Review of Partial Derivatives

Definition of Partial Derivatives. Let $z = f(x,y)$ be a function. Then the partial derivative of f with respect to x is defined as

$$\frac{\partial f}{\partial x} \equiv f_x(x,y) \equiv \lim_{h\to 0} \frac{f(x+h,y)-f(x,y)}{h}$$

and the partial derivative of f with respect to y is defined as

$$\frac{\partial f}{\partial y} \equiv f_y(x,y) \equiv \lim_{k\to 0} \frac{f(x,y+k)-f(x,y)}{k}$$

Partial derivatives may be computed algebraically; all of the same rules that applied to regular derivatives also apply to partial derivatives. **The only trick to remember when taking a partial derivative is hold all other variables (besides the one we are differentiating with respect to) constant.**

Example. Let $f(x,y) = \frac{x^2}{1+y}$. Then

$$f_x(x,y) = \frac{\partial}{\partial x}\frac{x^2}{1+y} = \frac{1}{1+y}\frac{\partial}{\partial x}x^2 = \frac{2x}{1+y}$$

$$f_y(x,y) = \frac{\partial}{\partial y}\frac{x^2}{1+y} = x^2\frac{\partial}{\partial y}\frac{1}{1+y} = x^2\frac{\partial}{\partial y}(1+y)^{-1} = \frac{-x^2}{(1+y)^2}$$

Example. Find the partial derivatives of $f(x,y) = y^2 e^{3x}$

$$f_x(x,y) = \frac{\partial}{\partial x}\left(y^2 e^{3x}\right) = y^2\frac{\partial}{\partial x}\left(e^{3x}\right) = y^2\cdot 3e^{3x} = 3y^2 e^{3x}$$

$$f_y(x,y) = \frac{\partial}{\partial y}\left(y^2 e^{3x}\right) = e^{3x}\frac{\partial}{\partial y}\left(y^2\right) = e^{3x}\cdot 2y = 2ye^{3x}$$

Example. Find the partial derivative of $f(x,y,z) = \frac{x^2y^3}{z}$

$$f_x(x,y) = \frac{\partial}{\partial x}\left(\frac{x^2y^3}{z}\right) = \frac{y^3}{z}\frac{\partial}{\partial x}\left(x^2\right) = \frac{y^3}{z}\cdot 2x = \frac{2xy^3}{z}$$

$$f_y(x,y) = \frac{\partial}{\partial y}\left(\frac{x^2y^3}{z}\right) = \frac{x^2}{z}\frac{\partial}{\partial y}\left(y^3\right) = \frac{x^2}{z}\cdot 3y^2 = \frac{3x^2y^2}{z}$$

$$f_z(x,y) = \frac{\partial}{\partial z}\left(\frac{x^2y^3}{z}\right) = x^2y^3\frac{\partial}{\partial z}\left(\frac{1}{z}\right) = x^2y^3\frac{\partial}{\partial z}\left(z^{-1}\right) = x^2y^3\left(-z^{-2}\right) = -\frac{x^2y^3}{z^2}$$

Second Order Partial Derivatives

We define higher order partial derivatives in much the same way as we did in single-variable calculus.

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

$$f_{yy} = \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)$$

With partial derivatives, we can also combine the variables, so there are more derivatives at each order. For example, we can differentiate f_x with respect to *y* and we can differentiate f_y with respect to *x*:

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

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

We can, of course, combine higher order derivatives in any order we like. For example:

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

There are not as many partials as you might think, however, because of the following theorem:

The order of the partials can be reversed: $\boxed{f_{xy} = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = f_{yx}}$

Example. Find all the second-order partial derivatives of $f(x,y) = xy^2 + 3x^2e^y$ and show that $f_{xy} = f_{yx}$ by taking partials in both orders.

$$f_x = y^2 + 6xe^y \;\Rightarrow\; f_{xx} = 6e^y,\;\; f_{yx} = \frac{\partial}{\partial y}\left(y^2 + 6xe^y\right) = 2y + 6xe^y$$

$$f_y = 2xy + 3x^2e^y \;\Rightarrow\; f_{yy} = 2x + 3x^2e^y,\;\; f_{xy} = \frac{\partial}{\partial x}\left(2xy + 3x^2e^y\right) = 2y + 6xe^y = f_{yx}$$

Example. Repeat the above example for $f(x,y) = xe^y$

$$f_x = e^y \;\Rightarrow\; f_{xx} = 0,\quad f_{yx} = e^y$$

$$f_y = xe^y \;\Rightarrow\; f_{yy} = xe^y,\quad f_{xy} = e^y = f_{yx}$$

Chain Rule

If $z(t) = f\big(x(t),\ y(t)\big)$ then the chain rule is

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

In general, if *z* is a function of any number of variables $x(t),\ y(t),\ z(t),\ w(t),\ \ldots$, each of which can be expressed as a function of only *t* (and not of any other parameter),

$$\frac{d}{dt} f\big(x,y,z,w,\ldots\big) = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} + \frac{\partial f}{\partial w}\frac{dw}{dt} + \ldots$$

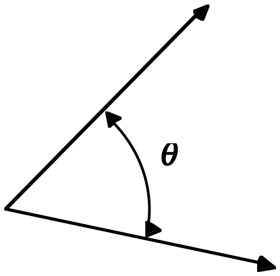
<u>Example.</u> Suppose $f(x,y) = x\sin y$, where $x = t^2$ and $y = 2t+1$. Then by the chain rule we have

$$\begin{aligned} f'(t) &= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \\ &= \frac{\partial}{\partial x}\big(x\sin y\big)\frac{d}{dt}\big(t^2\big) + \frac{\partial}{\partial y}\big(x\sin y\big)\frac{d}{dt}(2t+1) \\ &= \big(\sin y\big)(2t) + \big(x\cos y\big)(2) \\ &= 2t\sin(2t+1) + 2t^2\cos(2t+1) \end{aligned}$$

Vector Products

There are two types of products between vectors, one of which produces a vector and the other produces a scalar

- The dot product $\vec{v}\cdot\vec{w}\longrightarrow \mathit{scalar}$
- The cross product $\vec{v}\times\vec{w}\longrightarrow \mathit{vector}$



The Dot Product is defined geometrically

$$\vec{v}\cdot\vec{w}=|v||w|\cos\theta$$

where q is the angle between the two vectors as shown in the figure. Algebraically, if

$$\vec{v}=\vec{i}\,v_1+\vec{j}\,v_2+\vec{k}\,v_3 \quad \text{and} \quad \vec{w}=\vec{i}\,w_1+\vec{j}\,w_2+\vec{k}\,w_3$$

Then $\vec{v}\cdot\vec{w}=\vec{w}\cdot\vec{v}=v_1w_1+v_2w_2+v_3w_3$

Example. Suppose that $\vec{u}=3\vec{i}+4\vec{j}+5\vec{k}$ and $\vec{v}=7\vec{i}+8\vec{j}+9\vec{k}$

Then $\vec{u}\cdot\vec{v}=(3)(7)+(4)(8)+(5)(9)=21+32+45=98$

Properties of the dot product

1. $\vec{v}\cdot\vec{w}=\vec{w}\cdot\vec{v}$
2. $\vec{v}\cdot(a\vec{w})=(a\vec{v})\cdot\vec{w}=a(\vec{v}\cdot\vec{w})$
3. $(\vec{v}+\vec{u})\cdot\vec{w}=\vec{v}\cdot\vec{w}+\vec{u}\cdot\vec{w}$
4. $\vec{v}$ and $\vec{w}$ are perpendicular only if $\vec{v}\cdot\vec{w}=0$.

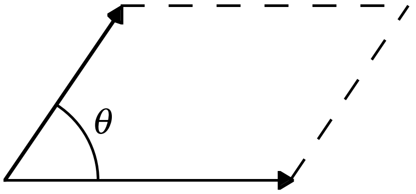
The Cross Product

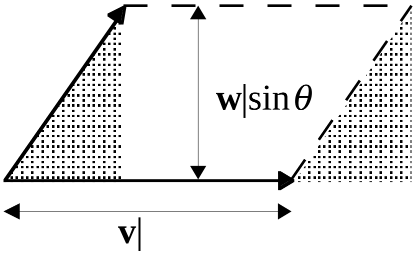
The cross product is a product between vectors that results in a vector. It is defined as a vector with the following properties:

- Its length is equal to $|\vec{v}\times\vec{w}|=|\vec{v}||\vec{w}|\sin\theta$
- direction is perpendicular to the plane that contains v and w
- Its orientation (up vs. down) is according to the right hand rule

Right-Hand Rule: Place $\vec{u}$ and $\vec{v}$ so that their tails coincide and curl the fingers of your right hand from through the angle from $\vec{u}$ to $\vec{v}$. Your thumb is pointing in the direction of $\vec{u}\times\vec{v}$

The cross product gives the area of the parallelogram formed by the two vectors:





We can also calculate the cross product algebraically from the components of the individual vectors if we use determinants.

$$\vec{v}\times\vec{w}=\begin{vmatrix}\vec{i} & \vec{j} & \vec{k}\\ v_1 & v_2 & v_3\\ w_1 & w_2 & w_3\end{vmatrix}=\vec{i}\begin{vmatrix}v_2 & v_3\\ w_2 & w_3\end{vmatrix}-\vec{j}\begin{vmatrix}v_1 & v_3\\ w_1 & w_3\end{vmatrix}+\vec{k}\begin{vmatrix}v_1 & v_3\\ w_1 & w_3\end{vmatrix}$$

$$=\vec{i}\left(v_2w_3-v_3w_2\right)-\vec{j}\left(v_1w_3-v_3w_1\right)+\vec{k}\left(v_1w_2-v_2w_1\right)$$

Determinant of a Matrix

$$\det\begin{pmatrix}a & b\\ c & d\end{pmatrix}\equiv\begin{vmatrix}a & b\\ c & d\end{vmatrix}\equiv ad-bc$$

$$\det\begin{pmatrix}a & b & c\\ d & e & f\\ g & h & i\end{pmatrix}=\begin{vmatrix}a & b & c\\ d & e & f\\ g & h & i\end{vmatrix}=a\begin{vmatrix}e & f\\ h & i\end{vmatrix}-b\begin{vmatrix}d & f\\ g & i\end{vmatrix}+c\begin{vmatrix}d & e\\ g & h\end{vmatrix}$$

$$=a(ei-fh)-b(di-fg)+c(dh-eg)$$

Properties of the Cross Product

(1) $\vec{w}\times\vec{v}=-\vec{v}\times\vec{w}$

(2) $(a\vec{v})\times\vec{w}=a(\vec{v}\times\vec{w})=\vec{v}\times(a\vec{w})$

(3) $\vec{u}\times(\vec{v}+\vec{w})=\vec{u}\times\vec{v}+\vec{u}\times\vec{w}$

(4) $\vec{u}\times\vec{v}=\mathbf{0}$ if and only if the vectors are parallel (assuming that $\vec{u},\vec{v}\neq\vec{0}$)

(5) $\vec{i}\times\vec{j}=\vec{k},\ \vec{j}\times\vec{k}=\vec{i},\ \vec{k}\times\vec{i}=\vec{j}$ (cyclic cross products)

(6) $\vec{j}\times\vec{i}=-\vec{k},\ \vec{k}\times\vec{j}=-\vec{i},\ \vec{i}\times k=-\vec{j}$ (acyclic cross products)

(7) $\vec{v}\times\vec{v}=0$

Vector Operators: Gradient, Divergence and Curl

The **gradient of a function of three variables** is the vector

$$\text{grad } f\equiv\nabla f(x,y,z)\equiv\vec{i}\,\frac{\partial f}{\partial x}+\vec{j}\,\frac{\partial f}{\partial y}+\vec{k}\,\frac{\partial f}{\partial z}$$

We can think of the gradient symbol ∇ as a operator that behaves exactly like a vector with the exception of the fact that it operates on the single item immediately to its right.

In this sense, we can think of ∇ as vector version of $\frac{\partial}{\partial x}$.

Specifically, ∇ represents the “vector” $\vec{\nabla}=\vec{i}\,\frac{\partial}{\partial x}+\vec{j}\,\frac{\partial}{\partial y}+\vec{k}\,\frac{\partial}{\partial z}$

Example. $f = xe^{y^2+z}$

Then

$$\begin{aligned}\vec{\nabla} f &= \vec{\nabla}\left(xe^{y^2+z}\right) = \left(\vec{i}\,\frac{\partial}{\partial x} + \vec{j}\,\frac{\partial}{\partial y} + \vec{k}\,\frac{\partial}{\partial z}\right)\left(xe^{y^2+z}\right)\\ &= \vec{i}\,\frac{\partial}{\partial x}\left(xe^{y^2+z}\right) + \vec{j}\,\frac{\partial}{\partial y}\left(xe^{y^2+z}\right) + \vec{k}\,\frac{\partial}{\partial z}\left(xe^{y^2+z}\right)\\ &= \vec{i}\,e^{y^2+z} + \vec{j}\,2xye^{y^2+z} + \vec{k}xe^{y^2+z}\end{aligned}$$

Since we are treating ∇ "like a vector," we can subject to all of the usual vector operations, such as dot product and cross product. We will define the following operations using ∇

Operation	Name of Operator	Input	Output
∇	gradient	scalar	vector
$\nabla\cdot$	divergence (dot product) read as "del dot ..."	vector	scalar
$\nabla\times$	curl (cross product) read as "del cross ..."	vector	vector
$\nabla\cdot\nabla$ or ∇^2	Laplacian read as "del squared ..."	scalar	scalar

Both the divergence and curl operate on vector functions. A vector function is a function

$$\vec{F}(x,y,z) = \vec{i}\,F_1(x,y,z) + \vec{j}F_2(x,y,z) + \vec{k}F_3(x,y,z)$$

where the functions $F_1,\ F_2,\ F_3$ are all real valued functions.

Example of a Vector Function. $\vec{F}(x,y,z) = x\vec{i} + \left(y^2x+z\right)\vec{j} + e^x\vec{k}$

The Divergence Operator is defined as "del dot a vector function",

$$\begin{aligned}\text{div }\vec{F} = \vec{\nabla}\cdot\vec{F} &= \left(\vec{i}\,\frac{\partial}{\partial x} + \vec{j}\,\frac{\partial}{\partial y} + \vec{k}\,\frac{\partial}{\partial z}\right)\cdot\left(\vec{i}\,F_1 + \vec{j}F_2 + \vec{k}F_3\right)\\ &= \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}\end{aligned}$$

Example. Find the divergence of $\vec{F}(x,y,z) = x\vec{i} + \left(y^2x+z\right)\vec{j} + e^x\vec{k}$

$$\begin{aligned}\vec{\nabla}\cdot\vec{F} &= \left(\vec{i}\,\frac{\partial}{\partial x} + \vec{j}\,\frac{\partial}{\partial y} + \vec{k}\,\frac{\partial}{\partial z}\right)\cdot\left(x\vec{i} + \left(y^2x+z\right)\vec{j} + e^x\vec{k}\right)\\ &= \frac{\partial x}{\partial x} + \frac{\partial}{\partial y}\left(y^2x+z\right) + \frac{\partial}{\partial z}e^x = 1 + 2xy + 0 = 1 + 2xy\end{aligned}$$

The Curl Operator is defined as “del cross f”

$$\text{curl } \vec{F} = \vec{\nabla} \times \vec{F} = \left(\vec{i}\,\frac{\partial}{\partial x} + \vec{j}\,\frac{\partial}{\partial y} + \vec{k}\,\frac{\partial}{\partial z} \right) \times \left(\vec{i}\,F_1 + \vec{j}F_2 + \vec{k}F_3 \right)$$

$$= \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{vmatrix} = \vec{i}\left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z} \right) + \vec{j}\left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x} \right) + \vec{k}\left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right)$$

Example. Suppose $\vec{F} = e^{y^2}\vec{i} \;+\; 2xye^{y^2}\vec{j} \;+\; \vec{k}$

$$\vec{\nabla} \times \vec{F} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ e^{y^2} & 2xye^{y^2} & 1 \end{vmatrix}$$

$$= \vec{i}\left(\frac{\partial(1)}{\partial y} - \frac{\partial}{\partial z}[2xye^{y^2}] \right) + \vec{j}\left(\frac{\partial}{\partial z}e^{y^2} - \frac{\partial(1)}{\partial x} \right) + \vec{k}\left(\frac{\partial}{\partial x}[2xye^{y^2}] - \frac{\partial e^{y^2}}{\partial y} \right)$$

$$= \vec{i}\,(0-0) + \vec{j}\,(0-0) + \vec{k}\left(2ye^{y^2} - 2ye^{y^2} \right) = \mathbf{0}$$

The Laplacian Operator is the dot product of ∇ with itself, or the divergence of the gradient. It is sometimes read as “del squared”.

$$\nabla^2 f(x,y,z) = \vec{\nabla} \cdot \left(\vec{\nabla} f(x,y,z) \right)$$

$$= \left(\vec{i}\,\frac{\partial}{\partial x} + \vec{j}\,\frac{\partial}{\partial y} + \vec{k}\,\frac{\partial}{\partial z} \right) \cdot \left[\left(\vec{i}\,\frac{\partial}{\partial x} + \vec{j}\,\frac{\partial}{\partial y} + \vec{k}\,\frac{\partial}{\partial z} \right) f \right]$$

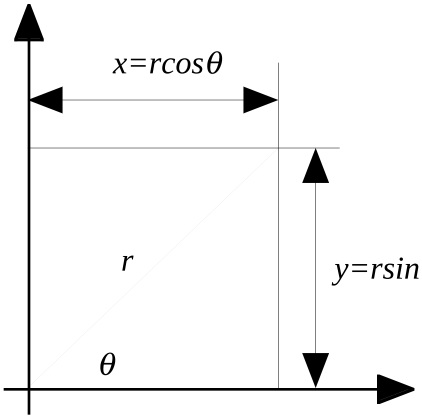
$$= \left(\vec{i}\,\frac{\partial}{\partial x} + \vec{j}\,\frac{\partial}{\partial y} + \vec{k}\,\frac{\partial}{\partial z} \right) \cdot \left(\vec{i}\,\frac{\partial f}{\partial x} + \vec{j}\,\frac{\partial f}{\partial y} + \vec{k}\,\frac{\partial f}{\partial z} \right)$$

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

The Laplacian operates on a scalar and its output is a scalar.

Polar Coordinates

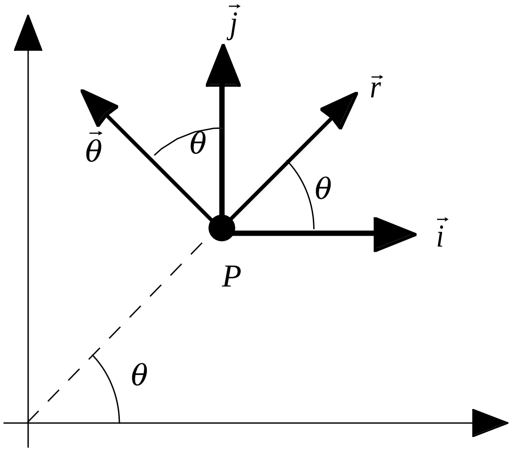
In polar coordinates, instead of using the distances x and y from the origin to locate a point, we use a single distance r and an orientation angle with respect to the x-axis.



The two coordinate systems are related to each other as follows:

$$x = r\cos\theta \qquad r=\sqrt{x^2+y^2}$$
$$y = r\sin\theta \qquad \theta = \arctan\frac{y}{x}$$

We can also define unit vectors $\vec{r}$ and $\vec{\theta}$ at any given point in space; they can be related to the cartesian unit vectors $\vec{i}$ and $\vec{j}$ from the following geometry.



Since all the vectors shown are unit vectors, we observe that:

x component of $\hat{r}$ is $\cos\theta$; $\qquad$ x component of $\hat{\theta}$ is $-\sin\theta$

y component of $\hat{r}$ is $\sin\theta$; $\qquad$ y component of $\hat{\theta}$ is $\cos\theta$

Therefore

$$\hat{r} = \cos\theta\hat{i} + \sin\theta\hat{j}$$
$$\hat{\theta} = -\sin\theta\hat{i} + \cos\theta\hat{j}$$

Exercise: Find expressions for $\hat{i}$ and $\hat{j}$ in terms of $\hat{r}$ and $\hat{\theta}$.

Solution. Multiply the first expression by $\sin\theta$ and the second expression by $\cos\theta$

$$\hat{r}\sin\theta = \sin\theta\cos\theta\hat{i} + \sin^2\theta\hat{j}$$
$$\hat{\theta}\cos\theta = -\cos\theta\sin\theta\hat{i} + \cos^2\theta\hat{j}$$

Adding the two expressions,

$$\hat{r}\sin\theta + \hat{\theta}\cos\theta = \sin^2\theta\hat{j} + \cos^2\theta\hat{j} = \hat{j}$$

which gives an expression for $\hat{j}$ in terms of $\hat{r}$ and $\hat{\theta}$.

To get a similar expressionf or $\hat{i}$, multiply the expression for $\hat{r}$ by $\cos\theta$ and the expression for $\hat{\theta}$ by $\sin\theta$,

$$\hat{r}\cos\theta = \cos^2\theta\hat{i} + \cos\theta\sin\theta\hat{j}$$
$$\hat{\theta}\sin\theta = -\sin^2\theta\hat{i} + \cos\theta\sin\theta\hat{j}$$

Subtract the bottom expression from the top,

$$\hat{r}\cos\theta - \hat{\theta}\sin\theta = \cos^2\theta\hat{i} + \sin^2\theta\hat{i} = \hat{i}$$

Summarizing the conversion expressions, we have

$$\begin{array}{lcl} i = \hat{r}\cos\theta - \hat{\theta}\sin\theta & & \vec{r} = \cos\theta\vec{i} + \sin\theta\vec{j} \\ & \text{and} & \\ \hat{j} = \hat{r}\sin\theta + \hat{\theta}\cos\theta & & \vec{\theta} = -\sin\theta\vec{i} + \cos\theta\vec{j} \end{array}$$

Example. Find an expression for the vector field $\vec{F} = 3x\hat{i} + 4x^2y\hat{j}$ in polar coordinates.

Solution.

$$\begin{aligned} &\vec{F} = 3x\hat{i} + 4x^2y\hat{j} \\ &= 3(r\cos\theta)(\hat{r}\cos\theta - \hat{\theta}\sin\theta) + 4(r\cos\theta)^2(r\sin\theta)(\hat{r}\cos\theta + \hat{\theta}\sin\theta) \\ &= 3r\cos^2\theta\hat{r} - 3r\cos\theta\sin\theta\hat{\theta} + 4r^3\cos^3\theta\sin\theta\hat{r} + 4r^3\cos^2\theta\sin^2\theta\hat{\theta} \\ &= (3r\cos^2\theta + 4r^3\cos^3\theta\sin\theta)\hat{r} + (4r^3\cos^2\theta\sin^2\theta - 3r\cos\theta\sin\theta)\hat{\theta} \\ &= r\cos^2\theta(3 + 4r^2\cos\theta\sin\theta)\hat{r} + r\cos\theta\sin\theta(4r^2\cos\theta\sin\theta - 3)\hat{\theta} \end{aligned}$$

Exercise: Recall that the gradient in two dimensions in cartesian coordinates is $\nabla f(x,y) = \hat{i}\dfrac{\partial f}{\partial x} + \hat{j}\dfrac{\partial f}{\partial y}$. Find an expression for the gradient in polar coordinates, i .e., find the functions *g* and *h* so that $\nabla f(r,\theta) = \hat{r}g(r,\theta) + \hat{\theta}h(r,\theta)$. Hint: you must use the chain rule and the expressions for transformation of coordinates.

Exercise: Find an expression for the gradient $\nabla f(x,y) = \hat{i}\,\frac{\partial f}{\partial x} + \hat{j}\,\frac{\partial f}{\partial y}$ in polar coordinates, i .e., find the functions *g* and *h* so that $\nabla f(r,\theta) = \hat{r}g(r,\theta) + \hat{\theta}h(r,\theta)$.

Solution: From the chain rule,

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial f}{\partial \theta}\frac{\partial \theta}{\partial x} \quad \text{and} \quad \frac{\partial f}{\partial y} = \frac{\partial f}{\partial r}\frac{\partial r}{\partial y} + \frac{\partial f}{\partial \theta}\frac{\partial \theta}{\partial y}$$

So we need expressions for $\partial r/\partial x,\ \partial r/\partial y,\ \ \partial\theta/\partial x,\ \ \partial\theta/\partial y$. But since $r^2 = x^2 + y^2$ we have by implicit differentiation that

$$2r\frac{\partial r}{\partial x} = 2x \quad \Rightarrow \quad \frac{\partial r}{\partial x} = \frac{x}{r} = \frac{r\cos\theta}{r} = \cos\theta$$

$$2r\frac{\partial r}{\partial y} = 2y \quad \Rightarrow \quad \frac{\partial r}{\partial y} = \frac{y}{r} = \frac{r\sin\theta}{r} = \sin\theta$$

To get the partial derivatives of θ with respect to x and y we use $\theta = \tan^{-1}(y/x)$

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

$$\frac{\partial\theta}{\partial y} = \frac{1}{1+(y/x)^2}\left(\frac{1}{x}\right) = \frac{1}{x+y^2/x} = \frac{x}{x^2+y^2} = \frac{r\cos\theta}{r^2} = \frac{\cos\theta}{r}$$

Therefore

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial f}{\partial \theta}\frac{\partial \theta}{\partial x} = \cos\theta\frac{\partial f}{\partial r} - \frac{\sin\theta}{r}\frac{\partial f}{\partial\theta}$$

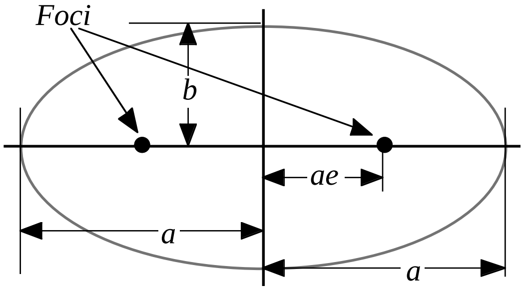
$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial r}\frac{\partial r}{\partial y} + \frac{\partial f}{\partial \theta}\frac{\partial \theta}{\partial y} = \sin\theta\frac{\partial f}{\partial r} + \frac{\cos\theta}{r}\frac{\partial f}{\partial\theta}$$

Thus the gradient is

$$\begin{aligned}
\nabla f(x,y) &= \hat{i}\,\frac{\partial f}{\partial x} + \hat{j}\,\frac{\partial f}{\partial y} \\
&= (\hat{r}\cos\theta - \hat{\theta}\sin\theta)\left(\cos\theta\frac{\partial f}{\partial r} - \frac{\sin\theta}{r}\frac{\partial f}{\partial\theta}\right) + (\hat{r}\sin\theta + \hat{\theta}\cos\theta)\left(\sin\theta\frac{\partial f}{\partial r} + \frac{\cos\theta}{r}\frac{\partial f}{\partial\theta}\right) \\
&= \hat{r}\cos^2\theta\frac{\partial f}{\partial r} - \hat{r}\frac{\cos\theta\sin\theta}{r}\frac{\partial f}{\partial\theta} - \hat{\theta}\sin\theta\cos\theta\frac{\partial f}{\partial r} + \hat{\theta}\frac{\sin^2\theta}{r}\frac{\partial f}{\partial\theta} \\
&\quad + \hat{r}\sin^2\theta\frac{\partial f}{\partial r} + \hat{r}\frac{\sin\theta\cos\theta}{r}\frac{\partial f}{\partial\theta} + \hat{\theta}\cos\theta\sin\theta\frac{\partial f}{\partial r} + \hat{\theta}\frac{\cos^2\theta}{r}\frac{\partial f}{\partial\theta} \\
&= \hat{r}\cos^2\theta\frac{\partial f}{\partial r} + \hat{\theta}\frac{\sin^2\theta}{r}\frac{\partial f}{\partial\theta} + \hat{r}\sin^2\theta\frac{\partial f}{\partial r} + \hat{\theta}\frac{\cos^2\theta}{r}\frac{\partial f}{\partial\theta} \\
&= \hat{r}\left(\cos^2\theta\frac{\partial f}{\partial r} + \sin^2\theta\frac{\partial f}{\partial r}\right) + \hat{\theta}\left(\frac{\sin^2\theta}{r}\frac{\partial f}{\partial\theta} + \frac{\cos^2\theta}{r}\frac{\partial f}{\partial\theta}\right) \\
&= \hat{r}\frac{\partial f}{\partial r}\left(\cos^2\theta + \sin^2\theta\right) + \hat{\theta}\frac{1}{r}\frac{\partial f}{\partial\theta}\left(\sin^2\theta + \cos^2\theta\right) \\
&= \hat{r}\frac{\partial f}{\partial r} + \hat{\theta}\frac{1}{r}\frac{\partial f}{\partial\theta}
\end{aligned}$$

Elliptical geometry

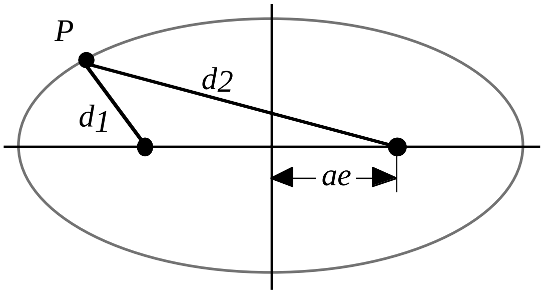
Consider an ellipse centered at the origin whose ***semi-major axis*** a is aligned with the x-axis and whose ***semi-minor axis*** b is aligned with the y-axis, as illustrated in the following figure.



Geometrically , an ellipse is defined as the locus of all points such that the sum of the distances from two particular points, the ***foci*** (singular: ***focus***) is a constant, i.e.,

$$d_1 + d_2 = C \tag{1}$$

in the following figure.

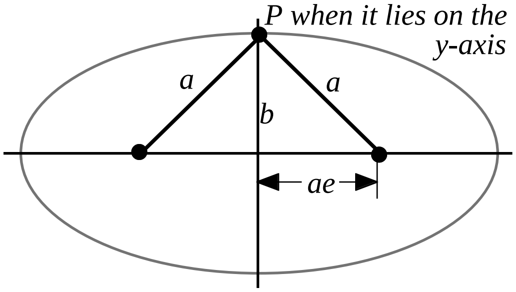


The ***eccentricity***, e, is defined so that each focus lies a distance ae from the origin.

By symmetry, when P lies along the x-axis, we see that

$$d_1 + d_2 = C = 2a \tag{2}$$

Similarly, when P lies along the y-axis, we have $d_1 = d_2$, and hence by (2) the distance from each focus to the intersection of the ellipse with the y-axis is a, as illustrated below.



By the Pythagorean theorem, we have

$$a^2 = b^2 + a^2e^2 \tag{3}$$

Subtracting b^2 from both sides of the equation,

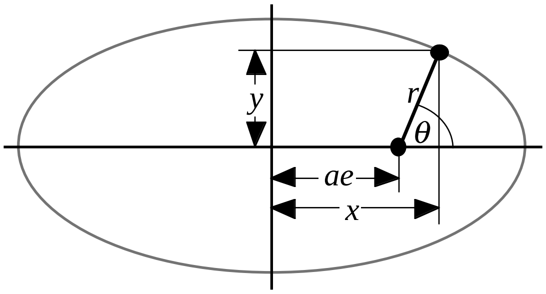
$$a^2 - b^2 = a^2e^2 \tag{4}$$

Solving for the eccentricity,

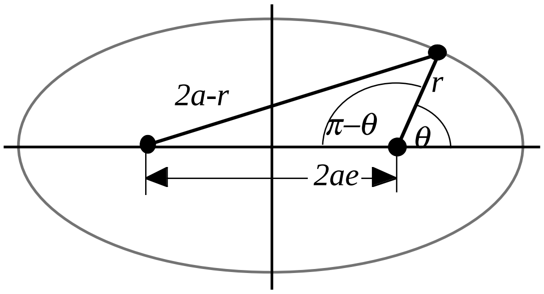
$$e^2 = \frac{a^2 - b^2}{a^2} \quad \text{or} \quad e = \sqrt{1 - (b/a)^2} \tag{5}$$

Equation (5) is sometimes taken as the definition of the eccentricity.

Now lets consider a second coordinate system, this one with origin at one of the foci. We will arbitrarily pick the focus on the right. We want to find an equation for the ellipse in polar coordinates about this focus.



We will do this by constructing the triangle illustrated in the following figure.



By the law of cosines, we have

$$(2a - r)^2 = r^2 + (2ae)^2 - 2(r)(2ae)\cos(\pi - \theta) \tag{6}$$

Using the fact that $\cos(\pi - \theta) = -\cos\theta$ and expanding all the squares,

$$4a^2 - 4ar + r^2 = r^2 + 4a^2e^2 + 4aer\cos\theta \tag{7}$$

Canceling the common term of r^2 on both sides of the equation gives

$$4a^2 - 4ar = 4a^2e^2 + 4aer\cos\theta \tag{8}$$

Equation (8) has a common factor $4a$ which can be factored and then canceled, giving

$$a - r = ae^2 + er\cos\theta \tag{9}$$

Collecting all of the terms that contain an *r* on the left hand side of the equation and collecting the remaining terms on the right hand side of the equation gives us

$$-er\cos\theta - r = ae^2 - a \tag{10}$$

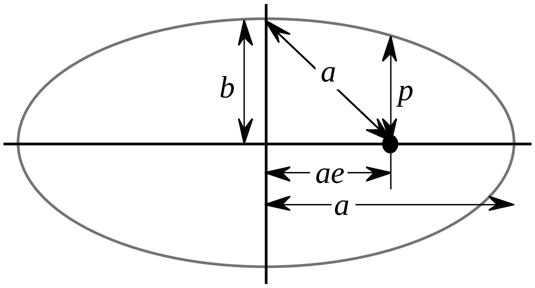
Solving for *r* we have

$$r = \frac{a(1-e^2)}{1+e\cos\theta} = \frac{p}{1+e\cos\theta} \tag{11}$$

The quantity

$$p = a(1-e^2) \tag{12}$$

is called the ***semiparameter*** of the ellipse. From equation (11) we see that when $\theta = \pi / 2$ we have $r = p$, as illustrated in the following figure.



Using similar arguments, and taking advantage of equation (5), it is thence possible to derive the usual equation of an ellipse,

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \tag{13}$$

where x and y are as illustrated in the figure following equation (5), although this derivation is somewhat lengthy algebraically. Equation (13) is almost never used in the study of Keplerian motion; instead, equation (11) is taken as the standard equation of an ellipse.

Newton’s Law of Gravity Leads to Elliptical Orbits

According to ***Newton’s Law of Gravity***, two force between two bodies of mass m_1 and m_2 is proportional to the inverse of the square of the distance between them, is attractive, and is exerted on the line between them. Put algebraically, if the two bodies are at positions $\vec{r}_1$ and $\vec{r}_2$, respectively, then the force on the first body is given by

$$m_1\vec{r}_1'' = -\frac{Gm_1m_2}{\left|\vec{r}_1-\vec{r}_2\right|^2}\frac{\vec{r}_1-\vec{r}_2}{\left|\vec{r}_1-\vec{r}_2\right|} = -\frac{Gm_1m_2\left(\vec{r}_1-\vec{r}_2\right)}{\left|\vec{r}_1-\vec{r}_2\right|^3} \tag{1}$$

while the force on the second body is given by an identical equation with the indices reversed,

$$m_2\vec{r}_2'' = -\frac{Gm_1m_2}{\left|\vec{r}_2-\vec{r}_1\right|^2}\frac{\vec{r}_2-\vec{r}_1}{\left|\vec{r}_2-\vec{r}_1\right|} = -\frac{Gm_1m_2\left(\vec{r}_2-\vec{r}_1\right)}{\left|\vec{r}_1-\vec{r}_2\right|^3} \tag{2}$$

The number G is a constant, call ***Newton’s Universal Gravitational Constant***. The second fraction in the center term of equation (1) represents a vector pointing from $\vec{r}_2$ to $\vec{r}_1$**,** while the corresponding term in equation (2) represents the same vector pointing in the opposite direction. Dividing the first equation by m_1 and the second equation by m_2 gives

$$\vec{r}_1'' = -\frac{Gm_2\left(\vec{r}_1-\vec{r}_2\right)}{\left|\vec{r}_1-\vec{r}_2\right|^3} \tag{3}$$

$$\vec{r}_2'' = -\frac{Gm_1\left(\vec{r}_2-\vec{r}_1\right)}{\left|\vec{r}_1-\vec{r}_2\right|^3} \tag{4}$$

Subtracting equation (4) from equation (1),

$$\vec{r}_1''-\vec{r}_2'' = -\frac{Gm_2\left(\vec{r}_1-\vec{r}_2\right)}{\left|\vec{r}_1-\vec{r}_2\right|^3}+\frac{Gm_1\left(\vec{r}_2-\vec{r}_1\right)}{\left|\vec{r}_1-\vec{r}_2\right|^3} = -\frac{G(m_2+m_1)\left(\vec{r}_1-\vec{r}_2\right)}{\left|\vec{r}_1-\vec{r}_2\right|^3} \tag{5}$$

If we define a new vector

$$\vec{r} = \vec{r}_1-\vec{r}_2 \tag{6}$$

then equation (5) becomes

$$\vec{r}'' = -\frac{G(m_2+m_1)\vec{r}}{r^3} \tag{7}$$

It is typical to define the constant

$$\mu = G(m_1+m_2) \tag{8}$$

When we are dealing with an artificial satellite orbiting the earth, we can take

$$m_1 = m_{Earth} \approx 5.9742\times10^{24}\,kg \tag{9}$$

$$m_2 = m_{satellite} \tag{10}$$

Typical satellite weights range from ≈100 *kg* (light) to ≈10,000 *kg* (heavy), so we can usually assume that

$$\mu = Gm_{earth} \approx 3.986 \times 10^{14} m^3 / sec^2 \tag{11}$$

The ***fundamental equation of motion*** is then

$$\vec{r}'' = -\mu\vec{r} / r^3 = -\mu\hat{r} / r^2 \tag{12}$$

where $\hat{r}$ is a unit vector in the direction $\vec{r}$.

Taking the cross product of (12) with $\vec{r}$ gives

$$\vec{r} \times \vec{r}'' = \vec{r} \times \left(\frac{-\mu\vec{r}}{r^3} \right) = -\frac{\mu}{r^3} \vec{r} \times \vec{r} = 0 \tag{13}$$

because the cross product of a vector with itself is zero. Furthermore, by the product rule,

$$\frac{d}{dt}(\vec{r} \times \vec{r}') = \vec{r} \times \vec{r}'' + \vec{r}' \times \vec{r}' = 0 + 0 = 0 \tag{14}$$

where the first zero follows from equation (13) and the second zero follows because the cross product of a vector with itself is zero. Whenever the derivative of a function is zero, then that function must be constant. Thus equation (14) tells us that $(\vec{r} \times \vec{r}')$ is a constant, i.e., it will not change as a function to time. Such a constant is called ***a constant of motion***; we will call this particular one $\vec{h}$**,**

$$\vec{h} = \vec{r} \times \vec{r}' \tag{15}$$

The constant $\vec{h}$ is called ***the angular momentum per unit mass***. Equation (15) is called the ***Law of Conservation of Angular Momentum***.

Taking the cross product of the fundamental equation of motion (12) with $\vec{h}$ gives

$$\vec{r}'' \times \vec{h} = -\mu \frac{\vec{r}}{r^3} \times (\vec{r} \times \vec{r}') = -\frac{\mu}{r^3} \vec{r} \times (\vec{r} \times \vec{r}') \tag{16}$$

Using the ***vector triple product*** identity,

$$\vec{a} \times (\vec{b} \times \vec{c}) = (\vec{a} \cdot \vec{c})\vec{b} - (\vec{a} \cdot \vec{b})\vec{c} \tag{17}$$

equation (16) can be rewritten as

$$\vec{r}'' \times \vec{h} = -\frac{\mu}{r^3} \vec{r} \times (\vec{r} \times \vec{r}') = -\frac{\mu}{r^3} [(\vec{r} \cdot \vec{r}')\vec{r} - (\vec{r} \cdot \vec{r})\vec{r}'] \tag{18}$$

Since

$$\vec{r} \times \vec{r} = r^2 \tag{19}$$

this becomes

$$\vec{r}'' \times \vec{h} = -\frac{\mu}{r^3} \left[(\vec{r} \cdot \vec{r}')\vec{r} - r^2 \vec{r}' \right] \tag{20}$$

By the quotient rule,

$$\frac{d}{dt}\left(\frac{\vec{r}}{r} \right) = \frac{r\vec{r}' - \vec{r}r'}{r^2} = \frac{1}{r^3} \left[r^2 \vec{r}' - r\vec{r}r' \right] \tag{21}$$

But

$$\hat{r}\cdot\vec{r}' = \text{component of } \frac{d}{dt}\vec{r} \text{ parallel to } \vec{r} = \frac{dr}{dt} = r' \tag{22}$$

Thus (multiply equation (22) through by *r*)

$$\vec{r}\cdot\vec{r}' = r\hat{r}\cdot\vec{r}' = rr' \tag{23}$$

Using (23) in (21) tells us that

$$\frac{d}{dt}\left(\frac{\vec{r}}{r}\right) = \frac{1}{r^3}\left[r^2\vec{r}' - \vec{r}(\vec{r}\cdot\vec{r}')\right] \tag{24}$$

Comparing (20) with (24) gives

$$\vec{r}''\times\vec{h} = \mu\frac{d}{dt}\left(\frac{\vec{r}}{r}\right) \tag{25}$$

Writing $\vec{r}'' = \frac{d}{dt}\vec{r}'$ equation (25) becomes

$$\left(\frac{d}{dt}\vec{r}'\right)\times\vec{h} = \mu\frac{d}{dt}\left(\frac{\vec{r}}{r}\right) \tag{26}$$

Since

$$\vec{a}\times\vec{b} = -\vec{b}\times\vec{a} \tag{27}$$

for any two vectors, we can reverse the order of the cross product in (26)

$$\vec{h}\times\left(\frac{d}{dt}\vec{r}'\right) = -\mu\frac{d}{dt}\left(\frac{\vec{r}}{r}\right) \tag{28}$$

Multiplying both sides of the equation by dt and integrating,

$$\int\vec{h}\times\left(\frac{d}{dt}\vec{r}'\right)dt = -\int\mu\frac{d}{dt}\left(\frac{\vec{r}}{r}\right)dt \tag{29}$$

Bringing the constants outside the integrals

$$\vec{h}\times\int\left(\frac{d}{dt}\vec{r}'\right)dt = -\mu\int\frac{d}{dt}\left(\frac{\vec{r}}{r}\right)dt \tag{30}$$

Since *for any function f(t),*

$$\int\left(\frac{d}{dt}f(t)\right)dt = f(t) + C \tag{31}$$

or in terms of vectors,

$$\int\left(\frac{d}{dt}\vec{f}(t)\right)dt = \vec{f}(t) + \vec{C} \tag{32}$$

Therefore we can evaluate the integrals in (30),

$$-\vec{h}\times\vec{r}' = \mu\frac{\vec{r}}{r} + \vec{C} = \mu\left[\frac{\vec{r}}{r} + \frac{\vec{C}}{\mu}\right] = \mu\left[\frac{\vec{r}}{r} + \vec{e}\right] \tag{33}$$

where

$$\vec{e} = \vec{C} / \mu \tag{34}$$

is a constant of integration. Reversing the cross product in (33),

$$\vec{r}' \times \vec{h} = \mu\left[\frac{\vec{r}}{r} + \vec{e}\right] \tag{35}$$

Taking the dot product of (35) with $\vec{r}$ is

$$(\vec{r}' \times \vec{h}) \cdot \vec{r} = \mu\left[\frac{\vec{r}}{r} + \vec{e}\right] \cdot \vec{r} = \mu\left[\frac{\vec{r} \cdot \vec{r}}{r} + \vec{r} \cdot \vec{e}\right] = \mu(r + \vec{r} \cdot \vec{e}) \tag{36}$$

But from the definition of $\vec{h}$ (see equation 15),

$$(\vec{r}' \times \vec{h}) \cdot \vec{r} = (\vec{r} \times \vec{r}') \cdot \vec{h} = \vec{h} \cdot \vec{h} = h^2 \tag{37}$$

where the first equality follows from the vector identity $(\vec{a} \times \vec{b}) \cdot \vec{c} = (\vec{c} \times \vec{a}) \cdot b$. Using (37) in (36),

$$h^2 = \mu(r + \vec{r} \cdot \vec{e}) \tag{38}$$

Letting θ be the angle between $\vec{r}$ and $\vec{e}$, called the ***true anomaly***, we have

$$h^2 = \mu(r + \vec{r} \cdot \vec{e}) = \mu(r + re\cos\theta) = \mu r(1 + e\cos\theta) \tag{39}$$

Solving for *r,*

$$r = \frac{h^2 / \mu}{1 + e\cos\theta} \tag{40}$$

which is the equation of an ellipse if *e* is the eccentricity and the orbital semi-parameter is

$$p = h^2 / \mu \tag{41}$$

The fact that the orbit is an ellipse – which we have derived from Newton's laws of motion to arrive at equation (40) – is sometimes called ***Kepler's First Law of Motion***.

Since $p = a(1 - e^2)$ equation (41) is equivalent to

$$a(1 - e^2) = h^2 / \mu \tag{42}$$

or

$$h^2 = \mu a(1 - e^2) \tag{43}$$

Equation (43) is true at all points around the orbit; in particular, at perigee ($\theta = 0$) $\vec{r}$ is perpendicular to $\vec{r}'$, and therefore at perigee

$$\vec{h} = \vec{r} \times \vec{r}' = a(1 - e)V_{perigee} \tag{44}$$

where $V_{perigee}$ is the speed at perigee. Using (44) in (43),

$$a^2(1 - e)^2 V_{perigee}^2 = h^2 = \mu a(1 - e^2) \tag{45}$$

Solving for the speed at perigee,

$$V_{perigee}^2 = \frac{\mu}{a}\frac{1-e^2}{(1-e)^2} = \frac{\mu}{a}\frac{(1-e)(1+e)}{(1-e)^2} = \frac{\mu}{a}\left(\frac{1+e}{1-e}\right) \tag{46}$$

Thus

$$V_{perigee} = \sqrt{\frac{\mu}{a}}\sqrt{\frac{1+e}{1-e}} \tag{47}$$

By a similar argument, the radius vector and velocity vector are also perpendicular at apogee (which occurs at $\theta = 180^\circ$, leading to

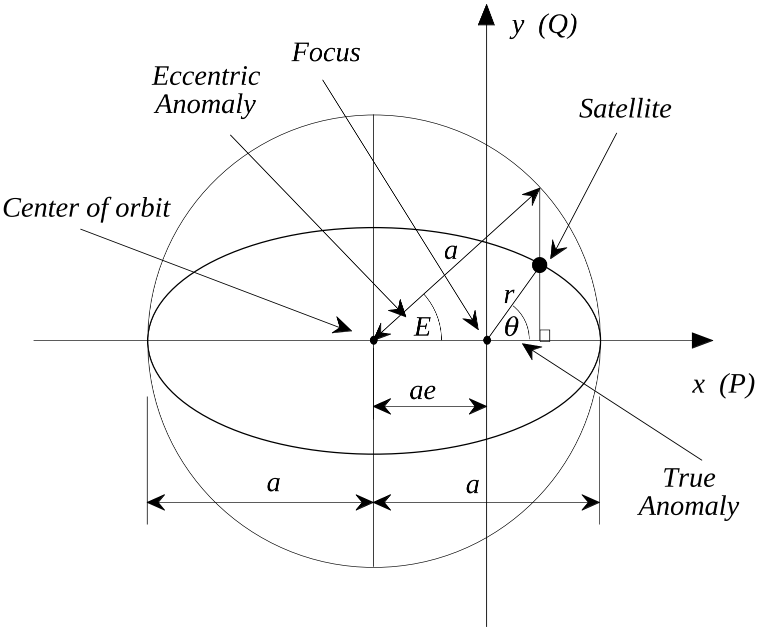
$$V_{apogee} = \sqrt{\frac{\mu}{a}}\sqrt{\frac{1-e}{1+e}} \tag{48}$$

When the eccentricity is zero, the orbit is precisely circular, and

$$V_{circular} = \sqrt{\frac{\mu}{a}} \tag{49}$$

Motion in the plane of the orbit

The motion of a satellite moving in an elliptical orbit is illustrated in the following figure. By definition, the motion is counter-clockwise. The x-axis and y-axis in the figure correspond to the $\vec{P}$ and $\vec{Q}$ axes in the PQW frame



Two special angles are defined, θ, the ***true anomaly***, and E, the ***eccentric anomaly***, as illustrated in the figure. The x and y coordinates of the satellite in this reference frame are

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \end{aligned} \qquad (1)$$

In terms of the eccentric anomaly,

$$a\cos E = ae + x \qquad (2)$$

Solving for x,

$$x = a(\cos E - e) \qquad (3)$$

From the equation of an ellipse, we also know that

$$r = \frac{a(1-e^2)}{1+e\cos\theta} \qquad (4)$$

Using the expression for $\cos\theta = x / r$ from equation (1) in equation (4),

$$r = \frac{a(1-e^2)}{1+ex/r} \qquad (5)$$

Cross-multiplying leads to

$$r + ex = a(1-e^2) \qquad (6)$$

Solving for r,

$$r = a(1-e^2) - ex \qquad (7)$$

Substituting equation (3) into equation (7)

$$\begin{aligned} r &= a(1-e^2) - ea(\cos E - e) \\ &= a - ae^2 - ae\cos E + ae^2 \\ &= a - ae\cos E \\ &= a(1 - e\cos E) \end{aligned} \tag{8}$$

By the Pythagorean theorem,

$$r^2 = x^2 + y^2 \tag{9}$$

so that we can solve for y to get

$$y = \sqrt{r^2 - x^2} \tag{10}$$

Using the last line of equation (8) and equation (3) in equation (10) and simplifying,

$$\begin{aligned} y &= \sqrt{a^2(1-e\cos E)^2 - a^2(\cos E - e)^2} \\ &= a\sqrt{(1-e\cos E)^2 - (\cos E - e)^2} \\ &= a\sqrt{1 - 2e\cos E + e^2\cos^2 E - \cos^2 E + 2e\cos E - e^2} \\ &= a\sqrt{1 + e^2\cos^2 E - \cos^2 E - e^2} \\ &= a\sqrt{1 - \cos^2 E + e^2(\cos^2 E - 1)} \\ &= a\sqrt{(1-e^2)(1-\cos^2 E)} \\ &= a\sqrt{1-e^2}\sin E \end{aligned} \tag{11}$$

because, in the last line, $1 - \cos^2 E = \sin^2 E$.

Summarizing, the coordinates of the satellite in terms of both the true and the eccentric anomaly are

$$\begin{aligned} r &= a(1 - e\cos E) \\ x &= r\cos\theta = a(\cos E - e) \\ y &= r\sin\theta = a(1-e^2)^{1/2}\sin E \end{aligned} \tag{12}$$

Differentiating with respect to time,

$$\frac{dx}{dt} = \frac{d}{dt}\left[a(\cos E - e)\right] = -a\sin E\frac{dE}{dt} \tag{13}$$

$$\frac{dy}{dt} = \frac{d}{dt}\left[a(1-e^2)^{1/2}\sin E\right] = a(1-e^2)^{1/2}\cos E\frac{dE}{dt} \tag{14}$$

$$\frac{dr}{dt} = \frac{d}{dt}\left[a(1-e\cos E)\right] = ae\sin E\frac{dE}{dt} \tag{15}$$

From the defiition of angular momentum,

$$\vec{h} = \vec{r} \times \frac{d\vec{r}}{dt} = \begin{pmatrix} \vec{P} & \vec{Q} & \vec{W} \\ a(\cos E - e) & a(1-e^2)^{1/2}\sin E & 0 \\ -a\sin E\frac{dE}{dt} & a(1-e^2)^{1/2}\cos E\frac{dE}{dt} & 0 \end{pmatrix} \tag{16}$$

Expanding out the cross product,

$$\begin{aligned}\vec{h} &= \left[a^2(\cos E - e)(1-e^2)^{1/2}\cos E\frac{dE}{dt} + a^2(1-e^2)^{1/2}\sin^2 E\frac{dE}{dt}\right]\vec{W}\\ &= a^2(1-e^2)^{1/2}\frac{dE}{dt}\left[(\cos E - e)\cos E + \sin^2 E\right]\vec{W}\\ &= a^2(1-e^2)^{1/2}\frac{dE}{dt}\left[\cos^2 E - e\cos E + \sin^2 E\right]\vec{W}\\ &= a^2(1-e^2)^{1/2}\frac{dE}{dt}[1 - e\cos E]\vec{W}\end{aligned} \tag{17}$$

Thus

$$h^2 = \vec{h}\cdot\vec{h} = a^4(1-e^2)\left(\frac{dE}{dt}\right)^2(1-e\cos E)^2 \tag{18}$$

In an earlier lecture we found that

$$a(1-e^2) = p = h^2/\mu \tag{19}$$

or equivalently,

$$h^2 = \mu a(1-e^2) \tag{20}$$

Equating the two expressions for h^2 (equations 18 and 20),

$$\mu a(1-e^2) = a^4(1-e^2)\left(\frac{dE}{dt}\right)^2(1-e\cos E)^2 \tag{21}$$

Canceling common factors

$$\mu = a^3\left(\frac{dE}{dt}\right)^2(1-e\cos E)^2 \tag{22}$$

Dividing by a^3 and taking the (positive) square root,

$$\frac{\sqrt{\mu}}{a^{3/2}} = (1-e\cos E)\frac{dE}{dt} \tag{23}$$

If we multiply across by dt and integrate, from the time of perigee (time $T_{Perigee}$), to some later time *T*, at which point the eccentric anomaly is assumed to be E_T, we have

$$\frac{\sqrt{\mu}}{a^{3/2}}\int_{T_{perigee}}^{T} dt = \int_0^{E_T}(1-e\cos E)dE \tag{24}$$

Evaluating the integrals,

$$\frac{\sqrt{\mu}}{a^{3/2}}(T - T_{Perigee}) = (E - e\sin E)\Big|_0^{E_T} = E_T - e\sin E_T \tag{25}$$

Using the usual notation of a lower case *t* for time and $E = E(t)$ be the eccentric anomaly,

$$n(t - T_P) = E - e\sin E \tag{26}$$

where

$$n = \frac{\sqrt{\mu}}{a^{3/2}} \tag{27}$$

is called the ***mean motion***.. Rearranging equation (26),

$$t - T_P = \frac{E - e\sin E}{n} \tag{28}$$

If we let τ be the ***period*** (or more technically, the ***anomalistic period***) of the orbit, i.e., the time it takes for the satellite to go from one perigee to the next, then in equation (28), $\tau = t - T_P$ and $E = 2\pi$, so that

$$\tau = \frac{2\pi - e \sin 2\pi}{n} = \frac{2\pi}{n} \tag{29}$$

Solving (29) for $n = 2\pi / \tau$ and substituting it into (27),

$$\frac{2\pi}{\tau} = n = \frac{\sqrt{\mu}}{a^{3/2}} \tag{30}$$

Cross multiplying,

$$2\pi a^{3/2} = \tau \sqrt{\mu} \tag{31}$$

Squaring equation (31),

$$4\pi^2 a^3 = \tau^2 \mu \tag{33}$$

which is sometimes call ***Kepler's Third Law of Motion***: that the square of the period is proportional to the cube of the semi-major axis.

Equation 26 is sometimes used to ***define*** a third angle, *M*, the ***mean anomaly***, as

$$M = n(t - T_P) = E - e \sin E \tag{34}$$

It is important to remember that the Mean Anomaly while it has units of radians, has no geometrical meaning in terms of the picture of the elliptical orbit. What it does have is a physical meaning, given by the first equality of equation (34), that by definition,

$$M = n(t - T_P) \tag{35}$$

The mean anomaly is the equivalent angle that a satellite with the same semi-major axis, but moving in a circle, would move through, in the same amount of time that the satellite moves through an eccentric anomaly *E* starting at perigee. This construction is very useful because *M* increases linearly with time. Dropping the middle of equation (34) we have ***Kepler's Equation***,

$$M = E - e \sin E \tag{36}$$

Using equations (35) and (36) it is possible to predict the position of the satellite at any time *t* as follows:

(1) Use (35) to compute *M*
(2) Given *M*, solve (36) numerically for *E*.
(3) Given *E*, use equations (12) to determine *x* and *y* in the PQW frame
(4) Use a coordinate transformation to convert back to the XYZ frame

Step (2) is the only difficult part because equation (36) can not be solved analytically for *E*. The most common method is to use Newton's Method to find the root of

$$f(E) = E - e \sin E - M \tag{37}$$

According to Newton's method, we first make an educated guess, call it E_1, and then proceed to compute better guesses as

$$E_n = E_{n-1} - \frac{f(E_{n-1})}{f'(E_{n-1})} \tag{38}$$

Differentiating (37), this means

$$E_n = E_{n-1} - \frac{E_{n-1} - e \sin E_{n-1} - M}{1 - e \cos E_{n-1}} \tag{39}$$

A relatively good first guess is to use

$$E_1 = M \tag{40}$$

The Vis-Viva Equation

The potential energy for a satellite of mass *m* in a planetary gravitational field is

$$E_{potential} = -\frac{GMm}{r} = -\frac{\mu m}{r} \tag{1}$$

where *M* is the mass of the planet and $\mu = GM$. The kinetic energy is

$$E_{kinetic} = \frac{1}{2} m V^2 \tag{2}$$

where *V* is the speed of the satellite, i.e., $V^2 = \vec{r}\,' \cdot \vec{r}\,'$ or $V = |\vec{r}\,'|$ (3)

By the ***law of conservation of total energy***, the sum of the kinetic and potential energies does not change as the satellite moves about its orbit:

$$E_{potential} + E_{kinetic} = \frac{1}{2} m V^2 - \frac{\mu m}{r} = Constant \tag{4}$$

What this means is that if r_1 and V_1 are the position and speed of the satellite at one point of the orbit, and if r_2 and V_2 are the position and speed at a second point of the orbit, then

$$\frac{1}{2} m V_1^2 - \frac{\mu m}{r_1} = \frac{1}{2} m V_2^2 - \frac{\mu m}{r_2} \tag{5}$$

Canceling out the common factor of *m*,

$$\frac{1}{2} V_1^2 - \frac{\mu}{r_1} = \frac{1}{2} V_2^2 - \frac{\mu}{r_2} \tag{6}$$

This equation is true for any two points we choose along the orbit. If we let the position and speed at point 1 be *r* and *V* (instead of r_1 and V_1), and if we choose point 2 to be perigee, then equation (6) becomes

$$\frac{1}{2} V^2 - \frac{\mu}{r} = \frac{1}{2} V_{perigee}^2 - \frac{\mu}{r_{perigee}} \tag{7}$$

But we know already that

$$r_{perigee} = a(1-e) \text{ and } V_{perigee} = \sqrt{\frac{\mu}{a}}\sqrt{\frac{1+e}{1-e}} \qquad (8), (9)$$

Substituting equations (8) and (9) into equation (7),

$$\frac{1}{2} V^2 - \frac{\mu}{r} = \frac{1}{2} \frac{\mu}{a} \frac{1+e}{1-e} - \frac{\mu}{a(1-e)} \tag{10}$$

Factoring out the common $\mu / a(1-e)$ on the right hand side of (10)

$$\frac{1}{2} V^2 - \frac{\mu}{r} = \frac{\mu}{a(1-e)}\left[\frac{1}{2}(1+e) - 1\right] = \frac{\mu}{a(1-e)}\left[\frac{1}{2} + \frac{e}{2} - \frac{2}{2}\right] = -\frac{\mu}{2a} \tag{11}$$

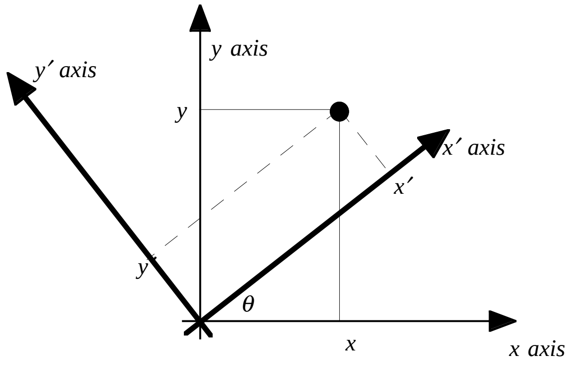
Solving for V^2 we find that

$$V^2 = \mu\left(\frac{2}{r} - \frac{1}{a}\right) \tag{12}$$

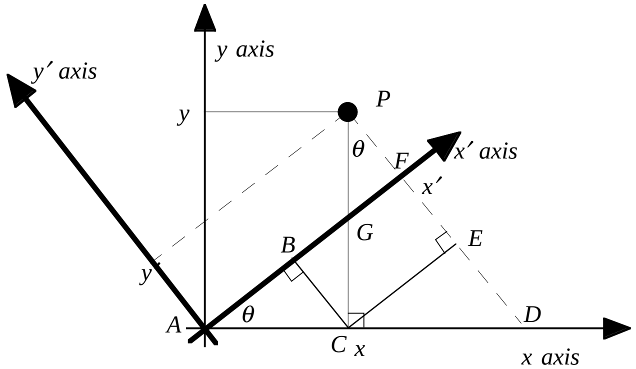
Equation (12) is called the ***Vis-Viva Equation***.

Rotation Matrices

Consider a rotation of the coordinate axes of the x-y plane about the origin by an angle θ. We are interested in finding the coordinates (x', y') of a point as measured in the new coordinate frame.



We determine the transformation by the following constructions:



The new x-coordinate is

$$x' = AB + BF \tag{1}$$

The old x-coordinate

$$x = AC \tag{2}$$

By definition of the cosine,

$$\cos\theta = \frac{AB}{AC} = \frac{AB}{x} \tag{3}$$

so that

$$AB = x\cos\theta \tag{4}$$

Thus (use 4 in 1)

$$x' = x\cos\theta + BF \tag{5}$$

Consider the right triangle *ACG*. Then because all of the angles in a triangle sum to π,

$$angle\ \ AGC = \pi / 2 - \theta \tag{6}$$

Thus in the right triangle *BCG*

$$angle\ \ GCB = \pi / 2 - angle(AGC) = \pi / 2 - (\pi / 2 - \theta) = \theta \tag{7}$$

Since *BC* is parallel to *PD*

$$angle(EPC) = angle(GCB) = \theta \tag{8}$$

From triangle EPC,

$$\sin\theta = \frac{CE}{PC} = \frac{CE}{y} = \frac{BF}{y} \tag{9}$$

Thus

$$BF = y\sin\theta \tag{10}$$

Substituting (10) into (5),

$$x' = x\cos\theta + y\sin\theta \tag{11}$$

Exercise: Show that

$$y' = y\cos\theta - x\sin\theta \tag{12}$$

Combining equations (11) and (12) we have

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x\cos\theta + y\sin\theta \\ y\cos\theta - x\sin\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

The matrix

$$R(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

is called a rotation matrix.

Exercise: Using a geometric argument, what is the rotation matrix that gives (x, y) in terms of (x', y'), i.e., try to find S so that $\begin{pmatrix} x \\ y \end{pmatrix} = S\begin{pmatrix} x' \\ y' \end{pmatrix}$?

Exercise: What is $R(\theta)R(-\theta)$

Exercise: What is R^{-1} (the inverse of $R(\theta)$)?

Now consider a point $P = (x, y, z)$ in 3-dimensional space. Suppose we rotate the x-y plane around the z-axis by an angle q and call the new coordinates (x', y', z'). Since we are rotating about the z-axis, the value of z does not change, and the values of x' and y' are exactly the same as the ones we calculated above! Hence

$$x' = x\cos\theta + y\sin\theta$$

$$y' = y\cos\theta - x\sin\theta$$

$$z' = z$$

We can also write this in matrix form as

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = R_z(\theta)\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

where

$$R_z(\theta) = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

is the rotation matrix for an angle θ about the *z* axis.

Exercise. Convince yourself that the rotation matrices for rotations about the x and y axes are given correctly by

$$R_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}$$

$$R_y(\theta) = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}$$

Now consider the effects of compound rotations: suppose that we first rotate by α about the *z* axis,

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = R_z(\alpha) \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

we then rotate by β about the x' axis, giving us a third set of coordinates (x'', y'', z'').

$$\begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix} = R_{x'}(\beta) \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$$

What is the relationship between the first and third set of coordinates?

$$\begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix} = R_{x'}(\beta) \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = R_{x'}(\beta) R_z(\alpha) \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

In general, the rotation matrix for a any sequence of rotations is just the product of the rotations in the reverse order. Fortunately, due to a theorem by Euler, we rarely have to multiply more than three rotation matrices together.

Euler's Theorem.

(1) The most general rotation of a rigid body (vector) with one point fixed is a rotation about some axis.

(2) Any sequence of rotations about the coordinate axes can be described by a single rotation matrix which is the product of the individual rotation matrices.

(3) Any rotation can be described by a product of at most three rotations about successive coordinate axes. The three angles are called ***Euler Angles***.

Remark. There sequence of rotations is not unique, i.e., you can get the same rotation in several different ways:

(a) A 3-1-3 rotation: rotate by α about *z*, β about x' and γ about z''

(b) A 1-2-3 rotation: rotate by α' about x, β' about y', and γ' about z''

etc.

The Kepler Elements and the PQW Frame

It is necessary to specify an set of initial conditions to completely determine the trajectory of an orbit. Since Newton's Law of gravity

$$m\frac{d^2\vec{r}}{dt^2} = -\frac{\mu\hat{r}}{r^3}$$

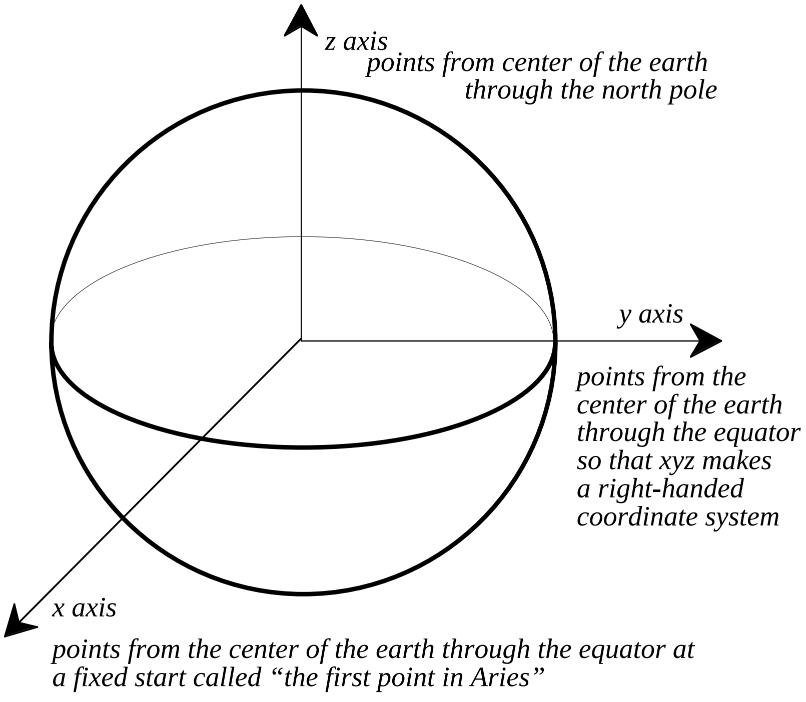
is a second order differential equation in the position, we need to specify both the position and velocity (i.e., both $\vec{r}(0)$ and $\vec{r}\,'(0) = \vec{v}(0)$). Since there are three dimensions, we need to specify a total of six parameters, i.e.,

$$x_0, y_0, z_0, v_{x,0}, v_{y,0}, v_{z,0}$$

where $\vec{r} = x\hat{i} + y\hat{j} + z\hat{k}$ is the position and $\vec{v} = v_x\hat{i} + v_y\hat{j} + v_z\hat{k}$ is the velocity (the subscripts on the velocity refer to the components and are not partial derivatives). Alternatively, it is possible to specify six other numbers that give the same information; for example, we might give the information in polar coordinates.

It turns out that there is a very natural description of the orbit within the orbit plane, that is given in terms of the three parameters a (the semimajor axis), e (the eccentricty), and θ (the true anomaly). Furthermore, we can specify the orientation of the plane of the orbit using three additional angles that we will define shortly.

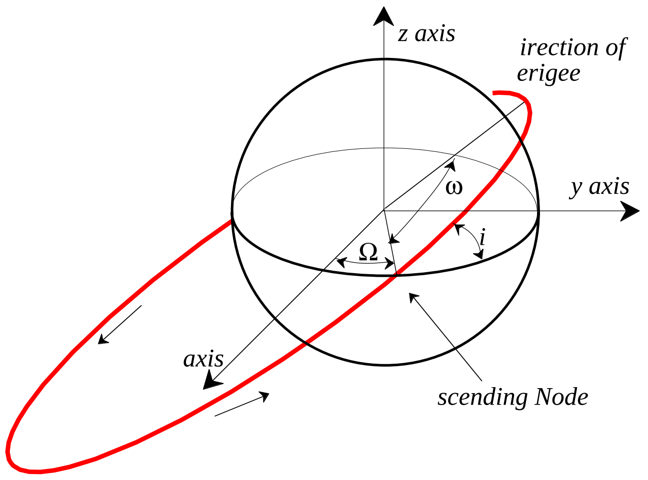
The so called "Earth Centered Inertial" coordinate frame is illustrated in the following figure.



In this coordinate system the *x*-axis points at a fixed spot in space called the ***first point in Aries***, and sometimes called the ***Vernal Equinox***. Technically, the vernal equinox is a point in time, and not a point in space. The reason for the this confusion is that the *x*-axis points from the center of the earth to the sun at the time of the vernal equinox. At this point in time, the sun is entering the constellation Aires (or it least it was some 3000 years ago when the constellations were named, it actually drifts with a period of around 24,000 years). It is usually easier to just think of the x-axis as pointing to a fixed star (which really doesn't exist but we know where, or at least in what direction, it should be).

These axes (of the XYZ frame) are fixed in the center of the Earth but their directions are fixed in space. That means that every point on the Earth (excepting the poles) rotate 360˚ around the *z*-axis every day! Furthermore, although the directions of the axes are fixed, the center of the coordinate system is not, as it moves around the sun every 365 days. Thus the XYZ system is not really inertial ("fixed" in space) but is a moving coordinate system. For this reason it is called an ***Earth Centered Inertial*** coordinate frame – it is sort of fixed – so long as we stay very close to the Earth. When we are dealing with Earth orbiting satellites virtually no accuracy at all is lost by treating this coordinate system as truly fixed in space; however, when studying interplanetary trajectories, we must account for the Earth's motion.

The plane of the orbit intersects the Earth's equator at an angle i, called the ***inclination*** of the orbit.



The orientation of the orbit with respect to the XYZ frame is defined in terms of three angles:

i, the ***inclination***, the angle between the Earth's equation and the plane of the orbit.

Ω, the ***right ascension of ascending node***, the angle measured in the equator between the *x* axis and the intersection of the satellite's orbit and equatorial plane as it moves from the southern hemisphere to the northern hemisphere. This point of intersection is called the ***ascending node*** of the orbit.

ω, the ***argument of perigee***, the angle measured in the plane of the orbit between the ascending node and the perigee.

We can specify the initial conditions in terms of the following set of elements:

$$(a,\ e,\ i,\ \Omega,\ \omega,\ \theta)$$

More commonly the mean anomaly is used instead of the true anomaly:

$$(a,\ e,\ i,\ \Omega,\ \omega,\ M)$$

Either of these sets of elements are called the **Kepler** (or **Keplerian**) **Orbital Elements**.

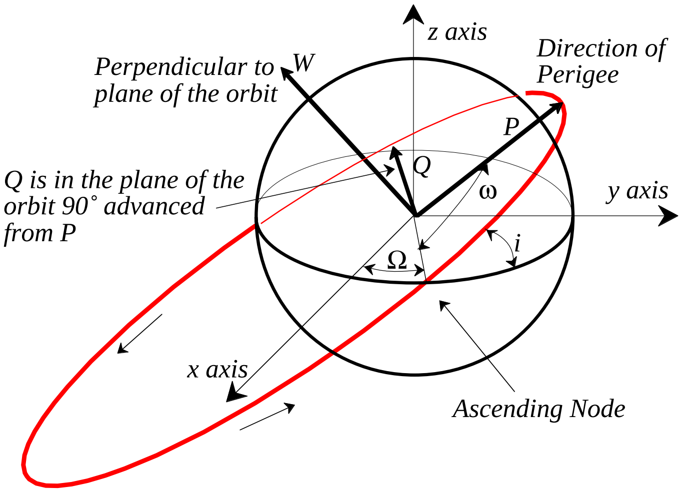
A new frame, called the PQW frame is defined in terms of these angles, as follows:

$\vec{P}$ points from the center of the Earth towards the direction of perigee, in the plane of the orbit.

$\vec{Q}$ lies in the plane of the orbit and points from the center of Earth in a direction 90° advanced from the direction of perigee.

$\vec{W}$ points from the center of the Earth in a direction perpendicular to the plane of the orbit in such a way that the triad $\vec{P},\vec{Q},\vec{W}$ forms a right-handed coordinate system, i.e., $\vec{W} = \vec{P} \times \vec{Q}$

This set of coordinate axes is illustrated in the following figure.



Although they are drawn in the figure with different lengths, **the vectors $\vec{P},\vec{Q},\vec{W}$ are all unit vectors.** We can use the angles shown in the figure to determine the components of the vectors $\vec{P},\vec{Q},\vec{W}$ in terms of the vectors $\vec{i},\vec{j},\vec{k}$.

The easiest way to derive the coordinate transformation is via rotation matrices. To rotate the x axis onto the P axis, while at the same time rotating the y axis onto the Q axis, and the z axis on the W axis, we would perform the following three rotations, in succession:

(1) Rotate by Ω about the z axis.

(2) Rotate by i about the new x axis.

(3) Rotate by ω about the newer z axis.

$$\begin{pmatrix} \vec{P} \\ \vec{Q} \\ \vec{W} \end{pmatrix} = R_Z(\omega) R_X(i) R_Z(\Omega) \begin{pmatrix} \vec{i} \\ \vec{j} \\ \vec{k} \end{pmatrix}$$

$$= \begin{pmatrix} \cos\omega & \sin\omega & 0 \\ -\sin\omega & \cos\omega & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos i & \sin i \\ 0 & -\sin i & \cos i \end{pmatrix} \begin{pmatrix} \cos\Omega & \sin\Omega & 0 \\ -\sin\Omega & \cos\Omega & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \vec{i} \\ \vec{j} \\ \vec{k} \end{pmatrix}$$

$$= \begin{pmatrix} \cos\omega & \sin\omega & 0 \\ -\sin\omega & \cos\omega & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\Omega & \sin\Omega & 0 \\ -\cos i \sin\Omega & \cos i \cos\Omega & \sin i \\ \sin i \sin\Omega & -\sin i \cos\Omega & \cos i \end{pmatrix} \begin{pmatrix} \vec{i} \\ \vec{j} \\ \vec{k} \end{pmatrix}$$

$$= \begin{pmatrix} \cos\omega\cos\Omega - \sin\omega\cos i \sin\Omega & \cos\omega\sin\Omega + \sin\omega\cos i\cos\Omega & \sin\omega\sin i \\ -\sin\omega\cos\Omega - \cos\omega\cos i\sin\Omega & -\sin\omega\sin\Omega + \cos\omega\cos i\cos\Omega & \sin i\cos\Omega \\ \sin i\sin\Omega & -\sin i\cos\Omega & \cos i \end{pmatrix} \begin{pmatrix} \vec{i} \\ \vec{j} \\ \vec{k} \end{pmatrix}$$

Reading off each component, we have

$$P_X = \cos\omega\cos\Omega - \sin\omega\cos i\sin\Omega$$
$$P_y = \cos\omega\sin\Omega + \sin\omega\cos i\cos\Omega$$
$$P_Z = \sin\omega\sin i$$

$$Q_X = -\sin\omega\cos\Omega - \cos\omega\cos i\sin\Omega$$
$$Q_y = -\sin\omega\sin\Omega + \cos\omega\cos i\cos\Omega$$
$$Q_Z = \sin i\cos\Omega$$

$$W_X = \sin i\sin\Omega$$
$$W_y = -\sin i\cos\Omega$$
$$W_Z = \cos i$$

Recipes for Conversion of Orbital Elements

1. Cartesian ⇒ Kepler: Given $\vec{r},\ \vec{v}$, determine $\{a, e, i, \Omega, \omega, M\}$

The alogorithm proceeds by following the numbered equations in order. Some additional text describes where these equations come frome

Eccentricity (e) In our proof that Newton's law of gravity leads to elliptical motion we were left with a vector constant of integration C that points from the center of the earth towards perigee and has magnitude equal to μe

$$\vec{C} = \vec{v} \times \vec{h} - \mu \frac{\vec{r}}{r}$$

We will use this to define the ***eccentricity vector*** ,

$$\vec{e} = \frac{1}{\mu}(\vec{v} \times \vec{h}) - \frac{\vec{r}}{r}$$

Then ***the algorithm to calculate the eccentricity*** is as follows

(1) $\vec{h} = \vec{r} \times \vec{v}$

(2) $\vec{e} = \frac{1}{\mu}(\vec{v} \times \vec{h}) - \frac{\vec{r}}{r}$

(3) $e = \sqrt{\vec{e} \cdot \vec{e}}$

Semi-major Axis (a) Since the orbital semi-parameter p satisfies

$$p = a(1-e^2) = h^2 / \mu$$

we can ***calculate the semi-major axis***,

(4) $a = \dfrac{\vec{h} \cdot \vec{h}}{\mu(1-e^2)}$

Inclination (i)Since the inclination is defined by either of the two (equivalent) defintions as (a) the angle between the orbit plane and the equatorial plane of the earth and (b) the angle between the orbit normal and the *z*-axis of the ECI coordinate system (i.e., the north polar axis), we conclude that since the angular momentum vector $\vec{h}$ is perpendicular to the plane of the orbit , the inclination is also the angle between the *z*-axis and angular momentum vector , so that

$$\hat{k} \cdot \hat{h} = \cos i$$

where the "hat" symbol denotes a unit vector and $\hat{k}$ is the standard unit vector oriented along the *z*-axis. Therefore

(5) $i = \cos^{-1}\left(\dfrac{\hat{k} \cdot \vec{h}}{h}\right)$ where $h = \sqrt{\vec{h} \cdot \vec{h}}$

The inclination is always chosen so that $0 \le i < \pi$,

Right Ascension of Ascending Node (Ω). You should convince yourself that the following ***node vector*** points from the center of the earth to towards the ascending node in the plane of the equator:

(6) $\vec{n} = \hat{k} \times \vec{h}$

and therefore Ω is the angle between $\vec{n}$ and the *x*-axis. Letting $\hat{i}$ be the usual unit vector parallel to the *x*-axis,

(7) $\Omega = \cos^{-1}\left(\dfrac{\hat{i}\cdot\vec{n}}{|\vec{n}|}\right)$ where $|\vec{n}| = \sqrt{\vec{n}\cdot\vec{n}}$

(8) If $\vec{n}\cdot\hat{j} < 0,\ \Omega \to 2\pi - \Omega$ (because the ArcCos function gives an angle between 0 and π; and if the line of nodes has a negative y-component, we need to get an angle between π and 2π.

Argument of Perigee (ω). Since the argument of perigee is defined as the angle between the line of ascending nodes and the perigee, it is also the angle between the vectors $\vec{n}$ and $\vec{e}$, and therefore

(9) $\omega = \cos^{-1}\left(\dfrac{\vec{n}\cdot\vec{e}}{|n|e}\right)$

(10) If $\vec{e}\cdot\hat{k} < 0,\ \omega \to 2\pi - \omega$ because the argument of perigee is between the angles of 0 and π only of the eccentricity vector is in the upper half plane.

True Anomaly (θ) The true anomaly is the angle measured in the orbit plane from perigee; thus it is the angle between the vectors $\vec{r}$ and $\vec{e}$:

(11) $\theta = \cos^{-1}\left(\dfrac{\vec{e}\cdot\vec{r}}{er}\right)$

(12) If $\vec{r}\cdot\vec{v} > 0,\ \theta \to 2\pi - \theta$ (convince yourself that this is true).

Eccentric Anomaly (E) Recall that in our derivation of Kepler's equation we showed that

$$r = a(1 - e\cos E)$$

Also, the equation of an ellipse is polar coordinates is

$$r = \frac{a(1-e^2)}{1+e\cos\theta}$$

Equating these two equations

$$1 - e\cos E = \frac{1-e^2}{1+e\cos\theta}$$

$$e\cos E = 1 - \frac{1-e^2}{1+e\cos\theta} = \frac{1+e\cos\theta - 1 + e^2}{1+e\cos\theta} = \frac{e\cos\theta + e^2}{1+e\cos\theta}$$

$$\cos E = \frac{\cos\theta + e}{1+e\cos\theta}$$

Therefore

(13) $E = \cos^{-1}\left(\dfrac{e+\cos\theta}{1+e\cos\theta}\right)$

(14) If $\pi < \theta < 2\pi,\ E \to 2\pi - E$

Mean Anomaly (M). By Kepler's equation

(15) $M = E - e\cos E$

2. Kepler ⇒ Cartesian: Given $\{a, e, i, \Omega, \omega, M\}$ compute $\vec{r}, \vec{v}$

(1) Determine E by solving Kepler's equation $M = E - e\sin E$ for E. By Newtons' method the procedure is:

$E_0 = M$

$\varepsilon = something\ \ small\ \ like\ \ 10^{-15}$

$i = 0$

Repeat the following:

$$E_{i+1} = E_i - \frac{E_i - e\sin E_i - M}{1 - e\cos E_i}$$

$$i \to i + 1$$

Until $\left|E_i - E_{i-1}\right| < \varepsilon$

E is the final value of E_i that you have calculated

(2) Compute the unit vectors $\vec{P}$ and $\vec{Q}$ along the axes of the PQW frame. From the notes on Kepler Elements, these are defined as

$$P_x = \cos\omega\cos\Omega - \sin\omega\cos i\sin\Omega$$
$$P_y = \cos\omega\sin\Omega + \sin\omega\cos i\cos\Omega$$
$$P_z = \sin\omega\sin i$$

$$Q_x = -\sin\omega\cos\Omega - \cos\omega\cos i\sin\Omega$$
$$Q_y = -\sin\omega\sin\Omega + \cos\omega\cos i\cos\Omega$$
$$Q_z = \sin i\cos\Omega$$

$$\vec{P} = P_x\hat{i} + P_y\hat{j} + P_z\hat{k}$$
$$\vec{Q} = Q_x\hat{i} + Q_y\hat{j} + Q_z\hat{k}$$

(3) In our derivation of Kepler's equation we determined that

$$\vec{r} = a(\cos E - e)\vec{P} + a\sqrt{1 - e^2}\,\sin E\vec{Q}$$
$$\vec{v} = -ae(\sin E)\dot{E}\vec{P} + a\sqrt{1 - e^2}\,(\cos E)\dot{E}\vec{Q}$$

To determine $\dot{E}$ we solve the expression that we had

$$\sqrt{\frac{\mu}{a^3}} = (1 - e\cos E)\dot{E}$$

and therefore

$$\vec{v} = -ae(\sin E)\frac{\sqrt{\mu}}{a^{3/2}(1 - e\cos E)}\vec{P} + a\sqrt{1 - e^2}\,(\cos E)\frac{\sqrt{\mu}}{a^{3/2}(1 - e\cos E)}\vec{Q}$$
$$= \frac{1}{1 - e\cos E}\sqrt{\frac{\mu}{a}}\left(-e\sin E\vec{P} + \sqrt{1 - e^2}\,\cos E\vec{Q}\right)$$