**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 $f(\vec{r}(t))$. The chain rule is

$$\frac{d}{dt} f(\vec{r}(t)) = \nabla f(\vec{r}(t)) \cdot \vec{r}'(t).$$

**WRITING IT OUT.** Writing the dot product out gives

$$\frac{d}{dt} f(x(t), y(t)) = f_x(x(t), y(t)) \frac{dx}{dt}(t) + f_y(x(t), y(t)) \frac{dy}{dt}(t).$$

**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 $z_y = 2 \cos(x + 2y)$. Then $\frac{dz}{dt} = e^t \cdot \frac{dz}{dt} = \cos(t)$ and $\frac{dz}{dt} = \cos(x + 2y) = 2 \cos(x + 2y) \sin(t)$.

**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 of the temperature $\vec{r}'(t)$ of the spider as well as the temperature gradient $\nabla f$ 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 $T(x, y) = x^2 - y^4$. Find the rate of change of the temperature, “Nabla” experiences.

**SOLUTION.** $\nabla T(x, y) = (\partial T/\partial x, \partial T/\partial y) = (2x, -4y^3)$. $\vec{r}'(t) = (\cos(t), \sin(t))$.

Then $\frac{\partial f}{\partial x} = 2(x(t), -4y(t)^3)(\cos(t), \sin(t)) = 2x \cos(t) - 4y \sin(t)$. The rate of change of the temperature, “Nabla” experiences is.

**APPLICATION ENERGY CONSERVATION.** If $\vec{H}(x, y)$ is the energy of a system, the system moves satisfies the equations $\vec{x}'(t) = \nabla H(\vec{x}(t))$. For example, if $\vec{H}(x, y) = y^2/2 - V(x)$ is a sum of kinetic and potential energy, then $\vec{x}'(t) = y(t, \vec{x})(\vec{x}') = \nabla V(\vec{x})$. In the case of the Kepler problem, we had $V(x) = Gm/x$. The energy $H$ is conserved.

**PROOF.** The chain rule shows that $d/\partial t H(\vec{x}(t), y(t)) = \nabla H(\vec{x}(t), y(t)) \cdot \vec{x}'(t) = H_x(x(t), y(t)) \frac{dx}{dt}(t) + H_y(x(t), y(t)) \frac{dy}{dt}(t) - H_x(x(t)) \frac{dx}{dt}(t) - H_y(x(t)) \frac{dy}{dt}(t) = 0$.

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

$$y' = -f_x/f_y.$$

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

**APPLICATION: DIFFERENTIATION RULES.** One ring of the chain, rules them all

$$f(x, y) = x + y, x = u(t), y = v(t), d/\partial t (x + y) = f_x(1)u' + f_y(1)v' = x' + v'.$$

$$f(x, y) = x + y, x = u(t), y = v(t), d/\partial t (x + y) = f_x(1)u' + f_y(1)v' = x' + v'.$$

$$f(x, y) = x + y, x = u(t), y = v(t), d/\partial t (x + y) = f_x(1)u' + f_y(1)v' = x' + v'. $$

**DIETERICIEQUATION.** In thermodynamics the temperature $T$, the pressure $p$ and the volume $V$ of a gas are related. One refinement of the ideal gas law $pV = RT$ is the Dieteric equation $g(p, V, T) = pV - b(x)p^{n-1}V^{n-1} - RT = 0$. The constant $b$ depends on the volume of the molecules and $n$ depends on the interaction of the molecules. (A different variation of the ideal gas law is van der Waal’s law)

**Problem:** compute $V_T$

If $V = V(T, p)$, the chain rule says $\frac{d}{dt} V_P + \frac{d}{dt} p = 0$, so that $V_T = f_T(f_T) = -\frac{V_T}{p V_T^n R T^n} = \left(p V_T^2 / R T^2 - R T / (p V_T^n R T^n) = \left(p V_T^2 / R T^2 - R T / (p V_T^n R T^n)ight) \right)$. (This could be simplified to $(R + u V_T / (RT^2 - R T / (p V_T^n R T^n))$).

**DERIVATIVE.** If $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a map, its derivative $f'$ is the $m \times n$ matrix $[f']_{ij} = \frac{d}{dt} f_j(x(t))$.

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

| $f : \mathbb{R}^n \rightarrow \mathbb{R}$ | $f' : \mathbb{R}^n \rightarrow \mathbb{R}^n$ | $f'' : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ | $f''' : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n \times n}$ | $f^\ast$ : \mathbb{R}^n \rightarrow \mathbb{R}^n |
---|---|---|---|---|
$\nabla$ | $\nabla f$ | $\nabla^2 f$ | $\nabla^3 f$ | $\nabla f$ |

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

$$\frac{d}{dt} f(g(x)) = \sum_i \frac{df_i}{dx_j} \frac{dx_i}{dt} \frac{dg_j}{dt}$$

Both $f'(g(x))$ and $g'(x)$ are matrices and 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 $\vec{\nabla} f = (f_r, f_\theta)$. Using the chain rule, we can relate this to the usual gradient: $\vec{\nabla} f(x(t), y(t), \theta(t)) = f_x(x, y, \theta) \cos(\theta) + f_y(x, y, \theta) \sin(\theta)$ and $d/\partial \theta f(x(t), y(t), \theta(t)) = -f_x(x, y, \theta) \sin(\theta) + f_y(x, y, \theta) \cos(\theta)$ means that the length of $\vec{\nabla} f$ is the same in both coordinate systems.

**PROOFS OF THE CHAIN RULE.**

1. **Proof.** Near a point $L$, we can approximate $f$ by a linear function $f$. 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 $\frac{d}{dt} f(\vec{r}(t)) = \frac{d}{dt}(g(x_0 + tu) + b(y_0 + tv)) = au + bv$ and this is the dot product of $\vec{\nabla} f(x_0, y_0)$ 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 $u_t + u u_x$ 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 $f$.
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 \( \vec{r}(t) = (1 + t, \sin(t)) \).

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

Find the rate of change of the height \( \frac{df(\vec{r}(t))}{dt} \) at the point \( t = 0 \) using the chain rule.
GRADIENT AND DIRECTIONAL DERIVATIVE

**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
\[
\frac{d}{dt}f(x(t),y(t),z(t)) = f_x(x(t),y(t),z(t)) \cdot x'(t) + f_y(x(t),y(t),z(t)) \cdot y'(t) + f_z(x(t),y(t),z(t)) \cdot z'(t).
\]
This looks like the 1D chain rule, but the derivative \( f' \) is replaced with the gradient and the derivative of \( t \) is the velocity. Written out in detail the chain rule is
\[
\frac{df}{dt} = \frac{\partial f}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dt} + \frac{\partial f}{\partial z} \cdot \frac{dz}{dt}.
\]
Every curve \( \vec{r}(t) \) on the level curve or level surface satisfies \( \nabla f(\vec{r}(t)) = 0 \). By the chain rule, \( \nabla f(\vec{r}(t)) \) is perpendicular to the tangent vector \( \vec{r}'(t) \).

**DIRECTINAL 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 \( \nabla f \cdot \vec{v} \) or \( D_{\vec{v}}f \).

**DERIVATIVE DIRECTIONAL DERIVATIVE.** The rate of change in all directions is zero if and only if \( \nabla f = 0 \).

**ZERO DIRECTIONAL DERIVATIVE.** The rate of change in all directions is zero if and only if \( \nabla f(x,y) = 0 \) is perpendicular to the level curve or level surface.

**PROPERTIES DIRECTIONAL DERIVATIVE.** Properties of Gradient
\[
D_{\vec{v}}(\lambda f) = \lambda D_{\vec{v}}f, \quad D_{\vec{v}}(f + g) = D_{\vec{v}}f + D_{\vec{v}}g, \quad D_{\vec{v}}(fg) = fD_{\vec{v}}g + gD_{\vec{v}}f.
\]

**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 \( \phi \)-direction is \( \nabla p(x,y) \cdot \vec{v} = 2\cos(\phi) + 4\sin(\phi) \). This is maximal for \(-2\sin(\phi) + 4\cos(\phi) = 0 \) which means \( \tan(\phi) = 1/2 \).

**EXAMPLE.** 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.

The Van der Waals law \( pV = nRT \) is obtained when \( a, b \) are set to 0. The level curves or **isotherms** \( T(p, V) = \text{const} \) tell much about the properties of the gas. The so-called **reduced van der Waals law** \( T(p, V) = (p + 3V^2)/(3V - 1)/8 \) is obtained by scaling \( p, V, T \) depending on the gas. Calculate the directional derivative of \( T(p, V) \) at the point \((1,1)\) into the direction \( \vec{v} = (1, 2) \).

We have \( T_p(1, 1) = (3V - 1)/8 \) and \( T_v(1, 1) = 3p/8 - (9/8)/V^2 - 3/(4V^2) \). Therefore, \( \nabla T(1,1) = (1/4, 0) \) and \( D_{\vec{v}}T(1,1) = 1/5 \).

**TANGENT LINE.** Because \( \vec{n} = \nabla f(x_0, y_0) = (a, b) \) is perpendicular to the level curve \( f(x, y) = c \) through \((x_0,y_0)\), the equation for the tangent line is
\[
ax + by = d = f_x(x_0, y_0) = f(x_0, y_0), \quad d = ax_0 + by_0.
\]

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

**EXAMPLE (2D)** Find the linear approximation of the function \( f(x, y) = \sin(xy) \) at the point \((1, 1)\). The graph of \( f \) is tangent to the graph of \( L \).

\[
L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)
\]

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

**EXAMPLE (3D)** Find the linear approximation to \( G(\mathbf{n}) \) at \((1, 2, 3)\). The gradient of \( G \) is \( \nabla G = \langle 2, 3, -1 \rangle \). The linear approximation is \( L(x, y, z) = f(x_0, y_0, z_0) + f_x(x_0, y_0, z_0)(x - x_0) + f_y(x_0, y_0, z_0)(y - y_0) + f_z(x_0, y_0, z_0)(z - z_0) \). 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

\[
L(\mathbf{z}) = f(\mathbf{z}_0) + \nabla f(\mathbf{z}_0) \cdot (\mathbf{z} - \mathbf{z}_0)
\]

**EXAMPLE (3D)** Use the best linear approximation to \( f(x, y, z) = e^{x+y} \) 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 of \( f \) is \( \nabla f(x, y, z) = (e^{x+y}, ye^{x+y}, xe^{x+y}) \). At the point \((x_0, y_0, z_0) = (0.25, 1)\) the gradient is the vector \((5, 10, 5)\). The linear approximation is \( L(x, y, z) = f(x_0, y_0, z_0) + f_x(x_0, y_0, z_0)(x - x_0, y - y_0, z - z_0) = 5 + (5, 10, 5)(x - 0.25, y - 1, z - 0.5) = 5x + 10y + 5z - 2.5 \). We can approximate \( f(0.01, 24.8, 1.02) \) by \( 5(0.01) + 10(24.8) + 5(1.02) \) or \( 5(0.01) + 248 + 5(1.02) \) which is \( 5 + 0.05 = 5.05 \). The actual value is \( f(0.01, 24.8, 1.02) = 5.1306 \), so the estimate is correct.

**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, \( f_{xy} = f_{yx} \), \( f_{yz} = f_{zy} \), and \( f_{xz} = f_{zx} \), the matrix \( f'' \) is symmetric (a reflection at the diagonals it invariant).

**QUADRATIC APPROXIMATION.** (informal)

If \( f \) is a function of several variables \( x \) and \( y \) is a point, then

\[
Q(f) = f(x_0) + \nabla f(x_0) \cdot (x - x_0) + \frac{1}{2} (x - x_0)^T H(x_0)(x - x_0)
\]

is called the quadratic approximation of \( f \).

It generalizes the quadratic approximation \( L(x) = f(x_0) + f_x(x_0)(x - x_0) + f''(x_0)/2 \) of a function of one variable.

**EXAMPLE.** If the height of a hill area is given by \( f(x, y) = 4000 - \sin(x^2 + y^2) \), find the quadratic approximation of \( f \) at \((0,0)\).

\[
Q(f) = f(0,0) + \nabla f(0,0) \cdot (x - 0, y - 0) + \frac{1}{2} (x - 0)^2 + (y - 0)^2
\]

is the quadratic approximation of \( f \). The actual value is \( f(0,0) = 4000 \), the linear approximation of \( f \) at \((0,0)\) is \( Q(f) = 4000 \). The graph of \( f \) is a plane tangent to the graph of \( F \).

**EXAMPLE: REDUCED VAN DER WAALS LAW**

If we take Taylor’s theorem that the error \( f(x, y) - L(x, y) \) is in a region \( R \) near \((x_0, y_0)\), then \( f(x, y) \) is smaller or equal to \( M ||x - x_0| + |y - y_0||^{1/2} \), where \( M \) is the maximum value of all the matrix entries \( f''(x, y) \) in that region \( R \).

**THERMAL ERROR OF APPROXIMATION.**

Athermalizing the equation \( f(x, y) = f(x, y) - L(x, y) \) in a region \( R \) near \((x_0, y_0)\), we can estimate it with the linear approximation which is \( L(x_0, y_0) = f(x_0, y_0) + \Delta f + \Delta f \). 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 \( \Delta f \) 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 \( \Delta f(x_0, y_0)/f(x_0, y_0) \).
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 understanding 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 $\epsilon$. How precise do you know the values $f(a, b, c)$? Because $f(a, b, c) = f(a_0, b_0, c_0) + \nabla f(a_0, b_0, c_0) \cdot (a - a_0, b - b_0, c - c_0)$, the answer is that we know $F$ up to accuracy $|\nabla f(a_0, b_0, c_0) \cdot (\epsilon, \epsilon, \epsilon)|$.

POWER LAWS.

Some laws in physics are given by functions of the form $g(x, y) = x^a y^b$. An example is the Cobb-Douglas formula in economics. Such dependence on $x$ or $y$ is called power law behavior. If we consider $f = \log(y)$, 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.

ENGINEERING.

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 $\ddot{x} = -g x$, 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 $\ddot{x}(t)$, the acceleration of a particle is proportional to the force $F$ which acts on the mass point: $\ddot{x}(t) = F/m$. For a constant force and zero initial velocity this implies $\ddot{x}(t) = F/m$. This law 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 $d/dt(\dot{x}(t)m(t)) = F$, where the mass $m(t)$ depends on the velocity. This gives $\ddot{x}(t) = (F/m_0)\frac{1}{\sqrt{1 - v^2/c^2}}$ Linearization at $t = 0$ produces the classical law $\ddot{x}(t) = F/m_0$.

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.
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 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 \( \{P, Q, R\} = \{(0, -0.02, 0.02), (0.29, 0.65, 0.02), (0.4, 0.53, 0.08)\} \). 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.

```mathematica
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 need that

Gradients are orthogonal to the level surfaces.

```mathematica
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;
   V = Map[NormalVector, {P1, P2}];
   Show[Graphics3D[{U, V}], Boxed -> False, ViewPoint -> {0, -3, 0}]
```
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

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

\[
ax + by = d, \quad a = f_x(x_0, y_0), \quad b = f_y(x_0, y_0), \quad d = ax_0 + by_0.
\]

Example: Find the tangent to the graph of the function \( f(x) = x^2 \) 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 \( (a, b) = \nabla f(2, 4) = (-g'(2), 1) = (-4, 1) \) and forming \(-4x + y = d\), where \( d = -4 \times 2 + 1 \times 4 = -4 \). The answer is \(-4x + y = -4\) which is the line \( y = 4x - 4 \) of slope 4. Graphs of 1D functions are curves in the plane, and 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 \( \nabla U \) 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 \((3 + 4t)(-1 + x^3 + y^3 + z^3)(-2 + 1 + x^3 + y^3 + z^3)(8(x^2 - 4y^2)(x^4 + y^4 + z^4)(y^4 + z^4 + x^4) - 2x^2y^2 + y^4 - 2x^2z^2 - 2y^2z^2 + z^4) = 0, \) 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:

\[
\text{The gradient } \nabla f(x, y, z) \text{ is normal to the level surface } f(x, y, z).
\]

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 us that \( \nabla f(x, y, z) \) is perpendicular to the velocity vector \( \vec{r}'(t) \). Having \( \nabla f \) 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 \((x, y, z) = (2x, 4y, 2z)\). It illustrates well that going into the direction of the gradient increases the value of the function.

TANGENT PLANE. Because \( \vec{n} = \nabla f(x_0, y_0, z_0) = (a, b, c) \) 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

\[
ax + by + cz = d, \quad (a, b, c) = \nabla f(x_0, y_0, z_0), \quad d = ax_0 + by_0 + cz_0.
\]

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. Let’s look instead at the quartic surface

\[
f(x, y, z) = x^4 - x^3 + y^3 + z^2 = 0
\]

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 \((a, b, c) = (20, 4, 4)\) 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?