CHAPTER 2
PARTIAL DERIVATIVES

2.1 *Introduction*

Any text on thermodynamics is sure to be liberally sprinkled with partial derivatives on almost every page, so it may be helpful here to give a brief summary of some of the more useful formulas involving partial derivatives that we are likely to use in subsequent chapters.

2.2 *Partial Derivatives*

The equation
$$z = z(x, y) \qquad 2.2.1$$

represents a two-dimensional surface in three-dimensional space. The surface intersects the plane y = constant in a plane curve in which z is a function of x. One can then easily imagine calculating the slope or gradient of this curve in the plane y = constant. This slope is $\left(\frac{\partial z}{\partial x}\right)_y$ - the partial derivative of z with respect to x, with y being held constant. For example, if

$$z = y \ln x, \qquad 2.2.2$$

then
$$\left(\frac{\partial z}{\partial x}\right)_y = \frac{y}{x}, \qquad 2.2.3$$

y being treated as though it were a constant, which, in the plane y = constant, it is. In a similar manner the partial derivative of z with respect to y, with x being held constant, is

$$\left(\frac{\partial z}{\partial y}\right)_x = \ln x . \qquad 2.2.4$$

When you have only three variables – as in this example – it is usually obvious which of them is being held constant. Thus $\partial z / \partial y$ can hardly mean anything other than at constant x. For that reason, the subscript is often omitted. In thermodynamics, there are often more than three variables, and it is usually (I would say always) essential to indicate by a subscript which quantities are being held constant.

In the matter of pronunciation, various attempts are sometimes made to give a special pronunciation to the symbol ∂. (I have heard “day”, and “dye”.) My own preference is just to say “partial dz by dy”.

Let us suppose that we have evaluated z at (x , y). Now if you increase x by δx, what will the resulting increase in z be? Obviously, to first order, it is $\frac{\partial z}{\partial x}\delta x$. And if y increases by δy, the increase in z will be $\frac{\partial z}{\partial y}\delta y$. And if both x and y increase, the corresponding increase in z, to first order, will be

$$\delta z = \frac{\partial z}{\partial x}\delta x + \frac{\partial z}{\partial y}\delta y . \qquad 2.2.5$$

No great and difficult mathematical proof is needed to "derive" this; it is just a plain English statement of an obvious truism. The increase in z is equal to the rate of increase of z with respect to x times the increase in x plus the rate of increase of z with respect to y times the increase in y.

Likewise if x and y are increasing with time at rates $\frac{dx}{dt}$ and $\frac{dy}{dt}$, the rate of increase of z with respect to time is

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt} . \qquad 2.2.6$$

2.3 *Implicit Differentiation*

Equation 2.2.5 can be used to solve the problem of differentiation of an implicit function. Consider, for example, the unlikely equation

$$\ln(xy) = x^2 y^3 . \qquad 2.3.1$$

Calculate the derivative dy/dx.

It would be easy if only one could write this in the form y = something; but it is difficult (impossible as far as I know) to write y *explicitly* as a function of x. Equation 2.3.1 *implicitly* relates y to x. How are we going to calculate dy/dx?

The curve $f(x, y) = 0$ might be considered as being the intersection of the surface $z = f(x, y)$ with the plane $z = 0$. Seen thus, the derivative dy/dx can be thought of as the limit as δx and δy approach zero of the ratio $\delta y / \delta x$ within the plane $z = 0$; that is, keeping z constant and hence δz equal to zero. Thus equation 2.2.5 gives us that

$$\frac{dy}{dx} = -\left(\frac{\partial f}{\partial x}\right) \Big/ \left(\frac{\partial f}{\partial y}\right). \tag{2.3.2}$$

For example, show that, for equation 2.3.1,

$$\frac{dy}{dx} = \frac{y(2x^2y^3 - 1)}{x(1 - 3x^2y^3)}. \tag{2.3.3}$$

2.4 *Product of Three Partial Derivatives*

Suppose x, y and z are related by some equation and that, by suitable algebraic manipulation, we can write any one of the variables explicitly in terms of the other two. That is, we can write

$$x = x(y, z), \tag{2.4.1}$$

or

$$y = y(z, x), \tag{2.4.2}$$

or

$$z = z(x, y). \tag{2.4.3}$$

Then

$$\delta x = \frac{\partial x}{\partial y}\delta y + \frac{\partial x}{\partial z}\delta z, \tag{2.4.4}$$

$$\delta y = \frac{\partial y}{\partial z}\delta z + \frac{\partial y}{\partial x}\delta x \tag{2.4.5}$$

and

$$\delta z = \frac{\partial z}{\partial x}\delta x + \frac{\partial z}{\partial y}\delta y. \tag{2.4.6}$$

Eliminate δy from equations 2.4.4 and 2.4.5:

$$\delta x\left(1 - \frac{\partial x}{\partial y}\cdot\frac{\partial y}{\partial x}\right) = \delta z\left(\frac{\partial x}{\partial z} + \frac{\partial x}{\partial y}\cdot\frac{\partial y}{\partial z}\right), \tag{2.4.7}$$

and δz from equations 2.4.4 and 2.4.6:

$$\delta x\left(1 - \frac{\partial x}{\partial z}\cdot\frac{\partial z}{\partial x}\right) = \delta y\left(\frac{\partial x}{\partial y} + \frac{\partial x}{\partial z}\cdot\frac{\partial z}{\partial y}\right). \tag{2.4.8}$$

Since z and x can be varied independently, and x and y can be varied independently, the only way in which equations 2.4.7 and 2.4.8 can always be true is for all of the expressions in parentheses to be zero. Equating the left-hand parentheses to zero shows that

$$\frac{\partial x}{\partial y} = 1 \Big/ \frac{\partial y}{\partial x} \tag{2.4.9}$$

and

$$\frac{\partial x}{\partial z} = 1 \Big/ \frac{\partial z}{\partial x} \, . \tag{2.4.10}$$

These results may seem to be trivial and "obvious" – and so they are, *provided that the same quantity is being kept constant in the derivatives of both sides of each equation.* In thermodynamics we are often dealing with more variables than just x, y and z, and we must be careful to specify which quantities are being held constant. If, for example, we are dealing with several variables, such as u, v, w, x, y, z, it is not in general true that $\frac{\partial u}{\partial y} = 1 \Big/ \frac{\partial y}{\partial u}$, unless the same variables are being held constant on both sides of the equation.

Return now to equation 2.4.7. The right hand parenthesis is zero, and this, together with equation 2.4.10, results in the important relation:

$$\left(\frac{\partial x}{\partial y}\right)_z \cdot \left(\frac{\partial y}{\partial z}\right)_x \cdot \left(\frac{\partial z}{\partial x}\right)_y = -1. \tag{2.4.11}$$

2.5 *Second Derivatives and Exact Differentials*

If $z = z(x, y)$, we can go through the motions of calculating $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$, and we can then further calculate the second derivatives $\frac{\partial^2 z}{\partial x^2}$, $\frac{\partial^2 z}{\partial y^2}$, $\frac{\partial^2 z}{\partial x \, \partial y}$ and $\frac{\partial^2 z}{\partial y \, \partial x}$. It will usually be found that the last two, the mixed second derivatives, are equal; that is, it doesn't matter in which order we perform the differentiations. *Example*: Let $z = x \sin y$. Show that $\frac{\partial^2 z}{\partial x \, \partial y} = \frac{\partial^2 z}{\partial y \, \partial x} = \cos y$.

We examine in this section what conditions must be satisfied if the mixed derivatives are to be equal.

Figure II.1 depicts z as a "well-behaved" function of x and y. By "well-behaved" in this context I mean that z is everywhere single-valued (that is, given x and y there is just one value of z), finite and continuous, and that its derivatives are everywhere continuous (that is, no sudden discontinuities in either the function itself or its slope). "Good behaviour" in this sense is the sufficient condition that the mixed second derivatives are equal.

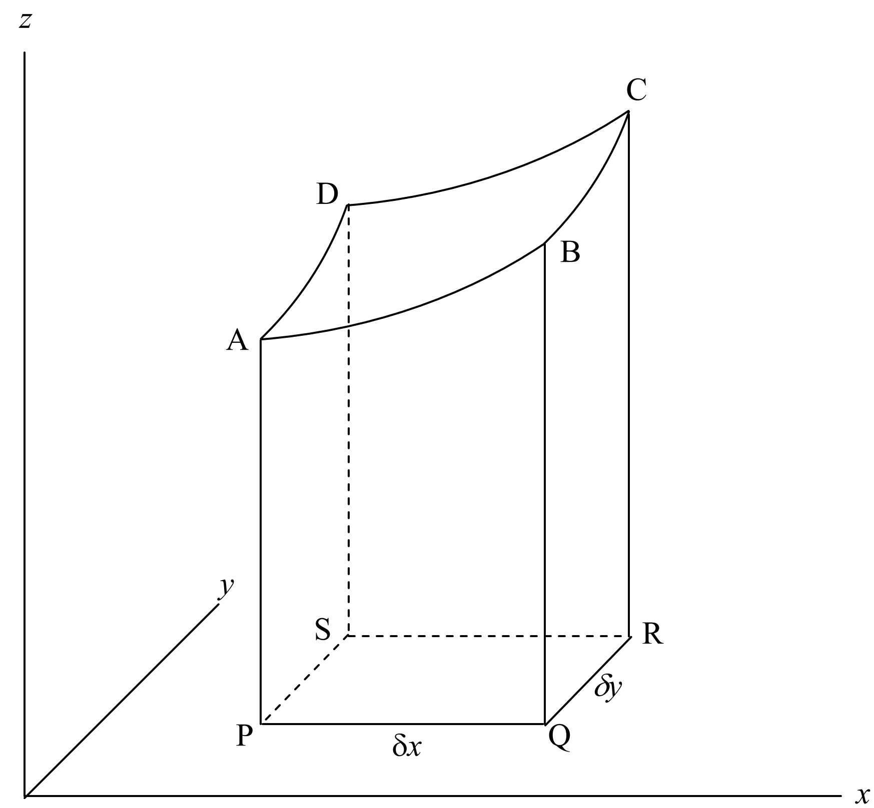

FIGURE II.1

Let us calculate the difference δz in the heights of A and C. We can go from A to C via B or via D, and δz is route-independent. That is, to first order,

$$\delta z = \left(\frac{\partial z}{\partial x}\right)_y^{(A)} \delta x + \left(\frac{\partial z}{\partial y}\right)_x^{(B)} \delta y = \left(\frac{\partial z}{\partial y}\right)_x^{(A)} \delta y + \left(\frac{\partial z}{\partial x}\right)_y^{(D)} \delta x. \qquad 2.5.1$$

Here the superscript (A) means "evaluated at A".

Divide both sides by $\delta x \, \delta y$:

$$\frac{\left(\frac{\partial z}{\partial y}\right)_x^{(B)} - \left(\frac{\partial z}{\partial y}\right)_x^{(A)}}{\delta x} = \frac{\left(\frac{\partial z}{\partial x}\right)_y^{(D)} - \left(\frac{\partial z}{\partial x}\right)_y^{(A)}}{\delta y} \,. \qquad 2.5.2$$

If we now go to the limit as δx and δy approach zero (the equation now becomes exact rather than merely "to first order"), this becomes:

$$\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x} \,. \qquad 2.5.3$$

A further property of a function that is well-behaved in the sense described is that if the differential dz can be written in the form

$$dz = A(x,y)dx + B(x,y)dy, \qquad 2.5.4$$

then equation 2.5.3 implies that

$$\frac{\partial A}{\partial y} = \frac{\partial B}{\partial x} \,. \qquad 2.5.5$$

A differential dz is said to be *exact* if the following conditions are satisfied: The integral of dz between two points is route-independent, and the integral around a closed path (i.e. you end up where you started) is zero, and if equations 2.5.3 and 2.5.5 are satisfied.

If a differential such as 2.5.4 is exact – i.e., if it is found to satisfy the conditions for exactness – then it should be possible to integrate it and determine $z(x,y)$. Let us look at an example. Suppose that

$$dz = (4x - 3y - 1)dx + (-3x + 2y + 4)dy. \qquad 2.5.6$$

It is readily seen that this is exact. The problem now, therefore, is to find $z(x,y)$.

Let $$u = \int (4x - 3y - 1)dx$$

So that $$u = 2x^2 - 3yx - x + g(y). \qquad 2.5.7$$

Note that we are treating y as constant. The "constant" of integration depends on the value of y – i.e. it is an arbitrary function of y.

Of course u is not the same as z – unless we can find a particular function $g(y)$ such that u indeed *is* the same as z.

Now $du = \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy$; that is,

$$du = (4x - 3y - 1)dx + \left(-3x + \frac{dg}{dy}\right)dy. \qquad 2.5.8$$

Then $du = dz$ (and $u = z$ plus an arbitrary constant) provided that $\frac{dg}{dy} = 2y + 4$. That is,

$$g(y) = y^2 + 4y + \text{ constant.} \qquad 2.5.9$$

Thus $$z = 2x^2 - 3xy + y^2 - x + 4y + \text{constant.} \qquad 2.5.10$$

The reader should verify that this satisfies equation 2.5.6. The reader should also try letting

$$\upsilon = -3xy + y^2 + 4y + f(x) \qquad 2.5.11$$

(where did this come from?) and go through a similar argument to arrive again at equation 2.5.10.

Consider another example

$$dz = 3\ln y\, dx + \frac{x}{y}dy. \qquad 2.5.12$$

You should immediately find that this differential is *not* exact, and, to emphasize that, I shall use the symbol $đz$, the special symbol $đ$ indicating an inexact differential. However, given an inexact differential $đz$, it is very often possible to find a function $H(x, y)$ such that the differential $dw = H(x,y)\ đz$ *is* exact, and dw can then be integrated to find w as a function of x and y. The function $H(x,y)$ is called an *integrating factor*. There may be more than one possible integrating factor; indeed it may be possible to find one simply of the form $F(x)$ or maybe $G(y)$. There are several ways for finding an integrating factor. We'll do a simple and straightforward one. Let us try and find an integrating factor for the inexact differential $đz$ above. Thus, let $dw = F(x)dz$, so that

$$dw = 3F\ln y\, dx + \frac{xF}{y}dy. \qquad 2.5.13$$

For dw to be exact, we must have

$$\frac{\partial}{\partial y}(3F\ln y) = \frac{\partial}{\partial x}\left(\frac{xF}{y}\right). \qquad 2.5.14$$

That is,
$$\frac{3F}{y} = \frac{1}{y}\left(F + x\frac{dF}{dx}\right). \qquad 2.5.15$$

Upon integration and simplification we find that

$$F = x^2, \quad \text{or any multiple thereof,} \qquad 2.5.16$$

is an integrating factor, and therefore

$$dw = 3x^2 \ln y\, dx + \frac{x^3}{y}dy \qquad 2.5.17$$

is an exact differential. The reader should confirm that this is an exact differential, and from there show that

$$w = x^3 \ln y + \text{constant}. \qquad 2.5.18$$

To anticipate – what has this to do with thermodynamics? To give an example, the *state* of many simple thermodynamical systems can be specified by giving the values of three *intensive state variables*, P, V and T, the pressure, molar volume and temperature. That is, the state of the system can be represented by a point in PVT space. Often, there will be a known relation (known as the *equation of state*) between the variables; for example, if the substance involved is an *ideal gas*, the variables will be related by $PV = RT$, which is the equation of state for an ideal gas; and the point representing the state of the system will then be represented by a point that is constrained to lie on the two-dimensional surface $PV = RT$ in three-dimensional PVT space. In that case it will be necessary to specify only two of the three variables. On the other hand, if the equation of state of a particular substance is unknown, you will have to give the values of all three variables.

Now there are certain quantities that one meets in thermodynamics that are *functions of state*. Two that come to mind are *entropy S* and *internal energy U*. By *function of state* it is meant that S and U are uniquely determined by the state (i.e. by P, V and T). If you know P, V and T, you can calculate S and U or any other *function of state*. In that case, the differentials dS and dU are *exact differentials*.

The internal energy U of a system is defined in such a manner that when you add a quantity dQ of heat **to** a system and also do an amount of work dW **on** the system, the **increase** dU in the internal energy of the system is given by $dU = dQ + dW$. Here dU is an exact differential, but dQ and dW are clearly not. You can achieve the same increase in internal energy by any combination of heat and work, and the heat you add to the system and the work you do on it are clearly not functions of the state of the system.

Some authors like to use a special symbol, such as đ, to denote an inexact differential (but beware, I have seen this symbol used to denote an *exact* differential!). I shall not in general do this, because there are many contexts in which the distinction is not important, or, if it is, it is obvious from the context whether a given differential is exact or not. If, however, there is some context in which the distinction is important (and there are many) and in which it may not be obvious which is which, I may, with advance warning, use a special đ for an inexact differential, and indeed I have already done so earlier in this section.

2.6 *Euler's Theorem for Homogeneous Functions*

There is a theorem, usually credited to Euler, concerning homogenous functions that we might be making use of.

A homogenous function of degree *n* of the variables *x*, *y*, *z* is a function in which all terms are of degree *n*. For example, the function

$$f(x,y,z) = Ax^3 + By^3 + Cz^3 + Dxy^2 + Exz^2 + Fyz^2 + Gyx^2 + Hzx^2 + Izy^2 + Jxyz,$$

is a homogeneous function of *x*, *y*, *z*, in which all terms are of degree three.

The reader will find it easy to evaluate the partial derivatives $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial x}$ and equally easy (if slightly tedious) to evaluate the expression $x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} + z\frac{\partial f}{\partial z}$. Tedious or not, I do urge the reader to do it. You should find that the answer is

$$3Ax^3 + 3By^3 + 3Cz^3 + 3Dxy^2 + 3Exz^2 + 3Fyz^2 + 3Gyx^2 + 3Hzx^2 + 3Izy^2 + 3Jxyz.$$

In other words, $x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} + z\frac{\partial f}{\partial z} = 3f$. If you do the same thing with a homogenous function of degree 2, you will find that $x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} + z\frac{\partial f}{\partial z} = 2f$. And if you do it with a homogenous function of degree 1, such as $Ax + By + Cz$, you will find that $x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} + z\frac{\partial f}{\partial z} = f$. In general, for a homogeneous function of $x, y, z, \ldots$ of degree *n*, it is always the case that

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} + z\frac{\partial f}{\partial z} + \ldots = nf. \qquad 2.6.1$$

This is Euler's theorem for homogeneous functions.

2.7 *Undetermined Multipliers*

Let $\psi(x, y, z)$ be some function of *x*, *y* and *z*. Then if *x*, *y* and *z* are independent variables, one would ordinarily understand that, where ψ is a maximum, the derivatives are zero:

$$\frac{\partial\psi}{\partial x} = \frac{\partial\psi}{\partial y} = \frac{\partial\psi}{\partial z} = 0. \qquad 2.7.1$$

However, if *x*, *y* and *z* are not completely independent, but are related by some constraining equation such as $f(x, y, z) = 0$, the situation is slightly less simple. (In a thermodynamical context, the three variables may be, for example, three "intensive state variables", *P*, *V* and *T*, and ψ might be the entropy, which is a function of state. However the intensive state variables may not be completely independent, since they are related by an "equation of state", such as $PV = RT$.)

If we move by infinitesimal displacements *dx*, *dy*, *dz* from a point where ψ is a maximum, the corresponding changes in ψ and *f* will both be zero, and therefore both of the following equations must be satisfied.

$$d\psi = \frac{\partial\psi}{\partial x}dx + \frac{\partial\psi}{\partial y}dy + \frac{\partial\psi}{\partial z}dx = 0, \qquad 2.7.2$$

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dx = 0. \qquad 2.7.3$$

Consequently any linear combination of ψ and f, such as $\Phi = \psi + \lambda f$, where λ is an arbitrary constant, also satisfies a similar equation. The constant λ is sometimes called an "undetermined multiplier" or a "Lagrangian multiplier", although often some additional information in an actual problem enables the constant to be identified.

In summary, the conditions that ψ is a maximum (or minimum or saddle point), if *x*, *y* and *z* are related by a functional constraint $f(x, y, z) = 0$, are

$$\frac{\partial\Phi}{\partial x} = 0, \quad \frac{\partial\Phi}{\partial y} = 0, \quad \frac{\partial\Phi}{\partial z} = 0, \qquad 2.7.4$$

where

$$\Phi = \psi + \lambda f. \qquad 2.7.5$$

Of course, if ψ is a function of many variables x_1, x_2, $x_3 \ldots$, and the variables are subjected to several constraints, such as $f = 0$, $g = 0$, $h = 0$, etc., where *f*, *g*, *h*, etc., are functions connecting all or some of the variables, the conditions for ψ to be a maximum (etc.) are

$$\frac{\partial \psi}{\partial x_i} + \lambda \frac{\partial \psi}{\partial x_i} + \mu \frac{\partial \psi}{\partial x_i} + \nu \frac{\partial \psi}{\partial x_i} + \ldots = 0, \qquad i = 1, 2, 3, \ldots \qquad 2.7.6$$

2.8 *Dee and Delta*

We have discussed the special meanings of the symbols ∂ and $đ$, but we also need to be clear about the meanings of the more familiar differential symbols Δ, δ and d. It is often convenient to use the symbol Δ to indicate an increment (not necessarily a particularly small increment) in some quantity. We can then use the symbol δ to mean a *small* increment. We can then say that if, for example, $y = x^2$, and if x were to increase by a small amount δx, the corresponding increment in y would be given approximately by

$$\delta y \cong 2x\,\delta x, \qquad 2.8.1$$

That is,

$$\frac{\delta y}{\delta x} \cong 2x. \qquad 2.8.2$$

This doesn't become exact until we take the limit as δx and δy approach zero. We write this limit as $\frac{dy}{dx}$, and then it is *exactly* true that

$$\frac{dy}{dx} = 2x. \qquad 2.8.3$$

There is a valid point of view that would argue that you cannot write dx or dy alone, since both are zero; you can write only the ratio $\frac{dy}{dx}$. It would be wrong, for example, to write

$$dy = 2x\,dx, \qquad 2.8.4$$

or at best it is tantamount to writing 0 = 0. I am not going to contradict that argument, but, at the risk of incurring the wrath of some readers, I am often going to write equations such as equation 2.7.4, or, more likely, in a thermodynamical context, equations such as $dU = T\,dS - P\,dV$, even though you may prefer me to say that, for small increments, $\delta U \cong T\,\delta S - P\,\delta V$. I am going to argue that, in the limit of infinitesimal increments, it is exactly true that $dU = T\,dS - P\,dV$. After all, the smaller the increments, the closer it becomes to being true, and, in the limit when the increments are infinitesimally small, it is exactly true, even if it does just mean that zero equals zero. I hope this does not cause too many conceptual problems.