Required for: MATH40004 Calculus and Applications

Based on the lectures of Demetrios Papageorgiou and Vahid Shahrezaei, Imperial College London

1 Limits of Functions

1.0.1 Definition: as $x \to x_0$, $f(x) \to l$

For a function f, defined over an interval containing x_0 but not necessarily at x_0, $l \in \mathbb{R}$ is the limit of $f(x)$ as x approaches x_0 if $\forall\,\varepsilon > 0$, $\exists\,\delta > 0$ such that $|x - x_0| < \delta \implies |f(x) - l| < \varepsilon$.

Written $\lim_{x\to x_0} f(x) = l$.

1.0.2 Definition: as $x \to \infty$, $f(x) \to l$

For a function f, defined over an interval (a, ∞), $l \in \mathbb{R}$ is the limit of $f(x)$ as x approaches ∞ if $\forall\,\varepsilon > 0$, $\exists\,A > a$ such that $x > A \implies |f(x) - l| < \varepsilon$.

Written $\lim_{x\to\infty} f(x) = l$ [Similarly for $\lim_{x\to-\infty} f(x) = l$].

1.0.3 Definition: as $x \to x_0$, $f(x) \to \infty$

For a function f, defined over an interval containing x_0 but not necessarily at x_0, the limit of $f(x)$ as x approaches x_0 is ∞ if $\forall\,B \in \mathbb{R}_{>0}$, $\exists\,\delta > 0$ such that $|x - x_0| < \delta \implies f(x) > B$.

Written $\lim_{x\to x_0} f(x) = \infty$ [Similarly for $\lim_{x\to x_0} f(x) = -\infty$].

1.0.4 Definition: one-sided limits

For a function f, defined over an interval to the right [Resp. left] of x_0, $l \in \mathbb{R}$ is the limit of $f(x)$ as x approaches x_0 from the right [Resp. left] if $\forall\,\varepsilon > 0$, $\exists\,\delta > 0$ such that $x_0 < x < x + \delta \implies |f(x) - l| < \varepsilon$ [Resp. $x - \delta < x < x_0 \implies |f(x) - l| < \varepsilon$].

Written $\lim_{x\to x_0+} f(x) = l$ [Resp. $\lim_{x\to x_0-} f(x) = l$].

1.1 Properties of Limits

1.1.1 Theorem

$$\lim_{x\to x_0} f(x) + g(x) = \lim_{x\to x_0} f(x) + \lim_{x\to x_0} g(x).$$

$$\lim_{x\to x_0} f(x)g(x) = \lim_{x\to x_0} f(x) \lim_{x\to x_0} g(x).$$

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \frac{\lim_{x\to x_0} f(x)}{\lim_{x\to x_0} g(x)}.$$

If $h(x)$ is continuous at $\lim_{x\to x_0} f(x)$, then

$$\lim_{x\to x_0} h(f(x)) = h\left(\lim_{x\to x_0} f(x)\right).$$

1.1.2 Theorem: Comparison Test

If $\lim_{x\to x_0} f(x) = 0$ and $|g(x)| \leq |f(x)|\ \forall\,x$ approaching x_0, then $\lim_{x\to x_0} g(x) = 0$ (applies to $\lim_{x\to\infty}$). Similarly, if $\lim_{x\to x_0} g(x) = \infty$, then $\lim_{x\to x_0} f(x) = \infty$.

2 Differentiation

2.0.1 Definition

For a function f, the derivative of f at x is

$$\frac{df}{dx} = \frac{d}{dx}[f(x)] = f'(x) = \lim_{h\to 0}\frac{f(x+h)-f(x)}{h}.$$

If the limit exists, f is differentiable.

2.0.2 Theorem

$$\frac{d}{dx}[(f+g)(x)] = f'(x) + g'(x).$$

$$\frac{d}{dx}[(fg)(x)] = f'(x)g(x) + f(x)g'(x).$$

$$\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2}.$$

$$\frac{d}{dx}[(f\circ g)(x)] = f'\circ g(x)g'(x).$$

2.0.3 Note

For two functions x and y of a parameter t, the derivative $\frac{dy}{dx} = \frac{\left(\frac{dy}{dt}\right)}{\left(\frac{dx}{dt}\right)}$.

2.1 Polynomials

2.1.1 Theorem

Let n be an integer, $n > 1$, and let $f(x) = x^n$. $f'(x) = \frac{df}{dx} = nx^{n-1}$.

2.1.2 Theorem

Let a be any real number and let $f(x) = x^a$. For $x > 0$, $f'(x) = \frac{df}{dx} = ax^{a-1}$.

2.2 Maxima, Minima and Continuity

2.2.1 Definition

For a function f defined at m, m is a maximum of f if $f(m) \geq f(x)\ \forall x$ and a minimum if $f(m) \leq f(x)\ \forall x$.

2.2.2 Theorem

For a function f defined and differentiable over an interval (a, b),

$$c \text{ is a maximum or a minimum} \implies f'(c) = 0$$

2.2.3 Definition

In the context of calculus, $f(x)$ is continuous over the interval $[a, b]$ if $\lim_{h\to 0} f(x+h) = f(x)\ \forall\, a \leq x \leq b$.

2.2.4 Theorem

For a function f defined and continuous over an interval $[a, b]$, $\exists\, x_{max}, x_{min}$ such that $f(x_{max}) \geq f(x)$ and $f(x_{min}) \leq f(x)\ \forall x \in [a, b]$.

2.2.5 Theorem: Rolle's Theorem

For a function f defined and continuous over an interval $[a, b]$ and differentiable over (a, b), $f(a) = f(b) \implies \exists\, c \in (a, b)$ such that $f'(c) = 0$.

2.2.6 Theorem: Mean Value Theorem

For a function f defined and continuous over an interval $[a, b]$ and differentiable over (a, b), $\exists\, c \in (a, b)$ such that

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

(there is a point in the interval at which the slope of the curve is equal to the slope between the two points).

2.2.7 Theorem: Cauchy Mean Value Theorem

For two functions f, g defined and continuous over an interval $[a, b]$ and differentiable over (a, b), where $g(a) \neq g(b)$, $\exists\, c \in (a, b)$ such that

$$\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

2.2.8 Definition

For a function f defined over an interval, $\forall\, x_1, x_2$ in the interval, f is
increasing if $x_1 < x_2 \implies f(x_1) \leq f(x_2)$,
decreasing if $x_1 < x_2 \implies f(x_1) \geq f(x_2)$,
strictly increasing if $x_1 < x_2 \implies f(x_1) < f(x_2)$,
strictly decreasing if $x_1 < x_2 \implies f(x_1) > f(x_2)$.

2.2.9 Theorem

For a function f defined and continuous over an interval $[a, b]$ and differentiable over (a, b),
$f'(x) = 0 \quad \forall\, x \in (a, b) \implies f$ is constant over (a, b),
$f'(x) > 0 \quad \forall\, x \in (a, b) \implies f$ is strictly increasing over (a, b),
$f'(x) < 0 \quad \forall\, x \in (a, b) \implies f$ is strictly decreasing over (a, b).

2.2.10 Theorem: Intermediate Value Theorem

For a function f defined and continuous over an interval $[a, b]$, $\forall\, y* \in [f(a), f(b)]\ \exists\, x* \in [a, b]$ such that $y* = f(x*)$.

3 Inverse Functions

3.0.1 Definition

For a function f defined on an interval, We require a unique x_0 in the domain for each y_0 in the codomain such that $y_0 = f(x_0)$ in order to define $g(y) = x$. This inverse function is often notated $x = f^{-1}(y)$.

3.0.2 Theorem

If a function is continuous and is strictly increasing or decreasing, the inverse exists.

3.0.3 Theorem

Specifically, for a function f continuous over an interval $[a, b]$ and strictly increasing or decreasing, the inverse is defined over the interval $[f(a), f(b)]$.

3.1 Derivatives of Inverse Functions

3.1.1 Theorem

For a function f differentiable and strictly increasing or decreasing over an interval, ${f^{-1}}'(x) = \frac{1}{f'(x)}$.

3.2 Logarithms

3.2.1 Definition

$\log x$ is the area under the curve $\frac{1}{x}$ between 1 and x for $x \geq 1$ or negative the area between 1 and x for $0 < x \leq 1$. $\log 1 := 0$.

3.2.2 Theorem

The derivative of $\log x$ exists and is equal to $\frac{1}{x}$.

3.2.3 Theorem

For $a, b > 0$, $\log(ab) = \log a + \log b$.

3.2.4 Corollary

For $a > 0, n \in \mathbb{Z}$, $\log(a^n) = \log(\underbrace{a \cdot a \cdot a \cdot ... \cdot a}_{n}) = n \log a$.

3.2.5 Corollary

For $a, b > 0$, $\log(\frac{a}{b}) = \log(ab^{-1}) = \log a + \log b^{-1} = \log a - \log b$.

3.2.6 Theorem

$\log x$ (defined for $x > 0$) is strictly increasing and its range is $(-\infty, \infty)$.

3.3 The Exponential Function

3.3.1 Definition

Define $\exp(x)$ as the inverse of $\log x$, which exists. Define $e = \exp(1)$. Now $\exp(x) = e^x$ (provable by induction). I will write e^x from now on.

3.3.2 Theorem

The derivative of e^x exists and is equal to e^x.

3.3.3 Theorem

$\frac{d}{dx} a^x = a^x (\log a)$.

4 Finding Limits

4.0.1 Theorem

For $a \in \mathbb{R}_{>0}$, $\lim_{n\to\infty} \frac{(1+a)^n}{n} \to \infty$.

4.0.2 Corollary

$\lim_{n\to\infty} \frac{e^n}{n} \to \infty$.

4.0.3 Theorem

$f(x) = \frac{e^x}{x}$ is strictly increasing for $x > 1$ and $\lim_{x\to\infty} f(x) = \infty$.

4.0.4 Corollary

$\lim_{x\to\infty} x - \log x = \infty$.

4.0.5 Corollary

$\lim_{x\to\infty} \frac{x}{\log x} = \infty$.

4.0.6 Corollary

$\lim_{x\to\infty} x^{\frac{1}{x}} = 1$.

4.0.7 Theorem

$f(x) = \frac{e^x}{x^m}$ $(m \in \mathbb{R}_{>0})$ is strictly increasing for $x > m$ and $\lim_{x\to\infty} f(x) = \infty$.

4.0.8 Theorem: L'Hôpital's Rule

When $\frac{f(x_0)}{g(x_0)}$ is of an indeterminant form ($\frac{0}{0}$ or $\frac{\infty}{\infty}$),

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \lim_{x \to x_0} \frac{f'(x)}{g'(x)}.$$

The rule holds for limits where x_0 is ∞ or $-\infty$.

5 Integration

5.0.1 Definition

The indefinite integral of a function f, defined over some interval, is F such that $F'(x) = f(x)$, defined over the same interval. The integral is written

$$F = \int f \; dx$$

which is not unique, since if $G(x) := F(x) + c$, $\frac{d}{dx}\big[F - G\big] = 0 \implies G'(x) = F'(x) = f(x)$.

5.0.2 Definition

The definite integral from a to b of a function f, defined over $[a, b]$, is the area under the curve between $x = a$ and $x = b$, written

$$\int_a^b f \; dx$$

(which <u>is</u> unique, since the constants from $F(a)$ and $F(b)$ are equal and cancel - see 5.1.4 fundamental theorem of calculus). Note any area beneath the x-axis is negative in the definite integral.

5.0.3 Theorem

Define a function for the area under the curve between a and some x,

$$F_a(x) = \int_a^x f(t) \; dt$$

$F_a(x)$ is differentiable with derivative $f(x)$ (for any a). ($f(x)$, which is the height of the curve at x, can be thought of as the rate of change of the area under the curve at x, which does not depend on the lower limit a).

5.1 The Riemann Sum

5.1.1 Definition

For a function f defined over $[a, b]$, take a partition $x_i = a + ih$, $i = 0, ..., n$ (where $h = \frac{b-a}{n}$). For each subinterval, let $x_i^* \in [x_{i-1}, x_i]$. The Riemann Sum is defined

$$\sum_{i=1}^{n} f(x_i^*)h.$$

Three useful cases are

- $x_i^* = x_i$, the right-hand Riemann Sum.
- $x_i^* = x_{i-1}$, the left-hand Riemann Sum.
- $x_i^* = \frac{1}{2}(x_i + x_{i-1})$, the midpoint Riemann Sum.

5.1.2 Theorem

$$\lim_{n \to \infty} \left(\sum_{i=1}^{n} f(x_i^*)h \right) = \int_a^b f \; dx$$

5.1.3 Theorem

$$\int_a^b cf\ dx = c\int_a^b f\ dx.$$

$$\int_a^b f + g\ dx = \int_a^b f\ dx + \int_a^b g\ dx.$$

$$\int_a^b f\ dx = -\int_b^a f\ dx.$$

For $c \in (a, b)$,

$$\int_a^b f\ dx = \int_a^c f\ dx + \int_c^b f\ dx.$$

If $f(x) \leq g(x)\ \forall\, x \in [a, b]$,

$$\int_a^b f\ dx \leq \int_a^b g\ dx.$$

5.1.4 Theorem: The Fundamental Theorem of Calculus

For a function F differentiable on $[a, b]$ and with F' integrable on $[a, b]$,

$$\int_a^b F'\ dx = F(b) - F(a).$$

5.1.5 Theorem

$$\frac{d}{dx}\left[\int_a^{g(x)} f(t)\ dt\right] = f(g(x)) \cdot g'(x).$$

5.2 Improper Integrals

5.2.1 Definition

$\int_a^b f\ dx$ is an improper integral if $a = -\infty$ or $b = \infty$ or $f(x) \to \pm\infty$ in (a, b). Improper integrals can be found by taking limits of proper integrals. If the limits are finite, the integral is convergent, otherwise it is divergent.

5.2.2 Theorem: Comparison Test

If $|g(x)| \leq f(x)\ \forall\, x \geq a$, then

1. $\int_a^\infty f\ dx$ is convergent $\implies \int_a^\infty g\ dx$ is convergent.
2. $\int_a^\infty g\ dx$ is divergent $\implies \int_a^\infty f\ dx$ is divergent.

5.2.3 Theorem

$$\int_1^\infty \frac{1}{x^p}\ dx \quad \begin{cases} \text{converges if } p > 1 \\ \text{diverges if } p \leq 1 \end{cases}$$

5.2.4 Theorem

$$\int_0^1 \log x\ dx \quad \text{is convergent}$$

5.2.5 Theorem: Integral Mean Value Theorem

For functions f and g continuous on $[a, b]$ with $g(x) \geq 0$ for $x \in [a, b]$, $\exists\, x_0 \in (a, b)$ such that

$$\int_a^b fg\ dx = f(x_0)\int_a^b g\ dx.$$

6 Applications of Integration

6.0.1 Theorem

The length of a curve between $x = a$ and $x = b$ is

$$L = \int_a^b ds = \int_a^b \sqrt{1 + \left(\frac{df}{dx}\right)^2}\, dx = \int_{t_a}^{t_b} \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt$$

where $x(t)$ and $y(t)$ are a parametrisation of the curve given by f.

6.0.2 Theorem

Let P_x be a family of planes with common axis x, where a solid V lies between planes P_a and P_b. If a function for the area of V cut by P_x is $A(x)$ then the volume is

$$V = \int_a^b A\ dx.$$

6.0.3 Theorem

The solid produced by revolving the area under a curve $f(x)$ between $x = a$ and $x = b$ about the x-axis is

$$V = \pi \int_a^b f^2\ dx.$$

The solid produced by revolving the area under a curve $f(x)$ between $x = a$ and $x = b$ about the y-axis is

$$V = 2\pi \int_a^b xf\ dx.$$

6.0.4 Theorem

The surface area produce by revolving the curve $f(x)$ between $x = a$ and $x = b$ about the x-axis is

$$S = 2\pi \int_a^b f\ ds = 2\pi \int_a^b f \sqrt{1 + \left(\frac{df}{dx}\right)^2}\ dx.$$

6.0.5 Theorem

In polar coordinates, where r is a function of θ, the region bounded by a curve and the half-lines $\theta = \alpha$ and $\theta = \beta$ is

$$R = \frac{1}{2} \int_\alpha^\beta r^2\ d\theta.$$

6.0.6 Theorem

The length of a curve between $\theta = \alpha$ and $\theta = \beta$ is

$$L = \int_\alpha^\beta \sqrt{\left(\frac{dr}{d\theta}\right)^2 + r^2}\ d\theta.$$

6.1 Centre of Mass

6.1.1 One-Dimensional Discrete Centre of Mass

For a set of discrete masses with coordinates x_i and masses m_i, the centre of mass $\bar{x}$ must be such that there is a zero total moment, so it satisfies

$$\sum_i m_i(\bar{x} - x_i) = 0 \qquad \Longleftrightarrow \qquad \bar{x} = \frac{\sum_i m_i x_i}{\sum_i m_i}.$$

6.1.2 Two-Dimensional Discrete Centre of Mass

For a set of discrete masses with coordinates (x_i, y_i) and masses m_i, the centre of mass $(\bar{x}, \bar{y})$ must be such that there is a zero total moment in both dimensions, so it satisfies

$$\sum_i m_i(\bar{x} - x_i) = 0 \quad \text{and} \quad \sum_i m_i(\bar{y} - y_i) = 0$$

$$\iff \quad (\bar{x}, \bar{y}) = \left(\frac{\sum_i m_i x_i}{\sum_i m_i}, \frac{\sum_i m_i y_i}{\sum_i m_i} \right).$$

6.1.3 Two-Dimensional Continuous Centre of Mass

For a continuous region bounded by a curve f and $x = a$ and $x = b$, take a partition $x_i = a + ih$, $i = 0, ..., n$ (where $h = \frac{b-a}{n}$). For a rectangle R_i, centre of mass $(x_i^*, y_i^*) = (\frac{1}{2}(x_i + x_{i-1}), \frac{1}{2}f(x_i))$, the moments of R_i about the x and y-axes are

$$\begin{aligned} M_x(R_i) &= m_i \cdot d_{x\text{-axis}} = \rho f(x_i^*)\Delta x \cdot \frac{1}{2} f(x_i^*) \\ \text{and} \quad M_y(R_i) &= m_i \cdot d_{y\text{-axis}} = \rho f(x_i^*)\Delta x \cdot x_i^*. \end{aligned}$$

Therefore, in the limiting partition, the moments of the union of the rectangles about the x and y-axes are

$$\begin{aligned} M_x &= \lim_{n\to\infty} \sum_{i=1}^{n} \rho \cdot \frac{1}{2} f(x_i^*)^2 \Delta x = \rho \cdot \frac{1}{2} \int_a^b f^2 \; dx \\ \text{and} \quad M_y &= \lim_{n\to\infty} \sum_{i=1}^{n} \rho \cdot x_i^* f(x_i^*)\Delta x = \rho \cdot \int_a^b xf \; dx. \end{aligned}$$

Note that the total mass of the region is $\rho \int_a^b f \; dx$, and so

$$\bar{x} = \frac{\int_a^b xf \; dx}{\int_a^b f \; dx} \qquad \bar{y} = \frac{\frac{1}{2}\int_a^b f^2 \; dx}{\int_a^b f \; dx}$$

(so we can assume $\rho = 1$ without loss of generality).

6.1.4 Theorem: Theorem of Pappus

Let R be a region entirely to one side of a line l. Let A be the area of R, V the volume of revolution of R about l and d the circumference of the circle of revolution of the centre of mass of R about l.

$$V = Ad.$$

7 Series, Power Series and Taylor's Theorem

7.0.1 Definition

For a sequence of real numbers $(a_n)_{n\geq 1}$, the sequence of N^{th} partial sums (series) is $(S_N)_{N\geq 1}$, where

$$S_N = \sum_{n=1}^{N} a_n.$$

7.0.2 Definition

If $S_N \to S$ as $n \to \infty$, the series (S_N) converges to

$$S = \lim_{N\to\infty} \sum_{n=1}^{N} a_n := \sum_{n=1}^{\infty} a_n.$$

If $S_N \to \infty$ as $n \to \infty$, the series diverges.

7.0.3 Definition

A geometric series is a sequence of partial sums of a sequence of the form $a_n = a_0 r^{n-1}$.

7.0.4 Theorem

For a geometric series,

$$S_N = \frac{a_0(1 - r^N)}{1 - r}$$

7.0.5 Corollary

For a geometric series with $|r| < 1$,

$$S_\infty = \frac{a_0}{1 - r}$$

and so the series converges. If $|r| \geq 1$, the series diverges.

7.1 Series of Positive or Negative Terms

7.1.1 Theorem

If a sequence of partial sums (S_N) has only positive terms and is bounded above, the series converges. If the sequence of partial sums is not bounded above, the series diverges (analogous theorem for negative terms).

7.1.2 Theorem

$$\sum_{n=1}^{\infty} \frac{1}{n} \to \infty.$$

7.1.3 Lemma

$$\sum_{n=1}^{\infty} a_n \text{ and } \sum_{n=1}^{\infty} b_n \text{ converge} \implies \sum_{n=1}^{\infty} (\alpha a_n + \beta b_n) \text{ converges.}$$

7.1.4 Theorem

$$\sum_{n=1}^{\infty} a_n \text{ converges} \implies a_n \to 0 \text{ as } n \to \infty.$$

7.1.5 Corollary

$$\sum_{n=1}^{\infty} a_n \text{ converges} \implies \sum_{n=N}^{\infty} a_n \to 0 \text{ as } N \to \infty.$$

7.2 Series and Cauchy Sequences

7.2.1 Definition

A sequence (a_n) is Cauchy if

$$\forall \varepsilon, \ \exists N \in \mathbb{N}, \ m, n \geq N \implies |a_m - a_n| < \varepsilon.$$

The convergent property of Cauchy sequences is necessary for the next few theorems (see Required for: Analysis I).

7.2.2 Theorem: Alternating Series Test

For a decreasing sequence of positive numbers (a_n), where $a_n \to 0$ as $n \to \infty$,

$$\sum_{n=1}^{\infty} (-1)^{n-1} a_n \text{ converges.}$$

7.2.3 Theorem: Comparison Test

If $|b_n| \leq a_n \; \forall \, n$, then

1. $\sum_{n=1}^{\infty} a_n$ converges $\implies \sum_{n=1}^{\infty} b_n$ converges.
2. $\sum_{n=1}^{\infty} b_n$ diverges $\implies \sum_{n=1}^{\infty} a_n$ diverges.

7.3 Absolute and Conditional Convergence

7.3.1 Definition

$\sum_{n=1}^{\infty} a_n$ is absolutely convergent if $\sum_{n=1}^{\infty} |a_n|$ is convergent. A series which is convergent but not absolutely convergent is conditionally convergent.

7.3.2 Theorem

Every absolutely convergent series is convergent.

7.4 The Integral Test

7.4.1 Theorem

For a function f, defined $\forall \, x \geq 1$ and both positive and decreasing,

$$\sum_{n=1}^{\infty} f(n) \text{ converges } \iff \int_{1}^{\infty} f \; dx \text{ converges.}$$

7.4.2 Theorem

$$\sum_{n=1}^{\infty} \frac{1}{n^p} \quad \begin{cases} \text{converges if } p > 1 \\ \text{diverges if } p \leq 1 \end{cases}$$

7.5 The Ratio Test

7.5.1 Theorem

For a sequence (a_n), let $\lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| = L$.

1. $L < 1 \implies \sum_{n=1}^{\infty} a_n$ is absolutely convergent.
2. $L > 1 \implies \sum_{n=1}^{\infty} a_n$ is divergent.
3. $L = 1$ is an inconclusive test.

7.6 The Root Test

7.6.1 Theorem

For a sequence (a_n), let $\lim_{n \to \infty} |a_n|^{\frac{1}{n}} = L$.

1. $L < 1 \implies \sum_{n=1}^{\infty} a_n$ is absolutely convergent.
2. $L > 1 \implies \sum_{n=1}^{\infty} a_n$ is divergent.
3. $L = 1$ is an inconclusive test.

7.7 Power Series

7.7.1 Definition

For a sequence of real numbers $(a_n)_{n\geq 0}$, the power series for $(a_n)_{n\geq 0}$ is the function

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

and we denote the N^{th} partial sum

$$f_N(x) = \sum_{n=0}^{N} a_n x^n$$

which is a degree-N polynomial.

7.7.2 Definition

The radius of convergence of a power series is R such that the series converges for $|x| < R$. (Note that the power series could be centered about a point x_0, in which case R is such that we have convergence for $|x - x_0| < R$).

7.7.3 Theorem

For a power series $f(x) = \sum_{n=0}^{\infty} a_n x^n$ with radius of convergence R, if $|x| < R$ (the series converges), then

1. $f'(x)$ exists and is equal to $\sum_{n=1}^{\infty} n a_n x^{n-1}$.
2. $\int f\ dx$ exists and is equal to $\sum_{n=0}^{\infty} \frac{1}{n+1} a_n x^{n+1}$ (+ a constant).

7.7.4 Theorem: Algebraic Operations

Suppose $f(x) = \sum_{n=0}^{\infty} a_n x^n$ and $g(x) = \sum_{n=0}^{\infty} b_n x^n$ have radii of convergence R_f and R_g respectively.

1. Similarly to Lemma 7.1.3, $\sum_{n=0}^{\infty} (\alpha a_n + \beta b_n) x^n$ has radius of convergence $R = \min\{R_f, R_g\}$.
2. $f(x)g(x) = \left(\sum_{n=0}^{\infty} a_n x^n\right)\left(\sum_{n=0}^{\infty} b_n x^n\right) = \sum_{n=0}^{\infty} \left(\sum_{m=0}^{n} a_m b_{n-m}\right) x^n$.

7.8 Taylor Series

7.8.1 Note: Obtaining the Maclaurin Series

Consider a power series $f(x)$. Over values of x for which the series is convergent, we can differentiate the series term by term, and obtain that:

$$\begin{aligned} f(0) &= a_0 \\ f'(0) &= a_1 \\ f''(0) &= 2a_2 \\ &\vdots \\ f^{(k)}(0) &= k! a_k. \end{aligned}$$

So we can write

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k.$$

7.8.2 Theorem: Taylor's Theorem

Let f be a function defined on $[x_0, x]$ and continuously differentiable $(n+1)$ times on the interval, then

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + ... + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + R_n$$

(where $R_n = \int_{x_0}^{x} \frac{(x-t)^n}{n!} f^{(n+1)}(t)\ dt$ is the remainder term) for which $n \to \infty$ gives a convergent sum provided $\lim_{n\to\infty} R_n = 0$. Note that I have assumed $x_0 < x$, but the proof is easily adapted if this is not the case.

7.8.3 Note

By the integral mean value theorem (5.2.5), $\exists\, c \in [x_0, x]$ such that

$$\int_{x_0}^{x} \frac{(x-t)^n}{n!} f^{(n+1)}(t)\ dt = f^{(n+1)}(c) \int_{x_0}^{x} \frac{(x-t)^n}{n!}\ dt = \frac{f^{(n+1)}(c)}{(n+1)!}(x-x_0)^{n+1}$$

so we have a form for R_n which is similar to the other terms of the Taylor series.

7.8.4 Theorem: L'Hôpital's Rule

L'Hôpital's Rule can now be proved with Taylor's theorem. Note: Many limits involving fractions are easier to evaluate with Taylor Expansions.

8 Trigonometric Series and Fourier Series

8.1 Prerequisites

8.1.1 Definition: Orthogonal Function Spaces

Let $S = \{\phi_0, \phi_1, ..., \phi_n\}$ (a set of functions) where all ϕ_i are Riemann integrable on $[a, b]$. Define the inner product

$$(\phi_m \cdot \phi_n) = \int_a^b \phi_m(x)\phi_n(x)\ dx.$$

and define ϕ_m and ϕ_n to be orthogonal with respect to $[a, b]$ if $(\phi_m \cdot \phi_n) = 0$. Define the norm $||\phi_m|| = (\phi_m \cdot \phi_m)^{\frac{1}{2}}$ (provided $\int_a^b \phi_m(x)\phi_m(x)\ dx \geq 0$) and define ϕ_i to be normal if $||\phi_i|| = 1$. If $(\phi_m \cdot \phi_n) = 0\ \forall\, m, n$ and $||\phi_i|| = 1\ \forall\, i$, S is orthonormal with respect to $[a, b]$.

8.1.2 Definition: The Trigonometric Orthonormal Set

$\left\{\phi \mid \phi_0 = \frac{1}{\sqrt{2\pi}},\ \phi_{2n-1} = \frac{\cos nx}{\sqrt{\pi}},\ \phi_{2n} = \frac{\sin nx}{\sqrt{\pi}}\right\}$ is an orthonormal set with respect to $[-\pi, \pi]$ since

$$\int_{-\pi}^{\pi} \frac{\sin px}{\sqrt{\pi}} \cdot \frac{\sin qx}{\sqrt{\pi}}\ dx = \int_{-\pi}^{\pi} \frac{\cos px}{\sqrt{\pi}} \cdot \frac{\cos qx}{\sqrt{\pi}}\ dx = \begin{cases} 0, & \text{if } p \neq q \\ 1, & \text{if } p = q \neq 0 \end{cases}$$

$$\text{and} \quad \int_{-\pi}^{\pi} \frac{\sin px}{\sqrt{\pi}} \cdot \frac{\cos qx}{\sqrt{\pi}}\ dx = 0 \quad \forall\, p, q.$$

8.1.3 Definition

A real function f is periodic with period $P \neq 0 \iff f(x) = f(x + P)\ \forall\, x \in \mathbb{R}$.

8.1.4 Definition

A periodic extension of a real function f defined over a chosen interval of length P is such that

$$f(x \pm nP) := f(x)\ \forall\, n \in \mathbb{Z}\ \forall\, x \in \text{the chosen interval.}$$

At any given point of discontinuity $x = \xi$, define

$$f(\xi) = \frac{1}{2}\Big(f(\xi-) + f(\xi+)\Big)$$

(where $\xi- = \lim_{x\to\xi-} f(x)$ and $\xi+ = \lim_{x\to\xi+} f(x)$).

8.1.5 Lemma

Let f be a function with period P.

$$\int_a^b f\ dx = \int_{a\pm P}^{b\pm P} f\ dx.$$

8.2 Trigonometric Polynomials

Define the trigonometric polynomial

$$S_N(x) = \frac{1}{2}a_0 + \sum_{n=1}^{N} a_n \cos\omega nx + b_n \sin\omega nx$$

so that a sum of trigonometric functions with periods $\frac{2\pi}{\omega}$ can now be represented by a list of $2N+1$ coefficients a_0, a_n and b_n. In complex form (with $\omega = 1$),

$$\begin{aligned}
&\frac{1}{2}a_0 + \sum_{n=1}^{N} a_n \cos nx + b_n \sin nx \\
&= \frac{1}{2}a_0 + \sum_{n=1}^{N} a_n \left(\frac{e^{inx} + e^{-inx}}{2}\right) + b_n \left(\frac{e^{inx} - e^{-inx}}{2i}\right) \\
&= \frac{1}{2}a_0 + \sum_{n=1}^{N} \frac{e^{inx}}{2}(a_n - ib_n) + \frac{e^{-inx}}{2}(a_n + ib_n) \\
&:= \sum_{n=-N}^{N} \gamma_n e^{inx}
\end{aligned}$$

where $\gamma_0 := \frac{1}{2}a_0$, $\gamma_{-n} := \frac{1}{2}(a_n + ib_n)$ and $\gamma_n := {\gamma_{-n}}^* = \frac{1}{2}(a_n - ib_n)$. Note that the sum is still real, necessarily by the fact that $\gamma_n = {\gamma_{-n}}^* \iff S_N(x) = S_N(x)^*$.

8.3 Fourier Series

8.3.1 Note: Obtaining the Trigonometric Coefficients

For a function f, assume $\exists\ a_0,\ a_n,\ b_n$ such that $f(x) = S_N(x)$. By integrating on $[-\pi, \pi]$, we obtain by trigonometric orthogonality that

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\ dt$$

and, multiplying by $\cos nx$ or $\sin nx$, that

$$\begin{aligned}
\int_{-\pi}^{\pi} \cos(nx)f(x)\ dx = \int_{-\pi}^{\pi} a_n \cos^2(nx)\ dx &\iff a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos(nt)f(t)\ dt \\
\int_{-\pi}^{\pi} \sin(nx)f(x)\ dx = \int_{-\pi}^{\pi} b_n \sin^2(nx)\ dx &\iff b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \sin(nt)f(t)\ dt
\end{aligned}$$

8.3.2 Lemma

$$\frac{1}{2} + \cos x + \cos 2x + ... + \cos nx = \frac{\sin(n+\frac{1}{2})x}{2\sin\frac{x}{2}} := c_n(x)$$

provided $c_n(x)\Big|_{x=2k\pi} := n + \frac{1}{2}$.

8.3.3 Lemma: Riemann-Lebesgue Lemma

For a function f, integrable and differentiable on $[a, b]$,

$$\lim_{\lambda\to\infty}\int_a^b f(x)\sin\lambda x\ dx = 0.$$

8.3.4 Lemma

$$\int_0^\infty \frac{\sin x}{x}\ dx = \frac{\pi}{2}.$$

8.3.5 Theorem: Convergence of Fourier Series

For a function f which is piecewise continuous, has piecewise continuous first and second derivatives and has period 2π, the Fourier series $\frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx)$, with a_0, a_n, and b_n defined as in 8.3.1, converges to $f(x)$ and to $f(\xi)$ at points of discontinuity.

8.3.6 Theorem: Parseval's Theorem

Let f be a function with Fourier series $\frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx)$.

$$\frac{1}{\pi}\int_{-\pi}^{\pi}\Big(f(x)\Big)^2\,dx = \frac{1}{2}a_0^2 + \sum_{n=1}^{\infty}(a_n^2 + b_n^2).$$

8.3.7 Note: Fourier Series in the Complex Form

By theorem 8.3.5, $f(x)$ is also the limit of $\sum_{n=-\infty}^{\infty}\gamma_n e^{inx}$, and the γ_n are equal to

$$\gamma_{-n} = \frac{1}{2}(a_n + ib_n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\Big(\cos nt + i\sin nt\Big)\,dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{int}\,dt$$
$$\gamma_n = \gamma_{-n}^* = \frac{1}{2}(a_n - ib_n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\Big(\cos nt - i\sin nt\Big)\,dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\,dt$$

so in the complex form, only one formula is required for γ_n and we have, by the theorem,

$$f(x) = \sum_{n=-\infty}^{\infty}\gamma_n e^{inx}$$

$$\text{where}\quad \gamma_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\,dt.$$

8.4 Fourier Transforms

The orthonormal set in 8.1.2 can be adapted for $2L$-periodic functions:
$\Big\{\phi \mid \phi_0 = \frac{1}{\sqrt{2L}},\ \phi_{2n-1} = \frac{1}{\sqrt{L}}\cos(\frac{n\pi x}{L}),\ \phi_{2n} = \frac{1}{\sqrt{L}}\sin(\frac{n\pi x}{L})\Big\}$ is an orthonormal set with respect to $[-L, L]$ and, setting $\omega = \frac{\pi}{L}$, $f(x)$ is equal to the Fourier series

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

$$\text{where}\quad a_n = \frac{1}{L}\int_{-L}^{L}\cos\left(\frac{n\pi t}{L}\right)f(t)\,dt \quad\text{and}\quad b_n = \frac{1}{L}\int_{-\pi}^{\pi}\sin\left(\frac{n\pi t}{L}\right)f(t)\,dt$$

or, in complex form, the Fourier series

$$f(x) = \sum_{n=-\infty}^{\infty}\gamma_n e^{i\frac{n\pi x}{L}}$$

$$\text{where}\quad \gamma_n = \frac{1}{2L}\int_{-L}^{L} f(t)e^{-i\frac{n\pi t}{L}}\,dt.$$

8.4.1 $L \to \infty$

As $L \to \infty$, $h := \frac{\pi}{L} \to 0$, so for

$$\sum_{n=-\infty}^{\infty}\gamma_n e^{inhx} = \sum_{n=-\infty}^{\infty}\left(\frac{h}{2\pi}\int_{-L}^{L} f(t)e^{-inht}\,dt\right)e^{inhx},$$

an appropriate limit is $\omega_n := nh$ which is the frequency. h can be thought of as $\omega_{n+1} - \omega_n = \delta\omega$, so $\sum_{-\infty}^{\infty} G(\omega_n)h$ will behave as the Riemann integral $\int_{-\infty}^{\infty} G(\omega_n)\,d\omega$. So in the limit,

$$\sum_{n=-\infty}^{\infty}\left(\frac{h}{2\pi}\int_{-\infty}^{\infty} f(t)e^{-i\omega t}\,dt\right)e^{i\omega x} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(t)e^{-i\omega t}\,dt\right)e^{i\omega x}\,d\omega.$$

8.4.2 Theorem: The Fourier Transform Pair

On infinite domains, a representation of the function f is thus given by

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \widehat{f}(\omega)e^{i\omega x}\ d\omega$$

$$\widehat{f}(\omega) = \int_{-\infty}^{\infty} f(x)e^{-i\omega x}\ dx$$

which is analogous to the Fourier series in 8.3.7. Note that in the above two equations, f can be obtained from $\widehat{f}$ and $\widehat{f}$ can be obtained from f, so we say that 'the Fourier transform of f' is equal to the function $\widehat{f}$, written $\mathcal{F}\{f(x)\} = \widehat{f}(\omega)$, and 'the inverse Fourier transform of $\widehat{f}$' is equal to the function f, written $\mathcal{F}^{-1}\{\widehat{f}(\omega)\} = f(x)$. In the second line, x has replaced t so as to resemble a transform of $f(x)$, rather than just f.

8.4.3 Definition

The Fourier cosine transform of $f(x)$ is $\mathcal{F}_c\{f(x)\} := \widehat{f}_c(\omega) := \int_0^\infty f(t)\cos\omega t\ dt$.

8.4.4 Lemma

If f is even about 0, $\widehat{f}(\omega) = 2\widehat{f}_c(\omega)$.

8.4.5 Definition

The Fourier sine transform of $f(x)$ is $\mathcal{F}_s\{f(x)\} := \widehat{f}_s(\omega) := \int_0^\infty f(t)\sin\omega t\ dt$.

8.4.6 Lemma

If f is odd about 0, $\widehat{f}(\omega) = -2i\widehat{f}_s(\omega)$.

8.4.7 Theorem

Using the above results, the Fourier cosine transform of even f and Fourier sine transform of odd f are given similarly to 8.4.2, with their inversion formulae, by

$$\widehat{f}_c(\omega) = \int_0^\infty f(t)\cos\omega t\ dt \qquad \widehat{f}_s(\omega) = \int_0^\infty f(t)\sin\omega t\ dt$$

$$f(x) = \frac{2}{\pi}\int_0^\infty \widehat{f}_c(\omega)\cos\omega x\ d\omega \qquad f(x) = \frac{2}{\pi}\int_0^\infty \widehat{f}_s(\omega)\sin\omega x\ d\omega$$

8.4.8 Theorem: Linearity of the Fourier Transform

1. $\mathcal{F}\{af(x) + bg(x)\} = a\widehat{f}(\omega) + b\widehat{g}(\omega)\ \forall\, a, b \in \mathbb{R}$.
2. $\mathcal{F}^{-1}\{a\widehat{f}(\omega) + b\widehat{g}(\omega)\} = af(x) + bg(x)\ \forall\, a, b \in \mathbb{R}$.

8.4.9 Theorem

$\mathcal{F}\{f(ax)\} = \frac{1}{a}\widehat{f}(\frac{\omega}{a})\ \forall\, a > 0$.

8.4.10 Theorem

$\mathcal{F}\{f(-x)\} = \widehat{f}(-\omega)$.

8.4.11 Theorem

1. $\mathcal{F}\{f(x - x_0)\} = e^{-i\omega x_0}\widehat{f}(\omega)$ (shift in domain space).
2. $\mathcal{F}\{e^{i\omega_0 x}f(x)\} = \widehat{f}(\omega - \omega_0)$ (shift in transform space).

8.4.12 Theorem: Symmetry Formula

$\mathcal{F}\{\widehat{f}(x)\} = 2\pi f(-\omega)$.

8.4.13 Theorem

$\mathcal{F}\left\{\frac{d^n f}{dx^n}\right\} = (i\omega)^n \widehat{f}(\omega)$ (assuming that all derivatives of $f \to 0$ as $x \to \pm\infty$).

8.4.14 Theorem

$\mathcal{F}\{xf(x)\} = i\widehat{f}\,'(\omega)$.

8.4.15 Theorem

Assuming that all derivatives of $f \to 0$ as $x \to \pm\infty$,

1. $\mathcal{F}_c\{f'(x)\} = -f(0) + \omega \widehat{f}_s(\omega)$.
2. $\mathcal{F}_s\{f'(x)\} = -\omega \widehat{f}_c(\omega)$.
3. $\mathcal{F}_c\{f''(x)\} = -f'(0) + \omega^2 \widehat{f}_c(\omega)$.
4. $\mathcal{F}_s\{f''(x)\} = \omega f(0) - \omega^2 \widehat{f}_s(\omega)$.

8.4.16 Theorem

Let $f^*(x)$ be the conjugate of the complex value $f(x)$. $\mathcal{F}\{f^*(x)\} = \widehat{f}^*(-\omega)$.

8.4.17 Definition

The convolution of two functions f and g is

$$f(x) * g(x) = \int_{-\infty}^{\infty} f(x-u)g(u)\ du.$$

8.4.18 Theorem: Convolution Theorem

$$\mathcal{F}\{f(x) * g(x)\} = \widehat{f}(\omega)\widehat{g}(\omega).$$

8.4.19 Theorem: Energy Theorem

For a real valued function f,

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \left|\widehat{f}(\omega)\right|^2\ d\omega = \int_{-\infty}^{\infty} \left(f(x)\right)^2\ dx.$$

8.5 The Dirac Delta Function

8.5.1 Definition

For the purpose of defining thr Dirac delta function, consider the function $f_k(x) = \begin{cases} \frac{k}{2}, & |x| < \frac{1}{k} \\ 0, & |x| > \frac{1}{k} \end{cases}$ and note that $\int_{-\infty}^{\infty} f_k(x)\ dx = 1$.

8.5.2 Definition

The Dirac delta function is

$$\delta(x) = \lim_{k\to\infty} f_k(x).$$

8.5.3 Theorem: Sifting Property

For any continuous function g, defined over $\mathbb{R}$,

$$\int_{-\infty}^{\infty} g(x-a)\delta(x)\ dx = g(a).$$

8.5.4 Theorem

$\mathcal{F}\{\delta(x)\} = 1$ and so $\delta(x)$ can be written $\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\pm i\omega x}\ d\omega$.

9 Ordinary Differential Equations

9.0.1 Definition

A function f is differentiable to order $k \iff$

$$\exists \lim_{h \to 0} \frac{f^{(i)}(x+h) - f^{(i)}(x)}{h} = f^{(i+1)}(x) \quad \forall\, i \leq k-1.$$

9.0.2 Definition

'Ordinary' refers to the presence of unknown functions in only one independent variable (here x); an ordinary differential equation is of the form

$$G\left(x, f, \frac{df}{dx}, ..., \frac{d^k f}{dx^k}\right) = 0.$$

The order of the ODE is the order of the highest derivative of f present.
The degree of the ODE is the exponent to which this highest derivative is raised.
An ODE is linear if G is a linear function of f and <u>all</u> its derivatives, i.e.

$$G\left(x, f, \frac{df}{dx}, ..., \frac{d^k f}{dx^k}\right) = g_0(x)f + g_1(x)\frac{df}{dx} + ... + g_k(x)\frac{d^k f}{dx^k} + g(x) = 0$$

where $g_0, ..., g_k$ and g are arbitrary and not necessarily linear functions, and f is the unknown function of x.

9.0.3 Definition

$$G\left(x, f, \frac{df}{dx}, ..., \frac{d^k f}{dx^k}\right) = 0 \quad \text{is the implicit form of an ODE (as above).}$$

$$\frac{d^k f}{dx^k} = F\left(x, f, \frac{df}{dx}, ..., \frac{d^{k-1} f}{dx^{k-1}}\right) \quad \text{is the explicit form of an ODE.}$$

9.0.4 Definition

To solve an ODE is to find f such that the ODE is satisfied over its domain. f_{PI} is a particular integral or particular solution of an ODE if $G\left(x, f_{PI}, \frac{df_{PI}}{dx}, ..., \frac{d^k f_{PI}}{dx^k}\right) = 0$.
f_{GS} is a general solution of an ODE of order k if $f_{GS}(x; c_1, ..., c_k),\ c_1, ..., c_k \in \mathbb{R}$ is a family of solutions for which $G\left(x, f_{GS}, \frac{df_{GS}}{dx}, ..., \frac{d^k f_{GS}}{dx^k}\right) = 0$. The constants of integration $\{c_i\}$ are fixed by conditions, giving particular solutions.

9.1 First and Second Order ODEs: Specific Cases

First Order ODEs

$$\text{Implicit form:} \quad G\left(x, y, \tfrac{dy}{dx}\right) = 0, \qquad \text{Explicit form:} \quad \tfrac{dy}{dx} = F(x, y).$$

9.1.1 Separable First Order ODEs

$$\frac{dy}{dx} = F_1(y)F_2(x) \iff \int \frac{dy}{F_1(y)} = \int F_2(x)\ dx + c_1.$$

9.1.2 Linear First Order ODEs

$$\begin{aligned} & F_1(x)\frac{dy}{dx} + F_2(x)y - F_3(x) = 0 \\ \iff & I(x)\frac{dy}{dx} + I(x)\frac{F_2(x)}{F_1(x)}y = I(x)\frac{F_3(x)}{F_1(x)} \iff \frac{d}{dx}\Big[I(x)y\Big] = I(x)\frac{F_3(x)}{F_1(x)} \\ \iff & y = \frac{1}{I(x)}\left(\int I(x)\frac{F_3(x)}{F_1(x)}\ dx + c_1\right). \end{aligned}$$

Letting $p = \frac{F_2}{F_1}$, The result above requires that

$$\frac{d[Iy]}{dx} = I\frac{dy}{dx} + Ipy \iff \frac{dI}{dx} = Ip \overset{\text{(separable)}}{\iff} I(x) = Ae^{\int p\ dx}.$$

9.1.3 Homogeneous First Order ODEs

$$\begin{aligned}
&\frac{dy}{dx} = F\left(\frac{y}{x}\right)\\
\iff u + x\frac{du}{dx} = F(u) \iff &\int \frac{du}{F(u) - u} = \int \frac{dx}{x} + c_1 = \ln|x| + c_1
\end{aligned}$$

(where $u = \frac{y}{x} \implies \frac{dy}{dx} = u + x\frac{du}{dx}$). Once u_{GS} is found, y_{GS} can be found by $y(x) = xu(x)$.

9.1.4 Bernoulli ODEs

$$\begin{aligned}
&\frac{dy}{dx} + F_1(x)y = F_2(x)y^n\\
\iff &y^{-n}\frac{dy}{dx} + F_1(x)y^{1-n} = F_2(x) \iff \frac{du}{dx} + (1-n)F_1(x)u = (1-n)F_2(x)
\end{aligned}$$

(where $u = y^{1-n} \implies \frac{du}{dx} = (1-n)y^{-n}\frac{dy}{dx}$). If the first derivative has an arbitrary coefficient function, it can be manipulated as in 9.1.2. u_{GS} can be found from this linear first order ODE and y_{GS} can then be found by $y(x) = \big(u(x)\big)^{\frac{1}{1-n}}$.

Second Order ODEs

Implicit form: $G\left(x, y, \frac{dy}{dx}, \frac{d^2y}{dx^2}\right) = 0$, Explicit form: $\frac{d^2y}{dx^2} = F\left(x, y, \frac{dy}{dx}\right)$.

9.1.5 F Depending Only on x

$$\begin{aligned}
&\frac{d^2y}{dx^2} = F(x)\\
\iff &\frac{dy}{dx} = \int F(x)\ dx + c_1 \iff y = \int\left(\int F(x)\ dx + c_1\right)\ dx + c_2.
\end{aligned}$$

9.1.6 F Depending Only on x and $\frac{dy}{dx}$

ODEs of the form

$$\frac{d^2y}{dx^2} = F\left(x, \frac{dy}{dx}\right)$$

often require substitution, for example:

9.1.7 Definition

The radius of curvature of a curve $y(x)$ is $\dfrac{\left(1 + \left(\frac{dy}{dx}\right)^2\right)^{\frac{3}{2}}}{\frac{d^2y}{dx^2}}$.

9.1.8 Theorem

The family of curves with constant radius of curvature R is a set of circles with radius R.

9.1.9 F Depending Only on y

$$\begin{aligned}
&\frac{d^2y}{dx^2} = F(y)\\
\iff &\frac{du}{dx} = F(y) \iff u\frac{du}{dy} = F(y) \overset{\text{(separable)}}{\iff} \frac{1}{2}u^2 = \frac{1}{2}\left(\frac{dy}{dx}\right)^2 = \int F(y)\ dy + c_1\\
\overset{\text{(separable)}}{\iff} &\int \frac{dy}{\pm\sqrt{2\left(\int F(y)\ dy + c_1\right)}} = \int dx + c_2
\end{aligned}$$

(where $u = \frac{dy}{dx} \implies \frac{d^2y}{dx^2} = \frac{du}{dx} = \frac{du}{dy}\frac{dy}{dx} = u\frac{du}{dy}$).

9.1.10 F Depending Only on y and $\frac{dy}{dx}$

$$\frac{d^2y}{dx^2} = F\left(y, \frac{dy}{dx}\right) \iff \frac{d}{dy}\left[\frac{1}{2}u^2\right] = F(y,u)$$

(where $u = \frac{dy}{dx} \implies \frac{d^2y}{dx^2} = \frac{du}{dx} = \frac{du}{dy}\frac{dy}{dx} = u\frac{du}{dy} = \frac{d}{dy}\left[\frac{1}{2}u^2\right]$). If u_{GS} is a solution to this first order ODE then y_{GS} can be found from the first order ODE $\frac{dy}{dx} = u_{GS}(y, c_1)$.

9.2 Linear ODEs of Order k (and Specific Cases)

As in 9.0.2, the general form of a linear ODE with unknown function f is

$$g_0(x)f + g_1(x)\frac{df}{dx} + ... + g_k(x)\frac{d^kf}{dx^k} + g(x) = 0.$$

9.2.1 Definition

In the above form, a linear ODE is homogeneous $\iff$ $g(x) = 0$ and it is inhomogeneous otherwise.

9.2.2 Definition

An operator acts on a function. A linear operator is an operator (e.g. $\mathcal{D}[f] \equiv \frac{d}{dx}[f]$) such that $\mathcal{D}[\lambda_1 f_1 + \lambda_2 f_2] = \lambda_1\mathcal{D}[f_1] + \lambda_2\mathcal{D}[f_2]$.

9.2.3 Note

Linear ODEs can be associated to the linear operator $\mathcal{L}[f] = \sum_{i=0}^{k} g_i(x)\mathcal{D}^i[f]$, such that the ODE is then represented by $\mathcal{L}[f] = g(x)$, and the homogeneous case is $\mathcal{L}[f] = 0$. We then have also that $\mathcal{L}[\lambda_1 f_1 + \lambda_2 f_2] = \lambda_1\mathcal{L}[f_1] + \lambda_2\mathcal{L}[f_2]$.

9.2.4 Definition

A set of functions $\{f_i\}_{i=1}^k$ is linearly independent $\iff$

$$\alpha_1 f_1 + ... + \alpha_k f_k = 0 \implies \alpha_1 = ... = \alpha_k = 0.$$

9.2.5 Note

We can think of functions as vectors and of the set of solutions to a homogeneous linear ODE of order k as a k-dimensional vector space. By the property in 9.2.3, also notice that a linear combination of two solutions to $\mathcal{L}[f] = 0$ is also a solution. The general solution of the homogeneous linear ODE can thus be written

$$f_{GS}^H(x; c_1, ..., c_k) = c_1 f_1 + ... + c_k f_k$$

where $\{f_i\}_{i=1}^k$ is a set of linearly independent solutions which form a basis of the solution space.

9.2.6 Theorem

$\{f_i\}_{i=1}^k$ is linearly independent if the Wronskian $W(\{f_i\}_{i=1}^k)$ (the determinant of the Wronskian matrix $\mathbb{W}$) $\neq 0$:

$$W\left(\{f_i\}_{i=1}^k\right) = \det\left(\mathbb{W}\right) = \det\left([f_j^{(i-1)}]_{k\times k}\right) = \begin{vmatrix} f_1 & f_2 & \cdots & f_k \\ \frac{df_1}{dx} & \frac{df_2}{dx} & & \frac{df_k}{dx} \\ \vdots & & \ddots & \vdots \\ \frac{d^{k-1}f_1}{dx^{k-1}} & \frac{d^{k-1}f_2}{dx^{k-1}} & \cdots & \frac{d^{k-1}f_k}{dx^{k-1}} \end{vmatrix} \neq 0.$$

9.2.7 Theorem

For an inhomogeneous linear ODE $\mathcal{L}[f] = g(x)$, let the general solution to the corresponding homogeneous linear ODE $\mathcal{L}[f] = 0$ be $f_{GS}^H(x; c_1, ..., c_k) = \sum_{i=1}^k c_i f_i$, known as the complementary function f_{CF}, and let f_{PI} be any particular integral such that $\mathcal{L}[f_{PI}] = g(x)$.

$$\left.\begin{array}{l} \mathcal{L}[f_{CF}] = 0 \\ \mathcal{L}[f_{PI}] = g(x) \end{array}\right\} \implies \mathcal{L}[f_{CF} + f_{PI}] = \mathcal{L}[f_{CF}] + \mathcal{L}[f_{PI}] = 0 + g(x)$$

so let the general solution be $f_{GS} = f_{CF} + f_{PI}$, satisfying $\mathcal{L}[f_{GS}] = g(x)$ as above.

9.2.8 Corollary

An inhomogeneous linear ODE of the form $\mathcal{L}[f] = h_1(x) + h_2(x) + h_3(x) + ...$ has a general solution $f_{GS} = f_{CF} + \sum f_{PIi}$ where f_{PIi} are particular integrals satisfying $\mathcal{L}[f_{PIi}] = h_i(x)$.

9.2.9 Note

An ansatz is an educated guess about a desired solution. It may be made based on the predicted effect of the operator $\mathcal{L}$ on a certain function. For example, suppose $\mathcal{L}$ is known to give only constant coefficients. If $g(x) = e^{ax}$, then the particular integral may be of the form $f_{PI} = Ae^{ax}$. f_{PI} can be substituted and the values of the undetermined coefficients found. If f_{PI} is part of the complementary function, it will give $\mathcal{L}[f_{PI}] = 0$, so we might instead try $A(x)e^{ax}$ for example, where an undertermined function of x is to be found. This is called the method of variation of parameters.

9.2.10 First Order Linear ODEs With Constant Coefficients

The solution may follow from the method in 9.1.2.

9.2.11 Second Order Linear ODEs With Constant Coefficients

Consider an ODE of the form

$$\alpha_2 \frac{d^2 f}{dx^2} + \alpha_1 \frac{df}{dx} + \alpha_0 f = g(x).$$

By 9.2.5, the complementary function will be a linear combination of two functions. To find it, we can use the ansatz $f^H = e^{\lambda x}$, which gives $\alpha_2 \lambda^2 + \alpha_1 \lambda + \alpha_0 = 0$, the characteristic equation, giving in turn $\lambda_{1,2} = \frac{-\alpha_1 \pm \sqrt{\alpha_1^2 - 4\alpha_2\alpha_0}}{2\alpha_2}$. We now have a pair of candidate functions $(e^{\lambda_1 x}, e^{\lambda_2 x})$ for a basis of the complementary solution space, and must consider the following cases:

1. If $\lambda_1 \neq \lambda_2$, $W(\{e^{\lambda_1 x}, e^{\lambda_2 x}\}) = e^{(\lambda_1+\lambda_2)x}(\lambda_2 - \lambda_1) \neq 0$, and so $f_{CF} = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}$.

2. If $\lambda_1 = \lambda_2$, then $f_1 = e^{\lambda_1 x}$ is one solution, and we propose $f_2 = A(x)e^{\lambda_1 x}$. Substituting gives $\frac{d^2 A}{dx^2} = 0 \implies A(x) f_1 = (B_1 x + B_2) f_1$. $B_2 f_1$ will be absorbed into $c_1' f_1$, and $W(\{e^{\lambda_1 x}, xe^{\lambda_1 x}\}) = e^{2\lambda_1 x} \neq 0$, and so $f_{CF} = c_1 e^{\lambda_1 x} + c_2 x e^{\lambda_1 x}$.

3. If $\lambda_{1,2} \in \mathbb{C} \setminus \mathbb{R}$, write $\lambda_{1,2} = -\frac{\alpha_1}{2\alpha_2} \pm i\omega$, where $\omega^2 = \left|\frac{\alpha_1^2 - 4\alpha_2\alpha_0}{4\alpha_2^2}\right|$. Using case 1, the complementary function is

$$\begin{aligned} f_{CF} &= e^{-\frac{\alpha_1}{2\alpha_2}x}(c_1' e^{i\omega x} + c_2' e^{-i\omega x}) = e^{-\frac{\alpha_1}{2\alpha_2}x}((c_1' + c_2')\cos\omega x + i(c_1' - c_2')\sin\omega x) \\ &:= e^{\psi x}(c_1 \cos\omega x + c_2 \sin\omega x) \end{aligned}$$

   where ψ and ω are the original real and imaginary coefficients pf $\lambda_{1,2}$, and it is easy to check that the Wronskian of sin and cos is $\neq 0$. While this form is useful, notice that since c_1, c_2 must be real, $c_1' = c_2'^*$, and so their sum and i times their difference can be written $A\cos\phi$ and $A\sin\phi$, giving further the form $f_{CF} = e^{\psi x} A\cos(\omega x - \phi)$.

Once f_{PI} is also found (which may require variation of parameters), it may contain terms similar to terms of f_{CF}, but the coefficients can be combined as usual.

9.2.12 Note

It is important to refer to new, combined constants differently. For example, in case 2 above I let the original coefficient of f_1 be c_1' so I could then let $B_2 + c_1' := c_1$, and in case 3 I wrote the linear combination with c_1', c_2' so their combinations could be c_1, c_2.

9.2.13 k^{th} Order Linear ODEs With Constant Coefficients

For ODEs of the form $\mathcal{L}[f] = \sum_{i=0}^{k} \alpha_i \mathcal{D}^i[f] = g(x)$, to find f_{CF} we can again use the ansatz $f^H = e^{\lambda x}$ which gives the characteristic equation $\sum_{i=0}^{k} \alpha_i \lambda^i = 0$. If the λ_i can be obtained, a candidate basis for the complementary solution space is thus found to be $\{e^{\lambda_1 x}, ..., e^{\lambda_k x}\}$.

1. The Wronskian is $e^{\sum \lambda_i} \prod_{1 \le i < j \le k}(\lambda_j - \lambda_i)$ (Vandermonde), which is clearly $\neq 0 \iff$ all the λ_i are distinct ($\implies$ the candidate basis is linearly independent).

2. If the λ_i are not all distinct, we have seen at a low level that multiplying a repeated function by successive powers of x gives linearly independent functions, and this in fact works generally: suppose λ_r is a root that appears d times, replacing the subset $\{e^{\lambda_r x}, ..., e^{\lambda_r x}\} \to \{e^{\lambda_r x}, ..., x^{d-1}e^{\lambda_r x}\}$ for each repeated root in the candidate basis gives a linearly independent basis.

Finally, the particular integral must be found, and if $g(x)$ is of the form e^{bx}, f_{PI} is known immediately to be $Ax^d e^{bx}$, where b appears d times in the set $\{\lambda_i\}$ (so $b \neq \lambda_i \; \forall \, i \implies f_{PI} = Ax^0 e^{bx} = Ae^{bx}$).

9.2.14 Euler-Cauchy ODEs

Euler-Cauchy equations are a specific example of linear ODEs with non-constant coefficients:

$$\alpha_k x^k \frac{d^k f}{dx^k} + \alpha_{k-1} x^{k-1} \frac{d^{k-1} f}{dx^{k-1}} + ... + \alpha_1 x \frac{df}{dx} + \alpha_0 f = g(x)$$

(i.e. $g_i(x) = \alpha_i x^i$). The change of variable $x = e^z$ transforms an Euler-Cauchy equation into a linear ODE with constant coefficients.

9.3 Systems of Ordinary Differential Equations

9.3.1 Definition

A system of ODEs is of the form

$$G_1\left(x, f_1, ..., f_n, \frac{df_1}{dx}, ..., \frac{df_n}{dx}, ..., \frac{d^k f_1}{dx^k}, ..., \frac{d^k f_n}{dx^k}\right) = 0$$

$$\vdots$$

$$G_n\left(x, f_1, ..., f_n, \frac{df_1}{dx}, ..., \frac{df_n}{dx}, ..., \frac{d^k f_1}{dx^k}, ..., \frac{d^k f_n}{dx^k}\right) = 0.$$

The system is still ordinary because the unknown functions $f_1, ..., f_n$ are in only one independent variable x. In single ODEs, the unknown function was $f : \mathbb{R} \longrightarrow \mathbb{R}$. To solve the system will be to find

$$\vec{f} = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}$$

($\vec{f} : \mathbb{R} \longrightarrow \mathbb{R}^n$) such that $\vec{f}$ satisfies the system.

9.3.2 Theorem

Any system of ODEs can be written as a system of first order ODEs, and the number of equations in the new system will be equal to the sum of the original number of equations and the total number of derivatives of $f_1, ..., f_n$ beyond their first that appear in the original system.

9.3.3 Corollary

Any single ODE of order $k > 1$ can be solved as a system of k first order ODEs.

9.3.4 Systems of First Order Linear ODEs With Constant Coefficients

By 9.3.2 we can, without loss of generality, discuss only systems of first order ODEs, and we will assume that each equation can be written in explicit form. A linear system with constant coefficients is then of the form

$$\tfrac{df_1}{dx} = \left(\textstyle\sum_{i=1}^n \alpha_{1i} f_i\right) + g_1(x)$$

$$\vdots$$

$$\tfrac{df_n}{dx} = \left(\textstyle\sum_{i=1}^n \alpha_{ni} f_i\right) + g_n(x)$$

which can be written as the matrix equation

$$\begin{bmatrix} \frac{df_1}{dx} \\ \vdots \\ \frac{df_n}{dx} \end{bmatrix} = \underbrace{\begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{bmatrix}}_{(:=A)} \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix} + \underbrace{\begin{bmatrix} g_1(x) \\ \vdots \\ g_n(x) \end{bmatrix}}_{(:=\vec{g}(x))} \iff \frac{d\vec{f}}{dx} = A\vec{f} + \vec{g}(x).$$

So if we define the linear operator $\mathcal{L}[\vec{f}\,] := \frac{d\vec{f}}{dx} - A\vec{f}$, then the entire system can be written

$$\mathcal{L}[\vec{f}\,] = \vec{g}(x).$$

The solutions to $\mathcal{L}[\vec{f}\,] = 0$ will form an n-dimensional vector space (since each of the n unknown functions in the vector is given by first order ODEs), and a set of linearly independent solutions $\{\vec{f_i}\}_{i=1}^n$ will form a basis. The complementary function is thus $\vec{f}_{CF} = \sum_{i=1}^n c_i\vec{f_i}$, and $\vec{f}_{GS} = \vec{f}_{CF} + \vec{f}_{PI}$ (where $\vec{f}_{PI}$ satisfies $\mathcal{L}[\vec{f}_{PI}] = \vec{g}(x)$).
To actually solve the homogeneous problem (find f_{CF}), we'll first consider the case where A is diagonalisable, so $\exists\, V$ (with columns equal to linearly independent eigenvectors of A: $\vec{v}_i$) such that $V^{-1}AV = \Lambda$, where Λ is diagonal with entries λ_i (the eigenvalues of A). Then

$$\begin{aligned}\mathcal{L}[\vec{f}_{CF}] = 0 \implies \tfrac{d\vec{f}_{CF}}{dx} = A\vec{f}_{CF} &\implies V^{-1}\tfrac{d\vec{f}_{CF}}{dx} = V^{-1}AVV^{-1}\vec{f}_{CF} \\ &\implies \tfrac{d}{dx}[V^{-1}\vec{f}_{CF}] = [V^{-1}AV][V^{-1}\vec{f}_{CF}] \\ &\implies \tfrac{d\vec{z}}{dx} = \Lambda\vec{z}\end{aligned}$$

(where $V^{-1}\vec{f}_{CF} := \vec{z}$). Since Λ is diagonal, we now have a simple system of ODEs of the form $\frac{dz_i}{dx} = \lambda_i z_i$ which all solve to give $z_i = c_i e^{\lambda_i x}$, and $\vec{f}_{CF} = V\vec{z} \implies$

$$\vec{f}_{CF} = c_1 e^{\lambda_1 x}\vec{v}_1 + ... + c_n e^{\lambda_n x}\vec{v}_n$$

and so a basis for the complementary solution space is $\{\vec{f_i}\} = \{e^{\lambda_i x}\vec{v}_i\}$.

9.3.5 Note

Do not confuse notation here. In section 9.2, $\{f_i\}_{i=1}^k$ was the basis of the complementary solution space for a single k^{th} order linear ODE. In 9.3.4, $\{\vec{f_i}\}_{i=1}^n$ is the basis of the system's complementary solution space, while $\{f_i\}_{i=1}^n$ is simply the set of unknown functions in the system (notice the difference in the cardinality of the sets, n and k).

9.3.6 (9.3.4 Cont.)

If A is not diagonalisable, we can instead find W such that $W^{-1}AW := J$ is Jordan - a matrix whose only non-zero off-diagonal entries lie on the superdiagonal above and adjacent to identical diagonal entries, and are equal to 1. In the context of eigenvalues, J will have 1s 'connecting' repeated eigenvalues of A, although for each additional dimension of an eigenspace, one fewer 1 will appear with that eigenvalue in J, so note that diagonalisation (Λ) is a special case of the Jordan canonical form (J). In solving the homogeneous problem, we again have

$$\begin{aligned}\mathcal{L}[\vec{f}_{CF}] = 0 \implies \tfrac{d\vec{f}_{CF}}{dx} = A\vec{f}_{CF} &\implies W^{-1}\tfrac{d\vec{f}_{CF}}{dx} = W^{-1}AWW^{-1}\vec{f}_{CF} \\ &\implies \tfrac{d}{dx}[W^{-1}\vec{f}_{CF}] = [W^{-1}AW][W^{-1}\vec{f}_{CF}] \\ &\implies \tfrac{d\vec{z}}{dx} = J\vec{z}\end{aligned}$$

(where $W^{-1}\vec{f}_{CF} := \vec{z}$). While J is not diagonal, the last ODE of this new system contains only one unknown function and can be solved and substituted upwards into successive inhomogeneous linear first order ODEs which, due to the structure of a Jordan matrix, contain only two unknown functions each.

9.4 Qualitative Analysis of ODEs

9.4.1 Definition

$\vec{f}^*$ is a fixed point or equilibrium point of a system of first order ODEs if $\vec{f}(t_0) = \vec{f}^* \implies \vec{f}(t) = \vec{f}^*\ \forall\, t > t_0$, i.e.

$$\left.\frac{d\vec{f}}{dt}\right|_{\vec{f}=\vec{f}^*} = 0$$

9.4.2 Definition

In the phase plane, a fixed point $\vec{f}^*$ is Lyapunov stable $\iff$

$$\forall\, \varepsilon > 0,\ \exists\, \delta > 0 \text{ such that } ||\vec{f}(0) - \vec{f}^*|| < \delta \implies ||\vec{f}(t) - \vec{f}^*|| < \varepsilon\ (\forall\, t \geq 0)$$

and asymptotically stable if additionally, $\exists\, \delta > 0$ such that $||\vec{f}(0) - \vec{f}^*|| < \delta \implies \lim_{t\to\infty} ||\vec{f}(t) - \vec{f}^*|| = 0$. Stability is essentially the idea that if the initial conditions ($t = 0$) are close to a stable fixed point, they will stay close to it, or in the asymptotic case converge to it, as time progresses.

9.4.3 Definition

For linear systems which can be written in matrix form $\frac{d\vec{f}}{dt} = A\vec{f}$, supposing $\vec{v}$ is an eigenvector of A, we have that $\frac{d\vec{v}}{dt} = A\vec{v} = \lambda\vec{v}$ for some λ. So initial conditions on the line in the phase plane defined by $\vec{v}$ will stay on it. The line is invariant.

9.4.4 Definition

A phase portrait of a system is a geometrical representation of all distinct solutions with qualitatively different trajectories represented across the phase plane, for example, eigenvector lines of the system matrix will appear in the phase portrait with arrows pointing in or out depending on the sign of their eigenvalue.

9.4.5 Theorem: Catalogue of Phase Portraits for the Two-Dimensional Linear System With Constant Coefficients

Consider the general two-dimensional system

$$\frac{d\vec{f}}{dt} = A\vec{f}, \quad A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad (\vec{f} = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}).$$

The eignvalues of A are given by $(a-\lambda)(d-\lambda)-bc = 0 \iff ad-bc-(a+d)\lambda+\lambda^2$, so letting $\tau = \text{Tr}(A) = a+d$ and $\Delta = \det(A) = ad - bc$, we have $\lambda_{1,2} = \frac{\tau\pm\sqrt{\tau^2-4\Delta}}{2}$. The $\tau\Delta$-plane can be split into case regions which give rise to distinct behaviours in the xy-phase plane. Each case is drawn and discussed in lectures, but in terms of stability, they can be grouped as follows: in the upper left quadrant, including the positive Δ-axis and negative τ-axis, the fixed points of the system are Lyapunov stable and also asymptotically stable aside from on the positive Δ-axis, and they are unstable elsewhere.

9.4.6 Note

Notice that in the above phase portraits, stability occurred when $\tau \leq 0,\ \Delta \geq 0$, which ensured that $\text{Re}\{\{\lambda_1\}, \{\lambda_2\}\} \leq 0$. In an n-dimensional system, i.e. where $A \in \mathbb{R}^{n\times n}$ and $\vec{f} \in \mathbb{R}^n$, we similarly require that $\text{Re}\{\lambda_i\} \leq 0\ \forall\, i \in \{1, ..., n\}$ for stability of the system.

9.5 Bifurcations

9.5.1 Definition

A bifurcation in a dynamical system is a qualitative change in behaviour brought about by change in parameters. This may cause a previously stable system to become unstable, for example as in 9.4.5-6, by making some eigenvalues of A have non-negative real parts, or in the one-dimensional system

$$\frac{dy}{dt} = ky$$

where a negative k gives rise to a stable system and a positive k to an unstable system, both clearly with a fixed point at $y = 0$.

9.5.2 Definition

A bifurcation diagram plots the changing values of the fixed points of a system against the system's varying parameter. Convention is to use $\bullet$ for a stable fixed point and $\circ$ for an unstable fixed point in graphical representations.

9.5.3 Definition

The following list gives some example ODE forms for the common types of bifurcations named beside them. Some of their names arise from their bifurcation diagrams, which are not shown here.

$\frac{dy}{dt} = r + y^2$ Saddle-node bifurcation†.

$\frac{dy}{dt} = ry - y^2$ Transcritical bifurcation.

$\frac{dy}{dt} = ry - y^3$ Supercritical pitchfork bifurcation.

$\frac{dy}{dt} = ry + y^3$ Subcritical pitchfork bifurcation.

† A saddle-node bifurcation is in fact any collision and disapperance of two equilibria.

9.5.4 Definition

Another type of special point is a singularity, at which the equation is undefined. Singularities may also be stable or unstable, as in the example $\frac{dy}{dt} = \frac{k}{y}$, for $k < 0$ or $k > 0$.

10 Introduction to Multivariate Calculus

So far, we have considered functions of single independent variables, of the form

$$f : \mathbb{R} \longrightarrow \mathbb{R}$$

$$\left(\text{or, in systems, } \vec{f} = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} : \mathbb{R} \longrightarrow \mathbb{R}^n, \text{ where } f_1, ..., f_n : \mathbb{R} \longrightarrow \mathbb{R}\right).$$

All the functions f and $f_1, ..., f_n$ above are called ordinary functions, and differential equations involving them are ordinary differential equations. In this section, we look at multivariate functions, which are of the form

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}.$$

10.0.1 Definition

Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function in n independent variables $x_1, ..., x_n$. The partial derivative of f with respect to $x_i \in \{x_1, ..., x_n\}$ is

$$\frac{\partial f}{\partial x_i} = \lim_{h \to 0} \frac{f(x_1, ..., x_i + h, ..., x_n) - f(x_1, ..., x_i, ..., x_n)}{h}.$$

10.0.2 Note

In actually finding partial derivative formulae, the formula of the function in all its variables can be differentiated with respect to one as if all others were constant. This is reflected in the fact that, taking for example the function with output $f(x, y)$, we often write the partial derivative with respect to one variable as being 'evaluated' at the other variable: $\left.\frac{\partial f}{\partial x}\right|_y$ or $\left(\frac{\partial f}{\partial x}\right)_y$.

Considering the same example function in two variables, let $g_x(x, y) = \frac{\partial f}{\partial x}$ and $g_y(x, y) = \frac{\partial f}{\partial y}$. These partial derivatives can be partially differentiated again, still with respect to either variable:

$$\begin{aligned} \frac{\partial g_x}{\partial x} &= \frac{\partial^2 f}{\partial x^2}, \qquad & \frac{\partial g_x}{\partial y} &= \frac{\partial}{\partial y}\left[\frac{\partial f}{\partial x}\right], \\ \frac{\partial g_y}{\partial y} &= \frac{\partial^2 f}{\partial y^2}, \qquad & \frac{\partial g_y}{\partial x} &= \frac{\partial}{\partial x}\left[\frac{\partial f}{\partial y}\right]. \end{aligned}$$

10.0.3 Corollary: Symmetry of Mixed Derivatives

Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ have continuous first and second partial derivatives with respect to two of its independent variables x, y, then $\frac{\partial}{\partial y}\left[\frac{\partial f}{\partial x}\right] = \frac{\partial}{\partial x}\left[\frac{\partial f}{\partial y}\right]$ and this derivative is written $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$.

10.0.4 Theorem: The Total Derivative

For the function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, let $\Delta f = f(x_1 + \Delta x_1, ..., x_n + \Delta x_n) - f(x_1, ..., x_n)$ and let $df = \lim_{\Delta x_i \to 0 \ \forall i} \Delta f$.

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i$$

where dx_i is the infinitesimal $\Delta x_i \to dx_i$. The total derivative df evaluates the infinitesimal change of $f(\vec{x})$ as all independent variables x_i change infinitesimally.

10.0.5 Corollary: Chain Rule for Multivariate Functions

Suppose $f : \mathbb{R} \longrightarrow \mathbb{R}$ is a function of t indirectly through the intermediate variables $x_1, ..., x_n$, i.e. f maps $t \longmapsto f(x_1(t), ..., x_n(t))$.

$$\frac{df}{dt} = \left(\frac{\partial f}{\partial x_i}\right)\left(\frac{dx_i}{dt}\right) + ... + \left(\frac{\partial f}{\partial x_n}\right)\left(\frac{dx_n}{dt}\right).$$

10.0.6 Note

As detailed in 10.0.2, when dealing with chain rule examples in multiple interdependent variables, it is important to state explicitly which variables are being kept constant in a partial derivative (as is apparent in the following two examples).

10.0.7 Corollary: Dependence on Another Set of Coordinates

Let f be a function in x, y and suppose x, y are in turn functions in u, v. By substituting the total derivatives dx and dy into the total derivative df and comparing with the total derivative formula for df in terms of u, v, the two partial derivatives below can be found.

$$\left(\frac{\partial f}{\partial u}\right)_v = \left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial x}{\partial u}\right)_v + \left(\frac{\partial f}{\partial y}\right)_x \left(\frac{\partial y}{\partial u}\right)_v \quad \text{and} \quad \left(\frac{\partial f}{\partial v}\right)_u = \left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial x}{\partial v}\right)_u + \left(\frac{\partial f}{\partial y}\right)_x \left(\frac{\partial y}{\partial v}\right)_u .$$

10.0.8 Corollary: Differentiation Using the Implicit Form

Given any function f in the variables $\vec{x}$, letting $y = f(\vec{x})$ (which is the explicit form) we can obviously find an implicit form: $F(\vec{x}, y) := y - f(\vec{x}) = 0$. Taking the total derivative dF and rearranging $dF = 0$ (below left) gives the relationship on the right:

$$dF = \frac{\partial F}{\partial y}dy + \sum_{i=1}^{n} \frac{\partial F}{\partial x_i}dx_i = 0, \qquad dy = -\left(\frac{\left(\frac{\partial F}{\partial x_1}\right)}{\left(\frac{\partial F}{\partial y}\right)}dx_1 + ... + \frac{\left(\frac{\partial F}{\partial x_n}\right)}{\left(\frac{\partial F}{\partial y}\right)}dx_n \right)$$

and so comparing with the total derivative $dy = \frac{\partial y}{\partial x_1}dx_1 + ... + \frac{\partial y}{\partial x_n}dx_n$ we obtain the individual relationship

$$\frac{\partial y}{\partial x_i} = -\frac{\left(\frac{\partial F}{\partial x_i}\right)}{\left(\frac{\partial F}{\partial y}\right)} \quad \forall\, 1 \le i \le n.$$

10.1 Multivariate Taylor Expansion

As before, let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function of variables $\vec{x} = (x_1, ..., x_n)^T$. Letting $\vec{\Delta x} = (\Delta x_1, ..., \Delta x_n)^T$, we can find the Taylor expansion for $f(\vec{x} + \vec{\Delta x})$ about $\vec{x}$ by first expanding about x_1:

$$\begin{aligned} f(\vec{x} + \vec{\Delta x}) &= f(x_1 + \Delta x_1, x_2 + \Delta x_2, ...) \\ &= f(x_1, x_2 + \Delta x_2, ...) + \left(\frac{\partial f}{\partial x_1}\right)_{x_1, x_2 + \Delta x_2, ...} \Delta x_1 + \frac{1}{2!}\left(\frac{\partial^2 f}{\partial^2 x_1}\right)_{x_1, x_2+\Delta x_2, ...} (\Delta x_1)^2 + ... \end{aligned}$$

and upon expanding each term about x_2, and in turn expanding each term in those expansions about x_3 etc., we eventually obtain an entire expansion which can be written, up to the second order, in the form

$$f(\vec{x} + \vec{\Delta x}) = f(\vec{x}) + \begin{bmatrix} \left(\frac{\partial f}{\partial x_1}\right)_{\vec{x}} & \cdots & \left(\frac{\partial f}{\partial x_n}\right)_{\vec{x}} \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \vdots \\ \Delta x_n \end{bmatrix} + \frac{1}{2!} \begin{bmatrix} \Delta x_1 & \cdots & \Delta x_n \end{bmatrix} \begin{bmatrix} \left(\frac{\partial^2 f}{\partial x_1^2}\right)_{\vec{x}} & \cdots & \left(\frac{\partial^2 f}{\partial x_1 \partial x_n}\right)_{\vec{x}} \\ \vdots & \ddots & \vdots \\ \left(\frac{\partial^2 f}{\partial x_n \partial x_1}\right)_{\vec{x}} & \cdots & \left(\frac{\partial^2 f}{\partial x_n^2}\right)_{\vec{x}} \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \vdots \\ \Delta x_n \end{bmatrix} + ...$$

and so defining the Hessian matrix associated with f at $\vec{x}$ and the gradient of f at $\vec{x}$ as

$$H(\vec{x}) = \left[\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{\vec{x}}\right]_{n \times n} \quad \text{and} \quad \vec{\nabla} f_{\vec{x}} = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}_{\vec{x}}$$

respectively, we have a Taylor expansion up to the second order in the form

$$f(\vec{x} + \vec{\Delta x}) = f(\vec{x}) + (\vec{\nabla} f_{\vec{x}})^T \Delta\vec{x} + \frac{1}{2}\Delta\vec{x}^{\,T} H(\vec{x}) \Delta\vec{x} + ...$$

for the multivariate function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$.