Mathematics 3

Lecture notes from the second-year undergraduate course *Matematika 3* (Mathematics 3) given by doc. dr. Oliver Dragičević at the Faculty of Mathematics and Physics at the University of Ljubljana in the academic year 2019-20. The course covers vector calculus and ordinary differential equations, and also gives an introduction to the calculus of variations, Hilbert spaces, and Fourier series. Credit for the material covered in these notes is due to professor Dragičević, while the voice, typesetting, and translation to English in this document are my own. Note: these notes do not include proofs of the propositions and theorems, and are thus incomplete.

Disclaimer: This document will inevitably contain some mistakes—both simple typos and legitimate errors. Keep in mind that these are the notes of an undergraduate student in the process of learning the material himself, so take what you read with a grain of salt. If you find mistakes and feel like telling me, I will be grateful and happy to hear from you, even for the most trivial of errors. You can reach me by email, in English, Slovene, or Spanish, at ejmastnak@gmail.com.

For more notes like this, visit ejmastnak.github.io/fmf

Elijan J. Mastnak
Winter Semester 2019-20
Last update: February 14, 2022
Faculty of Mathematics and Physics, University of Ljubljana

Contents

1 Parameter-Dependent Integrals

1.1 Introduction to Parameter-Dependent Integral

1.1.1 Definition: Parameter-Dependent Integrals

Let $a, b, c, d \in \mathbb{R}$ be real numbers and let $f : [a, b] \times [c, d] \to \mathbb{R}$ be an integrable scalar function of two variables. The parameter-dependent integral F is the scalar function $F : [c, d] \to \mathbb{R}$ defined by

$$F(t) = \int_a^b f(x, t) \, \mathrm{d}x$$

where the variable t is called the parameter.

1.1.2 Proposition: Continuity of Parameter-Dependent Integrals

Let $a, b, c, d \in \mathbb{R}$ be real numbers and $f : [a, b] \times [c, d] \to \mathbb{R}$ be a continuous function. In this case, the function $F : [c, d] \to \mathbb{R}$ defined by

$$F(t) = \int_a^b f(x, t) \, \mathrm{d}x$$

is also continuous.

1.1.3 Proposition: Differentiability of Parameter-Dependent Integrals I

Let $f(x, t) : P = [a, b] \times [c, d] \to \mathbb{R}$ be a continuous function and continuously differentiable with respect to the second variable (i.e. $\frac{\partial f}{\partial t}$ is continuous everywhere on P). In this case, the function

$$F(t) = \int_a^b f(x, t) \, \mathrm{d}x$$

is differentiable on $[c, d]$ and

$$F'(t) = \int_a^b \frac{\partial f}{\partial t} \, \mathrm{d}x$$

1.1.4 Proposition: Differentiability of Parameter-Dependent Integrals II

This proposition extends the previous one to include functions in the limits of the integral. Let $f(x, t) : P = [a, b] \times [c, d] \to \mathbb{R}$ be a continuous function and continuously differentiable with respect to the second variable (i.e. $\frac{\partial f}{\partial t}$ exists everywhere on P). Also, let $\alpha, \beta : [c, d] \to [a, b]$ be two differentiable functions. In this case, the function

$$F(t) := \int_{\alpha(t)}^{\beta(t)} f(x, t) \, \mathrm{d}x$$

is differentiable on $[c, d]$ and

$$F'(t) = \frac{\mathrm{d}F(t)}{\mathrm{d}t} = \int_{\alpha(t)}^{\beta(t)} \frac{\partial f}{\partial t} \, \mathrm{d}x + f(\beta(t), t)\beta'(t) - f(\alpha(t), t)\alpha'(t)$$

1.1.5   Simplified (Restricted) Fubini's Theorem for Double Integration

Define the finite constants $a, b, c, d \in \mathbb{R}$ and let $f(x,t) : P = [a,b] \times [c,d] \to \mathbb{R}$ be a continuous function. Note that the rectangle P on which f is defined is bounded. In this case

$$\int_a^b \left(\int_c^d f(x,y)\,\mathrm{d}y \right) \mathrm{d}x = \int_c^d \left(\int_a^b f(x,y)\,\mathrm{d}x \right) \mathrm{d}y$$

The general form of Fubini's theorem, given later in this section, extends to functions with an unbounded domain.

1.2   Improper Parameter-Dependent Integrals

1.2.1   Definition: Uniform Convergence of an Improper Parameter-Dependent Integral

Let $f : [a, \infty) \times [c,d] \to \mathbb{R}$ be a continuous function. In this case, the parameter-dependent integral

$$F(t) = \int_a^\infty f(x,t)\,\mathrm{d}x$$

converges uniformly on the interval $[c,d]$ if for all $\epsilon > 0$ there exists $M_0 \in \mathbb{R}$; $M_0 \geq a$ such that

$$\left| \int_M^\infty f(x,t)\,\mathrm{d}x \right| < \epsilon$$

for all $t \in [c,d]$ and for all $M > M_0$.

1.2.2   Proposition: Condition for Uniform Convergence

Let $f : [a, \infty) \times [c,d] \to \mathbb{R}$ be a continuous function. Assume also there exists an integrable function $\phi : [a, \infty) \times [0, \infty)$ such that $|f(x,t)| \leq \phi(x)$ for all $x \geq a$ and for all $t \in [c,d]$. In this case the parameter-dependent integral

$$F(t) := \int_a^\infty f(x,t)\,\mathrm{d}x$$

converges uniformly on the interval $[c,d]$.

1.2.3   Proposition: Uniform Convergence Implies Continuity

Let $f : [a, \infty) \times [c,d] \to \mathbb{R}$ be a continuous function and let the parameter-dependent integral

$$F(t) := \int_a^\infty f(x,t)\,\mathrm{d}x$$

converge uniformly on the interval $[c,d]$. In this case F is continuous on $[c,d]$.

1.2.4   Proposition: Uniform Convergence Implies Integrability

Let $f : [a, \infty) \times [c,d] \to \mathbb{R}$ be a continuous function and let the parameter-dependent integral

$$F(t) := \int_a^\infty f(x,t)\,\mathrm{d}x$$

converge uniformly on the interval $[c,d]$. In this case F is integrable on $[c,d]$ and

$$\int_c^d \left(\int_a^\infty f(x,t)\,\mathrm{d}x \right) \mathrm{d}t = \int_a^\infty \left(\int_c^d f(x,t)\,\mathrm{d}t \right) \mathrm{d}x$$

1.2.5 Example: An Identity for $\ln\left(\frac{d}{c}\right)$ Using Double Integration

Let $F(t)$ be the parameter-dependent integral given by

$$F(t) := \int_0^\infty e^{-tx}\,\mathrm{d}x \quad \text{where} \quad 0 < t < \infty$$

and define the constants $c, d \in \mathbb{R}$ such that $0 < c \le t \le d < \infty$.

Without proof, the F converges uniformly on $[c, d]$ by the definition of uniform convergence. Because F converges uniformly, it is also continuous and integrable on $[c, d]$. Using analysis techniques from Matematika 1 we can calculate

$$F(t) = \frac{1}{t} \qquad \text{and} \qquad \int_c^d F(t)\,\mathrm{d}t = \int_c^d \frac{1}{t}\,\mathrm{d}t = \ln(t)\Big|_c^d = \ln\left(\frac{d}{c}\right)$$

(Because t is positive, we drop the absolute value in the logarithm). By applying the previous proposition and making use of double integration, we can show that

$$\int_c^d F(t)\,\mathrm{d}t = \int_c^d \int_0^\infty e^{-tx}\,\mathrm{d}x\,\mathrm{d}t = \int_0^\infty \int_c^d e^{-tx}\,\mathrm{d}t\,\mathrm{d}x$$

$$= \int_0^\infty \frac{e^{-tx}}{-x}\Big|_{t=c}^d \mathrm{d}x = \int_0^\infty \frac{e^{-dx} - e^{-cx}}{-x}\,\mathrm{d}x$$

$$= \int_0^\infty \frac{e^{-cx} - e^{-dx}}{x}\,\mathrm{d}x$$

Note that in the derivation, we used the indefinite integral identity

$$\int e^{-tx}\,\mathrm{d}t = \frac{e^{-tx}}{-x} + C$$

It follows from the above discussion that

$$\int_0^\infty \frac{e^{-cx} - e^{-dx}}{x}\,\mathrm{d}x = \ln\left(\frac{d}{c}\right) \quad \forall \quad c, d > 0$$

1.2.6 Theorem: Differentiability of Improper Parameter-Dependent Integrals

Let $f : [a, \infty) \times [c, d] \to \mathbb{R}$ be a continuous function such that:

1. f's partial derivative with respect to t, given by $\frac{\partial f}{\partial t} : [a, \infty) \times [c, d] \to \mathbb{R}$ is continuous on $[c, d]$.

2. For all $t \in [c, d]$ the function $F(t)$ given by

$$F(t) := \int_a^\infty f(x, t)\,\mathrm{d}x$$

   exists. Note that the existence of $F(t)$ is not necessarily trivial.

3. The function $G(t)$ given by

$$G(t) := \int_a^\infty \frac{\partial f(x, t)}{\partial t}\,\mathrm{d}x$$

   converges uniformly on the interval $[c, d]$.

In this case F is differentiable on $[c, d]$ and

$$F'(t) = \frac{\mathrm{d}F(t)}{\mathrm{d}t} = G(t)$$

or, more thoroughly,

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_a^\infty f(x, t) \, \mathrm{d}x = \int_a^\infty \frac{\partial f(x, t)}{\partial t} \, \mathrm{d}x$$

1.2.7 Theorem: General Form of Fubini's Theorem

Define the constants $a, b, c, d \in \mathbb{R} \cup \{-\infty, \infty\}$ and let f be a continuous function such that either $f(x, t) \geq 0$ for all $x, t \in \mathbb{R}$ or

$$\int_a^b \int_c^d |f(x, t)| \, \mathrm{d}t \, \mathrm{d}x < \infty$$

In this case,

$$\int_a^b \int_c^d |f(x, t)| \, \mathrm{d}t \, \mathrm{d}x = \int_c^d \int_a^b |f(x, t)| \, \mathrm{d}x \, \mathrm{d}t$$

1.2.8 Example: Evaluating the Dirichlet integral

We aim to evaluate the Dirichlet integral

$$I = \int_0^\infty \frac{\sin x}{x} \, \mathrm{d}x$$

using double integration and the ideas of the past sections. First, define a new function of $I(t)$ given by:

$$I(t) := \int_0^\infty \frac{\sin x}{x} e^{-tx} \, \mathrm{d}x \tag{1.1}$$

Without proof, we note that for $t > 0$ the integral represented by $I(t)$ converges absolutely, permitting us to differentiate with respect to t inside the integral for values of $t > 0$.

$$I'(t) = \int_0^\infty \frac{\partial}{\partial t} \frac{\sin x}{x} e^{-tx} \, \mathrm{d}x = -\int_0^\infty e^{-tx} \sin x \, \mathrm{d}x$$

Evaluating the integral by parts yields

$$I'(t) = \left. \frac{e^{-tx}(t \sin x + \cos x)}{t^2 + 1} \right|_{x=0}^{x=\infty} = -\frac{1}{t^2 + 1}$$

Thus

$$I(t) = \int I'(t) = \int -\frac{1}{t^2 + 1} \, \mathrm{d}t = -\arctan t + C$$

$$I(t) = -\arctan t + C \tag{1.2}$$

From Equation 1.1 we see that

$$\lim_{t \to \infty} I(t) = 0$$

Meanwhile from Equation 1.2 we see that

$$\lim_{t \to \infty} I(t) = \lim_{t \to \infty} \left(-\arctan t + C \right)$$

From which follows that

$$0 = -\frac{\pi}{2} + C$$
$$C = \frac{\pi}{2}$$
$$I(t) = -\arctan t + \frac{\pi}{2}$$
$$I(0) = \frac{\pi}{2} \tag{1.3}$$

Referring back to Equation 1.1, we also see that

$$I(0) = \int_0^\infty \frac{\sin x}{x} e^{-(0)x} \, \mathrm{d}x = \int_0^\infty \frac{\sin x}{x} \, \mathrm{d}x$$

Plugging in the value from Equation 1.3 gives

$$\int_0^\infty \frac{\sin x}{x} \, \mathrm{d}x = \frac{\pi}{2}$$

Note, again without proof, that because of the uniform convergence of $I(t)$ on a small interval $[0, \tau), \tau > 0$ around $t = 0$ we are justified in extending our derivation of $I = \frac{\pi}{2}$ from values of $t > 0$ to value of $t \geq 0$.

1.2.9   Note: Weierstrass Test for Uniform Convergence

Let $[c, d] \in \mathbb{R}$ be finite constants such that $c < d$ and let $F(y)$ be the function given by:

$$F(y) = \int_a^\infty f(x, y) \, \mathrm{d}x \quad y \in [c, d]$$

More so, let $W : [a, \infty] \to [0, \infty]$ be a non-negative function of x such that the improper integral

$$\int_0^\infty W(x) \, \mathrm{d}x$$

converges. In this case, if

$$|f(x, y)| \leq W(x) \quad \forall \quad y \in [c, d]$$

then $F(y)$ converges uniformly for all $y \in [c, d]$.

1.3   The Gamma and Beta Functions

1.3.1   Definition of the Gamma Function

For $t > 0$ we define the gamma function Γ as:

$$\Gamma(t) := \int_0^\infty x^{t-1} e^{-x} \, \mathrm{d}x$$

The gamma function is locally uniformly convergent for all x in the interval $(0, \infty)$.

1.3.2   Properties of the Gamma Function

1. The gamma function is a smooth function, meaning all of its derivatives with respect to t are continuous functions i.e. $\Gamma \in C^\infty$

2. $\Gamma(1) = 1 \qquad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

3.
$$\Gamma(t+1) = \int_0^\infty x^t e^{-x}\,\mathrm{d}x = -x^t e^{-x}\Big|_{x=0}^{x=\infty} + t\int_0^\infty x^{t-1}e^{-x}\,\mathrm{d}x = t\,\Gamma(t)$$

4. It follows from $\Gamma(t+1) = t\,\Gamma(t)$ that $\Gamma(1) = 1$ that:

$$\Gamma(2) = 1 \cdot \Gamma(1) = 1$$
$$\Gamma(3) = 2 \cdot \Gamma(2) = 2$$
$$\Gamma(4) = 3 \cdot \Gamma(3) = 6$$
$$\vdots$$
$$\Gamma(n) = (n-1)! \quad \forall \quad n \in \mathbb{N}^+$$

1.3.3   Definition of the Beta Function

For real numbers $p, q > 0$ we define the beta function B as

$$B(p,q) := \int_0^1 x^{p-1}(1-x)^{q-1}\,\mathrm{d}x$$

1.3.4   Proposition: Trigonometric Definition of the Beta Function

For all real numbers $\alpha, \beta > 0$, we can alternatively define the beta function with the relationship:

$$\frac{1}{2}B\left(\frac{\alpha}{2}, \frac{\beta}{2}\right) = \int_0^{\pi/2} (\sin x)^{\alpha-1} (\cos x)^{\beta-1}\,\mathrm{d}x$$

or equivalently

$$B(p,q) = 2\int_0^{\pi/2} (\sin x)^{2p-1} (\cos x)^{2q-1}\,\mathrm{d}x$$

1.3.5   Proposition: Identity for the Beta Function

$$B(p,q) = \int_0^\infty \frac{u^{p-1}}{(1+u)^{p+q}}\,\mathrm{d}y$$

Proof: By applying the change of variables

$$x = \frac{u}{1+u} \qquad 1-x = \frac{1}{1+u} \qquad \mathrm{d}x = \frac{\mathrm{d}u}{1+u}$$
$$u(x=0) = 0 \qquad u(x=1) = \infty$$

to the definition of the beta function we obtain

$$B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}\,\mathrm{d}x = \int_0^\infty \frac{u^{p-1}}{(1+u)^{p-1}}\frac{1}{(1+u)^{q-1}}\frac{1}{(1+u)^2}$$

$$= \int_0^\infty \frac{u^{p-1}}{(1+u)^{(p-1)+(q-1)+2}} = \int_0^\infty \frac{u^{p-1}}{(1+u)^{p+q}}\,\mathrm{d}y$$

1.3.6   Theorem: Relationship Between the Gamma and Beta Function

The gamma and beta functions are related by the relationship:

$$B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

1.3.7   Consequence: The Identity $\Gamma(1/2) = \sqrt{\pi}$

Applying the above theorem with $p = q = 1/2$ and referring to the trigonometric definition of the beta function we obtain

$$\frac{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2})}{\Gamma(1)} = B\left(\frac{1}{2}, \frac{1}{2}\right) = 2\int_0^{\pi/2} (\sin x)^{(2\frac{1}{2}-1)}(\cos x)^{(2\frac{1}{2}-1)} \, \mathrm{d}x$$

$$= 2\int_0^{\pi/2} \mathrm{d}x = 2\left(\frac{\pi}{2}\right) = \pi$$

It follows that

$$\frac{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2})}{\Gamma(1)} = \left(\Gamma\left(\frac{1}{2}\right)\right)^2 = \pi$$

$$\Gamma(1/2) = \sqrt{\pi}$$

Without proof, a more general identity is

$$B(p, 1-p) = \Gamma(p)\Gamma(1-p) = \frac{\pi}{\sin(p\pi)}$$

1.3.8   Theorem: Stirling's Formula

Stirling's formula gives a good approximation for the factorial function and provides a measure of the approximation's accuracy. Additionally, Stirling's formula generalizes the factorial function from the positive integers to the positive real numbers.

$$\lim_{x\to\infty} \frac{x!}{\left(\frac{x}{e}\right)^x \sqrt{x}} = \sqrt{2\pi}$$

so

$$x! = \Gamma(x+1) \sim \left(\frac{x}{e}\right)^x \sqrt{2\pi x}$$

2 Multiple Riemann Integration

2.1 Introduction to Riemann Integrability

2.1.1 Definition: Partition of a Rectangle P

Let P be a rectangle defined by

$$P := [a, b] \times [c, d] \in \mathbb{R}^2$$

and $f : P \to [0, \infty)$ be a bounded, continuous function.

To form a partition of P, we split the interval $[a, b]$ into n parts given by $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$ and split the interval $[c, d]$ into m parts given by $c = y_0 < y_1 < \cdots < y_{m-1} < y_m = d$. Next, we create an arbitrary rectangle P_{jk}, which is a sub-rectangle of P, defined by

$$P_{jk} := [x_{j-1}, x_j] \times [y_{k-1}, y_k]$$

In this case, $D = \{P_{jk}; j \in \{1, \ldots, n\} \, k \in \{1, \ldots, m\}\}$, the set of all sub-rectangles P_{jk}, is a partition of P.

2.1.2 Definition: Upper and Lower Sums

Let P be a rectangle defined by

$$P := [a, b] \times [c, d] \in \mathbb{R}^2$$

and $f : P \to [0, \infty)$ be a bounded function and let D be a partition of P.

Let $m_{jk} = \inf f(x, y)$, $(x, y) \in P_{jk}$, written shorthand as $\inf_{P_{jk}} f$, be the infimum of f on the rectangle P_{jk} and let $M_{jk} = \sup f(x, y)$, $(x, y) \in P_{jk}$, written shorthand as $\sup_{P_{jk}} f$, be the supremum of f on the rectangle P_{jk}. In this case the lower and upper sums of f on the partition D, denoted by $s(f, D)$ and $S(f, d)$ respectively, are:

$$s(f, D) := \sum_{j,k} m_{jk} A\left(P_{jk}\right)$$

$$S(f, D) := \sum_{j,k} M_{jk} A\left(P_{jk}\right)$$

where $A\left(P_{jk}\right)$ denotes the area of the rectangle P_{jk}.

2.1.3 Definition: Riemann Sum

Let P be a rectangle defined by $P := [a, b] \times [c, d] \in \mathbb{R}^2$, let $f : P \to [0, \infty)$ be a bounded function, and let D be an arbitrary partition of P.

Next, for each sub-rectangle P_{jk} we choose a point $\xi_{jk} \in P_{jk}$ inside the rectangle, and let $\xi := \{\xi_{jk}; j, k\}$ be the set of all such points ξ_{jk}. In this case, the Riemann sum of f for the partition D and set of points ξ, denoted by $R(f, D, \xi)$, is defined as:

$$R(f, D, \xi) := \sum_{j,k} f(\xi_{jk}) A\left(P_{jk}\right)$$

where $A\left(P_{jk}\right)$ denotes the area of the rectangle P_{jk}.

2.1.4 Proposition: Effect of Refinement on Upper and Lower Sums

Let P be a rectangle defined by $P := [a, b] \times [c, d] \in \mathbb{R}^2$, let $f : P \to [0, \infty)$ be a bounded function, and let D and D' be partitions of P such that D' is finer than D. In this case,

$$s(f, D) \leq s(f, D') \leq S(f, D') \leq S(f, D)$$

In other words, refinement of a partition causes the upper and lower sums S and s to become closer together in value.

2.1.5 Consequence: Relationship Between Upper and Lower Sums

Let P be a rectangle defined by $P := [a, b] \times [c, d] \in \mathbb{R}^2$, let $f : P \to [0, \infty)$ be a bounded function, and let D_1 and D_2 be arbitrary partitions of P. In this case,

$$s(f, D_1) \leq S(f, D_2)$$

In other words, every lower sum s is less than or equal to every upper sum S, regardless of partition. Proof: Let D' be a finer partition than both D_1 and D_2. Applying the above proposition, we see that:

$$s(f, D_1) \leq s(f, D') \leq S(f, D') \leq S(f, D_2)$$

so $s(f, D_1) \leq S(f, D_2)$.

2.1.6 Definition: Riemann Integrability

Let P be a rectangle defined by $P := [a, b] \times [c, d] \in \mathbb{R}^2$, let $f : P \to [0, \infty)$ be a bounded function, and let D be an arbitrary partition of P.

More so, let $s(f) := \sup_D s(f, D)$ and $S(f) := \inf_D S(f, D)$ be the largest lower sum and smallest upper sum, respectively, of f over all possible partitions D. In this case, f is Riemann-integrable if

$$s(f) = S(f) \quad \text{(condition for Riemann integrability)}$$

Epsilon-delta definition: f is Riemann-integrable on the rectangle P if for each $\epsilon > 0$ there exists $\delta > 0$ such that for each partition D of P where the area of each partition $A(D_{jk}) < \delta$ for each j, k it follows that for each point $\xi_{jk} \in D_{jk}$ forming the set $\xi = \{\xi_{jk}; j, k\}$ the Riemann sum

$$|R(f, D, \xi) - I| < \epsilon$$

2.1.7 Definition: Riemann Sum

Let P be a rectangle defined by $P := [a, b] \times [c, d] \in \mathbb{R}^2$, let $f : P \to [0, \infty)$ be a bounded function, and let D be an arbitrary partition of P.

If f is Riemann-integrable by the previous definition, the $s(f) = S(f)$ are f's Riemann sum on the partition D. It follows that Riemann integrability on a partition implies the existence of the Riemann sum on that partition.

If f is Riemann-integrable on the partition D of the rectangle f, we write the Riemann integral as

$$\iint_P f(x, y) \, dx \, dy \quad \text{or more concisely} \quad \iint_P f \, dS$$

Note: For the remainder of this chapter, integrability is assumed to mean Riemann-integrability.

2.1.8 Lemma: Condition for Integrability

Let P be a rectangle defined by $P := [a, b] \times [c, d] \in \mathbb{R}^2$, let $f : P \to [0, \infty)$ be a bounded function, and let D be an arbitrary partition of P. In this case, f is integrable on P if for each $\epsilon > 0$ there exists partition D of P such that

$$S(f, D) - s(f, D) < \epsilon$$

2.1.9 Theorem: Continuity Implies Integrability

Let P be a rectangle defined by $P := [a, b] \times [c, d] \in \mathbb{R}^2$ and let $f : P \to [0, \infty)$ be a bounded function. In this case, if f is continuous, f is also integrable.

2.1.10 Theorem: Properties of Integrable Functions

Let P be a rectangle defined by $P := [a, b] \times [c, d] \in \mathbb{R}^2$ and let $f, g : P \to [0, \infty)$ be integrable functions. In this case:

1. The function $f + g$ is integrable.

2. The function cf is integrable for each $c \in \mathbb{R}$.

Additionally, if $f \leq g$ for all $x, y \in P$ then

$$\iint_P f(x, y) \, \mathrm{d}x \, \mathrm{d}y \leq \iint_P g(x, y) \, \mathrm{d}x \, \mathrm{d}y$$

Finally, assuming f is integrable, $|f|$ is also integrable and

$$\left| \iint_P f(x, y) \, \mathrm{d}x \, \mathrm{d}y \right| \leq \iint_P |f(x, y)| \, \mathrm{d}x \, \mathrm{d}y$$

2.2 Volume

2.2.1 Definition: Volume for a Function on a Rectangle in $\mathbb{R}$

Let P be a rectangle defined by $P := [a, b] \times [c, d] \in \mathbb{R}^2$ and let $f : P \to [0, \infty)$ be a continuous, bounded function. In this case, the volume V of the three dimensional body bounded by the graph of f, given by the set of points $\{(x, y, z); (x, y) \in P, 0 \leq z \leq f(x, y)\}$ is defined as

$$V = \iint_P f(x, y) \, \mathrm{d}x \, \mathrm{d}y$$

2.2.2 Note: Volume for a Function on a Rectangle in $\mathbb{R}^n$

This proposition generalizes the previous definition to hyper-rectangles of arbitrary dimension. Let $P \subset \mathbb{R}^n$ be a bounded hyper-rectangle in $\mathbb{R}^n$ and let $f : P \to [0, \infty)$ be a continuous, bounded function. In this case, the volume V of the body bounded by the graph of f is defined as

$$V = \overset{n}{\iint \ldots \int_P} f(x_1, x_2, \ldots, x_n) \, \mathrm{d}x_1 \, \mathrm{d}x_2 \ldots \mathrm{d}x_n$$

2.2.3 Proposition: Volume for a Function on an Arbitrary Set

Let $A \subset \mathbb{R}^n$ be an arbitrary bounded subset of $\mathbb{R}^n$ and let $P \subset \mathbb{R}^n$ be an arbitrary bounded hyper-rectangle containing A i.e. $A \subset P$. More so, let $f : A \to [0, \infty)$ be a continuous, bounded function defined on A and let $\tilde{f} : P \to [0, \infty)$ be a generalization of f given by

$$\tilde{f} := \begin{cases} f(x) & x \in A \\ 0 & x \in P \setminus A \end{cases}$$

In this case, assuming the integrals exist,

$$\int_A f(x) \, \mathrm{d}x = \int_P \tilde{f}(x) \, \mathrm{d}x$$

where $x \in \mathbb{R}^n$ is an arbitrary element of $\mathbb{R}^n$.

2.2.4 Definition: n-Dimensional Volume

By definition, the bounded set $A \subset \mathbb{R}^n$ has n-dimensional volume if the function 1 is integrable on A. In this case

$$V(A) := \int_A 1 \, \mathrm{d}x$$

where $x \in \mathbb{R}^n$ is an arbitrary element of $\mathbb{R}^n$.

2.2.5 Theorem: Condition for n-Dimensional Volume

The bounded set $A \subset \mathbb{R}^n$ has n-dimensional volume if the volume of its boundary, denoted by ∂A, is zero. That is, A has n-dimensional volume if $V(\partial A) = 0$.

2.2.6 Note: Condition for Zero n-Dimensional Volume

The bounded set $A \subset \mathbb{R}^n$ has n-dimensional volume equal to zero if for each $\epsilon > 0$

1. There exist hyper-rectangles $P_1, P_2, \ldots, P_n$ such that the union of the hyper-rectangles contains A i.e. $A \subset \bigcup P_j \quad j \in \{1, 2, \ldots, n\}$ and

2. The sum of the volumes of the hyper-rectangles is less than ϵ, written

$$\sum_j V(P_j) < \epsilon$$

2.2.7 Proposition: Finite Union of Sets with Zero Volume

Any bounded set consisting of a finite union of bounded subsets with n-dimensional volume equal to zero has n-dimensional volume equal to zero. More formally, let $m \in \mathbb{Z}$ be a finite integer, let

$$A_1, A_2, \ldots, A_m \subset \mathbb{R}^n \quad | \quad V(A_i) = 0$$

be m bounded sets with n-volume equal to zero and let $A \subset \mathbb{R}^n$ be the bounded set given by

$$A = A_1 \cup A_2 \cup \cdots \cup A_m$$

In this case $V(A) = 0$.

2.2.8 Proposition: The Graph of a Function Has Zero n-Volume

Let $K \subset \mathbb{R}^n$ be a bounded subset of $\mathbb{R}^n$ and let $f : K \to \mathbb{R}$ be an integrable function. In this case, the graph $\mathcal{G}(f)$ of f, given by

$$\mathcal{G}(f) \coloneqq \{(x, f(x)) \in \mathbb{R}^{n+1}; \; x \in K\} \subset \mathbb{R}^n \times \mathbb{R} \equiv \mathbb{R}^{n+1}$$

has zero volume in $\mathbb{R}^{n+1}$.

2.2.9 Integral Over a Set with Zero Volume

Let $K \subset \mathbb{R}^n$ be a bounded hyper-rectangle in $\mathbb{R}^n$, let $A \subset K$ be a subset of K such that A has zero n-volume i.e. $V(A) = 0$, and let $f : K \to \mathbb{R}$ be a bounded function whose restriction to the set $K \setminus A$ is zero everywhere on its domain, i.e.

$$f|_{K \setminus A} = 0$$

In this case, the integral of f over the set K equals zero, i.e.

$$\int_K f(x)\, \mathrm{d}x = 0$$

where $x \in \mathbb{R}^n$ is an element of $\mathbb{R}^n$.

2.2.10 Implication: Integral Over a Set with Zero Volume

Let $K \subset \mathbb{R}^n$ be a bounded hyper-rectangle in $\mathbb{R}^n$, let $f : K \to \mathbb{R}$ be a bounded, integrable function, and let $A \subset K$ be a subset of K with zero n-volume, i.e. $V(A) = 0$. More so, define the function $\tilde{f}$ given by

$$\tilde{f} = \begin{cases} f(x) & x \in K \setminus A \\ \phi(x) & x \in A \end{cases}$$

where $\phi(x) \to \mathbb{R}$ is an arbitrary bounded function that can take on any finite value in $\mathbb{R}$. In this case

$$\int_K f(x)\, \mathrm{d}x = \int_K \tilde{f}(x)\, \mathrm{d}x$$

where $x \in \mathbb{R}^n$ is an element of $\mathbb{R}^n$.

2.3 Zero Measure

2.3.1 Definition: Set with Zero Measure

Let $A \subset \mathbb{R}^n$ be an arbitrary subset of $\mathbb{R}^n$, noting that A is not necessarily bounded. In this case the subset A has measure zero on $\mathbb{R}^n$, denoted by $m(A) = 0$, if for each $\epsilon > 0$ there exist countably many hyper-rectangles A_j such that

$$A \subset A_1 \cup A_2 \cup \cdots \cup A_j$$

and

$$\sum_j V(A_j) < \epsilon$$

.

2.3.2 Note: Zero n-Volume Implies Zero Measure

Let $A \subset \mathbb{R}^n$ be an arbitrary subset of $\mathbb{R}^n$ such that A has zero n-volume. In this case, the measure of A on $\mathbb{R}^n$ is zero. In symbols,

$$V(A) = 0 \implies m(A) = 0$$

2.3.3 Proposition: Union of Sets with Zero Measure

Any set consisting of a countable union of subsets with measure zero on $\mathbb{R}^n$ also has zero measure on $\mathbb{R}^n$. More formally, let $m \in \mathbb{Z} \cup \{\infty\}$ be a countable number, let

$$A_1, A_2, \ldots, A_m \subset \mathbb{R}^n \quad | \quad m(A_i) = 0$$

be m sets with zero measure on $\mathbb{R}^n$ and let $A \subset \mathbb{R}^n$ be the set given by

$$A = A_1 \cup A_2 \cup \cdots \cup A_m$$

In this case $m(A) = 0$.

2.3.4 Theorem: Measure and Integrability

Let $A \subset \mathbb{R}^n$ be a bounded subset of $\mathbb{R}^n$ and define the function $f : A \to \mathbb{R}$. In this case, f is integrable on A if, and only if, the set of all of f's discontinuities has measure zero on $\mathbb{R}^n$.

2.3.5 Note: Almost Everywhere

A statement is said to hold *almost everywhere* (abbreviated *a.e.*) on a set $K \subset \mathbb{R}^n$ if the statement holds everywhere on K except possibly on a subset of K with zero measure.

2.3.6 Proposition: Compact Set and Zero Measure

Let $A \subset \mathbb{R}^n$ be a subset of $\mathbb{R}^n$. If A is compact, then A having zero measure is equivalent to A having zero volume.

$$m(A) = 0 \iff V(A) = 0 \quad (A \text{ is compact})$$

2.4 Integrals Over Bounded Sets

2.4.1 Theorem: Properties of Integrals Over Bounded Sets

Let $A \subset \mathbb{R}^n$ be a bounded subset of $\mathbb{R}^n$, let $f, g : A \to \mathbb{R}$ be two bounded functions, and let $x \in A$ denote an element of $A \subset \mathbb{R}^n$. In this case:

- If f and g are integrable on A, then $f + g$ is also integrable on A and:

$$\int_A (f+g)(x)\,\mathrm{d}x = \int_A f(x)\,\mathrm{d}x + \int_A g(x)\,\mathrm{d}x$$

- If f and g are integrable on A and $f \leq g$ for each $x \in A$ then:

$$\int_A f(x)\,\mathrm{d}x \leq \int_A g(x)\,\mathrm{d}x$$

- If f is integrable on A then $|f|$ is also integrable on A and:

$$\left| \int_A f(x)\,\mathrm{d}x \right| \leq \int_A |f(x)\,\mathrm{d}x|$$

- If f is integrable on A and has A has n-volume in $\mathbb{R}^n$ then:

$$\left| \int_A f(x)\,\mathrm{d}x \right| \leq \sup_{x \in A} |f(x)| V(A)$$

- Let $A, B \subset \mathbb{R}^n$ be two bounded subsets of $\mathbb{R}^n$ such that $V(A \cap B) = 0$ and let the function $f : A \cup B \to \mathbb{R}$ be bounded and integrable on both A and B. In this case, the integral of f over $A \cup B$ also exists and:

$$\int_{A \cup B} f(x)\,\mathrm{d}x = \int_A f(x)\,\mathrm{d}x + \int_B f(x)\,\mathrm{d}x$$

2.4.2 Note: Average Value of a Function on a Set

Let $A \subset \mathbb{R}^n$ be a bounded subset of $\mathbb{R}^n$ such that $V(A) > 0$ and let $f : A \to \mathbb{R}$ be a bounded, integrable function. In this case, the average value $\bar{f}$ of f on A is given by:

$$\bar{f} = \frac{1}{V(A)} \int_A f(x)\,\mathrm{d}x$$

2.4.3 Proposition: Integral Over a Set Equal to Zero

Let $A \subset \mathbb{R}^n$ be a bounded subset of $\mathbb{R}^n$ and let $f : A \to [0, \infty)$ be a function such that

$$\int_A f(x)\,\mathrm{d}x = 0$$

In this case $f(x) = 0$ almost everywhere on A.

2.4.4 Theorem: Fubini's Theorem III

Let $P = [a, b] \times [c, d] \in \mathbb{R}^2$ be a rectangle in $\mathbb{R}^2$, let $f : P \to \mathbb{R}$ be an integrable function, and let $f(x, \bullet) : [c, d] \to \mathbb{R}$ mapping $y \mapsto f(x, y)$ be integrable on $[c, d]$ for each $x \in [a, b]$. In this case:

$$\iint_P f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \int_a^b \int_c^d f(x, y)\,\mathrm{d}y\,\mathrm{d}x$$

2.4.5 Implication: Iterated Integral Over a Rectangle in $\mathbb{R}^2$

Let $P = [a, b] \times [c, d] \in \mathbb{R}^2$ be a rectangle in $\mathbb{R}^2$, let $f : P \to \mathbb{R}$ be an integrable function, and let $f(x, \bullet) : [c, d] \to \mathbb{R}$ mapping $y \mapsto f(x, y)$ be integrable on $[c, d]$ for each $x \in [a, b]$. If f is continuous on P, then:

$$\int_c^d \int_a^b f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \iint_P f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \int_a^b \int_c^d f(x, y)\,\mathrm{d}y\,\mathrm{d}x$$

2.4.6   Proposition: Iterated Integral Over an Arbitrary Set in $\mathbb{R}^2$

Let $\alpha, \beta : [a, b] \to \mathbb{R}$ be two continuous functions such that $\alpha(x) < \beta(x)$ for each $x \in [a, b]$, let $A \subset \mathbb{R}$ be the set given by:

$$A = \{(x, y) \in \mathbb{R}^2 \quad | \quad a \leq x \leq b, \, \alpha(x) \leq y \leq \beta(x)\}$$

and let $f : A \to \mathbb{R}$ be a continuous function. In this case

$$\iint_A f(x, y) \, \mathrm{d}x \, \mathrm{d}y = \int_a^b \int_{\alpha(x)}^{\beta(x)} f(x, y) \, \mathrm{d}y \, \mathrm{d}x$$

2.4.7   Theorem: Iterated Integral Over a Hyper-Rectangle in $\mathbb{R}^3$

Let $K = [a_1, b_1] \times [a_2, b_2] \times [a_1, b_2] \in \mathbb{R}^3$ be a hyper-rectangle in $\mathbb{R}^3$, let $f : K \to \mathbb{R}$ be a continuous function, and let $P = [a_1, b_1] \times [a_2, b_2] \subset \mathbb{R}^2$. In this case:

$$\iiint_K f(x, y, z) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z = \iint_P \left(\int_{a_3}^{b_3} f(x, y, z) \, \mathrm{d}z \right) \mathrm{d}x \, \mathrm{d}y$$

$$= \int_{a_1}^{b_1} \int_{a_2}^{b_2} \int_{a_3}^{b_3} f(x, y, z) \, \mathrm{d}z \, \mathrm{d}y \, \mathrm{d}x$$

2.4.8   Proposition: Iterated Integral Over an Arbitrary Set in $\mathbb{R}^3$

Let $A \in \mathbb{R}^2$ be a set in $\mathbb{R}$ with definite area, let $\alpha, \beta : A \to \mathbb{R}$ be two continuous functions such that $\alpha(x, y) < \beta(x, y)$ for each $(x, y) \in A$, let $B \subset \mathbb{R}^3$ be the set given by:

$$B = \{(x, y, z) \in A \times \mathbb{R} \quad | \quad \alpha(x, y) \leq z \leq \beta(x, y)\}$$

and let $f : B \to \mathbb{R}$ be a continuous function. In this case:

$$\iiint_B f(x, y, z) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z = \iint_A \left(\int_{\alpha(x,y)}^{\beta(x,y)} f(x, y, z) \, \mathrm{d}z \right) \mathrm{d}x \, \mathrm{d}y$$

2.4.9   Example: Iterated Integral in $\mathbb{R}^3$

Let $T : \mathbb{R}^3$ be the 3-dimensional body bounded by the planes $x = 0$, $y = 0$ and $z = 2$, and by the elliptic paraboloid $z = x^2 + y^2$, such that T is located in the first octant i.e. $(x, y, z) \geq 0$. The problem is to calculate the integral

$$I = \iiint_T x \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z$$

We will solve the problem by recursively reducing the degree of the iterated integral. First, to convert the triple integral into a double integral, we define the set $\Omega \in \mathbb{R}^2$ such that

$$\Omega = \mathrm{pr}_{\mathbb{R}^2}^{\perp}(T)$$

is the orthogonal projection of the set T onto the Cartesian plane $\mathbb{R}^2$ (i.e. the xy plane). We see that

$$\Omega = \left\{(x, y) \in \mathbb{R}^2 \, | \, x, y \geq 0 \quad \text{and} \quad x^2 + y^2 \leq 2 \right\}$$

Our original triple integral becomes

$$I = \iiint_T x \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z = \iint_\Omega \left(\int_{z=x^2+y^2}^{2} x \, \mathrm{d}z \right) \mathrm{d}x \, \mathrm{d}y$$

$$= \iint_\Omega \left(2x - x^3 - xy^2 \right) \mathrm{d}x \, \mathrm{d}y$$

Next, we simplify the double integral over Ω and apply the substitution $u = 2 - x^2$ to obtain

$$I = \iint_\Omega \left(2x - x^3 - xy^2 \right) \mathrm{d}x \, \mathrm{d}y = \int \left(\int_{y=0}^{\sqrt{2-x^2}} \left(2x - x^3 - xy^2 \right) \mathrm{d}y \right) \mathrm{d}x$$

$$= \int_0^{\sqrt{2}} \frac{2}{3} (2 - x^2)^{\frac{3}{2}} x \, \mathrm{d}x \overset{(u=2-x^2)}{=} \frac{1}{3} \int_0^2 u^{\frac{3}{2}} \, \mathrm{d}u = \frac{8\sqrt{2}}{15}$$

2.4.10 Definition: Jacobian Matrix and Determinant

Let $A \subset \mathbb{R}^n$ be an open, bounded set with definite volume and let $\boldsymbol{\Phi} : A \to \mathbb{R}^n$ be a vector-valued function of differentiability class C^1 with n scalar components

$$\boldsymbol{\Phi} = (\Phi_1, \ldots, \Phi_n) \quad \Phi_i : A \to \mathbb{R}$$

The function $\boldsymbol{\Phi}$'s Jacobian matrix $\mathbf{J}_{\boldsymbol{\Phi}}(x)$ at the point $x \in A$ is the $n \times n$ matrix:

$$\mathbf{J}_{\boldsymbol{\Phi}}(x) = \begin{bmatrix} \dfrac{\partial \boldsymbol{\Phi}}{\partial x_1} & \cdots & \dfrac{\partial \boldsymbol{\Phi}}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial \Phi_1}{\partial x_1} & \cdots & \dfrac{\partial \Phi_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial \Phi_n}{\partial x_1} & \cdots & \dfrac{\partial \Phi_n}{\partial x_n} \end{bmatrix}$$

The function $\boldsymbol{\Phi}$'s Jacobian determinant $\det(\mathbf{J}_{\boldsymbol{\Phi}}(x))$ at the point $x \in A$ is

$$\det(\mathbf{J}_{\boldsymbol{\Phi}}(x)) = \det \begin{bmatrix} \dfrac{\partial \boldsymbol{\Phi}}{\partial x_1} & \cdots & \dfrac{\partial \boldsymbol{\Phi}}{\partial x_n} \end{bmatrix} = \det \begin{bmatrix} \dfrac{\partial \Phi_1}{\partial x_1} & \cdots & \dfrac{\partial \Phi_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial \Phi_n}{\partial x_1} & \cdots & \dfrac{\partial \Phi_n}{\partial x_n} \end{bmatrix}$$

2.4.11 Theorem: Change of Variables for Multiple Integration

Let $A \subset \mathbb{R}^n$ be an open, bounded set with definite volume and let $\boldsymbol{\Phi} : A \to \mathbb{R}^n$ be an injective function of differentiability class C^1 such that $\boldsymbol{\Phi}$ has a non-zero Jacobian determinant $\det(\mathbf{J}_{\boldsymbol{\Phi}}(x))$ for all $x \in A$ and the function $\mathbf{J}_{\boldsymbol{\Phi}} : x \mapsto \det(\mathbf{J}_{\boldsymbol{\Phi}}(x))$ mapping $x \in A$ to the Jacobian determinant $\det(\mathbf{J}_{\boldsymbol{\Phi}}(x))$ is bounded for all $x \in A$. If the set $\boldsymbol{\Phi}(A) \subset \mathbb{R}^n$ is open and has definite n-volume, then the function $f : \boldsymbol{\Phi}(A) \to \mathbb{R}$ is integrable, the function with the mapping $x \mapsto f(\boldsymbol{\Phi}(x)) \left| \det(\mathbf{J}_{\boldsymbol{\Phi}}(x)) \right|$ is integrable, and:

$$\int_{\boldsymbol{\Phi}(A)} f(x) \, \mathrm{d}x = \int_A f(\boldsymbol{\Phi}(t)) \left| \det(\mathbf{J}_{\boldsymbol{\Phi}}(t)) \right| \mathrm{d}t$$

2.5 Integration in Polar, Cylindrical, and Spherical Coordinates

2.5.1 Note: Change of Variables to Polar Coordinates

Let $\boldsymbol{\Phi} : \mathbb{R}^+ \times [0, 2\pi) \to \mathbb{R}^2$ be the function transforming between polar coordinates (r, ϕ) and Cartesian coordinates (x, y) with the scalar components:

$$x(r, \phi) = r \cos \phi$$
$$y(r, \phi) = r \sin \phi$$

The corresponding Jacobian matrix $\mathbf{J_\Phi}$ for transformation from polar to Cartesian coordinates is

$$\mathbf{J_\Phi}(r, \phi) = \begin{bmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \phi} \\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \phi} \end{bmatrix} = \begin{bmatrix} \cos \phi & -r \sin \phi \\ \sin \phi & r \cos \phi \end{bmatrix}$$

The Jacobian determinant $\det \mathbf{J_\Phi}$ is

$$\det \mathbf{J_\Phi} = r \cos^2 \phi + r \sin^2 \phi = r(\cos^2 \phi + \sin^2 \phi) = r$$

If $A \subset \mathbb{R}^2$ is an open, bounded set with definite volume and $f : \mathbb{R}^2 \to \mathbb{R}$ is an integrable function, we convert between integration in Cartesian and polar coordinates using the formula:

$$\iint_{\boldsymbol{\Phi}(A)} f(x, y) \, \mathrm{d}x \, \mathrm{d}y = \iint_A f(r \cos \phi, r \sin \phi) \, r \, \mathrm{d}r \, \mathrm{d}\phi$$

2.5.2 Note: Change of Variables to Cylindrical Coordinates

Let $\boldsymbol{\Phi} : \mathbb{R}^+ \times [0, 2\pi) \times \mathbb{R} \to \mathbb{R}^3$ be the function transforming between cylindrical coordinates (r, ϕ, z) and Cartesian coordinates (x, y, z) with the scalar components:

$$x(r, \phi, z) = r \cos \phi$$
$$y(r, \phi, z) = r \sin \phi$$
$$z(r, \phi, z) = z$$

The corresponding Jacobian matrix $\mathbf{J_\Phi}$ for transformation from cylindrical to Cartesian coordinates is:

$$\mathbf{J}_\phi = \begin{bmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \phi} & \dfrac{\partial x}{\partial z} \\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \phi} & \dfrac{\partial y}{\partial z} \\ \dfrac{\partial z}{\partial r} & \dfrac{\partial y}{\partial \phi} & \dfrac{\partial z}{\partial z} \end{bmatrix} = \begin{bmatrix} \cos \phi & -r \sin \phi & 0 \\ \sin \phi & r \cos & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The Jacobian determinant $\det \mathbf{J_\Phi}$ is:

$$\det \mathbf{J_\Phi} = r(\cos^2 \phi + \sin^2 \phi) = r$$

If $A \subset \mathbb{R}^3$ is an open, bounded set with definite volume and $f : \mathbb{R}^3 \to \mathbb{R}$ is an integrable function, we convert between integration in Cartesian and cylindrical coordinates using the formula:

$$\iiint_{\boldsymbol{\Phi}(A)} f(x, y, z) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z = \iiint_A f(r \cos \phi, r \sin \phi, z) \, r \, \mathrm{d}r \, \mathrm{d}\phi \, \mathrm{d}z$$

2.5.3   Note: Change of Variables to Spherical Coordinates

Let $\mathbf{\Phi} : \mathbb{R}^+ \times [0, \pi] \times [0, 2\pi) \to \mathbb{R}^3$ be the function transforming between polar coordinates (r, θ, ϕ) (where θ is the azimuthal angle and ϕ is the polar angle) and Cartesian coordinates (x, y, z) with the scalar components:

$$x(r, \theta, \phi) = r \sin \theta \cos \phi$$
$$y(r, \theta, \phi) = r \sin \theta \sin \phi$$
$$z(r, \theta, \phi) = r \cos \theta$$

The corresponding Jacobian matrix $\mathbf{J_\Phi}$ for transformation from spherical to Cartesian coordinates is:

$$\mathbf{J}_\phi = \begin{bmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \theta} & \dfrac{\partial x}{\partial \phi} \\[2mm] \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \theta} & \dfrac{\partial y}{\partial \phi} \\[2mm] \dfrac{\partial z}{\partial r} & \dfrac{\partial z}{\partial \theta} & \dfrac{\partial z}{\partial \phi} \end{bmatrix} = \begin{bmatrix} \sin \theta \cos \varphi & r \cos \theta \cos \varphi & -r \sin \theta \sin \varphi \\ \sin \theta \sin \varphi & r \cos \theta \sin \varphi & r \sin \theta \cos \varphi \\ \cos \theta & -r \sin \theta & 0 \end{bmatrix}$$

The Jacobian determinant $\det \mathbf{J_\Phi}$ is:

$$\det \mathbf{J_\Phi} = r^2 (\cos^2 \phi \cos^2 \theta \sin \theta + \sin^3 \theta \sin^2 \phi + \sin^2 \phi \cos^2 \theta \sin \theta + \sin^3 \theta \cos^2 \phi)$$

$$= r^2 \left(\cos^2 \theta \sin \theta (\cos^2 \phi + \sin^2 \phi) + \sin^3 \theta (\sin^\phi + \cos^2 \phi) \right)$$

$$= r^2 \sin \theta (\sin^2 \theta + \cos^2 \theta) = r^2 \sin \theta$$

If $A \subset \mathbb{R}^3$ is an open, bounded set with definite volume and $f : \mathbb{R}^3 \to \mathbb{R}$ is an integrable function, we convert between integration in Cartesian and spherical coordinates using the formula:

$$\iiint_{\mathbf{\Phi}(A)} f(x, y, z) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z =$$

$$\iiint_A f\left(r \sin \theta \cos \phi, r \sin \theta \sin \phi, r \cos \theta\right) r^2 \sin \theta \, \mathrm{d}r \, \mathrm{d}\theta \, \mathrm{d}\phi$$

2.6   Applications of Multiple Integration

For this subsection, let $\mathrm{d}V = \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z$ be a differential volume element in $\mathbb{R}^3$.

2.6.1   Mass

Let $Q \subset \mathbb{R}^3$ be a continuous 3-dimensional body in Euclidean space with an associated density $\rho(x, y, z)$. The mass M of the body Q is given by

$$M = \iiint_Q \rho(x, y, z) \, \mathrm{d}V$$

2.6.2   Center of Mass

Let $Q \subset \mathbb{R}^3$ be a continuous 3-dimensional body in Euclidean space with an associated density $\rho(\boldsymbol{r})$ and total mass M. The center of mass $\boldsymbol{r}^*$ of the body Q is given by:

$$\boldsymbol{r}^* = \frac{1}{M} \iiint_Q \rho(\boldsymbol{r}) \boldsymbol{r} \, \mathrm{d}V$$

In terms of (x, y, z) components, if $\boldsymbol{r}^* = (x^*, y^*, z^*)$ and density is given by $\rho(x, y, z)$, the center of mass is:

$$x^* = \frac{1}{M} \iiint_Q \rho(x, y, z) \, x \, \mathrm{d}V$$

$$y^* = \frac{1}{M} \iiint_Q \rho(x, y, z) \, y \, \mathrm{d}V$$

$$z^* = \frac{1}{M} \iiint_Q \rho(x, y, z) \, z \, \mathrm{d}V$$

2.6.3   Moment of Inertia

Let $Q \subset \mathbb{R}^3$ be a continuous 3-dimensional body in Euclidean space with an associated density $\rho(\boldsymbol{r})$. The moment of inertia I_a of body Q about a fixed axis of rotation a is given by:

$$I_a = \iiint_Q \rho(\boldsymbol{r}) \|\boldsymbol{d}(\boldsymbol{r})\|^2 \, \mathrm{d}V$$

where

$$\boldsymbol{d}(\boldsymbol{r}) = \mathrm{pr}_a^\perp(\boldsymbol{r})$$

is the orthogonal projection of the point $\boldsymbol{r}$ onto the axis of rotation a and $\|\boldsymbol{d}\|$ is the perpendicular distance from the axis of rotation to the point $\boldsymbol{r}$.

2.7   Improper Integrals in $\mathbb{R}^n$

The concern of this sub-section is to find a way to generalize the notion of improper integrals to integrals over arbitrary subsets of $\mathbb{R}^n$.

2.7.1   Review: Improper Integrals in $\mathbb{R}$

We consider two cases: an integral of an unbounded function on a bounded interval and the integral of a bounded function on an unbounded interval. The notions of the improper integral in these two cases are defined below.

1. Integral of an unbounded function on a bounded interval. Let $a, b \in \mathbb{R}$ be two constants such that $a < b$ and let $g : (a, b] \to \mathbb{R}$ be unbounded at $x = a$, i.e.

$$\lim_{x \to a} g(x) = \infty$$

   In this case, we define the improper integral:

$$\int_a^b g(x) \, \mathrm{d}x = \lim_{\epsilon \to a} \int_\epsilon^b g(x) \, \mathrm{d}x$$

   provided the limit exists.

2. Integral of a bounded function on an unbounded interval. Let $a \in \mathbb{R}$ be a constant and let $g : [a, \infty) \to \mathbb{R}$ be a bounded function. In this case, we define the improper integral:

$$\int_a^\infty g(x)\,\mathrm{d}x = \lim_{M \to \infty} \int_a^M g(x)\,\mathrm{d}x$$

provided the limit exists.

2.7.2   Lemma: Improper Integral of a Bounded, Non-negative Function on an Unbounded Set

Let $D \subset \mathbb{R}^n$ be an unbounded subset of $\mathbb{R}^n$ and let $f : D \to \mathbb{R}^+$ be a bounded, non-negative function, i.e. $f(x) \geq 0$ for all $x \in D$. To define the improper integral of f on D, define the sequence of bounded sets D_j such that

$$D_j := D \cap [-j, j]^n \quad j \in \mathbb{N}$$

where $[-j, j]^n$ is a hyper-rectangle in $\mathbb{R}^n$. More so, assume the sequence of integrals I_j given by:

$$I_j := \int_{D_j} f(x)\,\mathrm{d}x$$

exists, noting that because $f \geq 0$, $I_1 \leq I_2 \leq I_3 \leq \ldots$ and I_j is an increasing sequence. In this case, we define the improper integral of f over D using the sequence I_j with the relation:

$$\int_D f(x)\,\mathrm{d}x := \lim_{j \to \infty} I_j$$

provided the limit exists.

2.7.3   Theorem: Improper Integral of a Bounded Function on an Unbounded Set

This theorem generalizes the previous lemma, which considered only non-negative functions. Let $D \subset \mathbb{R}^n$ be an unbounded subset of $\mathbb{R}^n$ and let $f : D \to \mathbb{R}$ be a bounded function. To make use of the previous lemma, we split f into two non-negative functions $f^+, f^- : D \to \mathbb{R}^+$ such that:

$$f^+ := \max(f, 0)$$
$$f^- := \max(-f, 0) = (-f)^+$$
$$f = f^+ - f^-$$

Next, we apply the previous lemma to f^+ and f^- separately. The definition for the improper integral of f over D that naturally follows is:

$$\int_D f(x)\,\mathrm{d}x := \int_D f^+(x)\,\mathrm{d}x - \int_D f^-(x)\,\mathrm{d}x$$

provided the integrals of both f^+ and f^- over D exist.

2.7.4   Note: Absolute Integrability On an Unbounded Set

Let $D \subset \mathbb{R}^n$ be an unbounded subset of $\mathbb{R}^n$ and define the function $f : D \to \mathbb{R}$ such that the improper integral of f over D

$$\int_D f(x)\,\mathrm{d}x = \int_D f^+(x)\,\mathrm{d}x - \int_D f^-(x)\,\mathrm{d}x$$

exists. The functions f^+ and f^- are defined above. In this case, the expression

$$\int_D f^+(x)\,\mathrm{d}x + \int_D f^-(x)\,\mathrm{d}x \in \mathbb{R}$$

also exists and is finite. More so,

$$\int_D f^+(x)\,\mathrm{d}x + \int_D f^-(x)\,\mathrm{d}x = \int_D |f(x)|\,\mathrm{d}x \in \mathbb{R}$$

In this case, we say that f is absolutely integrable on D.

2.7.5   Example: An Identity for $\sqrt{\pi}$

Let I be the integral given by

$$I = \iint_{\mathbb{R}^2} e^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y$$

though a change of from Cartesian to polar coordinates, the integral becomes

$$I = \int_0^{2\pi} \int_0^\infty e^{-r^2} r\,\mathrm{d}r\,\mathrm{d}\phi$$

$$= 2\pi \int_0^\infty e^{-r^2} r\,\mathrm{d}r$$

through a change of variables $u = r^2$ we obtain

$$I = \pi \int_0^\infty e^{-u}\,\mathrm{d}u = \pi \left(-e^{-u} \Big|_0^\infty \right) = \pi \tag{2.1}$$

On the other hand, in Cartesian coordinates, by applying the basic property of the exponential function $e^a e^b = e^{a+b}$ we obtain

$$I = \iint_{\mathbb{R}^2} e^{-x^2} e^{-y^2}\,\mathrm{d}x\,\mathrm{d}y = \int_{-\infty}^\infty \left(e^{-y^2} \int_{-\infty}^\infty e^{-x^2}\,\mathrm{d}x \right) \mathrm{d}y$$

$$= \int_{-\infty}^\infty e^{-y^2}\,\mathrm{d}y \int_{-\infty}^\infty e^{-x^2}\,\mathrm{d}x$$

$$= \left(\int_{-\infty}^\infty e^{-t^2}\,\mathrm{d}t \right)^2 \tag{2.2}$$

By equating Equations 2.1 and 2.2 we obtain

$$\left(\int_{-\infty}^\infty e^{-t^2}\,\mathrm{d}t \right)^2 = \pi$$

$$\int_{-\infty}^\infty e^{-t^2}\,\mathrm{d}t = \sqrt{\pi}$$

2.7.6 Lemma: Improper Integral of an Unbounded, Non-negative Function on a Bounded Set

Let $D \subset \mathbb{R}^n$ be a bounded subset of $\mathbb{R}^n$ and let $f : D \to \mathbb{R}^+ \cup \{\infty\}$ be an unbounded, non-negative function, i.e. $f(x) \geq 0$ for all $x \in D$. To define the improper integral of f on D, define the sequence of founded functions f_j such that:

$$f_j(x) = \begin{cases} f(x), & 0 \leq f(x) \leq j \\ j, & f(x) > j \end{cases}$$

noting that

$$f_1 \leq f_2 \leq \cdots \leq \lim_{j \to \infty} f_j = f$$

is an increasing sequence of bounded functions. More so, assume the sequence of integrals I_j given by:

$$I_j := \int_D f_j(x)\, \mathrm{d}x$$

exists, noting that because $f \geq 0$, $I_1 \leq I_2 \leq I_3 \leq \ldots$ and I_j is an increasing sequence. In this case, we define the improper integral of f over D using the sequence I_j with the relation:

$$\int_D f(x)\, \mathrm{d}x := \lim_{j \to \infty} I_j$$

provided the limit exists.

2.7.7 Theorem: Improper Integral of a an Unbounded Function on a Bounded Set

This theorem generalizes the previous lemma, which considered only non-negative functions. Let $D \subset \mathbb{R}^n$ be a bounded subset of $\mathbb{R}^n$ and let $f : D \to \mathbb{R} \cup \{\pm\infty\}$ be an unbounded function. To make use of the previous lemma, we split f into two non-negative functions $f^+, f^- : D \to \mathbb{R}^+$ such that:

$$f^+ := \max(f, 0)$$
$$f^- := \max(-f, 0) = (-f)^+$$
$$f = f^+ - f^-$$

Next, we apply the previous lemma to f^+ and f^- separately. The definition for the improper integral of f over D that naturally follows is:

$$\int_D f(x)\, \mathrm{d}x := \int_D f^+(x)\, \mathrm{d}x - \int_D f^-(x)\, \mathrm{d}x$$

provided the integrals of both f^+ and f^- over D exist.

2.7.8 Note: Absolute Integrability of an Unbounded Function

Let $D \subset \mathbb{R}^n$ be a bounded subset of $\mathbb{R}^n$ and let $f : D \to \mathbb{R} \cup \{\pm\infty\}$ be an unbounded function such that the improper integral of f over D

$$\int_D f(x)\, \mathrm{d}x = \int_D f^+(x)\, \mathrm{d}x - \int_D f^-(x)\, \mathrm{d}x$$

exists. The functions f^+ and f^- are defined above. In this case, the expression

$$\int_D f^+(x)\,\mathrm{d}x + \int_D f^-(x)\,\mathrm{d}x \in \mathbb{R}$$

also exists and is finite. More so,

$$\int_D f^+(x)\,\mathrm{d}x + \int_D f^-(x)\,\mathrm{d}x = \int_D |f(x)|\,\mathrm{d}x \in \mathbb{R}$$

In this case, we say that f is absolutely integrable on D.

3 Curves and Surfaces in Space

3.1 Introduction to Curves in Space

A curve is the image of a continuous function from an interval (e.g. $[a, b] \subset \mathbb{R}$) to a topological space (e.g. $\mathbb{R}^n$). The function that defines the curve is called a parametrization.

3.1.1 Definition: Curve and Parameterization

Let $a, b \in \mathbb{R}$ be two real numbers such that $a < b$, let $I = [a, b]$ be an interval in $\mathbb{R}$, and let $\boldsymbol{r} : I \to \mathbb{R}^n$ be a continuous, vector-valued function mapping values in I to vectors in $\mathbb{R}^n$. The image Γ of $\boldsymbol{r}$ given by $\Gamma = \boldsymbol{r}(I)$ is a curve (or space curve) in $\mathbb{R}^n$. The function $\boldsymbol{r}$ is a parameterization of the curve Γ.

3.1.2 Definition: Smooth and Regular Curve

Let $I \subset \mathbb{R}$ be a non-negative, real interval and let the continuous function $\boldsymbol{r} : I \to \mathbb{R}^n$ be a parameterization of the curve $\Gamma = \boldsymbol{r}(I)$.

1. For the purpose of this text, we say that $\boldsymbol{r}$ is smooth if $\boldsymbol{r}$ is of differentiability class C^1 i.e. if $\boldsymbol{r}$'s first derivative exists and is continuous for all $t \in I$. More generally, the smoothness of $\boldsymbol{r}$ is characterized by its differentiability class C^n. Note the definition of smoothness is not universal; other texts often require $\boldsymbol{r} \in C^n$ or $\boldsymbol{r} \in C^\infty$ for $\boldsymbol{r}$ to be smooth.

2. The function $\boldsymbol{r}$ is regular if, in addition to being smooth, $\dot{\boldsymbol{r}}(t) \neq \boldsymbol{0}$ or equivalently $\|\dot{\boldsymbol{r}}(t)\| \neq 0$ for all $t \in I$.

3.1.3 Length of a Curve

Let $a, b \in \mathbb{R}$ be two real numbers such that $a < b$, let $I = [a, b]$, let $\boldsymbol{r} : I \to \mathbb{R}^n$ be a smooth function and let $\Gamma = \boldsymbol{r}(I)$ be a curve in $\mathbb{R}^n$. In this case, the length ℓ of the curve Γ is given by:

$$\ell(\Gamma) := \int_a^b \|\dot{\boldsymbol{r}}(t)\| \, \mathrm{d}t$$

3.1.4 Theorem: Curve Length is Independent of Parameterization

The length of a curve $\Gamma \subset \mathbb{R}^n$ is independent of the choice of parameterization (assuming the parameterization is regular).

More formally, if $\Gamma \subset \mathbb{R}^n$ is a curve in $\mathbb{R}^n$ and $r : [a, b] \to \Gamma$ and $\rho : [\alpha, \beta] \to \Gamma$ are two regular parameterizations of the same curve Γ, then:

$$\int_a^b \|\dot{r}(t)\| \, \mathrm{d}t = \int_\alpha^\beta \|\dot{\rho}(t)\| \, \mathrm{d}t$$

Proof: Because r and ρ are regular and parameterize the same curve, it follows that $r(a) = \rho(\alpha)$ and $r(b) = \rho(\beta)$. More so, for each $t \in [a, b]$ there exists $\tau \in [\alpha, \beta]$ such that $r(t) = \rho(\tau)$. As a result, there exists a function $\varphi : [\alpha, \beta] \to [a, b]$ such that $\varphi(\tau) = t$ and

$$\rho(\tau) = r(\varphi(\tau))$$

Additionally, since the paths are regular, it follows that

$$\rho'(\tau) = (r(\varphi(\tau)))' = r'(\varphi(\tau))\varphi'(\tau)$$

where we use the chain rule for derivatives to get the second equality. Applying the definition of curve length, we see that

$$\ell(\Gamma) = \int_\alpha^\beta \left\| \rho'(\tau) \right\| \mathrm{d}\tau = \int_\alpha^\beta \left\| r'(\varphi(\tau))\varphi'(\tau) \right\| \mathrm{d}\tau \tag{3.1}$$

Note that φ is a scalar function. More so, because of the regularity of r and ρ it follows that φ is increasing on $[\alpha, \beta]$, so $\varphi'(\tau)$ is positive for all $\tau \in [\alpha, \beta]$. As a result we can more $\varphi'(\tau)$ out of the norm in Equation 3.1 to obtain:

$$\ell(\Gamma) = \int_\alpha^\beta \left\| r'(\varphi(\tau)) \right\| \varphi'(\tau) \, \mathrm{d}\tau \tag{3.2}$$

Next, we make the change of variables $t = \varphi(\tau)$, $\mathrm{d}t = \varphi'(\tau) \, \mathrm{d}\tau$, $\varphi(\alpha) = a$, and $\varphi(\beta) = b$. Plugging this change of variables into Equation 3.2 gives

$$\ell(\Gamma) = \int_\alpha^\beta \left\| r'(\varphi(\tau)) \right\| \varphi'(\tau) \, \mathrm{d}\tau = \int_a^b \left\| r'(t) \right\| \mathrm{d}t \tag{3.3}$$

Comparing Equations 3.1 and 3.3 gives the equality

$$\int_\alpha^\beta \left\| \rho'(\tau) \right\| \mathrm{d}\tau = \ell(\Gamma) = \int_a^b \left\| r'(t) \right\| \mathrm{d}t$$

Since the choice of parameterizations r and ϕ was arbitrary, it follows that the length of a curve is independent of parameterization.

3.1.5   Arc Length Parameterization

Let $a, b \in \mathbb{R}$ be two real numbers such that $a < b$, let $I = [a, b]$, and let the function $\boldsymbol{r} : I \to \mathbb{R}^n$ be a regular parameterization of the curve $\Gamma = \boldsymbol{r}(I)$. The *arc length parameter* $s(t)$ of the curve Γ is defined to be:

$$s(t) := \int_a^t \left\| \boldsymbol{r}'(\tau) \right\| \mathrm{d}\tau \quad \text{for} \quad t \in [a, b]$$

and the parameterization $\tilde{\boldsymbol{r}}(s) = \boldsymbol{r}(t(s))$, where $t(s)$ is the inverse of $s(t)$, is called the *arc length parameterization* of Γ. The arc length parameterization is useful because of the property:

$$\left\| \tilde{\boldsymbol{r}}'(s(t)) \right\| = 1 \quad \forall \quad t \in [a, b]$$

meaning that the arc length parameterization $\tilde{\boldsymbol{r}}$ traverses the curve Γ at unit speed.

3.2   Tangent, Normal, and Binormal Vectors

Notation: For the entirety of this subsection, let $a, b \in \mathbb{R}$ be two real numbers such that $a < b$, let $I = [a, b]$, let the function $\boldsymbol{r} : I \to \mathbb{R}^3$ be a regular parameterization of the curve $\Gamma = \boldsymbol{r}(I)$ in terms of the general parameter t and let $\boldsymbol{r}(s) = \boldsymbol{r}(t(s))$ be the arc length parameterization of Γ, where

$$s(t) = \int_0^t \left\| \boldsymbol{r}'(\tau) \right\| \mathrm{d}\tau$$

is the arc length.

3.2.1 Definition: Unit Tangent Vector

The unit tangent vector (or simply tangent vector) $\mathbf{T}$ to the curve Γ at t is the vector given by:

$$\mathbf{T}(t) = \frac{\dot{\boldsymbol{r}}(t)}{\|\dot{\boldsymbol{r}}(t)\|}$$

As the name implies, the tangent vector points tangent to the curve Γ at t.

3.2.2 Alternate Definition of the Tangent Vector

Because the general parameterization $\boldsymbol{r}(t)$ and the arc length s are related by

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}s} = \frac{\dot{\boldsymbol{r}}}{\|\dot{\boldsymbol{r}}\|}$$

it follows that the tangent vector can be defined in terms of the arc length s as

$$\mathbf{T} = \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}s} \quad \text{or more formally} \quad \mathbf{T}(s) = \frac{\mathrm{d}[\boldsymbol{r}(t(s))]}{\mathrm{d}s}$$

3.2.3 Definition: Unit Normal Vector

The unit normal vector (or simply normal vector) $\mathbf{N}$ to the curve Γ at s is the unit vector given by:

$$\mathbf{N}(s) = \frac{\mathbf{T}'(s)}{\|\mathbf{T}'(s)\|}$$

where s is the arc-length parameter and $\mathbf{T}'(s) = \frac{\mathrm{d}\mathbf{T}}{\mathrm{d}s}$.

3.2.4 Note: Alternate Definition of the Normal Vector

The unit normal vector $\mathbf{N}$ to the curve Γ at time t is also given in terms of the vector product:

$$\mathbf{N}(t) = \frac{\dot{\boldsymbol{r}}(t) \times \big(\dot{\boldsymbol{r}}(t) \times \ddot{\boldsymbol{r}}(t)\big)}{\big\|\dot{\boldsymbol{r}}(t) \times \big(\dot{\boldsymbol{r}}(t) \times \ddot{\boldsymbol{r}}(t)\big)\big\|}$$

Note that this formulation is valid for any parameter t, while the previous formula applied only to the arc length parameter s.

Additionally, we can define the unit normal vector $\mathbf{N}$ in terms of the cross product

$$\mathbf{N}(t) = \mathbf{B}(t) \times \mathbf{T}(t)$$

3.2.5 Definition: Binormal Vector

The binormal vector $\mathbf{B}$ is the unit vector defined as:

$$\mathbf{B} = \mathbf{T} \times \mathbf{N}$$

and the binormal vector to the curve Γ at t is given by

$$\mathbf{B}(t) = \mathbf{T}(t) \times \mathbf{N}(t)$$

3.2.6   Note: Alternate Definition of the Binormal Vector

The binormal vector $\mathbf{B}$ to the curve Γ at time t is also given by the vector product:

$$\mathbf{B}(t) = \frac{\dot{\boldsymbol{r}}(t) \times \ddot{\boldsymbol{r}}(t)}{\|\dot{\boldsymbol{r}}(t) \times \ddot{\boldsymbol{r}}(t)\|}$$

This formulation is valid for any parameter t, not just for the arc length parameter s.

3.2.7   Definition: Normal Plane

Let $a, b \in \mathbb{R}$ be two real numbers such that $a < b$, let $I = [a, b]$, and let the function $\boldsymbol{r} : I \to \mathbb{R}^3$ be a regular parameterization of the curve $\Gamma = \boldsymbol{r}(I)$. In this case, the normal plane to Γ at the point $\boldsymbol{r}(t)$ is the plane through $\boldsymbol{r}(t)$ that is perpendicular to the tangent vector $\mathbf{T}(t)$. Equivalently, the normal plane is the set of all points $\boldsymbol{p} = (x, y, z)$ such that:

$$(\boldsymbol{p} - \boldsymbol{r}(t)) \cdot \mathbf{T}(t) = 0$$

It follows that the normal plane at $\boldsymbol{r}(t)$ is the plane containing both $\mathbf{N}(t)$ and $\mathbf{B}(t)$.

3.2.8   Definition: Osculating Plane

Let $a, b \in \mathbb{R}$ be two real numbers such that $a < b$, let $I = [a, b]$, and let the function $\boldsymbol{r} : I \to \mathbb{R}^3$ be a regular parameterization of the curve $\Gamma = \boldsymbol{r}(I)$. In this case, the osculating plane to Γ at the point $\boldsymbol{r}(t)$ is the plane through $\boldsymbol{r}(t)$ that is perpendicular to the binormal vector $\mathbf{B}(t)$. Equivalently, the osculating plane is the set of all points $\boldsymbol{p} = (x, y, z)$ such that:

$$(\boldsymbol{p} - \boldsymbol{r}(t)) \cdot \mathbf{B}(t) = 0$$

It follows that the osculating plane at $\boldsymbol{r}(t)$ is the plane containing both $\mathbf{T}(t)$ and $\mathbf{N}(t)$.

3.2.9   Frenet-Serret Frame

Let Γ be a curve in $\mathbb{R}^3$ and let $\boldsymbol{r} : [a, b] \to \mathbb{R}^3$ be a regular parameterization of Γ. The triad of vectors $(\mathbf{T}, \mathbf{N}, \mathbf{B})$ (i.e. the unit tangent, normal, and binormal vectors to the curve Γ at a point $\boldsymbol{r}(t)$) are called the *TNB* or *Frenet-Serret frame* of the curve Γ at the point $\boldsymbol{r}(t)$. The TNB frame at any point $\boldsymbol{r}(t)$ forms an orthonormal basis of the Euclidean space $\mathbb{R}^3$.

3.3   Curvature

Notation: For the entirety of this subsection, let $a, b \in \mathbb{R}$ be two real numbers such that $a < b$, let $I = [a, b]$, let the function $\boldsymbol{r} : I \to \mathbb{R}^3$ be a regular parameterization of the curve $\Gamma = \boldsymbol{r}(I)$ in terms of the general parameter t and let $\boldsymbol{r}(s) = \boldsymbol{r}(t(s))$ be the arc length parameterization of Γ, where

$$s(t) = \int_0^t \left\| \boldsymbol{r}'(\tau) \right\| \mathrm{d}\tau$$

is the arc length.

3.3.1   Curvature in Terms of Arc Length Parameter and Acceleration

The curvature of a space curve at a point on the curve equals the *magnitude of the acceleration* of a particle moving along the curve at unit speed at that point. Because the definition requires unit speed, we often describe curvature in terms of the arc length parameter s, which results in traversal of a curve at unit speed.

If Γ is a space curve and $\boldsymbol{r}(s)$ is the arc length parameterization, the acceleration of the curve is given by $\boldsymbol{r}''(s)$, and the curvature at s is given by

$$\kappa(s) = \left\|\boldsymbol{r}''(s)\right\|$$

This expression applies only if the curve Γ is parameterized in terms of the arc length parameter s.

3.3.2 Curvature in Terms Arc Length Parameter and Tangent Vector

For a arc length parameterization, the unit tangent vector $\mathbf{T}(s)$ is given by $\boldsymbol{r}'(s)$, the acceleration $\boldsymbol{r}''(s)$ is given by $\mathbf{T}'(s)$, and the curvature at s in terms of the tangent vector is given by

$$\kappa(s) = \left\|\boldsymbol{r}''(s)\right\| = \left\|\mathbf{T}'(s)\right\|$$

Once again, this expression applies only if the curve Γ is parameterized in terms of the arc length parameter s.

3.3.3 Curvature in Terms of General Parameter

In terms of the general parameter t, the curvature κ of the curve Γ at t is given by

$$\kappa(t) = \frac{\left\|\dot{\boldsymbol{r}}(t) \times \ddot{\boldsymbol{r}}(t)\right\|}{\left\|\dot{\boldsymbol{r}}(t)\right\|^3}$$

where $\dot{\boldsymbol{r}}$ denotes differentiation with respect to the general parameter t. This formulation of curvature applies to any choice of parameter, while the previous one is valid only for the arc length parameter s.

3.3.4 Radius of Curvature

The radius of curvature ρ of the curve Γ at s is given by

$$\rho = \frac{1}{\kappa}$$

where $\kappa(s)$ is the curvature of the curve Γ at s. The radius of curvature is equal to the radius of the osculating circle at s.

3.3.5 Curvature in Terms of General Parameter and Tangent Vector

For a general parameterization $\boldsymbol{r}(t)$ of the curve Γ, we can describe the curvature on a curve at t in terms of the tangent vector $\mathbf{T}$ with the formula:

$$\kappa(t) = \frac{\left\|\dot{\mathbf{T}}(t)\right\|}{\left\|\dot{\boldsymbol{r}}(t)\right\|}$$

This formulation of curvature applies to any choice of parameter t. Incidentally, it follows that

$$\frac{\left\|\dot{\mathbf{T}}(t)\right\|}{\left\|\dot{\boldsymbol{r}}(t)\right\|} = \frac{\left\|\dot{\boldsymbol{r}}(t) \times \ddot{\boldsymbol{r}}(t)\right\|}{\left\|\dot{\boldsymbol{r}}(t)\right\|^3}$$

3.3.6 Relationship Between Curvature, T′, and N

Curvature κ is related to the tangent vector $\mathbf{T}$ and normal vector $\mathbf{N}$ by the relationship:

$$\mathbf{T}' = \kappa \mathbf{N}$$

Where $\mathbf{T}'$ denotes differentiation with respect to the arc length parameter s.

More thoroughly, if $\boldsymbol{r}(s)$ is the arc length parameterization of the curve Γ, then the acceleration $\boldsymbol{r}''(s) = \mathbf{T}'(s)$ at s is related to the normal vector $\mathbf{N}(s)$ at s by the relationship

$$\mathbf{T}'(s) = \kappa(s)\mathbf{N}(s)$$

where $\kappa(s)$ is the curvature of the curve Γ at s.

3.3.7 Note: Geometric Interpretation of Zero Curvature

If the curvature κ equals zero for all points on a curve, then $\mathbf{T}'(s) = 0$ and the tangent to the curve is constant, meaning the curve is a straight line.

$$\kappa = 0 \text{ for all points on } \Gamma \qquad \Longleftrightarrow \qquad \Gamma \text{ is a straight line}$$

3.4 Torsion

The torsion of a curve describes how sharply the curve twists out of the plane of curvature. In this respect, torsion gives a sense of how three-dimensional a plane curve is.

3.4.1 Relationship Between Torsion, B′, and N

Torsion τ is related to the binormal vector $\mathbf{B}$ and normal vector $\mathbf{N}$ by the relationship:

$$\mathbf{B}' = -\tau \mathbf{N}$$

Where $\mathbf{B}'$ denotes differentiation with respect to the arc length parameter s.

More thoroughly, if $\boldsymbol{r}(s)$ is the arc length parameterization of the curve Γ, then the derivative of the binormal vector $\mathbf{B}'(s)$ at s is related to the normal vector $\mathbf{N}(s)$ at s by the relationship

$$\mathbf{B}'(s) = \tau(s)\mathbf{N}(s)$$

where $\tau(s)$ is the torsion of the curve Γ at s.

3.4.2 Note: Geometric Interpretation of Zero Torsion

If the curvature τ equals zero for all points on a curve, then $\mathbf{B}'(s) = 0$ and the binormal to the curve is constant, meaning the curve is two-dimensional and contained in a single plane.

$$\tau = 0 \text{ for all points on } \Gamma \qquad \Longleftrightarrow \qquad \Gamma \text{ is a plane curve}$$

3.4.3 Formula for Torsion in Terms of Arc Length Parameter

If $\boldsymbol{r}(s)$ is the arc length parameterization of a space curve Γ, we describe the torsion τ of Γ at s with the formula:

$$\tau(s) = \frac{[\boldsymbol{r}'(s), \boldsymbol{r}''(s), \boldsymbol{r}'''(s)]}{\|\boldsymbol{r}''(s)\|^2} = \frac{(\boldsymbol{r}'(s) \times \boldsymbol{r}''(s)) \cdot \boldsymbol{r}'''(s)}{\kappa^2}$$

where we have used the definition of curvature $\kappa = \|\boldsymbol{r}''(s)\|$ and $[\boldsymbol{r}'(s), \boldsymbol{r}''(s), \boldsymbol{r}'''(s)]$ denotes the scalar triple product operation. Note that this formula applies only to an arc length parameterization $\boldsymbol{r}(s)$.

3.4.4 Formula for Torsion in Terms of General Parameter

If $\boldsymbol{r}(t)$ is a general parameterization of a space curve Γ, we describe the torsion τ of Γ at t with the formula:

$$\tau(t) = \frac{[\dot{\boldsymbol{r}}(t), \ddot{\boldsymbol{r}}(t), \dddot{\boldsymbol{r}}(t)]}{\|\dot{\boldsymbol{r}}(t) \times \ddot{\boldsymbol{r}}(t)\|^2} = \frac{(\dot{\boldsymbol{r}}(t) \times \ddot{\boldsymbol{r}}(t)) \cdot \dddot{\boldsymbol{r}}(t)}{\|\dot{\boldsymbol{r}}(t) \times \ddot{\boldsymbol{r}}(t)\|^2}$$

where $[\dot{\boldsymbol{r}}(t), \ddot{\boldsymbol{r}}(t), \dddot{\boldsymbol{r}}(t)]$ denotes the scalar triple product operation. Note that this formula holds for any parameter t, while the previous one applies only to the arc length parameter s.

3.5 Frenet-Serret Formulae

The Frenet-Serret formulas succinctly relate the $\mathbf{T}, \mathbf{N}$, and $\mathbf{B}$ vectors to the curvature κ and torsion τ of a curve.

Let $\boldsymbol{r}(s)$ be the arc length parameterization of a space curve Γ. In terms of s, the Frenet-Serret formulas are:

$$\mathbf{T}' = \kappa \mathbf{N}$$
$$\mathbf{N}' = -\kappa \mathbf{T} + \tau \mathbf{B}$$
$$\mathbf{B}' = -\tau \mathbf{N}$$

where $'$ denotes differentiation with respect to the arc length parameter s. In this form, the formulas hold only for the arc length parameterization $\boldsymbol{r}(s)$.

3.5.1 Matrix Form of the Frenet-Serret Formulas

In matrix form, the Frenet-Serret formulas are:

$$\begin{bmatrix} \mathbf{T}' \\ \mathbf{N}' \\ \mathbf{B}' \end{bmatrix} = \begin{bmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{bmatrix} \begin{bmatrix} \mathbf{T} \\ \mathbf{N} \\ \mathbf{B} \end{bmatrix}$$

Once again, $'$ denotes differentiation with respect to the arc length parameter s, and the formulas in this form hold only for the arc length parameterization $\boldsymbol{r}(s)$.

3.5.2 Frenet-Serret Formulas in Terms of General Parameter

In terms of a general parameterization $\boldsymbol{r}(t)$, the Frenet-Serret formulas are:

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} \mathbf{T} \\ \mathbf{N} \\ \mathbf{B} \end{bmatrix} = \begin{bmatrix} \dot{\mathbf{T}} \\ \dot{\mathbf{N}} \\ \dot{\mathbf{B}} \end{bmatrix} = \|\mathbf{r}'(t)\| \begin{bmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{bmatrix} \begin{bmatrix} \mathbf{T} \\ \mathbf{N} \\ \mathbf{B} \end{bmatrix}$$

where the extra factor $\|\mathbf{r}'(t)\|$ results because of the change of variables from s to t.

3.6 Surfaces in Euclidean Space

Throughout this section, unless otherwise stated, the word surface is taken to be a surface in Euclidean space $\mathbb{R}^3$. Surfaces may be given in explicit, implicit, and parametric form, with implicit and parametric form being most popular.

3.6.1 Explicitly Defined Surface

In explicit form, a surface is the graph of a continuous scalar function of two variables defined on a connected subset of $\mathbb{R}^{\not\in}$. More formally, if $D \subset \mathbb{R}^2$ is a connected subset of $\mathbb{R}^2$ and $f : D \to \mathbb{R}$ is a continuous function, then the graph of f

$$\mathcal{G}(f) = \{(x, y, z) \,|\, (x, y) \in D, z = f(x, y)\}$$

is an explicitly defined surface.

3.6.2 Implicitly Defined Surface

An implicit surface is the set of all zeros of a continuously differentiable scalar function of three variables. More formally, let $\Omega \subset \mathbb{R}^3$ be a subset of $\mathbb{R}^3$ and let $f : \Omega \to \mathbb{R}$ be a continuously differentiable function such that $\nabla f \neq 0$ for all $(x, y, z) \in \Omega$ (i.e the gradient of f given by $\nabla f = (f_z, f_y, f_z)$) is nonzero over f's domain. In this case, the set $\mathcal{S}$ of all points satisfying the equation $f(x, y, z) = 0$, given by:

$$\mathcal{S} = \{(x, y, z) \,|\, f(x, y, z) = 0\}$$

is an implicitly defined surface.

3.6.3 Parametric Surface

Let $D \subset \mathbb{R}^2$ be a connected surface in $\mathbb{R}^2$, and let $\boldsymbol{r} : D \to \mathbb{R}^3$ be a continuously differentiable vector function (i.e. $\boldsymbol{r} \in C^1$) with scalar components $\boldsymbol{r}(u, v) = (r_1(u, v), r_2(u, v), r_3(u, v))$. More so, let

$$\boldsymbol{r}_u(u, v) \times \boldsymbol{r}_v(u, v) \neq 0$$

for all points $(u, v) \in D$. This is equivalent to requiring that the Jacobian matrix

$$\mathbf{J}_{\boldsymbol{r}}(u, v) = \begin{bmatrix} \dfrac{\partial r_1}{\partial u} & \dfrac{\partial r_1}{\partial v} \\[2mm] \dfrac{\partial r_2}{\partial u} & \dfrac{\partial r_2}{\partial v} \\[2mm] \dfrac{\partial r_3}{\partial u} & \dfrac{\partial r_3}{\partial v} \end{bmatrix}$$

has rank 2 (maximum rank for a 3×2 matrix) for all $(u, v) \in D$ i.e.

$$\text{rank}\,(\mathbf{J}_{\boldsymbol{r}}(u, v)) = 2 \quad \forall \quad (u, v) \in D$$

In this case, the set

$$\mathcal{S} = \{\boldsymbol{r}(u, v) \,|\, (u, v) \in D\}$$

is a parametrically defined surface.

3.6.4 Normal Vector to a Parametric Surface

The normal vector $\boldsymbol{n}$ to a parametric surface $\mathcal{S}$ parameterized by the vector function $\boldsymbol{r} : \mathbb{R}^2 \to \mathbb{R}^3$ mapping $(u, v) \mapsto \boldsymbol{r}(u, v)$ at the point $\boldsymbol{r}_0 = \boldsymbol{r}(u_0, v_u)$ is

$$\boldsymbol{n}(u_0, v_0) = \boldsymbol{r}_u(u_0, v_0) \times \boldsymbol{r}_v(u_0, v_0)$$

where $\boldsymbol{r}_u$ and $\boldsymbol{r}_v$ are $\boldsymbol{r}$'s partial derivatives with respect to u and v, respectively.

3.6.5 Tangent Plane of a Parametric Surface

The tangent plane $\mathcal{T}$ to a parametric surface $\mathcal{S}$ parameterized by the vector function $\boldsymbol{r} : \mathbb{R}^2 \to \mathbb{R}^3$ mapping $(u, v) \mapsto \boldsymbol{r}(u, v)$ at the point $\boldsymbol{r}_0 = \boldsymbol{r}(u_0, v_u)$ is the set of all points $\boldsymbol{p} = (x, y, z)$ satisfying the equation:

$$(\boldsymbol{p} - \boldsymbol{r}_0) \cdot (\boldsymbol{r}_u \times \boldsymbol{r}_v) = 0$$

where $\boldsymbol{r}_u$ and $\boldsymbol{r}_v$ are $\boldsymbol{r}$'s partial derivatives with respect to u and v, respectively, evaluated at (u_0, v_0). We can write the equation for $\mathcal{T}$ at (u_0, v_0) more thoroughly as:

$$\mathcal{T} = \left\{ (x, y, z) \in \mathbb{R}^3 \,\middle|\, \big((x, y, z) - \boldsymbol{r}(u_0, v_u)\big) \cdot \big(\boldsymbol{r}_u(u_0, v_0) \times \boldsymbol{r}_v(u_0, v_0)\big) = 0 \right\}$$

3.6.6 Tangent Plane of an Implicit Surface

Let $\Omega \subset \mathbb{R}^3$ be a subset of $\mathbb{R}^3$, let $f : \Omega \to \mathbb{R}$ be a continuously differentiable function such that $\nabla f \neq 0$ for all $(x, y, z) \in \Omega$, and let $\mathcal{S}$ be the implicit surface given as the solution set of the equation $f(x, y, z) = 0$, i.e.

$$\mathcal{S} = \left\{ (x, y, z) \in \mathbb{R}^3 \,\middle|\, f(x, y, z) = 0 \right\}$$

In this case, the equation of the tangent plane $\mathcal{T}$ of the surface $\mathcal{S}$ at the point $\boldsymbol{r}_0 = (x_0, y_0, z_0)$ is the set of all points $\boldsymbol{p} = (x, y, z)$ satisfying the equation

$$(\boldsymbol{p} - \boldsymbol{r}_0) \cdot \nabla f(\boldsymbol{r}_0) = 0$$

where $\nabla f(\boldsymbol{r}_0)$ denotes the gradient of f evaluated at $\boldsymbol{r}_0 = (x_0, y_0, z_0)$. We can write the equation for $\mathcal{T}$ at (x_0, y_0, z_0) more thoroughly as:

$$\mathcal{T} = \left\{ (x, y, z) \in \mathbb{R}^3 \,\middle|\, \big((x, y, z) - (x_0, y_0, z_0)\big) \cdot \nabla f(x_0, y_0, z_0)\big) = 0 \right\}$$

3.7 Surfaces of Revolution

Surfaces of revolution are surfaces formed by rotating a curve around an axis of revolution in Euclidean space. This section considers only the simple case of a plane curve in the xz plane rotated about the z-axis.

3.7.1 Matrix for Rotation About the z-Axis

Rotation of a vector in the xz plane about the z-axis through an angle φ in $\mathbb{R}^3$ is a linear transformation described by the matrix R_φ

$$R_\varphi = \begin{bmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

3.7.2 Parametrization of a Surface of Rotation

A surface of rotation formed by rotating a plane curve in the xz plane about the z axis may be parameterized with the parameters by the function $\boldsymbol{r} : \mathbb{R}^2 \to \mathbb{R}^3$ mapping $(t, \phi) \mapsto \boldsymbol{r}(t, \phi)$, where $I \subset \mathbb{R}$ is an interval in $\mathbb{R}$, $t \in I$, $\varphi \in [0, 2\pi)$, and $\boldsymbol{r}$ is given by:

$$\boldsymbol{r}(t, \phi) = R_\varphi \big(x(t), y(t), z(t)\big)^T = \begin{bmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x(t) \\ 0 \\ z(t) \end{bmatrix}$$

where $y(t) \equiv 0$ because the curve is in the xz plane.

3.7.3 Parameterization of Common Surfaces of Revolution

1. A cylinder of radius a whose z coordinate ranges from α to β is described by the parameters z and polar angle φ. The parameterization $\boldsymbol{r}(z, \varphi)$ is:

$$\boldsymbol{r}(z, \varphi) = \begin{bmatrix} a \cos \varphi \\ a \sin \varphi \\ z \end{bmatrix}$$

   where $z \in [\alpha, \beta]$ and $\varphi \in [0, 2\pi)$.

2. A cone of radius R and height h is parameterized by the x coordinate x and the polar angle φ. The parameterization $\boldsymbol{r}(x, \varphi)$ is:

$$\boldsymbol{r}(x, \varphi) = \begin{bmatrix} x \cos \varphi \\ x \sin \varphi \\ h \left(1 - \frac{x}{R}\right) \end{bmatrix}$$

   where $x \in [0, R]$ and $\varphi \in [0, 2\pi)$.

3.8 Area of Surfaces

3.8.1 Definition of Surface Area in Euclidean Space

Let $D \subset \mathbb{R}^2$ be a connected subset of $\mathbb{R}^2$ and let $\boldsymbol{r} : D \to \mathbb{R}^3$ mapping $(u, v) \mapsto \boldsymbol{r}(u, v)$ be a continuously differentiable vector function parametrizing the surface $\mathcal{S}$. In this case, the area A of the surface $\mathcal{S}$ is defined as:

$$A(\mathcal{S}) \coloneqq \iint_D \|\boldsymbol{r}_u \times \boldsymbol{r}_v\| \, \mathrm{d}u \, \mathrm{d}v$$

where $\boldsymbol{r}_u$ and $\boldsymbol{r}_v$ are $\boldsymbol{r}$'s partial derivatives with respect to u and v, respectively.

3.8.2 Alternate Formulation of Area of a Surface

By applying the vector identity

$$\|\boldsymbol{r}_u \times \boldsymbol{r}_v\|^2 = \|\boldsymbol{r}_u\|^2 \|\boldsymbol{r}_v\|^2 - (\boldsymbol{r}_u \cdot \boldsymbol{r}_v)^2$$

to the definition of the area of the surface $\mathcal{S}$, we obtain can obtain the alternate formulation

$$A(\mathcal{S}) = \iint_D \sqrt{\|\boldsymbol{r}_u\|^2 \|\boldsymbol{r}_v\|^2 - (\boldsymbol{r}_u \cdot \boldsymbol{r}_v)^2} \, \mathrm{d}u \, \mathrm{d}v$$

3.8.3 Theorem: Surface Area is Independent of Parameterization

The area of a surface is independent of the surface's parameterization.

More formally, if $\mathcal{S}$ is a surface in Euclidean space, $D, \Delta \subset \mathbb{R}^2$ are connected subsets of $\mathbb{R}^2$, and the continuously differentiable vector functions $\boldsymbol{r} : D \to \mathbb{R}^3$ mapping $(u, v) \mapsto \boldsymbol{r}(u, v)$ and $\boldsymbol{\rho} : \Delta \to \mathbb{R}^3$ mapping $(\mu, \nu) \mapsto \boldsymbol{\rho}(\mu, \nu)$ are two parameterizations of $\mathcal{S}$, then

$$\iint_D \sqrt{\|\boldsymbol{r}_u\|^2 \|\boldsymbol{r}_v\|^2 - (\boldsymbol{r}_u \cdot \boldsymbol{r}_v)^2} \, \mathrm{d}u \, \mathrm{d}v = \iint_\Delta \sqrt{\|\boldsymbol{\rho}_\mu\|^2 \|\boldsymbol{\rho}_\nu\|^2 - (\boldsymbol{\rho}_\mu \cdot \boldsymbol{\rho}_\nu)^2} \, \mathrm{d}\mu \, \mathrm{d}\nu$$

Proof: Let $\mathbf{\Phi} : \Delta \to D$ be a bijective mapping between the sets D and Δ, so that the functions $\boldsymbol{r}$ and $\boldsymbol{\rho}$ are related by the composition $\boldsymbol{\rho} = \boldsymbol{r} \circ \mathbf{\Phi}$, written more formally as

$$\boldsymbol{\rho}(\mu, \nu) = \boldsymbol{r}\big(\mathbf{\Phi}(\mu, \nu)\big)$$

The function $\mathbf{\Phi}$ may be written in terms of the scalar components U and (μ, ν) where:

$$\mathbf{\Phi}(\mu, \nu) = \big(U(\mu, \nu), V(\mu, \nu)\big)$$

In terms of the scalar components U and V, the function $\boldsymbol{\rho}$ becomes

$$\boldsymbol{\rho}(\mu, \nu) = \boldsymbol{r}\big(U(\mu, \nu), V(\mu, \nu)\big)$$

. Using the chain rule for the derivative of a composition, we calculate $\boldsymbol{\rho}$'s partial derivatives $\boldsymbol{\rho}_\mu$ and $\boldsymbol{\rho}_\nu$ to be:

$$\boldsymbol{\rho}_\mu = \boldsymbol{r}_u U_\mu + \boldsymbol{r}_v V_\mu$$
$$\boldsymbol{\rho}_\nu = \boldsymbol{r}_u U_\nu + \boldsymbol{r}_v V_\nu$$

Next, we calculate the vector product

$$\begin{aligned}\boldsymbol{\rho}_\mu \times \boldsymbol{\rho}_\nu &= (\boldsymbol{r}_u U_\mu + \boldsymbol{r}_v V_\mu) \times (\boldsymbol{r}_u U_\nu + \boldsymbol{r}_v V_\nu) \\ &= \boldsymbol{r}_u \times \boldsymbol{r}_v \, (U_\mu V_\nu - V_\mu U_\nu)\end{aligned}$$

Next, taking the norm of both sides gives:

$$\|\boldsymbol{\rho}_\mu \times \boldsymbol{\rho}_\nu\| = \|\boldsymbol{r}_u \times \boldsymbol{r}_v\| \, (U_\mu V_\nu - V_\mu U_\nu)$$

in which we recognize that the term $(U_\mu V_\nu - V_\mu U_\nu)$ to be the determinant of the Jacobian matrix

$$\mathbf{J}_\mathbf{\Phi} = \begin{bmatrix} U_\mu & U_\nu \\ V_\mu & V_\nu \end{bmatrix}$$

encoding the change of coordinates between (u, v) and (μ, ν). From the equality:

$$\|\boldsymbol{\rho}_\mu \times \boldsymbol{\rho}_\nu\| = \|\boldsymbol{r}_u \times \boldsymbol{r}_v\| \, \det \mathbf{J}_\mathbf{\Phi}(\mu, \nu)$$

and Theorem 2.4.11 describing the change of coordinates during multiple integration, it follows that

$$A(\mathcal{S}) = \iint_D \sqrt{\|\boldsymbol{r}_u\|^2 \|\boldsymbol{r}_v\|^2 - (\boldsymbol{r}_u \cdot \boldsymbol{r}_v)^2} \, \mathrm{d}u \, \mathrm{d}v$$

$$= \iint_\Delta \sqrt{\|\boldsymbol{\rho}_\mu\|^2 \|\boldsymbol{\rho}_\nu\|^2 - (\boldsymbol{\rho}_\mu \cdot \boldsymbol{\rho}_\nu)^2} \, \mathrm{d}\mu \, \mathrm{d}\nu = A(\mathcal{S})$$

Since the parameterizations $\boldsymbol{r}$ and $\boldsymbol{\rho}$ were arbitrary, if follows that surface area is independent of parameterization.

4 Vector Analysis

4.1 Line Integrals

A line integral is an integral of a function (or field) evaluated along a curve.

4.1.1 Line Integral of a Scalar Field

Let $I = [a, b] \subset \mathbb{R}$ be an interval and let $\boldsymbol{r} : I \to \mathbb{R}^n$ be a smooth parameterization of the curve Γ. More so, let $f : \mathbb{R}^n \to \mathbb{R}$ be a scalar field. In this case, the line integral of f on Γ is:

$$\int_\Gamma f(\boldsymbol{r}) \, \mathrm{d}s = \int_a^b f(\boldsymbol{r}(t)) \|\boldsymbol{r}'(t)\| \, \mathrm{d}t$$

4.1.2 Proposition: Line Integral of a Scalar Field is Independent of Parameterization

The integral of a scalar field f along a curve is independent of the curve's parameterization.

More formally, let $f : \mathbb{R}^n \to \mathbb{R}$ be a scalar field and let $\Gamma \subset \mathbb{R}^n$ be a space curve. More so, let $I = [a, b]$ and $J = [\alpha, \beta]$ be connected intervals in $\mathbb{R}$, and let the smooth functions $\boldsymbol{r} : I \to \Gamma$ mapping $t \mapsto \boldsymbol{r}(t)$ and $\boldsymbol{\rho} : J \to \Gamma$ mapping $\tau \mapsto \boldsymbol{\rho}(\tau)$ be two arbitrary parameterizations of Γ. In this case:

$$\int_\Gamma f(\boldsymbol{r}) \, \mathrm{d}s = \int_a^b f(\boldsymbol{r}(t)) \|\boldsymbol{r}'(t)\| \, \mathrm{d}t = \int_\alpha^\beta f(\boldsymbol{\rho}(t)) \|\boldsymbol{\rho}'(\tau)\| \, \mathrm{d}\tau = \int_\Gamma f(\boldsymbol{\rho}) \, \mathrm{d}s$$

4.1.3 Line Integral of a Vector Field

Let $I = [a, b] \subset \mathbb{R}$ be an connected interval and let $\boldsymbol{r} : I \to \mathbb{R}^n$ be a smooth parameterization of the curve Γ. More so, let $\boldsymbol{F} : \mathbb{R}^n \to \mathbb{R}^n$ be a vector field. In this case, the line integral of $\boldsymbol{F}$ on Γ is:

$$\int_\Gamma \boldsymbol{F}(\boldsymbol{r}) \cdot \mathrm{d}\boldsymbol{r} = \int_a^b \boldsymbol{F}(\boldsymbol{r}(t)) \cdot \boldsymbol{r}'(t) \, \mathrm{d}t$$

4.1.4 Proposition: Line Integral of a Vector Field is Independent of Parameterization

The integral of a vector field $\boldsymbol{F}$ along a curve is independent of the curve's parameterization.

More formally, let $\boldsymbol{F} : \mathbb{R}^n \to \mathbb{R}^n$ be a vector field and let $\Gamma \subset \mathbb{R}^n$ be a space curve. More so, let $I = [a, b]$ and $J = [\alpha, \beta]$ be connected intervals in $\mathbb{R}$, and let the smooth functions $\boldsymbol{r} : I \to \Gamma$ mapping $t \mapsto \boldsymbol{r}(t)$ and $\boldsymbol{\rho} : J \to \Gamma$ mapping $\tau \mapsto \boldsymbol{\rho}(\tau)$ be two arbitrary parameterizations of Γ. In this case:

$$\int_\Gamma \boldsymbol{F}(\boldsymbol{r}) \cdot \mathrm{d}\boldsymbol{r} = \int_a^b \boldsymbol{F}(\boldsymbol{r}(t)) \cdot \boldsymbol{r}'(t) \, \mathrm{d}t = \int_\alpha^\beta \boldsymbol{F}(\boldsymbol{\rho}(\tau)) \cdot \boldsymbol{\rho}'(\tau) \, \mathrm{d}\tau = \int_\Gamma \boldsymbol{F}(\boldsymbol{\rho}) \cdot \mathrm{d}\boldsymbol{\rho}$$

4.1.5 Proposition: Line Integral of a Vector Field and Orientation

Changing the orientation of a curve over which we evaluate the line integral of a vector field changes the sign of the integral.

More formally, let $\boldsymbol{F} : \mathbb{R}^n \to \mathbb{R}^n$ be a vector field, let $I = [a, b]$ be a connected interval in $\mathbb{R}$, and let the smooth function $\boldsymbol{r} : I \to \mathbb{R}^n$ parameterize the curve $\Gamma \subset \mathbb{R}^n$. Finally, let Γ_- be the same curve Γ, but with opposite orientation. In this case

$$\int_{\Gamma_-} \boldsymbol{F}(\boldsymbol{r}) \cdot \mathrm{d}\boldsymbol{r} = -\int_{\Gamma} \boldsymbol{F}(\boldsymbol{r}) \cdot \mathrm{d}\boldsymbol{r}$$

4.1.6 Proposition: Line Integral of the Sum of Vector Fields

Let $\boldsymbol{F}, \boldsymbol{G} : \mathbb{R}^n \to \mathbb{R}^n$ be two vector fields and let $\Gamma \subset \mathbb{R}^n$ be a curve parameterized by the smooth function $\boldsymbol{r} : \mathbb{R} \to \mathbb{R}^n$. In this case

$$\int_{\Gamma} (\boldsymbol{F} + \boldsymbol{G})(\boldsymbol{r}) \cdot \mathrm{d}\boldsymbol{r} = \int_{\Gamma} \boldsymbol{F}(\boldsymbol{r}) \cdot \mathrm{d}\boldsymbol{r} + \int_{\Gamma} \boldsymbol{G}(\boldsymbol{r}) \cdot \mathrm{d}\boldsymbol{r}$$

4.2 Vector Fields

4.2.1 Definition: Gradient

Let $u : D \subset \mathbb{R}^n \to \mathbb{R}$ be a scalar function of n variables $x_1, x_2, \ldots, x_n$ and let $p \in \mathbb{R}^n$ be a point in $\mathbb{R}$. The gradient of u, denoted by ∇u, is the vector field given by

$$\nabla u(p) = \left(\frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \ldots, \frac{\partial u}{\partial x_n} \right)$$

mapping the point $p \in \mathbb{R}^n$ to the vector $\nabla u(p) \in \mathbb{R}^n$.

4.2.2 Definition: Conservative Vector Field

A conservative vector field is a vector field that is the gradient of a multivariable scalar function (or scalar field). More formally, the vector field $\boldsymbol{F} : D \subset \mathbb{R}^n \to \mathbb{R}^n$ is a conservative field if there exists scalar function $u : D \to \mathbb{R}$ such that

$$\boldsymbol{F} = \nabla u$$

4.2.3 Theorem: The Gradient Theorem for Line Integrals

The gradient theorem generalizes the fundamental theorem of calculus. Let $\boldsymbol{F} : D \subset \mathbb{R}^n \to \mathbb{R}^n$ be a conservative vector field and let $u : D \to \mathbb{R}$ be a continuously differentiable scalar function such that $\boldsymbol{F} = \nabla u$. More so, let $I = [a, b]$ be an interval in $\mathbb{R}$, let the smooth function $\boldsymbol{r} : I \to \mathbb{R}^n$ parameterize the curve Γ, and let $\boldsymbol{\alpha} = \boldsymbol{r}(a)$ and $\boldsymbol{\beta} = \boldsymbol{r}(b)$ be the endpoints of Γ. In this case,

$$\int_{\Gamma} \boldsymbol{F}(\boldsymbol{r}) \cdot \mathrm{d}\boldsymbol{r} = u(\boldsymbol{\beta}) - u(\boldsymbol{\alpha})$$

4.2.4 Proposition: Equivalent Properties of Vector Fields

If $\boldsymbol{F} : D \subset \mathbb{R}^n \to \mathbb{R}^n$ is a vector field, the following properties are equivalent:

1. $\boldsymbol{F}$ is a conservative field.

2. The line integral of $\boldsymbol{F}$ over any closed path is equal to zero.

3. The line integral of $\boldsymbol{F}$ between any two points is independent of the path taken and depends only on the starting and ending point of the curve.

4.3 Surface Integrals

A surface integral is an integral of a function (or field) evaluated on a surface.

4.3.1 Surface Integral of a Scalar Field

Let $D \subset \mathbb{R}^2$ be a connected subset of $\mathbb{R}^2$ and let $\boldsymbol{r} : D \to \mathbb{R}^3$ mapping $(u, v) \mapsto \boldsymbol{r}(u, v)$ be a continuously differentiable vector function parametrizing the surface $\mathcal{S}$. Finally, let $f : \mathbb{R}^3 \to \mathbb{R}$ be a scalar field. In this case, the surface integral of f over the surface $\mathcal{S}$ is given by:

$$\iint_{\mathcal{S}} f \, \mathrm{d}\mathcal{S} = \iint_{D} f(\boldsymbol{r}(u, v)) \|\boldsymbol{r}_u \times \boldsymbol{r}_v\| \, \mathrm{d}u \, \mathrm{d}v$$

where $\boldsymbol{r}_u$ and $\boldsymbol{r}_v$ are $\boldsymbol{r}$'s partial derivatives with respect to u and v, respectively.

4.3.2 Proposition: The Surface Integral of a Scalar Field is Independent of Parametrization

The surface integral of a scalar field is independent of the surface's parametrization.

More formally, if $\mathcal{S}$ is a surface in Euclidean space, $D, \Delta \subset \mathbb{R}^2$ are connected subsets of $\mathbb{R}^2$, and the continuously differentiable vector functions $\boldsymbol{r} : D \to \mathbb{R}^3$ mapping $(u, v) \mapsto \boldsymbol{r}(u, v)$ and $\boldsymbol{\rho} : \Delta \to \mathbb{R}^3$ mapping $(\mu, \nu) \mapsto \boldsymbol{\rho}(\mu, \nu)$ are two parameterizations of $\mathcal{S}$, then

$$\iint_{D} f(\boldsymbol{r}(u, v)) \|\boldsymbol{r}_u \times \boldsymbol{r}_v\| \, \mathrm{d}u \, \mathrm{d}v = \iint_{\Delta} f(\boldsymbol{\rho}(\mu, \nu)) \|\boldsymbol{\rho}_\mu \times \boldsymbol{\rho}_\nu\| \, \mathrm{d}\mu \, \mathrm{d}\nu$$

4.3.3 Definition: Surface Integral of a Vector Field

Let $\mathcal{S} \subset \mathbb{R}^3$ be a surface and let $\boldsymbol{N} : \mathcal{S} \to \mathbb{R}^3$ be the vector function mapping points on $\mathcal{S}$ to the unit normal vector $\mathbf{N}$ at that point. Finally, let $\boldsymbol{F} : \mathcal{S} \to \mathbb{R}^3$ be a vector field. In this case:

$$\iint_{\mathcal{S}} \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{\mathcal{S}} = \iint_{\mathcal{S}} \boldsymbol{F} \cdot \boldsymbol{N} \, \mathrm{d}\mathcal{S}$$

4.3.4 Note: Alternate Formulation of the Surface Integral of a Vector Field

Let $D \subset \mathbb{R}^2$ be a connected subset of $\mathbb{R}^2$, let $\boldsymbol{r} : D \to \mathcal{S} \subset \mathbb{R}^3$ mapping $(u, v) \mapsto \boldsymbol{r}(u, v)$ be a continuously differentiable vector function parametrizing the surface $\mathcal{S}$. Finally, let $\boldsymbol{F} : \mathcal{S} \to \mathbb{R}^3$ be a vector field. In this case:

$$\iint_{\mathcal{S}} \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{\mathcal{S}} = \iint_{D} \boldsymbol{F}(\boldsymbol{r}(u, v)) \cdot (\boldsymbol{r}_u \times \boldsymbol{r}_v) \, \mathrm{d}u \, \mathrm{d}v$$

4.4 Vector Operators

4.4.1 Definition: Curl in Euclidean Space

Let $\boldsymbol{F} : D \subset \mathbb{R}^3 \to \mathbb{R}^3$ be a vector field in Euclidean space with scalar components (F_x, F_y, F_z). The curl of $\boldsymbol{F}$, denoted by $\nabla \times \boldsymbol{F}$, is defined as

$$\nabla \times \boldsymbol{F} = \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \right) \hat{\boldsymbol{\imath}} + \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x} \right) \hat{\boldsymbol{\jmath}} + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) \hat{\boldsymbol{k}}$$

or, in matrix form

$$\nabla \times \boldsymbol{F} = \begin{bmatrix} \dfrac{\partial F_z}{\partial y} - \dfrac{\partial F_y}{\partial z} \\[2ex] \dfrac{\partial F_x}{\partial z} - \dfrac{\partial F_z}{\partial x} \\[2ex] \dfrac{\partial F_y}{\partial x} - \dfrac{\partial F_x}{\partial y} \end{bmatrix}$$

4.4.2   Proposition: Curl of a Conservative Vector Field

The curl of a conservative vector field is zero. More formally, if $\boldsymbol{F} : D \subset \mathbb{R}^3 \to \mathbb{R}^3$ is a conservative vector field and $\boldsymbol{p} = (x, y, z)$ denotes a point in Euclidean space, then

$$(\nabla \times \boldsymbol{F})(\boldsymbol{p}) = \boldsymbol{0} \quad \forall \quad \boldsymbol{p} \in D$$

where $\boldsymbol{0}$ denotes the zero vector.

4.4.3   Definition: Divergence in Euclidean Space

Divergence is an vector operator that operates on a vector field to produce a scalar field. If $\boldsymbol{F} : \Omega \subset \mathbb{R}^3 \to \mathbb{R}^3$ with scalar components $\boldsymbol{F} = (F_x, F_y, F_z)$, the divergence of $\boldsymbol{F}$, denoted by $\nabla \cdot F$, is the scalar field $\nabla \cdot F : \Omega \to \mathbb{R}$ given by:

$$\nabla \cdot F = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}$$

4.4.4   Definition: Laplacian in Euclidean Space

The Laplacian, or Laplace operator, is a differential operator giving the divergence of a scalar field's gradient. More formally, let $f : \Omega \subset \mathbb{R}^3 \to \mathbb{R}$ be a twice differentiable scalar field. The Laplacian of f, denoted by Δf or $\nabla^2 f$, is the scalar field $\Delta F : \Omega \to \mathbb{R}$ given:

$$\Delta f = \nabla \cdot \nabla f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

4.4.5   Note: Harmonic Functions

Scalar functions whose Laplacian is zero are called harmonic functions. More formally, if $f : \Omega \subset \mathbb{R}^3 \to \mathbb{R}$ is a twice continuously differentiable scalar function such that

$$\Delta f = 0$$

then f is a harmonic function.

4.4.6   Proposition: Properties of Vector Operators

Let $f, g : \mathbb{R}^3 \to \mathbb{R}$ be two continuously differentiable scalar fields and let $\boldsymbol{F}, \boldsymbol{G} : \mathbb{R}^3 \to \mathbb{R}^3$ be two vector fields. In this case:

1. $\nabla(fg) = g\nabla f + f\nabla g$

2. $\nabla \cdot (f\boldsymbol{F}) = \langle f, \nabla \boldsymbol{F} \rangle + \langle \nabla f, \boldsymbol{F} \rangle$

3. $\nabla \cdot (\boldsymbol{F} \times \boldsymbol{G}) = \langle \boldsymbol{G}, \nabla \times \boldsymbol{F} \rangle - \langle \boldsymbol{F}, \nabla \times \boldsymbol{G} \rangle$

where $\langle \cdot, \cdot \rangle$ denotes the standard scalar Euclidean product.

4.5 Important Theorems in Vector Analysis

4.5.1 Gauss's Theorem (Divergence Theorem)

Let $V \subset \mathbb{R}^3$ be a closed and bounded (or, more generally, compact) subset of $\mathbb{R}^3$ with a piecewise smooth boundary $S = \partial V$ where S is a piecewise smooth surface with outward orientation. More so, let $\boldsymbol{F}$ be a continuously differentiable vector field defined on a neighborhood of V. In this case:

$$\iint_S \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{S} = \iiint_V \nabla \cdot \boldsymbol{F} \, \mathrm{d}V$$

4.5.2 Green's Theorem

Let $D \subset \mathbb{R}^2$ be a bounded subset of $\mathbb{R}^2$ with a closed, piecewise continuous border $\partial D = C$ such that C is positively oriented. More so, let $\boldsymbol{F}$ be a continuously differentiable vector field defined on a neighborhood of D with scalar components given by $\boldsymbol{F}(x, y) = (X(x, y), Y(x, y))$. In this case:

$$\oint_C X \, \mathrm{d}x + Y \, \mathrm{d}y = \iint_D \left(\frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y} \right)$$

where the direction of integration in the integral $\oint_C$ is in the positive direction of the curve C.

4.5.3 Stoke's Theorem in $\mathbb{R}^3$

Let $\Sigma \subset \mathbb{R}^3$ be a bounded, outward oriented, piecewise smooth surface in $\mathbb{R}^3$ with piecewise continuous border $\partial \Sigma$. More so, let $\boldsymbol{F}$ be a continuously differentiable vector field defined on a neighborhood of Σ. In this case:

$$\int_{\partial \Sigma} \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{s} = \iint_\Sigma \nabla \times \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{S}$$

Where $\mathrm{d}\boldsymbol{s}$ denotes a line element of $\partial \Sigma$ and $\mathrm{d}\boldsymbol{S}$ denotes an area element of Σ.

4.5.4 Green's First Two Identities

Let $\Omega \subset \mathbb{R}^3$ be an open, bounded subset of $\mathbb{R}^3$ with a piecewise continuous border $\partial \Omega$, let u, v be smooth scalar fields defined on a neighborhood of $\bar{\Omega}$, and let $\boldsymbol{n}$ denote the outward-pointing normal vector to border $\partial \Omega$. In this case, Green's first two identities are:

$$\iint_{\partial \Omega} u \frac{\partial v}{\partial \boldsymbol{n}} \, \mathrm{d}S = \iiint_\Omega (u \Delta v + \nabla u \cdot \nabla v) \, \mathrm{d}V$$

$$\iint_{\partial \Omega} \left(u \frac{\partial v}{\partial \boldsymbol{n}} - v \frac{\partial u}{\partial \boldsymbol{n}} \right) \mathrm{d}S = \iiint_\Omega (u \Delta v - v \Delta u) \, \mathrm{d}V$$

where $\frac{\partial v}{\partial \boldsymbol{n}} = \nabla v \cdot \boldsymbol{n}$ denotes the directional derivative of v in the direction of $\boldsymbol{n}$.

4.5.5 Conservative Vector Field on a Star Domain

Let $S \subset \mathbb{R}^3$ be a closed, bounded star domain in $\mathbb{R}^3$, and let $\boldsymbol{F} : S \to \mathbb{R}^3$ be a vector field defined on the set S such that $\nabla \times \boldsymbol{F} = 0$ for all $\boldsymbol{x} \in S$. In this case, $\boldsymbol{F}$ is a conservative vector field and there exists scalar field $f : S \to \mathbb{R}$ such that $\boldsymbol{F} = \nabla f$.

4.6   Vector Operators in Various Coordinate Systems

TODO

5   Differential Equations

5.1   Ordinary First Order Differential Equations

TODO Add remaining types of ordinary first-order differential equations.

5.1.1   Homogeneous Differential Equation

A homogeneous differential equation is a differential equation of the form

$$y' = f(x, y)$$

where f is a homogeneous function of order 0, meaning that $f(tu, tv) = f(u, v)$ for $t > 0$.

To solve a homogeneous DE, we start by applying the 0th order homogeneity.

$$f(x, y) = f\left(x, \frac{y}{x}x\right)$$

We introduce a new variable $z = \frac{y}{x}$ with $y' = z'x + z$ and our DE takes the form

$$z'x + z = f(1, z)$$
$$z'x = f(1, z) - z$$
$$\frac{\mathrm{d}x}{x} = \frac{\mathrm{d}z}{f(1, z) - z}$$
$$\frac{\mathrm{d}x}{x} = \frac{\mathrm{d}z}{g(z)}$$

where $g(z) := f(1, z) - z$. This is an equation with separable variables, which can be solved for z by integrating both sides. Once z is found, we plug in $y = zx$.

5.1.2   Definition: First Integral

A first integral of the first-order differential equation $y' = f(x, y)$ is a function $u = u(a, b)$ such that
$$u(x, y(x)) = C \qquad C \in \mathbb{R}$$

for each solution $y = y(x)$ of the differential equation. Note that the set of all (a, b) such that $u(a, b) = C$ is a level set.

5.1.3   Definition: Exact Differential Equation

The differential equation
$$P(x, y)\,\mathrm{d}x + Q(x, y)\,\mathrm{d}y = 0$$

is exact if any of the equivalent conditions hold:

1. The vector field $\boldsymbol{F} = (P, Q)$ is conservative

2. There exists a scalar function $u(x, y)$ such that $\boldsymbol{F} = (P, Q) = \nabla u$

3. $P_y = Q_x$

5.1.4 Integrating Factor

Lets have a differential equation, not necessarily exact, of the form

$$P(x, y)\, \mathrm{d}x + Q(x, y)\, \mathrm{d}y = 0$$

For the special cases where either $\frac{P_y - Q_x}{Q}$ is independent of y or $\frac{Q_x - P_y}{P}$ is independent of x, we have the integrating factors

$$\mu(x) = \exp\left(\int \frac{P_y - Q_x}{Q}\, \mathrm{d}x\right)$$

$$\mu(y) = \exp\left(\int \frac{Q_x - P_y}{P}\, \mathrm{d}x\right)$$

5.1.5 Parametric Solution Case 1

Our goal is to describe the family of curves solving the differential equation

$$F(x, y') = 0$$

parametrically. We denote $y' = \frac{\mathrm{d}y}{\mathrm{d}x}$, $\dot{y} = \frac{\mathrm{d}y}{\mathrm{d}t}$, and $\dot{x} = \frac{\mathrm{d}x}{\mathrm{d}t}$. Next, we introduce the parametric functions

$$x(t) = \psi(t)$$

$$y'(t) = \frac{\mathrm{d}y}{\mathrm{d}x} = \theta(t)$$

Our goal is to find an expression for $y(t)$. We apply the chain rule to the composition $y \circ x$ to get

$$\frac{\mathrm{d}[y \circ x]}{\mathrm{d}t} = \frac{\mathrm{d}y}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t} \qquad \Longrightarrow \qquad \frac{\mathrm{d}[y \circ x]}{\mathrm{d}t} \equiv \dot{y}(t) = \frac{\mathrm{d}y}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t}$$

It follows that

$$\dot{y}(t) = \theta(t)\dot{\psi}(t) \qquad \Longrightarrow \qquad y(t) = \int \theta(t)\dot{\psi}(t)\, \mathrm{d}t = \int y'(t)\dot{x}(t)\, \mathrm{d}t$$

In this case, the functions $y(t)$ and $x(t)$ parameterize the solution to the differential equation.

5.1.6 Parametric Solution Case II

Our goal is to describe the family of curves solving the differential equation

$$F(y, y') = 0$$

parametrically. We denote $y' = \frac{\mathrm{d}y}{\mathrm{d}x}$, $\dot{y} = \frac{\mathrm{d}y}{\mathrm{d}t}$, and $\dot{x} = \frac{\mathrm{d}x}{\mathrm{d}t}$. Next, we introduce the parametric functions

$$y(t) = \psi(t)$$

$$y'(t) = \frac{\mathrm{d}y}{\mathrm{d}x} = \theta(t)$$

Our goal is to find an expression for $x(t)$. We apply the chain rule to the composition $y \circ x$ to get

$$\frac{\mathrm{d}[y \circ x]}{\mathrm{d}t} = \frac{\mathrm{d}y}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t} \qquad \Longrightarrow \qquad \frac{\mathrm{d}[y \circ x]}{\mathrm{d}t} \equiv \dot{y}(t) = \frac{\mathrm{d}y}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t}$$

It follows that

$$\dot{\psi}(t) = \theta(t)\dot{x}(t) \quad \Longrightarrow \quad x(t) = \int \frac{\dot{\psi}(t)}{\theta(t)}\,\mathrm{d}t = \int \frac{\dot{y}(t)}{y'(t)}\,\mathrm{d}t$$

In this case, functions $y(t)$ and $x(t)$ parameterize the solution to the differential equation.

5.1.7 Clairaut Differential Equation

The Clairaut differential equation has the form

$$y = xy' + f(y')$$

where $f \in C^1$ is a continuously differentiable function.

We solve the equation by differentiating with respect to x to get:

$$y' = y' + xy'' + f'(y')y''$$
$$0 = y''(x + f'(y'))$$

It follows that either $y'' = 0$, which gives the solution

$$y = Cx + f(C)$$

which is a family of linear functions parameterized by the constant C. This solution is called the general solution.

The alternative option

$$x + f'(y') = 0$$

can be written as $F(x, y') = 0$ and solved parametrically using previous techniques. The resulting solution for $y(x)$ is called the singular solution. The graph of the singular solution is the envelope of the graphs of the general solutions.

5.2 Metric Spaces and Picard's Existence Theorem

5.2.1 Review: Metric Spaces

A metric space (M, d) is a set M equipped with a metric $d : M \times M \to \mathbb{R}^+$ that measures the distance elements of M. For each $u, v, w \in M$, the metric d must satisfy:

$$d(u, v) \leq 0 \quad \text{and} \quad d(u, v) = 0 \iff v = u$$
$$d(u, v) = d(v, u)$$
$$d(u, w) \leq d(u, v) + d(v, w)$$

5.2.2 Definition: Lipschitz Continuity

Let (M, d_m) and (N, d_n) be two metric spaces. The function $\phi : M \to N$ mapping from M to N is said to be Lipschitz continuous if there exists a non-negative real constant $\gamma \in \mathbb{R}^+$ such that for each $m_1, m_2 \in M$

$$d_n\big(\phi(m_1), \phi(m_2)\big) \leq \gamma d_m(m_1, m_2)$$

If the special case of a real function $f : \mathbb{R} \to \mathbb{R}$ equipped with the standard metric $d(x_1, x_2) = |x_2 - x_1|$, the function f is Lipschitz continuous if there exists a non-negative real constant $\gamma \in \mathbb{R}^+$ such that for all $x_1, x_2 \in \mathbb{R}$

$$|f(x_1) - f(x_2)| \leq \gamma |x_1 - x_2|$$

5.2.3 Review: Banach's Fixed-Point Theorem

Let (M, d) be a complete metric space and let $\phi : M \to M$ be a contraction mapping (i.e. a Lipschitz continuous function from M to M with Lipschitz constant $\gamma < 1$). In this case, there exists exactly one fixed point $a \in M$ such that $\phi(a) = a$. More so, for the sequence $x_n = (\phi \circ \overset{n}{\ldots} \circ \phi)(x)$ defined as n compositions of the function ϕ, we have:

$$\lim_{n \to \infty} x_n(x) = a$$

for each $x \in M$.

5.2.4 Theorem: Picard's Existence Theorem

Let $x_0, y_0 \in \mathbb{R}$ be real numbers and let $P \coloneqq [x_0 - a, x_0 + a] \times [y_0 - b, y_0 + b] \subset \mathbb{R}^2$ be a rectangle in $\mathbb{R}^2$. Let $f : P \to \mathbb{R}$ be a continuous function such that f is uniformly Lipschitz continuous in the second variable y, meaning there exists Lipschitz constant $\gamma > 0$ such that for each $(x, y_1), (x, y_2) \in P$,

$$|f(x, y_1) - f(x, y_2)| \leq \gamma |y_2 - y_1|$$

In this case, there exists exactly one solution to the initial-value problem

$$y'(x) = f(x, y(x)), \qquad y(x_0) = x_0.$$

5.3 System of Linear Differential Equations

5.3.1 Definition: Homogeneous System of LDEs with Constant Coefficients

Define the function $\boldsymbol{y} : I \to \mathbb{R}^n$ and the constant matrix of coefficients $\mathbf{A} \in \mathbb{R}^{n \times n}$. In this case, the matrix equation

$$\boldsymbol{y}'(x) = \mathbf{A}\boldsymbol{y}(x)$$

is a homogeneous system of linear differential equations with constant coefficients.

5.3.2 Definition: Non-Homogeneous System of LDEs with Constant Coefficients

Define the functions $\boldsymbol{y} : I \to \mathbb{R}^n$ and $\boldsymbol{b} : I \to \mathbb{R}^n$ and the constant matrix of coefficients $\mathbf{A} \in \mathbb{R}^{n \times n}$. In this case, the matrix equation

$$\boldsymbol{y}'(x) = \mathbf{A}\boldsymbol{y}(x) + \boldsymbol{b}(x)$$

is a non-homogeneous system of linear differential equations with constant coefficients.

5.3.3 Solution of a Homogeneous System of LDEs with Constant Coefficients

Define the function $\boldsymbol{y} : I \to \mathbb{R}^n$ and the constant matrix of coefficients $\mathbf{A} \in \mathbb{R}^{n \times n}$ and let

$$\boldsymbol{y}'(x) = \mathbf{A}\boldsymbol{y}(x)$$

be a homogeneous system of linear differential equations with constant coefficients. The solution to this homogeneous system can be written:

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

where $\boldsymbol{c} \in \mathbb{R}^n$ is a constant vector.

5.3.4 Definition: Homogeneous and Non-Homogeneous System of First-Order LDEs with Variable Coefficients

Let $I \subset \mathbb{R}$ be an interval on the real line, and let $a_{ij}, b_j : I \to \mathbb{R}$ be continuous functions with indexes in the range $i, j \in \{1, 2, \ldots, n\}$. Let $\mathbf{A} : I \to \mathbb{R}^{n \times n}$ be a matrix function defined as:

$$\mathbf{A}(x) := \begin{bmatrix} a_{11}(x) & \ldots & a_{1n}(x) \\ \vdots & \ddots & \vdots \\ a_{n1}(x) & \ldots & a_{nn}(x) \end{bmatrix} \in \mathbb{R}^{n \times n}$$

and let $\boldsymbol{b} : I \to \mathbb{R}^n$ and $\boldsymbol{y} : I \to \mathbb{R}^n$ be a vector functions of one variable defined as:

$$\boldsymbol{b}(x) := \begin{bmatrix} b_1(x) \\ \vdots \\ b_n(x) \end{bmatrix} \in \mathbb{R}^n \quad \text{and} \quad \boldsymbol{y}(x) := \begin{bmatrix} y_1(x) \\ \vdots \\ y-_n(x) \end{bmatrix} \in \mathbb{R}^n$$

In this case, the matrix equations

$$\boldsymbol{y}'(x) = \mathbf{A}(x)\boldsymbol{y}(x) \qquad \text{and} \qquad \boldsymbol{y}'(x) = \mathbf{A}(x)\boldsymbol{y}(x) + \boldsymbol{b}(x)$$

are homogeneous and nonhomogeneous systems of first-order linear differential equations with constant coefficients, respectively.

5.3.5 Proposition: Equivalent Properties of Systems of Linear Differential Equations

Let $I \subset \mathbb{R}$ be an interval on the real line and let the vector-valued functions $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k : I \to \mathbb{R}^n$ be solutions of the homogeneous system $\boldsymbol{y}' = A\boldsymbol{y}$ on the interval I. In this case, the following are equivalent:

1. The vector functions $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k$ are linearly independent.

2. There exists $x_0 \in I$ such that the vectors $\boldsymbol{y}_1(x_0), \boldsymbol{y}_2(x_0), \ldots, \boldsymbol{y}_k(x_0)$ are linearly independent in $\mathbb{R}^n$.

3. The vectors $\boldsymbol{y}_1(x), \boldsymbol{y}_2(x), \ldots, \boldsymbol{y}_k(x)$ are linearly independent in $\mathbb{R}^n$ for all $x \in I$.

5.3.6 Theorem: Dimension of Solution Set

Let $I \subset R$ be an interval on the real line, let $\boldsymbol{y} : I \to \mathbb{R}^n$ be a vector valued function with n scalar components $(y_1, y_2, \ldots, y_n)$, and let $\mathbf{A} : I \to \mathbb{R}_{n \times n}$ be a matrix function. In this case, the solution set to the homogeneous system of first-order linear differential equations

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

is n-dimensional.

5.3.7 Definition: Fundamental Matrix

Let $I \subset \mathbb{R}$ be an interval on the real line, let $\boldsymbol{y} : I \to \mathbb{R}^n$ be a vector-valued function and let $\mathbf{A} : I \to \mathbb{R}^{n \times n}$ be a matrix-valued function, and let

$$\boldsymbol{y}'(x) = \mathbf{A}(x)\boldsymbol{y}(x)$$

be a system of n homogeneous linear differential equations. Finally, let $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n$ be n linearly independent solutions to the homogeneous system

$$\boldsymbol{y}'(x) = \mathbf{A}(x)\boldsymbol{y}(x)$$

In this case, the fundamental matrix $\mathbf{Y} : I \to \mathbb{R}^{n \times n}$ is a matrix-valued function whose columns are the linearly independent solutions to the homogeneous system. In other words

$$\mathbf{Y} = [y_1, y_2, \ldots, y_n]$$

5.3.8   Proposition: Solution in Terms of Fundamental Matrix

Let $I \subset \mathbb{R}$ be an interval on the real line, let $\boldsymbol{y} : I \to \mathbb{R}^n$ be a vector-valued function, let $\mathbf{A} : I \to \mathbb{R}^{n \times n}$ be a matrix-valued function, and let

$$\boldsymbol{y}'(x) = \mathbf{A}(x)\boldsymbol{y}(x)$$

be a system of n homogeneous linear differential equations with the fundamental matrix $\mathbf{Y} : I \to \mathbb{R}^{n \times n}$. In this case, every solution to the system can be written in the form

$$\boldsymbol{y}(x) = \mathbf{Y}(x)\boldsymbol{c}$$

where $\boldsymbol{c} \in \mathbb{R}^n$ is a constant vector.

5.3.9   Proposition: Condition for Fundamental Matrix

Let $I \subset \mathbb{R}$ be an interval on the real line, let $\boldsymbol{y} : I \to \mathbb{R}^n$ be a vector-valued function, let $\mathbf{A} : I \to \mathbb{R}^{n \times n}$ be a matrix-valued function, and let

$$\boldsymbol{y}'(x) = \mathbf{A}(x)\boldsymbol{y}(x)$$

be a system of n homogeneous linear differential equations. In this case, the matrix-valued function $\mathbf{Y} : I \to \mathbb{R}^{n \times n}$ is a fundamental matrix of the system if and only if

$$\mathbf{Y}'(x) = \mathbf{A}(x)\mathbf{Y}(x)$$

and $\mathbf{Y}(x)$ is invertible for all x in I.

5.3.10   Proposition: Particular Solution of a Non-Homogeneous System of LDEs with Variable Coefficients

Let $I \subset \mathbb{R}$ be an interval on the real line, let $\boldsymbol{y} : I \to \mathbb{R}^n$ and $\boldsymbol{b} : I \to \mathbb{R}^n$ be a vector-valued functions, let $\mathbf{A} : I \to \mathbb{R}^{n \times n}$ be a matrix-valued function, and let

$$\boldsymbol{y}'(x) = \mathbf{A}(x)\boldsymbol{y}(x) + \boldsymbol{b}(x)$$

be a non-homogeneous system of n linear differential equations with variable coefficients such that the corresponding homogeneous system

$$\boldsymbol{y}'(x) = \mathbf{A}(x)\boldsymbol{y}(x)$$

has the fundamental matrix $\mathbf{Y} : I \to \mathbb{R}^{n \times n}$ and the homogeneous solution $\boldsymbol{y}_h : I \to \mathbb{R}^n$ given by

$$\boldsymbol{y}_h(x) = \mathbf{Y}(x)\boldsymbol{c}$$

where $\boldsymbol{c} \in \mathbb{R}^n$ is a vector of constants. In this case, the particular solution $\boldsymbol{y}_p(x)$ can be written in the form

$$\boldsymbol{y}_p(x) = \mathbf{Y}(x)\boldsymbol{z}(x)$$

where $\mathbf{Y} = \mathbf{Y}(x)$ is the fundamental matrix and $\boldsymbol{z} : I \to \mathbb{R}^n$ is a vector-valued function given by:

$$\boldsymbol{z}(x) = \int \mathbf{Y}^{-1}(x)\boldsymbol{b}(x)\,\mathrm{d}x$$

Short Derivation: Differentiate $\boldsymbol{y}_p(x) = \mathbf{Y}(x)\boldsymbol{z}(x)$ with respect to x to get:

$$\boldsymbol{y}_p' = \mathbf{Y}'\boldsymbol{z} + \mathbf{Y}\boldsymbol{z}' = \mathbf{A}\mathbf{Y}\boldsymbol{z} + \mathbf{Y}\boldsymbol{z}' = \mathbf{A}\boldsymbol{y}_p + \mathbf{Y}\boldsymbol{z}'$$

Comparing this to the general equation $\boldsymbol{y}' = \mathbf{A}\boldsymbol{y} + \boldsymbol{b}$ and in particular comparing the last terms it follows that from the particularity of the solution that:

$$\boldsymbol{b} = \mathbf{Y}\boldsymbol{z}' \qquad \Longrightarrow \qquad \boldsymbol{z}' = \mathbf{Y}^{-1}\boldsymbol{b} \qquad \Longrightarrow \qquad \boldsymbol{z} = \boldsymbol{z}(x) = \int \mathbf{Y}^{-1}(x)\boldsymbol{b}(x)\,\mathrm{d}x$$

5.3.11   Summary: Particular Solution of a Nonhomogeneous System of LDEs with Variable Coefficients

The particular solution $\boldsymbol{y}_p$ to the nonhomogeneous system

$$\boldsymbol{y}'(x) = \mathbf{A}(x)\boldsymbol{y}(x) + \boldsymbol{b}(x)$$

is

$$\boldsymbol{y}_p(x) = \mathbf{Y}(x)\int \mathbf{Y}^{-1}(x)\boldsymbol{b}(x)\,\mathrm{d}x$$

where $\mathbf{Y}$ is the fundamental matrix of the homogeneous system $\boldsymbol{y}'(x) = \mathbf{A}(x)\boldsymbol{y}(x)$.

5.3.12   Proposition: General Solution of a Nonhomogeneous System of LDEs with Variable Coefficients

Let $I \subset \mathbb{R}$ be an interval on the real line, let $\boldsymbol{y}, \boldsymbol{b} : I \to \mathbb{R}^n$ be vector-valued functions, let $\mathbf{A} : I \to \mathbb{R}^{n \times n}$ be a matrix-valued function, and let

$$\boldsymbol{y}'(x) = \mathbf{A}(x)\boldsymbol{y}(x) + \boldsymbol{b}(x)$$

be a non-homogeneous system of n linear differential equations with variable coefficients with homogeneous solution $\boldsymbol{y}_h$ and particular solution $\boldsymbol{y}_p$. In this case, the system has the general solution $\boldsymbol{y}_g : I \to \mathbb{R}^n$

$$\boldsymbol{y}_g = \boldsymbol{y}_h + \boldsymbol{y}_p$$

5.3.13   Definition: Wronskian Determinant

Let $I \subset \mathbb{R}$ be an interval on the real line, let $\boldsymbol{y} : I \to \mathbb{R}^n$ be a vector-valued function, let $\mathbf{A} : I \to \mathbb{R}^{n \times n}$ be a matrix-valued function, and let

$$\boldsymbol{y}'(x) = \mathbf{A}(x)\boldsymbol{y}(x) + \boldsymbol{b}(x)$$

be a homogeneous system of n linear differential equations with variable coefficients with fundamental matrix $\mathbf{Y} : I \to \mathbb{R}^{n \times n}$. In this case, the system's Wronskian determinant, often called the Wronskian and denoted by $W(x)$, is the determinant:

$$W(x) := \det \mathbf{Y}(x)$$

5.3.14   Proposition: Wronskian Determinant and Trace

Let $I \subset \mathbb{R}$ be an interval on the real line, let $\boldsymbol{y} : I \to \mathbb{R}^n$ be a vector-valued function, let $\mathbf{A} : I \to \mathbb{R}^{n \times n}$ be a matrix-valued function, and let

$$\boldsymbol{y}'(x) = \mathbf{A}(x)\boldsymbol{y}(x) + \boldsymbol{b}(x)$$

be a homogeneous system of n linear differential equations with variable coefficients with fundamental matrix $\mathbf{Y} : I \to \mathbb{R}^{n \times n}$ and Wronskian $W(x) = \det \mathbf{Y}(x)$. In this case:

$$W'(x) = \operatorname{tr}\big(\mathbf{A}(x)\big)W(x)$$

and

$$W(x) = Ce^{\int_{x_0}^{x} \operatorname{tr}(\mathbf{A}(\chi))W(\chi)\,\mathrm{d}\chi} \quad \text{for each } x, x_0 \in I$$

where $C \in \mathbb{R}$ is a constant.

5.4   Higher-Order Linear Differential Equations with Variable Coefficients

Note: This section is a medley of content from LDEs with constant and not constant coefficients. It stills needs to be cleaned up and may not be correct.

5.4.1   Definition: Homogeneous and Non-Homogeneous nth-Order LDE with Constant Coefficients

Let $I \in \mathbb{R}$ be an interval on the real line, let $y : I \to \mathbb{R}$ be a real function of differentiability class C^n, let $b : I \to \mathbb{R}$ be a real function, and let $a_0, a_1, \dots, a_n \in \mathbb{R}$ be real constants. In this case, the equations

$$a_0 y + a_1 y' + \cdots + a^n y^{(n)} = 0 \quad \text{or} \quad \sum_{i=0}^{n} a_i y^{(i)} = 0 \quad \text{and}$$

$$a_0 y + a_1 y' + \cdots + a^n y^{(n)} = b(x) \quad \text{or} \quad \sum_{i=0}^{n} a_i y^{(i)} = b(x)$$

are homogeneous and non-homogeneous nth order linear differential equations with constant coefficients, respectively.

5.4.2   Proposition: Matrix Representation of an nth Order LDE

Let $I \in \mathbb{R}$ be an interval on the real line, let $y : I \to \mathbb{R}$ be a real function of differentiability class C^n, let $b : I \to \mathbb{R}$ be a real function, and let $a_0, a_1, \dots, a_n \in \mathbb{R}$ be real constants and let

$$\sum_{i=0}^{n} a_n y^{(n)} = b(x)$$

be a non-homogeneous nth order linear differential equation with constant coefficients. Define the functions $z_1, z_2, \dots, z_n : I \to \mathbb{R}$ given by:

$$
\begin{aligned}
z_1 &:= y & z_1' &= z_2 \\
z_2 &:= y' & z_2' &= z_3 \\
z_n &:= y^{(n-1)} & z_{n-1}' &= z_n
\end{aligned}
$$

In this case, the functions $z_1, z_2, \ldots, z_n$ satisfy the relationships:

$$\sum_{i=0}^{n} a_i z_{i+1} + a_n z_n' = b \qquad\qquad z_n' = \frac{1}{a_n}\left(b - \sum_{i=1}^{n} a_{i-1} z_i\right)$$

And the nth order equation $\sum_{i=0}^{n} a_i y^{(i)} = b(x)$ is equivalently represented by the matrix system:

$$\begin{bmatrix} z_1' \\ \vdots \\ z_n' \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \ddots & \ldots & \ddots & 0 \\ 0 & 0 & 0 & \ldots & 1 \\ -\frac{a_0}{a_n} & -\frac{a_1}{a_n} & -\frac{a_2}{a_n} & \ldots & -\frac{a_{n-1}}{a_n} \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \frac{b}{a_n} \end{bmatrix}$$

5.4.3 Proposition: Dimensions of Solution Set

The solution set of the homogeneous nth order linear differential equation

$$\sum_{i=0}^{n} a_i y^{(i)} = 0$$

is a n-dimensional vector space. In particular, this implies that the solution set consists of n linearly independent functions $y_1, y_2, \ldots, y_n$.

5.4.4 Definition: Fundamental Matrix of an nth Order LDE

Let $\sum_{i=0}^{n} a_i y^{(i)} = 0$ be homogeneous nth order linear differential equation with the linearly independent solutions $y_1, y_2, \ldots, y_n$. In this case, the equation's fundamental matrix $\mathbf{Y}$ is the $n \times n$ matrix:

$$\mathbf{Y} = \begin{bmatrix} y_1 & \cdots & y_n \\ \vdots & \ddots & \vdots \\ y_1^{(n-1)} & \cdots & y_n^{(n-1)} \end{bmatrix}$$

5.4.5 Definition: Wronskian of an nth Order LDE

Let $\mathbf{Y}$ be the fundamental matrix for the homogeneous nth order linear differential equation $\sum_{i=0}^{n} a_i y^{(i)} = 0$. In this case, we define the Wronskian as:

$$W = \det \mathbf{Y}$$

5.4.6 Proposition: Wronskian and Trace

Let $\sum_{i=0}^{n} a_i y^{(i)} = 0$ be homogeneous nth order linear differential equation defined on the real interval $I \subset \mathbb{R}$ with fundamental matrix $\mathbf{Y}$ and Wronskian $W(x)$. In this case, for all $x, x_0 \in I$

$$W(x) = W(x_0) e^{-\int_{x_0}^{x} \operatorname{tr} \mathbf{A}(\zeta)\, \mathrm{d}\zeta}$$

where the matrix $\mathbf{A}$ is defined as above

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \dots & 1 \\ -\frac{a_0}{a_n} & -\frac{a_1}{a_n} & -\frac{a_2}{a_n} & \dots & -\frac{a_{n-1}}{a_n} \end{bmatrix}$$

Note that the trace $\operatorname{tr} \mathbf{A}(x)$ is simply equal to the bottom right matrix element:

$$\operatorname{tr} \mathbf{A}(x) = \frac{-a_{n-1}(x)}{a_n(x)}$$

5.4.7   Variation of Parameters for an nth Order LDE

Let $\sum_{i=0}^{n} a_i(x)y^{(i)}(x) = b(x)$ be a non-homogeneous nth order LDE and let $y_1, y_2, \dots, y_n$ be the solutions to the homogeneous equation $\sum_{i=0}^{n} a_i(x)y^{(i)}(x) = 0$.

In this case, the particular solution y_p of the equation is:

$$y_p = c_1 y_1 + \dots + c_n y_n = \sum_{i=1}^{n} c_i y_i$$

To reach the same answer in matrix form, we define

$$\mathbf{Y}(x) = \begin{bmatrix} y_1 & \dots & y_n \\ \vdots & \ddots & \vdots \\ y_1^{(n-1)} & \dots & y_n^{(n-1)} \end{bmatrix} \quad \boldsymbol{b}(x) = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \frac{b(x)}{a_n} \end{bmatrix} \quad \boldsymbol{y}(x) = \begin{bmatrix} y_1(x) \\ \vdots \\ y_n(x) \end{bmatrix}$$

In this case the particular solution y_p is given by the dot product

$$y_p = \boldsymbol{y}(x) \cdot \left[\int \mathbf{Y}^{-1}(x)\boldsymbol{b}(x) \, \mathrm{d}x \right]$$

5.4.8   General Solution to a Non-Homogeneous nth Order LDE

Let

$$\sum_{i=0}^{n} a_i(x)y^{(i)}(x) = b(x)$$

be a non-homogeneous nth order linear differential equation with the homogeneous solution y_h and the particular solution y_p. In this case, the general solution y_g to the equation is

$$y_g = y_h + y_p$$

5.5   Higher-Order LDEs with Constant Coefficients

5.5.1   Definition: LDE with Constant Coefficients

Let $n \in \mathbb{N}$ be an integer, let $a_n, \ldots, a_0 \in \mathbb{R}$ be real constants, and let $b : \mathbb{R} \to \mathbb{R}$ be a function. In this case, an nth order linear differential equation with coefficients is a differential equation of the form

$$a_n y^{(n)} + \cdots + a_1 y' + a_0 y = b(x)$$

If additionally $b(x) = 0$, the equation is called a *homogeneous* linear differential equation with constant coefficients.

$$a_n y^{(n)} + \cdots + a_1 y' + a_0 y = 0 \qquad \text{(homogeneous)}$$

5.5.2   Note: Operator Notation for LDE with Constant Coefficients

Let $n \in \mathbb{N}$ be an integer, let $a_n, \ldots, a_0 \in \mathbb{R}$ be real constants, and let $b : \mathbb{R} \to \mathbb{R}$ be a function and let L be the differential operator

$$L = \sum_{i=0}^{n} a_i \frac{\mathrm{d}^i}{\mathrm{d}x^i} = a_0 + a_1 \frac{\mathrm{d}}{\mathrm{d}x} + \cdots + a_n \frac{\mathrm{d}^n}{\mathrm{d}x^n}$$

In terms of L, a linear differential equation with constant coefficients is written

$$Ly = b(x) \iff \sum_{i=0}^{n} a_i y^{(i)} = a_n y^{(n)} + \cdots + a_1 y' + a_0 y = b(x)$$

5.5.3   Trial Solution and Characteristic Polynomial

To solve the homogeneous linear differential equation

$$Ly \equiv a_n y^{(n)} + \cdots + a_1 y' + a_0 y = 0$$

we use the trial solution $y(x) = e^{\lambda x}$. When plugged into the equation, the trial solution gives the equation:

$$L(e^{\lambda x}) = (a_n \lambda^n + \cdots + a_1 \lambda + a_0) e^{\lambda x} = 0$$

The complex polynomial $p : \mathbb{R} \to \mathbb{C}$ given by:

$$p(\lambda) = a_n \lambda^n + \cdots + a_1 \lambda + a_0$$

is called the characteristic polynomial of the differential equation.

5.5.4   Solution in Terms of Characteristic Polynomial

Let $p : \mathbb{R} \to \mathbb{C}$ be the characteristic polynomial of the differential equation with constant coefficients

$$a_n y^{(n)} + \cdots + a_1 y' + a_0 y = 0$$

and let p have the factorized form

$$p(\lambda) = a_n (\lambda - \lambda_1)^{k_1} \cdots (\lambda - \lambda_m)^{k_m}$$

where $\lambda_i \in \mathbb{C}, m \in \mathbb{N}, k_1, \ldots, k_n \in \mathbb{N}, k_1 + \cdots + k_m = n$.

In this case, the exponential functions

$$y = e^{\lambda_i x} \quad j \in \{1, 2, \ldots, m\}$$

solve the linear differential equation.

5.5.5   Theorem: Complete Solution Set to LDE with Constant Coefficients

Let $a_0, a_1, \ldots, a_n \in \mathbb{R}$ be real constants, let $m, k_1, k_2, \ldots, k_m$ be (not necessarily equal) integers, let $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{C}$ by complex constants and let $p(\lambda)$ be the complex polynomial

$$p(\lambda) = a_n \lambda^n + \cdots + a_1 \lambda + a_0 = a_n (\lambda - \lambda_1)^{k_1} \cdots (\lambda - \lambda_m)^{k_m}$$

In this case, the family of n exponential functions

$$\mathcal{B} = \left\{ \begin{array}{c} e^{\lambda_1 x}, xe^{\lambda_1 x}, \ldots, x^{k_1-1}e^{\lambda_1 x}, \\ e^{\lambda_2 x}, xe^{\lambda_2 x}, \ldots, x^{k_2-1}e^{\lambda_2 x}, \\ \vdots \\ e^{\lambda_m x}, xe^{\lambda_m x}, \ldots, x^{k_m-1}e^{\lambda_m x} \end{array} \right\}$$

forms a basis of the solution set to the homogeneous linear differential equation with constant coefficients

$$a_n y^{(n)} + \cdots + a_1 y' + a_0 y = 0$$

and

$$k_1 + k_2 + \cdots + k_m = n$$

5.5.6   Proposition: Real Components of a Complex Solution

If the complex function $y : \mathbb{R} \to \mathbb{C}$ solves the homogeneous linear differential equation with constant coefficients

$$a_n y^{(n)} + \cdots + a_1 y' + a_0 y = \sum_{i=0}^{n} a_i y^{(n)} = 0$$

then the real and imaginary components $\operatorname{Re} y$ and $\operatorname{Im} y$ also solve the equation.

5.6   LDE with Variable Coefficients

5.6.1   Euler-Cauchy Equation

The Euler-Cauchy equation is the linear differential equation

$$a_n x^n y^{(n)} + \cdots + a_1 x y' + a_0 y = b(x)$$

where $a_n, \ldots, a_1, a_0 \in \mathbb{R}$ are real constants.

The equation is solved with the trial solution $y = x^\lambda$. When plugged in, the trial solution gives the polynomial equation

$$x^\lambda \sum_{i=0}^{n} a_i \lambda (\lambda - 1) \cdots (\lambda - i + 1) = 0$$

5.7   Introduction to the Calculus of Variations

5.7.1   Definition: Functional

A functional is a mapping from a set of functions to the real numbers $\mathbb{R}$.

5.7.2   Functionals and Lagrangian

Let $L : \mathbb{R}^3 \to \mathbb{R}$ be a differentiable scalar function of 3 real variables and let I be a functional mapping functions y to the real numbers $\mathbb{R}$ with the relationship:

$$I : y \mapsto \int_a^b L(x, y, y') \, dx$$

The motivation of the calculus of variations is to find the function $y : \mathbb{R} \to \mathbb{R}$ minimizing the functional I.

Typical problems involve:

- Find the shortest path connecting to points

- Find the path minimizing the time taken to travel between two points

5.7.3   Example: Minimizing Distance Between Points

Let $A, B \in \mathbb{R}^2$ be two points in a plane with coordinates $A = (a_1, a_2)$ and $B = (b_1, b_2)$ and let γ be a curve connecting the points A, B. More so, let γ be the graph of a differentiable function $f : \mathbb{R} \to \mathbb{R}$.

In this case, the length ℓ of the curve between A and B (i.e. the arc-length of γ) is a functional of f given by:

$$\ell(f) = \int_{a_1}^{b_1} \sqrt{1 + [f'(x)]^2} \, dx$$

For such a problem, the Lagrangian is $L(u, v, w) = \sqrt{1 + \omega^2}$ and the goal is to find the function f minimizing the functional ℓ.

5.7.4   Example: Brachistochrone Problem

Let $A, B \in \mathbb{R}^2$ be two points in a plane with coordinates $A = (a_1, a_2)$ and $B = (b_1, b_2)$ where $b_1 > a_1$ and let γ be a curve connecting the points A, B. More so, let γ be the graph of a differentiable function $f : \mathbb{R} \to \mathbb{R}$.

$$s(t) = \int_{a_1}^{x(t)} \sqrt{1 + [f(\tau)]^2} \, d\tau$$

$$s'(t) = v(t) = \sqrt{1 + [f(x(t))]^2}(x'(t))$$

Equating potential energy

$$E_p = mgh = mg(a_2 - f(x))$$

and kinetic energy

$$E_k = \frac{mv^2}{2} = \frac{m}{2} \left(\sqrt{1 + [f(x(t))]^2} \left(\frac{dx}{dt} \right) \right)^2$$

gives

$$\sqrt{2g(a_2 - f(x))} = \sqrt{1 + [f(x)]^2} \frac{\mathrm{d}x}{\mathrm{d}t}$$

$$t = \int \mathrm{d}t = \int_{a_1}^{b_1} \frac{1}{\sqrt{2g}} \sqrt{\frac{1 + [f(x)]^2}{a_2 - f(x)}} \, \mathrm{d}x$$

In this case, the linear functional is $t = t(f)$ and the Lagrangian is

$$L = L(u, v, w) = \frac{1}{\sqrt{2g}} \sqrt{\frac{1 + w^2}{a_2 - v}}$$

5.7.5 Formalizing the Calculus of Variation Problem

Let $L : \mathbb{R}^3 \to \mathbb{R}$ be a scalar function of three variables, let $I = [a, b] \in \mathbb{R}$ be an interval on the real line, let X be the space of differentiable real functions mapping from I to $\mathbb{R}$, and let $y : I \to \mathbb{R}$ be a differentiable function. For a given Lagrangian $L(u, v, w)$ we want to find the extrema of the functional

$$J(y) = \int_a^b L(x, y, y') \, \mathrm{d}x$$

mapping from X to the real numbers.

The extrema of the functional J is the function y for which J has an extrema in all directions. More formally, y is the function for which for each smooth function $\eta : I \to R$ that is zero at the endpoints ($\eta(a) = \eta(b) = 0$) the function $\epsilon I \to \mathbb{R}$ mapping to $I(y + \epsilon\eta)$ has an extrema at 0.

5.7.6 Critical Point of a Functional

The function y is a critical point of the functional J if J has directional derivatives

$$\frac{\partial}{\partial \epsilon} J(y + \epsilon\eta)\big|_{\epsilon=0}$$

equal to zero for all smooth functions $\eta : I \to \mathbb{R}$.

5.7.7 Lemma: Integral Over a Closed Interval

Let $I = [a, b]$ be an interval on the real line, let $f : I \to \mathbb{R}$ be a continuous function, and let

$$\int_a^b f(x)\eta(x) \, \mathrm{d}x = 0$$

for all continuously differentiable functions $\eta : I \to \mathbb{R}$ for which $\eta(a) = \eta(b) = 0$. In this case, $f(x) = 0$ for all $x \in I$.

5.7.8 Generalization of Previous Lemma

Let $I = [a, b]$ be an interval on the real line, let $M, N : I \to \mathbb{R}$ be continuous functions and let

$$\int_a^b (M\eta + N\eta') \, \mathrm{d}x = 0$$

for all continuously differentiable functions $\eta : I \to \mathbb{R}$ for which $\eta(a) = \eta(b) = 0$. In this case, N is differentiable and $N' = M$.

5.7.9 Euler-Lagrange Equation

For a Lagrangian $L(x, y, y')$, the Euler-Lagrange equation is

$$L_y - \frac{\mathrm{d}}{\mathrm{d}x}\left[L_{y'}\right] = 0$$

More formally, for a Lagrangian $L(u, v, w)$ where u, v, w are real functions, we define the function $V : \mathbb{R} \to \mathbb{R}^3$ with the relationship

$$V = V(x) = (x, y(x), y'(x))$$

In this case, the Euler-Lagrangian equation reads

$$\frac{\partial}{\partial v}\left[L(V(x))\right] - \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{\partial}{\partial w}\left[L(V(x))\right]\right]$$

5.7.10 Necessary Condition for Extrema of a Functional

Let $I = [a, b]$ be an interval on the real line, let $y : I \to \mathbb{R}$ be a continuously differentiable function of the variable x, and let $L : \mathbb{R}^3 \to R$ with the relationship $L = L(x, y, y')$ be the Lagrangian for the functional J defined as:

$$J(y) = \int_a^b L(x, y, y')\,\mathrm{d}x$$

If y is an extrema of the functional J, then y satisfies the Euler-Lagrange equation.

5.7.11 Note: Functional with Variable Endpoints

5.7.12 Note: Lagrangian Independent of y

Let $L = L(x, y')$ be a Lagrangian for a CV problem such that L is independent of y. In this case, the Euler-Lagrange equation simplifies to

$$L_{y'} = C \qquad\qquad \text{(Euler-Lagrange equation if } L = L(x, y'))$$

5.7.13 Note: Lagrangian Independent of x

Let $L = L(y, y')$ be a Lagrangian for a CV problem such that L is independent of x. In this case, the Euler-Lagrange equation simplifies to

$$L - y'L_{y'} = C \qquad\qquad \text{(Euler-Lagrange equation if } L = L(y, y'))$$

6   Introduction to Hilbert Spaces and Fourier Series

6.1   Basics of Hilbert Spaces

For the entirety of this section, let V be a vector space over the scalar field $\mathbb{F}$, where in general $\mathbb{F}$ is taken to be the complex numbers $\mathbb{C}$.

6.1.1   Definition: Inner Product

An inner product on the vector space V is a mapping $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ mapping $(x, y) \mapsto \langle x, y \rangle$ with the following properties:

1. $\langle x, x \rangle \geq 0$ for all $x \in V$

2. $\langle x, x \rangle = 0 \iff x = 0$

3. $\langle x, y \rangle$ is linear in the first factor and anti-linear in the second factor. In terms of the scalars $\lambda, \mu \in \mathbb{F}$, this is written:

$$\langle \lambda x + \mu y, z \rangle = \lambda \langle x, z \rangle + \mu \langle y, z \rangle$$
$$\langle z, \lambda x + \mu y \rangle = \overline{\lambda} \langle z, x \rangle + \overline{\mu} \langle z, y \rangle$$

   where $\overline{\lambda}$ denotes the complex conjugate of λ.

4. $\langle x, y \rangle = \overline{\langle y, x \rangle}$

6.1.2   Definition: Inner Product Space

An inner product space is a vector space V over the scalar field $\mathbb{F}$ equipped with an inner product $\langle \cdot, \cdot \rangle$.

6.1.3   Definition: Norm

Let V be an inner product space and let $x \in V$ be an element of V. A norm on the inner product space V is a mapping $\|\cdot\| : V \to [0, \infty)$ mapping $x \mapsto \|x\|$ with the following properties:

1. $\|x\| \geq 0$ for all $x \in V$

2. $\|x\| = 0 \iff x = 0$

3. $\|\lambda x\| = |\lambda| \|x\|$ for all $\lambda \in \mathbb{F}$.

4. $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in V$

6.1.4   Theorem: Cauchy-Schwartz Inequality

Let V be an inner product space with the norm $\|\cdot\|$ and let $x, y \in V$ be a elements of V. For each $x, y \in V$:

$$|\langle x, y \rangle| \leq \|x\| \|y\|$$

6.1.5   Definition: Metric

Let $x, y \in V$ be elements of the inner product space V. A metric on the inner product space V is the mapping $d : V \times V \to [0, \infty)$ mapping $(x, y) \mapsto \|x - y\|$.

$$d(x, y) = \|x - y\|$$

6.1.6   Definition: Completeness of a Vector Space

A vector space V is complete if every Cauchy sequence has its limit in V.

6.1.7   Definition: Hilbert Space

A Hilbert space is a complete inner product space with respect to the metric induced by the inner product.

6.1.8   Definition: The $\|\cdot\|_1$ and $\|\cdot\|_2$ Norms

Let $I = [a, b]$ be an interval on the real line and let $f : I \to \mathbb{R}$ be an integrable function. The $\|f\|_1$ and $\|f\|_2$ norms are defined as

$$\|f\|_1 = \int_a^b |f| \qquad\qquad \|f\|_2 = \sqrt{\int_a^b |f|^2}$$

6.1.9   Definition: L^2 Space

Let $I = [a, b]$ be an interval on the real line. The Hilbert space $L^2(I)$ is the set of all functions $f : I \to \mathbb{C}$ for which

$$\int_a^b |f|^2 < \infty$$

where the norm is defined as

$$\|f\|_2 = \sqrt{\int_a^b |f|^2}$$

6.1.10   Proposition: Integrating a Complex Function

Let $I = [a, b]$ be an interval on the real line, let $u, v : I \to \mathbb{R}$ be real, integrable functions and let $f : I \to \mathbb{C}$ be the complex function given by $f = u + iv$. In this case, f is integrable and

$$\int_a^b f = \int_a^b u + i \int_a^b v$$

6.1.11   Comparison of the $\|\cdot\|_1$ and $\|\cdot\|_2$ Norms

Let $I = [a, b]$ be an interval on the real line and let $f \in L^2(I)$ be a function in $L^2(I)$. In this case:

$$\|f\|_1 = \int_a^b |f| = \int_a^b |f| \cdot 1 = \langle |f|, 1 \rangle \leq \| |f| \|_2 \|1\|_2 =$$

$$= \sqrt{\int_a^b |f|^2} \sqrt{\int_a^b 1} = \|f\|_2 \sqrt{b - a}$$

The general result is that $\|f\|_1 \leq \|f\|_2$, which implies that $L^2(I) \subset L^1(I)$.

6.1.12   Definition: Orthogonality

Let V be an inner product space. The elements $x, y \in V$ are orthogonal if $\langle x, y \rangle = 0$.

6.1.13 Definition: Orthogonal Complement

Let V be an inner product space and let $A \subset V$ be a subset of V. In this case, A's orthogonal complement $A^\perp$ is the subset $A^\perp \subset V$ consisting of all $y \in V$ for which $\langle x, y \rangle = 0$ for all $x \in A$. In set builder notation:

$$A^\perp = \{ y \in V \mid \langle x, y \rangle = 0 \quad \forall \quad x \in A \}$$

Note that zero is always a member of any orthogonal complement because $\langle x, 0 \rangle \equiv 0$ for all x.

6.1.14 Proposition: Closure of the Orthogonal Complement

Let V be an inner product space, let $A \subset V$ be a subset of V, and let $A^\perp$ be A's orthogonal complement. In this case, $A^\perp$ is a closed set.

6.1.15 Theorem: Pythagorean Theorem

Let V be an inner products space and let $x_1, x_2, \ldots, x_n \in V$ be mutually orthogonal elements of V. In this case:

$$\left\| \sum_{i=1}^{n} x_i \right\|^2 = \sum_{i=1}^{n} \|x_i\|^2$$

6.1.16 Definition: Orthogonal Projection

Let V be an inner product space and let $U \subset V$ be a linear subspace of V. If there exists $u \in U$ such that, for a given $v \in V$, the vector difference $v - u \perp U$, then u is the orthogonal projections of v onto the subspace U and is denoted by $\mathrm{Pr}_U(v)$.

6.1.17 Proposition: Properties of the Orthogonal Projection

Let V be an inner product space, let $U \subset V$ be a linear subspace of V and define the vectors $v \in V$ and $u \in U$ such that $u = \mathrm{Pr}_U(v)$ is the orthogonal projection of v onto the subspace U.

1. u is unique.

2. u is the closest vector in U to the vector $v \in V$.

6.1.18 Proposition: Calculating the Orthogonal Projection

Let V be a finite-dimensional inner product, let $U \subset V$ be a subspace of V, and let $\mathcal{B} = \{e_1, e_2, \ldots, e_n\}$ be an orthonormal basis of the subspace U. In this case, the orthogonal projection $\mathrm{Pr}_U(v)$ exists for all $v \in V$ and is equal to

$$\mathrm{Pr}_U(v) = \sum_{i=1}^{n} \langle v, e_i \rangle \, e_i$$

6.1.19 Proposition: Bessel Inequality

Let H be a Hilbert space with the orthonormal basis $\mathcal{E} = \{e_1, e_2, \ldots, e_n\}$ where $n \in \{\mathbb{N} \cup \infty\}$ and let $x \in H$ be an element of H. In this case,

$$\|x\|^2 \geq \sum_{i=1}^{n} |\langle x, e_i \rangle|^2$$

for all $x \in H$.

6.1.20   Proposition: Sequences in Hilbert Spaces

Let H be a Hilbert space with orthonormal basis $\mathcal{E} = \{e_1, e_2, \dots\}$ and let $(c_n)_n \in \mathbb{F}$ be a sequence in the scalars such that

$$\sum_{i=1}^{n} |c_i|^2 < \infty$$

In this case, the series

$$x := \sum_{i=1}^{\infty} c_i e_i$$

converges and $c_i = \langle x, e_i \rangle$ for all $i \in \mathbb{N}$.

Note: Interpreting the Previous Proposition The sequence

$$(s_n)_n := \sum_{i=1}^{n} c_i e_i$$

converges in H. The confirm this, we refer to the Cauchy criterion. Let $m, n \in \mathbb{N}$ be integers such that $m > n$. In this case

$$s_m - s_n = \sum_{i=n+1}^{m} c_i e_i$$

Next,

$$\|s_m - s_n\|^2 = \left\| \sum_{i=n+1}^{m} c_i e_i \right\|^2 = \sum_{i=n+1}^{m} |c_j|^2 < \epsilon \text{ for } m, n > M_0 \in \mathbb{N}$$

We can make the final estimate because because the $c \in l^2$ so the partial sums $\sum_{i=1}^{n} c_j$ satisfy the Cauchy criterion in $\mathbb{F}$.

The result is that $(s_n)_n$ the Cauchy criterion for convergence in H. Because H is complete, there exists

$$x := \lim_{n \to \infty} s_n$$

(compare to previous proposition).

Note: On Coefficients

$$\lim_{n \to \infty} \langle s_n, e_i \rangle = \langle x, e_i \rangle = c_j$$

First, we show that

$$|\langle x, e_i \rangle - \langle s_n, e_i \rangle| = |\langle x - s_n \rangle, e_i| \le \|x - s_n\| \|e_j\| = \|x - s_n\|$$
$$\lim_{n \to \infty} \|x - s_n\| = \|x - x\| = 0$$

For $n \ge j$ we have:

$$\langle s_n, e_j \rangle = \sum_{j=1}^{n} c_j \langle e_j, e_i \rangle = c_j$$

The result here is that

$$\langle x, e_j \rangle = c_j$$

Quick Review of Terms:

H is a Hilbert space with orthonormal basis $\mathcal{E} = \{e_i\}_1^n$.

- $c_n \in \mathbb{F}$ is a scalar and $(c_n)_n$ is a sequence of scalars. All s, x and e terms are vectors (i.e. elements of H); e terms are orthonormal basis vectors.

- $s_n = \sum_{i=1}^n c_i e_i$ is a finite series of elements of H and an element of H itself.

- $(s_n)_n$ is a sequence of finite series i.e. a sequence of s_n terms.

- $x_n = \lim_{n \to \infty} s_n$ is a generalization of s_n to an infinite-dimensional space.

6.2 Introduction to Fourier Series

6.2.1 Definition: Fourier Series

Let H be a Hilbert space with orthonormal basis $\mathcal{E} = \{e_i\}_1^\infty$ and define $x \in H$. The series

$$\sum_{i=1}^\infty \langle x, e_i \rangle e_i$$

is called the Fourier series of x for the chosen basis $\mathcal{E}$ and the scalars $\langle x, e_j \rangle \in \mathbb{F}$ are called the Fourier coefficients of the Fourier series.

6.2.2 Definition: Completeness of a Orthonormal Basis

Let H be a Hilbert space with orthonormal basis $\mathcal{E} = \{e_i\}_1^\infty$ and define $x \in H$. The orthonormal basis $\mathcal{E}$ is complete if

$$x = \sum_{i=1}^\infty \langle x, e_i \rangle e_i \quad \text{for all} \quad x \in H$$

6.2.3 Theorem: Equivalent Properties of Fourier Series

Let H be a Hilbert space with orthonormal basis $\mathcal{E} = \{e_i\}_1^\infty$. The following properties are equivalent:

1. $\mathcal{E}$ is a complete orthonormal basis for H

2. For each $x, y \in H$
$$\langle x, y \rangle = \sum_{i=1}^\infty \langle x, e_i \rangle \overline{\langle y, e_i \rangle}$$

3. Parseval's equality: for all $x \in H$
$$\|x\|^2 = \sum_{i=1}^\infty |\langle x, e_i \rangle|^2$$

4. $\mathcal{E}$ is not contained in any orthonormal system with cardinality greater than $\mathcal{E}$.

5. The orthogonal complement $\mathcal{E}^\perp$ consists only of the zero vector, i.e. $\mathcal{E}^\perp = \{0\}$.

6.2.4 Note: The $L^2(-\pi, \pi)$ Space

The Hilbert space $L^2(-\pi, \pi)$ is the space given by:

$$\left\{ f : [-\pi, \pi] \to \mathbb{F} \,\middle|\, \|f\|_2 < \infty \right\}$$

where the norm $\|f\|_2$ is given by:

$$\|f\|_2 = \sqrt{\int_{-\pi}^{\pi} |f(x)|^2 \, \mathrm{d}x}$$

and the inner product is

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)\overline{g(x)} \, \mathrm{d}x \quad \forall \quad f, g \in L^2(-\pi, \pi)$$

6.2.5 Proposition: Complete Orthonormal Basis of $L^2(-\pi, \pi)$

Let $e_n : [-\pi, \pi] \to \mathbb{F}$ be the family of functions with the relation

$$e_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx} \qquad\qquad x \in [-\pi, \pi]; \; n \in \mathbb{Z}$$

and let $\mathcal{E}$ be the set given by $\mathcal{E} = \left\{ e_n \,\middle|\, n \in \mathbb{Z} \right\}$. In this case, $\mathcal{E}$ is a complete orthonormal basis of the Hilbert space $L^2(-\pi, \pi)$.

6.2.6 Proposition: Complete Orthonormal Basis of Real $L^2(-\pi, \pi)$

Let $L^2_{\mathbb{R}}(-\pi, \pi)$ denote the real component of the Hilbert space $L^2(-\pi, \pi)$; that is, the space given by:

$$\left\{ f : [-\pi, \pi] \to \mathbb{R} \,\middle|\, \|f\|_2 < \infty \right\}$$

where the norm and inner product are defined as for $L^2(-\pi, \pi)$. Let $e_{n_c}, e_{n_s} : [-\pi, \pi] \to \mathbb{R}$ be the family of functions with the relation

$$e_{n_c}(x) = \frac{1}{\sqrt{\pi}} \cos(nx) \qquad e_{n_s}(x) = \frac{1}{\sqrt{\pi}} \sin(nx) \qquad x \in [-\pi, \pi]; \; n \in \mathbb{Z} \setminus \{0\}$$

and let $\mathcal{F}$ be the set given by $\mathcal{F} = \left\{ \frac{1}{\sqrt{2\pi}}, e_{n_c}, e_{n_s} \,\middle|\, n \in \mathbb{Z} \setminus \{0\} \right\}$. In this case, $\mathcal{F}$ is a complete orthonormal basis of $L^2_{\mathbb{R}}(-\pi, \pi)$.

6.2.7 Proposition: Fourier Coefficients and Fourier Series

Let $f \in L^2_{\mathbb{R}}(-\pi, \pi)$ be a function. The Fourier coefficients $a_n, b_n \in \mathbb{R}$ for the function f with respect to the basis $\mathcal{F}$ are given by

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) \, \mathrm{d}x \qquad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) \, \mathrm{d}x \qquad n \in \mathbb{N}$$

and Fourier series for f is

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos(nx) + b_n \sin(nx)] \qquad \text{where} \qquad a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \, \mathrm{d}x$$

Note that in general the Fourier coefficients, and thus the Fourier series, for a given function $f \in L^2_{\mathbb{R}}(-\pi, \pi)$ depend on the choice of the basis for the space $L^2_{\mathbb{R}}(-\pi, \pi)$.

6.2.8 Proposition: Fourier Coefficients Fall with Increasing n

Let $f \in L^2(-\pi, \pi)$ be a function with Fourier coefficients a_n, b_n and define $x \in\in [-\pi, \pi]$. In this case

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} f(x)\cos(nx)\,\mathrm{d}x = \lim_{n\to\infty} \int_{-\pi}^{\pi} f(x)\sin(nx)\,\mathrm{d}x = 0$$

In other words:

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = 0$$

6.2.9 Definition: L_2 Convergence of Fourier Series

Let $f \in L_{\mathbb{R}}^2(-\pi, \pi)$ be a function and let $s_n(x)$ be the sequence given by

$$s_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n} [a_k \cos(kx) + b_k \sin(kx)]$$

The sequence s_n converges to f in the L_2 sense if

$$\lim_{n\to\infty} \|s_n - f\|_2 = 0$$

6.2.10 Proposition: L_2 Convergence of Fourier Series

Let $f \in L_{\mathbb{R}}^2(-\pi, \pi)$ be a function and let $s_n(x)$ be defined as above. In this case, s_n converges to f with respect to the L_2 norm (i.e. in the L_2 sense described above) for all $f \in L_{\mathbb{R}}^2(-\pi, \pi)$.

6.2.11 Proposition: Identity for Fourier Coefficients

Let $f \in L_{\mathbb{R}}^2(-\pi, \pi)$ be a function with Fourier coefficients a_0, a_n, b_n. In this case

$$\frac{1}{\pi} \int_{-\pi}^{\pi} |f(x)|^2 \,\mathrm{d}x = \frac{a_0^2}{2} + \sum_{n=1}^{\infty} \left(a_n^2 + b_n^2\right)$$

for all $x \in [-\pi, \pi]$.

6.2.12 Definition: Pointwise Convergence of Fourier Series

Let $f \in L_{\mathbb{R}}^2(-\pi, \pi)$ be a function and let $s_n(x)$ be the sequence given by

$$s_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n} [a_k \cos(kx) + b_k \sin(kx)]$$

The sequence s_n converges pointwise to f if

$$\lim_{n\to\infty} s_n(x) = f(x) \quad \forall \quad x \in [-\pi, \pi]$$

Pointwise convergence is in general a stronger requirement than convergence in the L_2 sense. Unlike for L_2 convergence, the sequence s_n *does not* converge pointwise to f for all $f \in L_{\mathbb{R}}^2(-\pi, \pi)$ but only for special cases of f.

6.2.13 Theorem: Pointwise Convergence of Fourier Series for Continuous Functions

Let $f \in L^2(-\pi, \pi)$ be a continuous and differentiable function with Fourier coefficients a_0, a_n, b_n. In this case, f's Fourier series converges pointwise to f for all $x \in [-\pi, \pi]$. In other words:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos(nx) + b_n \sin(nx)]$$

6.2.14 Theorem: Pointwise Convergence of Fourier Series for Piecewise Continuous Functions

Let $f \in L^2(-\pi, \pi)$ be a piecewise continuous and piecewise differentiable function with Fourier coefficients a_0, a_n, b_n. More so, let $f\_(x)$ and $f_+(x)$ denote the left and right limits, respectively of f at x, written in symbols as:

$$f\_(x) = \lim_{t \to x^-} f(t) \qquad\qquad f_+(x) = \lim_{t \to x^+} f(t)$$

In this case, f's Fourier series converges pointwise to f for all $x \in [-\pi, \pi]$. In other words:

$$\frac{f\_(x) + f_+(x)}{2} = \frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos(nx) + b_n \sin(nx)]$$