14: Chain rule

If f and g are functions of t, then the **single variable chain rule** tells

$$\frac{d}{dt} f(g(t)) = f'(g(t))g'(t) .$$

For example, $d/dt \sin(\log(t)) = \cos(\log(t))/t$. This **chain rule** can be proven by linearising the functions f and g and verifying the chain rule in the linear case. The rule is useful for finding derivatives like $\arccos'(x)$: write $1 = d/dx \cos(\arccos(x)) = -\sin(\arccos(x)) \arccos'(x) = -\sqrt{1-\sin^2(\arccos(x))} \arccos'(x) = \sqrt{1-x^2} \arccos'(x)$ so that $\arccos'(x) = -1/\sqrt{1-x^2}$.

1 Find the derivative d/dx $\arctan(x)$. **Solution.** We have $\sin' = \cos$ and $\cos' = -\sin$ and from $\cos^2(x) + \sin^2(x) = 1$. follows $1 + \tan^2(x) = 1/\cos^2(x)$. Therefore $d/dx \tan(\arctan(x)) = 1/\cos^2(\arctan(x)) \tan'(x) = x$ Now $1/\cos^2(x) = 1/(1 + \tan^2(x)$ so that $\tan'(x) = 1/(1 + x^2)$.

Define the **gradient** $\nabla f(x, y) = [f_x(x, y), f_y(x, y)]$ or $\nabla f(x, y, z) = [f_x(x, y, z), f_y(x, y, z), f_z(x, y, z)]$.

If $\vec{r}(t)$ is curve and f is a function of several variables we can build a function $t \mapsto f(\vec{r}(t))$ of one variable. Similarly, If $\vec{r}(t)$ is a parametrization of a curve in the plane and f is a function of two variables, then $t \mapsto f(\vec{r}(t))$ is a function of one variable.

The **multi-variable chain rule** is

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

Proof. When written out in two dimensions, it is

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

Now, the identity

$$\frac{f(x(t+h),y(t+h))-f(x(t),y(t))}{h} = \frac{f(x(t+h),y(t+h))-f(x(t),y(t+h))}{h} + \frac{f(x(t),y(t+h))-f(x(t),y(t))}{h}$$

holds for every $h > 0$. The left hand side converges to $\frac{d}{dt} f(x(t), y(t))$ in the limit $h \to 0$ and the right hand side to $f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t)$ using the single variable chain rule twice. Here is the proof of the later, when we differentiate f with respect to t and y is treated as a constant:

$$\frac{f(\text{ x(t+h) }) - f(x(t))}{h} = \frac{[f(\text{ x(t) + (x(t+h)-x(t)) }) - f(x(t))]}{\text{[x(t+h)-x(t)]}} \cdot \frac{\text{[x(t+h)-x(t)]}}{h} .$$

Write $H(t) =$ x(t+h)-x(t) in the first part on the right hand side.

$$\frac{f(x(t+h)) - f(x(t))}{h} = \frac{[f(x(t) + H) - f(x(t))]}{H} \cdot \frac{x(t+h) - x(t)}{h} .$$

As $h \to 0$, we also have $H \to 0$ and the first part goes to $f'(x(t))$ and the second factor to $x'(t)$.

2 We move on a circle $\vec{r}(t) = [\cos(t), \sin(t)]$ on a table with temperature distribution $f(x,y) = x^2 - y^3$. Find the rate of change of the temperature $\nabla f(x,y) = [2x, -3y^2]$, $\vec{r}'(t) = [-\sin(t), \cos(t)]$ $d/dt f(\vec{r}(t)) = \nabla T(\vec{r}(t)) \cdot \vec{r}'(t) = [2\cos(t), -3\sin(t)^2] \cdot [-\sin(t), \cos(t)] = -2\cos(t)\sin(t) - 3\sin^2(t)\cos(t)$.

From $f(x,y) = 0$, we 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 get

Implicit differentation: $y' = -f_x/f_y$.

Even so, we do not know $y(x)$, we can compute its derivative! Implicit differentiation works also in three variables. The equation $f(x,y,z) = c$ defines a surface. Near a point where f_z is not zero, the surface can be described as a graph $z = z(x,y)$. We can compute the derivative z_x without actually knowing the function $z(x,y)$. To do so, we consider y a fixed parameter and compute using the chain rule $f_x(x,y,z(x,y))1 + f_z(x,y)z_x(x,y) = 0$. This leads to the following

Implicit differentiation:
$z_x(x,y) = -f_x(x,y,z)/f_z(x,y,z)$ $\quad$ $z_y(x,y) = -f_y(x,y,z)/f_z(x,y,z)$

3 The surface $f(x,y,z) = x^2 + y^2/4 + z^2/9 = 6$ is an ellipsoid. Compute $z_x(x,y)$ at the point $(x,y,z) = (2,1,1)$. **Solution:** $z_x(x,y) = -f_x(2,1,1)/f_z(2,1,1) = -4/(2/9) = -18$.

4 How does the chain rule relate to other differentiation rules? **Answer.** The chain rule is universal: it implies single variable differentiation rules like the addition, product and quotient rule in one dimensions:
$f(x,y) = x + y, x = u(t), y = v(t), d/dt(x+y) = f_x u' + f_y v' = u' + v'$.
$f(x,y) = xy, x = u(t), y = v(t), d/dt(xy) = f_x u' + f_y v' = vu' + uv'$.
$f(x,y) = x/y, x = u(t), y = v(t), d/dt(x/y) = f_x u' + f_y v' = u'/y - v'u/v^2$.

5 Can one prove the chain rule from linearization and just verifying it for linear functions? **Solution**. Yes, as in one dimensions, the chain rule follows from linearization. If f is a linear function $f(x,y) = ax + by - c$ and if the curve $\vec{r}(t) = [x_0 + tu, y_0 + tv]$ parametrizes 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 $\vec{r}\,'(t) = (u,v)$. Since the chain rule only refers to the derivatives of the functions which agree at the point, the chain rule is also true for general functions.

6 Mechanical systems are determined by the energy function $H(x,y)$, which is a function of two variables. The first variable x is the position and the second variable y is the momentum. The equations of motion for the curve $\vec{r}(t) = [x(t), y(t)]$ are called **Hamilton equations**:

$$\begin{aligned} x'(t) &= H_y(x,y) \\ y'(t) &= -H_x(x,y) \end{aligned}$$

In a homework you verify that the energy of a Hamiltonian system is preserved: for every path $\vec{r}(t) = [x(t), y(t)]$ solving the system, we have $H(x(t), y(t)) = \mathit{const}$.