2.3.1 Definition of partial derivative for a function of two variables

Definition: The first partial derivatives of the function $f(x, y)$ with respect to the variables x and y are given by

$$\frac{\partial f}{\partial x} = \lim_{h\to 0} \frac{f(x+h, y) - f(x, y)}{h}, \tag{2.42}$$

$$\frac{\partial f}{\partial y} = \lim_{h\to 0} \frac{f(x, y+h) - f(x, y)}{h}. \tag{2.43}$$

Note: Notice that $\partial f/\partial x$ is nothing but the standard first derivative of $f(x, y)$ when considered as a function of x only, regarding y as a constant parameter. Similarly $\partial f/\partial y$ is the standard first derivative of $f(x, y)$ when considered as a function of y only, regarding x as a constant parameter.

Example: Let

$$f(x, y) = x^2 \cos y + 2xy, \tag{2.44}$$

then, according to the definition above, the partial derivatives of $f(x, y)$ with respect to x and y are

$$\frac{\partial f}{\partial x} = \cos y \frac{dx^2}{dx} + 2y\frac{dx}{dx} = 2x \cos y + 2y, \tag{2.45}$$

$$\frac{\partial f}{\partial y} = x^2 \frac{d \cos y}{dy} + 2x\frac{dy}{dy} = -x^2 \sin y + 2x. \tag{2.46}$$

Geometric interpretation: Partial derivatives of functions of two variables admit a similar geometrical interpretation as for functions of one variable. For a function of one variable $f(x)$, the first derivative with respect to x is defined as

$$\frac{df}{dx} = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}, \tag{2.47}$$

and geometrically it measures the slope of the curve $f(x)$ at the point x. This is illustrated in figure 5.

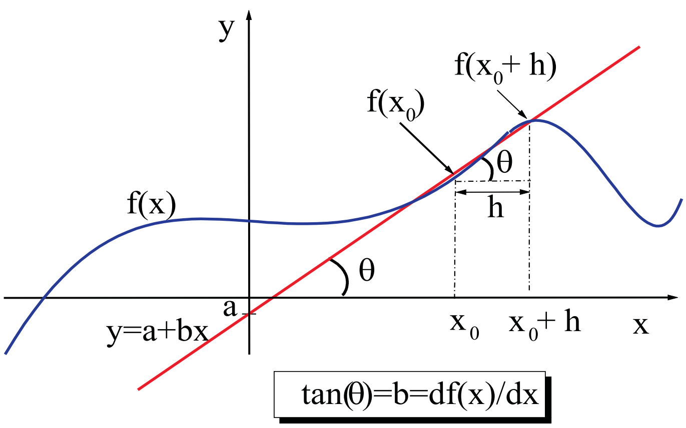


Figure 5: *Geometrical picture of df/dx at the point $x = x_0$. Applying the definition (2.47), the derivative is nothing but the slope of the line $y = a + bx$ which is tangent to $f(x)$ at the point x_0, namely $b = \frac{df}{dx}|_{x_0}$.*

Recalling the definition we gave in the note above, we can now interpret geometrically the first-order partial derivatives of a function of two variables in a completely analog fashion:

Geometrical definition of f_x and f_y: The partial derivative $\partial f/\partial x$ at a certain point (x_0, y_0) is nothing but the slope of the curve of intersection of the function $f(x, y)$ and the vertical plane $y = y_0$ at $x = x_0$. Likewise, the partial derivative $\partial f/\partial y$ at a certain point (x_0, y_0) is nothing but the slope of the curve of intersection of the function $f(x, y)$ and the horizontal plane $x = x_0$ at $y = y_0$. Graphically:

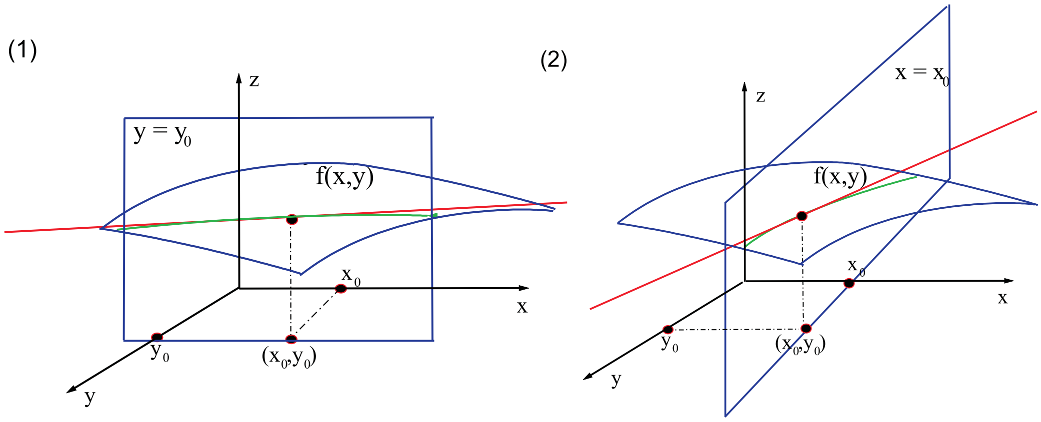


Figure 6: *Geometrical picture of $\partial f/\partial x$ (1) and $\partial f/\partial y$ (2) at the point (x_0, y_0). In (1) $\partial f/\partial x$ is the slope of the red line which is tangent to the green curve resulting from the intersection of $f(x, y)$ and the plane $y = y_0$. In (2) $\partial f/\partial y$ can be identified in a similar fashion.*

Notation: From now on we will employ the following shorter notation for the partial derivatives of $f(x, y)$

$$\frac{\partial f}{\partial x} = f_x, \qquad \frac{\partial f}{\partial y} = f_y. \tag{2.48}$$

We will denote by $f_x(x_0, y_0)$ and $f_y(x_0, y_0)$ the partial derivatives at the point (x_0, y_0).

Definition: Given a function $f(x,y)$, the function is said to be **differentiable** if f_x and f_y exist. If the function is differentiable its first-order derivatives can be differentiated again and we can define the **second-order partial derivatives** of $f(x,y)$ as follows:

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} = f_{xx}, \tag{2.49}$$

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y} = f_{xy}, \tag{2.50}$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2} = f_{yy}, \tag{2.51}$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x} = f_{yx}. \tag{2.52}$$

Note: Notice that f_{xy} and f_{yx} are in principle different functions. Using the definitions (2.42)-(2.43) twice we can see that the second derivatives f_{xy} and f_{yx} are given by the double limits

$$\begin{aligned} f_{xy} &= \lim_{p\to 0} \frac{f_y(x+p,y) - f_y(x,y)}{p} \qquad (2.53)\\ &= \lim_{p\to 0}\left[\lim_{h\to 0} \frac{f(x+p,y+h) - f(x+p,y) - f(x,y+h) + f(x,y)}{hp}\right], \\ f_{yx} &= \lim_{h\to 0} \frac{f_x(x,y+h) - f_x(x,y)}{h} \qquad (2.54)\\ &= \lim_{h\to 0}\left[\lim_{p\to 0} \frac{f(x+p,y+h) - f(x+p,y) - f(x,y+h) + f(x,y)}{hp}\right]. \end{aligned}$$

Therefore, the only difference between f_{xy} and f_{yx} is the order in which the limits are taken. **It is not guaranteed that the limits commute**.

Example 1: Let us compute the first- and second-order partial derivatives of the function

$$f(x,y) = e^{xy} + \ln(x^2+y). \tag{2.55}$$

We start with the 1st derivatives,

$$f_x = ye^{xy} + \frac{2x}{x^2+y}, \tag{2.56}$$

$$f_y = xe^{xy} + \frac{1}{x^2+y}. \tag{2.57}$$

Now we can compute the 2nd derivatives,

$$f_{xy} = \frac{\partial f_y}{\partial x} = e^{xy} + xye^{xy} - \frac{2x}{(x^2+y)^2}, \tag{2.58}$$

$$f_{yx} = \frac{\partial f_x}{\partial y} = e^{xy} + xye^{xy} - \frac{2x}{(x^2+y)^2}\,. \tag{2.59}$$

So, in this case $f_{xy} = f_{yx}$.

Example 2: As we said at the beginning of this section, all definitions for functions of two variables extend easily to functions of 3 or more variables. In this example let us consider the function of three variables

$$g(x,y,z) = e^{x-2y+3z}, \tag{2.60}$$

and compute its 1st and 2nd order partial derivatives. In this case we have 3 1st order derivatives

$$g_x = e^{x-2y+3z}, \qquad g_y = -2e^{x-2y+3z}, \qquad g_z = 3e^{x-2y+3z}. \tag{2.61}$$

Now we have in total 9 possible different 2nd order derivatives,

$$g_{xx} = e^{x-2y+3z}, \qquad g_{yx} = -2e^{x-2y+3z}, \qquad g_{zx} = 3e^{x-2y+3z}, \tag{2.62}$$
$$g_{xy} = -2e^{x-2y+3z}, \qquad g_{yy} = 4e^{x-2y+3z}, \qquad g_{zy} = -6e^{x-2y+3z}, \tag{2.63}$$
$$g_{xz} = 3e^{x-2y+3z}, \qquad g_{yz} = -6e^{x-2y+3z}, \qquad g_{zz} = 9e^{x-2y+3z}. \tag{2.64}$$

Definition: Consider a function of two variables $f(x,y)$ and let f,f_x, f_y, f_{xx}, f_{yy}, f_{xy} and f_{yx} exist and be **continuous** in a neighbourhood of a point (x_0,y_0). Then

$$f_{xy}(x_0,y_0) = f_{yx}(x_0,y_0). \tag{2.65}$$

A nice proof of this theorem is given in chapter 13 of the book "Calculus" by R. Adams. The key idea is that the continuity of all 1st and 2nd order derivatives of f allows us to proof that the order of the limits in (2.53) and (2.54) is irrelevant for the final result.

2.3.2 Chain rules

Definition: The chain rule for functions of one variable is a formula that gives the derivative of the composition of two functions f and g (you have used this last year):

$$\frac{df(g(x))}{dx} = \lim_{h\to 0}\frac{f(g(x+h)) - f(g(x))}{h} = \frac{df}{dg}\frac{dg}{dx} = f'(g(x))g'(x). \tag{2.66}$$

Example: Let $f(x) = 1/t + t^2$ and $x = t^2 + t + 1$, then

$$\frac{df}{dx} = \frac{df}{dt}\frac{dt}{dx} = \left(-\frac{1}{t^2} + 2t\right)(2t+1)^{-1}$$

Definition: The chain rule for functions of two variables becomes considerably more complicated than for functions of one variable. Suppose we have a function

$$z = f(u(t), v(t)) = g(t), \tag{2.67}$$

and we want to know how the function f changes with respect to the variable t. The way to evaluate that change is to compute the derivative dg/dt, which according to the definition is

$$\begin{aligned}
\frac{dg}{dt} &= \lim_{h\to 0}\frac{g(t+h)-g(t)}{h} = \lim_{h\to 0}\frac{f(u(t+h),v(t+h))-f(u(t),v(t))}{h} \\
&= \lim_{h\to 0}\underbrace{\frac{f(u(t+h),v(t+h))-f(u(t),v(t+h))}{h}}_{\text{increment with } v(t+h) \text{ fixed}} \\
&+ \lim_{h\to 0}\underbrace{\frac{f(u(t),v(t+h))-f(u(t),v(t))}{h}}_{\text{increment with } u(t) \text{ fixed}} \\
&= \frac{\partial f}{\partial u}\frac{du}{dt}+\frac{\partial f}{\partial v}\frac{dv}{dt} \equiv f_u u_t + f_v v_t. \qquad (2.68)
\end{aligned}$$

In the 2th and 3th line of (2.68) we have managed to separate the 1st line into the sum of two quotients by adding zero to the 1st line in an smart way. The first of these quotients (2th line) involves changes only on the first variable $u(t)$ whereas in the second quotient (3th line) we have changes only in the second variable $v(t)$. Now we can use the chain rule for functions of one variable (2.66) to obtain the final expression in the 4th line.

Generalizations of the chain rule: The formula above can be also generalized to the case of functions

$$z = f(x(s,t), y(s,t)) = g(s,t), \tag{2.69}$$

that is, when the variables x and y are functions of two other variables s and t. In this case the chain rule tells us that the partial derivatives f_s and f_t can be obtained as

$$g_s = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s}+\frac{\partial f}{\partial y}\frac{\partial y}{\partial s} \equiv f_x x_s + f_y y_s, \tag{2.70}$$

$$g_t = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t}+\frac{\partial f}{\partial y}\frac{\partial y}{\partial t} \equiv f_x x_t + f_y y_t, \tag{2.71}$$

provided that f_x and f_y are continuous functions. Notice that these equations can be written in matrix form

$$\left(\frac{\partial f}{\partial s}, \frac{\partial f}{\partial t}\right) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \begin{pmatrix} \frac{\partial x}{\partial s} & \frac{\partial x}{\partial t} \\ \\ \frac{\partial y}{\partial s} & \frac{\partial y}{\partial t} \end{pmatrix}. \tag{2.72}$$

This matrix form is very convenient for generalizations to functions of more than 2 variables. For example, given a function of n variables $f(x_1, \ldots x_n)$ such that

$$x_i = x_i(y_1, \ldots, y_m) \quad \forall \quad i = 1, \ldots, n, \tag{2.73}$$

the chain rule can be written as

$$\left(\frac{\partial f}{\partial y_1}, \ldots, \frac{\partial f}{\partial y_m}\right) = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right) \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_m} \end{pmatrix}, \tag{2.74}$$

or equivalently

$$\frac{\partial f}{\partial y_j} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\frac{\partial x_i}{\partial y_j} \quad \forall \quad i = 1, \ldots, n, \quad \text{and} \quad j = 1, \ldots, m. \tag{2.75}$$

provided that the 1st order partial derivatives $\partial f/\partial x_i$ are continuous. The matrix in (2.74) is called the **Jacobian matrix of the variable transformation**.

Let us consider now several examples:

Example 1: Consider the function

$$f(x, y) = \sin(x + y) \quad \text{with} \quad x = st^2 \quad \text{and} \quad y = s^2 + 1/t. \tag{2.76}$$

Compute f_s and f_t in two possible ways.
Two possible ways in which we can compute these partial derivatives are

- by using the chain rule,
- or by replacing x and y in $f(x, y)$ by their expressions in terms of s and t and then computing f_s and f_t directly.

If we use the chain rule we will need the following partial derivatives:

$$f_x = \cos(x + y), \qquad f_y = \cos(x + y), \tag{2.77}$$
$$x_s = t^2, \qquad x_t = 2ts, \tag{2.78}$$
$$y_s = 2s, \qquad y_t = -1/t^2. \tag{2.79}$$

Then f_s and f_t are simply given by

$$\begin{aligned} f_s &= f_x x_s + f_y y_s = \cos(x+y)(2s+t^2) \\ &= \cos(st^2+s^2+\frac{1}{t})(2s+t^2), \end{aligned} \tag{2.80}$$

$$\begin{aligned} f_t &= f_x x_t + f_y y_t = \cos(x+y)(2ts-\frac{1}{t^2}) \\ &= \cos(st^2+s^2+\frac{1}{t})(2ts-\frac{1}{t^2}). \end{aligned} \tag{2.81}$$

The second way to compute these derivatives is to substitute x and y in terms of s and t in $f(x,y)$. By doing that we obtain

$$f(x(s,t),y(s,t)) = \sin(st^2+s^2+1/t). \tag{2.82}$$

Now we can obtain f_s and f_t directly as

$$f_s = (t^2+2s)\cos(st^2+s^2+1/t), \tag{2.83}$$
$$f_t = (2st-1/t^2)\cos(st^2+s^2+1/t). \tag{2.84}$$

Notice that here we have called $f(x,y)$ and $f(x(s,t),y(s,t))$ both f (before, we have been using different names). We will keep doing this in the future, as it makes things simpler.

Example 2: Consider now a function of three variables $f(x,y,z)$ with $x=g(z)$ and $y=h(z)$. How can we compute the derivative df/dz?

We can again apply the chain rule now for a function of three variables which in this case are all functions of the same variable z. We need only to use our general formula (2.75) with $n=3$ and $m=1$, that is

$$\frac{df}{dz} = \frac{\partial f}{\partial x}\frac{dx}{dz} + \frac{\partial f}{\partial y}\frac{dy}{dz} + \frac{\partial f}{\partial z}. \tag{2.85}$$

Example 3: Suppose that the temperature T of a certain liquid varies with the depth of the liquid z and the time t as $T(z,t)=e^{-t}z$. What is the rate of change of the temperature with respect to the time at a point that is moving through the liquid in such a way that at time t its depth is $z=f(t)$? What is this rate if $f(t)=e^t$? What is happening in this case?

Here we have an example of a function of two variables $T(z,t)$ and they tell us to compute $\partial T/\partial t$ for a point such that $z=f(t)$. This is a clear case when we can use the chain rule

$$\frac{dT}{dt} = \frac{\partial T}{\partial z}\frac{dz}{dt} + \frac{\partial T}{\partial t} = e^{-t}f'(t) - ze^{-t} = e^{-t}f'(t) - f(t)e^{-t}. \tag{2.86}$$

If in particular $f(t) = e^t$, then the previous formula gives

$$\frac{dT}{dt} = e^{-t} f'(t) - f(t)e^{-t} = 1 - 1 = 0. \qquad (2.87)$$

In this case the decrease in temperature due to the increase of depth and the decrease in temperature due to the increase of time are perfectly balanced in such a way that the temperature does not change with time.

2.3.3 Definition of differential

As in previous sections, it is useful to start this section by recalling the definition of differential for functions of one variable:

Definition: Given a function $f(x)$ and assuming that its total derivative df/dx exists at a certain point x, the total differential df of the function is given by

$$df = \left(\frac{df}{dx}\right) dx = f'(x)dx. \qquad (2.88)$$

The quantity df can be interpreted as the infinitesimal change on the value of the function $f(x)$ when x changes by the infinitesimal amount dx. A mathematical way of expressing this is to define

$$\Delta f = f(x + \Delta x) - f(x), \qquad (2.89)$$

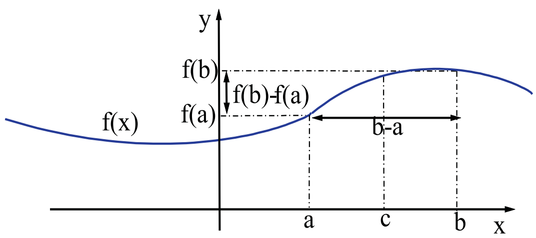
where Δf and Δx are finite increments of f and x. Then one needs to prove that

$$\Delta f \to df \qquad \text{as} \qquad \Delta x \to dx.$$

To prove this we can use the mean value theorem† which allows us to rewrite (2.89) as

$$\Delta f = f'(x + \theta\Delta x)\Delta x \quad \text{with} \quad \theta \in (0,1). \qquad (2.93)$$

†Given a function $f(x)$ which is continuous and has continuous 1st total derivative $f'(x)$ the mean value theorem tells us that if a, b are points at which the function takes values $f(a)$ and $f(b)$ and $a < b$, then a third point c exists, $c \in (a,b)$ such that



$$\frac{f(b) - f(a)}{b - a} = f'(c). \qquad (2.90)$$

Now we can take the limit when $\Delta x \to dx$ which implies $\Delta f \to df$ and dx, df being infinitesimal increments. Since dx is very small we can use $dx \ll x$, to obtain

$$df = f'(x)dx. \tag{2.94}$$

which is nothing but (2.88).

Definition: A similar quantity can be defined for functions of more than one variable. For example, let us consider now a function of two variables $f(x, y)$ with continuous first order partial derivatives f_x and f_y. We define the total differential of f,

$$df = \left(\frac{\partial f}{\partial x}\right) dx + \left(\frac{\partial f}{\partial y}\right) dy = f_x dx + f_y dy, \tag{2.95}$$

as the small variation experienced by f when the variables x and y are changed by infinitesimal amounts dx and dy respectively. As before we can define

$$\begin{aligned} \Delta f &= f(x+\Delta x, y+\Delta y) - f(x, y) \\ &= \underbrace{f(x+\Delta x, y+\Delta y) - f(x, y+\Delta y)}_{\text{increment with } y+\Delta y \text{ fixed}} + \underbrace{f(x, y+\Delta y) - f(x, y)}_{\text{increment with } x \text{ fixed}}, \end{aligned} \tag{2.96}$$

now we have managed to split Δf into two pieces, each of which involves a variation only in x and only in y, respectively. These two terms are now analogous to the definition (2.89) for a function of one variable and this allows us to write

$$\Delta f = f_x(x+\theta_1\Delta x, y+\Delta y)\Delta x + f_y(x, y+\theta_2\Delta y)\Delta y, \quad \text{with} \quad \theta_1, \theta_2 \in (0, 1). \tag{2.97}$$

Now, as for functions of one variable, in the limit $\Delta x \to dx$ and $\Delta y \to dy$ with dx, dy being infinitesimal increments ($dx \ll x$ and $dy \ll y$), then $\Delta f \to df$ and we recover the result (2.95).

The differential is a very useful concept when we want to obtain approximate values of a function nearby a point at which the value of the function and its partial derivatives are known. A good example of this is example 2 below. Example 1 is a practical example of how to compute the differential of a function of two variables:

Example 1: The fundamental equation which characterizes an ideal gas is

$$RT = PV, \tag{2.98}$$

If we define

$$\theta = \frac{c-a}{b-a}, \tag{2.91}$$

we have $\theta \in (0, 1)$ and we can rewrite (2.90) as

$$f(b) - f(a) = f'(a + \theta(b-a))(b-a). \tag{2.92}$$

where R is a universal constant and P, V, T are the three state variables (pressure, volume and temperature of the gas). Obtain the change in the pressure of the gas due to a small change of its volume and temperature.

This is a typical exercise in which they ask us to compute the differential of the pressure dP as a function of the differential of volume dV and temperature dT. We only need to use the general formula (2.95) and the relation

$$P(V,T) = R\frac{T}{V}, \tag{2.99}$$

which follows from (2.98) and we obtain

$$dP = \left(\frac{\partial P}{\partial V}\right)_T dV + \left(\frac{\partial P}{\partial T}\right)_V dT = R\left(\frac{-T}{V^2}\right) dV + \left(\frac{R}{V}\right) dT. \tag{2.100}$$

The sub-indices V and T in the partial derivatives above only indicate which variable remains constant. For example $(\partial P/\partial V)_T$ is the partial derivative of the pressure with respect to the volume at constant temperature.

Example 2: Use differentials to estimate the value of $\sqrt{27}\sqrt[3]{1021}$.

Let us start by defining the function

$$f(x,y) = \sqrt{x}\sqrt[3]{y}. \tag{2.101}$$

We can easily obtain the value of this function at the point $(x,y) = (25, 1000)$,

$$f(25, 1000) = 5 \times 10 = 50. \tag{2.102}$$

We can now exploit the fact that the point at which we want to compute the value of $f(x,y)$ is very close to $(25, 1000)$. In other words we can define

$$(27, 1021) = (x + \Delta x, y + \Delta y) = (25 + 2, 1000 + 21), \tag{2.103}$$

hence identifying $\Delta x \simeq dx = 2$ and $\Delta y \simeq dy = 21$. Employing now our formula (2.95) we have that

$$df = f_x dx + f_y dy = \frac{1}{2}\frac{\sqrt[3]{y}}{\sqrt{x}} dx + \frac{1}{3}\frac{\sqrt{x}}{y^{2/3}} dy, \tag{2.104}$$

and evaluating this formula with $x = 25$, $y = 1000$, $dx = 2$ and $dy = 21$ we obtain

$$df = \frac{1}{2}\frac{10}{5}2 + \frac{1}{3}\frac{5}{100}21 = 2 + \frac{7}{20} = 2.35. \tag{2.105}$$

Therefore, the approximate value of $f(27, 1021)$ is given by

$$f(27, 1021) \simeq f(25, 1000) + df = 52.35. \tag{2.106}$$

Now we can check if this approximation is good by calculating exactly the value

$$f(27, 1021) = \sqrt{27}\sqrt[3]{1021} = 52.3227\ldots \simeq 52.32, \tag{2.107}$$

and so the approximation is actually very good!

It is now easy to generalize the concept of differential to functions of n variables:

<u>**Definition:**</u> The total differential df of a function of n variables $f(x_1, \ldots, x_n)$ with continuous 1st order partial derivatives $f_{x_1} \ldots f_{x_n}$ is given by

$$df = f_{x_1}dx_1 + f_{x_2}dx_2 + \ldots + f_{x_n}dx_n. \tag{2.108}$$

The object df is frequently called the **1st-order differential of the function** f. The reason is that, we can define **higher order differentials** in the following way:

<u>**Definition:**</u> Consider a function $f(x, y)$ of two real variables with continuous 1st- and 2nd-order partial derivatives. We define the 2nd-order differential of f and we write it as d^2f as

$$\begin{aligned} d^2f &= d\left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy\right) = d(f_x dx + f_y dy) \\ &= (df_x)dx + (df_y)dy = (f_{xx}dx + f_{yx}dy)dx + (f_{xy}dx + f_{yy}dy)dy \\ &= f_{xx}dx^2 + 2f_{xy}dxdy + f_{yy}dy^2, \end{aligned} \tag{2.109}$$

where we assumed $f_{xy} = f_{yx}$. An equivalent way to write this result is

$$d^2f = \left(\frac{\partial}{\partial x}dx + \frac{\partial}{\partial y}dy\right)^2 f, \tag{2.110}$$

where we identify

$$\left(\frac{\partial}{\partial x}\right)^2 = \frac{\partial^2}{\partial x^2} \quad \text{and} \quad \left(\frac{\partial}{\partial y}\right)^2 = \frac{\partial^2}{\partial y^2}. \tag{2.111}$$

Now we can generalize the previous formula to the n-th differential of a function $f(x, y)$ as

Definition: The n-th differential of a function of two variables $f(x, y)$ with continuous n-th order partial derivatives is given by

$$d^n f = \left(\frac{\partial}{\partial x} dx + \frac{\partial}{\partial y} dy \right)^n f. \tag{2.112}$$

This formula can be proven by induction (which means that assuming it works for $d^{n-1}f$ we can prove that it works for $d^n f$). For that proof it is important to notice that the operator above admits the binomial expansion

$$\left(\frac{\partial}{\partial x} dx + \frac{\partial}{\partial y} dy \right)^n = \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) \left(\frac{\partial^k}{\partial x^k} \right) \left(\frac{\partial^{n-k}}{\partial y^{n-k}} \right) dy^{n-k} dx^k, \tag{2.113}$$

with

$$\left(\begin{array}{c} n \\ k \end{array} \right) = \frac{n!}{k!(n-k)!}. \tag{2.114}$$

2.3.4 Functions of several variables defined implicitly

As usual let us start by looking at the easiest example, namely functions of one variable.

Definition: We say that the function $y = f(x)$ is defined implicitly if y and x are related by an equation of the type

$$\Phi(x, y) = 0, \tag{2.115}$$

and there is no possibility of obtaining $y = f(x)$ explicitly from the constraint (2.115).

Note: Notice however that the equation (2.115) allows us to obtain the value of y for a given value of x (at least numerically), even though it does not allow us to know $f(x)$ for arbitrary x.

It is easier to understand this definition with some examples:

Example 1: Let

$$\Phi(x, y = f(x)) = \log(x + y) - \sin(x + y) = 0. \tag{2.116}$$

The constraint $F(x, y) = 0$ gives us a relation between x and y, however we are not able to obtain y as a function of x from this equation. Therefore the function $y = f(x)$ is defined implicitly through the constraint $\Phi(x, y) = 0$.