

# Math 222: Calculus 3

## Final Exam Revision Notes

### 1 Things we should've remembered from Calc 1, Calc 2, and high school

#### 1.1 Trig Identities

- $\sin(a + b) = \sin(a) \cos(b) + \cos(a) \sin(b)$
- $\cos(a + b) = \cos(a) \cos(b) - \sin(a) \sin(b)$

#### Double-Angle Formulae

- $\sin(2\theta) = 2 \sin(\theta) \cos(\theta) = \frac{2 \tan(\theta)}{1 + \tan(\theta)^2}$
- $\cos(2\theta) = 2 \cos(\theta)^2 - 1 = 1 - 2 \sin(\theta)^2$

#### Half-Angle Formulae

- $\sin(\frac{1}{2}\theta) = \frac{1}{2}(1 - \cos(\theta))$
- $\cos(\frac{1}{2}\theta) = \frac{1}{2}(1 + \cos(\theta))$

#### Power-reduction formulae

- $\sin^2(\theta) = \frac{1}{2}(1 - \cos(2\theta))$
- $\cos^2(\theta) = \frac{1}{2}(1 + \cos(2\theta))$

#### 1.2 Derivatives

$$\begin{aligned}\frac{d}{dx}c^u &= (\ln(c))c^u \frac{du}{dx} \\ \frac{d}{dx} \arcsin(u) &= \frac{1}{\sqrt{1-u^2}} \frac{du}{dx} \\ \frac{d}{dx} \arctan(u) &= \frac{1}{1+u^2} \frac{du}{dx}\end{aligned}$$

#### 1.3 Integrals

- Fundamental Theorem of Calculus:

$$\int_{a(x)}^{b(x)} f(x) dx = a'(x)f(a(x)) - b'(x)f(b(x)) \quad (1)$$

- Integration by parts:

$$\int f(x)g'(x) dx = f(x)g(x) - \int g(x)f'(x) dx \quad (2)$$

- Integrals to memorize:

$$\int \tan(x)dx = \ln(|\sec(x)|) + C \quad (3)$$

$$\int \sec(x)dx = \ln(|\sec(x) + \tan(x)|) + C \quad (4)$$

- Trig substitution

$$\sqrt{a^2 - x^2} \rightarrow x = a \sin(\theta) \quad (5)$$

$$\sqrt{a^2 + x^2} \rightarrow x = a \tan(\theta) \quad (6)$$

$$\sqrt{x^2 - a^2} \rightarrow x = a \sec(\theta) \quad (7)$$

## 2 Infinite Sequences and Series

(Some of this stuff is from my Math 141 notes since there's some overlap).

### 2.1 Sequences

**Definition 2.1.** A **sequence** is a mapping from the natural numbers to the real numbers. We say that a **sequence** converges if the limit of the sequence as you go to infinity is a finite number. Otherwise, we say that it diverges.

#### 2.1.1 Limit Laws

- If  $a_n \rightarrow a$ , then any tail sequence will also converge to  $a$ .
- $a_n$  converging to  $a$  is the same thing as  $|a_n - a| \rightarrow 0$ .
- If  $a_n \rightarrow a$  and  $b_n \rightarrow b$  and  $\lambda, \mu \in \mathbb{R}$ , then the linear combination  $\lambda a_n + \mu b_n$  converges to  $\lambda a + \mu b$ .
- $a_n b_n \rightarrow ab$ .
- If  $b \neq 0$ , then,  $\frac{a_n}{b_n} \rightarrow \frac{a}{b}$ .
- If  $f$  is continuous at  $a$ , then,  $f(a_n) \rightarrow f(a)$ .

#### 2.1.2 Facts

**Definition 2.2.** We say that a sequence is **increasing** if  $a_n < a_{n+1} \forall n \geq 1$ . We say that a function is **decreasing** if the reverse inequality holds. If a function is either increasing or decreasing, we say that it is **monotonic**.

**Definition 2.3.** A **bounded** sequence is one that is either bounded below or bounded above.

- A **bounded below** sequence is one where there exists a real number  $m$  such that  $a_n > m \forall n \geq 1$ .
- A **bounded above** sequence is one where there exists a real number  $n$  such that  $a_n < n, \forall n \geq 1$ .

**Theorem 1** (Monotonic Sequence Theorem). *Every bounded and monotonic sequence is convergent.*

**Theorem 2** (Squeeze Theorem). *If  $a_n \leq b_n \leq c_n$  for some  $n \geq n_0$  and if*

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} c_n = L$$

*then,  $\lim_{n \rightarrow \infty} b_n = L$ .*

## 2.2 Series

A **series** is the sum of all terms of a sequence. It is something of the form

$$a_1 + a_2 + a_3 + \dots$$

An important concept is the idea of a **partial sum** which is given by:

$$s_n = \sum_{k=1}^n a_k = a_1 + a_2 + \dots + a_n$$

If the partial sums converge, then we say that the series converges.

**Example 2.1.** The **geometric series** is the simplest example. It is given by:

$$G = \sum_{n=1}^{\infty} ar^{n-1} = a + ar + ar^2 + ar^3 + \dots + ar^{n-1}$$

- This series converges  $\iff$  the common ratio  $|r| < 1$ . In that case, the sum of the series would be given by:

$$\sum_{i=1}^{\infty} ar^{i-1} = \frac{a}{1-r}$$

**Theorem 3** (Test for Divergence). *If the series  $\sum_{n=1}^{\infty} a_n$  is convergent, then,  $a_n \rightarrow 0$  as  $n \rightarrow \infty$ . Note that the converse is completely false (just because a limit goes to 0 does not guarantee convergence, just see the harmonic series).*

## 2.3 Comparison Tests

**Idea:** to compare a given series with a series that we already know is convergent or divergent. There are two versions of this: (1) the regular comparison test and (2) the limit comparison test.

**Definition 2.4.** The **comparison test** says: suppose that we have two series  $\sum a_n$  and  $\sum b_n$  with only positive terms. Then:

- If  $\sum b_n$  is convergent and  $a_n \leq b_n$  for all  $n$  near infinity, then  $\sum a_n$  is also convergent.
- If  $\sum b_n$  is divergent and  $a_n \geq b_n$  for all  $n$  near infinity, then  $\sum a_n$  is also divergent.

You need a known series to work with. Generally, the two more useful types are geometric and p-series.

**Definition 2.5.** The **limiting comparison test** says: you can compare series without having one that is bigger or smaller than the other. This also requires terms to be positive. If:

$$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = c$$

where  $c$  is a positive, non-zero, finite constant, then there are two possibilities: (1) they either both converge or (2) they either both diverge. **We can basically conclude that the series have the same behaviour as  $n$  tends to infinity.** Advice: take  $b_n$  to be the main terms of both the denominator and numerator.

## 2.4 Integral Test

**Definition 2.6.** The **integral test** says: suppose that  $f$  is continuous, positive, and decreasing on the interval  $[1, \infty[$  and let  $a_n = f(n)$ . Then, the series converges  $\iff$  the corresponding integral is convergent.

This proves a very important lemma:

**Lemma 4.** The  **$p$ -series test**:  $P$ -series are something of the form:

$$\sum_{n=0}^{\infty} \frac{1}{n^p}$$

This series converges  $\iff p > 1$ .

**Useful Fact:** The following converges  $\iff p > 1$ :

$$\sum_{n=3}^{\infty} \frac{1}{n(\ln(n))^p}$$

## 2.5 Absolute Convergence, Root, and Ratio Tests

### 2.5.1 Absolute Convergence

**Motivation:** for any given series  $\sum a_n$ , we can consider the corresponding series:

$$\text{sum}_{n=1}^{\infty} |a_n| = |a_1| + |a_2| + \dots$$

whose terms are the **absolute value** of the terms in the original sequence.

**Definition 2.7.** A series is called **absolutely convergent** if the series of absolute values are convergent.

- A series is called **conditionally convergent** if the series is convergent but not absolutely convergent. For example: alternating harmonic series.
- All absolutely convergent series are also convergent.

### 2.5.2 Ratio Test

This test is useful when determining if a series is absolutely convergent. Good to use when there are factorials. There are 3 cases:

1. Series is absolutely convergent and therefore convergent when:

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = L < 1$$

2. Series is divergent when:

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = L > 1 \text{ or } L = \infty$$

3. The worst one - you cannot draw any conclusions. If this is the case, then the ratio test is completely inconclusive and you need to find another test.

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = L = 1$$

### 2.5.3 Root Test

This test is good to use when we have powers of  $n$ .

1. The series is absolutely convergent when:

$$\lim_{n \rightarrow \infty} |a_n|^{\frac{1}{n}} = L < 1$$

2. The series is divergent when:

$$\lim_{n \rightarrow \infty} |a_n|^{\frac{1}{n}} = L > 1 \text{ or } = \infty$$

3. Completely inconclusive when:

$$\lim_{n \rightarrow \infty} |a_n|^{\frac{1}{n}} = L = 1$$

## 2.6 Alternating Series Test

**Definition 2.8.** An **alternating series** is one where the terms alternate between positive and negative terms. They are in the form of:

$$a_n = (-1)^n b_n$$

where  $b_n$  is the absolute value of the sequence. It can also be present in trigonometric functions as well:

$$\cos(n\pi) \equiv \sin\left(\frac{\pi}{2} + n\pi\right) \equiv (-1)^n$$

If the following rules are satisfied, then the series is convergent. If they are not satisfied, then we cannot say anything and need to find another test:

- $b_n$  is decreasing
- the limit of  $b_n$  as  $n$  tends to infinity is zero.

## 2.7 Estimating Sums

The error involved in using the partial sums  $s_n$  to approximate a convergent series: if  $s \approx s_n$ , then, the error given by the remainder term  $R_n = s - s_n$

**Theorem 5** (Alternating Series Estimation). *If  $s = \sum (-1)^{n+1} b_n$ , where  $b_n > 0$  is the sum of an alternating series that satisfies the alternating series test rules, then,*

$$|R_n| = |s - s_n| \leq b_{n+1}$$

## 2.8 Strategy for Testing Series

1. Classify according to form: if we are asked to compute the sum, then we know that the series must be one that can be studied using a telescopic series or a geometric series (for now).
2. Always use the test for divergence at the beginning.
3. P-series-type series will be rendered inconclusive by the root and ratio tests.
4. Use the integral test when the summand is easy to integrate.

## 2.9 Power Series

**Definition 2.9.** A **power series** is a series of the form:

$$\sum_{n=0}^{\infty} c_n x^n = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + \dots$$

Power series converge for some values of  $x$  and diverge for other values of  $x$ . The **radius of convergence** is a quantity  $r$  in the range  $[0, \infty[$  with the following property: the series converges for  $|x - c| < r$ . It fails to converge when the reverse holds. Anything can happen when  $|x - c| = r$ .

Steps for obtaining the interval of convergence of a power series:

1. First obtain the radius of convergence. You generally do this using the root or ratio test.
2. Then, you need to check the end-points (where equality holds). You need to use a test other than root or ratio.

## 2.10 Representations of Functions as Power Series

Helpful rules for determining the radius of convergence for some power series compositions. Define the following:

$$f(x) = \sum_{n=0}^{\infty} \alpha_n x^n \quad r > 0$$

$$g(x) = \sum_{n=0}^{\infty} \beta_n x^n \quad s > 0$$

Then,  $f(x)$  and  $g(x)$  can be combined in various ways:

- Linear combinations:  $\mu f(x) + \lambda g(x) = \sum (\mu \alpha_n + \lambda \beta_n)$ 
  - Radius of convergence  $\geq \min\{r, s\}$ .
- Multiplication of  $f(x)$  and  $g(x)$ :  $r \geq \min\{r, s\}$ .
- Division: suppose  $g(0) = \beta_0 \neq 0$ . Then:  $\frac{f(x)}{g(x)}$ . All we can say is that  $r$  is strictly positive. To compute the coefficients, equate the coefficients and recursively solve for them.
- You can term-by-term integrate and differentiate: you'd obtain the same radius of convergence, but the interval of convergence may change.
- Compositions:  $g(f(x))$ .  $r > 0$  unless  $f(x)$ 's radius of convergence is infinite. Then, all we can say is that the radius will be at least  $r$ .
- **Useful theorem:** The radius of convergence of a function centred at  $a$  is determined by the distance to the nearest “blow-up” (function undefined) of the function in both  $\mathbb{R}$  and  $\mathbb{C}$ .

**Definition 2.10.** Suppose that  $f$  is  $n + 1$ -times differentiable at  $c$ . Then, it has the finite expansion of the form:

$$f(x) = f(c) + f'(c)(x - c) + \dots + \frac{1}{n!} f^{(n)}(c)(x - c)^n + R_n$$

For some remainder  $R_n$ . The function minus the remainder is called the **Taylor Polynomial** of degree  $P_n$ . The **Taylor series** is given by:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(c)}{n!} (x - c)^n$$

### 2.10.1 Useful Taylor Series Expansions about $x = 0$ :

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{1}{2}x^2 + \frac{1}{3!}x^3 + \frac{1}{4!}x^4 + \dots$$

$$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \frac{1}{6!}x^6 + \dots$$

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 + \dots$$

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + x^4 + \dots$$

$$\cosh(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} = 1 + \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \frac{1}{6!}x^6 + \dots$$

$$\sinh(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = x + \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + \frac{1}{7!}x^7 + \dots$$

$$(1+x)^n = 1 + nx + \frac{(n)(n-1)}{2!}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 + \dots = \sum_{n=0}^{\infty} \binom{n}{k} x^k$$

## 3 Vector-Valued Functions and Space Curves

**Vector functions** are functions of the form:

$$t \mapsto (u_1(t), u_2(t), u_3(t)) = u(t)$$

An example of a vector function are **space curves**. For example,

$$t \mapsto (x(t), y(t), z(t)) = r(t)$$

is a curve in 3-space or t-time. We can obtain the location of the particle at time  $t$  from  $r(t)$ .

**Definition 3.1.** The **velocity** of any space-curve is the derivative of  $r(t)$ :

$$r'(t) = \frac{dr}{dt} = \left( \frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} \right) \quad (8)$$

**Definition 3.2.** The **speed** of any space curve is the magnitude of  $r'(t)$ :

$$v = \sigma = \|r'(t)\| \quad (9)$$

**Definition 3.3.** The **arc-length** of the trajectory is gives us the distance travelled between  $t = a$  and  $t = b$ . It is given by:

$$\int_a^b \|r'(t)\| dt = \int_a^b \sigma dt \quad (10)$$

There are two parameterizations of a curve: with respect to time  $t$  and with respect to arc-length  $s$ . Obtaining an arc-length parametrization is nice, because it makes the speed one, which makes computation easier. If  $\left| \frac{dr}{ds} \right| = 1$ , then,  $s$  is an arc-length parameter.

### 3.1 Frenet-Serret Frame

The **Frenet-Serret frame** consist of three unit vectors: the unit tangent vector  $\hat{T}$ , the unit normal vector  $\hat{N}$ , and the unit binormal vector  $\hat{B}$ .

**Definition 3.4.** The **unit tangent vector**  $\hat{T}$  points in the direction of the curve. There are two ways we can obtain this vector, depending on the parameterization that we have:

$$\hat{T} = \frac{dr}{ds} \quad (11)$$

$$\hat{T} = \frac{r'(t)}{\|r'(t)\|} \quad (12)$$

The first is when we have a unit/arc-length parameterization; the second is when we are given an arbitrary parameterization.

**Definition 3.5.** The **normal unit vector** is the derivative of  $\hat{T}$  with respect to arc-length, divided by its length.

$$\hat{N} = \hat{B} \times \hat{T} \quad (13)$$

$$\frac{d\hat{T}}{ds} = \kappa \hat{N} \quad (14)$$

Note that this vector is only defined if  $\kappa \neq 0$ .

**Definition 3.6.** The **binormal unit vector** is the cross-product of  $\hat{T}$  and  $\hat{N}$ .

$$\hat{B} = \frac{r'(t) \times r''(t)}{\|r'(t) \times r''(t)\|} \quad (15)$$

$$\hat{B} = \hat{T} \times \hat{N} \quad (16)$$

**Definition 3.7.** The **curvature**  $\kappa$  of a curve measures the rate at which the curve is turning away from a tangent line at any point. It measures the curve's failure to be a line. It is defined as:

$$\kappa = \kappa(s) = \left\| \frac{d\hat{T}}{ds} \right\| \quad (17)$$

$$\kappa = \frac{\|r'(t) \times r''(t)\|}{\|r'(t)\|^3} \quad (18)$$

If the curve is planar, then curvature is given by

$$\kappa = \frac{\|f''(t)\|}{(1 + f'(t)^2)^{3/2}} \quad (19)$$

**Definition 3.8.** The **torsion**  $\tau$  of a curve is given by:

$$\frac{d\hat{N}}{ds} = -\tau \hat{B} - \kappa \hat{T} \quad (20)$$

$$\tau = \frac{(r'(t) \times r''(t)) \cdot r'''(t)}{\|r'(t) \times r''(t)\|^2} \quad (21)$$

**Definition 3.9.** The **normal plane** is given by  $(x - r) \cdot \hat{T} = 0$ . It is the plane that passes through  $r$  and is normal to  $\hat{T}$  and is therefore normal to the curve.



**Definition 3.10.** The **osculating plane** is given by  $(x - r) \cdot \hat{B} = 0$ . It is the plane that passes through  $r$  and contains  $\hat{T}$  and  $\hat{N}$ .

**Definition 3.11.** The **radius of curvature** is given by

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

**Definition 3.12.** The **centre of curvature** is given by

$$r + \kappa^{-1} \hat{N} \quad (22)$$

It lies on the osculating plane.

Together, the TNB vectors and the scalars above make up the **Frenet-Serret Formulas**, which are given by:

$$\begin{bmatrix} \hat{T}' \\ \hat{N}' \\ \hat{B}' \end{bmatrix} = \begin{bmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{bmatrix} \begin{bmatrix} \hat{T} \\ \hat{N} \\ \hat{B} \end{bmatrix} \quad (23)$$

if we have an arc-length parameterization and

$$\frac{d}{dt} = \|r'(t)\| \begin{bmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{bmatrix} \begin{bmatrix} \hat{T} \\ \hat{N} \\ \hat{B} \end{bmatrix} \quad (24)$$

for an arbitrary parameterization with respect to  $t$ .

### 3.1.1 Acceleration

Acceleration is the second derivative  $r''(t)$  of  $r(t)$ . There are two components of the acceleration vector:

$$a(t) = a_T(t) \hat{T}(t) + a_N(t) \hat{N}(t) \quad (25)$$

- The **tangential component**, which scales the unit tangent vector. Two ways to compute this:

$$a_T = \frac{d\sigma}{dt} \quad (26)$$

$$a_T = \frac{r'(t) \cdot r''(t)}{\|r'(t)\|} \quad (27)$$

- The **normal component**, which scales the unit normal vector

$$a_N = \kappa \sigma^2 \quad (28)$$

$$a_N = \frac{\|r'(t) \times r''(t)\|}{\|r'(t)\|} \quad (29)$$

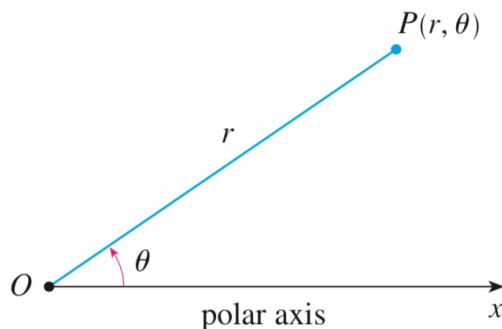


Figure 1: Polar coordinates

## 4 Parametric Equations & Polar Coordinates

### 4.1 Useful Parameterizations to Know

- A **circle** can be parameterized by  $x = \cos(t)$ ,  $y = \sin(t)$ ,  $0 \leq t \leq 2\pi$ .
- A **circle** can also be parameterized like this:  $x = \sin(2t)$ ,  $y = \cos(2t)$ ,  $0 \leq t \leq 2\pi$ .
- A general parameterization of a **circle** with centre  $(h, k)$  and radius  $r$  is given by:

$$x = h + r \cos(t) \quad (30)$$

$$y = k + r \sin(t) \quad (31)$$

$$0 \leq t \leq 2\pi \quad (32)$$

- A **cycloid** can be parameterized by:

$$x = r(\theta - \sin(\theta)) \quad (33)$$

$$y = r(1 - \cos(\theta)) \quad (34)$$

$$\theta \in \mathbb{R} \quad (35)$$

### 4.2 Polar Coordinates

In the polar coordinate system, each point has many representations. The connection between polar coordinates and rectangular coordinates is:

$$x = r \cos(\theta) \quad (36)$$

$$y = r \sin(\theta) \quad (37)$$

To find  $r$  and  $\theta$  when  $x$  and  $y$  are unknown:

$$r^2 = x^2 + y^2 \quad (38)$$

$$\tan(\theta) = \frac{y}{x} \quad (39)$$

#### 4.2.1 Rules of polar coordinates

- $r = a$  generally describes a circle centred at 0 with radius  $|a|$ .
  - $r = 2a \cos(\theta)$  is a circle of radius  $|a|$  with centre  $(a, 0)$ .

- $r = 2b \sin(\theta)$  is a circle of radius  $|b|$  with centre  $(0, b)$ .
- $r = 2a \cos(\theta) + 2b \sin(\theta)$  is a circle with radius  $\sqrt{a^2 + b^2}$  with centre  $(a, b)$ .
- $r = a \mp a \cos(\theta)$  and  $r = a \pm \sin(\theta)$  describes a cardioid. Heart-shaped graph that always contain the origin.
- $\theta = \beta$  is a line that goes through the origin and makes an angle of  $\beta$  with the positive  $x$ -axis; alternatively, it is a line through the origin with slope  $\tan(\beta)$ .
- $r = \cos(\theta) = a$  describes the vertical line  $x = a$ .
- $r = \sin(\theta) = b$  describes the horizontal line  $y = b$ .

### 4.2.2 Areas and Lengths of Polars

The area of a general polar region  $R$  bounded by the polar curve  $r = f(\theta)$  and by the rays  $\theta = a$  and  $\theta = b$ . The area  $A$  over the polar region is given by:

$$A = \int_a^b \frac{1}{2} [f(\theta)]^2 d\theta \quad (40)$$

It's useful to think of area as being swept out by a rotating ray through 0 that starts with an angle  $a$  and ends with  $b$ .

## 5 Multivariable Functions and Extrema

### 5.1 Limits and Continuity

**Definition 5.1.** Let  $f : D \rightarrow \mathbb{R}$  be a function with  $D \subseteq \mathbb{R}^2$  and fix a point  $(a, b) \in D$ . We say that the **limit** of  $f(x, y)$  as  $(x, y)$  approaches  $(a, b)$  is  $L$  if for every  $\epsilon > 0$  we can find a  $\delta > 0$  such that if  $0 < \sqrt{(x-a)^2 + (y-b)^2} < \delta$  and  $(x, y) \in D$ , then:

$$|f(x, y) - L| < \epsilon \quad (41)$$

The limit is denoted by:

$$\lim_{(x,y) \rightarrow (a,b)} f(x, y) = L \quad (42)$$

We say that the limit exists at a given point  $(a, b)$  if the limit equals that value for every possible direction we can approach  $(a, b)$  in.

**Definition 5.2.** We say that a function is **continuous** at  $(a, b)$  if

$$\lim_{(x,y) \rightarrow (a,b)} f(x, y) = f(a, b) \quad (43)$$

### 5.2 Partial Derivatives

**Definition 5.3.** Let  $f : D \rightarrow \mathbb{R}$  be a function with domain  $D \subseteq \mathbb{R}^2$  and fix  $(a, b) \in \mathbb{R}^2$ . Then, the **partial derivative** of  $f$  with respect to  $x$  at  $(a, b)$  is given by the limit:

$$\frac{\partial f}{\partial x}(a, b) = \lim_{h \rightarrow 0} \frac{f(a+h, b) - f(a, b)}{h} \quad (44)$$

The partial derivative with respect to  $y$  is given by

$$\frac{\partial f}{\partial y}(a, b) = \lim_{h \rightarrow 0} \frac{f(a, b+h) - f(a, b)}{h} \quad (45)$$

These are called the **first partial derivatives**. If we take further partial derivatives, we can obtain the **second-partial derivatives** of  $f$ .

**Theorem 6** (Clairaut's Theorem). *Suppose that the mixed second partial derivatives  $\frac{\partial^2 f}{\partial x \partial y}$  and  $\frac{\partial^2 f}{\partial y \partial x}$  exist and are continuous in the open neighbourhood of  $(a, b)$ . Then, the second partial derivatives are equal at  $(a, b)$  for nice functions.*

### 5.3 Linear Approximations and the Tangent Plane

**Definition 5.4.** The **tangent plane** is given by:

$$f(x + \delta x, y + \delta y) = f(x, y) + \frac{\partial f}{\partial x}(x, y)\delta x + \frac{\partial f}{\partial y}(x, y)\delta y + ERROR$$

If the error term tends to 0, then  $f$  can be locally approximated by a linear plane. In  $\mathbb{R}^3$ , this is given by:

$$f(x, y) = f(x_0, y_0) + \begin{bmatrix} \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} + ERROR \quad (46)$$

We can also find the tangent plane using the gradient:

$$\nabla F(a, b, c) \cdot \begin{bmatrix} x - a \\ y - b \\ x - c \end{bmatrix} = 0 \quad (47)$$

### 5.4 The Chain Rule and the Total Derivative

#### 5.4.1 "What the hell is a Jacobian?" - Drury

**Motivation:** the Jacobian matrix is important  $f$  because if  $f$  is differentiable at a point  $x$ , then, the Jacobian **defines a linear map from  $\mathbb{R}^m \rightarrow \mathbb{R}^n$ , which is the best best linear approximation of the function  $f$  near the point  $x$ .**

**Definition 5.5.** The **Jacobian Matrix** is the matrix containing all first-order partial derivatives of a vector-valued function. It comes from linearizing the function. Supposed  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is a function which takes as input the vector  $x \in \mathbb{R}^n$  and produces as output the vector  $f(x) \in \mathbb{R}^m$ . then, the Jacobian matrix  $Jac(F)$  is an  $m \times n$  matrix and is defined as:

$$Jac(F) = \begin{bmatrix} \frac{\partial \mathbf{f}}{\partial x_1} & \dots & \frac{\partial \mathbf{f}}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \quad (48)$$

#### 5.4.2 Chain Rule

We need to use the chain rule whenever we have compositions of functions, i.e.,  $f(g(x))$ . In Calc 3, we need to deal with compositions of vector-valued functions. Each vector-valued function has its own Jacobian matrix associated with it, which describes the linear map that the function induces. Let  $g : \mathbb{R}^p \rightarrow \mathbb{R}^q$  be a function of  $p$  variables and let  $f : \mathbb{R}^q \rightarrow \mathbb{R}^r$  be a

function of  $q$  variables. Then, we can write out the Jacobian matrices for each of these maps:

$$g'(t) = \begin{bmatrix} \frac{\partial x_1}{\partial t_1} & \cdots & \frac{\partial x_1}{\partial t_p} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_q}{\partial t_1} & \cdots & \frac{\partial x_q}{\partial t_p} \end{bmatrix}$$

$$f'(x) = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_q} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_r}{\partial x_1} & \cdots & \frac{\partial y_r}{\partial x_q} \end{bmatrix}$$

Since we are composing linear maps, and since compositions of maps for matrices is the same as matrix multiplication, to obtain the Jacobian matrix of the transformation, we just multiply the Jacobians:

$$h'(t) = f'(g(t))g'(t) \tag{49}$$

$$\begin{bmatrix} \frac{\partial y_1}{\partial t_1} & \cdots & \frac{\partial y_1}{\partial t_p} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_r}{\partial t_1} & \cdots & \frac{\partial y_r}{\partial t_p} \end{bmatrix} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_q} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_r}{\partial x_1} & \cdots & \frac{\partial y_r}{\partial x_q} \end{bmatrix} \begin{bmatrix} \frac{\partial x_1}{\partial t_1} & \cdots & \frac{\partial x_1}{\partial t_p} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_q}{\partial t_1} & \cdots & \frac{\partial x_q}{\partial t_p} \end{bmatrix} \tag{50}$$

### 5.4.3 Implicit Functions

**Theorem 7** (Implicit Function Theorem). *If*

1.  $F(a, b) = 0$  and
2.  $\frac{\partial F}{\partial y}(a, b) \neq 0$

*then, the equation  $F(x, y) = 0$  defines  $y$  as a function of  $x$  near the point  $(a, b)$  and the derivative of  $y$  as a function of  $x$  near  $a$  is given by:*

$$y'(x) = \frac{dy}{dx}(x) = -\frac{\frac{\partial F}{\partial x}(x, y)}{\frac{\partial F}{\partial y}(x, y)} \tag{51}$$

There are 3 cases of this theorem:

1. Case 1:  $x$  and  $y$  are both scalars:  $x \in \mathbb{R}$ ,  $y \in \mathbb{R}$ ,  $f(x, y) = 0$ . Then:

$$\frac{\partial F}{\partial y}(x, y) \neq 0 \Rightarrow y = y(x) \tag{52}$$

2. Case 2:  $x$  is a vector,  $y$  is a scalar:  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ ,  $y \in \mathbb{R}$ ,  $F(x_1, \dots, x_n, y) = 0$ . Then, if

$$\frac{\partial F}{\partial y}(x_1 \ \dots \ x_n \ y) \neq 0 \Rightarrow y = y(x_1 \ \dots \ x_n) \tag{53}$$

3. Case 3: both  $x$  and  $y$  are vectors:  $t = (t_1 \ \dots \ t_p) \in \mathbb{R}^p$ ,  $x = (x_1 \ \dots \ x_n) \in \mathbb{R}^n$ . The condition then becomes:

$$F_1(t_1, \dots, t_p, x_1, \dots, x_n) = 0$$

$$\vdots$$

$$F_n(t_1, \dots, t_p, x_1, \dots, x_n) = 0$$

Therefore,  $F = (F_1, \dots, F_n)$ . The Jacobian becomes:

$$Jac(F) = F' = F_t|F_x = \begin{bmatrix} \frac{\partial F_1}{\partial t_1} & \cdots & \frac{\partial F_1}{\partial t_p} & \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial t_1} & \cdots & \frac{\partial F_n}{\partial t_p} & \frac{\partial F_n}{\partial x_1} & \cdots & \frac{\partial F_n}{\partial x_n} \end{bmatrix}$$

The condition here is that  $\det(F_x) \neq 0$  so that we can invert the matrix. Then, we can write:

$$x(t) = \begin{bmatrix} x_1(t_1, \dots, t_p) \\ \vdots \\ x_n(t_1, \dots, t_p) \end{bmatrix} \Rightarrow x'(t) = \begin{bmatrix} \frac{\partial x_1}{\partial t_1}(t) & \cdots & \frac{\partial x_1}{\partial t_p}(t) \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial t_1}(t) & \cdots & \frac{\partial x_n}{\partial t_p}(t) \end{bmatrix} = -F_x^{-1} F_t$$

## 5.5 The Gradient and Directional Derivatives

**Definition 5.6.** Let  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  be a scalar function  $f(x, y, z)$ . Then, the **gradient** of  $f$ , denoted by  $\nabla f$ , is defined as:

$$\nabla f = \text{grad}(f) = \left( \frac{\partial f}{\partial x} \quad \frac{\partial f}{\partial y} \quad \frac{\partial f}{\partial z} \right) \quad (54)$$

**Definition 5.7.** Suppose  $u$  is a unit vector and  $f(x, y, z)$  is a scalar function. Then, the quantity

$$D_u f(x, y, z) = \nabla f(x, y, z) \cdot u \quad (55)$$

is called the **directional derivative** of  $f$  in the direction  $u$ .

- $D_u f$  takes its maximum value when  $u$  is in the direction of  $\nabla f$  and the maximum value is equal to  $\|\nabla f\|$ .
- The minimum value is  $-\|\nabla f\|$  and it happens when  $u = -\|\nabla f\|^{-1} \nabla f$ .
- $\nabla f$  points in the direction of greatest increase of  $f$ .
- The greatest rate of increase is equal to  $\|\nabla f\|$ .
- $\nabla f$  is normal to the level surfaces of  $f$ .

## 5.6 Extreme Values

A **critical point** of a scalar-valued function  $f$  is a point where  $\nabla f = 0$ .

**Definition 5.8.** The **Hessian matrix** is a matrix that contains the second-order partial derivatives of a scalar-valued function. The mixed derivatives are the entries off the main diagonal. Because of Clairaut's Theorem, this matrix is symmetric. In the 2-by-2 case, this is:

$$H(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} \quad (56)$$

We use the Hessian to conduct a second-derivative test, which we need to do to classify critical points:

1.  $\det(H) < 0 \rightarrow$  saddle

2.  $\det(H) > 0 \rightarrow$ 
  - (a)  $\frac{\partial^2 f}{\partial x^2} > 0 \rightarrow$  minimum
  - (b)  $\frac{\partial^2 f}{\partial x^2} < 0 \rightarrow$  maximum
3.  $\det(H) = 0 \rightarrow$  inconclusive; beyond scope of course.

## 5.7 Maxima and Minima on Restricted Domains

**Theorem 8.** *If  $f$  is continuous and defined on a closed, bounded set, then, it necessarily will attain a maximum and minimum. If you are interested in finding maximum and minimum values and where they are located, then, there are three places that you need to look:*

1. *Points where  $f$  is not differentiable.*
2. *Points where  $f'(x) = 0$ .*
3. *Boundary points.*

## 5.8 Lagrange Multipliers

**Motivation:** to introduce a better way to optimize a function subject to a restricted domain; we want to find **constrained extrema**. The idea is to introduce a new variable – the **lagrange multiplier**. We typically are given an objective function  $f$  to optimize subject to a constraint  $g$ . Then, the solution to the following system of equations plus the constraint  $g$  will give us our critical points:

$$\nabla f = \lambda \nabla g \quad (57)$$

If we are given two constraints:  $g$  and  $h$ , then, the previous equation becomes:

$$\nabla f = \lambda \nabla g + \mu \nabla h \quad (58)$$

# 6 Multiple Integrals

## 6.1 Double Integral

In Calc 2, integration happened over a line. Here, double integrals occur over some rectangle or general region.

**Definition 6.1.** The **double integral** of  $f$  over the rectangle  $R = [a, b] \times [c, d]$  is

$$\iint_R f(x, y) da = \int_{x=a}^{x=b} \int_{y=c}^{y=d} f(x, y) dy dx \quad (59)$$

### 6.1.1 Properties of Double Integrals

1. **Linearity:** you can break up linear combinations of double integrals like in Calc 2.
2. **Positivity:** if  $f(x)$  is positive in  $R$ , then the double integral is also positive in  $R$ .
3. **One:**

$$\iint_R 1 dA = \text{Area of } A$$

4. If  $R$  is a union of two disjoint pieces, then:

$$\iint_{R_1 \cup R_2} f dA = \iint_{R_1} f dA + \iint_{R_2} f dA$$

**Theorem 9** (Fubini's Theorem). *If  $R$  is a bounded, continuous rectangle  $[a, b] \times [c, d]$  and  $f$  is also bounded, then you can reverse the order of integration without changing the solution:*

$$\int_{x=a}^b \int_{y=c}^d f(x, y) dy dx = \int_{y=c}^d \int_{x=a}^b f(x, y) dx dy \quad (60)$$

The **average value** of the area in a region  $D$  is given by:

$$Avg = \frac{1}{area(D)} \iint_D f(x, y) dA \quad (61)$$

## 6.2 Volumes with Double Integrals

With volume questions, we are basically subtracting areas under surfaces –  $z_u$  and  $z_l$

$$Vol = \iint_R (z_u(x, y) - z_l(x, y)) dA \quad (62)$$

## 6.3 Double Integrals in Polar Coordinates

To obtain the area bounded by two angles and two functions  $r$ , we use the following formula:

$$\int_{\theta=\theta_1}^{\theta_2} \int_{r=0}^{r=r(\theta)} r dr d\theta \quad (63)$$

The  $r$  is what we use when we have a change of variables; it's the Jacobian.

## 6.4 Surface Area

**Aim:** compute the surface area of a piece  $z = z(x, y)$  that lies above the region  $R$  in the  $xy$ -plane. The formula that we use is:

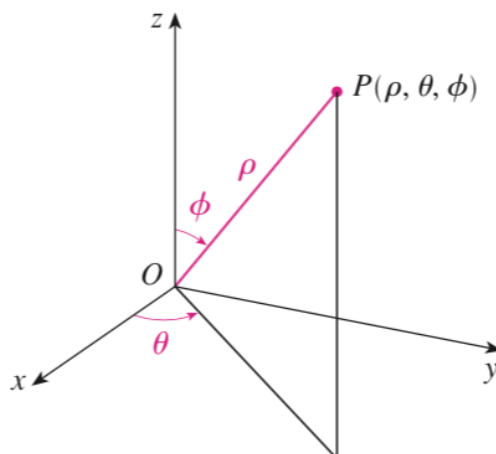
$$\iint_R \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2} dA \quad (64)$$

## 6.5 Spherical Coordinates

In spherical coordinates, we have:

- $\rho$  is the distance from the origin to the point. Must be positive or zero.
- $\theta$  is the same angle we say in polar/cylindrical coordinates; it is the angle between the positive  $x$  axis and the line  $r$  from polars and cylindrical. No restrictions on it.
- $\phi$  is the angle between the positive  $z$ -axis and the line from the origin to the point  $0 \leq \phi \leq \pi$ .





### 6.5.1 Conversions: Spherical to Rectangular

- $x = \rho \sin(\phi) \cos(\theta)$ .
- $y = \rho \sin(\phi) \sin(\theta)$ .
- $z = \rho \cos(\theta)$ .

Helpful patterns:

- $\rho = a$  describes a sphere of radius  $a$  centred at the origin.
- $\phi = \alpha$  describes a cone that makes an angle of  $\alpha$  with the positive  $z = axis$ .
- $\theta = \beta$  describes a vertical plane that makes an angle of  $\beta$  with the positive  $x$ -axis.

### 6.5.2 Setting up triple integrals with sphericals

You can use triple integrals to obtain solids; you integrate over a volume  $E$  with this.

$$V = \int \int \int_E \rho^2 \sin(\phi) d\rho d\phi d\theta$$

$\rho^2 \sin(\phi)$  is the change of variable Jacobian.

### 6.5.3 Quadratic Solids

Knowing these can be useful for figuring out regions of integration

- Ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

- Cone

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z^2}{c^2}$$

This cone will open along the positive  $z$ -axis.

- Cylinder

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

The cross-section here is an ellipse. It is centred on the axis corresponding to the variable that does not appear in the equation.

- Hyperboloid of one sheet

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

This is centred along the axis with the negative in front of it.