THEORY OF ODE

DEVANSH TRIPATHI

Abstract. We shall learn some theory related to ordinary differential equations.

Part 1.
Preliminaries

1. Preliminaries from Real Analysis

Definition 1 (Pointwise convergence)**.** Let I be any interval in $\mathbb{R}$. Let $f_n : I \to \mathbb{R}, n = 1, 2, \ldots,$ be a sequence of functions. We say that the $\{f_n\}$ converges pointwise to a function $f : I \subset \mathbb{R} \to \mathbb{R}$ if the sequence $\{f_n(x)\}$ converges to $f(x)$ for every $x \in I$.

Remark 1. Uniform convergence preserves continuity, interchange of limit and integral.

Theorem 1. *Uniform limit of the sequence of continuous function is continuous.*

Remark 2. Converse of the above theorem is not always true.

Theorem 2 (Cauchy Criterion)**.** *Let $\{f_n\}_{n\geq 1}$ a sequence of function defined on a metric space (X, d_X) with values in a complete metric space (Y, d_Y). Then there exists a function $f : X \to Y$ such that*

$$f_n \to f \text{ uniformly on } X$$

if and only if the following condition is satisfied: For every $\varepsilon > 0$, there exists an integer n_0 such that

$$m, n \geq n_0 \text{ implies } d_Y(f_m(x), f_n(x)) < \varepsilon$$

for every $x \in X$.

Proof. ($\implies$) Assume that the sequence $\{f_n\}$ converges uniformly on X. For given $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that $\forall m > n_0$ and for all $x \in X$

$$d_Y(f_m, f) \leq \frac{\varepsilon}{2}$$

and there exists $n_1 \in \mathbb{N}$ such that $\forall n \geq n_1$

$$d_Y(f_n, f) \leq \frac{\varepsilon}{2}$$

Take $n_3 = \min\{n_0, n_1\}$ then for all $m, n \geq n_3$

$$\begin{aligned} d_Y(f_m, f_n) &\leq d_Y(f_m, f) + d_Y(f, f_n) \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &\leq \varepsilon \text{ for all } x \in X \end{aligned}$$

$(\Longleftarrow)$ Conversely, suppose that $m, n \geq n_0$ implies that $d_Y(f_m, f_n) < \varepsilon$ for all $x \in X$. Then for each $x \in X$, the sequence $\{f_n(x)\}_{n\geq 1}$ is cauchy in a complete space Y and therefore converges in Y. Let $f(x) = \lim\limits_{n\to\infty} f_n(x)$ for each $x \in X$. For $k > 0$ then

$$d_Y(f_n(x), f_{n+k}(x)) < \frac{\varepsilon}{2}$$

for every $k = 0, 1, \dots$ and every $x \in X$.

$$d_Y(f_n(x), f(x)) = \lim_{n\to\infty} d_Y(f_n(x), f_{n+k}(x)) \leq \frac{\varepsilon}{2} < \varepsilon$$

$$d_Y(f_n(x), f(x)) < \varepsilon$$

for every $x \in X$ and $f_n \to f$ uniformly over X. □

Theorem 3 (Weierstrass M-test)**.** *Let $f_1, f_2, \dots$ be a sequence of real valued functions defined on a set X and suppose that*

$$|f_n(x)| \leq M_n$$

for all $x \in X$ and all $n = 1, 2, \dots$ If $\sum\limits_{n=1}^{\infty} M_n$ converges, then $\sum\limits_{n=1}^{\infty} f_n$ converges uniformly.

Proof. If $\sum\limits_{n=1}^{\infty} M_n$ converges, then for given $\varepsilon > 0$,

$$\left|\sum_{k=n}^{m} f_k(x)\right| \leq \sum_{k=n}^{m} |f_k(x)| \leq \sum_{k=n}^{m} M_k < \varepsilon$$

for every $x \in X$ and provided that m, n are sufficiently large. Now, by cauchy criterion, uniform convergence follows. □

Remark 3. For Uniform boundedness and equicontinuity, the underline space is assumed to be compact hence equicontinuity here is same as uniform equicontinuity.

Remark 4. Some authors define pointwise and uniform equicontinuity separately and then on compact space they prove the equivalence.

Definition 2 (Uniform Boundedness)**.** A sequence of functions $\{f_n\}$ defined on I(compact) is said to be uniformly bounded if there exists a constant $M > 0$ such that $|f_k(x)| \leq M$, for all $x \in I$, for all $k \in \mathbb{N}$.

Definition 3 (Equicontinuity)**.** A sequence of functions $\{f_k\}$ defined on I(compact) is said to be equicontinuity on I, if for every $\varepsilon > 0$, there exists $\delta > 0$ such that $|f_k(x) - f_k(y)| < \varepsilon$ whenever $x, y \in I$ and $|x - y| < \delta$, and for all k.

Remark 5. If the family of functions is equicontinuous, then each member in the family is uniformly continuous. Converse may not be true always. For e.g. $f_k(x) = x^k, 0 \leq x \leq 1, k = 1, 2, \dots$ is not an equicontinuous family; however each member is uniformly continuous function (continuous on closed and bounded interval).

Example 1. Each finite set of functions defined on compact set is equicontinuous. If the set is singleton then its trivial. For two element set, take minimum of the δ's and then for finite set take minimum of all δ's .

Theorem 4. *For the sequence of functions f_n and f defined on compact set E and $f_n \to f$ uniformly then the family of functions $A = \{f_n \colon n \in \mathbb{N}\}$ is equicontinuous.*

Proof. Since $f_n \to f$ uniformly, by cauchy criterion of uniform continuity $\exists n_0 \in \mathbb{N}$ such that $\forall m, n \geq n_0$

$$|f_m(x) - f_n(x)| < \frac{\varepsilon}{3}$$

Look at the family of functions $B = \{f_1, f_2, \dots, f_{n_0}\}$, since it is finite, it is equicontinuous. For $k \in \mathbb{N}$

$$|f_k(x) - f_k(y)| < \frac{\varepsilon}{3}$$

whenever $|x - y| < \delta$ for all $f_k \in B$ and for all $x, y \in E$. Now, for each $f_n \in A$

$$\begin{aligned}|f_n(x) - f_n(y)| &\leq |f_n(x) - f_{n_0}(x)| + |f_{n_0}(x) - f_{n_0}(y)| + |f_{n_0}(y) - f_n(y)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\ &< \varepsilon \text{ whenever } |x - y| < \delta\end{aligned}$$

□

Theorem 5 (Arzela-Ascoli)**.** *Let $\{f_k\}$ be a sequence of functions in $C[a, b]$ which is uniformly bounded and equicontinuous. Then, there exists a subsequence $\{f_{k_n}\}$ of $\{f_k\}$ such that $\{f_{k_n}\}$ converges uniformly to a function $f \in C[a, b]$.*

Proof. TODO □

Definition 4 (Lipschitz continuity)**.** A function $f\colon D \subset \mathbb{R} \to \mathbb{R}$, is said to be *locally* Lipschitz in D if for any $x_0 \in D$, there exists a neighbourhood N_{x_0} of x_0 and an $\alpha = \alpha(x_0) > 0$ such that

$$|f(x) - f(y)| \leq \alpha(x_0)|x - y|, \text{ for all } x, y \in N_{x_0}$$

The function $f\colon D \subset \mathbb{R} \to \mathbb{R}$ is said to be *globally* Lipschitz in D if there exists $\alpha > 0$ such that

$$|f(x) - f(y)| \leq \alpha|x - y|, \text{ for all } x, y \in D$$

Remark 6. The smallest α satisfying the above equation is called *Lipschitz constant* of f. It should be finite.

Remark 7. If f is globally Lipschitz, then it is uniformly continuous.(Take $\delta = \frac{\varepsilon}{\alpha}$).

Theorem 6 (Sufficient condition for Lipschitz continuity)**.** *Suppose D is an open interval in $\mathbb{R}$ and $f\colon D \to \mathbb{R}$ is differentiable on D and $\alpha = \sup_{x \in D} |f'(x)| < \infty$. Then, f is Lipschitz with a Lipschitz constant less than or equal to α.*

Proof. Use mean value theorem:

$$\frac{|f(b) - f(a)|}{|b - a|} = f'(c) \leq \alpha$$

And Lipschitz constant is by definition the smallest α satisfying above inequality. Therefore, Lipschitz constant $\leq \alpha$. □

Example 2. Example of Lipschitz continuous functions: polynomials, polynomials of sine and cosine, exponential functions etc.

Definition 5 (Lipschitz continuity for vector valued map)**.** A function $\mathbf{f}(t, \mathbf{y})\colon (a, b) \times D \to \mathbb{R}^n$ is said to be Lipschitz continuous (globally) with respect to $\mathbf{y}$ if there exists $\alpha > 0$ such that

$$|\mathbf{f}(t, \mathbf{y_1}) - \mathbf{f}(t, \mathbf{y_1})| \leq \alpha|\mathbf{y_1} - \mathbf{y_2}|$$

for all $(t, \mathbf{y_1})$ and $(t, \mathbf{y_2})$ in $(a, b) \times D$. α should be finite.

Theorem 7 (Sufficient condition for Lipschitz continuity of $\mathbf{f}(t, \mathbf{y})$)**.** *Let $\mathbf{f}\colon (a, b) \times D \to \mathbb{R}^n$ be a C^1 vector valued function, where D is a convex domain in $\mathbb{R}^n$ such that*

$$\sup_{(t,\mathbf{y}) \in (a,b) \times D} \left| \frac{\partial f_i}{\partial y_j}(t, \mathbf{y}) \right| = \alpha < \infty,$$

for $i, j = 1, 2, \dots, n$. Then , $\mathbf{f}(t, \mathbf{y})$ is Lipschitz continuous on $(a, b) \times D$ with respect to $\mathbf{y}$ having a Lipschitz constant less than or equal to a multiple of α.

Remark 8 (Convex Domain)**.** A set is called a convex domain if it contains all the line segments between any two points of the set.

Remark 9. Lipschitz continuity is a smoothness property stronger than continuity, but weaker than differentiability, locally. There can be functions which are not differentiable but still they can be lipschitz continuous.

Theorem 8 (Calculus Lemma)**.** *Let (a, b) be a finite or infinite interval and $h\colon (a, b) \to \mathbb{R}$ satisfy either*

(i) h is bounded above and non-decreasing or
(ii) h is bounded below and non-increasing,

then, $\lim\limits_{t\to b} h(t)$ exists.

Proof. Let $\alpha = \sup\limits_{t\in(a,b)} h(t)$. For some $\varepsilon > 0$, $\alpha - \varepsilon$ is not a supremum. Hence there exists t_0 such that $\alpha - \varepsilon < h(t_0) < \alpha$. For some $t \in (a, b)$ such that $t \geq t_0$ then

$$\alpha - \varepsilon < h(t_0) \leq h(t) < \alpha$$

Then $\alpha - h(t) < \varepsilon$. Also above can be written as $\alpha < h(t_0) + \varepsilon \leq h(t) + \varepsilon < \alpha + \varepsilon$. Then $h(t) + \varepsilon < \alpha + \varepsilon \implies h(t) < \alpha + \varepsilon \implies h(t) - \alpha < \varepsilon$. Therefore

$$|h(t) - \alpha| < \varepsilon \ \forall t \geq t_0 \implies \lim_{t\to b} h(t) = \alpha$$

□

Theorem 9 (Change of Variable Formula)**.** *Let $g\colon [c, d] \to \mathbb{R}$ be a C^1 function and let $[a, b]$ be any interval containing the image of g, that is $g[c, d] \subset g[a, b]$. If $f\colon [a, b] \to \mathbb{R}$ is a continuous function, then*

$$\int_c^d f(g(t))g'(t)dt = \int_{g(c)}^{g(d)} f(x)dx.$$

Theorem 10 (Generalized Leibnitz Formula)**.** *Let $\alpha, \beta\colon [a, b] \to \mathbb{R}$ be differentiable functions and c, d be real numbers satisfying*

$$c \leq \alpha(t), \beta(t) \leq d, \text{ for all } t \in [a, b].$$

Let $f\colon [a, b] \times [c, d] \to \mathbb{R}$ be a continuous function such that $\frac{\partial f}{\partial t}(t, s)$ is also continuous. Define

$$F(t) = \int_{\alpha(t)}^{\beta(t)} f(t, s)ds.$$

Then, F is differentiable and

$$\frac{dF}{dt} = \int_{\alpha(t)}^{\beta(t)} \frac{\partial f}{\partial t}(t,s)ds + f(t,\beta(t))\frac{d\beta}{dt} - f(t,\alpha(t))\frac{d\alpha}{dt}.$$

Definition 6 (Banach space)**.** A normed linear space which is a complete metric space (metric space by the norm) is called a Banach space.

Example 3 (Examples of Banach space)**.** $\mathbb{R}^n$ under supremum norm and p-norm. Function space $C[a,b]$ with supremum norm. However, it is *not* a complete space with repect to $\|\cdot\|_1$ norm.

Remark 10. A sequence $\{f_n\} \subset C[a,b]$, $f_n \to f$ in supremum norm is equivalent to saying that f_n converges uniformly to f ($\exists \delta > 0$ which works in supremum norm for f, will work for all f in the sequence.)

Theorem 11 (Banach Fixed Point theorem)**.** *Suppose (X,d) is a complete metric space and $T\colon X \to X$ is a contraction, that is, there exists an $\alpha \in (0,1)$ such that*

$$d(Tx,Ty) \leq \alpha d(x,y)$$

for all $x,y \in X$. Then, T has a unique fixed point $x^* \in X$. Further, the sequence $\{x_k\}$ defined by $x_k = Tx_{k-1}, x_0 \in X$ is arbitrary and $k = 1,2,\dots,$ converges to x^*.

Remark 11. If we omit the contraction condition then $f(x) = x+1$ show there does not exists any fixed point. If completion condition is droped then $f\colon (0,1) \to (0,1)$ defined by $f(x) = mx, 0 < m < 1$ is a contraction with $\alpha = m$ but no fixed point.

Proof. Choose any $x_0 \in X$ and define the sequence $x_1 = Tx_0, x_2 = T^2x_0, \cdots, x_k = T^kx_0, \cdots$. We need to show that sequence $\{x_k\}$ is a cauchy sequence. To see this

$$d(x_{n+1},x_n) = d(Tx_n,Tx_{n-1}) \leq \alpha d(x_n,x_{n-1})$$

By induction,

$$d(x_{n+1},x_n) \leq \alpha^n d(x_1,x_0)$$

Consider, for $m < n$ and using triangle inequality

$$d(x_n,x_m) \leq d(x_n,x_{n-1}) + \cdots + d(x_{m+1},x_m) \leq \alpha^m \frac{1-\alpha^{n-m-1}}{1-\alpha} d(x_1,x_0)$$

and the rhs $\to 0$ as $n,m \to \infty$. Hence the sequence is cauchy and by completeness, there exists $x^* \in X$ such that $x_k \to x^*$. By continuity of T, we get $Tx_k \to Tx^*$. By continuity of T, we get $Tx_k \to Tx^*$. But $Tx_k = x_{k+1} \to x^*$. Thus, $Tx^* = x^*$. If there exists y^* with same properties as x^* then by uniqueness of limits $x^* = y^*$. □

Remark 12. Here, fixed point can be constructed with any desired accuracy as limiting value is fixed point. And secondly, any point can be taken as an initial guess.

Corollary 1. *Let $T\colon X \to X$ where X is a complete metric space, be such that T^k is a contraction for some $k \geq 1$. Then, T has a unique fixed point.*

Proof. Since, T^k is contraction map hence it has a unique fixed point (by Banach fixed point theorem). Let x^* be that point i.e. $T^k x^* = x^*$. Applying T, we get $T^k(Tx^*) = Tx^*$ implies Tx^* is the fixed point for $T^k \implies Tx^* = x^*$. Therefore, x^* is the fixed point of T. For uniqueness, assume x_1 is another fixed point for $T \implies Tx_1 = x_1$. Applying T repeatedly, we get $T^k x_1 = x_1$ but fixed point for T^k is $x^*(unique) \implies x^* = x_1$. $\square$

2. Preliminaries from Linear Algebra

Definition 7 (Normed Linear Space)**.** A norm, denoted by $\|\cdot\|$ on a vector space or a linear space X is a mapping from $X \to \mathbb{R}$ that satisfies

- $\|x\| \geq 0$; $\|x\| = 0$ if and only if $x = 0$,
- $\|ax\| = |a|\|x\|$,
- (Triangle inequality) $\|x + y\| \leq \|x\| + \|y\|$,

for all $x, y \in X$ and scalar a.

Example 4. $\mathbb{R}^n, \mathbb{C}^n$

Remark 13. Every normed linear space is a metric space with metric induced by the norm. A metric space can be equipped with different norms which are fundamentally different. $C[0,1]$ with sup norm is complete while with the integral norm $\left(\int_0^1 |f(x)|\, dx\right)$ is not complete.

Definition 8 (Vector Field)**.** A vector field on a space(most commonly Euclidean space) is a function $\vec{F}$ that assigns a vector to each point of the space.

Definition 9 (Matrix norm)**.** $|A|$ should satisfy the following criterion

- $|A| \geq 0$; $|A| = 0$ if and only if $A = 0$,
- $|aA| = |a||A|$,
- (Triangle inequality) $|A + B| \leq |A| + |B|$,
- $|AB| \leq |A||B|$

for all $A, B \in M_n(\mathbb{R})$ and scalars a.

Remark 14. $M_n(\mathbb{R})$ is a complete metric space.

2.1. **Matrix Exponential e^A.** Let $A \in \mathbb{M}_n\mathbb{R}$, define the sequence of matrices

$$S_k = I + A + \frac{A^2}{2!} + \cdots + \frac{A^k}{k!}.$$

For $k > l$,

$$|S_k - S_l| \leq \sum_{j=l+1}^{k} \frac{|A|^j}{j!} \to 0 \text{ as } l, k \to \infty.$$

Thus $\{S_k\}$ is a Cauchy sequence and converges to some $S \in M_n(\mathbb{R})$.

Definition 10. Given $A \in M_n(\mathbb{R})$, e^A is defined as

$$e^A = S$$

where $S = \lim\limits_{k\to\infty} \sum\limits_{j=0}^{k} \frac{A^j}{j!}$.

Remark 15. $e^A \in M_n(\mathbb{R})$. Also, $|e^A| \leq e^{|A|}$ (substitute the values to see this.)

Remark 16. For diagonal matrix $A = diag(\lambda_1, \ldots, \lambda_n) \implies A^j = diag(\lambda_1^j, \ldots, \lambda_n^j)$ and $e^A = \sum\limits_{j=0}^{\infty} \frac{A^j}{j!}$ then

$$e^A = \sum_{j=0}^{\infty} \frac{1}{j!}\, diag(\lambda_1^j, \ldots, \lambda_n^j) \implies e^A = diag(\sum_{j=0}^{\infty} \frac{\lambda_1^j}{j!}, \ldots, \sum_{j=0}^{\infty} \frac{\lambda_n^j}{j!}),$$

therefore,

$$e^A = diag(e^{\lambda_1}, \ldots, e^{\lambda_n}).$$

Here are important observations:

(1) If $A \sim B \implies \exists$ non-singular matrix P such that $B = PAP^{-1}$. Then $e^A \sim e^B$ ($\because B^j = PA^jP^{-1} \implies e^B = P\left(\sum\limits_{j=0}^{\infty} \frac{A^j}{j!}\right)P^{-1} \implies e^B = Pe^AP^{-1}$).

(2) If A represents a block diagonal matrix $A = diag(A_1, \ldots, A_k)$ where all $A's$ are square matrix on the diagonal (their size maybe different) then

$$e^A = diag(e^{A_1}, e^{A_2}, \ldots, e^{A_k}).$$

Remark 17. (a) $|A| \leq \max\{|A_1|, \ldots, |A_k|\} \implies$
(b) $|e^A| \leq \max\{e^{|A_1|}, \ldots, e^{|A_k|}\}$. ((a)suspicious identity to me!!)
(c) Equality in above cases holds for Euclidean norm.

Computing e^A. If A is diagonalizable then we look for matrix P so that $PAP^{-1} = B$ is diagonal matrix hence computing e^A is easy.

If A is not diagonalizable then we look for P so that PAP^{-1} is a block diagonal, with easily computable e^{A_i}.

If T is a linear transformation and it is invariant on all coordinate axes then A is diagonalizable with respect to standard basis.

If usual coordinate axes are invariant under T then we look for n distinct directions, if possible, which are invariant under T and then take these directions as new bases. The matrix with respect to this basis will be diagonal matrix.

Remark 18. The set of all eigenvalues of A is known as *spectrum* of A. It is denoted by $\sigma(A)$.

The eigenvalues of A are the roots of the *characteristics polynomials* $\det(\lambda I - A)$ which is a real polynomial in λ of degree n. The roots maybe real or complex. If the eigenvalue is real then there exists a corresponding real eigenvector and if it complex then there exists a corresponding complex eigenvector.

3. Doubts

Ques 1. How to prove theorem 7? Is there a mean value theorem for vector valued functions that I can apply here?
Ques 2. Proof for change of varible formula in Theorem 9?
Ques 2. Proof for generalized Leibnitz formula in Theorem 10?

Part 2.
First and Second order linear equations

4. First Order equations

4.1. **General form of IVP and BVP.** General form of IVP of first order ODE is

$$\left.\begin{array}{l}\dot{y} = f(t, y)\\ y(t_0) = y_0\end{array}\right\}$$

General form of IVP of second order ODE is

$$\left.\begin{array}{l}\ddot{y} = f(t, y, \dot{y})\\ y(t_0) = y_0\\ \dot{y}(t_0) = y_1\end{array}\right\}$$

General form of BVP

$$\left.\begin{array}{l}\ddot{y} = f(t, y, \dot{y}) \text{ for } t \in (a, b)\\ \alpha_1 y(a) + \beta_1 \dot{y}(a) = \gamma_1\\ \alpha_2 y(b) + \beta_2 \dot{y}(b) = \gamma_2\end{array}\right\}$$

Definition 11 (Solution of ODE)**.** Let f be defined on the rectangle in $R := (a, b) \times (c, d)$ containing the initial data (t_0, y_0). A solution to the IVP is a function $y : (\bar{a}, \bar{b}) \to \mathbb{R}$ which is differentiable and satisfies the IVP together with initial condition.

Remark 19. Interval $(\bar{a}, \bar{b})$ is referred as interval of existence of solution. For all $t \in (\bar{a}, \bar{b}) \subset (a, b), y(t) \in (c, d)$. If $(\bar{a}, \bar{b}) = (a, b)$ then y is called the global solution otherwise it is a local solution.

For vector valued functions, definition of solution can be extended. Let $\mathbf{f} : (a, b) \times \Omega \to \mathbb{R}^n$ be a vector valued continuous function so that $\mathbf{f} = (f_1, \ldots, f_n)$ and each f_i is a real valued continuous (we assume this throughout the notes) function, where $\Omega \subset \mathbb{R}^n$ is open domain. For a given initial value $y_0 \in \Omega$, the IVP is given by

$$\left.\begin{array}{l}\dot{\mathbf{y}} = \mathbf{f}(t, \mathbf{y})\\ \mathbf{y}(t_0) = \mathbf{y}_0\end{array}\right\}$$

4.2. **First Order linear equations.** A general first order ODE can be written as

$$f(t, y, \dot{y}) = h(t)$$

where $h(t)$ is the function of t only. From linear algebra it can be shown that f takes the form (TODO)

$$f(t, y, \dot{y}) = p_0(t)\dot{y} + p_1(t)y$$

Thus the linear differential operator is given by $L = p_0(t)\frac{d}{dt} + p_1(t)$. General linear and homogeneous ODE is given by

$$Ly = 0$$

and non-homogeneous ODE is given by

$$Ly = q(t)$$

where p_0, p_1 and q are the given functions of t.

Remark 20. Due to the linear structure of the operator, it follows the superposition principle i.e. if y_1 and y_2 are the two solutions of the homogeneous equation then $\alpha y_1 + \beta y_2$ is also a solution of the homogeneous equation.

Definition 12 (Singular equations)**.** The equations in which coefficient of the highest order term vanished at one or more points are called singular equations. E.g. Bessel's eqn, Lagrange eqn, Legendre eqn.

Definition 13 (Regular equations)**.** The coefficient of highest order term is never 0. Their general form is given by $Ly := \dot{y} + p(t)y = q(t)$.

Remark 21. A continuous function defined on a interval in $\mathbb{R}$ whose modulus is a constant, then f itself is constant.

For non-homogeneous equation

$$\dot{y} + p(t)y = q(t)$$

we can find a function $h(t)$ such that $\dot{h}(t) = \dot{y} + p(t)y$ then we $\dot{h}(t) = q(t)$ which can be integrated easily to find the solution. Let a differentiable function $\mu(t)$ such that $d(\mu y)/dt = \mu\dot{y} + \dot{\mu}y$. Multiplying both sides by $\mu(t)$ in main non-homogeneous equation and then comparing with the derivative of $\mu(t)$, we get $\dot{\mu}(t) = \mu(t)p(t)$. Here $\mu(t)$ is called *integrating*

$$\mu(t)\dot{y}(t) + \mu(t)p(t)y(t) = \mu(t)q(t). \tag{3.1.16}$$

If μ is positive, then any solution of (3.1.15) is a solution of (3.1.16) and vice versa. The term on the left hand side of (3.1.16) can be written as $\frac{d}{dt}(\mu y)$, provided μ satisfies $\dot{\mu}(t) - p(t)\mu(t) = 0$. Thus, (3.1.16) is exact. Note that the equation satisfied by μ is a homogeneous linear DE in μ and hence, $\mu(t) = \exp\left(\int^t p(\tau)d\tau\right)$ is a solution and it is positive. Thus, (3.1.16) becomes

FIGURE 1. why μ has to be positive? Check this.

factor associated with homogeneous part.

4.3. **Exact Differential equations.**

Definition 14 (Exact Differential equations)**.** If the differential equation $\dot{y} = f(t, y)$ can be written as $\frac{d}{dt}\phi(t, y(t)) = 0$ for a two variable function ϕ in a domain in the $t-y$ plane, then the differential equation is said to be an exact differential equation.

Necessary condition for DE to be exact. Consider the differential form $Mdt + Ndy$, this can be written as general first order equation as

$$M(t, y) + N(t, y)\dot{y} = 0$$

Above DE is exact *if and only if* there exists ϕ such that $\frac{d}{dt}\phi(t, y) = 0$. Therefore,

$$\frac{\partial}{\partial t}\phi(t, y) + \frac{\partial}{\partial y}\phi(t, y)\dot{y} = M(t, y) + N(t, y)\dot{y}$$

and this implies

$$M = \frac{\partial \phi}{\partial t} \text{ and } N = \frac{\partial \phi}{\partial y}.$$

Assuming ϕ is double differentiable, we get

$$\frac{\partial M}{dy} = \frac{\partial N}{dt}$$

Theorem 12. *Assume M, N are defined on a rectangle $D = (a, b) \times (c, d)$ and $M, N \in C^1(D)$. Then, there exists a function ϕ defined in D, such that $M = \frac{\partial \phi}{\partial t}$ and $N = \frac{\partial \phi}{\partial y}$ if and only if $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial t}$.*

Proof. ($\implies$) Assume there exists a function ϕ following the needed conditions given in theorem. The first relation $M = \frac{\partial \phi}{\partial t}$ suggests that $\phi(t, y) = \int Mdt + h(y)$ for some h. Hence,

$$N = \frac{\partial \phi}{\partial t} = \int \frac{\partial M}{\partial y}dt + \frac{\partial h(y)}{\partial y} \implies \frac{\partial h}{\partial y} = N - \int \frac{\partial M}{\partial y}dt$$

Since, LHS is a function of y alone hence RHS should be a function of y alone. Therefore,

$$\frac{\partial}{\partial t}\left(N - \int \frac{\partial M}{\partial y}dt\right) = 0 \implies \frac{\partial N}{\partial t} - \frac{\partial M}{\partial y} = 0$$

($\Longleftarrow$) Assume M and N satisfies $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial t}$. Let a point $(t_0, y_0) \in D$. Define $\phi(t,y) = \int_{t_0}^{t} M(s,y)ds + h(y)$ where h has to be determined.

$$\begin{aligned}\frac{\partial \phi(t,y)}{\partial y} &= \int_{t_0}^{t} \frac{\partial M(s,y)}{\partial y} ds + \frac{\partial h(y)}{\partial y} \\ &= \int_{t_0}^{t} \frac{\partial N(s,y)}{\partial t} ds + \frac{\partial h(y)}{\partial y} \\ &= N(t,y) - N(t_0,y) + \frac{\partial h(y)}{\partial y}\end{aligned}$$

If we want $\frac{\partial \phi}{\partial y}$ to be equal to $N(t,y)$ then $\frac{\partial h(y)}{\partial y} = N(t_0,y) \implies h(y) = \int_{y_0}^{y} N(t_0,\xi)d\xi$. Therefore,

$$\phi(t,y) = \int_{t_0}^{t} M(s,y)ds + \int_{y_0}^{y} N(t_0,\xi)d\xi$$

ϕ is determined upto a constant. Hence on changing the point (t_0, y_0) the constant term in ϕ will change. We may discard all the constant in expression of ϕ. $\square$

Definition 15 (Alternate definition of Exact ODE)**.** The DE, $M(t,y) + N(t,y)\dot{y} = 0$ is said to be exact if $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial t}$.

Solution of exact DEs. If a DE is exact then it can be written as $\frac{\partial \phi}{\partial t}(t,y) = 0$ and the solution of this DE can be written as $\phi = constant$.

5. Second order linear equations

A second order linear differential equation is given by

$$Ly \equiv \ddot{y} + p(t)\dot{y} + q(t)y = r(t).$$

The operator $Ly = L(t, y, \dot{y}, \ddot{y})$ is a multi-linear operator (linear in each of its variable). The IVP for second order linear ODE is given by

$$\left.\begin{array}{l}\ddot{y} + p(t)\dot{y} + q(t)y = r(t), t \in I \\ y(t_0) = y_0, \dot{y}(t_0) = y_1\end{array}\right\}$$

Theorem 13. *Let p, q, r be continuous functions defined in a compact interval $I(t_0)$ and y_0, y_1 be any real numbers. Then, the IVP above has a unique solution y defined in $I(t_0)$ satisfying $y(t_0) = y_0$ and $\dot{y}(t_0) = y_1$*

Lemma 1. *Wronskian $W(z,w) \equiv 0$ if and only if z and w are dependent. Alternatively, $W \not\equiv 0$ if and only if z and w are independent.*

Proof. Let $t_0 \in I$ be a point in interval I on which z and w are defined. Then the Wronskian is given by

$$W = \begin{vmatrix} z(t_0) & w(t_0) \\ \dot{z}(t_0) & \dot{w}(t_0) \end{vmatrix}$$

This comes from the linear system (for y_0 and y_1 check next theorem)

$$\begin{cases} z(t_0)\alpha + w(t_0)\beta = y_0 \\ \dot{z}(t_0)\alpha + \dot{w}(t_0)\beta = y_1 \end{cases}$$

$(\Longleftarrow)$ If $z = kw$ or $w = kz$ then the $W \equiv 0$.
$(\Longrightarrow)$ Assume that $W \equiv 0 \implies z\dot{w} - \dot{z}w = 0$

$$z\frac{dw}{dt} = \frac{dz}{dt}w$$
$$\int \frac{dw}{wdt} \times dt = \int \frac{dz}{zdt} \times dt$$
$$\ln w = \ln z + c$$
$$w = e^{\ln z} e^c$$
$$w = ze^c$$

Therefore, w and z are dependent. □

Theorem 14. *Let z, w be two solutions of the linear homogeneous second order DE. Then, for any $\alpha, \beta \in \mathbb{R}$, the function $y = \alpha z + \beta w$ is also a solution.*
Further, if z and w are two linearly independent solution of the homogeneous DE in an interval $I(t_0)$, then every solution can be written as a linear combination of z and w.

Proof. First part of the theorem is trivial to verify. Let y be a solution of the homogeneous DE. For some $t_0 \in I$, $y(t_0) = y_0$ and $\dot{y}(t_0) = y_1$. We want to show that there exists α and β such that $y(t) = \alpha z(t) + \beta w(t)$. Since z and w are independent, $W \not\equiv 0$ then there exists a unique solution to the linear system

$$\begin{cases} z(t_0)\alpha + w(t_0)\beta = y_0 \\ \dot{z}(t_0)\alpha + \dot{w}(t_0)\beta = y_1 \end{cases}$$

Hence, there exists α and β. Here, z and w are basis elements for the solution space of second order linear homogeneous DE. Therefore, all the solution for this equation can be written in this form. □

Theorem 15. $\dim(S) = 2$ *where S is the solution space of second order linear homogeneous DE.*

Proof. TODO □

Remark 22. For n-th order linear homogeneous DE, the solution space is n-dimensional.