

Analysis in Many Variables Course Notes

Isaac Holt

March 17, 2023

Contents

1	Summary notes	3
1.1	The implicit function theorem	6
1.2	The implicit function theorem for surfaces	6
2	Differentiability of Vector Fields	7
2.1	Differentiable maps from \mathbb{R}^n to \mathbb{R}^n	7
2.2	Diffeomorphisms and the inverse function theorem	7
3	Volume, Line and Surface Integrals	8
3.1	Fubini's theorem	8
3.2	Line integrals	9
3.3	Surface integrals I: defining a surface	9
3.4	Surface integrals II: evaluating the integral	10
4	Green's, Stokes' and divergence theorems	12
4.1	The Big 3 theorems	12
4.2	Path independence of line integrals	13
5	Non-cartesian systems	15
5.1	Change of variables in surface and volume integrals	15
5.2	Differential operators in polar and spherical polar coordinates	15
5.3	Polar coordinates	15
6	Generalised functions	17
6.1	Birth of generalised functions	17
6.2	Test functions and distributions	17
6.3	Regular and singular distributions	19
6.4	Operations on distributions	20
6.5	The derivative of a distribution	22
6.6	Leibniz rule for differentiation	23
6.7	The delta distribution in polar/spherical coordinates	25
7	Sturm-Liouville Theory	26
7.1	Hilbert spaces: definitions and examples	26
7.2	Inner products defined by integrals	27
7.3	Bounded and unbounded linear operators	27
7.4	Sturm-Liouville operators	29
7.5	Eigenfunction method for differential operators	31
7.6	Spectral properties of regular Sturm-Liouville operators	36
8	Green's functions	39
8.1	The importance of Green's functions	39
8.2	Green's functions for ODEs	40

1 Summary notes

- **Scalar field:** maps $\mathbb{R}^n \rightarrow \mathbb{R}$.
- **Vector field:** maps $\mathbb{R}^n \rightarrow \mathbb{R}^n$.
- **Curve:** maps $\mathbb{R} \rightarrow \mathbb{R}^n$.
- A **tangent** to a curve $\underline{x}(t)$ is given by $\frac{d\underline{x}}{dt}$.
- The **arc-length** parameterisation of a curve \underline{x} is such that

$$\left| \frac{d\underline{x}(s)}{ds} \right| = 1 \quad \forall s$$

- **Partial derivatives:**

$$\frac{\partial f(\underline{x})}{\partial x_a} = \lim_{h \rightarrow 0} \frac{f(\underline{x} + h \underline{e}_a) - f(\underline{x})}{h}$$

- **Chain rule:** for a scalar field $f(\underline{x})$ and curve $\underline{x}(t) = x_1(t)\underline{e}_1 + \cdots + x_n(t)\underline{e}_n$,

$$\frac{df(\underline{x}(t))}{dt} = \sum_{i=1}^n \frac{\partial f(\underline{x})}{\partial x_i} \frac{dx_i}{dt}$$

Here $F(t) := f(\underline{x}(t))$ is the restriction of $f(\underline{x})$ to the curve $\underline{x}(t)$.

- **Differential operator:** maps functions to functions, e.g.

$$\frac{d}{dt} = \sum_{i=1}^n \frac{\partial}{\partial x_i} \frac{dx_i}{dt}$$

- Let $f, g : \mathbb{R} \rightarrow \mathbb{R}$, then

- $f(x) \frac{d}{dx}$ is a differential operator. It acts on $g(x)$ to give $f(x) \frac{dg(x)}{dx}$.
- $\frac{d}{dx} f(x)$ is a differential operator. It acts on $g(x)$ to give $\frac{d}{dx}(f(x)g(x))$.
- $(\frac{d}{dx} f(x))$ is an differential operator. It acts on $g(x)$ to give $\frac{df(x)}{dx} g(x)$.

- **del (or nabla):**

$$\underline{\nabla} = \sum_{i=1}^n \frac{\partial}{\partial x_i} \underline{e}_i$$

so $\frac{d}{dt} = \underline{\nabla} \cdot \frac{d\underline{x}(t)}{dt}$.

- **gradient** of a scalar field $f : \mathbb{R}^n \rightarrow \mathbb{R}$:

$$\underline{\nabla} f \equiv \text{grad}(f) = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \underline{e}_i$$

- **Directional derivative** of $f : \mathbb{R}^n \rightarrow \mathbb{R}$ in direction of a unit tangent $\hat{n} = \frac{d\underline{x}(s)}{ds}$ to a curve $\underline{x} : \mathbb{R} \rightarrow \mathbb{R}^n$:

$$\frac{df(\underline{x}(s))}{ds} = \hat{n} \cdot \underline{\nabla} f \equiv \frac{df}{d\hat{n}}$$

where \underline{x} is parameterised in terms of arc-length s .

- $\underline{\nabla}f$ at a point \underline{p} is orthogonal to curves contained in level set of f at \underline{p} .
- $\underline{\nabla}f$ points in the direction where f increases fastest.
- **Properties of the gradient:** let $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$, $a, b \in \mathbb{R}$, $\phi : \mathbb{R} \rightarrow \mathbb{R}$, then

$$\begin{aligned} - \underline{\nabla}(af + bg) &= a\underline{\nabla}f + b\underline{\nabla}g \\ - \underline{\nabla}(fg) &= f\underline{\nabla}g + g\underline{\nabla}f \\ - \underline{\nabla}\phi(f) &= (\underline{\nabla}f) \frac{d\phi}{df} \end{aligned}$$

- **Divergence** of a vector field $\underline{v}(\underline{x}) = \sum_{i=1}^n v_i(\underline{x})\underline{e}_i$:

$$\underline{\nabla} \cdot \underline{v} \equiv \text{div}(\underline{v}) = \sum_{i=1}^n \frac{\partial v_i}{\partial x_i}$$

Note that the formula will be different in other coordinates systems. Also $\underline{\nabla} \cdot \underline{v} \neq \underline{\nabla} \cdot \underline{v}$.

- Considering a vector field as a fluid, if the divergence at a point is positive the vector field acts as a **source** at that point (more fluid leaving than entering), if the divergence is negative the vector field acts as a **sink** at that point (more fluid entering than leaving). The magnitude of the divergence is the rate of flow and the direction of the divergence is the direction of flow.

- **Properties of div:** for $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $\underline{v}, \underline{w} : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $a, b \in \mathbb{R}$,

$$\begin{aligned} - \underline{\nabla} \cdot (a\underline{v} + b\underline{w}) &= a\underline{\nabla} \cdot \underline{v} + b\underline{\nabla} \cdot \underline{w} \\ - \underline{\nabla} \cdot (f\underline{v}) &= (\underline{\nabla}f) \cdot \underline{v} + f\underline{\nabla} \cdot \underline{v} \end{aligned}$$

- **Curl** of $\underline{v} : \mathbb{R}^n \rightarrow \mathbb{R}^n$:

$$\underline{\nabla} \times \underline{v} \equiv \text{curl}(\underline{v}) = \begin{vmatrix} \underline{e}_1 & \underline{e}_2 & \underline{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ v_1 & v_2 & v_3 \end{vmatrix} = \underline{e}_1 \left(\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z} \right) - \underline{e}_2 \left(\frac{\partial v_3}{\partial x} - \frac{\partial v_1}{\partial z} \right) + \underline{e}_3 \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} \right)$$

- Considering a vector field as a fluid, the magnitude of the curl at a point corresponds to the rotational speed of the fluid, and the direction of the curl corresponds to which axis the fluid is rotating around, determined using the right-hand rule (fingers represent rotation of the fluid, thumb points in direction of curl).

- **Properties of curl:** for $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, $\underline{v}, \underline{w} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $a, b \in \mathbb{R}$,

$$\begin{aligned} - \underline{\nabla} \times (a\underline{v} + b\underline{w}) &= a\underline{\nabla} \times \underline{v} + b\underline{\nabla} \times \underline{w} \\ - \underline{\nabla} \times (f\underline{v}) &= (\underline{\nabla}f) \times \underline{v} + f\underline{\nabla} \times \underline{v} \end{aligned}$$

- **Laplacian** of $f : \mathbb{R}^n \rightarrow \mathbb{R}$:

$$\Delta f \equiv \underline{\nabla}^2 f := \underline{\nabla} \cdot (\underline{\nabla}f) = \text{div}(\text{grad}(f)) = \sum_{i=1}^n \frac{\partial^2 f}{\partial (x_i)^2}$$

Note this formula is only valid for **cartesian** coordinates.

- **Einstein summation convention:** in an expression involving a summation, then index of summation always appears twice. The convention is that the summation sign is removed, and whenever an index appears twice, it is summed over.

- **Dummy indices:** repeated indices. They can be renamed without changing the expression.
- **Free indices:** non-repeated indices. They must match on both sides of an equation.
- An index can't be repeated more than twice, so $(\underline{u}.\underline{v})^2 = u_i v_i u_j v_j \neq u_i v_i u_i v_i$.

- **Kronecker delta:**

$$\delta_{ij} := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} = \frac{\partial x_i}{\partial x_j}$$

- If δ_{ij} has a dummy index i , then remove the δ_{ij} and replace the dummy index i by j in the rest of the expression.

- **Levi-Cevita symbol:**

$$\begin{aligned} \epsilon_{ijk} &= -\epsilon_{jik} = -\epsilon_{ikj} \quad (\text{antisymmetry}) \\ \epsilon_{123} &= 1 \end{aligned}$$

- **Properties of ϵ_{ijk} :**

- $\epsilon_{ijk} = -\epsilon_{kji}$
- $\epsilon_{ijk} = 0$ if $i = j$ or $j = k$ or $k = i$
- If ϵ_{ijk} is zero then $(i \ j \ k)$ is a permutation of $(1 \ 2 \ 3)$.
- $\epsilon_{ijk} = 1$ if $(i \ j \ k)$ is an even permutation of $(1 \ 2 \ 3)$ (even number of swaps).
- $\epsilon_{ijk} = -1$ if $(i \ j \ k)$ is an odd permutation of $(1 \ 2 \ 3)$ (odd number of swaps).
- $\epsilon_{ijk} = \epsilon_{jki} = \epsilon_{kij}$ (cyclic permutation).

- The cross product $\underline{C} = \underline{A} \times \underline{B}$ can be written as $C_i = \epsilon_{ijk} A_j B_k$.

- **Very useful ϵ_{ijk} formula:**

$$\sum_{k=1}^3 \epsilon_{ijk} \epsilon_{klm} = \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}$$

- Notation: $\partial_i := \frac{\partial}{\partial x_i}$.
- $\underline{\nabla}.\underline{v} = \frac{\partial v_i}{\partial x_i} = \partial_i v_i$.
- $(\underline{\nabla} \times \underline{v})_i = \epsilon_{ijk} \frac{\partial}{\partial x_j} v_k = \epsilon_{ijk} \partial_j v_k$.

Theorem 1.0.1. If $f(x) : U \rightarrow \mathbb{R}$ is differentiable with U an open subset of \mathbb{R}^n and if x is a function of u_1, \dots, u_m , such that the partial derivatives $\frac{\partial x_i}{\partial u_j}$ exist for all $1 \leq i \leq n$ and all $1 \leq j \leq m$, and if $F(u_1, \dots, u_m) = f(x(u_1, \dots, u_m))$ then $\frac{\partial F}{\partial u_b} = \frac{\partial x_i}{\partial u_b} \frac{\partial f}{\partial x_i}$

Proof. We have $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $x : \mathbb{R}^m \rightarrow \mathbb{R}^n$, $F : \mathbb{R}^m \rightarrow \mathbb{R}$ with $F = f(x(u))$.

let $a = x(u_1, \dots, u_b, \dots, u_m)$, $a + h(k) = x(u_1, \dots, u_b + k, \dots, u_m)$. then

$$\frac{\partial F}{\partial u_b} = \lim_{k \rightarrow 0} \frac{F(u_1, \dots, u_b + k, \dots, u_m) - F(u_1, \dots, u_b, \dots, u_m)}{k} = \lim_{k \rightarrow 0} \frac{f(x(u_1, \dots, u_b + k, \dots, u_m)) - f(x(u_1, \dots, u_b, \dots, u_m))}{k}$$

□

1.1 The implicit function theorem

$y = g(x)$ gives y as an explicit function of x . $f(x, y) = 0$ gives y as an implicit function of x .

To go from an explicit function to an implicit function, set

$$f(x, y) = y - g(x) = 0$$

Suppose the level curve $f(x, y) = c$ can be written as $y = g(x)$. Then $f(x, g(x)) = c$. Differentiating this using the chain rule:

$$\frac{d}{dx}f(x, g(x)) = \frac{\partial f}{\partial x} + \frac{dg}{dx} \frac{\partial f}{\partial y} = \frac{d}{dx}c = 0$$

Hence $\frac{dg}{dx} = \frac{-\partial f / \partial x}{\partial f / \partial y}$

Theorem 1.1.1. (Implicit Function Theorem or IFT): If $f(x, y) : U \rightarrow \mathbb{R}$ with $U \subseteq \mathbb{R}^2$ open is differentiable on the level curve $f(x, y) = c$, at which $\frac{\partial f}{\partial y} \neq 0$, then a differentiable function $g(x)$ exists in a neighbourhood of x_0 , satisfying $g(x_0) = y_0$.

Remark. At points when $\frac{\partial f}{\partial y} = 0$, if $\frac{\partial f}{\partial x} \neq 0$, we can use the IFT to find $h(y)$ such that $x = h(y)$.

Remark. If there is a point Q on a level curve $f(x, y) = c$ at which $\nabla f = 0$ (this is a critical point), then the value of c is called a critical value (otherwise it is a regular value) and the level curve cannot be written either as $y = g(x)$ or as $x = h(y)$ (g and h are differentiable) in a neighbourhood of Q .

1.2 The implicit function theorem for surfaces

The level sets of scalar fields on \mathbb{R}^3 generally define surfaces, and we therefore have the IFT for surfaces.

Theorem 1.2.1. Let $f(x, y, z) : U \rightarrow \mathbb{R}$ for $U \subseteq \mathbb{R}^3$ open be differentiable. Let $(x_0, y_0, z_0) \in U$ be a point of the level set $f(x, y, z) = c$ so $f(x_0, y_0, z_0) = c$.

If $\frac{\partial f}{\partial z}(x_0, y_0, z_0) \neq 0$ then the equation $f(x, y, z) = c$ implicitly defines a surface $z = g(x, y)$ in a neighbourhood of (x_0, y_0, z_0) if the following hold:

1. $f(x, y, g(x, y)) = c$ with $g(x_0, y_0) = z_0$

2. $\frac{\partial g}{\partial x} = \frac{-\frac{\partial f}{\partial x}(x_0, y_0, z_0)}{\frac{\partial f}{\partial z}(x_0, y_0, z_0)}$

3. $\frac{\partial g}{\partial y} = \frac{-\frac{\partial f}{\partial y}(x_0, y_0, z_0)}{\frac{\partial f}{\partial z}(x_0, y_0, z_0)}$

As in the IFT for curves, 2. and 3. must hold for $g(x, y)$ if it exists, since $f(x, y, g(x, y)) = c$ for all (x, y) in some neighbourhood of (x_0, y_0) . If we partially differentiate with respect to x and use the chain rule:

$$\begin{aligned} 0 &= \frac{\partial f}{\partial x}(x_0, y_0, z_0) = \frac{\partial x}{\partial x} \frac{\partial f}{\partial x} + \frac{\partial y}{\partial x} \frac{\partial f}{\partial y} + \frac{\partial g}{\partial x} \frac{\partial f}{\partial z} \\ &= \frac{\partial f}{\partial x} + \frac{\partial g}{\partial x} \frac{\partial f}{\partial z} \end{aligned}$$

Recall: ∇f at (x_0, y_0, z_0) is normal to the tangent plane of the surface $z = g(x, y)$ at (x_0, y_0) , so the normal line is given in parametric form as $\underline{x}(t) = \underline{x}_0 + t\nabla f$ and the tangent plane is given by $(\underline{x} - \underline{x}_0) \cdot \nabla f = 0$.

2 Differentiability of Vector Fields

2.1 Differentiable maps from \mathbb{R}^n to \mathbb{R}^n

Definition 2.1.1. For a vector field $\underline{F}(\underline{x}) : U \rightarrow \mathbb{R}^n$, with $U \subseteq \mathbb{R}^n$ open, F is differentiable at $\underline{a} \in U$ if there is a linear function $\underline{L} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that $\underline{F}(\underline{a} + \underline{h}) - \underline{F}(\underline{a}) = \underline{L}(\underline{h}) + \underline{R}(\underline{h})$ with $\lim_{\underline{h} \rightarrow 0} \frac{\underline{R}(\underline{h})}{|\underline{h}|} = 0$.

Now linear maps $\mathbb{R}^n \rightarrow \mathbb{R}^n$ are given by matrices. To see what matrix, use standard basis $\underline{F}(\underline{x}) = F_1(\underline{x})\underline{e}_1 + \cdots + F_n(\underline{x})\underline{e}_n$, $\underline{L}(\underline{h}) = L_1(\underline{h})\underline{e}_1 + \cdots + L_n(\underline{h})\underline{e}_n$, $\underline{R}(\underline{h}) = R_1(\underline{h})\underline{e}_1 + \cdots + R_n(\underline{h})\underline{e}_n$

So the j th component of A and B is $F_j(\underline{a} + \underline{h}) - F_j(\underline{a}) = L_j(\underline{h}) + R_j(\underline{h})$ with $\lim_{\underline{h} \rightarrow 0} \frac{R_j(\underline{h})}{|\underline{h}|} = 0$

So $L_j(\underline{h}) = \underline{h} \cdot \nabla F_j(\underline{a}) = h_1 \frac{\partial F_j}{\partial x_1} + \cdots + h_n \frac{\partial F_j}{\partial x_n}$

So L as a column vector is the matrix product $J\underline{h}$ where $J_{i,j} = \frac{\partial F_i}{\partial x_j}$. J is the Jacobian matrix or the differential of $\underline{F}(\underline{x})$ at $\underline{x} = \underline{a}$. It is written as $D\underline{F}(\underline{a})$.

Definition 2.1.2. The determinant of the differential $\det(D\underline{V}) = |D\underline{V}|$ is called the Jacobian of \underline{V} , $J(\underline{V})$

2.2 Diffeomorphisms and the inverse function theorem

We can think of a vector field $\underline{V}(\underline{x}) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ as a coordinate transformation on \mathbb{R}^n .

If we think of the components of \underline{h} as coordinates of $\underline{x} = \underline{a} + \underline{h}$ relative to an origin \underline{a} , the components of $\underline{V}(\underline{a} + \underline{h}) - \underline{V}(\underline{a})$ are the transformed coordinates of \underline{x} relative to the transformed origin $\underline{V}(\underline{a})$.

Theorem 2.2.1. (Inverse function theorem) Let $\underline{v} : U \rightarrow \mathbb{R}^n$, $U \subseteq \mathbb{R}^n$ open, be differentiable with continuous partial derivatives, and let $\underline{a} \in U$. Then if $J(\underline{v}(\underline{a})) \neq 0$, for some open set $u = \tilde{U} \subseteq U$ containing \underline{a} :

1. $\underline{v}(\tilde{U})$ is open.
2. the mapping \underline{v} from \tilde{U} to $\underline{v}(\tilde{U})$ has a differentiable inverse, i.e. there exists a differentiable $\underline{w} : \underline{v}(\tilde{U}) \rightarrow \mathbb{R}^n$ such that $\underline{w}(\underline{v}(\underline{x})) = \underline{x}$ and $\underline{v}(\underline{w}(\underline{y})) = \underline{y}$.

Definition 2.2.2. A mapping $\underline{v} : \tilde{U} \rightarrow V \subset \mathbb{R}^n$ satisfying 1. and 2. is called a diffeomorphism of \tilde{u} onto $\tilde{v} = \underline{v}(\tilde{U})$. We say \tilde{U} and \tilde{V} are diffeomorphic.

More generally, a mapping $\underline{v} : U \rightarrow V$ is a **local diffeomorphism** if for every point $\underline{a} \in U$, there is an open set $\tilde{U} \subseteq U$ containing \underline{a} such that $\underline{v} : \tilde{U} \rightarrow \underline{v}(\tilde{U})$ is a diffeomorphism.

Proof. Let $\underline{v} : U \rightarrow V \subseteq \mathbb{R}^n$, $\underline{w} : V \rightarrow W \subseteq \mathbb{R}^n$, u, v, w open in \mathbb{R}^n and $\underline{v}, \underline{w}$ differentiable.

Then $\underline{w}(\underline{v}(\underline{x}))$ is a map $U \rightarrow W \subseteq \mathbb{R}^n$ and its differential can be calculated using the chain rule: $D\underline{w}(\underline{v}(\underline{x})) = D\underline{w}(\underline{v})D\underline{v}(\underline{x})$.

In the particular case when \underline{v} is a local diffeomorphism and \underline{w} is its inverse $\underline{w}(\underline{v}(\underline{x})) = \underline{x}$, $D\underline{w}D\underline{v} = D\underline{w}(\underline{v}(\underline{x})) = D\underline{x}(\underline{x}) = I_n$.

Similarly, $\underline{v}(\underline{w}(\underline{y})) = \underline{y}$ so $D\underline{v}D\underline{w} = D\underline{v}(\underline{w}(\underline{y})) = I_n$.

So $D\underline{v}$ is invertible with inverse $(D\underline{v})^{-1} = D\underline{w}$ and by taking determinants, $J(\underline{w}) = 1/J(\underline{v})$ and $J(\underline{v}) \neq 0$. \square

Definition 2.2.3. Such a \underline{v} is **orientation-preserving** if $J(\underline{v}) > 0$ and **orientation-reversing** if $J(\underline{v}) < 0$.

3 Volume, Line and Surface Integrals

3.1 Fubini's theorem

Given a scalar field on \mathbb{R}^2 , $f(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}$ which is continuous on $R \subset \mathbb{R}^2$, then the double integral $\int_R f(x, y) dA$ (or equivalently $\iint_R f(x, y) dA$) is defined by partitioning R into smaller areas ΔA_k and then defining the integral as the limit of the Riemann sum (which should be independent of the partition):

$$\int_R f(x, y) dA = \lim_{N \rightarrow \infty} \sum_{k=1}^N f(x_k^*, y_k^*) \Delta A_k$$

where (x_k^*, y_k^*) lies in the base of the k th region.

If we choose the small areas ΔA_k to be rectangles on a rectangular grid, then $\Delta A_k = \Delta x_i \Delta y_j$ where $\Delta x_i = x_{i+1} - x_i$ and $\Delta y_j = y_{j+1} - y_j$ and x and y are partitioned as $x_0 < x_1 < \dots < x_n$ and $y_0 < y_1 < \dots < y_m$.

We then get

$$\int_R f(x, y) dA = \lim_{n \rightarrow \infty, m \rightarrow \infty} \sum_{i=0}^n \sum_{j=0}^m f(x_i^*, y_j^*) \Delta x_i \Delta y_j$$

where $x_i^* \in [x_i, x_{i+1}]$ and $y_j^* \in [y_j, y_{j+1}]$.

If we assume we can take the limit as $m \rightarrow \infty$ first and the limit as $n \rightarrow \infty$ afterwards, we get

$$\begin{aligned} \lim_{n \rightarrow \infty, m \rightarrow \infty} \sum_{i=0}^n \sum_{j=0}^m f(x_i^*, y_j^*) \Delta x_i \Delta y_j &= \lim_{n \rightarrow \infty} \sum_{i=0}^n \left(\lim_{m \rightarrow \infty} \sum_{j=0}^m f(x_i^*, y_j^*) \Delta y_j \right) \Delta x_i \\ &= \lim_{n \rightarrow \infty} \sum_{i=0}^n \left(\int_{y_0}^{y_1} f(x_i^*, y) dy \right) \Delta x_i = \int_{x_0}^{x_1} \int_{y_0}^{y_1} f(x, y) dy dx \end{aligned}$$

If we take $n \rightarrow \infty$ first, we exchange the order of integration:

$$\int_R f(x, y) dA = \int_{y_0}^{y_1} \int_{x_0}^{x_1} f(x, y) dx dy$$

Theorem 3.1.1. (Fubini's theorem): If $f(x, y)$ is continuous on a closed and bounded region of \mathbb{R}^2 (a region of integration R), then the double integral $\int_R f(x, y) dA$ can be written as an iterated integral, with the integrals in either order:

$$\int_R f(x, y) dA = \int_y \int_x f(x, y) dx dy = \int_x \int_y f(x, y) dy dx$$

Remark. If the region and/or function is unbounded, Fubini's theorem still holds if the double integral is absolutely convergent, i.e. if the integral of the $|f(x, y)|$ is finite. If this is not the case, Fubini's theorem doesn't necessarily hold.

Remark. If the region R is not rectangular, it is more complicated.

e.g. if $R = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, y_0(x) \leq y \leq y_1(x)\}$, then

$$\int_R f(x, y) dA = \int_a^b \int_{y_0(x)}^{y_1(x)} f(x, y) dy dx$$

If $f(x, y)$ is continuous then by Fubini's theorem we can change the order of integration. To calculate the integral over x first in this case, we need to split R into sub-regions where I can write the x -limits as functions of y .

3.2 Line integrals

Definition 3.2.1. A **regular arc** $C \subset \mathbb{R}^n$ is a parameterised curve $\underline{x}(t)$ whose cartesian components $x_a(t)$, $a \in \{1, \dots, n\}$ are continuous with continuous first derivative, where t lies in some (possibly infinite) interval.

Definition 3.2.2. A **regular curve** consists of a finite number of regular arcs joined end-to-end.

Given $\underline{v}(\underline{x}) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ its restriction to a regular arc $\underline{v}(\underline{x}(t))$ is a vector function of t and its scalar product with the tangent vector $\frac{d\underline{x}(t)}{dt}$ is a scalar function of t .

We can therefore integrate it along the arc to get a real number, this is called the **line integral** of \underline{v} along the arc $C : t \rightarrow \underline{x}(t)$ between $t = \alpha$ and $t = \beta$.

$$\int_C \underline{v} d\underline{x} = \int_{\alpha}^{\beta} \underline{v}(\underline{x}(t)) \cdot \frac{d\underline{x}}{dt} dt$$

This is independent of the choice of parameterisation. This can be proven using the chain rule:

let $t = t(u)$, then $dt = \frac{dt}{du} du$.

$$\int_{t^{-1}(\alpha)}^{t^{-1}(\beta)} \underline{v}(\underline{x}(t(u))) \cdot \frac{d\underline{x}(t(u))}{du} du = \int_{t^{-1}(\alpha)}^{t^{-1}(\beta)} \underline{v}(\underline{x}(t(u))) \cdot \frac{d\underline{x}}{dt} \frac{dt}{du} du = \int_{\alpha}^{\beta} \underline{v}(\underline{x}(t)) \cdot \frac{d\underline{x}}{dt} dt$$

If C is a regular curve, made up of regular arcs, the line integral $\int_C \underline{v} d\underline{x}$ is the sum of the line integral over the arcs.

If the integral is calculated over a regular closed curve (the endpoints join), then it is often written as

$$\oint_C \underline{v} d\underline{x}$$

3.3 Surface integrals I: defining a surface

Given a 2D surface $S \subset \mathbb{R}^3$, a 3D vector field can be integrated over the surface S to give a double integral analogue of the line integral.

There are two methods for specifying the surface:

1. Give the surface in parametric form $\underline{x}(u, v)$ where the real parameters u, v lie in some region $U \subseteq \mathbb{R}^2$ called the parameter domain.

In general, $\frac{\partial \underline{x}}{\partial u} \times \frac{\partial \underline{x}}{\partial v}$ is a normal vector to S at $\underline{x}(u, v)$, and so

$$\hat{n} = \left(\frac{\partial \underline{x}}{\partial u} \times \frac{\partial \underline{x}}{\partial v} \right) / \left| \frac{\partial \underline{x}}{\partial u} \times \frac{\partial \underline{x}}{\partial v} \right|$$

is a unit normal at $\underline{x}(u, v)$.

Note: if u and v are swapped we get a unit normal to S , but one pointing in the opposite direction.

- Express the surface as (part of) a level surface of a scalar field f , i.e. given implicitly as $f(x, y, z) = c$. Then ∇f is normal to the level surface S , and $\hat{n} = \frac{\nabla f}{|\nabla f|}$ is a unit normal.

Note: as with method 1, the negative of the unit normal found is also a valid unit normal.

3.4 Surface integrals II: evaluating the integral

Definition 3.4.1. We define the surface integral as a Riemann sum. Let $\underline{F}(\underline{x}) : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be a vector field. Let $S \subset \mathbb{R}^3$ be a surface with parameterised position vector $\underline{x}(u, v)$ with $u, v \in U$, U is the parameter domain.

Assume that the partial derivatives of \underline{x} exist and are continuous and the unit normal $\hat{n}(u, v)$ is continuous (so S is orientable).

The surface integral is defined as

$$\int_S \underline{F} d\underline{A} = \lim_{\Delta A_k \rightarrow 0} \sum_k \underline{F}(\underline{x}_k^*) \cdot \hat{n}_k \Delta A_k$$

Methods for computing:

- The surface S given parametrically as $\underline{x}(u, v)$, construct ΔA_k by approximating as parallelograms, by partitioning S using lines of constant u and v .

Let A_k be the area element with vertices $(u_i, v_j), \underline{x}(u_i + \Delta u_i, v_j), \underline{x}(u_i, v_j + \Delta v_j), \underline{x}(u_i + \Delta u_i, v_j + \Delta v_j)$. Then $\hat{n}_k \Delta A_k = (\underline{x}(u_i + \Delta u_i, v_j) - \underline{x}(u_i, v_j)) \times (\underline{x}(u_i, v_j + \Delta v_j) - \underline{x}(u_i, v_j)) \approx \Delta u_i \frac{\partial \underline{x}}{\partial u} \times \Delta v_j \frac{\partial \underline{x}}{\partial v}$.

Substituting this into the surface integral,

$$\int_S \underline{F} d\underline{A} = \lim_{\Delta u_i, \Delta v_j \rightarrow 0} \sum_{i,j} \underline{F}(\underline{x}_{i,j}^*) \cdot \left(\frac{\partial \underline{x}}{\partial u} \times \frac{\partial \underline{x}}{\partial v} \right) \Delta u_i \Delta v_j$$

Taking the limit, this becomes

$$\int_S \underline{F} d\underline{A} = \int_u \underline{F}(\underline{x}(u, v)) \cdot \left(\frac{\partial \underline{x}}{\partial u} \times \frac{\partial \underline{x}}{\partial v} \right) du dv$$

- Let S be given as (part of) a level set of a scalar field $f(x, y, z)$ and assume $\frac{\partial f}{\partial z} \neq 0$ on S . Then by the implicit function theorem, the points of S can be written as $(x, y, g(x, y))$ for some differentiable function g , where (x, y) range over some region A of the x, y plane (A is the projection of S onto the x, y plane).

We can then apply method 1, with $u = x$ and $v = y$. So $\underline{x}(x, y) = x\underline{e}_1 + y\underline{e}_2 + g(x, y)\underline{e}_3$, and $\frac{\partial \underline{x}}{\partial x} = \underline{e}_1 + \frac{\partial g}{\partial x}\underline{e}_3$, $\frac{\partial \underline{x}}{\partial y} = \underline{e}_2 + \frac{\partial g}{\partial y}\underline{e}_3$.

Using the implicit function theorem, noting that $f(x, y, g(x, y))$ is constant, $0 = \frac{\partial F}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial z} \frac{\partial g}{\partial x} \Rightarrow \frac{\partial g}{\partial x} = \frac{-\partial f / \partial x}{\partial f / \partial z}$ and similarly, $\frac{\partial g}{\partial y} = \frac{-\partial f / \partial y}{\partial f / \partial z}$.

So $\frac{\partial \underline{x}}{\partial x} \times \frac{\partial \underline{x}}{\partial y} = \frac{\nabla f}{\underline{e}_3 \cdot \nabla f}$. Then

$$\int_S \underline{F} d\underline{A} = \int_A \frac{\underline{F} \cdot \nabla f}{\underline{e}_3 \cdot \nabla f} dx dy$$

Remark. In the formula above, the z component of the normal $= 1$ so this corresponds to the upwards (positive z component) choice of normal. If we wanted the downward direction instead, we simply negate the formula.

If $\frac{\partial f}{\partial x} \neq 0$ we can project onto the y, z plane. Similarly, if $\frac{\partial f}{\partial y} \neq 0$ we can project onto the x, y plane.

4 Green's, Stokes' and divergence theorems

4.1 The Big 3 theorems

Theorem 4.1.1. (Green's theorem) Let $P(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}$ and $Q(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}$ be continuously differentiable scalar fields on \mathbb{R}^2 , and let c be a simple closed curve traversed in anti-clockwise direction (the positive direction) which is the boundary of a region A . Then

$$\oint_C (P(x, y)dx + Q(x, y)dy) = \int_A \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dxdy$$

We can also write this in vector form by embedding the xy -plane into \mathbb{R}^3 as the $z = 0$ plane and setting

$$\underline{F}(x, y, z) = (P(x, y), Q(x, y), R)$$

the Green's theorem can be written as

$$\oint_C \underline{F} \cdot d\underline{x} = \int_A (\nabla \times \underline{F}) \cdot \underline{e}_3 dA$$

Proof. Not examinable □

Theorem 4.1.2. (Stokes' theorem) - this generalises the vector form of Green's theorem to arbitrary surfaces in \mathbb{R}^3) Let $\underline{F}(x, y, z) : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ and let S be a surface in \mathbb{R}^3 with area elements $d\underline{A} = \hat{n}dA$ and let $C = \delta S$ be the boundary of S . Then

$$\oint_C \underline{F} \cdot d\underline{x} = \int_S (\nabla \times \underline{F}) \cdot d\underline{A}$$

We need to ensure that the orientations of S and of $C = \delta S$ match, which we can do with the **right-hand rule**:

Curl the fingers of your right hand and extend your thumb. If you placed your hand on the surface near the boundary with your thumb pointing in the direction of the surface normal, then your fingers curl in the direction of the orientation of the boundary.

Proof. Not examinable □

Theorem 4.1.3. (The divergence theorem) Let $V \subset \mathbb{R}^3$ be a volume bounded by S and $\underline{F} : V \rightarrow \mathbb{R}^3$ be continuously differentiable. Then

$$\int_S \underline{F} \cdot d\underline{A} = \int_V (\nabla \cdot \underline{F}) dV$$

where $d\underline{A} = \hat{n}dA$ where \hat{n} is the outward unit normal

Proof. Not examinable □

Remark. These three theorems can be seen as higher dimensional analogues of the fundamental theorem of calculus:

$$\int_a^b \frac{df}{dx} dx = f(b) - f(a)$$

4.2 Path independence of line integrals

In general, line integrals depend on the path between the end points. However, there is a type of vector field for which the line integral is **path independent**, known as a **conservative** vector field.

Example 4.2.1. Calculate the integral $\int_C \underline{F} \cdot d\underline{x}$ for $\underline{F} = (y \cos x, \sin x)$ between $(0, 0)$ and $(1, 1)$ on the paths C_1 , the straight line from $(0, 0)$ to $(1, 1)$ and C_2 , the straight line from $(0, 0)$ to $(1, 0)$ and then to $(1, 1)$.

C_1 is parameterised as $\underline{x}(t) = (t, t)$ for $0 \leq t \leq 1$ so $\frac{d\underline{x}}{dt} = (1, 1)$. $\underline{F}(\underline{x}(t)) = (t \cos t, \sin t)$ so

$$\int_{C_1} \underline{F} \cdot d\underline{x} = \int_0^1 \underline{F}(\underline{x}(t)) \cdot \frac{d\underline{x}}{dt} dt = \sin(1)$$

To calculate $\int_{C_2} \underline{F} \cdot d\underline{x}$, split C_2 into 2 arcs C_{21} and C_{22} . C_{21} is parameterised by $\underline{x}(t) = (t, 0)$ for $0 \leq t \leq 1$ and C_{22} is parameterised by $\underline{x}(t) = (1, t)$ $0 \leq t \leq 1$. So

$$\int_{C_2} \underline{F} \cdot d\underline{x} = \int_{C_{21}} \underline{F} \cdot d\underline{x} + \int_{C_{22}} \underline{F} \cdot d\underline{x} = \int_0^1 \sin(t) dt = \sin(1)$$

Which is the same result as for the integral along C_1 . But we have only checked this for two paths, not infinitely many.

Theorem 4.2.2. Let \underline{F} be continuously differentiable on an open $D \subseteq \mathbb{R}^3$ and let C_1 and C_2 be any two paths from \underline{a} to \underline{b} in D . If $\underline{\nabla} \times \underline{F} = \underline{0}$ then

$$\int_{C_1} \underline{F} \cdot d\underline{x} = \int_{C_2} \underline{F} \cdot d\underline{x}$$

and the line integral only depends on the end points: it is **path independent** and \underline{F} is **conservative**.

Proof. Let

$$\Delta I := \int_{C_1} \underline{F} \cdot d\underline{x} - \int_{C_2} \underline{F} \cdot d\underline{x}$$

For path independence, we need $\Delta I = 0$.

Let C_2 be parameterised in t for $ta \leq t \leq tb$ and let $\overline{C_2}$ be the path along C_2 taken in the opposite direction.

$$\int_{\overline{C_2}} \underline{F} \cdot d\underline{x} = \int_{tb}^{ta} \underline{F}(\underline{x}(t)) \cdot \frac{d\underline{x}}{dt} dt = \int_{tb}^{ta} \underline{F}(\underline{x}(t)) \cdot \frac{d\underline{x}}{dt} dt = - \int_{C_2} \underline{F} \cdot d\underline{x}$$

Then

$$\Delta I = \int_{C_1} \underline{F} \cdot d\underline{x} - \int_{C_2} \underline{F} \cdot d\underline{x} = \int_{C_1} \underline{F} \cdot d\underline{x} + \int_{\overline{C_2}} \underline{F} \cdot d\underline{x} = \oint_C \underline{F} \cdot d\underline{x}$$

where C is the closed path consisting of C_1 followed by $\overline{C_2}$. If C is the boundary of a surface S in D then by Stoke's theorem

$$\Delta I = \oint_C \underline{F} \cdot d\underline{x} = \int_S (\underline{\nabla} \times \underline{F}) \cdot d\underline{A}$$

so $\Delta I = 0$ if $\underline{\nabla} \times \underline{F} = \underline{0}$ throughout D , which implies path independence.

For this to work, we need that every closed curve C is the boundary of a surface in D , which is true if D is **simply connected**, which means that any closed curve in D can be continuously shrunk to a point.

e.g. a sphere is simply connected, a torus is not simply connected. □

Corollary 4.2.3. In a simply connected region D , $\underline{\nabla} \times \underline{F} = \underline{0} \Leftrightarrow$ path independence of $\int_C \underline{F} \cdot d\underline{x} \Leftarrow$ for some scalar field ϕ , $\underline{F} = \underline{\nabla} \phi$.

5 Non-cartesian systems

Definition 5.0.1. For a non-cartesian coordinate system in three dimensions, we use the letters u, v, w to describe the coordinates.

Definition 5.0.2. A change of variable (change of coordinate system) is denoted by

$$\begin{aligned}x &= g(u, v, w) \\y &= h(u, v, w) \\z &= k(u, v, w)\end{aligned}$$

where g, h, k are smooth and invertible functions, and

$$\begin{aligned}u &= \tilde{g}(x, y, z) \\v &= \tilde{h}(x, y, z) \\w &= \tilde{k}(x, y, z)\end{aligned}$$

where $\tilde{g}, \tilde{h}, \tilde{k}$ are smooth and invertible functions.

5.1 Change of variables in surface and volume integrals

5.2 Differential operators in polar and spehrical polar coordinates

Definition 5.2.1. In two dimensions, $P : (x, y)$ denotes a point P of coordinates $(x, y) \in \mathbb{R}^2$. The **position vector** of P is

$$\underline{r} = \underline{OP} = x\underline{e}_1 + y\underline{e}_2$$

where $\{\underline{e}_1, \underline{e}_2\}$ is an orthonormal basis and \underline{e}_1 and \underline{e}_2 do not depend on x or y .

Definition 5.2.2. The **gradient operator** is defined as

$$\underline{\nabla} = \underline{e}_1 \partial_x + \underline{e}_2 \partial_y$$

Definition 5.2.3. The Laplacian operator is defined as

$$\begin{aligned}\underline{\nabla} \cdot \underline{\nabla} &= (\underline{e}_1 \partial_x + \underline{e}_2 \partial_y) \cdot (\underline{e}_1 \partial_x + \underline{e}_2 \partial_y) \\&= \partial_x^2 + \partial_y^2\end{aligned}$$

5.3 Polar coordinates

Definition 5.3.1. A point P in polar coordinates is $P : (r, \theta)$ where $\underline{r} = r \cos(\theta) \underline{e}_1 + r \sin(\theta) \underline{e}_2$, $r \in [0, \infty)$, $\theta \in [0, 2\pi)$

$$|\underline{r}| = r \text{ and } \partial_r \underline{r} = \cos(\theta) \underline{e}_1 + \sin(\theta) \underline{e}_2 =: \tilde{e}_r, \partial_\theta \underline{r} = -r \sin(\theta) \underline{e}_1 + r \cos(\theta) \underline{e}_2 =: \tilde{e}_\theta$$

Definition 5.3.2. Let $x = g(u, v)$ and $y = h(u, v)$ be a change of variables. The norms of the partial derivatives of the position vector \underline{r} are caled the **scale factors** for the mapping (g, h) .

Definition 5.3.3. $h_u = |\partial_u \underline{r}| = |\underline{e}_u|$, $h_v = |\partial_v \underline{r}| = |\underline{e}_v|$.

So in polar coordinates, $h_r = 1$, $h_\theta = r$, $J(r, \theta) = h_r h_\theta$.

Definition 5.3.4. We define \underline{e}_r and \underline{e}_θ to be unit vectors, so $\underline{e}_r = \cos(\theta) \underline{e}_1 + \sin(\theta) \underline{e}_2$, $\underline{e}_\theta = -\sin(\theta) \underline{e}_1 + \cos(\theta) \underline{e}_2$.

Definition 5.3.5. Let $f(r, \theta)$ be a scalar function.

$$\begin{aligned}
df &= \underline{\nabla} f \cdot d\underline{r} \\
&= (\partial_r f) dr + (\partial_\theta f) d\theta \\
&= (\partial_r f) \underline{e}_r \cdot d\underline{r} + (\partial_\theta f) \frac{1}{r} (\underline{e}_\theta d\underline{r}) \\
&= (\underline{e}_r \partial_r + \underline{e}_\theta \frac{1}{r} \partial_\theta f) \cdot d\underline{r}
\end{aligned}$$

So $\underline{\nabla} = \underline{e}_r \partial_r + \frac{1}{r} \underline{e}_\theta \partial_\theta$

Definition 5.3.6. (Divergence in polar) Let $\underline{A}(r, \theta) = A_r \underline{e}_r + A_\theta \underline{e}_\theta$, where A_r and A_θ are functions of r and θ .

Note that $\partial_r \underline{e}_r = 0$, $\partial_r \underline{e}_\theta = 0$, $\partial_\theta \underline{e}_r = \underline{e}_\theta$, $\partial_\theta \underline{e}_\theta = -\underline{e}_r$. Using these results, we get

$$\underline{\nabla} \cdot \underline{A} = \partial_r A_r + \frac{1}{r} A_r + \frac{1}{r} \partial_\theta A_\theta$$

Definition 5.3.7. The Laplacian for polar coordinates is

$$\underline{\nabla} \cdot \underline{\nabla} = \partial_r^2 + \frac{1}{r} \partial_r + \frac{1}{r^2} \partial_\theta^2$$

Proposition 5.3.8. Let $x = g(u, v)$, $y = h(u, v)$ be a change of variables. The Jacobian matrix $A = [[\partial_u x, \partial_v x], [\partial_u y, \partial_v y]]$ is the inverse of the matrix $B = [[\partial_x u, \partial_y u], [\partial_x v, \partial_y v]]$.
TODO: write as matrices.

Proof. $dx = \partial_u x du + \partial_v x dv$, $dy = \partial_u y du + \partial_v y dv$, $du = \partial_x u dx + \partial_y u dy$.

Substituting the second equation into the first:

$$dx = \partial_u x du + \partial_v x (\partial_x v dx + \partial_y v dy)$$

If x is constant, then $0 = \partial_u x du + \partial_v x \partial_y v dy$ or $0 = \partial_u x du \partial_y u + \partial_v x \partial_y v$.

If y is constant, then $dx = \partial_u x du + \partial_v x \partial_x v dx$ or $1 = \partial_u x \partial_x u + \partial_v x \partial_x v$.

Now substituting the fourth equation into the third: If x is constant, $1 = \partial_u y \partial_y u + \partial_v y \partial_y v$,
if y is constant, $0 = \partial_u y \partial_x u + \partial_v y \partial_x v$.

Multiplying A and B gives $AB = I$, the identity matrix. □

6 Generalised functions

6.1 Birth of generalised functions

Definition 6.1.1. The **unit step function** is defined as

$$\Theta(t - t_0) = \begin{cases} 1 & \text{if } t > t_0 \\ 0 & \text{if } t \leq t_0 \end{cases}$$

Definition 6.1.2. The **delta function** is defined as the derivative of the unit step function:

$$\delta(t) = \Theta'(t)$$

6.2 Test functions and distributions

Definition 6.2.1. A **topological space** X is a set of points with a set of neighbourhoods obeying certain axioms.

Definition 6.2.2. An **open set** $O \subset X$ is a set which is a neighbourhood of all its points.

Definition 6.2.3. A set B is a **closed set** if $X \setminus B$ is an open set.

Definition 6.2.4. For $A \subset X$, p is a **limit point** of A if every neighbourhood of p contains at least one element of $A - \{p\}$.

Definition 6.2.5. The **closure** of a set A is the union of A and all its limit points.

Example 6.2.6. In $A = [0, 1)$, 1 is a limit point.

Definition 6.2.7. The **support** of Ψ , denoted $\text{supp}\Psi$, is defined as

$$\text{supp}\Psi := \overline{\{x \in \Omega : \Psi(x) \neq 0\}}$$

is the closure of the set of points $\{x \in \Omega : \Psi(x) \neq 0\}$.

Definition 6.2.8. Let $\Omega \subset \mathbb{R}^n$ be an open set. $\Psi : \mathbb{R}^n \rightarrow \mathbb{R}$ (or \mathbb{C}) is a **test function** if it satisfies:

1. $\Psi \in C^\infty(\Omega)$ i.e. Ψ has finite derivatives of all orders on Ω .
2. $\text{supp}\Psi$ is compact.

Definition 6.2.9. The **space of test functions** is a vector space denoted $\mathcal{D}(\Omega)$ when the test functions are defined on Ω . $\mathcal{D}(\Omega)$ is infinite dimensional, its elements are functions.

Proposition 6.2.10. If $\Psi \in \Omega(\mathbb{R}^n)$, then

1. $\Psi(\underline{x} + \underline{\xi}) \in \mathcal{D}(\mathbb{R}^n)$ for $\underline{\xi} \in \mathbb{R}^n$.
2. $\Psi(-\underline{x}) \in \mathcal{D}(\mathbb{R}^n)$ and $\Psi(a\underline{x}) \in \mathcal{D}(\mathbb{R}^n)$ for $a \in \mathbb{R} - \{0\}$.
3. $g(x)\Psi(x) \in \mathcal{D}(\mathbb{R}^n)$ for $g \in C^\infty(\mathbb{R}^n)$.

Remark. For a vector space V with norm $\|\cdot\|$, a sequence of functions $(f_m)_{m \in \mathbb{N}}$ converges to a limit function f if

$$\lim_{m \rightarrow \infty} \|f_m - f\| = 0$$

This is pointwise convergence. This requires $f_n(x)$ and $f(x)$ to be close for all $m > M_x$, but depending on the point x .

Definition 6.2.11. A sequence of test functions $(\Psi_m)_{m \in \mathbb{N}} \subseteq \mathcal{D}(\Omega)$ converges to a test function $\Psi \in \mathcal{D}(\Omega)$ as $m \rightarrow \infty$ if

1. For some compact set $K \subset \Omega$, $\text{supp} \Psi_m \subset K \quad \forall m \in \mathbb{N}$.
2. $(\Psi_m)_{m \in \mathbb{N}} \rightarrow \Psi$ with uniform convergence.
- 3.

$$(D_k \Psi_m) := \left(\frac{\partial^{k_1}}{\partial (x_1)^{k_1}} \frac{\partial^{k_2}}{\partial (x_2)^{k_2}} \cdots \frac{\partial^{k_n}}{\partial (x_n)^{k_n}} \right) \Psi_m \rightarrow D^k \Psi \text{ as } m \rightarrow \infty$$

with uniform convergence. We need uniform convergence of $D^k \Psi_m$ because the derivatives of all orders of test functions are also test functions.

These criteria are called \mathcal{D} -convergence.

Definition 6.2.12. We can summarise the notion of \mathcal{D} -convergence through the use of the **sup norm**. It is defined as

$$\|f\|_\infty := \sup\{|f(x)| : x \in \Omega\}$$

Here, the supremum is the maximum since K is bounded.

Definition 6.2.13. The sequence (Ψ_m) of test functions in $\mathcal{D}(\Omega)$ converges to $\Psi \in \mathcal{D}(\Omega)$ if

1. For some compact $K \subset \Omega$, $\text{supp} \Psi_m \subset K \quad \forall m$.
2. For all multi-indices $k = (k_1, \dots, k_n)$,

$$\|D^k \Psi_m - D^k \Psi\|_\infty \rightarrow 0 \text{ as } m \rightarrow \infty$$

Definition 6.2.14. Let $\Omega \subset \mathbb{R}^n$ be an open set. A **distribution** is linear continuous map $T : \mathcal{D}(\Omega) \rightarrow \mathbb{R}$. It satisfies

1. **Linearity:** $\forall \Psi \in \mathcal{D}(\Omega), \forall \phi \in \mathcal{D}(\Omega), \forall a, b \in \mathbb{R}^n$,

$$T(a\Psi + b\phi) = aT(\Psi) + bT(\phi)$$

2. **Continuity:** $\forall \Psi \in \mathcal{D}(\Omega), \forall (\Psi_m)_{m \in \mathbb{N}} \subseteq \mathcal{D}(\Omega)$ with $(\Psi_m) \rightarrow \Psi$,

$$T(\Psi_m) \rightarrow T(\Psi) \text{ as } m \rightarrow \infty$$

Definition 6.2.15. The vector space of distributions defined via test functions is denoted by $\mathcal{D}'(\Omega)$ and is also infinite-dimensional.

Example 6.2.16. The Dirac delta function is a distribution

$$\delta : \mathcal{D}(\mathbb{R}^n) \rightarrow \mathbb{R}, \delta(\Psi) := \Psi(\underline{0})$$

Proof.

1. Linearity: $\delta(a\Psi + b\phi) = (a\Psi + b\phi)(\underline{0}) = a\Psi(\underline{0}) + b\phi(\underline{0}) = a\delta(\Psi) + b\delta(\phi)$.
2. Continuity: $\forall \Psi \in \mathcal{D}(\Omega), \forall (\Psi_m)_{m \in \mathbb{N}} \subseteq \mathcal{D}(\Omega)$ with $(\Psi_m) \rightarrow \Psi$, but uniform convergence in \mathcal{D} -convergence implies pointwise convergence. So $(\Psi_m(\underline{x})) \rightarrow \Psi(\underline{x})$ as $m \rightarrow \infty \quad \forall \underline{x} \in \mathbb{R}^n$. Since $\underline{0} \in \mathbb{R}^n$,

$$\delta(\Psi_m) = (\Psi_m(\underline{0})) \rightarrow \Psi(\underline{0}) = \delta(\Psi) \text{ as } m \rightarrow \infty$$

□

Example 6.2.17. Continuous functions on \mathbb{R}^n are distributions. Let $f \in C^0(\mathbb{R})$. To treat f as a distribution, we must define how this distribution acts on $\mathcal{D}(\Omega)$.

Define $T_f : \mathcal{D}(\mathbb{R}^n) \rightarrow \mathbb{R}$,

$$T_f(\psi) = \int_{\mathbb{R}^n} f(\underline{x})\psi(\underline{x})d\underline{x}$$

Then T_f is a distribution.

- Linearity:

$$\begin{aligned} T_f(a\psi + b\phi) &= \int_{\mathbb{R}^n} f(\underline{x})(a\psi(\underline{x}) + b\phi(\underline{x}))d\underline{x} \\ &= a \int_{\mathbb{R}^n} f(\underline{x})\psi(\underline{x})d\underline{x} + b \int_{\mathbb{R}^n} f(\underline{x})\phi(\underline{x})d\underline{x} \\ &= aT_f(\psi) + bT_f(\phi) \end{aligned}$$

- Continuity: $\forall \psi \in \mathcal{D}(\mathbb{R}^n), \forall (\psi_m)_{m \in \mathbb{N}} \subseteq \mathcal{D}(\mathbb{R}^n)$ with $(\psi_m) \rightarrow \psi$, we have in particular uniform convergence. Therefore we can interchange lim and \int , so

$$\begin{aligned} \lim_{m \rightarrow \infty} T_f(\psi_m) &= \lim_{m \rightarrow \infty} \int_{\mathbb{R}^n} f(\underline{x})\psi_m(\underline{x})d\underline{x} \\ &= \int_{\mathbb{R}^n} \lim_{m \rightarrow \infty} f(\underline{x})\psi_m(\underline{x})d\underline{x} \\ &= \int_{\mathbb{R}^n} f(\underline{x})\psi(\underline{x})d\underline{x} \\ &= T_f(\psi) \end{aligned}$$

So

$$T_{f_m}(\psi) = \int_{\mathbb{R}^n} f_m(\underline{x})\psi(\underline{x})d\underline{x} \rightarrow \psi(0) \int_{\mathbb{R}^n} f_m(\underline{x})d\underline{x} = \psi(\underline{0}) \quad \text{as } m \rightarrow \infty$$

6.3 Regular and singular distributions

Definition 6.3.1. A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is called **locally integrable** if for every compact set $K \subset \mathbb{R}^n$,

$$\int_K |f(\underline{x})|d\underline{x} < \infty$$

Example 6.3.2. $f(x) = x^2$ is not integrable since

$$\int_{\mathbb{R}} f(x)dx$$

is not finite but it is locally integrable.

Definition 6.3.3. The space of locally integrable functions is called L^1_{loc} .

Definition 6.3.4. $T \in \mathcal{D}'(\mathbb{R}^n)$ is called a **regular distribution** if for some locally integrable function f ,

$$T(\psi) = T_f(\psi) = \int_{\mathbb{R}^n} f(\underline{x})\psi(\underline{x})d\underline{x}, \quad \psi \in \mathcal{D}(\mathbb{R}^n)$$

In general, if two locally integrable functions only differ by a finite amount at isolated points, they define the same distribution.

Example 6.3.5. Let $f : \mathbb{R} \rightarrow \mathbb{R}$,

$$f(x) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{otherwise} \end{cases}$$

then $f \in L^1_{\text{loc}}(\mathbb{R})$, and

$$T_f(\psi) = \int_{\mathbb{R}} f(x)\psi(x)dx = 0$$

Definition 6.3.6. If there is no $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ such that a distribution T can be written as T_f , then this distribution is called **singular**. We write

$$T(\psi) = \int_{\mathbb{R}^n} T(\underline{x})\psi(\underline{x})d\underline{x} = \langle T, \psi \rangle$$

To specify the distribution T , we must define how it acts on test functions.

Example 6.3.7. δ is not a regular distribution: there is no $f \in L^1_{\text{loc}}$ such that

$$\delta(\psi) = \int_{\mathbb{R}^n} f(\underline{x})\psi(\underline{x})d\underline{x}$$

However, we can write symbolically

$$\delta(\psi) = \int_{\mathbb{R}^n} \delta(\underline{x})\psi(\underline{x})d\underline{x} = \langle \delta, \psi \rangle := \psi(\underline{0})$$

This is called the **sifting property** of the δ distribution.

Remark. Important: $\delta(\underline{x})$ is **not** a function, we just use it as a function in the integral symbolically.

Remark. More generally, with $\Omega \subset \mathbb{R}^n$,

$$\delta(\psi) = \int_{\mathbb{R}^n} \delta(\underline{x})\psi(\underline{x})d\underline{x} = \langle \delta, \psi \rangle = \begin{cases} \psi(\underline{0}) & \text{if } \underline{0} \in \Omega \\ 0 & \text{otherwise} \end{cases}$$

Definition 6.3.8. In one dimension, we write $\delta(\underline{x}) = \delta(x)$. In n dimensions, we write $\delta(x_1, \dots, x_n) = \delta(x_1) \cdots \delta(x_n)$

6.4 Operations on distributions

Definition 6.4.1. $\forall \psi \in \mathcal{D}(\mathbb{R}^n)$, $\forall T_1, T_2 \in \mathcal{D}'(\mathbb{R}^n)$, we define the following operations:

1. **Addition:** $(T_1 + T_2)(\psi) = T_1(\psi) + T_2(\psi)$.
2. **Multiplication by a constant:** $(cT)(\psi) = cT(\psi)$ for every constant c .
3. **Shifting:** for $\underline{\xi} \in \mathbb{R}^n$,

$$\begin{aligned} T_{\underline{\xi}}(\psi(\underline{x})) &:= \int_{\mathbb{R}^n} T(\underline{x} - \underline{\xi})\psi(\underline{x})d\underline{x} \\ &= \int_{\mathbb{R}^n} T(\underline{y})\psi(\underline{y} + \underline{\xi})d\underline{y} \\ &= T(\psi(\underline{y} + \underline{\xi})) \end{aligned}$$

4. **Transposition:**

$$\begin{aligned} T^t(\psi(\underline{x})) &:= \int_{\mathbb{R}^n} T(-\underline{x})\psi(\underline{x})d\underline{x} \\ &= \int_{\mathbb{R}^n} T(\underline{y})\psi(-\underline{y})d\underline{y} \\ &= T(\psi(-\underline{x})) \end{aligned}$$

5. **Dilation:**

$$\begin{aligned} T_{(\alpha)}(\psi(x)) &= \int_{\mathbb{R}^n} T(\alpha\underline{x})\psi(\underline{x})d\underline{x} \\ &= \frac{1}{|\alpha|^n} \int_{\mathbb{R}^n} T(\underline{y})\psi\left(\frac{\underline{y}}{\alpha}\right) d\underline{y} \\ &= \frac{1}{|\alpha|^n} T\left(\psi\left(\frac{\underline{y}}{\alpha}\right)\right) \end{aligned}$$

6. **Multiplication by a smooth function** $\phi \in C^\infty(\mathbb{R}^n)$:

$$(\phi T)(\psi) = T(\phi\psi)$$

Remark. The above rules are natural for regular distributions, since the integrals are meaningful in the classical sense. For singular distributions, we extend the formalism and use the (symbolic) notations

$$T(\psi) = \int_{\mathbb{R}^n} T(\underline{x})\psi(\underline{x})d\underline{x}$$

Remark. Important: If $\Omega \subset \mathbb{R}$, with Ω open,

$$\delta(\xi)(\psi(x)) = \begin{cases} \psi(\xi) & \text{if } \xi \in \Omega \\ 0 & \text{otherwise} \end{cases}$$

This is called the **sifting** property of the delta distribution.

Example 6.4.2. Calculating with $\delta \in \mathcal{D}'(\mathbb{R})$:

- Shifting:

$$\begin{aligned} \delta_\xi(\psi(x)) &= \int_{\mathbb{R}} \delta(x - \xi)\psi(x)dx \\ &= \int_{\mathbb{R}} \delta(y)\psi(y + \xi)dy \\ &= \int_{\mathbb{R}} \delta(x)\psi(x + \xi)dx \\ &= \psi(\xi) \\ &= \delta(\psi(x + \xi)) \end{aligned}$$

- $\phi\delta_\xi(\psi)$ for smooth ϕ :

$$\begin{aligned} \phi\delta_\xi(\psi) &= \int_{\mathbb{R}} \phi(x)\delta(x - \xi)\psi(x)dx \\ &= \int_{\mathbb{R}} \delta(x - \xi)(\phi(x)\psi(x))dx \\ &= \phi(\xi)\psi(\xi) \\ &= \phi(\xi) \int_{\mathbb{R}} \delta(x - \xi)\psi(x)dx \\ &= \int_{\mathbb{R}} \phi(\xi)\delta(x - \xi)\psi(x)dx \end{aligned}$$

- $\delta_{(\alpha)}(\psi)$:

$$\begin{aligned}\delta_{(\alpha)}(\psi) &= \int_{\mathbb{R}} \delta(\alpha x) \psi(x) dx \\ &= \frac{1}{\alpha} \int_{\mathbb{R}} \delta(y) \psi\left(\frac{y}{\alpha}\right) dy\end{aligned}$$

For $\alpha < 0$, $(-\alpha) > 0$, so $dy = \alpha dx = -|a|dx$ so the limits of integrations swap from $\int_{-\infty}^{\infty}$ to $\int_{\infty}^{-\infty}$ but the minus sign in $-|a|dx$ compensates for this.

- For x_1, \dots, x_n are zeros of f .

$$\int_{\Omega} \delta(f(x)) \psi(x) dx = \sum_{i=1}^n \delta((x - x_i) f'(x_i)) \psi(x) dx = \sum_{i=1}^n \frac{1}{|f'(x_i)|} \psi(x_i)$$

by the sifting property of the delta distribution.

Example 6.4.3. For $b > 0$, let $f(x) = x^2 - b^2$, $f : \mathbb{R} \rightarrow \mathbb{R}$, then $f \in C^1(\mathbb{R})$, $f(b) = f(-b) = 0$, with $x_1, x_2 = b, -b$ as the zeroes. $f'(x) = 2x$ so $f'(b) = 2b$ and $f'(-b) = -2b$. Then

$$\int_{\mathbb{R}} \delta(x^2 - b^2) \psi(x) dx = \frac{1}{2|b|} (\psi(b) + \psi(-b))$$

Definition 6.4.4. T_1 and $T_2 \in \mathcal{D}'(\Omega)$ are **equal** if

$$\int_{\Omega} T_1(x) \psi(x) dx = \int_{\Omega} T_2(x) \psi(x) dx \quad \forall \psi \in \mathcal{D}(\Omega)$$

Example 6.4.5. For which value of $A \in \mathbb{R}$ are $T_1 = x^2 \delta(x^3)$ and $T_2 = A \delta(x)$ equal?

$$\int_{\mathbb{R}} x^2 \delta(x^3) \psi(x) dx = \frac{1}{3} \int_{\mathbb{R}} \delta(y) \psi(y^{1/3}) dy = \frac{1}{3} \psi(0) \int_{\mathbb{R}} A \delta(x) \psi(x) dx = A \psi(0)$$

so $A = 1/3$.

6.5 The derivative of a distribution

Definition 6.5.1. For a distribution $T \in \mathcal{D}'(\Omega)$, its **derivative** is defined as

$$T'(\psi) = -T(\psi') \quad \forall \psi \in \mathcal{D}(\Omega)$$

or symbolically, $\langle T', \psi \rangle = -\langle T, \psi' \rangle$. The **n th derivative** of T is defined as

$$T^{(n)}(\psi) = (-1)^n T(\psi^{(n)})$$

or symbolically, $\langle T^{(n)}, \psi \rangle = (-1)^n \langle T, \psi^{(n)} \rangle$.

Example 6.5.2. The unit step function $\Theta(t)$ is a locally integrable function, so defined a regular distribution:

$$\Theta(\psi) = T_{\Theta}(\psi) = \int_{-\infty}^{\infty} \Theta(t) \psi(t) dt = \int_0^{\infty} \psi(t) dt < \infty$$

because ψ has compact support. So

$$\begin{aligned}\Theta'(t) &= -\Theta(\psi') = -\int_{-\infty}^{\infty} \Theta(t) \psi'(t) dt \\ &= -\int_0^{\infty} \psi'(t) dt = -\psi(\infty) + \psi(0) = \psi(0) = \delta(\psi)\end{aligned}$$

because ψ has compact support. So we have, in the sense of distributions,

$$\Theta'(t) = \delta(t)$$

Example 6.5.3. The derivative of the delta distribution is $\delta'(\psi) = -\delta(\psi') = -\psi'(0)$ and generally, $\delta^{(n)}(\psi) = (-1)^n \delta(\psi^{(n)}) = (-1)^n \psi^{(n)}(0)$.

Example 6.5.4. Show that, in the sense of distributions, $\delta'(cx) = \frac{1}{c^2} \delta'(x)$ for a constant $c > 0$.

$$\begin{aligned} \int_{-\infty}^{\infty} \delta'(cx) \psi(x) dx &= \frac{1}{c} \int_{-\infty}^{\infty} \delta'(y) \psi(y/c) dy = -\frac{1}{c} \int_{-\infty}^{\infty} \delta(y) (\psi(y/c))' dy \\ &= -\frac{1}{c^2} \int_{-\infty}^{\infty} \delta(y) \psi'(y/c) dy = \frac{1}{c^2} \psi'(0) = -\frac{1}{c^2} \int_{-\infty}^{\infty} \delta(x) \psi'(x) dx \\ &= \frac{1}{c^2} \int_{-\infty}^{\infty} \delta'(x) \psi(x) dx \end{aligned}$$

6.6 Leibniz rule for differentiation

Proposition 6.6.1. For a distribution T and a smooth function ϕ , the product rule holds:

$$(\phi T)' = \phi' T + \phi T'$$

and more generally, the Leibniz rule holds:

$$(\phi T)^{(n)} = \sum_{k=0}^n \binom{n}{k} \phi^{(k)} T^{(n-k)}$$

Example 6.6.2. Let $\psi \in \mathcal{D}(\Omega)$. The symbolic representation of $(\phi \delta_a)^{(n)}(\psi)$ is given by

$$\begin{aligned} (\phi \delta_a)^{(n)}(\psi) &= \int_{\Omega} (\phi(x) \delta(x-a))^{(n)} \psi(x) dx \\ &= (-1)^n \int_{\Omega} (\phi(x) \delta(x-a)) \psi(x)^{(n)} dx = (-1)^n \phi(a) \psi^{(n)}(a) \end{aligned}$$

We can write this in inner product symbolic notation as

$$\langle (\phi \delta_a)^{(n)}, \psi \rangle = (-1)^n = \langle \phi \delta_a, \psi^{(n)} \rangle = (-1)^n \langle \delta_a, \phi \psi^{(n)} \rangle = (-1)^n \phi(a) \psi^{(n)}(a)$$

Alternatively, by the sifting property of the delta distribution, we can write

$$(\phi \delta_a)^{(n)}(\psi) = \int_{\Omega} (\phi(x) \delta(x-a))^{(n)} \psi(x) dx = \phi(a) \int_{\Omega} \delta^{(n)}(x-a) \psi(x) dx$$

before using the definition of the generalised derivative on the δ distribution. So a smooth factor of the delta function may be replaced by a constant and taken outside the integral.

TODO: notes up to end of section 2.4.1

Definition 6.6.3. A function f is **piecewise continuous** for $x \in (a, b)$ if (a, b) can be divided into a finite number of sub intervals and

1. f is continuous on each sub interval.
2. f tends to a finite limit on the boundaries of each sub interval as approached from the interior of these sub intervals.

Example 6.6.4.

$$f(x) = \begin{cases} \sin(x) & \text{if } -3\pi < x\pi \\ \frac{1}{2} \cos(x) & \text{if } \pi \leq x < 2\pi \\ \frac{3}{2} & \text{if } 2\pi \leq x < 3\pi \end{cases}$$

on the interval $(a, b) = (-3\pi, 3\pi)$. The sub intervals are $(-3\pi, \pi), (\pi, 2\pi), (2\pi, 3\pi)$. f is piecewise continuous.

Definition 6.6.5. A function f is **piecewise smooth** if

1. f is piecewise continuous.
2. f has piecewise continuous derivatives.

Example 6.6.6.

$$f(x) = \begin{cases} x & \text{if } -\pi < x \leq 0 \\ \pi & \text{if } 0 < x < \pi \end{cases}$$

on the interval $(-\pi, \pi)$. Then

$$f'(x) = \begin{cases} 1 & \text{if } -\pi < x \leq 0 \\ 0 & \text{if } 0 < x < \pi \end{cases}$$

which is piecewise continuous.

Example 6.6.7. \tan is not a piecewise smooth function.

Definition 6.6.8. To calculate the derivative of a piecewise smooth function which can be split into f_1, f_2, \dots on sub intervals $(a, x_1), (x_1, x_2), \dots, (x_n, b)$:

1. Define $\tilde{f}(x) = f_1(x) + (f_2(x) - f_1(x))\Theta(x - x_1) + (f_3(x) - f_2(x))\Theta(x - x_2) + \dots$. But $\tilde{f}(x) = f(x)$ on (a, b) .
2. Compute $f'(x) = \tilde{f}'(x)$.

Example 6.6.9.

$$f(x) = \begin{cases} \sin(x) & \text{if } -3\pi < x\pi \\ \frac{1}{2} \cos(x) & \text{if } \pi \leq x < 2\pi \\ \frac{3}{2} & \text{if } 2\pi \leq x < 3\pi \end{cases}$$

Then

$$f(x) = \tilde{f}(x) = \sin(x) + \left(\frac{1}{2} \cos(x) - \sin(x)\right) \Theta(x - \pi) + \left(\frac{3}{2} - \frac{1}{2} \cos(x)\right) \Theta(x - 2\pi)$$

so

$$\begin{aligned} f'(x) &= \cos(x) + \left(-\frac{1}{2} \sin(x) - \cos(x)\right) \Theta(x - \pi) + \left(\frac{1}{2} \cos(x) - \sin(x)\right) \delta(x - \pi) + \frac{1}{2} \sin(x) \Theta(x - 2\pi) + \\ &= \cos(x) - \left(\frac{1}{2} \sin(x) + \cos(x)\right) \Theta(x - \pi) + \left(-\frac{1}{2} - 0\right) \delta(x - \pi) + \frac{1}{2} \sin(x) \Theta(x - 2\pi) + \left(\frac{3}{2} - \frac{1}{2}\right) \delta(x - 2\pi) \\ &= -\frac{1}{2} \delta(x - \pi) + \delta(x - 2\pi) + \begin{cases} \cos(x) & \text{if } 3\pi < x \leq \pi \\ -\frac{1}{2} \sin(x) & \text{if } \pi < x \leq 2\pi \\ 0 & \text{if } 2\pi < x < 3\pi \end{cases} \end{aligned}$$

Definition 6.6.10. The **sign function** is defined as

$$\text{sgn}(x) = \Theta(x) - \Theta(-x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}$$

Example 6.6.11.

$$\begin{aligned} \Theta'(-x)(\psi) &= \int_{-\infty}^{\infty} \Theta'(-x)\psi(x)dx \\ &= - \int_{-\infty}^{\infty} \Theta(-x) \frac{d}{d(-x)} \psi(x)dx \\ &= \int_{-\infty}^{\infty} \Theta(-x) \frac{d}{dx} \psi(x)dx \\ &= \int_{-\infty}^{\infty} \frac{d}{dx} \psi(x)dx = \psi(0) - \psi(-\infty) = \psi(0) = \delta(\psi) \end{aligned}$$

So the derivative of the sign function is

$$\text{sgn}'(x) = \frac{d}{dx}\Theta(x) - \frac{d}{dx}\Theta(-x) = \delta(x) - \frac{d}{dx}\Theta(-x) = 2\delta(x)$$

6.7 The delta distribution in polar/spherical coordinates

Definition 6.7.1. For a change of coordinates the delta function is defined as

$$\delta(\underline{x} - \underline{x}_0) = \frac{1}{|J|} \delta(\underline{\xi} - \underline{\xi}_0)$$

where $\underline{x} \in \mathbb{R}^n, \underline{\xi} \in \mathbb{R}^n$, $\underline{x}, \underline{\xi}$ are constant vectors of coordinates $(x_1, \dots, x_n), (\xi_1, \dots, \xi_n)$. J is the Jacobian of the change of variable from \underline{x} to $\underline{\xi}$.

Example 6.7.2. If the problem at hand does not depend on θ , then

$$\delta(\underline{x} - \underline{x}_0) = \frac{1}{2\pi r} \delta(\underline{r} - \underline{r}_0)$$

7 Sturm-Liouville Theory

TODO: notes from lecture

Theorem 7.0.1. Let $f : [a, b] \rightarrow \mathbb{R}$ be bounded. Then f is Riemann integrable iff f is continuous ‘almost everywhere’ (it is continuous except on sets with zero measure).

7.1 Hilbert spaces: definitions and examples

Definition 7.1.1. A **Hilbert space** \mathbb{H} is a real or complex vector space which satisfies:

1. \mathbb{H} has a symmetric (or hermitian for a complex vector space) inner product, $\langle \cdot, \cdot \rangle : \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{C}$, which is a continuous function of two variables. It associates a complex number $\langle u, v \rangle$ to every pair (u, v) in $\mathbb{H} \times \mathbb{H}$. $\langle \cdot, \cdot \rangle$ and satisfies:
 - (a) **Hermiticity or symmetry:** $\langle u, v \rangle = \overline{\langle v, u \rangle}$.
 - (b) **Anti-linearity in the first entry:** $\forall a \in \mathbb{C}, \langle a(u + v), w \rangle = \bar{a} \langle u + v, w \rangle = \bar{a} \langle u, w \rangle + \bar{a} \langle v, w \rangle$.
 - (c) **Positivity:** $\langle u, v \rangle \geq 0$ and $\langle u, u \rangle = 0 \iff u = 0$.
2. \mathbb{H} is **complete** for the norm induced by the inner product, which is defined as $\|u\| = \sqrt{\langle u, u \rangle}$. (**Complete** means that every Cauchy sequence in \mathbb{H} has a limit in \mathbb{H}).

Remark. $\overline{\langle a(u + v), w \rangle} = \langle w, a(u + v) \rangle = a \overline{\langle u, w \rangle} + a \overline{\langle v, w \rangle} = a \langle w, u \rangle + a \langle w, v \rangle$. This is linearity in the second entry, or **sesquilinearity**.

Definition 7.1.2. A vector space with an inner product and induced norm $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$ is called a **inner product space** (or **pre-Hilbert space**).

Definition 7.1.3. If V is an inner product space, it can be viewed as a **metric space** with metric d , where

$$d(u, v) := \|u - v\| = \sqrt{\langle u - v, u - v \rangle}$$

Then a sequence $(v_n)_{n \in \mathbb{N}} \in (V, d)$ converges to v **in norm** if

$$\lim_{n \rightarrow \infty} \|v_n - v\| = 0$$

or equivalently,

$$\forall \epsilon > 0, \exists n \in \mathbb{N}, \forall n \geq N, \quad \|v_n - v\| < \epsilon$$

Definition 7.1.4. A sequence $(v_n)_{n \in \mathbb{N}} \in (V, d)$ is called a **Cauchy sequence** if

$$\forall \epsilon > 0, \exists n \in \mathbb{N}, \forall m \geq N, \forall n \geq N, \quad d(v_m, v_n) < \epsilon$$

Remark. Every Cauchy sequence is convergent in \mathbb{R} , since \mathbb{R} is complete, but this is not true in all spaces.

Definition 7.1.5. A metric space (V, d) is called **complete** if every Cauchy sequence in (V, d) converges in V .

7.2 Inner products defined by integrals

Definition 7.2.1. Consider the vector space V of complex-valued functions, defined on an interval $[a, b]$ and let $\omega : [a, b] \rightarrow \mathbb{R}^+$ be a non-negative function with finitely many zeroes on $[a, b]$. ω is called a **weight** function.

The inner product with weight ω for every pair of functions $(u, v) \in V^2$ is defined as

$$\langle u, v \rangle_\omega = \int_a^b \bar{u}(x)v(x)\omega(x)dx$$

Remark. If $\omega = 1$, then $\langle u, v \rangle_1$ is just written as $\langle u, v \rangle$.

Remark. $\langle u, v \rangle_\omega = \langle u, v\omega \rangle_1 = \langle u, v\omega \rangle = \langle u\omega, v \rangle$, since ω is a real function.

Example 7.2.2. The space $C^0(\omega, [a, b])$ of continuous functions on $[a, b]$ with inner product $*$ and norm

$$\|u\|_{L^2(\omega, [a, b])} := \sqrt{\int_a^b \bar{u}(x)u(x)\omega(x)dx} = \sqrt{\int_a^b |u(x)|^2\omega(x)dx}$$

is not complete for the norm L^2 . Its completion is $L^2(\omega, [a, b])$, the space of square-integrable functions on $[a, b]$ with inner product $\langle u, v \rangle_\omega$.

Example 7.2.3. Consider $C^0(1, [0, 1])$, written as $C^0([a, b])$. Let

$$f_n(x) = \begin{cases} 1 & \text{if } x > \frac{1}{2} \\ 0 & \text{if } x \leq \frac{1}{2} - \frac{1}{n} \\ n(x - \frac{1}{2} - \frac{1}{n}) & \text{if } \frac{1}{2} - \frac{1}{n} < x < \frac{1}{2} \end{cases}$$

$(f_n)_{n \in \mathbb{N}}$ is a Cauchy sequence, but it does not converge to a function in $C^0([0, 1])$.

7.3 Bounded and unbounded linear operators

Definition 7.3.1. Let (V, d) be a metric space. $W \subset V$ is called **dense** in V if

$$\forall v \in V, \forall \epsilon > 0, \exists w \in W, \quad d(v, w) < \epsilon$$

So W is dense in V if every point in V has points in W that are arbitrarily close.

Definition 7.3.2. A **linear Operator** on \mathbb{H} is a pair $(L, D(L))$ which satisfies

1. $D(L) \subseteq \mathbb{H}$ and $D(L)$ is dense in \mathbb{H} .
2. $L : D(L) \rightarrow \mathbb{H}$ is linear, so it satisfies

$$L(au + bv) = aL(u) + bL(v)$$

for every $a, b \in \mathbb{C}$, for every $u, v \in D(L)$.

Example 7.3.3. Let $L = -\frac{\partial^2}{\partial x^2}$.

1. The pair $(L, D_N(L))$ with $D_N(L) := \{u \in C^2([a, b]) \subset L^2([a, b]) : u'(a) = u'(b) = 0\}$ is a linear Operator on $L^2([a, b])$. $D_N(L)$ is dense in $L^2([a, b])$ since $C^\infty([a, b]) \subset C^2([a, b])$ is dense in $L^2([a, b])$. Here, D_N stands for Neumann boundary conditions.

2. The pair $(L, D_D(L)) := \{u \in C^2([a, b]) \subset L^2([a, b]) : u(a) = u(b) = 0\}$ is a linear Operator. Here, D_D stands for Dirichlet boundary conditions. This linear Operator is different from $(L, D_N(L))$, since for example, if $\lambda = 0$, then $L(u) = \lambda u$ and $L(u) = -\frac{\partial^2 u}{\partial x^2} = 0$, so $u(x) = Ax + B$ for constants A and B . But if $u(a) = u(b) = 0$, then $u(x) = 0$ for all $x \in [a, b]$. Whereas if $u'(a) = u'(b) = 0$, then $A = 0$ so $u(x) = B$.

Definition 7.3.4. Let $(\mathbb{H}_1, \|\cdot\|_{\mathbb{H}_1})$ and $(\mathbb{H}_2, \|\cdot\|_{\mathbb{H}_2})$ be Hilbert spaces. A linear operator $L : \mathbb{H}_1 \rightarrow \mathbb{H}_2$ is called **bounded** if for some $M > 0$,

$$\forall u \in \mathbb{H}_1, \quad \|L(v)\|_{\mathbb{H}_2} \leq M\|v\|_{\mathbb{H}_1}$$

If no such M exists, then L is called **unbounded**.

Definition 7.3.5. The **norm** of L is defined as

$$\|L\| := \inf\{M : \|L(v)\|_{\mathbb{H}_2} \leq M\|v\|_{\mathbb{H}_1} \text{ for every } v \in \mathbb{H}_1\}$$

Example 7.3.6. $I : \mathbb{H} \rightarrow \mathbb{H}$ defined as $I(v) = v$ is bounded, since

$$\|I(v)\|_{\mathbb{H}} = \|v\|_{\mathbb{H}} \leq M\|v\|_{\mathbb{H}}$$

for every $M \geq 1$. So $\|I\| = 1$ as $M = 1$ is the lower bound.

Example 7.3.7. Differential operators are usually unbounded. For example,

$$A : L^2([a, b]) \rightarrow L^2([a, b]), \quad A(x(t)) = x'(t)$$

The domain, $D(A) = \{x : x \in L^2([-\pi, \pi]), x \in C^1([a, b])\}$. Consider a family of functions $x_n(t)$ in $L^2([-\pi, \pi])$, for example, $x_n(t) = \cos(nt)$ then $A(x_n) = -n \sin(nt)$. Now,

$$\begin{aligned} \|x_n\|_{L^2} &= \sqrt{\int_{-\pi}^{\pi} \overline{\cos(nt)} \cos(nt) dt} = \sqrt{\pi} \\ \|A(x_n)\|_{L^2} &= \sqrt{\int_{-\pi}^{\pi} \overline{-n \sin(nt)} (-n \sin(nt)) dt} = \sqrt{\pi n^2} \end{aligned}$$

Hence

$$\frac{\|A(x_n)\|_{L^2}}{\|x_n\|_{L^2}} = n$$

which is unbounded.

Definition 7.3.8. Let $\mathbb{H}_1, \mathbb{H}_2$ be Hilbert spaces, with $\langle \cdot, \cdot \rangle^{(1)}$ and $\langle \cdot, \cdot \rangle^{(2)}$. Let $L : D(L) \subset \mathbb{H}_1 \rightarrow \mathbb{H}_2$ be an unbounded linear Operator on \mathbb{H}_1 , with $D(L)$ dense in \mathbb{H}_1 .

The **adjoint** $(L^*, D(L^*))$ of L is defined as the linear Operator with $L^* : D(L^*) \subset \mathbb{H}_2 \rightarrow \mathbb{H}_1$ defined by

$$\langle L(v_1), v_2 \rangle^{(2)} = \langle v_1, L^*(v_2) \rangle^{(1)}, \quad v_1 \in \mathbb{H}_1$$

where

$$D(L^*) = \{v_2 \in \mathbb{H}_2 : \exists v_2^* \in \mathbb{H}_1, \langle v_1, v_2^* \rangle^{(1)} = \langle L(v_1), v_2 \rangle^{(2)} \forall v_1 \in \mathbb{H}_1\}$$

For every $v_2 \in D(L^*)$, $v_2^* \in D(L)$ is unique and $v_2^* = L^*(v_2)$.

7.4 Sturm-Liouville operators

Proposition 7.4.1. (Green's formula) Let $(L, D(L))$ be a linear differential Operator and define

$$L^* := \overline{p_0}(x)d_x^2 + (2\overline{p_0}'(x) - \overline{p_1}(x))d_x + (\overline{p_0}''(x) - \overline{p_1}'(x) + \overline{p_2}(x))$$

Then

$$\forall u, v \in C^2([a, b]), \quad \langle L(u), v \rangle - \langle u, L^*(v) \rangle = [\overline{p_0}(v\overline{u}' - v'\overline{u}) + (\overline{p_1} - \overline{p_0}')v\overline{u}]_a^b$$

Proof. Omitted. □

Definition 7.4.2. The operator L^* is called the **formal adjoint** of L (since $D(L^*)$ has not yet been specified in order to determine the boundary terms in Green's formula).

TODO: define BVP or IVP

Definition 7.4.3. Consider a BVP or IVP involving $(L, D(L))$. The adjoint $(L^*, D(L^*))$ is given by L^* , defined in Green's formula and its domain $D(L^*)$ consists of all functions v whose boundary conditions ensure that the boundary terms in Green's formula vanish, i.e.

$$[\overline{p_0}(v\overline{u}' - v'\overline{u}) + (\overline{p_1} - \overline{p_0}')uv] = 0$$

If this is true, then $\langle L(u), v \rangle - \langle u, L^*(v) \rangle = 0$ by Green's formula and $(L^*, D(L^*))$ is the adjoint of $(L, D(L))$.

Definition 7.4.4. The BVP $L(u) = f$, $B_1(u) = B_2(u) = 0$ is called **self-adjoint** if $L^* = L$ and $D(L^*) = D(L)$.

Example 7.4.5. Let $L = \frac{\partial^2}{\partial x^2}$, $L(u) = f$, $a < x < b$, $B_1(u) := u(a) = 0$, $B_2(u) := u(b) = 0$ (Dirichlet boundary conditions). Then

$$D(L) = \{u : u \in C^2([a, b]) : u(a) = u(b) = 0\}$$

Also,

$$L^* := \overline{p_0}(x)d_x^2 + (2\overline{p_0}'(x) - \overline{p_1}(x))d_x + (\overline{p_0}''(x) - \overline{p_1}'(x) + \overline{p_2}(x))$$

Here, $p_0 = 1, p_1 = p_2 = 0, \overline{p_0} = p_0, p_0' = 0$. So

$$L^* = d_x^2 = L$$

so L is formally self-adjoint. By Green's formula, the boundary terms are

$$\begin{aligned} [\overline{p_0}(v\overline{u}' - v'\overline{u}) + (\overline{p_1} - \overline{p_0}')v\overline{u}]_a^b &= [v\overline{u}' - v'\overline{u}]_a^b \\ &= v(b)\overline{u}'(b) - v'(b)\overline{u}(b) - v(a)\overline{u}'(a) + v'(a)\overline{u}(a) \\ &= v(b)\overline{u}'(b) - v(a)\overline{u}'(a) \\ &= 0 \text{ iff } v(a) = v(b) = 0 \end{aligned}$$

since $u(a) = 0 \implies \overline{u}(a) = 0$ and $u(b) = 0 \implies \overline{u}(b) = 0$. So $D(L^*) = D(L)$ and the BVP is self-adjoint.

$$L = p_0 d_x^2 + p_1 d_x + p_2, \quad L^* = \overline{p_0} d_x^2 + (2\overline{p_0}' - \overline{p_1})d_x + (\overline{p_0}'' - \overline{p_1}' + \overline{p_2})$$

so $L = L^*$ if $p_0 = \overline{p_0}$ (so p_0 is a real-valued function), $p_1 = 2\overline{p_0}' - \overline{p_1}$, and $p_2 = \overline{p_0}'' - \overline{p_1}' + \overline{p_2}$. So $p_1 + \overline{p_1} = 2\text{Re}(p_1) = 2p_0'$. Also $p_2 - \overline{p_2} = 2\text{Im}(p_2) = p_0'' - \overline{p_1}'$ so $2\text{Im}(p_2) = \text{Im}(p_1')$.

Hence $\{p_0, \text{Im}(p_0), \text{Re}(p_2)\}$ characterises a formally self-adjoint operator.

Definition 7.4.6. Let L be a second-order linear operator with real valued coefficients. If $p'_0 = p_1$ then

$$L^* = p_0 d_x^2 + p'_0 d_x + p_2 = \frac{d}{dx} \left(p_0 \frac{d}{dx} \right) + p_2 = L$$

Operators of this form are called **Sturm-Liouville** operators.

Example 7.4.7. (Problems class) Let

$$L = \frac{d}{dx} \left(r(x) \frac{d}{dx} \right) + s(x)$$

be a differential operator on $[a, b]$, with $r(x) \in C^1([a, b])$ and $s \in C^0([a, b])$ and r and s real-valued functions. Assume that $r(x) \geq c$ on $[a, b]$ for some constant c . Consider the BVP

$$L(u(x)) = f(x), \quad x \in (a, b), u \in C^2([a, b])$$

where $f(x)$ is a source term and sufficiently smooth. The boundary conditions are $B_1(u) := u'(a) - \alpha u(a) = 0$ and $B_2(u) = u'(b) + \beta u(b) = 0$ for some constants $\alpha, \beta \in \mathbb{R}$. Find the adjoint boundary conditions $B_1^*(v) = 0$ and $B_2^*(v) = 0$ and determine whether the BVP is self-adjoint.

A BVP is self-adjoint if $L = L^*$ (L is formally self-adjoint) and $D(L) = D(L^*)$. L is a Sturm-Liouville operator so $L = L^*$. $D(L) = \{u \in C^2([a, b]) : u'(a) - \alpha u(a), u'(b) + \beta u(b) = 0\}$. Using Green's formula,

$$\begin{aligned} \langle L(u), v \rangle - \langle u, L^*(v) \rangle &= (\langle L(u), v \rangle - \langle u, L(v) \rangle) \\ &= [\overline{p_0}(v\overline{u}' - v'\overline{u}) + (\overline{p_1} - \overline{p_0}')v\overline{u}]_a^b \\ &= [r(x)v(x)u'(x) - r(x)v'(x)\overline{u}(x)]_a^b \\ &= r(b)v(b)\overline{u}'(b) - r(b)v'(b)\overline{u}(b) - r(a)v(a)\overline{u}'(a) + r(a)v'(a)\overline{u}(a) \end{aligned}$$

Now, $\overline{u}'(a) = \alpha \overline{u}(a)$ (since $u'(a) = \alpha u(a)$ and $\alpha \in \mathbb{R}$) and $\overline{u}'(b) = -\beta \overline{u}(b)$ (since $u'(b) = -\beta u(b)$ and $\beta \in \mathbb{R}$). So we get the boundary terms

$$\begin{aligned} &r(b)v(b)(-\beta \overline{u}(b)) - r(b)v'(b)\overline{u}(b) - r(a)v(a)(\alpha \overline{u}(a)) + r(a)v'(a)\overline{u}(a) \\ &= r(b)\overline{u}(b)(-\beta v(b) - v'(b)) + r(a)\overline{u}(a)(v'(a) - \alpha v(a)) \\ &= 0 \text{ iff } v'(b) = -\beta v(b) \text{ and } v'(a) = \alpha v(a) \end{aligned}$$

Hence the adjoint boundary conditions are

$$B_1^*(v) = v'(a) - \alpha v(a) = 0 \text{ and } B_2^*(v) = v'(b) + \beta v(b) = 0$$

and

$$D(L^*) = D(L) = \{v \in C^2([a, b]) : B_1(v) = 0, B_2(v) = 0\}$$

So the BVP is self-adjoint.

Example 7.4.8. (Problems class) Is the operator

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

where v is a constant and $-1 \leq x \leq 1$ formally self-adjoint of Sturm-Liouville type? If not, convert L into a formally self-adjoint operator of Sturm-Liouville type.

L is not a Sturm-Liouville operator since $p_0 = 1 - x^2$, $p_1 = -3x \neq p_0'$. We can convert L into a Sturm-Liouville operator by multiplying L by a factor

$$\rho(x) = \frac{1}{p_0(x)} \exp \left(\int \frac{p_1(x)}{p_0(x)} \right) = \frac{1}{1 - x^2} \exp \left(\int \frac{-3x}{1 - x^2} \right) = (1 - x^2)^{1/2}$$

so

$$(\rho L)(x) = (1 - x^2)^{3/2} \frac{d^2}{dx^2} - 3x(1 - x^2)^{1/2} \frac{d}{dx} + v(v + 2)(1 - x^2)^{1/2}$$

Remark. More generally, for an operator

$$L = p_0 \frac{d^2}{dx^2} + p_1 \frac{d}{dx} + p_2$$

we have

$$\rho L = \rho p_0 \frac{d^2}{dx^2} + \rho p_1 \frac{d}{dx} + \rho p_2$$

The adjoint of L is

$$L^* = \overline{p_0} \frac{d^2}{dx^2} + (2\overline{p_0}' - \overline{p_1}) \frac{d}{dx} + \overline{p_0}'' - \overline{p_1}' + \overline{p_2} = p_0 \frac{d^2}{dx^2} + (2p_0' - p_1) \frac{d}{dx} + p_0'' - p_1' + p_2$$

since p_i are real-valued. Hence, with $p_i \rightarrow \rho p_i$,

$$(\rho L)^* = \rho p_0 \frac{d^2}{dx^2} + (2(\rho p_0)' - (\rho p_1)) \frac{d}{dx} + (\rho p_0)'' - (\rho p_1)' + \rho p_2 = \rho L$$

So $2(\rho p_0)' - \rho p_1 = \rho p_1$ so

$$\int \frac{\rho'}{\rho} dx = \int \frac{p_1 - p_0'}{p_0} dx$$

which gives

$$\rho = \frac{1}{p_0} \exp \left(\int \frac{p_1}{p_0} dx \right)$$

Example 7.4.9. Let $L = d_x^2 - 2xd_x - 2\alpha$, where α is a constant. Here, $p_0 = 1$, $p_1 = -2x \neq 0 = p_0'$ so not L is not Sturm-Liouville. So

$$\rho = \frac{1}{1} \exp \left(\int \frac{-2x}{1} dx \right) = \exp(-x^2)$$

gives ρL is a Sturm-Liouville operator.

7.5 Eigenfunction method for differential operators

Definition 7.5.1. A function u_n satisfying

$$Lu_n(x) = \lambda_n u_n(x)$$

for a differential operator L is called an **eigenfunction of L** and λ_n is called an **eigenvalue** associated with u_n .

Remark. Eigenfunctions are unique up to scalar multiplication.

Definition 7.5.2. A function u_n satisfying

$$Lu_n(x) = \lambda_n w(x) u_n(x)$$

for a differential operator L and a weight function $w(x)$ is called an **eigenfunction of L with respect to w** , with eigenvalue λ_n .

Example 7.5.3. The heat flow in a non-uniform rod (with no internal heat source) can be modelled by the BVP

$$\begin{aligned} U_t - \frac{1}{w(x)}((p_0(x)U_x)_x + p_2(x)U) &= 0, \quad a < x < b, t > 0 \\ B_1(U) = \alpha_1 U(a, t) + \beta_1 U_x(a, t) &= 0, \quad t \geq 0 \\ B_2(U) = \eta_2 U(b, t) + \kappa_2 U_x(b, t) &= 0, \quad t \geq 0 \\ U(x, 0) &= f(x) \end{aligned}$$

To solve this PDE, we use the method of **separation of variables**: let

$$U(x, t) = T(t)u(x)$$

Substituting this into the BVP gives

$$\frac{T'(t)}{T(t)} = \frac{1}{w(x)u(x)}(p_0(x)d_x^2 + p'_0(x)d_x + p_2(x))u(x) \frac{(p_0(x)u'(x))_x + p_2(x)u(x)}{w(x)u(x)}$$

Note that $p_0(x)d_x^2 + p'_0(x)d_x + p_2(x)$ is a Sturm-Liouville operator. Now the LHS depends only on t and the RHS depends only on x . Hence for some constant λ ,

$$\frac{T_t}{T} = \frac{(p_0 u_x)_x + p_2 u}{w u} = -\lambda$$

This gives the ODES

$$(p_0 u')_x + p_2 u + \lambda w u = 0, \quad a < x < b, T' = -\lambda T$$

By assumption, $U \neq 0$,

$$B_1(U) = 0 = B_1(u) \quad \text{and} \quad B_2(U) = 0 = B_2(u)$$

Definition 7.5.4. The BVP $\mathcal{L}u + \lambda w u = 0$, $B_1(u) = B_2(u) = 0$, where \mathcal{L} is a Sturm-Liouville operator is called a **Sturm-Liouville eigenvalue problem**.

Remark. A Sturm-Liouville operator given by $L := -\mathcal{L}$, where $-\mathcal{L}$ is an SL operator, then Green's formula becomes

$$\langle -\mathcal{L}u, v \rangle - \langle u, -\mathcal{L}v \rangle = -[\overline{p_0}(v\overline{u}') - v'\overline{u}]_a^b$$

If the problem is restricted to the subspace

$$\mathcal{H} := \{g \in C^2([a, b]) \subset L^2(w, [a, b]) : B_1(g) := \alpha_1 g(a) + \beta_1 g'(a) = 0, B_2(g) := \eta_2 g(b) + \kappa_2 g'(b) = 0\} \\ |\alpha_1| + |\beta_1| > 0, |\eta_2| + |\kappa_2| > 0, \alpha_1, \beta_1, \eta_2, \kappa_2 \in \mathbb{R}$$

Then $D(\mathcal{L}) = D(-\mathcal{L})$ and $(-\mathcal{L}, D(-\mathcal{L}))$.

Definition 7.5.5. A **regular Sturm-Liouville (SL) eigenvalue problem** is a BVP on a finite interval $[a, b]$ given by

$$\begin{aligned}\mathcal{L}u + \lambda wu &= 0, \quad a < x < b, \lambda \in \mathbb{C} \\ B_1(u) &= \alpha_1 u(a) + \beta_1 u'(a) = 0 \\ B_2(u) &= \eta_2 u(b) + \kappa_2 u'(b) = 0 \\ \alpha_1, \beta_1, \eta_2, \kappa_2 &\in \mathbb{R}, |\alpha_1| + |\beta_1| > 0, |\eta_2| + |\kappa_2| > 0\end{aligned}$$

where \mathcal{L} is a Sturm-Liouville operator with $p_0 \in C^1([a, b])$, $p_2 \in C^0([a, b])$, $\forall x \in [a, b]$, $p_0(x) > 0$ and $w(x) > 0$.

Solving this BVP is equivalent to finding the set of eigenvalues $\{\lambda_n\}$ of the SL operator \mathcal{L} with respect to w , and the corresponding eigenfunctions $\{u_n\}$.

Example 7.5.6. Consider the BVP on $[a, b]$ given by

$$\mathcal{L}u + \lambda u = 0, \quad a < x < b, u'(a) = 0 = u'(b)$$

where $\mathcal{L} = d_x^2$. The boundary conditions $\alpha_1 = \eta_2 = 0$ and $\beta_1 = \kappa_2 = 1$ are Neumann boundary conditions.

Definition 7.5.7. A Sturm-Liouville eigenvalue problem is called **singular** if either

- Either p_0 or w vanish at either a or b .
- p_0 or w is not continuous.
- The problem is defined on an infinite interval.

Example 7.5.8. The Bessel equation of order ν is given by

$$r^2 y''(r) + r y'(r) + (r^2 - \nu^2) y(r) = 0, \quad r > 0, y(r)$$

where ν is a constant. This is a singular SL eigenvalue problem on $[0, L]$. To show this, let $r = \sqrt{\lambda}x$ then let $u(x) = y(r) = \sqrt{\lambda}y(x)$. Then

$$\begin{aligned}u'(x) &= \frac{d}{dx}y(r) = \frac{d}{dr}y(r) \cdot \frac{dr}{dx} = \sqrt{\lambda}y'(r) \\ &= u''(x) = \frac{d^2}{dx^2}y(r) = \lambda y''(r)\end{aligned}$$

This gives the Bessel equation as

$$x^2 u''(x) + x u'(x) + (\lambda x^2 - \nu^2) u(x) = 0$$

And since $x \neq 0$ (since $x \in [0, L]$), dividing by x gives

$$x d_x^2 + d_x - \frac{\nu^2}{x} + \lambda x u = 0$$

Here, the weight function is $w(x) = x$. If we were given boundary conditions, for example the Dirichlet boundary conditions $u(L) = 0$ and $|u(0)| \leq \infty$, then we obtain a Sturm-Liouville BVP. It is singular, since $p_0(x) = w(x)$ vanishes at $x = 0$ which is an endpoint of $[0, L]$.

Definition 7.5.9. An interval $[a, b]$ is called the **natural interval** of a BVP if $p_0(a) = p_0(b) = 0$ and $p_0(x) > 0$ for every $x \in (a, b)$. In this case, $\langle (-\mathcal{L}u, v) \rangle_w = \langle u, (\mathcal{L})v \rangle_w$, no matter what the boundary conditions on u and v are.

Example 7.5.10. The Hermite equation is defined as

$$Lu = 0, \quad L = d_x^2 - 2xd_x + 2\nu$$

where $x \in \mathbb{R}$ and $\nu \in \mathbb{R}$ is a constant. L is not an SL operator, but multiplying L by

$$\rho(x) = \exp\left(\int -2xdx\right) = e^{-x^2}$$

which makes

$$\mathcal{L} = \rho L = e^{-x^2} d_x^2 - 2xe^{-x^2} d_x + 2\nu e^{-x^2}$$

an SL operator. Then the Hermite equation becomes

$$\mathcal{L}u = d_x(e^{-x^2} d_x)u + 2\nu e^{-x^2} u = 0$$

Define

$$\tilde{\mathcal{L}} = d_x(e^{-x^2} d_x)$$

which is also an SL operator. Then

$$\tilde{\mathcal{L}}u + 2\nu e^{-x^2} u = 0$$

so here the eigenvalue is $\lambda = 2\nu$ and the weight function is $w(x) = e^{-x^2}$. $\lim_{x \rightarrow \pm\infty} p_9(x) = 0$ and $p_0(x) > 0$ on $(-\infty, \infty)$ so the natural interval is $[-\infty, \infty]$.

Definition 7.5.11. A periodic Sturm-Liouville eigenvalue problem on $[a, b]$ is defined as

$$\mathcal{L}u(x) + \lambda w(x)u(x) = 0, \quad a < x < b, u(a) = u(b), u'(a) = u'(b)$$

where $L = d_x(p_0 d_x) + p_2$ and p_0, p_2, w are periodic with period $b - a$.

Example 7.5.12. Consider the periodic Sturm-Liouville eigenvalue problem on $[0, L]$ given by

$$\mathcal{L}u + \lambda u = 0, \quad 0 < x < L, u(0) = u(L), u'(0) = u'(L)$$

where $\mathcal{L} = d_x^2$ and restrict the problem to the subspace of functions

$$\mathcal{H} := \{g : g \in C^2([0, L]), g(0) = g(L), g'(0) = g'(L)\}$$

which means that $(-\mathcal{L}, D(-\mathcal{L}))$ is self-adjoint and so has real eigenvalues. Let $u_n(x) = e^{\mu x}$, then $\mu^2 + \lambda = 0$ so $\mu = \pm\sqrt{-\lambda}$.

1. If $\lambda < 0$, then $u_n(x) = A_n e^{\sqrt{-\lambda}x} + B_n e^{-\sqrt{-\lambda}x} = (A_n + B_n) \cosh(\sqrt{-\lambda}x) + (A_n - B_n) \sinh(\sqrt{-\lambda}x)$ is an eigenfunction.
2. If $\lambda = 0$, then $u_n(x) = A_n x + B_n$ is an eigenfunction.
3. If $\lambda > 0$, then $u_n(x) = (\alpha_n + \beta_n) \cos(\sqrt{\lambda}x) + i(\alpha_n - \beta_n) \sin(\sqrt{\lambda}x)$ is an eigenfunction, with $\alpha, \beta \in \mathbb{C}$.

Now we impose the boundary conditions:

1. If $u(0) = u(L)$, then $A + B = A e^{L\sqrt{-\lambda}} + B e^{-L\sqrt{-\lambda}}$ and $u'(x) = A\sqrt{\lambda}e^{x\sqrt{-\lambda}} - B\sqrt{\lambda}e^{-x\sqrt{-\lambda}}$. $u'(0) = u'(L)$ so $A - B = A e^{L\sqrt{-\lambda}} - B e^{-L\sqrt{-\lambda}}$. So $A = B = 0$ gives the trivial solution $u(x) = 0$.

2. If $u(0) = u(L)$, then $A + B = AL + B$ so $A = 0$. So $u(x) = B$.
3. If $u(0) = u(L)$, then $A = A \cos(L\sqrt{\lambda}) + B \sin(L\sqrt{\lambda})$ and $u'(x) = -A\sqrt{\lambda} \sin(x\sqrt{\lambda}) + B\sqrt{\lambda} \cos(x\sqrt{\lambda})$. $u'(0) = u'(L)$. $u'(0) = u'(L)$ so $B = -A \sin(L\sqrt{\lambda}) + B \cos(L\sqrt{\lambda})$. Hence $A^2 + B^2 = (A^2 + B^2) \cos(L\sqrt{\lambda})$ so $\cos(L\sqrt{\lambda}) = 1$ so $\lambda = (2n\pi/L)^2$ for $n \in \mathbb{N}$. Hence $u(x) = A \cos(\frac{2\pi n}{L}x) + B \sin(\frac{2\pi n}{L}x)$ for $n \in \mathbb{N}$.

This means for a fixed eigenvalue $\lambda_n = (2n\pi/L)^2$, we have two linearly independent solutions: $u_n(x) = \cos(\frac{2\pi n}{L}x)$ and $u_m(x) = \sin(\frac{2\pi m}{L}x)$.

Example 7.5.13. (Problems class) Expand the function $f(x) = x$ for $0 \leq x \leq 1$ in terms of the normalised eigenfunctions $\hat{u}_n(x)$ of the Sturm-Liouville eigenvalue problem $u''(x) + \lambda u(x) = 0$, $0 < x < 1$ with the boundary conditions $u(0) = 0$ and $u'(1) + u(1) = 0$.

First we find $\{\lambda_n\}$ and $\{\hat{u}_n\}$ from $u''_n + \lambda_n u_n = 0$, $u_n(0) = 0$, $u'_n(1) + u_n(1) = 0$. We have $\mathcal{L} = d_x^2$, with weight function $w(x) = 1$. The characteristic equation is $m^2 + \lambda_n = 0$ so $m = \pm\sqrt{-\lambda_n}$.

- If $\lambda_n < 0$, there are no non-trivial solutions.
- If $\lambda_n = 0$, there are no non-trivial solutions.
- If $\lambda_n > 0$, then $u_n(x) = A_n \cos(\sqrt{\lambda_n}x) + B_n \sin(\sqrt{\lambda_n}x)$ is a solution. $u_n(0) = 0$ so $A_n = 0$. $u'_n(x) = B_n \sqrt{\lambda_n} \cos(\sqrt{\lambda_n}x)$ so $u'_n(1) + u_n(1) = B_n \sqrt{\lambda_n} \cos(\sqrt{\lambda_n}) + B_n \sin(\sqrt{\lambda_n}) = 0$. If $B_n \neq 0$, we need $\cos(\sqrt{\lambda_n}) + \sin(\sqrt{\lambda_n}) = 0$. Let $\lambda_n = k_n^2$ and rewrite the boundary conditions as $-k_n = \tan(k_n)$ so there are an infinite number of eigenvalues.

To find the normalised eigenfunctions,

$$\begin{aligned}
\langle u_n, u_n \rangle &= B_n^2 \int_0^1 \sin(\sqrt{\lambda_n}x)^2 dx = B_n^2 \frac{1}{2} \int_0^1 (1 - \cos(2\sqrt{\lambda_n}x)) dx \\
&= \frac{1}{2} B_n^2 \left(1 - \frac{1}{2\sqrt{\lambda_n}} 2 \sin(\lambda_n) \cos(\lambda_n) \right) \\
&= \frac{1}{2} B_n^2 (1 + \cos(\sqrt{\lambda_n})^2) \\
&\Rightarrow \hat{u}_n(x) = \sqrt{\frac{2}{1 + \cos(\sqrt{\lambda_n})^2}} \sin(\sqrt{\lambda_n}x)
\end{aligned}$$

Now we expand f in the basis $\{\hat{u}_n\}$. Let us denote the generalised Fourier expansion of f as

$$F_f(x) := \sum_{n=1}^{\infty} f_n \hat{u}_n(x)$$

where f_n are the generalised Fourier coefficients, given by

$$\begin{aligned}
f_n &= \langle f, \hat{u}_n \rangle_{w=1} \\
&= \int_0^1 f(y) \hat{u}_n(y) dy \\
&= \int_0^1 y B_n \sin(\sqrt{\lambda_n} y) dy \\
&= B_n \left[-\frac{y}{\sqrt{\lambda_n}} \cos(\sqrt{\lambda_n} y) \right]_0^1 + \int_0^1 1 \cdot \frac{\cos(\sqrt{\lambda_n} y)}{\sqrt{\lambda_n}} dy \\
&= -\frac{1}{\sqrt{\lambda_n}} B_n \cos(\sqrt{\lambda_n}) + \left[\frac{B_n}{\sqrt{\lambda_n}} \sin(\sqrt{\lambda_n} y) \right]_0^1 \\
&= \frac{2}{\lambda_n} \left(\frac{2}{1 + \cos(\sqrt{\lambda_n})^2} \right)^{1/2} \sin(\sqrt{\lambda_n})
\end{aligned}$$

So

$$F_f(x) = \sum_{n=1}^{\infty} \frac{4}{\lambda_n (1 + \cos(\sqrt{\lambda_n})^2)} \sin(\sqrt{\lambda_n} x) \sin(\sqrt{\lambda_n})$$

Note f is continuous with $f'(x) = 1$. Then $F_f(x)$ converges to $\frac{1}{2}(f(x^+) + f(x^-)) = \frac{1}{2} \lim_{\epsilon \rightarrow 0} (f(x + \epsilon) + f(x - \epsilon)) = f(x)$. However, it does not converge uniformly on $[0, 1]$ since $f(0) = 0$ but $f'(1) + f(1) = 1 + 1 \neq 0$.

7.6 Spectral properties of regular Sturm-Liouville operators

Proposition 7.6.1. The eigenvalues λ_n of the regular Sturm-Liouville eigenvalue problem are real.

Proof. Let u be a non-zero eigenfunction with eigenvalue λ , so $\mathcal{L}u + \lambda wu = 0$. Then

$$\lambda \langle u, u \rangle_w = \langle u, \lambda u \rangle_w = \langle u, \lambda wu \rangle = -\langle u, \mathcal{L}u \rangle = -\langle \mathcal{L}u, u \rangle = \langle \lambda wu, u \rangle = \bar{\lambda} \langle u, u \rangle_w$$

□

Proposition 7.6.2. The eigenfunctions u_m and u_n corresponding to distinct eigenvalues λ_m and λ_n are orthogonal with respect to the inner product.

Proof. By Proposition 7.6.1, λ_m and λ_n are real, hence

$$\lambda_n \langle u_m, u_n \rangle_w = \langle u_m, \lambda_n u_n \rangle_w = -\langle u_m, \mathcal{L}u_n \rangle = -\langle \mathcal{L}u_m, u_n \rangle = \lambda_m \langle u_m, u_n \rangle_w$$

and $\lambda_m \neq \lambda_n$ so $\langle u_m, u_n \rangle = 0$.

□

Proposition 7.6.3. The following properties hold for regular Sturm-Liouville problems:

1. The eigenfunctions of a regular or periodic Sturm-Liouville problem can be chosen to be real.
2. The eigenvalues have multiplicity one.
3. There exists a smallest eigenvalue but not a largest one. The set of eigenvalues is countably infinite and is a monotonically increasing sequence, so

$$\lambda_1 < \lambda_2 < \cdots < \lambda_n < \cdots$$

and $\lim_{n \rightarrow \infty} \lambda_n = \infty$.

Proposition 7.6.4. The orthonormal set $\{\hat{u}_n\}_{n=0}^\infty$ of all eigenfunctions of a regular Sturm-Liouville eigenvalue problem is complete in the inner product space $E(w, [a, b])$ of all piecewise continuous functions on $[a, b]$, with inner product $\langle \cdot, \cdot \rangle_w$. So

$$\lim_{k \rightarrow \infty} \|u - \sum_{n=0}^k \langle u, \hat{u}_n \rangle_w \hat{u}_n\|_w = 0 \quad \forall g \in E(w, [a, b])$$

We write

$$g(x) = \sum_{n=0}^{\infty} \langle \hat{u}_n, g \rangle_w \hat{u}_n(x), \quad \hat{u}_n \in C^2([a, b])$$

Proposition 7.6.5. Let $\{\hat{u}_n\}_{n=0}^\infty$ be an orthonormal set of eigenfunctions of a regular Sturm-Liouville eigenvalue problem.

1. Let f be piecewise smooth on $[a, b]$. Then for every $x \in (a, b)$, the eigenfunction expansion of f with respect to $\{\hat{u}_n\}_{n=0}^\infty$ converges to

$$\frac{1}{2}(f(x_+) + f(x_-)) := \frac{1}{2} \left(\lim_{\epsilon \rightarrow 0} f(x + \epsilon) + \lim_{\epsilon \rightarrow 0} f(x - \epsilon) \right)$$

So we have

$$\frac{1}{2}(f(x_+) + f(x_-)) = \sum_{n=0}^{\infty} \langle \hat{u}_n, f \rangle_w \hat{u}_n(x)$$

2. Let f be continuous and f' be piecewise continuous. If f satisfies the boundary conditions of the given Sturm-Liouville eigenvalue problem, then the eigenfunction expansion of f with respect to $\{\hat{u}_n\}_{n=0}^\infty$ converges uniformly to f on $[a, b]$.

Example 7.6.6. Let $\mathcal{L}u + \lambda u = 0$, $x \in (0, \pi)$, $u(0) = u(\pi) = 0$, where $\mathcal{L} = \frac{d^2}{dx^2}$ (assume $(-\mathcal{L}, D(-\mathcal{L}))$). Find the eigenvalues λ_n and eigenfunctions $u_n(x)$.

We have $u_n''(x) + \lambda_n u_n(x) = 0$, so the characteristic equation is $m^2 + \lambda_n = 0$ so $m = \pm \sqrt{-\lambda_n}$. The eigenfunctions are therefore

$$u_n(x) = Ae^{\sqrt{-\lambda_n}x} + Be^{-\sqrt{-\lambda_n}x}$$

There are only non-trivial solutions when $\lambda_n > 0$. Since $u_n(0) = u_n(\pi) = 0$, $A = 0$ and $\lambda_n = n^2$. So the eigenfunctions are $u_n(x) = B_n \sin(nx)$. Now we find the normalised eigenfunctions \hat{u}_n :

$$\langle u_n, u_n \rangle = \int_0^\pi B_n^2 \sin^2(nx) dx = 1 = B_n^2 \frac{\pi}{2}$$

so $B_n = \sqrt{2/\pi}$. We can therefore write the constant function $f(x) = 1$ as

$$\begin{aligned} f(x) = 1 &= \sum_{n=1}^{\infty} \langle 1, \hat{u}_n \rangle \hat{u}_n(x) \\ &= \sum_{n=1}^{\infty} \int_0^\pi 1 \cdot \sqrt{\frac{2}{\pi}} \sin(nx) dx \hat{u}_n(x) \\ &= \sum_{n=1}^{\infty} \sqrt{\frac{2}{\pi}} \left[\frac{-\cos(nx)}{n} \right]_0^\pi \hat{u}_n(x) \\ &= \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{1 - (-1)^n}{n} \sin(nx) \end{aligned}$$

f is constant on $[0, \pi]$, is piecewise smooth and $\forall x \in (0, \pi)$, the expansion of f converges to $\frac{1}{2}(f(x_+) + f(x_-)) = \frac{1}{2}(1 + 1) = 1$. However, the convergence is not uniform as f does not satisfy the boundary conditions $f(0) = f(\pi) = 0$.

Proposition 7.6.7. (Completeness of eigenfunctions) Given a Sturm-Liouville eigenvalue problem with weight w , we can expand any function $f(x)$ as an eigenfunction series

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} \langle f, \hat{u}_n \rangle_w \hat{u}_n(x) \\ &= \sum_{n=0}^{\infty} \left(\int_a^b f(y) \hat{u}_n(y) w(y) dy \hat{u}_n(x) \right) \\ &= \int_a^b \left(\sum_{n=0}^{\infty} \hat{u}_n(y) \hat{u}_n(x) w(y) \right) f(y) dy \\ &= w(x) \left(\int_a^b f(y) \hat{u}_n(y) dy \right) \hat{u}_n(x) \end{aligned}$$

Since this expansion holds for every f ,

$$\sum_{n=0}^{\infty} \hat{u}_n(y) \hat{u}_n(x) w(y) = \delta(x - y)$$

by the sifting property of the δ distribution.

8 Green's functions

8.1 The importance of Green's functions

Definition 8.1.1. The impulse of a force, $f(t)$, over the interval $[\tau, \tau + \Delta\tau]$ is given by

$$\int_{\tau}^{\tau+\Delta\tau} f(t)dt$$

Definition 8.1.2. We can describe an instantaneous force at $t = \tau$ as the distribution $\delta(t - \tau)$, since

$$\lim_{\Delta\tau} \int_{\tau}^{\tau+\Delta\tau} \delta(t - \tau)dt = 1$$

Definition 8.1.3. Consider the equation

$$Lx(t) := x''(t) + w^2x(t) = f(t), \quad x(0) = x'(0) = 0 \leq t < \infty$$

We can solve this equation for $x(t)$ by viewing $f(t)$ has a continuous sum of successive instantaneous forces. Instead of solving $Lx(t) := x''(t) + w^2x(t) = f(t)$, we solve

$$\frac{\partial^2 G(t, \tau)}{\partial t^2} + w^2 G(t, \tau) = \delta(t - \tau), \quad G(0, \tau) = \frac{\partial G(0, \tau)}{\partial t} = 0$$

where $G(t, \tau)$ some function called the **Green's function for the dynamical system**. G can be interpreted as the response of the system at time t to a unit impulse at τ .

Remark. $\partial_t G(t, \tau)$ should have a jump discontinuity at $t = \tau$.

Proposition 8.1.4. The solution to the equation in Definition 8.1.3 is given as

$$x(t) = \int_0^\infty f(\tau)G(t, \tau)d\tau$$

Proof. We can write $f(t)$ as a continuous sum of impulses as follows:

$$f(t) = \int_0^\infty f(\tau)\delta(\tau - t)d\tau$$

(by the sifting property of δ). Check by inserting $x(t)$ into the original equation:

$$\begin{aligned} Lx(t) &= \frac{d^2}{dt^2}x(t) + w^2x(t) = \int_0^\infty \left(\frac{d^2}{dt^2}G(t, \tau)d\tau + w^2 \int_0^\infty G(t, \tau) \right) f(\tau) \\ &= \int_0^\infty \delta(t - \tau)f(\tau)d\tau = f(t) \end{aligned}$$

□

Example 8.1.5. Consider the PDE

$$\Delta^2 \phi(\underline{r}) + k^2 \phi(\underline{r}) = \rho(\underline{r})$$

where $\rho(\underline{r})$ is a source function. This PDE is associated to electrostatic problems in 3D. The Green's function $G(\underline{r}, \underline{r}')$ is the potential at \underline{r} due to a point charge at \underline{r}' , represented by $\delta(\underline{r} - \underline{r}')$. G satisfies

$$\Delta^2 G(\underline{r}, \underline{r}') + k^2 G(\underline{r}, \underline{r}') = \delta(\underline{r} - \underline{r}')$$

Example 8.1.6. Green's functions can be used to model partial/wave scattering. For a particle with energy E and mass m , we have

$$-\frac{\hbar^2}{2m}\Delta^2\psi(r) + V(r)\psi(r) = E\psi(r)$$

This gives a solution

$$\psi(r) = \psi_0(r) + \int_V G(r, r') \frac{2m}{\hbar^2} V(r') \psi_0(r') dV'$$

8.2 Green's functions for ODEs

Definition 8.2.1. An **initial value problem (IVP)** is defined by the following equation and boundary conditions:

$$Lu(t) = f(t), \quad u(a) = \gamma_1, u'(a) = \gamma_2$$

Definition 8.2.2. An **IN/IN IVP** is an IVP of the form $Lu(t) = f(t)$, $u(a) = \gamma_1 \neq 0, u'(a) = \gamma_2 \neq 0$.

Definition 8.2.3. An **IN/HOM IVP** is an IVP of the form $Lu(t) = f(t)$, $u(a) = u'(a) = 0$.

Definition 8.2.4. A **HOM/IN IVP** is an IVP of the form $Lu(t) = 0$, $u(a) = \gamma_1 \neq 0, u'(a) = \gamma_2 \neq 0$.

Definition 8.2.5. A **boundary value problem (BVP)** is defined by

$$Lu(x) = f(x)$$

with boundary conditions

$$\begin{aligned} B_1(u) &= \alpha_1 u(a) + \beta_1 u'(a) + \eta_1 u(b) + \kappa_1 u'(b) = \gamma_1 \\ B_2(u) &= \alpha_2 u(a) + \beta_2 u'(a) + \eta_2 u(b) + \kappa_2 u'(b) = \gamma_2 \end{aligned}$$

If $\gamma_1 = \gamma_2 = 0$, the boundary conditions are called **homogenous**. There are similar definitions for IN/IN, IN/HOM, and HOM/IN BVPs.

Definition 8.2.6. The **Wronskian** of two differentiable functions u_1 and u_2 is defined as

$$W(u_1, u_2)(x) := \begin{vmatrix} u_1(x) & u_2(x) \\ u_1'(x) & u_2'(x) \end{vmatrix} = u_1(x)u_2'(x) - u_1'(x)u_2(x)$$

Theorem 8.2.7. Let $u_1, u_2 \in C^1([a, b])$ and $W(u_1, u_2)(x_0) \neq 0$ for some $x_0 \in [a, b]$. Then u_1 and u_2 are linearly independent on $[a, b]$

Theorem 8.2.8. Let u_1 and u_2 be linearly dependent, then the Wronskian $W(u_1, u_2)$ is zero at all points in the domain.

Remark. Note that for arbitrary functions u_1 and u_2 , the Wronskian being zero does not imply that the functions are linearly independent.

Lemma 8.2.9. Let u_1 and u_2 be solutions of the homogenous differential equation $u''(x) + p(x)u'(x) + q(x)u(x) = 0$ on an interval $[a, b]$. Then their Wronskian is either always zero or never zero on $[a, b]$.

Lemma 8.2.10. Let u_1 and u_2 be solutions of the homogenous differential equation $u''(x) + p(x)u'(x) + q(x)u(x) = 0$ on an interval $[a, b]$. Then u_1 and u_2 are linearly independent if and only if $W(u_1, u_2) \neq 0$ on $[a, b]$.

Theorem 8.2.11. Let $u_p(x)$ be a particular solution of the inhomogenous second order linear ODE $Lu(x) = f(x)$ and let $\{u_1, u_2\}$ be a basis for the vector space of solutions of the homogenous equation $Lu(x) = 0$. Then the general solution of $Lu(x) = f(x)$ is given by

$$u_g(x) = u_p(x) + c_1u_1(x) + c_2u_2(x)$$

where c_1 and c_2 are constants.

Theorem 8.2.12. (Existence and uniqueness of solutions to an IVP) If the function p_0, p_1, p_2 and f satisfy the conditions (lecture notes) on an interval $I \subseteq \mathbb{R}$ containing a , then there exists a unique solution to the IVP in 4.9 (lecture notes) in some neighbourhood of a .

Proposition 8.2.13. (Properties of the Green's function of an IVP)

1. $G(t, \tau)$ is continuous at $t = \tau$:

$$\lim_{\epsilon \rightarrow 0^+} G(\tau + \epsilon, \tau) = \lim_{\epsilon \rightarrow 0^+} G(\tau - \epsilon, \tau)$$

2. $\partial_t G(t, \tau)$ has a jump discontinuity at $t = \tau$:

$$\lim_{\epsilon \rightarrow 0^+} (\partial_t G(\tau + \epsilon, \tau) - \partial_t G(\tau - \epsilon, \tau)) = \frac{1}{p_0(\tau)}$$

Proof. TODO. □

Theorem 8.2.14. The Green's function $G(t, \tau)$ for the IVP

$$L_t G(t, \tau) = \delta(t - \tau), \quad G(0, \tau) = 0, \partial_t G(0, \tau) = 0, \quad L_t := p_0(t)\partial_t^2 + p_1(t)\partial_t + p_2(t)$$

is given by

$$G(t, \tau) = \frac{u_1(\tau)u_2(t) - u_2(\tau)u_1(t)}{p_0(\tau)W(u_1, u_2)(\tau)} \Theta(t - \tau) = \begin{cases} 0 & \text{if } 0 \leq t < \tau \\ \frac{u_1(\tau)u_2(t) - u_2(\tau)u_1(t)}{p_0(\tau)W(u_1, u_2)(\tau)} & \text{if } \tau \leq t \end{cases}$$

where u_1 and u_2 are two linearly independent solutions of $Lu = 0$.

Proof. TODO. □

Example 8.2.15. (Problems class) Consider the BVP

$$((x^2 - 15x + 27)d_x^2 + (2x - 15)d_x + 18x + 9)u = f(x)$$

on $[3, 9]$, $u(9) + \frac{1}{3}u(3) + \frac{1}{3}u'(3) = 0$ and $u'(9) + 2u(3) + 3u'(3) = 0$. Find the boundary conditions of the adjoint operator.

The operator L is Sturm-Liouville so L is formally self-adjoint. To find $D(L^*)$, we want the boundary terms in Green's formula to vanish:

$$\begin{aligned} B(x) &:= [p_0(x)(\bar{u}'(x)v(x) - \bar{u}(x)v'(x))]_3^9 \\ &= [(x^2 - 15x + 27)(\bar{u}'(x)v(x) - \bar{u}(x)v'(x))]_3^9 \\ &= -27(\bar{u}'(9)v(9) - \bar{u}(9)v'(9)) - (-9)(\bar{u}'(3)v(3) - \bar{u}(3)v'(3)) \end{aligned}$$

Now we use the boundary conditions to replace $\bar{u}(9)$ and $\bar{u}'(9)$ with $\bar{u}(9) = -\frac{1}{3}\bar{u}(3) - \frac{1}{3}\bar{u}'(3)$ and $\bar{u}'(9) = -2\bar{u}(3) - 3\bar{u}'(3)$:

$$\begin{aligned} B(x) &= -27[(-2\bar{u}(3) - 3\bar{u}'(3))v(9) + \frac{1}{3}(\bar{u}(3) + \bar{u}'(3))v'(9)] + 9[\bar{u}'(3)v(3) - \bar{u}(3)v'(3)] \\ &= \bar{u}(3)(54v(9) - 9v'(9) - 9v'(3)) + \bar{u}'(3)(81v(9) - 9v'(9) + 9v(3)) = 0 \end{aligned}$$

Hence, since $\bar{u}(3)$ and $\bar{u}'(3)$ are arbitrary, we have

$$\begin{aligned} 6v(9) - v'(9) - v'(3) &= 0 \\ 9v(9) - v'(9) + v(3) &= 0 \end{aligned}$$

First we eliminate $v(9)$, giving $v'(9) = -3v'(3) - 2v(3)$, and then eliminate $v'(9)$, giving $v(9) = \frac{1}{3}v(3) + \frac{1}{3}v'(3)$.

Example 8.2.16. (Problems class) Solve the IN/IN BVP on $[1, 3]$, given by

$$d_x^2 u - \frac{1}{x} d_x u = 16x, \quad u(1) = 1, u(3) = 9$$

First we solve the associated HOM/IN problem:

$$\begin{aligned} u''_{\text{hom}}(x) - \frac{1}{x} u'_{\text{hom}}(x) &= 0, \quad u_{\text{hom}}(1) = 1, u_{\text{hom}}(3) = 0 \\ \implies x^2 u''_{\text{hom}}(x) - x u'_{\text{hom}}(x) &= 0 \end{aligned}$$

which is a Cauchy-Euler equation. Using the variable change $x = e^t$, $u(x) = u(e^t) =: v(t)$, so $u'(x) = d_x v(t) = d_t v(t) \frac{dt}{dx} = \frac{1}{x} v'(t)$, we get

$$\begin{aligned} u''(x) &= d_x u'(x) = d_x \left(\frac{1}{x} v'(t) \right) \\ &= -\frac{1}{x^2} v'(t) + \frac{1}{x} \frac{d}{dt} v'(t) \frac{dt}{dx} \\ &= \frac{1}{x^2} (v''(t) - v'(t)) \\ \implies v''(t) - 2v'(t) &= 0 \end{aligned}$$

The characteristic equation for this is $m^2 - 2m = m(m - 2) = 0$ so $m = 0$ or $m = 2$. Hence

$$v(t) = A + Be^{2t} \implies u_{\text{hom}}(x) = A + Bx^2$$

Now imposing the boundary conditions, $u_{\text{hom}}(1) = 1 = A + B$ and $u_{\text{hom}}(3) = 9 = A + 9B$, so $B = 1$ and $A = 0$, so $u_{\text{hom}}(x) = x^2$. Now we solve the associated IN/HOM BVP:

$$Lu_p(x) = 16x, \quad L = d_x^2 - \frac{1}{x} d_x, \quad u_p(1) = 0, u_p(3) = 0$$

We construct the Green's function $G(x, \xi)$, which satisfies

$$\partial_x^2 G(x, \xi) - \frac{1}{x} \partial_x G(x, \xi) = \delta(x - \xi)$$

and $G(1, \xi) = G(3, \xi) = 0$. So for $1 < x < \xi$,

$$\partial_x^2 G(x, \xi) - \frac{1}{x} \partial_x G(x, \xi) = 0, \quad G(1, \xi) = 0$$

and for $\xi < x < 3$,

$$\partial_x^2 G(x, \xi) - \frac{1}{x} \partial_x G(x, \xi) = 0, \quad G(3, \xi) = 0$$

To construct the ansatz, we need two linear independent solutions of $Lu = 0$, which we have already solved. So we take $u_1(x) = 1$ and $u_2(x) = x^2$. So ansatz 1 is

$$G(x, \xi) = A_1(\xi) \cdot 1 + B_1(\xi) \cdot x^2, \quad G(1, \xi) = 0$$

and ansatz 2 is

$$G(x, \xi) = A_2(\xi) \cdot 1 + B_2(\xi) \cdot x^2, \quad G(3, \xi) = 0$$

$G(1, \xi) = 0 \implies A_1(\xi) = -B_1(\xi)$ and $G(3, \xi) = 0 \implies A_2(\xi) = -9B_2(\xi)$. So

$$G(x, \xi) = \begin{cases} B_1(\xi)(x^2 - 1) & \text{if } 1 < x < \xi \\ B_2(\xi)(x^2 - 9) & \text{if } \xi < x < 3 \end{cases}$$

Using the continuity of $G(x, \xi)$ at $x = \xi$, we have $B_1(\xi)(\xi^2 - 1) = B_2(\xi)(\xi^2 - 9)$ (since $x^2 - 9$ and $x^2 - 1$ are polynomials so continuous).

Now using the jump discontinuity of $\partial_x G(x, \xi)$ at $x = \xi$, since $\partial_x G(x, \xi) = 2xB_1(\xi)$ for $x < \xi$ and $\partial_x G(x, \xi) = 2xB_2(\xi)$ for $x > \xi$,

$$1 = 2\xi B_2(\xi) - 2\xi B_1(\xi)$$

Now we obtain

$$\begin{aligned} B_1(\xi) &= \frac{\xi^2 - 9}{16\xi} = -A_1(\xi) \\ B_2(\xi) &= \frac{\xi^2 - 1}{16\xi} = -\frac{1}{9}A_2(\xi) \end{aligned}$$

So

$$G(x, \xi) = \begin{cases} \frac{9-\xi^2}{16\xi}(1-x^2) & \text{if } 1 < x < \xi \\ \frac{1-\xi^2}{16\xi}(9-x^2) & \text{if } \xi < x < 3 \end{cases}$$

which gives

$$\begin{aligned} u_p(x) &= \int_1^3 G(x, \xi) f(\xi) d\xi \\ &= \int_x^3 (9 - \xi^2)(1 - x^2) d\xi + \int_1^x (1 - \xi^2)(9 - x^2) d\xi \end{aligned}$$

Finally, the full solution is

$$u(x) = u_{\text{hom}}(x) + u_p(x) = \frac{16}{3}x^3 - 16x^2 + 12$$