Multivariable Calculus – Lecture #6 Notes

This lecture focuses on five things: (a) the application of gradients to find normal vectors to curves and surfaces; (b) the generalization of the Basic Chain Rule to more general compositions of functions of several variables; (c) a new and more general approach to implicit differentiation; (d) partial derivatives in the context of non-independent variables and internal constraints; and (e) the definition and meaning of 2nd derivatives of functions of several variables.

Rate of change of a function along a parameterized curve and the Basic Chain Rule
We showed in the previous lecture that if a function $f(x,y)$ is given and a parameterized curve is described by $\mathbf{r}(t)=\langle x(t),y(t)\rangle$, then we can determine the rate of change of the function $f(x,y)$ as we travel along this parameterized curve by the Basic Chain Rule.

If we think of this as a composition, we have:

$$t \to (x(t),y(t)) \to f(x(t),y(t))$$

We showed that $\boxed{\frac{d}{dt}\left[f(x(t),y(t))\right]=\frac{\partial f}{\partial x}\frac{dx}{dt}+\frac{\partial f}{\partial y}\frac{dy}{dt}}$. This is the **Basic Chain Rule**.

The same construction can be done with a differentiable function $f(x,y,z)$ and a parameterized curve $\mathbf{r}(t)=\langle x(t),y(t),z(t)\rangle$ to give the rate of change $\boxed{\frac{d}{dt}\left[f(x(t),y(t),z(t))\right]=\frac{\partial f}{\partial x}\frac{dx}{dt}+\frac{\partial f}{\partial y}\frac{dy}{dt}+\frac{\partial f}{\partial z}\frac{dz}{dt}}$ as the Basic Chain Rule in this context.

In either case (or in an even more general context, we see that $\boxed{\frac{df}{dt}=\overrightarrow{\nabla f}\cdot\mathbf{v}}$ where the **gradient vector** $\overrightarrow{\nabla f}=\left\langle\frac{\partial f}{\partial x},\frac{\partial f}{\partial y}\right\rangle$ in the former case and $\overrightarrow{\nabla f}=\left\langle\frac{\partial f}{\partial x},\frac{\partial f}{\partial y},\frac{\partial f}{\partial z}\right\rangle$ in the latter case. It gives a vector at every point, i.e. a **vector field**.

We investigated the geometry of the gradient and found that at every point <u>the gradient vector will be perpendicular to the level set of the given function passing through any given point</u>. This gives us a remarkably simple way to determine normal vectors to curves and surfaces and, with this, a simple way to determine equations for tangent lines to curves and tangent planes to surfaces.

Example 1: Find an equation for the tangent line to the curve defined by the equation $x^2y+2xy^3=8$ at the point $(2,1)$.

Solution: If we let $f(x,y)=x^2y+2xy^3$, then this curve is, in fact, the $f=8$ level set (level curve or *contour*). (You may want to verify that $f(2,1)=8$.). We calculate the gradient vector $\overrightarrow{\nabla f}=\langle 2xy+2y^3, x^2+6xy^2\rangle$. At the point $(2,1)$ this gives the vector $\overrightarrow{\nabla f}(2,1)=\langle 6,16\rangle=2\langle 3,8\rangle$, and we know that this must be perpendicular to the $f=8$ level set at this point. We can therefore take $\mathbf{n}=\langle 3,8\rangle$ as a normal vector to the line tangent to this level set. Using the relation $\mathbf{n}\cdot(\mathbf{x}-\mathbf{x}_0)=0$, we have $\langle 3,8\rangle\cdot\langle x-2,y-1\rangle=0$ or $3(x-2)+8(y-1)=0$ or $3x+8y=14$.

Example 2: Find an equation for the tangent plane to the surface defined by the equation $xyz+2xz^3=10$ at the point $(2,3,1)$.

Solution: If we let $f(x,y,z) = xyz + 2xz^3$, then this surface is, in fact, the $f = 10$ level set (level surface). (You may want to verify that $f(2,3,1) = 10$.). We calculate the gradient vector $\overrightarrow{\nabla f} = \left\langle yz + 2z^3, xz, xy + 6xz^2 \right\rangle$. At the point $(2,3,1)$ this gives the vector $\overrightarrow{\nabla f}(2,3,1) = \langle 5,2,18 \rangle$, and we know that this must be perpendicular to the $f = 10$ level set at this point. We can therefore take $\mathbf{n} = \langle 5,2,18 \rangle$ as a normal vector to the plane tangent to this level set. Using the relation $\mathbf{n} \cdot (\mathbf{x} - \mathbf{x}_0) = 0$, we have $\langle 5,2,18 \rangle \cdot \langle x-2, y-3, z-1 \rangle = 0$ or
$5(x-2) + 2(y-3) + 18(z-1) = 0$ or $5x + 2y + 18z = 34$.

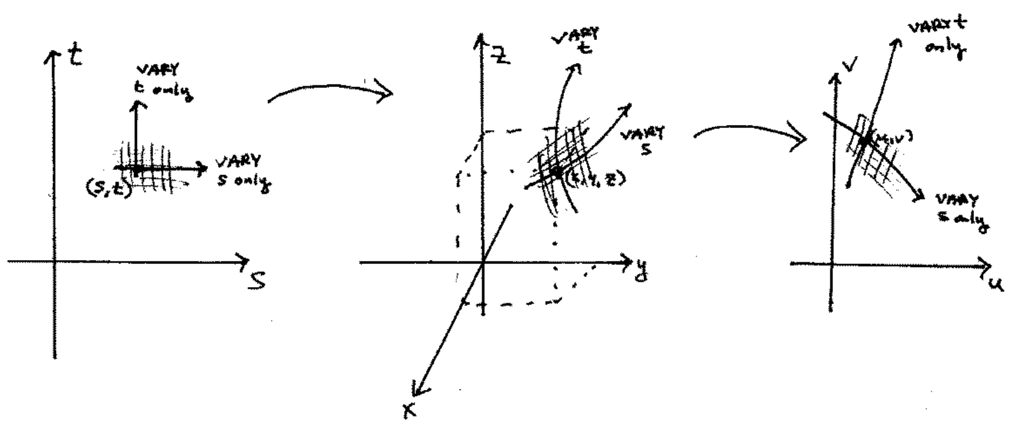
Note: In order to use this method to find normal vectors to curves or surfaces, you may have to transpose any variables in a given equation to one side of the equation leaving only a constant on the other side before defining a function by the variable expression on one side of this equation.

The General Chain Rule
In general, the chain rule is an algebraic rule that describes how to calculate rates of change of functions built from other functions through composition. For example, in a first semester calculus course we learn that if $y = y(u)$ and $u = u(x)$, then we can calculate $\dfrac{dy}{dx}$ by the chain rule: $\dfrac{dy}{dx} = \dfrac{dy}{du}\dfrac{du}{dx}$. In a multivariable setting, we might have $z = z(x,y)$ and $x = x(t),\ y = y(t)$. We then have $\dfrac{dz}{dt} = \dfrac{\partial z}{\partial x}\dfrac{dx}{dt} + \dfrac{\partial z}{\partial y}\dfrac{dy}{dt}$ by the basic chain rule.

The chain rule gets more interesting when you apply it to situations where there are more input variables and output variables. For example, let us suppose we have a situation where there are two parameters, *s* and *t*, and that for any *s* and *t* we have equations giving $\left\{ \begin{matrix} x = x(s,t) \\ y = y(s,t) \\ z = z(s,t) \end{matrix} \right\}$. Let us further suppose that for any choices of the variables *x*, *y*, and *z* we have two other variables, *u* and *v*, defined by equations $\left\{ \begin{matrix} u = u(x,y,z) \\ v = v(x,y,z) \end{matrix} \right\}$.

In this case we can think of this functionally as:

$$(s,t) \xrightarrow{G} (x,y,z) \xrightarrow{F} (u,v).$$



If we vary *s* only (holding *t* constant) and only focus on how the output variable *u* will change, the Basic Chain Rule gives that $\dfrac{\partial u}{\partial s} = \dfrac{\partial u}{\partial x}\dfrac{\partial x}{\partial s} + \dfrac{\partial u}{\partial y}\dfrac{\partial y}{\partial s} + \dfrac{\partial u}{\partial z}\dfrac{\partial z}{\partial s}$. Note that all of the derivatives are now partial derivatives.

We can do the same by selectively varying either *s* or *t* and focusing selectively on the output variables *u* or *v*.

We calculate:

$$\frac{\partial u}{\partial s}=\frac{\partial u}{\partial x}\frac{\partial x}{\partial s}+\frac{\partial u}{\partial y}\frac{\partial y}{\partial s}+\frac{\partial u}{\partial z}\frac{\partial z}{\partial s} \qquad \frac{\partial u}{\partial t}=\frac{\partial u}{\partial x}\frac{\partial x}{\partial t}+\frac{\partial u}{\partial y}\frac{\partial y}{\partial t}+\frac{\partial u}{\partial z}\frac{\partial z}{\partial t}$$

$$\frac{\partial v}{\partial s}=\frac{\partial v}{\partial x}\frac{\partial x}{\partial s}+\frac{\partial v}{\partial y}\frac{\partial y}{\partial s}+\frac{\partial v}{\partial z}\frac{\partial z}{\partial s} \qquad \frac{\partial v}{\partial t}=\frac{\partial v}{\partial x}\frac{\partial x}{\partial t}+\frac{\partial v}{\partial y}\frac{\partial y}{\partial t}+\frac{\partial v}{\partial z}\frac{\partial z}{\partial t}$$

There is nothing especially difficult about these seemingly complicated relationships. We simply selectively focus on a particular output variable (*u* or *v*), and then calculate partial derivatives (with respect to either *s* or *t*) by treating the other one as though constant. Unlike the Basic Chain Rule, all derivatives are now partial derivatives because all functions are functions of several variables. In each case there are as many terms as there are variables in the middle of the composition.

These equations can be organized into a statement about the <u>Jacobian matrices</u> of the two functions and of their composition. A Jacobian matrix may be thought of simply as an array of (partial) derivatives of the various output variables with respect to the various input variables, where the outputs are listed from top to bottom and the inputs are listed from left to right. If you know about matrix multiplication, we have:

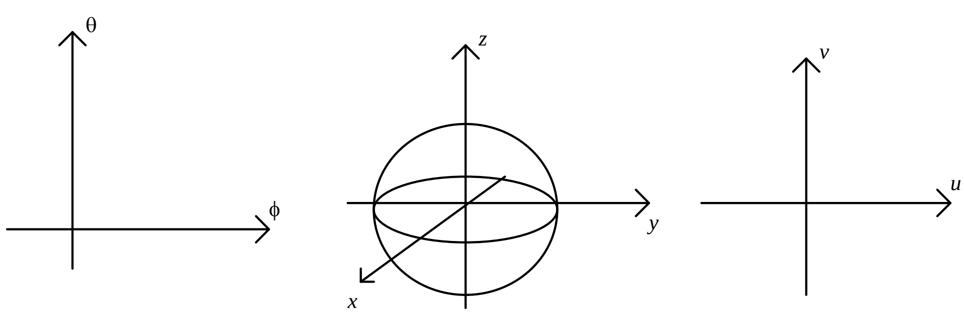
$$\begin{bmatrix} \frac{\partial u}{\partial s} & \frac{\partial u}{\partial t} \\ \frac{\partial v}{\partial s} & \frac{\partial v}{\partial t} \end{bmatrix} = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \end{bmatrix} \begin{bmatrix} \frac{\partial x}{\partial s} & \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial s} & \frac{\partial y}{\partial t} \\ \frac{\partial z}{\partial s} & \frac{\partial z}{\partial t} \end{bmatrix} \text{ or, more succinctly, } \boxed{\mathbf{J}_{F\circ G}=\mathbf{J}_F\mathbf{J}_G}.$$

Note: It's worth mentioning that the rows of each matrix look like *gradient* vectors, and the columns look like *velocity* vectors. This view of the Chain Rule can be explained in terms of how incremental vectors or tangent vectors in the original domain are transformed to their counterparts in the image spaces. It is really a statement of how the composition differentiable functions can be approximated by a composition of the linear transformations defined by the respective Jacobian matrices.

To picture what this is telling us, let's specifically look at the situation where ϕ and θ represent latitude and longitude with the minor change that latitude will be measured from the north pole as 0°, the equator as 90°, and the south pole as 180°. We can then describe a sphere of radius *R* by the parametric equations

$$\begin{Bmatrix} x = R\cos\theta\sin\phi \\ y = R\sin\theta\sin\phi \\ z = R\cos\phi \end{Bmatrix}.$$

Let us further suppose that the variables *u* and *v* measure, for example, temperature and barometric pressure at any point (x, y, z) in $\mathbf{R}^3$ and, in particular, at points on this parametrized sphere in $\mathbf{R}^3$.



We might ask questions about how temperature would vary as we change latitude or longitude, or how barometric pressure would vary as we change latitude or longitude. These are the quantities in the Jacobian matrix $\mathbf{J}_{F\circ G} = \begin{bmatrix} \frac{\partial u}{\partial \phi} & \frac{\partial u}{\partial \theta} \\ \frac{\partial v}{\partial \phi} & \frac{\partial v}{\partial \theta} \end{bmatrix}$. The rows of the Jacobian matrix $\mathbf{J}_F = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \end{bmatrix}$ are just the gradient vectors (in $\mathbf{R}^3$) of the temperature and barometric pressure functions. (Note that these are functions defined on $\mathbf{R}^3$ and not just on the spherical surface.)

The two columns of the Jacobian matrix $\mathbf{J}_G = \begin{bmatrix} \frac{\partial x}{\partial \phi} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial \phi} & \frac{\partial y}{\partial \theta} \\ \frac{\partial z}{\partial \phi} & \frac{\partial z}{\partial \theta} \end{bmatrix}$ represent "velocity" vectors tangent to the longitudes (ϕ varying) and latitudes (θ varying). These two column vectors are tangent to curves lying in the sphere and are therefore tangent to the sphere. They are, essentially, the "south vector" and the "east vector" at any point of the sphere (except at the poles). You might further observe that their cross product will be normal to this spherical surface at any given point – a fact which will be useful later in this course when we look at surface integrals.

The two columns of the Jacobian matrix $\mathbf{J}_{F\circ G} = \begin{bmatrix} \frac{\partial u}{\partial \phi} & \frac{\partial u}{\partial \theta} \\ \frac{\partial v}{\partial \phi} & \frac{\partial v}{\partial \theta} \end{bmatrix}$ represent vectors in the (u, v) plane and indicate the directions of change if we slightly vary the latitude or the longitude.

Implicitly Defined Functions and Implicit Differentiation

Often it is the case that an equation (or several equations) relate some variables and we wish to consider one variable (or several) as depending on the rest. For example, given the equation of a circle

$$x^2 + y^2 = 16$$

we may wish to consider $y = y(x)$. If we solve explicitly, we get either $y = \sqrt{16 - x^2}$ or $y = -\sqrt{16 - x^2}$ whose graphs are, respectively, the upper and lower semicircles. Though we could calculate the derivatives directly, there is an alternate approach. We can think of x as a parameter and use it to parametrize either one of the semicircles as $x \to (x, y(x))$, where the dependence of y on x is defined implicitly by the given curve (semi-circle). If we let $F(x, y) = x^2 + y^2$, then we can view the circle as just the $F = 16$ contour, or level set, of the function F. Composing these functions, we have:

$$x \to (x, y(x)) \to F(x, y(x)) = \text{constant}$$

Applying the chain rule (and using F_x and F_y to denote the partial derivatives of F), we have:

$$\frac{d}{dx} F(x, y(x)) = F_x 1 + F_y \frac{dy}{dx} = 0$$

Here we used the fact that $\frac{dx}{dx}=1$ and that the composite function was constant everywhere on this level set.

Solving for $\frac{dy}{dx}$, we get that $\boxed{\frac{dy}{dx}=-\frac{F_x}{F_y}}$. So, as long as we avoid those places where $F_y=0$ (where the two semicircles meet), we have a valid formula for calculating $\frac{dy}{dx}$. In the above example, this gives $\frac{dy}{dx}=-\frac{2x}{2y}=-\frac{x}{y}$. This may be used for either the upper or the lower semicircle.

Note: The expression $\frac{dy}{dx}=-\frac{F_x}{F_y}$ can also be derived <u>geometrically</u> by noting that $\overrightarrow{\nabla F}=\left\langle F_x,F_y\right\rangle$ will, at any point, give a vector perpendicular to the level set, so the slope of the normal line will be $\frac{F_y}{F_x}$ and the slope of the tangent line will therefore be given by its negative reciprocal, i.e. $\frac{dy}{dx}=-\frac{F_x}{F_y}$.

This formulation will be valid whenever we have a relation of the form $F(x,y)=\text{constant}$, where F is a differentiable function and where we can consider $y=y(x)$ as being implicitly defined by the equation. The only exception is at those points where $F_y=0$, i.e. at points where the tangent line to the relation is vertical.

This same approach can be used for relations of the form $F(x,y,z)=\text{constant}$, where we may wish to consider one of the variables as being dependent on the others. For example, if we choose to think of $z=z(x,y)$ defined implicitly by the given relation, then it is useful to consider x and y as <u>parameters</u> and to formulate the situation as

$$(x,y)\to(x,y,z(x,y))\to F(x,y,z(x,y))=\text{constant}$$

Here we can think of the relation as a surface in $\mathbf{R}^3$, and what this is saying is that by choosing (x,y) we may find one point (or several points) on the graph. We can apply the chain rule to calculate the partial derivatives of the composition with respect to the parameters x and y. What makes this a bit tricky is the fact that x and y are playing dual roles as parameters and as coordinates in $\mathbf{R}^3$. Nonetheless, we have

$$\begin{aligned}\frac{\partial}{\partial x}F(x,y,z(x,y))&=F_x\cdot 1+F_y\cdot 0+F_z\frac{\partial z}{\partial x}=0\\ \frac{\partial}{\partial y}F(x,y,z(x,y))&=F_x\cdot 0+F_y\cdot 1+F_z\frac{\partial z}{\partial y}=0\end{aligned}\qquad (*)$$

which enable us to solve for $\boxed{\frac{\partial z}{\partial x}=-\frac{F_x}{F_z}}$ and $\boxed{\frac{\partial z}{\partial y}=-\frac{F_y}{F_z}}$. These expressions will be valid wherever F is differentiable and where $F_z\neq 0$. It should be relatively clear that this same formulation could be done for relations with any number of variables and would give analogous expressions for the partial derivatives of the implicitly defined functions.

Note: The equations (*) above can also be interpreted in terms of dot products and perpendicularity. They give that $\left\langle F_x, F_y, F_z \right\rangle \cdot \left\langle 1, 0, \frac{\partial z}{\partial x} \right\rangle = \overrightarrow{\nabla F} \cdot \mathbf{v}_x = 0$ and $\left\langle F_x, F_y, F_z \right\rangle \cdot \left\langle 0, 1, \frac{\partial z}{\partial y} \right\rangle = \overrightarrow{\nabla F} \cdot \mathbf{v}_y = 0$ where you should recognize the vectors $\mathbf{v}_x$ and $\mathbf{v}_y$ as tangent vectors to the $z = z(x, y)$ graph surface. Once again, we see that the gradient vector $\overrightarrow{\nabla F}$ is perpendicular to tangent vectors to the $F(x, y, z) = \text{constant}$ level surface and therefore perpendicular to this level surface at any given point on the surface.

Note: Had we instead chosen to define $x = x(y, z)$ as being defined implicitly by this relation, we would have similarly obtained the expressions $\boxed{\frac{\partial x}{\partial y} = -\frac{F_y}{F_x}}$ and $\boxed{\frac{\partial x}{\partial z} = -\frac{F_z}{F_x}}$, and these expressions will be valid wherever F is differentiable and where $F_x \neq 0$. We might also have chosen to define $y = y(x, z)$ as being defined implicitly by this relation, and we would then obtain the expressions $\boxed{\frac{\partial y}{\partial x} = -\frac{F_x}{F_y}}$ and $\boxed{\frac{\partial y}{\partial z} = -\frac{F_z}{F_y}}$. These expressions will be valid wherever F is differentiable and where $F_y \neq 0$.

Partial Derivatives in the case of Non-Independent Variables and Internal Constraints

We occasionally are confronted with situations where variable quantities are subject to internal constraints, i.e. they are interrelated via one or more equations. In this context, calculating partial derivatives and rates of change are more complicated and, in fact, even the notation of partial derivatives becomes ambiguous.

Example: Consider the function defined algebraically by $w = f(x, y, z) = x^2 + y^2 - 2xz$ where the variables are related by the equation (constraint) $y + z^2 = 12$.

If we ignore the constraint, we calculate the following partial derivatives: $\left\{ \begin{array}{l} w_x = \frac{\partial w}{\partial x} = 2x - 2z \\ w_y = \frac{\partial w}{\partial y} = 2y \\ w_z = \frac{\partial w}{\partial z} = -2x \end{array} \right\}$.

However, if we absorb the constraint by writing $y = 12 - z^2$ and substitute to get $w = f(x, y(x, z), z) = x^2 + (12 - z^2)^2 - 2xz$, we might then calculate the (two) partial derivates to get $\frac{\partial w}{\partial x} = 2x - 2z$ and $\frac{\partial w}{\partial z} = 2(12 - z^2)(-2z) - 2x = -48z + 4z^3 - 2x = -4yz - 2x$ if we use $y = 12 - z^2$. Note that in the first instance we obtained $\frac{\partial w}{\partial z} = -2x$ and in the second instance we obtained $\frac{\partial w}{\partial z} = -4yz - 2x$. These are contradictory, so there seems to be either some inconsistency or perhaps some ambiguity in the notation.

This mystery can be unraveled if we think more carefully about composition of functions and the Chain Rule and if we also amend our notation somewhat to handle circumstances like this.

For clarity, let's modify the notation using subscripts to indicate which variables are being treated as constants when taking the partial derivative. That is, let's write $w_x = \left(\frac{\partial w}{\partial x} \right)_{y,z}$, $w_y = \left(\frac{\partial w}{\partial y} \right)_{x,z}$, and $w_z = \left(\frac{\partial w}{\partial z} \right)_{x,y}$. In these we have not absorbed the constraint to reduce this to a function of just two variables.

In the case where we have absorbed the constraint to get $w = f(x, y(x,z), z) = x^2 + (12 - z^2)^2 - 2xz = w(x,z)$, we'll write the two partial derivatives as $\left(\frac{\partial w}{\partial x}\right)_z$ and $w_z = \left(\frac{\partial w}{\partial z}\right)_x$.

There are at least two good ways to relate these partial derivatives: (a) the **Chain Rule**; and (b) **differentials**. As a composition of functions, we have: $(x,z) \to (x, y(x,z), z) \to w$. **Applying the Chain Rule**, we have:

$$\left\{\begin{aligned} \left(\frac{\partial w}{\partial x}\right)_z &= \left(\frac{\partial w}{\partial x}\right)_{y,z} \cdot 1 + \left(\frac{\partial w}{\partial y}\right)_{x,z} \cdot \left(\frac{\partial y}{\partial x}\right)_z + \left(\frac{\partial w}{\partial z}\right)_{x,y} \cdot 0 = w_x + w_y \left(\frac{\partial y}{\partial x}\right)_z = (2x - 2z) + (2y)(0) = 2x - 2z \\ \left(\frac{\partial w}{\partial z}\right)_x &= \left(\frac{\partial w}{\partial x}\right)_{y,z} \cdot 0 + \left(\frac{\partial w}{\partial y}\right)_{x,z} \cdot \left(\frac{\partial y}{\partial z}\right)_x + \left(\frac{\partial w}{\partial z}\right)_{x,y} \cdot 1 = w_z + w_y \left(\frac{\partial y}{\partial z}\right)_x = (-2x) + (2y)(-2z) = -2x - 4yz \end{aligned}\right\}$$

These agree with the calculations above where we explicitly made the substitution $y = 12 - z^2$ prior to taking derivatives.

Differentials are in some ways a simpler way to address the potential ambiguities. If we write $w = w(x,y,z)$ in the case where we do not involve the internal constraint, and $w = w(x,z)$ in the case where we have absorbed the constraint, then the more universal notion of differentials allows us to write:

$dw = \left(\frac{\partial w}{\partial x}\right)_{y,z} dx + \left(\frac{\partial w}{\partial y}\right)_{x,z} dy + \left(\frac{\partial w}{\partial z}\right)_{x,y} dz = w_x dx + w_y dy + w_z dz$ and $dw = \left(\frac{\partial w}{\partial x}\right)_z dx + \left(\frac{\partial w}{\partial z}\right)_x dz$. We can separately look at the constraint $y + z^2 = 12$ and write $dy + 2z dz = 0$ or $dy = -2z dz$. Substitution gives:

$$dw = (2x - 2z)dx + (2y)(-2z dz) - 2x dz = (2x - 2z)dx + (-4yz - 2x)dz = \left(\frac{\partial w}{\partial x}\right)_z dx + \left(\frac{\partial w}{\partial z}\right)_x dz$$

So we can conclude that $\left(\frac{\partial w}{\partial x}\right)_z = 2x - 2z$ and $\left(\frac{\partial w}{\partial z}\right)_x = -2x - 4yz$.

This situation of non-independent variables comes up in **thermodynamics and physical chemistry** where the pressure (*P*), volume (*V*), and absolute temperature (*T*) of an ideal gas is related through the Ideal Gas Law, $\boxed{PV = nRT}$, where *n* is the number of moles of gas and *R* is the ideal gas constant. There are a number of thermodynamical quantities of interest, including the entropy (*S*) of the gas where $S = S(P,V,T)$. We may be interested in knowing the rate of change of entropy as we vary each of the three thermodynamical variables <u>which cannot vary independently due to the Ideal Gas Law</u>.

We can differentiate to relate the differentials for the Ideal Gas Law to get $\boxed{PdV + VdP = nRdT}$ from which we can easily solve for any one of the differential quantities in terms of the other two.

(a) If we choose to think of volume and temperature as varying independently and the pressure then being determined by the Ideal Gas Law, we might consider the rates $\left(\frac{\partial S}{\partial V}\right)_T$ and $\left(\frac{\partial S}{\partial T}\right)_V$. If so, using differentials we can write

$$dS = S_P dP + S_V dV + S_T dT = S_P\left(\frac{nRdT - PdV}{V}\right) + S_V dV + S_T dT = \left(S_V - \frac{P}{V}S_P\right)dV + \left(S_T + \frac{nR}{V}S_P\right)dT.$$

So $\left(\frac{\partial S}{\partial V}\right)_T = S_V - \frac{P}{V}S_P$ and $\left(\frac{\partial S}{\partial T}\right)_V = S_T + \frac{nR}{V}S_P$.

(b) If we choose to think of pressure and temperature as varying independently and the volume then being determined by the Ideal Gas Law, we might consider the rates $\left(\dfrac{\partial S}{\partial P}\right)_T$ and $\left(\dfrac{\partial S}{\partial T}\right)_P$. If so, using differentials we can write

$$dS = S_P dP + S_V dV + S_T dT = S_P dP + S_V\left(\frac{nRdT - VdP}{P}\right) + S_T dT = \left(S_P - \frac{V}{P}S_V\right)dP + \left(S_T + \frac{nR}{P}S_V\right)dT .$$

So $\left(\dfrac{\partial S}{\partial P}\right)_T = S_P - \dfrac{V}{P}S_V$ and $\left(\dfrac{\partial S}{\partial T}\right)_P = S_T + \dfrac{nR}{P}S_V$.

(c) If we choose to think of pressure and volume as varying independently and the temperature then being determined by the Ideal Gas Law, we might consider the rates $\left(\dfrac{\partial S}{\partial P}\right)_V$ and $\left(\dfrac{\partial S}{\partial V}\right)_P$. If so, using differentials we can write

$$dS = S_P dP + S_V dV + S_T dT = S_P dP + S_V dV + S_T\left(\frac{PdV + VdP}{nR}\right) = \left(S_P + \frac{V}{nR}S_T\right)dP + \left(S_V + \frac{P}{nR}S_T\right)dV .$$

So $\left(\dfrac{\partial S}{\partial P}\right)_V = S_P + \dfrac{V}{nR}S_T$ and $\left(\dfrac{\partial S}{\partial V}\right)_P = S_V + \dfrac{P}{nR}S_T$.

Note: All of the above derivatives may be expressed in alternate ways using the relation $PV = nRT$.

Second derivatives and higher order derivatives of functions of several variables

If we think of partial derivatives as rates of change, then we can say that for a function $f(x,y)$,

$f_x = \dfrac{\partial f}{\partial x} =$ "rate of change of the values of *f* with respect to (increasing) *x*" $\leftrightarrow$ "*x*-slope"

$f_y = \dfrac{\partial f}{\partial y} =$ "rate of change of the values of *f* with respect to (increasing) *y*" $\leftrightarrow$ "*y*-slope"

Continuing with these interpretations, we can define (and interpret) 2nd derivatives as follows:

$f_{xx} = \dfrac{\partial}{\partial x}\left(\dfrac{\partial f}{\partial x}\right) = \dfrac{\partial^2 f}{\partial x^2} =$ "rate of change of the *x*-slopes with respect to (increasing) *x*"

$f_{xy} = \dfrac{\partial}{\partial y}\left(\dfrac{\partial f}{\partial x}\right) = \dfrac{\partial^2 f}{\partial y \partial x} =$ "rate of change of the *x*-slopes with respect to (increasing) *y*"

$f_{yx} = \dfrac{\partial}{\partial x}\left(\dfrac{\partial f}{\partial y}\right) = \dfrac{\partial^2 f}{\partial x \partial y} =$ "rate of change of the *y*-slopes with respect to (increasing) *x*"

$f_{yy} = \dfrac{\partial}{\partial y}\left(\dfrac{\partial f}{\partial y}\right) = \dfrac{\partial^2 f}{\partial y^2} =$ "rate of change of the *y*-slopes with respect to (increasing) *y*"

If we organize these 2nd derivatives into a 2×2 matrix, we call this the **Hessian matrix** of this function:

$$H_f = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$$

The first and last of these 2[nd] derivatives are relatively simple to interpret as concavities of the cross-sectional curves in the graph surface – just as was the case with 2[nd] derivatives of functions of a single variable. The "mixed partial derivatives" f_{xy} and f_{yx} are a little more difficult to interpret. You may find it helpful to imagine a graph surface and use a straight object to represent a tangent vector and imagine how the *x*-slope might change as you move laterally in the *y*-direction, e.g. if there was a little twist in the graph surface. This is what the mixed partial derivative f_{xy} would measure. Similarly, you can interpret f_{yx} by imagining how the *y*-slope might change as you move laterally in the *x*-direction. In fact, these two mixed partial derivatives are generally equal, though this is certainly not obvious. Indeed, this is the essence of Clairaut's Theorem.

Clairaut's Theorem: If a function of two or more variables is differentiable and if its first and second derivatives are continuous, then mixed partial derivatives are equal.

In the case of a function of two variables, this simply means that $f_{xy} = f_{yx}$.

For functions of three variables, we can define 9 second derivatives and organize them into a 3×3 Hessian matrix:

$$H_f = \begin{bmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{bmatrix}$$

Clairaut's Theorem in this case gives that $f_{xy} = f_{yx}$, $f_{xz} = f_{zx}$, and $f_{yz} = f_{zy}$.

In either this case or the previous case, Clairaut's Theorem means that the Hessian matrix is a *symmetric* matrix.

We could also consider higher order derivatives, and Clairaut's Theorem would continue to apply (assuming all derivatives are continuous functions. For example, for a function $f(x,y)$, there would be 8 third partial derivatives: f_{xxx}, f_{xxy}, f_{xyx}, f_{yxx}, f_{yyx}, f_{yxy}, f_{xyy}, and f_{yyy}. However, by Clairaut's Theorem, we would have $f_{xxy} = f_{xyx} = f_{yxx}$ and $f_{xyy} = f_{yxy} = f_{yyx}$. Try calculating some of these to see Clairaut's Theorem in action.

For a function of three variables $f(x,y,z)$, there would be 3 first partial derivatives, 9 second partial derivatives, and 27 third partial derivatives, but by Clairaut's Theorem the mixed partial derivatives would be equal.

Example: For the function $f(x,y) = x^2y + 2xy^3$, the two first partial derivatives are just the components of the gradient vector $\overrightarrow{\nabla f} = \left\langle f_x, f_y \right\rangle = \left\langle 2xy + 2y^3, x^2 + 6xy^2 \right\rangle$. The Hessian matrix will then be:

$$H_f = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} = \begin{bmatrix} 2y & 2x + 6y^2 \\ 2x + 6y^2 & 12xy \end{bmatrix}$$

Example: For the function $f(x,y,z) = xyz + 2xz^3$, the three first partial derivatives are just the components of the gradient vector $\overrightarrow{\nabla f} = \left\langle f_x, f_y, f_z \right\rangle = \left\langle yz + 2z^3, xz, xy + 6xz^2 \right\rangle$. The Hessian matrix will then be:

$$H_f = \begin{bmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{bmatrix} = \begin{bmatrix} 0 & z & y + 6z^2 \\ z & 0 & x \\ y + 6z^2 & x & 12xz \end{bmatrix}$$

Notes by Robert Winters