CHEM1047 – Week 3 Lecture 2 – Partial derivatives

☐ *Section 9.3 of Steiner, "The Chemistry Maths Book", 2nd edition.*

☐ *Section 4.5 of Cockett and Doggett, "Maths for Chemistry", Volume 1.*

1. Partial derivatives

A *partial derivative* of a function is a derivative with respect to just one of its arguments. The same differentiation rules apply, all other arguments should be treated as constants. Examples:

$$\frac{\partial}{\partial y}(xyz) = xz \qquad \frac{\partial}{\partial x}\left(\frac{1-x}{\sin y}\right) = \frac{-1}{\sin y}$$

Higher derivatives, such as $\partial^2 f/\partial x^2$, and mixed derivatives, such as $\partial^2 f/\partial x \partial y$, are obtained by sequential partial differentiation. For well-behaved functions, the value of mixed derivatives does not depend on the order of partial differentiation.

Partial derivatives have the same limit definition as the derivative of a univariate function:

$$\frac{\partial f(x,y,z,\ldots)}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x, y, z,\ldots) - f(x,y,z,\ldots)}{\Delta x} \tag{1}$$

The arguments are often collected into a vector for convenience, *e.g.*:

$$f(x,y,z) = f(\vec{r}), \qquad \vec{r} = [x \quad y \quad z] \tag{2}$$

Example 1:

$$z = \sin(xy)$$

$$\frac{\partial z}{\partial x} = y\cos(xy) \qquad \frac{\partial z}{\partial y} = x\cos(xy) \qquad \frac{\partial^2 z}{\partial x^2} = y^2\cos(xy)$$

$$\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial}{\partial y}\left[y\cos(xy)\right] = \cos(xy) + xy\cos(xy)$$

Example 2:

$$z = x^2 + y^2$$

$$\frac{\partial z}{\partial x} = 2x \qquad \frac{\partial z}{\partial y} = 2y \qquad \frac{\partial^2 z}{\partial x^2} = 2 \qquad \frac{\partial^2 z}{\partial x \partial y} = 0$$

2. Chain rules for multivariate functions

In situations when a function depends on multiple arguments that are themselves functions of other arguments, the chain rule for differentiation must be extended appropriately. Several situations exist:

1. A function of one variable that depends on multiple other variables:

$$\frac{\partial}{\partial x} f\left[g(x,y,z,\ldots)\right] \tag{3}$$

We can use the definition of the derivative twice to write out Δf when the argument x is incremented by Δx, and then take the limit:

$$\begin{gathered}\Delta f = \frac{df}{dg}\Delta g + O\left[\Delta g^2\right] = \frac{df}{dg}\left(\frac{\partial g}{\partial x}\Delta x + O\left[\Delta x^2\right]\right) + O\left[\Delta g^2\right] \\ \frac{\partial f}{\partial x} = \lim_{\Delta x\to 0}\frac{\Delta f}{\Delta x} = \frac{df}{dg}\frac{\partial g}{\partial x}\end{gathered} \tag{4}$$

We therefore arrive at the following chain rule:

$$\frac{\partial}{\partial x} f\left[g\left(x,y,z,\ldots\right)\right] = \frac{df}{dg}\frac{\partial g}{\partial x} \tag{5}$$

2. A function of multiple variables that each depend on the same variable:

$$\frac{d}{dt} f\left[x(t), y(t), \ldots\right] \tag{6}$$

This is a full derivative because the function ultimately only depends on one argument. Using the definition of a partial derivative followed by the definition of univariate derivative, we get:

$$\begin{gathered}\Delta f = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + \ldots + O\left[\Delta x^2\right] + O\left[\Delta y^2\right] + \ldots = \\ = \frac{\partial f}{\partial x}\frac{dx}{dt}\Delta t + \frac{\partial f}{\partial y}\frac{dy}{dt}\Delta t + \ldots + O\left[\Delta x^2\right] + O\left[\Delta y^2\right]\end{gathered} \tag{7}$$

After using the limit definition of the derivative with respect to t, we can conclude that:

$$\frac{d}{dt} f\left[x(t), y(t), \ldots\right] = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \ldots \tag{8}$$

3. A function of multiple variables that each depend on multiple other variables. This is a combination of the two situations above and the corresponding rule may be derived in a similar fashion:

$$\frac{\partial}{\partial \alpha} f\left[x(\alpha,\beta,\ldots), y(\alpha,\beta,\ldots),\ldots\right] = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \alpha} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha} + \ldots \tag{9}$$

Example 3:

$$z = \sin(x), \qquad x = u^2 v^2$$

$$\frac{\partial z}{\partial u} = \frac{dz}{dx}\frac{\partial x}{\partial u} = \left[\cos(x)\right]\left[2uv^2\right] = 2uv^2\cos\left(u^2v^2\right)$$

Example 4:

$$z = x^2 y - y^2, \qquad x = t^2, \qquad y = 2t$$

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt} = [2xy][2t] + \left[x^2 - 2y\right][2] = \ldots = 10t^4 - 8t$$

Example 5:

$$z = e^{x^2 y}, \qquad x(u,v) = \sqrt{uv}, \qquad y(u,v) = 1/v$$

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u} = \left[2xye^{x^2y}\right]\left[\frac{\sqrt{v}}{2\sqrt{u}}\right] + \left[x^2 e^{x^2 y}\right][0] = \ldots = e^u$$

3. Gradient vector

An important theorem in physics states that the force acting on a particle freely moving in some potential is equal to the partial derivatives of that potential:

$$E = U(x, y, z) \quad \Rightarrow \quad \vec{F} = \begin{bmatrix} \frac{\partial U}{\partial x} & \frac{\partial U}{\partial y} & \frac{\partial U}{\partial z} \end{bmatrix} \tag{10}$$

Because force is a vector, it is often convenient to store all partial derivatives of a function in a vector as shown in Equation (10). Such a vector is called the *gradient* of the function:

$$\vec{\nabla} U = \begin{bmatrix} \frac{\partial U}{\partial x} & \frac{\partial U}{\partial y} & \frac{\partial U}{\partial z} & \cdots \end{bmatrix} \tag{11}$$

Gradient indicates the direction of steepest growth of the function and has many uses in physical sciences – for example, one of the best methods for decreasing the value of the function (called *gradient descent*) is to move in the direction opposite to its gradient.

4. Hessian matrix

Consider a weight of mass m attached to a spring of stiffness k and equilibrium length x_0. From elementary mechanics we know that the total energy is the sum of kinetic and potential energies:

$$E = \frac{mv^2}{2} + \frac{k(x - x_0)^2}{2} \tag{12}$$

For a molecule with multiple atoms connected by multiple bonds, the expression is similar:

$$E = \frac{1}{2}\sum_n m_n |\vec{v}_n|^2 + \frac{1}{2}\sum_{n>m} k_{nm}\left(|\vec{r}_{nm}| - |\vec{r}_{nm}^{\,\mathrm{eq}}|\right)^2 \tag{13}$$

where m_n is the mass of *n*-th atom, $\vec{v}_n$ is its velocity vector, $\vec{r}_{nm}$ is the distance vector between atom *n* and atom *m*, and the "eq" superscript refers to the equilibrium value. It is easy to see that the second derivatives of the energy with respect to the inter-atomic distances are the force constants:

$$\frac{\partial^2 E}{\partial |\vec{r}_{nm}|^2} = k_{nm} \tag{14}$$

This is very useful in practice because it allows vibrational frequencies to be calculated from derivatives of the total energy. The matrix of second derivatives of a function, for example:

$$\mathbf{H}\left[f(x, y)\right] = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} \tag{15}$$

is called the *Hessian matrix*. It contains information about the local curvature of the function.