1 The Chain Rule

Recall that, for single-variable functions, the method we used to differentiate compositions of functions was called the chain rule. In general, we had

\[
\frac{d}{dx} f(g(x)) = f'(g(x))g'(x).
\]

To make things more explicit, we keep track of the dependence of the variables. My preferred way of keeping track of the variable dependence is with a tree diagram. In this case, we have the following dependences:

\[
\begin{align*}
&f \\
| \\
g \\
| \\
x
\end{align*}
\]

Figure 1: \(f\) depends on \(g\), and \(g\) in turn depends on \(x\).

We notice that each variable depends \emph{explicitly} on only a single other variable (here we do not consider the fact that the second variable may depend explicitly on another variable). The chain rule as we have stated it is sufficient to handle such derivatives.

The question arises, however, as to what we do when one (or more) variable depend on \emph{multiple} variables.

Consider the example of climbing a hill and asking how quickly you are rising at a particular point in time. The height of the hill (say \(H\)) depends on \emph{two} variables (say \(x\) and \(y\)). The function \(H(x, y)\) gives the elevation on the hill at the point \((x, y)\).

The position you are at, however, depends on the path you take. We can imagine parametrizing this by a time variable (say \(t\)) and introducing
the dependences \( x(t) \) and \( y(t) \). In other words, \( x(t) \) and \( y(t) \) tell you where you are on the hill at time \( t \). It follows that the height along your path (say \( h(t) \)) is given by \( h(t) = H(x(t), y(t)) \). Represented as a tree diagram, the variable dependences for this case are:

\[
\begin{align*}
H & \quad x \quad y \\
\quad t & \quad \quad t
\end{align*}
\]

Figure 2: \( H \) depends on \( x \) and \( y \), and both \( x \) and \( y \) depend on \( t \).

How do we answer the question of how quickly we are ascending the hill at a particular point in time? The single-dimensional chain rule is not sufficient because \( H \) depends on two variables. The question is, however, still sensible in terms of the formal definition of a derivative. We have

\[
h'(t) = \lim_{\Delta t \to 0} \frac{h(t + \Delta t) - h(t)}{\Delta t}
= \lim_{\Delta t \to 0} \frac{H(x(t + \Delta t), y(t + \Delta t)) - H(x(t), y(t))}{\Delta t}
= \lim_{\Delta t \to 0} \frac{H(x(t + \Delta t), y(t + \Delta t)) - H(x(t), y(t + \Delta t))}{\Delta t}
+ \lim_{\Delta \tilde{t} \to 0} \frac{H(x(\tilde{t} + \Delta \tilde{t}), y(\tilde{t} + \Delta \tilde{t})) - H(x(\tilde{t}), y(\tilde{t}))}{\Delta \tilde{t}}
\]

where we have set \( \tilde{t} = t + \Delta t \) and \( \Delta \tilde{t} = -\Delta t \).

We recognize this as two one-dimensional chain rule problems! We suspect, based on this and our rules for taking partial derivatives, that

\[
h'(t) = H_x(x(t), y(t))x'(t) + H_y(x(t), y(t))y'(t).
\]
This can be rewritten more explicitly in terms of partial derivatives as

\[ \frac{d}{dt} H(x(t), y(t)) = \frac{\partial H}{\partial x} \frac{dx}{dt} + \frac{\partial H}{\partial y} \frac{dy}{dt}. \]

This is, in fact, exactly the right formula! In general, we have

\[ \frac{d}{dt} H(x_1(t), x_2(t), \ldots, x_n(t)) = \frac{\partial H}{\partial x_1} \frac{dx_1}{dt} + \cdots + \frac{\partial H}{\partial x_n} \frac{dx_n}{dt}. \]

**Note:** The easiest way to remember this rule is by considering the tree diagram of variable dependences. If we write all the variable dependences as a tree diagram, the branches which lead from root to the independent variable will give the relevant partial derivatives we will have to take!

![Figure 3](image)

Figure 3: To remember the chain rule formula, write out the tree diagram and circle all the paths (from top to bottom) which take the dependent variable to the independent variable. Those branches correspond to the sequence of partial derivatives which must be taken! The result from each branch is then added together to give the overall derivative.

**Example 1:** Suppose \( z = x^2 + y^2, \ x(t) = t \cos(t), \) and \( y(t) = t \cos(t). \) Find \( \frac{dz}{dt} \) using the multivariate chain rule, then determine it is correct by substituting the functions in and taking the derivative directly.

**Solution:** We have

\[ \frac{\partial z}{\partial x} = 2x, \quad \frac{\partial z}{\partial y} = 2y, \quad \frac{\partial x}{\partial t} = \cos(t) - t \sin(t) \]

and \( \frac{\partial y}{\partial t} = \cos(t) + t \sin(t). \)
It follows that
\[
\frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt} = (2x)(\cos(t) - t \sin(t)) + 2y(\sin(t) + t \cos(t))
\]
\[
= 2t \cos(t)(\cos(t) - t \sin(t)) + 2t \sin(t)(\sin(t) + t \cos(t))
\]
\[
= 2t \cos^2(t) - 2t^2 \cos(t) \sin(t) + 2t \sin(t) + 2t^2 \sin(t) \cos(t)
\]
\[
= 2t(\cos^2(t) + \sin^2(t)) = 2t.
\]

We can verify this result by substituting the functions in directly. We have
\[
z = x^2 + y^2 = (t \cos(t))^2 + (t \sin(t))^2 = t^2(\cos^2(t) + \sin^2(t)) = t^2.
\]

It follows immediately that
\[
\frac{dz}{dt} = 2t.
\]

(In this case it was easier to plug in directly and then take the derivative. That is not the general trend! It was very nice in this case that we could easily check whether the chain rule produced a valid results—we can see that it did.)

**Example 2:** Suppose \( f(x,y,z) = xyz \), \( x(t) = t^2 \), \( y(t) = \sin(t) \), and \( z(t) = e^t \). Determine \( \frac{df}{dt} \).

**Solution:** From the chain rule formula, we have
\[
\frac{df}{dt}(x,y,z) = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} + \frac{\partial f}{\partial z} \frac{dz}{dt}.
\]

We have
\[
\frac{\partial f}{\partial x} = yz = \sin(t)e^t
\]
\[
\frac{\partial f}{\partial y} = xz = t^2e^t
\]
\[
\frac{\partial f}{\partial z} = xy = t^2 \sin(t)
\]
\[
\frac{dx}{dt} = 2t
\]
\[
\frac{dy}{dt} = \cos(t)
\]
\[
\frac{dz}{dt} = e^t.
\]
It follows that
\[
\frac{d}{dt} f(x, y, z) = 2t \sin(t)e^t + t^2 \cos(t)e^t + t^2 \sin(t)e^t = te^t(2 \sin(t) + t \cos(t) + t \sin(t)).
\]

2 The Gradient

There is a more convenient way to write the chain rule that depends on notions from linear algebra. We notice that
\[
\frac{d}{dt} f(x(t), y(t)) = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt}
\]
can be written
\[
\frac{d}{dt} f(x(t), y(t)) = \left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) \cdot \left( \frac{dx}{dt}, \frac{dy}{dt} \right)
\]
where by \((\cdot, \cdot)\) we denote the dot product operation.

The vector of first partial derivatives arises so often (as a result of the multivariate chain rule, usually) that it is given a special name. It is called the gradient.

**Definition 2.1.** The gradient of a multivariate function \(f(x_1, x_2, \ldots, x_n)\) is the vector composed component-wise of the first partial derivatives with respect to the indexed internal variable. It is denoted by
\[
\text{grad}(f) = \nabla f = \left( \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right).
\]

It follows from this definition that the chain rule can be written in the more compact algebraic form
\[
\frac{d}{dt} f(x(t)) = \nabla f(x) \cdot \frac{dx}{dt}
\]
where \(x = (x_1, x_2, \ldots, x_n)\) and \(\frac{dx}{dt} = \left( \frac{dx_1}{dt}, \ldots, \frac{dx_n}{dt} \right)\).

The gradient has the following properties:

1. The gradient at a point, \(\nabla f(a, b) = (f_x(a, b), f_y(a, b))\) is perpendicular to the level curve through the point \((a, b)\).
2. At any point \((a, b)\) the direction \(\nabla f(a, b)\) corresponds to the **direction of steepest ascent**. The magnitude of the (instantaneous) ascend is 
\[
|\nabla f(a, b)| = \sqrt{f_x(a, b)^2 + f_y(a, b)^2}.
\]

3. At any point \((a, b)\) the direction \(-\nabla f(a, b)\) corresponds to the **direction of steepest descent**. The magnitude of the (instantaneous) descent is 
\[
|\nabla f(a, b)| = \sqrt{f_x(a, b)^2 + f_y(a, b)^2}.
\]

**Example 1:** Verify that the gradient of a function at a point is perpendicular to the level curve through the point for the function \(f(x, y) = x^2 + y^2\).

**Solution:** We have that the level curves of \(f(x, y)\) are given by 
\[
x^2 + y^2 = C
\]
which are circles of radius \(\sqrt{C}\) centred at \((0, 0)\). We can find the slope of the tangent line at a point by taking the derivative of this expression **implicitly** with respect to \(x\). We have
\[
2x + 2yy' = 0 \implies y' = -\frac{x}{y}.
\]
At the point \((a, b)\) we can parametrize a vector with this slope in two-dimensions by the vector \((1, -a/b)\), or \((b, -a)\). Now consider the gradient. We have
\[
\nabla f(x, y) = (f_x(x, y), f_y(x, y)) = (2x, 2y) = 2(x, y)
\]
which implies that at the point \((a, b)\) we have \(\nabla f(a, b) = 2(a, b)\). We can clearly see that, at any point \((a, b)\), we have
\[
2(a, b) \cdot (b, -a) = 2ab - 2ab = 0.
\]
This implies that the gradient is perpendicular to the tangent line at every point \((a, b)\). This is perhaps seen most dramatically, however, through the picture contained in Figure 4.

**Example 2:** Determine the direction and magnitude of steepest ascend for the function \(f(x, y) = \sin(xy)\) at the point \((x, y) = (1, \pi)\).

**Solution:** The direction is steepest ascent is in the direction of the gradient. We have
\[
\nabla f(x, y) = (f_x(x, y), f_y(x, y)) = (y \cos(xy), x \cos(xy)).
\]
Figure 4: The level curves of $f(x, y) = x^2 + y^2$ and the gradient are perpendicular to one another.

At the point $(x, y) = (1, \pi)$ we have

$$\nabla f(1, \pi) = (\pi \cos(\pi), \cos(\pi)) = (-\pi, -1).$$

It follows that the direction of steepest ascent is in the direction $(-\pi, -1)$ and the magnitude of the ascent is

$$\sqrt{f_x(a, b)^2 + f_y(a, b)^2} = \sqrt{(-\pi)^2 + (-1)^2} = \sqrt{\pi^2 + 1}.$$

### 3 Directional derivatives

We can use this information to answer a question we posed in an earlier lecture. Suppose we are on a hill and want to know how quickly we are ascending/descending if we travel in an arbitrary direction.

Recall that we already know how to handle this problem is we move only in the $x$ or $y$ direction—this is given by the partial derivative with respect to $x$ or $y$, respectively. Now we are moving in a direction which is some combination of the $x$ and $y$ direction. How do we approach this problem?

We can set it up as a chain rule problem! Imagine the question of how quickly we are ascending/descending if we move from a point $(a, b)$ in the
direction of \((1, 2)\). We can set this up as the derivative

\[
\frac{d}{dt} f(a + t, b + 2t).
\]

This is a chain rule problem where \(f\) depends on \(x\) and \(y\), and \(x\) and \(y\) both depend on \(t\). We can compute this by the chain rule formula to get

\[
\frac{d}{dt} f(a + t, b + 2t) = \nabla f(x, y) \cdot \left(\frac{dx}{dt}, \frac{dy}{dt}\right) = (f_x(a, b), f_y(a, b)) \cdot (1, 2).
\]

In order words, we just need to find the first partial derivatives and take the dot product with the direction we are headed!

**Note:** This is not quite right! Consider taking the directional derivative along \((2, 4)\). This is in the same direction as \((1, 2)\), so the slope has to be the same. We notice, however, that

\[
(f_x(a, b), f_y(a, b)) \cdot (2, 4) \neq (f_x(a, b), f_y(a, b)) \cdot (1, 2).
\]

What have we done wrong?

The answer comes from interpretation. When we took the partial derivatives with respect to \(x\) and \(y\) (and standard derivatives as well), our interpretation was always that the derivative gave us the change in the dependent variable per unit change in the independent variable. But what is the per unit change for an arbitrary direction in space?

The answer comes from linear algebra. We need to consider the unit vector in the direction of the given vector. If \(v = (v_1, v_2, \ldots, v_n)\) is the direction in question, the unit vector in the direction of \(v\) is given by the formula

\[
\frac{v}{|v|} = \frac{1}{\sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}}(v_1, v_2, \ldots, v_n).
\]

For our example, we have that

\[
\frac{(2, 4)}{\sqrt{(2)^2 + (4)^2}} = \frac{(1, 2)}{\sqrt{(1)^2 + (2)^2}} = \left(1 \sqrt{5}, 2 \sqrt{5}\right)
\]

so that the directional derivative at the point \((a, b)\) in the direction of either \((1, 2)\) or \((2, 4)\) is given by

\[
(f_x(a, b), f_y(a, b)) \cdot \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right).
\]
**Example 1:** Find the directional derivative of the function \( f(x, y) = xy^2 \) at the point \((1, 1)\) in the direction of \((3, -4)\).

**Solution:** We have \( f_x(x, y) = y^2 \) and \( f_y(x, y) = 2xy \) so that \( f_x(1, 1) = 1 \) and \( f_y(1, 1) = 2 \). The unit vector in the direction of \((3, -4)\) is given by

\[
\frac{(3, -4)}{|(3, -4)|} = \frac{1}{\sqrt{(3)^2 + (-4)^2}} (3, -4) = \left( \frac{3}{5}, -\frac{4}{5} \right).
\]

It follows that the directional derivative at \((1, 1)\) in the direction of \((3, -4)\) is given by

\[
(f_x(1, 1), f_y(1, 1)) \cdot \frac{(3, -4)}{|(3, -4)|} = (1, 1) \cdot \left( \frac{3}{5}, -\frac{4}{5} \right) = \frac{3}{5} - \frac{4}{5} = -\frac{1}{5}.
\]

That is to say, at the point \((1, 1)\), the tangent line to the curve \( f(x, y) = xy^2 \) in the direction of \((3, -4)\) has a slope of \(-1/5\) (with respect to unit changes in that direction).

**Example 2:** Consider the curve \( f(x, y) = 1 - x^2 - y^2 \). In which directions does the tangent line at \((1, 1)\) produce no rise or fall in \( f(x, y) \)? Give the corresponding unit vectors.

**Solution:** We need to find a direction \( \mathbf{v} = (v_1, v_2) \) such that the directional derivative at \((-1, 1)\) is zero, that is to say, values \((v_1, v_2)\) such that

\[
(f_x(-1, 1), f_y(-1, 1)) \cdot \frac{(v_1, v_2)}{\sqrt{v_1^2 + v_2^2}} = 0.
\]

We don’t have to worry about the unit scaling \( \sqrt{v_1^2 + v_2^2} \) since any scaling of the unit vector will still give zero.

We have \( f_x(x, y) = -2x \) and \( f_y(x, y) = -2y \) so that \( f_x(-1, 1) = 2 \) and \( f_y(-1, 1) = -2 \). It follows that we need

\[
(2, -2) \cdot (v_1, v_2) = 0 \quad \Rightarrow \quad 2v_1 - 2v_2 = 0 \quad \Rightarrow \quad v_2 = v_1.
\]

If we set \( v_2 = t \), we can see that any vector satisfying \( t(1, 1) \) for \( t \in \mathbb{R} \) works. The unit vectors in this direction (positive and negative) are

\[
\mathbf{v}^{(1)} = \left( \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) \quad \text{and} \quad \mathbf{v}^{(2)} = \left( -\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right).
\]
4 Multiple independent variables and the Jacobian

So far we have only considered problems which have a single independent variable. That is to say, when we constructed our tree diagram, we found the same variable at the base of every branch in our tree. What changes when there are more than one independent variable?

The answer is... nothing! We apply the chain rule in exactly the same way. In terms of the tree diagram of variable dependence, we find all the branches leading from the dependent to the independent variable. These branches give us the sequence of partial derivatives. We then add up the result for each of the branches.

**Example 1:** Let \( f(x, y) = \sin(xy) \), \( x(s, t) = \frac{s^2}{t} \), and \( y(s, t) = \frac{t^2}{s} \). Find \( \frac{\partial f}{\partial s} \) and \( \frac{\partial f}{\partial t} \).

**Solution:** We consider the variable dependences. We have:

\[
\frac{\partial f}{\partial s} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial s} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial s}
\]
\[
\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial t}.
\]

Figure 5: \( f \) depends on \( x \) and \( y \), which both depend on \( s \) and \( t \).
We have
\[
\frac{\partial f}{\partial x} = y \cos(xy) = \frac{t^2}{s} \cos(st), \quad \frac{\partial f}{\partial y} = x \cos(xy) = \frac{s^2}{t} \cos(st),
\]
\[
\frac{\partial x}{\partial s} = 2 \frac{s}{t}, \quad \frac{\partial x}{\partial t} = -\frac{s^2}{t^2},
\]
\[
\frac{\partial y}{\partial s} = -\frac{t^2}{s^2}, \quad \text{and} \quad \frac{\partial y}{\partial t} = \frac{2t}{s}.
\]
It follows that
\[
\frac{\partial f}{\partial s} = \frac{t^2}{s} \cos(st) \left( 2 \frac{s}{t} \right) + \frac{s^2}{t} \cos(st) \left( -\frac{t^2}{s^2} \right)
\]
\[
= t \cos(st)
\]
\[
\frac{\partial f}{\partial t} = \frac{t^2}{s} \cos(st) \left( -\frac{s^2}{t^2} \right) + \frac{s^2}{t} \cos(st) \left( 2 \frac{t}{s} \right)
\]
\[
= s \cos(st).
\]
Again, this can be given more compactly in terms of linear algebra. The operations involved can be given as the linear system
\[
\begin{pmatrix}
\frac{\partial f}{\partial s} & \frac{\partial f}{\partial t}
\end{pmatrix} = \begin{pmatrix}
\frac{\partial f}{\partial x} & \frac{\partial f}{\partial y}
\end{pmatrix} \begin{pmatrix}
\frac{\partial x}{\partial s} & \frac{\partial x}{\partial t} \\
\frac{\partial y}{\partial s} & \frac{\partial y}{\partial t}
\end{pmatrix}
\]
\[
= \begin{pmatrix}
\frac{t^2}{s} \cos(st) & \frac{s^2}{t} \cos(st)
\end{pmatrix} \begin{pmatrix}
2 \frac{s}{t} & -\frac{s^2}{t^2} \\
-\frac{t^2}{s^2} & \frac{2t}{s}
\end{pmatrix}
\]
\[
= \begin{pmatrix}
t \cos(st) & s \cos(st)
\end{pmatrix}.
\]
The matrix
\[
\begin{pmatrix}
\frac{\partial x}{\partial s} & \frac{\partial x}{\partial t} \\
\frac{\partial y}{\partial s} & \frac{\partial y}{\partial t}
\end{pmatrix}
\]
comes up very often when dealing with multivariate functions. It is called the **Jacobian** matrix. For the general system of relationships
\[
y_1 = f_1(x_1, \ldots, x_n) \\
y_2 = f_2(x_1, \ldots, x_n) \\
\vdots \\
y_m = f_m(x_1, \ldots, x_n),
\]
or (more compactly in vector notation)

\[ y = f(x), \]

we have that the Jacobian is given by

\[
Df(x) = \begin{pmatrix}
\frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\
\frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n}
\end{pmatrix}.
\]

The Jacobian allows us to write the chain rule for multiple variable (dependent and independent) functions more compactly. Given the function dependences \( f(x) = f(y) \) where \( y = g(x) \), we have

\[
\nabla f(x) = \nabla f(y) \cdot Dg(x).
\]

**Example 2:** Let \( f(x, y, t) = t(x + y)^2 \), \( x(t) = 5t^2 \), and \( y(t) = t - 1 \). Determine the value of \( \frac{df}{dt} \) when \( t = 1 \).

**Solution:** We have the tree of dependence given in Figure 6.

![Figure 6: f depends on x, y, and t, and x and y depend on t.](image)

It follows by the chain rule that

\[
\frac{df}{dt} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} + \frac{\partial f}{\partial t}.
\]

We can find the derivative at \( t = 1 \) by plugging \( t = 1 \) into each of these expressions.
We have
\[
\frac{\partial f}{\partial x} = 2t(x + y) = 2t(5t^2 + t - 1) \quad \overset{t = 1}{\longrightarrow} \quad 2(1)(5(1)^2 + (1) - 1) = 10
\]
\[
\frac{\partial f}{\partial y} = 2y(x + y) = 2t(5t^2 + t - 1) \quad \overset{t = 1}{\longrightarrow} \quad 10
\]
\[
\frac{\partial f}{\partial t} = (x + y)^2 = (5t^2 + t - 1)^2 \quad \overset{t = 1}{\longrightarrow} \quad (5(1)^2 + 1 - 1)^2 = 25
\]
\[
\frac{dx}{dt} = 10t \quad \overset{t = 1}{\longrightarrow} \quad 10(1) = 10
\]
\[
\frac{dy}{dt} = 1 \quad \overset{t = 1}{\longrightarrow} \quad 1.
\]

It follows that \( \frac{df}{dt} \) when \( t = 1 \) is
\[
\left( \frac{df}{dt} \right)_{t=1} = (10)(10) + (10)(1) + (25) = 135.
\]

5 Higher-Order Derivatives and the Chain Rule

We are often interested in taking derivatives of functions which do not depend directly on the variables with which we are taking the derivatives.

For instance, consider being asked to determine \( \frac{d^2 f}{dt^2} \) given that \( f(x, y) = x^2 + y^2 \) and \( x(t) = e^t \) and \( y(t) = \sin(t) \). We can determine easily from the chain rule that
\[
\frac{df}{dt} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt}
\]
\[
= (2x)(e^t) + (2y)(\cos(t))
\]
\[
= 2e^{2t} + 2\sin(t)\cos(t).
\]

In order to determine the second derivative, we can simply differentiate this again with respect to \( t \). In general principle, however, we notice that each partial derivative in the first-order chain rule expression can depend on \( t \) so that in applying an addition differential operator \( \frac{\partial}{\partial x} \), we arrive at the expression
\[
\frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 f}{\partial x^2} \left( \frac{dx}{dt} \right)^2 + \frac{\partial f}{\partial x} \left( \frac{d^2 x}{dt^2} \right) + \frac{\partial^2 f}{\partial y^2} \left( \frac{dy}{dt} \right)^2
\]
\[
+ \frac{\partial f}{\partial y} \left( \frac{d^2 y}{dt^2} \right) + 2 \frac{\partial^2 f}{\partial x \partial y} \left( \frac{dx}{dt} \right) \left( \frac{dy}{dt} \right).
\]

13
We can verify that this expression works. Directly differentiating (1) we have
\[ \frac{d^2 f}{dt^2} = 4e^{2t} + 2\cos^2(t) - 2\sin^2(t). \]
We also have that
\[ \frac{\partial f}{\partial x} = 2x, \quad \frac{\partial f}{\partial y} = 2y, \quad \frac{\partial^2 f}{\partial x^2} = 2, \quad \frac{\partial^2 f}{\partial y^2} = 2, \quad \frac{\partial f}{\partial x\partial y} = 0, \]
\[ \frac{dx}{dt} = e^t, \quad \frac{d^2 x}{dt^2} = e^t, \quad \frac{dy}{dt} = \cos(t), \quad \frac{d^2 y}{dt^2} = -\sin(t). \]
It follows by the formula, without even computing the first partial derivative with respect to \( t \), that we have
\[ \frac{d^2 f}{dt^2} = 2(e^t)^2 + 2x(e^t) + 2\cos(t) - 2y\sin(t) \]
\[ = 2e^{2t} + 2e^t + 2\cos(t) - 2\sin^2(t) \]
\[ = 4e^{2t} + 2\cos^2(t) - 2\sin^2(t). \]

Notes:

1. A key application of this topic is **change of variable** of partial differential equations. Many physical problems (e.g. heat distribution, vibrations in a beam, wave motion, etc.) are modeled by partial differential equations. It turns out that change their variable system can often lead to very powerful insight into their behaviours.

2. This methodology can be generalized to more complicated chains of variable dependent by keeping in mind that each partial derivative in itself depends on all of the same variables as the original function.

**Example 1:** Suppose that \( u(x, y) \) satisfies the **one-dimensional wave equation**
\[ \frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}. \] (2)
Use the variable transformations \( r = x + ct \) and \( s = x - ct \) to show that \( u(r, s) \) satisfies the simpler form
\[ \frac{\partial^2 u}{\partial r\partial s} = 0. \] (3)
Show that \( u(r, s) = f(r) + g(s) \) where \( f(r) \) and \( g(s) \) are arbitrary functions is a solution to (3) and consequently \( u(x, y) = f(x + ct) + g(x - ct) \) is a
Solution: The equation (2) can be derived (although not in this class) from physical assumptions consistent with wave motion. Corresponding, the solutions are given by two waves (the functions $f$ and $g$) travelling in the left and right direction at a fixed velocity $c$.

To show the equivalence of the two systems (2) and (3), we evaluate the derivatives $\frac{\partial^2 u}{\partial t^2}$ and $\frac{\partial^2 u}{\partial x^2}$ using the intermediate variable dependences. We have

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial r} \frac{\partial r}{\partial x} + \frac{\partial u}{\partial s} \frac{\partial s}{\partial x} = \frac{\partial u}{\partial r} + \frac{\partial u}{\partial s}$$

and

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial r} \frac{\partial r}{\partial t} + \frac{\partial u}{\partial s} \frac{\partial s}{\partial t} = c \frac{\partial u}{\partial r} - c \frac{\partial u}{\partial s}.$$ 

Keeping in mind that each partial derivatives itself depends on $r$ and $s$, which depend on $x$ and $t$, when applying another derivative with respect to $x$ and $t$ we must apply the chain rule again. (For example,

$$\frac{\partial}{\partial x} \frac{\partial u}{\partial s} = \left[ \frac{\partial}{\partial s} \frac{\partial u}{\partial s} \right] \frac{\partial s}{\partial x} + \left[ \frac{\partial}{\partial t} \frac{\partial u}{\partial s} \right] \frac{\partial t}{\partial x} = \frac{\partial^2 u}{\partial s^2} \frac{\partial x}{\partial x} + \frac{\partial^2 u}{\partial s \partial r} \frac{\partial r}{\partial x}.$$ 

It follows that we have

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 u}{\partial r^2} + 2 \frac{\partial u}{\partial r} \frac{\partial s}{\partial x} + \frac{\partial^2 u}{\partial s^2}$$

and

$$\frac{\partial^2 u}{\partial t^2} = c^2 \left( \frac{\partial^2 u}{\partial r^2} - 2 \frac{\partial u}{\partial r} \frac{\partial s}{\partial s} + \frac{\partial^2 u}{\partial s^2} \right).$$

It follows that

$$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = -4c^2 \frac{\partial^2 u}{\partial r \partial s} = 0 \implies \frac{\partial^2 u}{\partial r \partial s} = 0.$$ 

This is the resulting partial differential equation in the new variable system.

We can clearly see that choosing $u(r, s) = f(r) + g(s)$ gives

$$\frac{\partial u}{\partial r} = f'(r) \implies \frac{\partial^2 u}{\partial r \partial s} = 0$$

or (alternatively)

$$\frac{\partial u}{\partial s} = g'(s) \implies \frac{\partial^2 u}{\partial s \partial r} = 0.$$
It follows that $u(r, s) = f(r) + g(s)$ satisfies (2).

The variable dependences $r = x + ct$ and $s = x - ct$ suggest that $u(x, y) = f(x + ct) + g(x - ct)$ are solutions of (2). Indeed, by the chain rule, we have

$$\frac{\partial u}{\partial x} = f'(x + ct) + g'(x - ct) \implies \frac{\partial^2 u}{\partial x^2} = f''(x + ct) + g''(x - ct)$$

and

$$\frac{\partial u}{\partial t} = cf'(x + ct) - g'(x - ct) \implies \frac{\partial^2 u}{\partial x^2} = c^2 f''(x + ct) + c^2 g''(x - ct).$$

It follows that $u(x, t) = f(x + ct) + g(x - ct)$ satisfies (2).