Analysis in Many Variables Course Notes

Isaac Holt

February 9, 2023

Contents

Theorem 0.0.1. If $f(x) : U \to \mathbb{R}$ is differentiable with U an open subset of $\mathbb{R}^n$ and if x is a function of $u_1, \ldots, u_m$, 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 $F(u_1, \ldots, u_m) = f(x(u_1, \ldots, u_m))$ then $\frac{\partial F}{\partial u_b} = \frac{\partial x_i}{\partial u_b}\frac{\partial f}{\partial x_i}$

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

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

$$\frac{\partial F}{\partial u_b} = \lim k \to 0 \frac{F(u_1, \ldots, u_b + k, \ldots, u_m) - F(u_1, \ldots, u_b, \ldots, u_m)}{k} = \lim k \to 0 \frac{f(x(u_1, \ldots, u_b + k, \ldots, u_r}$$

□

0.1 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 0.1.1. (Implicit Function Theorem or IFT): If $f(x, y) : U \to \mathbb{R}$ 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 $\frac{\partial f}{\partial x} = 0$, 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 $\underline{\nabla} f = 0$ (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.

0.2 The implicit function theorem for surfaces

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

Theorem 0.2.1. Let $f(x, y, z) : U \to \mathbb{R}$ 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 surfaces $z = g(x, y)$ 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:

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

$$= \frac{\partial f}{\partial x} + \frac{\partial g}{\partial x}\frac{\partial f}{\partial z}$$

Recall: $\underline{\nabla} f$ 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 $\underline{x}(t) = x_0 + t\underline{\nabla} f$ and the tangent plane is given by $(\underline{x} - x_0).\underline{\nabla} f = 0$.

1 Differentiability of Vector Fields

1.1 Diffentiable maps from $\mathbb{R}^n$ to $\mathbb{R}^n$

Definition 1.1.1. For a vector field $\underline{F}(\underline{x}) : U \to \mathbb{R}^n$, with $U \subseteq \mathbb{R}^n$ open, F is differentiable at $\underline{a} \in U$ if there is a linear function $\underline{L} : \mathbb{R}^n \to \mathbb{R}^n$ such that $\underline{F}(\underline{a}+\underline{h}) - \underline{F}(\underline{a}) = \underline{L}(\underline{h}) + \underline{R}(\underline{h})$ with $\lim_{\underline{h}\to\underline{0}} \frac{\underline{R}(\underline{h})}{|\underline{h}|} = \underline{0}$.

Now linear maps $\mathbb{R}^n \to \mathbb{R}^n$ are given by matrices. To see what matrix, use standard basis $\underline{F}(\underline{x}) = F_1(\underline{x})\underline{e_1} + \cdots + F_n(\underline{x})\underline{e_n}$, $\underline{L}(\underline{h}) = L_1(h)\underline{e_1} + \cdots + L_n(\underline{h})\underline{e_n}$, $\underline{R}(\underline{h}) = R_1(\underline{h})\underline{e_1} + \cdots + R_n(\underline{h})\underline{e_n}$

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

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

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

1.2 Diffeomorphisms and the inverse function theorem

We can think of a vector field $\underline{V}(\underline{x}) : \mathbb{R}^n \to \mathbb{R}^n$ 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 1.2.1. (Inverse function theorem) Let $\underline{v} : U \to \mathbb{R}^n$, $U \subseteq \mathbb{R}^n$ 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 $\underline{w} : \underline{v}(\tilde{u}) \to \mathbb{R}^n$ such that $\underline{w}(\underline{v})(\underline{x})) = \underline{x}$ and $\underline{v}(\underline{w}(\underline{y})) = y$.

Definition 1.2.2. A mapping $\underline{v} : \tilde{U} \to V \subset \mathbb{R}^n$ 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 $\underline{v} : U \to V$ 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 $\underline{v} : \tilde{U} \to \underline{v}(\tilde{U})$ is a diffeomorphism.

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

Then $\underline{w}(\underline{v}(\underline{x}))$ is a map $U \to W \subseteq \mathbb{R}^n$ 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 $J(\underline{v}) \neq 0$. □

Definition 1.2.3. Such a $\underline{v}$ is **orientation-preserving** if $J(\underline{v}) > 0$ and **orientation-reversing** if $J(\underline{v}) < 0$.

2 Volume, Line and Surface Integrals

2.1 Fubini's theorem

Given a scalar field on $\mathbb{R}^2$, $f(x,y) : \mathbb{R}^2 \to \mathbb{R}$ which is continuous on $R \subset \mathbb{R}^2$, then the double integral $\int_R f(x,y)dA$ (or equivalently $\int\int_R f(x,y)dA$) 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):

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

where (x_k^*, y_k^*) lies in the base of the kth 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 $\Delta y_i = y_{i+1} - y_i$ and x and y are partitioned as $x_0 < x_1 < \cdots < x_n$ and $x_0 < x_1 < \cdots < x_n$.

We then get

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

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

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

$$\lim_{n\to\infty, m\to\infty} \sum_{i=0}^{n}\sum_{j=0}^{m} f(x_i^*, y_j^*)\Delta x_i \Delta y_j = \lim_{n\to\infty} \sum_{i=0}^{n} (\lim_{m\to\infty} \sum_{j=0}^{m} f(x_i^*, y_j^*)\Delta y_j)\Delta y_j$$

$$= \lim_{n\to\infty} \sum_{i=0}^{n} (\int_y f(x_i^*, y)dy)\Delta x_i = \int_x \int_y f(x,y)dydx$$

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

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

Theorem 2.1.1. (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:

$$\int_R f(x,y)dA = \int_y \int_x f(x,y)dxdy = \int_x \int_y f(x,y)dydx$$

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 $R = \{(x,y) \in \mathbb{R}^2 : a \le x \le b, y_0(x) \le y \le y_1(x)\}$, then

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

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.

2.2 Line integrals

Definition 2.2.1. A **regular arc** $C \subset \mathbb{R}^n$ is a parameterised curve $\underline{x}(t)$ whose cartesian components $x_a(t)$, $a \in \{1, \ldots, n\}$ are continuous with continuous first derivative, where t lies in some (possibly infinite) interval.

Definition 2.2.2. A **regular curve** consists of a finite number of regular arcs joined end-to-end.

Given $\underline{v}(\underline{x}) : \mathbb{R}^n \to \mathbb{R}^n$ 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 $C : t \to \underline{x}(t)$ 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

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

2.3 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

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

   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.

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

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

2.4 Surface integrals II: evaluating the integral

Definition 2.4.1. We define the surface integral as a Riemann sum. Let $\underline{F}(\underline{x}) : \mathbb{R}^3 \to \mathbb{R}^3$ 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 $\hat{\underline{n}}(u,v)$ is continuous (so S is orientable).

The surface integral is defined as

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

Methods for computing:

1. The surface S given parametrically as $\underline{x}(u,v)$, construct ΔA_k by approximating as parallelograms, by partioning 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 $\hat{\underline{n}}_k \Delta A_k = (\underline{x}(u_i + \Delta u_i, v_j)) - \underline{x}(u_i, v_j)) \times (\underline{x}(u_i, v_j + \Delta v_j) - \underline{x}(u_i, v_j) \approx \Delta u_i \frac{\partial \underline{x}}{\partial u} \times \Delta v_j \frac{\partial \underline{x}}{\partial v}$.

   Substituting this into the surface integral,

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

   Taking the limit, this becomes

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

2. 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 $\frac{\partial \underline{x}}{\partial x} \times \frac{\partial \underline{x}}{\partial y} = \frac{\underline{\nabla} f}{\underline{e_3} \cdot \underline{\nabla} f}$. Then

$$\int_S \underline{F} d\underline{A} = \int_A \frac{\underline{F} \cdot \underline{\nabla} f}{\underline{e_3} \cdot \underline{\nabla} f} dx dy$$

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.

3 Green's, Stokes' and divergence theorems

3.1 The Big 3 theorems

Theorem 3.1.1. (**Green's theorem**) Let $P(x,y) : \mathbb{R}^2 \to \mathbb{R}$ and $Q(x,y) : \mathbb{R}^2 \to \mathbb{R}$ 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

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

Proof. Not examinable □

Theorem 3.1.2. (**Stokes' theorem** - this generalises the vector form of Green's theorem to arbitrary surfaces in $\mathbb{R}^3$) Let $\underline{F}(x,y,z) : \mathbb{R}^3 \to \mathbb{R}^3$ and let S be a surface in $\mathbb{R}^3$ with area elements $d\underline{A} = \underline{\hat{n}} dA$ and let $C = \delta S$ be the boundary of S. Then

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

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

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

where $d\underline{A} = \underline{\hat{n}} dA$ where $\underline{\hat{n}}$ 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:

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

3.2 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 3.2.1. 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 stragiht 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 3.2.2. 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 F is **conservative**.

Proof. Let

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

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.

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

Then

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

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

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

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 3.2.3. 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 $\phi, \underline{F} = \underline{\nabla}\phi$.

4 Non-cartesian systems

Definition 4.0.1. For a non-cartesian coordinate system in three dimensions, we use the letters u, v, w to describe the coordinates.

Definition 4.0.2. 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.

4.1 Change of variables in surface and volume integrals

4.2 Differential operators in polar and spehrical polar coordinates

Definition 4.2.1. 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 4.2.2. The **gradient operator** is defined as

$$\underline{\nabla} = \underline{e_1}\partial_x + \underline{e_2}\partial_y$$

Definition 4.2.3. The Laplacian operator is defined as

$$\begin{aligned} \underline{\nabla}.\underline{\nabla} &= (\underline{e_1}\partial_x + \underline{e_2}\partial_y).(\underline{e_1}\partial_x + \underline{e_2}\partial_y) \\ &= \partial_x^2 + \partial_y^2 \end{aligned}$$

4.3 Polar coordinates

Definition 4.3.1. A point P in polar coordinates is $P : (r, \theta)$ where $\underline{r} = r\cos(\theta)\underline{e_1} + r\sin(\theta)\underline{e_2}$, $r \in [0, \infty)$, $\theta \in [0, 2\pi)$

$|\underline{r}| = r$ 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 4.3.2. 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 4.3.3. $h_u = |\partial_u \underline{r}| = |\underline{e_u}|$, $h_v = |\partial_v \underline{r}| = |\underline{e_v}|$.

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

Definition 4.3.4. We define $\underline{e_r}$ and $\underline{e_\theta}$ 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 4.3.5. Let $f(r,\theta)$ be a scalar function.

$$\begin{aligned} df &= \underline{\nabla} f.d\underline{r} \\ &= (\partial_r f)dr + (\partial_\theta f)d\theta \\ &= (\partial_r f)\underline{e_r}.d\underline{r} + (\partial_\theta f)\frac{1}{r}(\underline{e_\theta} d\underline{r}) \\ &= (\underline{e_r}\partial_r + \underline{e_\theta}\frac{1}{r}\partial_\theta f).d\underline{r} \end{aligned}$$

So $\underline{\nabla} = \underline{e_r}\partial_r + \frac{1}{r}\underline{e_\theta}\partial_\theta$

Definition 4.3.6. (**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

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

Definition 4.3.7. The Laplacian for polar coordinates is

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

Proposition 4.3.8. 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. $dx = \partial_u x du + \partial_v x dv$, $dv = \partial_x v dx + \partial_y v dy$, $dy = \partial_u y du + \partial_v y dv$, $du = \partial_x u dx + \partial_y u dy$.

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 $0 = \partial_u x du \partial_y u + \partial_v x \partial_y dv$.

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 euqation 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. □

5 Generalised functions

5.1 Birth of generalised functions

Definition 5.1.1. 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 5.1.2. The **delta function** is defined as the derivative of the unit step function:

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

5.2 Test functions and distributions

Definition 5.2.1. A **topological space** X is a set of points with a set of neighbourhoods obeying certain axioms.

Definition 5.2.2. An **open set** $O \subset X$ is a set which is a neighbourhood of all its points.

Definition 5.2.3. A set B is a **closed set** if $X \backslash B$ is an open set.

Definition 5.2.4. For $A \subset X$, p is a **limit point** of A is every neighbourhood of p contains at least one element of $A - \{p\}$.

Definition 5.2.5. The **closure** of a set A is the union of A and all its limit points.

Example 5.2.6. In $A = [0, 1)$, 1 is a limit point.

Definition 5.2.7. The **support** of Ψ, denoted suppΨ, 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 5.2.8. Let $\Omega \subset \mathbb{R}^n$ be an open set. $\Psi : \mathbb{R}^n \to \mathbb{R}(\text{ or } \mathbb{C})$ 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. suppΨ is compact.

Definition 5.2.9. 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 5.2.10. If $\Psi \in \Omega(\mathbb{R}^n)$, then

1. $\Psi(\underline{x} + \underline{\xi}) \in \mathcal{D}(\mathbb{R}^n)$ for $\xi \in \mathbb{R}^n$.
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 $||.||$, a sequence of functions $(f_m)_{m\in\mathbb{N}}$ converges to a limit function f if

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

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 5.2.11. 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 $m \to \infty$ if

1. For some compact set $K \subset \Omega$, $\text{supp}\Psi_m \subset K \quad \forall m \in \mathbb{N}$.
2. $(\Psi_m)_{m\in\mathbb{N}} \to \Psi$ with uniform convergence.
3. $$(D_k\Psi_m) := \left(\frac{\partial^{k_1}}{\partial(x_1)^{k_1}}\frac{\partial^{k_2}}{\partial(x_2)^{k_2}}\cdots\frac{\partial^{k_n}}{\partial(x_n)^{k_n}}\right)\Psi_m \to D^k \text{ as } m \to \infty$$
   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 5.2.12. 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 5.2.13. 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 $k = (k_1, \ldots, k_n)$,
$$||D^k - \Psi_m - D^k\Psi||_\infty \to 0 \text{ as } m \to \infty$$

Definition 5.2.14. Let $\Omega \subset \mathbb{R}^n$ be an open set. A **distribution** is linear continuous map $T : \mathcal{D}(\Omega) \to \mathbb{R}$. 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 $(\Psi_m) \to \Psi$,
$$T(\Psi_m) \to T(\Psi) \text{ as } m \to \infty$$

Definition 5.2.15. The vector space of distributions defined via test functions is denoted by $\mathcal{D}'(\Omega)$ and is also infinite-dimensional.

Example 5.2.16. The Dirac delta function is a distribution

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

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: $\forall\Psi \in \otimes, \forall(\Psi_m)_{m\in\mathbb{N}} \subseteq \mathcal{D}(\Omega)$ with $(\Psi_m) \to \Psi$, but uniform convergence in $\mathcal{D}$-convergence implies pointwise convergence. So $(\Psi_m(\underline{x})) \to \Psi(\underline{x})$ as $m \to \infty \quad \forall\underline{x} \in \mathbb{R}^n$. Since $\underline{0} \in \mathbb{R}^n$,
$$\delta(\Psi_m) = (\Psi_m(\underline{0})) \to \Psi(\underline{0}) = \delta(\Psi) \text{ as } m \to \infty$$

□

Example 5.2.17. 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 $T_f : \mathcal{D}(\mathbb{R}^n) \to \mathbb{R}$,

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

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 $(\psi_m) \to \psi$, we have in particular uniform convergence. Therefore we can interchange lim and $\int$, so

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

  So

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

5.3 Regular and singular distributions

Definition 5.3.1. A function $f : \mathbb{R}^n \to \mathbb{R}$ is called **locally integrable** if for every compact set $K \subset \mathbb{R}^n$,

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

Example 5.3.2. $f(x) = x^2$ is not integrable since

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

is not finite but it is locally integrable.

Definition 5.3.3. The space of locally integrable functions is called L^1_{loc}.

Definition 5.3.4. $T \in \mathcal{D}'(\mathbb{R}^n)$ 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 5.3.5. Let $f : \mathbb{R} \to \mathbb{R}$,

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

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

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

Definition 5.3.6. 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 5.3.7. 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 5.3.8. In one dimension, we write $\delta(\underline{x}) = \delta(\underline{x})$. In n dimensions, we write $\delta(x_1, \ldots, x_n) = \delta(x_1) \cdots \delta(x_n)$

5.4 Operations on distributions

Definition 5.4.1. $\forall \psi \in \mathcal{D}(\mathbb{R}^n),\ \forall T_1, T_2 \in \mathcal{D}'(\mathbb{R}^n)$, 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 5.4.2. 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 $x_1, \ldots, x_n$ 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 5.4.3. For $b > 0$, let $f(x) = x^2 - b^2$, $f : \mathbb{R} \to \mathbb{R}$, 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 5.4.4. 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 5.4.5. For which value of $A \in \mathbb{R}$ are $T_1 = x^2\delta(x^3)$ and $T_2 = A\delta(x)$ equal?

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

so $A = 1/3$.

5.5 The derivative of a distribution

Definition 5.5.1. 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 5.5.2. 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 5.5.3. 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 5.5.4. Show that, in the sense of distributions, $\delta'(cx) = \frac{1}{c^2}\delta'(x)$ for a constant $c > 0$.

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

5.6 Leibniz rule for differentiation

Proposition 5.6.1. 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 5.6.2. 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 5.6.3. 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 5.6.4.

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

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 5.6.5. A function f is **piecewise smooth** if

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

Example 5.6.6.

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

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

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

which is piecewise continuous.

Example 5.6.7. tan is not a piecewise smooth function.

Definition 5.6.8. 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 $\tilde{f}(x) = f_1(x) + (f_2(x) - f_1(x))\Theta(x - x_1) + (f_3(x) - f(2))\Theta(x - x_2) + \cdots$. But $\tilde{f}(x) = f(x)$ on (a, b).
2. Compute $f'(x) = \tilde{f}'(x)$.

Example 5.6.9.

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

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

$$\begin{aligned} f'(x) &= \cos(x) + \left(-\frac{1}{2}\sin(x) - \cos(x)\right)\Theta(x - \pi) + \left(\frac{1}{2}\cos(x) - \sin(x)\right)\delta(x - \pi) + \frac{1}{2}\sin(x)\Theta(x - 2\pi) + \\ &= \cos(x) - \left(\frac{1}{2}\sin(x) + \cos(x)\right)\Theta(x - \pi) + \left(-\frac{1}{2} - 0\right)\delta(x - \pi) + \frac{1}{2}\sin 9x)\Theta(x - 2\pi) + \left(\frac{3}{2} - \frac{1}{2}\right) \\ &= -\frac{1}{2}\delta(x - \pi) + \delta(x - 2\pi) + \begin{cases} \cos(x) & \text{if } 3\pi < x \le \pi \\ -\frac{1}{2}\sin(x) & \text{if } \pi < x \le 2\pi \\ 0 & \text{if } 2\pi < x < 3\pi \end{cases} \end{aligned}$$

Definition 5.6.10. The **sign function** is defined as

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

Example 5.6.11.

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

So the derivative of the sign function is

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

5.7 The delta distribution in polar/spherical coordinates

Definition 5.7.1. 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 $(x_1, \ldots, x_n), (\xi_1, \ldots, \xi_n)$. J is the Jacobian of the change of variable from $\underline{x}$ to $\underline{\xi}$.

Example 5.7.2. If the problem at hand does not depend on θ, then

$$\delta(\underline{x} - \underline{x_0}) = \frac{1}{2\pi r}\delta(\underline{r} - \underline{r_0})$$

6 Lebesgue Integration

TODO: notes from lecture

Theorem 6.0.1. Let $f : [a, b] \to \mathbb{R}$ be bounded. Then f is Riemann integrable iff f is continuous 'almost everywhere" (it is continuous except on sets with zero measure).

6.1 Hilbert spaces: definitions and examples

Definition 6.1.1. 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, $\langle \cdot, \cdot \rangle : \mathbb{H} \times \mathbb{H} \to \mathbb{C}$, 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**: $\forall a \in \mathbb{C}, \langle a(u + v), w \rangle = \overline{a}\langle u + v, w \rangle = \overline{a}\langle u, w \rangle + \overline{a}\langle v, w \rangle$.
   (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 6.1.2. 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 6.1.3. 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

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

or equivalently,

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

Definition 6.1.4. 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 6.1.5. A metric space (V, d) is called **complete** if every Cauchy sequence in (V, d) converges in V.

6.2 Inner products defined by integrals

Definition 6.2.1. Consider the vector space V of complex-valued functions, defined on an interval $[a, b]$ and let $\omega : [a, b] \to \mathbb{R}^+$ 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 6.2.2. The space $C^0(\omega, [a, b])$ of continuous functions on $[a, b]$ with inner product $*$ and norm

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

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 6.2.3. 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 $C^0([0, 1])$.

6.3 Bounded and unbounded linear operators

Definition 6.3.1. 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 6.3.2. 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. $L : D(L) \to \mathbb{H}$ 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 6.3.3. 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 6.3.4. Let $(\mathbb{H}_1, ||\cdot||_{H_1})$ and $(\mathbb{H}_2, ||\cdot||_{H_2})$ be Hilbert spaces. A linear operator $L : \mathbb{H}_1 \to \mathbb{H}_2$ is called **bounded** if for some $M > 0$,

$$\forall u \in \mathbb{H}_1, \quad ||L(v)||_{H_2} \leq M||v||_{H_1}$$

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

Definition 6.3.5. The **norm** of L is defined as

$$||L|| := \inf\{M : ||L(v)||_{H_2} \leq M||v||_{H_1} \text{ for every } v \in \mathbb{H}_1\}$$

Example 6.3.6. $I : \mathbb{H} \to \mathbb{H}$ 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 6.3.7. Differential operators are usually unbounded. For example,

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

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,

$$||x_n||_{L^2} = \sqrt{\int_\pi^\pi \overline{\cos(nt)}\cos(nt)dt} = \sqrt{\pi}$$

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

Hence

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

which is unbounded.

Definition 6.3.8. Let $\mathbb{H}_1, \mathbb{H}_2$ be Hilbert spaces, with $\langle\cdot,\cdot\rangle^{(1)}$ and $\langle\cdot,\cdot\rangle^{(2)}$. Let $L : D(L) \subset \mathbb{H}_1 \to \mathbb{H}_2$ 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 $L* : D(L^*) \subset \mathbb{H}_2 \to \mathbb{H}_1$ 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 $v_2^* = L*(v_2)$.

6.4 Sturm-Liouville operators

Proposition 6.4.1. (**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 6.4.2. 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 6.4.3. 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 coniditions 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 6.4.4. 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 6.4.5. 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 $p_1 + \overline{p_1} = 2\mathrm{Re}(p_1) = 2p_0'$. Also $p_2 - \overline{p_2} = 2\mathrm{Im}(p_2) = p_0'' - \overline{p_1}'$ so $2\mathrm{Im}(p_2) = \mathrm{Im}(p_1')$.

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

Definition 6.4.6. 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 6.4.7. (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 6.4.8. (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

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

we have

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

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 $p_i \to \rho p_i$,

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

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

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