On gradient

Notation: $\Delta(\ )$ denotes the increase of the variable ( ). We omit the higher order terms in $\Delta(\ )$, because they are comparatively neglectable when $\Delta(\ )$ tends to vanish (which is the only case that concerns us).

Let $h\binom{x}{y} = 10-\left(\frac{x^2}{2}+y^2\right)$ be the height of a hill at the point $\binom{x}{y}$ on a map. Then $h\binom{x}{y}$ forms a scalar field on the plane and the figure of $10-\left(\frac{x^2}{2}+y^2\right)=z$ in the xyz coordinate system depicts the hill (See Fig. 1).

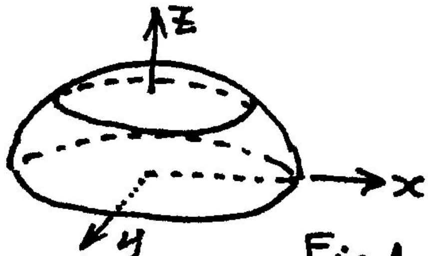

Fig. 1

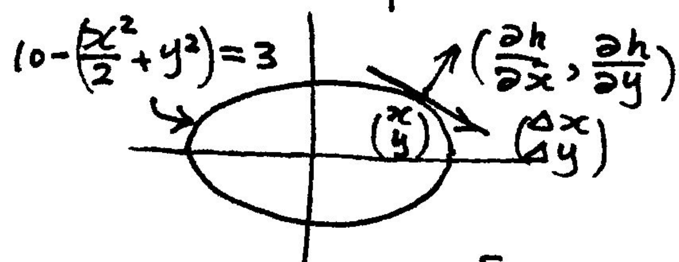

Fig. 2

For a general point $\binom{x}{y}$ on the equal-height curve in Fig. 2:

$$10-\left(\frac{x^2}{2}+y^2\right)=3$$

$$\Delta\left(10-\left(\frac{x^2}{2}+y^2\right)\right) = -x\Delta x - 2y\Delta y = 0 \qquad \frac{\partial h}{\partial x}\Delta x + \frac{\partial h}{\partial y}\Delta y = 0$$

$$\text{i.e. } \begin{pmatrix} -x \\ -2y \end{pmatrix} \cdot \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = 0 \qquad \begin{pmatrix} \frac{\partial h}{\partial x} \\ \frac{\partial h}{\partial y} \end{pmatrix} \cdot \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = 0$$

i.e. $\begin{pmatrix} \frac{\partial h}{\partial x} \\ \frac{\partial h}{\partial y} \end{pmatrix} \perp \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \rightarrow$ having the same direction as the touching line of the equal-height curve at a general point.

Thus $\begin{pmatrix} \frac{\partial h}{\partial x} \\ \frac{\partial h}{\partial y} \end{pmatrix}$ is a vector perpendicular to the (touching line of) equal height curve (at a general point), called the gradient of $h\binom{x}{y}$, denoted by $\begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix} h$, or simply ∇h, where ∇ (read nabla) is the operation $\begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix}$. At every point, the gradient ∇h is a [redacted] vector, and the totality of ∇h is a vector field, which is perpendicular to the equal-height curve at any point.

The function $h\binom{x}{y}$ of two variables has different rate of change on different directions, but the gradient vector ∇h characterizes the rate of change —— ∇h has the direction along which the rate of change of $h\binom{x}{y}$ reaches its maximum, which is the length of ∇h.
(This is intuitively clear, because it has no component on the direction of the touching line, along which the height does not change.)

In fact, the rate of change of $h\binom{x}{y}$ along a given direction $\binom{a}{b}$ is:
(unit vector)

$$\frac{h\binom{x+ta}{y+tb}-h\binom{x}{y}}{t}, \text{ as } t\to 0$$

$=$ the rate of change of $h\binom{x+ta}{y+tb}$ with respect to t, at $t=0$.

$$\begin{tikzcd} & x+ta \arrow[dr] & \\ t \arrow[ur] \arrow[dr, "\text{drives}"'] & & h\binom{x+ta}{y+tb} \\ & y+tb \arrow[ur, "\text{drives}"'] & \end{tikzcd}$$

so

$$\begin{tikzcd} & \Delta(x+ta) \arrow[dr, "\times \left.\frac{\partial h(\cdot\,*)}{\partial\,\cdot}\right|_{t=0} = \frac{\partial h}{\partial x}"] & \\ \Delta t \arrow[ur, "\times a"] \arrow[dr, "\times b"'] & & \Delta h\binom{x+ta}{y+tb} \\ & \Delta(y+tb) \arrow[ur, "\times \left.\frac{\partial h(\cdot\,*)}{\partial *}\right|_{t=0} = \frac{\partial h}{\partial y}"'] & \end{tikzcd}$$

Thus the total increase of $h\binom{x+ta}{y+tb}$ caused by Δt is:

$$\Delta t\times\left(a\cdot\frac{\partial h}{\partial x}+b\cdot\frac{\partial h}{\partial y}\right),$$

and $a\cdot\frac{\partial h}{\partial x}+b\cdot\frac{\partial h}{\partial y}=\begin{pmatrix}\frac{\partial h}{\partial x}\\ \frac{\partial h}{\partial y}\end{pmatrix}\cdot\begin{pmatrix}a\\ b\end{pmatrix}$ is the required rate of change

$=\nabla h\cdot\binom{a}{b}=$ the directional rate of change of h along $\binom{a}{b}$, which reaches its maximum value $|\nabla h|$ when $\binom{a}{b}$ turns to the same direction as ∇h.

Notice that $\nabla h\cdot\binom{a}{b}\Delta\ell$ is the increase of the height h when one moves a distance $\Delta\ell$ along $\binom{a}{b}$ at a general point $\binom{x}{y}$.

$\nabla h\cdot\binom{a}{b}\Delta\ell$

$=\nabla h\cdot\overrightarrow{\Delta\ell}$ (Fig.3)

$\downarrow$ sum up along ℓ from A to B (Fig.3)

$\int_{P\in\ell}\nabla h\cdot\overrightarrow{\Delta\ell}=$ total increase of height from A to B

$=h(B)-h(A)$

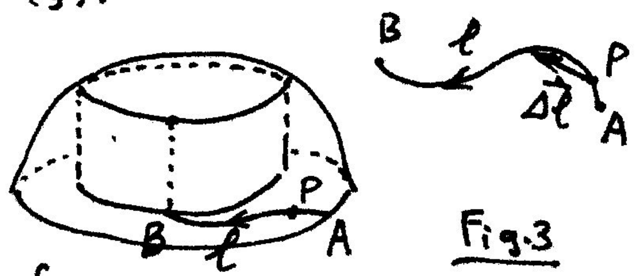

Fig.3

For any oriented curve ℓ connecting from A to B, the above sum always assumes the same value —— the difference of height from A to B. (When one climbs the hill, the change of height does not depend on the path.)

In particular, the height at a point B could be expressed by the vector field ∇h as a sum:

$$h(B) = \int_{P\in\ell} \nabla h \cdot \vec{\Delta\ell} \quad \cdots\cdots (*)$$ where the fixed initial point A is of height 0.

An analogous situation is the work done by the attractive force of the sun from A to B (Fig.4): $\int_{P\in\ell} \vec{G}\cdot\vec{\Delta\ell}$, which also does not depend on the oriented path ℓ from A to B. —— it is determined only by the position of the two extremities of the path (See Appendix I).

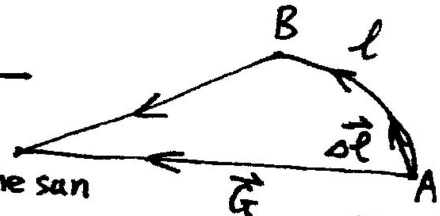

Fig.4.

Thus, just like the equality $(*)$ express the height in terms of the vector field ∇h, a scalar field ψ could be defined from the vector field $\vec{G}$:

$$\psi(B) = \int_{P\in\ell} \vec{G}\cdot\vec{\Delta\ell}$$ where ℓ is an oriented path connecting the fixed initial point A to the point B. $(\psi(A)=0)$.

By the correspondance, we should have $\vec{G} = \nabla\psi$, which could be verified by computation.

The scalar function $\psi(B)$ is called the potential function of $\vec{G}$, because an object situating at B would have a potential energy $\psi(B)$ —— when the object moves to the base point A, the energy of it would increase by $\psi(B)$, due to the work done by the force $\vec{G}$ on it.

Appendix I. The work done by the attractive force of the sun.

In Fig.5: work $= \int_{\substack{P \text{ runs} \\ \text{along } \ell}} \vec{G}\cdot\Delta\vec{r} \quad (\vec{\Delta\ell} = \vec{r}(t+\Delta t) - \vec{r}(t) = \Delta\vec{r})$

$(\dagger)$: $\vec{r} = r\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}$

$\Delta\vec{r} = \Delta r\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix} + r\begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix}\Delta\theta$

$\vec{r}\cdot\Delta\vec{r} = r\cdot\Delta r\cdot 1 + 0$

a critical fact, which is obtained by calculus, but not so obviously seen in the geometric relations in Fig.5.

$$= \int_{\cdots} -\frac{GmM}{r^2}\frac{\vec{r}}{r}\cdot\Delta\vec{r}$$

involves a vector variable $\vec{r}$

where r is the magnitude of $\vec{r}$.

both $\vec{r}$ and r vary as time elapses

$$\overset{(\dagger)}{=} \int_{\cdots} -\frac{GmM}{r^2}\Delta r$$

only involves a scalar variable r

$$\propto \int_{\cdots} -\frac{\Delta r}{r^2} = \int_{r_A}^{r_B}\left(-\frac{\Delta r}{r^2}\right) = \frac{1}{r_B} - \frac{1}{r_A}.$$

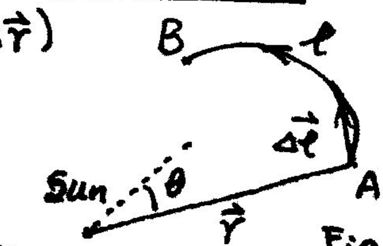

Fig.5