MATH2921 Vector Calculus & Differential Equations

Jerry Ye Xu

March 2020

Abstract

Hi there! This is the set of Vector Calc. and Diff. Eqs. notes I wrote during my self study. Feel free to send me feedback via email.

I write about math and programming on my personal website, and share non-technical snippets and other opinions through Medium.

Contents

MATH2921 Vector Calc. & ODEs

1. Partial Differentiation

Introduction to Partial Differentiatives

1.1 Introduction to Partial Differentiatives

1.1.1 Quick definition and Example

Definition 1.1 *(Partial derivative)*

When dealing with functions of multiple variables, the partial derivative with respect to a single variable is its ordinary derivative whilst keeping the other variables constant.

Suppose you have $f(x,y) = xy + e^x \cos(y)$. Then

$$\begin{aligned}\frac{\delta f}{\delta x} &= f_x \\ &= y + e^x \cos(y) \\ \frac{\delta f}{\delta y} &= f_y \\ &= x - e^x \sin(y)\end{aligned}$$

Of course, we can extend to any general function $f(x_1, x_2, ..., x_n)$.

We can also take higher order derivatives;

$$\begin{aligned}\frac{\delta^2 f}{\delta x \delta y} &= f_{xy} \\ &= -e^x \sin(y) \\ \frac{\delta^2 f}{\delta y \delta x} &= f_{yx} \\ &= 1 - e^x \sin(y) \\ \frac{\delta^2 f}{\delta x^2} &= f_{xx} \\ &= e^x \cos(y) \\ \frac{\delta^2 f}{\delta y^2} &= f_{yy} \\ &= -e^x \cos(y)\end{aligned}$$

and this can be neatly represented in a matrix

$$\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yx} \end{bmatrix}$$

1.1.2 Equality of Mixed Partials

Definition 1.2 *(Equality of Mixed Partials)*

If $f(x,y)$ is continuous and twice differentiable, then the mixed partial derivatives of $f(x,y)$ are equal, that is

$$\frac{\delta^2 f}{\delta x \delta y} = \frac{\delta^2 f}{\delta y \delta x}$$

(Sketch of Proof) Given the definition of a derivative, we have

$$\frac{\delta f}{\delta x} = \lim_{h \to 0} \frac{f(x+h,y) - f(x,y)}{h}$$
$$\frac{\delta f}{\delta y} = \lim_{k \to 0} \frac{f(x,y+k) - f(x,y)}{k}$$

"Substituting" one into the other leads to

$$\frac{\delta^2 f}{\delta x \delta y} = \lim_{h \to 0} \lim_{k \to 0} \left[\frac{f(x+h,y+k) - f(x,y+k) - (f(x+h,y) - f(x,y))}{h \cdot k} \right]$$

This is what we are looking for.

Now we apply the Mean value Theorem twice, which states that

$$\frac{\delta}{\delta x} f(x + \lambda h) = \frac{f(x+h) - f(x)}{h}$$

for some $\lambda \in (0,1)$, to a function

$$\begin{aligned}
\psi(x) &= \frac{f(x,y+k) - f(x,y)}{k} \\
\implies \frac{\delta}{\delta x}\psi(x+\lambda h) &= \frac{\psi(x+h) - \psi(x)}{h} \\
&= \frac{f(x+h,y+k) - f(x+h,y) - (f(x,y+k) - f(x,y))}{h \cdot k} \\
&= \frac{f_x(x+\lambda h, y+k) - f_x(x+\lambda h, y)}{k} \\
&= \frac{\phi(y+k) - \phi(y)}{k} \\
& \textit{where } \ \phi(y) = \frac{f_x(x+\lambda h, y)}{k} \\
& \textit{and once you realise that the RHS function is equivalent to} \\
& \textit{the LHS after you apply mean-value theorem, then we have} \\
\implies \frac{\delta}{\delta y}\left(\phi(y+\gamma k)\right) &= \frac{\delta}{\delta y} f_x(x+\lambda h, y+\gamma k) \\
&= f_{yx}(x+\lambda h, y+\gamma k)
\end{aligned}$$

Notice how this is equivalent with the long expression we obtained earlier.

A similar expression can be obtained by using the same ideas in the reverse direction.

Finally, you can say that because the functions are continuous, we have that $(x+h, y+k) \to (x,y)$ as $(h,k) \to (0,0)$ regardless of the order in which you take the limit.

Hence the mixed partials are equal.

1.1.3 Linear Approximations and Tangent Planes

Let's look at linear approximations first. For the simplest case, we have that

$$f(x) = f(x_0) + f'(x_0)(x - x_0)$$

Extending this to planes, we have the linear approximation

$$f(x, y) \approxeq L(x, y) = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

This might seem familiar yet different. Let's have a look at the linear algebra version of this, which was what I had learnt first.

Recall that to find the equation of a plane, we'll need a normal vector and 2 'points' on a plane, let's say $\boldsymbol{n} = (a, b, c)$, $\boldsymbol{r} = (x, y, z)$ and $\boldsymbol{r'} = (x_0, y_0, z_0)$.

We know that $\boldsymbol{n} \cdot (\boldsymbol{r} - \boldsymbol{r'}) = 0$ and this we can obtain

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

where you can rearrange to get

$$ax + by + cz = d$$

Now in familiar territory, we can simply rearrange some things to obtain

$$z - z_0 = -\frac{a}{c}(x - x_0) - \frac{b}{c}(y - y_0)$$

Suppose we hold either x or y fixed by letting $y = y_0$ for example.

This turns out to be the equation that satisfies the tangent line of a 2D curve, and thus $-\frac{a}{c} = f_x(x_0, y_0)$, and similar for $-\frac{b}{c}$.

Hence that is why the **normal vector of the plane** is

$$(-f_x, -f_y, 1)$$

To make it more explicit, we can write

$$z - f(x_0, y_0) = -f_x(x_0, y_0)(x - x_0) - f_y(x_0, y_0)(y - y_0)$$

1.1.4 Curves on Surfaces

We show here that we can find the tangent plane of a surface at an arbitrary point if we are given a curve on the graph. Let's start with an example.

Suppose you have

$$\begin{gathered} z = x^2 - y^2 \\ \text{with curve on the graph} \\ \gamma(t) = (\exp^t, \exp^{-t}, 2\sinh(2t)) = (x, y, z) \end{gathered}$$

We know that the equation of the tangent plane at t_0 is

$$
\begin{aligned}
z &= -f_x(x_0, y_0)(x - x_0) - f_y(x_0, y_0)(y - y_0) + f(x_0, y_0) \\
&= 2e^{t_0}(x - e^{t_0}) - 2e^{-t_0}(y - e^{-t_0}) + 2\sinh(2t_0) \\
&= 2e^{t_0}x - 2e^{-t_0}y - 4\left(\frac{e^{2t_0} - e^{-2t_0}}{2}\right) + 2\sinh(2t_0) \\
&= 2e^{t_0}x - 2e^{-t_0}y - 2\sinh(2t_0)
\end{aligned}
$$

since $\frac{e^t + e^{-t}}{2} = \cosh(t)$, and now we use $\frac{e^t - e^{-t}}{2} = \sinh(t)$ to show that

$$
\begin{aligned}
\gamma'(t_0) \cdot \boldsymbol{n} &= (e^{t_0}, -e^{-t_0}, 4\cosh(2t_0)) \cdot (2e^{t_0}, -2e^{-t_0}, 1) \\
&= \mathbf{0}
\end{aligned}
$$

Hence, we see that if $\gamma(t)$ is a curve on some graph then $\gamma(t_0)$ lies on the tangent plane to the graph. Combined with the second idea, we can now find the tangent plane.

Definition 1.3 *(Tangent Plane to a Graph)*

Let G be the graph of a differentiable function $f(x_1, ..., x_n)$, then the tangent plane to the graph G at the point P is

$$T_P(G) = \{\boldsymbol{v} \mid \boldsymbol{v} = \gamma'(0), \;\; \textit{for} \;\; \gamma(t) \in G, t \in (-\epsilon, \epsilon), \gamma(0) = P\}$$

1.1.5 Chain Rule

Definition 1.4 *(Chain Rule)*

A formula to compute the derivative of a composite function. That is, in the simplest case, given $h = f \circ g$ where $g(x)$ is a function of x, $f(y)$ depends on g and h depends on f,

$$
\begin{aligned}
(f \circ g)' &= (f' \circ g) \cdot g' \\
&= f'(g(x))g'(x) \\
\frac{\delta h}{\delta x} &= \frac{\delta h}{\delta y} \cdot \frac{\delta y}{\delta x}
\end{aligned}
$$

Definition 1.5 *(Derivative Matrix)*

We want to generalise partial derivatives. Suppose we have a bunch of functions $u_1 = f_1(x_1, x_2, ..., x_n), u_2 = f_2(x_1, x_2, ..., x_n), ..., u_k = f_k(x_1, x_2, ..., x_n)$. Then we can write the derivatives of every possible combination in a matrix.

$$
\frac{\delta(u_1, u_2, ..., u_k)}{\delta(x_1, x_2, ..., x_n)} = \begin{pmatrix}
\frac{\delta u_1}{\delta x_1} & \frac{\delta u_1}{\delta x_2} & \cdots & \frac{\delta u_1}{\delta x_n} \\
\frac{\delta u_2}{\delta x_1} & \frac{\delta u_2}{\delta x_2} & \cdots & \frac{\delta u_2}{\delta x_n} \\
\vdots & \vdots & & \vdots \\
\frac{\delta u_k}{\delta x_1} & \frac{\delta u_k}{\delta x_2} & \cdots & \frac{\delta u_k}{\delta x_n}
\end{pmatrix}
$$

Definition 1.6 *We want to generalise the chain rule. Suppose we have a bunch of functions* $u_1 = f_1(x_1, x_2, ..., x_n), u_2 = f_2(x_1, x_2, ..., x_n), ..., u_k = f_k(x_1, x_2, ..., x_n)$*, and let* $x_1 = g_1(t_1, ..., t_q), x_2 = g_2(t_1, ..., t_q), ..., x_n = g_n(t_1, ..., t_q)$.

Then we can concisely write this as

$$\frac{\delta(u_1, u_2, ..., u_k)}{\delta(t_1, t_2, ..., t_q)} = \frac{\delta(u_1, u_2, ..., u_k)}{\delta(x_1, x_2, ..., x_n)} \frac{\delta(x_1, x_2, ..., x_n)}{\delta(t_1, t_2, ..., t_q)}$$

Feel free to expand this out into the derivative matrix you saw above... and then once you do that, you can see that for an individual partial derivative, we have

$$\frac{\delta u_i}{\delta t_j} = \frac{\delta u_i}{\delta x_1} \cdot \frac{\delta x_1}{\delta t_j} + \frac{\delta u_i}{\delta x_2} \cdot \frac{\delta x_2}{\delta t_j} + ... + \frac{\delta u_i}{\delta x_n} \cdot \frac{\delta x_n}{\delta t_j}$$

You can also use the change run to switch between polar and Cartesian coordinates etc. Suppose you have a function $f(x, y)$ and you wanted to know what $f_r(x, y)$ is. Then by using polar coordinates $x = r\cos(\theta), y = r\sin(\theta), r^2 = x^2 + y^2, \theta = \arctan(\frac{y}{x})$, we can get

$$\frac{\delta f}{\delta r} = \frac{\delta f}{\delta x} \cdot \frac{\delta x}{\delta r} + \frac{\delta f}{\delta y} \cdot \frac{\delta y}{\delta r}$$

MATH2921 Vector Calc. & ODEs

2. Gradients and Friends

Gradients, Maxima and Minima

2.2 Gradients, Maximia and Minima

2.2.1 Gradients and Directional Derivatives

Definition 2.7 *(Gradient Vector Field)*

Suppose you have $u = f(x, y, z)$. Then the gradient vector field, denoted ∇f, is

$$\nabla f(x, y, z) = f_x(x, y, z)\boldsymbol{i} + f_y(x, y, z)\boldsymbol{j} + f_z(x, y, z)\boldsymbol{k}$$

*The gradient vector field contains the **vector form** of all partial derivatives of f.*

Definition 2.8 *(Vector Field)*

A function $\phi(x, y, z)$ which assigns a vector in space to each point of the function (x, y, z) is called a vector field.

Note: If the mixed partials are not equal, then the vector field is not a gradient vector field.

Now suppose we have a function $u = f(x, y, z)$ and a curve $\boldsymbol{\sigma}(t)$. If we take the derivative w.r.t to t , this can be described as "the derivative of f along the curve. We have

$$\begin{aligned} \frac{\delta}{\delta t} f(\boldsymbol{\sigma}(t)) &= \nabla f(\boldsymbol{\sigma}(t))\boldsymbol{\sigma}'(t) \\ &= \|f'(\boldsymbol{\sigma}(t))\|\|\boldsymbol{\sigma}'(t)\| \cos(\theta) \end{aligned}$$

Definition 2.9 *(Directional Derivative)*

Let $f(x, y, z)$ be a function and $\boldsymbol{r}$ a point in it's domain $\boldsymbol{d}$ a unit vector. Define the parametric curve $\boldsymbol{\sigma}(t) = \boldsymbol{r} + t\boldsymbol{d}$. Then

$$\begin{aligned} \frac{\delta}{\delta t} f(\boldsymbol{\sigma}(t))_{|t=0} &= \nabla f(\boldsymbol{r}) \cdot \boldsymbol{d} \\ &= \|f'(\boldsymbol{r})\| \cos(\theta) \end{aligned}$$

is defined as the directional derivative at $\boldsymbol{r}$ in the direction of $\boldsymbol{d}$.

The maximum occurs when $\boldsymbol{d}$ points in the same direction as $\nabla f(\boldsymbol{r})$, and the minimum occurs when $\boldsymbol{d}$ points in the direction of -$\nabla f(\boldsymbol{r})$

Definition 2.10 *(Tangent Plane)*

*If S is a surface in space, and s_0 a point of S. The plane that contains all the possible tangelt lines at s_0 to all curves through s_0 in S is called the **tangent plane** to S at s_0. The normal to this tangent plane is perpendicular to S.*

Definition 2.11 *(Level Set/Curve)*

Define a function $z = f(x, y)$. A level curve is a function $f(x, y) = k$, where k is an arbitrary constant.

Remark 2.12 *The gradient vector $\nabla f(x_0, y_0)$ is orthogonal to the level curve at $f(x, y) = k$ at the point (x_0, y_0).*

Now suppose you have a level set $M := f(x, y, z) = C$ and a curve $\gamma(t) = (x(t), y(t), z(t))$ that is on the level set. We have

$$\begin{aligned} \frac{\delta}{\delta t}(f(x(t), y(t), z(t)) &= \nabla f(\gamma(t)) \cdot \gamma'(t) \\ &= \mathbf{0} \end{aligned}$$

With this, we can define the tangent space $T_P(M)$ as the null space of the linear map defined by the inner product with the gradient vector.

Of course, you can understand the equation of the tangent line at (x_0, y_0) to the curve $f(x, y) = k$ to be

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) = 0$$

2.2.2 Maxima and minima

We consider cases with functions of at least 2 variables.

Furthermore, we **assume that the function is at least twice differentiable**, denoted C^2.

Definition 2.13 *(Local Maxima and minima)*

A local max (min) is a point (x_0, y_0) s.t. $\exists$ a disc of positive radius about it s.t. $f(x, y) \leq f(x_0, y_0) (\geq f(x_0, y_0))$ $\forall (x, y)$ in the disc.

Remark 2.14 *Let $f(x, y)$ be C^2 i.e. twice differentiable. Then the tangent plane of the surface $z = f(x, y)$ at a local max (or min) is "horizontal".*

(Sketch of Proof)

Consider the directional derivative of $f(x, y)$ in the direction $\boldsymbol{d} = (d_1, d_2)$ at the local maximum (x_0, y_0).

we know that

$$\nabla f(x_0, y_0) \cdot \boldsymbol{d} = |\nabla f(x_0, y_0)| \cos(\theta)$$

but since $f(x_0, y_0)$ is a local max, it's derivative must be zero. Since this holds for all possible values of $\boldsymbol{d}$, it must be that

$$\begin{aligned} \nabla f(x_0, y_0) &= \mathbf{0} \\ \Rightarrow f_x &= f_y = 0 \end{aligned}$$

And we know that the normal vector, $\boldsymbol{n}$ at (x_0, y_0) is equal to $(-f_x, -f_y, 1) = (0, 0, 1)$.

As we understand the orientation of x, y, z axis - z is actually "vertical" which means that the tangent plane must be "horizontal".

Definition 2.15 *(Critical Points)*

Points (x_0, y_0) *are critical points if* $\nabla f(x_0, y_0) = \mathbf{0}$.

Definition 2.16 *(Saddle Points)*

Intuitively, a saddle point occurs when a critical point is a local 'maximum' in one direction but a local 'minimum' in another.

Suppose (x_0, y_0) *is a critical point of* $f(x, y)$. *Then the critical point is a saddle if* $\forall$ *discs of positive radius* $\exists(x^*, y^*)$ *in the disc s.t.* $f(x^*, y^*) \geq f(x_0, y_0)$ *and* $\exists(x', y')$ *in the same disc s.t.* $f(x', y') \leq f(x_0, y_0)$

Consider an example $f(x, y) = x^2 - y^2$.

You have a critical point at $(0, 0)$, and now we check the nearby points by varying x and y.

Suppose we fix $x = 0$ and vary y, then $f(0, y) < 0$ BUT if we fix $y = 0$ and vary x, then $f(x, 0) > 0$.

What is happening here is that the two directives we are moving in contradict each other in terms of whether the critical point is a maximum or minimum!

2.2.3 Classifying Maxima and Minima

The textbook suggests a method and the course itself teaches a method that is based on eigenvalues.

I think both are valid but the eigenvalue approach using Hessian matrices are more generalisable.

Remark 2.17 *(The Second Derivative Test)*

Let $f(x, y)$ *be* C^2 *with* (x_0, y_0) *as the critical point.*

Let

$$\begin{aligned} A &= f_{xx}(x_0, y_0) \\ B &= f_{xy}(x_0, y_0) \\ C &= f_{yy}(x_0, y_0) \end{aligned}$$

Then we have that

$$\begin{aligned} A > 0, AC - B^2 > 0 &\Rightarrow (x_0, y_0) \textit{ is a local minimum} \\ A < 0, AC - B^2 > 0 &\Rightarrow (x_0, y_0) \textit{ is a local maximum} \\ AC - B^2 < 0 &\Rightarrow (x_0, y_0) \textit{ is a saddle point} \\ AC - B^2 = 0 &\Rightarrow (x_0, y_0) \;\; \textit{inconclusive} \end{aligned}$$

Remark 2.18 *(The Hessian Second Derivative Test)*

Let $f(x, y)$ *be* C^2 *with* (x_0, y_0) *as the critical point.*

Define the Hessian matrix

$$\boldsymbol{H} = \boldsymbol{D}^2 f = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$$

Define λ_1, λ_2 *as the eigenvalues of* $\boldsymbol{D}^2 f$ *at the critical point* (x_0, y_0)*. Then*

$$\begin{aligned}
\lambda_1, \lambda_2 > 0 &\Rightarrow (x_0, y_0) \textit{ is a local minimum} \\
\lambda_1, \lambda_2 < 0 &\Rightarrow (x_0, y_0) \textit{ is a local maximum} \\
\lambda_1 > 0 > \lambda_2 &\Rightarrow (x_0, y_0) \textit{ is a saddle point} \\
\lambda_1 = 0 \textit{ or } \lambda_2 = 0 &\Rightarrow (x_0, y_0) \textit{ inconclusive}
\end{aligned}$$

With this method, you can generalise to higher dimensions much easier.

2.2.4 Constrained Maxima and Minima

For a function $f(x, y)$, you may want to constraint the domain by considering only a particular region D. In this case, how we would go about finding the critical points?

Let's walk through an example. Suppose you have $z = f(x, y) = x^2 + 2y^2$ that is constrained by the disc $D = x^2 + y^2 \leq 1$.

First we look at the critical points and filter the ones that are within the domain.

$$\begin{aligned}
\frac{\delta z}{\delta x} &= 2x_{|x=0} \\
\frac{\delta z}{\delta y} &= 4y_{|y=0}
\end{aligned}$$

We have $(x, y) = (0, 0)$ as the critical points. Using the second derivative test we find that it is a minimum point.

Now we look at the boundary of D, a circle. Let's parametrise this as $h(t) = f(\cos(t), \sin(t)) = \cos^2(t) + 2\sin^2(t) = 1 + \sin^2(t)$.

$$\begin{aligned}
h'(t) &= 2\sin(t)\cos(t) \\
&\Rightarrow t = 0 \text{ for a critical point}
\end{aligned}$$

For this to work, $t = 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$.

This means that

$$\begin{aligned}
(\cos(0), \sin(0)) &= (1, 0) \\
(\cos(\pi), \sin(\pi)) &= (-1, 0) \\
(\cos(\frac{\pi}{2}), \sin(\frac{\pi}{2})) &= (0, 1) \\
(\cos(\frac{3\pi}{2}), \sin(\frac{3\pi}{2})) &= (0, 1)
\end{aligned}$$

Now evaluating all of this means that

$$\begin{aligned}
f(1, 0) = f(-1, 0) &= 1 \\
f(0, 1) = f(0, -1) &= 2
\end{aligned}$$

which gives us a maximum at $f(0, 1)$ and $f(0, -1)$.

Suppose f and g are functions of two variables with C^2. Then suppose f is restricted to a level curve C defined by $g(x, y) = C$ and it has a local extremum at (x_0, y_0) and $\nabla g(x_0, y_0) \neq \mathbf{0}$. Then $\exists \lambda$ s.t.

$$\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0)$$

Let's think about this for a second. Construct a parametrisation $(x, y) = \boldsymbol{\sigma}(t))$ where $\boldsymbol{\sigma}(0)) = (x_0, y_0)$ and $\boldsymbol{\sigma'}(0)) \neq 0$. We need the non-zero constraint as you will see later.

Now, $h(t) = f(\boldsymbol{\sigma}(t))) = h(x_0, y_0)$ at $t = 0$ so $h'(0) = 0$.

$$h'(0) = \nabla f(x_0, y_0) \cdot \boldsymbol{\sigma'}(0)$$

We know that $\nabla f(x_0, y_0)$ is perpendicular to $\boldsymbol{\sigma'}(0)$, and since $\nabla g(x_0, y_0)$ is the gradient of a level curve, we know that it's derivative is perpendicular to the tangent vector $\boldsymbol{\sigma'}(0)$.

Because $\boldsymbol{\sigma'}(0)$ is not a zero vector, the 2 vectors normal to a vector must be parallel to each other.

In effect, this means that

$$\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0)$$

2.2.5 The Method of Lagrange Multipliers

Definition 2.19 *(Method of Lagrange Multipliers)*

To find the constrained extreme of a function $f(x, y)$ subject to the constraint $g(x, y) = C$, solve the system

$$\begin{aligned} \nabla f(x_0, y_0) &= \lambda \nabla g(x_0, y_0) \\ g(x, y) &= C \end{aligned}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$

Let's look at an example.

1) $g(x, y) = \frac{(x-1)^2}{16} + \frac{(y)^2}{3} = 1$, with $f(x, y) = x^2 + y^2 = z^2$

Writing the equations we get

$$\begin{aligned} 2x &= \lambda \cdot \frac{x-1}{8} \\ 2y &= \lambda \cdot \frac{2y}{3} \end{aligned}$$

Notice that $y = 0$ or $\lambda = 3$. For now let's ignore $y = 0$.

$$\begin{aligned} 2x &= 3 \cdot \frac{x-1}{8} \\ 16x &= 3(x-1) \\ \Rightarrow x &= -\frac{3}{13} \end{aligned}$$

$$\begin{aligned} y^2 &= 3\left(1 - \frac{16^2}{13^2 \cdot 16}\right) \\ &= 3\left(1 - \frac{16}{13^2}\right) \\ \Rightarrow y &\pm \frac{3\sqrt{51}}{13} \end{aligned}$$

Now, what happens if $y = 0$. Then we have

$$\frac{(x-1)^2}{16} = 1$$
$$\Rightarrow x = 5, -3$$

If $x = 5$, then

$$f(5,0) = 5^2 + 0$$
$$= 25 \text{ is a maximum}$$

If $x = -3$, what could it be? It isn't a maximum since $25 > 9$. First let's find λ.

$$-6 = \lambda \cdot -\frac{4}{8} \Rightarrow \lambda = 12$$

and so we have $(x, y, \lambda) = (-3, 0, 12)$

Now we write $g(x, y)$ as $g(x(y), y)$

$$g^*(y) = \left(1 - 4\left(1 - \frac{y^2}{3}\right)^{\frac{1}{2}}\right) + y^2$$

differentiating twice and setting $y = 0$ gives

$$\frac{\delta}{\delta y^2} g^*(0) = -6$$

And thus at $x = -3$ this is a constrained local maxima.

Hence the extreme points are $\left(-\frac{3}{13}, \pm\frac{3\sqrt{51}}{13}\right), (5, 0)$

Definition 2.20 *(Lagrangian Function)*

Consider the function $f(x, y)$ subject to the constraint $g(x, y) = C$. Define a function

$$\mathcal{L}(x, y) = f(x, y) - \lambda g(x, y)$$

*L(x, y) is the **Lagrangian function**.*

In order to solve the same problem using this approach, we first derive the Hessian, $D^2\mathcal{L}(P)$ at the critical point P.

Given a tangent vector $\boldsymbol{v} = (x_0, y_0)$ passing P, we examine

$$\boldsymbol{v}^T D^2\mathcal{L}(P)\boldsymbol{v} > 0 \Rightarrow (x_0, y_0) \text{ is a local minimum}$$
$$\boldsymbol{v}^T D^2\mathcal{L}(P)\boldsymbol{v} < 0 \Rightarrow (x_0, y_0) \text{ is a local maximum}$$
$$\text{if sign indefinite } \Rightarrow (x_0, y_0) \text{ is a saddle point}$$
$$\text{if } \boldsymbol{v} = \mathbf{0} \text{ for some 2nd derivative } \Rightarrow (x_0, y_0) \text{ inconclusive}$$

$\forall \boldsymbol{v} \neq \mathbf{0} \in T_P(g(x, y) = c)$

Let's use the previous example to illustrate this idea. Consider the triplet $(x, y, \lambda) = (-3, 0, 12)$ again.

This time, we have

$$D^2\mathcal{L}(P) = \begin{pmatrix} 2-\frac{\lambda}{8} & 0 \\ 0 & 2-\frac{2\lambda}{3} \end{pmatrix}_{|\lambda=12} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -6 \end{pmatrix}$$

Now, if we find the tangent plane of $f(x, y)$ as normal, then we have

$$\begin{aligned} f_x(-3,0) &= \frac{2\cdot -3+1}{16} \\ f_y(-3,0) &= \frac{2\cdot 0}{3} \end{aligned}$$

which gives

$$\begin{aligned} f_x(-3,0)(x+3) + f_y(-3,0)(y-0) &= -\frac{7}{16}(x+3) + 0(y-0) \\ &= -\frac{7}{16}(x+3) \end{aligned}$$

This implies that y is under no constraint, which means that the tangent space to the ellipse at $P = (-3, 0)$ are vectors of the form $[0, t]^T$, $t \in \mathbb{R} \setminus \{0\}$.

$$\boldsymbol{v}^T D^2\mathcal{L}(P)\boldsymbol{v} = \begin{pmatrix} 0 & t \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -6 \end{pmatrix} \begin{pmatrix} 0 \\ t \end{pmatrix} = -6t^2 < 0$$

Hence it is a constrained maximum.

MATH2921 Vector Calc. & ODEs

3. Integration of Multiple Integrals

Multiple Integral

3.3 Multiple Integral

3.3.1 Double Integral and Iterated Integral

First let's rewind time back to Riemann integrals.

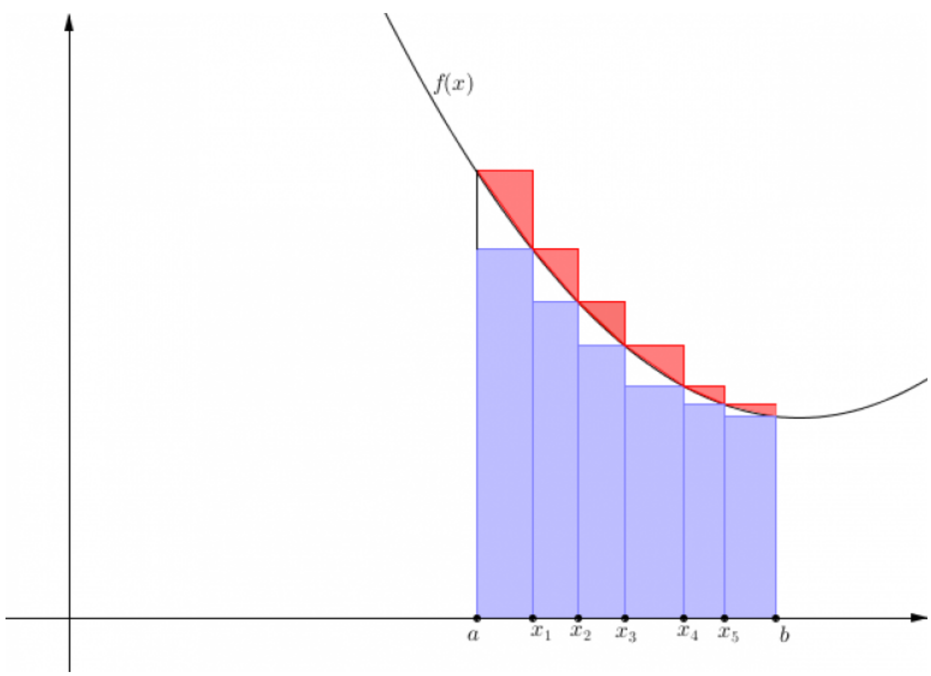
Definition 3.21 *(Riemann Integral)*

we define the Riemann integral as the the limit of Riemann sums of a function as the partitions of a function with a domain $[a, b]$ become finer.

We define Δx_i as a partition and we can construct upper and lower bound rectangles.

$$L_p = \sum_i f(x_i^L)\Delta x_i$$
$$U_p = \sum_i f(x_i^U)\Delta x_i$$
$$\Rightarrow L_p \leq f_a^b f(x)\ \delta x \leq U_p$$

for all partitions $\Delta x_i \in [a, b]$



Properties of the Riemann integral include

1. Linearity
2. It's independent of its partitions
3. Viewed as the "limit" as $\Delta x_i \to 0$.
4. Signed area under the curve (definite integral)
5. Fundamental Theorem of Calculus (we discuss later)

So we have a "interval" partition in the for an integral in one dimension. We can apply the intuition to two dimension as well.

Definition 3.22 *(Double Integral)*

For two dimensions, suppose we have a closed rectangular region $[a, b] \times [c, d]$. We can draw rectangles in this region by using $g(x, y)$, a step function, with partitions $a = t_0 < t_1 < ... < t_n = b$ and $c = s_0 < s_1 < ... < s_m = d$. Each of the open rectangles, denoted $r_{ij} := (t_{i-1}, t_i) \times (s_{j-1}, s_k)$ when put into the function $g(x, y) = k_{ij}$.

We define the double integral as the sum of the heights of the rectangles

$$\begin{aligned}\int\int_D g(x,y)\ \delta x\delta y &= \sum_{i=1}^{n}\sum_{i=j}^{m}(k_{ij})r_{ij}\\ &= \sum_{i=1}^{n}\sum_{i=j}^{m}(k_{ij})(\Delta t_i)(\Delta s_j)\end{aligned}$$

for a region D.

Alternatively, the textbook defines the integral using upper and lower bounds again. We say that f is integrable on D if $\exists S_0$ s.t. $S < S_0$ is the lower bound for f on D and every $S^* > S_0$ is an upper sum.

Note: Right now we are only dealing with definite integrals! Integrating indefinite integrals would require anti-derivatives to exist, but it is not obvious what it is right now.

We still have not discussed the computation of such integrals. We will need the notion of iterated integrals to do this.

Definition 3.23 *(Iterated integrals)*

If $f(x, y)$ is continuous on $\mathbb{R}$in the region $[a, b] \times [c, d]$ then the iterated integral is

$$\int\int_D f(x,y)\ \delta A = \int_a^b\int_c^d f(x,y)\ \delta y\delta x \tag{3.1}$$

$$= \int_a^b\int_c^d f(x,y)\ \delta x\delta y \tag{3.2}$$

You can think of iterated integrals as computing the area of one of the faces of the rectangle and the extending that to the 3rd dimension. Since we know that the integral in the one dimension case is an area, the iterated integral is equal to the double integral if $f(x, y)$ is integrable.

Remark 3.24 *If the iterated integral $f(x, y)$ is integrable on the region $D = [a, b] \times [c, d]$, then the iterated integral is equal to the double integral.*

3.3.2 Definite Integrals on General Regions

In general, we can integrate on general regions if we close our bounds correctly. For example, if we have a region bounded by 2 functions and the domain $x \in [a, b]$ is identical, then we would use

$$\int\int_D f(x,y)\ \delta x \delta y = \int_a^b \int_{\phi(x)}^{\phi(x)} f(x,y)\ \delta y \delta x$$

If the region is bounded by 2 functions by the range $y \in [c, d]$ is identical, then we would use

$$\int\int_D f(x,y)\ \delta x \delta y = \int_c^d \int_{\psi(y)}^{\psi(y)} f(x,y)\ \delta x \delta y$$

You can of course, compute the appropriate the bounds and swap the order of variable integration around.

Consider

$$\begin{aligned}
\int_0^1 \int_0^{x^2} (x+y)\ \delta y \delta x &= \int_0^1 xy + \frac{1}{2}y^2 \mid_0^{x^2}\ \delta x \\
&= \int_0^1 x^3 + \frac{1}{2}y^4\ \delta x \\
&= \frac{1}{4} + \frac{1}{10} = \frac{7}{20}
\end{aligned}$$

$$\begin{aligned}
\int_0^1 \int_{\sqrt{y}}^1 (x+y)\ \delta x \delta y &= \int_0^1 xy + \frac{1}{2}x^2 \mid_{\sqrt{y}}^1\ \delta y \\
&= \int_0^1 y + \frac{1}{2} - (y\sqrt{y} + \frac{1}{2}y)\ \delta y \\
&= \frac{1}{2}y^2 + \frac{1}{2}y - \frac{2}{5}y^{\frac{5}{2}} - \frac{1}{4}y^2 \mid_{\sqrt{y}}^1 \\
&= 1 - \frac{8}{20} - \frac{5}{20} = \frac{7}{20}
\end{aligned}$$

Now consider the following region

$$\begin{aligned}
\int_0^2 \int_x^1 xy\ \delta y \delta x &= \int_0^2 \frac{x}{2} - \frac{x^3}{2}\ \delta x \\
&= \frac{2^2}{4} - \frac{2^4}{8} = -1
\end{aligned}$$

How did we end up with a negative area? Well, it turns out that geometrically we are evaluating the one part of the integral that underneath a line. Hence we're getting the 'negative' value of the area.

Now, you can break the integral up as well with

$$\begin{aligned}
\int_0^1 \int_0^x xy\ \delta y \delta x + \int_1^2 \int_x^1 xy\ \delta y \delta x &= \int_0^1 \frac{1}{2}x^3\ \delta x + \int_1^2 \frac{1}{2}x - \frac{1}{2}x^3\ \delta x \qquad (3.3) \\
&= -1 \qquad (3.4)
\end{aligned}$$

Suppose we want to integral the x variable first. Then in order to get the same answer as above we take the negative of the first integral (region underneath the line)

$$\int_0^2 \int_y^2 xy\ \delta x \delta y = -\int_0^1 \int_0^y xy\ \delta x \delta y + \int_1^2 \int_y^2 xy\ \delta x \delta y \tag{3.5}$$

$$\tag{3.6}$$

which gets us our answer.

Remark 3.25 *Dealing with triple integrals is a similar story.*

Think of a unit ball

$$B := \{(x, y, z)\ |x^2 + y^2 + z^2 \leq 1\}$$

with integral

$$\int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \int_{\sqrt{1-x^2-y^2}}^{-\sqrt{1-x^2-y^2}} \delta z \delta y \delta x \tag{3.7}$$

which you can solve in Cartesian coordinates with a trig substitution.

3.3.3 Change of Variable

Proposition 3.26 *Change-of-variable for one dimension is relatively straightforward. You have*

$$\begin{aligned} x = g(u) &\longrightarrow \delta x = g'(u)\ \delta u \\ x = a &\longrightarrow u = g^{-1}(a) \\ x = b &\longrightarrow u = g^{-1}(b) \\ \Rightarrow \int_a^b f(x)\ \delta x &= \int_{g^{-1}(a)}^{g^{-1}(b)} f(g(u))g^{'}(u)\delta u \end{aligned}$$

Theorem 3.27 *(Change-of-variables for multiple variables)*

Let $x = x(u, v)$ and $y = y(u, v)$. Define a 1-1 mapping of a region $\mathcal{R}^*$ in the (u, v) plane onto a region $\mathcal{R}$ in the (x, y) plane so that the Jacobian determinant

$$\mathcal{J}(u, v) = \begin{pmatrix} \frac{\delta x}{\delta u} & \frac{\delta x}{\delta v} \\ \frac{\delta y}{\delta u} & \frac{\delta y}{\delta v} \end{pmatrix} \tag{3.8}$$

is never 0 in $\mathcal{R}^*$. The change of variable formula is

$$\int\int_{\mathcal{R}} f(x, y)\ \delta x \delta y = \int\int_{\mathcal{R}^*} f(x(u, v), y(u, v))\ |\det(J)_{u,v}|\ \delta u \delta v \tag{3.9}$$

Remark 3.28 *Different texts define the Jacobian differently. For this set of notes, we define the Jacobian as the matrix itself, the Jacobian determinant as the determinant.*

Let's look at a classic example.

$$\int_{-\infty}^{\infty} e^{-x^2}\ \delta x$$

The trick is to multiple the integral with $\int_{-\infty}^{\infty} e^{-y^2}\ \delta y$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-x^2-y^2}\ \delta x \delta y = \int_{0}^{2\pi}\int_{0}^{\infty} e^{-r^2} r\ \delta r \delta \theta$$

and integrate normally.

Remark 3.29 *The change-of-variables formula requires a 1-1 function in more than 1 variable. In one dimension, 1-1 is sufficient (but not necessary). The reason for this is that you need to be able to invert your* $g(u) = x$.

Theorem 3.30 *(Change-of-variable to polar coordinates) For* $(x, y, z) \to (r, \theta, z)$ *we have*

$$\begin{aligned} x &= r\cos\theta & r &= \sqrt{x^2+y^2} \\ y &= r\sin\theta & \theta &= \arctan\left(\frac{y}{x}\right) \\ \det(J)_{r,\theta} &= r \end{aligned}$$

The region is $W = [0, \infty) \times [0, 2\pi]$ *for* (r, θ).

Theorem 3.31 *(Change-of-variable to cylindrical coordinates) For* $(x, y) \to (r, \theta)$ *we have*

$$\begin{aligned} x &= r\cos\theta & r &= \sqrt{x^2+y^2} \\ y &= r\sin\theta & \theta &= \arctan\left(\frac{y}{x}\right) \\ z &= z & z &= z \\ \det(J)_{r,\theta,z} &= r \end{aligned}$$

The region is $W = [0, \infty) \times [0, 2\pi] \times (-\infty, \infty)$ *for* (r, θ, z).

Theorem 3.32 *(Change-of-variable to spherical coordinates) For* $(x, y, z) \to (r, \theta, z)$ *we have*

$$\begin{aligned} x &= \rho\cos\theta\sin\phi & \rho^2 &= x^2+y^2+z^2 \\ y &= \rho\sin\theta\sin\phi & \theta &= \arctan\left(\frac{y}{x}\right) \\ z &= \rho\cos\phi & \phi &= \arccos\left(\frac{z}{\sqrt{x^2+y^2+z^2}}\right) \\ \det(J)_{\rho,\theta,\phi} &= \rho^2\sin\phi \end{aligned}$$

The region is $W = [0, R] \times [0, 2\pi] \times [0, \pi]$ *for* (r, θ, ϕ).

Theorem 3.33 *(Change-of-variable of linear equations) If we have an matrix of equations* $\mathbf{x} = \mathbf{AU}$, *then the change of variable from* x *to* u

$$\left|\frac{\delta \mathbf{x}}{\delta \mathbf{u}}\right| = |\det(\mathbf{A})|$$

MATH2921 Vector Calc. & ODEs

4. Applications of Multiple Integrals

Applications

4.4 Applications

4.4.1 Centre of Mass

Definition 4.34 *(Density)*

We define the density of a substance over a region $\mathcal{R}$ as $\rho(x, y)$. This can be extended to three dimension. The density can be constant, which we denote $\rho(x, y) = c$.

Definition 4.35 *(Mass)*

The mass of the substance is defined as the integral of the density over the region $\mathcal{R}$, that is,

$$M = \int_{\mathcal{R}} \rho(\mathbf{x})\ \delta\mathbf{x}$$

Definition 4.36 *(Centre of Mass)*

We define the centre of mass as $\bar{\mathbf{x}} = (\bar{x_1}, \bar{x_2})$, where

$$\bar{x}_i = \frac{1}{M} \int_{\mathcal{R}} x_i \rho(\mathbf{x})\ \delta\mathbf{x} \tag{4.10}$$

Let's do a quick example. Suppose you have a square plate with side length s with the bottom left corner situated at the origin. Let the density be $\rho(x, y) = x^n + y^n$, $\in \mathbb{Z} \setminus 0$.

We have

$$\begin{aligned}
\int_0^s \int_0^s x^n + y^n\ \delta x \delta y &= \int_0^s x^n y + \frac{1}{n+1} y^{n+1}\Big|_0^s\ \delta y \\
&= \int_0^s x^n s + \frac{1}{n+1} s^{n+1}\ \delta y \\
&= \frac{1}{n+1} x^{n+1} s + \frac{1}{n+1} s^{n+1} x\Big|_0^2 \\
&= \frac{2}{n+1} s^{n+1}
\end{aligned}$$

Then

$$\begin{aligned}
\bar{x} &= \frac{1}{M}\int_0^s \int_0^s x(x^n + y^n)\ \delta x \delta y \\
&= \frac{1}{M}\int_0^s \frac{1}{n+2} s^{n+2} + \frac{1}{2} s^2 y^n)\ \delta y \\
&= \frac{1}{M} s^{n+3} \left(\frac{1}{n+2} + \frac{1}{2(n+1)} \right) \\
&= \frac{s(3n+4)}{4(n+2)}
\end{aligned}$$

Since the funciton is symmetrical,

$$\bar{y} = \frac{s(3n+4)}{4(n+2)}$$

This gives $(\bar{x}, \bar{y}) = \left(\frac{s(3n+4)}{4(n+2)}, \frac{s(3n+4)}{4(n+2)} \right)$

Note that if $s = 1$, then $(\bar{x}, \bar{y}) \to \left(\frac{3}{4}, \frac{3}{4}\right)$ and $\rho\left(\frac{3}{4}, \frac{3}{4}\right) = \left(\frac{3}{4}\right)^n + \left(\frac{3}{4}\right)^n \to 0$.

We can interpret this as the centre of mass tending to a point of zero density on the plate.

4.4.2 Surface Area of a Graph

Definition 4.37 *(Surface area)*

Let $z = f(x, y)$ be a function. The surface area of a graph over a region $\mathcal{R}$ is

$$SA = \int \int_{\mathcal{D}} \sqrt{1 + f_x^2 + f_y^2}\ \delta x \delta y$$

Definition 4.38 *(Arc length)*

It turns out that a graph $y = f(x)$ over a region $[a, b]$, the one dimension case, has arc length

$$\ell = \int_a^b \sqrt{1 + f'(x)^2}\ \delta x$$

So the formula for surface area of a curve is the 'same' for arc-length!

Note: The same integral is also the integral of the length of the normal vector over the region!

Consider the upper hemisphere of a unit sphere $z = \sqrt{1 - x^2 - y^2}$. We have

$$\begin{aligned}
f_x &= \frac{x}{\sqrt{1 - x^2 - y^2}} \\
f_y &= \frac{y}{\sqrt{1 - x^2 - y^2}}
\end{aligned}$$

and thus

$$\int\int_{\mathcal{D}} \sqrt{1+f_x^2+f_y^2}\ \delta x\delta y = \int\int_{\mathcal{D}} \sqrt{1+\frac{x^2}{1-x^2-y^2}+\frac{y^2}{1-x^2-y^2}}\ \delta x\delta y \tag{4.11}$$

$$= \int\int_{\mathcal{D}} \sqrt{\frac{1}{1-x^2-y^2}}\ \delta x\delta y \tag{4.12}$$

switching to polar coordinates yields

$$\int_0^{2\pi}\int_0^1 \frac{r}{\sqrt{1-r^2}}\ \delta r\delta\theta = \int_0^{2\pi} 1\ \delta\theta$$
$$= pi$$

4.4.3 Average Value of a Function

Definition 4.39 *(Average value of a function)*

We have the area (or volume in three dimension) $\int_R 1\ \delta\mathbf{x}$ *and*

$$Avg = \frac{\int_R f(\mathbf{x})\ \delta\mathbf{x}}{\int_R 1\ \delta\mathbf{x}}$$

Consider this example. We have a unit square in two dimension, and the distance squared from the origin to a point in the square is $d^2 = x^2 + y^2$. The average distance is

$$1\cdot\int_0^1\int 0^1 x^2+y^2\ \delta x\delta y = 1\cdot\int_0^1 \frac{1}{3}x^3+xy^2\ \delta y$$
$$= \frac{1}{3}+\frac{1}{3} = \frac{2}{3}$$

for three-dimension we have

$$1\cdot\int_0^1\int 0^1 x^2+y^2+z^2\ \delta x\delta y\delta z = 1\cdot\int_0^1 \frac{1}{3}x^3+xy^2+xz^2\ \delta y\delta z$$
$$= \frac{1}{3}+\frac{1}{3}+\frac{1}{3} = 1$$

for n-dimension, we have

$$1\cdot\int_0^1\int_0^1 x_1^2+x_2^2+\cdots+x_n^2\ \delta x_1\delta x_2\ldots\delta x_n = \frac{1}{3}+\frac{1}{3}+\cdots+\frac{1}{3} = \frac{n}{3}$$

We this is counterintuitive. We have distance is increasing with dimension, but the volume of the n-dimensional hypercube is actually constant!

This is known as the curse of dimensionality.

Another example. Suppose you have a circle of radius R and a point $P = (0,0,D)$. What is the average distance-squared of a point (x,y,z) in the ball from the point P?. The distance function is $d^2(x,y,z) = x^2+y^2+(z-D)^2$.

We have

$$\int\int\int_B \delta x\delta y\delta z = \frac{4}{3}\pi R^3.$$

For the other integral, let's switch to spherical coordinates.

$$\begin{aligned}
\int\int\int_B x^2 + y^2 + (z-D)^2\ \delta x\delta y\delta z &= \int_0^R \int_0^{2\pi} \int_0^{\pi} (\rho^2 - 2D\rho\cos(\phi) + D^2)\rho^2 \sin(\phi)\ \delta\phi\delta\theta\delta\rho \\
&\Rightarrow \int_0^{\pi} \sin(\phi)\ \delta\phi = 2 \text{ and } \int_0^{\pi} \sin(2\phi)\ \delta\phi = 0 \\
&= \int_0^R \int_0^{2\pi} 2\rho^4 + 2D^2\rho^2\ \delta\theta\delta\rho \\
&= 4\pi\left(\frac{R^5}{5} + \frac{D^2R^2}{3}\right)
\end{aligned}$$

Hence the average distance squared is

$$\frac{3}{5}R^2 + D^2$$

4.4.4 Curves in Space

Our previous material dealt with functions that map from $f : \mathbb{R}^n \to \mathbb{R}$. For curves in space, these functions $\sigma : \mathbb{R} \to \mathbb{R}^n$. These are called parametric curves.

Think about $g(t) = (\sin(t), 2\cos(t))$. What does this trace out? An ellipse with Cartesian formula $x^2 + \frac{y^2}{4} = 1$!

In general, a parametric curve in space consists of functions $(x(t), y(t), z(t)) = (f(t), g(t), h(t))$. Consider the parametric curve $(x(t), y(t), z(t) = (3t + 2, 8t - 1, t)$. What kind of curve is it?

Well, fix the point $t = 0$, we get $(2, -1, 0)$. Hence we have $P = (2, -1, 0) + t(3, 81)$ which is a straight line in the direction of $(3, 8, 1)$.

$\sigma(t) = (x(t), y(t))$ is called a **vector function**.

Definition 4.40 *(Vector function)*

A vector function of one variable is a rule σ that associates a vector $\tau = \sigma(t)$ in space $\forall r \in [a, b] \subseteq \mathbb{R}$.

The derivative of $\sigma(t)$ is $\sigma'(t) = (x'(t), y'(t), z'(t))$, a tangent vector to the curve $\sigma(t)$.

Definition 4.41 *(Arc length using parametrised curve)*

If $\sigma(t)$ is a parametric curve $\sigma : [a, b] \in \mathbb{R}$, the length of the curve is

$$L = \int_a^b \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}\ \delta t$$

The arc length is probably painful to compute (given the square root), but this concept is good to know for theoretical purposes.

Proposition 4.42 *Suppose $\sigma'(t) \neq 0\ \forall t$. You can re-parametrise $\sigma(t)$ with $t = t(s) \in [a,b]$ as $\sigma(t(s))$. This is a curve that covers the same set of points but moves through the points at different speeds (depending on $t(s)$.*

We can choose $t(s)$ to be arbitrary but we can actually choose a function makes the re-parametrisation move at unit speed along the curve.

Using the chain rule, we have

$$\frac{\delta}{\delta s}\sigma(t(s)) = \sigma'(t)t'(s) \tag{4.13}$$

Hence the length of the new vector is

$$|\sigma'(t)||t'(s)| = \left|\sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}\right|\left|\frac{\delta t}{\delta s}\right| \tag{4.14}$$

Theoretically you can choose any $t(s)$ but we pick $s(t)$ instead. In this case let's define it as

$$\begin{aligned} s &= \int_a^t \sqrt{x'(\tau)^2 + y'(\tau)^2 + z'(\tau)^2}\ \delta\tau \\ \Rightarrow \frac{\delta s}{\delta t} &= \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2} \end{aligned}$$

Unfortunately you can't just flip the "fraction", because it's not a fraction.

So in order to obtain an expression for $\frac{\delta t}{\delta s}$, we define $s(t(s)) = s$ and hence

$$\begin{aligned} \frac{\delta}{\delta s}\left[s(t(s))\right] &= s'(t(s))t'(s) \\ \frac{\delta}{\delta s}s &= \frac{\delta s}{\delta t}\cdot\frac{\delta t}{\delta s} \\ &= 1 \\ \Rightarrow \frac{\delta t}{\delta s} &= \frac{1}{\frac{\delta s}{\delta t}} \\ &= \frac{1}{\sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}} \end{aligned}$$

So at the end we get

$$\begin{aligned} |\sigma'(t)||t'(s)| &= \left|\sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}\right| \frac{1}{\left|\sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}\right|} \\ &= 1 \\ \Rightarrow |\sigma'(s)| &= 1 \end{aligned}$$

That is, we have re-parametrised the curve from $\sigma(t)$ to $\sigma(s)$ and now $\sigma(s)$ has unit speed 1.

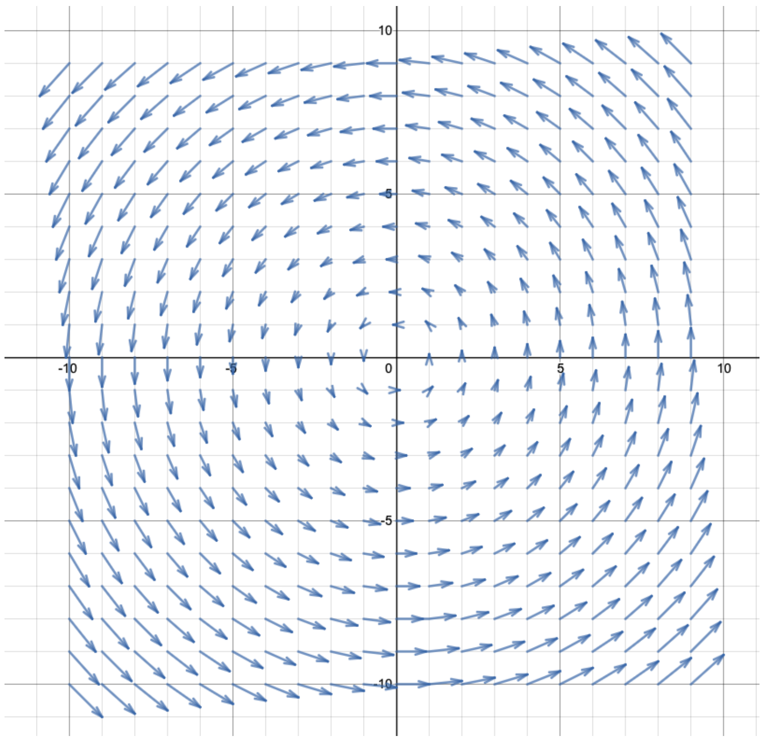
4.4.5 Vector Fields

So what happens if you're dealing with a curve that is more than one variables?

So instead of $\sigma(t) = (x(t), y(t))$ you have $\mathbb{F}(x,y) = (f(x,y), g(x,y))$. Well at each point you'd have a vector $(f(x,y), g(x,y))$.

We call this a vector field. Graphically, we can represent a vector field by plotting a series of arrows at each point. The arrows at (x, y) represent the direction (and magnitude) the vector is pointing to at that point.

Let's look at an example. $\mathbb{F} = (-y, x)$.



Definition 4.43 *(Line integral of a vector field) The line integral, or path integral of* $\mathbb{F}$ *along* $\sigma(t)$ *from* a *to* b *is defined as*

$$\int_a^b \mathbb{F}(\sigma(t)) \cdot \sigma'(t) \; \delta t$$

The intuition is this: if you have position dependent vector field and a curve. The line integral is the work done by that vector field along the curve.

So take a curve and integral w.r.t to the vector field. The resulting line integral shows the relative amount the vector in $\mathbb{F}$ is pointing in the direction of the tangent vector of $\sigma'(t)$. Think of the geometric representation of the dot product in linear algebra. What does that formula look like?

Let's do a few examples.

Let $\mathbb{F} = (y - y^2, 0)$.

Suppose you have $\sigma_1 : [0, 1] \to S$, $\sigma_t(t) = (0, t)$ and $\sigma_2 : [0, 1] \to S$, $\sigma_t(t) = (t, t)$.

For σ_1, we have

$$\begin{aligned}\int_0^1 \mathbb{F}(\sigma_1(t)) \cdot \sigma_1'(t) &= \int_0^1 \begin{pmatrix} t - t^2 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \ \delta t \\ &= \int_0^1 0 \ \delta t = 0\end{aligned}$$

For σ_2, we have

$$\begin{aligned}\int_0^1 \mathbb{F}(\sigma_2(t)) \cdot \sigma_2'(t) &= \int_0^1 \begin{pmatrix} t - t^2 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \ \delta t \\ &= \int_0^1 t - t^2 \ \delta t \\ &= \frac{1}{2} - \frac{1}{3} = \frac{1}{6}\end{aligned}$$

Remember how we said before that the vector field "helps" the curve.

Suppose we have $\sigma : [0, T] \to R$, with $\mathbb{F} = \begin{pmatrix} -y \\ x \end{pmatrix}$ and $\sigma(t) = (\cosh(t), \sinh(t))$. Then with $\sigma(0) = (1, 0)$ and $\sigma(T) = (\cosh(T), \sinh(T))$ we have

$$\int_0^T \begin{pmatrix} -\sinh(t) \\ \cosh(t) \end{pmatrix} \cdot \begin{pmatrix} \sinh(t) \\ \cosh(t) \end{pmatrix} \ \delta t = \int_0^T \ \delta t \tag{4.15}$$

$$= T \tag{4.16}$$

Consider $\sigma(0) = (1, 0)$ at then $\mathbb{F}(\sigma(0)) = (0, 1) = \sigma'(0)$. Hence at $t = 0$ the vectors are the same, and thus the projection of $\mathbb{F}$ in that direction is just the length of the curve σ'.

For any other point where the vectors aren't parallel, you can take the dot product to get the projection of $\mathbb{F}$ in the direction of σ'.

MATH2921 Vector Calc. & ODEs

5. Path Integrals and Vector Fields

Path Integrals

5.5 Path Integrals

5.5.1 Reparametrising Path Integrals

Recall that a path integral of $\mathbb{F}$ along $\sigma(t)$ from a to b is defined as

$$\int_a^b \mathbb{F}(\sigma(t)) \cdot \sigma'(t)\ \delta t$$

Now, a particular geometric curve can remain the same even if you change the parametrisation of that curve, think $\sigma(t) = (\sin(t), \cos(t))$ and $\sigma^* = (\sin(2t), \cos(2t))$ on $t \in [0, 2\pi]$.

However, their parametrisation is different. So the question is, to what extent does its parametrisation affect the path integral?

We will need to look change-of-variables.

Theorem 5.44 *The line integral of a vector field $\mathbb{F}$ along a curve σ is independent of its parametrisation.*

(Sketch of proof)

Suppose you have a differentiable parametrised curve

$$\int_{t_1}^{t_2} \mathbb{F}(\sigma(t)) \cdot \sigma'(t)\ \delta t$$

where $\mathbb{F}$ is a vector field defined on some region $\mathcal{D} \subseteq \mathbb{R}^n$ with $\sigma : [t_1, t_2] \to \mathcal{D}$.

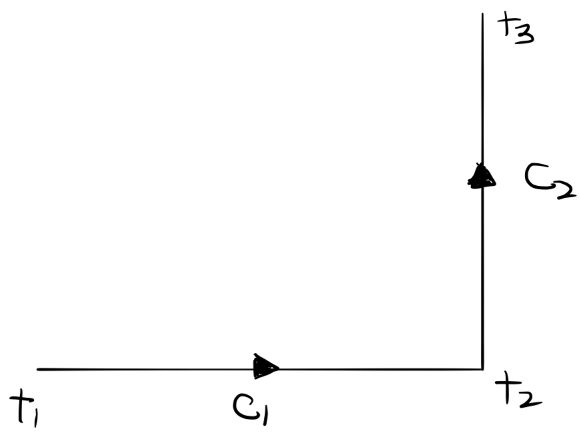
Choose an arbitrary one-to-one function $t = g(u)$ with $\delta t = g'(u)\ \delta u$ and start and endpoints $g_{u_1} = t_1$ and $g_{u_2} = t_2$. We see that

$$\begin{aligned}
\int_{t_1}^{t_2} \mathbb{F}(\sigma(t)) \cdot \sigma'(t)\ \delta t &= \int_{u_1}^{u_2} \mathbb{F}(\sigma(g(u))) \cdot \sigma'(g(u)) \cdot g'(u)\ \delta u \\
\textit{and reparametrise } \alpha(u) = \sigma(g(u)) \\
\Rightarrow &= \int_{u_1}^{u_2} \mathbb{F}(\alpha(u)) \cdot \alpha'(u)\ \delta u
\end{aligned}$$

The line integral of a vector field $\mathbb{F}$ along a curve σ is **independent** of the parametrisation of σ. Thus we are able to define a line integral along a geometric curve.

This result allows us to define line integrals on curves that aren't necessarily differentiable everywhere.

For example, consider the line integral below.



For the path integral $\sigma_3(t)$ we have $t \in [t_1, t_3]$, the derivative $\sigma_3(t)$ exists except at t_2. We can rewrite it as $t \in [t_1, t_2) \cup (t_2, t_3]$ in such a way that we have $\sigma_3' = \sigma_1'(t)$ for $t \in [t_1, t_2)$ and $\sigma_3' = \sigma_1'(2)$ for $t \in (t_2, t_3]$.

The integral is well-defined;

$$\int_{t_1}^{t_3} \mathbb{F}(\sigma_3(t))\sigma_3'(t)\ \delta t = \int_{t_1}^{t_2} \mathbb{F}(\sigma_1(t))\sigma_1'(t)\ \delta t + \int_{t_2}^{t_3} \mathbb{F}(\sigma_2(t))\sigma_2'(t)\ \delta t$$

The curves C_1 and C_2 is called the **image** of $\sigma_3(t)$ from $[t_1, t_3]$.

Definition 5.45 *(Piecewise smooth)*

A curve C is piecewise smooth if it can be written as the union of a finite number of smooth curves $C_1, C_2, ..., C_n$.

We refine this later in terms of functions but for now just think of it as a union of lines.

Remark 5.46 *If we want to evaluate the path integral travelling in the opposite direction, we simply negate the integral*

$$\int_{-C} \mathbb{F}(\sigma(t)) \cdot \sigma'(t)\ \delta t = -\int_{C} \mathbb{F}(\sigma(t)) \cdot \sigma'(t)\ \delta t$$

Definition 5.47 *(Differential one form)*

Suppose you have a vector field $\mathbb{F} = (f_1(x_1, ..., x_n), f_2(x_1, ..., x_n), ..., f_n(x_1, ..., x_n))$ on a domain $\mathcal{D} \subseteq \mathbb{R}^n$ with a piecewise smooth curve C parametrised by $\sigma(t) = (x_1(t), x_2(t), ..., x_n(t)$, for $\sigma : [t_1, t_2] \to \mathcal{D}$. Then $\sigma'(t) = (\frac{\delta x_1}{\delta t}, \frac{\delta x_2}{\delta t}, ..., \frac{\delta x_n}{\delta t})$.

we define the differential one form as

$$\int_C f_1\delta x_1 + f_2\delta x_2 + ... + f_n\delta x_n = \int_{t_1}^{t_2} \left[f_1 \cdot \frac{\delta x_1}{\delta t} + f_2 \cdot \frac{\delta x_2}{\delta t} + ... + f_n \cdot \frac{\delta x_n}{\delta t}\right] \delta t$$
$$= \int_{t_1}^{t_2} \mathbb{F}(\sigma(t)) \cdot \sigma'(t)\ \delta t$$

5.5.2 Conservative Vector Fields

Definition 5.48 *(Path independent integral)*

A path independent integral is an integral that only requires the endpoints to be evaluated.

Geometrically, this means you can draw any arbitrary path from the start and end point (say t_1 and t_2) and its path integral will be identical.

Definition 5.49 *(Convservative vector field)*

A vector field $\mathbb{F}$ is called conservative if any two path integrals of two different curves in the domain $\mathbb{F}$ with identical endpoints are equal. That is,

$$\int_{C_1} \mathbb{F}(r) \cdot\ \delta r = \int_{C_2} \mathbb{F}(r) \cdot\ \delta r$$

Proposition 5.50 *Loosely speaking, Path independent $\iff$ conservative. The idea of path independent refers to path integrals whereas conservative concerns vector fields.*

Lemma 5.51 *A vector field $\mathbb{F}$ is conservative $\iff$ the path integral of $\mathbb{F}$ along any closed curve is zero.*

(Sketch of proof)

Let C be a closed curve and pick any two arbitrary points A and B. Let $C = C_1 - C_2$ where C_1 and C_2 both go from A to B.

"$\Rightarrow$" Since the vector field is conservative, we have

$$\int_{C_1} \mathbb{F}(r) \cdot \delta r = \int_{C_2} \mathbb{F}(r) \cdot \delta r$$

Now, the closed curve starting and ending at A is

$$\int_{C_1} \mathbb{F}(r) \cdot \delta r + \int_{-C_2} \mathbb{F}(r) \cdot \delta r = \int_{C_1} \mathbb{F}(r) \cdot \delta r - \int_{C_2} \mathbb{F}(r) \cdot \delta r$$
$$= 0$$

"$\Leftarrow$" Suppose the path integral of all closed curves in $\mathbb{F}$ are zero. Then, we have

$$\int_{C_1} \mathbb{F}(r) \cdot \delta r + \int_{-C_2} \mathbb{F}(r) \cdot \delta r = 0$$

Then, we have

$$\int_{C_1} \mathbb{F}(r) \cdot \delta r = \int_{C_2} \mathbb{F}(r) \cdot \delta r$$

as required.

Remark 5.52 *We consider only geometric curves whose line integrals pass through each point once (and there are no crossovers).*

Definition 5.53 *(Gradient vector field)*

A vector field $\mathbb{F}$ *is called a gradient vector field if* $\exists f : \mathbb{R}^n \to \mathbb{R}^n$ *s.t.*

$$\mathbb{F} = \nabla \mathbf{f}$$

Theorem 5.54 *Gradient vector fields are conservative.*

(Sketch of proof)

Let C *be a closed curve and* $\mathbb{F} = \nabla \mathbf{f}$ *be a gradient vector field. Let* $\sigma(t)$ *be a parametrisation of* C $t \in [t_1, t_2]$*. We have*

$$\begin{aligned} \int_{t_1}^{t_2} \mathbb{F}(\sigma(t)) \cdot \sigma'(t)\ \delta t &= \int_{t_1}^{t_2} \nabla \mathbf{f}(\sigma(t)) \cdot \sigma'(t)\ \delta t \\ &= \int_{t_1}^{t_2} \frac{\delta}{\delta t} \left(\mathbf{f}(\sigma(t))\right)\ \delta t \\ &= \mathbf{f}(\sigma(t_2)) - \mathbf{f}(\sigma(t_1)) \\ &= 0 \ \text{ because } C \text{ is a closed curve.} \end{aligned}$$

Corollary 5.55 *If* C *is a piecewise smooth curve from* A *to* B *and* $\mathbb{F}$ *is a gradient vector field, then*

$$\begin{aligned} \int_C \mathbb{F}(r) \cdot \delta r &= \int_C \nabla \mathbf{f}(r) \cdot \delta r \\ &= \mathbf{f}(B) - \mathbf{f}(A) \end{aligned}$$

In general, if C *is any piecewise smooth path from* A *to* B *and* $\mathbb{F}$ *is a gradient vector field, then*

$$\begin{aligned} \int_C \mathbb{F}\ \delta r &= \int_C \nabla f(r)\ \delta r \\ &= f(B) - f(A) \end{aligned}$$

The theorem is nice but how would we know if a given vector field is a gradient vector field? We give a necessary condition for answering this question.

Definition 5.56 *(Scalar potential of* $\mathbb{F}$*)*

The scalar potential of $\mathbb{F}$*, or anti-derivative, primitive, is a function* f *s.t.* $\mathbb{F} = \nabla f$*.*

Remark 5.57 *In multiple variable calculus, only gradient vector fields can have an anti-derivative. It turns out all gradient vector fields have one.*

Now, we show that not only are gradient vector fields conservative BUT all conservative vector fields are gradient vector fields!

Theorem 5.58 *Conservative vector fields are gradient vector fields.*

(Sketch of proof)

Let $\mathbb{F}$ be a smooth vector field defined on a domain $\mathcal{D}$. Let $O \in \mathcal{D}$. For any $P \in \mathcal{D}$ define a curve $\sigma : [0,1] \to \mathcal{D}$ s.t. $\sigma(0) = O$ and $\sigma(1) = P$. Define

$$f(P) = \int_\sigma \mathbb{F}(r)\ \delta r$$

Since $\mathbb{F}$ is conservative, the integral is well defined (and thus so is the function). Also note that by definition we have $\mathbb{F} = \nabla f$. We claim that

$$\int_C \mathbb{F}(r)\ \delta r = \int_C \nabla f(r)\ \delta r \tag{5.17}$$

Let M be a 3rd point in the domain $\mathcal{D}$. Define C_1 to be a curve from M to O, C_2 a curve from M to P and C_* be the curve from O to P. Using the fact that $\mathbb{F}$ is conservative, we have

$$\begin{aligned} 0 &= \int_{C_1+C-C_2} \delta r \\ &= \int_C \mathbb{F}(r)\ \delta r + \int_{C_1} \mathbb{F}(r)\ \delta r + \int_{-C_2} \mathbb{F}(r)\ \delta r \\ \Rightarrow \int_C \mathbb{F}(r)\ \delta r &= \int_{C_2} \mathbb{F}(r)\ \delta r - \int_{C_1} \mathbb{F}(r)\ \delta r \\ &= f(P) - f(O) \end{aligned}$$

So it's nice to know this, but how do we actually test whether a given vector field is actually a gradient field?

5.5.3 Some Misc. Definitions

Before continuing, we define a few different types of paths and regions.

Definition 5.59 *(Simple path)*

A path C is simple if it does not cross itself.

Definition 5.60 *(Closed path)*

A path C is closed if its initial and final points are the same point.

Definition 5.61 *(Open region)*

A region D is open if it does not contain its boundary points.

Definition 5.62 *(Connected region)*

A region D is connected if any two points can be connected by a path that lies completely in the region in the region

Definition 5.63 *(Simply connected region)*

A region D is simply connected if it is connected and contains no holes.

5.5.4 Cross Derivatives Test

So how do we actually test whether a given vector field is actually a gradient field? We use the cross-derivatives test.

Theorem 5.64 *(Cross derivatives test)*

For simplicity, we state the idea for the two dimension case. Let $\mathbb{F}$ be a smooth vector field defined on $\mathbb{R}^2$. Suppose

$$\mathbb{F} = \begin{pmatrix} p(x,y) \\ q(x,y) \end{pmatrix}$$

Then $\mathbb{F}$ is a gradient vector field $\iff$

$$p_y = q_x \iff \frac{\delta p}{\delta y} = \frac{\delta q}{\delta x}$$

(Sketch of proof)

"$\Rightarrow$" Suppose $\mathbb{F} = \nabla \mathbf{f}$. Then by the definition of both functions, we have

$$\begin{aligned} p &= f_x \\ q &= f_y \end{aligned}$$

with the equality of mixed partials for continuous functions up to second derivative (we assume this)

$$\begin{aligned} p_y &= f_{xy} \\ q_x &= f_{yx} \\ \Rightarrow p_y &= q_x \end{aligned}$$

"$\Leftarrow$" Suppose that $p_y = q_x$. Then we can construct a function

$$\bar{f}(x,y) = \int_o^x p(t,y)\ \delta t$$

since we know that $p(x,y) = f_x$. Now, the (scalar) potential $f(x,y)$ must satisfy

$$f(x,y) = \bar{f}(x,y) + g(y)$$

where $g(y)$ is the constant of integration w.r.t to x.

$$\begin{aligned} f_y &= \bar{f}_y + g'(y) \\ &= q(x,y) \\ \Rightarrow g'(x,y) &= q(x,y) - \bar{f}_y(x,y) \end{aligned}$$

now we need to show that this is independent of x, i.e. the derivative w.r.t. to x equals zero. Once we show this, we can integrate w.r.t. to y to obtain $g(y)$ which would give a concrete expression for $f(x, y)$.

$$\begin{aligned}
\frac{\delta}{\delta x}(g'(y)) &= \frac{\delta}{\delta x}(q(x,y) - \bar{f}_y(x,y)) \\
&= q_x - \frac{\delta}{\delta x}\frac{\delta}{\delta y}\int_o^x p(t,y)\ \delta t \\
&= q_x - \frac{\delta}{\delta x}\int_o^x p_y(t,y)\ \delta t \\
&= q_x - p_y \\
&= 0 \ \ \textit{by assumption}
\end{aligned}$$

To show that the derivative can be moved inside the integral requires some analysis but we will not delve into that here. Hence we have

$$\begin{aligned}
g'(y) &= q(x,y) - \bar{f}_y(x,y) \\
&= h(y) \\
g(y) &= \int h(y)
\end{aligned}$$

and we are done.

Definition 5.65 *(Exact differential form)*

Recall we had the differential one form for $\mathbb{F} = \begin{pmatrix} P \\ Q \end{pmatrix}$ where $\int_C f_1 \delta x_1 + f_2 \delta x_2$.

If $\mathbb{F}$ is a gradient vector field with (scalar) potential f s.t. $f_x = P$ and $f_y = Q$ then for a vector

$$P \cdot \delta x + Q \cdot \delta y$$

is called the exact differential form.

Definition 5.66 *(Scalar curl)*

Given a vector field $\mathbb{F}$ and the standard basis of linear directions ∇, define the (scalar) curl for three dimension as

$$\begin{aligned}
\nabla \times \mathbb{F} &= (\delta x\ \mathbf{i}, \delta y\ \mathbf{j}, \delta z\ \mathbf{k}) \times (P, Q, R) \\
&= (R_y - Q_z, P_z - R_x, Q_x - P_y)
\end{aligned}$$

For two dimension, we have $\nabla \times \mathbb{F} = Q_x - P_y$.

What is the intuition of a (scalar) curl? Suppose you place a paddle at any point in the vector field. The curl tells you how much the vector field would spin the paddle at that point.

Remark 5.67 *If $\mathbb{F}$ is a gradient vector field that is twice differentiable and continuous, then we obtain equality of mixed partials and hence*

$$\nabla \times (\nabla f) = 0$$

That is, the curl of the gradient vector is zero.

5.5.5 Green's Theorem

Theorem 5.68 *Suppose $\mathcal{D}$ is a region in the plane $\mathbb{R}^2$ that is simply-connected. Suppose further that $\mathcal{D}$ has a boundary denoted by C which is oriented anti-clockwise. Then for a vector field $\mathbb{F} : \mathcal{D} \to \mathbb{R}^2$, $\mathbb{F} = (P, Q)$ we have*

$$\int\int_{\mathcal{D}} \left(\frac{\delta Q}{\delta x} - \frac{\delta P}{\delta y} \right) \delta x \delta y = \oint_C P\delta x + Q\delta y$$

In words, Green's theorem says that the double integral of the (scalar) curl of a planar vector field over a simply-connected region is equal to the line integral of that vector field along the boundary of $\mathcal{D}$, oriented anti-clockwise.

Definition 5.69 *(Vector area)*

In three-dimensional geometry, the vector area is defined as the unit normal scaled by the finite scalar area A. That is,

$$\mathbf{A} = \mathbf{n} \cdot \mathcal{A}$$

where $\mathcal{A}$ is a set of flat facet areas.

Definition 5.70 *(Flux of a vector field across the δA)*

Define a vector field $\mathbb{F} = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$. Suppose we have the tangent vector $\sigma'(t) = (x'(t), y'(t))$ and we construct a outward pointing normal vector to $\sigma'(t)$ as follows:

$$\begin{pmatrix} y'(t) \\ -x'(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix}$$

So here we have $\delta\mathcal{A}$ is the vector area of the surface of $\mathcal{A}$ and the normal vector $\mathbf{n}$ is the normal

Then the area of the region is

$$\begin{aligned} \mathcal{A} &:= \frac{1}{2} \oint_{\delta A} x \; \delta y - y \; \delta x \\ &= \frac{1}{2} \oint_{\delta A} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \cdot \begin{pmatrix} y'(t) \\ -x'(t) \end{pmatrix} \; \delta t \\ &= \frac{1}{2} \oint_{\delta A} \mathbb{F} \cdot \mathbf{n} \; \delta \underline{s} \end{aligned}$$

where

$$\mathbf{n} := \frac{1}{\sqrt{(x')^2 + (y')^2}} (y', -x')$$

and

$$\delta \underline{s} := \sqrt{(x')^2 + (y')^2} \; \delta t \tag{5.18}$$

The integral is called the flux of $\mathbb{F}$ across δA.

Remark 5.71 *(Flux vs. work of $\mathbb{F}$ across C)*

Previously, we've mentioned the intuition of vector fields being how much "help" you get when travelling in a certain direction. It turns out that flux and work are related. The flux is

$$\int_C \mathbb{F} \cdot \mathbf{n}\, \delta \underline{s}$$

while work is defined as

$$\int_C \mathbb{F} \cdot \mathbf{T}\, \delta \underline{s}$$

where

$$\begin{aligned} \mathbf{T} &:= \frac{1}{\sqrt{(x')^2 + (y')^2}}(x', y') \\ \Rightarrow \mathbf{n} &:= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \mathbf{T} \end{aligned}$$

Remark 5.72 *Using Green's theorem which equates the (scalar) curl with the boundary, we see that the work done by $\mathbb{F}$ along the boundary is equal to the (scalar) curl of $\mathbb{F}$ over the region.*

We can reformulate flux using Green's theorem. Suppose we have $\mathbb{F}^* = \begin{pmatrix} M \\ N \end{pmatrix}$

$$\begin{aligned} \oint_{\delta \mathcal{D}} \mathbb{F}^* \cdot \mathbf{n}\, \delta \underline{s} &= \oint_{\delta \mathcal{D}} \begin{pmatrix} M \\ N \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} \delta t \\ &= \oint_{\delta \mathcal{D}} M \delta y - N \delta x \\ &= \int\int_{\mathcal{D}} \left(\frac{\delta M}{\delta x} + \frac{\delta N}{\delta y} \right) \delta x \delta y \end{aligned}$$

Definition 5.73 *(Divergence of F)*

The divergence of a vector field $\mathbb{F}$ is defined as

$$\nabla \cdot \mathbb{F}$$

It is pronounced "del dot $\mathbb{F}$". For the case above

$$\begin{aligned} \nabla \cdot \mathbb{F} &= (\delta x, \delta y) \cdot (M, N) \\ &= M_x + N_y \end{aligned}$$

Corollary 5.74 *From the concepts mentioned above, we can see that the flux formulation of Green's theorem states that the flux of $\mathbb{F}$ across $\delta \mathcal{D}$ is equal to the divergence of $\mathcal{F}$ over $\mathcal{D}$.*

Remark 5.75 *What happens when divergence equals zero?*

$$\begin{aligned} \nabla \cdot \mathbb{F} &= M_x + N_y \\ &= 0 \\ \Rightarrow M\delta y - N\delta x &= 0 \end{aligned}$$

This is an exact differential form and thus $-N = g_x$ and $M = g_y$.

Definition 5.76 *(Stream function)*

In the above remark, $g(x, y)$ is called the stream function because

$$\begin{aligned} \mathbb{F} \cdot \nabla g &= (M, N) \cdot (-N, M) \\ &= 0 \end{aligned}$$

Definition 5.77 *(Conservative and divergence free vector fields)*

$$\begin{aligned} \mathbb{F} &= \nabla \mathbf{f} \\ \nabla \cdot \mathbb{F} = 0 \Rightarrow \nabla \cdot \nabla \mathbf{f} &= f_{xx} + f_{yy} = 0 \end{aligned}$$

Definition 5.78 *(Laplace's equation)*

The following function is called Laplace's equation.

$$\begin{aligned} 0 &= f_{xx} + f_{yy} \\ &= \nabla^2(f) \\ &= \Delta f \end{aligned}$$

Definition 5.79 *(Cauchy-Riemann equations) If $\nabla \mathbf{f} = (f_x, f_y) = (M, N)$ is divergence free then it'll have a stream function*

$$\begin{aligned} f_x &= g_y = M \\ f_y &= -g_x = N \end{aligned}$$

This set of equations is called the Cauchy-Riemann equations, where $f = Re(f)$ and $g = Im(f)$.

When $f = c$ a constant, the lines are equipotentials.

When $g = c$ a constant, the lines are called streamlines.

MATH2921 Vector Calc. & ODEs

6. Stokes Theorem and Applications

Stokes Theorem and Applications

6.6 Stokes Theorem and Applications

6.6.1 The Surface Integral Revisited

We can define the surface integral using

$$\int\int_{\mathcal{D}} \sqrt{1+(f_x)^2+f_y)^2}\,\delta x\delta y = \int\int_{\mathcal{D}} \delta S$$

where $\delta S := \sqrt{1+(f_x)^2+f_y)^2}\,\delta x\delta y$.

Proposition 6.80 *Suppose you have a parametrisation of the surface* $(x(u,v), y(u,v), z(u,v))$ *where* $f : \mathcal{D} \to \mathcal{S}$*, where* $\mathcal{D} \subseteq \mathbb{R}^2$ *and* $\mathcal{S} \subseteq \mathbb{R}^3$*.*

Define

$$T_u := \frac{\delta x}{\delta u} = (x_u, y_u, z_u)$$
$$T_v := \frac{\delta x}{\delta v} = (x_v, y_v, z_v)$$

Then the surface area of $\mathcal{S}$ *is*

$$\int\int_{\mathcal{D}} \delta S = \int\int_{\mathcal{D}} |T_u \times T_v|\,\delta u\delta v$$

Theorem 6.81 *The flux of a vector field* $\mathbb{F} = (P,Q,R)$ *across a surface* $z = f(x,y) \iff (x(u,v), y(u,v), z(u,v))$ *is the aggregate projection of* $\mathbb{F}$ *onto the outward normal direction of* $\mathcal{S}$*. That is,*

$$\begin{aligned}\int\int_{\mathcal{S}} \mathbb{F}\cdot\mathbf{n}\,\delta S &:= \int\int_{\mathcal{D}} (P,Q,R)\cdot\frac{(-f_x,-f_y,1)}{\sqrt{1+(f_x)^2+f_y)^2}}\cdot\sqrt{1+(f_x)^2+f_y)^2}\,\delta x\delta y\\ &= \int\int_{\mathcal{D}} (-fxP - f_yQ + R)\,\delta x\delta y\end{aligned}$$

where $\mathcal{S} := \{(x,y,f(x,y)\ |(x,y)\in\mathcal{D}\}$*.*

When you parametrise the surface using $(x(u,v), y(u,v), z(u,v))$*, the formula becomes*

$$\begin{aligned}\int\int_{\mathcal{S}} \mathbb{F}\cdot\mathbf{n}\,\delta S &:= \int\int_{\mathcal{D}} (P,Q,R)\cdot\frac{(T_u\times T_v)}{|T_u\times T_v|}\cdot|T_u\times T_v|\,\delta u\delta v\\ &= \int\int_{\mathcal{D}} (P,Q,R)\cdot|T_u\times T_v|\,\delta u\delta v\end{aligned}$$

Remark 6.82 *You always need to check that $(T_u \times T_v)$ is pointing outwards. It might require you to actual compute the value and see.*

Remark 6.83 *The work done by a vector field $\mathbb{F}$ along a curve C is denoted as $\int_C \mathbb{F} \cdot \mathbf{T}\, \delta \underline{s}$ as we have seen, but this doesn't exactly have a single generalisation. However, the flux of a vector field $\mathbb{F}$ does general to surfaces in $\mathbb{R}^n$.*

What this means is that there is actually only a single outward direction you can point to, which defined in three-dimension is what we saw earlier.

Definition 6.84 *(Surface integral of $\mathbb{F}$ over S)*

The flux of $\mathbb{F}$ across S is also known as the surface integral of $\mathbb{F}$ over S.

6.6.2 Stokes Theorem

Theorem 6.85 *(Stokes theorem)*

Let $\mathcal{D}$ be a region in the plane to which Green's theorem applies. Let $\mathcal{S}$ be the surface $z = f(x, y)$ where $f \in C^2(\mathcal{D})$. Let δS be the boundary of $\mathcal{S}$ and $\delta\mathcal{D}$ be the boundary traversed anti-clockwise. If $\mathbb{F}$ is a continuously differentiable vector field in space, then

$$\int_{\delta \mathcal{S}} \mathbb{F}(r) \cdot \delta r = \int \int_{\mathcal{S}} (\nabla \times \mathbb{F}) \cdot \mathbf{n}\, \delta S$$

Recall that $\nabla \times \mathbb{F}$ is the (scalar) curl of $\mathbb{F}$.

In words, the path integral (work) along the boundary of $\mathcal{S}$ is equal the the surface integral of the curl of $\mathbb{F}$ over $\mathcal{S}$.

6.6.3 Divergence Theorem

Theorem 6.86 *(Divergence theorem)*

Let $\mathbb{F} = (P, Q, R)$ be a differentiable vector field defined on a solid region $\mathcal{W} \subseteq \mathbb{R}^3$. Let $\mathbf{n}$ be the normal vector pointing outward on $\delta\mathcal{W}$ a closed surface. Then divergence theorem states that

$$\int \int_{\delta \mathcal{W}} \mathbb{F} \cdot \mathbf{n}\, \delta S = \int \int \int_{\mathcal{W}} \nabla \cdot \mathbb{F}\, \delta x \delta y \delta z$$

where $\nabla \cdot \mathbb{F}$ is the flux of $\mathbb{F}$. The divergence theorem states that the flux of a vector field $\mathbb{F}$ across the boundary of a solid is equal to the triple integral of the divergence over the solid.

6.6.3.1 Conservation law

We show an application of the divergence theorem. Suppose you have a region of space and there's matter moving inside. For our sake let's say heat is moving inside the region of space.

We define $\rho(x, y, z, t)$ be the density at time t at the point (x, y, z) and let the velocity at time t be $v(x, y, z, t)$.

Suppose we have conservation of mass, and thus none of the heat is created or destroyed.

What is the rate at which the amount of heat in some volume $\mathcal{W}$ is changing?

At time t the amount of stuff in $\mathcal{W}$ is

$$\int\int\int_{\mathcal{W}} \rho(x,y,z,t)\,\delta\mathcal{W} \Rightarrow \frac{\delta}{\delta t}\left(\int\int\int_{\mathcal{W}} \rho(x,y,z,t)\,\delta\mathcal{W}\right) = \int\int\int_{\mathcal{W}} \frac{\delta\rho}{\delta t}\,\delta\mathcal{W}$$

Now we look at the boundary of $\mathcal{W}$, let's call it S. The rate at which heat flows through S is

$$\int\int_{S} \rho v\cdot\mathbf{n}\,\delta S \Rightarrow \int\int\int_{\mathcal{W}} \frac{\delta\rho}{\delta t}\,\delta\mathcal{W} = -\int\int_{S} \rho v\cdot\mathbf{n}\,\delta S \tag{6.19}$$

which we claim these two expressions are equal. Just believe for now. The reason why we have a negative sign is because the heat flows out of the region, and the RHS is a surface integral. The heat in the region can only escape via the boundary (i.e. it has to flow across $\partial\mathcal{W}$.

Now we apply Divergence Theorem to obtain

$$\begin{aligned}
\int\int_{S} \rho v\cdot\mathbf{n}\,\delta S &= \int\int\int_{\mathcal{W}} \nabla\cdot(\rho v)\,\delta\mathcal{W}\\
\Rightarrow \int\int\int_{\mathcal{W}} \frac{\delta\rho}{\delta t}\,\delta\mathcal{W} &= -\int\int\int_{\mathcal{W}} \nabla\cdot(\rho v)\,\delta\mathcal{W}\\
\Rightarrow \rho_t &= -\nabla\cdot(\rho v)
\end{aligned}$$

Defining $J=\rho v$, we obtain what is known as the continuity equation $\frac{\delta\rho}{\delta t}+\nabla\cdot(J)=0$.

6.6.4 Jacobian of Vector Field

Very quicklly, we have $\mathbb{F}=(P,Q,R)$. The Jacobian $D\mathbb{F}$ is

$$D\mathbb{F} = \begin{pmatrix} P_x & P_y & P_z\\ Q_x & Q_y & Q_z\\ R_x & R_y & R_z \end{pmatrix}$$

Note that

$$\nabla\cdot\mathbb{F} = tr(\mathbb{F})$$

Definition 6.87 *(Symmetric and anti-symmetric matrix)*

A square matrix can be written as the sum of a symmetric and anti-symmetric matrix.

$$\begin{aligned}
A &= \frac{1}{2}(A+A^T)\ \ \textit{symmetric}\\
A &= \frac{1}{2}(A-A^T)\ \ \textit{anti-symmetric}
\end{aligned}$$

We have

$$\frac{1}{2}(D\mathbb{F}-D\mathbb{F}^T) = \frac{1}{2}\begin{pmatrix} 0 & P_y-Q_x & P_z-R_x\\ Q_x-P_y & 0 & Q_z-R_y\\ R_x-P_z & R_y-Q_z & 0 \end{pmatrix} \tag{6.20}$$

See if you can find the elements of $\nabla\times\mathbb{F}$ in that matrix.

If $\mathbb{F}=\nabla\cdot\mathbb{F}$ then $D\mathbb{F}$ is symmetric and $\frac{1}{2}(D\mathbb{F}-D\mathbb{F}^T)=0$.

MATH2921 Vector Calc. & ODEs

7. Expressing Second Order ODEs as First Order Linear Systems

First Order Linear Systems

7.7 First Order Linear Systems

7.7.1 Exponential Matrix

Before we discuss how to represent linear ODEs as first order linear systems, there are some important theoretical concepts we should cover.

Definition 7.88 *(Matrix Exponential)*

Define a matrix $\mathbf{A} \in \mathbb{K}^{n\times n}$. Define e^A as

$$\exp(\mathbf{A}) = \sum_{k=0}^{\infty} \frac{1}{k!}\mathbf{A}^k \tag{7.21}$$

$$= \frac{1}{0!}\mathbf{I} + \frac{1}{1!}\mathbf{A} + \frac{1}{2!}\mathbf{A}^2 + \frac{1}{3!}\mathbf{A}^3 + \dots \tag{7.22}$$

provided that (7.21) converges.

This naturally leads to the question of "what matrix would converge in this power series?

Clearly the diagonal matrix will converge.

Lemma 7.89 *If $\mathbf{D} \in \mathbb{K}^{n\times n}$, then*

$$e^{\mathbf{D}} = \begin{pmatrix} e^{\lambda_1} & 0 & \dots & 0 \\ 0 & e^{\lambda_2} & \dots & 0 \\ \vdots & \ddots & \vdots & \\ 0 & \dots & e^{\lambda_n} & \end{pmatrix}$$

Lemma 7.90 *If $\mathbf{A} \in \mathbb{K}^{n\times n}$ is diagonalisable, then*

$$\mathbf{A}^n = \mathbf{P}\mathbf{D}^n\mathbf{P}^{-1}$$

Note: This lemma holds $\forall n \in \mathbb{Q}$. By convention, we take positive eignvalues.

(Sketch of proof)

Use induction.

Corollary 7.91 *If $\mathbf{A} \in \mathbb{K}^{n\times n}$ is diagonalisable, then*

$$e^{\mathbf{A}} = \mathbf{P}e^{\mathbf{D}}\mathbf{P}^{-1} \tag{7.23}$$

We illustrate the idea with an important matrix exponential. Define $\mathbf{J} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Then using power series we are able to separate the powers into sin and cos. Hence we have

$$\begin{aligned} e^{\mathbf{J}} &= \cos(1)\mathbf{I} + \sin(1)\mathbf{J} \\ &= \begin{pmatrix} \cos(1) & \sin(1) \\ -\sin(1) & \cos(1) \end{pmatrix} \end{aligned}$$

You can also obtain this result by diagonalising $\mathbf{J}$ and calculating the eigenvalues and eigenvectors.

Define $\mathbf{J}x = \begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix}$. Using the same methods we obtain

$$e^{\mathbf{J}x} = \begin{pmatrix} \cos(x) & \sin(x) \\ -\sin(x) & \cos(x) \end{pmatrix}$$

Lemma 7.92 *We have* $\mathbf{Ax} = \mathbf{PDP}^{-1}$*. Where* $\mathbf{A}$ *is diagonalisable. Then*

$$\frac{\delta}{\delta \mathbf{x}} e^{\mathbf{Ax}} = \mathbf{A} e^{\mathbf{Ax}}$$

(Sketch of proof)

Clearly

$$\frac{\delta}{\delta \mathbf{x}} e^{\mathbf{Dx}} = \mathbf{D} e^{\mathbf{Dx}}$$

Now we have

$$\begin{aligned} \frac{\delta}{\delta \mathbf{x}} e^{\mathbf{Ax}} &= \mathbf{PD} e^{\mathbf{Dx}} \mathbf{P}^{-1} \\ &= \mathbf{PDP}^{-1}\mathbf{P} e^{\mathbf{Dx}} \mathbf{P}^{-1} \\ &= \mathbf{PD} e^{\mathbf{Dx}} \mathbf{P}^{-1} \\ &= \mathbf{A} e^{\mathbf{Ax}} \end{aligned}$$

Lemma 7.93 *Suppose* $]mathbf{A}$ *and* $\mathbf{B}$ *are two* $\mathbb{K}^{n\times n}$ *matrices. If* $\mathbf{AB} = \mathbf{BA}$*. then we have*

$$e^{\mathbf{A}+\mathbf{B}} = e^{\mathbf{A}} e^{\boldsymbol{B}} = e^{\mathbf{B}} e^{\boldsymbol{A}}$$

Lemma 7.94 *If it exists, the matrix* $e^{\mathbf{A}x}$ *is invertible, with inverse* $e^{-\mathbf{A}x}$ $\forall x \in \mathbb{R}$*.*

Lemma 7.95 *If* $\mathbf{A} \in \mathbb{K}^{2\times 2}$ *with eigenvalue* λ*. Then*

$$(\mathbf{A} - \lambda \mathbf{I})^2 = \mathbf{0} \tag{7.24}$$

(Sketch of proof)

By direct calculation.

7.7.2 Representing ODEs as a First Order System

Definition 7.96 *(First Order Linear System)*

$$\mathbf{y}' = \mathbf{A}\mathbf{y}$$

Let $\mathbf{v}$ *be an eigenvector of* $\mathbf{A}$ *and* λ *be the corresponding eigenvalue. Then*

$$\mathbf{y}(x) = e^{\lambda x}\mathbf{v}$$

is a solution to the linear system.

(Sketch of proof)

We have

$$\begin{aligned}\mathbf{y}'(x) &= \frac{\delta}{\delta x}\left(e^{\lambda x}\mathbf{v}\right)\\ &= \lambda\left(e^{\lambda x}\mathbf{v}\right)\\ &= \lambda\mathbf{y}(x)\end{aligned}$$

On the RHS we have

$$\begin{aligned}\mathbf{A}\mathbf{y} &= \mathbf{A}\left(e^{\lambda x}\mathbf{v}\right)\\ &= \left(e^{\lambda x}\mathbf{A}\mathbf{v}\right)\\ &= \lambda\left(e^{\lambda x}\mathbf{v}\right)\\ &= \lambda\mathbf{y}(x)\\ &= \mathbf{y}'\end{aligned}$$

Lemma 7.97 *If* $\mathbf{A} \in \mathbb{R}^{2\times 2}$ *with distinct eigenvalues, then we have a pair of linearly independent solutions*

$$\begin{aligned}\mathbf{y}_1(x) &= e^{\lambda_1 x}\mathbf{v}_1\\ \mathbf{y}_2(x) &= e^{\lambda_2 x}\mathbf{v}_2\end{aligned}$$

and this will span the space of solutions to the system. Then their linear combination is the general solution to the system. That is,

$$\begin{aligned}\mathbf{y}(x) &= \mathbf{c}_1\mathbf{y}_1(x) + \mathbf{c}_2(x)\mathbf{y}_2(x)\\ &= \mathbf{c}_1 e^{\lambda_1 x}\mathbf{v}_1 + \mathbf{c}_2 e^{\lambda_1 x}\mathbf{v}_2\\ &= \begin{pmatrix} | & | \\ \mathbf{v}_1 & \mathbf{v}_2 \\ | & | \end{pmatrix}\begin{pmatrix} e^{\lambda_1 x} & 0 \\ 0 & e^{\lambda_2 x}\end{pmatrix}\begin{pmatrix} c_1 \\ c_2\end{pmatrix}\end{aligned}$$

The first matrix is the $\mathbf{P}$ *from* (7.23) *and the second is the exponential of* $\mathbf{D}$ *from* (7.23).

Proposition 7.98 *To solve for ODEs that have initial conditions,* $\mathbf{y}(0)$ *we have*

$$\begin{aligned}\mathbf{y}(0) &= \begin{pmatrix} | & | \\ \mathbf{v}_1 & \mathbf{v}_2 \\ | & | \end{pmatrix}\begin{pmatrix} e^{\lambda_1 0} & 0 \\ 0 & e^{\lambda_2 0}\end{pmatrix}\begin{pmatrix} c_1 \\ c_2\end{pmatrix}\\ &= \mathbf{P}\begin{pmatrix} c_1 \\ c_2\end{pmatrix}\end{aligned}$$

Then the initial conditions will be

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \mathbf{P}^{-1} \begin{pmatrix} y_1(0) \\ y_2(0) \end{pmatrix}$$

Corollary 7.99 *We initial conditions, we can rewrite the solution to system as*

$$\begin{aligned}
\mathbf{y}(x) &= \begin{pmatrix} | & | \\ \mathbf{v}_1 & \mathbf{v}_2 \\ | & | \end{pmatrix} \begin{pmatrix} e^{\lambda_1 \cdot 0} & 0 \\ 0 & e^{\lambda_2 \cdot 0} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \\
&= \begin{pmatrix} | & | \\ \mathbf{v}_1 & \mathbf{v}_2 \\ | & | \end{pmatrix} \begin{pmatrix} e^{\lambda_1 0} & 0 \\ 0 & e^{\lambda_2 0} \end{pmatrix} \begin{pmatrix} | & | \\ \mathbf{v}_1 & \mathbf{v}_2 \\ | & | \end{pmatrix}^{-1} \\
&= \mathbf{P} e^{\mathbf{D}x} \mathbf{P}^{-1} \mathbf{y}(0) \\
&= e^{\mathbf{A}x} \mathbf{y}(0)
\end{aligned}$$

Note: We never mentioned anything about requiring 2 distinct eigenvalues. All we need is $\mathbf{A}$ *to be diagonalisable!*

Theorem 7.100 *The unique solution to the initial value problem*

$$\mathbf{y}' = \mathbf{A}\mathbf{y}$$

with $\mathbf{y}(0) = \mathbf{y}_0$*, and* $\mathbf{A}$ *diagonalisable is*

$$\mathbf{y}(x) = e^{\mathbf{A}x} \mathbf{y}_0$$

Proposition 7.101 *If* $\mathbf{A} \in \mathbb{K}^{2\times 2}$ *is not diagonalisable, we can write*

$$\begin{aligned}
\mathbf{A} &= \lambda \mathbf{I} + (\mathbf{A} - \lambda \mathbf{I}) \\
&= \lambda \mathbf{I} + \mathbf{N}
\end{aligned}$$

Taking the power series of the exponential of $\mathbf{A}$ *and noting* $\mathbf{N}^2 = 0$*, from (7.24), we have*

$$\begin{aligned}
e^{\mathbf{A}x} &= e^{\lambda \mathbf{I} x} e^{\mathbf{N}x} \\
&= e^{\lambda x} (\mathbf{I} + \mathbf{N}x)
\end{aligned}$$

MATH2921 Vector Calc. & ODEs

8. Solving Second Order ODEs

Solving Second Order ODEs

8.8 Solving Second Order ODEs

8.8.1 Recap: solving first order ODEs

Definition 8.102 *(First order linear ODEs)*

Any first order linear ODE has the form

$$\dot{y} + p(x)y = f(x)$$

If $f(x) = 0$, we have an homogeneous first order ODE and we can solve this by separating the functions. That is,

$$\begin{aligned} \frac{\dot{y}}{y} &= -p(x) \\ \Rightarrow y &= \exp\left[\int^x -p(t)\ \delta t + C\right] \\ &= A\exp\left[\int^x -p(t)\ \delta t\right] \end{aligned}$$

Otherwise we multiply by the **integrating factor**

$$\mu(t) = \exp\left[\int^x p(t)\ \delta t\right]$$

which allows us to simply the LHS using the product rule

$$\frac{\delta}{\delta x}(\mu(x)y(x)) = \mu(x)p(x)$$

and integrating both sides we get

$$y(x) = [\mu(x)]^{-1}\int^x \mu(t)p(t)\ \delta t + C$$

where C is a constant.

For **exact differential equations**, it must have the form $M(x,y) + N(x,y)\frac{\delta y}{\delta x} = 0$. Then a solution exists if M and N are twice differentiable and continuous for both itself and its derivatives. That is, we test if

$M_y = N_x$, a similar idea to the cross-derivatives test we saw earlier. If a solution exists, we can write

$$\begin{aligned} M(x,y) + N(x,y)\frac{\delta y}{\delta x} &= \Phi_x + \Phi_y \frac{\delta y}{\delta x} \\ \Rightarrow \frac{\delta}{\delta x}(\Phi(x, y(x))) &= 0 \\ \Rightarrow \Phi(x, y(x)) &= C \end{aligned}$$

where C is a constant.

Finally we look at Bernoulli equations, which are of the form

$$\begin{aligned} \dot{y} + p(x)y &= q(x)y^n \\ \Rightarrow y^{-n} + p(x)y^{1-n} &= q(x) \end{aligned}$$

Make a substitution $v = y^{1-n}$ yields

$$\frac{1}{n-1}\dot{v} + p(x)v = q(x)$$

which is a first order differential equation we can solve.

Okay that should be enough revision for now. Now let's move onto second order linear ODEs.

8.8.2 Key Concepts for Second Order ODEs

Definition 8.103 *(Second order linear ODEs)*

The complete form of second order linear ODEs is

$$y'' + p(x)y' + q(x)y = F(x)$$

where $F(x)$ is called the forcing function. We require that $p(x)$ and $q(x)$ be continuous on the interval (a, b) and that $x_0 \in (a, b)$, where x_0 is the initial condition.

If we are given initial conditions

$$\begin{aligned} y(x_0) &= k_1 \\ y'(x_0) &= k_1 \end{aligned}$$

then the initial value problem (IVP) has a unique solution on the interval (a, b).

We discuss the **principle of superposition**.

Definition 8.104 *(Principle of superposition)*

If $y_1(x)$ and $y_2(x)$ are a fundamental set of solutions to a second order linear ODE on an interval (a, b) then every solution to the system can be written as a linear combination of these solutions. This scales to the general, n^{th}-order case as well. For second order we have

$$y(x) = c_1 y_1(x) + c_2 y_2(x)$$

How do we tell if a set of solutions is linearly independent? We use the **Wronskian determinant**.

Definition 8.105 *(Wronskian determinant)*

For a set of solutions $y_1(x)$ and $y_2(x)$ to a second order ODE, the Wronskian determinant is defined as

$$W(x) = \det \begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix}$$

which is similarly applicable to n^{th}-order ODEs.

Theorem 8.106 *For the set of solutions to be independent, $W(x)$ must $\neq 0$. If you write the solutions as a linear combination, we would have*

$$\begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Then if the solutions are linearly independent we require c_1 and c_2 not to be zero (trivial solution) whilst y_1 and y_2 to be zero (by the definition of linear independence).

Definition 8.107 *(Derivative of the Wronskian)*

We define the derivative of the Wronskian (determinant), which is a matrix of functions, as the sum of the determinants where the derivative is taken once in each row. Then when we work everything out we get

$$W'(x) = -p_{n-1} W(x)$$

where p_{n-1} is the coefficient of the $y^{n-1}(x)$ from the equation. This notation indicates how we apply this to higher dimension.

Definition 8.108 *(Abel's Identity)*

Let y_1 and y_2 be solutions to a second order ODE. Abel's identity can be derived from the Wronskian

$$W(x) = W(x_0) \exp \left[\int_{x_0}^{x} p(t)\ \delta t \right]$$

where $W'(x) = -p(x)W(x)$ solves the first order ODE.

Definition 8.109 *(Principal fundamental solution matrix)*

The ***principal fundamental solution matrix*** *is defined as*

$$\Phi(x) = \begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix}$$

with $\Phi(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, and is a matrix that solves the first order system

$$\Phi'(x) = \mathbf{A}(x)\Phi(x) \tag{8.25}$$

Without the initial condition, the matrix is called the ***fundamental solution matrix****.*

Proposition 8.110 *Using the fact that*

$$\Phi(x)\cdot\left[\Phi^{-1}(x)\right]=\mathbf{I}$$

we have

$$\begin{aligned}\frac{\delta}{\delta x}\mathbf{I}&=\frac{\delta}{\delta x}\Phi(x)\cdot\left[\Phi^{-1}(x)\right]\\&=\Phi'(x)\cdot\left[\Phi^{-1}(x)\right]+\Phi(x)\cdot\left[\Phi^{-1}(x)\right]'\\\Rightarrow\Phi^{-1}(x)\Phi'(x)\Phi^{-1}(x)&=\Phi^{-1}(x)\mathbf{A}(x)\Phi(x)\Phi^{-1}(x)\\\Rightarrow-\Phi^{-1}(x)\Phi(x)\left[\Phi^{-1}(x)\right]'&=\Phi^{-1}(x)\mathbf{A}(x)\\\Rightarrow\left[\Phi^{-1}(x)\right]'&=-\Phi^{-1}(x)\mathbf{A}(x)\end{aligned}\tag{8.26}$$

when you apply (8.25) *in addition to the result here.*

Remark 8.111 *We will need* (8.26) *for the method of variation of parameters.*

8.8.3 Homogeneous Constant Coefficient Second Order ODEs

Here we deal with equations of the form

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

a homogeneous second order ODE with constant coefficients.

So we have a few cases.

Real roots

The solution is then

$$y(x)=c_1e^{\lambda_1 x}+c_2e^{\lambda_2 x}$$

Repeated roots

$$y(x)=c_1e^{\lambda x}+c_2xe^{\lambda x}$$

Complex roots

We obtain

$$y(x)=c_1e^{(\lambda+i\mu)x}+c_2e^{(\lambda-i\mu)x}$$

and with a bit of re-arranging and using Euler's identity $e^{i\theta}=\cos(\theta)+i\sin(\theta)$ we obtain

$$y(x)=c_1e^{\lambda x}\cos(\mu t)+c_2e^{\lambda x}\sin(\mu t)$$

8.8.4 Inhomogeneous Second Order ODEs

Let's begin by discussing the 'constant coefficient' version.

$$y''+ay'+by=F(x)$$

Here the fundamental set of solutions consist of the homogeneous solution $y_h(x)$ and the particular solution $y_p(x)$. Combining the two together gives the general solution. In order to solve for $y_p(x)$ we obtain need to first obtain the homogeneous solution.

There are a few ways of solving these problems. The homogeneous solution part has been discussed already, and thus let's consider how to find the particular solution.

Method 1: Take a good guess - anstaz.

If $F(x)$ has the form

1. ce^{nx} - take a guess Ae^{nx}
2. $a_0 + a_1 x + ...a_n x^n$ - guess $A + Bx + ... + Cx^n$
3. $c_1 \sin(x) + c_2 \cos(x)$ - guess $A \sin(x) + B \cos(x)$

Remember that you should always guess the appropriate coefficient for the power of exp, up to the highest power for the polynomial and both sin and cos.

The principle of superposition means that if $F(x)$ is a combination then your guess should also be the appropriate combination. This method of solving the equations means you'll need to substitute and solve for the unknown coefficients, so sometimes this is called the method of unknown coefficients. Note: ansatz just sounds smarter =D.

Remark 8.112 *If your homogeneous solution contains a solution, then guessing the same form for the particular solution will not lead to a linearly independent solution. So in this case you would want to guess $y_p(x) = x^s e^{nx}$ where s is the smallest power for x such that you don't have clashing linearly dependent solution.*

Method 2: Reduction of Order

Sometimes the form of the forcing function doesn't allow for a good guess or it is not one of the forms mentioned above. In this case we require another method. **This method is very useful if you are given one of the solutions to the problem.**

Suppose you are given a solution $y_1(x)$. Then you suppose that the second solution is of the form $y_2(x) = \mu(x)y_1(x)$. Substituting the appropriate values into the original problem eventually leads to

$$\begin{aligned} 0 &= \mu'' y_1 + \mu'(2y_1' + py_1) \\ &= \mu'' + \mu' p(x) \end{aligned}$$

which looks like a first order equation. Then you can solve this by multiplying through by an integrating factor

$$\mu(x) = \int^x \frac{\exp\left[-\int^t p(s)\ \delta s\right]}{y_1^2(t)}\ \delta t$$

You can verify linear independence quickly for the Wronskian determinant.

You can also apply the Wronskian determinant to obtain your reduction of order. So suppose you are given that $y_1(x) = e^x$. Then

$$W(x) = \det \begin{pmatrix} e^x & y_2 \\ e^x & y_2' \end{pmatrix}$$

and expanding this out you get

$$(y_2' - y_2)e^x = W(x_0) \exp\left[\int^x p(t)\ \delta t\right]$$

as you would expect from the definition of Abel's identity.

Remark 8.113 *This method often leads directly to the general solution of the problem.*

Remark 8.114 *We can use **reduction of order** to solve for non-constant coefficient problems as well!*

Method 3: Method of Variation of Parameters

Here we attempt to find a solution to the matrix problem using a first order linear system. So we have

$$\mathbf{y}'(x) = \mathbf{A}(x)\mathbf{y}(x) + \mathbf{F}(x) \tag{8.27}$$

Multiply this by the (P)FSM to obtain

$$\begin{aligned}\frac{\delta}{\delta x}\left[\Phi^{-1}(x)\mathbf{y}(x)\right] &= \Phi^{-1}(x)\mathbf{F}(x) \\ \Rightarrow \mathbf{y}(x) &= \Phi(x)\int^x \Phi^{-1}(t)\mathbf{F}(t)\ \delta t + \Phi(x)\mathbf{c}\end{aligned}$$

where $\mathbf{x}$ is a vector constant.

For this method, we will require the actual (P)FSM $\Phi(x)$. One way to obtain it would be to solve the homogeneous equation and use it as $Phi(x)$.

You can always directly try multiplying by $\exp\left[-\mathbf{A}\cdot\mathbf{x}\right]$ and seeing whether it works or not. However, the exponential matrix may not be so easy to compute.

MATH2921 Vector Calc. & ODEs

9. Series Solutions to ODEs

Series Solutions to ODEs

9.9 Series Solutions to ODEs

9.9.1 Introduction

Suppose we have some ODE with an IVP of the form

$$\begin{aligned} y'(x) &= ay(x) \\ y(0) &= 1 \end{aligned}$$

and we wanted to derive a solution from 'first principles'. We could integrate the RHS from 0 to x.

$$\begin{aligned} y(x) - y(0) &= \int_0^x ay(t)\ \delta t \\ \Rightarrow y(x) &= \int_0^x ay(t)\ \delta t + 1 \end{aligned}$$

Define a series of equations

$$y_n = \int_0^x ay_n(t)\ \delta t + 1$$

with the initial value $y_0 = 1$. Then

$$\lim_{n\to\infty} y_n = \sum_{k=1}^{n} \frac{(ax)^k}{k!}$$

which is the power series!

Proposition 9.115 *For some ODE with IVP*

$$\begin{aligned} y'(x) &= f(x, y) \\ y(x_0) &= y_0 \end{aligned}$$

we have

$$y_n = y_0 + \int_{x_0}^x f_n(t, y_{n-1}(t))\ \delta t + 1$$

converging as a power series to a unique fixed point in the limit, say y_∞, if f is continuous and first differentiable in x and y.

Remark 9.116 *Since all ODEs are in essence 1st order linear systems, the above implies (and it is infact true) that there exists a solution to all ODEs as long as f is continuous and first differentiable!*

Note: We prove this in MATH3961 (Metric Spaces).

Definition 9.117 *(analytic function)*

A function f is real analytic on a set $D \in \mathbb{R}$ if $\forall x_0 \in D$ we can represent $f(x)$ as

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$$

where $a_n \in \mathbb{R}$ and the series is convergent to $f(x)$ for $x \in B(x, r)$, where $B(x, r)$ is an open ball with radius $r > 0$.

Definition 9.118 *(smooth function)*

A function is smooth if it is first-differentiable i.e. $f'(x)$ exists for $f(x)$ $\forall x \in D$ where D is the domain and continuous. These two properties combined is called ***continuously differentiable****.*

Note: I wouldn't worry about above formal definitions too much - analytic (and smooth) functions are discussed in more detailed in MATH2923 (Analysis).

9.9.2 Power Series Solution to N-th Order ODEs

Definition 9.119 *(Linear inhomogeneous n-th order ODE)*

It is fairly clear what the concept entails.

If $p_j(x)$ are smooth, analytic functions $\forall j$ then we have the n-th order ODE

$$\begin{aligned}\mathcal{L}y &= p_0(x)y^{(n)} + p_1(x)y^{(n-1)} + \cdots + p_0(x)y \\ &= F(x)\end{aligned}$$

where none of the $p_n(x)$ functions have a common factor (so you can't divide through basically).

Definition 9.120 *(Ordinary and singular point)*

x_0 from above is called an ordinary point if

$$p_0(x_0) \neq 0$$

Otherwise x_0 is called a singular point.

We can solve problems using this approach but it is very tedious. Let's look at a quick example.

$$(1 + x^2)y'' + xy' + y = 0$$

Let $x_0 = 0$. We have

$$y = \sum_{k=0}^{\infty} a_k x^k$$
$$y' = \sum_{k=1}^{\infty} k \cdot a_k x^{k-1}$$
$$y'' = \sum_{k=2}^{\infty} (k(k-1))_k x^{k-2}$$

You can expand this out and group the coefficients based on the powers of the polynomials.

For this question, since the RHS is equal to zero, the coefficients must equate to zero for each power of x since the polynomial is a basis.

Sometimes you can find a closed form of the power series and evaluate it. However, most of the time you probably would get stuck with a recurrence formula.

For this question, we have

$$y(x) = \sum_{n=0}^{\infty} x^n \left[(n+2)(n+1)a_{n+2} + (n(n-1) + n + 1)a_n\right]$$

And the recurrence formula.

Suppose you add two initial conditions $y(0) = 1$ and $y'(0) = 0$. Then we have

$$\begin{aligned} y'(0) &= \sum_{n=1}^{\infty} k \cdot a_k x^{k-1} \\ &= 0 \\ \Rightarrow 1 \cdot a_1 (0)^0 &= 0 \end{aligned}$$

Hence $a_1 = 0$. Since we have an even recurrence, it must be that all $a_{2k+1} = 0 \in \mathbb{N}$. In this case our solution becomes

$$y(x) = \sum_{n=0}^{\infty} x^{2n} \left[(2n+2)(2n+1)a_{2n+2} + (2n(2n-1) + 2n + 1)a_{2n}\right]$$

Without initial conditions, it might be difficult to pull out a single solution to the recurrence relation.

MATH2921 Vector Calc. & ODEs

10. Boundary Value Problems

Boundary Value Problems

10.10 Boundary Value Problems

10.10.1 The 5 Boundary Value Problems

Definition 10.121 *(Boundary Value Problems)*

A boundary value problem is an ODE (or a PDE) that requires solutions that satisfy conditions on the boundary of some region (an interval for 1-dimension).

Note: You can also view the boundary value problem as a linear map operator on functions.

The 5 boundary value problems we tackled in this course revolve around a single linear ODE:

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

This ODE can have the following 5 boundary conditions:

1. **Dirichlet:** $y(0) = 0,\ y(L) = 0$
2. **Neumann:** $y'(0) = 0,\ y'(L) = 0$
3. **Mixed Type One:** $y(0) = 0,\ y'(L) = 0$
4. **Mixed Type Two:** $y'(0) = 0,\ y(L) = 0$
5. **Periodic:** $y(-L) = y(L),\ y'(-L) = y'(L)$

where $L \neq 0$.

10.10.2 Solving the Dirichlet Example

Let's look at the first example and see how we can solve. Immediately we have the characteristic equation

$$r^2 + \lambda = 0$$

Case 1: If $\lambda < 0$ then let $w^2 = \lambda$ and we have

$$y(x) = c_1 e^{w_1 x} + c_2 e^{w_2 x}$$

Applying the boundary conditions we obtain

$$\begin{pmatrix} 1 & 1 \\ e^{wL} & e^{-wL} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Since det $= -2\sinh(wL)$, the matrix isn't able since by assumption $\lambda < 0$ and $w > 0$ and by definition $L \neq 0$.

Hence this is not invertible and the only solution is $y = 0$.

Case 2: If $\lambda = 0 \Rightarrow y = c_1 + c_2 x$ and by the boundary conditions we have $c_1 = c_2 = 0 \Rightarrow y = 0$ again.

Case 3: Now we reach the most interesting case, $\lambda > 0$.

$$r^2 + w^2 = 0$$

Implies that we have

$$y(x) = c_1 \sin(wx) + c_2 \cos(wx)$$

Applying the boundary conditions gives us

$$\begin{aligned} y(L) &= 0 \\ &= c_1 \sin(wL) \end{aligned}$$

Suppose $c_1 \neq 0$ (if $c_1 = 0$ then we have a zero solution).

Given our knowledge of the sin curve, we have

$$\begin{aligned} c_1 \sin(wL) = 0 &\iff w = \frac{k\pi}{L} \\ &\Rightarrow \lambda_k = \left(\frac{k\pi}{L}\right)^2 \end{aligned}$$

$\forall k \in \mathbb{N} \setminus \{0\}$. This ∞-sequence of eigenvalues has a corresponding ∞-sequence of eigenfunctions

$$\phi_k(x) = \sin\left(\frac{k\pi}{L} x\right)$$

defined up to a constant multiple.

10.10.3 Solving the Neumann Example

It turns out we can eliminate all non-negative λ values for the 5 problems by showing the following:

$$\begin{aligned} y''(x) + \lambda y(x) &= 0 \\ \Rightarrow y(x)y''(x) + \lambda y^2(x) &= 0 \end{aligned}$$

where we can take the integral and integrate by parts to get

$$\begin{aligned} 0 &= y'(x)y(x)\,\Big|_0^L - \int_0^L (y'(x))^2 \delta x + \lambda \int_0^L y^2(x)\,\delta x \\ &= \lambda \int_0^L y^2(x)\,\delta x - \int_0^L (y'(x))^2 \delta x \\ \Rightarrow \lambda \int_0^L y^2(x)\,\delta x &= \int_0^L (y'(x))^2 \delta x \end{aligned}$$

Here we have $(y'(x))^2 \geq 0$ and $y(x) \neq 0 \Rightarrow \lambda \geq 0$.

If $\lambda = 0$ we have $y''(x) = 0$ which means that $y'(x) = 0$ from the above. This gives $y(x) = c$ a constant.

So we deal with case 3 directly.

Given $\lambda > 0$ and let $\lambda = w^2$ we get complex roots and thus $y(x) = c_1 \sin(wx) + c_2 \cos(wx)$. Applying the initial conditions yields

$$\begin{aligned} 0 &= y'(L) \\ &= -wc_2 \sin(wL) \end{aligned}$$

As we saw in the first problem, this yields $\sqrt{\lambda_k} = w = \frac{k\pi}{L}$ again giving an ∞-sequence of eigenvalues and eigenfunctions.

$$\phi_k(x) = \cos\left(\frac{k\pi}{L}x\right)$$

with $\lambda_k = \left(\frac{k\pi}{L}\right)^2$

10.10.4 Solving the Mix Type One

We have boundary conditions $y'(L) = 0,\ y(0) = 0$. We know that $\lambda \geq 0$ with $f(x) = c$ a constant if $\lambda = 0$.

The general solution is $y(x) = c_1 \sin(wx) + c_2 \cos(wx)$ with $\lambda = w^2$. With the boundary conditions, we have

$$\begin{aligned} y(0) &= 0 \\ &= c_2 \end{aligned}$$

and

$$\begin{aligned} y'(L) &= 0 \\ &= wc_2 \cos(wx) \end{aligned}$$

Hence have $w = \frac{(2k-1)\pi}{2L}$ if $c_2 \neq 0$. This gives us

$$\lambda_k = \left(\frac{(2k-1)\pi}{2L}\right)^2$$

with the associated eigenfunction

$$\phi_k = \sin\left(\frac{(2k-1)\pi}{2L}x\right)$$

10.10.5 Solving the Mix Type Two

We have boundary conditions $y(L) = 0,\ y'(0) = 0$. We know that $\lambda \geq 0$ with $f(x) = c$ a constant if $\lambda = 0$.

The general solution is $y(x) = c_1 \sin(wx) + c_2 \cos(wx)$ with $\lambda = w^2$. With the boundary conditions, we have

$$\begin{aligned} y'(0) &= 0 \\ &= wc_1 \end{aligned}$$

and

$$y(L) = 0$$
$$= c_2 \cos(wx)$$

Hence have $w = \frac{(2k-1)\pi}{2L}$ if $c_2 \neq 0$. This gives us

$$\lambda_k = \left(\frac{(2k-1)\pi}{2L}\right)^2$$

with the associated eigenfunction

$$\phi_k = \cos\left(\frac{(2k-1)\pi}{2L}x\right)$$

10.10.6 Solving the Periodic Example

Now we have boundary conditions $y(-L) = y(L),\ y'(-L) = y'(L)$. Using the trick in problem 2, we obtain that $\lambda \geq 0$. If $\lambda = 0$, then $y(x) = c$ is a constant as from before.

So we are back at case 3. Let $\lambda = w^2$ we get complex roots and thus $y(x) = c_1 \sin(wx) + c_2 \cos(wx)$. Applying the initial conditions yields

$$\begin{pmatrix} 0 & -\sin(wL) \\ -\sin(wL) & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

So again, we have the boundary value problem has a non-zero solution $\iff$ det $= 0$ That is, det $= w \sin^2(wL) = 0$. $w = 0 \iff \lambda = 0$ so we want $\sin^2(wL) = 0$ which once again gives us $w = \frac{k\pi}{L}$.

However, this time dimension of the null space of this matrix is 2! And this matters. Because now we can construct a pair of linearly independent eigenfunctions for each eigenvalue except $\lambda = 0$.

So this time we have

$$\lambda = 0, \left(\frac{\pi}{L}\right)^2, \left(\frac{2\pi}{L}\right)^2, ..., \left(\frac{k\pi}{L}\right)^2$$

and for each λ we have

$$\phi_{1,k} = \sin\left(\frac{k\pi}{2}\right)$$
$$\phi_{2,k} = \cos\left(\frac{k\pi}{2}\right)$$

With $\phi_{0,k} = 1$ for $\lambda = 0$.

10.10.7 Orthogonal Basis

10.10.7.1 The boundary problem as a linear map

In fact, it is possible to represent the boundary value problem as

$$\partial_{xx} y = -\lambda y$$

Where ∂_{xx} is a linear map from the vector space of smooth functions to itself. Hence we can represent δ_{xx} as an infinite dimensional matrix.

From MATH2922 (Linear Algebra), we know that we should think about constructing a basis. We can take the polynomial basis

$$\{1, x, x^2, x^3, ..., \}$$

but this gives us a matrix

$$\mathcal{D} = \begin{pmatrix} 0 & 0 & 2 & 0 & 0 & \dots & 0 & \\ 0 & 0 & 0 & 6 & 0 & \dots & 0 & \\ 0 & 0 & 0 & 0 & 12 & \dots & 0 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots \end{pmatrix}$$

Where $\partial_{xx}(x) = 0$, $\partial_{xx}(x^3) = 6x$ etc.

We can also write the matrix as

$$d_{ij} = \begin{cases} (i+2)(i+1) & \text{if } j = i+2 \\ 0 & \text{otherwise} \end{cases}$$

This is inconvenient because it is not exactly a diagonal matrix.

It turns out that

$$\left\{ \sin\left(\frac{n\pi}{L}x\right), \cos\left(\frac{n\pi}{L}x\right) \right\}$$

is a span $\forall n \in \mathbb{N}$ and from this assumption we can write

$$f(x) = \sum_{n}^{\infty} a_n \cos\left(\frac{n\pi}{L}x\right) + b_n \sin\left(\frac{n\pi}{L}x\right)$$

where for some fixed $x_0 \in [L, -L]$ we have the coefficients $a_n, b_n \in \mathbb{R}$ such that the series converges pointwise to some limit.

With this basis, ∂_{xx} becomes diagonal and is negative semi-definite (non-positive eigenvalues).

So how far does the set of sin and cos functions span exactly?

Well... it turns out that as long as f is integrable on $[L, -L]$, then any function is good. This is a very weak condition considering we're not even throwing differentiability in here! Nor does it need to be continuous!

10.10.7.2 Orthogonal Basis

Suppose we have $\langle\, x, y \,\rangle$.

Definition 10.122 *(Orthogonal basis) An orthogonal basis of* $\{b_1, b_2, ..., b_n\}$

$$\langle\, b_i, b_j \,\rangle = \begin{cases} 0 & \textit{if } i \neq j \neq 0 \\ \textit{if } i = j & \end{cases}$$

Hence for any vector v in the basis we have

$$v = \sum_{i=1}^{n} v_i b_i$$

where

$$v_i = \frac{\langle\, v, b_i \,\rangle}{\langle\, b_i, b_i \,\rangle}$$

Lemma 10.123 *The inner product of functions can be evaluated as integrals. That is,*

$$\langle\, f, g \,\rangle := \int_{-L}^{L} f \cdot g \; \delta x$$

Proposition 10.124

$$\left\{ \sin\left(\frac{n\pi}{L} x\right), \cos\left(\frac{n\pi}{L} x\right) \right\}$$

is such an orthogonal set.

We have the following identities, for $m \neq n$,

1. $\int_{-L}^{L} \sin\left(\frac{m\pi}{L} x\right) \sin\left(\frac{n\pi}{L} x\right) \; \delta x = 0$
2. $\int_{-L}^{L} \cos\left(\frac{m\pi}{L} x\right) \cos\left(\frac{n\pi}{L} x\right) \; \delta x = 0$
3. $\int_{-L}^{L} \sin^2\left(\frac{n\pi}{L} x\right) \; \delta x = \cos^2\left(\frac{n\pi}{L} x\right) \; \delta x = L$
4. $\int_{-L}^{L} 1 \; \delta x = 2L$ if $n = 0$ for the previous.

10.10.8 Fourier Series

Definition 10.125 *(Fourier series)*

Now, the formal Fourier series of a function $f(x)$ on $[L, -L]$ is

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi}{L} x\right) + b_n \sin\left(\frac{n\pi}{L} x\right)$$

where

$$\begin{aligned} a_0 &= \frac{\langle\, 1, f \,\rangle}{\langle\, 1, 1 \,\rangle} \\ &= \frac{1}{L} \int_{-L}^{L} f(x) \; \delta x \end{aligned}$$

$$\begin{aligned} a_n &= \frac{\langle\, \cos\left(\frac{n\pi}{L} x\right), f \,\rangle}{\langle\, \cos\left(\frac{n\pi}{L} x\right), \cos\left(\frac{n\pi}{L} x\right) \,\rangle} \\ &= \frac{1}{L} \int_{-L}^{L} f(x) \cos\left(\frac{n\pi}{L} x\right) \; \delta x \end{aligned}$$

$$
\begin{aligned}
b_n &= \frac{\langle \sin\left(\frac{n\pi}{L}x\right), f \rangle}{\langle \sin\left(\frac{n\pi}{L}x\right), \sin\left(\frac{n\pi}{L}x\right) \rangle} \\
&= \frac{1}{L}\int_{-L}^{L} f(x)\sin\left(\frac{n\pi}{L}x\right)\,\delta x
\end{aligned}
$$

Remark 10.126 *This isn't exactly math-related but you should memorise this since it is very likely that this will come up in an assessment and you will be using this very often.*

You can compute some examples e.g. x^2, which after a fair amount of calculations yields

$$
x^2 = \frac{\pi^2}{3} + 4\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\cos(nx)
$$

on $[-\pi, \pi]$. You will have to prove that x^2 is an actual solution to the expression and not just a formal (Fourier series) solution.

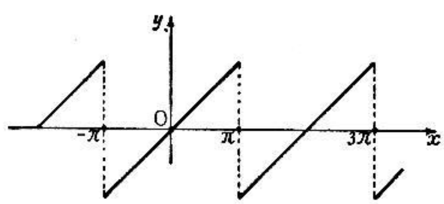
10.10.9 Periodic Extensions

Suppose calculate a Fourier series approximation for $f(x) = x$ on $[-\pi, \pi]$. It turns out that

$$
x \sim 2\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}\sin(nx)
$$

And if you sub in $x = \pi$ you get the RHS $= 0$! So what happened?

Well it turns out that the approximation is good at the values exclusive of the end points.



So at the boundary values the approximation takes the average of the two extremes which is why it gives zero.

The fact that we get an approximation on the endpoint with discontinuities is a very generic property of Fourier series approximations. A function like the above is called piecewise smooth.

Definition 10.127 *(Piecewise smooth)*

A function $f(x)$ is piecewise smooth on $[a, b]$ if

1. *f has at most finitely many jump discontinuities on $[a, b]$.*
2. *f' exists except at the discontinuities and has itself at most finitely many discontinuities*

Sometimes this is denoted as piecewise C^1.

Now we have a concrete example in mind, we can define the convergence of the Fourier series.

Theorem 10.128 *Let f is a piecewise smooth function on $[-L, L]$ and let the Fourier series*

$$F(x) := a_0 + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi}{L}x\right) + b_n \sin\left(\frac{n\pi}{L}x\right)$$

then we have that

$$F(x_0) := \frac{1}{2}\left[\lim_{x \to x_0^-} f(x) + \lim_{x \to x_0^+} f(x)\right]$$

$\forall x_0 \in [-L, L]$.

Remark 10.129 *If we define the discontinuities to be average value, then the Fourier series of f will converge to the periodic extension everywhere. We state this more formally below.*

Theorem 10.130 *If f is piecewise smooth on $[-L, L]$ then the Fourier series converges $\forall x \in [-L, L]$, with*

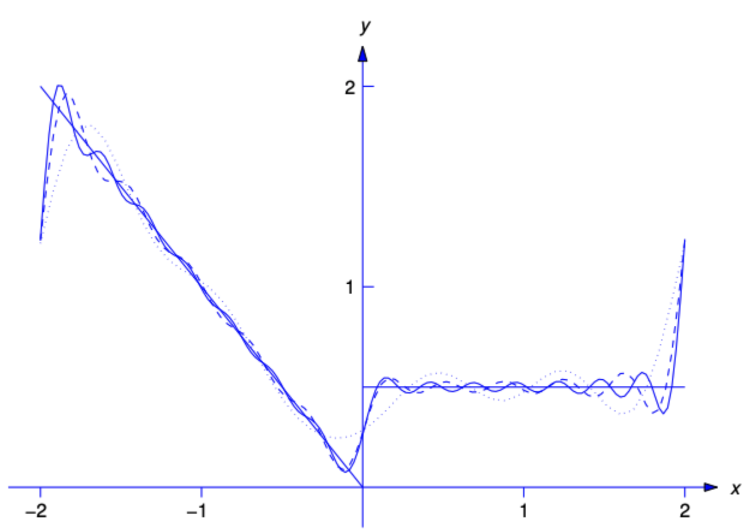
$$F(x) = \begin{cases} f(x) & \text{if } -L < x < L \text{ and } f \text{ is continuous at } x \\ \frac{f(x_-)+f(x_+)}{2} & \text{if } -L < x < L \text{ and } f \text{ is discontinuous at } x \\ \frac{f(L_-)+f(L_+)}{2} & \text{if } -L < x < L \text{ and } f \text{ is discontinuous at } x \end{cases}$$

Note: the value of the Fourier approximation at the endpoints is the average of the 2 endpoint values!

See this example.

$$f(x) = \begin{cases} -x & -2 < x < 0 \\ \frac{1}{2} & 0 < x < 2 \end{cases}$$

The Fourier series approximation is



Notice the endpoints.

Now we can discuss even and odd extensions.

Definition 10.131 *(Even extension)*

Suppose we have a function $f(x)$ defined on $[0, L]$. Then the even periodic extension of the function is

$$f_{even}(x) = \begin{cases} f(-x) & -L \leq x < 0 \\ f(x) & 0 \leq x < L \end{cases}$$

First we notice that

$$\begin{aligned} b_n &= \frac{1}{L} \int_{-L}^{L} f(x) \sin\left(\frac{n\pi}{L} x\right) \, \delta x \\ &= 0 \end{aligned}$$

Since f is even and sin *is odd and the integrand an odd function. Thus we only have* cos *components remaining. Then we get*

$$f_{even}(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi}{L} x\right)$$

with

$$a_n = \frac{2}{L} \int_{0}^{L} f(x) \cos\left(\frac{n\pi}{L} x\right) \, \delta x$$

Definition 10.132 *(Odd extension)*

Suppose we have a function $f(x)$ defined on $[0, L]$. Then the odd periodic extension of the function is

$$f_{odd}(x) = \begin{cases} -f(-x) & -L \leq x < 0 \\ f(x) & 0 \leq x < L \end{cases}$$

Again notice that

$$\begin{aligned} a_n &= \frac{1}{L} \int_{-L}^{L} f(x) \cos\left(\frac{n\pi}{L} x\right) \, \delta x \\ &= 0 \end{aligned}$$

Since f is odd and cos *is even and integrand becomes an odd function. Thus we only have* sin *components remaining. Then we get*

$$f_{odd}(x) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi}{L} x\right)$$

with

$$b_n = \frac{2}{L} \int_{0}^{L} f(x) \sin\left(\frac{n\pi}{L} x\right) \, \delta x$$

Remark 10.133 *If f has period π, then the Fourier series expansion will contain both* sin *and* cos *terms but only for odd values of n*

Remark 10.134 *If f has period 2π and alternatives on each half period, that is*

$$f(x+\pi) = -f(x)$$

then the Fourier series expansion will contain both sin *and* cos *terms but only for even values of n.*

MATH2921 Vector Calc. & ODEs

11. Partial Differential Equations

Partial Differential Equations

11.11 Partial Differential Equations

11.11.1 Heat/Diffusion Equation

11.11.1.1 Introduction to Heat/Diffusion Equation

Now we will apply the ideas of Fourier series and friends to solve partial differential equations.

First let's discuss the heat/diffusion equation. We can derive this from first principles using Fick's 1st Law (mass) or Fourier's Law (temperature).

From the divergence theorem, the law of conservation of mass is

$$u_t + \nabla \cdot \mathbf{J} = 0$$

where u_t is the density of temperature (or mass) $\mathbf{J}$ is the "local flux".

Using the aforementioned laws, we understand that the flux is proportional to the spatial gradient of the concentration of temperature and the heat moves from higher temperatures to lower temperatures. That is,

$$\begin{aligned} \mathbf{J} &= -a^2 \nabla u \\ \Rightarrow u_t &= \nabla(a^2 \nabla u) \\ &= a^2(\Delta u) \\ &= a^2(u_{xx} + u_{yy} + u_{zz}) \end{aligned}$$

and since we are assuming one dimension only, we are done. Therefore we have

$$u_t = a^2 u_{xx}$$

Here we represent time as t.

11.11.1.2 Solving the Heat/Diffusion Equation

For our problem, we assume that there is a tube of finite length, say $[0, L]$ and time is $t \geq 0$. We require boundary conditions and an initial profile.

We specify the following form:

$$u_t = a^2 u_{xx}$$

$$u(0, t) = u(L, t) = 0$$

$$u(x,0) = f(x)$$

In order to solve the equation, we will guess the solution to be separable. That is, we have

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

Substituting this into the equation and rearranging gives

$$\frac{T'(t)}{a^2T(t)} = \frac{X''(t)}{X(t)} = -\lambda$$

where λ is a constant. The reason this works is because the function is separable, and thus both the LHS and RHS are functions independent of the other variable. Hence they must be equal to some constant.

Now we can set up 2 equations, a first order and second order one. The first order one is

$$\begin{aligned} \frac{T'(t)}{a^2T(t)} &\Rightarrow T'(t) = -\lambda a^2 T(t) \\ &\Rightarrow T(t) = Ce^{-\lambda a^2 t} \end{aligned}$$

and we want $C \neq 0$. The second order equation looks to be very familiar to the previous chapter.

$$\begin{aligned} \frac{X''(t)}{X(t)} &= -\lambda \\ \Rightarrow X''(t) + \lambda X(t) &= 0 \end{aligned}$$

Using the boundary conditions at the start, we have

$$X(0)T(t) = X(L)T(t) = 0$$

If $T(t) = 0$ then we would have a zero solution so we assume

$$X(0) = X(L) = 0$$

which is the Dirichlet boundary condition $\Rightarrow \lambda_k = \left(\frac{k\pi}{L}\right)^2$ with $X_k = \sin\left(\frac{k\pi}{L}x\right)$. Finally we put all of this together to obtain

$$u(x,t) = e^{-a^2\lambda_k t} \cdot \sin\left(\frac{k\pi}{L}x\right)$$

$\forall k \in \mathbb{N} \setminus \{0\}$. You can check that this solution satisfies the heat/diffusion equation and the boundary values.

Lemma 11.135 *The PDE is linear w.r.t to $u(x,t)$ so we can take a linear combination of them and it will still be a solution, as you have seen before with ODEs.*

Finally we check that the initial profile condition is satisfied. Given the linearity, we take

$$\sum_{k=1}^{N} b_k u_k(x,0) = \sum_{k=1}^{N} b_k \sin\left(\frac{k\pi}{L}x\right)$$

which is suspiciously similar to a Fourier series. Hence if we add the "t" back in, we have the following theorem

Theorem 11.136 *The formal solution to the heat/diffusion equation with the initial boundary values and profiles*

$$u_t = a^2 u_{xx}$$

$$u(0,t) = u(L,t) = 0$$

$$u(x,0) = f(x)$$

is the fourier series

$$u(x,t) = \sum_{k=1}^{\infty} b_n e^{-a^2 \lambda_k t} \cdot \sin\left(\frac{k\pi}{L} x\right)$$

where $\lambda_k = \left(\frac{k\pi}{L}\right)^2$ *and*

$$b_n = \frac{2}{L} \int_0^L f(x) \sin\left(\frac{k\pi}{L} x\right) \, \delta x$$

$u(x,t)$ *is an actual solution if* $f(x)$ *equals the Fourier Sine series on the interval with continuity except at the end points. The boundary conditions should also be satisfied of course.*

Remark 11.137 *Notice that for the Fourier Sine series we initially looked at, we took a finite series. Indeed, the definition of linear combination technically only holds for finite dimension (elements).*

In this case, the infinite linear combination for the boundary conditions are fine but since the PDE involves differentiation we cannot, in general, argue that the infinite series is still a solution. That is, we cannot assume

$$\frac{\partial}{\partial x} \sum_{k=1}^{\infty} b_k u_k(x,t) = \sum_{k=1}^{\infty} \frac{\partial}{\partial x} b_k u_k(x,t)$$

It turns out that we need to test for uniform and absolute convergence (using Weierstrauss M-test for example). Only then can we assume that the derivative can be moved inside the series. This way, this series solution satisfies the PDE.

Remark 11.138 *(Physical Intuition of Boundary Conditions)*

For the Dirichlet boundary, you can think of it as blocking one end of the tube so that heat can only escape from one end. It is also called a zero-flux boundary condition.

Fourier's Law states that the temperature gradient dictates how the heat will flow through the bar and the condition ensures that the heat will flow out of the right end point.

Theorem 11.139 *The solution to the heat/diffusion equation for "mixed type one" boundary condition, that is* $X'(0) = X'(L) = 0$*, with initial profile* $u(x,0) = f(x)$ *is*

$$u(x,t) = \sum_{k=1}^{\infty} b_n e^{-a^2 \lambda_k t} \cdot \sin\left(\frac{(2k-1)\pi}{2L} x\right)$$

where $\lambda_k = \left(\frac{(2k-1)\pi}{2L}\right)^2$ *with*

$$b_n = \frac{2}{L} \int_0^L f(x) \sin\left(\frac{(2k-1)\pi}{2L} x\right) \, \delta x$$

Theorem 11.140 *The solution to the heat/diffusion equation for Neumann boundary conditions, that is $X'(0) = X'(L) = 0$, with initial profile $u(x, 0) = f(x)$ is*

$$u(x,t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_n e^{-a^2 \lambda_k t} \cdot \cos\left(\frac{k\pi}{2L} x\right)$$

where $\lambda_k = \left(\frac{(2k-1)\pi}{2L}\right)^2$ with

$$a_n = \frac{2}{L} \int_0^L f(x) \cos\left(\frac{(2k-1)\pi}{2L} x\right) \; \delta x$$

Here, we take note that $\lim_{t\to\infty} u(x,t) = a_0$. Physically, if we cap both ends and prevent heat from escaping, then the temperature will tend towards a constant profile which would be the average of the initial condition.

11.11.1.3 Inhomogeneous Heat/Diffusion Problems

Now let's look at the inhomogeneous version with Dirichlet boundary conditions.

$$\begin{aligned} u_t(x,t) &= a^2 u_{xx} + h(x) \\ u(0,t) &= u_0 \\ u(L,t) &= u_L \\ u(x,0) &= f(x) \end{aligned}$$

with $x \in [0, L]$.

In this setup, the physical intuitive is that we're holding the bar against an object that has non-ambient temperature and enforcing a heating element throughout the bar that is constant in time at each point in the space.

To solve this, we write

$$\begin{aligned} u(x,t) &= v(x,t) + q(x) \\ \Rightarrow u_t &= v_t \\ \Rightarrow u_{xx} &= v_{xx} + q'' \end{aligned} \tag{11.28}$$

Substituting this back into the original equation yields

$$\begin{aligned} v_t(x,t) &= a^2 v_{xx} + a^2 q''(x) + h(x) \\ v(0,t) &= u_0 - q(0) \\ v(L,t) &= u_L - q(L) \\ v(x,0) &= f(x) - q(x) \end{aligned}$$

This makes our life difficult but what if we were to suppose that

$$\begin{aligned} 0 &= a^2 q''(x) + h(x) \\ 0 &= u_0 - q(0) \\ 0 &= u_L - q(L) \end{aligned} \tag{11.29}$$

then this would transform the problem into something we can already solve.

$$\begin{aligned} v_t(x,t) &= a^2 v_{xx} \\ u(0,t) &= u(L,t) = 0 \\ v(x,0) &= f(x) - q(x) \end{aligned} \tag{11.30}$$

Now we want to solve (11.29) before we solve (11.30).

We have

$$\begin{aligned} q''(x) &= -\frac{h(x)}{a^2} \\ \Rightarrow q'(x) &= -\int_0^x \frac{h(t)}{a^2}\ \delta t + C \\ &= H(x)\ \delta t + C \end{aligned}$$

Then we can write

$$q(x) = \int_0^x H(t)\ \delta t + Cx + D$$

From (11.29) we want

$$\begin{aligned} q(0) &= u(0) \\ \Rightarrow u(0) &= D \end{aligned}$$

so we have found the first constant. The second constant is found using the other condition

$$\begin{aligned} u(L) &= q(L) \\ &= \int_0^L H(t)\ \delta t + C \cdot L + D \\ \Rightarrow C &= \frac{1}{L}\left(u(L) - u(0) - \int_0^L H(t)\ \delta t\right) \end{aligned}$$

and thus we have $q(x)$ which you can then use when you solve for (11.30) which then gives you (11.28).

Alright that's all for the heat/diffusion equation. Yay for PDEs =D!

11.11.2 The Linear Transport Equation

So the linear transport equation is

$$u_t(x,t) + q(x)u_x(x,t) = 0$$

We require an initial value $f(x)$ at time t_0 such that

$$u(x,t_0) = f(x)$$

is satisfied, $x \in \mathbb{R}$.

11.11.2.1 Uniform Transport

The uniform transport is the simplest case, with $q(x) = c$ being a constant. We also let $t_0 = 0$ be the initial time.

$$u_t(x,t) + cu_x(x,t) = 0 \tag{11.31}$$

We require an initial value $f(x)$ at time $t = 0$ such that

$$u(x,0) = f(x) \tag{11.32}$$

Time for a change of variables

$$\begin{aligned} \xi &= x - ct \\ \tau &= t \end{aligned}$$

Intuitively, this change of variables shifts the frame moving with velocity c to the right (left) if $c > 0$ ($c < 0$). Hence we have

$$\begin{aligned} u(x,t) &= v(x - ct, t) \\ &= v(\xi, \tau) \end{aligned} \tag{11.33}$$

Using the change rule we have partial derivatives

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{\partial v}{\partial \xi} \cdot \frac{\partial \xi}{\partial t} + \frac{\partial v}{\partial \tau} \cdot \frac{\partial \tau}{\partial t} \\ &= \frac{\partial v}{\partial \xi} \cdot (-c) + \frac{\partial v}{\partial \tau} \cdot (1) \end{aligned} \tag{11.34}$$

and

$$\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial v}{\partial \xi} \cdot \frac{\partial \xi}{\partial x} + \frac{\partial v}{\partial \tau} \cdot \frac{\partial \tau}{\partial x} \\ &= \frac{\partial v}{\partial \xi} \cdot (1) + \frac{\partial v}{\partial \tau} \cdot (0) \end{aligned} \tag{11.35}$$

Substituting these results back into (11.31) yields

$$0 = -cv_\xi + v_\tau + cv_\xi \tag{11.36}$$
$$= v_\tau \tag{11.37}$$

but wait $v_\tau = 0$ that means $v(\xi, \tau) = v(\xi)$. So in fact we have the following proposition.

Proposition 11.141 *Any $f \in C^1$ of the variable $\xi := x - ct$ solves the transport equation (11.31).*

(Sketch of proof)

Let $u(x,t) = f(x - ct) = f(\xi)$. Then

$$\begin{aligned} u_t &= -cf'(x - ct) \\ &= -cf(\xi) \\ u_x &= f'(x - ct) \\ &= f'(\xi) \\ \Rightarrow u_t + cu_x &= -cf'(x - ct) + cf'(x - ct) = 0 \end{aligned}$$

Remark 11.142 *The parameter (constant) c in (11.31) is called the **wavespeed**.*

Remark 11.143 *The variable $\xi = x - ct$ is called the **characteristic variable**.*

Since the solution comes out immediately, we can do a quick example.

$$
\begin{aligned}
0 &= u_t(x,t) + 2u_x(x,t) \\
u(x,0) &= \frac{1}{1+x^2}
\end{aligned}
$$

The solution is $u(x,t) = f(x-ct) = \frac{1}{1+(x-2t)^2}$.

If we have initial conditions

$$u(x,t_0) = f(x) \tag{11.38}$$

Then the solution will be

$$
\begin{aligned}
u(x,t) &= f(x - ct + ct_0) \\
&= f(\xi + ct_0)
\end{aligned}
$$

11.11.2.2 Transport with Decay/Growth

$$u_t(x,t) + cu_x(x,t) + au(x,t) = 0 \tag{11.39}$$

We require an initial value $f(x)$ at time $t = 0$ such that

$$u(x,t_0) = f(x)$$

Here, we have a constant a that governs the rate of decay/growth. We do exactly the same as in (11.33), (11.34), and (11.35) to obtain a 1st order equation and solve using an integrating factor

$$
\begin{aligned}
0 &= v_\tau + av(\xi,\tau) \\
\Rightarrow v &= e^{-a\tau}K(\xi)
\end{aligned}
$$

and solving for the initial conditions yields

$$
\begin{aligned}
u(x,t_0) &= v(x - ct + ct_0, t_0) \\
&= e^{-a(t-t_0)}K(x - ct + ct_0)
\end{aligned}
$$

which is the final solution to the problem.

Proposition 11.144 *We consider a few cases for the constants a and c.*

1. $a < 0,\ c > 0$: *growth and shift right*
2. $a > 0,\ c < 0$: *decay and shift left*
3. $a < 0,\ c < 0$: *growth and shift left*
4. $a > 0,\ c > 0$: *decay and shift right*

11.11.2.3 Non-uniform Linear Transport

We have

$$u_t(x,t) + c(x)u_x(x,t) = 0 \tag{11.40}$$

with initial condition

$$u(x,0) = f(x) \tag{11.41}$$

This time, we have a variable substitution

$$\begin{aligned} \xi &= g(x) - t \\ \tau &= t \end{aligned} \tag{11.42}$$

Then we have

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{\partial v}{\partial \xi} \cdot \frac{\partial \xi}{\partial t} + \frac{\partial v}{\partial \tau} \cdot \frac{\partial \tau}{\partial t} \\ &= \frac{\partial v}{\partial \xi} \cdot (-1) + \frac{\partial v}{\partial \tau} \cdot (1) \end{aligned} \tag{11.43}$$

and

$$\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial v}{\partial \xi} \cdot \frac{\partial \xi}{\partial x} + \frac{\partial v}{\partial \tau} \cdot \frac{\partial \tau}{\partial x} \\ &= \frac{\partial v}{\partial \xi} \cdot (g'(x)) + \frac{\partial v}{\partial \tau} \cdot (0) \end{aligned} \tag{11.44}$$

Applying (11.43) and (11.44) to (11.41) yields

$$\begin{aligned} 0 &= u_t + c(x)u_x \\ &= -v_\xi + v_\tau + c(x)v_\xi g'(x) \\ &= v_\xi(c(x)g'(x) - 1) + v_\tau \end{aligned}$$

As in (11.36) we would like v_ξ to vanish. In order to do that we set

$$\begin{aligned} g'(x)c(x) &= 1 \\ \Rightarrow g(x) &= \int^x \frac{\delta s}{c(s)} \end{aligned}$$

With this, we can solve

$$\xi = g(x) - t \tag{11.45}$$

$$= \int^x \frac{\delta s}{c(s)} - t \tag{11.46}$$

Using the initial condition, we have

$$\begin{aligned} u(x,0) &= v(\xi,0) \\ &= v(g(x) - 0, 0) \\ &= v(g(x)) \\ &= f(x) \end{aligned}$$

At $t = 0$ we have $x = g^{-1}(\xi)$ which gives us

$$\begin{aligned} v(g(x)) &= v(\xi) \\ &= f(g^{-1}(\xi)) \\ \Rightarrow u(x,t) &= f(g^{-1}(g(x) - t)) \end{aligned}$$

11.11.2.4 The Characteristic Curve

The characteristic curve(s) illustrate how the solution to the PDE varies in the (x,t)-plane. First, we parametrically define

$$x = x(t)$$

and set

$$\begin{aligned} \frac{\delta x}{\delta t} &= c(x) \\ \Rightarrow \delta t &= \frac{\delta x}{c(x)} \end{aligned} \tag{11.47}$$

The gradient of the characteristic curve at an arbitrary point (x,t) is the wave speed. Furthermore, if $c(x') = 0$ then x' is a fixed point for (11.47) and the horizontal line $x = x'$ is a stationary curve.

From (11.45) and the fact that (11.47) is separable, we have

$$\begin{aligned} g(x) &= \int^x \frac{\delta s}{c(s)} \\ &= \int \delta t \\ &= t + k \end{aligned} \tag{11.48}$$

where k is some constant. Then by taking the inverse we get

$$x(t) = g^{-1}(t+k)$$

Let's suppose $c(x) = k'$. From (11.48) we have

$$\begin{aligned} g(x) &= \int^x \frac{\delta s}{c(s)} \\ &= \int^x \frac{\delta s}{k'} \\ \Rightarrow \frac{1}{k'}x &= t + k \\ x - k't &= k^* \end{aligned}$$

which are straight lines. In solving this, we can actually show that

$$\begin{aligned} u(x,t) &= v(\xi, \tau) \\ &= v(\xi) \end{aligned}$$

where the solution is dependent only on $\xi = g(x) - t$

Remark 11.145 *The solution $u(x,t)$ is constant along the characteristic curves*

11.11.2.5 Solving Linear Transport using Characteristics

We look at two examples before indicating closing remarks.

$$u_t + \frac{1}{1+x^2}u_x = 0$$

We have

$$\begin{aligned} g(x) &= \int^x \frac{1}{c(s)} \\ &= \frac{1}{3}x^3 + x \\ &= t + k \end{aligned}$$

Hence we have $\xi = g(x) - t = \frac{1}{3}x^3 + x - t$. We see the wave movement and characteristic curves. For the

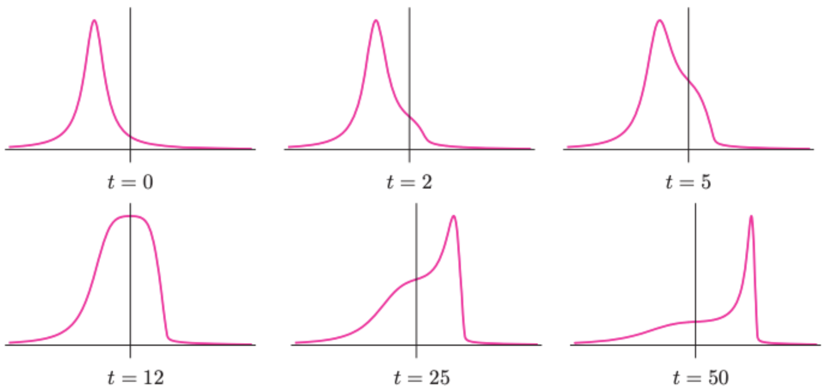


Figure 11.1: Solution to $u_t + \frac{1}{1+x^2}u_x$

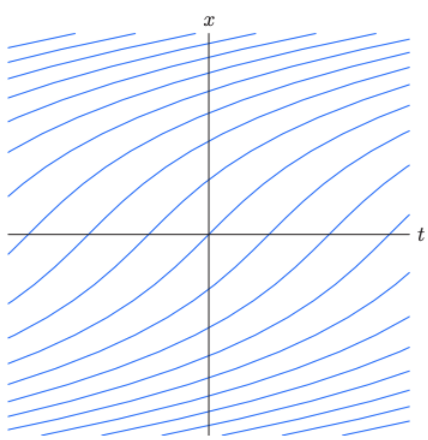


Figure 11.2: Characteristic curve of $\frac{1}{3}x^3 + x = t + k$

characteristic curve, if we specify the curve to intersect the x-axis at $(x^*, 0)$ then we have

$$\begin{aligned} u(x^*, 0) &= u(x, t) \\ &= f(x^*) \end{aligned}$$

since $u(x, t)$ is constant on the curve as we have remarked before.

The second example is also pretty cool.

$$u_t + (x^2 - 1)u_x = 0$$

We have

$$\begin{aligned} g(x) &= \int^x \frac{1}{x^2-1} \\ &= \frac{1}{2}\ln\left(\left|\frac{x-1}{x+1}\right|\right) \\ &= t+k \end{aligned}$$

Hence we have $\xi = g(x) - t = \frac{1}{2}\ln\left(\left|\frac{x-1}{x+1}\right|\right) - t$. Here the initial value is a Gaussian - $u(x,0) = e^{-x^2}$. We see the wave movement and characteristic curve below.

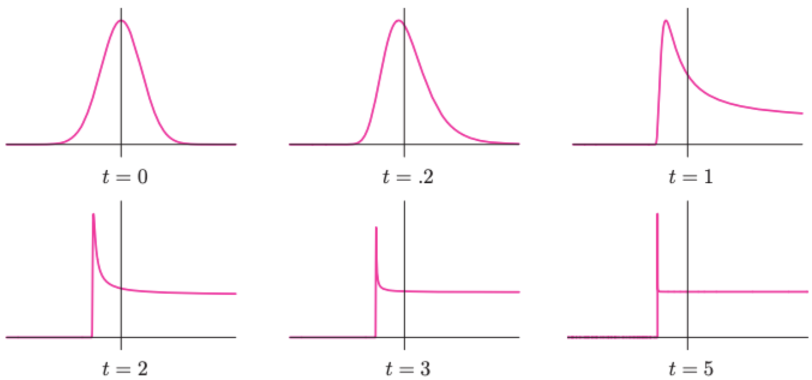


Figure 11.3: Solution to $u_t + (x^2-1)u_x$

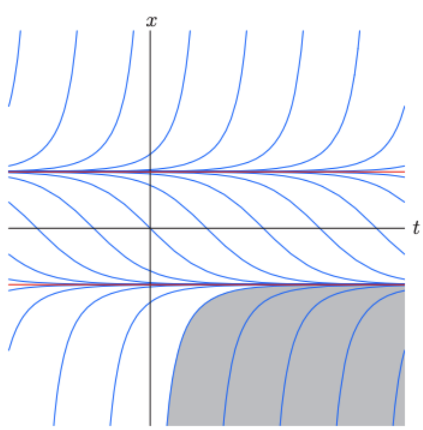


Figure 11.4: Characteristic curve of $\frac{1}{2}\ln\left(\left|\frac{x-1}{x+1}\right|\right) = t+k$

Remark 11.146 *There is a unique characteristic curve through each point of the (x,t)-plane.*

Remark 11.147 *The characteristic curves don't cross each other.*

Remark 11.148 *Each non-horizontal curve is a strictly monotone function. That is, each point on a wave always moves in the same direction and cannot reverse its direction of propagation.*

Remark 11.149 *As $t \to \infty$, we either have $x(t) \to x^*$ a fixed point with $c(x^*) = 0$ or it tends to either $\pm\infty$.*

11.11.3 The Wave Equation

11.11.3.1 Intuition of the Wave Equation

We will actually not derive the wave equation. However, for intuition we will include a diagram and explain the composition of the equation.

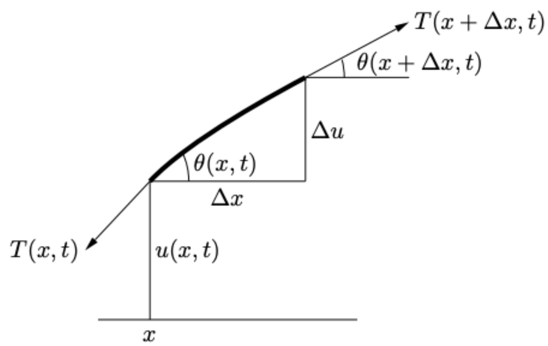


Figure 11.5: The dynamics of a tiny element of a string

So the 'full' wave equation looks something like

$$\rho(x)u_{tt}(x,t) = \frac{\partial}{\partial x}\left(\kappa(x)\frac{\partial u}{\partial x}\right) \tag{11.49}$$

Here we have

1. $u(x,t)$ is the displacement of the string at time t and position x.
2. $\rho(x) > 0$ being the density
3. $\kappa(x) > 0$ being the stiffness/tension of the string (independent of t)

In order to specify a unique solution to this PDE, we will require both its initial position and initial velocity.

11.11.3.2 One-dimension Wave Equation

From (11.49) we simplify by assuming $\rho(x)$, $\kappa(x)$ are constants to obtain

$$u_{tt}(x,t) = cu_{xx}(x,t) \tag{11.50}$$

where $c = \sqrt{\frac{\kappa}{\rho}}$ is the **wave speed**. For now, we assume initial position and velocity to be

$$\begin{aligned} u(x,0) &= f(x) \\ u_t(x,0) &= g(x) \end{aligned} \tag{11.51}$$

where $t_0 = 0$ for simplicity.

To solve this equation, notice that

$$\begin{aligned} u_{tt}(x,t) - cu_{xx}(x,t) &= \partial_{tt}u - c\partial_{xx}u \\ &= (\partial_t - c\partial_x)(\partial_t + c\partial_x) \end{aligned} \tag{11.52}$$

because of equality of partial derivatives.

Remark 11.150 *Observe that if u solves either the left or right transport equations then u also solves the wave equation. To see this let*

$$\begin{aligned} u(x,t) &= k_1 p(x+ct) + k_2 q(x-ct) \\ \Rightarrow u_{tt} &= c^2 k_1 p'' + c^2 k_2 q'' \\ \Rightarrow u_{xx} &= k_1 p'' + k_2 q'' \end{aligned}$$

Theorem 11.151 *Every $f \in C^2$ solution to the wave equation is a superposition of solutions travelling to the right with $p(x-ct)$, and left with $q(x+ct)$.*

Let $\xi = x - ct$ and $\eta = x + ct$. We want to show that

$$\begin{aligned} u(x,t) &= v(x-ct, x+ct) \\ &= v(\xi, \eta) \\ &= u\left(\frac{\eta+\xi}{2}, \frac{\eta-\xi}{2}\right) \\ &= p(\xi) + q(\eta) \end{aligned} \tag{11.53}$$

(Sketch of proof)

Applying the chain rule and differentiating twice, you obtain

$$\begin{aligned} u_{tt} &= c^2\left(v_{\xi\xi} - 2v_{\xi\eta} + v_{\eta\eta}\right) \\ u_{xx} &= v_{\xi\xi} + 2v_{\xi\eta} + v_{\eta\eta} \\ \Rightarrow 0 &= u_{tt} - c^2 u_{xx} = -4c^2 v_{\eta\xi} \end{aligned}$$

Since $c \neq 0$ then we have $v_{\eta\xi} = 0$. The order of differentiation matters so

$$\begin{aligned} 0 &= v_{\eta\xi} \\ &= \frac{\partial}{\partial\xi}(v_\eta) \end{aligned}$$

which if we define $q(\eta) = \int^\eta v_\eta(s)\,\partial s$ we obtain

$$\begin{aligned} v(\xi,\eta) &= \int^\eta v_\eta(s)\,\partial s + p(\xi) \\ &= q(\eta) + p(\xi) \end{aligned}$$

So now we know that every solution $u(x,t) = p(x-ct) + q(x+ct) = p(\xi) + q(\eta)$ we can use the initial conditions. Using (11.51) we obtain

$$\begin{aligned} f(x) &= p(x) + q(x) \\ g(x) &= -cp'(x) + cq'(x) \end{aligned} \tag{11.54}$$

doing some plus-minus we get

$$\begin{aligned} g(x) - cf'(x) &= -2cp'(x) \\ \Rightarrow p(x) &= \frac{1}{2}f(x) - \frac{1}{2c}\int_0^x g(s)\,\partial s + A \end{aligned}$$

and using we obtain

$$\begin{aligned} q(x) &= f(x) - p(x) \\ &= \frac{1}{2}f(x) + \frac{1}{2c}\int_0^x g(s)\,\partial s - A \end{aligned}$$

Putting all this together and using (11.53) we obtain

$$\begin{aligned} p(\xi) &= \frac{1}{2}f(\xi) - \frac{1}{2c}\int_0^\xi g(s)\,\partial s + A \\ q(\eta) &= \frac{1}{2}f(\eta) + \frac{1}{2c}\int_0^\eta g(s)\,\partial s - A \\ \Rightarrow u(x,t) &= \frac{1}{2}(f(\xi) + f(\eta) + \frac{1}{2c}\int_\xi^\eta g(s)\,\partial s \end{aligned}$$

The integrals were combined by swapping the signs and swapping the order of integration.

Remark 11.152 *(D'Alembert's solution)*

$$u(x,t) = \frac{1}{2}(f(\xi) + f(\eta) + \frac{1}{2c}\int_\xi^\eta g(s)\,\partial s$$

Is known as D'Alembert's solution to the wave equation.

11.11.3.3 One-dimension with Boundary Conditions

From (11.50) and (11.51) we now add in a boundary condition on a string of finite interval $[0, L]$. For example we have

$$\begin{aligned} u(0,t) &= f(x) \\ u(L,t) &= g(x) \end{aligned} \tag{11.55}$$

As with the heat equation in the previous section, we assume that the solution is separable

$$\frac{T''(t)}{c^2T(t)} = \frac{X''(x)}{X(x)} = -\lambda \tag{11.56}$$

The case for $X(x)$ is similar - we obtain, using the initial conditions,

$$\lambda_k = \left(\frac{n\pi}{L}\right)^2$$

and

$$X_n = \sin\left(\frac{n\pi}{L}x\right)$$

For $T(t)$, we also have something similar

$$T''(t) + c^2\frac{n^2\pi^2}{L^2}T(t) = 0$$

giving

$$T(t) = a_n \cos\left(c\frac{n\pi}{L}t\right) + b_n \sin\left(c\frac{n\pi}{L}t\right)$$

Hence a solution is

$$\begin{aligned} u(x,t) &= X(x)T(t) \\ &= a_n \cos\left(c\frac{n\pi}{L}t\right)\sin\left(\frac{n\pi}{L}x\right) + b_n \sin\left(c\frac{n\pi}{L}t\right)\sin\left(\frac{n\pi}{L}x\right) \end{aligned} \tag{11.57}$$

Summing the eigensolutions as an (possible) infinite linear combination, assuming that the required properties are satisfied, we obtain

$$u(x,t) = \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi}{L}ct\right)\sin\left(\frac{n\pi}{L}x\right) + b_n \sin\left(\frac{n\pi}{L}ct\right)\sin\left(\frac{n\pi}{L}x\right) \tag{11.58}$$

To solve for the coefficients, we utilise the initial conditions (11.55) to get

$$\begin{aligned} u(x,0) &= f(x) \\ &= \sum_{n=1}^{\infty} a_n \sin\left(\frac{n\pi}{L}x\right) \\ \Rightarrow a_n &= \frac{2}{L}\int_0^L f(x)\sin\left(\frac{n\pi}{L}x\right)\ \partial x \end{aligned}$$

and

$$\begin{aligned} u_t(x,0) &= g(x) \\ &= \sum_{n=1}^{\infty} b_n \left(c\frac{n\pi}{L}\right)\sin\left(\frac{n\pi}{L}x\right) \\ \Rightarrow b_n &= \frac{2}{cn\pi}\int_0^L g(x)\sin\left(\frac{n\pi}{L}x\right)\ \partial x \end{aligned}$$

because $b_n\left(c\frac{n\pi}{L}\right)$ is treated as the sin coefficient of $g(x)$.

So now we have a question: For what values of $t \in \mathbb{R}$ does the above series converge? And what happens to the series at $t = 0$? Okay sorry that's two questions.

Notice that $a_n \cos\left(c\frac{n\pi}{L}t\right)\sin\left(\frac{n\pi}{L}x\right)$ is the Fourier sine series with an additional cos component. It turns out that neither the sin nor cos component in (11.57) affects convergence.

Lemma 11.153 *If $h(x)$ (as in the solution to some wave equation) is piecewise smooth, then the a_n component converges $\forall x, t \in \mathbb{Z}$. a_n is defined as*

$$\frac{2}{L}\int_0^L f(x)\sin\left(\frac{n\pi}{L}x\right)\ \delta x$$

(Sketch of proof)

Using the trig identity $\cos(A)\sin(B) = \frac{1}{2}(\sin(A+B)+\sin(A-B))$ *we have*

$$\begin{aligned}\sum_{n=0}^{\infty} a_n \cos\left(c\frac{n\pi}{L}t\right)\sin\left(\frac{n\pi}{L}x\right) &= \frac{1}{2}\sum_{n=0}^{\infty} a_n \sin\left(\frac{n\pi}{L}(x+ct)\right) + a_n\sin\left(\frac{n\pi}{L}(x-ct)\right)\\ &= \frac{1}{2}\sum_{n=0}^{\infty} \widetilde{h}(x+ct)+\widetilde{h}(x-ct)\end{aligned}$$

where $\widetilde{h}(x)$ is the odd periodic extension of $h(x)$. So then we also have convergence at $t = 0$. As we hoped.

In a similar fashion, we can propose the next lemma.

Lemma 11.154 *If $h(x)$ (as in the solution to some wave equation) is piecewise smooth, then the b_n component converges $\forall x, t \in \mathbb{Z}$. b_n is defined as*

$$\frac{2}{nc\pi}\int_0^L g(x)\sin\left(\frac{n\pi}{L}x\right)\ \delta x = \frac{L}{nc\pi}\frac{2}{L}\int_0^L g(x)\sin\left(\frac{n\pi}{L}x\right)\ \delta x \tag{11.59}$$

(Sketch of proof) Use the trig identity $\sin(A)\sin(B) = \frac{1}{2}(\cos(A-b)+\cos(A+B)$.

$$\begin{aligned}\sum_{n=0}^{\infty} b_n \sin\left(\frac{n\pi}{L}ct\right)\sin\left(\frac{n\pi}{L}x\right) &= \sum_{n=0}^{\infty} \frac{b_n}{2}\cos\left(\frac{n\pi}{L}(x-ct)\right) - \frac{b_n}{2}\cos\left(\frac{n\pi}{L}(x+ct)\right)\\ &= \frac{1}{2C}\int_{x-ct}^{x+ct} \widetilde{b}_n \sin\left(\frac{n\pi}{L}z\right)\ \delta z\end{aligned}$$

where we re-write b_n using (11.59) and apply the Fundamental Theorem of Calculus. Finally, we define $\widetilde{b}_n$ as

$$\widetilde{b}_n = \frac{2}{L}\int_0^L g(x)\sin\left(\frac{n\pi}{L}x\right)\ \delta x$$

Here $\widetilde{b}_n$ is the Fourier sine coefficient of $g(x)$.

As you might have expected, both these expressions will converge to the D'Alembert's solution to the wave equation on the line with initial conditions equal to the Fourier sine series of $f(x)$ and $g(x)$, provided the convergence works out.

11.11.3.4 One-dimension with Neumann Boundary Conditions

From (11.50) and (11.51) we now add in the Neumann boundary condition on a string of finite interval $[0, L]$. For example we have

$$\begin{aligned}u_x(0,t) &= f(x)\\ u_x(L,t) &= g(x)\end{aligned} \tag{11.60}$$

Assume the solution is separable $u(x,t) = X(x)T(t)$. We then proceed as in the Dirichlet boundary condition case, and obtain

$$\lambda_0 = 0$$
$$\lambda_k = \left(\frac{n\pi}{L}\right)^2$$

and

$$x_0 = 1$$
$$x_n = \cos\left(\frac{n\pi}{L}x\right)$$

as consistent with the heat equation. For $T(t)$ we have

$$T''(t) = \lambda c^2 T(t) = 0$$
$$\Rightarrow T(t) = a_0 + b_0 t$$

if $\lambda = 0$. Otherwise we have for

$$T(t) = a_n \cos\left(\frac{nc\pi}{L}t\right) + b_n \sin\left(\frac{nc\pi}{L}t\right)$$

so we have

$$u(x,t) = a_0 + b_0 t + \sum_{n=1}^{\infty} a_n \cos\left(\frac{nc\pi}{L}t\right)\cos\left(\frac{n\pi}{L}x\right) + b_n \sin\left(\frac{nc\pi}{L}t\right)\cos\left(\frac{n\pi}{L}x\right) \tag{11.61}$$

Again, we observe the initial values

$$u_x(x,0) = f(x)$$
$$= a_0 + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi}{L}x\right)$$
$$\Rightarrow a_0 = \frac{1}{L}\int_0^L f(x)\ \delta x$$
$$\Rightarrow a_n = \frac{2}{L}\int_0^L f(x)\cos\left(\frac{n\pi}{L}x\right)\ \delta x$$

and for b_n we have

$$u_t(x,0) = g(x)$$
$$= b_0 + \sum_{n=1}^{\infty} b_n\left(\frac{nc\pi}{L}\right)\cos\left(\frac{n\pi}{L}x\right)$$
$$\Rightarrow b_0 = \frac{1}{L}\int_0^L g(x)\ \delta x$$
$$\Rightarrow b_n = \frac{2}{nc\pi}\int_0^L g(x)\cos\left(\frac{n\pi}{L}x\right)\ \delta x$$

Remark 11.155 *For the solution, (11.61), we take note that if $b_0 \neq 0$, then there is an unstable node and the solution grows towards infinity as time passes. However if it is equal to zero, the solution is time periodic and oscillates around the initial displacement.*

11.11.4 External Force of the Wave Equation

$$u_{tt}(x,t) = cu_{xx}(x,t) + F(x,t) \tag{11.62}$$

where $c = \sqrt{\frac{\kappa}{\rho}}$ is the **wave speed**. For now, we assume initial position and velocity to be

$$\begin{aligned} u(x,0) &= f(x) \\ u_t(x,0) &= g(x) \end{aligned} \tag{11.63}$$

where $t_0 = 0$ for simplicity.

Theorem 11.156 *The solution to the general initial value problem of the wave equation with external force is*

$$u(x,t) = \frac{1}{2}\left(f(x+ct) + f(x-ct)\right) + \frac{1}{2c}\int_{x-ct}^{x+ct} g(z)\ \delta z + \int_0^t \int_{x-c(t-s)}^{x+c(t-s)} F(y,s)\ \delta y \delta s \tag{11.64}$$

You can think of it as having a homogeneous solution and a particular solution as in the ODE case. It is a nice analogy.