CHAIN RULE

Maths21a, O. Knill

REMEMBER? If f and g are functions of one variable t, then $d/dt f(g(t)) = f'(g(t))g'(t)$. For example, $d/dt \sin(\log(t)) = \cos(\log(t))/t$.

GRADIENT. Define $\nabla f(x,y) = (f_x(x,y), f_y(x,y))$. It is called the gradient of f.

THE CHAIN RULE. If $\vec{r}(t)$ is curve in space and f is a function of three variables, we get a function of one variables $t \mapsto f(\vec{r}(t))$. The **chain rule** is

$$\boxed{d/dt f(\vec{r}(t)) = \nabla f(\vec{r}(t)) \cdot \vec{r}\,'(t)}$$

WRITING IT OUT. Writing the dot product out gives

$$\frac{d}{dt} f(x(t), y(t)) = f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t) \ .$$

EXAMPLE. Let $z = \sin(x + 2y)$, where x and y are functions of t: $x = e^t, y = \cos(t)$. What is $\frac{dz}{dt}$?

Here, $z = f(x,y) = \sin(x+2y)$, $z_x = \cos(x+2y)$, and $z_y = 2\cos(x+2y)$ and $\frac{dx}{dt} = e^t$, $\frac{dy}{dt} = -\sin(t)$ and $\frac{dz}{dt} = \cos(x+2y)e^x - 2\cos(x+2y)\sin(t)$.

EXAMPLE. If f is the temperature distribution in a room and $\vec{r}(t)$ is the path of the spider Shelob, then $f(\vec{r}(t))$ is the temperature, Shelob experiences at time t. The rate of change depends on the velocity $\vec{r}\,'(t)$ of the spider as well as the temperature gradient ∇f and the angle between gradient and velocity. For example, if the spider moves perpendicular to the gradient, its velocity is tangent to a level curve and the rate of change is zero.

EXAMPLE. A nicer spider called "Nabla" moves along a circle $\vec{r}(t) = (\cos(t), \sin(t))$ on a table with temperature distribution $T(x,y) = x^2 - y^3$. Find the rate of change of the temperature, "Nabla" experiences.

SOLUTION. $\nabla T(x,y) = (2x, -3y^2)$, $\vec{r}\,'(t) = (-\sin(t), \cos(t))$ $d/dt T(\vec{r}(t)) = \nabla T(\vec{r}(t)) \cdot \vec{r}\,'(t) = (2\cos(t), -3\sin(t)^2) \cdot (-\sin(t), \cos(t)) = -2\cos(t)\sin(t) - 3\sin^2(t)\cos(t)$.

APPLICATION ENERGY CONSERVATION. If $H(x,y)$ is the energy of a system, the system moves satisfies the equations $\boxed{x'(t) = H_y, y'(t) = -H_x}$. For example, if $H(x,y) = y^2/2 + V(x)$ is a sum of kinetic and potential energy, then $x'(t) = y, y'(t) = V'(x)$ is equivalent to $x''(t) = -V'(x)$. In the case of the Kepler problem, we had $V(x) = Gm/|x|$. $\boxed{\text{The energy } H \text{ is conserved}}$. Proof. The chain rule shows that $d/dt H(x(t), y(t)) = H_x(x,y)x'(t) + H_y(x,y)y'(t) = H_x(x,y)H_y(x,y) - H_y(x,y)H_x(x,y) = 0$.

APPLICATION: IMPLICIT DIFFERENTIATION.
From $f(x,y) = 0$ one can express y as a function of x. From $d/df(x, y(x)) = \nabla f \cdot (1, y'(x)) = f_x + f_y y' = 0$ we obtain $\boxed{y' = -f_x/f_y}$.

EXAMPLE. $f(x,y) = x^4 + x\sin(xy) = 0$ defines $y = g(x)$. If $f(x, g(x)) = 0$, then $g_x(x) = -f_x/f_y = -(4x^3 + \sin(xy) + xy\cos(xy))/(x^2\cos(xy))$.

APPLICATION: DIFFERENTIATION RULES. One ring of the chain, rules them all!
$f(x,y) = x + y, x = u(t), y = v(t), d/dt(x+y) = f_x u' + f_y v' = u' + v'$.
$f(x,y) = xy, x = u(t), y = v(t), d/dt(xy) = f_x u' + f_y v' = vu' + uv'$.
$f(x,y) = x/y, x = u(t), y = v(t), d/dt(x/y) = f_x u' + f_y v' = u'/y - v'u/v^2$.

DIETERICI EQUATION. In thermodynamics the temperature T, the pressure p and the volume V of a gase are related. One refinement of the ideal gas law $pV = RT$ is the **Dieterici equation** $f(p,V,T) = p(V-b)e^{a/RVT} - RT = 0$. The constant b depends on the volume of the molecules and a depends on the interaction of the molecules. (A different variation of the ideal gas law is van der Waals law). Problem: compute V_T.

If $V = V(T,p)$, the chain rule says $f_V V_T + f_T = 0$, so that $\boxed{V_T = -f_T/f_V} = -(-ap(V-b)e^{a/RVT}/(RVT^2) - R)/(pVe^{a/RVT} - p(V-b)e^{a/RVT}/(RV^2T))$. (This could be simplified to $(R + a/TV)/(RT/(V-b) - a/V^2)$).

DERIVATIVE. If $f : \mathbf{R}^n \to \mathbf{R}^m$ is a map, its **derivative** f' is the $m \times n$ matrix $[f']_{ij} = \frac{\partial}{\partial x_j} f_i$.

MORE DERIVATIVES. (The last three derivatives will only appear later)

$f : \mathbf{R} \to \mathbf{R}^3$	$f : \mathbf{R}^3 \to \mathbf{R}$	$f : \mathbf{R}^2 \to \mathbf{R}^3$	$f : \mathbf{R}^2 \to \mathbf{R}^2$	$f : \mathbf{R}^3 \to \mathbf{R}^3$
curve	scalar function	surface	coordinate change	gradient field
f' velocity vector.	f' gradient vector.	f' tangent matrix	f' Jacobean matrix.	f' Hessian matrix.

THE GENERAL CHAIN RULE. (for people who have seen some linear algebra). First of all, the $d/dt f(\vec{r}(t)) = \nabla f(r(t)) \cdot \vec{r}\,'(t)$ is true in any dimension. If f is vector valued, the same equation holds for each component $d/dt f_i(\vec{r}(t)) = \nabla f_i(r(t)) \cdot \vec{r}\,'(t)$. One can further assume that $\vec{r}$ depends on different variables. Then this formula holds for each variable x_i. Here is the general chain rule for the curious: If $f : \mathbf{R}^n \to \mathbf{R}^m$ and $g : \mathbf{R}^k \to \mathbf{R}^n$, we can compose $f \circ g$, which is a map from $\mathbf{R}^k$ to $\mathbf{R}^m$. The chain rule expresses the derivative of $f \circ g(x) = f(g(x))$ in terms of the derivatives of f and g.

$$\boxed{\frac{\partial}{\partial x_j} f(g(x))_i = \sum_k \frac{\partial}{\partial x_k} f_i(g(x)) \frac{\partial}{\partial x_j} g_k(x)} \quad \text{or short} \quad \boxed{(f \circ g)'(x) = f'(g(x)) \cdot g'(x)}$$

Both $f'(g(x))$ and $g'(x)$ are matrices and $\cdot$ is the matrix multiplication. The chain rule in higher dimensions looks like the chain rule in one dimension, only that the objects are matrices and the multiplication is matrix multiplication.

EXAMPLE. GRADIENT IN POLAR COORDINATES. In polar coordinates, the gradient is defined as $\nabla f = (f_r, f_\theta/r)$. Using the chain rule, we can relate this to the usual gradient: $d/dr f(x(r,\theta), y(r,\theta)) = f_x(x,y)\cos(\theta) + f_y(x,y)\sin(\theta)$ and $d/(rd\theta) f(x(r,\theta), y(r,\theta)) = -f_x(x,y)\sin(\theta) + f_y(x,y)\cos(\theta)$ means that the length of ∇f is the same in both coordinate systems.

PROOFS OF THE CHAIN RULE.

1. Proof. Near any point, we can approximate f by a linear function L. It is enough to check the chain rule for linear functions $f(x,y) = ax + by - c$ and if $\vec{r}(t) = (x_0 + tu, y_0 + tv)$ is a line. Then $\frac{d}{dt} f(\vec{r}(t)) = \frac{d}{dt}(a(x_0 + tu) + b(y_0 + tv)) = au + bv$ and this is the dot product of $\nabla f = (a,b)$ with $\vec{r}\,'(t) = (u,v)$.

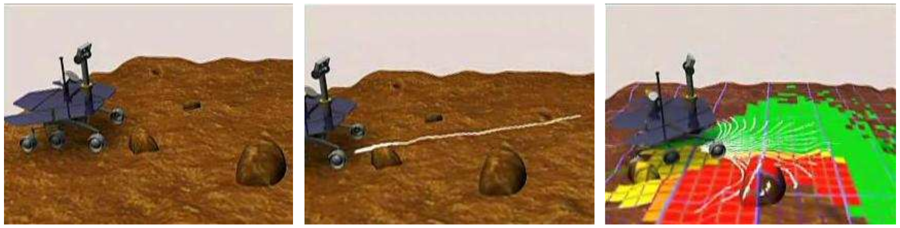
2. Proof. Plugging in the definitions of the derivatives and use limits.

WHERE IS THE CHAIN RULE NEEDED? (informal).

While the chain rule is useful in calculations using the composition of functions, the iteration of maps or in doing change of variables, it is also useful for **understanding** some theoretical aspects. Examples:

- In the proof of the fact that **gradients are orthogonal to level surfaces**. (see Wednesday).
- It appears in **change of variable** formulas.
- It will be used in the **fundamental theorem for line integrals** coming up later in the course.
- The chain rule illustrates also the **Lagrange multiplier** method which we will see later.
- In **fluid dynamics**, PDE's often involve terms $u_t + u\nabla u$ which give the change of the velocity in the frame of a fluid particle.
- In **chaos theory**, where one wants to understand what happens after iterating a map f.

ICE Chain rule Maths21a



The Mars surface on which "opportunity" drives has the height $f(x,y) = x + (2x^2 + 3y^2 - xy)$. The rover moves along the path $\vec{r}(t) = \langle (1+t), \sin(t) \rangle$.

Find the rate of change of the height $\frac{d}{dt} f(\vec{r}(t))$ at the point $t = 0$ by differentiating the function $t \mapsto f(\vec{r}(t))$.

Find the rate of change of the height $\frac{d}{dt} f(\vec{r}(t))$ at the point $t = 0$ using the chain rule.

GRADIENT AND DIRECTIONAL DERIVATIVE — Math21a

GRADIENT. Define the **gradient** $\nabla f(x,y) = (f_x(x,y), f_y(x,y))$ or $\nabla f(x,y,z) = (f_x(x,y,z), f_y(x,y,z), f_z(x,y,z))$ in three dimensions.

CHAIN RULE. The chain rule in multivariable calculus is $\boxed{\frac{d}{dt} f(\vec{r}(t)) = \nabla f(\vec{r}(t)) \cdot \vec{r}'(t)}$ It looks like the 1D chain rule, but the derivative f' is replaced with the gradient and the derivative of r is the velocity. Written out in detail the chain rule is

$$\frac{d}{dt} f(x(t),y(t),z(t)) = f_x(x(t),y(t),z(t))x'(t) + f_y(x(t),y(t),z(t))y'(t) + f_z(x(t),y(t),z(t))z'(t) \ .$$

DIRECTIONAL DERIVATIVE. If f is a function of several variables and $\vec{v}$ is a vector, then $\nabla f \cdot \vec{v}$ is called the **directional derivative** of f in the direction $\vec{v}$. One writes $\nabla_{\vec{v}} f$ or $D_{\vec{v}} f$.

$$\boxed{D_v f(x,y,z) = \nabla f(x,y,z) \cdot v}$$

Unlike done in some calculus books, we do not insist that $\vec{v}$ is a unit vector. The chain rule gives $\frac{d}{dt} D_{\vec{v}} f = \frac{d}{dt} f(x + t\vec{v})$.

EXAMPLE. PARTIAL DERIVATIVES ARE SPECIAL DIRECTIONAL DERIVATIVES.

If $\vec{v} = (1,0,0)$, then $D_{\vec{v}} f = \nabla f \cdot v = f_x$.
If $\vec{v} = (0,1,0)$, then $D_{\vec{v}} f = \nabla f \cdot v = f_y$.
If $\vec{v} = (0,0,1)$, then $D_{\vec{v}} f = \nabla f \cdot v = f_z$.

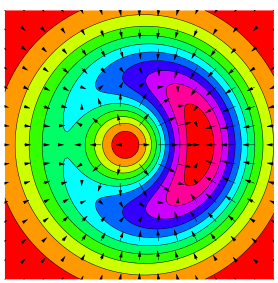
The directional derivative is a generalization of the partial derivatives. Like the partial derivatives, it is a **scalar**.

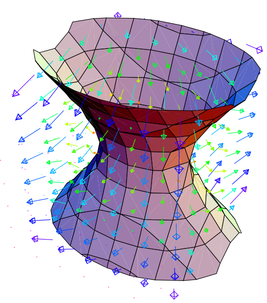
EXAMPLE. DIRECTIONAL DERIVATIVE ALONG A CURVE.
If f is the temperature in a room and $\vec{r}(t)$ is a curve with velocity $\vec{r}\,'(t)$, then $\nabla f(\vec{r}(t)) \cdot \vec{r}\,'(t)$ is the temperature change, one measures on the point moving on a curve $\vec{r}(t)$ experiences: the chain rule told us that this is $d/dt f(\vec{r}(t))$.

GRADIENTS AND LEVEL CURVES/SURFACES.

Gradients are orthogonal to level curves and level surfaces.

Every curve $\vec{r}(t)$ on the level curve or level surface satisfies $\frac{d}{dt} f(\vec{r}(t)) = 0$. By the chain rule, $\nabla f(\vec{r}(t))$ is perpendicular to the tangent vector $\vec{r}'(t)$.





STEEPEST DECENT. The directional derivative satisfies $\boxed{|D_{\vec{v}} f| \leq |\nabla f||\vec{v}|}$ because $\nabla f \cdot \vec{v} = |\nabla f||\vec{v}||\cos(\phi)| \leq |\nabla f||\vec{v}|$. The direction $\vec{v} = \nabla f$ is the direction, where f **increases** most, the direction $-\nabla f$ is the direction where f **decreases most**. It is the direction of steepest decent.

IN WHICH DIRECTION DOES f INCREASE? If $\vec{v} = \nabla f$, then the directional derivative is $\nabla f \cdot \nabla f = |\nabla f|^2$. This means that f **increases**, if we move into the direction of the gradient!

EXAMPLE. You are on a trip in a air-ship at $(1,2)$ and want to avoid a thunderstorm, a region of low pressure. The pressure is given by a function $p(x,y) = x^2 + 2y^2$. In which direction do you have to fly so that the pressure change is largest?

Parameterize the direction by $\vec{v} = (\cos(\phi), \sin(\phi))$. The pressure gradient is $\nabla p(x,y) = (2x, 4y)$. The directional derivative in the ϕ-direction is $\nabla p(x,y) \cdot v = 2\cos(\phi) + 4\sin(\phi)$. This is maximal for $-2\sin(\phi) + 4\cos(\phi) = 0$ which means $\tan(\phi) = 1/2$.

ZERO DIRECTIONAL DERIVATIVE. The rate of change in all directions is zero if and only if $\nabla f(x,y) = 0$: if $\nabla f \neq \vec{0}$, we can choose $\vec{v} = \nabla f$ and get $D_{\nabla f} f = |\nabla f|^2$.

We will see later that points with $\nabla f = \vec{0}$ are candidates for **local maxima** or **minima** of f. Points (x,y), where $\nabla f(x,y) = (0,0)$ are called **stationary points** or **critical points**. Knowing the critical points is important to understand the function f.

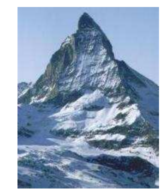
PROPERTIES DIRECTIONAL DERIVATIVE.

$$\begin{aligned} D_v(\lambda f) &= \lambda D_v(f) \\ D_v(f+g) &= D_v(f) + D_v(g) \\ D_v(fg) &= D_v(f)g + f D_v(g) \end{aligned}$$

PROPERTIES GRADIENT

$$\begin{aligned} \nabla(\lambda f) &= \lambda \nabla(f) \\ \nabla(f+g) &= \nabla(f) + \nabla(g) \\ \nabla(fg) &= \nabla(f)g + f\nabla(g) \end{aligned}$$

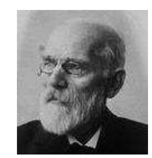
THE MATTERHORN is a popular climbing mountain in the Swiss alps. Its height is 4478 meters (14,869 feet). It is quite easy to climb with a guide. There are ropes and ladders at difficult places. Even so, about 3 people die each year from climbing accidents at the Matterhorn, this does not stop you from trying an ascent. In suitable units on the ground, the height $f(x,y)$ of the Matterhorn is approximated by $f(x,y) = 4000 - x^2 - y^2$. At height $f(-10,10) = 3800$, at the point $(-10,10,3800)$, you rest. The climbing route continues into the north-east direction $v = (1,-1)$. Calculate the rate of change in that direction. We have $\nabla f(x,y) = (-2x,-2y)$, so that $(20,-20) \cdot (1,-1) = 40$. This is a place, with a ladder, where you climb 40 meters up when advancing 1m forward.



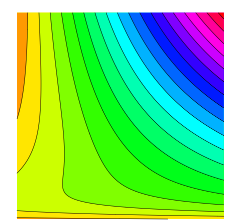
THE VAN DER WAALS (1837-1923) equation for real gases is

$$(p + a/V^2)(V - b) = RT(p,V) \ ,$$

where $R = 8.314 J/K$mol is a constant called the **Avogadro number**. This law relates the pressure p, the volume V and the temperature T of a gas. The constant a is related to the molecular interactions, the constant b to the finite rest volume of the molecules.



The **ideal gas** law $pV = nRT$ is obtained when a, b are set to 0. The level curves or **isotherms** $T(p,V) = const$ tell much about the properties of the gas. The so called **reduced van der Waals law** $T(p,V) = (p + 3/V^2)(3V - 1)/8$ is obtained by scaling p, T, V depending on the gas. Calculate the directional derivative of $T(p,V)$ at the point $(p,V) = (1,1)$ into the direction $v = (1,2)$. We have $T_p(p,V) = (3V-1)/8$ and $T_V(p,V) = 3p/8 - (9/8)1/V^2 - 3/(4V^3)$. Therefore, $\nabla T(1,1) = (1/4, 0)$ and $D_v T(1,1) = 1/5$.



TANGENT LINE. Because $\vec{n} = \nabla f(x_0, y_0) = \langle a, b \rangle$ is perpendicular to the level curve $f(x,y) = c$ through (x_0, y_0), the equation for the tangent line is

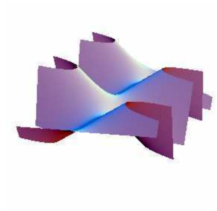
$$\boxed{ax + by = d, \ a = f_x(x_0,y_0), \ b = f_y(x_0,y_0), \ d = ax_0 + by_0}$$

EXAMPLE. The isotherm in the previous example through $(1,1)$ has there the tangent $(1/4)x + 0 \cdot y = (1/4)1 + 0 \cdot 1 = 1/4$ which is the horizontal line $x = 1$.

LINEAR APPROXIMATION — Maths21a, O. Knill

LINEAR APPROXIMATION.

$\boxed{\text{1D:}}$ The **linear approximation** of a function $f(x)$ at a point x_0 is the linear function $\boxed{L(x) = f(x_0) + f'(x_0)(x - x_0)}$. The graph of L is tangent to the graph of f.



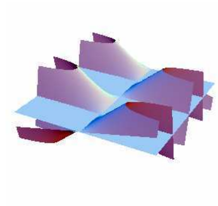
$\boxed{\text{2D:}}$ The **linear approximation** of a function $f(x, y)$ at (x_0, y_0) is

$$\boxed{L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)}$$

The level curve of g is tangent to the level curve of f at (x_0, y_0). The graph of L is tangent to the graph of f.

$\boxed{\text{3D:}}$ The **linear approximation** of a function $f(x, y, z)$ at (x_0, y_0, z_0) by

$$\boxed{L(x, y, z) = f(x_0, y_0, z_0) + f_x(x_0, y_0, z_0)(x - x_0) + f_y(x_0, y_0, z_0)(y - y_0) + f_z(x_0, y_0, z_0)(z - z_0)}.$$



The level surface of L is tangent to the level surface of f at (x_0, y_0, z_0).

In vector form, the linearization can be written as

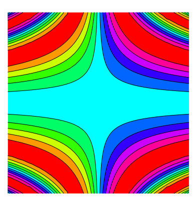
$$\boxed{L(\vec{x}) = f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)}$$

JUSTIFYING THE LINEAR APPROXIMATION. 3 ways to see it:
1) We know the tangent plane to $g(x, y, z) = z - f(x, y)$ at $(x_0, y_0, z + 0 = f(x_0, z_0)$ is $-f_x x - f_y y + z = -f_x x_0 - f_y y_0 + z_0$. This can be read as $z = z_0 + +f_x(x - x_0) + f_y(y - y_0)$. Calling the right hand side $L(x, y)$ shows that the graph of L is tangent to the graph of f at (x_0, y_0).

2) The higher dimensional case can be reduced to the one dimensional case: if $y = y_0$ is fixed and x, then $f(x, y_0) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0)$ is the linear approximation of the function. Similarly, if $x = x_0$ is fixed y is the single variable, then $f(x_0, y) = f(x_0, y_0) + f_y(x_0, y_0)(y - y_0)$. So along two directions, the linear approximations are the best. Together we get the approximation for $f(x, y)$.

3) An other justification is that $\nabla f(\vec{x}_0)$ is orthogonal to the level curve at $\vec{x}_0$. Because $\vec{n} = \nabla f(\vec{x}_0)$ is orthogonal to the plane $\vec{n} \cdot (\vec{x} - \vec{x_0}) = d$ also, the graphs of $L(x, y)$ and $f(x, y)$ have the same normal vector at $(x_0, y_0, f(x_0, y_0)$.

EXAMPLE (2D) Find the linear approximation of the function $f(x, y) = \sin(\pi x y^2)$ at the point $(1, 1)$. The gradient is $\nabla f(x, y) = (\pi y^2 \cos(\pi x y^2), 2y\pi \cos(\pi x y^2))$. At the point $(1, 1)$, we have $\nabla f(1, 1) = (\pi \cos(\pi), 2\pi \cos(\pi)) = (-\pi, 2\pi)$. The linear function approximating f is $L(x, y) = f(1, 1) + \nabla f(1, 1) \cdot (x - 1, y - 1)$ $= 0 - \pi(x - 1) - 2\pi(y - 1) = -\pi x - 2\pi y + 3\pi$. The level curves of G are the lines $x + 2y = const$. The line which passes through $(1, 1)$ satisfies $x + 2y = 3$.



Application: $-0.00943407 = f(1{+}0.01, 1{+}0.01) \sim g(1{+}0.01, 1{+}0.01) = -\pi 0.01 - 2\pi 0.01 + 3\pi = -0.00942478$.

EXAMPLE (3D) Find the linear approximation to $f(x, y, z) = xy + yz + zx$ at the point $(1, 1, 1)$.

We have $f(1, 1, 1) = 3$, $\nabla f(x, y, z) = (y + z, x + z, y + x), \nabla f(1, 1, 1) = (2, 2, 2)$. Therefore $L(x, y, z) = f(1, 1, 1) + (2, 2, 2) \cdot (x - 1, y - 1, z - 1) = 3 + 2(x - 1) + 2(y - 1) + 2(z - 1) = 2x + 2y + 2z - 3$.

EXAMPLE (3D). Use the best linear approximation to $f(x, y, z) = e^x \sqrt{y} z$ to estimate the value of f at the point $(0.01, 24.8, 1.02)$.

Solution. Take $(x_0, y_0, z_0) = (0, 25, 1)$, where $f(x_0, y_0, z_0) = 5$. The gradient is $\nabla f(x, y, z) = (e^x \sqrt{y} z, e^x z/(2\sqrt{y}), e^x \sqrt{y})$. At the point $(x_0, y_0, z_0) = (0, 25, 1)$ the gradient is the vector $(5, 1/10, 5)$. The linear approximation is $L(x, y, z) = f(x_0, y_0, z_0) + \nabla f(x_0, y_0, z_0)(x - x_0, y - y_0, z - z_0) = 5 + (5, 1/10, 5)(x - 0, y - 25, z - 1) = 5x + y/10 + 5z - 2.5$. We can approximate $f(0.01, 24.8, 1.02)$ by $5 + (5, 1/10, 5) \cdot (0.01, -0.2, 0.02) = 5 + 0.05 - 0.02 + 0.10 = 5.13$. The actual value is $f(0.01, 24.8, 1.02) = 5.1306$, very close to the estimate.

SECOND DERIVATIVE.
If $f(x, y)$ is a function of two variables, then the matrix $f''(x, y) = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$ is called the **second derivative** or the **Hessian** of f.

For functions of three variables, the Hessian is the $3x3$ matrix $f''(x, y, z) = \begin{bmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{bmatrix}$ Because for smooth functions, $f_{xy} = f_{yx}, f_{yz} = f_{zy}$, the matrix f'' is **symmetric** (a reflection at the diagonal leaves it invariant).

QUADRATIC APPROXIMATION. (informal)
If F is a function of several variables $\vec{x}$ and $\vec{x}_0$ is a point, then

$$\boxed{Q(\vec{x}) = f(\vec{x}_0) + \nabla f(\vec{x}_0)(\vec{x} - \vec{x}_0) + [f''(\vec{x}_0)(\vec{x} - \vec{x}_0)] \cdot (\vec{x} - \vec{x}_0)/2}$$

is called the **quadratic approximation** of $\vec{x}$.

It generalizes the quadratic approximation $L(x) = f(x_0) + f'(x_0)(x - x_0) + f''(x_0)(x - x_0)^2/2$ of a function of one variables.

EXAMPLE. If the height of a hilly region is given by $f(x, y) = 4000 - \sin(x^2 + y^2)$, find the quadratic approximation of F at $(0, 0)$.
$\nabla f(x, y) = (2x, 2y)\cos(x^2 + y^2)$ so that $\nabla f(0, 0) = (0, 0)$. The linear approximation of F at $(0, 0)$ is $G(x, y) = f(0, 0) = 4000$. The graph of G is a plane tangent to the graph of F.

$$f''(x, y) = \cos(x^2 + y^2) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

so that

$$f''(0, 0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The quadratic approximation at $(0, 0)$ is $Q(x, y) = 4000 - x^2 - y^2$. The graph of F is an inverted paraboloid.

EXAMPLE: REDUCED VAN DER WAALS LAW $T(p, V) = (p + 3/V^2)(3V - 1)/8$ Find the quadratic approximation of T at $(p, V) = (1, 1)$. We had $T_p(p, V) = (3V - 1)/8$ and $T_V(p, V) = 3p/8 - (9/16)1/V^2 - 3/(4V^3)$ and $\nabla T(1, 1) = (1/4, 0)$. Now, $T_{pp}(p, V) = 0$, $T_{pV} = 3/8$, $T_{Vp} = 3/8$ and $T_{VV} = (9/8)1/V^3 + (9/4)1/V^4$ so that $T''(1, 1) = \begin{bmatrix} 0 & 3/8 \\ 3/8 & 27/8 \end{bmatrix}$. The quadratic approximation at $T = (1, 1)$ is $Q(p, T) = T(1, 1) + T'(1, 1) \cdot (p - 1, T - 1) + [T''(1, 1)(p - 1, T - 1)] \cdot (p - 1, T - 1)/2$.

ERROR OF APPROXIMATION. It follows from Taylor's theorem that the error $f(x, y) - L(x, y)$ in a region R near (x_0, y_0) is smaller or equal to $M(|x - x_0| + |y - y_0|)^2/2$, where M is the maximal value of all the matrix entries $f''(x, y)$ in that region R.

TOTAL DIFFERENTIAL. Aiming to estimate the change $\Delta f = f(x, y) - f(x_0, y_0)$ of f for points $(x, y) = (x_0, y_0) + (\Delta x, \Delta y)$ near (x_0, y_0), we can estimate it with the linear approximation which is $L(\Delta x, \Delta y) = f_x(x_0, y_0)\Delta x + f_y \Delta y$. In an old-fashioned notation, one writes also $df = f_x dx + f_y dy$ and calls df the **total differential**. One can **totally avoid** the notation of the **total differential**.

RELATIVE AND ABSOLUTE CHANGE. The **relative change** Δf depends on the magnitude of f. One also defines $\Delta f(x_0, y_0)/f(x_0, y_0)$, the **relative change** of f. A good estimate is $\Delta L(x_0, y_0)/L(x_0, y_0)$.

PHYSICAL LAWS.

Many physical laws are in fact linear approximations to more complicated laws. One could say that a large fraction of physics consists of understand nature with linear laws.

LINEAR STABILITY ANALYSIS.

In physics, complicated situations can occur. Usually, many unknown parameters are present and the only way to analyze the situation theoretically is to assume that things depend linearly on these parameters. The analysis of the linear situation allows then to predict for example the stability of the system with respect to perturbations. Sometimes, the stability of the linearized system will imply the stability of the perturbation.

ERROR ANALYSIS.

Error analysis is based on linear approximation. Assume, you make a measurement of a function $f(a,b,c)$, where a,b,c are parameters. Assume, you know the numbers a,b,c up to accuracy ϵ. How precise do you know the values $f(a,b,c)$? Because $f(a_0+\epsilon_a, b_0+\epsilon_b, c_0+\epsilon_c)$ is about $f(a_0,b_0,c_0)+\nabla f(a_0,b_0,c_0)\cdot(\epsilon_a,\epsilon_b,\epsilon_c)$, the answer is that we know F up to accuracy $|\nabla f(a_0,b_0,c_0)|\epsilon$.

POWER LAWS.

Some laws in physics are given by functions of the form $g(x,y)=x^\alpha y^\beta$. An example is the Cobb-Douglas formula in economics. Such dependence on x or y is called **power law behavior**. If we consider $f=\log(g)$, and introduce $a=\log(x), b=\log(y)$, then this becomes $f(a,b)=\log(g(x,y))=a\alpha+b\beta$. Power laws become linear laws in a logarithmic scale. But they usually are linear approximations to more complicated nonlinear relations.

ELECTRONICS.

If we apply a voltage difference U at the ends of a resistor R, then a current I flows. The relation $U=RI$ is called **Ohms law**. In logarithmic coordinates $\log(U)=\log(R)+\log(I)$, this is a linear law. In reality, the relation between current, voltage and resistance is more complicated. For example, if the resistor heats up, then its characteristics begin to change. Nonlinear resistors are used for example in synthesizers or in radars. While Ohm's law works **extremely well**, the nonlinear behavior can have important consequences for example to stabilize systems or to protect equipment against over-voltages.

THERMODYNAMICS.

If $l(T)$ is the length of an object with temperature T, then $l(T)=l(T_0)+c(T-T_0)$, where the expansion coefficient c depends on the material. (Trick question: What happens if you heat a ring, does the inner ring become smaller or bigger?). The volume of a hot air balloon and therefore its lift capacity grows like $c(T-T_0)^3$. The law of expansion is only an approximation.

OCEANOGRAPHY.

For oceanographers, it is important to know the water density $\rho(T,S)$ in dependence on the **temperature** T (Kelvin) and the **salinity** S (psu). If we would include the pressure P (Bar), then we had a function $\rho(T,S,P)$ of three variables. Near a specific point (T,S,P) the density can be approximated by a linear function giving a law which is precise enough.

ENGENEERING.

Hooks law tells that the force of a spring is proportional to the length with which it is pulled: $F(l)=c(l-l_0)$, where l_0 is the length when the spring is relaxed. This allows to measure weights or to cushion shocks. However, this law is only good in a certain range. If the spring is pulled too strongly, then more force is needed. Such a nonlinear behavior is needed for example in shock absorbers.

MECHANICS

For small amplitudes, the pendulum motion $\ddot{x}=-g\sin(x)$ can be approximated by $\ddot{x}=-gx$, the harmonic oscillator. Nonlinear (partial) differential equations like $u_{xx}+u_{yy}+u_{zz}=F(x,y,z)$ are often approximated by linear differential equations.

CARTOGRAPHY.

It was well known already to the Greeks that we live on a sphere. On a sphere a triangle however the sum of its angles adds up to more than 180 degrees and every straight line (great circles) crosses every other line at least twice. Despite this, a city map can perfectly assume that the coordinate system is Cartesian. When drawing a plan of a house, an architect can assume that the house stands on a plane (the level curve of the linearization $G(x,y)$ of $F(x,y)$ defining the surface of the earth.)

RELATIVITY

Newton's law tells that $r''(t)$, the acceleration of a particle is proportional to the force F which acts on the mass point: $r''(t)=F/m$. For a constant force and zero initial velocity this implies $r'(t)=tF/m$. This law can not apply for all times, because we can not reach the speed of light with a massive body. In special relativity, the Newton axiom is replaced with $d/dt(r'(t)m(t))=F$, where the mass $m(t)$ depends on the velocity. This gives $v(t)=(tF/m_0)\frac{1}{\sqrt{1+F^2t^2/(c^2m_0^2)}}$. Linearization at $t=0$ produces the classical law $v(t)=tF/m_0$.

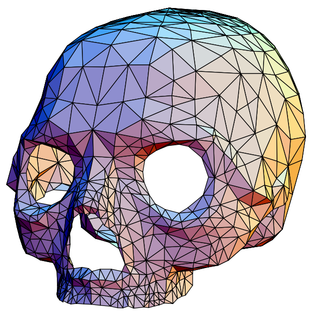
ECONOMICS.

The mathematician Charles W. Cobb and the economist Paul H. Douglas found in 1928 empirically a formula $F(L,K)=bL^\alpha K^\beta$ giving the total production F of an economic system as a function of the amount of labor L and the capital investment K. This is a linear law in logarithmic coordinates. The formula actually had been found by linear fit of empirical data. In general, the production depends in a more complicated way on labor and capital investment. For example, with increase of labor and investment, logistic constraints will become relevant.

CHEMISTRY.

The ideal gas law $PV=RT$ relates the pressure, the volume and the temperature of an ideal gas using a constant R called the Avogadro number. This law $T=f(P,V)$ is linear in logarithmic scales. This law is only an approximation and has to be replaced by the van der Waals law, which takes into account the molecular interactions as well as the volume of the molecules.

THE SKULL SURFACE, Oliver Knill, Maths21a



Surfaces are usually represented in a triangulated form, where a few points on the surface are given and triples of points form triangles. This illustrates the concept of **linear approximation**. Every triangle is close to the actual surface.

HOW TO PRODUCE THE SKULL SURFACE.

We found on the web a file **skull.dat**, which contains data

```
0 -0.02 0.02
0.29 0.65 0.08
0.4 0.53 0.06
0 -0.02 0.02
0.4 0.53 0.06
0.53 0.41 0.02
0 -0.02 0.02
...
...
```

These data encode points in space: every line is a point, three points in a row represent a triangle in space.

So, the first 3 lines from the datafile represent the triangle $\{P, Q, R\} = \{(0, -0.02, 0.02), (0.29, 0.65, 0.02), (0.4, 0.65, 0.08)\ \}$. In total, the file contains 3729 points which means that there are 1243 triangles encoded.

PLOTTING THE SURFACE.

To the right is a small Mathematica program which produces a surface from these data. The data are first read, then packed into groups of 3 to get the points, then packed into groups of three to get the triangles, then each triple of points is made into a polygon. All these polygons are finally displayed.

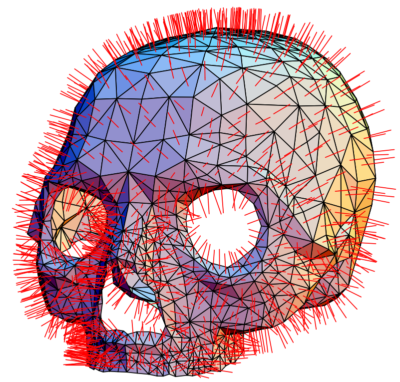
```
A=ReadList[”skull.dat”,Number];
B=Partition[A,3];
P=Partition[B,3];
U=Map[Polygon,P];
Show[Graphics3D[U],Boxed -> False, ViewPoint -> {0, −3, 0}]
```

PLOTTING THE GRADIENTS.

We can get the gradient vector on each triangle $\{P_1, P_2, P_3\}$ by taking the cross product $n = (P_1 - P_2) \times (P_3 - P_2)$ and adding lines connecting $Q = (P_1 + P_2 + P_3)$ with the point $Q + n$. We used that

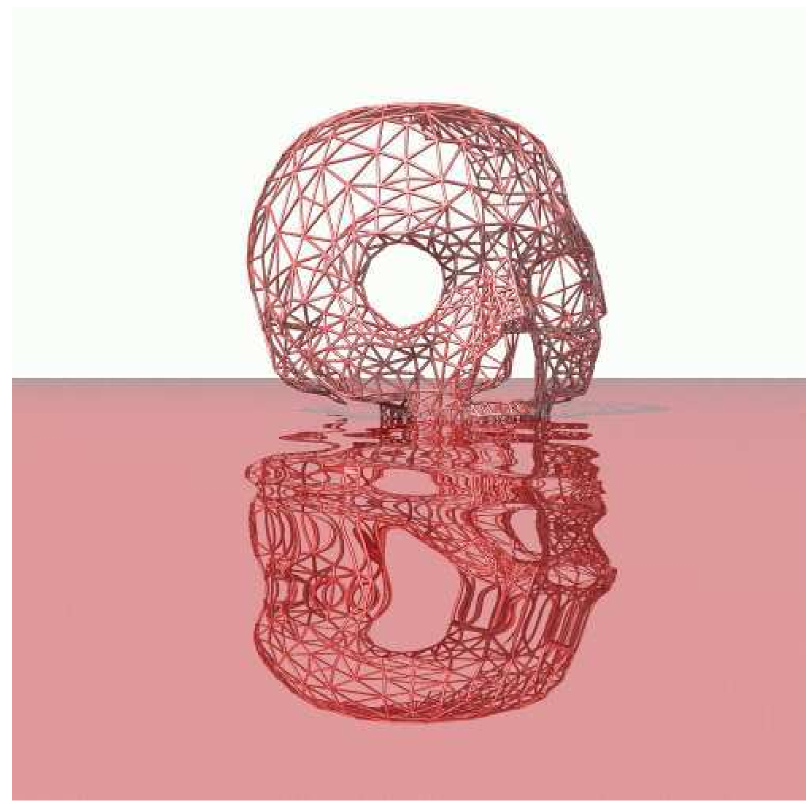
Gradients are orthogonal to the level surfaces.

```
A=ReadList[”skull.dat”,Number];
B=Partition[A,3];
P=Partition[B,3];
U=Map[Polygon,P];
NormedCross[a_,b_]:=Module[{},u=Cross[a,b]; u/Sqrt[u.u]];
NormalVector[{ a_,b_,c_ }]:=Module[{},P1=(a+b+c)/3;
P2=P1-NormedCross[b-a,c-a]/3;Line[{P1,P2}]];
V=Map[NormalVector,P];
Show[Graphics3D[{U,V}],Boxed-> False,ViewPoint-> {0, −3, 0}]
```



Here is the output of the above routine. Fortunately, the points of each triangle were given in such a way that we also know the direction of the normal vector. We have chosen the direction pointing outwards. Through every triangle, one gradient vector is displayed.

Once we have the data in Mathematica, we can also export a file, which a raytracer can read. We could also export to a VRML (Virtual Reality Markup Language) and fly around or through it. The ultimate Halloween experience!



ICE Linear Approximation/Estimates **Maths21a O.Knill**

1) Find the linear approximation $L(x, y, z)$ to $f(x, y, z) = xyz + yz + zx$ at the point $(1, 0, 1)$.

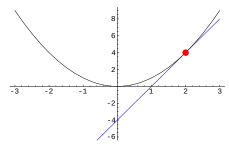
2) Estimate $f(1.001, -0.001, 0.999)$.

TANGENT LINES AND PLANES — Maths21a

TANGENT LINE. Because $\vec{n} = \nabla f(x_0, y_0) = \langle a, b\rangle$ is perpendicular to the level curve $f(x,y) = c$ through (x_0, y_0), the equation for the tangent line is

$$\boxed{ax + by = d,\ a = f_x(x_0, y_0),\ b = f_y(x_0, y_0),\ d = ax_0 + by_0}$$

Example: Find the tangent to the graph of the function $g(x) = x^2$ at the point $(2, 4)$. Solution: the level curve $f(x,y) = y - x^2 = 0$ is the graph of a function $g(x) = x^2$ and the tangent at a point $(2, g(2)) = (2, 4)$ is obtained by computing the gradient $\langle a, b\rangle = \nabla f(2, 4) = \langle -g'(2), 1\rangle = \langle -4, 1\rangle$ and forming $-4x + y = d$, where $d = -4 \cdot 2 + 1 \cdot 4 = -4$. The answer is $\boxed{-4x + y = -4}$ which is the line $y = 4x - 4$ of slope 4. Graphs of 1D functions are curves in the plane, you have computed tangents in single variable calculus.

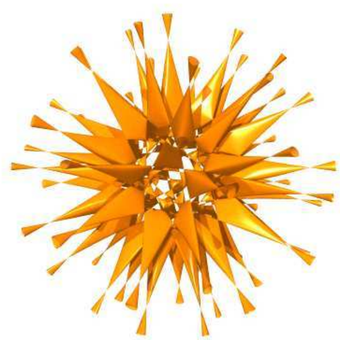


GRADIENT IN 3D. If $f(x,y,z)$ is a function of three variables, then $\nabla f(x,y,z) = (f_x(x,y,z), f_y(x,y,z), f_z(x,y,z))$ is called the **gradient** of f.

POTENTIAL AND FORCE. Force fields F in nature often are gradients of a function $U(x,y,z)$. The function U is called a **potential** of F or the potential energy.

EXAMPLE. If $U(x,y,z) = 1/|x|$, then $\nabla U(x,y,z) = -x/|x|^3$. The function $U(x,y,z)$ is the **Coulomb potential** and ∇U is the **Coulomb force**. The gravitational force has the same structure but a different constant. While much weaker, it is more effective because it only appears as an attractive force, while electric forces can be both attractive and repelling.

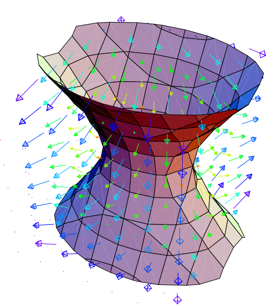
LEVEL SURFACES. If $f(x,y,z)$ is a function of three variables, then $f(x,y,z) = C$ is a surface called a level surface of f. The picture to the right shows the Barth surface $(3 + 5t)(-1 + x^2 + y^2 + z^2)^2\ (-2 + t + x^2 + y^2 + z^2)^2\ 8(x^2 - t^4y^2)(-t^4x^2 + z^2)\ (y^2 - t^4z^2)\ (x^4 - 2x^2y^2 + y^4 - 2x^2z^2 - 2y^2z^2 + z^4) = 0$, where $t = (\sqrt{5} - 1)/2$ is the golden ratio.



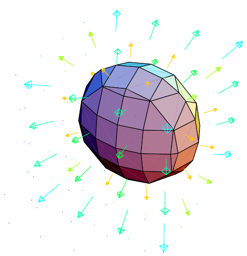
ORTHOGONALITY OF GRADIENT. We have seen that the gradient $\nabla f(x,y)$ is normal to the level curve $f(x,y) = c$. This is also true in 3 dimensions:

> The gradient $\nabla f(x,y,z)$ is normal to the level surface $f(x,y,z)$.

The argument is the same as in 2 dimensions: take a curve $\vec{r}(t)$ on the level surface. Then $\frac{d}{dt} f(\vec{r}(t)) = 0$. The chain rule tells from this that $\nabla f(x,y,z)$ is perpendicular to the velocity vector $\vec{r}\,'(t)$. Having ∇f tangent to all tangent velocity vectors on the surfaces forces it to be orthogonal.

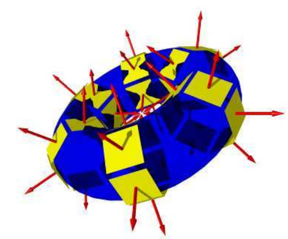


EXAMPLE. The gradient of $f(x,y,z) = x^2 + 2y^2 + z^2$ at a point (x,y,z) is $(2x, 4y, 2z)$. It illustrates well that going into the direction of the gradient **increases** the value of the function.



TANGENT PLANE. Because $\vec{n} = \nabla f(x_0, y_0, z_0) = \langle a, b, c\rangle$ is perpendicular to the level surface $f(x,y,z) = C$ through (x_0, y_0, z_0), the equation for the tangent plane is

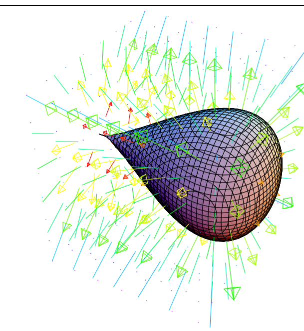
$$\boxed{ax + by + cz = d,\quad (a,b,c) = \nabla f(x_0, y_0, z_0),\ d = ax_0 + by_0 + cz_0.}$$



EXAMPLE. Find the general formula for the tangent plane at a point (x,y,z) of the Barth surface. Just kidding ... Note however that computing this would be no big deal with the help of a computer algebra system like Mathematica. Lets look instead at the quartic surface

$$f(x,y,z) = x^4 - x^3 + y^2 + z^2 = 0$$

which is also called the "piriform" or "pair shaped surface". What is the equation for the tangent plane at the point $P = (2, 2, 2)$? We get $\langle a, b, c\rangle = \langle 20, 4, 4\rangle$ and so the equation of the plane $20x + 4y + 4z = 56$.



EXAMPLE. An important example of a level surface is $g(x,y,z) = z - f(x,y)$ which is the graph of a function of two variables. The gradient of g is $\nabla f = (-f_x, -f_y, 1)$. This allows us to find the equation of the tangent plane at a point.

Quiz: What is the relation between the gradient of f in the plane and the gradient of g in space?