Contents

1 Partial Derivatives

1.1 Notation

Partial derivatives denote a derivative of a function (or functions) of more than one variable with respect to a single variable. The notation is as follows

$$\lim_{dx\to 0}\frac{f(x+dx,y,z,...)-f(x,y,z,...)}{dx}=\left(\frac{\partial f}{\partial x}\right)_{y,z,...} \tag{1}$$

If it is clear which variables are kept constant, we drop the brackets and the lower index and have just $\frac{\partial f}{\partial x}$.

1.2 Second Order Derivatives

For partial derivatives of higher orders, similar rules as for total derivatives follow

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)=\frac{\partial^2 f}{\partial x^2} \tag{2}$$

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)=\frac{\partial^2 f}{\partial x\partial y} \tag{3}$$

But, the most used theorem in partial differentiation, which is now presented, tells us

$$\frac{\partial^2 f}{\partial x\partial y}=\frac{\partial^2 f}{\partial y\partial x} \tag{4}$$

This means that it does not matter in which order are the partial derivatives taken, the resultant function is always the same.

1.3 Total Differential

The total differential of function f is the change of this f with respect to all possible variables, i. e.

$$df(x_1,x_2,...,x_N)=\sum_{i=1}^{N}\frac{\partial f}{\partial x_i}dx_i \tag{5}$$

Hence the total derivative of f with respect to one of these variables is

$$\frac{df}{dx_j}=\sum_{i=1}^{N}\frac{\partial f}{\partial x_i}\frac{dx_i}{dx_j}=\frac{\partial f}{\partial x_j}+\sum_{i\neq j}^{N}\frac{\partial f}{\partial x_i}\frac{dx_i}{dx_j} \tag{6}$$

1.4 Inexact differentials

We defined how to get the total differential from a certain function. However, not expressions of form $\sum_i a_i dx_i$ have a corresponding function f for which $df=\sum_i a_i dx_i$. Such expressions are called inexact differentials, and are often denoted as $\text{đ}f$. There is a simple way to test whether a differential is exact or inexact if its form is known. If the differential is exact, then

$$df=\sum_{i}^{N}a_i dx_i=\sum_{i}^{N}\frac{\partial f}{\partial x_i}dx_i$$

And due to equation (4)

$$\forall i:\left(\prod_{j\neq i}^{N}\frac{\partial}{\partial x_j}\right)a_i=\left(\prod_{j\neq i}^{N}\frac{\partial}{\partial x_j}\right)\frac{\partial f}{\partial x_i}=\frac{\partial^N f}{\partial x_1\partial x_2...\partial x_{i-1}\partial x_{i+1}...\partial x_N\partial x_i}=\frac{\partial^N f}{\partial x_1...\partial x_N} \tag{7}$$

But, since the expression is independent of i, it is the same expression for each coefficient. Therefore, the differential is exact if and only if the partial derivative of a coefficient of dx_i with respect to all other variables is the same for all dx_i.

1.5 Chain Rule

Further remark is to be made about the total differential. It might be the case that the total change can be expressed as both function of several variables (which are somehow handy to use) and just a single variable (which is a bit problematic to use). Such situation often occurs when integrating in spaces that are subspaces of the integration space - integrating over lines in planes, over lines in spaces, over planes in spaces etc.
Lets now take the example of integrating over line. We can parametrize the line by single parameter s - the distance along the line from a certain point. Then, in a plane, $x = x(s)$ and $y = y(s)$. And therefore

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy = \frac{\partial f}{\partial x}\frac{dx}{ds}ds + \frac{\partial f}{\partial y}\frac{dy}{ds}ds$$

$$\frac{df}{ds} = \frac{\partial f}{\partial x}\frac{dx}{ds} + \frac{\partial f}{\partial y}\frac{dy}{ds} \tag{8}$$

Similarly, if $f = f(x_1, x_2, ..., x_N)$ and $\forall i : x_i = x_i(s_1, s_2, ..., s_M)$

$$df = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}dx_i = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}\sum_{j=1}^{M}\frac{\partial x_i}{\partial s_j}ds_j$$

$$\frac{df}{ds_k} = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}\sum_{j=1}^{M}\frac{\partial x_i}{\partial s_j}\frac{ds_j}{ds_k} \tag{9}$$

and so on. For mutually orthogonal variables $\frac{ds_j}{ds_k} = \delta_{jk}$, where δ_{jk} is the Kronecker delta

$$\frac{df}{ds_k} = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}\frac{\partial x_i}{\partial s_k} \tag{10}$$

2 Scalar Multivariable Calculus

2.1 Directional Derivatives

The gradient of a scalar function of several variables is defined as

$$\nabla f = \sum_i \frac{\partial f}{\partial x_i}\hat{x}_i \tag{11}$$

where x_i are components in the cartesian coordinate system and $\hat{x}_i$ is the unit vector of the cartesian coordinate system. The directional derivative of f along direction defined by vector $\vec{u}$ is

$$\nabla_u f = \nabla f \cdot \hat{u} = \nabla f \cdot (\frac{\vec{u}}{|\vec{u}|}) \tag{12}$$

In order to translate gradient into different coordinate systems, it is sufficient to apply the chain rule

$$\nabla = \sum_i \hat{x}_i \frac{\partial}{\partial x_i} = \sum_i \hat{x}_i(\hat{s}_1, \hat{s}_2, ...) \sum_j \frac{\partial s_j}{\partial x_i}\frac{\partial}{\partial s_j} \tag{13}$$

where s_j and $\hat{s}_j$ are coordinates and unit directions, respectively, in the new coordinate system. We than only need to now inverse transformations of the coordinate system and unit directions.

2.2 Laplacian

Laplacian (or Laplace's operator) ∇^2 is a differential operator of form (in cartesian coordinate system)

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{14}$$

It is a scalar operator - returns a scalar from a scalar function. It has a vector alternative, which has form

$$\nabla^2 \vec{f} = \nabla^2 f_x \hat{x} + \nabla^2 f_y \hat{y} + \nabla^2 f_z \hat{z} \tag{15}$$

where f_i is the component of $\vec{f}$ in $\hat{i}$ direction.
Forms for Laplacian in other coordinate systems can be calculated directly by applying the chain rule twice and keeping track of derivative order, but are not stated here, as the form is calculated later on from the vector differential operators.

2.3 Integral elements

Commonly in integration, we integrate over certain volume or surface or line. The integration elements (dV, dS and dl) are best determined geometrically in most cases, but there is a direct way, using the Jacobian matrix.
Jacobian matrix is defined as

$$\mathbf{J} = \begin{pmatrix} \frac{\partial x_1}{\partial s_1} & \frac{\partial x_1}{\partial s_2} & \dots & \frac{\partial x_1}{\partial s_N} \\ \frac{\partial x_2}{\partial s_1} & \frac{\partial x_2}{\partial s_2} & \dots & \frac{\partial x_2}{\partial s_N} \\ \dots & \dots & \dots & \dots \\ \frac{\partial x_M}{\partial s_1} & \frac{\partial x_M}{\partial s_2} & \dots & \frac{\partial x_M}{\partial s_N} \end{pmatrix} \tag{16}$$

We will consider one cases - square $\mathbf{J}$.
For square $\mathbf{J}$, we might want to transfer the integral from cartesian $dxdy$ to $dsdt$, or $dxdydz$ to $dsdtdu$. This is done by

$$dS = dxdy = ||\mathbf{J}||dsdt \tag{17}$$

where $||\mathbf{J}||$ is the absolute value of the determinant of the Jacobian matrix between (x, y) in rows and (s, t) in columns. Similarly

$$dV = dxdydz = ||\mathbf{J}||dsdtdu \tag{18}$$

The line integration is substitution is best understood geometrically - element of length dl is the Euclidian sum of dx and dy:

$$dl = \sqrt{dx^2 + dy^2} = dx\sqrt{1 + \left(\frac{dy}{dx}\right)^2} \tag{19}$$

where x is the coordinate in some direction and y is the corresponding coordinate on the line.
Now follows list of surface/volume/line elements in several coordinate systems.

2.3.1 Polar Planar Coordinates

In polar planar coordinates, the position of point P is given by its distance from the origin r and angle ϕ between the line connecting P and origin and the x axis. In previous years, we have shown that

$$x = r\cos(\phi)$$

$$y = r\sin(\phi)$$

$$\hat{e}_r = \cos(\phi)\hat{i} + \sin(\phi)\hat{j}$$

$$\hat{e}_\phi = -\sin(\phi)\hat{i} + \cos(\phi)\hat{j}$$

Hence we see that the unit vectors are orthogonal.
Polar planar coordinates have only one surface element to consider - element in the plane. One side of the element corresponds to change in coordinate r and has length dr, second component corresponds to change in ϕ coordinate and has length $rd\phi$. Hence, the size of the element is $dS = rdrd\phi$. There is no volume element to consider.
The line element here can be in two general directions. First is the line element in the increasing r direction, with length dr. So, the element is $d\vec{l} = dr\hat{e}_r$. The second line element is along the azimuthal direction (direction of increasing ϕ). This has length $rd\phi$, and therefore is $d\vec{l} = rd\phi\hat{e}_\phi$.
Both the surface element and line elements are indicated in the Fig. 1.
In the Jacobian formalism

$$\mathbf{J} = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \phi} \end{pmatrix} = \begin{pmatrix} \frac{\partial(r\cos(\phi))}{\partial r} & \frac{\partial(r\cos(\phi))}{\partial \phi} \\ \frac{\partial(r\sin(\phi))}{\partial r} & \frac{\partial(r\sin(\phi))}{\partial \phi} \end{pmatrix} = \begin{pmatrix} \cos(\phi) & -r\sin(\phi) \\ \sin(\phi) & r\cos(\phi) \end{pmatrix}$$

$$|\mathbf{J}| = r\cos^2(\phi) + r\sin^2(\phi) = r$$

as expected.

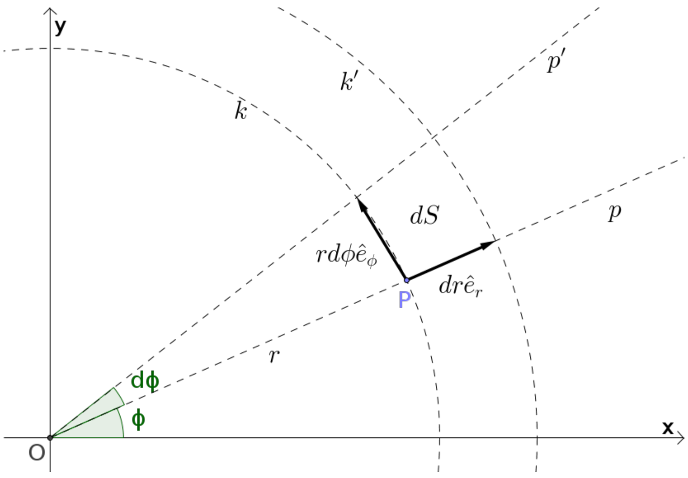

Figure 1: Line and surface elements in the planar polar coordinates. Circles k and k' are the contours of constant r, lines p and p' are contours of constant ϕ. The line elements are indicated. The surface element has magnitude of the product of lengths of the line elements, as the $\hat{e}_r$ and $\hat{e}_\phi$ vectors are perpendicular to each other.

2.3.2 Hyperbolic coordinates

Here, I am using the definition used for the first time in lectures. It shows that for the second time in lectures and anywhere on the internet, the definitions of this coordinate system are always different. I only include these as an illustration of the principles used before.

In hyperbolic coordinates, the position of point P is determined by two coordinates, which satisfy following relations

$$\frac{y}{x} = t$$
$$yx = s$$

The contours of t are lines, the contours of s are hyperbolae (hence the name of the coordinate system). To find unit vectors, we take the small change of the position vector $\vec{r}$ with small change in s or t coordinate, respectively.

For s coordinate (keeping t constant)

$$d\vec{r} = d\vec{r}_s = \vec{r}(s + ds, t) - \vec{r}(s, t)$$

This situation is represented in Fig. 2.

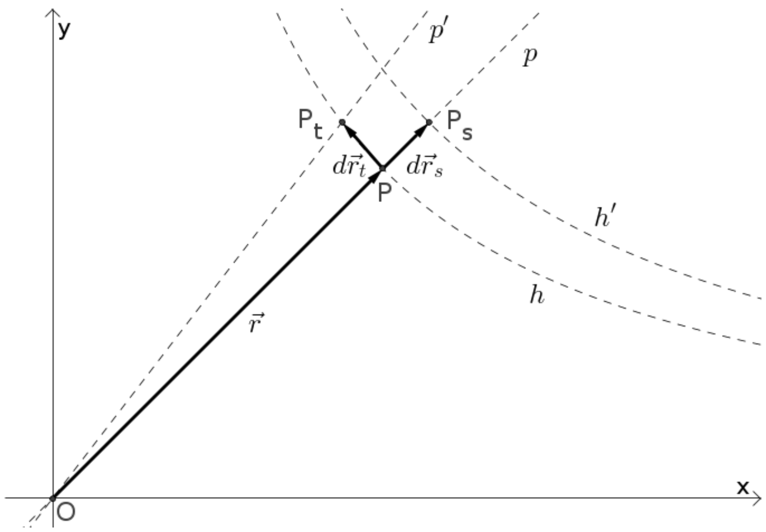

Figure 2: Change of the position vector upon changing the s and t variable by small amounts ds and dt, respectively. Lines p and p' are contours of t and $t + dt$, hyperbolae h and h' are contours of s and $s + ds$.

Point P has coordinates x and y, which translate into s and t, as we defined. To find the inverse transforms, we note that $y = tx$. Therefore

$$tx^2 = s$$
$$x = \sqrt{\frac{s}{t}}$$
$$y = tx = \sqrt{st}$$

For point P_s, the relations are

$$\frac{y_s}{x_s} = t$$
$$y_s x_s = s + ds$$

Hence, we can find that again $y_s = tx_s$, and hence

$$tx_s^2 = s + ds$$
$$x_s = \sqrt{\frac{s+ds}{t}} = \sqrt{\frac{s}{t}}\sqrt{1+\frac{ds}{s}} = x\sqrt{1+\frac{ds}{s}}$$
$$y_s = \sqrt{st}\sqrt{1+\frac{ds}{s}} = y\sqrt{1+\frac{ds}{s}}$$

The vector $d\vec{r}_s$ is then

$$d\vec{r}_s = (x_s - x)\hat{i} + (y_s - y)\hat{j} = x\left(\sqrt{1+\frac{ds}{s}} - 1\right)\hat{i} + y\left(\sqrt{1+\frac{ds}{s}} - 1\right)\hat{j}$$

Because ds is small

$$d\vec{r}_s \approx x\frac{ds}{2s}\hat{i} + y\frac{ds}{2s}\hat{j} = \frac{ds}{2s}\vec{r}$$

The magnitude of the change in $\vec{r}$ is then

$$|d\vec{r}_s| = \frac{ds}{2s}|\vec{r}| = \frac{ds}{2s}\sqrt{x^2+y^2} = \frac{ds}{2s}\sqrt{\frac{s}{t} + st}$$

And the unit vector of s direction is then

$$\hat{e}_s = \frac{d\vec{r}_s}{|d\vec{r}_s|} = \frac{\vec{r}}{|\vec{r}|} = \frac{\sqrt{\frac{s}{t}}}{\sqrt{\frac{s}{t}+st}}\hat{i} + \frac{\sqrt{st}}{\sqrt{\frac{s}{t}+st}}\hat{j} = \frac{1}{\sqrt{1+t^2}}\hat{i} + \frac{1}{\sqrt{1+\frac{1}{t^2}}}\hat{j}$$

Similarly for the change in t, the coordinates of P_t are

$$\frac{y_t}{x_t} = t + dt$$
$$y_t = (t+dt)x_t$$
$$y_t x_t = s$$
$$(t+dt)x_t^2 = s$$
$$x_t = \sqrt{\frac{s}{t+dt}} = \sqrt{\frac{s}{t}}\sqrt{\frac{t}{t+dt}} = x\sqrt{\frac{1}{1+\frac{dt}{t}}}$$
$$y_t = \sqrt{s(t+dt)} = \sqrt{st}\sqrt{1+\frac{dt}{t}} = y\sqrt{1+\frac{dt}{t}}$$
$$d\vec{r}_t = (x_t - x)\hat{i} + (y_t - y)\hat{j} = x\left(\sqrt{\frac{1}{1+\frac{dt}{t}}} - 1\right)\hat{i} + y\left(\sqrt{1+\frac{dt}{t}} - 1\right)\hat{j}$$

Again, because dt is small

$$d\vec{r}_t \approx x\left(\sqrt{1-\frac{dt}{t}} - 1\right)\hat{i} + y\frac{dt}{2t}\hat{j} \approx x\frac{-dt}{2t}\hat{i} + y\frac{dt}{2t}\hat{j}$$

The magnitude is

$$|d\vec{r}_t| = \frac{dt}{2t}\sqrt{x^2+y^2} = \frac{dt}{2t}|\vec{r}| = \frac{dt}{2t}\sqrt{\frac{s}{t}+st}$$

And the unit vector is

$$\hat{e}_t = \frac{d\vec{r}_t}{|d\vec{r}_t|} = \frac{-\sqrt{\frac{s}{t}}}{\sqrt{\frac{s}{t}+st}}\hat{\imath} + \frac{\sqrt{st}}{\sqrt{\frac{s}{t}+st}}\hat{\jmath} = -\frac{1}{\sqrt{1+t^2}}\hat{\imath} + \frac{1}{\sqrt{1+\frac{1}{t^2}}}\hat{\jmath}$$

The magnitude of the surface element dS is given by the magnitude of the vector product of these two vectors, $d\vec{r}_s$ and $d\vec{r}_t$. Since they both lie in the xy plane, the only component of the vector product will be in the direction of z. It will be

$$dS = |d\vec{r}_s \times d\vec{r}_t| = |(d\vec{r}_s)_x(d\vec{r}_t)_y - (d\vec{r}_s)_y(d\vec{r}_t)_x| =$$

$$\left|x\frac{ds}{2s}y\frac{dt}{2t} - y\frac{ds}{2s}\left(-x\frac{dt}{2t}\right)\right| = \left|2xy\frac{ds}{2s}\frac{dt}{2t}\right| = \left|s\frac{dsdt}{2st}\right| = \left|\frac{dsdt}{2t}\right| = \left|\frac{1}{2t}\right|dsdt$$

In the Jacobian formalism

$$\mathbf{J} = \begin{pmatrix} \frac{\partial x}{\partial s} & \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial s} & \frac{\partial y}{\partial t} \end{pmatrix} = \begin{pmatrix} \frac{\partial(\sqrt{\frac{s}{t}})}{\partial s} & \frac{\partial(\sqrt{\frac{s}{t}})}{\partial t} \\ \frac{\partial(\sqrt{st})}{\partial s} & \frac{\partial(\sqrt{st})}{\partial t} \end{pmatrix} = \begin{pmatrix} \frac{1}{2\sqrt{st}} & \frac{-\sqrt{s}}{2\sqrt{t^3}} \\ \frac{\sqrt{t}}{2\sqrt{s}} & \frac{\sqrt{s}}{2\sqrt{t}} \end{pmatrix}$$

$$|\mathbf{J}| = \frac{1}{2\sqrt{st}}\frac{\sqrt{s}}{2\sqrt{t}} - \frac{\sqrt{t}}{2\sqrt{s}}\left(-\frac{\sqrt{s}}{2\sqrt{t^3}}\right) = \frac{1}{4t} + \frac{1}{4t} = \frac{1}{2t}$$

and hence

$$dS = dxdy = ||\mathbf{J}||dsdt = \left|\frac{1}{2t}\right|dsdt$$

which is what we expected.

2.3.3 Cylindrical coordinates

Cylindrical coordinates are extension of the polar coordinates to 3D. They do this by simply adding the cartesian z coordinate to the radial and azimuthal coordinate.

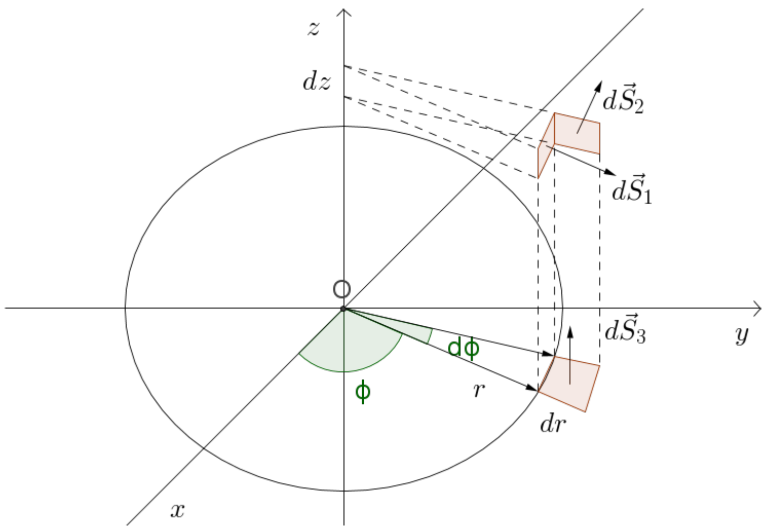

Figure 3: Surface elements in cylindrical coordinates in directions of unit vectors. The unit vectors in cylindrical coordinates are the same as in planar polar plus the cartesian $\hat{k}$ vector.

There are three surface elements to consider in the cylindrical coordinates. First is the element in the direction of the radial vector (in the direction means that the normal vector of the surface element is in certain direction). One edge of the element has length dz, second edge has length $rd\phi$. Hence, the element

itself has surface area $dS_1 = rd\phi dz$ (all unit vectors are orthogonal, as $\hat{e}_r$ and $\hat{e}_\phi$ are orthogonal, as we have already shown, and they are both orthogonal to $\hat{e}_z = \hat{k}$).
The second surface element, $d\vec{S}_2$, is in the azimuthal direction, and has size $dS_2 = drdz$.
The last element has direction of $\hat{k}$ and size $dS_3 = rdrd\phi$.
The volume element is the volume of the parallelopiped created by these surface elements, which is given by the size of one of the surface elements multiplied by the length of the line element in the remaining direction (surface element is given by two directions). So

$$dV = dS_1 dr = rdrd\phi dz = dS_2 rd\phi = dS_3 dz$$

2.3.4 Spherical Coordinates

In spherical coordinates, point P is given by its distance from the origin r and two angles. First angle ϕ is the angle between plane A and the x axis, where plane A is a plane perpendicular to xy plane that contains the point P. Second angle θ is the angle between the line L, connecting the origin and P, and the z axis. The illustration is in the Fig. 4

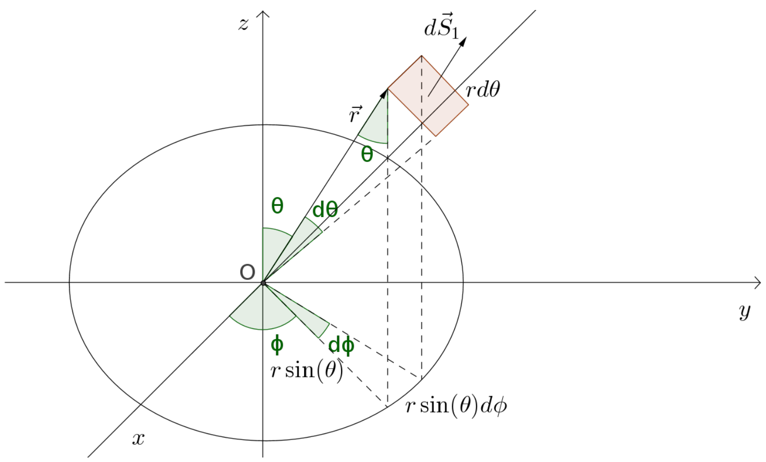

Figure 4: First surface element in spherical coordinates

The transform to cartesian coordinates is as follows.

$$\vec{r} = r\cos\phi\sin\theta\hat{i} + r\sin\phi\sin\theta\hat{j} + r\cos\theta\hat{k}$$

The unit vector in the direction of change of ϕ is

$$\hat{e}_\phi = \frac{d\vec{r}_\phi}{|d\vec{r}_\phi|} = \frac{\vec{r}(r,\phi+d\phi,\theta) - \vec{r}(r,\phi,\theta)}{|d\vec{r}_\phi|} = \frac{-r\sin\phi\sin\theta d\phi\hat{i} + r\cos\phi\sin\theta d\phi\hat{j} + 0\hat{k}}{\sqrt{r^2\sin^2\phi\sin^2\theta + r^2\cos^2\phi\sin^2\theta}d\phi} =$$

$$= \frac{-r\sin\phi\sin\theta\hat{i} + r\cos\phi\sin\theta\hat{j}}{r\sqrt{\sin^2\theta(\sin^2\phi + \cos^2\phi)}} = \frac{r\sin\theta(-\sin\phi\hat{i} + \cos\phi\hat{j})}{r\sin\theta} = -\sin\phi\hat{i} + \cos\phi\hat{j}$$

The unit vector in the direction of change of θ is

$$\hat{e}_\theta = \frac{d\vec{r}_\theta}{|d\vec{r}_\theta|} = \frac{r\cos\phi\cos\theta d\theta\hat{i} + r\sin\phi\cos\theta d\theta\hat{j} - r\sin\theta d\theta\hat{k}}{rd\theta\sqrt{\cos^2\phi\cos^2\theta + \sin^2\phi\cos^2\theta + \sin^2\theta}} = \cos\phi\cos\theta\hat{i} + \sin\phi\cos\theta\hat{j} - \sin\theta\hat{k}$$

Importantly

$$\hat{e}_\phi \cdot \hat{e}_\theta = -\sin\phi\cos\phi\cos\theta + \cos\phi\sin\phi\cos\theta + 0 = 0$$

which means that these unit vectors are perpendicular. Finally

$$\hat{e}_r = \frac{d\vec{r}_r}{|d\vec{r}_r|} = \frac{dr\cos\phi\sin\theta\hat{i} + dr\sin\phi\sin\theta\hat{j} + dr\cos\theta\hat{k}}{dr\sqrt{\cos^2\phi\sin^2\theta + \sin^2\phi\sin^2\theta + \cos^2\theta}} = \cos\phi\sin\theta\hat{i} + \sin\phi\sin\theta\hat{j} + \cos\theta\hat{k}$$

$$\hat{e}_r \cdot \hat{e}_\phi = \cos\phi\sin\theta(-\sin\phi) + \sin\phi\sin\theta\cos\phi + 0 = 0$$

$$\hat{e}_r \cdot \hat{e}_\theta = \cos\phi\sin\theta\cos\phi\cos\theta + \sin\phi\sin\theta\sin\phi\cos\theta + \cos\theta(-\sin\theta) = \cos\theta\sin\theta(\cos^2\phi + \sin^2\phi - 1) = 0$$

Hence all unit vectors are mututally perpendicular.
Therefore, the surface first surface element with direction of $\hat{e}_r$ has area

$$dS_1 = r\sin\theta d\phi r d\theta = r^2\sin\theta d\theta d\phi$$

The other two surface elements are along the two remaining directions - $\hat{e}_\phi$ and $\hat{e}_\theta$. $d\vec{S}_2$ along $\hat{e}_\phi$ has size $dS_2 = dr r d\theta = r dr d\theta$, $d\vec{S}_3$ along $\hat{e}_\theta$ has size $dS_3 = dr r\sin\theta d\phi = r\sin\theta dr d\phi$.
We can determine the size of the volume element again from the size of dS_1 and change in the perpendicular direction, which is dr. Therefore

$$dV = r^2\sin\theta dr d\theta d\phi$$

2.4 Pappus' Theorems

Pappus' theorems are connected to integrals of revolution - calculating surfaces and volumes of objects formed by full rotation of some curve/surface around a solid axis. The basic idea is to split the volume into layers that form cylinders with different radii and common axis.

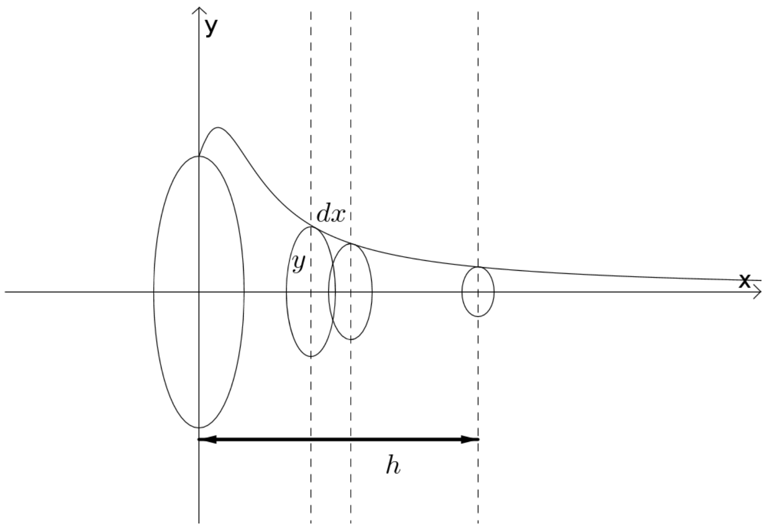

Figure 5: Illustration for infinitesimal cylinder formed by revolution of some function $y(x)$ around the x axis.

To find the surface area of the object, we need to integrate over all small cylindrical surfaces, which are created by line elements of function y between x and $x + dx$. This line element has length

$$ds = \sqrt{(dy)^2 + (dx)^2} = dx\sqrt{1 + \left(\frac{dy}{dx}\right)}$$

So the cylindrical surface has area

$$2\pi y ds = 2\pi y\sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx$$

And the area is

$$S = 2\pi\int_0^h y\sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx = 2\pi\int_0^l y ds$$

where l is the total length curve y. This resembles a centroid integral in 1D.
Centroid integral can be described as

$$\bar{y} = \frac{\int_V y dV}{\int_V dV} = \frac{\int_V y dV}{V} \tag{20}$$

where V is some volume and y is some coordinate. In our 1D case

$$\bar{y} = \frac{\int_l yds}{l} = \frac{\int_0^l yds}{l}$$

So our integral for surface becomes

$$S = 2\pi l\bar{y} \tag{21}$$

where $\bar{y}$ is the **1D centroid of the line** y and l is the length of the line y. This is called Pappus' first theorem.
To find the volume of the created object, we integrate over all the volume elements in cylindrical coordinates. This gives us

$$V = \int_0^h \int_0^y \int_0^{2\pi} rd\phi drdx$$

This can be reordered as

$$V = \int_0^{2\pi} d\phi \int_0^h \int_0^y rdrdx$$

But, the inner integral has a familiar form. Centroid $\bar{y}$ in 2D is a point defined by

$$\bar{y} = \frac{\iint_S ydS}{\iint_S dS}$$

where S is some surface and y is some coordinate. Hence, in our case

$$\bar{r} = \frac{\int_0^h \int_0^y rdrdx}{\int_0^h \int_0^y drdx} = \frac{\int_0^h \int_0^y rdrdx}{\int_0^h ydx}$$

$$\int_0^h \int_0^y rdrdx = \bar{r}\int_0^h ydx$$

And therefore

$$V = \int_0^2 \pi\bar{r}\int_0^h ydxd\phi = 2\pi\bar{r}\int_0^h ydx = 2\pi\bar{r}A \tag{22}$$

where $\bar{r}$ is the **2D centroid of the area under the curve** y, A is the area under y. This is called the second Pappus' theorem.

3 Differential Vector Calculus

3.1 Differential Operators

Following are the definitions and basic forms of several differential operators. Important notice is that each operator can be viewed as independent of certain coordinate system - the coordinate system only gives the operator some specific form. I usually start with cartesian form and then try to derive other forms as well.

3.1.1 Gradient

Gradient is a differential operator on a scalar function. As such, we know how to transform it into different coordinate system.

Cartesian coordiantes Gradient in cartesian coordinates is simply

$$\nabla f = \hat{i}\frac{\partial f}{\partial x} + \hat{j}\frac{\partial f}{\partial y} + \hat{k}\frac{\partial f}{\partial z} \tag{23}$$

Cylindrical coordinates The cylindrical coordinates follow

$$x = r\cos\phi$$

$$y = r\sin\phi$$

Hence the inverse transforms from cylindrical to cartesian coordinates are

$$r = \sqrt{x^2+y^2} \tag{24}$$

$$\phi = \tan^{-1}\left(\frac{y}{x}\right) \tag{25}$$

The unit vectors in direction of r and ϕ are

$$\hat{e}_r = \cos\phi\hat{i} + \sin\phi\hat{j}$$

$$\hat{e}_\phi = -\sin\phi\hat{i} + \cos\phi\hat{j}$$

Hence the inverse transforms of unit directions are

$$\hat{i} = \cos\phi\hat{e}_r - \sin\phi\hat{e}_\phi \tag{26}$$

$$\hat{j} = \sin\phi\hat{e}_r + \cos\phi\hat{e}_\phi \tag{27}$$

Hence, the gradient operator is (as $\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y} = 0$)

$$\nabla f = \hat{i}\frac{\partial f}{\partial x} + \hat{j}\frac{\partial f}{\partial y} + \hat{k}\frac{\partial f}{\partial z} =$$

$$= (\cos\phi\hat{e}_r - \sin\phi\hat{e}_\phi)\left(\frac{\partial r}{\partial x}\frac{\partial f}{\partial r} + \frac{\partial\phi}{\partial x}\frac{\partial f}{\partial\phi}\right) + (\sin\phi\hat{e}_r + \cos\phi\hat{e}_\phi)\left(\frac{\partial r}{\partial y}\frac{\partial f}{\partial r} + \frac{\partial\phi}{\partial y}\frac{\partial f}{\partial\phi}\right) + \hat{k}\frac{\partial f}{\partial z}$$

From the inverse transforms

$$\frac{\partial r}{\partial x} = \frac{\partial(\sqrt{x^2+y^2})}{\partial x} = \frac{1}{2\sqrt{x^2+y^2}}2x = \frac{x}{\sqrt{x^2+y^2}}$$

From the forward transforms

$$\frac{x}{\sqrt{x^2+y^2}} = \frac{r\cos\phi}{r} = \cos\phi$$

So $\frac{\partial r}{\partial x} = \cos\phi$. Similarly

$$\frac{\partial r}{\partial y} = \sin\phi$$

$$\frac{\partial\phi}{\partial x} = \frac{\partial\left(\tan^{-1}\left(\frac{y}{x}\right)\right)}{\partial x} = \frac{1}{1+\frac{y^2}{x^2}}\frac{-y}{x^2} = \frac{-y}{y^2+x^2} = \frac{-\sin\phi}{r}$$

$$\frac{\partial\phi}{\partial y} = \frac{\cos\phi}{r}$$

So, we have

$$\nabla f = (\cos\phi\hat{e}_r - \sin\phi\hat{e}_\phi)\left(\cos\phi\frac{\partial f}{\partial r} - \frac{\sin\phi}{r}\frac{\partial f}{\partial\phi}\right) + (\sin\phi\hat{e}_r + \cos\phi\hat{e}_\phi)\left(\sin\phi\frac{\partial f}{\partial r} + \frac{\cos\phi}{r}\frac{\partial f}{\partial\phi}\right) + \hat{k}\frac{\partial f}{\partial z}$$

Now, I will group together the parts containing $\hat{e}_r$

$$(\nabla f)_r = \cos^2\phi\frac{\partial f}{\partial r} - \frac{\sin\phi\cos\phi}{r}\frac{\partial f}{\partial\phi} + \sin^2\phi\frac{\partial f}{\partial r} + \frac{\sin\phi\cos\phi}{r}\frac{\partial f}{\partial\phi} = \frac{\partial f}{\partial r}$$

Similarly for other directions

$$(\nabla f)_\phi = -\sin\phi\cos\phi\frac{\partial f}{\partial r} + \frac{\sin^2\phi}{r}\frac{\partial f}{\partial\phi} + \sin\phi\cos\phi\frac{\partial f}{\partial r} + \frac{\cos^2\phi}{r}\frac{\partial f}{\partial\phi} = \frac{1}{r}\frac{\partial f}{\partial\phi}$$

$$(\nabla f)_z = \frac{\partial f}{\partial z}$$

Hence, in cylindrical coordinates

$$\nabla f = \frac{\partial f}{\partial r}\hat{e}_r + \frac{1}{r}\frac{\partial f}{\partial\phi}\hat{e}_\phi + \frac{\partial f}{\partial z}\hat{k} \tag{28}$$

Spherical coordinates In spherical coordinates, following relations apply

$$x = r\cos\phi\sin\theta$$

$$y = r\sin\phi\sin\theta$$

$$z = r\cos\theta$$

The inverses are

$$r = \sqrt{x^2+y^2+z^2} \tag{29}$$

$$\phi = \tan^{-1}\left(\frac{y}{x}\right) \tag{30}$$

$$\theta = \tan^{-1}\left(\frac{\sqrt{x^2+y^2}}{z}\right) \tag{31}$$

The unit direction vectors are

$$\hat{e}_r = \cos\phi\sin\theta\hat{i} + \sin\phi\sin\theta\hat{j} + \cos\theta\hat{k}$$

$$\hat{e}_\phi = -\sin\phi\hat{i} + \cos\phi\hat{j}$$

$$\hat{e}_\theta = \cos\phi\cos\theta\hat{i} + \sin\phi\cos\theta\hat{j} - \sin\theta\hat{k}$$

The inverses are

$$\hat{i} = \cos\phi\sin\theta\hat{e}_r - \sin\phi\hat{e}_\phi + \cos\phi\cos\theta\hat{e}_\theta \tag{32}$$

$$\hat{j} = \sin\phi\sin\theta\hat{e}_r + \cos\phi\hat{e}_\phi + \sin\phi\cos\theta\hat{e}_\theta \tag{33}$$

$$\hat{k} = \cos\theta\hat{e}_r - \sin\theta\hat{e}_\theta \tag{34}$$

Therefore, we can determine the gradient as

$$\nabla f = \hat{i}\left(\frac{\partial r}{\partial x}\frac{\partial f}{\partial r} + \frac{\partial \phi}{\partial x}\frac{\partial f}{\partial \phi} + \frac{\partial \theta}{\partial x}\frac{\partial f}{\partial \theta}\right) + \hat{j}\left(\frac{\partial r}{\partial y}\frac{\partial f}{\partial r} + \frac{\partial \phi}{\partial y}\frac{\partial f}{\partial \phi} + \frac{\partial \theta}{\partial y}\frac{\partial f}{\partial \theta}\right) + \hat{k}\left(\frac{\partial r}{\partial z}\frac{\partial f}{\partial r} + \frac{\partial \phi}{\partial z}\frac{\partial f}{\partial \phi} + \frac{\partial \theta}{\partial z}\frac{\partial f}{\partial \theta}\right)$$

Here

$$\frac{\partial r}{\partial x} = \frac{x}{r} = \cos\phi\sin\theta$$

$$\frac{\partial r}{\partial y} = \sin\phi\sin\theta$$

$$\frac{\partial r}{\partial z} = \cos\theta$$

$$\frac{\partial \phi}{\partial x} = -\frac{\sin\phi}{r\sin\theta}$$

$$\frac{\partial \phi}{\partial y} = \frac{\cos\phi}{r\sin\theta}$$

$$\frac{\partial \phi}{\partial z} = 0$$

$$\frac{\partial \theta}{\partial x} = \frac{1}{1+\frac{x^2+y^2}{z^2}}\frac{1}{2z\sqrt{x^2+y^2}}2x = \frac{xz}{(x^2+y^2+z^2)\sqrt{x^2+y^2}} = \frac{\cos\phi\sin\theta\cos\theta}{r\sin\theta} = \frac{\cos\phi\cos\theta}{r}$$

$$\frac{\partial \theta}{\partial y} = \frac{\sin\phi\cos\theta}{r}$$

$$\frac{\partial \theta}{\partial z} = \frac{1}{1+\frac{x^2+y^2}{z^2}}\left(-\frac{\sqrt{x^2+y^2}}{z^2}\right) = -\frac{\sqrt{x^2+y^2}}{x^2+y^2+z^2} = -\frac{\sin\theta}{r}$$

Using expressions for $\hat{i}$, $\hat{j}$ and $\hat{k}$, we can group the terms by $\hat{e}_r$, $\hat{e}_\phi$ and $\hat{e}_\theta$:

$$(\nabla f)_r = \cos\phi\sin\theta\left(\cos\phi\sin\theta\frac{\partial f}{\partial r} - \frac{\sin\phi}{r\sin\theta}\frac{\partial f}{\partial \phi} + \frac{\cos\phi\cos\theta}{r}\frac{\partial f}{\partial \theta}\right) +$$

$$+\sin\phi\sin\theta\left(\sin\phi\sin\theta\frac{\partial f}{\partial r}+\frac{\cos\phi}{r\sin\theta}\frac{\partial f}{\partial\phi}+\frac{\sin\phi\cos\theta}{r}\frac{\partial f}{\partial\theta}\right)+$$
$$+\cos\theta\left(\cos\theta\frac{\partial f}{\partial r}-\frac{\sin\theta}{r}\frac{\partial f}{\partial\theta}\right)=\frac{\partial f}{\partial r}$$
$$(\nabla f)_\phi=-\sin\phi\left(\cos\phi\sin\theta\frac{\partial f}{\partial r}-\frac{\sin\phi}{r\sin\theta}\frac{\partial f}{\partial\phi}+\frac{\cos\phi\cos\theta}{r}\frac{\partial f}{\partial\theta}\right)+$$
$$+\cos\phi\left(\sin\phi\sin\theta\frac{\partial f}{\partial r}+\frac{\cos\phi}{r\sin\theta}\frac{\partial f}{\partial\phi}+\frac{\sin\phi\cos\theta}{r}\frac{\partial f}{\partial\theta}\right)=\frac{1}{r\sin\theta}\frac{\partial f}{\partial\phi}$$
$$(\nabla f)_\theta=\cos\phi\cos\theta\left(\cos\phi\sin\theta\frac{\partial f}{\partial r}-\frac{\sin\phi}{r\sin\theta}\frac{\partial f}{\partial\phi}+\frac{\cos\phi\cos\theta}{r}\frac{\partial f}{\partial\theta}\right)+$$
$$+\sin\phi\cos\theta\left(\sin\phi\sin\theta\frac{\partial f}{\partial r}+\frac{\cos\phi}{r\sin\theta}\frac{\partial f}{\partial\phi}+\frac{\sin\phi\cos\theta}{r}\frac{\partial f}{\partial\theta}\right)-\sin\theta\left(\cos\theta\frac{\partial f}{\partial r}-\frac{\sin\theta}{r}\frac{\partial f}{\partial\theta}\right)=$$
$$=\frac{1}{r}\frac{\partial f}{\partial\theta}$$

Therefore, in spherical coordinates

$$\nabla f=\frac{\partial f}{\partial r}\hat{e}_r+\frac{1}{r\sin\theta}\frac{\partial f}{\partial\phi}\hat{e}_\phi+\frac{1}{r}\frac{\partial f}{\partial\theta}\hat{e}_\theta \tag{35}$$

3.1.2 Divergence

Following derivation is pretty much taken from Feynmann's lectures on physics - I greatly recommend reading the chapters on vector calculus.
Consider a flux of $\vec{F}$ from some volume V. It can be calculated as a surface integral across the boundary of V - ∂V. It is then

$$\Phi=\iint_{\partial V}\vec{F}\cdot d\vec{S} \tag{36}$$

where $d\vec{S}$ is normal to the surface ∂V and pointing outwards from the surface.
Now, consider that we split the region by a plane P into two regions, V_1 and V_2. The boundary surfaces of these regions consist of the part of boundary of region V ($\partial V_1'$ and $\partial V_2'$, respectively) and of the plane of division, P. Now, consider that we want to calculate the net flux from these two regions separately. For region V_1

$$\Phi_1=\iint_{\partial V_1}\vec{F}\cdot d\vec{S}_1=\iint_{\partial V_1'}\vec{F}\cdot d\vec{S}+\iint_P\vec{F}\cdot d\vec{S}_1$$

as in the integral over the boundary that coincides with boundary of V the $d\vec{S}_1$ is identical with $d\vec{S}$.
For region V_2

$$\Phi_2=\iint_{\partial V_2}\vec{F}\cdot d\vec{S}_2=\iint_{\partial V_2'}\vec{F}\cdot d\vec{S}+\iint_P\vec{F}\cdot d\vec{S}_2=\iint_{\partial V_2'}\vec{F}\cdot d\vec{S}-\iint_P\vec{F}\cdot d\vec{S}_1$$

as the vectors $d\vec{S}_1$ and $d\vec{S}_2$ must be equal and opposite on the plane of division P. Hence, the total flux

$$\Phi=\iint_{\partial V}\vec{F}\cdot d\vec{S}=\iint_{\partial V_1'}\vec{F}\cdot d\vec{S}+\iint_{\partial V_2'}\vec{F}\cdot d\vec{S}=\Phi_1+\Phi_2 \tag{37}$$

Therefore, the total flux can be found as a sum of the fluxes of separate parts of the region.
We now consider a infinitesimal region of space, and study the flux out of it in the first approximation.

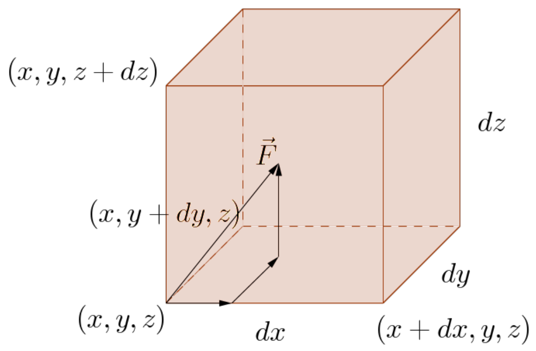

Figure 6: Divergence of an infinitesimal cube in cartesian coordinates

Cartesian coordinates The flux from an infinitesimal cube is approximately taking vector at each side of the cube and multiplying its normal component by the area of the side of the cube.
The flux therefore is

$$d\Phi = F_x(x+dx, y, z)dydz - F_x(x, y, z)dydz + F_y(x, y+dy, z)dxdz - F_y(x, y, z)dxdz +$$

$$+ F_z(x, y, z+dz)dxdy - F_z(x, y, z)dxdy = \left(\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}\right) dxdydz = \nabla \cdot \vec{F} dV$$

where $\nabla \cdot \vec{F}$ is called the divergence of $\vec{F}$ (sometimes also written as div$\vec{F}$). Hence, due to (37), flux from any volume can be written as

$$\Phi = \iiint_V d\Phi = \iiint_V \nabla \cdot \vec{F} dV$$

Comparing with (36), we have Gauss's law

$$\iint_{\partial V} \vec{F} \cdot d\vec{S} = \iiint_V \nabla \cdot \vec{F} dV \tag{38}$$

In other coordinate systems, we define the divergence so that the Gauss's law is met, i. e. we define it as the flux from an infinitesimal element. The equations defining divergence in spherical and cylindrical coordinates follow.

Cylindrical coordinates In cylindrical coordinates, the flux from infinitesimal element can be determined as follows.
Along the $\hat{e}_r$ direction, the flux flows into the element from the side of with lengths $rd\phi$ and dz. The we choose reference point for $\vec{F}$ at (r, ϕ, z). The flux out of the element in this direction is from the sidewith lengths $(r + dr)d\phi$ and dz. The reference point here is $(r + dr, \phi, z)$. Only the r element of the $\vec{F}$ matters, all other lie in the planes of the sides, and therefore do not contribute to the flux. Thus, the part of the flux due to these two sides is

$$d\Phi_r = F_r(r+dr, \phi, z)(r+dr)d\phi dz - F_r(r, \phi, z)rd\phi dz =$$

$$= (F_r(r+dr, \phi, z)r - F_r(r, \phi, z)r + drF_r(r+dr, \phi, z))\, d\phi dz =$$

$$= \left(F_r(r, \phi, z)r + \frac{\partial F_r}{\partial r}drr - F_r(r, \phi, z)r + drF_r(r, \phi, z) + (dr)^2\frac{\partial F_r}{\partial r}\right) d\phi dz$$

Since all changes are small, we can disregard the element of order $(dr)^2$.
Then

$$d\Phi_r \approx \left(\frac{\partial F_r}{\partial r}r + F_r\right) drd\phi dz = \frac{\partial (rF_r)}{\partial r}drd\phi dz = \frac{1}{r}\frac{\partial (rF_r)}{\partial r}rdrd\phi dz = \frac{1}{r}\frac{\partial (rF_r)}{\partial r}dV$$

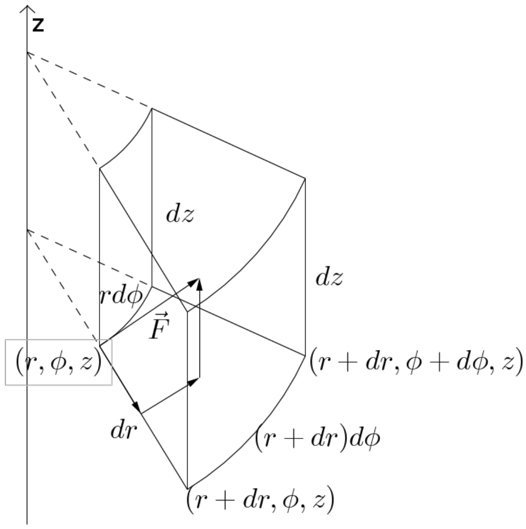

Figure 7: Divergence from small volume element in cylindrical coordinates.

Along the $\hat{e}_\phi$ direction, the analogous analysis leads to

$$d\Phi_\phi = F_\phi(r,\phi+d\phi,z)drdz - F_\phi(r,\phi,z)drdz = \frac{\partial F_\phi}{\partial \phi}drdzd\phi = \frac{1}{r}\frac{\partial F_\phi}{\partial \phi}rdrd\phi dz = \frac{1}{r}\frac{\partial F_\phi}{\partial \phi}dV$$

And finally along the $\hat{k}$ direction

$$d\Phi_z = F_z(r,\phi,z+dz)rdrd\phi - F_z(r,\phi,z)rdrd\phi = \frac{\partial F_z}{\partial z}rdrd\phi dz = \frac{\partial F_z}{\partial z}dV$$

And the total flux from the small element is

$$d\Phi_z = \left(\frac{1}{r}\frac{\partial(rF_r)}{\partial r} + \frac{1}{r}\frac{\partial F_\phi}{\partial \phi} + \frac{\partial F_z}{\partial z}\right)dV = \nabla\cdot\vec{F}dV$$

And therefore in cylindrical coordinates

$$\nabla\cdot\vec{F} = \frac{1}{r}\frac{\partial(rF_r)}{\partial r} + \frac{1}{r}\frac{\partial F_\phi}{\partial \phi} + \frac{\partial F_z}{\partial z} \tag{39}$$

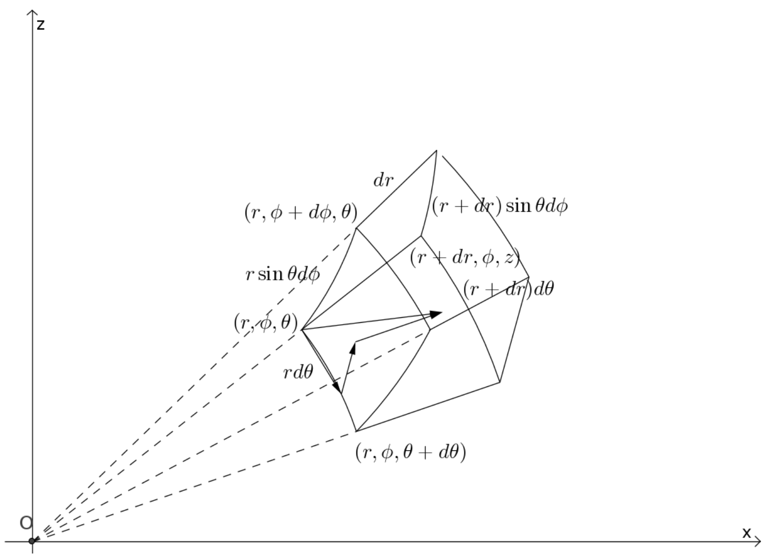

Figure 8: Divergence from small volume element in spherical coordinates.

Spherical cooridnates In spherical coordinates, we follow exactly analogous analysis

$$d\Phi_r = F_r(r+dr,\phi,\theta)(r+dr)\sin\theta d\phi(r+dr)d\theta - F_r(r,\phi,\theta)r\sin\theta d\phi r d\theta =$$

$$= \left(F_r(r,\phi,\theta)(r+dr)^2 + \frac{\partial F_r}{\partial r}dr(r+dr)^2 - F_r(r,\phi,\theta)r^2\right)\sin\theta d\phi d\theta =$$

$$= \left(F_r r^2 + 2F_r r dr + F_r(dr)^2 + \frac{\partial F_r}{\partial r}dr r^2 + 2\frac{\partial F_r}{\partial r}r(dr)^2 + \frac{\partial F_r}{\partial r}(dr)^3 - F_r r^2\right)\sin\theta d\phi d\theta =$$

$$= \left(2F_r r + F_r dr + \frac{\partial F_r}{\partial r}r^2 + 2\frac{\partial F_r}{\partial r}r dr + \frac{\partial F_r}{\partial r}(dr)^2\right)\sin\theta dr d\theta d\phi$$

Since dr is small

$$d\Phi_r \approx \left(2F_r r + \frac{\partial F_r}{\partial r}r^2\right)\sin\theta dr d\theta d\phi = \frac{\partial(r^2F_r)}{\partial r}\sin\theta dr d\theta d\phi =$$

$$= \frac{1}{r^2}\frac{\partial(r^2F_r)}{\partial r}r^2\sin\theta dr d\theta d\phi = \frac{1}{r^2}\frac{\partial(r^2F_r)}{\partial r}dV$$

$$d\Phi_\theta = F_\theta(r,\phi,\theta+d\theta)r\sin(\theta+d\theta)d\phi dr - F_\theta(r,\phi,\theta)r\sin\theta d\phi dr =$$

$$= F_\theta r(\sin\theta + \cos\theta d\theta)d\phi dr + \frac{\partial F_\theta}{\partial\theta}d\theta r(\sin\theta+\cos\theta d\theta)d\phi dr - F_\theta r\sin\theta d\phi dr =$$

$$= \left(F_\theta\cos\theta + \frac{\partial F_\theta}{\partial\theta}\sin\theta + \frac{\partial F_\theta}{\partial\theta}\cos\theta d\theta\right)r d\theta d\phi dr \approx \left(F_\theta\cos\theta + \frac{\partial F_\theta}{\partial\theta}\sin\theta\right)r d\theta d\phi dr =$$

$$= \frac{\partial(\sin\theta F_\theta)}{\partial\theta}r d\theta d\phi dr = \frac{1}{r\sin\theta}\frac{\partial(\sin\theta F_\theta)}{\partial\theta}r^2\sin\theta dr d\theta d\phi = \frac{1}{r\sin\theta}\frac{\partial(\sin\theta F_\theta)}{\partial\theta}dV$$

$$d\Phi_\phi = F_\phi(r,\phi+d\phi,\theta)r d\theta dr - F_\phi(r,\phi,\theta)r d\theta dr = \left(F_\phi + \frac{\partial F_\phi}{\partial\phi}d\phi - F_\phi\right)r d\theta dr =$$

$$= \frac{\partial F_\phi}{\partial\phi}r dr d\theta d\phi = \frac{1}{r\sin\theta}\frac{\partial F_\phi}{\partial\phi}r^2\sin\theta dr d\theta d\phi = \frac{1}{r\sin\theta}\frac{\partial F_\phi}{\partial\phi}dV$$

And therefore

$$d\Phi = \nabla\cdot\vec{F}dV = \left(\frac{1}{r^2}\frac{\partial(r^2F_r)}{\partial r} + \frac{1}{r\sin\theta}\frac{\partial(\sin\theta F_\theta)}{\partial\theta} + \frac{1}{r\sin\theta}\frac{\partial F_\phi}{\partial\phi}dV\right)$$

$$\nabla\cdot\vec{F} = \frac{1}{r^2}\frac{\partial(r^2F_r)}{\partial r} + \frac{1}{r\sin\theta}\frac{\partial(\sin\theta F_\theta)}{\partial\theta} + \frac{1}{r\sin\theta}\frac{\partial F_\phi}{\partial\phi} \tag{40}$$

3.1.3 Curl

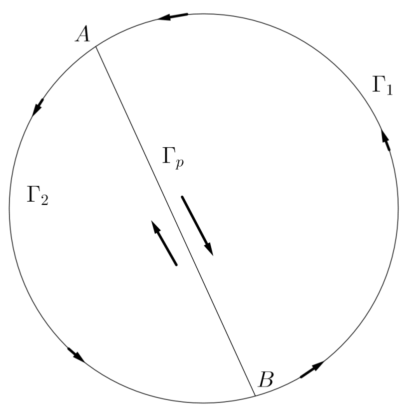

Figure 9: General case is illustrated for circular path Γ and line segment as Γ_p. The arrows indicate the direction of integration.

Consider cirulation of a vector along a certain path Γ.

$$C = \oint_\Gamma \vec{F} \cdot d\vec{l}$$

Now imagine that connect two points (A and B) of the path with intermediate path Γ_p. Therefore, we can now do integrals along to paths from A to B both along parts of Γ and along the intermediate path. Lets then do, rather arbitrarily, two closed loop integrals. One starting from A to B along the intermediate path and then returning to A along one part of Γ, which we denote as Γ_1. The second integral will start from B to A and then return to B via second part of Γ, Γ_2. This way, at points close to A and B that lie on Γ, the integration will have the same direction.
These integrals are

$$I_1 = \oint_{(A,B,\Gamma_1)} \vec{F} \cdot d\vec{l} = \int_A^B \vec{F} \cdot d\vec{l} + \int_{\Gamma_1} \vec{F} \cdot d\vec{l}$$

$$I_2 = \oint_{(B,A,\Gamma_2)} \vec{F} \cdot d\vec{l} = \int_B^A \vec{F} \cdot d\vec{l} + \int_{\Gamma_2} \vec{F} \cdot d\vec{l} = \int_{\Gamma_2} \vec{F} \cdot d\vec{l} - \int_A^B \vec{F} \cdot d\vec{l}$$

And therefore the sum of the integrals is

$$I_1 + I_2 = \int_{\Gamma_1} \vec{F} \cdot d\vec{l} + \int_{\Gamma_2} \vec{F} \cdot d\vec{l} = \oint_\Gamma \vec{F} \cdot d\vec{l} = C \tag{41}$$

Therefore, the circulation over a loop can be gained by summing the circulations over constituent loops. Important fact is that the choice of Γ_p is not important as long as it connects two distinct points on Γ.
We now imagine that we fill Γ with some open surface. Then, if we find the circulation of any small loop that lies in this surface, we can find the circulation over Γ just by integrating over this surface.
There is one more important observation to be made. The circulation of an infinitesimally small loop lying in the surface can be also calculated as a sum of over three closed loops, each of which lies in a plane normal to a unit direction in some coordinate system. These three loops are formed by projection of our general loop onto the planes normal to coordinate directions.

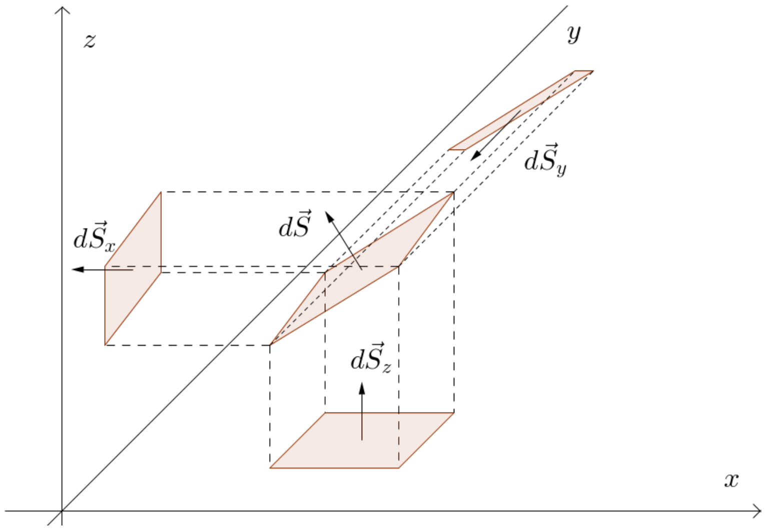

Figure 10: Decomposition of infinitesimal loop into 3 perpendicular loops. The surface vectors of these loops are the components of the loop surface vector $d\vec{S}$. The circulation along the loop $d\vec{S}$ can be obtained as a sum of three circulations, each over separate loop $d\vec{S}_x$, $d\vec{S}_y$ and $d\vec{S}_z$. The loops are distanced from the original loop, in order to make image clearer.

This can be done because small step along the loop can be always decomposed into sum of three small steps along coordinate directions, i. e.

$$d\vec{l} = l_1\hat{e}_1 + l_2\hat{e}_2 + l_3\hat{e}_3$$

Carrying out these three integrals along the three loops is then just decomposition of vectors. We therefore can the circulation along each of these loops and sum it together to get circulation of the initial loop.

Furthermore, we can find circulations of some very small unit loops in planes perpendicular to directions of coordinate system, and then sum their circulations, multiplied by components of $d\vec{S}$ in different directions. Then, we can talk about some form of unit circulation vector $\vec{C}_u$, which gives circulation around a small loop when scalar product with surface vector of this loop is taken, i.e. the small circulation dC along loop $d\vec{S}$ can be described as

$$dC = d\vec{S} \cdot \vec{C}_u$$

, where $\vec{C}_u$ consists of small unit loops' circulations (loops lying in planes perpendicular to unit vectors). We can also switch this and use unit vector in the direction of the infinitesimal loop $d\vec{S}$, which we denote as $\hat{n}$ and use infinitesimal unit loops for circulation $d\vec{C}_u$, so that

$$dC = \hat{n} \cdot d\vec{C}_u$$

Cartesian coordinates In cartesian coordinates, we have 3 planes to consider, plane perpendicular to $\hat{\imath}$, to $\hat{\jmath}$ and to $\hat{k}$. Lets start with the plane normal to $\hat{\imath}$, in order to x component of $d\vec{C}_u$, which is denoted as dC_{ux}.

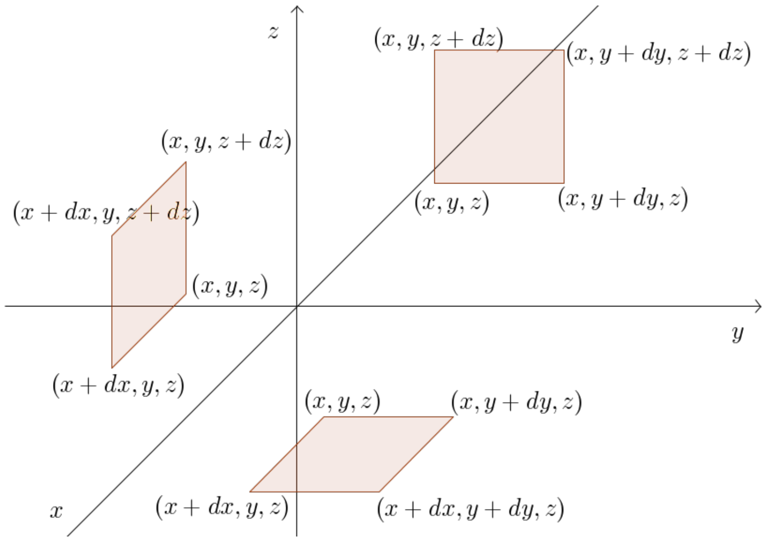

Figure 11: Different unit loops in cartesian coordinate system.

To determine the direction in which we take the circulation, we will use the right hand rule - when looking on the plane and $\hat{\imath}$ vector goes out of the plane, we integrate anticlockwise.
Therefore, we have four components of the circulation dC_{ux}. First is from (x, y, z) to $(x, y + dy, z)$. We choose the starting point as the reference point, and then approximate this component as

$$\int_{(x,y,z)}^{(x+dx,y,z)} \vec{F} \cdot d\vec{l} \approx \vec{F}(x, y, z) \cdot dy\hat{\jmath} = F_y(x, y, z)dy$$

The second part is from $(x, y + dy, z)$ to $(x, y + dy, z + dz)$.

$$\int_{(x,y+dy,z)}^{(x,y+dy,z+dz)} \vec{F} \cdot d\vec{l} \approx \vec{F}(x, y + dy, z) \cdot dz\hat{k} = F_z(x, y + dy, z)dz = F_z(x, y, z)dz + \frac{\partial F_z}{\partial y}dydz$$

The third part is from $(x, y + dy, z + dz)$ to $(x, y, z + dz)$

$$\int_{(x,y+dy,z+dz)}^{(x,y,z+dz)} \vec{F} \cdot d\vec{l} \approx \vec{F}(x, y + dy, z + dz) \cdot (-dy)\hat{\jmath} = -F_y(x, y, z)dy - \frac{\partial F_y}{\partial y}(dy)^2 - \frac{\partial F_y}{\partial z}dzdy$$

And the last part is

$$\int_{(x,y,z+dz)}^{(x,y,z)} \vec{F} \cdot d\vec{l} \approx \vec{F}(x, y, z + dz) \cdot (-dz)\hat{k} = -F_z(x, y, z)dz - \frac{\partial F_z}{\partial z}(dz)^2$$

So, the circulation is

$$dC_{ux} \approx \frac{\partial F_z}{\partial y}dydz - \frac{\partial F_y}{\partial y}(dy)^2 - \frac{\partial F_y}{\partial z}dzdy - \frac{\partial F_z}{\partial z}(dz)^2$$

If we take out the factor $dydz$

$$dC_{ux} = \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} - \frac{\partial F_y}{\partial y}\frac{dy}{dz} - \frac{\partial F_z}{\partial z}\frac{dz}{dy}\right) dydz$$

Since the unit directions are normal, $\frac{dy}{dz} = 0 = \frac{dz}{dy}$. Therefore, finally

$$dC_{ux} = \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}\right) dS_x$$

where $dS_x = dydz$ is the area of the unit plane perpendicular to x. Similarly, for other two directions

$$\int_{(x,y,z)}^{(x,y,z+dz)} \vec{F}\cdot d\vec{l} \approx F_z(x,y,z)dz$$

$$\int_{(x,y,z+dz)}^{(x+dx,y,z+dz)} \vec{F}\cdot d\vec{l} \approx F_x(x,y,z)dx + \frac{\partial F_x}{\partial z}dxdz$$

$$\int_{(x+dx,y,z+dz)}^{(x+dx,y,z)} \vec{F}\cdot d\vec{l} \approx -F_z(x,y,z)dz - \frac{\partial F_z}{\partial z}(dz)^2 - \frac{\partial F_z}{\partial x}dxdz$$

$$\int_{(x+dx,y,z)}^{(x,y,z)} \vec{F}\cdot d\vec{l} \approx -F_x(x,y,z)dx - \frac{\partial F_x}{\partial x}(dx)^2$$

$$dC_{uy} = \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}\right) dS_y$$

$$\int_{(x,y,z)}^{(x+dx,y,z)} \vec{F}\cdot d\vec{l} \approx F_x(x,y,z)dx$$

$$\int_{(x+dx,y,z)}^{(x+dx,y+dy,z)} \vec{F}\cdot d\vec{l} \approx F_y(x,y,z)dy + \frac{\partial F_y}{\partial x}dxdy$$

$$\int_{(x+dx,y+dy,z)}^{(x,y+dy,z)} \vec{F}\cdot d\vec{l} \approx -F_x(x,y,z)dx - \frac{\partial F_x}{\partial x}(dx)^2 - \frac{\partial F_x}{\partial y}dxdy$$

$$\int_{(x,y+dy,z)}^{(x,y,z)} \vec{F}\cdot d\vec{l} \approx -F_y(x,y,z)dy - \frac{\partial F_y}{\partial y}(dy)^2$$

$$dC_{uz} = \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right) dS_z$$

We can then used identity $d\vec{S}\cdot\vec{C}_u = \hat{n}\cdot d\vec{C}_u$ and knowledge that $d\vec{S} = dS_x\hat{\imath} + dS_y\hat{\jmath} + dS_z\hat{k}$ to determine

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

Hence, the total circulation of $\vec{F}$ along Γ is given by

$$\oint_\Gamma \vec{F}\cdot d\vec{l} = \iint_S dC = \iint_S \vec{C}_u\cdot d\vec{S} = \iint_S (\nabla\times\vec{F})\cdot d\vec{S} \tag{42}$$

where S is the surface that fills Γ. This is called the Stokes' theorem. The operation $\nabla\times\vec{F}$ is called the curl of $\vec{F}$.

Again, using the definition as vector circulations around unit small loops, we can determine the curl in alternative coordinate systems.

Cylindrical coordinates We start with the loop normal to radial direction. Parts of the line integral are

$$\int_{(r,\phi,z)}^{(r,\phi+d\phi,z)} \vec{F}\cdot d\vec{l} \approx F_\phi(r,\phi,z)rd\phi$$

$$\int_{(r,\phi+d\phi,z)}^{(r,\phi+d\phi,z+dz)} \vec{F}\cdot d\vec{l} \approx F_z(r,\phi,z)dz + \frac{\partial F_z}{\partial \phi}d\phi dz$$

$$\int_{(r,\phi+d\phi,z+dz)}^{(r,\phi,z+dz)} \vec{F}\cdot d\vec{l} \approx -F_\phi(r,\phi,z)rd\phi - \frac{\partial F_\phi}{\partial \phi}r(d\phi)^2 - \frac{\partial F_\phi}{\partial z}rd\phi dz$$

$$\int_{(r,\phi,z+dz)}^{(r,\phi,z)} \vec{F}\cdot d\vec{l} \approx -F_z(r,\phi,z)dz - \frac{\partial F_z}{\partial z}(dz)^2$$

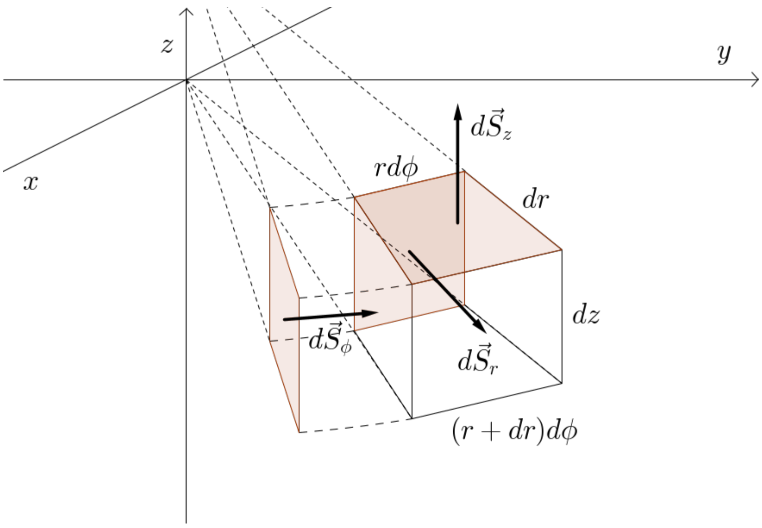

Figure 12: Curl loops in cylindrical coordinates

Hence the sum is

$$(\nabla\times\vec{F})_r dS_r = \left(\frac{1}{r}\frac{\partial F_z}{\partial \phi} - \frac{\partial F_\phi}{\partial z}\right) rd\phi dz$$

$$(\nabla\times\vec{F})_r = \frac{1}{r}\frac{\partial F_z}{\partial \phi} - \frac{\partial F_\phi}{\partial z}$$

For ϕ element ($\partial(dS_\phi)$ denotes boundary of small surface dS_ϕ)

$$\oint_{\partial(dS_\phi)} \vec{F}\cdot d\vec{l} \approx$$

$$\approx F_z(r,\phi,z)dz + F_r(r,\phi,z)dr + \frac{\partial F_r}{\partial z}drdz - F_z(r,\phi,z)dz - \frac{\partial F_z}{\partial r}dzdr - \frac{\partial F_z}{\partial z}(dz)^2 - F_r(r,\phi,z)dr - \frac{\partial F_r}{\partial r}(dr)^2 \approx$$

$$\approx \left(\frac{\partial F_r}{\partial z} - \frac{\partial F_z}{\partial r}\right) drdz$$

$$(\nabla\times\vec{F})_\phi = \frac{\partial F_r}{\partial z} - \frac{\partial F_z}{\partial r}$$

And finally, for the z element (dropping explicit point denotation of point (r,ϕ,z))

$$\oint_{\partial(dS_z)} \vec{F}\cdot d\vec{l} \approx F_r dr + F_\phi(r+dr)d\phi + \frac{\partial F_\phi}{\partial r}dr(r+dr)d\phi - F_r dr - \frac{\partial F_r}{\partial r}(dr)^2 - \frac{\partial F_r}{\partial \phi}d\phi dr - F_\phi rd\phi - \frac{\partial F_\phi}{\partial \phi}r(d\phi)^2$$

$$\oint_{\partial(dS_z)} \vec{F}\cdot d\vec{l} \approx \left(F_\phi + \frac{\partial F_\phi}{\partial r}r - \frac{\partial F_r}{\partial \phi}\right) drd\phi = \frac{1}{r}\left(\frac{\partial (rF_\phi)}{\partial r} - \frac{\partial F_r}{\partial \phi}\right) rdrd\phi$$

$$(\nabla \times \vec{F})_z = \frac{1}{r}\left(\frac{\partial(rF_\phi)}{\partial r} - \frac{\partial F_r}{\partial \phi}\right)$$

Hence the curl is

$$\nabla \times \vec{F} = \left(\frac{1}{r}\frac{\partial F_z}{\partial \phi} - \frac{\partial F_\phi}{\partial z}\right)\hat{e}_r + \left(\frac{\partial F_r}{\partial z} - \frac{\partial F_z}{\partial r}\right)\hat{e}_\phi + \frac{1}{r}\left(\frac{\partial(rF_\phi)}{\partial r} - \frac{\partial F_r}{\partial \phi}\right)\hat{e}_z \tag{43}$$

Spherical coordinates For illustration, consider again Figure 8.
First, we start with the r component of the curl

$$\oint_{\partial(dS_r)} \vec{F}\cdot d\vec{l} \approx F_\theta r d\theta + F_\phi r(\sin\theta + \cos\theta d\theta)d\phi + \frac{\partial F_\phi}{\partial \theta} d\theta r(\sin\theta + \cos\theta d\theta)d\phi -$$

$$-F_\theta r d\theta - \frac{\partial F_\theta}{\partial \theta} r(d\theta)^2 - \frac{\partial F_\theta}{\partial \phi} r d\theta d\phi - F_\phi r \sin\theta d\phi - \frac{\partial F_\phi}{\partial \phi} r\sin\theta(d\phi)^2 \approx$$

$$\approx \left(F_\phi\cos\theta + \frac{\partial F_\phi}{\partial\theta}\sin\theta - \frac{\partial F_\theta}{\partial\phi}\right) r d\theta d\phi = \frac{1}{r\sin\theta}\left(\frac{\partial(\sin\theta F_\phi)}{\partial\theta} - \frac{\partial F_\theta}{\partial\phi}\right) r^2 \sin\theta d\theta d\phi$$

$$(\nabla\times\vec{F})_r = \frac{1}{r\sin\theta}\left(\frac{\partial(\sin\theta F_\phi)}{\partial\theta} - \frac{\partial F_\theta}{\partial\phi}\right)$$

Now, the θ component

$$\oint_{\partial(dS_\theta)} \vec{F}\cdot d\vec{l} \approx F_\phi r\sin\theta d\phi + F_r dr + \frac{\partial F_r}{\partial\phi} dr d\phi - F_\phi(r+dr)\sin\theta d\phi - \frac{\partial F_\phi}{\partial\phi}(r+dr)\sin\theta(d\phi)^2 - \frac{\partial F_\phi}{\partial r}(r+dr)\sin\theta d\phi dr -$$

$$-F_r dr - \frac{\partial F_r}{\partial r}(dr)^2 \approx \left(\frac{\partial F_r}{\partial\phi} - F_\phi\sin\theta - \frac{\partial F_\phi}{\partial r} r\sin\theta\right) dr d\phi =$$

$$= \frac{1}{r\sin\theta}\left(\frac{\partial F_r}{\partial\phi} - \sin\theta\frac{\partial(F_\phi r)}{\partial r}\right) r\sin\theta dr d\phi$$

$$(\nabla\times\vec{F})_\theta = \frac{1}{r\sin\theta}\left(\frac{\partial F_r}{\partial\phi} - \sin\theta\frac{\partial(rF_\phi)}{\partial r}\right)$$

And the ϕ component

$$\oint_{\partial(dS_\phi)} \vec{F}\cdot d\vec{l} \approx F_r dr + F_\theta(r+dr)d\theta + \frac{\partial F_\theta}{\partial r}(r+dr)dr d\theta - F_r dr - \frac{\partial F_r}{\partial r}(dr)^2 - \frac{\partial F_r}{\partial\theta} dr d\theta - F_\theta r d\theta - \frac{\partial F_\theta}{\partial\theta} r(d\theta)^2 =$$

$$= \left(F_\theta + \frac{\partial F_\theta}{\partial r} r - \frac{\partial F_r}{\partial\theta}\right) dr d\theta = \frac{1}{r}\left(\frac{\partial(rF_\theta)}{\partial r} - \frac{\partial F_r}{\partial\theta}\right) r d\theta dr$$

$$(\nabla\times\vec{F})_\phi = \frac{1}{r}\left(\frac{\partial(rF_\theta)}{\partial r} - \frac{1}{r}\frac{\partial F_r}{\partial\theta}\right)$$

And therefore

$$\nabla\times\vec{F} = \frac{1}{r\sin\theta}\left(\frac{\partial(\sin\theta F_\phi)}{\partial\theta} - \frac{\partial F_\theta}{\partial\phi}\right)\hat{e}_r + \frac{1}{r\sin\theta}\left(\frac{\partial F_r}{\partial\phi} - \sin\theta\frac{\partial(rF_\phi)}{\partial r}\right)\hat{e}_\theta + \frac{1}{r}\left(\frac{\partial(rF_\theta)}{\partial r} - \frac{\partial F_r}{\partial\theta}\right)\hat{e}_\phi \tag{44}$$

3.1.4 Laplacian Operator

Laplacian operator ∇^2 is a scalar operator acting on a scalar function. The Laplacian operator is equal to the divergence of a gradient of the scalar function. Using derivations above, in different coordinate systems, we can find explicit forms of Laplacian.

Cartesian coordinates Gradient of f is

$$\nabla f = \frac{\partial f}{\partial x}\hat{i} + \frac{\partial f}{\partial y}\hat{j} + \frac{\partial f}{\partial z}\hat{k}$$

Divergence is

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

And therefore the divergence of gradient is

$$\nabla^2 f = \nabla \cdot (\nabla f) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) + \frac{\partial}{\partial z}\left(\frac{\partial f}{\partial z}\right) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

Cylindrical coordinates The gradient is

$$\nabla f = \frac{\partial f}{\partial r}\hat{e}_r + \frac{1}{r}\frac{\partial f}{\partial \phi}\hat{e}_\phi + \frac{\partial f}{\partial z}\hat{e}_z$$

The divergence is

$$\nabla \cdot \vec{F} = \frac{1}{r}\frac{\partial (rF_r)}{\partial r} + \frac{1}{r}\frac{\partial F_\phi}{\partial \phi} + \frac{\partial F_z}{\partial z}$$

Therefore

$$\nabla^2 f = \nabla \cdot (\nabla f) = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial f}{\partial r}\right) + \frac{1}{r}\frac{\partial}{\partial \phi}\left(\frac{1}{r}\frac{\partial f}{\partial \phi}\right) + \frac{\partial^2 f}{\partial z^2} = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial f}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 f}{\partial \phi^2} + \frac{\partial^2 f}{\partial z^2}$$

Spherical coordinates The gradient is

$$\nabla f = \frac{\partial f}{\partial r}\hat{e}_r + \frac{1}{r}\frac{\partial f}{\partial \theta}\hat{e}_\theta + \frac{1}{r\sin\theta}\frac{\partial f}{\partial \phi}$$

The divergence is

$$\nabla \cdot \vec{F} = \frac{1}{r^2}\frac{\partial (r^2 F_r)}{\partial r} + \frac{1}{r\sin\theta}\frac{\partial (\sin\theta F_\theta)}{\partial \theta} + \frac{1}{r\sin\theta}\frac{\partial F_\phi}{\partial \phi}$$

Therefore

$$\nabla^2 f = \nabla \cdot (\nabla f) = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial r}\right) + \frac{1}{r\sin\theta}\frac{\partial}{\partial \theta}\left(\frac{\sin\theta}{r}\frac{\partial f}{\partial \theta}\right) + \frac{1}{r\sin\theta}\frac{\partial}{\partial \phi}\left(\frac{1}{r\sin\theta}\frac{\partial f}{\partial \phi}\right) =$$

$$= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial r}\right) + \frac{1}{r^2\sin\theta}\left(\sin\theta\frac{\partial f}{\partial \theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial \phi^2}$$

Sometimes, laplacian also acts on a vector. In that case, laplacian applies to each component of the vector, i. e.

$$\nabla^2 \vec{F} = \nabla^2 F_1\hat{e}_1 + \nabla^2 F_2\hat{e}_2 + \nabla^2 F_3\hat{e}_3$$

3.2 Operator identities

The differential operators have properties that are independent of the coordinate system we choose. In following section, I list some properties of the differential operators and show them in the cartesian coordinate system, but they hold in any other coordinate system as well.

Curl/Div orthogonality Take a divergence of a curl of a vector in cartesian coordinates

$$\nabla \cdot (\nabla \times \vec{F}) = \frac{\partial}{\partial x}\left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}\right) + \frac{\partial}{\partial y}\left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}\right) + \frac{\partial}{\partial z}\left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right) =$$

$$= \frac{\partial^2 F_z}{\partial x \partial y} - \frac{\partial^2 F_z}{\partial y \partial x} + \frac{\partial^2 F_x}{\partial y \partial z} - \frac{\partial^2 F_x}{\partial z \partial y} + \frac{\partial^2 F_y}{\partial z \partial x} - \frac{\partial^2 F_y}{\partial x \partial z} = 0$$

$$\nabla \cdot (\nabla \times \vec{F}) = 0 \tag{45}$$

Scalar potential Take a curl of a gradient in cartesian coordinates

$$\nabla \times (\nabla f) = \left(\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial z}\right) - \frac{\partial}{\partial z}\left(\frac{\partial f}{\partial y}\right)\right)\hat{i} + \left(\frac{\partial}{\partial z}\left(\frac{\partial f}{\partial x}\right) - \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial z}\right)\right)\hat{j} + \left(\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) - \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\right)\hat{k}$$

$$\nabla \times (\nabla f) = \vec{0} \tag{46}$$

This leads to first two identical conditions for a conservative field $\vec{F}$. These conditions two conditions are $\vec{F} = \nabla f$ and $\nabla \times \vec{F} = \vec{0}$, where f is the so called scalar potential.

Gradient distribution Consider a product of two scalar functions. The gradient of the product is

$$\nabla(fg) = \frac{\partial}{\partial x}(fg)\hat{i} + \frac{\partial}{\partial y}(fg)\hat{j} + \frac{\partial}{\partial z}(fg)\hat{k} = \frac{\partial f}{\partial x}g\hat{i} + f\frac{\partial g}{\partial x}\hat{i} + \frac{\partial f}{\partial y}g\hat{j} + f\frac{\partial g}{\partial y}\hat{j} + \frac{\partial f}{\partial z}g\hat{k} + f\frac{\partial g}{\partial z}\hat{k} =$$

$$= f\left(\frac{\partial g}{\partial x}\hat{i} + \frac{\partial g}{\partial y}\hat{j} + \frac{\partial g}{\partial z}\hat{k}\right) + \left(\frac{\partial f}{\partial x}\hat{i} + \frac{\partial f}{\partial y}\hat{j} + \frac{\partial f}{\partial z}\hat{k}\right)g = f(\nabla g) + (\nabla f)g$$

$$\nabla(fg) = f(\nabla g) + (\nabla f)g \tag{47}$$

In a special case when f is some constant c in space

$$\nabla(cg) = c\nabla g$$

Divergence distribution Consider a vector $\vec{F}$ that is a multiplied by scalar function f. The divergence of this new vector is

$$\nabla\cdot(f\vec{F}) = \frac{\partial(fF_x)}{\partial x} + \frac{\partial(fF_y)}{\partial y} + \frac{\partial(fF_z)}{\partial z} = \frac{\partial f}{\partial x}F_x + \frac{\partial F_x}{\partial x}f + \frac{\partial f}{\partial y}F_y + \frac{\partial F_y}{\partial y}f + \frac{\partial f}{\partial z}F_z + \frac{\partial F_z}{\partial z}f = (\nabla f)\cdot\vec{F} + f(\nabla\cdot\vec{F})$$

$$\nabla \cdot (f\vec{F}) = (\nabla f) \cdot \vec{F} + f(\nabla \cdot \vec{F}) \tag{48}$$

In special case of constant vector $\vec{C}$ in space

$$\nabla \cdot (f\vec{C}) = (\nabla f) \cdot \vec{C}$$

Curl distribution Consider taking curl of vector $f\vec{F}$

$$\nabla \times (f\vec{F}) = \left(\frac{\partial(fF_z)}{\partial y} - \frac{\partial(fF_y)}{\partial z}\right)\hat{i} + \left(\frac{\partial(fF_x)}{\partial z} - \frac{\partial(fF_z)}{\partial x}\right)\hat{j} + \left(\frac{\partial(fF_y)}{\partial x} - \frac{\partial(fF_x)}{\partial y}\right)\hat{k} =$$

$$= \left(\frac{\partial f}{\partial y}F_z - \frac{\partial f}{\partial z}F_y\right)\hat{i} + f\left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}\right)\hat{i} + \left(\frac{\partial f}{\partial z}F_x - \frac{\partial f}{\partial x}F_z\right)\hat{j} + f\left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}\right)\hat{j} +$$

$$+ \left(\frac{\partial f}{\partial x}F_y - \frac{\partial f}{\partial y}F_x\right)\hat{k} + f\left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right)\hat{k} = (\nabla f) \times \vec{F} + f(\nabla \times \vec{F})$$

$$\nabla \times (f\vec{F}) = (\nabla f) \times \vec{F} + f(\nabla \times \vec{F}) \tag{49}$$

Again, in special case of constant vector $\vec{C}$

$$\nabla \times (f\vec{C}) = (\nabla f) \times \vec{C}$$

Curl of curl Consider taking a curl of curl of a vector $\vec{F}$

$$(\nabla\times(\nabla\times\vec{F}))_x = \frac{\partial}{\partial y}\left(\frac{\partial F_y}{\partial x}-\frac{\partial F_x}{\partial y}\right)-\frac{\partial}{\partial z}\left(\frac{\partial F_x}{\partial z}-\frac{\partial F_z}{\partial x}\right)=$$

$$=\frac{\partial}{\partial x}\left(\frac{\partial F_y}{\partial y}+\frac{\partial F_z}{\partial z}\right)-\left(\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)F_x=\frac{\partial}{\partial x}\left(\frac{\partial F_x}{\partial x}+\frac{\partial F_y}{\partial y}+\frac{\partial F_z}{\partial z}\right)-\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)F_x$$

$$\nabla\times(\nabla\times\vec{F})_y = \frac{\partial}{\partial z}\left(\frac{\partial F_z}{\partial y}-\frac{\partial F_y}{\partial z}\right)-\frac{\partial}{\partial x}\left(\frac{\partial F_y}{\partial x}-\frac{\partial F_x}{\partial y}\right)=$$

$$\frac{\partial}{\partial y}\left(\frac{\partial F_z}{\partial z}+\frac{\partial F_x}{\partial x}\right)-\left(\frac{\partial^2}{\partial z^2}+\frac{\partial^2}{\partial x^2}\right)F_y=\frac{\partial}{\partial y}\left(\frac{\partial F_x}{\partial x}+\frac{\partial F_y}{\partial y}+\frac{\partial F_z}{\partial z}\right)-\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)F_y$$

$$\nabla\times(\nabla\times\vec{F})_z = \frac{\partial}{\partial x}\left(\frac{\partial F_x}{\partial z}-\frac{\partial F_z}{\partial x}\right)-\frac{\partial}{\partial y}\left(\frac{\partial F_z}{\partial y}-\frac{\partial F_y}{\partial z}\right)=$$

$$\frac{\partial}{\partial z}\left(\frac{\partial F_x}{\partial x}+\frac{\partial F_y}{\partial y}\right)-\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)F_z=\frac{\partial}{\partial z}\left(\frac{\partial F_x}{\partial x}+\frac{\partial F_y}{\partial y}+\frac{\partial F_z}{\partial z}\right)-\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)F_z$$

Therefore, we have

$$\nabla\times(\nabla\times\vec{F}) = \left(\frac{\partial}{\partial x}(\nabla\cdot\vec{F})-\nabla^2F_x\right)\hat{\imath}+\left(\frac{\partial}{\partial x}(\nabla\cdot\vec{F})-\nabla^2F_y\right)\hat{\jmath}+\left(\frac{\partial}{\partial z}(\nabla\cdot\vec{F})-\nabla^2F_z\right)\hat{k}=$$

$$=\nabla(\nabla\cdot\vec{F})-\nabla^2\vec{F}$$

$$\nabla\times(\nabla\times\vec{F})=\nabla(\nabla\cdot\vec{F})-\nabla^2\vec{F} \tag{50}$$

3.3 Position Vectors and their Time Derivatives

Position vector is a special vector that points from the origin to a point where some object is located. The time derivative of this position vector is the velocity of the object, its second time derivative is the acceleration.
According to the partial derivatives and total derivatives chain rule, the total time derivative of a position vector is

$$\frac{d\vec{r}}{dt}=\frac{d}{dt}(r_1\hat{e}_1+r_2\hat{e}_2+r_3\hat{e}_3)=\frac{dr_1}{dt}\hat{e}_1+r_1\frac{d(\hat{e}_1)}{dt}+\frac{dr_2}{dt}\hat{e}_2+r_2\frac{d(\hat{e}_2)}{dt}+\frac{dr_3}{dt}\hat{e}_3+r_3\frac{d(\hat{e}_3)}{dt}$$

The $\frac{dr_i}{dt}$ terms have to be determined depending on the coordinate system. As unit vectors in certain directions only depend on position and not on the time, the total time derivative for unit vector $\hat{e}_i$ becomes

$$\frac{d(\hat{e}_i)}{dt}=\frac{\partial(\hat{e}_i)}{\partial t}+\frac{dr_1}{dt}\frac{\partial(\hat{e}_i)}{\partial r_1}+\frac{dr_2}{dt}\frac{\partial(\hat{e}_i)}{\partial r_2}+\frac{dr_3}{dt}\frac{\partial(\hat{e}_i)}{\partial r_3}=\sum_{j=1}^{3}\frac{dr_j}{dt}\frac{\partial(\hat{e}_i)}{\partial r_j}$$

Cartesian coordinates In cartesian coordiantes, position vector is determined by how far along each unit direction the object is. It therefore has the form

$$\vec{r}=x\hat{\imath}+y\hat{\jmath}+z\hat{k}$$

The unit vectors are independent of position and time, therefore the velocity vector is simply

$$\vec{v}=\frac{d\vec{r}}{dt}=\frac{dx}{dt}\hat{\imath}+\frac{dy}{dt}\hat{\jmath}+\frac{dz}{dt}\hat{k}$$

And the acceleration vector is

$$\vec{a}=\frac{d\vec{v}}{dt}=\frac{d^2x}{dt^2}\hat{\imath}+\frac{d^2y}{dt^2}\hat{\jmath}+\frac{d^2z}{dt^2}\hat{k}$$

Cylindrical coordinates The position in cylindrical coordinates is determined by the distance along the $\hat{e}_r$ vector and the $\hat{k}$ vector, and the position vector is therefore

$$\vec{r} = r\hat{e}_r + z\hat{k}$$

I will rewrite the $\hat{e}_r$ vector in terms of the cartesian unit vectors, which are independent of space. Then, the velocity vector is

$$\vec{v} = \frac{d\vec{r}}{dt} = \frac{dr}{dt}\hat{e}_r + r\frac{d}{dt}\left(\cos\phi\hat{i} + \sin\phi\hat{j}\right) + \frac{dz}{dt}\hat{k} = \frac{dr}{dt}\hat{e}_r + r\left(-\sin\phi\hat{i} + \cos\phi\hat{j}\right)\frac{d\phi}{dt} + \frac{dz}{dt}\hat{k} = \frac{dr}{dt}\hat{e}_r + r\frac{d\phi}{dt}\hat{e}_\phi + \frac{dz}{dt}\hat{k}$$

Here, we also determined that

$$\frac{d\hat{e}_r}{dt} = \frac{d\phi}{dt}\hat{e}_\phi$$

Similarly

$$\begin{aligned}\vec{a} = \frac{d\vec{v}}{dt} &= \frac{d^2r}{dt^2}\hat{e}_r + \frac{dr}{dt}\frac{d\phi}{dt}\hat{e}_\phi + \frac{dr}{dt}\frac{d\phi}{dt}\hat{e}_\phi + r\frac{d^2\phi}{dt^2}\hat{e}_\phi + r\frac{d\phi}{dt}\frac{d\hat{e}_\phi}{dt} + \frac{d^2z}{dt^2}\hat{k} = \\ &= \frac{d^2r}{dt^2}\hat{e}_r + \left(r\frac{d^2\phi}{dt^2} + 2\frac{dr}{dt}\frac{d\phi}{dt}\right)\hat{e}_\phi - r\left(\frac{d\phi}{dt}\right)^2\hat{e}_r + \frac{d^2z}{dt^2}\hat{k} = \\ &= \left(\frac{d^2r}{dt^2} - r\left(\frac{d\phi}{dt}\right)^2\right)\hat{e}_r + \left(r\frac{d^2\phi}{dt^2} + 2\frac{dr}{dt}\frac{d\phi}{dt}\right)\hat{e}_\phi + \frac{d^2z}{dt^2}\hat{k}\end{aligned}$$

Where I used

$$\frac{d\hat{e}_\phi}{dt} = -\frac{d\phi}{dt}\hat{e}_r$$

4 Integral Vector Calculus

4.1 Scalar difference

Consider integrating gradient of a scalar function f along some path Γ. What is the value of the integral I

$$I = \int_\Gamma (\nabla f)\cdot d\vec{l}$$

In cartesian coordinates

$$I = \int_\Gamma \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz = \int_{\Gamma_x}\frac{\partial f}{\partial x}dx + \int_{\Gamma_y}\frac{\partial f}{\partial y}dy + \int_{\Gamma_z}\frac{\partial f}{\partial z}dz = f(2) - f(1) = \Delta f$$

$$\Delta f = \int_\Gamma (\nabla f)\cdot d\vec{l} \tag{51}$$

where $f(2)$ is the value of the function at the beginning of the path and $f(1)$ is the value of the function at the end of the path and Γ_i are projections of Γ into the a direction. This applies in any coordinate system. We can therefore immidietely see that if the two points are the same, the value of the integral is zero. We also see that the integral only depends on the end points, and not on the path taken. These two are other two conditions for conservative fields.
The last condition that follows is that element $\vec{c}\cdot d\vec{r}$ is an exact differential for conservative field $\vec{c}$ (follows again from $\vec{c} = \nabla f$).

4.2 Green's theorem in plane

Green's theorem is special case of the Stoke's theorem in 2 dimensions. Because it is purely in 2 dimensions, we need a different way of derivation, as derivation we have done for Stoke's theorem required 3D vector field.
Consider we have some curve Γ which is entirely in xy plane. We integrate the vector filed in this plane along Γ. The vector field can be desribed as

$$\vec{F} = F_x(x,y)\hat{i} + F_y(x,y)\hat{j}$$

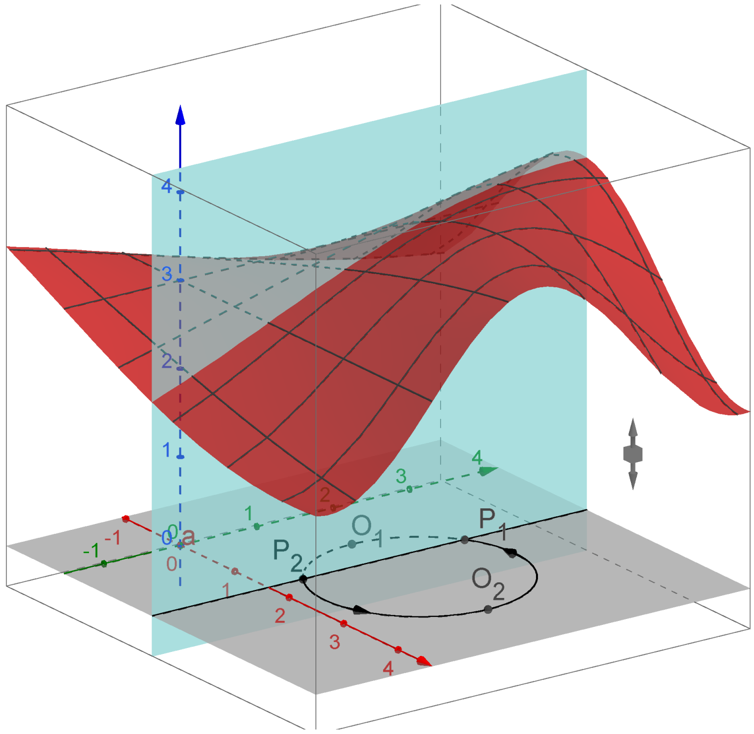

Figure 13: Illustration for Green's theorem in plane, with indication of direction of integration. The red surface is some scalar function of x and y.

The integral is

$$I = \oint_\Gamma \vec{F} \cdot d\vec{l} = \oint_\Gamma F_x dx + \oint_\Gamma F_y dy$$

This situation is presented for a circle in Fig. 13

We now split the curve Γ with points O_1 and O_2 in such a way so that path Γ_1 going from O_1 to O_2 along Γ and Γ_2 along Γ from O_2 to O_1 are both distinct functions of x. We suppose that the x coordinate of O_1 x_1 is smaller than that of O_2, which is x_2 (just dependent on our construction). We denote the function that traces Γ_1 as $y_1(x)$, the second one as $y_2(x)$. The first part of the integral I then becomes

$$\oint_\Gamma F_x dx = \int_{x_1}^{x_2} F_x(x, y_1(x))dx + \int_{x_2}^{x_1} F_x(x, y_2(x))dx = \int_{x_1}^{x_2} (F_x(x, y_1(x)) - F_x(x, y_2(x)))dx$$

But, F_x is simply a scalar function of two variables. The difference of such scalar function can be expressed in form of gradient integral

$$\Delta F_x = \int_{P_1}^{P_2} (\nabla F_x) \cdot d\vec{l}$$

where P_1 and P_2 are the points between which we take the difference. Taking the opposite of difference and writing the form explicitly

$$-\Delta F_x = F_x(x_1, y_1) - F_x(x_2, y_2) = \int_{(x_1,y_1)}^{(x_2,y_2)} -(\nabla F_x) \cdot d\vec{l}$$

Now, we apply this to our integral

$$F_x(x, y_1) - F_x(x, y_2) = \int_{(x,y_1)}^{(x,y_2)} -(\nabla F_x) \cdot d\vec{l} = \int_{y_1}^{y_2} -(\nabla F_x) \cdot (dy\hat{j}) = \int_{y_1}^{y_2} -\frac{\partial F_x}{\partial y} dy$$

$$\oint_\Gamma F_x dx = \int_{x_1}^{x_2} (-\Delta F_x)dx = \int_{x_1}^{x_2} \int_{y_1}^{y_2} \left(-\frac{\partial F_x}{\partial y}\right) dydx = \iint_R -\frac{\partial F_x}{\partial y} dS$$

where R is the region enclosed by Γ.

We can do similar analysis for the other part of the integral with the F_y function - we choose points O_1 and O_2 now in such a way that the parts of the paths are functions of y. But this time, the direction of integration of Γ that we choose (to satisfy the right hand rule), indicates that Γ_2 part goes from O_1 to O_2 (in direction of increasing y) and Γ_1 part goes in the opposite direction. Therefore

$$\oint_\Gamma F_y dy = \int_{y_2}^{y_1} F_y(x_1(y),y)dy + \int_{y_1}^{y_2} F_y(x_2(y),y) = \int_{y_1}^{y_2} (F_y(x_2,y) - F_y(x_1,y))dy =$$

$$= \int_{y_1}^{y_2}\int_{x_1}^{x_2} \frac{\partial F_y}{\partial x} dxdy = \iint_R \frac{\partial F_y}{\partial x} dS$$

The total integral then is

$$\oint_\Gamma \vec{F}\cdot d\vec{l} = \oint_\Gamma F_x dx + \oint F_y dy = \iint_R -\frac{\partial F_x}{\partial y} dS + \iint_R \frac{\partial F_y}{\partial x} dS = \iint_R \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right) dS$$

$$\oint_\Gamma \vec{F}\cdot d\vec{l} = \iint_R \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right) dS \qquad (52)$$

Notice that if $\vec{F}$ was 3D vector, the integrand would be just the z component of the curl of $\vec{F}$ - Green's theorem is clearly just a special case of the Stoke's theorem.

4.2.1 Calculating Areas

If we choose a vector field in the plane such that

$$\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} = 1$$

Then we have

$$\oint_\Gamma \vec{F}\cdot d\vec{l} = \iint_R \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right) dS = \iint_R dS = S_R$$

we can calculate the area of the region by integrating some vector field around its boundary.
Commont functions for this use are

$$\vec{F} = 0\hat{\imath} + x\hat{\jmath}$$

or

$$\vec{F} = \frac{-y}{2}\hat{\imath} + \frac{x}{2}\hat{\jmath}$$

This principle can be further extended into calculations of other 2D integrals. If the integrand function can be expressed as a difference of the derivatives of the coordinates x and y, the integral can be calculated as a vector integral around the boundary. Consider for example the moment of inertia integral for uniform bodies.

$$I = \iint_R r^2 dm = \rho\iint_R r^2 dS$$

Here, $r^2 = x^2 + y^2$. Can we find $\vec{F}$ such that

$$\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} = x^2 + y^2$$

One such function is for example $\vec{F} = -\frac{y^3}{3}\hat{\imath} + \frac{x^3}{3}\hat{\jmath}$. Then

$$\iint_R r^2 dS = \iint_R \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right) dS = \oint_\Gamma \vec{F}\cdot d\vec{l} = \oint_\Gamma \frac{x^3}{3}dy - \frac{y^3}{3}dx$$

4.2.2 Green's theorem in other coordinate systems

In other coordinate systems, we simply substitute for the correct component of curl in that coordinate system. For example, in a polar planar system, the integrand in the area integral becomes the z component of the curl in the cylindrical coordinate system. Green's theorem then reads

$$\oint_\Gamma \vec{F}\cdot d\vec{l} = \iint_R \frac{1}{r}\left(\frac{\partial(rF_\phi)}{\partial r} - \frac{\partial F_r}{\partial \phi}\right) dS$$

4.2.3 Divergence in 2D

Divergence theorem can be viewed as a consequence of Green's theorem in plane.
The integral along the boundary Γ taking the normal component of a vector is the 2D flux out of the region

$$\oint_\Gamma \vec{F} \cdot d\vec{n}$$

where $d\vec{n}$ has length of the element of the path along the boundary but is perpendicular to it. In order for $d\vec{n}$ to point out of the object when we are integrating according to the right hand rule, we have only one possibility for $d\vec{n}$

$$d\vec{n} = dy\hat{\imath} - dx\hat{\jmath}$$

Then

$$\oint_\Gamma \vec{F} \cdot d\vec{n} = \oint_\Gamma (-F_y)dx + F_x dy = \oint_\Gamma \vec{G} \cdot d\vec{l}$$

where $\vec{G} = (-F_y)\hat{\imath} + F_x\hat{\jmath}$. Due to Green's theorem

$$\oint_\Gamma \vec{G} \cdot d\vec{l} = \iint_R \left(\frac{\partial G_y}{\partial x} - \frac{\partial G_x}{\partial y} \right) dS = \iint_R \left(\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} \right) dS$$

And therefore

$$\oint_\Gamma \vec{F} \cdot d\vec{n} = \iint_R \left(\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} \right) dS \tag{53}$$

4.3 Total vector surface area

Due to Stoke's theorem, any integral along an unclosed surface can be transformed into integral along its boundary. This suggests that for some curve we can choose any surface that spans it and the integral should be the same for every surface.
To illustrate this, consider the total vector area of a surface. This is calculated as

$$\iint_S d\vec{S}$$

If the surface is closed, this is definitely zero, as for every element in one direction, there is an opposite element in the opposite direction. For an unclosed surface, the elements opposing the hole created by the boundary do not have an opposite element - the integral is some finite vector.
Further consider that the boundary (curve) lies in a plane. Then, the only component of each $d\vec{S}$ opposing the boundary that does not get subtracted by some other element opposing the boundary is the element normal to the plane the boundary lies in. Therefore, for any surface that spans this boundary, the total vector area is given by the integral across the plane the boundary lies in.
This is also the reason behind the invariance of the Stoke's theorem for any surface that spans any boundary.

4.4 Divergence Theorem for scalars

If we multiply a scalar function f by a constant vector $\vec{c}$, the divergence theorem for this scalar still applies. The surface integral is

$$\iint_{\partial V} f\vec{c} \cdot d\vec{S} = \iint_{\partial V} c_1 f dS_1 + c_2 f dS_2 + c_3 f dS_3 = c_1 \iint_{\partial V} f dS_1 + c_2 \iint_{\partial V} f dS_2 + c_3 \iint_{\partial V} f dS_3 =$$

$$= c_1\hat{\imath} \cdot \iint_{\partial V} f dS_1\hat{\imath} + c_2\hat{\jmath} \cdot \iint_{\partial V} f dS_2\hat{\jmath} + c_3\hat{k} \cdot \iint_{\partial V} f dS_3\hat{k} = \vec{c} \cdot \iint_{\partial V} f d\vec{S}$$

The volume integral becomes

$$\iiint_V (\nabla \cdot (f\vec{c}))dV = \iiint_V (\nabla f) \cdot \vec{c} dV = \vec{c} \cdot \iint_V \nabla f dV$$

Then, due to divergence theorem

$$\vec{c} \cdot \iint_{\partial V} f d\vec{S} = \vec{c} \cdot \iiint_V \nabla f dV$$

We could have chosen any constant vector c for this calculation and it would still be correct. If we than separately chose unit vectors of some coordinate system, we would have three equations for $i \in \{1, 2, 3\}$

$$\left(\iint_{\partial V} f d\vec{S}\right)_i = \left(\iiint_V \nabla f dV\right)_i$$

which leaves us with

$$\iint_{\partial V} f d\vec{S} = \iiint_V \nabla f dV \tag{54}$$

5 Partial Differential Equations

Partial differential equations describe a behaviour of some field inside a certain medium. They usually do so using partial derivatives of of the field with respect to space and time. In this part, I will only consider using the solution by method of separation of variables. There are other methods, such as Green's functions methods (solution somewhere in certain time is dependant on solution everywhere else in previous time) or eigenvalue methods.

The general recipe to use when using separation of variables is to start with writing the field (start with scalar for simplicity) as a product of two or more components, each depending on one variable considered. For example, a general wave amplitude $u(t, x, y, z)$ is rewritten as

$$u(t, x, y, z) = T(t)X(x)Y(y)Z(z)$$

This solution is then substituted into the equation to obtain total derivatives. Then, the resultant equation is manipulated so that on one side there is only function of one variable and on the other side the expression is function of other variables, but not the one on the origin side.

Then, as each side is function of different variable, the only way that these two sides are equal for each possible value of every variable is when both of these sides are together equal to some constant.

This process is then repeated several times until all variables are separated into differential equations of one variable.

The solution is then constructed by combination of solutions for different allowed values of separation constants. In cases mentioned here, the equations are linear and hence the combinations are simple linear combinations.

The boundary conditions are then used to further narrow the allowed values of separation constants.

Additional constants arise from solving the several differential equations for the separated functions of one variable, and these can be also eliminated by initial conditions or normalisation conditions.

Several different partial differential equations are now mentioned and their homogeneous solutions are derived (without source terms).

5.1 Diffusion equation

Assume there is some concentration function of some molecules in a medium which freely diffuse. We can guess that the molecules tend to diffuse from the more concentrated regions to the less concentrated regions. Therefore, they flow anti-parallel to the concentration gradient. Fick's law, which we will use further on, assumes that there is a direct proportionality of a form

$$\vec{j}(x, y, z, t) = -D\nabla c(x, y, z, t) \tag{55}$$

where D is some diffusion constant, real number greater than zero, and $\vec{j}(x, y, z, t)$ is the flux of the particles through infinitesimal area at (x, y, z, t) normal to $\vec{j}$ per unit time. The flux of the particle out of some small volume is then

$$d\Phi = (j_x(x, y, z, t) - j_x(x + dx, y, z, t))dydz + (j_y(x, y, z, t) - j_y(x, y + dy, z, t))dxdz+$$

$$+(j_z(x, y, z, t) - j_z(x, y, z + dz, t))dxdy = -\nabla \cdot \vec{j}dV$$

The change in number of particles per some time dt is then

$$dN = d\Phi dt = -\nabla \cdot \vec{j}dVdt$$

Here, $\frac{dN}{dV} = dc$ is the change of concentration at (x, y, z, t) per time dt. Therefore

$$\frac{\partial c}{\partial t} = -\nabla \cdot \vec{j}$$

This is a form of a continuity equation. Applying the Fick's law now gives us

$$\frac{\partial c}{\partial t} = -\nabla \cdot (-D\nabla c) = D\nabla \cdot (\nabla c) = D\nabla^2 c$$

$$\frac{\partial c}{\partial t} = D\nabla^2 c \tag{56}$$

Or for a special case in 1D

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2}$$

In case that there is some source of molecules per unit volume $s(x, y, z, t) = \frac{dN'}{dV dt}$, the continuity equation does not apply. Instead, the balance of molecules in small volume dV per time dt is

$$dN = d\Phi dt = -\nabla \cdot \vec{j} dV dt + s dV dt$$

And the equation becomes

$$\frac{\partial c}{\partial t} - D\nabla^2 c = s(x, y, z, t)$$

We now consider no source terms. The separation is

$$c = T(t)X(x)Y(y)Z(z)$$

Then

$$XYZ\frac{dT}{dt} = DT(t)\left(YZ\frac{d^2X}{dx^2} + XZ\frac{d^2Y}{dy^2} + XY\frac{d^2Z}{dz^2}\right)$$

$$\frac{1}{T}\frac{dT}{dt} = D\left(\frac{1}{X}\frac{d^2X}{dx^2} + \frac{1}{Y}\frac{d^2Y}{dy^2} + \frac{1}{Z}\frac{d^2Z}{dz^2}\right)$$

Left hand side must be equal to a constant. Lets label this constant K. For $K = 0$ we have the stationary state. For positive K, we have

$$\frac{dT}{dt} = KT$$

$$T = Ae^{Kt}$$

which leads to diverging concentrations at very late times - this is clearly not a physical solution. Therefore, we can relabel $K = -K'$, as K has to be negative. We can even just drop the prime, as no other K' than positive is possible, hence

$$\frac{dT}{dt} = -KT$$

$$T = Ae^{-Kt}$$

However, this is possible for any real negative number K. We can now continue the separation

$$K = D\left(\frac{1}{X}\frac{d^2X}{dx^2} + \frac{1}{Y}\frac{d^2Y}{dy^2} + \frac{1}{Z}\frac{d^2Z}{dz^2}\right)$$

$$K - \frac{D}{X}\frac{d^2X}{dx^2} = D\left(\frac{d^2Y}{dy^2} + \frac{d^2Z}{dz^2}\right)$$

Again, this has to be equal to a constant and so on. For now, I will further explore one dimensional case. There $Y = Z = 1$. So we have

$$\frac{d^2X}{dx^2} = \frac{K}{D}X$$

But, only possible solution occured for $K = -K$ (negative K). Therefore

$$\frac{d^2X}{dx^2} = -\frac{K}{D}X$$

$$X = Be^{i\sqrt{\frac{K}{D}}x} + Ce^{-i\sqrt{\frac{K}{D}}x}$$

One last class of solutions we have to discuss is solutions for case when $K = 0$, which are the solutions for the stationary state. In this case, the differential equation for X becomes

$$\frac{d^2X}{dx^2} = 0$$

$$X = A_0x + B_0$$

wher A_0 and B_0 are some constants.
Since the diffusion equation is linear, the full solution is constructed by superposition of all possible solutions.
The full solution is then

$$c(x,t) = A_0x + B_0 + \sum_K Ae^{-Kt}\left(Be^{i\sqrt{\frac{K}{D}}x} + Ce^{-i\sqrt{\frac{K}{D}}x}\right)$$

Relabeling $\sqrt{\frac{K}{D}} = \lambda$

$$c(x,t) = A_0x + B_0 + \sum_\lambda Ae^{-\lambda^2Dt}\left(Be^{i\lambda x} + Ce^{-i\lambda x}\right) \tag{57}$$

5.1.1 Boundary conditions

The tipical boundary conditions in 1D are setting either the concentration or the flux of the particles at the ends of some finite tube to 0.
Lets start with the case when the flux $j(t,x)$ on both ends of some tube of length L. We can place one end into the beginning of our coordinate system. Then, we require $j(t,0) = j(t,L) = 0$ (we drop the vector nature of j, as we are in one dimension).
j is given by Fick's law. Substituting from the general solution (57)

$$j = -DA_0 - D\frac{\partial c}{\partial x} = -DA_0 - D\frac{\partial}{\partial x}\sum_\lambda e^{-\lambda^2Dt}\left(Ae^{i\lambda x} + Be^{-i\lambda x}\right)$$

where I absorbed the constant A into constants B and C and relabeled the constants.
Now, applying the derivative (assuming that the sum converges)

$$j = -DA_0 - D\sum_\lambda e^{-\lambda^2Dt}\left(Ai\lambda e^{i\lambda x} - Bi\lambda e^{-i\lambda x}\right)$$

Our boundary conditions are

$$j(t,0) = -DA_0 - D\sum_\lambda e^{-\lambda^2Dt}(Ai\lambda - Bi\lambda) = 0$$

In order for this to be true at any time t, two conditions must both apply. First is $A_0 = 0$ and second is $A = B$. The condition on the second boundary is

$$j(t,L) = -D\sum_\lambda e^{-\lambda^2Dt}(Ai\lambda e^{i\lambda L} - Ai\lambda e^{-i\lambda L}) = 0$$

$$j(t,L) = -D\sum_\lambda e^{-\lambda^2Dt}Ai\lambda\left(e^{i\lambda L} - e^{-i\lambda L}\right) = 0$$

Since A can generaly depend on λ, we cannot take it out of the sum. However, we can rewritte the exponentials as a sine

$$j(t,L) = -D\sum_\lambda e^{-\lambda^2Dt}Ai\lambda 2i\sin(\lambda L) = 2D\sum_\lambda A\lambda e^{-\lambda^2Dt}\sin(\lambda L)$$

In order for this to be equal to zero for all times, $\sin(\lambda L) = 0$ (which also covers the case $\lambda = 0$) Therefore

$$\lambda L = n\pi$$

where n is an integer. And therefore

$$\lambda = \frac{n\pi}{L}$$

Therefore, the full solution becomes (as $A_0 = 0$)

$$c(t,x) = B_0 + \sum_{n=-\infty}^{\infty} e^{-\left(\frac{n\pi}{L}\right)^2 Dt} A \left(e^{i\frac{n\pi}{L}x} + e^{-i\frac{n\pi}{L}x}\right) = B_0 + \sum_{n=-\infty}^{\infty} 2Ae^{-\left(\frac{n\pi}{L}\right)^2 Dt} \cos\left(\frac{n\pi}{L}x\right)$$

Since cos is an even function, for some n and the opposite $n' = -n$

$$e^{-\left(\frac{n\pi}{L}\right)^2 Dt} \cos\left(\frac{n\pi}{L}x\right) = e^{-\left(\frac{n'\pi}{L}\right)^2 Dt} \cos\left(\frac{n'\pi}{L}x\right)$$

Hence, these two terms in the sum can be substituted by

$$2A'_n e^{-\left(\frac{n'\pi}{L}\right)^2 Dt} \cos\left(\frac{n'\pi}{L}x\right) + 2A_n e^{-\left(\frac{n\pi}{L}\right)^2 Dt} \cos\left(\frac{n\pi}{L}x\right) = (2A'_n + 2A_n) e^{-\left(\frac{n\pi}{L}\right)^2 Dt} \cos\left(\frac{n\pi}{L}x\right)$$

where we can further relabel $2A'_n + 2A_n$ as a single constant A_n. For $n = 0$, we don't have the second opposite term in the sum, but both cosine and the decaying exponential are equal to one, and this term can be simply added to the constant term B_0 outside of the sum. Therefore, the solution can be represented by a more familiar form.

$$c(t,x) = B_0 + \sum_{n=1}^{\infty} A_n e^{-\left(\frac{n\pi}{L}\right)^2 Dt} \cos\left(\frac{n\pi}{L}x\right) \tag{58}$$

From the initial conditions, we can determine the values of constants A_n and B_0. At time $t = 0$

$$c(0,x) = B_0 + \sum_{n=1}^{\infty} A_n \cos\left(\frac{n\pi}{L}x\right)$$

But this is a Fourier serie of some initial distribution of concentration. Since this is cosine series, the periodic extension of the function is even, and we can find A_n as

$$A_n = \frac{2}{L}\int_0^L c(0,x) \cos\left(\frac{n\pi}{L}x\right) dx$$

And B_0 as

$$B_0 = \frac{1}{L}\int_0^L c(0,x) dx$$

Now, I consider a more general case when the flux at the boundaries is constant but otherwise not determined (can be both positive, meaning matter travelling in positive x direction, and negative, for left travelling matter). Let the flux at $x = 0$ be j_0 and at $x = L$ the flux is j_L. Fick's law gives us

$$j = -DA_0 - D\sum_\lambda e^{-\lambda^2 Dt}(Ai\lambda - Bi\lambda) = j_0$$

In order for j_0 to be constant, we have again $A = B$. But, we furthemore have

$$j_0 = -DA_0$$

The second boundary condition is

$$j_L = -DA_0 - D\sum_\lambda e^{-\lambda^2 Dt} Ai\lambda \left(e^{i\lambda L} - e^{-i\lambda L}\right) = -DA_0 + D\sum_\lambda e^{-\lambda^2 Dt} 2A\sin(\lambda L)$$

j_L can be constant only when $\sin(\lambda L) = 0$, and hence we find that we require $j_L = j_0$ - there is some total net flux accross the pipe. The analysis for initial conditions is very similar as above.

5.2 Wave Equation

In 1D, the wave equation is (as many times observed and derived in other modules)

$$c^2 \frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 u}{\partial t^2}$$

where c is some real constant.

The separation of variables is as follows, and is again very similar to diffusion equation

$$u = X(x)T(t)$$

$$c^2 T \frac{d^2X}{dx^2} = X \frac{d^2T}{dt^2}$$

$$\frac{1}{X}\frac{d^2X}{dx^2} = \frac{1}{c^2 T}\frac{d^2T}{dt^2}$$

Again, both sides are separately functions of different variables, and hence must be together equal to a constant. Lets first assume that the constant is negative, and therefore denote it as $-k^2$. Then

$$\frac{1}{X}\frac{d^2X}{dx^2} = -k^2$$

$$\frac{d^2X}{dx^2} = -k^2 X$$

$$X = A\cos(kx) + B\sin(kx)$$

$$\frac{1}{c^2 T}\frac{d^2T}{dt^2} = -k^2$$

$$\frac{d^2T}{dt^2} = -k^2 c^2 T$$

$$T = C\cos(kct) + D\sin(kct)$$

$$u = XT = (A\cos(kx) + B\sin(kx))(C\cos(ckt) + D\sin(ckt))$$

This is the most common wave solution. The full solution then occurs for superposition of all possible k.
The other family of solutions occurs for separation constant being positive, k^2. Then

$$\frac{d^2X}{dx} = k^2 X$$

$$X = Ae^{kx} + Be^{-kx}$$

$$\frac{d^2T}{dt^2} = c^2 k^2 T$$

$$T = Ce^{ckt} + De^{-ckt}$$

$$u = XT = (Ae^{kx} + Be^{-kx})(Ce^{ckt} + De^{-ckt})$$

If the solution is bounded in space or time, these solutions can still be physical solutions - they represent the so called evanescent waves, which are for example present in description of wavefunction tunneling in Schrödinger equation.
The final family of solutions is for separation constant 0. Then

$$\frac{d^2X}{dx^2} = 0$$

$$X = A_0 x + B_0$$

$$\frac{d^2T}{dt^2} = 0$$

$$T = C_0 t + D_0$$

$$u = (A_0 x + B_0)(C_0 t + D_0)$$

This solution only represents the 1D wave medium (string for example) moving in the space (T component) or somehow tilted in the space (X component) and do not add any new physics. We usually only have to consider first two types of solutions. Specially, for the first solution, we can again use Fourier analysis to determine the displacement u depending on boundary and initial conditions.
There are many other solutions that can be derived using separation of variables (including the Schrödinger equation), but they are included in other modules.

6 Fourier Transforms

6.1 Complex Fourier Series

In analysis of partial differential equations, we many times used the Fourier series representation of functions on a finite length interval L

$$f(x) = A_0 + \sum_{n=1}^{\infty} A_n \cos\left(\frac{n\pi}{L}x\right) + B_n \sin\left(\frac{n\pi}{L}x\right) \tag{59}$$

Consider rewritting the cosines and sines in terms of complex exponentials. Then

$$f(x) = A_0 + \sum_{n=1}^{\infty} \frac{1}{2} A_n \left(e^{i\frac{n\pi}{L}x} + e^{-i\frac{n\pi}{L}x}\right) + \frac{1}{2i} B_n \left(e^{i\frac{n\pi}{L}x} - e^{i\frac{n\pi}{L}x}\right)$$

Lets name $\frac{n\pi}{L} = k_n$. The representation becomes

$$f(x) = A_0 + \sum_{n=1}^{\infty} \frac{A_n}{2}\left(e^{ik_nx} + e^{-ik_nx}\right) + \frac{B_n}{2i}\left(e^{ik_nx} - e^{-ik_nx}\right) = A_0 + \sum_{n=1}^{\infty}\left(\frac{A_n}{2} + \frac{B_n}{2i}\right)e^{ik_nx} + \sum_{n=1}^{\infty}\left(\frac{A_n}{2} - \frac{B_n}{2i}\right)e^{-ik_nx}$$

We can now notice that the second sum can be renumbered by $n \to -n$. This causes $k_n \to -k_n$, and the representation becomes

$$f(x) = A_0 + \sum_{n=1}^{\infty}\left(\frac{A_n}{2} + \frac{B_n}{2i}\right)e^{ik_nx} + \sum_{n=-\infty}^{-1}\left(\frac{A_{-n}}{2} - \frac{B_{-n}}{2i}\right)e^{ik_nx}$$

Since A_n is generally different for different positive n and B_n as well, we can define new constant dependent on n, $C_n = \frac{A_n}{2} + \frac{B_n}{2i}$ for positive n. For negative n, we can define this constant as $C_n = \frac{A_{-n}}{2} - \frac{B_{-n}}{2i}$ (as for positive n, A_n and B_n are both defined). Then we have

$$f(x) = A_0 + \sum_{n=1}^{\infty} C_n e^{ik_nx} + \sum_{n=-\infty}^{-1} C_n e^{ik_nx}$$

Finally, we note that we can rewrite A_0 as

$$A_0 = A_0 e^0 = A_0 e^{ik_0x}$$

And therefore if we define $C_0 = A_0$, we have

$$f(x) = \sum_{n=-\infty}^{-1} C_n e^{ik_nx} + C_0 e^{ik_0x} + \sum_{n=1}^{\infty} C_n e^{ik_nx} = \sum_{-\infty}^{\infty} C_n e^{ik_nx}$$

This representation of f is called the complex Fourier series. The coefficients C_n can be generally complex, although C_0 is always real (derived from A_0).
The coefficients are recovered from the function as follows

$$\int_{-L}^{L} f(x)e^{-ik_mx}dx = \int_{-L}^{L} e^{-ik_mx}\sum_{-\infty}^{\infty} C_n e^{ik_nx}dx = \sum_{-\infty}^{\infty} C_n \int_{-L}^{L} e^{i(k_n-k_m)x}dx$$

Remembering the definition of k_n

$$\int_{-L}^{L} e^{i(k_n-k_m)x}dx = \int_{-L}^{L} e^{i\frac{\pi}{L}(n-m)x}dx$$

If $n \neq m$

$$\int_{-L}^{L} e^{i\frac{\pi}{L}(n-m)x}dx = \frac{1}{i\frac{\pi}{L}(n-m)}\left(e^{i\frac{\pi}{L}(n-m)L} - e^{-i\frac{\pi}{L}(n-m)L}\right) = \frac{2L}{\pi(n-m)}\sin\left(\pi(n-m)\right)$$

And since $n - m$ is an integer, this is always zero. In the case when $n = m$

$$\int_{-L}^{L} e^{i0x}dx = 2L$$

Hence, for general n and m

$$\int_{-L}^{L} e^{i\frac{\pi}{L}(n-m)x}dx = \int_{-L}^{L} e^{i(k_n-k_m)x}dx = 2L\delta_{nm}$$

Where δ is the Kronecker delta. Hence

$$\int_{-L}^{L} f(x)e^{-ik_m x}dx = \sum_{-\infty}^{\infty} 2LC_n\delta_{nm} = 2LC_m$$

$$C_n = \frac{1}{2L}\int_{-L}^{L} f(x)e^{-ik_n x}dx$$

The reverse transforms to A_n and B_n can be obtained from manipulating C_n and C_{-n} for some positive n.

$$A_n = C_n + C_{-n}$$

$$B_n = i(C_n - C_{-n})$$

Exception is A_0, which is defined directly

$$A_0 = C_0$$

These also give us the conditions on complex C_n in order to be valid Fourier coefficients that give real A_n and B_n. From relation for A_n, $\mathrm{Im}(C_n) = -\mathrm{Im}(C_{-n})$. From the second reverse transform

$$B_n = i(\mathrm{Re}(C_n) + i\mathrm{Im}(C_n) - \mathrm{Re}(C_{-n}) - i\mathrm{Im}(C_{-n})) = i(\mathrm{Re}(C_n) - \mathrm{Re}(C_{-n})) + \mathrm{Im}(C_{-n}) - \mathrm{Im}(C_n)$$

In order for this to be real, $\mathrm{Re}(C_n) = \mathrm{Re}(C_{-n})$. Hence, the overall condition can be summarized as

$$C_n = \bar{C}_{-n}$$

, the coefficients with opposite indices have to be complex conjugates.

6.2 Fourier Transforms

Now, consider that we increase the length of the interval L to infinity. The coefficient C_n will clearly go to zero, but that might not be true for expression $2LC_n$. This expression goes to

$$\lim_{L\to\infty} 2LC_n = \int_{-\infty}^{\infty} f(x)e^{-ik_n x}dx$$

Another thing to notice is that the spacing between k_n and k_{n+1} decreases

$$k_{n+1} - k_n = \Delta k_n = \frac{\pi}{L}(n+1-n) = \frac{\pi}{L}$$

$$\lim_{L\to\infty}(k_{n+1} - k_n) = 0$$

Hence, we can view the k_n as continuous variable and drop the indexing. We than have

$$\lim_{L\to\infty} 2LC_n = \int_{-\infty}^{\infty} f(x)e^{-ikx}dx$$

The complex Fourier representation of $f(x)$ can be rewritten as

$$f(x) = \lim_{L\to\infty}\sum_{-\infty}^{\infty} C_n e^{ik_n x} = \lim_{L\to\infty}\sum_{-\infty}^{\infty}(2LC_n)\frac{\pi}{2L\pi}e^{ik_n x} = \lim_{L\to\infty}\frac{1}{2\pi}\sum_{-\infty}^{\infty}(2LC_n)\Delta k_n e^{ik_n x}$$

Again, dropping indexing for k_n, the sum converges to following integral

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}(2LC_n)e^{ikx}dk$$

We call the expression $(2LC_n) = \tilde{f}(k)$ the Fourier transform of f. It depends on k because C_n depends on n, but there is a one-to-one correspondence between n and k. Then we can rewrite the derived relations as

$$\tilde{f}(k) = \mathscr{F}\{f(x)\} = \int_{-\infty}^{\infty} f(x)e^{-ikx}dx \tag{60}$$

$$f(x) = \mathscr{F}^{-1}\{\tilde{f}(k)\} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{f}(k)e^{ikx}dk \tag{61}$$

6.3 Dirac Delta

When taking consequently the forward and reverse Fourier transform of a function f, we should arrive at the same function. Doing this explicitly

$$\mathscr{F}^{-1}\{\mathscr{F}\{f\}\} = f$$

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikx}\int_{-\infty}^{\infty} e^{-ikx'} f(x')dx'dk$$

(note the use of dummy variable for inner integration). Reversing the order of integration

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{ik(x-x')} f(x')dkdx' = \int_{-\infty}^{\infty} f(x')\left(\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ik(x'-x)}dk\right)dx'$$

The object in the brackets only depends on $(x' - x)$ term and is called the integral representation of Dirac delta function, denoted as $\delta(x' - x)$.
The Dirac delta function therefore has property (by the above definition from Fourier transform)

$$\int_{-\infty}^{\infty} f(x)\delta(x-s)dx = f(s) \tag{62}$$

where I switched x' to x and x to s. From this property, we can imediately determine the second property of Dirac delta function. If $f(x)$ is a constant function $f(x) = 1$

$$\int_{-\infty}^{\infty} f(x)\delta(x'-x)dx = \int_{-\infty}^{\infty}\delta(x'-x)dx = 1$$

Hence the Dirac delta is normalised to one. Other properties of Dirac delta are defined from the approximation of the integral representation. Assume that we do not integrate over the whole range from $infty$ to ∞, but just from some big $-L$ to some L. Dirac delta function is represented in this way in Fig. 14

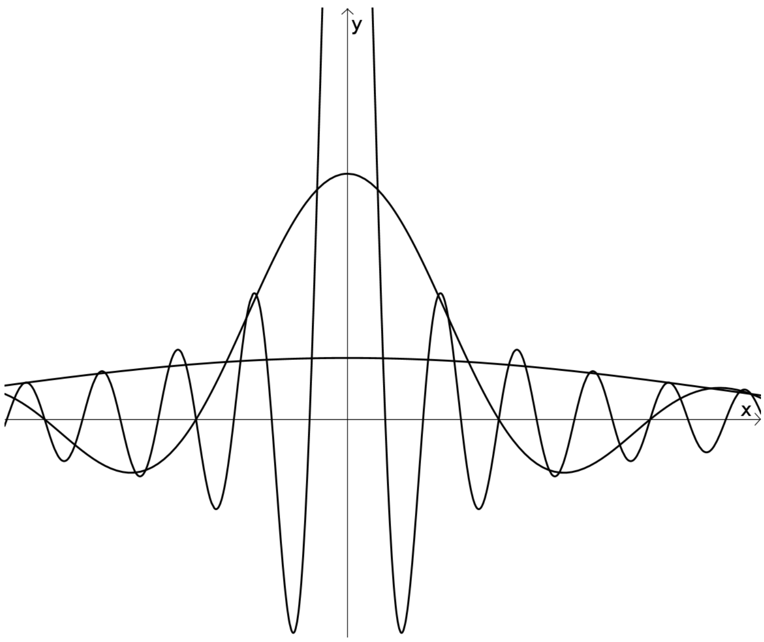

Figure 14: Dirac delta approximated as $\delta_L = \int_{-L}^{L} e^{-ik(x'-x)}dk$. The central peak gradualy increases in height for increasing L. For this example, $x = 0$ - the Dirac delta is centered on zero. Other parts of the function become insignificant when compared to the central peak.

There are two important properties to derive from this approximation. First is that Dirac delta is zero everywhere except at zero, i. e.

$$\forall x \neq s : \delta(x-s) = 0$$

Second property is that Dirac delta function is symmetric around the zero in a sense that

$$\int_{-\infty}^{s}\delta(x-s)dx = \int_{s}^{\infty}\delta(x-s)dx$$

Together with the normalisation of Dirac delta, this leads to

$$\int_{-\infty}^{s} \delta(x-s)dx = \int_{s}^{\infty} \delta(x-s)dx = \frac{1}{2}$$

Obvious consequence of the first property is also that integral of Dirac delta over any interval I which does not include the point where the argument of Dirac delta is zero goes to zero

$$\int_{I, s \notin I} \delta(x-s)dx = 0$$

Besides being a part of the Fourier transforms theory, Dirac delta function is also a very useful function to model any point-like sources/particles etc. For example, a point charge can be modeled as charge density ρ which takes form

$$\rho = Q\delta^3(\vec{r} - \vec{r}')$$

where Q is the charge and $\delta^3(\vec{r}-\vec{r}')$ is the 3D equivalent of Dirac delta and $\vec{r}'$ is the position of the charge. In cartesian coordinates, the 3D Dirac delta is simply

$$\delta^3(\vec{r} - \vec{r}') = \delta(x - x')\delta(y - y')\delta(z - z')$$

In other coordinate system, the normalisation condition must apply, which means that we must always have

$$\iiint_{V, \vec{r} \in V} \delta^3(\vec{r} - \vec{r}')dV = 1$$

Therefore, in cylindrical coordinates, we have

$$\delta^3(\vec{r} - \vec{r}') = \frac{1}{r}\delta(r - r')\delta(\phi - \phi')\delta(z - z')$$

And in spherical polar coordinates

$$\delta^3(\vec{r} - \vec{r}') = \frac{1}{r^2 \sin(\theta)}\delta(r - r')\delta(\theta - \theta')\delta(\phi - \phi')$$

Similarly, Dirac delta for other number of dimensions or for tilted planes/lines in different coordinate systems can be found, but this is not further explored here.

6.4 Some explicit Fourier transforms

Now I present some Fourier transforms of explicit functions $f(x)$, and also reverse Fourier transforms of the analogous functions $\tilde{f}(k)$

6.4.1 Dirac Delta

Let $f(x) = \delta(x - x')$. Then

$$\mathscr{F}\{\delta(x - x')\} = \int_{-\infty}^{\infty} \delta(x - x')e^{-ikx}\delta(x - x')dx = e^{-ikx'}$$

Let $\tilde{f}(k) = \delta(k - k')$. Then

$$\mathscr{F}^{-1}\{\delta(k - k')\} = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikx}\delta(k - k')dk = \frac{e^{ik'x}}{2\pi}$$

For special case when $x' = 0$ (or $k' = 0$), $\mathscr{F}\{\delta(x)\} = 1$ (or $\mathscr{F}^{-1}\{\delta(k)\} = \frac{1}{2\pi}$). Therefore, perfect, point-like localization for a function means its Fourier transform (or reverse Fourier transform) is perfectly delocalized (constant function). This is further explored in quantum mechanics.

6.4.2 Cosine and sine

Cosine and sine transform as follows

$$\mathscr{F}\{\cos(ax)\} = \int_{-\infty}^{\infty} e^{-ikx}\cos(ax)dx = \int_{-\infty}^{\infty} e^{-ikx}\frac{e^{iax}+e^{-iax}}{2}dx =$$

$$= \frac{1}{2}\left(\int_{-\infty}^{\infty} e^{-i(k-a)x}dx + \int_{-\infty}^{\infty} e^{-i(k+a)x}dx\right) = \pi(\delta(k-a)+\delta(k+a))$$

$$\mathscr{F}\{\sin(ax)\} = \int_{-\infty}^{\infty} e^{-ikx}\sin(ax)dx = \frac{1}{2i}\left(\int_{-\infty}^{\infty} e^{-i(k-a)x}dx - \int_{-\infty}^{\infty} e^{-i(k+a)x}dx\right) =$$

$$= \frac{\pi}{i}\left(\delta(k-a)-\delta(k+a)\right) = \pi i\left(\delta(k+a)-\delta(k-a)\right)$$

6.4.3 Exponential

Pure exponential is not transformable, because at some infinity, this function diverges and the integral diverges as well. However, function of type $e^{-|ax|}$ does not diverge. The transform is (assuming positive a)

$$\mathscr{F}\{e^{-|ax|}\} = \int_{-\infty}^{\infty} e^{-ikx}e^{-|ax|}dx = \int_{-\infty}^{0} e^{-ikx}e^{-a(-x)}dx + \int_{0}^{\infty} e^{-ikx}e^{-a(x)}dx =$$

$$= \int_{-\infty}^{0} e^{-x(ik-a)}dx + \int_{0}^{\infty} e^{-x(ik+a)}dx = \left[\frac{e^{-x(ik-a)}}{-(ik-a)}\right]_{-\infty}^{0} + \left[\frac{e^{-x(ik+a)}}{-(ik+a)}\right]_{0}^{\infty}$$

At the infinity, the real part always dominates the imaginary part, as the imaginary part is oscillatory and finite while the real part is infinitesimal. Therefore

$$\left[\frac{e^{-x(ik-a)}}{-(ik-a)}\right]_{-\infty}^{0} = \frac{1}{a-ik}$$

$$\left[\frac{e^{-x(ik+a)}}{-(ik+a)}\right]_{0}^{\infty} = \frac{1}{a+ik}$$

Therefore

$$\mathscr{F}\{e^{-|ax|}\} = \frac{1}{a-ik}+\frac{1}{a+ik} = \frac{a+ik+a-ik}{a^2+k^2} = \frac{2a}{a^2+k^2}$$

6.4.4 Gaussian

Gaussian is a function of type $f(x) = e^{-\frac{x^2}{a^2}}$. It tranforms as

$$\mathscr{F}\{e^{-\frac{x^2}{a^2}}\} = \int_{-\infty}^{\infty} e^{-ikx}e^{-\frac{x^2}{a^2}}dx = \int_{-\infty}^{\infty} e^{-\left(\frac{x^2}{a^2}+ikx\right)}dx = \int_{-\infty}^{\infty} e^{-\left(\frac{x^2}{a^2}+ikx+\left(\frac{iak}{2}\right)^2-\left(\frac{iak}{2}\right)^2\right)}dx =$$

$$= \int_{-\infty}^{\infty} e^{-\left(\frac{x}{a}+\frac{iak}{2}\right)^2+\left(\frac{iak}{2}\right)^2}dx = e^{-\frac{a^2k^2}{4}}\int_{-\infty}^{\infty} e^{-\left(\frac{x}{a}+\frac{iak}{2}\right)^2}dx$$

By substitution $z = \frac{x}{a}+\frac{iak}{2}$

$$\mathscr{F}\{e^{-\frac{x^2}{a^2}}\} = ae^{-\frac{k^2a^2}{4}}\int_{-\infty+\frac{iak}{2}}^{\infty+\frac{iak}{2}} e^{-z^2}dz$$

Because the offset is constant for the limits, the integration contour is just a line parallel to the real axis, adn therefore the we can transform the integral into real integral. Then

$$\mathscr{F}\{e^{-\frac{x^2}{a^2}}\} = ae^{-\frac{k^2a^2}{4}}\int_{-\infty}^{\infty} e^{-z^2}dz = a\sqrt{\pi}e^{-\frac{k^2a^2}{4}}$$

Therefore, the transform of a Gaussian is another Gaussian, multiplied by some constant and with inverse width $\frac{1}{a}$. This is the mathematical principle behind the Heisenberg Uncertainty Principle.

6.5 General Function Properties

Several (and perhaps more important) properties of a Fourier transform of a function can be derived independently of the specific function form. I now give few examples of those

6.5.1 Linearity

Assume we have a function $h(x) = af(x) + bg(x)$, where $f(x)$ and $g(x)$ are some Fourier-transformable functions and a and b are some constants. Then, the Fourier transform of h is

$$\tilde{h}(k) = \int_{-\infty}^{\infty} (af(x) + bg(x))e^{-ikx}dx = a\int_{-\infty}^{\infty} f(x)e^{-ikx}dx + b\int_{-\infty}^{\infty} g(x)e^{-ikx}dx = a\tilde{f}(k) + b\tilde{g}(k) \quad (63)$$

Therefore, the Fourier transform is a linear operation. This is property is inhereted from the linearity if the integration.

6.5.2 Translation

Assume we have a function $h(x) = f(x - x')$ (function translated to the right with respect to original function f). The Fourier transform is (using substitution $z = x - x'$

$$\tilde{h}(k) = \int_{-\infty}^{\infty} e^{-ikx}f(x - x')dx = \int_{-\infty}^{\infty} e^{-ik(z+x')}f(z)dz = e^{-ikx'}\int_{-\infty}^{\infty} e^{-ikz}f(z)dz = e^{-ikx'}\tilde{f}(k) \quad (64)$$

Therefore translation is transformed as multiplication by factor $e^{-ikx'}$, which is equivalent to rotation of a phasor by some phase angle.

6.5.3 Scaling

Assume we have a function $h(x) = f(ax)$. The Fourier transform is $(z = ax)$

$$\tilde{h}(k) = \int_{-\infty}^{\infty} e^{-ikx}f(ax)dx = \frac{1}{a}\int_{-\infty}^{\infty} e^{-ik\frac{z}{a}}f(z)dz = \frac{1}{a}\tilde{f}\left(\frac{k}{a}\right) \quad (65)$$

The principle here is similar to the transform of a Gaussian - wider function in real space $(f(x))$ is narrower in phase space $(\tilde{f}(k))$.

6.5.4 Derivatives

Let $h(x) = \frac{d^n f}{dx^n}$. The Fourier transform is

$$\tilde{h}(k)\int_{-\infty}^{\infty} e^{-ikx}\frac{d^n f}{dx^n}dx = \left[e^{-ikx}\frac{d^{n-1}f}{dx^{n-1}}\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty}(-ik)e^{-ikx}\frac{d^{n-1}f}{dx^{n-1}}$$

If we require $\frac{d^{n-1}f}{dx^{n-1}}|_{x\to\pm\infty} = 0$, the above becomes

$$\tilde{h}(k) = (ik)^1\int_{-\infty}^{\infty} e^{-ikx}\frac{d^{n-1}f}{dx^{n-1}}dx$$

We can continue this chain of integration by parts and requiring all derivatives of f to go to zero at infinities and we then arrive at

$$\tilde{h}(k) = (ik)^n\int_{-\infty}^{\infty} e^{-ikx}\frac{d^{n-n}f}{dx^{n-n}} = (ik)^n\int_{-\infty}^{\infty} e^{-ikx}f(x)dx = (ik)^n\tilde{f}(k) \quad (66)$$

6.5.5 Solving Differential Equations

We can use the properties of Fourier trasforms to solve some differential equations. For example, the Fourier transform of the diffusion equation (with respect to x) becomes

$$\mathscr{F}\left\{\frac{\partial u}{\partial t}\right\} = \mathscr{F}\left\{D\frac{\partial^2 u}{\partial x^2}\right\}$$

Let $\tilde{u} = \mathscr{F}\{u\}$. Then, because we can take the derivative with respect to t outside of the integral, the equation becomes

$$\frac{\partial \tilde{u}}{\partial t} = D(ik)^2 \tilde{u}(k,t)$$

$$\frac{\partial \tilde{u}}{\partial t} = -Dk^2 \tilde{u}(k,t)$$

$$\tilde{u}(k,t) = Ae^{-Dk^2 t} = Ae^{-\frac{k^2(\sqrt{4Dt})^2}{4}} = \frac{A}{\sqrt{4Dt\pi}}\sqrt{4Dt\pi}e^{-\frac{k^2(\sqrt{4Dt})^2}{4}}$$

Comparing this to the Fourier transform of a Gaussian, we can see that

$$u(x,t) = \frac{A}{\sqrt{4Dt\pi}}e^{-\frac{x^2}{4Dt}}$$

This is a specific solution to the diffusion equation, which corresponds to point source in time and space. This is the Green's function of the diffusion equation. This means that a solution at some later time t' can be gained from the solution in time $t = 0$ according to

$$u(x,t') = \int_{all x} \frac{A}{\sqrt{4Dt\pi}}e^{-\frac{x^2}{4Dt}} u(x,0)dx$$

6.5.6 Parseval's theorem

Consider calculating integral $\int_{-\infty}^{\infty} |f(x)|^2 dx$. This integral can be rewritten as

$$\int_{-\infty}^{\infty} f(x)\bar{f}(x)dx$$

We now express f using the reverse Fourier transform of $\tilde{f}$:

$$\int_{-\infty}^{\infty} dx \left(\frac{1}{2\pi}\int_{-\infty}^{\infty} dk e^{ikx}\tilde{f}(k)\right)\overline{\left(\frac{1}{2\pi}\int_{-\infty}^{\infty} dk' e^{ik'x}\tilde{f}(k')\right)}$$

Taking the complex conjugate inside the integral and staking all functions into the inner most integral, we transform the expression to

$$\int_{-\infty}^{\infty} dx \left(\int_{-\infty}^{\infty} dk \left(\int_{-\infty}^{\infty} dk' \frac{1}{4\pi^2} e^{i(k-k')x}\tilde{f}(k)\bar{\tilde{f}}(k')\right)\right)$$

Reversing the order of integration so that we do integration with respect to x first, the expression becomes

$$\int_{-\infty}^{\infty} dk \left(\int_{-\infty}^{\infty} dk' \left(\int_{-\infty}^{\infty} dx \frac{1}{4\pi^2} e^{i(k-k')x}\tilde{f}(k)\bar{\tilde{f}}(k')\right)\right) =$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk \left(\int_{-\infty}^{\infty} dk' \tilde{f}(k)\bar{\tilde{f}}(k')\left(\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i(k'-k)x}dx\right)\right)$$

We can recognize the innermost expression as Dirac delta integral representation, and therefore we can continue with transforming the expression to

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} dk \left(\int_{-\infty}^{\infty} dk' \tilde{f}(k)\bar{\tilde{f}}(k')\delta(k'-k)\right) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dk \tilde{f}(k)\bar{\tilde{f}}(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} |\tilde{f}(k)|^2 dk$$

Therefore, we have

$$\int_{-\infty}^{\infty} |f(x)|^2 dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} |\tilde{f}(k)|^2 dk \tag{67}$$

which is called the Parseval's theorem.

6.5.7 Convolutions

The convolution of two functions $f(x)$ and $g(x)$ is a function $h(x)$ defined as

$$h(x) = f * g(x) = \int_{-\infty}^{\infty} f(y)g(x-y)dy$$

This seems like a rather arbitrary definition, but this concept has several uses. For example, the translated function can be expressed as convolution with Dirac delta function

$$f(x-x') = \int_{-\infty}^{\infty} f(y)\delta((x-x')-y)dy$$

Again, convolutions are linear in both arguments, i. e.

$$(af+bg)*h(x) = \int_{-\infty}^{\infty} (af(y)+bg(y))h(x-y)dy = a\int_{-\infty}^{\infty} f(y)h(x-y)dy + b\int_{-\infty}^{\infty} g(y)h(x-y)dy = a(f*h)(x)+b(g*h)(x)$$

and similarly

$$f*(ag+bh)(x) = a(f*g)(x) + b(f*h)(x)$$

where a and b are constants and h is a function. This means that we can represent some repeating pattern by sum of convolutions of Dirac delta functions with the unit cell of the pattern. This will be covered explicitly in diffraction.
Now, we are more interested in the Fourier transform of the convolution.

$$\mathscr{F}\{f*g\} = \int_{-\infty}^{\infty} dx e^{-ikx} \int_{-\infty}^{\infty} dy f(y)g(x-y)$$

Using inverse Fourier transform to represent $g(x-y)$

$$\mathscr{F}\{f*g\} = \int_{-\infty}^{\infty} dx e^{-ikx} \int_{-\infty}^{\infty} dy f(y) \frac{1}{2\pi} \int_{-\infty}^{\infty} dk' e^{ik'(x-y)} \tilde{g}(k') =$$

$$= \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dk' \frac{1}{2\pi} f(y)\tilde{g}(k') e^{-ik'y} e^{-i(k-k')x}$$

Reversing the order of integration so that we integrate with respect to x first

$$\mathscr{F}\{f*g\} = \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dk' \int_{-\infty}^{\infty} dx \frac{1}{2\pi} f(y\tilde{g}(k')) e^{-ik'y} e^{-i(k-k')x} =$$

$$= \int_{-\infty}^{\infty} dy f(y) \int_{-\infty}^{\infty} dk' \tilde{g}(k') e^{-ik'y} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i(k-k')x} dx \right) = \int_{-\infty}^{\infty} dy f(y) \int_{-\infty}^{\infty} dk' \tilde{g}(k') e^{-ik'y} \delta(k-k') =$$

$$= \int_{-\infty}^{\infty} dy f(y) \tilde{g}(k) e^{-iky} = \tilde{g}(k) \int_{-\infty}^{\infty} dy f(y) e^{-iky} = \tilde{f}(k)\tilde{g}(k)$$

Therefore, the Fourier transform of a convolution is simply the product of the Fourier trasforms of the functions in convolution. This is called the convolution theorem.

$$\mathscr{F}\{f*g\} = \mathscr{F}\{f\}\mathscr{F}\{g\} \tag{68}$$

6.5.8 Moments of distributions

A general n-th moment of a distribution $f(x)$, $M_n(f)$ is given by integral

$$M_n(f) = \int_{-\infty}^{\infty} x^n f(x)dx$$

Consider the Fourier transform of $x^n f(x)$

$$\mathscr{F}\{x^n f(x)\} = \int_{-\infty}^{\infty} e^{-ikx} x^n f(x)dx$$

Clearly, $M_n(f) = \mathscr{F}\{x^n f(x)\}(k=0)$. Now consider the inverse Fourier transform of derivative of transform of f

$$\mathscr{F}^{-1}\left\{\frac{d^n\tilde{f}}{dk^n}\right\} = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikx}\frac{d^n\tilde{f}}{dk^n}dk$$

We can deconstruct this integral analogously as the one for the derivative of forward Fourier transform, assuming that all derivatives of $\tilde{f}(k)$ go to zero at infinities

$$\int_{-\infty}^{\infty} e^{ikx}\frac{d^n\tilde{f}}{dk^n} = \left[e^{ikx}\frac{d^{n-1}\tilde{f}}{dk^{n-1}}\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty}(ix)e^{ikx}\frac{d^{n-1}\tilde{f}}{dk^{n-1}}dk =$$

$$= \left(\frac{x}{i}\right)^1 \int_{-\infty}^{\infty} e^{ikx}\frac{d^{n-1}\tilde{f}}{dk^{n-1}}dk = \left(\frac{x}{i}\right)^n \int_{-\infty}^{\infty} e^{ikx}\tilde{f}(k)dk$$

Therefore

$$\mathscr{F}^{-1}\left\{\frac{d^n\tilde{f}}{dk^n}\right\} = \left(\frac{x}{i}\right)^n \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikx}\tilde{f}(k)dk = \left(\frac{x}{i}\right)^n f(x)$$

$$\mathscr{F}^{-1}\left\{i^n\frac{d^n\tilde{f}}{dk^n}\right\} = x^n f(x)$$

Taking Fourier transform of both sides, we find out that

$$\mathscr{F}\{x^n f(x)\} = i^n\frac{d^n\tilde{f}}{dk^n} \tag{69}$$

But, this can be used to calculate moments of distributions as

$$M_n(f) = \mathscr{F}\{x^n f(x)\}(k=0) = i^n\frac{d^n\tilde{f}}{dk^n}|_{k=0}$$

Summarizing

$$M_n(f) = i^n\frac{d^n\tilde{f}}{dk^n}|_{k=0} \tag{70}$$

6.6 Generalization to 3D

Fourier transforms in 3D in cartesian coordinates are simply taken with respect to all the coordinates sequentially. Hence

$$\tilde{f}(k_x,k_y,k_z) = \tilde{f}(\vec{k}) = \int_{-\infty}^{\infty} dx e^{-ik_x x}\int_{-\infty}^{\infty} dy e^{-ik_y y}\int_{-\infty}^{\infty} dz e^{-ik_z z} f(x,y,z) =$$

$$= \int_{-\infty}^{\infty} dx\int_{-\infty}^{\infty} dy\int_{-\infty}^{\infty} dz e^{-i(k_x x+k_y y+k_z z)} f(\vec{r}) = \int_{-\infty}^{\infty} dx\int_{-\infty}^{\infty} dy\int_{-\infty}^{\infty} dz e^{-i\vec{k}\cdot\vec{r}} f(\vec{r})$$

Therefore, the general formula (valid in other coordinate systems as well) is

$$\mathscr{F}\{f(\vec{r})\} = \iiint_{all\ space} e^{-i\vec{k}\cdot\vec{r}} f(\vec{r})dV \tag{71}$$

Similarly, the inverse Fourier transform is

$$\mathscr{F}^{-1}\{\tilde{f}(\vec{k})\} = \frac{1}{(2\pi)^3}\iiint_{all\ k-space} \tilde{f}(\vec{k})dV_k \tag{72}$$

where dV_k is the element of the k vector space volume. Denoting real space as V and k space as V_k, we can use these definitions to find the integral representation of Dirac delta in multiple dimensions

$$f(\vec{r}) = \frac{1}{(2\pi)^3}\iiint_{V_k} dV_k e^{i\vec{k}\cdot\vec{r}}\iiint_V e^{-i\vec{k}\cdot\vec{r}'} f(\vec{r}')dV' = \frac{1}{(2\pi)^3}\iiint_{V_k} dV_k\iiint_V dV' e^{-i\vec{k}\cdot(\vec{r}'-\vec{r})} f(\vec{r}')$$

Reversing the order of integration

$$f(\vec{r}) = \frac{1}{(2\pi)^3}\iiint_V dV'\iiint_{V_k} dV_k e^{-\vec{k}\cdot(\vec{r}'-\vec{r})} f(\vec{r}') = \iiint_V dV' f(\vec{r}')\left(\frac{1}{(2\pi)^3}\iiint_{V_k} e^{-i\vec{k}\cdot(\vec{r}'-\vec{r})}dV_k\right)$$

The last expression clearly plays the role of 3D Dirac delta. Therefore

$$\delta^3(\vec{r}-\vec{s}) = \frac{1}{(2\pi)^3}\iiint_{V_k} e^{-i\vec{k}\cdot(\vec{r}-\vec{s})} dV_k$$

Most of other results, importantly including the convolution theorem, have their 3D analogue. The convolution theorem is simply $\mathscr{F}\{f*g\}(\vec{k}) = \tilde{f}(\vec{k})\tilde{g}(\vec{k})$

7 Diffraction

7.1 Wave Equation

7.1.1 Decomposition of 1D wave

The wave equation in 1D has form

$$\frac{\partial^2 u}{\partial t^2} = c^2\frac{\partial^2 u}{\partial x^2}$$

Taking Fourier transform with respect to space first

$$\frac{\partial^2}{\partial t^2}\mathscr{F}_x\{u\} = -c^2k^2\mathscr{F}_x\{u\}$$

and then with respect to time

$$-\omega^2\mathscr{F}_t\{\mathscr{F}_x\{u\}\} = -c^2k^2\mathscr{F}_t\{\mathscr{F}_x\{u\}\}$$

For non-zero total Fourier transform (with repect to both t and x) of u, we can rewrite this as familiar dispersion relation

$$\omega^2 = k^2c^2$$

Or equivalently

$$\omega^2 - k^2c^2 = 0$$

Consider now taking the inverse Fourier transform of some function $\tilde{f}(\omega,k)$. In order to make sure it is a Fourier transform of a function satisfying the wave equation, we can take all possible functions of ω and k and only choose those for which dispersion relation applies using Dirac delta. This means that

$$\tilde{u}(\omega,k) = \tilde{f}(\omega,k)\delta(\omega^2-k^2c^2)$$

where $\tilde{f}$ is any function of ω and k. Now we can take the inverse Fourier transform in frequency

$$\mathscr{F}_t^{-1}\{\tilde{u}(\omega,k)\} = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega t}\tilde{f}(\omega,k)\delta(\omega^2-k^2c^2)$$

However, delta function here becomes zero at two points, at $\pm kc$. We are only able to handle delta function if it becomes zero once in the interval. Lets therefore split the integral so that it consists of two intervals, each with one root of argument of the delta function.

$$\mathscr{F}_t^{-1}\{\tilde{u}\} = \frac{1}{2\pi}\left(\int_{-\infty}^{0} d\omega e^{i\omega t}\tilde{f}(\omega,k)\delta(\omega^2-c^2k^2) + \int_{0}^{\infty} d\omega e^{i\omega t}\tilde{f}(\omega,k)\delta(\omega^2-c^2k^2)\right)$$

We now use substitution $z=\omega^2-c^2k^2$. This has general solution

$$\omega = \pm\sqrt{z+c^2k^2}$$

$$d\omega = \pm\frac{dz}{2\sqrt{z+c^2k^2}}$$

The negative pair corresponds to the first integral (only this pair can produce $\omega\to-\infty$). The positive pair corresponds to the second integral. Therefore the first integral becomes

$$\int_{\infty}^{-ck}\frac{-dz}{2\sqrt{z+c^2k^2}}e^{-i\sqrt{z+c^2k^2}t}\tilde{f}(-\sqrt{z+c^2k^2},k)\delta(z) =$$

$$= \int_{-ck}^{\infty}\frac{1}{2\sqrt{z+c^2k^2}}e^{-i\sqrt{z+c^2k^2}t}\tilde{f}(-\sqrt{z+c^2k^2},k)\delta(z)dz = \frac{1}{2ck}e^{-ickt}\tilde{f}(-ck,k)$$

The second integral becomes

$$\int_{-ck}^{\infty} \frac{1}{2\sqrt{z+c^2k^2}} e^{i\sqrt{z+c^2k^2}t} \tilde{f}(\sqrt{z+c^2k^2}, k)\delta(z)dz = \frac{1}{2ck} e^{ickt} \tilde{f}(ck, k)$$

So

$$\mathscr{F}_t^{-1}\{\tilde{u}\} = \frac{1}{4\pi ck} \left(e^{-ickt} \tilde{f}(-ck, k) + e^{ickt} \tilde{f}(ck, k) \right)$$

Taking the inverse Fourier transform with respect to wavenumber (space), we obtain the original amplitude function

$$\mathscr{F}_x^{-1}\{\mathscr{F}^{-1}\{\tilde{u}\}\} = u(x,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ikx} \frac{1}{4\pi ck} \left(e^{-ickt} \tilde{f}(-ck, k) + e^{ickt} \tilde{f}(ck, k) \right) =$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\tilde{f}(-ck,k)}{4\pi ck} e^{ik(x-ct)} dk + \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\tilde{f}(ck,k)}{4\pi ck} e^{ik(x+ct)} dk$$

These two integrals are just inverse Fourier transforms of some wavenumber functions. If we name these functions

$$\tilde{g}(k) = \frac{\tilde{f}(-ck,k)}{4\pi ck}$$

$$\tilde{h}(k) = \frac{\tilde{f}(ck,k)}{4\pi ck}$$

we then have

$$u(x,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{g}(k) e^{ik(x-ct)} dk + \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{h}(k) e^{ik(x+ct)} dk = g(x-ct) + h(x+ct) \tag{73}$$

Therefore, any wave in one dimension can be expressed as superposition of two waves travelling in opposite directions.

7.1.2 Radially Symmetric 3D waves

In 3D, the wave equation is more complicated and takes form

$$c^2 \nabla^2 u = \frac{\partial^2 u}{\partial t^2}$$

However, for spherically symmetric waves in spherical polar coordinate system, only derivatives with respect to radial direction will be non-zero. Therefore the equation becomes

$$c^2 \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial u}{\partial r} \right) = \frac{\partial^2 u}{\partial t^2}$$

In order to make further progress, we assume that the intensity of the wave, which is a square of the amplitude of the wave, drops as the inverse square of distance from the source, which lies at origin (centre of symmetry). This means that we assume $I \propto u^2 \propto \frac{1}{r^2}$. This can be expressed for some function $v(r,t)$ as

$$u = \frac{v(r,t)}{r}$$

Then, the equation becomes

$$\frac{c^2}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \left(\frac{v}{r} \right) \right) = \frac{1}{r} \frac{\partial^2 v}{\partial t^2}$$

$$\frac{c^2}{r^2} \frac{\partial}{\partial r} \left(r^2 \left(\frac{\frac{\partial v}{\partial r} r - v}{r^2} \right) \right) = \frac{1}{r} \frac{\partial^2 v}{\partial t^2}$$

$$\frac{c^2}{r} \frac{\partial}{\partial r} \left(\frac{\partial v}{\partial r} r - v \right) = \frac{\partial^2 v}{\partial t^2}$$

$$\frac{c^2}{r} \left(\frac{\partial^2 v}{\partial r^2} r + \frac{\partial v}{\partial r} - \frac{\partial v}{\partial r} \right) = \frac{\partial^2 v}{\partial t^2}$$

Which finally leads to

$$c^2\frac{\partial^2 v}{\partial r^2}=\frac{\partial^2 v}{\partial t^2}$$

This means that the v function behaves exactly like a wave in 1D. We can therefore express the spherically symmetric wave as superposition of left and right travelling waves in r, divided by r (the wave is the u field)

$$u(r,t)=\frac{g(r-ct)+h(r+ct)}{r} \tag{74}$$

7.1.3 Monochromatic Light

Now, we consider the effects that a monochromaticity of light has on our derived results. First, for monochromatic light, just one wavelength, adn therefore just one wavenumber describes the wave. This means that functions $\tilde{g}(k)$ and $\tilde{h}(k)$ in (73) become effectively Dirac delta functions $A\delta(k-k_0)$ and $B\delta(k-k_0)$, where k_0 is the wavenumber of the wave. (73) then becomes

$$u(x,t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}A\delta(k-k_0)e^{ik(x-ct)}dk+\frac{1}{2\pi}\int_{-\infty}^{\infty}B\delta(k-k_0)e^{ik(x+ct)}dk$$

Absorbing the $\frac{1}{2\pi}$ into the constants, we then have

$$u(x,t)=Ae^{ik_0(x-ct)}+Be^{ik_0(x+ct)} \tag{75}$$

Hence a monochromatic wave in 1D is always a combination of two plane waves travelling in the opposite direction.
Similarly, for the monochromatic radially symmetrical wave in 3D

$$u(r,t)=\frac{Ae^{ik_0(r-ct)}+Be^{ik_0(r+ct)}}{r} \tag{76}$$

7.2 Huygens' principle

It might seem that spherically symmetrical source is a very specific situation, but in fact the solution of this situation can be extended using the Huygens' principle.
This principle states that every point on a wavefront of a wave behaves like a independent point source. Because it is a point source and it is independent, it is spherically symmetrical. Therefore, progression of any wavefront can be expressed as combination of progression of spherically symmetrical point sources.
There are two last subtleties to take into account. Firtly, at the wavefront, sometimes know the direction of progression of the wavefront in some small distance around it. We then usually choose the constituent plane waves created by the point source so that this direction is obeyed. This means that if we know that the wave is right travelling at some point source, we only choose the $e^{ik_0(r-ct)}$ part, effectively setting B in (76) to zero.
Second subtlety is discusset in following section

7.2.1 Coherent Light

A complication with the Huygens' principle approach is that different points at the wavefront have different relative phases. This can be expressed by some phase difference function $\eta(\vec{r},t)$, so that we can describe any point on the wavefront correctly with respect to others as creating a wave proportional to $e^{i(k_0(r-ct)+\eta(\vec{r},t))}$. This phase difference greatly complicates any calculations, and when its variation is big enough (order of π across the considered points), the stable diffraction phenomena becomes not visible (it might be visible for some slow waves, but for light the pattern simply disappears).
Coherent light is light for which all the points on some wavefront have the same phase, which then enables us to set $\eta(\vec{r},t)=0$.

7.3 Fraunhofer Diffraction

Fraunhofer diffraction is a diffraction in a specific setup, illustrated in Fig. 15. The light incident on the screen from the left is coherent and only right travelling. Therefore, any point source on the screen creates a wave

$$u_p(\vec{r},t)=\frac{A_pe^{ik_0(r_p-ct)}}{r_p}$$

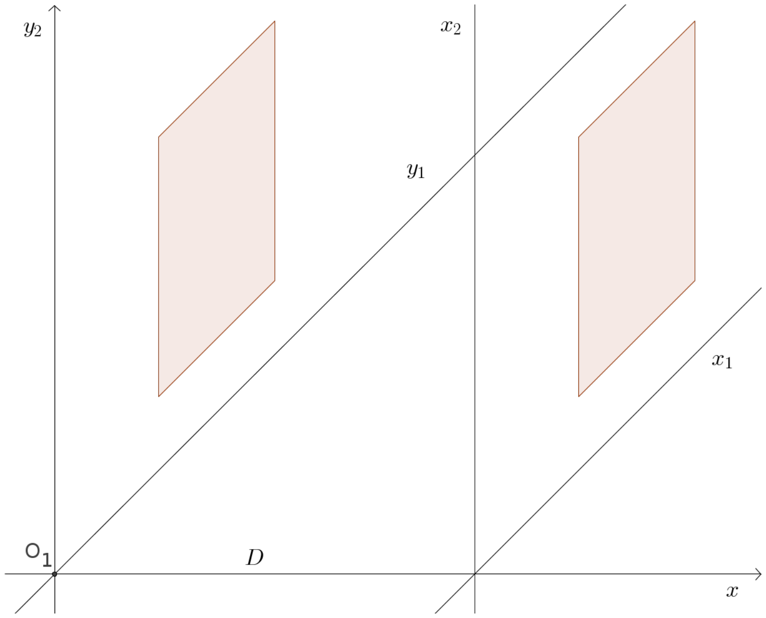

Figure 15: Setup for Fraunhofer diffraction. Light is incident on the plane y_2y_1 from the left and travels in the direction of x. The interference pattern is created on distant screen x_2x_1. The distance of the two screens is D. The origin of both coordinate systems is O_1 point.

where r_p is the distance of some point $\vec{r}$ from the given point source.
The amplitude A_p varies from point to point on the wavefront. It is perhaps better to describe it as a function that gives the intensity at some point surface in screen plane - y_2y_1 on which the light is incident. We call this function the aperture function $a(\vec{y})$, where $\vec{y} = (y_1, y_2, 0)$ is the cartesian coordinate vector on the screen plane. Then

$$A_p(\vec{y}) = a(\vec{y})dy_1dy_2 = a(\vec{y})d^2y$$

where d^2y is just a shorthand for dy_1dy_2.
We can make one more change to make the calculations easier - in order to be able to easily add the effects of several point sources, we need them to be unit point sources. This is achieved by adding a constant to the aperture function, which in this case is $\frac{1}{4\pi}$. Then, we have that from one point source on the screen, the wave is

$$u_p(\vec{r}, t) = \frac{a(\vec{y})e^{ik_0(|\vec{r}-\vec{y}|-ct)}}{4\pi|\vec{r}-\vec{y}|}d^2y$$

where $\vec{r} = (x_1, x_2, x)$ is the position vector on the image screen.
Then, the total interference pattern is constructed by adding together the influences of all independent point sources in the screen plane. The wave amplitude in the image plane is then

$$u_i(\vec{r}, t) = \iint_y a(\vec{y})e^{ik_0(|\vec{r}-\vec{y}|-ct)}\frac{1}{4\pi|\vec{r}-\vec{y}|}d^2y$$

where we integrate over the whole screen plane.
In the zeroth approximation, which is valid for the denominator of the integrand, for big distance D when compared to $|\vec{y}|$ for $\vec{y}$ where $a(\vec{y}) \neq 0$

$$|\vec{r} - \vec{y}| \approx D$$

However, we cannot do the same approximation for the expression in the exponent. Here, we have to be more careful, as the exponent can vary the value of u_i greatly. The first approximation is

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

This can be approximated. Present a new position vector that has origin in the image plane and has x_2 and x_1 axis parallel to those of y_2y_1 coordinate system. Then, the position on the image plane can be given by $\vec{x} = (x_1, x_2, 0)$. Then

$$|\vec{r} - \vec{y}| = \sqrt{(\vec{x} - \vec{y})^2 + x^2}$$

Since the screen is at distance $x = D$, which is big compared to distances $|\vec{x}|$ and $|\vec{y}|$, which we consider

$$|\vec{r}-\vec{y}| = D\sqrt{1+\frac{(\vec{x}-\vec{y})^2}{D^2}} \approx D\left(1+\frac{(|\vec{x}|^2+|\vec{y}|^2+2\vec{x}\cdot\vec{y})}{2D^2}\right) \approx D\left(1+\frac{\vec{x}\cdot\vec{y}}{D^2}\right) \approx D+\frac{\vec{x}\cdot\vec{y}}{D}$$

So, the interference pattern on the image screen is approximated by

$$u_i(\vec{x},t) \approx \frac{1}{4\pi D}\iint_y d^2y a(\vec{y}) e^{-ik_0\left(D+\frac{\vec{x}\cdot\vec{y}}{D}-ct\right)} = \frac{e^{-ik_0(D-ct)}}{4\pi D}\iint_y d^2y a(\vec{y}) e^{-i\frac{k_0\vec{x}}{D}\cdot\vec{y}}$$

Here, I put together $\frac{k_0\vec{x}}{D}$ because it represents a wavevector between two points on the image screen created due to one point source. Therefore, we can write that on the image screen due to one point source

$$\vec{k} = \frac{k_0\vec{x}}{D}$$

Hence, we have formula for the interference pattern

$$u_i(\vec{x},t) = \frac{e^{-ik_0(D-ct)}}{4\pi D}\iint_y a(\vec{y}) e^{-i\vec{k}\cdot\vec{y}} d^2y = \frac{e^{-ik_0(D-ct)}}{4\pi D}\tilde{a}(\vec{k}) \tag{77}$$

The amplitude on the image screen is therefore proportional to the Fourier transform of the aperture function.
Now I include a few examples of diffraction patterns.

7.3.1 Single Point Source

The aperture function for a single point source has a form of $a(\vec{y}) = A\delta^2(\vec{y}-\vec{y}')$ where $\vec{y}'$ is the position of the point source on the screen. The image on the image screen becomes

$$u_i(\vec{x},t) = \frac{e^{-ik_0(D-ct)}}{4\pi D}\iint_y A\delta^2(\vec{y}-\vec{y}') e^{-i\vec{k}\cdot\vec{y}} d^2y = \frac{e^{-ik_0(D-ct)}}{4\pi D} A e^{-i\vec{k}\cdot\vec{y}'}$$

And the observed intensity becomes

$$I = u_i\bar{u}_i = \frac{A^2}{16\pi^2 D^2}$$

which is constant with position on the image plane, $\vec{x}$.

7.3.2 Two Point Sources

Let there be two point sources on the screen so that aperture function is $a(\vec{y}) = A\delta(\vec{y}-\vec{y}_1) + A\delta(\vec{y}-\vec{y}_2)$. Then, the Fourier transform of this function is

$$\iint_y e^{-i\vec{k}\cdot\vec{y}} A(\delta^2(\vec{y}-\vec{y}_1)+\delta^2(\vec{y}-\vec{y}_2)) d^2y = A\iint_y e^{-i\vec{k}\cdot\vec{y}}\delta^2(\vec{y}-\vec{y}_1) d^2y + A\iint_y e^{-i\vec{k}\cdot\vec{y}}\delta^2(\vec{y}-\vec{y}_2) d^2y =$$

$$= Ae^{-i\vec{k}\cdot\vec{y}_1} + Ae^{-i\vec{k}\cdot\vec{y}_2}$$

Therefore, the intensity on the image screen is

$$I = u_i\bar{u}_i = \frac{1}{16\pi^2 D^2} A^2\left(e^{-i\vec{k}\cdot\vec{y}_1}+e^{-i\vec{k}\cdot\vec{y}_2}\right)\left(e^{i\vec{k}\cdot\vec{y}_1}+e^{i\vec{k}\cdot\vec{y}_2}\right) =$$

$$= \frac{A^2}{16\pi^2 D^2}\left(2+e^{i\vec{k}\cdot(\vec{y}_2-\vec{y}_1)}+e^{-i\vec{k}\cdot(\vec{y}_2-\vec{y}_1)}\right) = \frac{A^2}{16\pi^2 D^2}\left(2+2\cos\left(\vec{k}\cdot(\vec{y}_2-\vec{y}_1)\right)\right) =$$

$$= \frac{A^2}{16\pi^2 D^2} 4\cos^2\left(\frac{\vec{k}\cdot(\vec{y}_2-\vec{y}_1)}{2}\right) = \frac{A^2}{16\pi^2 D^2} 4\cos^2\left(\frac{k_0\vec{x}\cdot(\vec{y}_2-\vec{y}_1)}{2D}\right)$$

where I used the definition mentioned earlier $\vec{k} = \frac{k_0\vec{x}}{D}$. We can further express the k_0 using the wavelength to get more familiar form

$$I = \frac{A^2}{16\pi^2 D^2} 4\cos^2\left(\frac{\pi\vec{x}\cdot(\vec{y}_2-\vec{y}_1)}{\lambda D}\right)$$

Therefore, if we move along the direction of $\vec{y}_2 - \vec{y}_1$ along the image screen (changing vector x), we change the phase most rapidly. If we move in a direction normal to $\vec{y}_2 - \vec{y}_1$, we do not change the phase at all. This means that the pattern looks like familiar fringes that are periodically dim and birght and are all perpendicular to direction $\vec{y}_2 - \vec{y}_1$.
The distance between the maxima is the distance between maxima of $\cos^2$, which is distance from maxima of cos to minimum of cos. This means that for $\vec{x}$ along $\vec{y}_2 - \vec{y}_1$ which gives a maximum, the next maximum occurs at $\vec{x} + \Delta\vec{x}$ for which

$$\pi = \frac{\pi(\vec{x} + \Delta\vec{x}) \cdot (\vec{y}_2 - \vec{y}_1)}{\lambda D} - \frac{\pi\vec{x} \cdot (\vec{y}_2 - \vec{y}_1)}{\lambda D}$$

$$1 = \frac{\Delta x l}{\lambda D}$$

where l is the distance between the point sources ($\Delta\vec{x}$ is along $\vec{y}_2 - \vec{y}_1$)

$$\Delta x = \frac{D\lambda}{l}$$

Therefore, the smaller the distance between the sources, the bigger the distance between the maxima of the fringes. Higher order maxima are at multiples of this distance.

7.3.3 Rectangular Slit

Now, consider a case of rectangular shape apreture centered around the origin. If the aperture has width w (along y_1 direction) and height h (along the y_2 direction), the aperture function is

$$a(\vec{y}) = 0, \forall\vec{y} : |y_1| > \frac{w}{2} \vee |y_2| > \frac{h}{2}$$

$$a(\vec{y}) = A, \forall\vec{y} : |y_1| \leq \frac{w}{2} \wedge |y_2| \leq \frac{h}{2}$$

The Fourier transform is

$$\iint_y a(\vec{y}) e^{-i\vec{k}\cdot\vec{y}} d^2y = \int_{-\frac{w}{2}}^{\frac{w}{2}} \int_{-\frac{h}{2}}^{\frac{h}{2}} A e^{-i\vec{k}\cdot\vec{y}} d^2y = A \left(\int_{-\frac{w}{2}}^{\frac{w}{2}} e^{-ik_1 y_1} dy_1 \right) \left(\int_{-\frac{h}{2}}^{\frac{h}{2}} e^{-ik_2 y_2} dy_2 \right) =$$

$$= A\frac{1}{-ik_1} \left(e^{-ik_1\frac{w}{2}} - e^{ik_1\frac{w}{2}} \right) \frac{1}{-ik_2} \left(e^{-ik_2\frac{h}{2}} - e^{ik_2\frac{h}{2}} \right) = \frac{A}{k_1 k_2} 2\sin\left(k_1\frac{w}{2}\right) 2\sin\left(k_2\frac{h}{2}\right) =$$

$$= \frac{Awh}{\frac{k_1 k_2 wh}{4}} \sin\left(k_1\frac{w}{2}\right) \sin\left(k_2\frac{h}{2}\right) = Awh \operatorname{sinc}\left(k_1\frac{w}{2}\right) \operatorname{sinc}\left(k_2\frac{h}{2}\right)$$

The intensity is (substituting for $\vec{k}$)

$$I = \frac{1}{16\pi^2 D^2} A^2 w^2 h^2 \operatorname{sinc}^2\left(\frac{\pi x_1 w}{\lambda D}\right) \operatorname{sinc}^2\left(\frac{\pi x_2 h}{\lambda D}\right)$$

The most important feature in the sinc function is the central peak. It reaches from the value of argument $-\pi$ to π. Then, the central peak has approximate size Δx_1

$$2\pi = \frac{2\pi\Delta x_1 w}{\lambda D 2}$$

$$\Delta x_1 = \frac{2\lambda D}{w}$$

And the size in the other direction is

$$\Delta x_2 = \frac{2\lambda D}{h}$$

Therefore, the central slit is wider in horizontal direction than in vertical direction if the apreture is narrower in horizontal direction and wider in vertical direction. This leads to a rather bizzare prediction that for infinitely wide slit in vertical direction the image is slit that is infinitely wide in the horizontal direction. The reason why we do not observe this commonly in experiment is because the approximation that for all $\vec{r}$, $\vec{x}$ and $\vec{y}$ are small does not apply anymore, and the difraction pattern has to be calculated otherwise.

7.3.4 Two slits

To describe two slits, we need to somehow just replicate and displace the aperture function from previous problem for a single slit. This can be elegantly done using the convolution. The two slits at positions $\vec{y}_1$ and $\vec{y}_2$ can be described as

$$a_2(\vec{y}) = a * g(\vec{y})$$

where

$$g(\vec{y}) = \delta(\vec{y} - \vec{y}_1) + \delta(\vec{y} - \vec{y}_2)$$

The Fourier transform is

$$\mathscr{F}\{a * g\} = \mathscr{F}\{a\}\mathscr{F}\{g\}$$

We already calculated both these transforms. Hence the product is

$$\mathscr{F}\{a * g\} = Awh \operatorname{sinc}\left(\frac{\pi x_1 w}{\lambda D}\right) \operatorname{sinc}\left(\frac{\pi x_2 h}{\lambda D}\right)\left(e^{-i\vec{k}\cdot\vec{y}_2} + e^{-i\vec{k}\cdot\vec{y}_2}\right)$$

Hence the intensity on the image screen is

$$I = \frac{1}{16\pi^2 D^2} 16A^2 w^2 h^2 \operatorname{sinc}^2\left(\frac{\pi x_1 w}{\lambda D}\right) \operatorname{sinc}^2\left(\frac{\pi x_2 h}{\lambda D}\right) \cos^2\left(\frac{2\pi\vec{x}}{\lambda D}\cdot(\vec{y}_2 - \vec{y}_1)\right)$$

7.3.5 Diffraction Grating

Diffraction grating characteristic follows exactly the same analysis as above, but $g(\vec{y})$ is now

$$g(\vec{y}) = \sum_{n=-N}^{N} \delta(\vec{y} + n\vec{l})$$

where $\vec{l} = l\hat{y}_1$ is the displacement vector between two slits of the grating (I choose for all the slits to be ordered in the horizontal direction). Then

$$\tilde{g}(\vec{y}) = \sum_{n=-N}^{N} e^{i\vec{k}\cdot(n\vec{l})} = \sum_{n=-N}^{N} e^{ik_1 nl} = e^{-ik_1 Nl}\sum_{n=0}^{2N} e^{ik_1 nl}$$

This is a standard geometric series with sum equal to

$$\tilde{g}(\vec{k}) = e^{-ik_1 Nl}\frac{1 - e^{ik_1(2N+1)l}}{1 - e^{ik_1 l}}$$

The absolute value of this squared is

$$|\tilde{g}(\vec{k})|^2 = \frac{(1 - e^{ik_1(2N+1)l})(1 - e^{-ik_1(2N+1)l})}{(1 - e^{ik_1 l})(1 - e^{-ik_1 l})} = \frac{1 - (e^{-ik_1(2N+1)l} + e^{ik_1(2N+1)l}) + 1}{1 - (e^{ik_1 l} + e^{-ik_1 l}) + 1} =$$

$$= \frac{2 - 2\cos(k_1(2N+1)l)}{2 - 2\cos(k_1 l)} = \frac{\sin^2\left(\frac{k_1(2N+1)l}{2}\right)}{\sin^2\left(\frac{k_1 l}{2}\right)} = \frac{\sin^2\left(\frac{\pi x_1(2N+1)l}{\lambda D}\right)}{\sin^2\left(\frac{\pi x_1 l}{\lambda D}\right)}$$

This is a typical pattern with some big maxima with several smaller maxima in between.
Here, $2N+1$ is the total number of the slits. Therefore, the total diffraction intensity is

$$I = \frac{|\tilde{a}|^2|\tilde{g}|^2}{16\pi^2 D^2} = \frac{1}{16\pi^2 D^2} 4Awh \operatorname{sinc}^2\left(\frac{\pi x_1 w}{\lambda D}\right) \operatorname{sinc}^2\left(\frac{\pi x_2 h}{\lambda D}\right) \frac{\sin^2\left(\frac{\pi x_1(2N+1)l}{\lambda}\right)}{\sin^2\left(\frac{\pi x_1 l}{\lambda D}\right)}$$

7.3.6 Circular Aperture

The circular aperture has aperture function

$$a(\vec{y}) = A, \forall \vec{y} : |\vec{y}| \leq R$$

$$a(\vec{y}) = 0, \forall \vec{y} : |\vec{y}| > R$$

where R is the radius of the aperture, which is centered around the origin.
We now do the Fourier transform. To make the integration sensible, we do the integration in planar polar coordinates, with $r = |\vec{y}|$ and ϕ the angle from the y_1 axis. Then

$$\tilde{a}(\vec{k}) = \int_0^\infty \int_0^{2\pi} a(\vec{y}) e^{-i\vec{k}\cdot\vec{y}} r d\phi dr = \int_0^R r \int_0^{2\pi} e^{-i|\vec{k}||\vec{y}|\cos(\phi_k - \phi)} d\phi dr$$

where ϕ_k is the angle from x_1 axis to point $\vec{x}_1$. Substituting $\alpha = \phi_k - \phi$ and $z = |\vec{k}||\vec{y}| = \frac{2\pi r_x r}{\lambda D}$ where r_x is the distance of the point $\vec{x}$ from the origin projected on the image screen. Then

$$\tilde{a}(\vec{k}) = \int_0^R r dr \int_{\phi_k}^{\phi_k - 2\pi} -d\alpha e^{-iz\cos(\alpha)}$$

Since we complete a full circle in the inner integral, we can rewrite the limits (using the minus side inside the integral) as

$$\tilde{a}(\vec{k}) = \int_0^R r dr \int_0^{2\pi} e^{-iz\cos(\alpha)} d\alpha$$

The inner integral is the integral representation of the Bessel function of the first kind of order 0.

Bessel functions of the first kind Bessel function of the first kind of the zeroth order has integral representation

$$J_0(z) = \frac{1}{2\pi} \int_0^{2\pi} e^{-iz\cos(\alpha)} d\alpha \tag{78}$$

The Bessel functions of the first kind are connected via recursion relation

$$\frac{d}{dz}\left(z^n J_n(z)\right) = z^n J_{n-1}(z) \tag{79}$$

Also, Bessel functions are solutions of differential equation

$$\left(\frac{d^2}{dz^2} + \frac{1}{z}\frac{d}{dz} + \left(1 - \frac{n^2}{z^2}\right)\right) f(z) = 0$$

where n is the order of the Bessel function. This differential equation has two solutions, and the second one is the Bessel function of the second kind, which is often written as $Y_n(z)$.
In our problem, we then have

$$\tilde{a}(\vec{k}) = \int_0^R r dr 2\pi J_0(z) = 2\pi \int_0^R r J_0(|\vec{k}|r) dr = 2\pi \int_0^{|\vec{k}|R} \frac{z}{|\vec{k}|} J_0(z) \frac{dz}{|\vec{k}|} = \frac{2\pi}{|\vec{k}|^2} \int_0^{|\vec{k}|R} z J_0(z) dz$$

Using the recursion relation

$$\tilde{a}(\vec{k}) = \frac{2\pi}{|\vec{k}|^2} \int_0^{|\vec{k}|R} \frac{d}{dz}\left(z J_1(z)\right) dz = \frac{2\pi}{|\vec{k}|^2} \left[z J_1(z)\right]_0^{|\vec{k}|R}$$

Since J_1 does not diverge at zero

$$\tilde{a}(\vec{k}) = \frac{2\pi}{|\vec{k}|^2} |\vec{k}| R J_1(|\vec{k}|R) = 2\pi \frac{J_1\left(|\vec{k}|R\right)}{\frac{|\vec{k}|}{R}} = 2\pi \frac{J_1\left(\frac{2\pi r_x R}{\lambda D}\right)}{\frac{2\pi r_x}{\lambda D R}}$$

7.3.7 Angular Resolution

We can rewrite the above relation with angle from the centre of the aperture θ (for $r_x << D$) as ($k_0 = \frac{2\pi}{\lambda}$)

$$\tilde{a}(\vec{k}) = 2\pi \frac{J_1\left(k_0\theta R\right)}{\frac{k_0\theta}{R}}$$

We say that two objects are just resolved when the first diffraction pattern minimum lies in the zeroth maximum of the other object. The first minimum of Bessel function of the first kind and of the first order appears approximately at 1.22π. This means that the minimum angle separation between the objects has to be $\Delta\theta$ such that

$$1.22\pi = k_0\Delta\theta R$$

$$\Delta\theta = \frac{1.22\lambda}{2R} = 1.22\frac{\lambda}{d}$$

where d is the diameter of the circular aperture. Therefore, telescopes can resolve more if they are larger.