ODE & CALCULUS OF VARIATIONS (MTH2C09)

STUDY MATERIAL

II SEMESTER

CORE COURSE

M.SC. MATHEMATICS

(2019 ADMISSION ONWARDS)

UNIVERSITY OF CALICUT

SCHOOL OF DISTANCE EDUCATION

CALICUT UNIVERSITY P.O, MALAPPURAM KERALA, INDIA 673 635.

190559

UNIVERSITY OF CALICUT

SCHOOL OF DISTANCE EDUCATION

STUDY MATERIAL

SECOND SEMESTER

M.Sc. MATHEMATICS (2019 ADMISSION)

CORE COURSE : MTH2C09 : ODE & CALCULUS OF VARIATIONS

Prepared by : ***Fedric Joseph,***
Assistant Professor,
St.Aloysius College,
Elthuruth, Thrissur.

Scrutinized by : ***Sini P,***
Assistant Professor,
Dept. of Mathematics,
University of Calicut.

MODULE 1

CHAPTER I

POWER SERIES SOLUTIONS AND SPECIAL FUNCTIONS

I.1 Introduction

An Algebraic function is a polynomial, a rational function or any function y=f(x) that satisfies an equation of the form

$P_n(x)y^n+P_{n-1}(x)y^{n-1}$ + ………… + $P_1(x)y+P_0(x)=0$, where each $P_i(x)$ is a polynomial.

The elementary functions consists of the algebraic functions; the elementary transcendental (or non algebraic) functions occurring in calculus i.e. the trigonometric, inverse trigonometric, exponential and logarithmic functions; and all others that can be constructed from these by adding, substracting, multiplying, dividing or forming a function of a function.

Eg: $y=\tan\left[\dfrac{xe^{1/x}+\tan^{-1}(1+x^2)}{\sin x \operatorname{Cos} 2x-\sqrt{\log x}}\right]^{1/3}$ is an elementary function.

Beyond, the elementary functions lie the higher transcendental functions, or, as they are often called the special functions. Eg: The gamma function, the Riemann Zeta function, the elliptic functions etc.

<u>Power Series</u>

An infinite series of the form

$$\sum_{n=0}^{\infty} a_n (x-x_0)^n = a_0 + a_1 (x-x_0) + a_2(x-x_0)^2+\ldots\ldots.. \quad (1)$$

is a power series in $(x-x_0)$. Similarly,

$$\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1x + a_2x^2+\ldots\ldots.. \quad (2)$$

is called a power series in x.

Convergence

The series (2) is said to converge at a point x if the limit $\lim\limits_{m \to \infty} \sum_{n=0}^{m} a_n x^n$ exists, and in this case the sum of the series is the value of this limit. Obviously (2) always converges at the point x=0. Each series of the kind (2) in 'x' has a radius of convergence 'R', where $0 \leq R \leq \infty$, with the property that the series converges if $|x|<R$ and diverges if $|x|>R$.

It is customary to put R equal to 0 when the series converges only for x=0, and equal to ∞ when it converges for all x. So if R=0 then no x satisfies $|x|<R$, and if $R=\infty$ then no x satisfies $|x|>R$. If a series fails to converge, then it is said to diverge.

Eg: 1. The series $\sum_{n=0}^{\infty} x^n = 1+x+x^2+x^3+\ldots\ldots\ldots$

converges for $|x|<1$ and diverges for $|x|>1$. So for this series R=1.

2. The series $\sum_{n=0}^{\infty} x^n/n! = 1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\ldots\ldots\ldots\ldots$ converges for all x and so $R=\infty$.

3. The series $\sum_{n=0}^{\infty} n!x^n = 1+x+2!x^2+3!x^3+\ldots.$ diverges (fails to converge) for all $x \neq 0$ and the radius of convergence R=0.

Note:

1. If the raidus of convergence R is finite and nonzero, then it determines an interval of convergence $-R<x<R$ such that inside the interval the series converges and outside the interval it diverges. A power series may or may not converge at either endpoint of its interval of convergence.
2. Suppose that (2) converges for $|x|<R$ with R>0, and denote this sum by f(x);

$$f(x) = \sum_{n=0}^{\infty} a_n x^n = a_0+a_1x+a_2x^2+\ldots\ldots\ldots \quad (3)$$

Then f(x) is automatically continuous and has derivatives of all orders for $|x|<R$.

The series can be differentiated term wise in the sense that

$$f^{1}(x)= \sum_{n=1}^{\infty} na_n x^{n-1} = a_1 + 2a_2x + 3a_3x^2 + ... = \sum_{n=0}^{\infty} na_n x^{n-1}$$

$$f^{11}(x) = \sum_{n=1}^{\infty} n(n-1)\, a_n x^{n-2} = 2a_2+3.2a_3x + \ldots.., = \sum_{n=0}^{\infty} n(n-1)\, a_n x^{n-1}$$ and so on,

and each of the resulting series converges for $|x|<R$.

Further more (3) can be integrated term wise provided the limits of integration lie inside the interval of convergence.

If we have a second power series in x that converges to a function g(x) for $|x|<R$, so that

$$g(x) = \sum_{n=0}^{\infty} b_n x^n = b_0+b_1x + b_2x^2+\ldots.. \qquad (4)$$

then (3) and (4) can be added or subtracted term wise:

$$f(x) \pm g(x) = \sum_{n=0}^{\infty} (a_n \pm b_n)x^n = (a_0 \pm b_0)+(a_1 \pm b_1)x + \ldots\ldots\ldots..$$

(3) and (4) can be multiplied as if they were polynomials, in the sense that

$f(x)\, g(x) = \sum_{n=0}^{\infty} c_n x^n$, where

$$c_n = a_0b_n+a_1b_{n-1} + \ldots\ldots\ldots + a_nb_0 = \sum_{k=0}^{n} a_k b_{n-k} = \sum_{k=0}^{n} a_{n-k} b_k$$

3. A function f(x) with the property that a power series expansion of the form $f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$ is valid in some neighbourhood of the point x_0 is said to be analytic at x_0. The a_n are necessarily given by $a_n= \frac{f^{(n)}(x_0)}{n!}$ and $f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$ is called the Taylor series of f(x) at x_0.
    - Polynomials and the functions e^x, sin x and cos x are analytic at all points.

- If $f(x)$ and $g(x)$ are analytic at x_0, then $f(x) + g(x)$, $f(x)\, g(x)$ and $f(x) / g(x)$ [if $g(x_0) \neq 0$] are also analytic at x_0.
- If $g(x)$ is analytic at x_0 and $f(x)$ is analytic at $g(x_0)$, then $f(g(x))$ is analytic at x_0.
- The sum of a power series is analytic at all points inside the interval of convergence.
- If $f(x)$ is analytic at x_0 and $f^{-1}(x)$ is a continuous inverse, then $f(x)$ is analytic at $f(x_0)$ if $f^1(x_0) \neq 0$.

I.2 Series solutions of first order equations.

It may be recalled that many differential equations cannot be solved in terms of the elementary functions, and also that solutions for equations of this kind can be found in terms of power series. Consider the equation $y^1 = y$. (1)

We assume that this equation has a power series solution of the form

$$y=a_0 + a_1x + a_2x^2 + \dots + a_nx^n + \dots \qquad (2)$$

that converges for $|x|<R$ with $R>0$; that is, we assume that (1) has a solution that is analytic at the origin. A power series can be differentiated term by term in its interval of convergence, so

$$y^1 = a_1 + 2a_2x + 3a_3x^2 + \dots + (n+1)\, a_{n+1}x^n + \dots \qquad (3)$$

since $y^1 = y$. the series (2) and (3) must have the same coefficients:

$$a_1 = a_0,\ 2a_2 = a_1,\ 3a_3 = a_2, \dots, (n+1)\, a_{n+1} = a_n \dots$$

these equations enable us to express each a_n in terms of a_0:

$$a_1 = a_0,\quad a_2 = \frac{a_1}{2} = \frac{a_0}{2},\quad a_3 = \frac{a_3}{3} = \frac{a_0}{2\,.\,3}, \dots\dots, a_n = \frac{a_0}{n!}, \dots$$

When these coefficients are inserted in (2), we obtain our power series solution.

$$y = a_0 \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots . + \frac{x^n}{n!} + \ldots .\right) \qquad (4)$$

where no condition is imposed on a_0. It is essential to understand that so far this solution is only tentative, because we have no guarantee that (1) actually has a power series solution of the form (2). The above argument shows only that if (1) has such a solution, then that solution must be (4). However, it follows at once from the ratio test that the series in (4) converges for all x, so the term-by-term differentiation is valid and (4) really is a solution of (1). In this case we can easily recognize the series in (4) as the power series expansion of e^x, so (4) can be written as $y=a_0e^x$.

Note:

The power series expansion of a given function can be obtained as below: Find the differential equation satisfied by the function and then solve this equation by power series.

Example: consider the function $y = (1+x)^p$, (5)

where p is an arbitrary constant. It is easy to see that (5) is the indicated particular solution of the following differential equation:

$$(1 + x)y^1 = py, \quad y(0) = 1. \qquad (6)$$

As before, we assume that (6) has a power series solution

$$y = a_0 + a_1x + a_2x^2 + \ldots\ldots + a_nx^n + \ldots . \qquad (7)$$

with positive radius of convergence. It follows from this that

$$y^1 = a_1 + 2a_2x + 3a_3x^2 + \ldots + (n + 1)a_{n+1}\, x^n + \ldots . ,$$

$$xy^1 = \quad a_1x + 2a_2x^2 + \ldots + na_nx^n + \ldots ..,$$

$$py = pa_0 + pa_1x + pa_2x^2 + \ldots .+ pa_nx^n + \ldots .$$

By equation (6), the sum of the first two series must equal the third, so equating the coefficients of successive powers of x gives,

$$a_1 = pa_0,\ 2a_2 + a_1 = pa_1,\ 3a_3 + 2a_2 = pa_2,\ \ldots .$$

$$(n + 1)a_{n+1} + na_n = pa_n, \ldots.$$

The initial condition in (6) implies that $a_0 = 1$, so

$$a_1 = p,\ a_2 = \frac{a_1(p-1)}{2} = \frac{p(p-1)}{2},$$

$$a_3 = \frac{a_2(p-2)}{3} = \frac{p(p-1)(p-2)}{2\cdot 3}, \ldots,$$

$$a_n = \frac{p(p-1)(p-2)\ldots(p-n+1)}{n!}$$

With these coefficients, (7) becomes.

$$y=1+Px+\frac{p(p-1)}{2!}x^2+\frac{p(p-1)(p-2)}{3!}x^3+\ldots+\frac{p(p-1)(p-2)\ldots(p-n+1)}{n!}x^n \ \ldots. \qquad (8)$$

Since the series in (8) converges for |x| < |, (8) is the desired solution. On comparing the two solutions (5) and (8), and using the fact that (6) has only one solution, we have.

$$(1+x)^p = 1+px+\frac{p(p-1)}{2!}x^2+\ldots+\frac{p(p-1)\ldots(p-n+1)}{n!}x^n+\ldots \qquad (9)$$

for $|x|<1$. This expansion is called the binomial series.

Problem.

1. Express $\sin^{-1}x$ in the form of a power series $\Sigma a_n x^n$ by solving $y^1=(1-x^2)^{-1/2}$ in two ways. [Soln. $\sin^{-1}x = x+\sum_{n=1}^{\infty}\frac{1\cdot 3\ldots(2n-1)}{2\cdot 4\ldots(2n)}\frac{x^{2n+1}}{2n+1}$

I.3 Second order linear equations. Ordinary points.

Consider the general homogeneous second order linear equation $y^{11}+P(x)y^1+Q(x)y=0$ …………….. (1)

If both P(x) and Q(x) are analytic at a point x_0, (which means that each has a power series expansion valid in some neighborhood of x_0), then x_0 is called an ordinary point of (1). Any point that is not an ordinary point of (1) is called a singular point.

Theorem A: (Nature of solutions near ordinary points)

Let x_0 be an ordinary point of the differential equation (1) and let a_0 and a_1 be arbitrary constants. Then there exists a unique function y(x) that is analytic at x_0, is a solution of (1) in a certain neighbourhood of this point and satisfies the initial conditions $y(x_0)=a_0$ and $y^1(x_0)=a_1$. Furthermore, if the power series expansions of P(x) and Q(x) are valid on the interval $|x-x_0|<R$, $R>0$, then the power series expansion of this solution is also valid on the same interval.

Q: Find a power series solution of the differential equation $y^{11}+y=0$ (2)

Soln: The coefficient functions are P(x)=0 and Q(x)=1. These functions are analytic at all points, so we seek a solution of the form.

$$y=a_0 + a_1x + a_2x^2 + \quad + a_nx^n + \ldots.. \qquad (3)$$

Differentiating (3) yields

$$y^1 = a_1 + 2a_2x + 3a_3x^2 + \ldots. + (n+1)a_{n+1}x^n + \ldots \qquad (4)$$

and

$$y^{11} = 2a_2 + 2\cdot3a_3x + 3\cdot4a_4x^2 + \ldots + (n+1)\,(n+2)a_{n+2}x^n + \ldots \qquad (5)$$

If we substitute (5) and (3) into (2) and add the two series term by term, we get

$(2a_2 + a_0) + (2\cdot3a_3 + a_1)x + (3\cdot4a_4 + a_2)x^2 + (4\cdot5a_5 + a_3)x^3 +\ldots+ [(n+1)\,(n+2)a_{n+2} + a_n]x^n + \ldots = 0$; and equating to zero the coefficients of successive powers of x gives.

$$2a_2 + a_0 = 0, \quad 2\cdot3a_3 + a_1 = 0, \quad 3\cdot4a_4 + a_2 = 0, \quad 4\cdot5a_5 + a_3 = 0, \ldots, \quad (n+1)\,(n+2)a_{n+2} + a_n = 0, \ldots$$

By means of these equations we can express a_n in terms of a_0 or a_1 according as n is even or odd:

$$a_2 = -\frac{a_0}{2}, \quad a_3 = -\frac{a_1}{2\cdot3}, \quad a_4 = -\frac{a_2}{3\cdot4}=\frac{a_0}{2\cdot3\cdot4}, \quad a_5 = -\frac{a_3}{4\cdot5}=\frac{a_1}{2\cdot3\cdot4\cdot5},\ldots$$

with these coefficients, (3) becomes

$$y = a_0 + a_1x - \frac{a_0}{2}x^2 - \frac{a_1}{2\cdot3}x^3 + \frac{a_0}{2\cdot3\cdot4}x^4 + \frac{a_1}{2\cdot3\cdot4\cdot5}x^5 - \ldots$$

$$= a_0 \left(1-\frac{x^2}{2!}+\frac{x^4}{4!}-....\right)+a_1\left(x-\frac{x^3}{3!}+\frac{x^5}{5!}-...\right).$$

Note

In the above problem, the two series in parentheses are easily recognizable as the expansions of cos x and sin x and so the solution can be written in the form $y=a_0 \cos x + a_1 \sin x$. However most series solutions found in this way are quite impossible to identify and represent previously unknown functions.

Problems.

1. Find the general solution of $(1+x^2)\, y^{11} + 2xy^1 - 2y = 0$ in terms of power series in x. Can you express this solution by means of elementary functions.

   Soln: $y = a_0 \left(1+x^2-\frac{1}{3}x^4+\frac{1}{5}x^6-\frac{1}{7}x^8+...\right)+a_1x$ In terms of elementary functions, $y = a_0 (1+x \tan^{-1}x) + a_1x$.

2. Verify that the equation $y^{11} + y^1 - xy = 0$ has a three term recursion formula, and find its series solutions $y_1(x)$ and $y_2(x)$ such that (a) $y_1(0)=1$, $y_1'(0) = 1$ (b) $y_2(0) = 0$, $y_2^1(0)=1$.

   Soln: Recursion formula is $a_{n+2} = -\left[\frac{(n+1)a_{n+1}-a_{n-1}}{(n+1)(n+2)}\right]$

   a. $y_1(x) = 1+\frac{x^3}{2\cdot 3}-\frac{x^4}{2\cdot 3\cdot 4}+\frac{x^5}{2\cdot 3\cdot 4\cdot 5}+...;$

   b. $y_2(x) = x-\frac{x^2}{2}+\frac{x^3}{2\cdot 3}+\frac{x^4}{2\cdot 3\cdot 4}-\frac{4x^5}{2\cdot 3\cdot 4\cdot 5}+$

3. Chebyshev's equation is $(1-x^2)y^{11} - xy^1 + p^2y = 0$, where p is a constant. Find two linearly independent series solutions valid for |x|<|.

   Soln: $y_1(x) = 1-\frac{p.p}{2!}x^2+\frac{p(p-2)p(p+2)}{4!}x^4-...$

   $$y_2(x) = x-\frac{(p-1)(p+1)}{3!}x^3+\frac{(p-1)(p-3)(p+1)(p+3)}{5!}x^5...$$

I.4 Regular Singular Points

A point x_0 is a singular point of the differential equation

$$y^{11} + P(x)y^1 + Q(x)y=0 \qquad (1)$$

if one or the other (or both) of the coefficient functions P(x) and Q(x) fails to be analytic at x_0.

eg: Origin is a singular point of the differential equation $x^2y^{11}+xy^1-y=0$.

A singular point x_0 of (1) is said to be regular if the functions $(x-x_0)\,P(x)$ and $(x-x_0)^2\,Q(x)$ are analytic and irregular otherwise.

Eg: Consider the Legendre's equation $(1-x^2)y^{11} - 2xy^1 + p(p+1)\,y=0$. Rearranging in the form $y^{11}+P(x)y^1+Q(x)y=0$, the above equation becomes $y^{11} - \frac{2x}{1-x^2}y^1 + \frac{p(p+1)}{1-x^2}y = 0$, so that $P(x) = \frac{-2x}{1-x^2}$ and

$Q(x) = \frac{p(p+1)}{1-x^2}$. Clearly x=1 and x = -1 are singular points.

Now $(x-1)\,p(x) = \frac{2x}{x+1}$ and $(x-1)^2\,Q(x) = \frac{-(x-1)p(p+1)}{x+1}$ are analytic at x=1. Similarly $(x+1)\,p(x)$ and $(x+1)^2\,Q(x)$ are analytic at x=-1.

$\therefore$ x=1 and x=-1 are regular singular points of the Legendre's equation.

eg: Consider $x^2y^{11}+xy^1+(x^2-p^2)y = 0$ (2)

the Bessel's equation of order 'p', where 'p' is a non negative constant.

(2) expressed in the form $y^{11}+P(x)y^1+Q(x)y=0$, becomes $y^{11}+\frac{1}{x}y^1+\frac{x^2-p^2}{x^2}y=0$, so that $p(x) = 1/x$ and $Q(x) = \frac{x^2-p^2}{x^2}$. So P(x) and Q(x) are not analytic at x=0 and x=0 is a singular point of (2). But $x\,P(x) = 1$ and $x^2\,Q(x) = x^2 - p^2$ are analytic at x=0 and so the origin is a regular singular point of the differential equation (2).

I.5 Solutions of the differential equation near a regular singular point – Frobenius Series Method.

To simplify matters, we may assume that the singular point x_0 is located at the origin. Suppose that the origin is a regular singular point of the differential equation $y^{11} + P(x)y^1 + Q(x)y = 0$ (1)

We assume a series of the form $y = x^m (a_0 + a_1x + a_2x^2+\ldots.)$ (2)

where m may be a negative integer, a fraction or even an irrational real number, to be solution of (1). Series of the form (2) are called Frobenius series. The procedure described below for finding solutions of type (2) are called Frobenius method.

Eg: Consider the equation $2x^2y^{11}+x(2x+1)y^1-y=0$ (3)

If this is written in the more revealing form

$$y^{11} + \frac{1/2+x}{x}y^1 + \frac{-1/2}{x^2}y\ 0, \qquad (4)$$

Then we see at once that $xP(x) = \frac{1}{2} + x$ and $x^2Q(x) = -\frac{1}{2}$, so x=0 is a regular singular point. We now introduce our assumed Frobenius series solution.

$$y = x^m (a_0 + a_1x + a_2x^2 + \ldots)$$

$$= a_0x^m + a_1 x^{m+1} + a_2 x^{m+2} + \ldots., \qquad (5)$$

and its derivatives

$$y^1 = a_0mx^{m-1} + a_1 (m + 1)x^m + a_2(m + 2) x^{m+1} + \ldots$$

and

$$y^{11} = a_0m(m - 1)x^{m-2} + a_1(m + 1)mx^{m-1} + a_2 (m + 2) (m + 1) x^m + \ldots.$$

To find the coefficients in (5), we proceed in essentially the same way as in the case of an ordinary point, with the significant difference that now we must also find the appropriate value (or values) of the exponent m. When the three series above are inserted in (4), and the common factor x^{m-2} is canceled, the result is

$a_0m (m-1) + a_1(m + 1) mx + a_2 (m + 2) (m +1) x^2 + \ldots$

$$+\left(\frac{1}{2}+x\right)\left[a_0 m+a_1(m+1)x+a_2(m+2)x^2+\ldots\right]$$

$$-\frac{1}{2}\left(a_0+a_1x+a_2x^2+\ldots\right)=0.$$

By inspection, we combine corresponding powers of x and equate the coefficient of each power of x to zero. This yields the following system of equations:

$$a_0\left[m(m-1)+\frac{1}{2}m-\frac{1}{2}\right]=0,$$

$$a_1\left[(m+1)m+\frac{1}{2}(m+1)-\frac{1}{2}\right]=0, \tag{6}$$

$$a_2\left[(m+2)(m+1)+\frac{1}{2}(m+2)-\frac{1}{2}\right]+a_1(m+1)=0,$$

As we explained above, it is understood that $a_0 \neq 0$. It therefore follows from the first of these equation that

$$m(m-1)+\frac{1}{2}m-\frac{1}{2}=0. \tag{7}$$

This is called the indicial equation of the differential equation (3). Its roots are

$$m_1 = 1 \quad \text{and} \quad m_2 = -\frac{1}{2}.$$

and these are the only possible values for the exponent m in (5). For each of these values of m, we now use the remaining equations of (6) to calculate a_1, a_2, … in terms of a_0. For $m_1 = 1$, we obtain

$$a_1 = -\frac{a_0}{2\cdot 1+\frac{1}{2}\cdot 2-\frac{1}{2}}=-\frac{2}{5}a_0,$$

$$a_2 = -\frac{2a_1}{3\cdot 2+\frac{1}{2}\cdot 3-\frac{1}{2}}=-\frac{2}{7}a_1=\frac{4}{35}a_0,$$

And for $m_2 = -\frac{1}{2}$, we obtain

$$a_1 = \frac{\frac{1}{2}a_0}{\frac{1}{2}\left(-\frac{1}{2}\right)+\frac{1}{2}\cdot\frac{1}{2}-\frac{1}{2}} = -a_0,$$

$$a_2 = -\frac{\frac{1}{2}a_1}{\frac{3}{2}\cdot\frac{1}{2}+\frac{1}{2}\cdot\frac{3}{2}-\frac{1}{2}} = -\frac{1}{2}a_1 = \frac{1}{2}a_0,$$

We therefore have the following two Frobenius series solutions, in each of which we have put $a_0 = 1$:

$$y_1 = x\left(1-\frac{2}{5}x + \frac{4}{35}x^2 + \ldots\right) \tag{8}$$

$$y_2 = x^{-1/2}\left(1-x+\frac{1}{2}x^2 + \ldots\right) \tag{9}$$

These solutions are clearly independent for x>0, so the general solution of (3) on this interval is

$$y=c_1x\left(1-\frac{2}{5}x+\frac{4}{35}x^2+\ldots\right)+c_2x^{-1/2}\left(1-x+\frac{1}{2}x^2+\ldots\right).$$

Problems

1. For each of the following differential equations, locate and classify its singular points on the x-axis;
   a. $x^3(x-1)y^{11} - 2(x-1)y^1 + 3xy = 0$; <u>Soln.</u> x=0 irregular, x=1 regular
   b. $x^2(x^2-1)^2y^{11} - x(1-x)y^1 + 2y = 0$; <u>Soln.</u> x=0 and x=1 both regular, x=–1 irregular
   c. $x^2y^{11} + (2-x)y^1 = 0$; <u>Soln.</u> x=0 irregular.
   d. $(3x+1)xy^{11} - (x+1)y^1 + 2y = 0$; <u>Soln.</u> x=0 and $x=\frac{-1}{3}$ regular.
2. Determine the nature of the point x=0 for each of the following equation.
   a. $y^{11} + (\text{Sin } x)y=0$; <u>Soln.</u> ordinary point
   b. $xy^{11} + (\text{Sin } x)y=0$; <u>Soln.</u> ordinary point

c. x^2y^{11} + (Sin x) y=0; <u>Soln.</u> regular singular point

d. x^3y^{11} + (Sin x) y=0; <u>Soln.</u> regular singular point

e. x^4y^{11}+(Sin x) y=0; <u>Soln.</u> irregular singular point

3. Find the indicial equation and its roots for each of the following differential equations.

   a. $x^3y^{11} + (\cos 2x - 1)y^1 + 2xy = 0$. <u>Soln.</u> m(m–1) – 2m + 2=0; $m_1 = 2$, $m_2=1$

   b. $4x^2y^{11}+(2x^4-5x)y^1+(3x^2+2)y=0$.<u>Soln.</u> m(m–1) $-\frac{5}{4}m+\frac{1}{2}=0; m_1=2, m_2=1/4$

4. For each of the following equations verify that the origin is a regular singular point and calculate two independent Frobenius series solutions.

   a. $4xy^{11} + 2y^1 + y=0$; <u>Soln.</u> $\begin{cases} y_1(x)=x^{1/2}\left(1-\frac{x}{3!}+\frac{x^2}{5!}-\ldots\right)=\sin\sqrt{x} \\ y_2(x)=1-\frac{x}{2!}+\frac{x^2}{4!}-\ldots=\cos\sqrt{x} \end{cases}$

   b. $2x^2y^{11} + xy^1 - (x+1)y = 0$;

   <u>Soln</u> . $\begin{cases} y_1(x) = x\left(1+\frac{1}{5}x+\frac{1}{70}x^2+\cdots\right) \\ y_2(x) = x^{-1/2}\left(1-x-\frac{1}{2}x^2+\cdots\right) \end{cases}$

5. Show that $y=\sum_{n=0}^{\infty}\frac{(-1)^n}{2^{2n}(n!)^2}x^{2n}$ is the Frobenius series solution of the differential equation $x^2y^{11} + xy^1 + x^2y = 0$.

6. Show that the differential equation $x^2y^{11} + (3x-1)y^1 + y=0$ has x=0 as an irregular singular point. Also show that 0 is a root of the indicial equation and the corresponding Frobenius series solution is y=$\sum_{n=0}^{\infty} n!x^n$, which converges only at x=0. [This problem show that even when a Frobenius series formally satisfies a differential equation, it is not necessarily a valid solution]

I.6 Theoretical side of the solution near a regular singular point.

Theorem A

Assume that x=0 is a regular singular point of the differential equation $y^{11} + P(x)\, y^{1} + Q(x)\, y = 0 \ldots$ (1)

and that the power series expansions $x\,P(x) = \sum_{n=0}^{\infty} p_n x^n$ and $x^2 Q(x) = \sum_{n=0}^{\infty} q_n x^n$

are valid on an interval |x|<R with R>0. Let the indicial equation $m(m-1) + mp_0 + q_0 = 0$ have real roots m_1 and m_2 with $m_2 \leq m_1$. Then equation (1) has at least one solution $y_1 = x^{m_1} \sum_{n=0}^{\infty} a_n x^n$ $(a_0 \neq 0)$ on the interval 0<x<R, where the a_n are determined in terms of a_0 by the recursion formula

$a_n[(m+n)(m+n-1)+(m+n)p_0+q_0]+\sum_{k=0}^{n-1} a_k[(m+k)p_{n-k}+q_{n-k}]=0$

with m replaced by m_1, and the series $\sum a_n x^n$ converges for |x| <R: Furthermore, if $m_1 - m_2$ is not zero or a positive integer, then equation (1) has a second independent solution $y_2 = x^{m_2} \sum_{n=0}^{\infty} a_n x^n$ $(a_0 \neq 0)$ on the same interval, where in this case the a_n are determined in terms of a_0 by the same recursion formula with m replaced by m_2, and again the series $\sum a_n x^n$ converges for |x|<R.

Proof.

We attempt a formal calculation of any solutions of (1) that have the Frobenius form.

$$y = x^m (a_0 + a_1 x + a_2 x^2 + \ldots), \qquad (2)$$

where $a_0 \neq 0$ and *m* is a number to be determined.

We confine our attention to the interval x>0. The behavior of solutions on the interval x<0 can be studied by changing the variable to t = –x and solving the resulting equation for t>0.

Our hypothesis is that xP(x) and x^2Q(x) are analytic at x=0, and therefore have power series expansions.

$$xP(x)=\sum_{n=0}^{\infty} p_n x^n \quad \text{and} \quad x^2Q(x)=\sum_{n=0}^{\infty} q_n x^n \qquad (3)$$

Which are valid on an interval |x| < R for some R>0. We must find the possible values of *m* in (2); and then, for each acceptable *m*, we must calculate the corresponding coefficients a_0, a_1, a_2 …. If we write (2) in the form.

$$y = x^m \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a_n x^{m+n},$$

then differentiation yields

$$y^1 = \sum_{n=0}^{\infty} (m+n) x^{m+n-1}$$

and

$$y^{11} = x^{m-2} \sum_{n=0}^{\infty} a_n (m+n)(m+n-1) x^n.$$

The terms P(x)y^1 and Q(x)y in (1) can now be written as

$$P(x)y^1 = \frac{1}{x}\left(\sum_{n=0}^{\infty} P_n x^n\right)\left[\sum_{n=0}^{\infty} a_n (m+n) x^{m+n-1}\right]$$

$$= x^{m-2}\left(\sum_{n=0}^{\infty} P_n x^n\right)\left[\sum_{n=0}^{\infty} a_n (m+n) x^n\right]$$

$$= x^{m-2}\sum_{n=0}^{\infty}\left[\sum_{k=0}^{\infty} P_{n-k} a_k (m+k)\right] x^n$$

$$= x^{m-2}\sum_{n=0}^{\infty}\left[\sum_{k=0}^{n} P_{n-k} a_k (m+k) + p_0 a_n (m+n)\right] x^n$$

and $$Q(x)y = \frac{1}{x^2}\left(\sum_{n=0}^{\infty} q_n x^n\right)\left(\sum_{n=0}^{\infty} a_n x^{m+n}\right)$$

$$= x^{m-2}\left(\sum_{n=0}^{\infty} q_n x^n\right)\left(\sum_{n=0}^{\infty} a_n x^n\right)$$

$$= x^{m-2}\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} q_{n-k} a_k\right) x^n$$

$$= x^{m-2} \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n-1} q_{n-k}a_k + q_0 a_n\right)x^n.$$

When these expression for y^{11}, $P(x)y^1$, and $Q(x)$ y are inserted in (1) and the common factor x^{m-2} is canceled, then the differential equation becomes.

$$\sum_{n=0}^{\infty}\{a_n[(m+n)(m+n-1)+(m+n)P_0+q_0] + \sum_{k=0}^{n-1} a_k[(m+k)p_{n-k}+q_{n-k}]\}x^n = 0;$$

and equating to zero the coefficient of x^n yields the following recursion formula for the a_n:

$$a_n[(m+n)(m+n-1) + (m+n)p_0 + q_0] + \sum_{k=0}^{n-1} a_k[(m+k)P_{n-k}+q_{n-k}] = 0. \qquad (4)$$

On writing this out for the successive values of n, we get

$$a_0[m(m-1) + mp_0 + q_0] = 0,$$

$$a_1[(m+1)m + (m+1)p_0 + q_0] + a_0(mp_1 + q_1) = 0,$$

$$a_2[(m+2)(m+1) + (m+2)p_0 + q_0] + a_0(mp_2 + q_2) + a_1[(m+1)p_1 + q_1)] = 0,$$

$$\dots$$

$$a_n[(m+n)(m+n-1) + (m+n)p_0 + q_0]$$

$$+ a_0(mp_n + q_n) + \dots + a_{n-1}[(m+n-1)p_1 + q_1 = 0,$$

If we put $f(m) = m(m-1) + mp_0 + q_0$, then these equation become

$$a_0 f(m) = 0,$$

$$a_1 f(m+1) + a_0(mp_1 + q_1) = 0,$$

$$a_2 f(m+2) + a_0(mp_2 + q_2) + a_1[(m+1)p_1 + q_1] = 0,$$

$$\dots$$

$$a_n f(m+n) + a_0(mp_n + q_n) + \dots + a_{n-1}[(m+n-1)p_1 + q_1] = 0,$$

Since $a_0 \neq 0$, we conclude from the first of these equation that $f(m) = 0$ or, equivalently, that

$$m(m-1) + mp_0 + q_0 = 0. \qquad (5)$$

This is the indicial equation, and its roots m_1 and m_2 – which are possible values for m in our assumed solution (2) – are called the exponents of the differential equation (1) at the regular singular point x=0. The following equation give a_1 in terms of a_0, a_2 in terms of a_0 and a_1, and so on. The a_n are therefore determined in terms of a_0 for each choice of m unless f(m+n) = 0 for some positive integer n, in which case the process breaks off. Thus, if $m_1 = m_2 + n$ for some integer $n \geq 1$, the choice $m=m_1$ gives a formal solution but in general $m = m_2$ does not–since $f(m_2 + n) = f(m_1) = 0$. If $m_1 = m_2$ we also obtain only one formal solution. In all other cases where m_1 and m_2 are real numbers, this procedure yields two independent formal solutions.

Note:

Theorem A unfortunately fails to answer the question of how to find a second solution when the difference $m_1 - m_2$ is zero or a positive integer. In order to convey an idea of the possibilities here, we distinguish three cases.

Case A. If $m_1 = m_2$, there cannot exist a second Frobenius series solution.

The other two cases, in both of which $m_1 - m_2$ is a positive integer, will be easier to grasp if we insert $m = m_2$ in the recursion formula (4) and write it as

$$a_n f(m_2 + n) = -a_0 (m_2 p_n + q_n) - \dots - a_{n-1} [(m_2 + n - 1) p_1 + q_1]. \qquad (8)$$

As we know, the difficulty in calculating the a_n arises because $f(m_2 + n)=0$ for a certain positive integer n. The next two cases deal with this problem.

Case B. If the right side of (8) is not zero when $f(m_2 + n) = 0$, then there is no possible way of continuing the calculation of the coefficients and there cannot exist a second Frobenius series solution.

Case C. If the right side of (8) happens to be zero when $f(m_2 + n) = 0$, then a_n is unrestricted and can be assigned any value whatever. In particular, we can put $a_n = 0$ and continue to compute the coefficients without any further difficulties. Hence in this case there does exist a second Frobenius series solution.

The problems below will demonstrate that each of these three possibilities actually occurs.

The following calculations enable us to discover what form the second solution takes when $m_1 - m_2$ is zero or a positive integer. We begin by defining a positive integer k by $k = m_1 - m_2 + 1$. The indicial equation (5) can be written as

$$(m-m_1)(m-m_2) = m^2 - (m_1 + m_2)m + m_1m_2 = 0.$$

So equating the coefficients of m yields $p_0 - 1 = -(m_1 + m_2)$ or $m_2 = 1 - p_0 - m_1$, and we have $k=2m_1 + p_0$. We can find a second solution y_2 from the known solution $y_1 = x^{m_1}(a_0 + a_1x + \dots)$ by writing $y_2 = vy_1$, where

$$v^1 = \frac{1}{y_1^2} e^{-\int P(x)dx}$$

$$= \frac{1}{x^{2m_1}(a_0 + a_1x + \cdots)^2} e^{-\int\left(\left({p_0}/{x}\right) + p_1 + \cdots\right)dx}$$

$$= \frac{1}{x^{2m_1}(a_0 + a_1x + \dots)^2} e^{(-p_0 \log x - p_1 x - \dots)}$$

$$= \frac{1}{x^k(a_0 + a_1x + \dots)^2} e^{(-p_1 x - \dots)} = \frac{1}{x^k} g(x).$$

The function g(x) defined by the last equality is clearly analytic at x = 0, with g(0) = $1/a_0^2$, so in some interval about the origin we have

$$g(x) = b_0 + b_1x + b_2x^2 + \dots, \quad b_0 \neq 0. \qquad (9)$$

It follows that

$$v^1 = b_0x^{-k} + b_1k^{-k+1} + \dots + b_{k-1}x^{-1} + b_k + \dots,$$

so

$$v = \frac{b_0 x^{-k+1}}{-k+1} + \frac{b_1 x^{-k+2}}{-k+2} + \dots + b_{k-1}\log x + b_k x + \dots$$

and

$$y_2 = y_1 v = y_1 \left(\frac{b_0\, x^{-k+1}}{-k+1} + \ldots + b_{k-1} \log x + b_k x + \ldots \right)$$

$$= b_{k-1} y_1 \log x + x^{m_1} (a_0 + a_1 x + \ldots) \left(\frac{b_0 x^{-k+1}}{-k+1} + \ldots. \right)$$

If we factor x^{-k+1} out of the series last written, use $m_1 - k + 1 = m_2$, and multiply the two remaining power series, then we obtain.

$$y_2 = b_{k-1}\, y_1 \log x + x^{m_2} \sum_{n=0}^{\infty} c_n x^n \qquad (10)$$

as our second solution.

Formula (10) has only limited value as a practical tool; but it does yield several grains of information. First, if the exponents m_1 and m_2 are equal, then k=1 and $b_{k-1} = b_0 \neq 0$; so in this case – which is Case A above – the term containing log x is definitely present in the second solution (10). However, if $m_1 - m_2 = k - 1$ is a positive integer, then sometimes $b_{k-1} \neq 0$ and the logarithmic term is present (Case B), and sometimes $b_{k-1} = 0$ and there is no logarithmic term (Case C). The practical difficulty here is that we cannot readily find b_{k-1} because we have no direct means of calculating the coefficients in (9). In any event, we at least know that in Cases A and B, when $b_{k-1} \neq 0$ and the method of Frobenius is only partly successful, the general form of a second solution is

$$y_2 = y_1 \log x + x^{m_2} \sum_{n=0}^{\infty} c_n x^n \qquad (11)$$

where the c_n are certain unknown constants that can be determined by substituting (11) directly into the differential equation.

Problems

1. The equation $x^2y^{11} - 3xy^1 + (4x+4)y=0$ has only one Frobenius series solution. Find it.

   Soln. $y=x^2(1-4x+4x^2+\ldots)$

2. Find two independent Frobenius series solutions of each of the following equations.

   a. $xy^{11} + 2y^{1} + xy = 0$

      Soln. $y_1 = 1 - x^2/3! + x^4/5! - \ldots = x^{-1}$ Sin x; $y_2 = x^{-1}\left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots\right)$

      $= x^{-1}$Cos x

   b. $x^2y^{11} - x^2y^1 + (x^2 - 2)\, y = 0$

      Soln. $y_1 = x^2\left(1 + \frac{1}{2}x + \frac{1}{20}x^2 - \frac{1}{60}x^3 + \ldots.\right)$; $y_2 = x^{-1}\left(1 + \frac{1}{2}x + \frac{1}{2}x^2 - \frac{1}{8}x^4 + \ldots\right)$

   c. $xy^{11} - y^1 + 4x^3\, y = 0$

      Soln. $y_1 = x^2\left(1 - \frac{x^4}{3!} + \frac{x^8}{5!} - \ldots\right) = \sin x^2$; $y_2 = 1 - \frac{x^4}{2!} + \frac{x^8}{4!} - \ldots = Cos\, x^2$

3. Bessel's equation of order p=1 is $x^2\, y^{11} + xy^1 + (x^2 - 1)\, y = 0$. Show that $m_1 - m_2 = 2$ and that the equation has only one Frobenius series solution. Then find it.

   Soln. $y = x\left(1 - \frac{x^2}{2^2\, 2!} + \frac{x^4}{2^4\, 2!3!} - \ldots\right)$

4. Bessel's equation of order $p = \frac{1}{2}$ is $x^2y^{11} + xy^1 + \left(x^2 - 1/4\right)\, y = 0$. Show that $m_1 - m_2 = 1$, but that nevertheless the equation has two independent Frobenius series solutions. Then find them.

   Soln. $y_1 = x^{1/2}\left(1 - x^2/3! + x^4/5! - \ldots\right) = x^{-1/2}\sin x$

   $y_2 = x^{-1/2}\left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots\right) = x^{-1/2}\cos x$

I.7 Gauss's Hypergeometric Equation

This famous differential equation is

$$x(1-x)y^{11} + [c - (a + b + 1)\,x]y^{1} - aby = 0. \qquad (1)$$

where a,b, and c are constants. The coefficients of (1) may look rather strange, but we shall find that they are perfectly adapted to the use of its solutions in a wide variety of situations. The best way to understand this is to solve the equation for ourselves and see what happens.

We have

$$P(x) = \frac{c-(a+b+1)x}{x(1-x)} \text{ and } Q(x) = \frac{-ab}{x(1-x)},$$

so x = 0 and x = 1 are the only singular points on the x-axis. Also,

$$xP(x) = \frac{c-(a+b+1)x}{1-x} = [c-(a+b+1)x](1+x+x^2+\ldots)$$

$$= c+[c-(a+b+1)]x+\ldots$$

and

$$x^2\,Q(x) = \frac{-abx}{1-x} = -abx(1+x+x^2+\ldots)$$

$$= -abx - abx^2 - \ldots,$$

so x = 0 (and similarly x = 1) is a regular singular point. These expansions show that $p_0 = c$ and $q_0 = 0$, so the indicial equation is

$$m\,(m-1) + mc = 0 \quad \text{or} \quad m[m-(1-c)] = 0$$

and the exponents are $m_1 = 0$ and $m_2 = 1 - c$. If 1–c is not a positive integer, that is, if c is not zero or a negative integer, then Theorem I.6.A guarantees that (1) has a solution of the form

$$y = x^0\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1\,x + a_2\,x^2 + \ldots., \qquad (2)$$

where a_0 is a nonzero constant. On substituting this into (1) and equating to zero the coefficient of x^n, we obtain the following recursion formula for the a_n:

$$a_{n+1} = \frac{(a+n)(b+n)}{(n+1)(c+n)} a_n. \qquad (3)$$

We now set $a_0 = 1$ and calculate the other a_n in succession:

$$a_1 = \frac{ab}{1\cdot c}; \qquad a_2 = \frac{a(a+1)b(b+1)}{1\cdot 2c(c+1)},$$

$$a_3 = \frac{a(a+1)(a+2)b(b+1)(b+2)}{1\cdot 2\cdot 3c(c+1)(c+2)},\ldots$$

With these coefficients, (2) becomes

$$y = 1 + \frac{ab}{1\cdot c}x + \frac{a(a+1)b(b+1)}{1\cdot 2c(c+1)}x^2$$

$$+ \frac{a(a+1)(a+2)b(b+1)(b+2)}{1\cdot 2\cdot 3c(c+1)(c+2)}x^3 + \ldots$$

$$= 1 + \sum_{n=1}^{\infty} \frac{a(a+1)\ldots(a+n-1)b(b+1)\ldots(b+n-1)}{n!c(c+1)\ldots(c+n-1)}x^n. \qquad (4)$$

This is known as the hypergeometric series, and is denoted by the symbol F(a,b,c,x). It is called this because it generalizes the familiar geometric series as follows: when a = 1 and c=b, we obtain.

$$F(1,b,b,x) = 1 + x + x^2 + \ldots = \frac{1}{1-x}.$$

If a or b is zero or a negative integer, the series (4) breaks off and is a polynomial; otherwise the ratio test shows that it converges for $|x|<1$, since (3) gives.

$$\left|\frac{a_{n+1}x^{n+1}}{a_n x^n}\right| = \left|\frac{(a+n)(b+n)}{(n+1)(c+n)}\right||x| \to |x| \; as\; n \to \infty.$$

This convergence behaviour could also have been predicted from the fact that the singular point closest to the origin is x=1. Accordingly, when c is not zero or a negative integer, F(a, b, c, x) is an analytic function – called the hypergeometric function – on the interval $|x|<1$. It is the simplest particular solution of the hypergeometric equation. The hypergeometric function has a great many properties, of which the most obvious is that it is unaltered when a and b are interchanged : $F(a,b,c,x) = F(b, a, c, x)$.

If $1 - c$ is not zero or a negative integer – which means that c is not a positive integer – then Theorem 1.6 A also tells us that there is a second independent solution of (1) near $x = 0$ with exponent $m_2 = 1 - c$. This solution can be found directly, by substituting

$$y=x^{1-c} (a_0 + a_1x + a_2x^2 + \dots)$$

into (1) and calculating the coefficients. It is more instructive, however, to change the dependent variable in (1) from y to z by writing.

$$y = x^{1-c}z.$$

when the necessary computations are performed – students should do this work themselves – equation (1) becomes.

$$x(1-x)z^{11} + [(2-c) - ([a-c+1] + [b-c+1] + 1)x]z^1 - (a-c+1)(b-c+1)z=0. \qquad (5)$$

which is the hypergeometric equation with the constants a, b, and c replaced by $a-c+1$, $b-c+1$, and $2-c$. We already know that (5) has the power series solution.

$$z=F(a-c+1, b-c+1, 2-c,x)$$

near the origin, so our desired second solution is

$$y=x^{1-c}F(a-c+1, b-c+1, 2-c,x).$$

Accordingly, when c is not an integer, we have

$$y=c_1F(a,b,c,x) + c_2x^{1-c} F(a-c+1, b-c+1, 2-c,x) \qquad (6)$$

as the general solution of the hypergeometric equation near the singular point $x=0$.

In general, the above solution is only valid near the origin. We now solve (1) near the singular point $x=1$. The simplest procedure is to obtain this solution from the one already found, by introducing a new independent variable $t=1-x$. This makes $x=1$ correspond to $t=0$ and transforms (1) into

$$t(1-t)y^{11} + [(a+b-c+1) - (a+b+1)t]y^{1} - aby = 0,$$

where the primes signify derivatives with respect to t. Since this is a hypergeometric equation, its general solution near t=0 can be written down at once from (6), by replacing x by t and c by a+b−c+1; and when t is replaced by 1−x, we see that the general solution of (1) near x=1 is

$$y=c_1F(a,b,a+b-c+1, 1-x)$$

$$+ c_2(1-x)^{c-a-b}F(c-b, c-a, c-a-b+1, 1-x). \qquad (7)$$

Note:

Consider the general equation.

$$(x-A)(x-B)\, y^{11} + (C+Dx)y^{1} + Ey = 0, \qquad (8)$$

where $A \neq B$. If we change the independent variable from x to t by means of

$$t = \frac{x-A}{B-A},$$

then x=A corresponds to t = 0, and x=B to t=1. With a little calculation, equation (8) assumes the form.

$$t(1-t)\, y^{11} + (F + Gt)y^{1} + Hy = 0,$$

where F, G, and H are certain combinations of the constants in (8) and the primes indicate derivatives with respect to t. This is a hypergeometric equation with a, b, and c defined by

$$F=c,\ G = -(a+b+1),\ H = -ab,$$

and can therefore be solved near t=0 and t=1 in terms of the hypergeometric function. But this means that (8) can be solved in terms of the same functions near x=A and x=B.

Problems

1. Verify each of the following by examining the series expansions of the functions on the left sides:

   a. $(1+x)^p = F(-p,b,b,-x)$;

   b. $\log(1+x) = xF(1,1,2,-x)$;

   c. $\sin^{-1} x = xF\left(\frac{1}{2},\frac{1}{2},\frac{3}{2},x^2\right)$;

   d. $\tan^{-1} x = xF\left(\frac{1}{2},1,\frac{3}{2},-x^2\right)$.

   It is also true that

   e. $e^x = \lim_{b\to\infty} F\left(a,b,a,\frac{x}{b}\right)$;

   f. $\sin x = x\left[\lim_{b\to\infty} F\left(a,a,\frac{3}{2},\frac{-x^2}{4a^2}\right)\right]$;

   g. $\cos x = \lim_{a\to\infty} F\left(a,a,\frac{1}{2},\frac{-x^2}{4a^2}\right)$.Satisfy yourself of the validity of these statements without attempting to justify the limit processes involved.

2. Find the general solution of each of the following differential equations near the indicated singular point:

   a. $x(1-x)\,y^{11} + \left(\frac{3}{2}-2x\right)y^1 + 2y = 0$,x = 0

   b. $(2x^2 + 2x)y^{11} + (1+5x)y^1 + y = 0$, $x = 0$;

   c. $(x^2 - 1)y^{11} + (5x + 4)\,y^1 + 4y = 0$, $x = -1$;

   d. $(x^2 - x - 6)\,y^{11} + (5+3x)y^1 + y = 0$, $x = 3$.

Solution:

2(a) $y = c_1\,F(2, -1, 3/2, x) + c_2x^{-1/2}\,F(3/2, -3/2, ½, x)$

$$= c_1\left(1-\frac{4}{3}x\right) + c_2\,x^{-1/2}\,F(3/2, -3/2, ½, x)$$

2(b) $y = c_1\,F(1/2, 1, ½, -x) + c_2\,(-x)^{1/2}\,F(1, 3/2, 3/2, -x)$

$$= c_1\left(\frac{1}{1+x}\right) + c_2\left[\frac{(-x)^{1/2}}{1+x}\right]$$

2(c) $y = c_1 F\left(2,2,\tfrac{1}{2},\frac{x+1}{2}\right)+c_2\left(\frac{x+1}{2}\right)^{1/2} F\left(\frac{5}{2},\frac{5}{2},\frac{3}{2},\frac{x+1}{2}\right)$

2(d) $y = c_1 F\left(1,1,\frac{14}{5},\frac{3-x}{5}\right)+c_2\left(\frac{3-x}{5}\right)^{-9/5} F\left(-\tfrac{4}{5},-\tfrac{4}{5},-\tfrac{4}{5},\frac{3-x}{5}\right)$

3. Show that $F^1(a,b,c,x) = \frac{ab}{c}$ F(a+1, b+1, c+1, x)
4. Transform the chebyshev's equation $(1-x^2)y^{11} - xy^1 + p^2 y = 0$ in to a hypergeometric equation and show that its general solution near x=1 is $y=c_1 F\left(p,-p,\tfrac{1}{2},\frac{1-x}{2}\right)+c_2\left(\frac{1-x}{2}\right)^{1/2} F\left(p+\tfrac{1}{2},-p+\tfrac{1}{2},\tfrac{3}{2},\frac{1-x}{2}\right)$

I.8 The point at infinity

It is often desirable, in both physics and pure mathematics, to study the solutions of

$$y^{11} + P(x)y^1 + Q(x)y = 0 \tag{1}$$

for large values of the independent variable. For instance, if the variable is time, we may want to know how the physical system described by (1) behaves in the distant future, when transient disturbances have faded away.

We can adapt our previous ideas to this broader purpose by studying solutions near the point at infinity. The procedure is quite simple, for if we change the independent variable from x to

$$t = \frac{1}{x}, \tag{2}$$

then large x's correspond to small t's. Consequently, if we apply (2) to (1), solve the transformed equation near t=0, and then replace t by 1/x in these solutions, we have solutions of (1) that are valid for large values of x. To carry out this program, we need the formulas.

$$y^1 = \frac{dy}{dx} = \frac{dy}{dt}\frac{dt}{dx} = \frac{dy}{dt}\left(-\frac{1}{x^2}\right) = -t^2\frac{dy}{dt} \qquad (3)$$

and

$$y^{11} = \frac{d}{dx}\left(\frac{dy}{dx}\right) = \frac{d}{dt}\left(\frac{dy}{dx}\right)\frac{dt}{dx} = \left(-t^2\frac{d^2y}{dt^2} - 2t\frac{dy}{dt}\right)\left(-t^2\right) \qquad (4)$$

when these expressions are inserted in (1), and primes are used to denote derivatives with respect to t, then (1) becomes.

$$y^{11} + \left[\frac{2}{t} - \frac{P(1/t)}{t^2}\right]y^1 + \frac{Q(1/t)}{t^4}y = 0. \qquad (5)$$

We say that equation (1) has $x = \infty$ as an ordinary point, a regular singular point with exponents m_1 and m_2, or an irregular singular point, if the point t=0 has the corresponding character for the transformed equation (5).

As a simple illustration, consider the Euler equation

$$y^{11} + \frac{4}{x}y^1 + \frac{2}{x^2}y = 0. \qquad (6)$$

A comparison of (6) with (5) shows that the transformed equation is

$$y^{11} - \frac{2}{t}y^1 + \frac{2}{t^2}y = 0. \qquad (7)$$

It is clear that t=0 is a regular singular point for (7), with indicial equation.

$$m(m-1) - 2m + 2 = 0$$

and exponents $m_1 = 2$ and $m_2 = 1$. This means that (6) has $x=\infty$ as a regular singular point with exponents 2 and 1.

Consider the hypergeometric equation

$$x(1-x)\,y^{11} + [c-(a+b+1)x]y^1 - aby = 0. \qquad (8)$$

We already know that (8) has two finite regular singular points: x=0 with exponents 0 and 1 – c; and x=1 with exponents 0 and c – a – b. To determine the nature of the point x=∞, we substitute (3) and (4) directly into (8). After a little rearrangement, we find that the transformed equation is

$$y^{11} + \left[\frac{(1-a-b)-(2-c)t}{t(1-t)}\right]y^{1} + \frac{ab}{t^2(1-t)}\,y = 0. \tag{9}$$

This equation has t=0 as a regular singular point with indicial equation

$$m\,(m-1) + (1-a-b)\,m + ab = 0.$$

or

$$(m-a)(m-b) = 0.$$

This shows that the exponents of equation (9) at t=0 are a and b, so equation (8) has x=∞ as a regular singular point with exponents a and b. We conclude that the hypergeometric equation (8) has precisely three regular singular points: 0, 1, and ∞ with corresponding exponents 0 and 1–c, 0 and c–a–b, and a and b.

CHAPTER II

SPECIAL FUNCTIONS

II.1 Legendre Polynomials

Consider the Legendre's equation

$$(1-x^2)y^{11} - 2xy^1 + p(p+1)y=0, \ldots\ldots\ldots \qquad (1)$$

Where p is a constant. It is clear that the coefficient functions

$$P(x) = \frac{-2x}{1-x^2} \text{ and } Q(x)=\frac{p(p+1)}{1-x^2} \qquad (2)$$

are analytic at the origin. The origin is therefore an ordinary point, and we expect a solution of the form $y=\Sigma a_n x^n$. Since $y^1=\Sigma(n+1)a_{n+1} x^n$, we get the following expansions for the individual terms on the L.H.S of (1)

$$y^{11}=\Sigma(n+1)(n+2)a_{n+2}x^n, \ -x^2y^{11} = \Sigma^{-}(n-1)n\, a_n x^n,$$

$$-2xy^1 = \Sigma-2na_n x^n, \ p(p+1)\, y = \Sigma p\,(p+1)a_n x^n$$

By equation (1), the sum of these series is required to be zero, so the coefficient of x^n must be zero for every n and hence

$$(n+1)(n+2)a_{n+2} - (n-1)n\, a_n - 2na_n + p(p+1)a_n=0.$$

With a little manipulation this becomes,

$$a_{n+2} = -\frac{(p-n)(p+n+1)}{(n+1)(n+2)}a_n \qquad (3)$$

This recursion formula enables us to express a_n in terms of a_n or a_1 according as n is even or odd.

$$a_2 = \frac{-p(p+1)}{1\cdot 2}a_0, \ a_3 = \frac{-(p-1)(p+2)}{2\cdot 3}a_1,$$

$$a_4 = \frac{-(p-2)(p+3)}{3\cdot 4}a_2 = \frac{p(p-2)(p+1)(p+3)}{4!}a_0$$

$$a_5 = \frac{-(p-3)(p+4)}{4\cdot 5}a_3 = \frac{(p-1)(p-3)(p+2)(p+4)}{5!}a_1$$

$$a_6 = \frac{-(p-4)(p+5)}{5.6}a_4 = \frac{-p(p-2)(p-4)(p+1)(p+3)(p+5)}{6!}a_0$$

$$a_7 = \frac{-(p-5)(p+6)}{6\cdot 7}a_5 = \frac{(p-1)(p-3)(p-5)(p+2)(p+4)(p+6)}{7!}a_1$$

and so on. By inserting these coefficients in to the assumed solution $y=\Sigma a_n x^n$, we obtain

$$y=a_0\left[1-\frac{p(p+1)}{2!}x^2+\frac{p(p-2)(p+1)(p+3)}{4!}x^4-\frac{p(p-2)(p-4)(p+1)(p+3)(p+5)}{6!}x^6+\ldots\right]$$

$$+\ a_1\left[x-\frac{(p-1)(p+2)}{3!}x^3+\frac{(p-1)(p-3)(p+2)(p+4)}{5!}x^5\right.$$

$$\left.-\frac{(p-1)(p-3)(p-5)(p+2)(p+4)(p+6)}{7!}x^7+\ldots\right]$$

is our formal solution.

Each bracketed series is a particular solution and has radius of convergence R=1. Since it is clear that the functions defined by these series are linearly independent, (4) is the general solution of (1) on the interval $|x|<1$. The functions defined by (4) are called Legendre functions and in general they are not elementary. However, when p is a non negative integer, one of the series terminates and is thus a polynomial – (the first series if p is even and the second series if p is odd) – while the other does not and remains an infinite series. These polynomials are known as Legendre polynomials.

<u>Rodrigues' formula for the n^{th} Legendre polynomial.</u>

With p replaced by n, the Legendre's equation (1) takes the form

$$(1-x^2)\, y^{11} - 2xy^1 + n(n+1)y = 0 \qquad (5)$$

where n is understood to be a non negative integer.

The solutions most useful in the applications are those bounded near x=1, and for convenience in singling these out we change the independent variable from x to $t=\frac{1}{2}(1-x)$. This makes x=1 correspond to t=0 and transforms (5) in to

$$t(1-t)\,y'' + (1-2t)y' + n(n+1)\,y=0 \qquad (6)$$

where the primes signify derivatives with respect to t. This is a hypergeometric equation with a=−n, b=n+1, and c=1, so it has the following solution near

$$t=0;\ y_1 = F(-n, n+1, 1, t) \qquad (7)$$

Since the exponents of (6) at the origin are both zero ($m_1 = 0$ and $m_2 = 1 - c = 0$), second solution is $y_2 = \upsilon y_1$, where

$$\upsilon' = \frac{1}{y_1^2}e^{-\int P\,dt} = \frac{1}{y_1^2}e^{\int (2t-1)/t(1-t)\,dt}$$

$$= \frac{1}{y_1^2\,t(1-t)} = \frac{1}{t}\left[\frac{1}{y_1^2(1-t)}\right],$$

by an elementary integration. Since y_1^2 is a polynomial with constant term 1, the bracketed expression on the right is an analytic function of the form $1 + a_1t + a_2t^2 + \ldots\ldots$, and we have

$$\upsilon' = \frac{1}{t} + a_1 + a_2t + \ldots..$$

This yields $\upsilon = \log t + a_1t + \ldots.$, so

$$y_2 = y_1(\log t + a_1t + \ldots)$$

and the general solution of (6) near the origin is

$$y = c_1y_1 + c_2y_2 \qquad (8)$$

Because of the presence of the term log t in y_2, it is clear that (8) is bounded near t = 0 if and only if $c_2 = 0$. If we replace t in (7) by $\frac{1}{2}(1-x)$, it follows that the solutions of (1) bounded near x = 1 are precisely constant multiples of the polynomial $F\left[-n, n+1, 1, \frac{1}{2}(1-x)\right]$.

This brings us to the fundamental definition. The *n*th Legendre polynomial is denoted by $P_n(x)$ and defined by

$$P_n(x) = F\left[-n, n+1, 1, \frac{1}{2}(1-x)\right] = 1 + \frac{(-n)(n+1)}{(1!)^2}\left(\frac{1-x}{2}\right)$$

$$+\frac{(-n)(-n+1)(n+1)(n+2)}{(2!)^2}\left(\frac{1-x}{2}\right)^2 + \dots\dots$$

$$+\quad \frac{(-n)(-n+1)\dots[-n+(n-1)](n+1)(n+2)\dots(2n)}{(n!)^2}\left(\frac{1-x}{2}\right)^n$$

$$= 1 + \frac{n(n+1)}{(1!)^2 2}(x-1) + \frac{n\,(n-1)(n+1)(n+2)}{(2!)^2 2^2}(x-1)^2 + \dots\dots + \frac{(2n)!}{(n!)^2 2^n}(x-1)^n \qquad (9)$$

$P_n(x)$ is a polynomial of degree n that contains only even or only odd powers of x according as n is even or odd. It can therefore be written in the form

$P_n(x) = a_n x^n + a_{n-2} x^{n-2} + a_{n-4} x^{n-4} + \dots,$ (10)

where this sum ends with a_0 if *n* is even and $a_1 x$ if *n* is odd. It is clear from (9) that $P_n(1) = 1$ for every *n*, and in view of (10) we also have $P_n(-1) = (-1)^n$.

As it stands, formula (9) is a very inconvenient tool to use in studying $P_n(x)$, So we look for something simpler. We could expand each term in (9), collect like powers of x, and arrange the result in the form (10), but this would be unnecessarily laborious. What we shall do is notice from (9) that $a_n = (2n)!/(n!)^2 2^n$ and calculate a_{n-2}, a_{n-4},…. recursively in terms of a_n. What is needed here is formula (3) with p replaced by n and n by k - 2:

$$a_k = -\frac{(n-k+2)(n+k-1)}{(k-1)k} a_{k-2}$$

or

$$a_{k-2} = -\frac{k(k-1)}{(n-k+2)(n+k-1)} a_k$$

When k = n, n-2, …. this yields

$$a_{n-2} = -\frac{n(n-1)}{2(2n-1)} a_{n,}$$

$$a_{n-4} = -\frac{(n-2)(n-3)}{4(2n-3)} a_{n-2} \quad = \quad \frac{n(n-1)(n-2)(n-3)}{2.4(2n-1)(2n-3)} a_n,$$

and so on, so (6) becomes

$$P_n(x) = \frac{(2n)!}{(n!)^2 2^n} \left[x^n - \frac{n(n-1)}{2(2n-1)} x^{n-2} \right.$$

$$+ \frac{n(n-1)(n-2)(n-3)}{2.4(2n-1)(2n-3)} x^{n-4} + \ldots$$

$$\left. + (-1)^k \frac{n(n-1)\ldots(n-2k+1)}{2^k k!(2n-1)(2n-3)\ldots(2n-2k+1)} x^{n-2k} + \ldots \right] \qquad (11)$$

Since

n(n – 1) …. (n -2k +1) = $\frac{n!}{(n-2k)!}$

and

(2n – 2k + 1)(2n – 2k + 3) ……. (2n – 3) (2n –1)

$$= \frac{(2n - 2k + 1)\,(2n - 2k + 2)\,(2n - 2k + 3)\ldots(2n - 3)(2n-2)(2n-1)2n}{(2n-2k+2)\ldots(2n-2)2n}$$

$$= \frac{(2n)!}{(2n-2k)!} \frac{1}{2^k (n-k+1)\ldots(n-1)n} = \frac{(2n)!(n-k)!}{(2n-2k)! 2^k n!},$$

the coefficient of x^{n-2k} in (11) is

$$(-1)^k \frac{n!}{2^k k!(n-2k)!} \frac{(2n-2k)! 2^k n!}{(2n)!(n-k)!} = (-1)^k \frac{(n!)^2 (2n-2k)!}{k!(2n)!(n-k)!(n-2k)!}$$

This enables us to write (11) as

$$p_n(x) = \sum_{k=0}^{[n/2]} (-1)^k \frac{(2n-2k)!}{2^n k!(n-k)!(n-2k)!} x^{n-2k}, \qquad (12)$$

Where [n/2] is the usual symbol for the greatest integer ≤n/2. We continue toward an even more concise form by observing that

$$P_n(x) = \sum_{k=0}^{[n/2]} \frac{(-1)^k}{2^n k!(n-k)!} \frac{(2n-2k)!}{(n-2k)!} x^{n-2k}$$

$$= \sum_{k=0}^{[n/2]} \frac{(-1)^k}{2^n k!(n-k)!} \frac{d^n}{dx^n} x^{2n-2k}$$

$$= \frac{1}{2^n n!} \frac{d^n}{dx^n} \sum_{k=0}^{[n/2]} \frac{n!}{k!(n-k)!} (x^2)^{n-k} (-1)^k$$

If we extend the range of this sum by letting k vary from 0 to n- which changes nothing since the new terms are of degree<n and their n th derivatives are zero – then we get

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} \left[\sum_{k=0}^{n} \binom{n}{k} (x^2)^{n-k} (-1)^k \right]$$

and the binominal formula yields

$$p_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n. \tag{13}$$

This expression for $P_n(x)$ is called Rodrigues' formula. It provides a relatively easy method for computing the successive Legendre polynomials, of which the first few are

$P_0(x) = 1,$ $\qquad$ $P_1(x) = x,$

$P_2(x) = \frac{1}{2}(3x^2 - 1),$ $\qquad$ $P_3(x) = \frac{1}{2}(5x^3 - 3x)$

Problem

1. The function on the left side of $\frac{1}{\sqrt{1-2xt+t^2}} = p_0(x) + p_1(x)t + p_2(x)t^2 + \ldots$ $+ p_n(x)t^n + \ldots$ is called the generating function of the Legendre polynomials. Assume that this relation is true, and use it
   a. to verify that $P_n(1)=1$ and $P_n(-1) = (-1)^n$
   b. to show that $P_{2n+1}(0) = 0$ and $P_{2n}(0) = (-1)^n \frac{1 \cdot 3 \ldots (2n-1)}{2^n n!}$
   c. $(n+1) P_{n+1}(x) = (2n+1) x P_n(x) - n P_{n-1}(x)$

[Hint) Differentiate both sides of $\frac{1}{\sqrt{1-2xt+t^2}}=\sum_{n=0}^{\infty} p_n(x)t^n$, to show that (x-t) $\sum_{n=0}^{\infty} p_n(x)t^n =$(1–2xt + t²) $\sum_{n=1}^{\infty} n\, p_n(x)t^{n-1}$ and equate coefficient of t^n on both sides]

d. Using recursion formula in (c) above calculate p_2 (x), p_3 (x), p_4 (x), taking P_0 (x) = 1 andP_1 (x) = X.

II.2 Properties of Legendre Polynomials

Orthogonality. The most important property of the Legendre polynomials is the fact that

$$\int_{-1}^{1} P_m(x)P_n(x)dx=\begin{cases} 0 & \text{if } m\neq n, \\ \dfrac{2}{2n+1} & \text{if } m=n. \end{cases} \qquad (1)$$

This is often expressed by saying that the Legendre polynomials $P_0(x)$, $P_1(x)$, $P_2(x)$,… $P_n(x)$,.. is a sequence of orthogonal functions on the interval $-1 \leq x \leq 1$.

Proof.

Let f(x) be any function with at least n continuous derivatives on the interval $-1\leq x \leq 1$, and consider the integral.

$$I = \int_{-1}^{1} f(x)\, p_n(x)\, dx$$

Rodriguies' formula enables us to write this as

$$I=\frac{1}{2^n n!}\int_{-1}^{1} f(x)\frac{d^n}{dx^n}\left(x^2-1\right)^n dx.$$

and an integration by parts gives.

$$I = \frac{1}{2^n\, n!}\left[f(x)\, \frac{d^{n-1}}{dx^{n-1}}(x^2-1)^n\right]_{-1}^{1} - \frac{1}{2^n\, n!}\int_{-1}^{1} f^1(x)\frac{d^{n-1}}{dx^{n-1}}\left(x^2-1\right)^n dx;$$

The expression in brackets vanishes at both limits, so

$$I = -\frac{1}{2^n\, n!}\int_{-1}^{1} f^1(x)\frac{d^{n-1}}{dx^{n-1}}\left(x^2-1\right)^n dx;$$

and by continuing to integrate by parts, we obtain.

$$I = \frac{(-1)^n}{2^n n!}\int_{-1}^{1} f^{(n)}(x)(x^2-1)^n\,dx.$$

If f (x) = P_m (x) with m < n, then $f^{(n)}$ (x) = 0 and consequently I=0, which proves the first part of (1). To establish the second part, we put f(x) = P_n(x). Since $p_n^{(n)}(x)$ = (2n)! / 2^nn!, it follows that

$$I = \frac{(2n)!}{2^{2n}(n!)^2}\int_{-1}^{1}(1-x^2)^n\,dx$$

$$= \frac{2(2n)!}{2^{2n}(n!)^2}\int_{0}^{1}(1-x^2)^n\,dx \tag{2}$$

If we change the variable by writing x = sin θ, and recall the formula (proved by an integration by parts).

$$\int \cos^{2n+1}\theta\,d\theta = \frac{1}{2n+1}\cos^{2n}\theta\sin\theta + \frac{2n}{2n+1}\int\cos^{2n-1}\theta\,d\theta, \tag{3}$$

then the definite integral in (3) becomes

$$\int_0^{\pi/2}\cos^{2n+1}\theta\,d\theta = \frac{1}{2n+1}\int_0^{\pi/2}\cos^{2n-1}\theta\,d\theta$$

$$= \frac{2n}{2n+1}\;\frac{2n-2}{2n-1}\cdots\frac{2}{3}\int_0^{\pi/2}\cos\theta\,d\theta$$

$$= \frac{2^n n!}{1\cdot 3\ldots(2n-1)(2n+1)} = \frac{2^{2n}(n!)^2}{(2n)!(2n+1)}.$$

we conclude that in this case I=2/(2n+1), and the proof of (1) is complete.

Legendre series.

Formula (14) in previous section, tell us that

$$1 = P_0(x), \qquad x = P_1(x),\; x^2 = \frac{1}{3} + \frac{2}{3}P_2(x) = \frac{1}{3}P_0(x) + \frac{2}{3}P_2(x)$$

$$x^3 = \frac{3}{5}x + \frac{2}{5}P_3(x) = \frac{3}{5}P_1(x) + \frac{2}{5}P_3(x)$$

and it follows that any third-degree polynomial p(x) = $b_0 + b_1x + b_2x^2 + b_3x^3$ can be written as

$$p(x) = b_0P_0(x) + b_1P_1(x) + b_2\left[\frac{1}{3}P_0(x) + \frac{2}{3}P_2(x)\right] + b_3\left[\frac{3}{5}P_1(x) + \frac{2}{5}P_3(x)\right]$$

$$= \left(b_0 + \frac{b_2}{3}\right)P_0(x) + \left(b_1 + \frac{3b_3}{5}\right)P_1(x) + \frac{2b_2}{3}P_2(x) + \frac{2b_3}{5}P_3(x)$$

$$= \sum_{n=0}^{3} a_n P_n(x).$$

More generally, since $P_n(x)$ is a polynomial of degree n for every positive integer n, a simple extension of this procedure shows that x^n can always be expressed as a linear combination of P_0 (x), $P_1(x)$, …. $P_n(x)$, so any polynomial p(x) of degree k has an expansion of the form.

$$p(x) = \sum_{n=0}^{k} a_n P_n(x).$$

An obvious problem that arises from these remarks – and also from the demands of the applications – is that of expanding an “arbitrary” function f(x) in a so-called Legendre series:

$$f(x) = \sum_{n=0}^{\infty} a_n P_n(x). \tag{4}$$

It is clear that a new procedure is needed for calculating the coefficients a_n in (4), and the key lies in formulas (1).

If we throw mathematical caution to the winds, and multiply (4) by $P_m(x)$ and integrate term by term from –1 to 1, then the result is

$$\int_{-1}^{1} f(x)P_m(x)dx = \sum_{n=0}^{\infty} a_n \int_{-1}^{1} P_m(x)P_n(x)dx;$$

and in view of (1), this collapse to

$$\int_{-1}^{1} f(x)P_m(x)\,dx = \frac{2a_m}{2m+1}.$$

We therefore have the following formula for the a_n in (4):

$$a_n = \left(n + \frac{1}{2}\right) \int_{-1}^{1} f(x) P_n(x)\, dx. \qquad (5)$$

These manipulations are easy to justify if f(x) is known in advance to have a series expansion of the form (4) and this series is integrable term by term on the interval $-1 \leq x \leq 1$. Both conditions are obviously satisfied when f(x) is a polynomial.

Problems

1. Find the first three terms of the Legendre series of

   a. $f(x) = \begin{cases} 0, if\ -1 \leq x < 0 \\ x, if\ 0 \leq x \leq 1; \end{cases}$ b. $f(x) = e^x$

   Soln. (a) $f(x) = \frac{1}{4} p_0(x) + \frac{1}{2} p_1(x) + \frac{5}{16} p_2(x) + \ldots$

   (b) $f(x) = \frac{1}{2}\left(e - e^{-1}\right) p_0(x) + 3e^{-1}\, p_1(x) + \frac{1}{2}\left(5e - 35e^{-1}\right) p_2(x) + \ldots$

MODULE II

II.3. Bessel Function – The Gamma Function

The differential equation

$$x^2y^{11} + xy^1 + (x^2 - p^2)y = 0 \tag{1}$$

where p is a non negative constant, is called Bessel's equation and its solutions are known as Bessel functions.

Finding the Bessel function – $J_p(x)$

We begin our study of the solutions of (1) by noticing that after division by x^2 the coefficients of y^1 and y are $P(x) = 1/x$ and $Q(x) = (x^2 - p^2)/x^2$, so $xP(x) = 1$ and $x^2 Q(x) = - p^2 + x^2$. The origin is therefore a regular singular point, the indicial equation is $m^2 - p^2 = 0$, and the exponents are $m_1 = p$ and $m_2 = -p$. It follows from that equation (1) has a solution of the form

$$y = x^p \sum a_n x^n = \sum a_n x^{n+p}, \tag{2}$$

where $a_0 \neq 0$ and the power series $\sum a_n x^n$ converges for all x. To find this solution, we write

$$y^1 = \sum (n+p)\, a_n x^{n+p-1}$$

and

$$y^{11} = \sum (n+p-1)\,(n+p)\, a_n x^{n+p-2}.$$

These formulas enable us to express the terms on the left side of equation (1) in the form.

$$x^2\, y^{11} = \sum (n+p-1)(n+p) a_n x^{n+p},$$

$$xy^1 = \sum (n+p)\, a_n x^{n+p},$$

$$x^2 y = \sum a_{n-2}\, x^{n+p},$$

$$-p^2 y = \sum -p^2 a_n x^{n+p}.$$

If we add these series and equate to zero the coefficient of x^{n+p}, then after a little simplification we obtain the following recursion formula for the a_n:

$$n(2p + n)\, a_n + a_{n-2} = 0 \tag{3}$$

or

$$a_n = -\frac{a_{n-2}}{n(2p+n)}. \tag{4}$$

We known that a_0 is nonzero and arbitrary. Since $a_{-1} = 0$, (4) tells us that $a_1 = 0$; and repeated application of (4) yields the fact that $a_n = 0$ for every odd subscript n. The non zero coefficients of our solution (2) are therefore.

$$a_0, \qquad a_2 = -\frac{a_0}{2(2p+2)}$$

$$a_4 = -\frac{a_2}{4(2p+4)} = \frac{a_0}{2.4(2p+2)(2p+4)}$$

$$a_6 = -\frac{a_4}{6(2p+6)} = -\frac{a_0}{2\cdot 4\cdot 6(2p+2)(2p+4)(2p+6)}, \ldots$$

and the solution itself is

$$y = a_0x^p\left[1-\frac{x^2}{2^2\,(p+1)}+\frac{x^4}{2^4\,2!(p+1)(p+2)}-\frac{x^6}{2^6 3!(p+1)(p+2)(p+3)}+\ldots\right]$$

$$= a_0x^p \sum_{n=0}^{\infty}(-1)^n \frac{x^{2n}}{2^{2n}\,n!(p+1)\ldots(p+n)} \tag{5}$$

The Bessel function of the first kind of order p, denoted by J_P (x), is defined by putting $a_0 = 1/2^P p!$ in (5), so that

$$J_p(x) = \frac{x^p}{2^p\,p!}\sum_{n=0}^{\infty}(-1)^n \frac{x^{2n}}{2^{2n} n!(p+1)\ldots(p+n)}$$

$$= \sum_{n=0}^{\infty}(-1)^n \frac{(x/2)^{2n+p}}{n!(p+n)!} \tag{6}$$

$J_P(x)$ as defined by (6) is meaningless unless the non-negative real number p is an integer, since only in this case has any meaning been assigned to the factors (p+n)! in the denominators.

The gamma function. The purpose of this digression is to give a reasonable and useful meaning to p! [and more generally to (p+n)! for n=0, 1, 2, …] when the non-negative real number p is not an integer. We accomplish this by introducing the gamma function Γ(p), defined by

$$\Gamma(p) = \int_0^\infty t^{p-1} e^{-t}\, dt,\ p>0. \tag{9}$$

The factor $e^{-t} \to 0$ so rapidly as $t \to \infty$ that this improper integral converges at the upper limit regardless of the value of p. However, at the lower limit we have $e^{-t} \to 1$, and the factor $t^{p-1} \to \infty$ whenever $p<1$. The restriction that p must be positive is necessary in order to guarantee convergence at the lower limit.

It is easy to see that

$$\Gamma(p + 1) = p\Gamma(p); \tag{10}$$

for integration by parts yields.

$$\Gamma(p + 1) = \lim_{b \to \infty} \int_0^b t^p e^{-t}\, dt$$

$$= \lim_{b \to \infty} \left(-t^p e^{-t} \Big|_0^b + p \int_0^b t^{p-1} e^{-t}\, dt \right)$$

$$= p \left(\lim_{b \to \infty} \int_0^b t^{p-1} e^{-t}\, dt \right) = p\Gamma(p),$$

Since $b^p/e^b \to 0$ as $b \to \infty$. If we use the fact that

$$\Gamma(1) = \int_0^\infty e^{-t}\, dt = 1, \tag{11}$$

then (10) yields $\Gamma(2) = 1\,\Gamma(1) = 1$, $\Gamma(3) = 2\Gamma(2) = 2\cdot1$, $\Gamma(4) = 3\Gamma(3) = 3\cdot2\cdot1$, and in general.

$$\Gamma(n + 1) = n!. \tag{12}$$

for any integer $n \geq 0$.

We began our discussion of the gamma function under the assumption that p is non-negative, and we mentioned at the outset that the integral (9) does not exist if p=0. However, we can define $\Gamma(p)$ for many negative p's without the aid of this integral if we write (10) in the form

$$\Gamma(p) = \frac{\Gamma(p+1)}{p}. \tag{13}$$

This extension of the definition is necessary for the applications, and it begins as follows: If $-1<p<0$, then $0 < p + 1 < 1$, so the right side of equation (13) has a value and the left side of (13) is defined to have the value given by the right side. The next step is to notice that if $-2 < p < -1$, then $-1 < p + 1 < 0$, so we can use (13) again to define $\Gamma(p)$ on the interval $-2 <p<-1$ in terms of the values of $\Gamma(p+1)$ already defined in the previous step. It is clear that this process can be continued indefinitely. Furthermore, it is easy to see from (11) that

$$\lim_{P \to 0} \Gamma(p) = \lim_{P \to 0} \frac{\Gamma(p+1)}{p} = \pm \infty$$

according as $p \to 0$ from the right or left.

Since $\Gamma(p)$ never vanishes, the function $1/\Gamma(p)$ will be defined and well behaved for all values of p if we agree that $1/\Gamma(p) = 0$ for $p=0, -1, -2, \dots$

These ideas enable us to define p! by

$$P! = \Gamma(p+1)$$

for all values of p except negative integers, and by formula (12) this function has its usual meaning when p is a non-negative integer. Its reciprocal, $1/p! = 1/\Gamma(p+1)$, is defined for all p's and has the value 0 whenever p is a negative integer.

The general solution of Bessel's equation.

We have found a particular solution of (1) corresponding to the exponent $m_1 = p$, namely, J_p (x). In order to find the general solution, we must now construct a second independent solution – that is, one that is not a constant multiple of $J_p(x)$. Any such solution is called a Bessel function of the second kind. The natural procedure is to try the other exponent, $m_2 = -p$. But in doing so, we expect to encounter difficulties whenever the difference $m_1 - m_2 = 2p$ is zero or a positive integer, that is, whenever the non-negative constant p is an integer or half an odd integer. It turns out that the expected difficulties are serious only in the first case.

We therefore begin by assuming that p is not an integer. In this case we replace p by –p in our previous treatment, and it is easy to see that the discussion goes through almost without change. The only exception is that (3) becomes.

$$n(-2p + n)\, a_n + a_{n-2} = 0;$$

and if it happens that p =1/2, then by letting n = 1 we see that there is no compulsion to choose $a_1 = 0$. However, since all we want is a particular solution, it is certainly permissible to put $a_1=0$. The same problem arises when p = 3/2 and n=3, so on; and we solve it by putting $a_1 = a_3 = \ldots = 0$ in all cases. Everything else goes as before, and we obtain a second solution.

$$J_{-p}(x) = \sum_{n=0}^{\infty} (-1)^n \frac{(x/2)^{2n-p}}{n!(-p+n)!} \qquad (15)$$

The first term of this series is

$$\frac{1}{(-p)!}\left(\frac{x}{2}\right)^{-p}$$

So $J_{-p}(x)$ is unbounded near x=0. Since $J_p(x)$ is bounded near x=0, these two solutions are independent and

$$y=c_1J_p(x) + c_2J_{-p}(x),\ \text{p not an integer}, \qquad (16)$$

is the general solution of (1).

The solution is entirely different when p is an integer m$\geq$ 0. Formula (15) now becomes

$$J_{-m}(x) = \sum_{n=0}^{\infty} (-1)^n \frac{(x/2)^{2n-m}}{n!(-m+n)!}$$

$$= \sum_{n=m}^{\infty} \frac{(-1)^n (x/2)^{2n-m}}{n!(-m+n)!}$$

since the factors 1/(–m + n)! are zero when n = 0, 1, …., m–1. On replacing the dummy variable n by n + m and compensating by beginning the summation at n=0, we obtain.

$$J_{-m}(x) = \sum_{n=0}^{\infty} (-1)^{n+m} \frac{(x/2)^{2(n+m)-m}}{(n+m)!\,n!}$$

$$= (-1)^m \sum_{n=0}^{\infty} (-1)^n \frac{(x/2)^{2n+m}}{n!\,(m+n)!}$$

$$= (-1)^m J_m(x).$$

This show that $J_{-m}(x)$ is not independent of $J_m(x)$, so in this case

$$y = c_1 J_m(x) + c_2 J_{-m}(x)$$

is not the general solution of (1), and the search continues.

By known methods, $J_m(x) \int \frac{dx}{xJ_m(x)^2}$

is a second solution independent of $J_m(x)$. It is customary, however, to proceed somewhat differently, as follows. When p is not an integer, any function of the form (16) with $c_2 \neq 0$ is a Bessel function of the second kind, including $J_{-p}(x)$ itself. The standard Bessel function of the second kind is defined by

$$Y_p(x) = \frac{J_p(x)\cos p\pi - J_{-p}(x)}{\sin p\pi}. \qquad (17)$$

(16) can certainly be written in the equivalent form

$$y = c_1 J_p(x) + c_2 Y_p(x), \quad \text{p not an integer.} \qquad (18)$$

We still have the problem of what to do when p is an integer m, for (17) is meaningless in this case. It turns out after detailed analysis that the function defined by.

$$Y_m(x) = \lim_{p \to m} Y_p(x) \qquad (19)$$

exists and is a Bessel function of the second kind; and it follows that

$$y = c_1 J_p(x) + c_2 Y_p(x) \qquad (20)$$

is the general solution of Bessel's equation in all cases, whether p is an integer or not.

If we are interested only in solutions of Bessel's equation that are bounded near x=0, and this is often the case in the applications, then we must take $c_2 = 0$ in (20).

Note:

Bessel function of order '0' and '1'are respectively

$$J_0(x) = \sum_{n=0}^{\infty}(-1)^n \frac{1}{(n!)^2}\left(\frac{x}{2}\right)^{2n} = 1-\frac{x^2}{2^2}+\frac{x^4}{2^2\,4^2}-\frac{x^6}{2^2\,4^2\,6^2}+\ldots$$

and $$J_1(x) = \sum_{n=0}^{\infty}(-1)^n \frac{1}{n!(n+1)!}\left(\frac{x}{2}\right)^{2n+1} = \frac{x}{2}-\frac{1}{1!\,2!}\left(\frac{x}{2}\right)^3+\frac{1}{2!\,3!}\left(\frac{x}{2}\right)^5-\ldots$$

Problems

I. Show that (a) $\frac{d}{dx}J_0(x) = -J_1(x)$ and (b) $\frac{d}{dx}[xJ_1(x)] = xJ_0(x)$

II. Show that $\Gamma 1/2 = \sqrt{\pi}$

(Hint: by definition $\Gamma 1/2 = \int_0^{\infty} t^{-1/2} e^{-t}\,dt$

Change of variable $t=s^2$ leads to $\Gamma(1/2) = 2\int_0^{\infty} e^{-s^2}\,ds.$

Since s is a dummy variable, $\Gamma\left(\frac{1}{2}\right)^2 = 4\left(\int_0^{\infty} e^{-x^2}\,dx\right)\left(\int_0^{\infty} e^{-y^2}\,dy\right)$

$$= 4\int_0^{\infty}\int_0^{\infty} e^{-(x^2+y^2)}\,dx\,dy$$

Taking x=r cos θ, y = r sin θ and changing the integral in to polar coordinates,

$$\Gamma\left(\frac{1}{2}\right)^2 = 4\int_0^{\pi/2}\int_0^{\infty} e^{-r^2}\,r\,dr\,d\theta = \pi, \text{so that}\ \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

III. Prove that $\left(n+\frac{1}{2}\right)! = \frac{(2n+1)!}{2^{2n+1}\,n!}\sqrt{\pi}$

II.4 Properties of Bessel functions.

The Bessel function $J_p(x)$ has been defined for any real number p by

$$J_p(x) = \sum_{n=0}^{\infty}(-1)^n \frac{(x/2)^{2n+p}}{n!(p+n)!}. \qquad (1)$$

In this section we develop several properties of these functions that are useful in their applications.

Identities and the functions $J_{m+1/2}(x)$. We begin by considering the formulas.

$$\frac{d}{dx}\left[x^p J_p(x)\right] = x^p J_{p-1}(x) \qquad (2)$$

and

$$\frac{d}{dx}\left[x^{-p} J_p(x)\right] = -x^{-p} J_{p+1}(x) \qquad (3)$$

To establish (2), we simply multiply the series (1) by x^p and differentiate;

$$\frac{d}{dx}\left[x^P J_p(x)\right] = \frac{d}{dx}\sum_{n=0}^{\infty}\frac{(-1)^n x^{2n+2p}}{2^{2n+p}\, n!(p+n)!}$$

$$= \sum_{n=0}^{\infty}\frac{(-1)^n x^{2n+2p-1}}{2^{2n+p-1}\, n!\,(p+n-1)!}$$

$$= x^p \sum_{n=0}^{\infty} (-1)^n \frac{(x/2)^{2n+p-1}}{n!\,(p-1+n)!} = x^p J_{p-1}(x)$$

The verification of (3) is similar.

If the differentiations in (2) and (3) are carried out, and the results are divided by $x^{\pm p}$, then the formulas become,

$$J_p^1(x) + \frac{p}{x} J_p(x) = J_{p-1}(x) \qquad (4)$$

and

$$J_p^1(x) - \frac{p}{x} J_p(x) = J_{p+1}(x) \qquad (5)$$

If (4) and (5) are first added and then subtracted, the results are

$$2J_p^1(x) = J_{p-1}(x) - J_{p+1}(x) \qquad (6)$$

and

$$\frac{2p}{x} J_p(x) = J_{p-1}(x) + J_{p+1}(x) \qquad (7)$$

These formulas enable us to express Bessel functions and their derivatives in terms of other Bessel functions.

Orthogonality property of Bessel function.

Suppose that λ_n are the positive zeros of some fixed Bessel function $J_P(x)$ with $P \geq 0$. Then $\int_0^1 xJ_p(\lambda_m x)J_p(\lambda_n x)\,dx = \begin{cases} 0 & if\ m \neq n \\ \frac{1}{2}J_{p+1}(\lambda_n)^2 & if\ m = n \end{cases}$ (11)

Proof

To establish (11), we begin with the fact that $y = J_p(x)$ is a solution of

$$y^{11} + \frac{1}{x}y^1 + \left(1 - \frac{p^2}{x^2}\right)y = 0$$

If a and b are distinct positive constants, it follows that the functions $u(x) = J_p(ax)$ and $v(x) = J_p(bx)$ satisfy the equations.

$$u^{11} + \frac{1}{x}u^1 + \left(a^2 - \frac{p^2}{x^2}\right)u = 0 \qquad (12)$$

and

$$v^{11} + \frac{1}{x}v^1 + \left(b^2 - \frac{p^2}{x^2}\right)v = 0. \qquad (13)$$

We now multiply these equations by v and u, the subtract the results, to obtain.

$$\frac{d}{dx}(u^1 v - v^1 u) + \frac{1}{x}(u^1 v - v^1 u) = (b^2 - a^2)uv;$$

and after multiplication by x, this becomes

$$\frac{d}{dx}[x(u^1 v - v^1 u)] = (b^2 - a^2)xuv. \qquad (14)$$

When (14) is integrated from x = 0 to x = 1, we get

$$(b^2 - a^2)\int_0^1 xuv\,dx = [x(u^1 v - v^1 u)]_0^1$$

The expression in brackets clearly vanishes at x = 0, and at the other end of the interval we have $u(1) = J_p(a)$ and $v(1) = J_p(b)$. It therefore follows that the integral on the left is zero if a and b are distinct positive zeros λ_m and λ_n of $J_p(x)$; that is, we have

$$\int_0^1 xJ_p(\lambda_m x)\, J_p(\lambda_n x)\, dx = 0, \qquad (15)$$

which is the first part of (11).

Our final task is to evaluate the integral in (15) when m = n. If (12) is multiplied by $2x^2u^1$, it becomes

$$2x^2 u^1 u^{11} + 2xu^{1^2} + 2a^2 x^2 uu^1 - 2p^2uu^1 = 0$$

or

$$\frac{d}{dx}\left(x^2u^{1^2}\right) + \frac{d}{dx}(a^2x^2u^2) - 2a^2 xu^2 - \frac{d}{dx}(p^2u^2) = 0.$$

So on integrating from x = 0 to x = 1, we obtain

$$2a^2 \int_0^1 xu^2\, dx = [x^2u^{1^2} + (a^2 x^2 - p^2)\, u^2]_0^1 \qquad (16)$$

When x = 0, the expression in brackets vanishes; and since $u^1(1) = aJ_p^1(a)$,

(16) yields

$$\int_0^1 xJ_p(ax)^2\, dx = \frac{1}{2} J_p^1(a)^2 + \frac{1}{2}\left(1-\frac{p^2}{a^2}\right)J_p(a)^2.$$

We now put $a = \lambda_n$ and get

$$\int_0^1 xJ_p(\lambda_n x)^2\, dx = \frac{1}{2} J_p^1(\lambda_n)^2 = \frac{1}{2} J_{p+1}(\lambda_n)^2.$$

Where the last step makes use of (5), and the proof of (11) is complete.

Bessel series

Suppose that f(x) is a function defined on the interval $0 \leq x \leq 1$ and λ_n are the positive zeros of some fixed Bessel function $J_p(x)$ with $p \geq 0$.

The series $f(x) = \sum_{n=1}^{\infty} a_n J_p(\lambda_n x) = a_1 J_p(\lambda_1 x) + a_2 J_p(\lambda_2 x) + \ldots$ (17)

where the coefficient a_n are calculated from

$$a_n = \frac{2}{J_{p+1}(\lambda_n)^2} \int_0^1 x\, f(x)\, J_p(\lambda_n x)\, dx \qquad (18)$$

is called the Bessel series – or sometimes the Fourier – Bessel series – of the function f(x).

Proof.

Hint: If an expansion of the form (17) is assumed to be possible, then multiplying throughout by $xJ_p(\lambda_m x)$, formally integrating term by term from 0 to 1, and using (11) yields $\int_0^1 x\, f(x)\, J_p(\lambda_m x)\, dx = \frac{a_m}{2} J_{p+1}(\lambda_m)^2$ and on replacing m by n, we obtain the formula (18) for a_n.

Theorem A (Bessel expansion theorem – conditions under which series (17) converges and has the sum f(x))

Assume that f(x) and $f^1(x)$ have atmost a finite number of jump discontinuities on the interval $0 \leq x \leq 1$. If $0 < x < 1$, then Bessel series (17) converges to f(x) when x is a point of continuity of this function, and converges to

$\frac{1}{2}\left[f(x^-) + f(x^+)\right]$ when x is a point of discontinuity.

Note: At x = 1, the series (17) converges to zero regardless of the nature of the function because every $J_p(\lambda_n)$ is zero. The series also converges at x=0, to zero if p>0 and to f(0+) if p=0.

Eg: Compute the Bessel series of the function f(x) = 1 for the interval $0 \leq x \leq 1$ in terms of the functions $J_0(\lambda_n x)$, where λ_n are the positive zeros of $J_0(x)$.

Soln. Let $1 = \sum_{n=1}^{\infty} a_n J_0(\lambda_n x)$ be the required Bessel series. Then replacing p=0 and f(x)=1 in (18), $a_n = \frac{2}{J_1(\lambda_n)^2}\int_0^1 J_0(\lambda_n x)dx$ (19)

Now from (2), $\frac{d}{dx}[x^p J_p(x)] = x^p J_{p-1}(x)$ so that $\frac{d}{dx}(xJ_1(x)) = xJ_0(x)$

So $xJ_1(x) = \int xJ_0(x)\,dx + c$ or $\int xJ_0(x)\,dx = x J_1(x) + c$.

$\therefore \int_0^1 xJ_0(\lambda_n x)\,dx = \left[\frac{1}{\lambda_n} x J_1(\lambda_n x)\right]_0^1 = \frac{J_1(\lambda_n)}{\lambda_n}$. Substituting in (19),

$$a_n = \frac{2}{J_1(\lambda_n)^2}\,\frac{J_1(\lambda_n)}{\lambda_n} = \frac{2}{\lambda_n J_1(\lambda_n)}.$$

$\therefore$ Required Bessel series expansion is $1 = \sum_{n=1}^{\infty} \frac{2}{\lambda_n J_1(\lambda_n)} J_0(\lambda_n x)$.

Problems

1. Verify $\frac{d}{dx}[x^{-p} J_p(x)] = -x^{-p} J_{p+1}(x)$
2. Express $J_2(x)$, $J_3(x)$ and $J_4(x)$ in terms of $J_0(x)$ and $J_1(x)$.
3. If f(x) is defined by $f(x) = \begin{cases} 1 & 0 \leq x < 1/2 \\ 1/2 & x = 1/2 \\ 0 & 1/2 < x \leq 1 \end{cases}$ then show that

   $f(x) = \sum_{n=1}^{\infty} \frac{J_1(\lambda_n/2)}{\lambda_n J_1(\lambda_n)^2} J_0(\lambda_n x)$, where λ_n are the positive zeros of $J_0(x)$.
4. If $f(x) = x^p$ for the interval $0 \leq x < 1$, show that its Bessel series in the functions $J_p(\lambda_n x)$, where λ_n are positive zeros of $J_p(x)$, is $x^p = \sum_{n=1}^{\infty} \frac{2}{\lambda_n J_{p+1}(\lambda_n)} J_p(\lambda_n x)$

CHAPTER III

SYSTEMS OF FIRST ORDER EQUATIONS

III.1 Linear Systems

Let, x, y be variables depending on the independent variable 't'. Consider the following system of first order differential equations.

$$\begin{cases} dx/dt = F(t, x, y) \\ dy/dt = G(t, x, y) \end{cases} \tag{1}$$

We specialize even further, to linear systems, of the form.

$$\begin{cases} dx/dt = a_1(t)\,x + b_1(t)\,y + f_1(t) \\ dy/dt = a_2(t)\,x + b_2(t)\,y + f_2(t) \end{cases} \tag{2}$$

We shall assume in the present discussion and in the theorems stated below, that the functions $a_i(t)$, $b_i(t)$ and $f_i(t)$, i=1, 2, are continuous on a certain closed internal [a, b] of the t-axis.

If $f_1(t)$ and $f_2(t)$ are identically zero, then the system (2) is called homogeneous; otherwise it is said to be non homogeneous.

A solution of (2) on [a, b] is of course a pair of functions x(t) and y(t) that satisfy both equations of (2) throughout this interval.

Eg: x=$e^{3t}, y=e^{3t}$ and x = $e^{2t}, y=2e^{2t}$ are solutions of the system $\left.\begin{matrix} dx/dt = 4x - y \\ dy/dt = 2x + y \end{matrix}\right\}$ (3)

on any interval.

Theorem A. (Fundamental existence and uniqueness theorem)

If t_0 is any point of the interval [a, b], and if x_0 and y_0 are any numbers whatever, then (2) has one and only one solution $\begin{cases} x = x(t) \\ y = y(t) \end{cases}$

valid throughout [a, b] such that $x(t_0) = x_0$ and $y(t_0) = y_0$.

Theorem B

If the homogeneous system $\begin{cases} dx/dt = a_1(t)\,x + b_1(t)\,y \\ dy/dt = a_2(t)\,x + b_2(t)\,y \end{cases}$ (5)

has two solution $\begin{cases} x = x_1(t) \\ y = y_1(t) \end{cases}$ and $\begin{cases} x = x_2(t) \\ y = y_2(t) \end{cases}$ (6)

on [a,b], then $\begin{cases} x = c_1\,x_1(t) + c_2\,x_2(t) \\ y = c_1\,y_1(t) + c_2\,y_2(t) \end{cases}$ (7)

is also a solution on [a, b] for any constants c_1 and c_2.

Theorem C

If the two solutions (6) of the homogeneous system (5) have a wronskain,

w(t) = $\begin{vmatrix} x_1(t) & x_2(t) \\ y_1(t) & y_2(t) \end{vmatrix}$ that (8)

does not vanish on [a, b], then (7) is the general solution of (5) on this interval.

Theorem D

If W(t) is the Wronskian of the two solutions (6) of the homogeneous system (5), then W(t) is either identically zero or nowhere zero on [a, b].

<u>Proof.</u>

A simple calculation shows that W(t) satisfies the first order differential equation

$$\frac{dW}{dt} = [a_1(t) + b_2(t)]W, \quad (9)$$

from which it follows that

$$W(t) = ce^{\int [a_1(t)+b_2(t)]dt} \quad (10)$$

for some constant c. The conclusion of the theorem is now evident from the fact that the exponential factor in (10) never vanishes on [a,b].

Note:

The two solutions (6) are called linearly dependent on [a, b] if one is a constant multiple of the other in the sense that

$$\begin{aligned} x_1(t) &= kx_2(t) \\ y_1(t) &= ky_2(t) \end{aligned} \quad \text{or} \quad \begin{aligned} x_2(t) &= kx_1(t) \\ y_2(t) &= ky_1(t) \end{aligned}$$

for some constant k and all t in [a, b], and linearly independent if neither is a constant multiple of the other. It is clear that linear dependence is equivalent to the condition that there exist two constants c_1 and c_2, at least one of which is not zero, such that.

$$\begin{aligned} c_1\, x_2(t) + c_2\, x_2(t) &= 0 \\ c_1\, y_1(t) + c_2\, y_2(t) &= 0 \end{aligned} \tag{11}$$

for all t in [a, b].

Theorem E

If the two solutions (6) of the homogeneous system (5) are linearly independent on [a, b], then (7) is the general solution of (5) on this interval.

Proof.

In view of Theorems C and D, it suffices to show that the solutions (6) are linearly dependent if and only if their Wronskian W(t) is identically zero. We begin by assuming that they are linearly dependent, so that, say,

$$\begin{aligned} x_1(t) &= kx_2(t) \\ y_1(t) &= ky_2(t) \end{aligned} \tag{12}$$

Then

$$W(t) = \begin{vmatrix} x_1(t) & x_2(t) \\ y_1(t) & y_2(t) \end{vmatrix} = \begin{vmatrix} kx_2(t) & x_2(t) \\ ky_2(t) & y_2(t) \end{vmatrix}$$

$$= kx_2(t)\, y_2(t) - kx_2(t)\, y_2(t) = 0$$

for all t in [a, b]. The same argument works equally well if the constant k is on the other side of equations (12). We now assume that W(t) is identically zero, and show that the solutions (6) are linearly dependent in the sense of equations (11).

Let t_0 be a fixed point in [a, b]. Since $W(t_0) = 0$, the system of linear algebraic equations.

$$c_1 x_1(t_0) + c_2 x_2(t_0) = 0$$
$$c_1 y_1(t_0) + c_2 y_2(t_0) = 0$$

has a solution c_1, c_2 in which these numbers are not both zero. Thus, the solution of (5) given by

$$\begin{cases} x = c_1 x_1(t) + c_2 x_2(t) \\ y = c_1 y_1(t) + c_2 y_2(t) \end{cases} \tag{13}$$

equals the trivial solution at t_0. It now follows from the uniqueness part of Theorem A that (13) must equal the trivial solution throughout the interval [a, b], so (11) holds and the proof is complete.

Theorem F

If the two solutions (6) of the homogeneous system (5) are linearly independent on [a, b], and if

$$\begin{cases} x = x_p(t) \\ y = y_p(t) \end{cases}$$

is any particular solution of (2) on this interval, then

$$\begin{cases} x = c_1 x_1(t) + c_1 x_2(t) + x_p(t) \\ y = c_1 y_1(t) + c_2 y_2(t) + y_p(t) \end{cases} \tag{14}$$

is the general solution of (2) on [a, b].

Proof: It suffices to show that if

$$\begin{cases} x = x(t) \\ y = y(t) \end{cases}$$

is an arbitrary solution of (2), then

$$\begin{cases} x = x(t) - x_p(t) \\ y = y(t) - y_p(t) \end{cases}$$

is a solution of (5), and this we leave to the reader.

Problems:

1. a. Show that

$$\begin{cases} x = e^{4t} \\ y = e^{4t} \end{cases} \text{ and } \begin{cases} x = e^{-2t} \\ y = -e^{-2t} \end{cases}$$

are solution of the homogeneous system

$$\begin{cases} \dfrac{dx}{dt} = x + 3y \\ \dfrac{dy}{dt} = 3x + y \end{cases}$$

b. Show in two ways that the given solutions of the system in (a) are linearly independent on every closed interval, and write the general solution of this system.

<u>Soln:</u> General solution is $\begin{cases} x = c_1 e^{4t} + c_2 e^{-2t} \\ y = c_1 e^{4t} - c_2 e^{-2t} \end{cases}$

c. Find the particular solution.

$$\begin{cases} x = x(t) \\ y = y(t) \end{cases}$$

of this system for which x(0) = 5 and y (0) = 1.

<u>Soln:</u> $\begin{cases} x = 3e^{4t} + 2e^{-2t} \\ y = 3e^{4t} - 2e^{-2t} \end{cases}$

2. Show that

$$\begin{cases} x = 2e^{4t} \\ y = 3e^{4t} \end{cases} \text{ and } \begin{cases} x = e^{-t} \\ y = -e^{-t} \end{cases}$$

are solutions of the homogeneous system

$$\begin{cases} \dfrac{dx}{dt} = x + 2y \\ \dfrac{dy}{dt} = 3x + 2y \end{cases}$$

III.2. Homogeneous linear systems with constant coefficients.

We are now in a position to give a complete explicit solution of the simple system.

$$\begin{cases} \dfrac{dx}{dt} = a_1 x + b_1 y \\ \dfrac{dy}{dt} = a_2 x + b_2 y \end{cases} \qquad (1)$$

where a_1, b_1, a_2 and b_2 are given constants. Some of the problems at the end of the previous section illustrate a procedure that can often be applied to this case: differentiate one equation, eliminate one of the dependent variables, and solve the resulting second order linear equation. The method we now describe is based instead on constructing a pair of linearly independent solutions directly from the given system.

If we recall that the exponential function has the property that its derivatives are constant multiples of the function itself, then it is natural to work solutions of (1) having the form.

$$\begin{cases} x = Ae^{mt} \\ y = Be^{mt} \end{cases} \qquad (2)$$

If we substitute (2) into (1) we get

$$Ame^{mt} = a_1Ae^{mt} + b_1 Be^{mt}$$

$$Bme^{mt} = a_2Ae^{mt} + b_2 Be^{mt};$$

and dividing by e^{mt} yields the linear algebraic system,

$$(a_1 - m) A + b_1 B = 0$$

$$a_2A + (b_2 - m) B = 0 \qquad (3)$$

in the unknowns A and B. It is clear that (3) has the trivial solution A=B=0, which makes (2) the trivial solution of (1). Since we are looking for nontrivial solutions of (1), this is no help at all. However, we know that (3) has nontrivial solutions whenever the determinant of the coefficients vanishes, i.e., whenever

$$\begin{vmatrix} a_1 - m & b_1 \\ a_2 & b_2 - m \end{vmatrix} = 0.$$

When this determinant is expanded, we get the quadratic equation

$$m^2 - (a_1 + b_2)\, m + (a_1 b_2 - a_2 b_1) = 0 \tag{4}$$

for the unknown m. By analogy with our previous work, we call this the auxiliary equation of the system (1). Let m_1 and m_2 be the roots of (4). If we replace m in (3) by m_1, then we know that the resulting equation have a nontrivial solution A_1, B_1, so

$$\begin{cases} x = A_1 e^{m_1 t} \\ y = B_1 e^{m_1 t} \end{cases} \tag{5}$$

is a nontrivial solution of the system (1). By proceeding similarly with m_2, we find another nontrivial solution.

$$\begin{cases} x = A_2 e^{m_2 t} \\ y = B_2 e^{m_2 t} \end{cases} \tag{6}$$

In order to make sure that we obtain two linearly independent solutions – and hence the general solution – it is necessary to examine in detail each of the three possibilities for m_1 and m_2.

Distinct real roots. When m_1 and m_2 are distinct real numbers, then (5) and (6) are easily seen to be linearly independent (why?) and

$$\begin{cases} x = c_1 A_1 e^{m_1 t} + c_2 A_2 e^{m_2 t} \\ y = c_1 B_1 e^{m_1 t} + c_2 B_2 e^{m_2 t} \end{cases} \tag{7}$$

is the general solution of (1).

Example 1. In the case of the system

$$\begin{cases} \dfrac{dx}{dt} = x + y \\ \dfrac{dy}{dt} = 4x - 2y \end{cases} \tag{8}$$

(3) is $\quad (1-m)\,A + B = 0$

$$4A + (-2 - m)\,B = 0. \tag{9}$$

The auxiliary equation here is

$$m^2 + m - 6 = 0 \quad \text{or} \quad (m + 3)\,(m - 2) = 0,$$

So m_1 and m_2 are –3 and 2. With m = –3, (9) becomes

$$4A + B = 0$$

$$4A + B = 0.$$

A simple nontrivial solution of this system is A=1, B= –4, so we have

$$\begin{cases} x = e^{-3t} \\ y = -4e^{-3t} \end{cases} \tag{10}$$

as a nontrivial solution of (8). With m = 2, (9) becomes

$$-A + B = 0$$

$$4A - 4B = 0,$$

and a simple nontrivial solution is A = 1, B = 1. This yields

$$\begin{cases} x = e^{2t} \\ y = e^{2t} \end{cases} \tag{11}$$

as another solution of (8); and since it is clear that (10) and (11) are linearly independent,

$$\begin{cases} x = c_1 e^{-3t} + c_2 e^{2t} \\ y = -4c_1 e^{-3t} + c_2 e^{2t} \end{cases} \tag{12}$$

is the general solution of (8).

Distinct complex roots. If m_1 and m_2 are distinct complex numbers, then they can be written in the form a ± ib where a and b are real numbers and b ≠ 0. In this case we expect the A's and B's obtained from (3) to be complex numbers, and we have two linearly independent solution.

$$\begin{cases} x = A_1^* e^{(a+ib)t} \\ y = B_1^* e^{(a+ib)t} \end{cases} \text{ and } \begin{cases} x = A_2^* e^{(a-ib)t} \\ y = B_2^* e^{(a-ib)t} \end{cases} \tag{13}$$

However, these are complex – valued solutions, and to extract real-valued solutions we proceed as follows. If we express the numbers $A_1^* = A_1 + iA_2$ and $B_1^* = B_1 + iB_2$, and use Euler's formula, then the first of the solutions (13) can be written as

$$\begin{cases} x=(A_1+iA_2)e^{at}\,(\cos bt+i\sin bt) \\ y=(B_1+iB_2)e^{at}\,(\cos bt+i\sin bt) \end{cases}$$

or

$$\begin{cases} x=e^{at}[(A_1\cos bt-A_2\sin bt)+i(A_1\sin bt+A_2\cos bt)] \\ y=e^{at}[(B_1\cos bt-B_2\sin bt)+i(B_1\sin bt+B_2\cos bt)] \end{cases} \tag{14}$$

It is easy to see that if a pair of complex valued functions is a solution of (1), in which the coefficients are real constants, then their two real parts and their two imaginary parts are real valued solutions. It follows from this that (14) yields the two real valued solutions.

$$\begin{cases} x=e^{at}\,(A_1\cos bt-A_2\sin bt) \\ y=e^{at}\,(B_1\cos bt-B_2\sin bt) \end{cases} \tag{15}$$

and

$$\begin{cases} x=e^{at}\,(A_1\,Sin\,bt+A_2\cos bt) \\ y=e^{at}\,(B_1\,Sin\,bt+B_2\cos bt) \end{cases} \tag{16}$$

It can be shown that these solutions are linearly independent, so the general solution in this case is

$$\begin{cases} x=e^{at}\,[c_1\,(A_1\cos bt-A_2\sin bt)+c_2\,(A_1\sin bt+A_2\cos bt)] \\ y=e^{at}\,[c_1\,(B_1\cos bt-B_2\sin bt)+c_2\,(B_1\sin bt+B_2\cos bt)] \end{cases} \tag{17}$$

Since we have already found the general solution, it is not necessary to consider the second of the two solutions (13).

Equal real roots. When m_1 and m_2 have the same value m, then (5) and (6) are not linearly independent and we essentially have only one solution

$$\begin{cases} x=Ae^{mt} \\ y=Be^{mt} \end{cases} \tag{18}$$

Our experience would lead us to expect a second linearly independent solution of the form

$$\begin{cases} x = Ate^{mt} \\ y = Bte^{mt} \end{cases}$$

Unfortunately the matter is not quite as simple as this, and we must actually look for a second solution of the form

$$\begin{cases} x = (A_1 + A_2 t)e^{mt} \\ y = (B_1 + B_2 t)e^{mt} \end{cases} \tag{19}$$

so that the general solution is

$$\begin{cases} x = c_1 Ae^{mt} + c_2 (A_1 + A_2 t)e^{mt} \\ x = c_1 Be^{mt} + c_2 (B_1 + B_2 t)e^{mt} \end{cases} \tag{20}$$

The constants A_1, A_2, B_1, and B_2 are found by substituting (19) into the system (1). Instead of trying to carry this through in the general case, we illustrate the method by showing how it works in a simple example.

Example 2. In the case of the system

$$\begin{cases} \dfrac{dx}{dt} = 3x - 4y \\ \dfrac{dy}{dt} = x - y \end{cases} \tag{21}$$

(3) is $$(3 - m) A - 4B = 0$$

$$A + (-1 - m) B = 0. \tag{22}$$

The auxiliary equation is

$$m^2 - 2m + 1 = 0 \quad \text{or} \quad (m-1)^2 = 0,$$

Which has equal real roots 1 and 1. With m = 1, (22) becomes

$$2A - 4B = 0$$

$$A - 2B = 0.$$

A simple nontrivial solution of this system is A = 2, B = 1, so

$$\begin{cases} x = 2e^t \\ y = e^t \end{cases} \tag{23}$$

is a nontrivial solution of (21). We now seek a second linearly independent solution of the form.

$$\begin{cases} x = (A_1 + A_2 t)e^t \\ y = (B_1 + B_2 t)e^t \end{cases} \tag{24}$$

When this is substituted into (21), we obtain

$$(A_1 + A_2 t + A_2)e^t = 3(A_1 + A_2 t)e^t - 4(B_1 + B_2 t)e^t$$

$$(B_1 + B_2 t + B_2)e^t = (A_1 + A_2 t)e^t - (B_1 + B_2 t)e^t,$$

which reduces at once to

$$(2A_2 - 4B_2)t + (2A_1 - A_2 - 4B_1) = 0$$

$$(A_2 - 2B_2)t + (A_1 - 2B_1 - B_2) = 0.$$

Since these are to be identities in the variable t, we must have

$$2A_2 - 4B_2 = 0 \qquad 2A_1 - A_2 - 4B_1 = 0$$

$$A_2 - 2B_2 = 0, \qquad A_1 - 2B_1 - B_2 = 0.$$

The two equation on the left have $A_2 = 2$, $B_2 = 1$ as a simple nontrivial solution. With this, the two equation on the right become

$$2A_1 - 4B_1 = 2$$

$$A_1 - 2B_1 = 1,$$

so we may take $A_1 = 1$, $B_1 = 0$. We now insert these numbers into (24) and obtain

$$\begin{cases} x = (1 + 2t)e^t \\ y = te^t \end{cases} \tag{25}$$

as our second solution. It is obvious that (23)and (25) are linearly independent, so

$$\begin{cases} x = 2c_1 e^t + c_2 (1 + 2t)e^t \\ y = c_1 e^t + c_2 te^t \end{cases} \tag{26}$$

is the general solution of the system (21).

Problems

1. Use the methods described in this section to find the general solution of each of the following system:

(a) $\begin{cases} \frac{dx}{dt}=4x-2y \\ \frac{dy}{dt}=5x+2y; \end{cases}$

(b) $\begin{cases} \frac{dx}{dt}=5x+4y \\ \frac{dy}{dt}=-x+y; \end{cases}$

(c) $\begin{cases} \frac{dx}{dt}=4x-3y \\ \frac{dy}{dt}=8x-6y; \end{cases}$

(d) $\begin{cases} \frac{dx}{dt}=2x \\ \frac{dy}{dt}=3y; \end{cases}$

(e) $\begin{cases} \frac{dx}{dt}=-4x-y \\ \frac{dy}{dt}=x-2y; \end{cases}$

(f) $\begin{cases} \frac{dx}{dt}=7x+6y \\ \frac{dy}{dt}=2x+6y; \end{cases}$

(g) $\begin{cases} \frac{dx}{dt}=x-2y \\ \frac{dy}{dt}=4x+5y. \end{cases}$

Solution:I. a) $\begin{cases} x= e^{3t}[2c_1 \cos 3t +2c_2 \sin 3t] \\ y=e^{3t}[c_1(\cos 3t+3\sin 3t)+ c_2(\sin 3t-3\cos 3t)] \end{cases}$

$b)$ $\begin{cases} x = -2\, c_1 e^{3t} + c_2(1+2t)e^{3t} \\ y = c_1 e^{3t} - c_2 t e^{3t} \end{cases}$

c) $\begin{cases} x= 3c_1+c_2\, e^{-2t} \\ y=4c_1+2c_2\, e^{-2t} \end{cases}$

d) $\begin{cases} x= c_1\, e^{2t} \\ y=c_2\, e^{3t} \end{cases}$

e) $\begin{cases} x= \lfloor c_1\, e^{-3t}+c_2\,(1-t)e^{-3t} \\ y= \lfloor -c_1\, e^{-3t}+c_2\, te^{-3t} \end{cases}$

f) $\begin{cases} x=2c_1\, e^{10t}+3c_2\, e^{3t} \\ y=c_1\, e^{10t}-2c_2\, e^{3t} \end{cases}$

g) $x = e^{3t}\,(c_1 \cos 2t + c_2 \sin 2t)$

$y = e^{3t}\,(c_1\,(\sin 2t - \cos 2t) - c_2\,(\sin 2t + \cos 2t))$

CHAPTER IV

NON LINEAR EQUATIONS

IV.1. Autonomous systems. The phase plane and its phenomena.

If x is the angle of deviation of an undamped pendulum of length a whose bob has mass m, then its equation of motion is

$$\frac{d^2x}{dt^2}+\frac{g}{a}\sin x=0; \tag{1}$$

and if there is present a damping force proportional to the velocity of the bob, then the equation becomes.

$$\frac{d^2x}{dt^2}+\frac{c}{m}\frac{dx}{dt}+\frac{g}{a}\sin x=0. \tag{2}$$

In the usual linearlization we replace sin x by x, which is reasonable for small oscillations but amounts to a gross distortion when x is large. Equation (1) and (2) are examples for non linear differential equations. An example of a different type can be found in the theory of the vacuum tube, which leads to the important van der Pol equation.

$$\frac{d^2x}{dt^2}+\mu\left(x^2-1\right)\frac{dx}{dt}+x=0. \tag{3}$$

It will be seen later that each of these non linear equations has interesting properties not shared by the others.

Throughout this chapter we shall be concerned with second order nonlinear equations of the form

$$\frac{d^2x}{dt^2}=f\left(x,\frac{dx}{dt}\right). \tag{4}$$

which includes equations (1), (2), and (3) as special cases. If we imagine a simple dynamical system consisting of a particle of unit mass moving on the x-axis, and if f(x, dx/dt) is the force acting on it, then (4) is the equation of motion. The values of x (position) and dx/dt (velocity), which at each instant characterize the state of the system, are called its phases, and the plane of the variables x and dx/dt is called the phase plane. If we introduce the variable y = dx/dt, then (4) can be replaced by the equivalent system

$$\begin{cases} \dfrac{dx}{dt} = y \\ \dfrac{dy}{dt} = f(x, y). \end{cases} \tag{5}$$

We shall see that a good deal can be learned about the solutions of (4) by studying the solutions of (5). When t is regarded as a parameter, then in general a solution of (5) is a pair of functions x(t) and y(t) defining a curve in the xy-plane, which is simply the phase plane mentioned above. We shall be interested in the total picture formed by these curves in the phase plane.

More generally, we study systems of the form

$$\begin{cases} \dfrac{dx}{dt} = F(x, y) \\ \dfrac{dy}{dt} = G(x, y), \end{cases} \tag{6}$$

where F and G are continuous and have continuous first partial derivatives throughout the plane. A system of this kind, in which the independent variable t does not appear in the functions F and G on the right, is said to be autonomous. We now turn to a closer examination of the solution of such a system.

If t_0 is any number and (x_0, y_0) is any point in the phase plane, then there exists a unique solution.

$$\begin{cases} x = x(t) \\ y = y(t) \end{cases} \tag{7}$$

of (6) such that $x(t_0) = x_0$ and $y(t_0) = y_0$. If x(t) and y(t) are not both constant functions, then (7) defines a curve in the phase plane called a path of the system. It is clear that if (7) is a solution of (6), then

$$\begin{cases} x = x(t + c) \\ y = y(t + c) \end{cases} \tag{8}$$

is also a solution for any constant c. Thus each path is represented by many solutions, which differ from one another only by a translation of the parameter. Also, any path through the point (x_0, y_0) must correspond to a solution of the form (8). It follows from this that at most one path passes through each point of the phase plane. Furthermore, the direction of increasing t along a given path is the same for all solutions representing the path. A path is therefore a directed curve, and in our figures we shall use arrows to indicate the direction in which the path is traced out as t increases.

The above remarks show that in general the paths of (6) cover the entire phase plane and do not intersect one another. The only exceptions to this statement occur at points (x_0, y_0) where both F and G vanish:

$$F(x_0, y_0) = 0 \quad \text{and} \quad G(x_0, y_0) = 0.$$

These points are called critical points, and at such a point the unique solution is the constant solution $x = x_0$ and $y = y_0$. A constant solution does not define a path, and therefore no path goes through a critical point. In our work we will always assume that each critical point (x_0, y_0) is isolated, in the sense that there exists a circle centered on (x_0, y_0) that contains no other critical point.

In order to obtain a physical interpretation of critical points, let us consider the special autonomous system (5) arising from the dynamical equation (4). In this case a critical point is a point $(x_0, 0)$ at which $y=0$ and $f(x_0, 0) = 0$; that is, it corresponds to a state of the particle's motion in which both the velocity dx/dt and the acceleration $dy/dt = d^2x/dt^2$ vanish. This means that the particle is at rest with no force acting on it, and is therefore in a state of equilibrium. It is obvious that the states of equilibrium of a physical system are among its most important features, and this accounts in part for our interest in critical points.

The general autonomous system (6) does not necessarily arise from any dynamical equation of the form (4). What sort of physical meaning can be attached to the paths and critical points in this case? Here it is convenient to consider Fig. 1 and the two-dimensional vector field

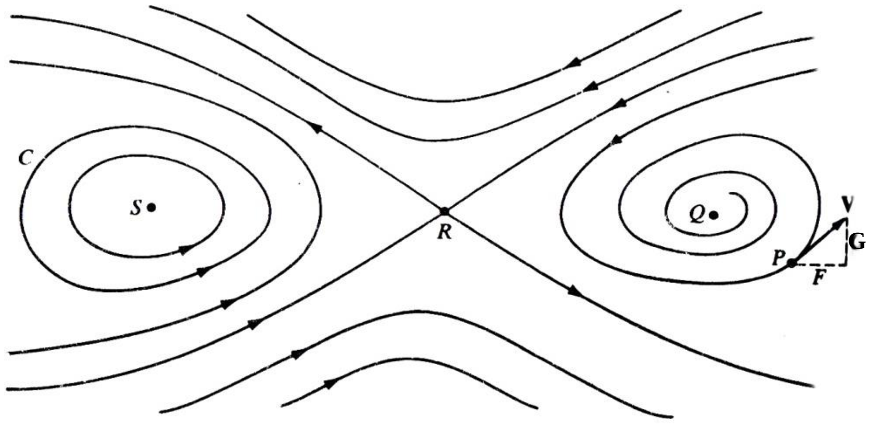


Fig. 1

defined by

$$V(x,y) = F(x,y)i + G(x,y)j,$$

which at a typical point P = (x,y) has horizontal component F(x,y) and vertical component G(x,y). Since dx/dt = F and dy/dt = G, this vector is tangent to the path at P and points in the direction of increasing t. If we think of t as time, then V can be interpreted as the velocity vector of a particle moving along the path. We can also imagine that the entire phase plane is filled with particles, and that each path is the trail of a moving particle preceded and followed by many others on the same path and accompanied by yet others on nearby paths. The situation can be described as a two-dimensional fluid motion; and since the system (6) is autonomous, which means that the vector V(x,y) at a fixed point (x,y) does not change with time, the fluid motion is stationary. The paths are the trajectories of the moving particles, and the critical points Q, R, and S are points of zero velocity where the particles are at rest (i.e., stagnation points of the fluid motion).

The most striking features of the fluid motion illustrated in fig. 1 are:

a. The critical points;
b. The arrangement of the paths near critical points;

c. The stability or instability of critical points, that is, whether a particle near such a point remains near or wanders off into another part of the plane;
d. Closed paths (like C in the figure), which correspond to periodic solutions.

These features constitute a major part of the phase portrait (or overall picture of the paths) of the system (6). Since in general nonlinear equations and systems cannot be solved explicitly, the purpose of the qualitative theory discussed in this chapter is to discover as much as possible about the phase portrait directly from the functions F and G. To gain some insight into the sort of information we might hope to obtain, observe that if x(t) is a periodic solution of the dynamical equation (4), then its derivative y(t) = dx/dt is also periodic and the corresponding path of the system (5) is therefore closed. Conversely, if any path of (5) is closed, then (4) has a periodic solution.

Problems

1. Describe the phase portrait of each of the following systems:

a. $\begin{cases} \dfrac{dx}{dt} = 0 \\ \dfrac{dy}{dt} = 0 \end{cases}$

c. $\begin{cases} \dfrac{dx}{dt} = 1 \\ \dfrac{dy}{dt} = 2; \end{cases}$

b. $\begin{cases} \dfrac{dx}{dt} = x \\ \dfrac{dy}{dt} = 0; \end{cases}$

d. $\begin{cases} \dfrac{dx}{dt} = -x \\ \dfrac{dy}{dt} = -y. \end{cases}$

2. The critical points and paths of equation (4) are by definition those of the equivalent system (5). Find the critical points of equations (1), (2), and (3).
3. Find the critical points of.

a. $\dfrac{d^2x}{dt^2} + \dfrac{dx}{dt} - (x^3 + x^2 - 2x) = 0;$

b. $\begin{cases} \dfrac{dx}{dt} = y^2 - 5x + 6 \\ \dfrac{dy}{dt} = x - y. \end{cases}$

Solutions:

1.

a. Every point is a critical point, and there are no paths

b. Every point on the y–axis is a critical point, and the paths are horizontal half – lines directed out to the left and right from the y–axis.

c. There are no critical points, and the paths are straight lines with slope 2 directed up to the right.

d. The point (0,0) is the only critical point and paths are half-lines of all possible slopes directed in toward the origin.

2. For equations (1) and (2) they are (0,0), ($\pm\pi$, 0), ($\pm 2\pi$, 0) ($\pm 3\pi$, 0)… and for equation (3), (0,0) is the only critical point.

3.

a. (-2, 0), (0,0), (1,0)

b. (2,2), (3,3)

IV. 2 TYPES OF CRITICAL POINTS. STABILITY

Consider an autonomous system

$$\begin{cases} \dfrac{dx}{dt} = F(x,y) \\ \dfrac{dy}{dt} = G(x,y). \end{cases} \qquad (1)$$

We assume, as usual, that the functions F and G are continuous and have continuous first partial derivatives throughout the xy-plane. The critical points of (1) can be found, at least in principle, by solving the simultaneous equations F(x,y) = 0 and G(x,y) = 0. There are four simple types of critical points that occur quite frequently, and our purpose in this section is to describe them in terms of the configurations of nearby paths.

Let (x_0, y_0) be an isolated critical point of (1). If C=[x(t), y(t)] is a path of (1), then we say that C approaches (x_0, y_0) as $t\to\infty$ if

$$\lim_{t \to \infty} x(t) = x_0 \text{ and } \lim_{t \to \infty} y(t) = y_0. \tag{2}$$

Geometrically, this means that if P = (x,y) is a point that traces out C in accordance with the equations x = x(t) and y = y(t), then $P \to (x_0, y_0)$ as $t \to \infty$. If it is also true that

$$\lim_{t \to \infty} \frac{y(t) - y_0}{x(t) - x_0} \tag{3}$$

exists, or if the quotient in (3) becomes either positively or negatively infinite as $t \to \infty$, then we say that C enters the critical point (x_0, y_0) as $t \to \infty$. The quotient in (3) is the slope of the line joining (x_0, y_0) and the point P with coordinates x(t) and y(t),so the additional requirement means that this line approaches a definite direction as $t \to \infty$. In the above definitions, we may also consider limits as $t \to -\infty$. It is clear that these properties are properties of the path C, and do not depend on which solution is used to represent this path.

It is sometimes possible to find explicit solutions of the system (1), and these solutions can then be used to determine the paths. In most cases, however, to find the paths it is necessary to eliminate t between the two equations of the system, which yields.

$$\frac{dy}{dx} = \frac{G(x,y)}{F(x,y)} \tag{4}$$

This first order equation gives the slope of the tangent to the path of (1) that passes through the point (x,y) provided that the functions F and G are not both zero at this point. In this case, of course, the point is a critical point and no path passes through it. The paths of (1) therefore coincide with the one-parameter family of integral curves of (4).

We now give geometric descriptions of the four main types of critical points. In each case we assume that the critical point under discussion is the origin O = (0,0).

Nodes. A critical point like that in Fig. 2 is called a node. Such a point is approached and also entered by each path as $t\to\infty$ (or as $t\to-\infty$). For the node shown in Fig. 2, there are four half-line paths, AO, BO, CO, and DO, which together with the origin make up the lines AB and CD. All other paths resemble parts of parabolas, and as each of these paths approaches O its slope approaches that of the line AB.

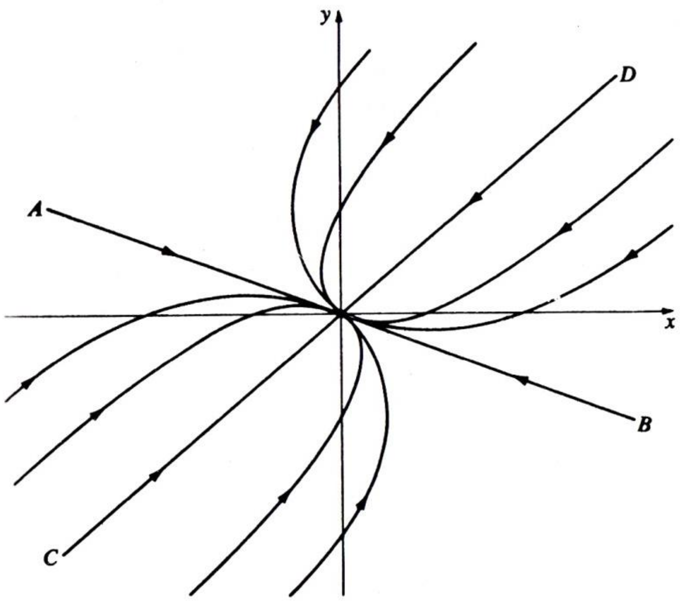


Fig. 2

Example 1. Consider the system

$$\begin{cases} \dfrac{dx}{dt} = x \\ \dfrac{dy}{dt} = -x + 2y. \end{cases} \qquad (5)$$

It is clear that the origin is the only critical point, and the general solution is

$$\begin{cases} x = c_1 e^t \\ y = c_1 e^t + c_2 e^{2t} \end{cases} \qquad (6)$$

When $c_1 = 0$, we have $x = 0$ and $y = c_2 e^{2t}$. In this case the path (Fig. 3) is the positive y-axis when $c_2>0$, and the negative y-axis when $c_2< 0$, and each path approaches and enters the origin as $t\to-\infty$. When $c_2 = 0$, we have $x = c_1 e^t$ and $y = c_1 e^t$. This path is the half-line $y = x$, $x>0$, when $c_1>0$, and the half line y = x, x<0, when $c_1 < 0$ and again both paths approach and enter the origin as $t \to -\infty$. When both c_1 and c_2 are $\neq 0$, the paths lie on the parabolas $y = x + (c_2 / c^2_1)\, x^2$, which go through the origin with slope 1. It should be understood that each of these paths consists of only part of a parabola, the part with $x > 0$ if $c_1 > 0$, and the part with $x < 0$ if $c_1 < 0$. Each of these paths also approaches and enters the origin as $t \to -\infty$; this can be seen at once from (6). If we proceed directly from (5) to the differential equation.

$$\frac{dy}{dx} = \frac{-x + 2y}{x}, \tag{7}$$

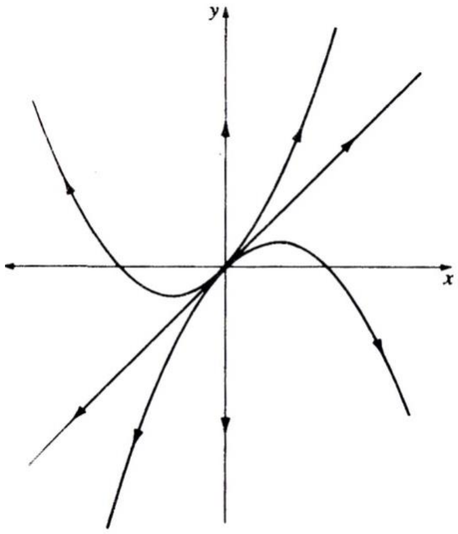


Fig. 3

giving the slope of the tangent to the path through (x,y) [provided $(x,y) \neq (0,0)$], then on solving (7) as a homogeneous equation, we find that $y = x + cx^2$. This procedure yields the curves on which the paths lie (except those on the y axis), but gives no information about the manner in which the paths are traced out. It is clear from this discussion that the critical point (0, 0) of the system (5) is a node.

Saddle points. A critical point like that in fig. 4 is called a saddle point. It is approached and entered by two half-line paths AO and BO as $t \to \infty$, and these two paths lie on a line AB. It is also approached and entered by two half – line paths CO and DO as $t\to-\infty$, and these two paths lie on another line CD. Between the four half-line paths there are four regions, and each contains a family of paths resembling hyperbolas. These paths do not approach O as $t\to\infty$ or as $t\to-\infty$, but instead are asymptotic to one or another of the half-line paths as $t\to\infty$ and as $t\to-\infty$.

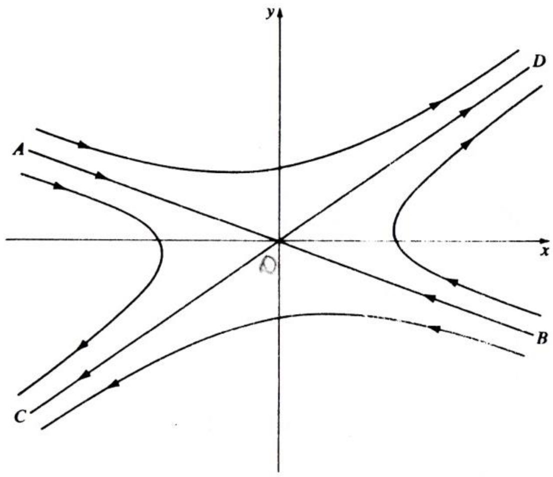


Fig. 4

Centers. A center (sometimes called a vortex) is a critical point that is surrounded by a family of closed paths. It is not approached by any path as $t\to\infty$ or as $t\to-\infty$.

Example 2. The system

$$\begin{cases} \dfrac{dx}{dt} = -y \\ \dfrac{dy}{dt} = x. \end{cases} \tag{8}$$

has the origin as its only critical point, and its general solution is

$$\begin{cases} x = -c_1 \sin t + c_2 \cos t \\ y = c_1 \cos t + c_2 \sin t. \end{cases} \tag{9}$$

The solution satisfying the conditions x(0) = 1 and y(0) = 0 is clearly

$$\begin{cases} x = \cos t \\ y = \sin t; \end{cases} \tag{10}$$

and the solution determined by x(0) = 0 y(0) = –1 is

$$\begin{cases} x = \sin t = \cos\left(t - \frac{\pi}{2}\right) \\ y = -\cos t = \sin\left(t - \frac{\pi}{2}\right) \end{cases} \tag{11}$$

These two different solutions define the same path C (Fig. 5), which is evidently the circle $x^2 + y^2 = 1$. Both (10) and (11) show that this path is traced out in the counterclockwise direction. If we eliminate t between the equations of the system, we get

$$\frac{dy}{dx} = -\frac{x}{y},$$

whose general solution $x^2 + y^2 = c^2$ yields all the paths (but without their directions). It is obvious that the critical point (0,0) of the system (8) is a center.

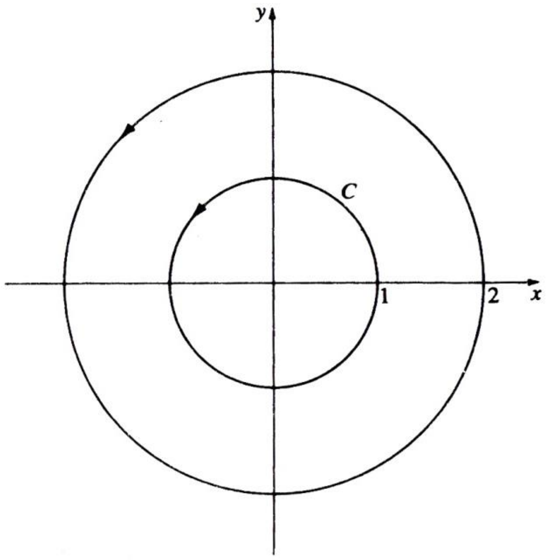


Fig. 5

Spirals. A critical point like that in Fig. 6 is called spiral (or sometimes a focus). Such a point is approached in a spiral-like manner by a family of paths that wind around it an infinite number of times as $t\to\infty$ (or as $t\to-\infty$). Note particularly that while the paths approach O, they do not enter it. That is, a point P moving along such a path approaches O as $t\to\infty$ (or as $t\to-\infty$), but the line OP does not approach any definite direction.

Example 3. If a is an arbitrary constant, then the system

$$\begin{cases} \dfrac{dx}{dt} = ax - y \\ \dfrac{dy}{dt} = x + ay. \end{cases} \tag{12}$$

has the origin as its only critical point (why?). The differential equation of the paths,

$$\frac{dy}{dx} = \frac{x + ay}{ax - y}, \tag{13}$$

is most easily solved by introducing polar coordinates r and θ defined by x = r cos θ and y = r sin θ. Since

$$r^2 = x^2 + y^2 \text{ and } \theta = \tan^{-1} \frac{y}{x},$$

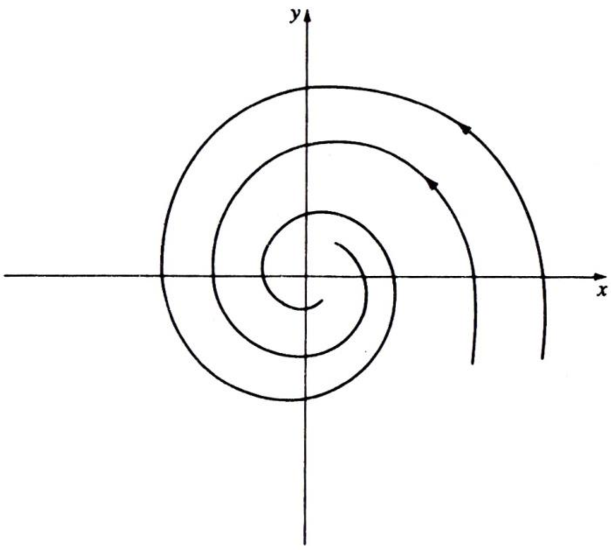


Fig. 6

we see that

$$r\frac{dr}{dx} = x + y\frac{dy}{dx} \quad \text{and} \quad r^2\frac{d\theta}{dx} = x\frac{dy}{dx} - y.$$

With the aid of these equations, (13) can easily be written in the very simple form.

$$\frac{dr}{d\theta} = ar,$$

so

$$r = ce^{a\theta} \tag{14}$$

is the polar equation of the paths. The two possible spiral configurations are shown in Fig. 7 and the direction in which these paths are traversed can be seen from the fact that $dx/dt = -y$ when $x = 0$. If $a = 0$, then (12) collapses to (8) and (14) becomes $r = c$, which is the polar equation of the family $x^2 + y^2 = c^2$ of all circles centered on the origin. This example therefore generalizes example 2; and since the center shown in Fig. 5 stands on the borderline between the spirals of Fig. 7, a critical point that is a center is often called a borderline case. We will encounter other borderline cases in the next section.

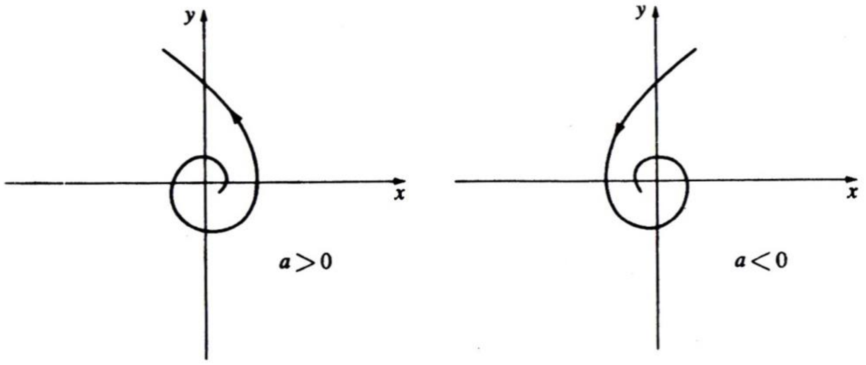


Fig. 7

Consider an isolated critical point of the system (1), and assume for the sake of convenience that this point is located at the origin $O = (0,0)$ of the phase plane. This critical point is said to be stable if for each positive number R there exists a positive number $r \leq R$ such that every path which is inside the circle $x^2 + y^2 = r^2$ for some $t = t_0$ remains inside the circle $x^2 + y^2 = R^2$ for all $t > t_0$ (Fig. 8). Loosely speaking, a critical point is stable if all paths that get sufficiently close to the point stay close to the point. Further, our critical point is said to be asymptotically stable if it is stable and there exists a circle $x^2 + y^2 = r_0^2$ such that every path which is inside this circle for some $t = t_0$ approaches the origin as $t \to \infty$. Finally, if our critical point is not stable, then it is called unstable.

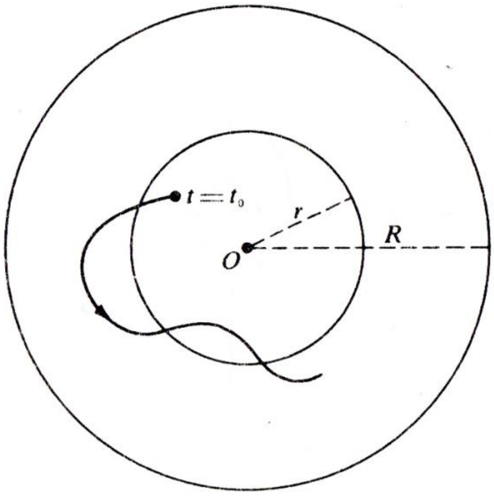


Fig. 8

As examples of these concepts, we point out that the node in Fig. 3 the saddle point in Fig. 4, and the spiral on the left in Fig. 7 are unstable, while the center in Fig. 5 is stable but not asymptotically stable. The node in Fig. 2, the spiral in Fig. 6, and the spiral on the right in Fig. 7 are asymptotically stable.

Problems

I. For each of the following non linear systems

i. Find the critical points
ii. Find the differential equation of the paths
iii. Solve this equation to find the paths

a. $\begin{cases} dx/dt = y(x^2+1) \\ dy/dt = 2xy^2 \end{cases}$

b. $\begin{cases} dx/dt = y(x^2+1) \\ dy/dt = -x(x^2+1) \end{cases}$

c. $\begin{cases} dx/dt = e^y \\ dy/dt = e^y \cos x \end{cases}$

d. $\begin{cases} dx/dt = -x \\ dy/dt = 2x^2 y^2 \end{cases}$

II. Each of the following linear systems has the origin as an isolated critical point.

i. Find the general solution

ii. Find the differential equation of the paths

iii. Solve the equation found in (ii)

iv. Discuss the stability of the critical point

a. $\begin{cases} dx/dt = x \\ dy/dt = -y \end{cases}$

b. $\begin{cases} dx/dt = -x \\ dy/dt = -2y \end{cases}$

c. $\begin{cases} dx/dt = 4y \\ dy/dt = -x \end{cases}$

Solution

I. a. (i) The critical points are the points on the x-axis; (ii) $\dfrac{dy}{dx} = \dfrac{2xy}{x^2+1}$

(iii) $y = c(x^2+1)$

b. (i) (0,0) (ii) $\dfrac{dy}{dx} = -x/y$ (iii) $x^2 + y^2 = c^2$

c. (i) There are no critical points (ii) $\dfrac{dy}{dx} = \cos x$ (iii) $y = \sin x + c$

d. (i) Critical points are the points on the y-axis (ii) $\dfrac{dy}{dx} = -2xy^2$

(iii) $y = \dfrac{1}{x^2+c}$ and $y=0$.

II. a. (i) $\begin{cases} x = c_1 e^{t} \\ y = c_2 e^{-t} \end{cases}$ (ii) $\dfrac{dy}{dx} = -\dfrac{y}{x}$ (iii) $xy=c$ (iv) unstable

b. (i) $\begin{cases} x = c_1 e^{-t} \\ y = c_2 e^{-2t} \end{cases}$ (ii) $\dfrac{dy}{dx} = \dfrac{2y}{x}$ (iii) $y=cx^2$ (iv) asymptotically stable.

c. (i) $\begin{cases} x = 2c_1 \cos 2t + 2c_2 \sin 2t \\ y = -c_1 \sin 2t + c_2 \cos 2t \end{cases}$ (ii) $\dfrac{dy}{dx} = \dfrac{-x}{4y}$ (iii) $\dfrac{x^2}{4c^2} + \dfrac{y^2}{c^2} = 1$ (iv) stable but not asymptotically stable.

IV.3 Critical Points And Stability For Linear Systems

Our goal in this chapter is to learn as much as we can about nonlinear differential equations by studying the phase portraits of nonlinear autonomous systems of the form.

$$\begin{cases} \dfrac{dx}{dt} = F(x, y) \\ \dfrac{dy}{dt} = G(x, y). \end{cases}$$

Under suitable conditions this problem can be solved for a given non linear system by studying a related linear system. We therefore devote this section to a complete analysis of the critical points of linear autonomous systems.

We consider the system

$$\begin{cases} \dfrac{dx}{dt} = a_1 x + b_1 y \\ \dfrac{dy}{dt} = a_2 x + b_2 y, \end{cases} \tag{1}$$

which has the origin (0,0) as an obvious critical point. We assume throughout this section that

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0, \tag{2}$$

so that (0,0) is the only critical point. (1) has a nontrivial solution of the form

$$\begin{cases} x = Ae^{mt} \\ y = Be^{mt} \end{cases}$$

whenever m is a root of the quadratic equation.

$$m^2 - (a_1 + b_2)m + (a_1 b_2 - a_2 b_1) = 0, \tag{3}$$

which is called the auxiliary equation of the system. Observe that condition (2) implies that zero cannot be a root of (3).

Let m_1 and m_2 be the roots of (3). We shall prove that the nature of the critical point (0,0) of the system (1) is determined by the nature of the numbers m_1 and m_2. It is reasonable to expect that three possibilities will occur, according as m_1 and m_2 are real and distinct real and equal, or conjugate complex. Unfortunately the situation is a little more complicated than this, and it is necessary to consider five cases, subdivided as follows.

Major cases:

Case A. The roots m_1 and m_2 are real, distinct, and of the same sign (node)

Case B. The roots m_1 and m_2 are real, distinct, and of opposite signs (saddle point).

Case C. The roots m_1 and m_2 are conjugate complex but not pure imaginary (spiral).

Borderline cases:

Case D. The roots m_1 and m_2 are real and equal (node)

Case E. The roots m_1 and m_2 are pure imaginary (center).

Case A. If the roots m_1 and m_2 are real, distinct, and of the same sign, then the critical point (0,0) is a node.

Proof. We begin by assuming that m_1 and m_2 are both negative, and we choose the notation so that $m_1 < m_2 < 0$. The general solution of (1) in this case is.

$$\begin{cases} x = c_1 A_1 e^{m_1 t} + c_2 A_2 e^{m_2 t} \\ x = c_1 B_1 e^{m_1 t} + c_2 B_2 e^{m_2 t}, \end{cases} \tag{4}$$

where the A's and B's are definite constants such that $B_1/A_1 \neq B_2/A_2$, and where the c's are arbitrary constants. When $c_2 = 0$, we obtain the solutions

$$\begin{cases} x = c_1 A_1 e^{m_1 t} \\ y = c_1 B_1 e^{m_1 t} \end{cases} \tag{5}$$

and when $c_1 = 0$, we obtain the solution

$$\begin{cases} x = c_2 A_2 e^{m_2 t} \\ y = c_2 B_2 e^{m_2 t} \end{cases} \tag{6}$$

For any $c_1 > 0$, the solution (5) represents a path consisting of half of the line $A_1 y = B_1 x$ with slope B_1/A_1; and for any $c_1 < 0$, it represents a path consisting of the other half of this line (the half on the other side of the origin). Since $m_1<0$, both of these half-line paths approach (0,0) as $t\to\infty$; and since $y/x = B_1/A_1$, both enter (0,0) with slop B_1/A_1 (fig. 9). In exactly the same way, the solutions (6) represent two half-line paths lying on the line $A_2y = B_2x$ with slope B_2/A_2. These two paths also approach (0,0) as $t\to\infty$, and enter it with slope B_2 / A_2.

If $c_1 \neq 0$ and $c_2 \neq 0$, the general solution (4) represents curved paths. Since $m_1 < 0$ and $m_2 < 0$, these paths also approach (0,0) as $t\to\infty$. Furthermore, since $m_1 - m_2 < 0$ and

$$\frac{y}{x}=\frac{c_1 B_1 e^{m_1 t}+c_2 B_2 e^{m_2 t}}{c_1 A_1 e^{m_1 t}+c_2 A_2 e^{m_2 t}} = \frac{(c_1 B_1 / c_2)e^{(m_1-m_2)t}+B_2}{(c_1 A_1 / c_2)e^{(m_1-m_2)t}+A_2},$$

it is clear that $y/x \to B_2/A_2$ as $t\to\infty$, so all of these paths enter (0,0) with slope B_2/A_2. Figure 9 presents a qualitative picture of the situation. It is evident that our critical point is a node, and that it is asymptotically stable.

If m_1 and m_2 are both positive, and if we choose the notation so that $m_1 > m_2 > 0$, then the situation is exactly the same except that all the paths now approach and enter (0,0) as $t\to-\infty$. The picture of the paths given in Fig. 9 is unchanged except that the arrows showing their direction are all reversed. We still have a node, but now it is unstable.

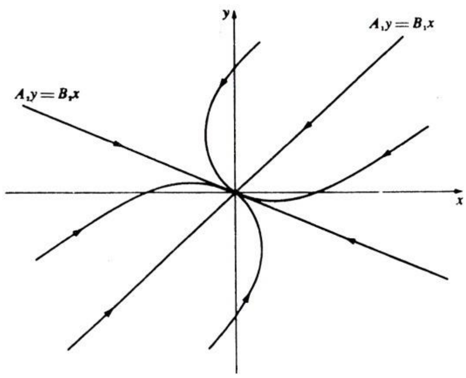


Fig. 9

Case B. If the roots m_1 and m_2 are real, distinct, and of opposite signs, then the critical point (0,0) is a saddle point.

Proof. We may choose the notation so that $m_1 < 0 < m_2$. The general solution of (1) can still be written in the form (4), and again we have particular solutions of the forms (5) and (6). The two half-line paths represented by (5) still approach and enter (0,0) as $t \to \infty$, but this time the two half-line paths represented by (6) approach and enter (0,0) as $t \to -\infty$. If $c_1 \neq 0$ and $c_2 \neq 0$, the general solution (4) still represents curved paths, but since $m_1 < 0 < m_2$, none of these paths approaches (0,0) as $t \to \infty$ or $t \to -\infty$. Instead, as $t \to \infty$, each of these paths is asymptotic to one of the half-line paths represented by (6); and as $t \to -\infty$, each is asymptotic to one of the half-line paths represented by (5). Figure 10 gives a qualitative picture of this behavior. In this case the critical point is a saddle point, and it is obviously unstable.

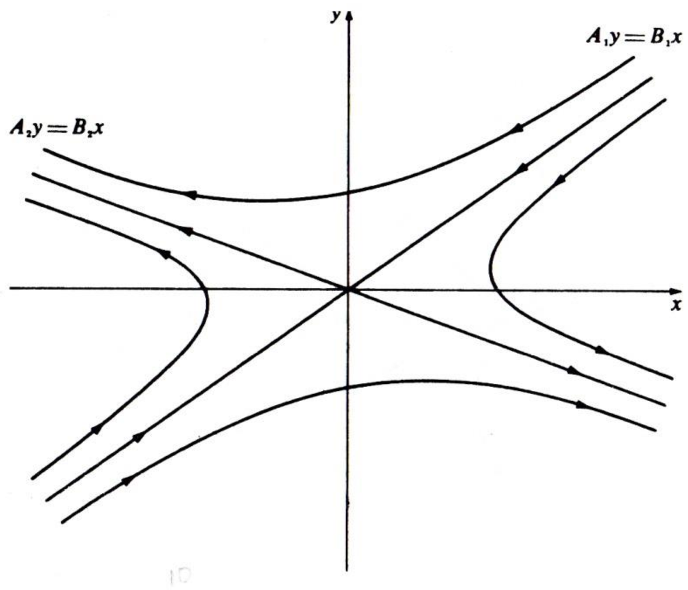


Fig. 10

Case C. If the roots m_1 and m_2 are conjugate complex but not pure imaginary, then the critical point (0,0) is a spiral.

Proof. In this case we can write m_1 and m_2 in the form $a \pm ib$ where a and b are nonzero real numbers. Also, for later use, we observe that the discriminant D of equation (3) is negative:

$$D = (a_1 + b_2)^2 - 4(a_1b_2 - a_2b_1)$$
$$= (a_1 - b_2)^2 + 4a_2b_1 < 0. \qquad (7)$$

The general solution of (1) in this case is

$$\begin{cases} x = e^{at}\left[c_1\left(A_1\cos bt - A_2\sin bt\right) + c_2\left(A_1\sin bt + A_2\cos bt\right)\right] \\ y = e^{at}\left[c_1\left(B_1\cos bt - B_2\sin bt\right) + c_2\left(B_1\sin bt + B_2\cos bt\right)\right] \end{cases} \qquad (8)$$

where the A's and B's are definite constant and the c's are arbitrary constants.

Let us first assume that $a < 0$. Then it is clear from formulas (8) that $x \to 0$ and $y \to 0$ as $t \to \infty$, so all the paths approach (0,0) as $t \to \infty$. We now prove that the paths do not enter the point (0,0) as $t \to \infty$, but instead wind around it in a spiral – like manner. To accomplish this we introduce the polar coordinate θ and show that, along any path, $d\theta/dt$ is either positive for all t or negative for all t. We begin with the fact that $\theta = \tan^{-1}(y/x)$, so

$$\frac{d\theta}{dt} = \frac{x\,dy/dt - y\,dx/dt}{x^2 + y^2}; \qquad (9)$$

and by using equations (1) we obtain $\dfrac{d\theta}{dt} = \dfrac{a_2x^2 + (b_2 - a_1)xy - b_1y^2}{x^2 + y^2}$

Since we are interested only in solutions that represent paths, we assume that $x^2 + y^2 \neq 0$. Now (7) implies that a_2 and b_1 have opposite signs. We consider the case in which $a_2 > 0$ and $b_1 < 0$. When $y = 0$, (9) yields $d\theta / dt = a_2 > 0$. If $y \neq 0$, $\dfrac{d\theta}{dt}$ cannot be 0; for if it were, then (9) would imply that

$$a_2x^2 + (b_2 - a_1)\,xy - b_1y^2 = 0$$

or

$$a_2 \left(\frac{x}{y}\right)^2 + (b_2 - a_1)\frac{x}{y} - b_1 = 0 \tag{10}$$

for some real number x/y–and this cannot be true because the discriminant of the quadratic equation (10) is D, which is negative by (7). This shows that dθ/dt is always positive when $a_2 > 0$, and in the same way we see that it is always negative when $a_2<0$. Since by (8), x and y change sign infinitely often as $t \to \infty$, all paths must spiral in to the origin (counterclockwise or clockwise according as $a_2>0$ or $a_2 < 0$). The critical point in this case is therefore a spiral, and it is asymptotically stable.

If a>0, the situation is the same except that the paths approach (0,0) as $t \to -\infty$ and the critical point is unstable. Figure 7 illustrates the arrangement of the paths when $a_2>0$.

Case D. If the roots m_1 and m_2 are real and equal, then the critical point (0,0) is a node.

Proof. We begin by assuming that $m_1 = m_2 = m < 0$. There are two subcases that require separate discussion: (i) $a_1 = b_2 \neq 0$ and $a_2 = b_1 = 0$; (ii) all other possibilities leading to a double root of equation (3).

We first consider the sub case (i). If a denotes the common value of a_1 and b_2, then equation (3) becomes $m^2 - 2am + a^2 = 0$ and m = a. The system (1)

is thus

$$\begin{cases} \dfrac{dx}{dt} = ax \\ \dfrac{dy}{dt} = ay, \end{cases}$$

and its general solution is

$$\begin{cases} x = c_1 e^{mt} \\ y = c_2 e^{mt} \end{cases} \tag{11}$$

where c_1 and c_2 are arbitrary constants. The paths defined by (11) are half-lines of all possible slopes (Fig. 11), and since m<0 we see that each path approaches and enters (0,0) as $t\to\infty$. The critical point is therefore a node, and it is asymptotically stable. If m>0, we have the same situation except that the paths enter (0,0) as $t\to-\infty$, the arrows in Fig. 11 are reversed, and (0,0) is unstable.

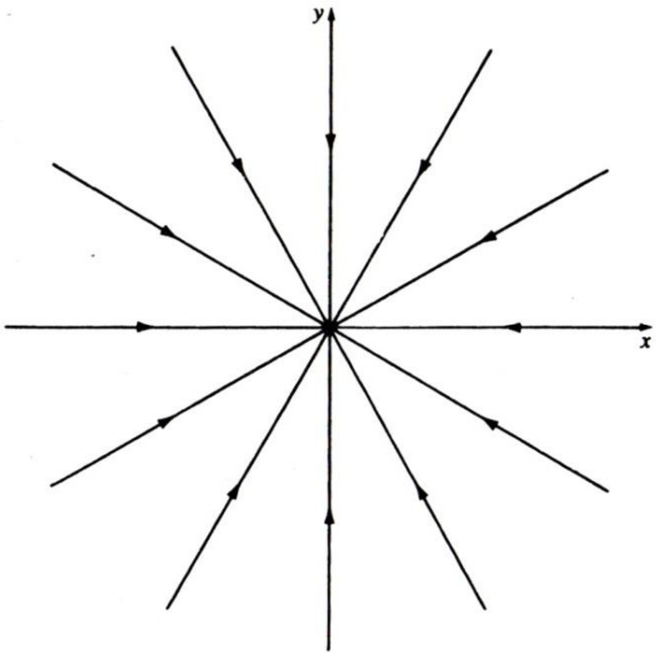


Fig. 11

We now discuss subcase (ii). The general solution of (1) can be written in the form.

$$\begin{cases} x = c_1 Ae^{mt} + c_2 (A_1 + At)e^{mt} \\ y = c_1 Be^{mt} + c_2 (B_1 + Bt)e^{mt}, \end{cases} \tag{12}$$

where the A's and B's are definite constants and the c's are arbitrary constants. When $c_2 = 0$, we obtain the solutions.

$$\begin{cases} x = c_1 Ae^{mt} \\ y = c_1 Be^{mt} \end{cases} \tag{13}$$

We known that these solution represent two half line paths lying on the line Ay = Bx with slope B/A, and since m<0 both paths approach (0, 0) as $t\to\infty$ (Fig. 12). Also, since y/x = B/A, both paths enter (0, 0) with slope B/A. If $c_2 \neq 0$, the solutions (12) represent curved paths, and since m<0 it is clear from (12) that these paths approach (0,0) as $t\to\infty$. Furthermore, it follows from

$$\frac{y}{x}=\frac{c_1 Be^{mt}+c_2(B_1+Bt)e^{mt}}{c_1 Ae^{mt}+c_2(A_1+At)e^{mt}}=\frac{c_1 B/c_2+B_1+Bt}{c_1 A/c_2+A_1+At}$$

that $y/x\to B/A$ as $t\to\infty$, so these curved paths all enter (0,0) with slope B/A. We also observe that $y/x\to B/A$ as $t\to-\infty$. Figure 12 gives a qualitative picture of the arrangement of these paths. It is clear that (0,0) is a node that is asymptotically stable. If m>0, the situation is unchanged except that the directions of the paths are reversed and the critical point is unstable.

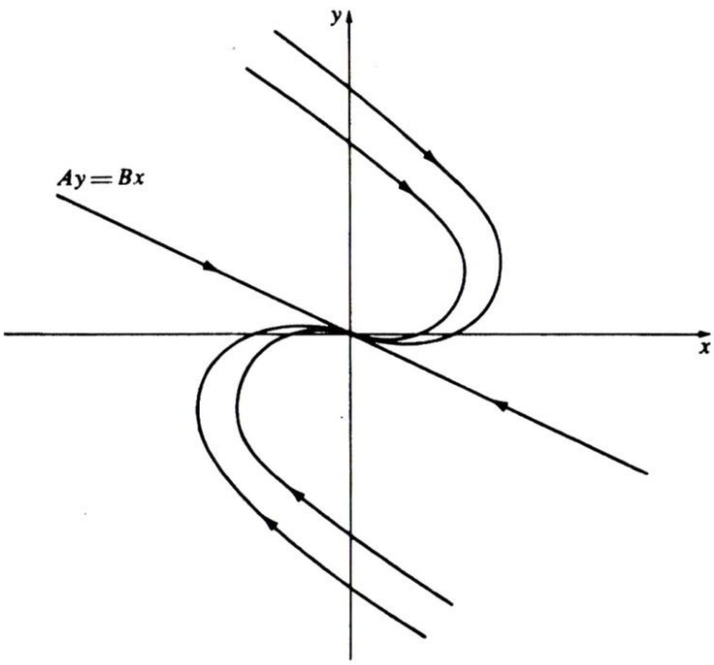


Fig. 12

Case E. If the roots m_1 and m_2 are pure imaginary, then the critical point (0,0) is a center.

Proof. It suffices here to refer back to the discussion of Case C, for now m_1 and m_2 are of the form $a \pm ib$ with $a = 0$ and $b\neq 0$. The general solution of (1) is therefore given by (8) with the exponential factor missing, so x(t) and y(t) are periodic and each path is a closed curve surrounding the origin. As Fig. 13 suggests, these curves are actually ellipses; this can be proved by solving the differential equation of the paths,

$$\frac{dy}{dx} = \frac{a_2 x + b_2 y}{a_1 x + b_1 y}. \tag{14}$$

Our critical point (0,0) is evidently a center that is stable but not asymptotically stable.

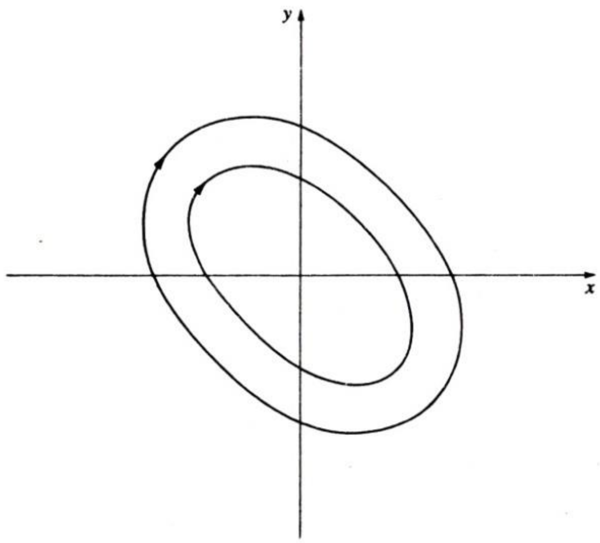


Fig. 13

In the above discussions we have made a number of statements about stability. It will be convenient to summarize this information as follows.

Theorem A. The critical point (0,0) of the linear system (1) is stable if and only if both roots of the auxiliary equation (3) have non-positive real parts, and it is asymptotically stable if and only if both roots have negative real parts.

If we now write equation (3) in the form

$$(m-m_1)\,(m-m_2) = m^2 + pm + q = 0, \tag{15}$$

so that $p = -(m_1 + m_2)$ and $q = m_1 m_2$, then our five cases can be described just as readily in terms of the coefficients p and q as in terms of the roots m_1 and m_2. In fact, if we interpret these cases in the pq plane, then we arrive at a striking diagram (Fig. 14) that displays at a glance the nature and stability properties of the critical point (0,0). The first thing to notice is that the p-axis q=0 is excluded, since by condition (2) we know that $m_1 m_2 \neq 0$. In the light of what we have learned about our five cases, all of the information contained in the diagram follows directly from the fact that

$$m_1, m_2 = \frac{-p \pm \sqrt{p^2 - 4q}}{2}.$$

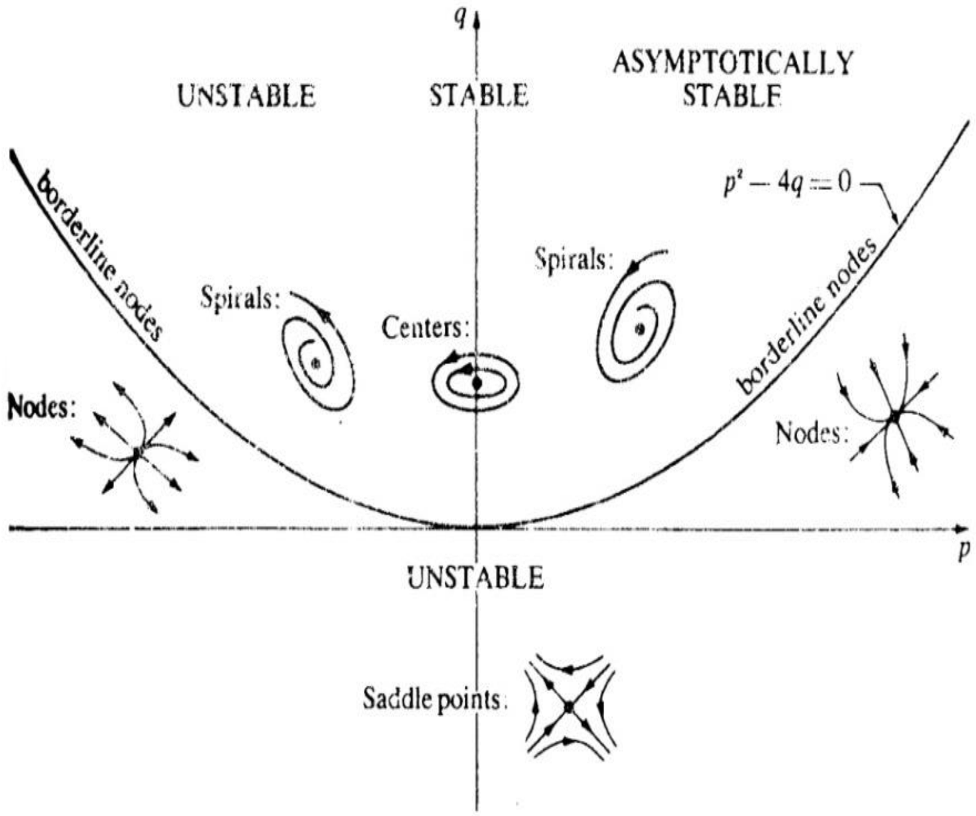


Fig. 14

Thus, above the parabola $p^2 - 4q = 0$, we have $p^2 - 4q < 0$, so m_1 and m_2 are conjugate complex numbers that are pure imaginary if and only if p=0; these are Cases C and E comprising the spirals and centers. Below the p-axis we have q<0, which means that m_1 and m_2 are real, distinct, and have opposite signs; this yields the saddle points of Case B. And finally, the zone between these two regions (including the parabola but excluding the p-axis) is characterized by the relations $p^2 - 4q \geq 0$ and $q > 0$, so m_1 and m_2 are real and of the same sign; here we have the nodes of Cases A and D. Furthermore, it is clear that there is precisely one region of asymptotic stability: the first quadrant. We state this formally as follows.

Theorem B. The critical point (0, 0) of the linear system (1) is asymptotically stable if and only if the coefficients $p = -(a_1 + b_2)$ and $q = a_1b_2 - a_2b_1$ of the auxiliary equation (3) are both positive.

Finally, it should be emphasized that we have studied the paths of our linear system near a critical point by analyzing explicit solutions of the system. In the next two sections we enter more fully into the spirit of the subject by investigating similar problems for nonlinear systems, which in general cannot be solved explicitly.

Problems

I. Determine the nature and stability properties of the critical point (0, 0) for each of the linear autonomous systems:

a. $\begin{cases} \dfrac{dx}{dt} = 2x \\ \dfrac{dy}{dt} = 3y; \end{cases}$

b. $\begin{cases} \dfrac{dx}{dt} = -x - 2y \\ \dfrac{dy}{dt} = 4x - 5y; \end{cases}$

c. $\begin{cases} \dfrac{dx}{dt} = -3x + 4y \\ \dfrac{dy}{dt} = -2x + 3y; \end{cases}$

d. $\begin{cases} \dfrac{dx}{dt} = 5x + 2y \\ \dfrac{dy}{dt} = -17x - 5y; \end{cases}$

e. $\begin{cases} \dfrac{dx}{dt} = -4x - y \\ \dfrac{dy}{dt} = x - 2y; \end{cases}$

f. $\begin{cases} \dfrac{dx}{dt} = 4x - 3y \\ \dfrac{dy}{dt} = 8x - 6y; \end{cases}$

g. $\begin{cases} \dfrac{dx}{dt} = 4x - 2y \\ \dfrac{dy}{dt} = 5x + 2y; \end{cases}$

Solutions

I.

a. Unstable node
b. Asymptotically stable spiral
c. Unstable saddle point
d. Stable but not asymptotically stable center
e. Asymptotically stable node
f. The critical point is not isolated
g. Unstable spiral

IV.4 Stability By Liapunov's Direct Method

If the total energy of a physical system has a local minimum at a certain equilibrium point, then that point is stable. This idea was generalized by Liapunov into a simple but powerful method for studying stability problems in a broader context. We shall discuss Liapunov's method and some of its application in this and the next section.

Consider an autonomous system

$$\begin{cases} \dfrac{dx}{dt} = F(x, y) \\ \dfrac{dy}{dt} = G(x, y), \end{cases} \qquad (1)$$

and assume that this system has an isolated critical point, which as usual we take to be the origin (0,0). Let C=[x(t), y(t)] be a path of (1), and consider a function E(x,y) that is continuous and has continuous first partial derivatives in a region containing this path. If a point (x,y) moves along the path in accordance with the equations x=x(t) and y=y(t), then E(x,y) can be regarded as a function of t along C[we denote this function by E(t)] and its rate of change is

$$\frac{dE}{dt} = \frac{\partial E}{\partial x}\frac{dx}{dt} + \frac{\partial E}{\partial y}\frac{dy}{dt}$$

$$= \frac{\partial E}{\partial x} F + \frac{\partial E}{\partial y} G. \qquad (2)$$

This formula is at the heart of Liapunov's ideas, and in order to exploit it we need several definition that specify the kinds of functions we shall be interested in.

Suppose that E(x,y) is continuous and has continuous first partial derivatives in some region containing the origin. If E vanishes at the origin, so that E(0,0) = 0, then it is said to be positive definite it E(x,y) > 0 for (x,y) ≠ (0,0), and negative definite it E(x,y)<0 for (x,y) ≠ (0,0). Similarly, E is called positive semidefinite if E(0,0) = 0 and E(x,y) ≥ 0 for (x,y) ≠ (0,0), and negative semidefinite if E(0,0) = 0 and E(x,y) ≤ 0 for (x,y) ≠ (0,0). It is clear that functions of the form

$ax^{2m} + by^{2n}$, where a and b are positive constants and m and n are positive integers, are positive definite. Since E(x,y) is negative definite if and only if $-$ E(x,y) is positive definite, fuctions of the form $ax^{2m} + by^{2n}$ with $a < 0$ and $b < 0$ are negative definite. The functions x^{2m}, y^{2m}, and $(x-y)^{2m}$ are not positive definite, but are nevertheless positive semidefinite.

A positive definite function E(x,y) with the property that

$$\frac{\partial E}{\partial x} F + \frac{\partial E}{\partial y} G. \tag{3}$$

is negative semidefinite is called a Liapunov function for the system (1). By formula (2), the requirement that (3) be negative semidefinite means that $dE/dt \leq 0$ – and therefore E is non increasing – along the paths of (1) near the origin. These functions generalize the concept of the total energy of a physical system. Their relevance for stability problems is made clear in the following theorem, which is Liapunov's basic discovery.

Theorem A. If there exists a Liapunov function E(x,y) for the system (1), then the critical point (0,0) is stable. Furthermore, if this function has the additional property that the function (3) is negative definite, then the critical point (0,0) is asymptotically stable.

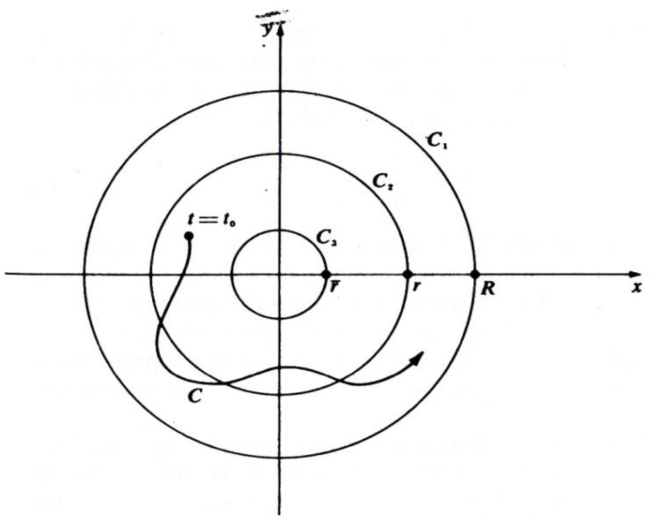


Fig. 15

Proof. Let C_1 be a circle of radius R>0 centered on the origin (Fig. 15), and assume also that C_1 is small enough to lie entirely in the domain of definition of the function E. Since E(x,y) is continuous and positive definite, it has a positive minimum m on C_1. Next, E(x,y) is continuous at the origin and vanishes there, so we can find a positive number r<R such that E(x,y) < m whenever (x,y) is inside the circle C_2 of radius r. Now let C=[x(t), y(t)] be any path which is inside C_2 for $t = t_0$. Then $E(t_0) < m$, and since (3) is negative semidefinite we have $dE/dt \leq 0$, which implies that $E(t) \leq E(t_0) < m$ for all $t > t_0$. It follows that the path C can never reach the circle C_1 for any $t > t_0$, so we have stability.

To prove the second part of the theorem, it suffices to show that under the additional hypothesis we also have $E(t) \to 0$, for since E(x,y) is positive definite this will imply that the path C approaches the critical point (0,0). We begin by observing that since dE/dt < 0, it follows that E(t) is a decreasing function; and since by hypothesis E(t) is bounded below by 0, we conclude that E(t) approaches some limit $L \geq 0$ as $t\to\infty$. To prove that $E(t) \to 0$ it suffices to show that L = 0, so we assume that L > 0 and deduce a contradiction.Choose a positive number $\overline{r} < r$ with the property that E(x,y) < L whenever (x,y) is inside the circle C_3 with radius $\overline{r}$. since the function (3) is continuous and negative definite, it has a negative maximum – k in the ring consisting of the circles C_1 and C_3 and the region between them. This ring contains the entire path C for $t \geq t_0$, so the equation

$$E(t) = E(t_0) + \int_{t_0}^{t} \frac{dE}{dt}\,dt$$

yields the inequality

$$E(t) \leq E(t_0) - k(t-t_0) \tag{4}$$

for all $t \geq t_0$. However, the right side of (4) becomes negatively infinite as $t\to\infty$, so $E(t) \to -\infty$ as $t\to\infty$. This contradicts the fact that $E(x,y) \geq 0$, so we conclude that L=0 and the proof is complete.

Example 1. Consider the equation of motion of a mass m attached to a spring:

$$m\frac{d^2x}{dt^2}+c\frac{dx}{dt}+kx=0. \tag{5}$$

Here c $\geq$ 0 is a constant representing the viscosity of the medium through which the mass moves, and k > 0 is the spring constant. The autonomous system equivalent to (5) is

$$\begin{cases} \dfrac{dx}{dt}=y \\ \dfrac{dy}{dt}=-\dfrac{k}{m}x-\dfrac{c}{m}y, \end{cases} \tag{6}$$

and its only critical point is (0,0). The kinetic energy of the mass is $my^2 / 2$, and the potential energy (or the energy stored in the spring) is

$$\int_0^x kx\,dx=\frac{1}{2}kx^2.$$

Thus the total energy of the system is

$$E(x,y) = \frac{1}{2}my^2+\frac{1}{2}kx^2. \tag{7}$$

It is easy to see that (7) is positive definite; and since

$$\frac{\partial E}{\partial x}F+\frac{\partial E}{\partial y}G=kxy+my\left(-\frac{k}{m}x-\frac{c}{m}y\right)$$

$$= -cy^2 \leq 0,$$

(7) is a Liapunov function for (6) and the critical point (0,0) is stable. When c>0 this critical point is asymptotically stable, but the particular Liapunov function discussed here is not capable of detecting this fact.

Example 2. The system

$$\begin{cases} \dfrac{dx}{dt}=-2xy \\ \dfrac{dy}{dt}=x^2-y^3, \end{cases} \tag{8}$$

has (0,0) as an isolated critical point. Let us try to prove stability by constructing a Liapunov function of the from $E(x,y) = ax^{2m} + by^{2n}$. It is clear that

$$\frac{\partial E}{\partial x}F + \frac{\partial E}{\partial y}G = 2\,\text{max}^{2m-1}(-2xy) + 2nby^{2n-1}(x^2 - y^3)$$

$$= (-4\text{max}^{2m}y + 2nbx^2y^{2n-1}) - 2nby^{2n+2}.$$

We wish to make the expression in parentheses vanish, and inspection shows that this can be done by choosing m=1, n=1, a=1, and b=2. With these choices we have $E(x,y)=x^2+2y^2$ (which is positive definite) and $(\partial E/\partial x)\,F+(\partial E\ \partial y)G= -4y^4$ (which is negative semidefinite). The critical point (0,0) of the system (8) is therefore stable. It is clear from this example that in complicated situations it may be very difficult indeed to construct suitable Liapunov functions. The following result is sometimes helpful in this connection.

Theorem B. The function $E(x,y) = ax^2 + bxy + cy^2$ is positive definite if and only if $a>0$ and $b^2 - 4ac < 0$, and is negative definite if and only if $a < 0$ and $b^2 - 4ac < 0$.

Proof. If $y = 0$, we have $E(x,0) = ax^2$, so $E(x,0) > 0$ for $x \neq 0$ if and only if $a > 0$. If $y \neq 0$, we have

$$E(x,y) = y^2\left[a\left(\frac{x}{y}\right)^2 + b\left(\frac{x}{y}\right) + c\right];$$

and when $a > 0$ the bracketed polynomial in x/y (which is positive for large x/y) is positive for all x/y if and only if $b^2 - 4ac < 0$. This proves the first part of the theorem, and the second part follows at once by considering the function $-E(x,y)$.

Problems

I. Determine whether each of the following functions is positive definite, negative definite or neither.

a. $x^2 - xy - y^2$;

b. $2x^2 - 3xy + 3y^2$;

c. $-2x^2 + 3xy - y^2$;

d. $-x^2 - 4xy - 5y^2$.

Solution

a. Neither

b. Positive definite

c. Neither

d. Negative definite

IV.5 Simple Critical Points of Non Linear Systems

Consider an autonomous system

$$\begin{cases} \dfrac{dx}{dt} = F(x,y) \\ \dfrac{dy}{dt} = G(x,y) \end{cases} \tag{1}$$

with an isolated critical point at (0,0). If F(x,y) and G(x,y) can be expanded in power series in x and y, then (1) takes the form

$$\begin{cases} \dfrac{dx}{dt} = a_1x + b_1y + c_1x^2 + d_1xy + e_1y^2 + \ldots \\ \dfrac{dy}{dt} = a_2x + b_2y + c_2x^2 + d_2xy + e_2y^2 + \ldots \end{cases} \tag{2}$$

When |x| and |y| are small – that is, when (x,y) is close to the origin – the terms of second degree and higher are very small. It is therefore natural to discard these non linear terms and conjecture that the qualitative behavior of the paths of (2) near the critical point (0,0) is similar to that of the paths of the related linear system

$$\begin{cases} \dfrac{dx}{dt} = a_1 x + b_1 y \\ \dfrac{dy}{dt} = a_2 x + b_2 y. \end{cases} \tag{3}$$

We shall see that in general this is actually the case. The process of replacing (2) by the linear system (3) is usually called linearization.

More generally, we shall consider systems of the form

$$\begin{cases} \dfrac{dx}{dt} = a_1 x + b_1 y + f(x,y) \\ \dfrac{dy}{dt} = a_2 x + b_2 y + g(x,y). \end{cases} \tag{4}$$

It will be assumed that

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0, \tag{5}$$

so that the related linear system (3) has (0,0) as an isolated critical point, that *f*(x,y) and g(x,y) are continuous and have continuous first partial derivatives for all (x,y); and that as (x,y) $\to$ (0,0) we have

$$\lim \frac{f(x,y)}{\sqrt{x^2+y^2}} = 0 \text{ and } \lim \frac{g(x,y)}{\sqrt{x^2+y^2}} = 0 \,. \tag{6}$$

Observe that conditions (6) imply that *f*(0,0) = 0 and g(0,0) = 0, so (0,0) is a critical point of (4); also, it is not difficult to prove that this critical point is isolated. With the restrictions listed above, (0,0) is said to be a simple critical point of the system (4).

Example 1. In the case of the system

$$\begin{cases} \dfrac{dx}{dt} = -2x + 3y + xy \\ \dfrac{dy}{dt} = -x + y - 2xy^2. \end{cases} \tag{7}$$

we have

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = \begin{vmatrix} -2 & 3 \\ -1 & 1 \end{vmatrix} = 1 \neq 0.$$

So (5) is satisfied. Furthermore, by using polar coordinates we see that

$$\frac{|f(x,y)|}{\sqrt{x^2+y^2}} = \frac{|r^2 \sin\theta\, \cos\theta|}{r} \le r \text{ and}$$

$$\frac{|g(x,y)|}{\sqrt{x^2+y^2}} = \frac{|2r^3 \sin^2\theta\, \cos\theta|}{r} \le 2r^2$$

So f(x,y)/r and g(x,y)/r$\to$0 as (x,y) $\to$ (0,0) (or as r$\to$0). This shows that conditions (6) are also satisfied, so (0,0) is a simple critical point of the system (7).

The main facts about the nature of simple critical points are given in the following theorem of Poincare, which we state without proof.

Theorem A. Let (0,0) be a simple critical point of the nonlinear system (4), and consider the related linear system (3). If the critical point (0,0) of (3) falls under any one of the three major cases described in Section IV.3, then the critical point (0,0) of (4) is of the same type.

As an illustration, we examine the nonlinear system (7) of Example 1, whose related linear system is

$$\begin{cases} \dfrac{dx}{dt} = -2x + 3y \\ \dfrac{dy}{dt} = -x + y. \end{cases} \tag{8}$$

The auxiliary equation of (8) is $m^2 + m + 1 = 0$, with roots

$$m_1, m_2 = \frac{-1 \pm \sqrt{3}i}{2}.$$

Since these roots are conjugate complex but not pure imaginary, we have Case C and the critical point (0,0) of the linear system (8) is a spiral. By Theorem A, the critical point (0,0) of the nonlinear system (7) is also a spiral.

It should be understood that while the type of the critical point (0,0) is the same for (4) as it is for (3) in the cases covered by the theorem, the actual appearance of the paths may be somewhat different. For example, Fig. 10 shows a typical saddle point for a linear system, whereas Fig. 16 suggests how a non linear saddle point might look. A certain amount of distortion is clearly present in the latter, but nevertheless the qualitative features of the two configuration are the same.

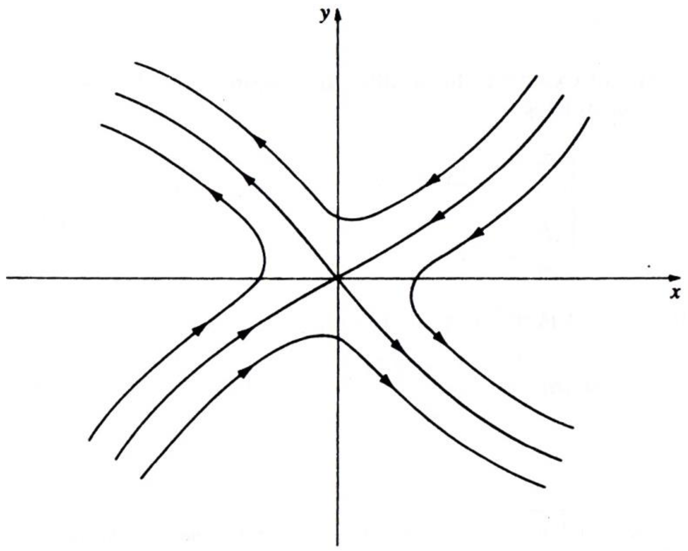


Fig. 16

It is natural to wonder about the two borderline cases, which are not mentioned in Theorem A. The facts are these: if the related linear system (3) has a borderline node at the origin (Case D), then the nonlinear system (4) can have either a node or a spiral; and if (3) has a center at the origin (Case E), then (4) can have either a center or a spiral. For example, (0,0) is a critical point for each of the nonlinear systems.

$$\begin{cases} \dfrac{dx}{dt} = -y - x^2 \\ \dfrac{dy}{dt} = x. \end{cases} \quad \text{and} \quad \begin{cases} \dfrac{dx}{dt} = -y - x^3 \\ \dfrac{dy}{dt} = x. \end{cases} \tag{9}$$

In each case the related linear system is

$$\begin{cases} \dfrac{dx}{dt} = -y \\ \dfrac{dy}{dt} = x. \end{cases} \tag{10}$$

It is easy to see that (0,0) is a center for (10). However, it can be shown that while (0,0) is a center for the first system of (9), it is a spiral for the second.

Theorem B. Let (0,0) be a simple critical point of the nonlinear system (4), and consider the related linear system (3). If the critical point (0,0) of (3) is asymptotically stable, then the critical point (0,0) of (4) is also asymptotically stable.

Proof. By Theorem A of section IV.3 it suffices to construct a suitable Liapunov function for the system (4), and this is what we do.

Theorem B of section IV.4 tells us that the coefficients of the linear system (3) satisfy the conditions.

$$p = -(a_1 + b_2) > 0 \text{ and } q = a_1b_2 - a_2b_1 > 0. \tag{11}$$

Now define

$$E(x,y) = \frac{1}{2}(ax^2 + 2bxy + cy^2)$$

by putting

$$a = \frac{a_2^2 + b_2^2 + (a_1 b_2 - a_2 b_1)}{D},$$

$$b = -\frac{(a_1 a_2 + b_1 b_2)}{D}$$

and

$$c = \frac{a_1^2 + b_1^2 + (a_1 b_2 - a_2 b_1)}{D},$$

where

$$D = pq = -(a_1 + b_2)(a_1b_2 - a_2b_1).$$

By (11), we see that D>0 and a>0. Also, an easy calculations shows that

$$\begin{aligned} D^2(ac-b^2) &= (a_2^2+b_2^2)(a_1^2+b_1^2) + (a_2^2+b_2^2+a_1^2+b_1^2)(a_1b_2-a_2b_1) \\ &\quad + (a_1b_2-a_2b_1)^2 - (a_1a_2+b_1b_2)^2 \\ &= (a_2^2+b_2^2+a_1^2+b_1^2)(a_1b_2-a_2b_1) + 2(a_1 b_2 - a_2 b_1)^2 \\ &> 0 \end{aligned}$$

so $b^2 - ac < 0$. Thus, by Theorem B of IV.4 we know that the function E(x,y) is positive definite. Furthermore, another calculation (whose details we leave to the reader) yields.

$$\frac{\partial E}{\partial x}(a_1 x + b_1 y) + \frac{\partial E}{\partial y}(a_2 x + b_2 y) = -(x^2 + y^2). \tag{12}$$

This function is clearly negative definite, so E(x,y) is a Liapunov function for the linear system (3).

We next prove that E(x,y) is also a Liapunov function for the nonlinear system (4). If F and G are defined by

$$F(x,y) = a_1 x + b_1 y + f(x,y)$$

and

$$G(x,y) = a_2 x + b_2 y + g(x,y),$$

then since E is known to be positive definite, it suffices to show that

$$\frac{\partial E}{\partial x} F + \frac{\partial E}{\partial y} G \tag{13}$$

is negative definite. If we use (12), then (13) becomes

$$-(x^2 + y^2) + (ax + by)\, f(x,y) + (bx + cy)\, g(x,y);$$

and by introducing polar coordinates we can write this as

$$-r^2 + r[(a \cos \theta + b \sin \theta)\, f(x,y) + (b \cos \theta + c \sin \theta)\, g(x,y)].$$

Denote the largest of the numbers |a|, |b|, |c| by K. Our assumption (6) now implies that

$$|f(x,y)| < \frac{r}{6K} \text{ and } |g(x,y)| < \frac{r}{6K}$$

for all sufficiently small r > 0, so

$$\frac{\partial E}{\partial x} F + \frac{\partial E}{\partial y} G < -r^2 + \frac{4Kr^2}{6K} = -\frac{r^2}{3} < 0$$

for these r's. Thus E(x,y) is a positive definite function with the property that (13) is negative definite. Theorem A of IV.4 now implies that (0,0) is an asymptotically stable critical point of (4), and the proof is complete.

To illustrate this theorem, we again consider the nonlinear system (7) of Example 1, whose related linear system is (8). For (8) we have $p=1>0$ and $q = 1 > 0$, so the critical point (0,0) is asymptotically stable, both for the linear system (8) and for the non linear system (7).

Example 2. We know from Section IV.1 that the equation of motion for the damped vibrations of a pendulum is

$$\frac{d^2x}{dt^2}+\frac{c}{m}\frac{dx}{dt}+\frac{g}{a}\sin x = 0.$$

where c is a positive constant. The equivalent nonlinear system is

$$\begin{cases} \dfrac{dx}{dt}=y \\ \dfrac{dy}{dt}=-\dfrac{g}{a}\sin x-\dfrac{c}{m}y. \end{cases} \tag{14}$$

Let us now write (14) in the form

$$\begin{cases} \dfrac{dx}{dt}=y \\ \dfrac{dy}{dt}=-\dfrac{g}{a}x-\dfrac{c}{m}y+\dfrac{g}{a}(x-\sin x). \end{cases} \tag{15}$$

It is easy to see that

$$\frac{x-\sin x}{\sqrt{x^2+y^2}}\to 0$$

as $(x,y) \to (0,0)$ for if $x \neq 0$, we have

$$\frac{|x-\sin x|}{\sqrt{x^2+y^2}} \leq \frac{|x-\sin x|}{|x|}=\left|1-\frac{\sin x}{x}\right|\to 0;$$

and since (0,0) is evidently an isolated critical point of the related linear system

$$\begin{cases} \dfrac{dx}{dt}=y \\ \dfrac{dy}{dt}=-\dfrac{g}{a}x-\dfrac{c}{m}y, \end{cases} \tag{16}$$

it follows that (0,0) is a simple critical point of (15). Inspection shows ($p = c/m > 0$ and $q = g/a > 0$) that (0,0) is an asymptotically stable critical point of (16), so by Theorem B it is also an asymptotically stable critical point of (15). This reflects the obvious physical fact that if the pendulum is slightly disturbed, then the resulting motion will die out with the passage of time.

MODULE – III

CHPATER – V

OSCILLATION THEORY OF BOUNDARY VALUE PROBLEMS

V.1 OSCILLATIONS AND THE STURM SEPARATION THEOREM

In this section we turn our attention to the problem of learning what we can do about the characteristics of the solution of $y^{11} + P(x)y^{1} + Q(x)y = 0$ (1)

by direct analysis of the equation itself, in the absence of formal expressions for these solutions.

Theorem A: Sturm separation Theorem

If $y_1(x)$ and $y_2(x)$ are two linearly independent solutions of $y^{11} + P(x)y^{1} + Q(x)y=0$,

then the zeros of these functions are distinct and occur alternately – in the sense that $y_1(x)$ vanishes exactly once between any two successive zeros of $y_2(x)$, and conversely.

Proof. The argument rests primarily on the fact that since y_1 and y_2 are linearly independent, their Wronskian

$$W(y_1, y_2) = y_1(x)\, y_2^{I}(x) - y_2(x)\, y_1^{I}(x)$$

does not vanish, and therefore – since it is continuous – must have constant sign. First, it is easy to see that y_1 and y_2 cannot have a common zero; for it they do, then the Wronskian will vanish at that point, which is impossible. We now assume that x_1 and x_2 are successive zeros of y_2 and show that y_1 vanishes between these points. The Wronskian clearly reduces to $y_1(x)\, y_2^{I}(x)$ at x_1 and x_2, so both factors $y_1(x)$ and $y_2^{I}(x)$ are $\neq 0$ at each of these points. Furthermore, $y_2^{I}(x_1)$ and $y_2^{I}(x_2)$ must have opposite signs, because if y_2 is increasing at x_1 it must be decreasing at x_2, and vice versa. Since the Wronskian has constant sign,

$y_1(x_1)$ and $y_1(x_2)$ must also have opposite signs, and therefore, by continuity, $y_1(x)$ must vanish at some point between x_1 and x_2. Note that y_1 cannot vanish more than once between x_1 and x_2; for if it does, then the same argument shows that y_2 must vanish between these zeros of y_1, which contradicts the original assumption that x_1 and x_2 are successive zeros of y_2.

Theorem B. If $q(x) < 0$, and if $u(x)$ is a nontrivial solution of $u^{11} + q(x)\, u = 0$, then $u(x)$ has at most one zero.

Proof. Let x_0 be a zero of $u(x)$, so that $u(x_0) = 0$. Since $u(x)$ is nontrivial (i.e., is not identically zero), $u^1(x_0) \neq 0$. For the sake of concreteness, we now assume that $u^1(x_0) > 0$, so that $u(x)$ is positive over some interval to the right of x_0. Since $q(x) < 0$, $u^{11}(x) = -q(x)\, u(x)$ is a positive function on the same interval. This implies that the slope $u^1(x)$ is an increasing function, so $u(x)$ cannot have a zero to the right of x_0, and in the same way it has none to the left of x_0. A similar argument holds when $u^1(x_0) < 0$, so $u(x)$ has either no zeros at all or only one, and the proof is complete.

Theorem C

Let $u(x)$ be any nontrivial solution of $u^{11}+q(x)u = 0$ (2)

where $q(x)>0$ for all $x>0$. If $\int_1^{\infty} q(x)\,dx = \infty$, (3)

Then $u(x)$ has infinitely many zeros on the positive x – axis.

Proof.

Assume the contrary, namely, that $u(x)$ vanishes at most a finite number of times for $0 < x < \infty$, so that a point $x_0 > 1$ exists with the property that $u(x) \neq 0$ for all $x \geq x_0$. We may clearly suppose, without any loss of generality, that $u(x) > 0$ for all $x \geq x_0$, since $u(x)$ can be replaced by its negative if necessary. Our purpose is to contradict the assumption by showing that $u^1(x)$ is negative somewhere to the right of x_0– for by the remarks below, this will imply that u(x) has a zero to the right of x_0. If we put

$$\nu(x) = -\frac{u'(x)}{u(x)}$$

for $x \geq x_0$, then a simple calculation shows that

$$\nu^1(x) = q(x) + \nu(x)^2;$$

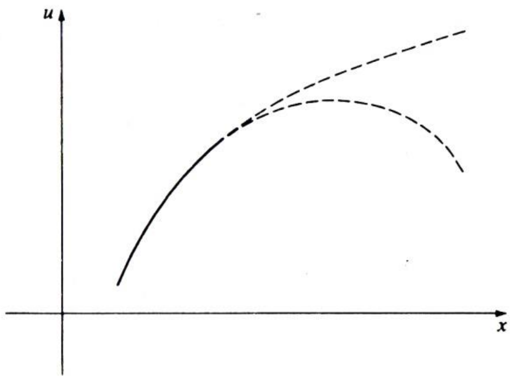
and on integrating this from x_0 to x, where $x > x_0$, we get

$$\nu(x) - \nu(x_0) = \int_{x_0}^{x} q(x)dx + \int_{x_0}^{x} \nu(x)^2\, dx.$$

We now use (3) to conclude that ν(x) is positive if x is taken large enough. This shows that u(x) and $u^1(x)$ have opposite signs if x is sufficiently large, so $u^1(x)$ is negative and the proof is complete.

Remarks

Let u(x) be a nontrivial solution of (2) with q(x)>0. If we consider a portion of the graph above the x-axis,



then, $u^{11}(x) = -q(x)\,u(x)$ is negative. So the graph is concave down and the slope $u^1(x)$ is decreasing. If this slope ever becomes negative, then the curve plainly crosses the x-axis somewhere to the right and we get a zero for u(x), when q(x) is constant. Alternative is that although $u^1(x)$ decreases, it never reaches zero and curve continues to rise. So u(x) will have zeros as x increases whenever q(x) does not decrease too rapidly.

V.2 The sturm Comparison Theorem

Theorem: A

Let y(x) be a nontrivial solution of the differential equation $y^{11}+q(x)\, y = 0$… (1) on a closed interval [a, b]. Then y(x) has at most a finite number of zeros in this interval.

Proof. We assume the contrary, namely, that y(x) has an infinite number of zeros in [a, b]. It follows from this that there exist in [a, b] a point x_0 and a sequence of zeros $x_n \neq x_0$ such that $x_n \to x_0$. Since y(x) is continuous and differentiable at x_0, we have

$$y(x_0) = \lim_{x_n \to x_0} y(x_n) = 0$$

and

$$y^1(x_0) = \lim_{x_n \to x_0} \frac{y(x_n) - y(x_0)}{x_n - x_0} = 0.$$

these statements imply that y(x) is the trivial solution of (1), and this contradiction completes the proof.

Theorem B:

Let y(x) and z(x) be nontrivial solutions of

$$y^{11} + q(x)\, y = 0$$

and

$$z^{11} + r(x)\, z = 0,$$

where q(x) and r(x) are positive functions such that q(x) > r(x). Then y(x) vanishes at least once between any two successive zeros of z(x).

Proof.

Let x_1 and x_2 be successive zeros of z(x), so that $z(x_1) = z(x_2) = 0$ and z(x) does not vanish on the open interval (x_1, x_2). We assume that y(x) does not vanish on (x_1, x_2), and prove the theorem by deducing a contradiction. It is clear that no loss of generality is involved in supposing that both y(x) and z(x) are positive on (x_1,

x_2), for either function can be replaced by its negative if necessary. If we emphasize that the Wronskian.

$$W(y,z) = y(x)\, z^1(x) - z(x)\, y^1(x)$$

is a function of x by writing it W(x), then

$$\frac{dW(x)}{dx} = yz^{11} - zy^{11}$$

$$= y(-rz) - z(-qy)$$

$$= (q-r)\, yz > 0$$

on (x_1, x_2). We now integrate both sides of this inequality from x_1 to x_2 and obtain.

$$W(x_2) - W(x_1) > 0 \text{ or } W(x_2) > W(x_1).$$

However, the Wronskian reduces to $y(x)z^1(x)$ at x_1 and x_2, so

$$W(x_1) \geq 0 \text{ and } W(x_2) \leq 0,$$

which is the desired contradiction.

Theorem: C.

Let $y_p(x)$ be a nontrivial solution of Bessel's equation on the positive x-axis. If $0 \leq p < \frac{1}{2}$, then every interval of length π contains at least one zero of $y_p(x)$; if $p = \frac{1}{2}$, then the distance between successive zeros of $y_p(x)$ is exactly π; and if $p > \frac{1}{2}$, then every interval of length π contains at most one zero of $y_p(x)$.

Proof

Suppose $q(x)>k^2>0$ in (1). Then from Theorem: B above, any solution must vanish between any two successive zeros of a solution $y(x) = \sin k(x-x_0)$ of the equation $y^{11}+ k^2y = 0$, and therefore must vanish in any interval of length π/k. If we Consider Bessel's equation $x^2y^{11}+xy^1+(x^2-p^2)y = 0$ in the normal form $u^{11}+ \left(1+\frac{1-4p^2}{4x^2}\right)u=0$, and comparing with $u^{11}+ u = 0$, we have the theorem.

CHAPTER –VI

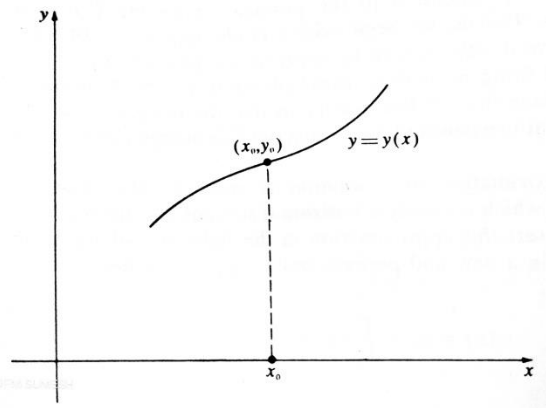
THE EXISTENCE AND UNIQUENESS OF SOLUTIONS

VI.1 The method of successive Approximations

Consider the initial value problem

$$y^1 = f(x,y), \quad y(x_0) = y_0, \tag{1}$$

where f(x,y) is an arbitrary function defined and continuous in some neighborhood of the point (x_0, y_0). In geometric language, our purpose is to devise a method for constructing a function y = y(x) whose graph passes through the point (x_0, y_0) and that satisfies the differential equation $y^1 = f(x,y)$ in some neighborhood of x_0 (Fig.). We are prepared for the idea that elementary procedures will not work and that in general some type of infinite process will be required.



The method we describe furnishes a line of attack for solving differential equations that is quite different from any the reader has encountered before. The key to this method lies in replacing the initial value problem (1) by the equivalent integral equation.

$$y(x) = y_0 + \int_{x_0}^{x} f[t, y(t)]dt. \tag{2}$$

This is called an integral equation because the unknown function occurs under the integral sign. To see that (1) and (2) are indeed equivalent, suppose that y(x) is a solution of (1). Then y(x) is automatically continuous and the right side of

$y^1(x) = f[x,y(x)]$

is a continuous function of x; and when we integrate this from x_0 to x and use $y(x_0) = y_0$, the result is (2). As usual, the dummy variable *t* is used in (2) to avoid confusion with the variable upper limit x on the integral. Thus any solution of (1) is a continuous solution of (2). Conversely, if y(x) is a continuous solution of (2), then $y(x_0) = y_0$ because the integral vanishes when $x = x_0$, and by differentiation of (2) we recover the differential equation $y^1(x) = f[x,y(x)]$. These simple arguments show that (1) and (2) are equivalent in the sense that the solutions of (1) – if any exist – are precisely the continuous solutions of (2) .In particular we automatically obtain a solution for (1) if we can construct a continuous solution for (2).

We now turn our attention to the problem of solving (2) by a process of iteration. That is, we begin with a crude approximation to a solution and improve it step by step by applying a repeatable operation which we hope will bring us as close as we please to an exact solution. The primary advantage that (2) has over (1) is that the integral equation provides a convenient mechanism for carrying out this process, as we now see.

A rough approximation to a solution is given by the constant function $y_0(x) = y_0$, which is simply a horizontal straight line through the point (x_0, y_0). We insert this approximation in the right side of equation (2) in order to obtain a new and perhaps better approximation $y_1(x)$ as follows:

$$y_1(x) = y_0 + \int_{x_0}^{x} f(t, y_0)\,dt.$$

The next step is to use $y_1(x)$ to generate another and perhaps even better approximation $y_2(x)$ in the same way:

$$y_2(x) = y_0 + \int_{x_0}^{x} f[t, y_1(t)]\,dt.$$

At the nth stage of the process we have

$$y_n(x) = y_0 + \int_{x_0}^{x} f(t, y_{n-1}(t))\,dt. \qquad (3)$$

This procedure is called Picard's method of successive approximations.

We show how it works by means of a few examples.

The simple initial value problem

$$y^1 = y,\ y(0) = 1$$

has the obvious solution $y(x) = e^x$. The equivalent integral equation is

$$y(x) = 1 + \int_0^x y(t)\,dt,$$

and (3) becomes, $y_n(x) = 1 + \int_0^x y_{n-1}(t)\,dt$

With $y_0(x) = 1$, it is easy to see that

$$y_1(x) = 1 + \int_0^x dt = 1 + x,$$

$$y_2(x) = 1 + \int_0^x (1+t)\,dt = 1 + x + \frac{x^2}{2},$$

$$y_3(x) = 1 + \int_0^x \left(1 + t + \frac{t^2}{2}\right)dt = 1 + x + \frac{x^2}{2} + \frac{x^3}{2\cdot 3},$$

and in general

$$y_n(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!}.$$

In this case it is very clear that the successive approximations do in fact converge to the exact solution, for these approximations are the partial sums of the power series expansion of e^x.

Let us now consider the problem.

$$y^1 = x + y,\ \ y(0) = 1.$$

This is a first order linear equation, and the solution satisfying the given initial condition is easily found to be $y(x) = 2e^x - x - 1$. The equivalent integral equation is

$$y(x) = 1 + \int_0^x [t + y(t)]dt,$$

and (3) is

$$y_n(x) = 1 + \int_0^x [t + y_{n-1}(t)]dt.$$

with $y_0(x) = 1$, Picard's method yields

$$y_1(x) = 1 + \int_0^x (t+1)dt = 1 + x + \frac{x^2}{2!},$$

$$y_2(x) = 1 + \int_0^x \left(1 + 2t + \frac{t^2}{2!}\right)dt = 1 + x + x^2 + \frac{x^3}{3!},$$

$$y_3(x) = 1 + \int_0^x \left(1 + 2t + t^2 + \frac{t^3}{3!}\right)dt$$

$$= 1 + x + x^2 + \frac{x^3}{3} + \frac{x^4}{4!},$$

$$y_4(x) = 1 + \int_0^x \left(1 + 2t + t^2 + \frac{t^3}{3} + \frac{t^4}{4!}\right)dt$$

$$= 1 + x + x^2 + \frac{x^3}{3} + \frac{x^4}{3 \cdot 4} + \frac{x^5}{5!},$$

and in general

$$y_n(x) = 1 + x + 2\left(\frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!}\right) + \frac{x^{n+1}}{(n+1)!}.$$

This evidently converges to

$$1 + x + 2(e^x - x - 1) + 0 = 2e^x - x - 1,$$

So again we have the exact solution.

Problems

1. Find the exact solution of the initial value problem

   $$y^1 = y^2, \quad y(0) = 1.$$

   Starting with $y_0(x) = 1$, apply Picard's method to calculate $y_1(x)$, $y_2(x)$, $y_3(x)$ and compare these results with the exact solution.
2. Find the exact solution of the initial value problem

$y^1 = 2x(1 + y)$, $y(0) = 0$.

Starting with $y_0(x) = 0$, calculate $y_1(x)$, $y_2(x)$, $y_3(x)$, $y_4(x)$ and compare these results with the exact solution.

3. It is instructive to see how Picard's method works with a choice of the initial approximation other than the constant function $y_0(x) = y_0$. Apply the method to the initial value problem (4) with
   a. $y_0(x) = e^x$;
   b. $y_0(x) = 1 + x$;
   c. $y_0(x) = \cos x$.

Solutions

1. $y = 1+x+x^2+\ldots.. = \frac{1}{1-x}, |x| < |$

$y_1(x) = 1+x$, $y_2(x) = 1+x+x^2+\frac{1}{3}x^3$, $y_3(x) = 1+x+x^2+x^3+\frac{2}{3}x^4+\frac{1}{3}x^5+\frac{1}{9}x^6+\frac{1}{63}x^7$

2. $y = e^{x^2} - 1$

$y_1(x) = x^2$, $y_2(x) = x^2 + \frac{x^4}{2}$, $y_3(x) = x^2 + \frac{x^4}{2} + \frac{x^6}{2\cdot 3}$, $y_4(x) = x^2 + \frac{x^4}{2} + \frac{x^6}{2\cdot 3} + \frac{x^8}{2\cdot 3.4}$

VI. 2 PICARD'S THEOREM

Theorem A. (Picard's theorem) Let $f(x, y)$ and $\partial f / \partial y$ be continuous functions of x and y on a closed rectangle R with sides parallel to the axes (Fig.) If (x_0, y_0) is any interior point of R, then there exists a number $h>0$ with the property that the initial value problem.

$$y^1 = f(x,y), \quad y(x_0) = y_0 \tag{1}$$

has one and only one solution $y = y(x)$ on the interval $|x - x_0| \leq h$.

Proof. The argument is fairly long and intricate, and is best absorbed in easy stages.

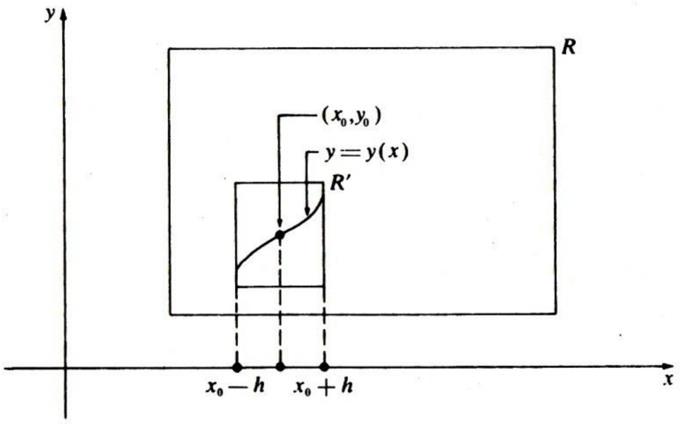


Fig.

First, we know that every solution of (1) is also a continuous solution of the integral equation.

$$y(x) = y_0 + \int_{x_0}^{x} f[t, y(t)]dt \qquad (2)$$

and conversely. This enables us to conclude that (1) has a unique solution on an interval $|x - x_0| \leq h$ if and only if (2) has a unique continuous solution on the same interval.

The sequence of functions $y_n(x)$ defined by

$$y_0(x) = y_0,$$

$$y_1(x) = y_0 + \int_{x_0}^{x} f[t, y_0(t)]dt,$$

$$y_2(x) = y_0 + \int_{x_0}^{x} f[t, y_1(t)]dt, \qquad (3)$$

$$\ldots$$

$$y_n(x) = y_0 + \int_{x_0}^{x} f[t, y_{n-1}(t)]dt,$$

$$\ldots$$

converges to a solution of (2). We next observe that $y_n(x)$ is the nth partial sum of the series of functions

$$y_0(x) + \sum_{n=1}^{\infty}[y_n(x) - y_{n-1}(x)] = y_0(x) + [y_1(x) - y_0(x)]$$

$$+ [y_2(x) - y_1(x)] + \ldots + [y_n(x) - y_{n-1}(x)] + \ldots, \qquad (4)$$

so the convergence of the sequence (3) is equivalent to the convergence of this series. In order to complete the proof, we produce a number $h>0$ that defines the interval $|x-x_0| \le h$, and then we show that on this interval the following statements are true: (i) the series (4) converges to a function $y(x)$; (ii) $y(x)$ is a continuous solution of (2); (iii) $y(x)$ is the only continuous solution of (2).

The hypotheses of the theorem are used to produce the positive number h, as follows. We have assumed that $f(x,y)$ and $\partial f / \partial y$ are continuous functions on the rectangle R. But R is closed (in the sense that it includes its boundary) and bounded, so each of these functions is necessarily bounded on R. This means that there exist constants M and K such that

$$|f(x,y)| \le M \qquad (5)$$

and

$$\left|\frac{\partial}{\partial y} f(x,y)\right| \le K \qquad (6)$$

for all points (x,y) in R. We next observe that if (x,y_1) and (x,y_2) are distinct points in R with the same x coordinate, then the mean value theorem guarantees that

$$|f(x,y_1) - f(x,y_2)| = \left|\frac{\partial}{\partial y} f(x,y^*)\right| |y_1 - y_2| \qquad (7)$$

for some number y^* between y_1 and y_2. It is clear from (6) and (7) that

$$|f(x,y_1) - f(x,y_2)| \le K|y_1 - y_2| \qquad (8)$$

for any points (x, y_1) and (x,y_2) in R (distinct or not) that lie on the same vertical line. We now choose h to be any positive number such that

$$Kh < 1 \qquad (9)$$

and the rectangle R^1 defined by the inequalities $|x-x_0| \le h$ and $|y-y_0| \le Mh$ is contained in R. Since (x_0, y_0) is an interior point of R, there is no difficulty in seeing that such an h exists. The reasons for these apparently bizarre requirements will of course emerge as the proof continues.

From this point on, we confine our attention to the interval $|x-x_0| \leq h$. In order to prove (i), it suffices to show that the series

$$|y_0(x)| + |y_1(x) - y_0(x)| + |y_2(x) - y_1(x)| + \ldots + |y_n(x) - y_{n-1}(x)| + \ldots \qquad (10)$$

converges; and to accomplish this, we estimate the terms $|y_n(x) - y_{n-1}(x)|$. It is first necessary to observe that each of the functions $y_n(x)$ has a graph that lies in R^1 and hence in R. This is obvious for $y_0(x) = y_0$, so the points $[t, y_0(t)]$ are in R^1, (5) yields $|f[t,y_0(t)]| \leq M$ and

$$|y_1(x) - y_0| = \left|\int_{x_0}^{x} f\left[t, y_0(t)\right]dt\right| \leq Mh,$$

which proves the statement for $y_1(x)$. It follows in turn from this inequality that the points $[t, y_1(t)]$ are in R^1, so $|f[t,y_1(t)]| \leq M$ and

$$|y_2(x) - y_0| = \left|\int_{x_0}^{x} f\left[t, y_1(t)\right]dt\right| \leq Mh,$$

Similarly,

$$|y_3(x) - y_0| = \left|\int_{x_0}^{x} f\left[t, y_2(t)\right]dt\right| \leq Mh,$$

and so on. Now for the estimates mentioned above,since a continuous function on a closed interval has a maximum, and $y_1(x)$ is continuous, we can define a constant a by $a = \max |y_1(x) - y_0|$ and write

$$|y_1(x) - y_0(x)| \leq a.$$

Next, the points $[t, y_1(t)]$ and $[t, y_0(t)]$ lie in R^1, so (8) yields

$$|f[t,y_1(t)] - f[t, y_0(t)]| \leq K|y_1(t) - y_0(t)| \leq Ka$$

and we have

$$|y_2(x) - y_1(x)| = \left|\int_{x_0}^{x} \left(f\left[t, y_1(t)\right] - f\left[t, y_0(t)\right]\right)dt\right|$$

$$\leq Kah = a(Kh)$$

Similarly.

$$|f[t,y_2(t)] - f[t, y_1(t)]| \leq K\, |y_2(t) - y_1(t)| \leq K^2 a h,$$

so

$$|y_3(x) - y_2(x)| = \left| \int_{x_0}^{x} \left(f[t, y_2(t)] - f[t, y_1(t)] \right) dt \right|$$

$$\leq (K^2 a h)h = a(Kh)^2.$$

By continuing in this manner, we find that

$$|y_n(x) - y_{n-1}(x)| \leq a(Kh)^{n-1}$$

for every n = 1, 2, ….. Each term of the series (10) is therefore less than or equal to the corresponding term of the series of constants

$$|y_0| + a + a(Kh) + a(Kh)^2 + \ldots + a(Kh)^{n-1} + \ldots.$$

But (9) guarantees that this series converges, so (10) converges by the comparison test, (4) converges to a sum which we denote by y(x), and $y_n(x) \to y(x)$. Since the graph of each $y_n(x)$ lies in R^1, it is evident that the graph of y(x) also has this property.

Now for the proof of (ii). The above argument shows not only that $y_n(x)$ converges to y(x) in the interval, but also that this convergence is uniform. This means that by choosing n to be sufficiently large, we can make $y_n(x)$ as close as we please to y(x) for all x in the interval; or more precisely, if $\in > 0$ is given, then there exists a positive integer n_0 such that if $n \geq n_0$ we have $|y(x) - y_n(x)| < \in$ for all x in the interval. Since each $y_n(x)$ is clearly continuous, this uniformity of the convergence implies that the limit function y(x) is also continuous. To prove that y(x) is actually a solution of (2), we must show that

$$y(x) - y_0 - \int_{x_0}^{x} f[t, y(t)]\, dt = 0. \tag{11}$$

But we know that

$$y_n(x) - y_0 - \int_{x_0}^{x} f[t, y_{n-1}(t)]\, dt = 0, \tag{12}$$

so subtracting the left side of (12) from the left side of (11) gives

$$y(x)-y_0-\int_{x_0}^{x} f[t,y(t)]dt = y(x)-y_n(x)+\int_{x_0}^{x}\left(f[t,y_{n-1}(t)]-f[t,y(t)]\right)dt,$$

and we obtain

$$\left|y(x)-y_0-\int_{x_0}^{x} f[t,y(t)]dt\right|$$

$$\leq |y(x)-y_n(x)| + \left|\int_{x_0}^{x}\left(f[t,y_{n-1}(t)]-f[t,y(t)]\right)dt\right|.$$

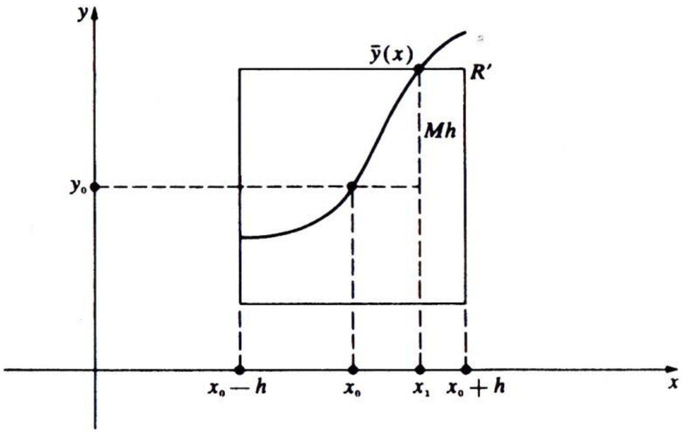
Since the graph of y(x) lies in R^1 and hence in R, (8) yields

$$\left|y(x)-y_0-\int_{x_0}^{x} f[t,y(t)]dt\right| \leq |y(x)-y_n(x)| + Kh \max |y_{n-1}(x)-y(x)|. \qquad (13)$$

The uniformity of the convergence of $y_n(x)$ to y(x) now implies that the right side of (13) can be made as small as we please by taking n large enough. The left side of (13) must therefore equal zero, and the proof of (11) is complete.

In order to prove (iii), we assume that $\bar{y}(x)$ is also a continuous solution of (2) on the interval $|x-x_0| \leq h$, and we show that $\bar{y}(x)$ = y(x) for every x in the interval. For the argument we give, it is necessary to know that the graph of $\bar{y}(x)$ lies in R^1 and hence in R, so our first step is to establish this fact. Let us suppose that the graph of $\bar{y}(x)$ leaves R^1 (Fig.below). Then the properties of this function [continuity and the fact that $\bar{y}(x_0)=y_0$] imply that there exists an x_1 such that $|x_1-x_0| < h$, $|\bar{y}(x_1)-y_0|$ = Mh, and $|\bar{y}(x)-y_0| < Mh$ if $|x-x_0| < |x_1-x_0|$. It follows that

$$\frac{|\bar{y}(x_1)-y_0|}{|x_1-x_0|} = \frac{Mh}{|x_1-x_0|} > \frac{Mh}{h} = M.$$



However, by the mean value theorem there exists a number x^* between x_0 and x_1 such that

$$\frac{\left|\bar{y}(x_1)-y_0\right|}{\left|x_1-x_0\right|}=\left|\bar{y}^1(x^*)\right|=\left|f\left[x^*,\bar{y}(x^*)\right]\right|\le M,$$

since the point $\left[x^*,\bar{y}(x^*)\right]$ lies in R^1. This contradiction shows that no point with the properties of x_1 can exist, so the graph of $\bar{y}(x)$ lies in R^1. To complete the proof of (iii), we use the fact that $\bar{y}(x)$ and y(x) are both solution of (2) to write

$$|\bar{y}(x) - y(x)| = \left|\int_{x_0}^{x}\left\{f\left[t,\bar{y}(t)\right]-f\left[t,y(t)\right]\right\}dt\right|.$$

Since the graphs of $\bar{y}(x)$ and y(x) both lie in R^1, (8) yields.

$$|\bar{y}(x)-y(x)| \le Kh \max |\bar{y}(x) - y(x)|,$$

so

$$\max |\bar{y}(x) - y(x)| \le Kh \max |\bar{y}(x) - y(x)|$$

This implies that $\max |\bar{y}(x) - y(x)| = 0$, for otherwise we would have $1 \le Kh$ in contradiction to (9). It follows that $\bar{y}(x) = y(x)$ for every x in the interval $|x - x_0| \le h$, and Picard's theorem is fully proved.

Remark 1. This theorem can be strengthened in various ways by weakening its hypotheses. For instance, our assumption that $\partial f/\partial y$ continuous on R is stronger than the proof requires, and is used only to obtain the inequality (8). We can therefore introduce this inequality into the theorem as an assumption that replaces the one about $\partial f/\partial y$. In this way we arrive at a stronger form of the theorem since there are many functions that lack a continuous partial derivative but nevertheless satisfy (8) for some constant K. This inequality, which says that the difference quotient.

$$\frac{f(x,y_1)-f(x,y_2)}{y_1-y_2}$$

is bounded on R, is called a Lipschitz condition in the variable y.

Theorem B.

Let f(x,y) be a continuous function that satisfies a Lipschitz condition

$|f(x,y_1) - f(x,y_2)| \leq K\,|y_1 - y_2|$

on a strip defined by $a \leq x \leq b$ and $-\infty < y < \infty$. If (x_0, y_0) is any point of the strip, then the initial value problem

$$y^1 = f(x,y), \quad y(x_0) = y_0 \tag{15}$$

has one and only one solution $y = y(x)$ on the interval $a \leq x \leq b$.

Proof. The argument is similar to that given for Theorem A, with certain simplifications permitted by the fact that the region under discussion is not bounded above or below. In particular, we start the proof in the same way and show that the series (4) – and therefore the sequence (3) – is uniformly convergent on the whole interval $a \leq x \leq b$. We accomplish this by using a somewhat different method of estimating the terms of the series (10).

First, we define M_0, M_1 and M by

$$M_0 = |y_0|, \quad M_1 = \max |y_1(x)|, \quad M = M_0 + M_1,$$

and we notice that $|y_0(x)| \leq M$ and $|y_1(x) - y_0(x)| \leq M$. Next, if $x_0 \leq x \leq b$, it follows that

$$|y_2(x) - y_1(x)| = \left| \int_{x_0}^{x} \{ f[t, y_1(t)] - f[t, y_0(t)] \} dt \right|$$

$$\leq \int_{x_0}^{x} \left| f[t, y_1(t)] - f[t, y_0(t)] \right| dt$$

$$\leq K \int_{x_0}^{x} \left| y_1(t) - y_0(t) \right| dt$$

$$\leq KM(x - x_0),$$

$$|y_3(x) - y_2(x)| = \left| \int_{x_0}^{x} \{ f[t, y_2(t)] - f[t, y_1(t)] \} \, dt \right|$$

$$\leq K \int_{x_0}^{x} \left| y_2(t) - y_1(t) \right| dt$$

$$\leq K^2 M \int_{x_0}^{x} (t - x_0) dt = K^2 M \frac{(x - x_0)^2}{2},$$

and in general

$$|y_n(x) - y_{n-1}(x)| \leq K^{n-1} M \frac{(x - x_0)^{n-1}}{(n-1)!}.$$

The same argument is also valid for $a \leq x \leq x_0$, provided only that $x - x_0$ is replaced by $|x - x_0|$, so we have

$$|y_n(x) - y_{n-1}(x)| \leq K^{n-1} M \frac{|x - x_0|^{n-1}}{(n-1)!}.$$

$$\leq K^{n-1} M \frac{(b - a)^{n-1}}{(n-1)!}.$$

for every x in the interval and n = 1, 2, …. We conclude that each term of the series (10) is less than or equal to the corresponding term of the convergent series of constants.

$$M + M + KM(b - a) + K^2 M \frac{(b - a)^2}{2!} + K^3 M \frac{(b - a)^3}{3!} + \ldots,$$

so (3) converges uniformly on the interval $a \leq x \leq b$ to a limit function y(x). Just as before, the uniformity of the convergence implies that y(x) is a solution of (15) on the whole interval, and all that remains is to show that it is the only such solution. We assume that $\bar{y}(x)$ is also a solution of (15) on the interval. Our strategy is to show that $y_n(x) \to \bar{y}(x)$ for each x as $n \to \infty$; and since we also have $y_n(x) \to y(x)$, it will follow that $\bar{y}(x) = y(x)$. We begin by observing that $\bar{y}(x)$ is continuous and satisfies the equation.

$$\bar{y}(x) = y_0 + \int_{x_0}^{x} f\left[t, \bar{y}(t)\right] dt.$$

If $A = \max |\bar{y}(x) - y_0|$, then for $x_0 \le x \le b$ we see that

$$|\bar{y}(x) - y_1(x)| = \left| \int_{x_0}^{x} \left\{ f\left[t, \bar{y}(t)\right] - f\left[t, y_0(t)\right] \right\} dt \right|$$

$$\le \int_{x_0}^{x} \left| f\left[t, \bar{y}(t)\right] - f\left[t, y_0(t)\right] \right| dt$$

$$\le K \int_{x_0}^{x} \left| \bar{y}(t) - y_0 \right| dt$$

$$\le KA(x - x_0),$$

$$\left| \bar{y}(x) - y_2(x) \right| = \left| \int_{x_0}^{x} \left\{ f\left[t, \bar{y}(t)\right] - f\left[t, y_1(t)\right] \right\} dt \right|$$

$$\le K \int_{x_0}^{x} \left| \bar{y}(t) - y_1(t) \right| dt$$

$$\le K^2 A \int_{x_0}^{x} (t - x_0)\, dt = K^2 A \frac{(x - x_0)^2}{2},$$

and in general

$$|\bar{y}(x) - y_n(x)| \le K^n A \frac{(x - x_0)^n}{n!}$$

A similar result holds for $a \le x \le x_0$, so for any x in the interval we have

$$|\bar{y}(x) - y_n(x)| \le K^n A \frac{|x - x_0|^n}{n!} \le K^n A \frac{(b - a)^n}{n!}.$$

Since the right side of this approaches zero as $n \to \infty$, we conclude that $\bar{y}(x) = y(x)$ for every x in the interval, and the proof is complete.

VI.3 SECOND ORDER LINEAR EQUATION

Consider, for example, the initial value problem consisting of the following pair of first order equations and initial conditions:

$$\begin{cases} \dfrac{dy}{dx} = f(x, y, z), & y(x_0) = y_0, \\ \dfrac{dz}{dx} = g(x, y, z), & z(x_0) = z_0, \end{cases} \tag{1}$$

where the right sides are continuous functions in some region of xyz space that contains the point (x_0, y_0, z_0). We use the differential notation here in order to emphasize that x is the independent variable. A solution of such a system is of course a pair of functions y = y(x) and z = z(x) which together satisfy the conditions imposed by (1) on some interval containing the point x_0. As in the case of a single first order equation, it is apparent that the system (1) is equivalent to the system of integral equations.

$$\begin{cases} y(x) = y_0 + \int_{x_0}^{x} f[t, y(t), z(t)]dt, \\ z(x) = z_0 + \int_{x_0}^{x} g[t, y(t), z(t)]dt, \end{cases} \tag{2}$$

in the sense that the solution of (1) – if any exist – are precisely the continuous solutions of (2). If we attempt to solve (2) by successive approximations beginning with the constant functions.

$$y_0(x) = y_0 \quad \text{and } z_0(x) = z_0,$$

then the Picard method proceeds exactly as before. At the first stage we have

$$\begin{cases} y_1(x) = y_0 + \int_{x_0}^{x} f[t, y_0(t), z_0(t)]dt, \\ z_1(x) = z_0 + \int_{x_0}^{x} g[t, y_0(t), z_0(t)]dt; \end{cases}$$

at the second stage we have

$$\begin{cases} y_2(x) = y_0 + \int_{x_0}^{x} f[t, y_1(t), z_1(t)]dt, \\ z_2(x) = z_0 + \int_{x_0}^{x} g[t, y_1(t), z_1(t)]dt; \end{cases}$$

and so on. This procedure generates two sequences of functions $y_n(x)$ and $z_n(x)$; and under suitable hypotheses, these sequences converge to a solution of (1) which exists and is unique on some interval $|x-x_0| \leq h$.

We now specialize to a linear system, in which the functions f(x,y,z) and g(x,y,z) in (1) are linear functions of y and z. That is, we consider an initial value problem of the form.

$$\begin{cases} \frac{dy}{dx} = p_1(x)y + q_1(x)z + r_1(x), \quad y(x_0) = y_0, \\ \frac{dz}{dx} = p_2(x)y + q_2(x)z + r_2(x), \quad z(x_0) = z_0, \end{cases} \tag{3}$$

where the six functions $p_i(x)$, $q_i(x)$, and $r_j(x)$ are continuous on an interval $a \leq x \leq b$ and x_0 is a point in this interval. Since each of these functions is bounded for $a \leq x \leq b$, there exists a constant K such that $|p_i(x)| \leq K$ and $|q_i(x)| \leq K$ for i = 1,2. It is now easy to see that the functions on the right sides of the differential equation in (3) satisfy Lipschitz conditions of the form.

$$| f(x,y,z_1) - f(x,y,z_2) | \leq K(|y_1 - y_2| + |z_1 - z_2|)$$

and

$$| g(x,y_1,z_1) - g(x,y_2,z_2) | \leq K(|y_1 - y_2| + |z_1 - z_2|).$$

These conditions can be used to show that (3) has a unique solution on the whole interval $a \leq x \leq b$.

Theorem A.

Let P(x), Q(x) and R(x) be continuous functions on an interval $a \leq x \leq b$. If x_0 is any point in this interval, and y_0 and y_0^1 are any numbers whatever, then the initial value problem.

$$\frac{d^2y}{dx^2} + P(x)\frac{dy}{dx} + Q(x)y = R(x), y(x_0) = y_0 \text{ and } y^1(x_0) = y_0^1, \tag{4}$$

has one and only one solution y = y(x) on the interval $a \le x \le b$.

Proof.

If we introduce the variable z = dy/dx, them it is clear that every solution of (4) yields a solution of the linear system

$$\begin{cases} \dfrac{dy}{dx} = z & y(x_0) = y_0, \\ \dfrac{dz}{dx} = -P(x)\,z - Q(x)\,y + R(x), & z(x_0) = y_0^1, \end{cases} \tag{5}$$

and conversely. We have seen that (5) has a unique solution on the internal $a \le x \le b$. so the same is true of (4).

Problem

Solve the following initial value problem by Picard's method, and compare the result with the exact solution:

$$\begin{cases} \dfrac{dy}{dx} = z, & y(0) = 1, \\ \dfrac{dz}{dx} = -y, & z(0) = 0. \end{cases}$$

Solution.

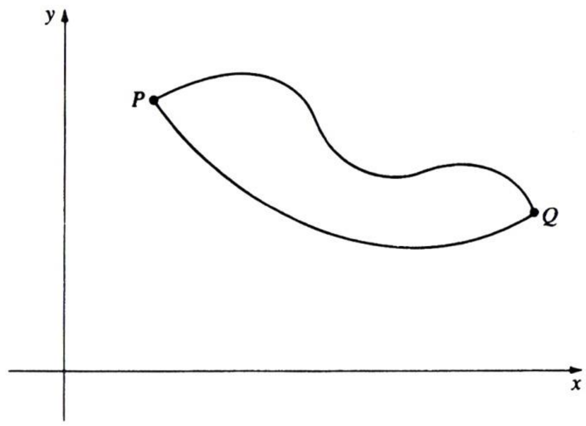
y = cos x, z = – sin x

CHPATER – VII

CALCULUS OF VARIATIONS

VII.1 INTRODUCTION. SOME TYPICAL PROBLEMS OF THE SUBJECT

The Calculus of variations has been one of the major branches of analysis for more than two centuries. It is a tool of great power that can be applied to a wide variety of problems in pure mathematics. It can also be used to express the basic principles of mathematical physics in form of the utmost simplicity and elegance. The flavor of the subject is easy to grasp by considering a few of its typical problems. Suppose that two points P and Q are given in a plane (Fig.). There are infinitely many curves joining these points, and we can ask which of these curves is the shortest. The initutive answer is of course a straight line. We can also ask which curve will generate the surface of revolution of smallest area when revolved about the x-axis, and in this case the answer is far from clear. If we think of a typical curve as a frictionless wire in a vertical plane, then another nontrivial problem is that of finding the curve down which a bead will slide from P to Q in the shortest time.



Every student of elementary calculus is familiar with the problem of finding points at which a function of a single variable has maximum or minimum values. The above problems show that in the calculus of variations we consider some quantity (arc length, surface area, time of descent) that depends on an entire curve, and we seek the curve that minimizes the quantity in question. The calculus of variations also deals with minimum problems depending on surface. For example, if a circular wire is bent in any manner and dipped into a soap solution, then the soap film spanning the wire will assume the shape of the surface of smallest area bounded by the wire. The mathematical problem is to find the surface from this minimum property and the known shape of the wire.

Observe that each of the problems described in this section is a special case of the following more general problem. Let P and Q have coordinates (x_1, y_1) and (x_2, y_2), and consider the family of functions

$$y = y(x) \tag{1}$$

that satisfy the boundary conditions $y(x_1) = y_1$ and $y(x_2) = y_2$–that is, the graph of (1) must join P and Q. Then we wish to find the function in this family that minimizes an integral of the form

$$I(y) = \int_{x_1}^{x_2} f\left(x, y, y^1\right) dx. \tag{2}$$

To see that this problem indeed contains the others, we note that the length of the curve (1) is $\int_{x_1}^{x_2} \sqrt{1+\left(y^1\right)^2}\, dx,$ (3)

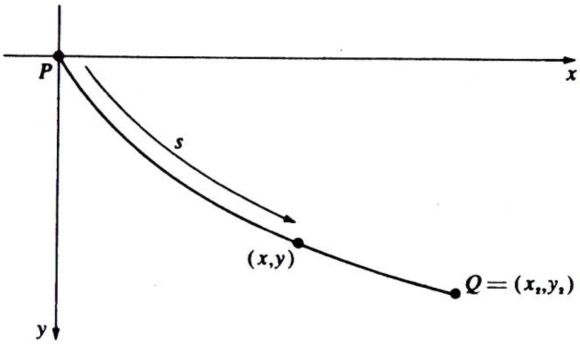
and that the area of the surface of revolution obtained by revolving it about the x-axis is

$$\int_{x_1}^{x_2} 2\pi y \sqrt{1+\left(y^1\right)^2}\, dx, \tag{4}$$

In the case of the curve of quickest descent, it is convenient to invert the coordinate system and take the point P at the origin, as in Figure below.

Since the speed $\nu = ds/dt$ is given by $\nu = \sqrt{2gy}$, the total time of descent is the integral of ds/ν and the integral to be minimized is

$$\int_{x_1}^{x_2} \frac{\sqrt{1+(y^1)^2}}{\sqrt{2gy}} dx \tag{5}$$



Accordingly, the function f(x,y,y[1]) occurring in (2) has the respective forms $\sqrt{1+(y')^2}$, $2\pi y\sqrt{1+(y')^2}$ and $\sqrt{1+(y')^2}\Big/\sqrt{2gy}$ in our three problems.

It is necessary to be somewhat more precise in formulating the basic problem of minimizing the integral (2). First, we will always assume that the function f(x,y,y[1]) has continuous partial derivatives of the second order with respect to x, y, and y[1]. The next question is, what types of functions (1) are to be allowed? The integral (2) is a well-defined real number whenever the integrand is continuous as a function of x, and for this it suffices to assume that y[1](x) is continuous. However, in order to guarantee the validity of the operations we will want to perform, it is convenient to restrict ourselves once and for all to considering only unknown functions y(x) that have continuous second derivatives and satisfy the given boundary conditions $y(x_1) = y_1$ and $y(x_2) = y_2$. Functions of this kind will be called admissible. We can imagine a competition in which only admissible functions are allowed to enter, and the problem is to select from this family the function or functions that yield the smallest value for I.

VII.2 Euler's Differential Equation for an Extremal.

Assuming that there exists an admissible function y(x) that minimizes the integral

$$I = \int_{x_1}^{x_2} f\left(x, y, y^1\right) dx, \qquad (1)$$

how do we find this function? We shall obtain a differential equation for y(x) by comparing the values of I that correspond to neighboring admissible functions. The central idea is that since y(x) gives a minimum value to I, I will increase if we "disturb" y(x) slightly. These disturbed functions are constructed as follows.

Let $\eta(x)$ be any function with the properties that $\eta^{11}(x)$ is continuous and

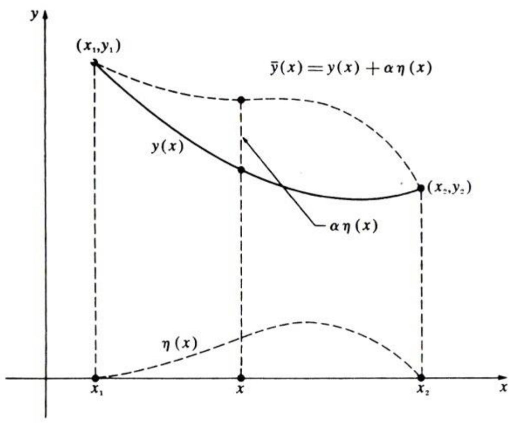
$$\eta(x_1) = \eta(x_2) = 0. \qquad (2)$$

If α is a small parameter, then

$$\bar{y}\,(x) = y(x) + \alpha\eta(x) \qquad (3)$$

represents a one-parameter family of admissible functions. The vertical deviation of a curve in this family from the minimizing curve y(x) is $\alpha\eta(x)$, as shown in Fig. The significance of (3) lies in the fact that for each family of this type, that is, for each choice of the function $\eta(x)$, the minimizing function y(x) belongs to the family and corresponds to the value of the parameter $\alpha=0$.

Now,with $\eta(x)$ fixed, we substitute $\bar{y}\,(x)=y(x)+\alpha\eta(x)$ and $\bar{y}^{\,1}(x)= y^1(x)+\alpha\eta^1(x)$ into the integral (1), and get a function of α,



$$I(\alpha) = \int_{x_1}^{x_2} f\left(x, \bar{y}, \bar{y}'\right)dx$$

$$= \int_{x_1}^{x_2} f\left[x, y(x)+\alpha\eta(x), y'(x)+\alpha\eta'(x)\right]dx. \qquad (4)$$

When α=0, formula (3) yields $\bar{y}$ (x) = y(x); and since y(x) minimizes the integral, we know that I(α) must have a minimum when α=0. By elementary calculus, a necessary condition for this is the vanishing of the derivative $I^1(\alpha)$ when α=0: $I^1(0) = 0$. The derivative $I^1(\alpha)$ can be computed by differentiating (4) under the integral sign, that is,

$$I^1(\alpha) = \int_{x_1}^{x_2} \frac{\partial}{\partial\alpha} f\left(x, \bar{y}, \bar{y}'\right)dx. \qquad (5)$$

By the chain rule for differentiating function of several variables, we have

$$\frac{\partial}{\partial\alpha} f\left(x, \bar{y}, \bar{y}'\right) = \frac{\partial f}{\partial x}\frac{\partial x}{\partial\alpha} + \frac{\partial f}{\partial \bar{y}}\frac{\partial \bar{y}}{\partial\alpha} + \frac{\partial f}{\partial \bar{y}'}\frac{\partial \bar{y}'}{\partial\alpha}$$

$$= \frac{\partial f}{\partial \bar{y}}\eta(x) + \frac{\partial f}{\partial \bar{y}'}\eta^1(x).$$

So (5) can be written as

$$I^1(\alpha) = \int_{x_1}^{x_2}\left[\frac{\partial f}{\partial y}\eta(x) + \frac{\partial f}{\partial \bar{y}'}\eta'(x)\right]dx. \qquad (6)$$

Now $I^1(0) = 0$, so putting α = 0 in (6) yields.

$$\int_{x_1}^{x_2}\left[\frac{\partial f}{\partial y}\eta(x) + \frac{\partial f}{\partial y'}\eta'(x)\right]dx = 0 \qquad (7)$$

In this equation the derivative $\eta^1(x)$ appears along with the function η(x). We can eliminate $\eta^1(x)$ by integrating the second term by parts, which gies

$$\int_{x_1}^{x_2}\frac{\partial f}{\partial y}\eta'(x)\,dx = \left[\eta(x)\frac{\partial f}{\partial y'}\right]_{x_1}^{x_2} - \int_{x_1}^{x_2}\eta(x)\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)dx$$

$$= -\int_{x_1}^{x_2}\eta(x)\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)dx$$

by virtue of (2). We can therefore write (7) in the form

$$\int_{x_1}^{x_2} \eta(x)\left[\frac{\partial f}{\partial y}-\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right]dx=0. \qquad (8)$$

Our reasoning up to this point is based on a fixed choice of the function $\eta(x)$. However, since the integral in (8) must vanish for every such function, we at once conclude that the expression in brackets must also vanish. This yields

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)-\frac{\partial f}{\partial y}=0. \qquad (9)$$

which is Euler's equation.

It is important to have a clear understanding of the exact nature of our conclusion: namely, if y(x) is an admissible function that minimizes the integral (1), then y satisfies Euler's equation. Suppose an admissible function y can be found that satisfies this equation. Does this mean that y minimizes I? Not necessarily. The situation is similar to that in elementary calculus, where a function g(x) whose derivative is zero at a point x_0 may have a maximum, a minimum, or a point of inflection at x_0. When no distinctions are made, these cases are often called stationary values of g(x), and the points x_0 at which they occur are stationary points. In the same way, the condition $I^1(0) = 0$ can perfectly well indicate a maximum or point of inflection for $I(\alpha)$ at $\alpha = 0$, instead of a minimum. Thus it is customary to call any admissible solution of Euler's equation a stationary function or stationary curve, and to refer to the corresponding value of the integral (1) as a stationary value of this integral – without committing ourselves as to which of the several possibilities actually occurs. Furthermore, solution of Euler's equation which are unrestricted by the boundary conditions are called extremals.

Euler's equation (9) is not very illuminating. In order to interpret it and convert it into a useful tool, we begin by emphasizing that the partial derivatives $\partial f/\partial y$ and $\partial f/\partial y^1$ are computed by treating x, y, and y^1 as independent variables. In general, however, $\partial f/\partial y^1$ is a function of x explicitly, and also implicitly through y and y^1, so the first term in (9) can be written in the expanded form

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y'}\right)+\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y'}\right)\frac{dy}{dx}+\frac{\partial}{\partial y'}\left(\frac{\partial f}{\partial y'}\right)\frac{dy'}{dx}.$$

Accordingly, Euler's equation is

$$f_{y^1y^1}\frac{d^2y}{dx^2}+f_{y^1y}\frac{dy}{dx}+\left(f_{y^1x}-f_y\right)=0. \tag{10}$$

This equation is of the second order unless $f_{y^1y^1}=0$, so in general the extremals–its solutions–constitute a two-parameter family of curves; and among these, the stationary functions are those in which the two parameters are chosen to fit the given boundary conditions. A second order nonlinear equation like (10) is usually impossible to solve, but fortunately many applications lead to special cases that can be solved.

CASE A. If x and y are missing from the functions *f*, then Euler's equation reduces to $f_{y^1y^1}\frac{d^2y}{dx^2}=0;$

and if $f_{y^1y^1}\neq 0$, we have $d^2y/dx^2 = 0$ and $y = c_1x + c_2$, so the extremals are all straight lines.

CASE B. If y is missing from the function *f*, then Euler's equation becomes.

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)=0,$$

and this can be integrated at once to yield the first order equation

$$\frac{\partial f}{\partial y'}=c_1$$

for the extremals.

CASE C. If x is missing from the function *f*, then Euler's equation can be integrated to

$$\frac{\partial f}{\partial y'}y'-f=c_1.$$

This follows from the identity

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}y' - f\right) = y'\left[\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y}\right] - \frac{\partial f}{\partial x},$$

since $\partial f/\partial x=0$ and the expression in brackets on the right is zero by Euler's equations.

Example 1. To find the shortest curve joining two points (x_1, y_1) and (x_2, y_2) – which we know intuitively to be a straight line – we must minimize the are length integral

$$I = \int_{x1}^{x2} \sqrt{1+(y')^2}\, dx.$$

The variables x and y are missing from $f(y^1) = \sqrt{1+(y')^2}$, so this problem falls under case A. Since

$$f_{y^1 y^1} = \frac{\partial^2 f}{\partial y'^2} = \frac{1}{\left[1+(y')^2\right]^{3/2}} \neq 0, \quad ,$$

Case A tells us that the extremals are the two-parameter family of straight lines $y = c_1 x + c_2$. The boundary conditions yield.

$$y - y_1 = \frac{y_2 - y_1}{x_2 - x_1}(x - x_1) \tag{11}$$

as the stationary curve, and this is of course the straight line joining the two points. It should be noted that this analysis shows only that if I has a stationary value, then the corresponding stationary curve must be the straight line (11). However, it is clear from the geometry that I has no maximizing curve but does have a minimizing curve, so we conclude in this way that (11) actually is the shortest curve joining our two points.

In this example we arrived at an obvious conclusion by analytical means. A much more difficult and interesting problem is that of finding the shortest curve joining two fixed points on a given surface and lying entirely on that surface. These curves are called geodesics, and the study of their properties is one of the focal points of the branch of mathematics known as differential geometry.

Example 2. To find the curve joining the points (x_1, y_1) and (x_2, y_2) that yields a surface of revolution of minimum area when revolved about the x-axis, we must minimize

$$I = \int_{x_2}^{x_2} 2\pi y\sqrt{1+(y')^2}\, dx. \tag{12}$$

The variable x is missing from $f(y, y^1) = 2\pi y\sqrt{1+(y')^2}$, so Case C tells us that Euler's equation becomes

$$\frac{y(y')^2}{\sqrt{1+(y')^2}} - y\sqrt{1+(y')^2} = c_1,$$

which simplifies to

$$c_1 y^1 = \sqrt{y^2 - c_1^2}.$$

On separating variables and integrating, we get

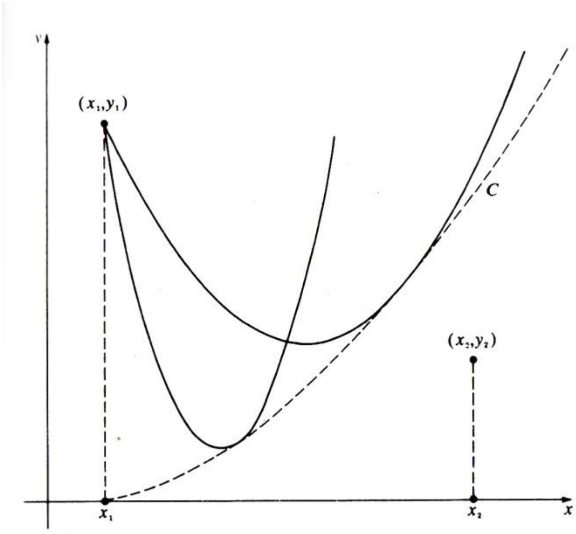
$$x = c_1\int\frac{dy}{\sqrt{y^2-c_1^2}} = c_1\log\left(\frac{y+\sqrt{y^2-c_1^2}}{c_1}\right) + c_2,$$

and solving for y gives

$$y = c_1 \cosh\left(\frac{x-c_2}{c_1}\right). \tag{13}$$

The extremals are therefore catenaries, and the required minimal surface – if it exists – must be obtained by revolving a catenary. The next problem is that of seeing whether the parameters c_1 and c_2 can indeed be chosen so that the curve (13) joins the points (x_1, y_1) and (x_2, y_2).

The choosing of these parameters turns out to be curiously complicated. If the curve (13) is made to pass through the first point (x_1, y_1) then one parameter is left free. Two members of this one-parameter family are shown in Fig. It can be proved that all such curves are tangent to the dashed curve C, so no curve in the family crosses C. Thus, when the second point (x_2, y_2) is below C, as in Fig. there is no catenary through both points and no stationary function exists. In this case it is found that smaller and smaller surfaces are generated by curves that approach the dashed line from (x_1, y_1) to $(x_1, 0)$ to $(x_2, 0)$ to (x_2, y_2), so no admissible curve can generate a minimal surface. When the second point lies above C, there are two catenaries through the points, and hence two stationary functions, but only the upper catenary generates a minimal surface. Finally, when the second point is on C, there is only one stationary function but the surface it generates is not minimal.



Example 3. To find the curve of quickest descent, we must minimize.

$$I = \int_{x_1}^{x_2} \frac{\sqrt{1+(y')^2}}{\sqrt{2gy}}\, dx.$$

Again the variable x is missing from the function $f(y,y^1) = \sqrt{1+(y')^2}\,/\sqrt{2gy}$, so by Case C, Euler's equation becomes

$$\frac{(y')^2}{\sqrt{y}\sqrt{1+(y')^2}} - \frac{\sqrt{1+(y')^2}}{\sqrt{y}} = c_1.$$

This reduces to $y[1+(y^1)^2] = c$,

The resulting stationary curve is the cycloid

$$x = a\,(\theta - \sin\theta) \quad \text{and} \quad y = a(1 - \cos\theta) \tag{14}$$

generated by a circle of radius a rolling under the x axis, where a is chosen so that the first inverted arch passes through the point (x_2, y_2). As before, this argument shows only that if I has a minimum, then the corresponding stationary curve must be the cycloid (14). However, it is reasonably clear from physical consideration that I has no maximizing curve but does have a minimizing curve, so this cycloid actually minimizes the time of descent.

Suppose we want to find conditions necessarily satisfied by two functions y(x) and z(x) that give a stationary value to the integral

$$I = \int_{x_1}^{x_2} f(x, y, z, y', z')dx, \tag{15}$$

where the boundary values $y(x_1)$, $z(x_1)$ and $y(x_2)$, $z(x_2)$ are specified in advance. Just as before, we introduce functions $\eta_1(x)$ and $\eta_2(x)$ that have continuous second derivatives and vanish at the end points. From these we form the neighboring functions $\bar{y}(x) = y(x) + \alpha\eta_1(x)$ and $\bar{z}(x) = z(x) + \alpha\eta_2(x)$, and then consider the functions of α defined by

$$I(\alpha) = \int_{x_1}^{x_2} f(x, y + \alpha\eta_1, z + \alpha\eta_2, y' + \alpha\eta_1', z' + \alpha\eta_2')dx. \tag{16}$$

Again, if y(x) and z(x) are stationary functions we must have $I^1(0) = 0$, so by computing the derivative of (16) and putting $\alpha = 0$ we get

$$\int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y}\eta_1 + \frac{\partial f}{\partial z}\eta_2 + \frac{\partial f}{\partial y'}\eta_1' + \frac{\partial f}{\partial z'}\eta_2' \right) dx = 0,$$

or, if the terms involving η_1' and η_2' are integrated by parts,

$$\int_{x_1}^{x_2} \left\{ \eta_1(x)\left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'} \right) \right] + \eta_2(x)\left[\frac{\partial f}{\partial z} - \frac{d}{dx}\left(\frac{\partial f}{\partial z'} \right) \right] \right\} dx = 0. \tag{17}$$

Finally, since (17) must hold for all choices of the functions $\eta_1(x)$ and $\eta_2(x)$, we are led at once to Euler's equations.

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'} \right) - \frac{\partial f}{\partial y} = 0 \text{ and } \frac{d}{dx}\left(\frac{\partial f}{\partial z'} \right) - \frac{\partial f}{\partial z} = 0. \tag{18}$$

Thus, to find the extremals of our problem, we must solve the system (18). Needless to say, a system of intractable equations is harder to solve than only one; but if (18) can be solved, then the stationary functions are determined by fitting the resulting solutions to the given boundary conditions. Similar considerations apply without any essential change to integrals like (15) which involve more than two unknown functions.

Problems

1. Find the extremals for the integral (1) if the integrand is

   a. $\dfrac{\sqrt{1+(y')^2}}{y}$;

   b. $y^2 - (y^1)^2$.

2. Find the stationary functions of

$$\int_0^4 \left[xy' - (y')^2\right] dx$$

   which is determined by the boundary conditions y(0) = 0 and y(4) = 3.

Soln.

1. a. $(x - c_2)^2 + y^2 = c_1^2$ b. $y = c_1 \sin(x - c_2)$
2. $y = \frac{1}{4}(x^2 - x)$

VII. 3 ISOPERIMETRIC PROBLEMS

The ancient Greeks proposed the problem of finding the closed plane curve of given length that encloses the largest area. They called this the isoperimetric problem, and were able to show in a more or less rigorous manner that the obvious answer – a circle – is correct. If the curve is expressed parametrically by x = x(t) and y = y(t), and is traversed once counter clockwise as t increases from t_1 to t_2, then the enclosed area is known to be

$$A = \frac{1}{2}\int_{t_1}^{t_2}\left(x\frac{dy}{dt} - y\frac{dx}{dt}\right)dt, \quad (1)$$

which is an integral depending on two unknown functions. Since the length of the curve is

$$L = \int_{t_1}^{t_2}\sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\,dt, \quad (2)$$

the problem is to maximize (1) subject to the side condition that (2) must have a constant value. The term isoperimetric problem is usually extended to include the general case of finding extremals for one integral subject to any constraint requiring a second integral to take on a prescribed value.

We will also consider finite side conditions, which do not involve integrals or derivatives. For example, if

$$G(x,y,z) = 0 \tag{3}$$

is a given surface, then a curve on this surface is determined parametrically by three functions x = x(t), y = y(t), and z = z(t) that satisfy equation (3), and the problem of finding geodesics amounts to the problem of minimizing the are length integral

$$\int_{t_1}^{t_2} \sqrt{\left(\frac{dx}{dx}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}\, dt \tag{4}$$

subject to the side condition (3).

Lagrange multipliers. It is necessary to begin by considering some problems in elementary calculus that are quite similar to isoperimetric problems. For example, suppose we want to find the points (x,y) that yield stationary values for a function z = *f*(x,y), where, however, the variables x and y are not independent but are constrained by a side condition.

$$g(x,y) = 0. \tag{5}$$

The usual procedure is to arbitrarily designate one of the variables x and y in (5) as independent, say x, and the other as dependent on it, so that dy/dx can be computed from

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

We next use the fact that since z is now a function of x alone, dz/dx = 0 is a necessary condition for z to have a stationary value, so

$$\frac{dz}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} = 0.$$

or

$$\frac{\partial f}{\partial x} - \frac{\partial f}{\partial y}\,\frac{\partial g/\partial x}{\partial g/\partial y} \;=\; 0 \qquad (6)$$

On solving (5) and (6) simultaneously, we obtain the required points (x,y).

One drawback to this approach is that the variables x and y occur symmetrically but are treated unsymmetrically. It is possible to solve the same problem by a different and more elegant method that also has many practical advantages. We form the function

$$F(x, y, \lambda) = f(x,y) + \lambda g(x.y)$$

and investigate its unconstrained stationary values by means of the necessary conditions

$$\begin{aligned} \frac{\partial F}{\partial x} &= \frac{\partial f}{\partial x} + \lambda\frac{\partial g}{\partial x} = 0, \\ \frac{\partial F}{\partial y} &= \frac{\partial f}{\partial y} + \lambda\frac{\partial g}{\partial y} = 0, \\ \frac{\partial F}{\partial \lambda} &= g\left(x, y\right) = 0 \end{aligned} \qquad (7)$$

If λ is eliminated from the first two of these equations, then the system clearly reduces to

$$\frac{\partial f}{\partial x} - \frac{\partial f}{\partial y}\,\frac{\partial g/\partial x}{\partial g/\partial y} \;=\; 0 \quad \text{and} \quad g(x,y) = 0,$$

and this is the system obtained in the above paragraph. It should be observed that this technique (solving the system (7) for x and y) solves the given problem in a way that has two major features important for theoretical work: it does not disturb the symmetry of the problems by making an arbitrary choice of the independent variable; and it removes the side condition at the small expense of introducing λ as another variable. The parameter λ is called a Lagrange multiplier, and this method is known as the method of Lagrange multipliers. This discussion extends in an obvious manner to problems involving functions of more than two variables with several side conditions.

Integral side conditions. Here we want to find the differential equation that must be satisfied by a functions y(x) that gives a stationary value to the integral

$$I = \int_{x_1}^{x_2} f(x, y, y')\, dx, \tag{8}$$

Where y is subject to the side condition

$$J = \int_{x_1}^{x_2} g(x, y, y')\, dx = c \tag{9}$$

and assumes prescribed values $y(x_1) = y_1$ and $y(x_2) = y_2$ at the endpoints. As before, we assume that y(x) is the actual stationary function and disturb it slightly to find the desired analytic condition. However, this problem cannot be attacked by our earlier method of considering neighboring functions of the form $\bar{y}(x) = y(x) + \alpha\eta(x)$, for in general these will not maintain the second integral J at the constant value c. Instead, we consider a two-parameter family of neighboring functions

$$\bar{y}(x) = y(x) + \alpha_1\eta_1(x) + \alpha_2\eta_2(x), \tag{10}$$

where $\eta_1(x)$ and $\eta_2(x)$ have continuous second derivatives and vanish at the endpoints. The parameter α_1 and α_2 are not independent, but are related by the condition that

$$J(\alpha_1, \alpha_2) = \int_{x_1}^{x_2} g\ (x, \bar{y}, \bar{y}')\ dx = c \tag{11}$$

Our problem is then reduced to that of finding necessary conditions for the function

$$I(\alpha_1, \alpha_2) = \int_{x_1}^{x_2} f\,(x, \bar{y}, \bar{y}')\ dx \tag{12}$$

to have a stationary value at $\alpha_1 = \alpha_2 = 0$, where α_1 and α_2 satisfy (11). This situation is made to order for the method of Lagrange multipliers. We therefore introduce the function

$$K(\alpha_1, \alpha_2, \lambda) = I\,(\alpha_1, \alpha_2) + \lambda J\,(\alpha_1, \alpha_2)$$

$$= \int_{x_1}^{x_2} F(x, \bar{y}, \bar{y}')\ dx \quad ' \tag{13}$$

where

$$F = f + \lambda g,$$

and investigate its unconstrained stationary value at $\alpha_1 = \alpha_2 = 0$ by means of the necessary conditions

$$\frac{\partial K}{\partial \alpha_1} = \frac{\partial K}{\partial \alpha_2} = 0 \quad \text{when } \alpha_1 = \alpha_2 = 0. \tag{14}$$

If we differentiate (13) under the integral sign and use (10), we get

$$\frac{\partial K}{\partial \alpha_i} = \int_{x_1}^{x_2} \left[\frac{\partial F}{\partial \bar{y}}\, \eta_i(x) + \frac{\partial F}{\partial \bar{y}\prime}\, \eta_i{}'(x)\right] dx \text{ for } i = 1, 2;$$

and setting $\alpha_1 = \alpha_2 = 0$ yields

$$\int_{x_1}^{x_2} \left[\frac{\partial F}{\partial y}\, \eta_i(x) + \frac{\partial F}{\partial y\prime}\, \eta_i{}'(x)\right] dx = 0$$

by virtue of (14). After the second term is integrated by parts, this becomes

$$\int_{x_1}^{x_2} \eta_i(x) \left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right] dx = 0 \tag{15}$$

Since $\eta_1(x)$ and $\eta_2(x)$ are both arbitrary, the two conditions embodied in (15) amount to only one condition, and as usual we conclude that the stationary functions y(x) must satisfy Euler's equation.

$$\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) - \frac{\partial F}{\partial y} = 0 \tag{16}$$

The solutions of this equation (the extremals of our problems) involve three undetermined parameters: two constants of integration, and the Lagrange multiplier λ. The stationary function is then selected from these extremals by imposing the two boundary conditions and giving the integral J its prescribed value c.

In the case of integrals that depend on two or more function, this result can be extended in the same way as in the previous section. For example, if

$$I = \int_{x_1}^{x_2} f(x, y, z, y', z')\, dx$$

has a stationary value subject to the side condition

$$J = \int_{x_1}^{x_2} g(x, y, z, y', z')\, dx = c,$$

then the stationary functions y(x) and z(x) must satisfy the system of equation

$$\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)-\frac{\partial F}{\partial y}=0 \quad \text{and} \quad \frac{d}{dx}\left(\frac{\partial F}{\partial z'}\right)-\frac{\partial F}{\partial z}=0,$$

where F = f + λg. The reasoning is similar to that already given, and we omit the details.

Example 1. We shall find the curve of fixed length L that joins the points (0,0) and (1,0), lies above the x-axis, and encloses the maximum area between itself and the x-axis. This is a restricted version of the original isoperimetric problem in which part of the curve surrounding the area to be maximized is required to be a line segment of length 1. Our problem is to maximize $\int_0^1$ y dx subject to the side condition

$$\int_0^1 \sqrt{1+(y')^2}\,dx=L$$

and the boundary conditions y(0) = 0 and y(1) = 0. Here we have

F = y+$\lambda\sqrt{1+(y')^2}$, so Euler's equation is

$$\frac{d}{dx}\left(\frac{\lambda y'}{\sqrt{1+(y')^2}}\right)-1=0, \tag{18}$$

or, after carrying out the differentiation,

$$\frac{y''}{\left[1+(y')^2\right]^{3/2}}=\frac{1}{\lambda}. \tag{19}$$

In this case no integration is necessary, since (19) tells us at once that the curvature is constant and equals 1/λ. It follows that the required maximizing curve is an arc of a circle (as might have been expected) with radius λ. As an alternate procedure, we can integrate (18) to get

$$\frac{y'}{\sqrt{1+(y')^2}}=\frac{x-c_1}{\lambda}.$$

On solving this for y^1 and integrating again, we obtain

$$(x-c_1)^2+(y-c_2)^2=\lambda^2, \tag{20}$$

which of course is the equation of a circle with radius λ.

Example 2. In Example 1 it is clearly necessary to have L>1. Also, it L> π/2, the circular arc determined by (20) will not define y>0 as a single-valued function of x. We can avoid these artificial issues by considering curves in parametric form x=x(t) and y=y(t) and by turning our attention to the original isoperimetric problem of maximizing.

$$\frac{1}{2}\int_{t_1}^{t_2}(x\dot{y}-y\dot{x})dt$$

(where $\dot{x}$ = dx/dt and $\dot{y}$ = dy/dt) with the side condition

$$\int_{t_1}^{t_2}\sqrt{\dot{x}^2+\dot{y}^2}\,dt=L.$$

Here we have

$$F = \frac{1}{2}(x\dot{y}+y\dot{x})+\lambda\sqrt{\dot{x}^2+\dot{y}^2},$$

so the Euler equations (17) are

$$\frac{d}{dt}\left(-\frac{1}{2}y+\frac{\lambda\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2}}\right)-\frac{1}{2}\dot{y}=0$$

and

$$\frac{d}{dt}\left(\frac{1}{2}x+\frac{\lambda\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2}}\right)+\frac{1}{2}\dot{x}=0.$$

These equations can be integrated directly, which yields

$$-y+\frac{\lambda\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2}}=c_1 \quad \text{and} \quad x+\frac{\lambda\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2}}=c_2.$$

If we solve for $x-c_2$ and $y-c_1$, square, and add, then the result is

$$(x - c_2)^2 + (y - c_1)^2 = \lambda^2,$$

so the maximizing curve is a circle. This result can be expressed in the following way: if L is the length of a closed plane curve that encloses an area A, then $A \leq L^2/4\pi$, with equality if and only if the curve is a circle. A relation of this kind is called an isoperimetric inequality.

Finite side conditions. At the beginning of this section we formulated the problem of finding geodesics on a given surface.

$$G(x,y,z) = 0. \tag{21}$$

We now consider the slightly more general problem of finding a space curve x = x(t), y = y(t), z = z(t) that gives a stationary value to an integral of the form.

$$\int_{t_1}^{t_2} f\left(\dot{x},\dot{y},\dot{z}\right)dt, \tag{22}$$

where the curve is required to lie on the surface (21).

Our strategy is to eliminate the side condition (21), and to do this we proceed as follow. There is no loss of generality in assuming that the curve lies on a part of the surface where $G_z \neq 0$. On this part of the surface we can solve (21) for z, which gives z = g(x,y) and

$$\dot{z} = \frac{\partial g}{\partial x}\dot{x} + \frac{\partial g}{\partial y}\dot{y}. \tag{23}$$

When (23) is inserted in (22), our problem is reduced to that of finding unconstrained stationary functions for the integral

$$\int_{t_1}^{t_2} f\left(\dot{x},\dot{y},\frac{\partial g}{\partial x}\dot{x} + \frac{\partial g}{\partial y}\dot{y}\right)dt.$$

We know from the previous section that the Euler equations for this problem are

$$\frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}} + \frac{\partial f}{\partial \dot{z}}\frac{\partial g}{\partial x}\right) - \frac{\partial f}{\partial \dot{z}}\frac{\partial \dot{z}}{\partial x} = 0,$$

and

$$\frac{d}{dt}\left(\frac{\partial f}{\partial \dot{y}} + \frac{\partial f}{\partial \dot{z}}\frac{\partial g}{\partial y}\right) - \frac{\partial f}{\partial \dot{z}}\frac{\partial \dot{z}}{\partial y} = 0.$$

It follows from (23) that

$$\frac{\partial \dot{z}}{\partial x} = \frac{d}{dt}\left(\frac{\partial g}{\partial x}\right) \quad \text{and} \quad \frac{\partial \dot{z}}{\partial y} = \frac{d}{dt}\left(\frac{\partial g}{\partial y}\right),$$

so the Euler equation can be written in the form

$$\frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}}\right) + \frac{\partial g}{\partial x}\frac{d}{dt}\left(\frac{\partial f}{\partial \dot{z}}\right) = 0 \quad \text{and} \quad \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{y}}\right) + \frac{\partial g}{\partial y}\frac{d}{dt}\left(\frac{\partial f}{\partial \dot{z}}\right) = 0.$$

If we now define a function $\lambda(t)$ by

$$\frac{d}{dt}\left(\frac{\partial f}{\partial \dot{z}}\right)=\lambda(t)G_z, \tag{24}$$

and use the relations $\partial g/\partial x = -G_x/G_z$ and $\partial g/\partial y = - G_y/G_z$, then Euler's equations become

$$\frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}}\right)=\lambda(t)G_x, \tag{25}$$

and

$$\frac{d}{dt}\left(\frac{\partial f}{\partial \dot{y}}\right)=\lambda(t)G_y, \tag{26}$$

Thus a necessary conditions for a stationary value is the existence of a function $\lambda(t)$ satisfying equations (24),(25) and (26).On eliminating $\lambda(t)$ we obtain the symmetric equations

$$\frac{(d/dt)(\partial f/\partial \dot{x})}{G_x}=\frac{(d/dt)(\partial f/\partial \dot{y})}{G_y}=\frac{(d/dt)(\partial f/\partial \dot{z})}{G_z}, \tag{27}$$

which together with (21) determine the extremals of the problem. It is worth remarking that equations (24), (25) and (26) can be regarded as the Euler equations for the problem of finding unconstrained stationary functions for the integral.

$$\int_{t_1}^{t_2} [f(\dot{x},\dot{y},\dot{z})+\lambda(t)G(x,y,z)]dt.$$

This is very similar to our conclusion for integral side conditions, except that here the multiplier is an undetermined function of t instead of an undetermined constant.

When we specialize this result to the problem of finding geodesics on the surface (21), we have

$$f=\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}.$$

The equations (27) become

$$\frac{(d/dt)(\dot{x}/f)}{G_x}=\frac{(d/dt)(\dot{y}/f)}{G_y}=\frac{(d/dt)(\dot{z}/f)}{G_z}, \tag{28}$$

and the problem is to extract information from this system.

Example 3. If we choose the surface (21) to be the sphere $x^2 + y^2 + z^2 = a^2$, then $G(x, y, z) = x^2 + y^2 + z^2 - a^2$ and (28) is

$$\frac{f\ddot{x}-\dot{x}\dot{f}}{2xf^2} = \frac{f\ddot{y}-\dot{y}\dot{f}}{2yf^2} = \frac{f\ddot{z}-\dot{z}\dot{f}}{2zf^2}$$

which can be rewritten in the form

$$\frac{x\ddot{y}-y\ddot{x}}{x\dot{y}-y\dot{x}}=\frac{\dot{f}}{f}=\frac{y\ddot{z}-z\ddot{y}}{y\dot{z}-z\dot{y}}.$$

If we ignore the middle term, this is

$$\frac{(d/dt)(x\dot{y}-y\dot{x})}{x\dot{y}-y\dot{x}}=\frac{(d/dt)(y\dot{z}-z\dot{y})}{y\dot{z}-z\dot{y}}.$$

One integration gives $x\dot{y}-y\dot{x} = c_1(y\dot{z}-z\dot{y})$ or

$$\frac{\dot{x}+c_1\dot{z}}{x+c_1z} = \frac{\dot{y}}{y}$$

and a second yields $x + c_1z = c_2y$. This is the equation of a plane through the origin, so the geodesics on a sphere are arcs of great circles.

****

* Please follow that y^1 means first derivative and y^{11} means the second derivative.

* Please treat both 'a' and 'a' as same.