0^1 r(x)f(x)\overline{g(x)}dx = \int_0^1 f(x)\,\overline{r(x)g(x)}dx$$

$$= f\cdot M_{r(x)}(g).)$$

(2) Let H be the set of twice differentiable functions vanishing at 0 and 1 i.e. $f \in H \Leftrightarrow f''$ exists on $[0, 1]$ and $f(0) = f(1) = 0$.
The operator $D^2 = \frac{d^2}{dx^2}$ is self-adjoint on H.

Proof. For any $f, g \in H$ we have

$$(D^2f)\cdot g = \int_0^1 f''(x)\overline{g(x)}dx$$

$$= \left[f'(x)\overline{g(x)}\right]_0^1 - \int_0^1 f'(x)\,\overline{g'(x)}dx \qquad \text{(integration by parts)}$$

$$= -\int_0^1 f'(x)\,\overline{g'(x)}dx \qquad (\text{as } g(0) = g(1) = 0)$$

$$= -[f(x)\overline{g'(x)}]_0^1 + \int_0^1 f(x)\overline{g''(x)}dx$$

$$= \int_0^1 f(x)\overline{g''(x)}dx \qquad \text{(as } f(0) = f(1) = 0)$$

$$= f\cdot(D^2 g).$$

The boundary value problem:

$$a(x)y'' + b(x)y' + c(x)y = -\lambda r(x)y \qquad (r(x), \text{ real})$$

with prescribed boundary conditions on [0, 1], is termed a self-adjoint problem if the operator L defined by

$$L(f) = af'' + bf' + cf$$

is self-adjoint on the vector space H of all complex valued twice differentiable functions satisfying the prescribed boundary conditions.

Thus from the above example the problem

$$y'' = -\lambda y$$
$$y(0) = y(1) = 0$$

is a self-adjoint boundary value problem.

One of the most difficult tasks in any application of Boundary value problems is establishing that the appropriate operators are self-adjoint. Fortunately many of the commonly encountered problems are covered by the following result:

THEOREM (Liouville) *A boundary value problem of the form*

$$\frac{d}{dx}\left(p(x)\frac{dy}{dx}\right) - q(x)y = -\lambda r(x)y$$

with the boundary conditions

$$\left[p(x)(y_1'(x)y_2(x) - y_1(x)y_2'(x))\right]_{x=0}^{x=1} = 0,$$

where y_1*,* y_2 *are any solutions of the equation for some values of* λ*,* [of type (3) including type (1)] *is a self-adjoint problem.*

Proof. Here H is the set of twice differentiable functions satisfying the boundary conditions and L is the operator defined by

$$L(f) = (pf')' - qf,$$

whence, for any f, g $\in$ H we have

$$L(f)\cdot g - f\cdot L(g) = \int_0^1 (pf')'\,\overline{g} - qf\overline{g} - \int_0^1 f\overline{(pg')'} - qf\overline{g}$$

$$= \int_0^1 (pf')'\,\overline{g} - f\overline{(pg')'} \qquad \text{(as } q = \overline{q}\text{)}$$

$$= [pf'\overline{g} - fp\overline{g'}]_0^1 - 0 \qquad \text{(integration by parts)}$$

$= 0$ by the boundary conditions satisfied by f and g. ∎

Boundary value problems of the form treated in the above theorem are known as <u>Sturm-Liouville problems</u>. Many commonly occurring boundary value problems are reducible to Sturm-Liouville problems.

Generally, the second order linear differential equation

$$a(x)y'' + b(x)y' + c(x)y = \lambda r(x)y$$

may be converted to the form

$$\frac{d}{dx}\left(p(x)\frac{dy}{dx}\right) - q(x)y = \lambda\rho(x)y$$

by means of an integrating factor.

Multiplying throughout by $\mu(x)$ gives

$$\mu ay'' + \mu by' + \mu cy = \lambda\mu ry$$

which will be of the form

$$(\mu ay')' + \mu cy = \lambda\mu ry$$

provided we select μ so that

$$(\mu a)' = \mu b$$

<u>i.e.</u> $$\mu' a + \mu a' = \mu b$$

or $$\frac{\mu'}{\mu} = \frac{b}{a} - \frac{a'}{a}$$

<u>i.e.</u> $$\ln \mu = \int^x \frac{b}{a} - \ln a$$

So $$\mu(x) = \frac{1}{a(x)}\, e^{\int^x \frac{b(t)}{a(t)}\,dt}.$$

EXERCISE

Show that by this means, the equations listed below may be reduced to the form shown, and so define self-adjoint boundary value problems for the specified domain and boundary conditions.

Equation	Reduced Form	Boundary Conditions
Bessel Equation $x^2y'' + xy' - n^2y = -\lambda x^2y$	$(xy')' - \frac{n^2}{x} y = -\lambda xy$	y finite at 0, $y(1) = 0$
Legendre Equation $(1 - x^2)y'' - 2xy' = -\lambda y$	$((1 - x^2)y')' = -\lambda y$	y finite at $x = \pm 1$
Laguerre Equation $xy'' + (1 - x)y' = -\lambda y$	$(xe^{-x} y')' = -\lambda e^{-x}y$	y finite at 0, and $\left\lvert \lim_{x\to\infty} y(x)\right\rvert < \infty$
Hermite Equation $y'' - xy' = -\lambda y$	$(e^{-x^2/2} y')' = -\lambda e^{-x^2/2} y$	$\left\lvert \lim_{x\to\pm\infty} y(x)\right\rvert < \infty$
Tchebychev Equation $(1 - x^2)y'' - xy' = -\lambda y$	$(\sqrt{1 - x^2}\, y')' = -\frac{\lambda}{\sqrt{1 - x^2}} y$	y finite at $x = \pm 1$

Theory of Self-adjoint Boundary Value Problems

Let H be the vector space of twice differentiable, complex valued functions satisfying prescribed boundary conditions and L a self-adjoint (second order) linear differential operator on H.

We then have the following *results for the self-adjoint Boundary value problem* -

$$L(y) = \lambda r(x)y, \quad y \in H,$$

where $$r(x) > 0$$

(with the possible exception of a finite number of points).

THEOREM. *The eigenvalues are real.*

Proof. Let $y = \phi(x)$ be an eigenfunction corresponding to the eigenvalue λ_0

i.e. $$\phi \in H \text{ and } L(\phi) = \lambda_0 r\phi.$$

Then,

$$\begin{aligned} \lambda_0(r\phi)\cdot\phi &= (\lambda_0 r\phi)\cdot\phi \\ &= (L(\phi))\cdot\phi \\ &= \phi\cdot L(\phi) \qquad \text{(L self-adjoint)} \\ = \phi\cdot(\lambda_0 r\phi) &= \overline{\lambda}_0(\phi\cdot(r\phi)) \\ &= \overline{\lambda}_0(r\phi)\cdot\phi \qquad \text{(as multiplication by r} \end{aligned}$$

defines a self-adjoint operator on H, see p.22).

Thus $$(\lambda_0 - \overline{\lambda}_0)(r\phi)\cdot\phi = 0,$$

now $$(r\phi\cdot\phi) = \int_0^1 r|\phi|^2 > 0 \quad \text{(by the assumption on r)}$$

and so $$\lambda_0 = \overline{\lambda}_0 \text{ or } \lambda_0 \text{ is real.}$$

■

COROLLARY. *The eigenfunctions may always be chosen to be real valued.*

Proof. If $\phi = u + iv$ is a eigenfunction corresponding to the (real) eigenvalue λ_0 we have

$$L(u + iv) = L(u) + iL(v) = \lambda_0 r(u + iv) = \lambda_0 ru + i\lambda_0 rv$$

and so equating real and imaginary parts

$$L(u) = \lambda_0 ru, \; L(v) = \lambda_0 rv,$$

i.e. the real valued functions u and v are also eigenfunctions corresponding to λ_0.

■

THEOREM (Orthogonality)

If ϕ_1, ϕ_2 are (real valued) eigenfunctions corresponding to distinct eigenvalues λ_1, λ_2 respectively, then

$$(r\phi_1)\cdot\phi_2 = 0 \quad (\text{or } \int_0^1 r(x)\phi_1(x)\phi_2(x)dx = 0)$$

and we say ϕ_1, ϕ_2 are orthogonal with respect to the weight function $r(x)$.

Proof.

$$\begin{aligned} \lambda_1(r\phi_1)\cdot\phi_2 &= (\lambda_1 r\phi_1)\cdot\phi_2 \\ &= L(\phi_1)\cdot\phi_2 \\ &= \phi_1\cdot L(\phi_2) \qquad \text{(L self-adjoint)} \\ &= \phi_1\cdot(\lambda_2 r\phi_2) \end{aligned}$$

$$= \overline{\lambda}_2\phi_1\cdot(r\phi_2)$$
$$= \lambda_2(r\phi_1)\cdot\phi_2$$

(as λ_2 is real, and r defines a self-adjoint operator on H).

Thus $(\lambda_1 - \lambda_2)\ (r\phi_1)\cdot\phi_2 = 0$ and since $\lambda_1 \neq \lambda_2$ $(r\phi_1)\cdot\phi_2 = 0$, as required. ∎

EXAMPLE.

The most trivial boundary value problem

$$y'' = -\lambda y$$

$$y(0) = y(1) = 0$$

is readily seen to be self-adjoint. It can also be solved explicitly.

The eigenvalues are $\lambda = \pi^2, \quad 4\pi^2, \quad \ldots\ldots, \quad n^2\pi^2, \quad \ldots\ldots$

with corresponding eigenfunctions $\sin \pi x, \quad \sin 2\pi x, \ldots\ldots, \quad \sin n\pi x, \ldots\ldots$

Here $r(x) \equiv 1$ and so, from the above theorem, we have the orthogonality relationship

$$\int_0^1 \sin n\pi x \sin m\pi x \, dx = 0, \text{ for } m \neq n,$$

basic to the construction of (odd) Fourier Series.

<u>Further results and discussion</u>

On finite dimensional vector spaces the eigenvectors of certain self-adjoint linear transformations form a (orthogonal) basis for the space. When this happens it is extremely useful.

Similarly, for certain (but <u>not</u> all) self-adjoint boundary value problems

$$L(y) = \lambda r y, \ y \in H,$$

the eigenfunctions form a basis for H (frequently for some vector space $H' \supseteq H$) in the sense that any $f \in H$ may be "expanded" as

$$f(x) = \sum_{n=1}^{\infty} a_n \ \phi_n(x)$$

where $\phi_1(x), \phi_2(x), \ldots, \phi_n(x), \ldots$ are the eigenfunctions corresponding to the countable set of eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n, \ldots$.

Fourier series are a special instance of this.

It can be shown that any self-adjoint problem has only a countable number of eigenvalues and that most of the commonly occurring boundary value problems (for example, those listed previously) have a set of eigenfunctions with this property (Hilbert-Schmidt Theorem).

In case f can be so expanded, the coefficients, a_n, are determined by the orthogonality relationship.

If
$$f = \sum_{n=1}^{\infty} a_n\phi_n$$

then
$$(rf)\cdot\phi_m = \left(r \sum_{n=1}^{\infty} a_n\phi_n\right)\cdot\phi_m = \sum_{n=1}^{\infty} ((ra_n\phi_n)\cdot\phi_m)$$
$$= \sum_{n=1}^{\infty} a_n((r\phi_n)\cdot\phi_m)$$
$$= a_m(r\phi_m\cdot\phi_m) \text{ as } (r\phi_n)\cdot\phi_m = 0$$
$$\text{unless} \quad n = m$$

Whence
$$a_m = \frac{(rf)\cdot\phi_m}{(r\phi_m)\cdot\phi_m} \qquad (m = 1,2,3,\ldots)$$

or
$$a_m = \frac{\int_0^1 r(x)f(x)\phi_m(x)dx}{\int_0^1 r(x)\phi_m^2(x)dx}$$

c.f. Fourier sine coefficients
$$\nu_n = \frac{\int_0^1 f(x)\sin m\pi x\, dx}{\int_0^1 \sin^2 m\pi x\, dx} .$$

Application to the non-homogeneous problem.

Consider $L[y] = \mu r(x)y + f(x)$ with boundary conditions, with respect to which L is self-adjoint. Let ϕ_n be an eigenfunction corresponding to the eigenvalue λ_n of the corresponding homogeneous problem $L[y] = \lambda r(x)y$ under the same boundary conditions.

Assume that f and r are such that f/r has the expansion $f/r = \sum_{k=1}^{\infty} a_k\,\phi_k$.

We seek a solution y which can be expressed as $y = \sum_{k=1}^{\infty} b_k\,\phi_k$, substituting this into the equation we have

$$L[y] = \sum_{k=1}^{\infty} b_k\, L[\phi_k] = \sum_{k=1}^{\infty} b_k\,\lambda_k\, r\phi_k$$
$$= \mu r \sum_{k=1}^{\infty} b_k\,\phi_k + f$$

Dividing by r and rearranging we obtain

$$\sum_{k=1}^{\infty}(\lambda_k - \mu)\, b_k\, \phi_k = {}^f/r = \sum_{k=1}^{\infty} a_k\, \phi_k.$$

So using the orthogonality of the ϕ_n we see that

$$(\lambda_k - \mu)b_k = a_k.$$

So provided $\mu \neq \lambda_n$ for any n, we have $b_k = \dfrac{a_k}{\lambda_k - \mu}$ and can write down the solution as

$$y = \sum_{n=1}^{\infty} \frac{a_n}{\lambda_n - \mu}\phi_n \quad \text{where } a_n = (f \cdot \phi_n)/(r\phi_n)\cdot\phi_n$$

Thus, by first solving the corresponding homogeneous problem, and then calculating the constants a_n, we can write down a series solution for the non-homogeneous problem.

<u>The Existence Problem</u>

Although we have established results about the eigenvalues and eigenfunctions of a boundary problem, we as yet have no guarantee that for a given problem there are any eigenvalues. Indeed some boundary value problems do not have satisfactory solutions.

We now investigate this situation for the relatively simple type of boundary value problem

$$\left.\begin{aligned} L(y) = y'' + a(x)y' + b(x)y = \lambda y \\ y(0) = y(1) = 0 \end{aligned}\right\} \qquad (1)$$

where b is bounded and a has a bounded derivative on [0, 1].

In this case, writing y = uv where $u = e^{-\frac{1}{2}\int^x a}$ we have, by the arguments used on p.8, for some function I,

$$L[y] = L[uv] = [v'' + I(x)v]u = \lambda uv$$

and so since $u(x) \neq 0$ for any $x \in [0, 1]$ we have

$$\left.\begin{aligned} v'' + I(x)v = \lambda v \\ v(0) = v(1) = 0. \end{aligned}\right\} \qquad (1^*)$$

while

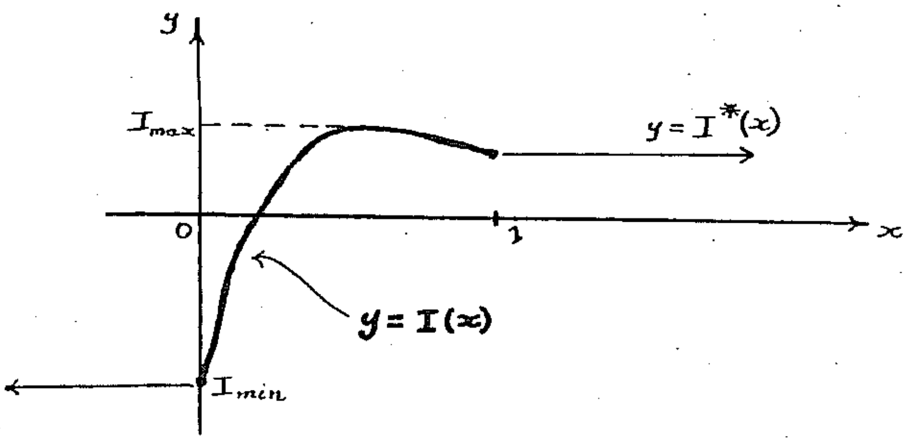
Further, if λ_0 is an eigenvalue of (1*) with corresponding eigenfunction v_0, then $y_0 \equiv uv_0$ is an eigenfunction of (1) and λ_0 is the corresponding eigenvalue. It is equally clear that the converse is also true.

Thus (1) and (1*) have the same eigenvalues and have eigenfunctions related by a factor u.

It is therefore sufficient to consider the boundary value problem (1*) which is clearly a Sturm-Liouville problem and so a self-adjoint problem.

Let I_{min} = minimum $\{I(x): 0 \leqslant x \leqslant 1\}$ (strictly we should use 'greatest lower bound' or 'infinum' in place of minimum) and I_{max} = maximum $\{I(x): 0 \leqslant x \leqslant 1\}$.

$$\text{Set } I^*(x) = \begin{cases} I(x) \text{ for } 0 \leqslant x \leqslant 1 \\ I(0) \text{ for } x < 0 \\ I(1) \text{ for } x > 1 \end{cases}$$



and for any λ let $v_\lambda(x)$ be a solution of the initial value problem

$$f'' + (I^*(x) - \lambda)v = 0, \; v(0) = 0 \; (v'(0) \neq 0)$$

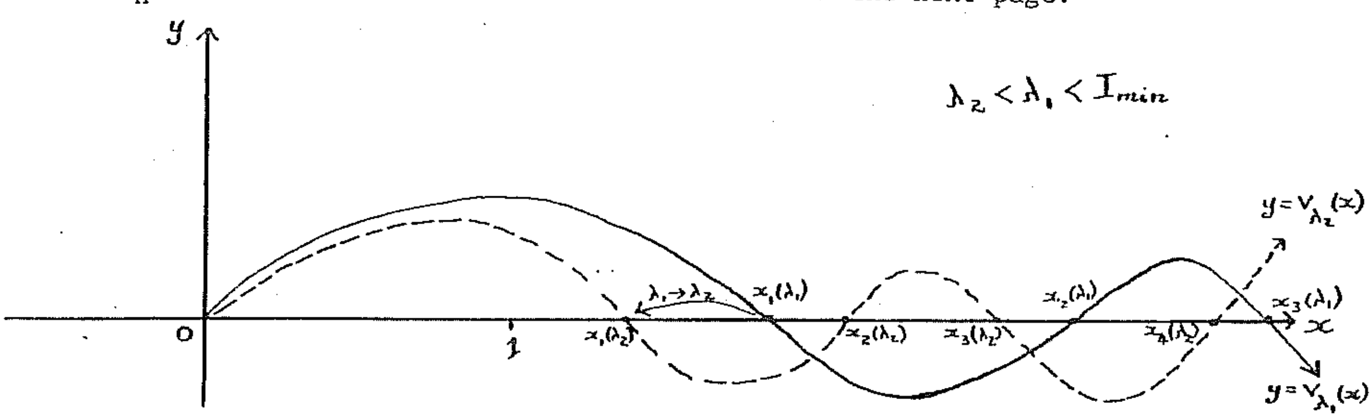
(that such a v may always be found follows from the 'existence theorem' for initial value problems - see B. diP. p.88).

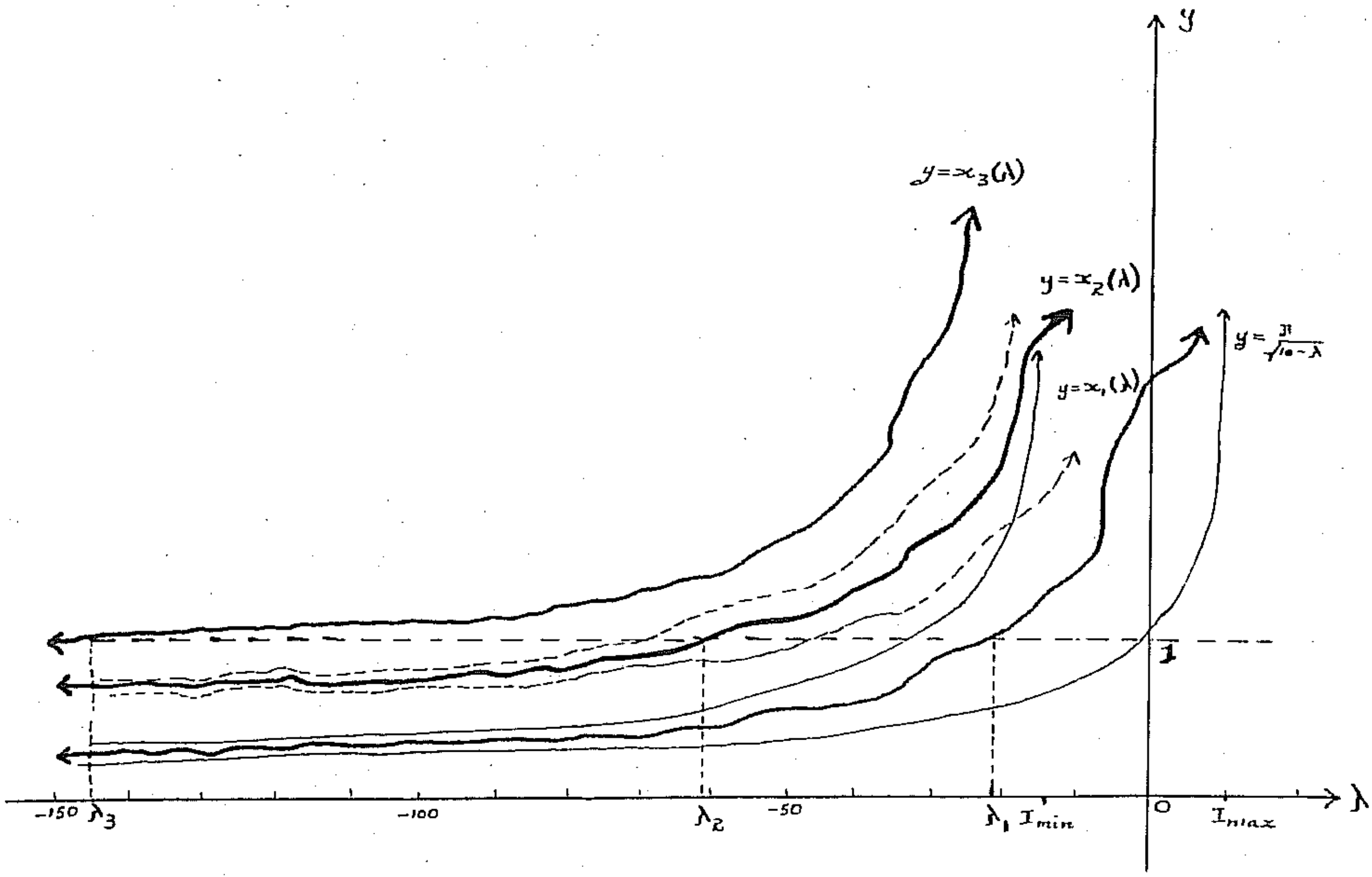
The following argument is motivated by the observation that v_{λ_0} is an eigenfunction of (1*) which corresponds to the eigenvalue λ_0 if v_{λ_0} has a zero at 1.

Let the strictly positive zeros of v_λ occur at the points $x_1(\lambda) < x_2(\lambda) < \ldots < x_n(\lambda) < \ldots$, where $x_n(\lambda) \to \infty$ as $n \to \infty$. (For $\lambda < I_{min}$, the existence of such a sequence of zeros follows from the corollary to Sturms theorem since, then $I^*(x) - \lambda \geqslant I_{min} - \lambda > 0$.)

For any n, $x_n(\lambda)$ will vary as we vary λ. We will <u>assume</u> that $x_n(\lambda)$ is a continuous function of λ. (This may be proved using arguments similar to those appearing in the proof of the existence theorem, and is a special case of the "continuous dependence on parameters theorem" - see Sachez p.136).

We now show that the graphs, in the y-λ plane, of the family of functions $y = x_n(\lambda)$, n = 1,2,3,... are of the form shown on the next page.





Since 0 is a zero of v_λ for all λ (by assumption) the spacing of zeros theorem gives

$$\pi/\sqrt{I_{max} - \lambda} < x_1(\lambda) < \pi/\sqrt{I_{min} - \lambda}$$

while for $n \geqslant 1$

$$\pi/\sqrt{I_{max} - \lambda} < x_{n+1}(\lambda) - x_n(\lambda) < \pi/\sqrt{I_{min} - \lambda}.$$

So $y = x_1(\lambda)$ is a continuous curve lying between the two curves $y = \pi/\sqrt{I_{max} - \lambda}$ and $y = \pi/\sqrt{I_{min} - \lambda}$ both of which tend to zero as $\lambda \to -\infty$, while the first tends to ∞ as $\lambda \to I_{max}-$ and the second as $\lambda \to I_{min}-$. So we conclude that $x_1(\lambda) \to 0$ as $\lambda \to -\infty$ and assumes arbitrarily large values for values of λ between I_{min} and I_{max}. Hence $y = x_1(\lambda)$ has a graph of the form illustrated, and from the intermediate value theorem for continuous functions we see that there is a λ_1 for which $x_1(\lambda_1) = 1$. Thus, from our earlier remark, λ_1 is an eigenvalue of (1*) with v_{λ_1} a corresponding eigenfunction. Further since $x_1(\lambda_1) = 1$ is the location of the smallest strictly positive zero of v_{λ_1} (by definition of $x_1(\lambda)$) we see that the eigenfunction v_{λ_1} corresponding to λ_1 has no zeros in the interval between 0 and 1.

EXERCISE. From the above arguments obtain bounds between which the eigenvalue λ_1 must lie.

Similarly $y = x_2(\lambda)$ is a continuous curve which, from above, lies between

$$y = x_1(\lambda) + \pi/\sqrt{I_{max} - \lambda} \text{ and } y = x_1(\lambda) + \pi/\sqrt{I_{min} - \lambda}.$$

Again both of these curves tend to zero as $\lambda \to -\infty$ and assume arbitrarily large values for λ sufficiently large. So $y = x_2(\lambda)$ has the form illustrated and there exists a value $\lambda_2 < \lambda_1$ at which $x_2(\lambda_2) = 1$.

Whence, λ_2 is an eigenvalue of (1*) with v_{λ_2} a corresponding eigenfunction which, from the definition of the $x_n(\lambda)$'s, has a zero at $x_1(\lambda_2)$ and so has one zero in the interval between 0 and 1.

Continuing in this way we see, by an inductive argument, that the graphs of $y = x_n(\lambda)$ $(n = 3,4,\ldots)$ are as illustrated, and that (1*) has eigenvalues $\lambda_1 > \lambda_2 > \lambda_3 > \ldots > \lambda_n > \ldots$ (where λ_n is such that $x_n(\lambda_n) = 1$). That $\lambda_n \to -\infty$ as $n \to \infty$ follows from the earlier observation that $x_n(\lambda) \to \infty$ as $n \to \infty$ for each λ. Further v_{λ_n} is an eigenfunction, corresponding to the eigenvalue λ_n, which has zeros at the points $x_1(\lambda_n), x_2(\lambda_n), \ldots, x_{n-1}(\lambda_n)$ in the interval between 0 and 1, and so has n-1 zeros in the interval between 0 and 1.

We have therefore proved

THEOREM. *Let* I *be a bounded function on the interval from* 0 *to* 1, *then the boundary value problem*

$$\upsilon'' + I(x)\upsilon = \lambda\upsilon$$

$$\upsilon(0) = \upsilon(1) = 0$$

has eigenvalues

$\lambda_1 > \lambda_2 > \lambda_3 > \ldots > \lambda_n > \ldots$, *with* $\lambda_n \to -\infty$ *as* $n \to \infty$. *Further corresponding to the eigenvalue* λ_n *there is an eigenfunction* υ_{λ_n} *with precisely* n-1 *zeros between* 0 *and* 1.

PROBLEMS

1. Show that the solution of the boundary value problem

$$y'' = -\lambda y$$
$$y(0) = y(1) = 0$$

is as stated in the text.

2. Find approximately the eigenvalues and corresponding eigenfunctions for the boundary value problem

$$y'' = -\lambda y$$
$$y(0) + y'(0) = 0 = y(1)$$

3. For the boundary value problem

$$y'' = -\lambda y$$
$$y(0) - y(1) = 0 = y'(0) - y'(1)$$

(a) Find the eigenvalues and corresponding eigenfunctions;
(b) Show the problem is self-adjoint, even though the boundary conditions are not separated and so the problem is not a Sturm-Liouville problem;
(c) Observe that to each eigenvalue there corresponds two linearly independent eigenfunctions.

*4. For a Sturm-Liouville problem with separated boundary conditions (of type 1) show that, if ϕ_1 and ϕ_2 are two eigenfunctions corresponding to the one eigenvalue, then they are linearly dependent.
(Hint: Show that it is sufficient to prove

$$W(\phi_1, \phi_2)(0) = \begin{vmatrix} \phi_1(0) & \phi_2(0) \\ \phi_1'(0) & \phi_2'(0) \end{vmatrix} = 0,$$ and then do so.)

This property of Sturm-Liouville problems is an important one for more advanced work, which is not necessarily true for a general self-adjoint boundary value problem (see problem 3).

5. Find a series solution for the problem

$$y'' + \lambda y = x$$
$$y(0) = y(1) = 0$$

in terms of the eigenfunctions of the corresponding homogeneous problem (see problem 1).

6. What is the import of our theory on the Hydrogenic ion model, considered in the introduction?

7. Observe that, if ϕ_n is an eigenfunction, of the self-adjoint boundary value problem

$$L(y) = -\lambda r y, \; y \in H,$$

corresponding to the n'th eigenvalue λ_n, then $k\phi_n$ is also an eigenfunction corresponding to λ_n, for any $k \neq 0$.
The eigenfunction, $k\phi_n$, is referred to as <u>normalized</u> if k is chosen so that $(rk\phi_n)\cdot(k\phi_n) = 1$

<u>i.e.</u> $$k = {}^{1}/(r\phi_n \cdot \phi_n)^{\frac{1}{2}}.$$

What are the normalized eigenfunctions for problem 1 above?

*8. GREEN'S FUNCTIONS (An alternative approach to boundary value problems)
Suppose $L(y) \equiv -[p(x)y']' + q(x)y = f(x)$ under

$$\left.\begin{aligned} a_1 y(0) + a_2 y'(0) = 0 \\ b_1 y(1) + b_2 y'(1) = 0 \end{aligned}\right\} \quad \ldots\ldots (*)$$

has the solution

$$y = \phi(x) = \int_0^1 G(x,u) f(u)\,du,$$ for some <u>Green's Function</u> $G(x, u)$.

Let $\phi_i(x)$ be the normalised (see problem 7) eigenfunction of $L(y) = \lambda r(x)y$ under (*) corresponding to the eigenvalue λ_i, then $y = \phi_i(x)$ is a solution of

$$L(y) = f(x) \equiv \lambda_i r(x)\phi_i(x) \text{ under } (*)$$

So $$\phi_i(x) = \int_0^1 G(x, u)\lambda_i r(u)\phi_i(u)\,du.$$

Use this to determine $G(x, u)$ by assuming

$$G(x, u) = \sum_{i=1}^{\infty} a_i(x)\phi_i(u)$$

*9. Consider the Legendre Boundary value problem

$$(1 - x^2)y'' - 2xy' = -\lambda y$$
$$y(-1),\ y(1) \text{ finite.}$$

Clearly any polynomial $y = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ satisfies the boundary conditions.

*9. (continued)

Show that for appropriate values of λ, the equation has such polynomial solutions for eigenfunctions (the orthogonal polynomials so arrived at are the "Legendre Polynomials").

(Hint: λ must have the form $m(m + 1)$, m a positive integer, for which the corresponding polynomial is of degree m. If you can do nothing more, at least find polynomial eigenfunctions for $m = 0,1,2$.)

10. On p.32 you were asked to estimate bounds for the first eigenvalue, λ_1, of the problem 1*. Using similar arguments estimate bounds between which the n'th eigenvalue λ_n must lie.

(Hint: Show $\frac{n\pi}{\sqrt{I_{max} - \lambda}} < x_n(\lambda) < \frac{n\pi}{\sqrt{I_{min} - \lambda}}$).

References:

The two books

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

and

Sánchez "Ordinary Differential Equations and Stability Theory" Freeman

have been referred to in the text. The first (B & diP) is a useful general reference.

I know of no references which parallel our treatment in a suitable way, however the following works do contain relevant material

Physical Background

Wallace "Mathematical Analysis of Physical Problems"

Oscillation Theory

Richard Bellman "Stability Theory of Differential Equations" - in particular chapter 6, section 10 (Dover)

Einar Hille "Lectures on Ordinary Differential Equations" (Addison-Wesley, 1969) gives a full account of Sturm's Oscillation Theory

Birkhoff and Rota "Ordinary differential Equations" (Blaisdell) chapter X is also highly relevant here.

Boundary Value Problems

Weinberger "Partial Differential Equations", Blaisdell, chapters IV, V and VII

Hochstadt "Differential Equations", Holt, Rinehart and Winston, chapter IV

Carrier and Pearson, "Ordinary Differential Equations", Blaisdell, Chapters VI and VIII.