

HOMEWORK. Section 11.4: 8,28,36,38,46.

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 variable $t \mapsto f(\vec{r}(t))$.The chain rule looks like the 1D chain rule, but where the derivative f' is replaced with the gradient $\nabla f = (f_x, f_y, f_z)$. $d/dt f(\vec{r}(t)) = \nabla f(\vec{r}(t)) \cdot \vec{r}'(t)$

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 $x'(t) = H_y$, $y'(t) = -H_x$. 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 $V(x) = Gm/|x|$. The energy H 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 $y' = -f_x/f_y$.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 (OSCAR NIGHT). 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 gas 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 $V_T = -f_T/f_V = -(-ap(V - b)e^{a/RVT}/(RV^2T^2) - R)/(pVe^{a/RVT} - p(V - b)e^{a/RVT}/(RV^2T))$. (This could be simplified to $(R + a/TV)/(RT/(V - b) - a/V^2)$).

DERIVATIVE. If $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$ 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)

$f : \mathbf{R} \rightarrow \mathbf{R}^3$ curve f' velocity vector.	$f : \mathbf{R}^3 \rightarrow \mathbf{R}$ scalar function f' gradient vector.	$f : \mathbf{R}^2 \rightarrow \mathbf{R}^3$ surface f' tangent matrix	$f : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ coordinate change f' Jacobean matrix.	$f : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ gradient field 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 $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$ and $g : \mathbf{R}^k \rightarrow \mathbf{R}^n$, 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 .

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

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 $\frac{d}{dt} f(\vec{r}(t)) = \frac{d}{dt} (a(x_0 + tu) + b(y_0 + tv) - c) = au + bv$ 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 .