# 1 Curved Lines and Surfaces

**Curved Lines**
An equation involving $x$ and $y$, gives a curve in the plane.
By replacing $x$ with $x - a$, the graph is shifted to the right by $a$.
By replacing $y$ with $y - b$, the graph is shifted upward by $b$.

The **parabola** (with vertex) at the origin is defined by

$$y = ax^2 \qquad x = by^2$$

$y = ax^2$ opens up for $a > 0$, and opens down for $a < 0$.
$x = by^2$ opens to the right for $b > 0$, and opens to the left for $b < 0$.

The **ellipse** (with center) at the origin is defined by

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1$$

with $(a, 0)$ being the right-most point on the x-axis,
and $(0, b)$ as the top-most point on the y-axis.
When $a = b$, an ellipse becomes a circle, with $a = b = R$ as the radius of the circle.

The **hyperbola** (with center) at the origin is defined by

$$\left(\frac{x}{a}\right)^2 - \left(\frac{y}{b}\right)^2 = 1 \qquad \left(\frac{y}{a}\right)^2 - \left(\frac{x}{b}\right)^2 = 1$$

A hyperbola has 2 curves, and a center line that is not crossed. The 2 curves goes in the opposite direction of the center.
Since the right side is positive, and a square is positive, the position of the negative sign determines the direction of the curves.
To determine the direction, we can either set $x = 0$ or $y = 0$.
For example, in the case of $x^2 - y^2 = 1$, by setting $x = 0$, we attain $-y^2 = 1$ is impossible, so the verticle line $x = 0$ is the center and is never crossed.
Thus the 2 curves opens toward right and left.

To find the **slant asymptotes** when $x$ and $y$ are both large, change the number 1 on the right side of the equation to 0.

(The above 3 curves together are sometimes called: Conic Sections)

**Curved Surfaces**
An equation involving $x$, $y$, $z$, gives a curved surface in 3D space.
In general, this surface is difficult to imagine and sketch in 3D.

The **3D Sphere** (with center) at the origin with radius $R$ is given by

$$x^2 + y^2 + z^2 = R^2$$

The **3D ellipsoid** (with center) at the origin is given by

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

$(a, 0, 0), (0, b, 0), (0, 0, c)$ are the 3 points on the ellipsoid, similar to the ellipse.
When $a = b = c$, an ellipsoid becomes a sphere, with $a = b = c = R$ as the radius.

**Surfaces Attained Through Rotation:** $r^2 = x^2 + y^2$
When the variables $x$ and $y$ appear together as $x^2 + y^2$, this signals that the surface is attained through rotation. Set a new variable $r$, with $r^2 = x^2 + y^2$.
We graph the equation involving $z$ and $r$ in the $r - z$ plane, with $r > 0$ only.
We then revolve the curve around the z-axis to attain the surface in 3D:
The r-axis stands for both x-axis and y-axis.

**Question 1**
Sketch the following curves.

1. $x = 2y^2$
2. $\dfrac{x^2}{2} + \dfrac{y^2}{9} = 1$
3. $x^2 + 2y^2 = 4$
4. $x^2 + 3y^2 + 2x - 12y + 10 = 0$
5. $y^2 - x^2 = 1$
6. $\dfrac{(x-2)^2}{4} - \dfrac{(y+2)^2}{9} = 1$

**Question 2**
Sketch the following surfaces.

1. $x^2 + y^2 = 1$
2. $z = x^2 + y^2$
3. $z^2 = x^2 + y^2$
4. $x^2 + y^2 + \dfrac{z^2}{4} = 1$
5. $z = (\sqrt{x^2 + y^2} - 1)^2$
6. $x^2 + y^2 + z^2 = 2z$

# 2 Vector and Matrix

**Vector**
A point in 2D needs 2 coordinates. It is typically written as, for example:

$$\vec{a} = (1, 2) \quad \text{or} \quad \vec{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

This is the point $x = 1$ and $y = 2$ on the 2D xy plane.

A point in 3D needs 3 coordinates. It is typically written as, for example:

$$\vec{b} = (0, 2, 1) \quad \text{or} \quad \vec{b} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$$

This is the point $x = 0$, $y = 2$, $z = 1$ in 3D space.

The point can also be interpreted as a **vector**, written in the same way.
We can also think of a vector as an **arrow** from the origin to the point.

Typically, it doesn't matter if we write the vector horizontally or vertically.

**Matrix**
A **matrix** is a block of numbers written in a rectangle (or square) in a specific order. For example, $A$ is a $2 \times 3$ (2 by 3)matrix, with 2 rows and 3 columns:

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 5 & 2 \end{bmatrix}$$

We may think of **vectors** as $1 \times n$ or $n \times 1$ matrix. (for n = 2 or 3)
We can **add** matrices, and **multiply** matrices by a **scalar** (number).

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 1+5 & 2+6 \\ 3+7 & 4+8 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}$$

$$c \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} c \cdot 1 & c \cdot 2 \\ c \cdot 3 & c \cdot 4 \end{bmatrix} = \begin{bmatrix} c & 2c \\ 3c & 4c \end{bmatrix}$$

Standard arithmetic rules apply. For matrices A, B, and scalar c:

$$A + B = B + A \qquad cA = Ac \qquad c(A + B) = cA + cB$$

# 3 Determinant

The determinant of a **square** matrix is a number.

Start at the first row, first position, write down this value, and remove this row and this column from the original matrix. Multiply this value to the determinant of the remaining matrix.

$$det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = 1 \cdot det \begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix} ...$$

Now move right, write down this value, and remove this row and this column from the original matrix. Multiply this value to the determinant of the remaining matrix with a **minus sign**.

$$det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = 1 \cdot det \begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix} - 2 \cdot det \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix} ...$$

Now move right again and repeat until we reached the last column. The positive and negative sign need to **alternate**.

$$det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = 1 \cdot det \begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix} - 2 \cdot det \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix} + 3 \cdot det \begin{bmatrix} 4 & 5 \\ 7 & 8 \end{bmatrix}$$

Each time we apply the algorithm, we end up with several new determinants to calculate, but with matrices of **smaller sizes**.

$$det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

Determinant does not behave well with addition and scalar multiplication:

$$det(A + B) \neq det(A) + det(B) \qquad det(cA) \neq c \cdot det(A)$$

**Question 3**

Find the determinant.

1. $\begin{bmatrix} 3 & -1 \\ 2 & 5 \end{bmatrix}$ 2. $\begin{bmatrix} 1 & 3 & -3 \\ -3 & -5 & 2 \\ -4 & 4 & -6 \end{bmatrix}$ 3. $\begin{bmatrix} 3 & 1 & 0 \\ 1 & 3 & 4 \\ 0 & 0 & 4 \end{bmatrix}$

# 4 Dot product

2 vectors can form a **dot product**, and the result is a **scalar** (number).

$$\vec{x} \cdot \vec{y} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ \vdots \end{bmatrix} = x_1 y_1 + x_2 y_2 + ...$$

Note that this dot product is **different** from scalar multiplication, as we are multiplying 2 vectors together, not a scalar with a vector.

The **length** (absolute value) of a vector using the Pythagoras Theorem:

$$|\vec{x}| = \sqrt{\vec{x} \cdot \vec{x}} = \sqrt{(x_1)^2 + (x_2)^2 + ...}$$

Similarly, define the distance between two points using pythagoras Theorem:

$$d(\vec{x}, \vec{y}) = |\vec{x} - \vec{y}| = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2 + ...}$$

Going from a point $a$ to point $c$, is always shorter than going to some other point $b$ first and then back to $c$. This is called the **Triangle Inequality**:

$$d(\vec{a}, \vec{c}) \leq d(\vec{a}, \vec{b}) + d(\vec{b}, \vec{c})$$

$$|\vec{x} + \vec{y}| \leq |\vec{x}| + |\vec{y}|$$

Let $\theta$ be the angle between $\vec{x}$ and $\vec{y}$, the dot product is also given by,

$$\vec{x} \cdot \vec{y} = |\vec{x}| \ |\vec{y}| \ cos\theta$$

Thus, 2 vectors are **orthogonal** (perpendicular) if the dot product is zero.

The dot product have the usual properties of multiplication:

$$\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}$$

$$\vec{a} \cdot (\vec{b} + \vec{c}) = \vec{a} \cdot \vec{b} + \vec{a} \cdot \vec{c}$$

$$c(\vec{a} \cdot \vec{b}) = (c\vec{a}) \cdot \vec{b} = \vec{a} \cdot (c\vec{b})$$

Note that since the dot product needs two vectors and produce a number, a quantity such as $\vec{a} \cdot \vec{b} \cdot \vec{c}$ is not well defined.

**Question 4**
Find the length of the vectors. Find $\vec{a} \cdot \vec{b}$ and the angle between them.
1. $\vec{a} = (4, 3)$, $\vec{b} = (2, -1)$.
2. $\vec{a} = (4, 0, 2)$, $\vec{b} = (2, -1, 0)$.

# 5 Cross product

2 vectors (in 3D) can form a **cross product**, and the result is a **vector** (in 3D).

Let $\vec{x} = (x_1, x_2, x_3)$, $\vec{y} = (y_1, y_2, y_3)$ in 3D.
We typically write the cross product using determinant:

$$\vec{x} \times \vec{y} = det \begin{bmatrix} i & j & k \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix}$$

$$\vec{x} \times \vec{y} = i(x_2y_3 - x_3y_2) - j(x_1y_3 - x_3y_1) + k(x_1y_2 - x_2y_1)$$

$$\vec{x} \times \vec{y} = (x_2y_3 - x_3y_2, -x_1y_3 + x_3y_1, x_1y_2 - x_2y_1)$$

The notations $i = (1, 0, 0)$, $j = (0, 1, 0)$, $k = (0, 0, 1)$ are very easy to use, where the quantity attached to $i$ is the first component of the vector, and the quantity attached to $j$ would be the second, and $k$ would be the third.

The vector produced by the cross product, $\vec{x} \times \vec{y}$, would be **orthogonal to both vectors** $\vec{x}$ and $\vec{y}$. Notice that in most cases, there would be 2 such vectors with this property. They are exactly

$$\vec{x} \times \vec{y} \qquad \text{and} \qquad -\vec{x} \times \vec{y}$$

In fact, we have

$$\vec{x} \times \vec{y} = -\vec{y} \times \vec{x}$$

(Notice that the order of a dot product does not matter, but the order of cross product matters up to a negative sign.)
The other usual properties of multiplication hold:

$$(c\vec{a} \times \vec{b}) = c(\vec{a} \times \vec{b}) = a \times (c\vec{b})$$

$$\vec{a} \times (\vec{b} + \vec{c}) = \vec{a} \times \vec{b} + \vec{a} \times \vec{c}$$

Similar to the dot product, we can also relate the angle between the 2 vectors,

$$\left|\vec{x} \times \vec{y}\right| = \left|\vec{x}\right| \left|\vec{y}\right| \sin\theta$$

Note that since $\vec{x} \times \vec{y}$ is a vector, we need to take its absolute value.
So 2 vectors are **parallel** if the cross product is zero.

**Question 5**
Let $\vec{x} = (3, -2, 1)$, $\vec{y} = (1, -1, 1)$. Find $\vec{x} \times \vec{y}$.

# 6 Lines and Planes

**Line**
For a line to be defined, we need a **direction** $\vec{v}$ and a point on the line $\vec{r_0}$.

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \vec{r} = t\vec{v} + \vec{r_0}$$

The vector $\vec{r}$ and the value t do not need to be determined.
$\vec{v}$ and $\vec{r_0}$ are not unique, any correct pair would give the right equation.)

**Plane**
For a plane to be defined, we need a **normal vector** $\vec{n} = (a, b, c)$ and a point on the plane $\vec{r_0}$.

$$\vec{n} \cdot \vec{r} = \vec{n} \cdot \vec{r_0}$$

$$ax + by + cz = \vec{n} \cdot \vec{r_0}$$

where $\vec{r}$ is the same as above and does not need to be determined.
The **normal vector** $\vec{n}$ is orthogonal to the plane.
$\vec{n}$ and $\vec{r_0}$ are not unique, any correct pair would give the right equation.)

**Plane Defined by 3 Points**
Consider 3 points $\vec{u}, \vec{v}, \vec{w}$. We can connect **any** 2 pairs of points together to create 2 direction vectors which are inside the plane. For example, we may take $\vec{x} = \vec{u} - \vec{v}$, and $\vec{y} = \vec{w} - \vec{v}$. From there we can form the **cross product** to attain a vector that is orthogonal to both vectors inside the plane, so the cross product would be the normal vector, which is orthogonal to the plane.

**Question 6**
Find the equation of the line or plane.

1. The line through $(-8, 0, 4)$ and $(3, -2, 4)$.
2. The plane through the origin and perpendicular to the vector $(1, 5, 2)$.
3. The plane through $(2, 4, 6)$ parallel to the plane $x + y - z = 5$.
4. The plane through $(0, 1, 1)$, $(1, 0, 1)$, $(1, 1, 0)$.
5. The plane equidistant from point $(3, 1, 5)$ and $(-2, 0, 0)$.

# 7 Multivariable Function

In first year, we study 1D functions: **functions of one variable**, as $f(x)$.
This can be plotted in 2D plane, with $y = f(x)$.
Now, we can study **functions of multiple variables**.

**2D Function**
A 2D function, $f(x, y)$, has 2 inputs being $x, y$, and 1 output.

Sometimes, we can label the output as $z = f(x, y)$.
We can plot this 2D function in 3D space, where the height $z$ above the point $(x, y)$ on the xy-plane is equal to $f(x, y)$.

This would give a graph, which is a surface.
Alternatively, $z = f(x, y)$ can be thought of as an equation which involves $x, y, z$, which would give a surface in 3D space.

However, similar to 1D functions, given an equation involving $x, y, z$, it is **not always possible** to isolate $z$ into $z = f(x, y)$.

Examples: $f(x, y) = x^2{+}2y$ $\quad f(x, y) = xe^{xy}$ $\quad f(x, y) = \sin(xy^2)$

**3D Function**
A 3D function, $f(x, y, z)$, has 3 inputs being $x, y, z$, and 1 output.

It is possible to label the output as $w = f(x, y, z)$, but this is typically not useful as this introduces a 4th variable. The ”plot” of this function would be in 4D space, which does not exist.

Examples: $f(x, y, z) = x^2 + 2z - 1$ $\quad f(x, y, z) = ye^{xz}$

**Level Set**
The level set of $f$ at $k$ is given by setting $f$ to be equal to the constant $k$.
This will reduce the dimension of $f$ by 1.

**Question 7**
1. Draw the level set of $f(x, y) = 4x^2 + y^2 + 1$ for $k = 2$, $k = 5$.
2. Draw the level set of $f(x, y, z) = x^2 + y^2 - z$ for $k = 0$ and $k = 2$.

# 8 Limits and Continuity

$f(x, y)$ is continuous at the 2D point $\vec{a}$ if

$$\lim_{(x,y)\to\vec{a}} f(x, y) = f(\vec{a})$$

**Every standard function we know are continuous in their domains. Any function transformations (such as $+, -, \times, /$, composition) of continuous functions, are also continuous. (except division by 0)**

As most functions are continuous, most limits can be obtained by putting $(x, y) = \vec{a}$ in the formula of f(x,y).

Different from 1D functions, computing limits in 2D is much more difficult. Most limit evaluation techniques from 1D functions does not work for 2D functions. In particular, there is **no L'hopital's Rule** for 2D functions.

Fortunately, since most functions are continuous except when dividing by 0, we typically only need to focus on the point where the function is dividing by 0 (and state the function is continuous everywhere else).

It turns out that it is much easier to show a limit does not exist.

**Test for limit that does not exist in $\mathbb{R}^2$**
Given $f(x, y)$, compute the limit as (x, y) approaches $\vec{a} = (0, 0)$.

1. Replace $y$ with a easy curve $y = g(x)$, passing through $\vec{a} = (0, 0)$, (such as $y = 0$, $y = x$, $y = x^2$), then let $x$ approach 0 to attain a (1D) limit.

2. Replace $x$ with a easy curve $x = h(y)$, passing through $\vec{a} = (0, 0)$, (such as $x = 0$, $x = y^2$), then let $y$ appraoch 0 to attain a (1D) limit.

3. Try with several different $g(x)$ and $h(y)$ to attain many (1D) limits.

4. If you find 2 **different** (1D) limits generated by 2 different curves, then the (2D) limit of $(x, y)$ approaches $\vec{a} = (0, 0)$ of $f(x, y)$ does not exist.

Note: If all the 1D limits are the same, that is **NOT** enough for you to conclude the 2D limit exists. To show the limit exists, we typically must use **Squeeze Theorem**.

Since most limit evaluation techniques does not work for 2D functions, it is a very fortunate fact that Squeeze Theorem does work for 2D functions. The formulation is almost the same as the 1D version.

**Squeeze Theorem**
To attain $\lim\limits_{(x,y)\to\vec{a}} f(x,y)$, can try to find $g(x,y)$ and $h(x,y)$ so that:

1. $g(x,y) \le f(x,y) \le h(x,y)$ near the point $\vec{a}$.
2. $\lim\limits_{(x,y)\to\vec{a}} g(x,y) = L = \lim\limits_{(x,y)\to\vec{a}} h(x,y)$

Then we conclude $\lim\limits_{(x,y)\to\vec{a}} f(x.y) = L$.

**Use Squeeze Theorem to show limit exist**
In practice, we typically want to show $\lim\limits_{(x,y)\to\vec{a}} f(x,y) = 0$.

Start at $|f(x,y)|$, create a **chain of inequality**, and simplify $|f(x,y)|$ to attain $|g(x,y)|$, which has limit 0 as $(x,y)$ approaches $\vec{a}$ :

$$|f(x,y)| \le ... \le |g(x,y)| \to 0 \qquad \text{i.e. } \lim_{(x,y)\to\vec{a}} |g(x,y)| = 0$$

Then we conclude $\lim\limits_{(x,y)\to\vec{a}} f(x,y) = 0$.

The typical strategy in constructing the above inequality is to remove positive quantities from the denominator of $f(x,y)$.

**Question 8**

$$f(x,y) = \frac{x^2y}{x^4+y^2} \qquad \text{and} \qquad f(0,0) = 0$$

Show the limit at $(x,y) = (0,0)$ does not exist.

**Question 9**

$$f(x,y) = \frac{xy}{\sqrt{x^4+y^2}} \qquad \text{and} \qquad f(0,0) = 0$$

Show the function is **continuous** at $(0,0)$.

**Question 10**

Find $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2 \sin^2 y}{2x^4+y^2}$

# 9 Derivative

Consider multivariable function $f(x,y)$ or $f(x,y,z)$.
Define the **derivative** to be:

$$\text{For } f(x,y) \qquad \nabla f = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y})$$

$$\text{For } f(x,y,z) \qquad \nabla f = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z})$$

where $\nabla f$ is called **'del'** f, or **gradient** of f.
We can interpret $\nabla f$ as a vector, even though $f(x,y)$ or $f(x,y,z)$ is a scalar.

The quantities in $\nabla f$ such as $\frac{\partial f}{\partial x}$, are called **partial derivatives**.
When taking **partial derivatives** with respect to a variable,
we regard **all other variables as constants** and take derivative as usual.
Sometimes, we use the notation: $\frac{\partial f}{\partial x} = f_x \quad \frac{\partial f}{\partial y} = f_y \quad \frac{\partial f}{\partial z} = f_z$

**Example**
Let $f(x,y) = x^2y^3 + x$.
To take the partial derivative with respect to $x$, $\frac{\partial f}{\partial x}$, we view $y$ as constant.

To take the partial derivative with respect to $y$, $\frac{\partial f}{\partial y}$, we view $x$ as constant.

$$\frac{\partial f}{\partial x} = 2xy^3 + 1 \qquad \frac{\partial f}{\partial y} = 3x^2y^2$$

Let $f(x,y,z) = x^2z + yz^2$.
To take the partial derivative with respect to $x$, $\frac{\partial f}{\partial x}$, we view **both** $y, z$ as constants.
Similarly, for other partial derivatives:

$$\frac{\partial f}{\partial x} = 2xz \qquad \frac{\partial f}{\partial y} = z^2 \qquad \frac{\partial f}{\partial z} = x^2 + 2yz$$

**Question 11**
Compute $\nabla f$.
1. $f(x,y) = (x^2-1)(y+1)$ 2. $f(x,y) = (xy-2)^2$ 3. $f(x,y) = \dfrac{2}{x+3y}$
4. $f(x,y) = e^{x+4y}$ 5. $f(x,y,z) = x^2 + 2z - 1$ 6. $f(x,y,z) = ye^{xz}$

# 10 Directional Derivative

Remember one dimensional derivatives in first year?

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$$

where $f$ is defined on the real line.
We start at the point $x = a$, and we ask, if we move just a tiny bit, away from $x = a$, how much does the function change?

We generalize this idea of moving a tiny bit away in higher dimensions.
In the real line, we can move only left or right.
In 2D, we can move in any direction we like (on the plane).

We can describe the direction of movement by a **unit vector**, similar to the direction vector of the equation of a line.
The **directional derivative** toward the direction $\vec{u}$ at the point $\vec{a}$:

$$D_{\vec{u}} f(\vec{a}) = \lim_{h \to 0} \frac{f(\vec{a} + h\vec{u}) - f(\vec{a})}{h}$$

where $\vec{u}$ is a **unit vector**.

This is also the **rate of change** of the function in the direction $\vec{u}$ at point $\vec{a}$.
Given a vector $\vec{v}$, we can turn it into a **unit vector**: $\vec{u} = \dfrac{\vec{v}}{|\vec{v}|}$
Note: This is a **1D limit**, as $h$ is a number (the length of the movement).

Facts:

1. Partial derivatives are directional derivatives with $\vec{u}$ being in a **coordinate direction**. Example, for $f(x, y)$:

   $$\frac{\partial f}{\partial x} = D_{\vec{u}} f \qquad\qquad \frac{\partial f}{\partial y} = D_{\vec{u}} f$$

   with $\vec{u} = (1, 0)$ with $\vec{u} = (0, 1)$

2. If $f$ is **differentiable**, then all directional derivative exist and

   $$D_{\vec{u}} f = \nabla f \cdot \vec{u}$$

   where $\vec{u}$ is a **unit vector**.

   This formula gives an easy way to calculate directional derivative.

3. We can interpret $\nabla f(\vec{a})$ as a vector.
   If we evaluate $\nabla f$ at a point $\vec{a}$, and if we interpret the vector $\nabla f(\vec{a})$ as a direction, then the directional derivative in the direction of $\nabla f(\vec{a})$ at the point $\vec{a}$ is largest, and the value of this maximum directional derivative is equal to $|\nabla f(\vec{a})|$.

4. Recall equation of plane: $\vec{n} \cdot \vec{r} = \vec{n} \cdot \vec{r_0}$
   For $f(x, y, z) = k$ for some constant k, gives a **level set** surface.
   Define the **tangent plane** at some fixed point $r_0 = (x_0, y_0, z_0)$ by the normal vector $\vec{n} = \nabla f(x_0, y_0, z_0)$.

**Question 12**
$f(x, y) = e^{2x}y + y^2 + 4$. At the point (0, 0), find the directional derivative along the direction given by $\vec{v} = (1, 3)$. (First turn $\vec{v}$ into a unit vector)

**Question 13**
$f(x, y, z) = x^2y + z$. At the point (2, 2, 1), find the maximum rate of change, and the direction with the maximum.
(The direction with maximum rate of change is $\nabla f$, and the value of this maximum rate of change is $|\nabla f|$.)

**Question 14**
$f(x, y, z) = xy^2e^z$. Let $f(x, y, z) = e$.
At the point (1, 1, 1), find the equation of the tangent plane.

**Question 15**
Find all directional derivatives at $(0, 0)$ (including the partials), if they exist. Is the function **continuous** at $(0, 0)$?

*a)*
$$f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

*b)*
$$f(x, y) = \sqrt{|xy|}$$

**Strategy:**
When the function has division by 0, or some other problems at the origin (such as absolute value or square root), we must use the definition of directional derivative, because $f$ **may not be differentiable** at $(0, 0)$ so the formula $\nabla f \cdot \vec{u}$ does not work. Set $\vec{a} = (0, 0)$ and $\vec{u} = (a, b)$, where $a^2 + b^2 = 1$.

# 11 Differentiation

$f$ is **differentiable** if the gradient vector $\nabla f(\vec{a})$ satisfies

$$\lim_{\vec{h}\to\vec{0}} \frac{f(\vec{a}+\vec{h}) - f(\vec{a}) - \nabla f(\vec{a})\cdot\vec{h}}{|\vec{h}|} = 0$$

Note: This is a **2D limit**. This limit is typically used when $\vec{a} = (0,0)$, so we may set $\vec{h} = (x, y)$, and show the limit is zero (using Squeeze Theorem).

Facts:

1. If $f$ is $C^1$ at $\vec{a}$, then f is differentiable at $\vec{a}$ ($\nabla f$ exists).
   $C^1$: All partial derivatives exist near $\vec{a}$ and including at $\vec{a}$,
   and the partial derivatives are **continuous** (as functions) at $\vec{a}$.
2. Being differentiable is a stronger condition than having directional derivative.
   If $f$ is differentiable, then all directional derivative is given by $D_{\vec{u}}f = \nabla f \cdot \vec{u}$
3. If a function is differentiable at $\vec{a}$, then $f$ is continuous at $\vec{a}$.

**Question 16**

$$f(x,y) = \begin{cases} \dfrac{x^2y}{x^4+y^2} & (x,y)\neq(0,0) \\ 0 & (x,y)=(0,0) \end{cases}$$

Show the directional derivatives exist at (0,0), but $f$ is **not differentiable** at (0,0).
**Strategy:**
There are 2 ways to approach the problem.
1. If the function is **not continuous** at $(0,0)$, then it is **not differentiable** at $(0,0)$. So we may try to show it is not continuous at $(0,0)$. However, some functions **are continuous** at $(0,0)$, so we can attain no conclusion in such cases.
2. Compute all the directional derivatives. If f is differentiable, then **every** directional derivative should satisfy $D_{\vec{u}}f = \nabla f\cdot\vec{u}$.
Check this equation for $\vec{u} = (1,0), \vec{u} = (0,1), \vec{u} = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ (easy unit vectors)

**Question 17**
Show whether the function is differentiable at $(0,0)$.

$a)$ $f(x,y) = \begin{cases} \dfrac{x|y|}{\sqrt{x^2+y^2}} & (x,y)\neq(0,0) \\ 0 & (x,y)=(0,0) \end{cases}$ $\quad b)$ $f(x,y) = \begin{cases} \dfrac{x^4+y^4}{x^2+y^2} & (x,y)\neq(0,0) \\ 0 & (x,y)=(0,0) \end{cases}$

# 12 Higher Order Derivative

We may take (partial) derivatives on top of partial derivatives.

In 2 dimensions, $f(x,y)$, there are 4 ways of taking second derivatives.
We put them into the **Hessian Matrix**:

$$H = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial xy} \\ \dfrac{\partial^2 f}{\partial yx} & \dfrac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial}{\partial x}\left(\dfrac{\partial f}{\partial x}\right) & \dfrac{\partial}{\partial x}\left(\dfrac{\partial f}{\partial y}\right) \\ \dfrac{\partial}{\partial y}\left(\dfrac{\partial f}{\partial x}\right) & \dfrac{\partial}{\partial y}\left(\dfrac{\partial f}{\partial y}\right) \end{bmatrix} = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$$

Facts:
If $f$ is $C^2$, then **mixed partial derivatives commute**:
The order of taking derivatives does not change the final answer: $f_{xy} = f_{yx}$
$\rightarrow$ Taking derivative in x then y, is the same as taking in y then x.

In other words, if $f$ is $C^2$, **the Hessian Matrix is symmetric**.

$C^2$: All **2nd order** partial derivatives exist near $\vec{a}$ and including at $\vec{a}$,
and the **2nd order** partial derivatives are **continuous** (as functions) at $\vec{a}$.
(In practice, a function can fail to be continuous when there is division by zero.)

We can also construct the Hessian for $f(x,y,z)$, which will be a $3\times 3$ matrix.
We can also construct higher order derivatives, such as 3rd order $f_{xyz}$ or $f_{xyy}$.
In general, if $f$ is $C^k$, then mixed partial derivatives commute (up to order k).

**Question 18**
Find the Hessian Matrix.
1. $f(x,y) = 3x^2 + 4xy + 5y^2$ 2. $f(x,y) = \cos(x+2y)$ 3.$f(x,y) = e^{2x+y}$
4. $f(x,y) = e^{x^2+y}$ 5. $f(x,y) = e^x \sin(y)$ 6. $f(x,y,z) = x^2y + xz + z^2$

**Question 19**

$$f(x,y) = \begin{cases} \dfrac{x^3y - xy^3}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

Find $f_x, f_y$ at $(x,y) = (0,0)$ **and** $(x,y) \neq (0,0)$ . Then show $f_{xy} \neq f_{yx}$ at $(0,0)$.
i.e. Taking 2nd derivatives in different order give different answers.
Is $f$ $C^2$ at $(0,0)$? (and have we reached a contradiction?)

# 13 Critical Points

In first year, we found critical points of functions by setting the derivative to zero. Then we check their second derivative to conclude whether critical points are max or min. In higher dimensions, it is very similar.

Let $f(x, y)$ be a 2D function (which is $C^2$).
We set $\nabla f(\vec{x}) = \vec{0}$. Solve $\vec{x}$, these are the **critical points**.
Then we check second derivatives: the **Hessian Matrix**.

$$H = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial xy} \\ \dfrac{\partial^2 f}{\partial yx} & \dfrac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

Note that $H$ can contain $x$ and $y$, so $H$ would be different for different points.
For each critical point $\vec{a}$,

if $Det(H(\vec{a})) > 0$, look at the top left entry:

if $\dfrac{\partial^2 f}{\partial x^2}(\vec{a}) > 0$, then it is a **local minimum**.

if $\dfrac{\partial^2 f}{\partial x^2}(\vec{a}) < 0$, then it is a **local maximum**.

if $Det(H(\vec{a})) < 0$, then it is a **saddle point**.

if $Det(H(\vec{a})) = 0$, then it is **inconclusive**.

To find the (absolute) maximum and minimum of $f$ in a region $D$:

1. Find all local extrema inside $D$ using critical points
2. Find all extrema on the **boundary** of $D$
3. Compare all the function values to attain the absolute extrema

**Question 20**
Find and classify all critical points.
1. $f(x, y) = 3x^2 + 4xy + 5y^2$
2. $f(x, y) = x^2 + 3y^4 + 4y^3 - 12y^2$
3. $f(x, y) = (x - 1)(x^2 - y^2)$
4. $f(x, y) = (x^2 + 2y^2)e^{-x^2-y^2}$

**Question 21**
Let $f(x, y) = x^2 + y^2 - 4x$.
Find the max and min inside the triangle $D$ defined by (0, -3), (0, 3), (3, 0).

# 14 Lagrange Multiplier

Sometimes, we want to maximim/minimize a function on some specific curve, or there are some constraints on the variables that the function takes in.

Take the function we want to maximize/minimize to be $f$.
Take the equation that describes the constraint to be $g = 0$.
(Perhaps we have more than 1 constraint, so the second constraint is $h = 0$)
The **Lagrange Multiplier** Formula is:

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

Each constraint will create an extra $\nabla$ term on the right side with some constant in front. (We need every $\nabla$ term on the right to be non-zero.)

$\lambda$ and $\mu$ are constants that you need to solve for, along with $n$ variables of $f$.
The equation is an equation for vectors, so each of the components are equal.
So, there are $n$ equations in the formula since $\nabla f$ has $n$ components.
For each constant like $\lambda$ or $\mu$, there is a equation of constraint for each.
So if there are 2 constraints, there are $n + 2$ variables that you need to solve, and you have $n + 2$ equations to do it.

You are not allowed to **divide by zero** when you cancel terms.
When attempting to divide a quantity $x$, you need to split into 2 cases:
Case 1: when $x \neq 0$ and you can divide
Case 2: when $x = 0$

**Question 22**
Find the max and min of the function $f(x, y)$ subject to constraints given.
$f(x, y) = 3x - 6y$ with constraint $4x^2 + 2y^2 = 25$

**Question 23**
Consider the curve in $\mathbb{R}^2$, $C : y = x^2$
Find the point $p$ on $C$ so that the distance between $p$ and (0, -1) is smallest.
**Strategy**: We minimize $f(x, y) = \text{distance}^2 = (x - 0)^2 + \big(y - (-1)\big)^2$

**Question 24**
Consider 2 curves in $\mathbb{R}^2$, $C_1 : y = x^2$ and $C_2 : y = x - 1$
Find points $p$ on $C_1$ and $q$ on $C_2$ so that the distance between them is smallest.

# 15 Chain Rule in 1D

We typically think of chain rule as a rule for computation:

$$f\big(g(x)\big)' = f'\big(g(x)\big) \cdot g'(x)$$

where we are given a function inside another function such as $y = \sin(e^x)$, and we take the derivative of the outside (not touching the inside), then multiply the derivative of the inside.

Alternatively, we can define a **variable dependence**.
For example, define the variables as $y = y(t)$, and $t = t(x)$.
(In the above example, we would set $y = \sin t$ and $t = e^x$.)
Then we may express the chain rule (in 1D) as:

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

Intuitively, $\dfrac{dy}{dx}$ asks how much would $y$ change, if there is a small change in $x$.
Given the variable dependence above, a small change in $x$ would cause a small change in $t$, which in turn would cause a small change in $y$.
Hence, we have the 2 "fractions" on the right.

**2nd Derivative**
We may use the above form of the chain rule to calculate the 2nd derivative.
We apply another differential operator $\dfrac{d}{dx}$ on top of $\dfrac{dy}{dx}$:

$$\frac{d^2y}{dx^2} = \frac{d}{dx}\Big(\frac{dy}{dx}\Big)$$

Plug in the above form of the chain rule,

$$= \frac{d}{dx}\Big(\frac{dy}{dt}\frac{dt}{dx}\Big)$$

We want to take the derivative of this product, so we need **product rule**,

$$= \frac{d}{dx}\Big(\frac{dy}{dt}\Big) \cdot \frac{dt}{dx} + \frac{dy}{dt} \cdot \frac{d}{dx}\Big(\frac{dt}{dx}\Big)$$

Notice that there are 2 types of derivatives here.

The 2nd term involves the expression $\frac{d}{dx}\left(\frac{dt}{dx}\right)$.
Recall that we assumed $t = t(x)$. (such as $t = e^x$)
So $\frac{dt}{dx}$ is **also a function of x**.
We want to take the derivative of such function **against x**.
This results in the **2nd derivative** of $t = t(x)$ against $x$:

$$\frac{d}{dx}\left(\frac{dt}{dx}\right) = \frac{d^2t}{dx^2}$$

The 1st term involves the expression $\frac{d}{dx}\left(\frac{dy}{dt}\right)$.
This term is more difficult. (and requires an **additional expansion**)
Recall that we assumed $y = y(t)$. (such as $y = \sin t$)
So $\frac{dy}{dt}$ is **also a function of t**. (such as $\frac{dy}{dt} = \cos t$)
We want to take the derivative of such function **against x**, on an expression involving $t$.

Notice that $\frac{dy}{dt}$ is an expression involving $t$, similar to $y = y(t)$.
So to take the derivative against $x$, on an expression involving $t$, we know that for $y = y(t)$:

$$\frac{d}{dx}\Big(y(t)\Big) = \frac{dy}{dx} = \frac{dy}{dt}\,\frac{dt}{dx} = \frac{d}{dt}\Big(y(t)\Big)\,\frac{dt}{dx}$$

Since $\frac{dy}{dt}$ is also an expression involving $t$, it would **behave in the same way as** $y(t)$ when taking derivative **against x**.
We can substitue $y(t)$ with $\frac{dy}{dt}$ in the above expression:

$$\frac{d}{dx}\left(\frac{dy}{dt}\right) = \frac{d}{dt}\left(\frac{dy}{dt}\right)\frac{dt}{dx}$$

Since $y = y(t)$ is a function of $t$, $\frac{dy}{dt}$ on the right is **also a function of t**.
We want to take the derivative of such function **against t**.
This results in the **2nd derivative** of $y = y(t)$ against $t$:

$$\frac{d}{dt}\left(\frac{dy}{dt}\right) = \frac{d^2y}{dt^2}$$

In conclusion, to take the **2nd derivative involving chain rule**:

$$\frac{d^2y}{dx^2} = \frac{d}{dx}\Big(\frac{dy}{dx}\Big) = \frac{d}{dx}\Big(\frac{dy}{dt}\frac{dt}{dx}\Big) = \frac{d}{dx}\Big(\frac{dy}{dt}\Big) \cdot \frac{dt}{dx} + \frac{dy}{dt} \cdot \frac{d}{dx}\Big(\frac{dt}{dx}\Big)$$

Taking an **additional expansion** in the first term:

$$\begin{aligned} &= \frac{d}{dt}\Big(\frac{dy}{dt}\Big)\frac{dt}{dx} \cdot \frac{dt}{dx} + \frac{dy}{dt} \cdot \frac{d^2t}{dx^2} \\ &= \frac{d^2y}{dt^2}\frac{dt}{dx} \cdot \frac{dt}{dx} + \frac{dy}{dt} \cdot \frac{d^2t}{dx^2} \end{aligned}$$

Notice that we get a square of $\dfrac{dt}{dx}$ in the first term.
Thus the **chain rule formula for 2nd derivative**:

$$\frac{d^2y}{dx^2} = \frac{d^2y}{dt^2}\Big(\frac{dt}{dx}\Big)^2 + \frac{dy}{dt} \cdot \frac{d^2t}{dx^2}$$

Note: $\Big(\dfrac{dt}{dx}\Big)^2 \neq \dfrac{d^2t}{dx^2}$.

Of course, given a function $y = y(x)$, one would not go through this long and complicated process just to take 2nd derivative. It is much easier to just take the derivative (twice) using the computation rules.

However, this formula shines when **at least one of the function is unknown**. This is exactly the case for **ODE**, where $y = y(x)$ is **unknown**. We may use this formula to carry out a **change of variables** and turn a complicated ODE into a simple one.

**ODE - Euler Equation**
Consider $y = y(x)$. For numbers $a, b, c$,

$$ax^2y'' + bxy' + cy = 0$$

This is a **2nd order linear equation**, but with **non-constant** coefficients. In general, this type of equation has no solution. But here, notice that we have exactly $x^2$ in front of $y''$, and $x$ in front of $y'$.
Due to this, we can **guess the solution to be** $y(x) = x^n$.
We plug in $y(x) = x^n$ into the ODE, and solve for $n$.

Typically, there will be 2 solutions for $n$. The general solution will be a linear combination of the 2 solutions (with constants $c_1, c_2$ in front).

**Example**

$$x^2y'' - xy' - 3y = 0$$

Take $y = x^n$, then $y' = nx^{n-1}$ and $y'' = n(n-1)x^{n-2}$.
Plugging into the ODE,

$$x^2n(n-1)x^{n-2} - xnx^{n-1} - 3x^n = 0$$

Notice that we can divide both sides by $x^n$.

$$n(n-1) - n - 3 = 0$$

$$n^2 - 2n - 3 = 0$$

We will always get a quadratic of this form. In this case, we get

$$(n-3)(n+1) = 0$$

Thus the general solution is given by

$$y(x) = c_1x^3 + c_2x^{-1}$$

However, notice that given an equation with different numbers, the quadratic we get may have only one (repeated) solution. In this case, we would be "missing" the second solution. We can instead find the solution by using the **chain rule formula for 2nd derivative**.

**Example**

$$x^2y'' - 3xy' + 4y = 0$$

If we proceed using the previous approach, we would get $(n-2)^2 = 0$, giving one (repeated) solution.

Instead, we can perform a **change of variable**, $t = \ln x$.
We get

$$y' = \frac{dy}{dx} = \frac{dy}{dt}\frac{dt}{dx} = \frac{dy}{dt}\frac{1}{x}$$

$$y'' = \frac{d^2y}{dt^2}\left(\frac{dt}{dx}\right)^2 + \frac{dy}{dt} \cdot \frac{d^2t}{dx^2} = \frac{d^2y}{dt^2}\left(\frac{1}{x}\right)^2 + \frac{dy}{dt} \cdot \frac{-1}{x^2}$$

Plugging the above into the ODE,

$$x^2y'' - 3xy' + 4y = 0$$

$$x^2\left(\frac{d^2y}{dt^2}\left(\frac{1}{x}\right)^2 + \frac{dy}{dt}\cdot\frac{-1}{x^2}\right) - 3x\left(\frac{dy}{dt}\,\frac{1}{x}\right) + 4y = 0$$

Notice that all the $x$ cancels out.

$$\frac{d^2y}{dt^2} - \frac{dy}{dt} - 3\frac{dy}{dt} + 4y = 0$$

$$\frac{d^2y}{dt^2} - 4\frac{dy}{dt} + 4y = 0$$

This is a **2nd order linear equation with constant coefficients**.
The solution is given by $y(t) = e^{rt}$.
Plugging into the equation,

$$r^2 - 4r + 4 = 0$$

$$(r-2)^2 = 0$$

This is a **repeated** root solution.
Thus the general solution for $y(t)$ is given by,

$$y(t) = c_1e^{2t} + c_2te^{2t}$$

Since $t = \ln x$, we get the general solution for $y(x)$ is given by,

$$y(x) = c_1e^{2\ln x} + c_2\ln xe^{2\ln x} = c_1\left(e^{\ln x}\right)^2 + c_2\ln x\left(e^{\ln x}\right)^2$$

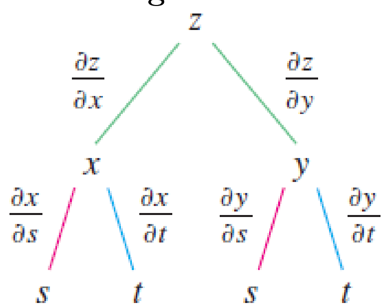
$$y(x) = c_1x^2 + c_2(\ln x)x^2$$

In other words, the "missing" second solution from the repeated root of $n = 2$, is attained by multiplying an extra factor of $\ln x$ in front of $x^n$.

**Question 25**
Find the general solution to the ODE, where $y = y(x)$.
1. $x^2y'' - 3xy' + 4y = 0$
2. $2x^2y'' - 5xy' + 5y = 0$
3. $x(1-x^2)^2y'' - (1-x^2)(1+4x^2)y' + 2x^3y = 0$
   with the change of variable $t = \dfrac{-1}{2}\ln(1-x^2)$.

# 16 Chain Rule in 2D and 3D

Consider an example: $z = z(x, y)$, and $x = x(s, t), y = y(s, t)$.
We can construct the **tree diagram of variable dependence**.




We start with the variable $z$ on the top, and for each variable that $z$ depends on, we get an extra branch below. In this case, $z = z(x, y)$ and we get 2 branches for $x$ and $y$. The 2 branches correspond to the partial derivatives $\dfrac{\partial z}{\partial x}$ and $\dfrac{\partial z}{\partial y}$ as we will see.

Similarly, we create branches for $x$ and $y$, and get 4 branches for $s$ and $t$. The branches correspond to the partial derivatives $\dfrac{\partial x}{\partial s}, \dfrac{\partial x}{\partial t}, \dfrac{\partial y}{\partial s}, \dfrac{\partial y}{\partial t}$.

To compute $\dfrac{\partial z}{\partial s}$:
Using the diagram, we start at $z$, and consider **all possible ways to reach** $s$.
The first way is, start from $z$, go to $x$, then from $x$ to $s$.
This gives (using multiply)

$$\frac{\partial z}{\partial x} \cdot \frac{\partial x}{\partial s}$$

The second way is, start from $z$, go to $y$, then from $y$ to $s$.
This gives (using multiply)

$$\frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial s}$$

The answer is the 2 ways **added** together

$$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x} \cdot \frac{\partial x}{\partial s} + \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial s}$$

Similarly,

$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x} \cdot \frac{\partial x}{\partial t} + \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial t}$$

Intuitively, $\dfrac{\partial z}{\partial s}$ asks how much would $z$ change, if there is a small change in $s$. Given the variable dependence above, a small change in $s$ would cause **both** a small change in $x$ and a small change in $y$. The small change in $x$ would in turn cause a small change in $z$, and the the small change in $y$ would also cause a small change in $z$. Thus the "total" small change in $z$ would be the **sum** of the 2 changes caused by $x$ and $y$.
Hence, we get the **sum** of 2 pairs of "fractions" multiplied on the right.

Of course, given the functions $z = z(x, y)$ and $x = x(s, t), y = y(s, t)$, one does not need go through this long and complicated process just to take the derivative. It is much easier to just plug in the expression of $x$ and $y$ into $z$, and take the derivative using the computation rules.

**Question 26**

1. $z = x - y, x = sin(t), y = 3t$. Find $\dfrac{\partial z}{\partial t}$.
2. $z = xy^2, x = s\cos(t), y = s\sin(t)$. Find $\dfrac{\partial z}{\partial s}, \dfrac{\partial z}{\partial t}$.
3. $u = xy + yz, x = s + t, y = s, z = 2t$. Find $\dfrac{\partial u}{\partial s}, \dfrac{\partial u}{\partial t}$.

Similar to the Chain Rule in 1D, the chain rule formula shines when **at least one of the function is unknown**.

**Question 27**
Suppose that $z = f(x, y), x = 2s + t$ and $y = s - 3t$. Assume:
$f(1, 1) = 2, f(3, -2) = 1, z_x(1, 1) = 2, z_x(3, -2) = 4, z_y(1, 1) = 6, z_y(3, -2) = 1$.
Find $z_s$ and $z_t$ at $(s, t) = (1, 1)$.

**PDE - Partial Differential Equation**
A Partial Differential Equation (PDE) is an equation, which **implicitly defines** a function $u(x, y)$ by mixing $u, x, y$, and **(partial) derivatives of u**.

Similar to ODE, the function $u(x, y)$ defined by an PDE often can not be solved explicitly. Different from an ODE, the general solution of a PDE would involves some **unknown function**, instead of **unknown constants**.

**First Order Linear Equation with Constant Coefficients**
Consider $u = u(x, y)$. Consider the example:

$$2u_x + 3u_y = 0$$

We perform the **change of variables**: $s = 2x + 3y$ and $t = 3x - 2y$.
(This change of variable works in general: $a = 2$, and $b = 3$ in this case.)
Using chain rule, $\big($where we think of $u$ as $u = u(s, t)$, $s = s(x, y)$, $t = t(x, y)\big)$

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial s} \cdot \frac{\partial s}{\partial x} + \frac{\partial u}{\partial t} \cdot \frac{\partial t}{\partial x} = \frac{\partial u}{\partial s}\,2 + \frac{\partial u}{\partial t}\,3$$

$$\frac{\partial u}{\partial y} = \frac{\partial u}{\partial s} \cdot \frac{\partial s}{\partial y} + \frac{\partial u}{\partial t} \cdot \frac{\partial t}{\partial y} = \frac{\partial u}{\partial s}\,3 + \frac{\partial u}{\partial t}\,(-2)$$

Plugging into the PDE, we get

$$2u_x + 3u_y = 0$$

$$2\Big(\frac{\partial u}{\partial s}\,2 + \frac{\partial u}{\partial t}\,3\Big) + 3\Big(\frac{\partial u}{\partial s}\,3 + \frac{\partial u}{\partial t}\,(-2)\Big) = 0$$

Notice that the terms involving $\dfrac{\partial u}{\partial t}$ cancels.

$$(2^2 + 3^2)\frac{\partial u}{\partial s} = 0$$

Thus the **partial derivative** of $u(s, t)$ against $s$ is 0.
However, since we are in 2D, this **does not** mean $u$ is constant.
For example, $u(s, t) = t^2$, would have $\dfrac{\partial u}{\partial s} = 0$.
Thus, $\dfrac{\partial u}{\partial s} = 0$ implies that $u$ must **not** depend on $s$, but can still depend on $t$.
To signal this dependence, we get

$$u(s, t) = f(t)$$

where $f$ is a 1D function only depending on $t$.

Since $t = 3x - 2y$, the **general solution** of $u$ is given by:

$$u(x, y) = f(3x - 2y)$$

Notice that the general solution involves an **unknown function** $f$.

**Initial Condition**
Indeed, one may check that $u(x,y) = f(3x-2y)$ satisfy the original PDE, without needing to know what the unknown function $f$ is.

However, one may impose an **initial condition** on $u$, which would allow one to solve for $f$. This is analogous to imposing an initial condition on an ODE, which would solve for the constants.

Suppose in the above example, we impose the initial condition: $u(x,0) = x^2$. Then since the **general solution** is $u(x,y) = f(3x-2y)$:

$$u(x,0) = f(3x) = x^2$$

We may solve for $f$ by taking a substitution $w = 3x$, so $x = \frac{w}{3}$.

$$f(w) = \frac{w^2}{9}$$

Thus, the **particular solution** is

$$u(x,y) = f(3x-2y) = \frac{(3x-2y)^2}{9}$$

**General Case**
For the equation $au_x + bu_y = 0$: Set $s = ax + by$ and $t = bx - ay$
The **general solution** is $u(x,y) = f(bx - ay)$.

**Question 28**
Consider $u = u(x,y)$. Find the solution to the PDE (general or particular).
1. $-u_x + u_y = 0$ with $u(x,0) = x$.
2. $au_x + bu_y = 0$
3. $u_x + u_y = 1$
4. $u_x - u_y + u = 0$.
5. $u_x + u_y + u = e^{x+2y}$ with $u(x,0) = 0$
6. $u_x + 2u_y + (2x-y)u = 2x^2 + 3xy - 2y^2$

**Question 29**
Consider $2u_x + yu_y = 0$.
(First order linear equation with **non-constant** coefficients: $a = 2$, $b = y$)
Find the solution with the change of variables: $s = x$, $t = e^{-x/2}y$.
($C = e^{-x/2}y$ is the solution to the ODE: $\frac{dy}{dx} = \frac{y}{2} = \frac{b}{a}$)

# 17 Higher Order Chain Rule

Chain rule with higher order derivatives are very complicated.
Consider a simple example,
$u = u(x, y), x = x(s, t), y = y(s, t)$
Using 1st order chain rule,

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial t}$$

then we can apply another differential operator to attain a second derivative, we use product rule:

$$\frac{\partial}{\partial t}(\frac{\partial u}{\partial t}) = \frac{\partial}{\partial t}(\frac{\partial u}{\partial x}\frac{\partial x}{\partial t}) + \frac{\partial}{\partial t}(\frac{\partial u}{\partial y}\frac{\partial y}{\partial t})$$

$$= \frac{\partial}{\partial t}(\frac{\partial u}{\partial x}) \cdot \frac{\partial x}{\partial t} + \frac{\partial u}{\partial x} \cdot \frac{\partial}{\partial t}(\frac{\partial x}{\partial t}) + \frac{\partial}{\partial t}(\frac{\partial u}{\partial y}) \cdot \frac{\partial y}{\partial t} + \frac{\partial u}{\partial y} \cdot \frac{\partial}{\partial t}(\frac{\partial y}{\partial t})$$

Since $u = u(x, y)$, $\frac{\partial u}{\partial x}$ and $\frac{\partial u}{\partial y}$ are **also** functions of $x$ and $y$, so we use the 1st order chain rule from above, except we **replace** $u$ with $\frac{\partial u}{\partial x}$ and $\frac{\partial u}{\partial y}$:

$$\frac{\partial}{\partial t}(\frac{\partial u}{\partial x}) = \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right)\frac{\partial x}{\partial t} + \frac{\partial}{\partial y}\left(\frac{\partial u}{\partial x}\right)\frac{\partial y}{\partial t}$$

$$\frac{\partial}{\partial t}(\frac{\partial u}{\partial y}) = \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial y}\right)\frac{\partial x}{\partial t} + \frac{\partial}{\partial y}\left(\frac{\partial u}{\partial y}\right)\frac{\partial y}{\partial t}$$

These 2 expressions go into the above formula and give a total of 6 terms for $\frac{\partial}{\partial t}(\frac{\partial u}{\partial t})$.

**Question 30**

Let $u = u(x, y), x = st + t^2, y = se^t$. Find $\frac{\partial^2 u}{\partial s \partial t}$

**Question 31**

We find the general solution $u = u(x, t)$ to the PDE, the **wave equation**:

$$u_{tt} - c^2 u_{xx} = 0$$

$c$ is a constant, called the "speed of the wave".
We consider the **change of variables**: $v = x + ct$ and $w = x - ct$

1. Find $u_t$ and $u_x$, according to the change of variables given.
2. Find $u_{tt}$, and $u_{xx}$.
3. Find the **general solution** to the **wave equation**, in terms of $u(x, t)$.

# 18 Regular Integral in 2D

**1 Dimensional Integration**
In 1 dimensional integration, we have a function that depends only on $x$.
We can plot the function in 2D, by letting the $y = f(x)$.
This is called the graph of the function.

When we integrate a function, the domain of integration needs to be an interval $[a, b]$. The interval can be interpreted as a 1 dimensional object.
**The dimension of the domain must match the dimension of the function.**

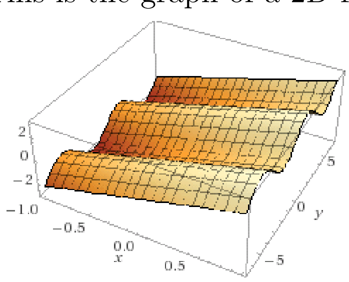
The result of the integral represents the (2 dimensional) area under the curve $y = f(x)$ above the interval $[a, b]$. (if $f(x) > 0$)

**2 Dimensional Integration**
In 2 dimensional integration, we have a function that depends on $x$ and $y$.
For example, $f(x, y) = xy + x^2$, $f(x, y) = e^x sin(xy + 1)$.

Similar to the case that $f(x) = 10$ can be considered as a 1D function,
$f(x, y) = 2x$, $f(x, y) = e^y$, $f(x, y) = 5$, can all be considered as a 2D function, even though they may only depend on one or none of the variables.

If we were to calculate the 2D integral of these functions by considering them as 2D functions, the answer would be **different** than their 1D integral (where we consider them as 1D functions).

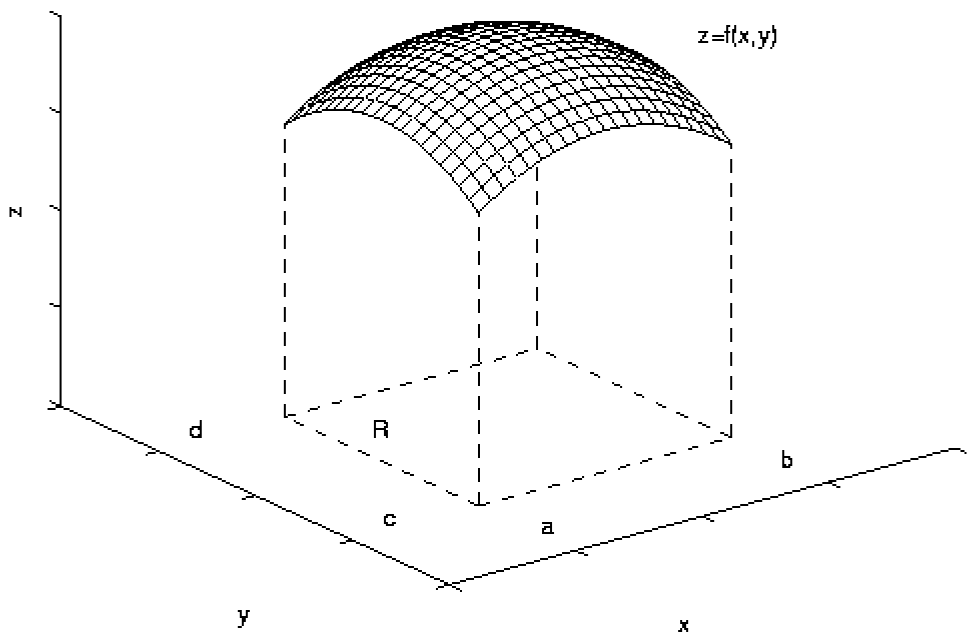
To plot a 2D function, we can let $z = f(x, y)$.
The value of the function will be the height of the point plotted in 3D space.
This is the graph of a 2D function. It will be a 2D surface in 3D.




$z = sin(y) + 2x$

When we integrate a 2D function $f(x, y)$, the domain of integration needs to be a patch of area on the xy-plane. Perhaps the domain is a square, or a circle, or some area that is irregularly shaped.
Regardless, the domain must be a **2 dimensional object** on the xy-plane.
**The dimension of the domain must match the dimension of the function.**

The value of the integral over the domain corresponds to the **3D volume** under the curved surface $f(x, y)$ **above the domain of integration**.




In the above case, the domain is the rectangle on the xy-plane,
with $x$ between $[a, b]$ and $y$ between $[c, d]$.
For the volume, the base is always flat, but the top will usually be curved, since the function $z = f(x, y)$ tends to create a curved surface in 3D.
If the domain had been a triangle, then the value of the integral would be the volume of a triangular prism, except the top will be curved.

Thus, when we try to integrate a 2D function $f(x, y)$ over any domain that is a **1D curve** in the xy-plane, we would be trying to find the volume above the 1D curve. By agreeing that a curve is infinitely thin, we conclude that the volume above the curve must be zero.
This is why our domain of integration for $f(x, y)$ must be a **2D area**.

**Compute 2D Integrals**
A 2D integral typically will look like this:

$$\int_D f(x,y)$$

where $D$ is some 2D domain on the xy-plane.

To compute a 2D integral, we don't plot the function in 3D space.
We split the 2D integral into 2 components, with $dx$ and $dy$.

1. Draw the domain in the xy-plane.

2. For all points inside the domain, determine the maximum and minimum values of $x$. These two values will be **constant** numbers $x_{min}$, $x_{max}$.

3. Pick a random value of $x$ between the minimum and maximum. Consider the **verticle line** that passes through the point $x$. Consider the maximum and minimum values of $y$ on the verticle line.

4. Repeat step 3 several times. For different values of $x$ chosen, the max and min values of $y$ would be different. Typically, all the minimum values of $y$ would lie on a smooth curve, $y_{min} = m(x)$. Similarly, the maximum values would be the same, $y_{max} = M(x)$. In other words, the max and min of $y$ is a **function** of $x$.

5. Thus the integral can be evaluated as

$$\int_D f(x,y) = \int_{x=x_{min}}^{x=x_{max}} \left( \int_{y=m(x)}^{y=M(x)} f(x,y)\,\mathbf{dy} \right) \quad \mathbf{dx}$$

   where the **1D integral** inside the bracket is evaluated first. When carrying out integration with $dy$, we pretend that the variable $x$ is constant in the integration. We use the results of the inner integral, to carry out the integral outside the brackets with $dx$ afterwards.

In the above steps, we first **fixed** $x$, and let $y$ vary between the lower bound curve and the upper bound curve while integrating. Then we add (integrate) all possible values of $x$.

We can also do the reverse by first attaining the min and max values of $y$ in the domain.

1. Draw the domain in the xy-plane.
2. For all points inside the domain, determine the maximum and minimum values of $y$. These two values will be **constant** numbers $y_{min}$, $y_{max}$.
3. Pick a random value of $y$ between the minimum and maximum. Consider the **horizontal line** that passes through the point $y$. Consider the maximum and minimum values of $x$ on the horizontal line.
4. Repeat step 3 several times. For different values of $y$ chosen, the max and min values of $x$ would be different. Typically, all the minimum values of $x$ would lie on a smooth curve, $x_{min} = m(y)$. Similarly, the maximum values would be the same, $x_{max} = M(y)$. In other words, the max and min of $x$ is a **function** of $y$.
5. Thus the integral can be evaluated as

$$\int_D f(x,y) = \int_{y=y_{min}}^{y=y_{max}} \left( \int_{x=m(y)}^{x=M(y)} f(x,y)\,\mathbf{dx} \right) \mathbf{dy}$$

   where the **1D integral** inside the bracket is evaluated first.

Notice in the inner integral, $x$ is integrated first, and $y$ is regarded as constant. Here, $y$ is integrated after $x$, whereas in the previous way, $x$ is integrated after $y$. This is sometimes called **changing the order of integration**.

This is not a good name since, when the order is changed, the limits of integrationg are almost **never the same**. In fact, the outer integral would always have limits that are constant numbers.

**Fubini's Theorem** tells us that the 2 results will always be the **same** regardless of the order we choose, so we are free to choose the order of integration to our convenience when calculating 2D integrals.
(Fubini's Theorem requires all relevant integrals to be integrable.
Typically no division by zero will ensure the function is integrable.)

Sometimes this may be used to our advantage, as some functions do not have 1D antiderivatives.

**Properties of Integrals**
For integrals over some domain $D$ in $\mathbb{R}^n$ with integrable functions:

$$\int_D (f+g) = \int_D f + \int_D g \qquad\qquad \int_D cf = c\int_D f$$

If $f \leq g$ on all of $D$, then

$$\int_D f \leq \int_D g$$

$|f|$ is integrable and

$$\left|\int_D f\right| \leq \int_D |f|$$

If domain $C$ is contained in $D$ and $f \geq 0$, then

$$\int_C f \leq \int_D f$$

For domain $C$ and $D$ (that we usually see)

$$\int_{C\cup D} f = \int_C f + \int_D f - \int_{C\cap D} f$$

Notice that all properties of integral are shared with the integral in 1D.

**Sharp points on the boundary of domain**
In the case where the domain has a sharp point in the boundary, we can split up domain at the sharp point into 2 pieces, and integrate the 2 pieces separately, then add up the answers.

**n-dimensional Volume**
For any set $D$ in $\mathbb{R}^n$, define the **n-dimensional volume** of $D$:

$$V = \int_D 1$$

(Note: here $f = 1$, the constant function)

If n=1, then V is the length of $D$. (1D interval with 1-dimensional volume)
If n=2, then V is the area of $D$. (2D area with 2-dimensional volume)
If n=3, then V is the volume of $D$. (3D volume with 3-dimensional volume)

**Example 1**
Let $f(x, y) = 10$. Let D be the unit square between 0 and 1.

By noticing that the function is constant, we can see that the plot in 3D space will be a flat surface.
The integral will represent the **volume of the rectangular prism**, which can be calculated as base area $\times$ height.
Thus, the answer will be $10 \cdot 1 = 10$.

We can also evaluate the integral using the above steps.
Drawing the domain, we see that the min and max values of $x$ is 0 and 1.
For any fixed value of $x$, we see that the min and max values of $y$ is 0 and 1.

$$\int_D f(x, y) = \int_D 10 = 10(\int_D 1) = 10 \int_{x=0}^{x=1} (\int_{y=0}^{y=1} 1\ dy)\ dx$$

Rules of 1D integration apply, we can move the constant 10 outside.

$$= 10 \int_{x=0}^{x=1} 1\ dx = 10$$

**Example 2**
Let $f(x, y) = x$. Let D be the triangle with vertices $(0, 0), (1, 0), (1, 2)$.

Drawing the domain, The min and max values of $x$ is 0 and 1.
For any fixed value of $x$, the min value of y is 0.
However, the max value of $y$ changes.
The max value of $y$ all lie on the curve $y = 2x = M(x)$.

$$\int_D f(x, y) = \int_D x = \int_{x=0}^{x=1} \left( \int_{y=0}^{\mathbf{y=2x}} x\ dy \right) dx = \int_{x=0}^{x=1} x \left( \int_{y=0}^{\mathbf{y=2x}} 1\ dy \right) dx$$

Notice that the bound inside the inner integral is a function of $x$.
Since we are integrating with $dy$, we regard $x$ as constant, and move it out of the integral.

$$= \int_{x=0}^{x=1} x \left( y \Big|_{y=0}^{\mathbf{y\ =\ 2x}} \right) dx = \int_{x=0}^{x=1} x \cdot 2x\ dx = \frac{2x^3}{3} \Big|_{x=0}^{x=1} = \frac{2}{3}$$

**Domain Defined Using Inequalities**

When domains are defined using an inequality, we take the inequality and make it equality. This will create a curve in the xy-plane which we can draw. We start on the curve, and move a little bit to the right of the curve. This increased the value of $x$, so one side of the equality has become larger.
Using this information, we can determine the area defined by the equality.

**Example 3**

Let $x > 0$, $y > 0$. Let D be the domain defined by $y > x^3$.

We start by taking the equality $y = x^3$.
By moving right, we increased $x$, so we would have $y < x^3$.
Since this is not what we want, we conclude D is the domain to the left of the curve $y = x^3$.

**Domain Bounded Between Curves**

Having curves in the xy-plane, would sometimes bound an area.
We draw the curves in xy-plane and calculate the intersections.

**Example 4**

Let $f(x, y)$ be a function. Let D be bounded by $y = x$ and $y = x^2$.

After drawing the curve, we see that the intersection is at $(0, 0)$ and $(1, 1)$.
We see that the min value of $y$ lie on the curve $y = x^2 = m(x)$.
The max value of $y$ is $y = x = M(x)$.

$$\int_D f(x, y) = \int_{x=0}^{x=1} \left( \int_{y=x^2}^{y=x} f(x, y)\, dy \right) dx$$

**Example 5**

Let D be bounded by $y = x$ and $y = -x + 2$, $y = 0$.

The max value of $y$ is not a smooth curve due to the sharp point at $x = 1$.
We can split the integral in 2 pieces, with $x \in [0, 1]$, and $x \in [1, 2]$.

$$\int_D f(x, y) = \int_{x=0}^{x=1} \left( \int_{y=0}^{y=x} f(x, y)\, dy \right) dx + \int_{x=1}^{x=2} \left( \int_{y=0}^{y=-x+2} f(x, y)\, dy \right) dx$$

We can also interchange the order of integration.

**Question 32**
Let D be bounded by $y = x$ and $y = -x + 2$, $y = 0$.
Compute the integral by changing the order of integration.
Compare with example 5, notice that we no longer need to split up the domain, because the sharp point is now the endpoint of the integral.

**Question 33**
Integrate $\int_D f(x, y)$.

1. Let D be bounded by $x = 0$, $y = 0$, $y = -2x + 2$.
2. Let D be bounded by $y = x + 1$, $y = -x^2 + 1$.
3. Let $D$ be the triangle defined by (0, 2), (0, -2), (2, 0).
4. Let $D = \{(x, y) \in \mathbb{R}^2 \mid x \in [0, 1], y \in [0, 1], y > x^2\}$

**Question 34**
Evaluate
$$\int_{y=0}^{y=1} \left( \int_{x=0}^{x=y^4} y \frac{\sin x}{(1 - \sqrt{x})} \, dx \right) dy$$
**Strategy:**
The function looks difficult. What is the region of integration?
Try changing the order of integration.

**Question 35**
Evaluate.
1.
$$\int_{x=0}^{x=8} \left( \int_{y=\sqrt[3]{x}}^{y=2} \frac{x}{y^7 + 1} \, dy \right) dx$$
2.
$$\int_{x=0}^{x=4} \left( \int_{y=\sqrt{x}}^{y=2} y \cos(y^4 + 3) \, dy \right) dx$$
3.
$$\int_{y=0}^{y=1} \left( \int_{x=y}^{x=\sqrt{y}} \frac{\sin x}{x} \, dx \right) dy$$
4.
$$\int_{y=0}^{y=1} \left( \int_{x=0}^{x=\arccos y} e^{2 \sin x} \, dx \right) dy$$

# 19 Regular Integral in 3D

To compute integral of $f(x, y, z)$ over a 3 dimensional domain $D$, it is similar.

1. Draw the domain in 3D. (This is difficult)
2. Take the **projection** of the 3D domain onto one of the 3 planes created by the axis. This creates a 2D domain.
   Suppose in this case, we take the projection onto the xy-plane.
3. Compute the limits of integration for the 2D domain in xy-plane (as a 2D integral). These will be the limits of the outer most integrals. Match $dx$ and $dy$ according to the order used.
4. Pick a random point $(x, y)$ in the projection. Consider the **vertical** line in the $z$ direction. Consider the maximum and minimum values of $z$.
5. Repeat step 4 several times. The min and max values of $z$ usually will be on a smooth **surface** that **depend** on the point $(x, y)$. Thus $z_{min} = m(x, y)$ and $z_{max} = M(x, y)$.
   These are the limits of the inner most integral, with $dz$.

$$\int_D f(x,y,z) = \int_{2D} \left( \int_{z=m(x,y)}^{z=M(x,y)} f(x,y,z) \ \ \mathbf{dz} \right) dx \ \ dy$$

where the order of $dxdy$ depends upon the order chosen for the 2D integral. In general, inner limits can depend on variables from outer integrals, but limits of outer integrals **never** depend on variables from inner integrals.

When an equation involves only 2 variables (in 3D), it gives a curve in 2D which extends infinitely toward the 3rd dimension, giving a cylinder-like surface in 3D. With 2 such curves, we will get at least 2 intersection points. Connecting the intersection points with a "smooth" curve typically gives the picture of the domain in 3D.

**Question 36**

Let $D$ be bounded by the given equations, in the first octant.
Compute the integral $\int_D f(x, y, z)$ in 6 ways:
(3 ways to project , 2 ways to compute the 2D integral of the projection.)

1. $z = 1 - x$, $y = 1$
2. $z = 1 - y$, $x = y^2$
3. $z = 1 - x^2$, $y = 1 - x$
4. $z = 1 - y^2$, $y = x$
5. $y^2 + z^2 = 9$, $y = 3x$
6. $2x + y + 2z = 1$

# 20 Vector Function

So far, we have considered **multi-variable functions**.
The function takes in 2 (or 3) inputs, and gives out 1 output:

$$f(x,y) : \mathbb{R}^2 \to \mathbb{R}$$

$$f(x,y,z) : \mathbb{R}^3 \to \mathbb{R}$$

We can also have **vector functions**, which gives **multiple outputs**.
In general, we can have

$$f : \mathbb{R}^m \to \mathbb{R}^n$$

$$f\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

It is called a vector function, since $f$ have many outputs, so the outputs as a whole can be regarded as a vector.

$$f(\vec{x}) = \vec{y}$$

Each output, may depend on **all of the inputs**, and so each output is a **coordinate function** that depends on all of the input variables.
In other words, a vector function is made up of many multi-variable functions.

$$f\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} f_1(x_1, x_2, ..., x_m) \\ f_2(x_1, x_2, ..., x_m) \\ \vdots \\ f_n(x_1, x_2, ..., x_m) \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

This makes extending the definitions from multi-variable functions to vector functions very easy:

A vector function $f$ is integrable/continuous/differentiable,
when all of its coordinate functions are integrable/continuous/differentiable.

If we focus only within 3 dimensions, there are (mainly) 2 new types of functions.

**Parametrization**
When $f$ goes from a lower dimensional space, to a higher dimensional space, we sometimes call $f$ a **parametrization**.

**1D Parametrization**

$$f(t) : \mathbb{R} \to \mathbb{R}^2 \qquad \textbf{OR} \qquad f(t) : \mathbb{R} \to \mathbb{R}^3$$

where we have used the variable $t$ as the input of $f$.
Thus for each $t$, $f(t) = (x, y)$ is a point in 2D, or $f(t) = (x, y, z)$ in 3D.

If we view $t$ as time, then we can view $f(t)$ as the **position** of an object.
Thus $f(t)$ describes the position of an object, as time goes on, which will traverse out a **curve**.
Since a curve is a 1D object, this is called a **1D parametrization**.

In fact, this is one of the most important ways to describe a curve in higher dimensions, and finding a parametrization for a given curve is an important skill in Calculus.

**Example**
Given a circle of radius $R$ at the origin, we can choose the parametrization to be:

$$f(t) = (R\cos t, R\sin t) = (x, y) \qquad t \in [0, 2\pi]$$

where $x = R\cos t$, $y = R\sin t$ in polar coordinates.
The parametrization starts at $t = 0$, which is the point $(R, 0)$. Then as $t$ increases, $f(t)$ traverses the circle **counter-clockwise**, as $t = \theta$ is the angle, and back to $(R, 0)$ when $t = 2\pi$.

However, parametrization is not unique. The same circle can be parametrized by:

$$f(t) = (R\sin t, R\cos t) = (x, y) \qquad t \in [0, 2\pi]$$

where $x = R\cos t$, $y = R\sin t$ is **not** the standard polar coordinates.
The parametrization starts at $t = 0$, which is the point $(0, R)$. Then as $t$ increases, $f(t)$ traverses the circle **clockwise**, since $t$ is now the angle between the point $(x, y)$ and the positive y-axis.
We still traverse the whole circle, but this is the less natural parametrization.

When we need a parametrization of a curce, we typically choose the easiest one.

**2D Parametrization**

$$f(u,v) : \mathbb{R}^2 \to \mathbb{R}^3$$

where we have used the variables $u, v$ as the inputs of $f$.
Thus for each $(u,v)$ in $2D$, $f(u,v) = (x,y,z)$ is a point in 3D.

Given a 2D region in the domain $\mathbb{R}^2$ (such as a square), $f(u,v)$ will create a **surface** in 3D.
Since a surface is a 2D object, this is called a **2D parametrization**.

Going from left to right:
We may think that $f$ **lifts up** a 2D region in $\mathbb{R}^2$ into 3D, and maybe stretches the region somewhat, depending on the definition of $f$.

Going from right to left:
Given a surface $S$ in 3D, we can try to find a parametrization for this surface $f(u,v) = (x,y,z)$. $f$ **pulls back** this (curved) surface $S$ in 3D with $(x,y,z)$ as variables, into a **flat** region in 2D with $(u,v)$ as variables.

We know that an equality involving $x, y, z$ gives a surface in 3D: $g(x,y,z) = 0$
This gives an alternative way of describing a surface in 3D. Finding a parametrization for a given surface is also an important skill in Calculus.

**Example**
If we consider the plane $z = x + y + 10$ that is above the unit square on the xy-plane, we may choose the parametrization to be:

$$f(u,v) = (u, v, u + v + 10) = (x,y,z) \qquad u \in [0,1] \quad v \in [0,1]$$

This is called the **natural parametrization**, as we have set $u = x$ and $v = y$.
For any point $(x,y) = (u,v)$ on the xy-plane, $f$ takes in this point, and gives out the same point with a new **3rd coordinate**, or **height**, with the 3rd component defined by $f$: $z = u + v + 10$.
This **lifts up** the unit square from the xy-plane, onto the plane $z = x+y+10$.
However, notice that the square is now slanted, and also stretched.
In this case, the plane $z = x + y + 10$ is flat, but in general, $f$ lifts up a flat region in 2D, into a curved surface in 3D, with some stretching.

In reverse, we can view $f$ pulls back a (possibly complicted) surface in 3D, (in this case $z = x + y + 10$), into a flat region in 2D (the unit square).

**Coordinate Transformation**
When $f$ goes between spaces of the same dimension, we sometimes call $f$ a **coordinate transformation**.

$$f(u,v) : \mathbb{R}^2 \to \mathbb{R}^2 \qquad \textbf{OR} \qquad f(u,v,w) : \mathbb{R}^3 \to \mathbb{R}^3$$

where $f(u,v) = (x,y)$, and $f(u,v,w) = (x,y,z)$.

Going from left to right:
Given a 2D region in the domain $\mathbb{R}^2$ (such as a square), $f(u,v) = (x,y)$ will produce another 2D region in $\mathbb{R}^2$, with some stretching as defined by $f$.

Going from right to left:
Given a (possibly complicated) region in 2D with variables $(x,y)$, we can try to find a coordinate transformation $f$ for this region, where $f(u,v)$ **transforms a simple region** in the uv-plane, onto the original region in the xy-plane.
**We have found a new set of coordinates (u,v), to describe the more complicated region in the xy-plane.**

Similarly, we can also do this for regions in 3D.

**Example**
Given a **solid** circle of radius $R$ at the origin, $x^2 + y^2 \leq R^2$, it is a region in 2D. We know that representing this region is complicated in $(x,y)$, where we get expressions such as $y = \sqrt{R^2 - x^2}$.

However, we may consider **polar coordinates**:

$$f(r,\theta) = (r\cos\theta, r\sin\theta) = (x,y) \qquad r \in [0,R] \quad \theta \in [0,2\pi]$$

where for each point $(r,\theta)$ in the "$r-\theta$" plane, $f(r,\theta) = (x,y)$ is a point in the solid circle in the xy-plane.
In other words, $f$ **transforms a solid rectangle** in the "$r-\theta$" plane, into the solid circle in the xy-plane. Instead of working with a circle, we now can work with a rectangle instead, as we have found **better coordinates** to describe the original region in the xy-plane.

For $f$ to turn a rectangle into a circle, it must stretch the rectangle at multiple places, at different rates, which is based on the definition of $f$.

# 21 Derivative Matrix

Given a vector function, for example $f(x, y, z)$,

$$f \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} f_1(x, y, z) \\ f_2(x, y, z) \\ f_3(x, y, z) \end{bmatrix}$$

The derivative of a vector function is $D_f$, called the derivative matrix.

For the first row of the matrix, we take the first coordinate function, and take the partial derivative with respect to all possible variables $(x, y, z)$ **in order**, horizontally. Then for each coordinate function, we do the same thing to get a new row of the matrix.

$$D_f = \begin{bmatrix} \frac{\partial}{\partial x} f_1 & \frac{\partial}{\partial y} f_1 & \frac{\partial}{\partial z} f_1 \\ \frac{\partial}{\partial x} f_2 & \frac{\partial}{\partial y} f_2 & \frac{\partial}{\partial z} f_2 \\ \frac{\partial}{\partial x} f_3 & \frac{\partial}{\partial y} f_3 & \frac{\partial}{\partial z} f_3 \end{bmatrix}$$

**Relation to Gradient $\nabla f$**
For a multi-variable function $f(x, y, z) : \mathbb{R}^3 \to \mathbb{R}$, it returns one value.
(For example, $f(x, y, z) = xyz^2$)
$D_f = \nabla f = (\frac{\partial}{\partial x} f, \frac{\partial}{\partial y} f, \frac{\partial}{\partial z} f)$
where the derivative matrix has only 1 row.

On the other hand, for a vector function: $f(x, y, z) = (f_1, f_2, f_3)$, each row of $D_f$ is the gradient of all the coordinate functions:

$$D_f = \begin{bmatrix} \text{—} \nabla f_1 \text{—} \\ \text{—} \nabla f_2 \text{—} \\ \text{—} \nabla f_3 \text{—} \end{bmatrix}$$

where $\nabla f_i$ are viewed as horizontal vectors in the matrix.

**Question 37**
$f(x, y, z) = (3x + y, e^y z, xyz)$
First, state the dimension of the vector $f$ takes in, and the dimension of the vector $f$ gives out. Then, compute $D_f$.

# 22 Change of Variables Formula

Suppose you need to integrate $f$ over a domain $B$.
But $B$ is very bad. You don't want to integrate over a bad domain.
Suppose you can find a function $g$ such that $g$ takes another region $A$ onto $B$:

$$g : A \to B \qquad g(A) = B$$

and somehow $A$ is a good domain.
So, instead of integrating $f$ over the bad domain $B$, you can integrate over the good domain $A$ instead.:

$$\int_B f = \int_A f \circ g \cdot |Det(D_g)|$$

$f \circ g$ means replacing the old variables of $f$ in $B$ with the new ones in $A$.
(We require that $g$ to be a one-to-one transformation, and $Det(D_g) \neq 0$.)
The natural question is of course: How do I find such $g$?

1. **Polar Coordinates**
Need to integrate $f(x, y)$

$g(r, \theta) = (rcos\theta, rsin\theta) = (x, y)$

$f \circ g = f(g(r, \theta)) = f(rcos\theta, rsin\theta) = f(x, y)$
$r^2 = x^2 + y^2$
$Det(D_g) = \mathbf{r}$

Useful when domain B is (partly) of circular shape, or function given is already in polar coordinates.

**Question 38**
Integrate $f(x, y) = x$ over $D = \{1 \leq x^2 + y^2 \leq 4\}$ in the 2nd quadrant.
**Question 39**
Show that $\displaystyle\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$
**Strategy**:
The normal distribution has no anti-derivative. Start with $(\displaystyle\int_{-\infty}^{\infty} e^{-x^2})^2$ and write it as a product of 2 same things. Then change the variable in one integral from $x$ to $y$, and use the polar coordinate transformation.

**Polar Curve**
$r$ represent the distance from the point $(x, y)$ to the origin,
$\theta$ represent the angle rotated counterclockwise from x-axis to the point.
Negative angles represent angles rotated clockwise.
Thus $\theta = \pi$ is the same as $\theta = -\pi$.

If $r$ is **negative**, it represent the point on the line with angle $\theta$ through the origin, except on the **opposite** side of the line, with radius being $|r|$.
**Example**
The point $(r, \theta) = (1, 3\pi/4)$ in polar coordinates represent $(x, y) = (-1/\sqrt{2}, 1/\sqrt{2})$.
The point $(r, \theta) = (-1, 3\pi/4)$ in polar coordinates has $r$ as negative. We go on the line of $3\pi/4$, which is toward the top left. Since $r$ is negative, we go on the opposite side of the line, and take the point with radius being $|r| = 1$.
Thus $(-1, 3\pi/4)$ in polar coordinates represent $(x, y) = (1/\sqrt{2}, -1/\sqrt{2})$.

Using the above convention, we can define **polar curve**: $r = h(\theta)$.
To draw the polar curve, we pick some values of $\theta$, and plot the value of $r$ specified by $r = h(\theta)$ on the line represented by $\theta$.

$r = a$ gives the circle with radius $a$ centered at the origin.
$r = acos\theta$ gives the circle to the right (if $a > 0$).
The circle is between the points $(a, 0)$ and the origin. The radius is $a/2$.
$r = bsin\theta$ gives the circle to the top (if $b > 0$).
The circle is between the points $(0, b)$ and the origin. The radius is $b/2$.

**Question 40**
$a$) Find the area enclosed by polar curve $h(\theta)$, but ouside of polar curve $k(\theta)$:
$h(\theta) = r = 1 + cos(\theta)$, $k(\theta) = r = 3cos(\theta)$
**Strategy:**
Plot the curve. Find the intersection between the polar curves. First choose the min and max value of $\theta$ as the outer integral, then $r$ would start from the smaller value (inner polar curve) to the larger value (outer polar curve).

$b$) Sketch the curve $r = 1 + 2cos\theta$ and find the area it encloses in the outer loop but outside the inner loop.
**Strategy:**
This shape is complicated and it crosses itself. We can not apply the formula directly , we need to use symmetry and subtract the inner area.

2. **Cylindrical Coordinates**
Need to integrate $f(x, y, z)$

$g(r, \theta, z) = (rcos\theta, rsin\theta, z) = (x, y, z)$

For any specific point $(x, y, z)$
it can be described by a radius, a height, and an angle:

$r$ is the distance between the point and the **z-axis**.
$\theta$ is the angle between the x-axis and the point (x, y, 0).
$(x, y, 0)$ is the projection of the point onto the xy-plane.
$z$ is the height.

$f \circ g = f(g(r, \theta, z)) = f(rcos\theta, rsin\theta, z) = f(x, y, z)$
$r^2 = x^2 + y^2$
$Det(D_g) = \mathbf{r}$

Useful when domain B is (partly) of cylindrical shape, or has circular projection, or function given is already in cylindrical coordinates, or the domain can be obtained through a rotation with the defining equations containing $x^2+y^2$.

**Graphing surfaces attained through rotation**
When the variables $x$ and $y$ in the defining equations can be written **completely** with $r^2 = x^2 + y^2$, this signals that the surface is attained through rotation. We graph the 2D curves involving $z$ and $r$ in the $r - z$ plane. We then revolve the curves around the z-axis to attain surfaces in 3D. To integrate the volume bounded by the surfaces, we need to use cylindrical coordinates and perform a **2D integration** in the $r - z$ plane, and then we add the $\theta-$integral on the outside.
**Don't forget to include the term** $det(D_g) = \mathbf{r}$**.**
We only integrate the portion with $r > 0$.

**Question 41**
$z = h(x, y) = a \cdot (\sqrt{(x^2 + y^2)} - 1)^2 \qquad a > 0$
Consider the volume:
inside the cylinder $x^2 + y^2 = 1$, above $z = 0$, and below $z = h(x, y)$
Find $a$ so that the volume is 1.

3. **Spherical Coordinates**
Need to integrate $f(x, y, z)$

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

For any specific point $(x, y, z)$, it can be described by the radius and 2 angles:

$\rho$ is the distance from the point to the **origin**, as the radius.
$\phi$ is the angle between the z axis and the point (x, y, z).
  This is called **inclination** angle.
$\theta$ is the angle between the x-axis and the point (x, y, 0).
  $(x, y, 0)$ is the projection of the point onto the xy-plane.
  This is called **azimuthal** angle.

Note: $r$ and $\rho$ both represent 'radius' in Spherical and Cylindrical Coordinate, but they stand for **different radius**.

$f \circ g = f(g(r, \theta, z)) = f(\rho sin\phi cos\theta, \rho sin\phi \sin\theta, \rho cos\phi)$

$\rho^2 = x^2 + y^2 + z^2$
$Det(D_g) = \rho^2 \mathbf{sin}\phi$

Useful when domain B is (partly) of spherical shape, or function given is already in spherical coordinates.

The maximum value of $\phi$ is 180 degrees, at most to the negative z-axis.
For a full sphere of radius R, the range of the variables are:
$\rho : 0 \to R$
$\phi : 0 \to \pi$
$\theta : 0 \to 2\pi$
If the sphere is **not** centered at the origin, spherical coordinates usually is not the choice.

**Question 42**
Find the 3D-volume of a solid hemisphere in $\mathbb{R}^3$ with radius R and $x > 0$.

**Strategy:**
Recall integrating the function $f = 1$ gives the volume of the domain.

**Question 43**

Consider the surface A characterized by $z = \sqrt{3}\sqrt{(x^2+y^2)} = \sqrt{3}\cdot r$, where $r$ is the radius in **cylindrical coordinates**.
Consider the surface B as the sphere of radius 2, $x^2+y^2+z^2=4$.
Find the volume bounded by surface A and B in 2 ways:

a) Use Spherical Coordinates.
**Strategy**:
Plot the curve in the $r-z$ plane, and rotate to a 3D volumne.
The volume looks like an ice-cream cone. What is the maximum inclination angle $\phi$? Remember that $r$ in both coordinates are not the same.

b) Use Cylindrical Coordinates. Which method do you think is more suited?
**Strategy**:
Solve the intersection of the surfaces by pluging a quantity of one equation into the other. Perform a 2D integral in $r-z$ plane, and add the $\theta-$integral on the outside.

c) Now consider surface C as the plane $z=\sqrt{3}$. What is the volume between surface A and C? Which method should you use?

**Question 44**

Compute the volume above the cone $z=\sqrt{x^2+y^2}$ and below the sphere $x^2+y^2+z^2=2z$.

$a$) Use Cylindrical Coordinates.
**Strategy:**
Why is it a sphere?
Plot the curve on the $r-z$ plane. For the sphere, complete the square in $z$.
Find the intersection and perform a 2D integral in $r-z$ plane.

$b$) Use Spherical Coordinates. Which way do you prefer?

**Strategy:**
Since the sphere is **not** centered at the origin, the range of $\rho$ is **NOT** 0 to 1.
Use the defining equation $x^2+y^2+z^2=2z$ and switch variables to spherical coordinates to attain $\rho = 2cos(\phi)$. Notice $\rho$ varies in terms of the value $\phi$ in 3D, similar to polar curves in 2D, so we **must** put it as the innermost integral.

4. **Ellipse Transformation**
To integrate over an ellipse, we write the equation of ellipse as

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1$$

Similar to polar coordinates, we use

$$g(r,\theta) = \big(a \cdot rcos(\theta), b \cdot rsin(\theta)\big)$$

$Det(D_g) =$ **abr**
Since the stretching factors of the ellipse are included in function $g$, the range of radius is **always** $r : 0 \to 1$.

$\theta$ does not exactly represent the angle from polar coordinates anymore.
It still works as normal for these angles: $\theta = 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}, 2\pi...$
A full ellipse is still $\theta : 0 \to 2\pi$.

Useful when a 2D domain is an ellipse, or if the 3D projection of a volume is of ellipse shape.

**Question 45**
Verify $Det(D_g) =$ **abr**.
Then find the area of the ellipse given by the standard equation above.

**Question 46**
Integrate $f(x, y) = \sin(2x^2 + 4y^2)$ over the region bounded by $2x^2 + 4y^2 = 4$.

**Question 47**
Compute the volume above $z = \sqrt{2x^2 + y^2}$ and below $z = 2$.

**Strategy:**
This surface is **not given by a circular rotation**, with $r^2 = x^2 + y^2$. However, it is similar. We may take the "radius" to be $r^2 = 2x^2 + y^2$, so we can still plot on the $r - z$ plane and do an **ellipse rotation**. Taking the projection of the volume onto the xy-plane gives an ellipse. Instead of integrating the 2D projection integral directly, we first perform the inner $z$ integral, and use ellipse transformation onto the 2D integral.

5. **Parallelogram Transformation**
Given a parallelogram in $\mathbb{R}^2$ with one vertex being origin.
It is generated by 2 points/vectors, $\vec{a}$ and $\vec{b}$.
We can turn the unit square into parallelogram by requiring that

$$g\begin{bmatrix}1\\0\end{bmatrix} = \vec{a} \qquad g\begin{bmatrix}0\\1\end{bmatrix} = \vec{b}$$

Since the parallelogram is linear, this is a **linear transformation**.
It is characterized by a matrix.

$$g\begin{bmatrix}u\\v\end{bmatrix} = A \cdot \begin{bmatrix}u\\v\end{bmatrix} = \begin{bmatrix}x\\y\end{bmatrix}$$

Due to the contraint on the unit vectors above, A must be

$$A = \begin{bmatrix}| & |\\ \vec{a} & \vec{b}\\ | & |\end{bmatrix}$$

If the paralellogram does not have one vertex at origin, we can perform a shift of the parallelogram, by subtracting the coordinates of one point from all points of the parallelogram. We take note of $\vec{a}$ and $\vec{b}$ that generate the **shifted** parallelogram. The transformation $g$ would take 2 steps. First the transformation of the square to the shifted parallelogram, then from the shifted parallelogram to the original one.

**Example**
Given the parallelogram $(1, 1), (2, 1), (2, 2), (3, 2)$ which is a standard 45 degrees parallelogram that is shifted.
The standard 45 degrees parallelogram is defined by the vectors $\begin{bmatrix}1\\0\end{bmatrix}$ and $\begin{bmatrix}1\\1\end{bmatrix}$.
We can use the above transformation, to turn the unit square into the standard parallelogram:

$$h\begin{bmatrix}u\\v\end{bmatrix} = \begin{bmatrix}1 & 1\\0 & 1\end{bmatrix} \cdot \begin{bmatrix}u\\v\end{bmatrix}$$

Then we need to shift the parallelogram upward and right by 1 unit. Thus

$$g\begin{bmatrix}u\\v\end{bmatrix} = \begin{bmatrix}1 & 1\\0 & 1\end{bmatrix} \cdot \begin{bmatrix}u\\v\end{bmatrix} + \begin{bmatrix}1\\1\end{bmatrix}$$

In both the non-shifted and the shifted case, $D_g = A$.

6. **Tetrahedron Transformation**
Generalizing the idea of parallelogram, the tetrahedron transformation is much more useful in practice.

Given a tetrahedron in $\mathbb{R}^3$ with one vectex being origin.
It is characterized by 3 points/vectors, $\vec{a}$, $\vec{b}$, $\vec{c}$.
We can turn the **unit tetrahedron** into the tetrahedron given by requiring that

$$g\begin{bmatrix}1\\0\\0\end{bmatrix}=\vec{a} \qquad g\begin{bmatrix}0\\1\\0\end{bmatrix}=\vec{b} \qquad g\begin{bmatrix}0\\0\\1\end{bmatrix}=\vec{c}$$

If we define the names of the variables of $g$ as $g(u,v,w)$, then our domain of integration is the unit tetrahedron, given by

$$\int_{w=0}^{w=1}\int_{v=0}^{v=1-w}\int_{u=0}^{u=1-v-w} f\circ g\cdot|Det(D_g)| \quad du \quad dv \quad dw$$

Of course, $g$ is again a linear transformation.

$$g\begin{bmatrix}u\\v\\w\end{bmatrix}=A\cdot\begin{bmatrix}u\\v\\w\end{bmatrix}=\begin{bmatrix}x\\y\\z\end{bmatrix}$$

where the matrix A, by the constraints above, is

$$A=\begin{bmatrix}| & | & |\\ \vec{a} & \vec{b} & \vec{c}\\ | & | & |\end{bmatrix}$$

If the tetrahedron does not have one vertex at origin, we perform a shift similar to parallelogram. Same as the parallelogram transformation, in both the non-shifted and the shifted case, $D_g = A$.

**Question 48**
Find the volume bounded by $x+2y+z=2$, $x=2y$, $x=0$, $z=0$.
**Strategy:**
(Take any 3 planes, they intersect at 1 point. 4 points form a tetrahedron.)

**Question 49**
Compute the integral $\int_D z-x-y$ where $D$ is the tetrahedron given by $(0,0,0),(1,2,3),(0,2,2),(-1,1,1)$.

7. **Domain Bounded by 2 Expressions**
Sometimes, a 2D domain B is defined by the region bounded by some curves given by: $f_1(x,y) = a, f_1(x,y) = b,\ f_2(x,y) = c, f_2(x,y) = d.$

If the curves only intersect at 4 points, we can transform the square onto the domain by using

$$h \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f_1(x,y) \\ f_2(x,y) \end{bmatrix} = \begin{bmatrix} u \\ v \end{bmatrix}$$

and since $u = f_1(x,y)$ and $v = f_2(x,y)$, the bounds of $\begin{bmatrix} u \\ v \end{bmatrix}$ is exactly the square $A = \{u \in (a,b),\ v \in (c,d)\}$.

However, the direction of $h$ is **opposite** to the direction neeeded in the change of variable formula. To use the change of variable formula, we technically should compute the inverse of $h$, giving

$$g \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

However, computing inverse is sometimes difficult.
Since we need $det(D_g)$ in the formula, and $g = h^{-1}$, we can instead use

$$det(D_g) = det(D_{h^{-1}}) = det((D_h)^{-1}) = det(D_h)^{-1} = \frac{1}{det(D_h)}$$

Thus we can simply compute $D_h$ which is given, and use it in the change of variable formula without needing to compute the inverse.
(We still need to compute the inverse sometimes.)
**Example**
Let B given by $xy = 1, xy = 2,\ xy^2 = 1, xy^2 = 2$. Then set $u = xy, v = xy^2$.

$$D_h = \begin{bmatrix} y & x \\ y^2 & 2xy \end{bmatrix} \qquad \Rightarrow \qquad det(D_h) = 2xy^2 - xy^2 = xy^2 = v$$

Suppose we want to integrate $f(x,y) = y$, then this correspond to $y = v/u$.

$$\int_B y = \int_A \frac{v}{u} \cdot |\frac{1}{v}| = \int_{u=1}^{2} \int_{v=1}^{2} \frac{1}{u}\, dv\, du$$

**Question 50**
1. $\int_D (x+y)(x-y)$ over $D$ bounded by $x-y=0$, $x-y=2$, $x+y=1$, $x+y=2$.
2. $\int_D \sin\frac{2x-y}{2x+y}$ over $D$ bounded by $x=0$, $y=0$, $2x+y=1$.