# Contents

# 1 Recap: H2 Math Vectors (inc. extension)

The relationship between points, lines and planes (i.e. distance between two planes) is considered trivial for this discussion. It will not be discussed.

## 1.1 Dot Product

The dot product of two vectors $\mathbf{a}$ and $\mathbf{b}$ is defined by $\mathbf{a} \bullet \mathbf{b} = |\mathbf{a}|\,|\mathbf{b}| \cos\theta$ where $0 \leq \theta \leq \pi$. Two vectors are said to be orthogonal if and only if the angle between them is $0°$. That is,

$$\mathbf{a} \bullet \mathbf{b} = 0 \Leftrightarrow \mathbf{a} \text{ and } \mathbf{b} \text{ are orthogonal.}$$

## 1.2 Triangle Inequality

The triangle inequality states that for vectors $\mathbf{a}$ and $\mathbf{b}$, then we have the following inequality:

$$|\mathbf{a} + \mathbf{b}| \leq |\mathbf{a}| + |\mathbf{b}|$$

with equality attained if and only if the triangle formed is degenerate (i.e. the three vertices are collinear).

*Proof:* Expand the dot product $(\mathbf{a} + \mathbf{b}) \bullet (\mathbf{a} + \mathbf{b})$ and note that $|\cos\theta| \leq 1$. □

Another form the Triangle Inequality, which uses the Reverse Triangle Inequality states that

$$|\mathbf{a} - \mathbf{b}| \geq ||\mathbf{a}| - |\mathbf{b}||\,.$$

*Proof:*

$$\begin{aligned}
|\mathbf{a}| &= |(\mathbf{a} - \mathbf{b}) + \mathbf{b}| \\
&\leq |\mathbf{a} - \mathbf{b}| + |\mathbf{b} \\
|\mathbf{a}| - |\mathbf{b}| &\leq |\mathbf{a} - \mathbf{b}|
\end{aligned}$$

On the other hand,

$$\begin{aligned}
|\mathbf{b}| &= |(\mathbf{b} - \mathbf{a}) + \mathbf{a}| \\
&\leq |\mathbf{b} - \mathbf{a}| + |\mathbf{a} \\
|\mathbf{b}| - |\mathbf{a}| &\leq |\mathbf{b} - \mathbf{a}|
\end{aligned}$$

Since $|\mathbf{a} - \mathbf{b}| = |\mathbf{b} - \mathbf{a}|$, then the result follows. □

## 1.3 Cauchy-Schwarz Inequality

The Cauchy-Schwarz Inequality states that

$$|\mathbf{a} \bullet \mathbf{b}| \leq |\mathbf{a}|\,|\mathbf{b}|$$

which can be easily proven using the definition of the dot product.

## 1.4 Cross Product and Determinant

For two vectors $\mathbf{a}$ and $\mathbf{b}$, where $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$and $\mathbf{b} = \langle b_1, b_2, b_3 \rangle$, the cross product $\mathbf{a} \times \mathbf{b}$ is given by the determinant of the following matrix:

$$\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} \mathbf{i} - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} \mathbf{j} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \mathbf{k}$$

### 1.4.1 Triple Product

We first discuss some properties of the scalar triple product for vectors $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$.
**(1):** The scalar triple product is unchanged under a circular shift of its three operands $(\mathbf{a}, \mathbf{b}, \mathbf{c})$:

$$\mathbf{a} \bullet (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \bullet (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \bullet (\mathbf{a} \times \mathbf{b})$$

**(2):** Swapping the positions of the operators without re-ordering the operands leaves the triple product unchanged. This follows from the preceding property and the commutative property of the dot product:

$$\mathbf{a} \bullet (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \bullet \mathbf{c}$$

**(3):** If any two vectors in the scalar triple product are equal, then its value is zero:

$$\mathbf{a} \bullet (\mathbf{a} \times \mathbf{b}) = \mathbf{a} \bullet (\mathbf{b} \times \mathbf{a}) = \mathbf{a} \bullet (\mathbf{b} \times \mathbf{b}) = \mathbf{b} \bullet (\mathbf{a} \times \mathbf{a}) = 0$$

Next, we discuss an important property of the vector triple product, also known as Lagrange's Formula:
**(1):** Lagrange's Formula

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \bullet \mathbf{c})\mathbf{b} - (\mathbf{a} \bullet \mathbf{b})\mathbf{c}$$

**(2):** Jacobi's Identity, which is a corollary of Lagrange's Formula

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = \mathbf{0}$$

# 2 Quadric Surfaces

## 2.1 General Equation of a Quadric Surface

A quadric surface is the graph of a second degree equation in $x$, $y$ and $z$. That is,

$$Ax^2 + By^2 + Cz^2 + Dxy + Eyz + Fxz + Gx + Hy + Iz + J = 0.$$

Using translation and rotation, the equation can be expressed in one of the following two standard forms:

$$Ax^2 + By^2 + Cz^2 + J = 0 \text{ and } Ax^2 + By^2 + Iz = 0$$

We will study a total of fifteen quadric surfaces in this section. Even though it may seem like there are many characteristics which we need to know such as the shape and critical points, it is crucial that we draw similarities between these quadric surfaces and the four conic sections - namely the circle, ellipse, parabola and hyperbola. Along the way, we will provide certain techniques to remember them.

## 2.2 Non-Degenerate Real Quadric Surfaces

### 2.2.1 Ellipsoid

The equation of an ellipsoid is given by

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

Recall that the above is similar to the equation of an ellipse centered at the origin with semi-major axis $a$ and semi-minor axis $b$ (that is $a > b$). If we have the equation

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

we get either an oblate or a prolate spheroid (special ellipsoid). From the diagrams, we observe that the lines $z = k$, where $k$ is a real constant cut the ellipsoid, forming concentric circles. That is, circles sharing the same centre. This is true because each concentric circle is lying on the $xy$-plane and the vector with **k**-component $k$ is a normal vector to the plane.

If we have the equation

$$x^2 + y^2 + z^2 = a^2,$$

we have a sphere (special spheroid). This should be easy to remember as it as strong semblance to the equation of a circle.

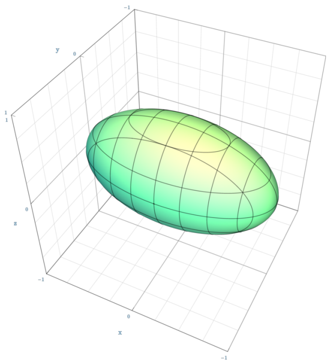

Figure 1: Ellipsoid

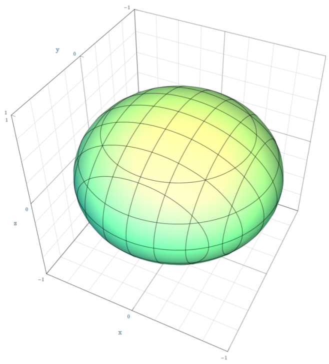

Figure 2: Oblate Spheroid

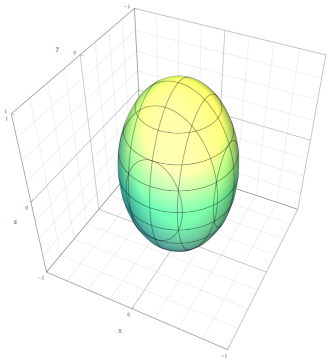

Figure 3: Prolate Spheroid

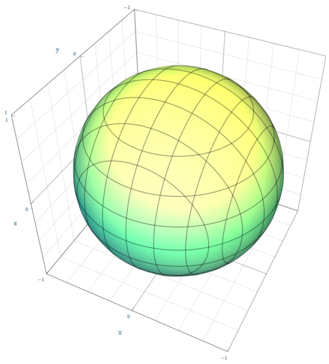

Figure 4: Sphere

### 2.2.2 Elliptic Paraboloid

The equation of an elliptic paraboloid is given by

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-z=0.$$

The way to remember this is by first making $z$ the subject, which yields

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

We note that the right side of the equation resembles that of an ellipse, which implies that any line $z=k$ that slices the elliptic paraboloid results in an ellipse to be obtained. As $x, y \geq 0$, it implies that when $x=y=0$, we obtain the minimum value of $z$, which is 0. Making reference to the $xy$-plane, as we go higher (i.e. $z$-coordinate increases), we observe that the semi-major axis and the semi-minor axis of the ellipses increase. This is a good way to remember how to sketch the elliptic paraboloid.

A circular paraboloid is a special case of an elliptic paraboloid. The former is given by the equation

$$\frac{x^2}{a^2}+\frac{y^2}{a^2}-z=0.$$

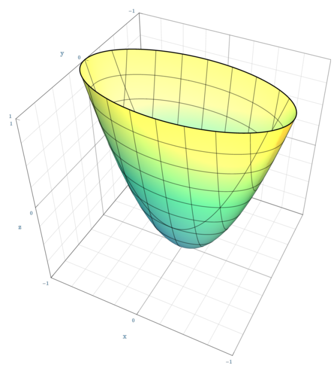

Figure 5: Elliptic Paraboloid

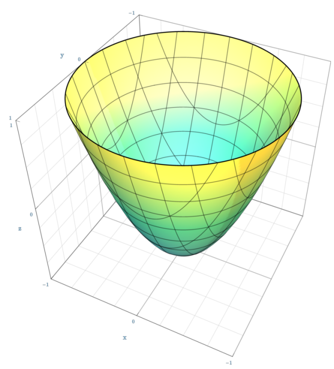

Figure 6: Circular Paraboloid

### 2.2.3 Hyperbolic Paraboloid

The equation of a hyperbolic paraboloid is given by

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - z = 0.$$

A fun fact before we proceed with the properties of the hyperbolic paraboloid is that potato chips (I love sour cream and onion flavour) have such a shape. There is a saddle point (will be discussed in due course) on the quadric surface which allows easier stacking of chips.

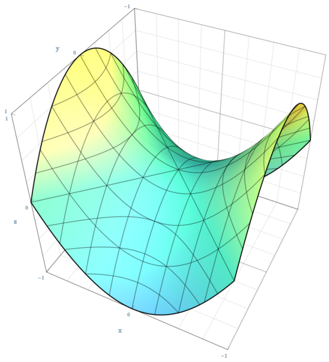

Figure 7: Hyperbolic Paraboloid

We first observe that

$$z = \frac{x^2}{a^2} - \frac{y^2}{b^2}.$$

Without a loss of generality, we set a line on the red arrow to be the $x$-axis and a line on the blue arrow to be the $y$-axis, where these two lines are of course orthogonal to each other and their intersection is the origin. The intersection coincides with the saddle point. Observe that as $x$ increases, $z$ increases and as $y$ increases, $z$ increases. Regardless of the polarity of $x$ and $y$, it would not affect $z$ since $x^2$ and $y^2$ will always be non-negative. From here, observe that the loci of points traced by the red arrow and blue arrow form two hyperbolas opening in different directions.

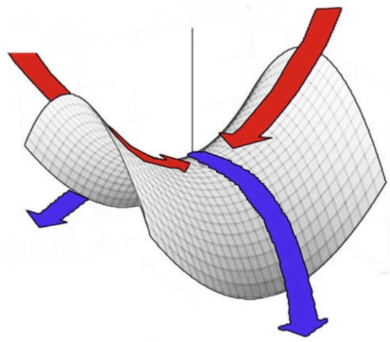

Figure 8: Features of Hyperbolic Paraboloid

### 2.2.4 Hyperboloid of One Sheet

The equation of a hyperboloid of one sheet or a hyperbolic hyperboloid is given by

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

Rearranging, we have

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

When $z = 0$, we obtain an ellipse on the $xy$-plane centered at the origin with semi-major axis $a$ and semi-minor axis $b$ (assmuing that $a > b$). As $z$ increases, then the semi-major and semi-minor axes increase as well, and as such, we obtain bigger ellipses. The same argument can be applied to the case where $z$ decreases for $z < 0$ because $z^2$ will increase too.

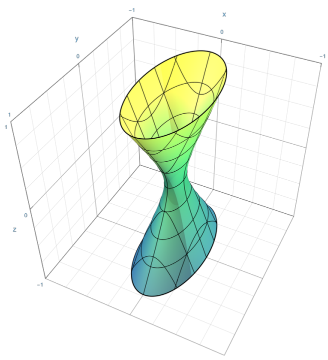

Figure 9: Hyperboloid of One Sheet

### 2.2.5 Hyperboloid of Two Sheets

The equation of a hyperboloid of two sheets or an elliptic hyperboloid is given by

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

Rearranging,

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

Note that

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}\geq 0 \implies z^2\geq c^2 \implies z\leq -c \text{ or } z\geq c.$$

Hence, we cannot sketch the quadric surface for $-c<z<c$. In a similar fashion to hyperboloids of one sheet, if we set the left side of the equation to be a constant $k$, then we obtain the equation of the ellipse, which semi-major and semi-minor axis increase for larger values of $k$. Note that $k$ is affected by $z^2$ so as $z^2$ increases, $k$ increases too.

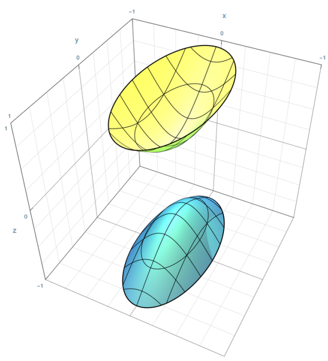

Figure 10: Hyperboloid of Two Sheets

## 2.3 Degenerate Real Quadric Surfaces

### 2.3.1 Elliptic Cone

The equation of an elliptic cone is given by

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=0.$$

A good way to remember this is by noting that the equation of an elliptic cone resembles that of an elliptic paraboloid.

A circular cone with equation

$$\frac{x^2}{a^2}+\frac{y^2}{a^2}-\frac{z^2}{b^2}=0$$

can be obtained from the previous one.

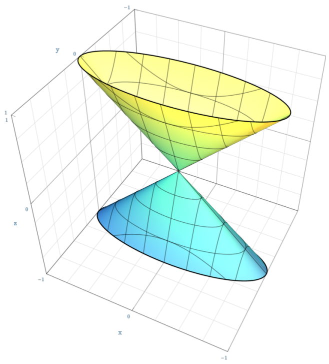

Figure 11: Elliptic Cone

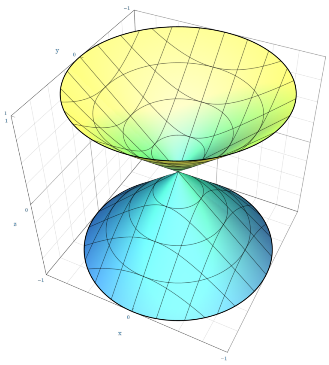

Figure 12: Circular Cone

### 2.3.2 Elliptic Cylinder

The equation of an elliptic cylinder is given by

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

and that of a circular cylinder is given by

$$x^2 + y^2 = a^2.$$

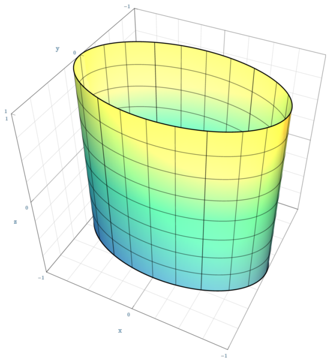

Figure 13: Elliptic Cylinder

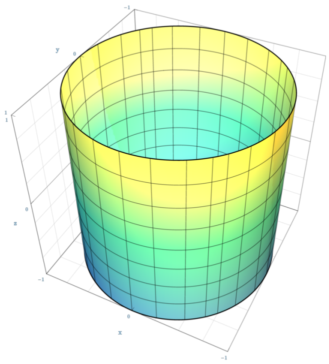

Figure 14: Circular Cylinder

Note that these resemble the ellipse and circle respectively, just that now we are dealing in three-dimensional space.

### 2.3.3 Hyperbolic Cylinder

The equation of a hyperbolic cylinder is given by

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

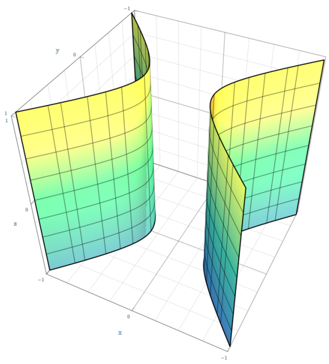

Figure 15: Hyperbolic Cylinder

### 2.3.4 Parabolic Cylinder

A parabolic cylinder is given by the equation

$$x^2 + 2ay = 0.$$

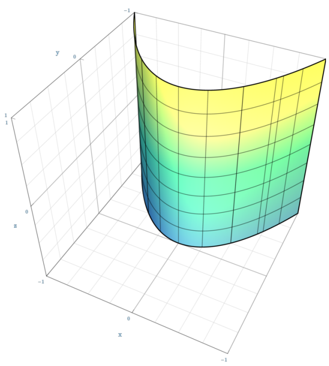

Figure 16: Parabolic Cylinder

# 3 Cylindrical and Spherical Coordinates

## 3.1 Polar Coordinates

The polar coordinates are given by $(r, \theta)$, where $r$ is known as the radius vector and $\theta$ is known as the vectorial angle.

**Conversion to Polar Coordinates**

The conversion formulae are $x = r\cos\theta$ and $y = r\sin\theta$, thus giving

$$x^2 + y^2 = r^2 \text{ and } \tan\theta = \frac{y}{x},$$

provided that $x \neq 0$.

Using a change of variables to polar coordinates, certain limits can be solved.
*Example:* For instance, we wish to compute the following limit by using polar coordinates:

$$\lim_{(x,y)\to(0,0)} \frac{x^3 + y^3}{x^2 + y^2}$$

*Solution:* Note that if we use Cartesian coordinates, we cannot obtain a solution since both the numerator and denominator tend to zero. Using polar coordinates, as $x, y \to 0$, then $r, \theta \to 0$. Hence,

$$\begin{aligned}
\lim_{(x,y)\to(0,0)} \frac{x^3 + y^3}{x^2 + y^2} &= \lim_{(r,\theta)\to(0,0)} \frac{r^3\cos^3\theta + r^3\sin^3\theta}{r^2\cos^2\theta + r^2\sin^2\theta} \\
&= \lim_{(r,\theta)\to(0,0)} \left(r\cos^3\theta + r\sin^3\theta\right) \\
&= 0
\end{aligned}$$

□

## 3.2 Cylindrical Coordinates

Suppose we want to convert Cartesian coordinates $(x, y, z)$ to cylindrical coordinates $(r, \theta, z)$.

**Conversion to Cylindrical Coordinates**

For the $x$- and $y$- components, they are similar to polar coordinates. Thus, we have

$$x = r\cos\theta, \ y = r\sin\theta \text{ and } z = z.$$

## 3.3 Spherical Coordinates

For spherical coordinates, it is slightly more complicated. This time, we will convert from $(x, y, z)$ to $(\rho, \theta, \phi)$, where in a similar fashion, $\rho$ is the radius and $\theta$ is the vectorial angle/inclination. The angle $\phi$ is known as the azimuthial angle, which can be commonly thought as the angle between the coordinate $P(x, y, z)$ and the $z$-axis.

**Conversion to Spherical Coordinates**

The conversion formulae state that

$$x = \rho\sin\phi\cos\theta, \ y = \rho\sin\phi\sin\theta \text{ and } z = \rho\cos\phi,$$

where $\rho \geq 0$, $0 \leq \theta \leq 2\pi$ and $0 \leq \phi \leq \pi$. Observe that three equations satisfy the identity $\rho^2 = x^2 + y^2 + z^2$.

One way as to how I like to remember this formula (but understand its derivation first) is that observe that $x$ and $y$ both contain the $\rho\cos\theta$ and $\rho\sin\theta$ components respectively, which resemble polar coordinates. Next, both $x$ and $y$ contain the term $\sin\phi$, which also appears in the calculation of its Jacobian determinant. If $x$ and $y$ contain $\sin\phi$, then $z$ must contain $\cos\phi$!

Just to jump the gun, the Jacobian determinant for spherical coordinates, $\det(\mathbf{J})$, is given by $\det(\mathbf{J}) = \rho^2\sin\phi$. This idea will be formally introduced later.

Some limits can be solved by using a change of variables to spherical coordinates.

*Example:* Suppose we wish to find the following limit:

$$\lim_{(x,y,z)\to(0,0,0)} \frac{xyz}{x^2+y^2+z^2}.$$

*Solution:* As $x, y, z \to 0$, then it is clear that $\rho, \theta, \phi \to 0$ too.

$$\begin{aligned}
\lim_{(x,y,z)\to(0,0,0)} \frac{xyz}{x^2+y^2+z^2} &= \lim_{(\rho,\theta,\phi)\to(0,0,0)} \frac{(\rho\sin\phi\cos\theta)(\rho\sin\phi\sin\theta)(\rho\cos\phi)}{(\rho\sin\phi\cos\theta)^2+(\rho\sin\phi\sin\theta)^2+(\rho\cos\phi)^2} \\
&= \lim_{(\rho,\theta,\phi)\to(0,0,0)} \frac{\rho^3\cos\phi\sin^2\phi\cos\theta\sin\theta}{\rho^2} \\
&= \lim_{(\rho,\theta,\phi)\to(0,0,0)} \rho\cos\phi\sin^2\phi\cos\theta\sin\theta \\
&= 0
\end{aligned}$$

□

# 4 Vector Functions

## 4.1 Definition of a Vector Function

A vector function $\mathbf{r}(t)$ is a function whose domain is a set of real numbers and whose range is a set of vectors. It can be written as

$$\mathbf{r}(t) = \langle f(t), g(t), h(t)\rangle.$$

As $t$ tends to $a$, then the limit of $\mathbf{r}(t)$ is defined by

$$\lim_{t\to a}\mathbf{r}(t) = \left\langle \lim_{t\to a} f(t), \lim_{t\to a} g(t), \lim_{t\to a} h(t)\right\rangle.$$

Standard properties like linearity and scalar multiplicativity are considered *trivial* in this context. One important property in relation to the cross product is

$$\frac{d}{dt}(\mathbf{u}(t)\times\mathbf{v}(t)) = \mathbf{u}(t)\mathbf{v}'(t) + \mathbf{u}'(t)\mathbf{v}(t)$$

whose proof is rather rigorous. Note that this formula is analogous to the product rule.

## 4.2 Derivative of a Vector Function

For a vector function $\mathbf{r}(t)$, the derivative is given by

$$\frac{d\mathbf{r}}{dt} = \mathbf{r}'(t) = \lim_{h\to 0}\frac{\mathbf{r}(t+h)-\mathbf{r}(t)}{h}$$

which is a consequence of First Principles.

# 5 Functions of Several Variables

A function $f$ of 2 variables is a rule that assigns to each ordered pair of real numbers $(x, y)$ in a set $D$ a unique real number denoted by $f(x, y)$. Here, $D$ is called the domain of $f$. The set of values that $f$ takes on is called the range of $f$. That is, $R_f = \{f(x, y) \mid (x, y) \in D\}$.

Note that the graph of $f$ is the set of all points $(x, y, z) \in \mathbb{R}^3$ such that $z = f(x, y)$.

For 3 or more variables, we construct a mapping from $\mathbb{R}^3$ to $R$. Let $f : D \subseteq \mathbb{R}^3 \to R$ be a function of 3 variables. We can describe $f$ by examining its *level surfaces*. These are surfaces in $\mathbb{R}^3$ given by the equations $f(x, y, z) = K$, where $K \in \mathbb{R}$.

## 5.1 Level Curves

The level curves of a function of 2 variables are the curves in the $xy$-plane with equation $f(x, y) = K$, where $K$ is a constant (i.e. $K$ is in the range of $f$).

## 5.2 Limits and Continuity

**$\varepsilon$-$\delta$ Definition of a Limit**

Let $f$ be a function of two variables whose domain $D$ includes points arbitrarily close to $(a, b)$. We say that the limit of $f(x, y)$ as $(x, y)$ approaches $(a, b)$ is $L$ and we write

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

if for any positive number $\varepsilon$, there is a corresponding positive number $\delta$ such that

$$(x, y) \in D \text{ and } 0 < \sqrt{(x-a)^2 + (y-b)^2} < \delta \Rightarrow |f(x, y) - L| < \varepsilon.$$

This definition can also be applied to a function which is defined parametrically. That is, $f(x, y) = f(g(t), h(t))$.

If the limit $L$ exists, then **(i)** its value is unique and **(ii)** $L$ is independent of the choice of any path approaching $(a, b)$. The latter is related to a test (known as the two path test) to determine if a limit exists.

To prove that the limit exists, we have to use the formal $\varepsilon$-$\delta$ definition. A common trick to simplify the inequality in a particular step would be to use the Triangle Inequality. On the other hand, to prove that the limit does not exist, we need to obtain a contradiction. This is where the two path test comes in handy. For example, the function might have two different values when approaching along the line $y = x$ and the curve $y = x^2$. This technique, though requires some intuition, is useful.

A function $f$ of 2 variables is said to be continuous at $(a, b)$ if

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

$f$ is continuous on $D \in \mathbb{R}^2$ if $f$ is continuous at each point $(a, b)$ in $D$.

### 5.2.1 Smoothness

Sidetrack to a bit of Analysis, the term *smoothness* is commonly associated with differentiability and continuity. Consider an open set on $\mathbb{R}$ and a function $f$ defined on that set with real values. Let $k \in \mathbb{N}$. The function $f$ is said to be of class $C^k$ if the derivatives $f', f'', \ldots, f^{(k)}$ exist and are continuous. $f$ is said to be infinitely differentiable, smooth, or of class $C^\infty$, if it has derivatives of all orders. $f$ is said to be of class $C^\omega$, or *analytic*, if $f$ is smooth and if its Taylor Series expansion around any point in its domain converges to the function in some neighborhood of the point. $C^\omega$ is thus strictly contained in $C^\infty$.

Simply said, $C^0$ is the class of functions for which $f$ is continuous, $C^1$ is the class of functions for which $f$ and $f'$ are continuous, $C^2$ is the class of functions for which $f, f'$ and $f''$ are continuous and so on...

### 5.2.2 Relationship between Differentiability and Continuity

Back in Junior College, a friend of mine, who was a Chinese scholar, sent me this meme. The foreground translates to 'falling implies continuity' and the background translates to 'continuity does not imply falling', which is similar when talking about the relationship between continuity and differentiability, for which if a function $f$ is differentiable, then it must be continuous. However, the converse is not necessarily true!

*Example:* For example, it is clear that $|x|$ is continuous on $\mathbb{R}$. However, it is not differentiable at $x = 0$.

*Proof:* $y = |x|$ can be defined by a piecewise function. For $x \geq 0$, $|x| = x$, whereas for $x < 0$, $|x| = -x$. Considering the left limit and right limit as $\delta x \to 0$,

$$\begin{aligned}\lim_{\delta x \to 0^-} \frac{f(0+\delta x) - f(0)}{\delta x} &= \lim_{\delta x \to 0^-} \frac{|\delta x|}{\delta x} \\ &= \lim_{\delta x \to 0^-} -1 \\ &= -1\end{aligned}$$

$$\begin{aligned}\lim_{\delta x \to 0^+} \frac{f(0+\delta x) - f(0)}{\delta x} &= \lim_{\delta x \to 0^+} \frac{|\delta x|}{\delta x} \\ &= \lim_{\delta x \to 0^+} 1 \\ &= 1\end{aligned}$$

Since the left limit and right limit are different, the result follows. □


Figure 17: Bicycle meme

## 5.3 Partial Derivatives

Let $f$ be a function of 2 variables, namely $x$ and $y$. The partial derivative of $f$ with respect to $x$ is denoted by $\frac{\partial f}{\partial x}$ or $f_x$ and that with respect to $y$ is denoted by $\frac{\partial f}{\partial y}$ or $f_y$. Using First Principles, we have

$$\frac{\partial f(a,b)}{\partial x} = \lim_{h \to \infty} \frac{f(a+h,b) - f(a,b)}{h} \text{ and } \frac{\partial f(a,b)}{\partial y} = \lim_{h \to \infty} \frac{f(a+h,b) - f(a,b)}{h}.$$

Take $f_x$ for instance. This means that we differentiate $f$ with respect to $x$, while treating everything else as a constant. This is unlike the conventional differentiation problems that we have been exposed to in Secondary School and Junior College where we treat other letters like $y$, $u$ and $v$ as variables. However, what is the geometric interpretation? It turns out that $f_x(a,b)$ measures the rate of change of $f$ in the direction of $\mathbf{i}$ at the point $(a,b)$.

We can also find expressions for higher order derivatives like $\frac{\partial^3 f}{\partial x^3}$. The list below shows how we can express them:

$$\begin{aligned}
f_{xx} &= \frac{\partial^2 f}{\partial x^2} \\
f_{xy} &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x} \\
f_{yx} &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y}
\end{aligned}$$

Some of us would have had exposure to ordinary differential equations (ODEs) in Junior College. Here, partial derivatives have their own type of differential equations too, known as partial differential equations (PDEs). This imposes relations between the various partial derivatives of a multivariable function.

## 5.4 Clairaut's Theorem

**Clairaut's Theorem**

Let $f$ be defined on an open disk containing the point $(a,b)$. If $f_{xy}$ and $f_{yx}$ are continuous at $(a,b)$, then

$$f_{xy}(a,b) = f_{yx}(a,b).$$

It is also known as the *symmetric property of the second derivatives.*

We provide a proof of Clairaut's Theorem using the Mean Value Theorem. It can also be proven using iterated integrals which involves the Fubini-Tonelli Theorem. The latter would be discussed in one of the upcoming sections.

*Proof:* First, assume that $f_{xy}$ and $f_{yx}$ are defined on a small open disk $D$ centered at $(a,b)$. Let $(x,y)$ be a point in $D$. Fix $x$ and consider the function

$$[f(x,y) - f(a,y)] - [f(x,b) - f(a,b)]$$

in $y$. Applying the Mean Value Theorem with respect to $y$ yields

$$[f(x,y) - f(a,y)] - [f(x,b) - f(a,b)] = [f_y(x,\zeta_1) - f_y(a,\zeta_2)]\,(y-b)$$

for some $\zeta_1$ between $y$ and $b$. Applying the Mean Value Theorem to $f_y(x,\zeta_1)$ with respect to $x$,

$$[f(x,y) - f(a,y)] - [f(x,b) - f(a,b)] = f_{yx}(\zeta_2,\zeta_1)(x-a)(y-b)$$

for some $\zeta_2$ between $x$ and $a$. Note that

$$[f(x,y) - f(a,y)] - [f(x,b) - f(a,b)] = [f(x,y) - f(x,b)] - [f(a,y) - f(a,b).]$$

Applying the Mean Value Theorem first with respect to $x$, then with respect to $y$, we have

$$\begin{aligned}
[f(x,y) - f(a,y)] - [f(x,b) - f(a,b)] &= [f(x,y) - f(x,b)] - [f(a,y) - f(a,b)] \\
&= f_{xy}(\zeta_3,\zeta_4)(y-b)(x-a)
\end{aligned}$$

where $\zeta_3$ is between $x$ and $a$, and $\zeta_4$ is between $y$ and $b$. Thus,

$$f_{xy}(\zeta_2,\zeta_1) = f_{yx}(\zeta_3,\zeta_4).$$

Since $f_{xy}$ and $f_{yx}$ are continuous on $(a,b)$, by taking the limit as $(x,y)$ tends to $(a,b)$, we obtain Clairaut's Theorem $\square$

## 5.5 Tangent Plane

Let $f$ be a function of two variables. The graph of $f$ is a surface in $\mathbb{R}^3$ with equation $z = f(x, y)$. Let $P(x_0, y_0, z_0)$ be a point on this surface. Thus, $z_0 = f(x_0, y_0)$. Assuming a tangent plane to the surface exists, its equation is given by

$$z = z_0 + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - x_0).$$

*Proof:* Any plane passing through $P(x_0, y_0, z_0)$ has the following equation:

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0$$

We assume that the plane is not vertical, implying that $C$ is non-zero. Then, the equation of the plane is

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

The tangent line to $C_1$ at $P$ in the $x$-direction has a gradient of $a$, which by the geometric interpretation of the partial derivative, can be represented by $a = f_x(x_0, y_0)$. Similarly, $b = f_y(x_0, y_0)$. Substituting $a$ and $b$ into the equation of the plane yields the result. □

## 5.6 Linear Approximation

Note that the tangent plane to the surface $z = f(x, y)$ at $P(x, y, z)$ is very close to the surface at least when it is near $P$. Hence, we may use the function defining the tangent plane as a linear approximation to f. At the point $(x, y) = (a, b)$, note that $P$ has coordinates $(a, b, f(a, b))$. The linear function $L$ whose graph is this tangent plane is given by

$$L(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

where $L$ is known as the linearisation of $f$ at $(a, b)$. Therefore, the approximation

$$f(x, y) \approx L(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

is called the linear approximation of $f$ at $(a, b)$.

## 5.7 The Differential

The term differential refers to an infinitesimal change in some varying quantity. In single variable Calculus, we know that if $y$ is a function of $x$, then the differential of $y$, known as $dy$, is related to $dx$ by the following formula:

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

In Multivariable Calculus, since we are dealing with partial derivatives, we have a similar formula for the differential, $dz$, where $z = f(x, y)$. It is defined to be

$$dz = f_x(x, y)\, dx + f_y(x, y)\, dy.$$

The actual change $\Delta z$ of $z$ satisfies the equation $\Delta z \approx dz$ as a consequence of the tangent plane approximation. $f$ is said to be differentiable at $(a, b)$ if

$$\Delta z = f_x(a, b)\, \Delta x + f_y(a, b)\, \Delta y + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y$$

where

$$\lim_{(\Delta x, \Delta y) \to (0,0)} \varepsilon_1 = 0 \text{ and } \lim_{(\Delta x, \Delta y) \to (0,0)} \varepsilon_2 = 0.$$

If $f_x(x, y)$ and $f_y(x, y)$ exist in an open disk containing $(a, b)$ and are continuous at $(a, b)$, then $f$ is differentiable at $(a, b)$.

## 5.8 Chain Rule

Suppose $z = f(x, y)$ is a differentiable function of $x$ and $y$, where $x = g(t)$ and $y = h(t)$ are both differentiable functions of $t$. Then $z$ is a differentiable function of $t$ and

$$\frac{dz}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}.$$

### 5.8.1 Implicit Differentiation

Suppose $F(x, y) = 0$ defines $y$ implicitly as a function of $x$. That is $y = f(x)$. Then $F(x, f(x)) = 0$. Using the chain rule to differentiate $F$ with respect to $x$, we have the following result:

$$\frac{dy}{dx} = -\frac{F_x}{F_y}$$

## 5.9 Directional Derivatives

Let $f$ be a function of $x$ and $y$. The directional derivative of $f$ at $(x_0, y_0)$ in the direction of a unit vector $\mathbf{u} = \langle a, b \rangle$ is

$$D_{\mathbf{u}} f\left(x_0, y_0\right) = \lim_{h \to 0} \frac{f\left(x_0 + ha, y_0 + hb\right) - f\left(x_0, y_0\right)}{h}$$

provided that the limit exists.

There are times where we are interested to compute the directional derivative of a function. Suppose $f$ is a function in terms of $x$ and $y$. Then, $f$ has a directional derivative in the direction of any unit vector $\mathbf{u} = \langle a, b \rangle$ and

$$D_{\mathbf{u}} f\left(x, y\right) = \left\langle f_x\left(x, y\right), f_y\left(x, y\right)\right\rangle \bullet \mathbf{u}.$$

## 5.10 Gradient Vector

The gradient of a scalar-valued differentiable function $f$ of several variables is the vector field $\nabla f$ whose value at a point $P$ is the vector whose components are the partial derivatives of $f$ at $P$. As such, it can be written as such:

$$\nabla f\left(x, y\right) = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j}$$

When we were talking about directional derivatives, observe that

$$\left\langle f_x\left(x, y\right), f_y\left(x, y\right)\right\rangle = \nabla f\left(x, y\right) = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j},$$

which implies that the gradient vector has an alternative (and useful) formula. That is,

$$D_{\mathbf{u}} f = \nabla f \bullet \mathbf{u}.$$

Suppose $P \in D_f$. Then, note that

$$\nabla f(P) \bullet \mathbf{u} = \left|\nabla f(P)\right| \left|\mathbf{u}\right| \cos\theta = \left|\nabla f(P)\right| \cos\theta$$

and the transition from the second last to the last step involves the fact that $\mathbf{u}$ is a unit vector. This equation implies that the maximum value of $D_{\mathbf{u}} f(P)$ is $\left|\nabla f\left(P\right)\right|$ and this occurs when $\mathbf{u}$ is acting in the same direction as the gradient vector $\nabla f(P)$. $\nabla f(P)$ can also be regarded as the direction and rate of fastest increase.

To summarise, we shall state the geometric properties of $\nabla f$. Suppose $f$ is a differentiable function at the point $(a, b)$ such that $\nabla f(a, b) \neq 0$. Then,
**(i)** The direction of $\nabla f(a, b)$ is perpendicular to the contour of $f$ through $(a, b)$ and it is in the direction of the maximum rate of increase of $f$ and
**(ii)** $|\nabla f|$ is the maximum rate of change of $f$ at that point and is large when the contours are close together and small when they are far apart.

# 6 Maxima and Minima

$f(x, y)$ has a local maximum at $(a, b)$ if $f(x, y) \leq f(a, b)$ for all points $(x, y)$ in some disk with centre $(a, b)$. The value $f(a, b)$ is called a *local maximum value*. Similarly, if $f(x, y) \geq f(a, b)$ for all points $(x, y)$ in some disk with centre $(a, b)$, then $f(a, b)$ is a *local minimum value*.

If $f$ has a local maximum or a local minimum at $(a, b)$ and $f_x(a, b)$ and $f_y(a, b)$ exist, then $f_x(a, b) = 0$ and $f_y(a, b) = 0$. That is, $\nabla f(a, b) = \mathbf{0}$. A point $(a, b)$ is called a critical point of $f$ if $f_x(a, b) = f_y(a, b) = 0$, or if one of these partial derivatives does not exist.

## 6.1 Saddle Point

A function $f$ is said to have a saddle point at $(a, b)$ if there is a disk centred at $(a, b)$ such that $f$ assumes its maximum value on one diameter of the disk only at $(a, b)$ and assumes its minimum value on another diameter of the disk only at $(a, b)$.

In other words, at the saddle point, the derivatives in orthogonal directions are zero but the point is not regarded as a local extremum of the function.

The following shows the graph of $z = x^2 - y^2$, which is a hyperbolic paraboloid that passes through the origin (0,0,0). However, at this point, it is not a local extremumm (i.e. a saddle point). This is because along the path $z = x^2$, we observe that the origin *seems like* a minimum point. However, along the path $z = -y^2$, the origin *seems like* a maximum point. Such an observation clearly implies that the origin is not a local extremum.

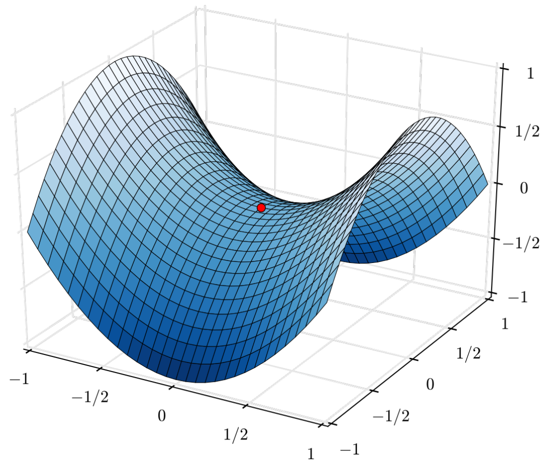


Figure 18: Graph of $z = x^2 - y^2$, which contains a saddle point

## 6.2 Second Derivative Test

**Second Derivative Test**

Suppose the partial derivatives $f_{xx}, f_{xy}, f_{yx}$ and $f_{yy}$ are continuous on a disk with centre $(a, b)$ and suppose $f_x(a, b) = f_y(a, b) = 0$. Let $D$ (which is actually called the determinant) be defined as such:

$$D(a, b) = f_{xx}(a, b)f_{yy}(a, b) - (f_{xy}(a, b))^2$$

The second derivative test states that if
**(a)**: $D > 0$ and $f_{xx}(a, b) > 0$, then $f$ has a local minimum at $(a, b)$;
**(b)**: $D > 0$ and $f_{xx}(a, b) < 0$, then $f$ has a local maximum at $(a, b)$;
**(c)**: $D < 0$, then $f$ has a saddle point at $(a, b)$ and
**(d)**: $D = 0$, then no conclusion can be drawn.

### 6.2.1 Hessian Matrix

To readers who are interested and have background knowledge on Linear Algebra, the Hessian Matrix, denoted by $\mathbf{H}$ is of concern in this section. If the second partial derivatives of $f$ exist and are continuous in its domain, then the Hessian Matrix of $f$ is the following $2 \times 2$ (for simplicity sake) matrix:

$$\mathbf{H} = \begin{pmatrix} f_{xx}(a, b) & f_{xy}(a, b) \\ f_{yx}(a, b) & f_{yy}(a, b) \end{pmatrix}$$

Observe that $\det(\mathbf{H}) = f_{xx}(a, b)f_{yy}(a, b) - (f_{xy}(a, b))^2$, which is the same as what was stated previously!

## 6.3 Sets in $\mathbb{R}^2$

A bounded set in $\mathbb{R}^2$ is one that is contained in some disk. A closed set in $\mathbb{R}^2$ is one that contains all its boundary points.

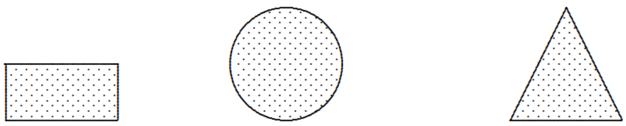

Figure 19: Bounded sets and closed sets

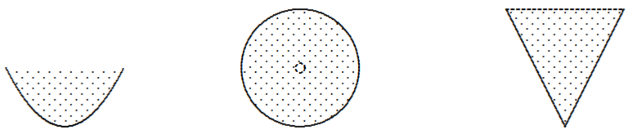

Figure 20: The left set is not bounded, and the middle and right sets are not closed

## 6.4 Extreme Value Theorem

If $f$ is continuous on a closed, bounded set $D$ in $\mathbb{R}^2$, then $f$ attains an absolute maximum value $f(x_1, y_1)$ and an absolute minimum value $f(x_2, y_2)$ at some points $(x_1, y_1)$ and $(x_2, y_2)$ in $D$.

To find the absolute maximum/absolute minimum of a function defined on a closed and bounded set, firstly, we have to find the values of $f$ at the critical points. That is, to find the coordinates $(x, y)$ such that $f_x(x, y) = f_y(x, y) = 0$. We then find the extreme values of $f$ on the boundary of $D$. Lastly, the largest of the values obtained from the earlier process yields the absolute maximum and vice versa.

## 6.5 Method of Lagrange Multipliers

The Method of Lagrange Multipliers involves maximising or minimising a function subject to a constraint. For example, the function could be $f(x,y)$ and the constraint could be $g(x,y)=0$.

The following diagram shows the level curves of $f(x,y)=\alpha$, where $8\le\alpha\le 12$ and the constraint $g(x,y)=0$. Suppose the extreme value of $f(x,y)$ subjected to the constraint is $k$ and it is attained at the point $(x_0,y_0)$. Firstly, the curve $g$ must touch the level curve at $f(x,y)=k$ so as to allow one to move the point along $g(x,y)=0$ in order to increase/decrease the value of $f$.

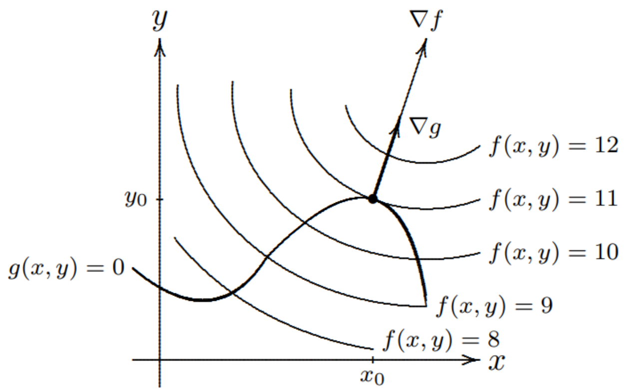


Figure 21: Geometric interpretation of the Lagrange Multiplier

As such, observe that the gradients of $f$ and $g$ are parallel at $(x_0,y_0)$. That is,

$$\nabla f\left(x_0,y_0\right)=\lambda\nabla g\left(x_0,y_0\right)\text{ if }\nabla g\left(x_0,y_0\right)\neq\mathbf{0}.$$

The Method of Lagrange Multipliers works for 2 constraints and 3 variables too! For such a case, suppose we want to maximise/minimise $f(x,y,z)$ subjected to the constraints $g(x,y,z)=h(x,y,z)=0$. If $f$ attains an extreme value at $(x_0,y_0,z_0)$, then

$$\nabla f\left(x_0,y_0,z_0\right)=\lambda\nabla g\left(x_0,y_0,z_0\right)+\mu\nabla h\left(x_0,y_0,z_0\right)\text{ if }\nabla g\left(x_0,y_0,z_0\right)\neq\mathbf{0}\text{ and }\nabla h\left(x_0,y_0,z_0\right)\neq\mathbf{0}.$$

To those who have substantial Olympiad experience, you would be familiar with some classical inequalities like Cauchy-Schwarz and AM-GM. Most extremum problems can be solved using these inequalities. However, there are some which cannot be solved using these methods. This is where the Method of Lagrange Multipliers comes in handy.

*Example:* Suppose $x,y,z\ge 0$ and they satisfy the equation

$$\frac{1}{x}+\frac{1}{y}+\frac{1}{z}=1.$$

We wish to solve for $x,y$ and $z$ which can minimise

$$f(x,y,z)=x+y+z^2$$

and also, solve for the minimum of $f(x,y,z)$.

*Solution:* We let $f(x,y,z)$ and $g(x,y,z)$ denote the following:

$$f(x,y,z)=x+y+z^2\text{ and }g(x,y,z)=\frac{1}{x}+\frac{1}{y}+\frac{1}{z}=1.$$

Setting $\nabla f=\lambda\nabla g$, we have

$$\langle 1,1,2z\rangle=\lambda\left\langle-\frac{1}{x^2},-\frac{1}{y^2},-\frac{1}{z^2}\right\rangle,$$

which implies that $x^2 = y^2 = 2z^2 = -\lambda$. Note that $x^2 = y^2 \implies x = y$ or $x = -y$. However, as $x$ and $y$ are non-negative, then the case where $x = -y$ does not make sense. We consider the case where $x = y$. Substituting into $g(x, y, z) = 0$ yields

$$\frac{1}{x} + \frac{1}{x} + \frac{2^{\frac{1}{3}}}{x^{\frac{2}{3}}} = 1.$$

We observe that $x = 4$ satisfies the equation, implying that $y = 4$ and $z = 2$. Hence,

$$f_{\min}(4, 4, 2) = 12.$$

□

*Example:* The famous AM-GM Inequality, where AM stands for arithmetic mean and GM stands for geometric mean, can be proven using the Method of Lagrange Multipliers too. Some other notable proofs are by using backward-forward induction (proven by Cauchy), using the exponential function $e^x$ (proven by Pólya), as well as Jensen's Inequality. The inequality states that

$$\frac{x_1 + x_2 + \ldots + x_n}{n} \geq (x_1 x_2 \ldots x_n)^{\frac{1}{n}},$$

where $x_1, x_2, \ldots, x_n$ are non-negative real numbers and equality is attained if and only if $x_1 = x_2 = \ldots = x_n$. We shall prove this inequality using the method of Lagrange Multipliers.

*Proof:* The problem can be regarded as asking how do we maximise $f(x_1, x_2, \ldots, x_n) = x_1 x_2 \ldots x_n$ subject to the constraint $g(x_1, x_2, \ldots, x_n) = x_1 + x_2 + \ldots + x_n = c$. Using $\nabla f = \lambda \nabla g$,

$$\begin{aligned} x_2 x_3 \ldots x_n &= \lambda \\ x_1 x_3 \ldots x_n &= \lambda \\ \vdots &= \vdots \\ x_1 x_2 \ldots x_{n-1} &= \lambda \end{aligned}$$

This is a system of $n$ equations. Multiplying all the left side of each equation together and equating them to the product of the right side of each equation,

$$(x_1 x_2 \ldots x_n)^{n-1} = \lambda^n.$$

Hence,

$$\begin{aligned} \frac{(x_1 x_2 \ldots x_n)^{n-1}}{(x_2 x_3 \ldots x_n)^{n-1}} &= \frac{\lambda^n}{\lambda^{n-1}} \\ x_1^{n-1} &= \lambda \\ x_1 &= \lambda^{\frac{1}{n-1}} \end{aligned}$$

It can be shown that

$$x_1 = x_2 = \ldots = x_n = \lambda^{\frac{1}{n-1}}$$

so

$$\begin{aligned} x_1 + x_2 + \ldots x_n = n\lambda^{\frac{1}{n-1}} &= c \\ \lambda &= \left(\frac{c}{n}\right)^{n-1} \\ x_1 = x_2 = \ldots = x_n &= \frac{c}{n} \end{aligned}$$

Now, note that $f(x_1, x_2, \ldots, x_n) = \frac{c^n}{n^n}$. As $g(x_1, x_2, \ldots, x_n) = c$, then

$$\frac{(x_1 + x_2 + \ldots + x_n)^n}{n^n} \geq x_1 x_2 \ldots x_n$$

so by taking the $n^{\text{th}}$ root on both sides, we obtain the AM-GM Inequality! □

# 7 Multiple Integrals

## 7.1 Double Integrals

Let $f$ be a function of two variables defined over a rectangle $R = [a, b] \times [c, d]$. We would like to define the double integral of $f$ over $R$ as the volume of the solid under the graph of $z = f(x, y)$ over $R$. Without a loss of generality, we let $f(x, y) \geq 0$ here.

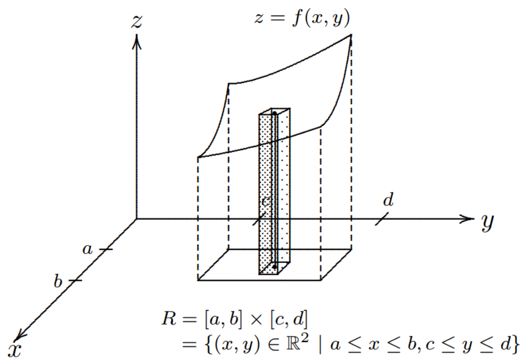


Figure 22: Geometric interpretation of the double integral

We subdivide $R$ into $mn$ small rectangles, each of area $\Delta A$, namely $R_{ij}$, where $1 \leq i \leq m$ and $1 \leq j \leq n$. The volume of an infinitesimally small rectangular solid erected over $R_{ij}$ is its height multiplied by its area, which is $f\left(x_{ij}^*, y_{ij}^*\right) \Delta A$.

Using ideas of a Riemann Sum, the double integral of $f$ over $R$ is defined by

$$\iint_R f(x, y)\ dA = \lim_{m,n\to\infty} \sum_{j=1}^{n} \sum_{i=1}^{m} f\left(x_{ij}^*, y_{ij}^*\right) \Delta A$$

if the limit exists. That is, the volume of the solid that lies above the rectangle $R$ and below the surface $z = f(x, y)$ is

$$\iint_R f(x, y)\ dA.$$

**REMARK:** If $f(x, y)$ is continuous on $R$, then $\iint_R f(x, y)\ dA$ always exists.

Some properties of double integrals are:

**(1)**:

$$\iint_D (\alpha f(x, y) + g(x, y))\ dA = \alpha \iint_D f(x, y)\ dA + \iint_D g(x, y)\ dA,$$

where $\alpha \in \mathbb{R}$. (1) simply means that the integral operator is a linear transformation.

**(2)**: If $f(x, y) \geq g(x, y)\ \forall x \in D$, then

$$\iint_D f(x, y)\ dA \geq \iint_D g(x, y)\ dA.$$

**(3)**:

$$\iint_D f(x, y)\ dA = \iint_{D_1} f(x, y)\ dA + \iint_{D_2} f(x, y)\ dA,$$

provided that $D = D_1 \cup D_2$ and $D_1$ and $D_2$ do not intersect except perhaps on their boundary. This uses the property that

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

**(4)**: If $f$, is bounded (i.e. $m \leq f(x, y) \leq M$), then

$$mA(D) \leq \iint_D f(x, y)\ dA \leq MA(D)$$

where $A(D)$ denotes the area of $D$.

## 7.2 Iterated Integrals and the Fubini-Tonelli Theorem

Let $f(x,y)$ be a function defined on $R=[a,b]\times[c,d]$. The integral $\int_c^d f(x,y)\ dy$ means that $x$ is regarded as a constant and $f(x,y)$ is integrated with respect to $y$ from $y=c$ to $y=d$. Thus, $\int_c^d f(x,y)\ dy$ is a function of $x$ and we can integrate it with respect to $x$ from $x=a$ to $x=b$. The resulting integral

$$\int_a^b \int_c^d f(x,y)\ dydx$$

is known as an *iterated integral.*

The Fubini-Tonelli Theorem allows the order of integration to be changed in certain iterated integrals. It states that if $f(x,y)$ is *absolutely convergent* and continuous on $R=[a,b]\times[c,d]$,

$$\iint_R f(x,y)\ dA=\int_a^b \int_c^d f(x,y)\ dydx=\int_c^d \int_a^b f(x,y)\ dxdy.$$

### 7.2.1 Absolute Convergence

As mentioned earlier, for Fubini's Theorem to be applied, $f$ must be an absolutely convergent integral. Similar to the absolute convergence of series, if an integral is said to be absolutely convergent, then

$$\int_R |f(x)|\ dx<\infty.$$

### 7.2.2 Gaussian Integral

One of the ways to evaluate the famous Gaussian Integral, which is

$$\int_{-\infty}^{\infty} e^{-x^2}\ dx=\sqrt{\pi},$$

involves Fubini's Theorem.

*Proof:* We will use polar coordinates. Let $I$ be the original integral. Then,

$$\begin{aligned}
I&=\int_{-\infty}^{\infty} e^{-x^2}\ dx=\int_{-\infty}^{\infty} e^{-y^2}\ dy\\
I^2&=\left(\int_{-\infty}^{\infty} e^{-x^2}\ dx\right)\left(\int_{-\infty}^{\infty} e^{-y^2}\ dy\right)\\
&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-x^2}e^{-y^2}\ dxdy \quad \because \text{Fubini's Theorem}\\
&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\left(x^2+y^2\right)}\ dxdy
\end{aligned}$$

We will do a change of variables from the Cartesian world to the polar world. We will establish the following result

$$dxdy=rdrd\theta$$

using the Jacobian of a suitable matrix. That is,

$$\mathrm{J}=\begin{pmatrix}\frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta}\\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta}\end{pmatrix}.$$

Since $dxdy=\det(\mathrm{J})\,drd\theta$, then the result follows. Hence, the integral can be transformed to

$$\begin{aligned}
I^2&=\int_0^{2\pi}\int_0^{\infty} re^{-r^2}\ drd\theta\\
&=\pi\\
I&=\sqrt{\pi}
\end{aligned}$$

□

### 7.2.3 Type 1 and Type 2 Regions

**Type 1 Region:** Consider a region, $D$, bounded by the vertical lines $x = a$ and $x = b$ and the curves $y = g_1(x)$ and $y = g_2(x)$, where $a < x < b$ and $g_1(x) < y < g_2(x)$. The double integral of $f$ over $D$ can be expressed as

$$\iint_D f(x, y)\ dA = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x, y)\ dydx.$$

**REMARK**

Note that $\int_{g_1(x)}^{g_2(x)} f(x, y)\ dy$ is a function of $x$.

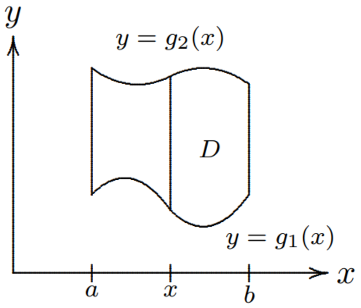


Figure 23: Type 1 Region

**Type 2 Region:** Consider a region, $D$, bounded by the horizontal lines $y = c$ and $y = d$ and the curves $y = h_1(y)$ and $y = h_2(y)$, where $c < y < d$ and $h_1(y)x < h_2(y)$. The double integral of $f$ over $D$ can be expressed as

$$\iint_D f(x, y)\ dA = \int_c^d \int_{h_1(y)}^{h_2(y)} f(x, y)\ dxdy.$$

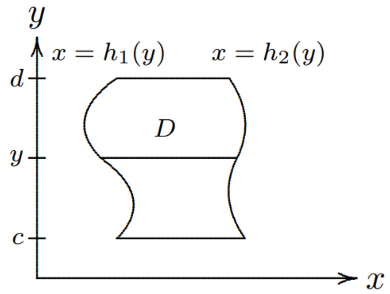


Figure 24: Type 2 Region

*Example:* We can evaluate

$$\int_0^8 \int_{\sqrt[3]{y}}^2 e^{x^4}\ dxdy$$

by changing the order of integration because $\int e^{x^4}\ dx$ does not have an anti-derivative.

*Solution:* We consider the graph of $x = \sqrt[3]{y}$, or by making $y$ the subject, $y = x^3$. Note that the green region satisfies both $\sqrt[3]{y} \le x \le 2 \implies y \le x^3 \le 8$ and $0 \le y \le 8$. Thus, if we were to change the order of integration such that we integrate with respect to $y$ first, then we have the inequalities $0 \le y \le x^3$ and $0 \le x \le 2$. Hence,

$$\int_0^8 \int_{\sqrt[3]{y}}^2 e^{x^4}\ dxdy = \int_0^2 \int_0^{x^3} e^{x^4}\ dydx = \int_0^2 \left[ye^{x^4}\right]_0^{x^3}\ dx = \frac{1}{4}\left(e^{16} - 1\right).$$

□

*Example:* We can also prove that

$$\int_0^1 \int_x^1 \sin\left(y^2\right)\ dydx = \frac{1}{2}\left(1-\cos 1\right)$$

by changing the order of integration too.

*Solution:* Note that $x \le y \le 1$ and $0 \le x \le 1$. Hence, if we were to integrate with respect to $x$ first, note that $0 \le x \le y$. Then, since $0 \le y \le 1$, we have

$$\int_0^1 \int_x^1 \sin\left(y^2\right)\ dydx = \int_0^1 \int_0^y \sin\left(y^2\right)\ dxdy = \int_0^1 \left[x\sin\left(y^2\right)\right]_0^y\ dy = \frac{1}{2}\left(1-\cos 1\right).$$

□

## 7.3 Double Integrals in Polar Coordinates

We first introduce the area differential in polar coordinates. Consider a point $(r,\theta)$ on a plane. Small increments in $r$ and $\theta$ are denoted by $dr$ and $d\theta$ respectively. Hence, by considering the area of the new sector is $dA = rdrd\theta$, which is known as the area differential.

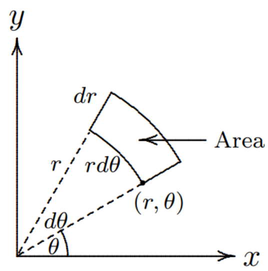


Figure 25: Area Differential in polar coordinates

Let $f$ be a continuous function defined on a polar rectangle $R$, where

$$R = \left\{(r,\theta)\,|0 \le a \le r \le b, \alpha \le \theta \le \beta\right\}$$

and $0 \le \beta - \alpha \le 2\pi$. The double integral of $f$ over $R$ can be expressed as

$$\iint_R f\left(x,y\right)\ dA = \int_\alpha^\beta \int_a^b f\left(r\cos\theta, r\sin\theta\right)\ rdrd\theta.$$

### 7.4 Surface Area

Let $f$ be a differentiable function of 2 variables defined on a domain $D$. We wish to find the surface area, $A(S)$, of the graph of $f$ over $D$. That is $\iint_D dS$. We wish to express the differential of the surface area, $dS$, in terms of the differential of the domain, $dA$. Note that $dA = |dxdy|$.

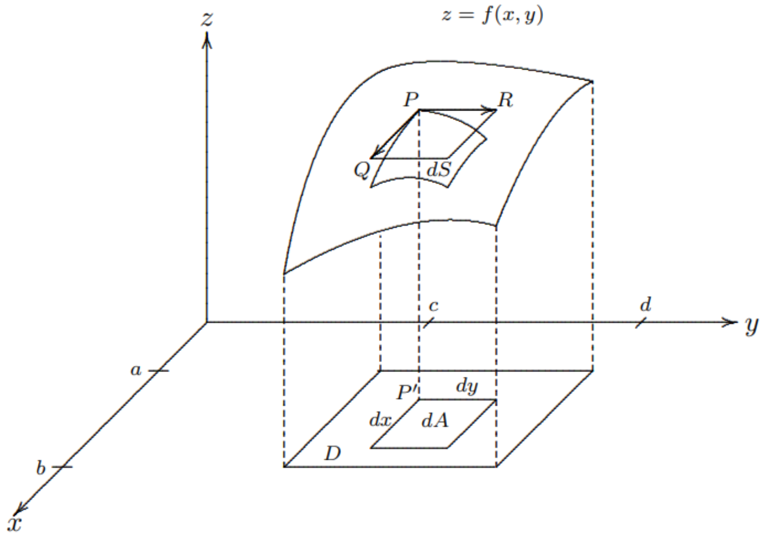


Figure 26: Derivation of the Surface Area Formula

From the figure above, let $\overrightarrow{PQ} = \langle dx, 0, f_x(x,y)\,dx\rangle$ and $\overrightarrow{PR} = \langle 0, y, f_y(x,y)\,dy\rangle$. Since $\overrightarrow{PQ}$ and $\overrightarrow{PR}$ are linearly independent and span the tangent plane, it is clear that the area of the tangent plane (which is in the shape of a parallelogram) formed by $\overrightarrow{PQ}$ and $\overrightarrow{PR}$ can be calculated by taking their cross product. That is,

$$\text{Area of tangent plane} = \left|\overrightarrow{PQ} \times \overrightarrow{PR}\right| = |\langle -f_x, -f_y, 1\rangle\, dxdy| = \sqrt{f_x^2 + f_y^2 + 1}\; dA$$

and hence, by the Fundamental Theorem of Calculus,

$$A(S) = \iint_D \sqrt{f_x^2 + f_y^2 + 1}\; dA.$$

### 7.5 Triple Integrals

Similar to the notion of using a Riemann Sum to explain the idea of double integrals, for triple integrals, we start off by considering a continuous function $f : B \subseteq \mathbb{R}^3 \to \mathbb{R}$, where $B$ is defined to be the following rectangular solid:

$$B = \{(x,y,z)\,|a \le x \le b, c \le y \le d, r \le z \le s\}$$

Dividing $[a,b]$, $[c,d]$ and $[r,s]$ into $l, m$ and $n$ equal subintervals respectively, it is clear that the triple integral of $f$ over $B$ is

$$\iiint_R f(x,y,z)\;dV = \lim_{l,m,n\to\infty} \sum_{i=1}^{l}\sum_{j=1}^{m}\sum_{k=1}^{n} f\left(x_{ijk}^*, y_{ijk}^*, z_{ijk}^*\right)\;\Delta V.$$

Note that Fubini's Theorem still applies for triple integrals. There are $3! = 6$ such iterated integrals involved and they are all equal.

The limits of integration are usually tricky to find. However, with sufficient practice, one should eventually be proficient at it.

*Example:* Suppose we wish to evaluate

$$\iiint_E xy\;dV,$$

where $E$ is the solid tetrahedron with vertices $(0,0,0)$, $(1,0,0)$, $(0,2,0)$ and $(0,0,3)$.

*Solution:* We find the equation of the plane that is formed by the vertices $(1,0,0),(0,2,0)$ and $(0,0,3)$ since the tetrahedron is in the first octant where all the $x,y$ and $z$ values are positive. The equation of the plane is $6x+3y+2z=6$. Hence, $E$ is the region that lies below the plane with $x$-, $y$-, and $z$-intercepts 1,2 and 3 respectively. Now, we find the bounds for $x$, $y$ and $z$.

Suppose we integrate in the order $dzdydx$, then note that as $6x+3y+2z=6$, then $z=\frac{1}{2}(6-6x-3y)$, which implies that $0\le z\le\frac{1}{2}(6-6x-3y)$. Next, $y=\frac{2}{3}(3-3x-z)$, so when $z=0$, we obtain an upper bound for $y$, which is $2-2x$. Hence, $0\le y\le 2-2x$. It is clear that $0\le x\le 1$. Combining everything together (I omitted the integration process since it is trivial),

$$\iiint_E xy\ dV=\int_0^1\int_0^{2-2x}\int_0^{\frac{1}{2}(6-6x-3y)} xy\ dzdydx=\frac{1}{10}.$$

□

### 7.5.1 Steinmetz Solid

A Steinmetz Solid is the solid obtained by the intersection of two or three cylinders of equal radius at right angles.

For a bicylinder (intersection of 2 cylinders) of radius $r$, its volume is given by $V=\frac{16}{3}r^3$ and its surface area is given by $A=16r^2$. For a tricylinder (intersection of 3 cylinders) of radius $r$, its volume is given by $V=8(2-\sqrt{2})r^3$ and its surface area is given by $A=24(2-\sqrt{2})r^2$.

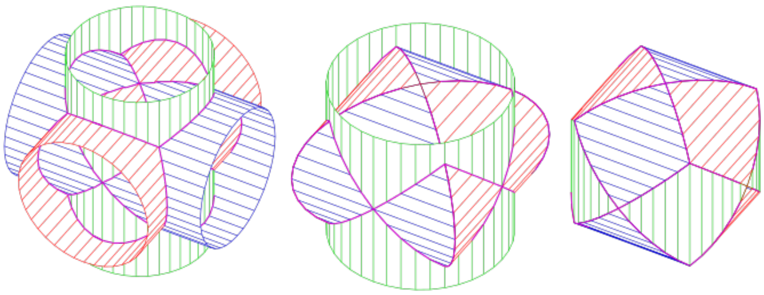

Figure 27: Forming a tricylinder

We can prove that the volume of a bicylinder is $V=\frac{16}{3}r^3$, for instance.
*Proof:* Without a loss of generality, suppose that the two cylinders have equations $x^2+y^2=r^2$ and $x^2+z^2=r^2$. As $z=\pm\sqrt{r^2-x^2}$ and $y=\pm\sqrt{r^2-x^2}$, then plugging into the formula gives

$$\begin{aligned}\iiint_R 1\ dV&=\int_{-r}^{r}\int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}}\int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} 1\ dzdydx\\&=\int_{-r}^{r}\int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} 2\sqrt{r^2-x^2}\ dydx\\&=\int_{-r}^{r} 4\left(r^2-x^2\right)\ dx\\&=\frac{16}{3}r^3\end{aligned}$$

□

### 7.5.2 Volumes of Classical 3D Shapes

The formulae for the volumes of classical 3D shapes, which include the cuboid, circular cylinder, sphere and cone can be derived through integration. Conversion to cylindrical or spherical coordinates are more helpful, where appropriate. You may visit the following link which directs you to a summary page of the volumes of 3D shapes using integration created by the University of Washington.

### 7.5.3 Type 1, Type 2 and Type 3 Regions

**Type 1 Region:** For a Type 1 region $E$, it is of the form $E = \{(x,y,z) \mid (x,y) \in D, u_1(x,y) \le z \le u_2(x,y)\}$. We can write the triple integral as such:

$$\iiint_E f(x,y,z)\ dV = \iint_D \left[\int_{u_1(x,y)}^{u_2(x,y)} f(x,y,z)\ dz\right]\ dA$$

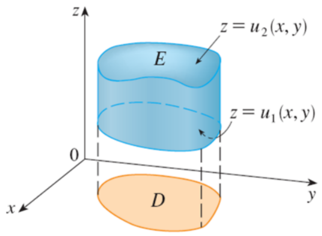


Figure 28: Type 1 Region

**Type 2 Region:** For a Type 2 region $E$, it is of the form $E = \{(x,y,z) \mid (y,z) \in D, u_1(y,z) \le x \le u_2(y,z)\}$. We can write the triple integral as such:

$$\iiint_E f(x,y,z)\ dV = \iint_D \left[\int_{u_1(y,z)}^{u_2(y,z)} f(x,y,z)\ dx\right]\ dA$$

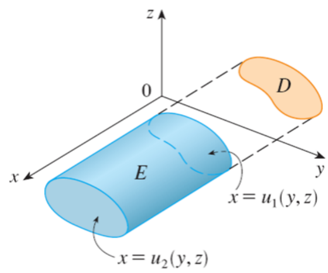


Figure 29: Type 2 Region

**Type 3 Region:** For a Type 3 region $E$, it is of the form $E = \{(x,y,z) \mid (x,z) \in D, u_1(x,z) \le y \le u_2(x,z)\}$. We can write the triple integral as such:

$$\iiint_E f(x,y,z)\ dV = \iint_D \left[\int_{u_1(x,z)}^{u_2(x,z)} f(x,y,z)\ dy\right]\ dA$$

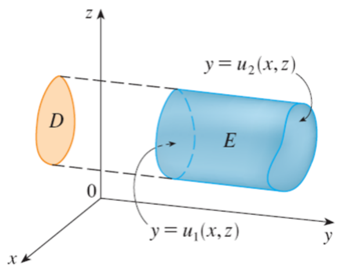


Figure 30: Type 3 Region

### 7.5.4 Triple Integrals in Cylindrical Coordinates

Consider the cylindrical rectangle as shown. The region $E$ (in blue) is defined by

$$E = \{(r, \theta, z) \,|\alpha \leq \theta \leq \beta, h_1(\theta) \leq r \leq h_2(\theta), u_1(r,\theta) \leq z \leq u_2(r,\theta)\}.$$

The triple integral of $f(x,y,z)$ over $E$ can be expressed as

$$\iiint_E f(x,y,z)\ dV = \int_\alpha^\beta \int_{h_1(\theta)}^{h_2(\theta)} \int_{u_1(\theta)}^{u_2(\theta)} f(r\cos\theta, r\sin\theta, z)\ r dz dr d\theta.$$

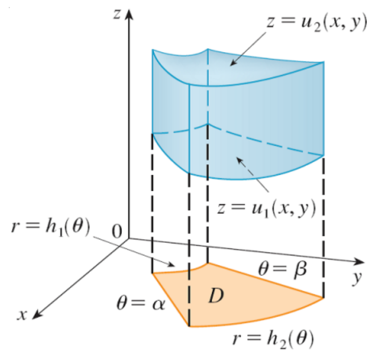


Figure 31: A cylindrical rectangle

### 7.5.5 Triple Integrals in Spherical Coordinates

We consider a spherical wedge $E$ where its spherical coordinates are bounded as such:

$$E = \{(\rho, \theta, \phi) \,| a \leq \rho \leq b, \alpha \leq \theta \leq \beta, c \leq \phi \leq d\}$$

The triple integral of $f$ over $E$ can be expressed as

$$\iiint_E f(x,y,z)\ dV = \int_c^d \int_\alpha^\beta \int_a^b f(\rho\sin\phi\cos\theta, \rho\sin\phi\sin\theta, \rho\cos\phi)\,\rho^2\sin\phi\ d\rho d\theta d\phi.$$

## 7.6 Change of Variables and the Jacobian

The figure below shows 2 regions, $S$ and $R$. $S$ and $R$ contain the points $(u_1, v_1)$ and $(x_1, y_1)$ respectively. Let $T$ be a $C^1$ transformation from the $uv$-plane to the $xy$-plane. The transformation $T$ maps the point $(u_1, v_1)$ in $S$ to $(x_1, y_1)$ in $R$. Assumming that $T^{-1}$ exists, the point $(x_1, y_1)$ in $R$ will get mapped to $(u_1, v_1)$ in $S$. This shows that $T$ is injective. $T$ is surjective too since $(u_1, v_1)$ and $(x_1, y_1)$ have pre-images, which are $(x_1, y_1)$ and $(u_1, v_1)$ respectively, implying that $T$ is bijective.

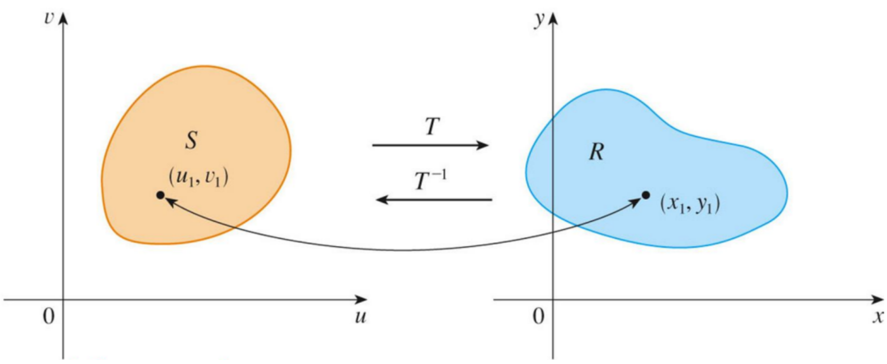


Figure 32: The transformation $T$ is said to be bijective

The Jacobian of the transformation $T$ given by $x = x(u, v), y = y(u, v)$ is

$$\frac{\partial (x, y)}{\partial (u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}.$$

The area differential, $dA$, can be written as

$$dA = \left| \frac{\partial (x, y)}{\partial (u, v)} \right| dudv.$$

Let $T(u, v)$ be a bijective $C^1$-transformation whose Jacobian is non-zero except possibly at a finite number of points. Suppose $T$ maps a region $S$ in the $uv$-plane onto a region $R$ of the $xy$-plane. Suppose $f$ is continuous on $R$. Then,

$$\iint_R f(x, y)\ dA = \iint_S f(x(u, v), y(u, v)) \left| \frac{\partial (x, y)}{\partial (u, v)} \right|\ dudv.$$

For the case of triple integrals, we have a completely analogous formula for change of variables.

# 8 Introduction to Vector Calculus

## 8.1 Vector Fields

Let $D \subseteq \mathbb{R}^2$. A vector field on $D$ is a function $\mathbf{F}$ that assigns to each point $(x, y)$ in $D$ a two dimensional vector $\mathbf{F}(x, y)$. $\mathbf{F}$ can be written in terms of its component functions. That is

$$\mathbf{F}(x, y) = \langle P(x, y), Q(x, y) \rangle$$

or $\mathbf{F} = P\mathbf{i} + Q\mathbf{j}$.

There is a variety of vector fields in $\mathbb{R}^2$. For example, there are radial and rotational vector fields. A radial field is one where all the vectors point towards or away from the origin and is rotationally symmetric. An example would be $\mathbf{F} = x\mathbf{i} + y\mathbf{j}$.

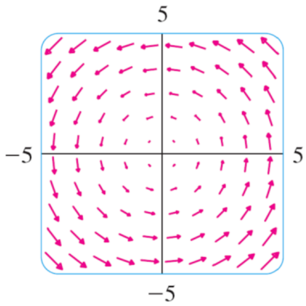


Figure 33: A rotational vector field which has the function $\mathbf{F} = -y\mathbf{i} + x\mathbf{j}$

The following diagram can be regarded as the electric field (an example of a vector field) of a negative point charge. Observe the magnitude of the electric field lines: in all directions, the lines become longer as they approach the point charge.

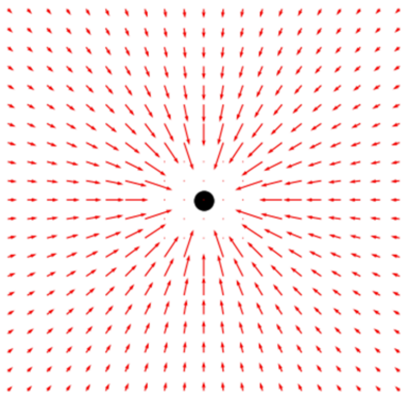

Figure 34: Electric force on a negative charge

We can also extend this idea to 3-dimensional vector fields. Let $E \subseteq \mathbb{R}^3$. A vector field on $E$ is a function $\mathbf{F}$ that assigns to each point $(x, y, z)$ in $E$ a three dimensional vector $\mathbf{F}(x, y, z)$. In terms of its component functions, $\mathbf{F}$ can be written as

$$\mathbf{F}(x, y, z) = \langle P(x, y, z), Q(x, y, z), R(x, y, z) \rangle$$

or $\mathbf{F} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$.

**REMARK**

For the vector field to be defined on its domain, $D$, each of its component vectors must be continuous on $D$.

## 8.2 Gradient Fields

If $f:\mathbb{R}^2\to\mathbb{R}$ is a differentiable function, then $\nabla f$ is a vector field on $\mathbb{R}^2$ and it is called the gradient vector field of $f$. Similarly, if $f:\mathbb{R}^3\to\mathbb{R}$ is a differentiable function, then $\nabla f$ is a vector field on $\mathbb{R}^3$ and it is called the gradient vector field of $f$.

Note that the gradient vectors are perpendicular to the level curves as proven using the chain rule.

A vector field $\mathbf{F}$ is called a conservative vector field if it is the gradient of some scalar function. That is there exists a differentiable function $f$ such that $\mathbf{F}=\nabla f$. In this situation, $f$ is called a potential function for $\mathbf{F}$.

## 8.3 Introduction to Line Integrals

**Line Integral**

Consider a plane curve $C$ with equation $\mathbf{r}(t)=x(t)\mathbf{i}+y(t)\mathbf{j}$, where $a\le t\le b$. Assume that $C$ is a smooth curve, meaning that $\mathbf{r}'(t)\neq 0$, and $\mathbf{r}'(t)$ is continuous $\forall t\in[a,b]$. Let $f(x,y)$ be a continuous function defined in a domain containing $C$. Using the idea of a Riemann Sum and the arc length differential, the line integral of $C$ is defined as

$$\int_C f(x,y)\ ds=\int_a^b f(x(t),y(t))\sqrt{\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2}\ dt.$$

Recall that

$$\int_a^b\sqrt{\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2}\ dt$$

gives the arc length of the curve from $a$ to $b$.

A piecewise smooth curve $C$ can be regarded as the union of a finite number of smooth curves, $C_i$, where $1\le i\le n$, or in other words

$$C=\bigcup_{i=1}^n C_i$$

where the initial point of $C_{i+1}$ is the terminal point of $C_i$. In Graph Theory terminologies, the above is the same as a *finite walk*. The line integral $f$ along $C$ can hence, be written as

$$\int_C f(x,y)\ ds=\sum_{i=1}^n\int_{C_i}f(x,y)\ ds.$$

The line integrals of $f$ along $C$ with respect to $x$ and $y$ respectively are

$$\int_C f(x,y)\ dx=\int_a^b f(x(t),y(t))\,x'(t)\ dt \text{ and } \int_C f(x,y)\ dx=\int_a^b f(x(t),y(t))\,y'(t)\ dt.$$

For a smooth space curve $C$ given by $\mathbf{r}(t)=x(t)\mathbf{i}+y(t)\mathbf{j}+z(t)\mathbf{j}$ where $a\le t\le b$, the formula for a line integral of a scalar field in $\mathbb{R}^3$ is

$$\int_C f(x,y,z)\ ds=\int_a^b f(x(t),y(t),z(t))\sqrt{\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2+\left(\frac{dz}{dt}\right)^2}\ dt.$$

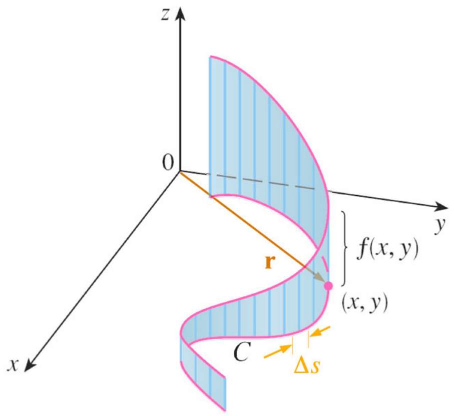


Figure 35: The line integral over a scalar field $f$ is the area under the curve $C$ along a surface $z = f(x, y)$

## 8.4 Line Integrals of Vector Fields

Let $\mathbf{F}$ be a continuous vector field defined on a domain containing a smooth curve $C$ given by a vector function $\mathbf{r}(t)$, where $a \le t \le b$. The line integral of $\mathbf{F}$ along $C$ is

$$\int_C \mathbf{F} \bullet d\mathbf{r} = \int_a^b \mathbf{F}\left(\mathbf{r}\left(t\right)\right) \bullet \mathbf{r}'\left(t\right)\ dt.$$

We can write $\mathbf{F}$ and $C$ in their component forms, where

$$\mathbf{F}(x, y, z) = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}$$

and

$$C : \mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}.$$

Thus,

$$\int_C \mathbf{F} \bullet d\mathbf{r} = \int_C P\ dx + Q\ dy + R\ dz.$$

### 8.4.1 Fundamental Theorem of Line Integrals

The Fundamental Theorem of Line Integrals, also known as the Gradient Theorem, is a generalisation of the Second Fundamental Theorem of Calculus to any curve in a plane or space.

For line integrals, let $C$ be a smooth curve given by $\mathbf{r}(t)$, where $a \le t \le b$. Let $f$ be a function of 2 or 3 variables whose gradient, denoted by $\nabla f$, is continuous. Then,

$$\int_C \nabla f \bullet d\mathbf{r} = f(\mathbf{r}(b)) - f(\mathbf{r}(a)).$$

### 8.4.2 Path Independence and Conservative Vector Fields

Let $\mathbf{F}$ be a continuous vector field with domain $D$. The line integral $\int_C \mathbf{F} \bullet d\mathbf{r}$ is independent of path in $D$ if $\int_{C_1} \mathbf{F} \bullet d\mathbf{r} = \int_{C_2} \mathbf{F} \bullet d\mathbf{r}$ for any 2 paths $C_1$ and $C_2$ in $D$ that have the same initial and terminal points.

A path is said to be *closed* if its terminal point coincides with its initial point. $\int_C \mathbf{F} \bullet d\mathbf{r}$ is independent of path in $D$ if and only if $\int_C \mathbf{F} \bullet d\mathbf{r} = 0$ for every closed path in $D$.

Aforementioned, we said that a vector field $\mathbf{F}$ is conservative if there exists a differentiable function $f$ such that $\mathbf{F} = \nabla f$. For example, in Physics, the gravitational field is said to be conservative. Consider the gravitational force between two objects of masses $m$ and $M$. By Newton's Law of Gravitation, we have the following result:

$$\mathbf{F} = -\frac{GMm}{|\mathbf{r}|^3}\mathbf{r}$$

where $G$ is the universal gravitational constant and $\mathbf{r} = \langle x, y, z \rangle$, implying that $|\mathbf{r}| = \sqrt{x^2 + y^2 + z^2}$. Note that a negative sign is present in the formula to indicate that the gravitational field is attractive. In this case, the potential function, $f$ (which is known as the gravitational potential energy), is given by

$$f = \frac{GMm}{|\mathbf{r}|}.$$

Note that the following statements are equivalent:

$$\mathbf{F} \text{ is conservative on } D \Leftrightarrow \int_C \mathbf{F} \cdot d\mathbf{r} \text{ is independent of path in } D \Leftrightarrow \int_C \mathbf{F} \cdot d\mathbf{r} = 0 \text{ for any closed path } C \text{ in } D$$

A set $D$ is said to be open if for every point $P$ in $D$, we can construct a disc about $P$ such that it lies entirely within $D$. A connected set is one such that any two points in $D$ can be joined by a path in $D$.

### 8.4.3 Test for Conservative Field

**Not conservative:** We can show that there are 2 paths with the same initial and terminal points but their line integrals are different.

**Conservative:** Let $\mathbf{F}(x, y) = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j}$ be a vector field in an *open* and *simply-connected* region $D \subset \mathbb{R}^2$, where $P$ and $Q$ are have continuous partial derivatives in $D$. $\mathbf{F}$ is conservative if and only if

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}.$$

By definition, a simple curve is one that does not intersect itself between its endpoints. A simply-connected region $D$ is one in which every simple closed curve in $D$ encloses only points within $D$.

# 9 Differential Operators

## 9.1 Green's Theorem

Green's Theorem gives the relationship between a line integral along a simple closed curve $C$ on the plane and the double integral over the plane region $D$ that $C$ bounds.

### 9.1.1 Simply Connected Regions

**Green's Theorem**

Let $C$ is a positively oriented (single counterclockwise transversal), piecewise-smooth and simple closed curve on the plane and $D$ is the region bounded by $C$. If $P(x, y)$ and $Q(x, y)$ have continuous partial derivatives on an open simply connected region that contains $D$, then

$$\int_C P dx + Q dy = \iint_D \left( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA.$$

### 9.1.2 Non Simply-Connected Regions

We can extend Green's Theorem to regions which are not simply-connected as they can be divided into regions which are simply connected. The line integrals along the divides end up being opposite in orientation and cancel out, and the result follows.

Consider the region $D$ where $\partial D = C_1 + C_2$. We may cut the region $D$ by 2 line segments $L_1$ and $L_2$ into 2 simply connected regions $D'$ and $D''$ respectively. Observe that $C_1$ and $C_2$ are traversing in opposite directions.

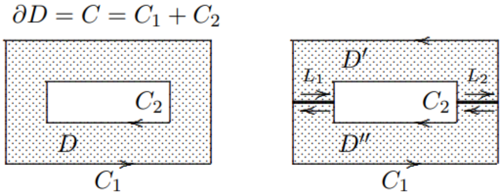


Figure 36: Green's Theorem for non a simply-connected region

As

$$\iint_D \left( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA = \iint_{D'} \left( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA - \left( - \iint_{D''} \left( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA \right),$$

then it can be verified that

$$\iint_D \left( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA = \iint_C P dx + Q dy,$$

which correlates with what was mentioned regarding Green's Theorem.

## 9.2 Curl

Let $\mathbf{F} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ be a vector curl in $\mathbb{R}^3$. The curl of $\mathbf{F}$ is defined by

$$\text{curl } \mathbf{F} = \left( \frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) \mathbf{i} + \left( \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) \mathbf{j} + \left( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \mathbf{k}.$$

This formula can be remembered easily by considering a suitable cross product. We first define the differential operator, $\nabla$, by

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}.$$

Note that

$$\text{curl } \mathbf{F} = \nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & R \end{vmatrix}$$

and upon expansion of the determinant yields the result as mentioned.

Some properties of the curl of $\mathbf{F}$ are as follows:
**(1):** If $f(x,y,z)$ has continuous second order partial derivatives, then curl $(\nabla f) = \mathbf{0}$.
*Proof:* By Clairaut's Theorem,

$$\text{curl } (\nabla f) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{vmatrix} = \left( \frac{\partial^2 f}{\partial y \partial z} - \frac{\partial^2 f}{\partial z \partial y} \right) \mathbf{i} - \left( \frac{\partial^2 f}{\partial x \partial z} - \frac{\partial^2 f}{\partial z \partial x} \right) \mathbf{j} + \left( \frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial x} \right) \mathbf{k} = \mathbf{0}.$$

**(2):** If $\mathbf{F}$ is conservative, then curl $\mathbf{F} = \mathbf{0}$.
**(3):** If $\mathbf{F}$ is a vector field on $\mathbb{R}^3$ whose component functions have continuous partial derivatives and curl $\mathbf{F} = \mathbf{0}$, then $\mathbf{F}$ is conservative.

## 9.3 Divergence

Let $\mathbf{F} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ be a vector field in $\mathbb{R}^3$. The divergence of $\mathbf{F}$ is defined by

$$\text{div } \mathbf{F} = \nabla \bullet \mathbf{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}.$$

If $P$, $Q$ and $R$ have continuous second order partial derivatives, then div curl $\mathbf{F} = 0$, which is a consequence of Clairaut's Theorem.

A nice property related to divergence and curl is that if $\mathbf{F}$ and $\mathbf{G}$ are vector fields in $R^3$, then

$$\text{div}\,(\mathbf{F} \times \mathbf{G}) = \mathbf{G}\,\text{curl}\,\mathbf{F} - \mathbf{F}\,\text{curl}\,\mathbf{G}.$$

*Proof:* Let $\mathbf{F} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ and $\mathbf{G} = A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$. Then,

$$\text{div}\,(\mathbf{F} \times \mathbf{G}) = \text{div} \left[ \begin{pmatrix} P \\ Q \\ R \end{pmatrix} \times \begin{pmatrix} A \\ B \\ C \end{pmatrix} \right] = \text{div} \begin{pmatrix} CQ - BR \\ AR - CP \\ BP - AQ \end{pmatrix}.$$

Expanding yields

$$\begin{aligned} \text{div}\,(\mathbf{F} \times \mathbf{G}) &= \frac{\partial (CQ)}{\partial x} - \frac{\partial (BR)}{\partial x} + \frac{\partial (AR)}{\partial x} - \frac{\partial (CP)}{\partial x} + \frac{\partial (BP)}{\partial x} - \frac{\partial (AQ)}{\partial x} \\ &= \underbrace{A \left( \frac{\partial R}{\partial z} - \frac{\partial Q}{\partial z} \right) - B \left( \frac{\partial R}{\partial x} - \frac{\partial P}{\partial z} \right) + C \left( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right)}_{\mathbf{G}\,\text{curl}\,\mathbf{F}} \\ &\quad \underbrace{+ P \left( \frac{\partial B}{\partial z} - \frac{\partial C}{\partial y} \right) - Q \left( \frac{\partial A}{\partial z} - \frac{\partial C}{\partial x} \right) + R \left( \frac{\partial A}{\partial y} - \frac{\partial B}{\partial x} \right)}_{-\mathbf{F}\,\text{curl}\,\mathbf{G}} \end{aligned}$$

Another differential operator occurs when we compute the divergence of a gradient vector field $\nabla f$. If $f$ is a function of three variables, then

$$\text{div}(\nabla f) = \nabla \bullet (\nabla f) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}.$$

$\text{div}(\nabla f)$ can also be written as $\nabla^2 f$, and we refer to this as the Laplace Operator or the Laplacian due to its connection with Laplace's Equation, an example of a partial differential equation. That is,

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = 0.$$

Consider a velocity vector field $\mathbf{F}$. Divergence measures the amount of flow radiating at a point. If the flow is uniform and without compression or expansion, then div $\mathbf{F} = 0$. Thus, if div $\mathbf{F} = 0$, then $\mathbf{F}$ is *incompressible*. As the curl measures the rotational effect of the vector field, if curl $\mathbf{F} = \mathbf{0}$, then $\mathbf{F}$ is *irrotational*.

If div $\mathbf{F} > 0$, there is a net outflow. If div $\mathbf{F} < 0$, there is a net inflow.

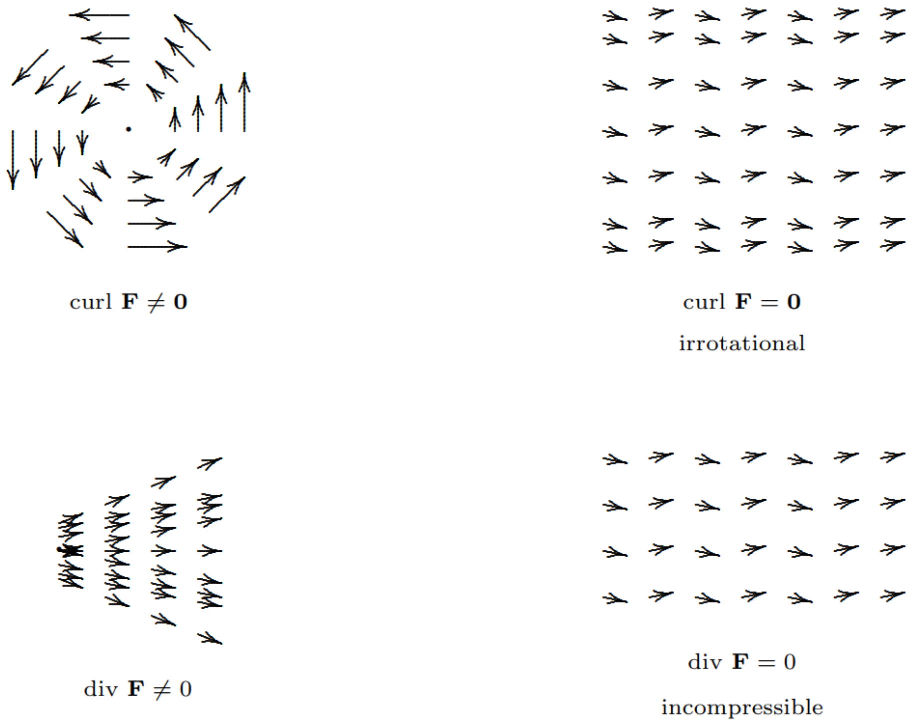


Figure 37: Examples of velocity vector fields and their characteristics

## 9.4 Green's Theorem in Vector Forms

Let $\mathbf{F} = P\mathbf{i} + Q\mathbf{j} + 0\mathbf{k}$ be a vector field in $\mathbb{R}^3$. Note that

$$\text{curl } \mathbf{F} = \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k},$$

so Green's Theorem can be rewritten as

$$\int_{\partial D} \mathbf{F} \bullet \mathrm{d}\mathbf{r} = \iint_D (\text{curl } \mathbf{F}) \bullet \mathbf{k} \; dA.$$

Suppose $\partial D$ can be parametrised by $\mathbf{r}(t) = \langle x(t), y(t) \rangle$, where $a \le t \le b$. Assuming the pararmetrisation gives the positive orientation of $\partial D$, the unit tangent vector is

$$\mathbf{T} = \frac{\mathbf{r}'(t)}{|\mathbf{r}'(t)|}.$$

Then, Green's Theorem can be also expressed as

$$\int_{\partial D} \mathbf{F} \bullet d\mathbf{r} = \int_{\partial D} \mathbf{F} \bullet \mathbf{T} \; ds,$$

where $ds = |\mathbf{r}'(t)|dt$ is the arc length differential.

We could also derive a formula involving the normal component of $\mathbf{F}$ along $\partial D$. In that way, Green's Theorem will be stated in terms of the divergence of the vector field $\mathbf{F}$. Using the above parametrisation of $C$, one can easily verify (by taking dot product with $\mathbf{T}$) that the outward unit normal vector to $\partial D$, $\mathbf{n}$, is given by

$$\mathbf{n}\,(t) = \left\langle \frac{y'\,(t)}{|\mathbf{r}'\,(t)|}, -\frac{x'\,(t)}{|\mathbf{r}'\,(t)|} \right\rangle.$$

Note that $\mathbf{n}$ is pointing outwards. Using Green's Theorem, it can be shown that

$$\int_{\partial D} \mathbf{F} \bullet \mathbf{n} \; ds = \iint_D \text{div } \mathbf{F} \; dA.$$

In summary, we have the following 3 integrals:

**(1):**

$$\int_{\partial D} \mathbf{F} \bullet d\mathbf{r} = \iint_D (\text{curl } \mathbf{F}) \bullet \mathbf{k} \; dA$$

**(2):**

$$\int_{\partial D} \mathbf{F} \bullet d\mathbf{r} = \int_{\partial D} \mathbf{F} \bullet \mathbf{T} \; ds$$

**(3):**

$$\int_{\partial D} \mathbf{F} \bullet \mathbf{n} \; ds = \iint_D \text{div } \mathbf{F} \; dA$$

# 10 Parametric Surfaces and their Areas

Let $\mathbf{r}(u,v)=\langle x(u,v),y(u,v),z(u,v)\rangle$ be a vector-valued function defined on a region $D$ in the $uv$-plane. Then,

$$S=\{(x,y,z)|x=x(u,v),y=y(u,v),z=z(u,v),(u,v)\in D\}$$

is a parametric surface. $x=x(u,v)$, $y=y(u,v)$ and $z=z(u,v)$ are the parametric equations of $S$.

## 10.1 Tangent Planes

Let $S$ be a parametric surface defined by $\mathbf{r}(u,v)=\langle x(u,v),y(u,v),z(u,v)\rangle$. The equation of the tangent plane to $S$ at a point $P_0$ with position vector $\mathbf{r}_0=\mathbf{r}(u_0,v_0)$ is

$$(\mathbf{r}-\mathbf{r}_0)\bullet(\mathbf{r}_u\times\mathbf{r}_v),$$

where

$$\mathbf{r}_u=\left\langle\frac{\partial x}{\partial u}(u_0,v_0),\frac{\partial y}{\partial u}(u_0,v_0),\frac{\partial z}{\partial u}(u_0,v_0)\right\rangle \text{ and } \mathbf{r}_v=\left\langle\frac{\partial x}{\partial v}(u_0,v_0),\frac{\partial y}{\partial v}(u_0,v_0),\frac{\partial z}{\partial v}(u_0,v_0)\right\rangle$$

as $\mathbf{r}_u$ and $\mathbf{r}_v$ lie on the tangent plane to $S$ at $P_0$, so their cross product gives the normal to the plane.

**REMARK**

$S$ is said to be smooth if $\mathbf{r}_u\times\mathbf{r}_v\neq\mathbf{0}\ \forall\ (x,y)\in D$.

## 10.2 Surface Area

Suppose a smooth parametric surface $S$ is injective except possibly on the boundary of $D$. The surface area of $S$ over $D$ is given by

$$A(S)=\iint_D|\mathbf{r}_u\times\mathbf{r}_v|\ dA.$$

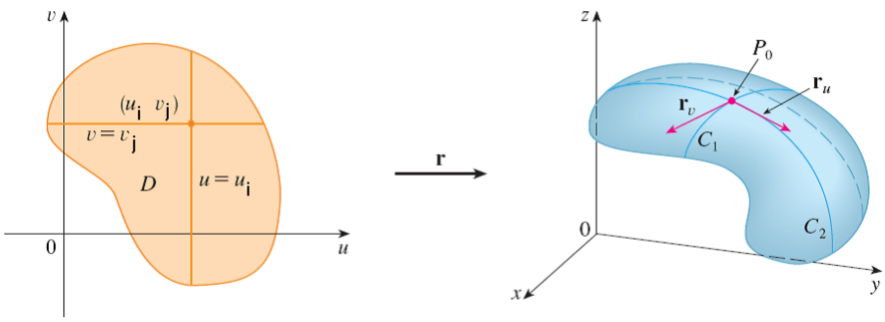


Figure 38: Geometric interpretation of the surface integral of a scalar field

## 10.3 Surface Area of the Graph of a Function

Let $S$ be a surface which is the graph of a function $f(x,y)$ defined on a domain $D\subset\mathbb{R}^2$. Recall that

$$A(S)=\iint_D\sqrt{f_x^2+f_y^2+1}\ dA.$$

**COROLLARY**

Consider a curve $y=f(x)$ where $a\le x\le b$, $f(x)\ge 0$ and $f'(x)$ is continuous. $S$ is the surface obtained by rotating the curve $2\pi$ radians about the $x$-axis. Then, the area of the surface of revolution is given by

$$2\pi\int_a^b f(x)\sqrt{1+f'(x)}\ dx.$$

The case where the curve $x=g(y)$ (for $c\le y\le d$) is rotated about the $y$-axis yields a similar formula.

### 10.4 Surface Integrals

A surface integral is related to surface area much like how a line integral is related to arc length.

Let $f(x,y,z)$ be a continuous function defined on $S$. The surface integral of $f$ over $S$ is

$$\iint_S f(x,y,z)\ dS = \iint_D f(\mathbf{r}(u,v))\,|\mathbf{r}_u \times \mathbf{r}_v|\ dA.$$

If $S$ is the graph of $z = g(x,y)$, then

$$\iint_S f(x,y,z)\ dS = \iint_D f(x,y,g(x,y))\sqrt{1+g_x^2+g_y^2}\ dA$$

## 11 Oriented Surfaces

A surface $S$ is said to be orientable if it is two-sided, otherwise it is non-orientable.

**Examples of orientable surfaces:** Sphere/ellipsoid, plane, cylinder, elliptic paraboloid
**Examples of non-orientable surfaces:** Möbius Strip, Klein Bottle

If $S$ is orientable, then it is possible to choose a unit normal vector $\mathbf{n}$ at every point $S$ so that $\mathbf{n}$ varies continuously over $S$. In that case, $S$ is an oriented surface and the choice of $\mathbf{n}$ is an orientation of $S$. There are only 2 orientations of S, namely one for each side of the surface which corresponds to the choice where all $\mathbf{n}$ point away from that side of the surface.

If $S$ is the graph of $z = g(x,y)$, then

$$n = \frac{\langle -g_x, -g_y, 1\rangle}{\sqrt{g_x^2+g_y^2+1}}$$

is the upward orientation of $S$ because the $\mathbf{k}$-component is positive. As such, the downward orientation is simply $-\mathbf{n}$.

If $S$ is written in parametric form $\mathbf{r} = \mathbf{r}(u,v)$, then

$$\mathbf{n} = \frac{\mathbf{r}_u \times \mathbf{r}_v}{|\mathbf{r}_u \times \mathbf{r}_v|}.$$

## 12 Surface Integrals of Vector Fields

Let $\mathbf{F}$ be a continuous vector field defined on an oriented surface $S$ with unit normal vector $\mathbf{n}$. The surface integral of $\mathbf{F}$ over $S$ is

$$\iint_S \mathbf{F} \bullet d\mathbf{S} = \iint_S \mathbf{F} \bullet \mathbf{n}\ dS = \iint_S \mathbf{F} \bullet \frac{\mathbf{r}_u \times \mathbf{r}_v}{|\mathbf{r}_u \times \mathbf{r}_v|}\ dS.$$

This integral is called the flux of $\mathbf{F}$ over $S$.

If $S$ is the graph of a function $z = g(x,y)$ over a region $D$ in the $xy$-plane (assuming upward orientation of $S$), and $\mathbf{F} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$, the above integral can be written as

$$\iint_S \mathbf{F} \bullet d\mathbf{S} = \iint_D (-Pg_x - Qg_y + R)\ dA.$$

**REMARK**

Always check the conditions before applying the formula:
**(1):** The surface $S$ is traced out by $\mathbf{r}(u,v)$, $(u,v) \in D$, where $D$ is the parameter domain.
**(2):** The orientation $\mathbf{n}$ given by the question is indeed the following expression:

$$\frac{\mathbf{r}_u \times \mathbf{r}_v}{|\mathbf{r}_u \times \mathbf{r}_v|}$$

# 13 Integral Theorems

## 13.1 Stokes' Theorem

**Stokes' Theorem**

Let $S$ be an oriented piecewise smooth surface that is bounded by a simple, closed, piecewise smooth boundary curve $C$ with positive orientation. Let $\mathbf{F}$ be a vector field whose components have continuous partial derivatives on an open region in $\mathbb{R}^3$ that contains $S$, where $\mathbf{F} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$. Then, Stokes' Theorem states that

$$\iint_S (\text{curl } \mathbf{F}) \bullet d\mathbf{S} = \iint_D -\left(R_y - Q_z\right) g_x - \left(P_z - R_x\right) g_y + \left(Q_x - P_y\right) \, dA.$$

**REMARK**

If $S_1$ and $S_2$ are oriented surfaces with the same oriented boundary curve $C$ and both satisfy the hypotheses of Stokes' Theorem, then

$$\iint_{S_1} (\text{curl } \mathbf{F}) \bullet d\mathbf{S} = \iint_C \mathbf{F} \bullet d\mathbf{r} = \iint_{S_2} (\text{curl } \mathbf{F}) \bullet d\mathbf{S}.$$

If curl $\mathbf{F} = \mathbf{0}$ on all of $\mathbb{R}^3$, then $\mathbf{F}$ is conservative.

### 13.1.1 Special Case of Stokes' Theorem

Consider the special case where the surface $S$ is flat and lies in the $xy$-plane with upward orientation. Then, the unit normal is $\mathbf{k}$ and the surface integral becomes a double integral. It can be shown that this is simply Green's Theorem.

## 13.2 The Divergence Theorem

**The Divergence Theorem**

Let $E$ be a solid region where the boundary surface $S$ of $E$ is piecewise smooth with positive orientation. Let $\mathbf{F}(x, y, z)$ be a vector field whose component functions have continuous partial derivatives on an open region that contains $E$. Then, by the Divergence Theorem,

$$\iint_S \mathbf{F} \bullet d\mathbf{S} = \iiint_E \text{div } \mathbf{F} \, dV.$$