Supplementary notes

Exact differentials

Introduction

These notes are a review of some parts of multi-variable calculus that are relevant to what we're doing in differential equations. One goal is to explain exact differentials. Another is to re-enforce the view that thinking in terms of differentials makes calculus more intuitive and easier to understand and use. The notation and calculations in this Introduction are explained in the following sections.

The *exact differentials* method for solving special ODEs depends on knowing when a pair of functions $f(x,y)$ and $g(x,y)$ can be the partial derivatives of a *potential* function $\psi(x,y)$. The *necessary* and *sufficient* condition is that

$$\frac{\partial f}{\partial y} = \frac{\partial g}{\partial x} \ . \tag{1}$$

Being *necessary* means that unless (1) is satisfied, there can be no ψ with

$$f = \frac{\partial \psi}{\partial x} \ , \quad g = \frac{\partial \psi}{\partial y} \ . \tag{2}$$

Being *sufficient* means that if (1) is satisfied then there is a ψ that satisfies (2). The fact that (1) is necessary comes from the fact that partial derivatives *commute* (see below).

The explanation of the fact that (1) is sufficient is *constructive*. The function ψ not only is shown to exist, but there is a way to "construct" it. You just integrate. The function you get works if the compatibility condition is satisfied. For example, suppose

$$\begin{aligned} f(x,y) &= 2xy^2 + y^3 + 2x \\ g(x,y) &= 2x^2y + 3xy^2 + 1 \ . \end{aligned}$$

The compatibility condition (1) is (notation explained below)

$$\begin{aligned} f = 2xy^2 + y^3 + 2x \quad &\xrightarrow{\partial_y} \quad f_y = 4xy + 3y^2 \\ g = 2x^2y + 3xy^2 + 1 \quad &\xrightarrow{\partial_x} \quad g_x = 4xy + 3y^2 \ . \end{aligned}$$

The calculation shows that (1) is satisfied in this example. Any ψ that satisfies $\psi_x = f$ is an anti-derivative with respect to x of f. Traditional informal calculus gives (integrating each term separately and treating y as a constant in this x integration)

$$\psi(x,y) = \int^x f(x,y)\,dx = \int^x \left(2xy^2 + y^3 + 2x \right) dx = x^2y^2 + xy^3 + x^2 + C \ .$$

The constant of integration, C could be different for each y, so we should write this as

$$\psi(x,y) = x^2y^2 + xy^3 + x^2 + C(y) \ .$$

On the other hand, $\psi_y = g(x,y)$ leads to (writing $D(x)$ for the constant of integration, which may depend on x)

$$\psi(x,y) = \int^y \left(2x^2y + 3xy^2 + 1 \right) dy = x^2y^2 + xy^3 + y + D(x) \ .$$

This leads to two equations for $\psi(x,y)$, if there is a ψ:

$$\psi(x,y) = x^2y^2 + xy^3 + x^2 + C(y) = x^2y^2 + xy^3 + y + D(x) \; .$$

If the second equality is true, then $C(y)$ must be equal to y (the only term on the right with no x dependence) and $D(x)$ must be equal to x^2. This implies that

$$\psi(x,y) = x^2y^2 + xy^3 + x^2 + y \; .$$

To be absolutely correct, you could add a constant, independent of x and y, to ψ and still satisfy $\psi_x = f$ and $\psi_y = g$. Strictly speaking, we could only infer that $C(y) = y + E$ (E being a constant independent of x and y) and similarly for $D(x)$.

Some classes explain these facts in vector form. The equations (2) are written in vector form (using various notations) as

$$(f,g) = \text{grad } \psi = \vec{\nabla}\psi = \nabla\psi \; .$$

The condition (1) is expressed, in various notations, as

$$\text{curl}\,(f,g) = \vec{\nabla}\times(f,g) = \nabla\times(f,g) = \partial_x g - \partial_y f = 0 \; .$$

It is a fact of vector calculus, in 2D as here, or in 3D, that a vector field is the gradient of a scalar (potential) field if and only if the curl of the vector field is zero. If you don't know about grad and curl, don't worry about it.

Partial derivatives

Suppose[1] $\psi(x,y)$ is a function of two variables x and y. There are several common notations for partial derivatives, including

$$\frac{\partial\psi(x,y)}{\partial x} = \partial_x\psi(x,y) = \psi_x(x,y) = \lim_{h\to 0}\frac{\psi(x+h,y)-\psi(x,y)}{h}$$
$$\frac{\partial\psi(x,y)}{\partial y} = \partial_y\psi(x,y) = \psi_y(x,y) = \lim_{h\to 0}\frac{\psi(x,y+h)-\psi(x,y)}{h}$$

You can take partial derivatives of partial derivatives, which are called *second partial derivatives*. These may be written in various ways, such as

$$\frac{\partial}{\partial y}\frac{\partial\psi}{\partial x} = \partial_y\partial_x\psi = \psi_{yx}$$
$$\frac{\partial}{\partial x}\frac{\partial\psi}{\partial x} = \partial_x\partial_x\psi = \partial_x^2\psi = \psi_{xx}$$

We sometimes talk about partial derivative *operators*, such as ∂_x or $\partial_x\partial_y$, etc. A partial differential operator "operates" on a function sort of in the way a matrix "operates" on a vector. For example, the operator ∂_x operates on the function ψ as

$$\partial_x\psi = \psi_x \; .$$

We might say that the operator ∂_x "acts on" the function ψ. This is sometimes written with arrows to indicate the operation, such as

$$x^2y^2 \;\xrightarrow{\partial_y}\; 2x^2y \; .$$

[1] This ψ is the Greek letter "psi", which is pronounced "sigh" in the US (but not in Greece). Make sure you ψ is not confused with your ϕ, which is the Greek letter "phi" pronounced like "fly" (without the "l'').

The operator $\partial_x \partial_y$ means "first differentiate with respect to y, then differentiate with respect to x:

$$\partial_x \partial_y \psi = \partial_x \left(\partial_y \psi \right) \ .$$

You differentiate with respect to y to evaluate the quantity in parentheses, then differentiate the result with respect to x. This is expressed with arrows as

$$\psi \ \xrightarrow{\partial_y} \ \psi_y \ \xrightarrow{\partial_x} \ \psi_{xy} \ .$$

This is a convenient way do practical calculations. For example

$$x^2 y^2 \ \xrightarrow{\partial_y} \ 2x^2 y \ \xrightarrow{\partial_x} \ 4xy \ .$$

The operator $\partial_x \partial_y$ acts on x^2y^2 to give $4xy$. The *order* of a partial differential operator is the number of partial derivatives it involves. Thus, ∂_x is first order, and ∂_x^2 and $\partial_x \partial_y$ are second order.

Partial derivatives *commute*. This is a term from algebra which means that the answer does not depend on the order in which you apply the operators. For example

$$\partial_x \partial_y \psi = \partial_y \partial_x \psi \ .$$

More simply,

$$\psi_{xy} = \psi_{yx} \ . \tag{3}$$

For example, doing the y derivative first gives

$$\sin(x) e^{2y} \ \xrightarrow{\partial_y} \ 2\sin(x) e^{2y} \ \xrightarrow{\partial_x} \ 2\cos(x) e^{2y} \ .$$

Doing the x derivative first gives

$$\sin(x) e^{2y} \ \xrightarrow{\partial_x} \ \cos(x) e^{2y} \ \xrightarrow{\partial_y} \ 2\cos(x) e^{2y} \ .$$

The intermediate results are different but the final result is the same.

You can understand the fact that partial derivatives commute by looking at *finite difference approximations*. The finite difference approximation to ψ_x is

$$\partial_x \psi(x,y) \approx \frac{\psi(x+\Delta x, y) - \psi(x,y)}{\Delta x} \ . \tag{4}$$

The finite difference approximation to $\partial_y(\,\partial_x \psi\,)$ is

$$\partial_y(\,\partial_x \psi(x,y)\,) \approx \frac{\psi_x(x, y+\Delta y) - \psi_x(x,y)}{\Delta y} \ . \tag{5}$$

If you use ψ_x approximation (4) in the $\partial_y \psi_x$ approximation (5), you get the complicated expression

$$\partial_y(\,\partial_x \psi(x,y)\,) \approx \frac{\dfrac{\psi(x+\Delta x, y+\Delta y) - \psi(x, y+\Delta y)}{\Delta x} - \dfrac{\psi(x+\Delta x, y) - \psi(x,y)}{\Delta x}}{\Delta y} \ .$$

This simplifies to

$$\partial_y(\,\partial_x \psi(x,y)\,) \approx \frac{\psi(x+\Delta x, y+\Delta y) - \psi(x, y+\Delta y) - \psi(x+\Delta x, y) + \psi(x,y)}{\Delta x \Delta y} \ .$$

On the right are values of ψ evaluated at the four corners of the rectangle starting at (x,y) and going to $(x+\Delta x, y+\Delta y)$. You would get ψ evaluated at the same four points, with the same signs, if you had done the partial derivatives in the other order $\partial_x \psi_y$. That's why $\partial_x \psi_y = \partial_y \psi_x$.

Integrating on lines in 2D

A *line integral* is a one dimensional integral along a *path* in two or more dimensions. There is a general version of this you might see in vector calculus. We only need a simple version of the general theory that integrates only in horizontal or vertical directions. Suppose a 2D path starts at (a,b) and ends at (c,d). You can find the difference between $\psi(c,d)$ and $\psi(a,b)$ by integrating partial derivatives of ψ along the path. We will use paths that have only horizontal and vertical pieces. You should draw your own diagrams to follow the description. The diagrams may assume $c > a$ and $d > b$ so the end point (c,d) is above and to the right of the starting point (a,b). The formulas are true even if the geometry is not like this.

The fundamental theorem of calculus says that

$$\int_{x_0}^{x_1} f'(x)\,dx = f(x_1) - f(x_0)\;. \tag{6}$$

This can be applied to functions of two variables. Suppose y is constant and $f(x) = \psi(x,y)$. Then $f'(x) = \psi_x(x,y)$. The integral theorem (6) with $x_0 = a$ and $x_1 = c$ gives

$$\int_{x=a}^{x=c} \psi_x(x,y)\,dx = \psi(c,y) - \psi(a,y)\;.$$

This is a *line integral* along the (horizontal) line segment from (a,y) to (c,y). You can also integrate on vertical segments (with x coordinate fixed), so

$$\int_{y=b}^{y=d} \psi_y(x,y)\,dy = \psi(x,d) - \psi(x,b)\;.$$

We write the limits of integration in the form $x = a$ or $y = d$ to indicate which variable is equal to a or d, etc.

We consider two paths that go from (a,b) to (c,d) using one horizontal and one vertical segment. One path moves first in the horizontal and then in the vertical directions. It goes from (a,b) to (c,b) (horizontal) and then from (c,b) to (c,d) (vertical). The other path moves first in the y direction from (a,b) to (a,d) and then in the x direction from (a,d) to (c,d). Of course, there are many other paths from (a,b) to (c,d).

You can express the difference between $\psi(c,d)$ and $\psi(a,b)$ using line integrals along either of these paths. Suppose you do the horizontal part first. The integral along this segment (where x goes from a to c while y stays at b) gives

$$\psi(c,b) - \psi(a,b) = \int_{x=a}^{x=c} \psi_x(x,b)\,dx\;.$$

Then the vertical part (y goes from b to d while x stays at c) gives

$$\psi(c,d) - \psi(c,b) = \int_{y=b}^{y=d} \psi_y(c,y)\,dy\;.$$

We can put these together to get

$$\psi(c,d) - \psi(a,b) = \int_{x=a}^{x=c} \psi_x(x,b)\,dx + \int_{y=b}^{y=d} \psi_y(c,y)\,dy\;. \tag{7}$$

The other path, which is vertical first then horizontal, gives ($y = d$ on the horizontal part and $x = a$ on the vertical part)

$$\psi(c,d) - \psi(a,b) = \int_{x=a}^{x=c} \psi_x(x,d)\,dx + \int_{y=b}^{y=d} \psi_y(a,y)\,dy\;. \tag{8}$$

We put the vertical integral after the horizontal one in the formula even though the vertical part comes first in the path to make it easier to compare the two expressions (7) and (8).

Two dimensional integrals

A two dimensional integral over a rectangle has the form

$$I = \int_{x=a}^{x=c} \int_{y=b}^{y=d} f(x,y)\, dxdy \; .$$

There are several ways to think of this integral. You can interpret it as an *iterated* integral, which means first integrating with respect to one of the variables and then integrating the result with respect to the other variable. In this case, you could take the *inner integral* to be

$$u(x) = \int_{y=b}^{y=d} f(x,y)\, dy \; . \tag{9}$$

Then the *outer* integral would be

$$I = \int_{x=a}^{x=c} u(x)\, dx \; . \tag{10}$$

This is written in one line as

$$I = \int_{x=a}^{x=c} \left(\int_{y=b}^{y=d} f(x,y)\, dy \right) dx \; . \tag{11}$$

The integral may be evaluated in the other order, with the inner integration with respect to x and the outer with respect to y. The result is the same

$$I = \int_{y=b}^{y=d} \left(\int_{x=a}^{x=c} f(x,y)\, dx \right) dy \; . \tag{12}$$

Finally, the integral can be written more abstractly without specifying an order of integration. Let R be the rectangle with corners (moving in the counter-clockwise direction starting at the lower left) $(a,b), (c,b), (c,d), (a,d)$.

$$R = \{\, (x,y) \text{ with } a \leq x \leq c \text{ and } b \leq y \leq d \,\} \; .$$

Then

$$I = \iint_R f(x,y)\, dxdy \; . \tag{13}$$

The three expressions (11), (12), and (13) are different formulas for the same number, I. The question is: Why do they give the same answer?

We saw that a one variable integral like

$$\int_a^b f(x)\, dx$$

can be thought of as a sum of small contributions $f(x)dx$ that you get by dividing the interval $[a,b]$ into small pieces of length dx. In the same way, the double integral (13) can be thought of as a sum of small contributions $f(x,y)dxdy$, with each contribution corresponding to a small $dx \times dy$ rectangular piece of the integration domain R.

There is more than one way to add up the contributions from the little patches. One way is to divide R into vertical columns of rectangles ranging from $y = b$ up to $y = d$. The x coordinates for all the boxes in this column are the interval $(x, x + dx)$. Adding up the contributions from one tower means adding up

$$f(x,y)\, dxdy$$

over y values ranging from b to d. The result is (remember that the integral sign "$\int$" and the sum sign "Σ" mean almost the same thing)

$$\int_{y=b}^{y=d} f(x,y)\, dxdy \; .$$

You can factor out the common factor dx and use the definition of the vertical integral (9) to get

$$\int_{y=b}^{y=d} f(x,y)\, dxdy = \left(\int_{y=b}^{y=d} f(x,y)\, dy \right) dx = u(x)dx \; .$$

This result is small, proportional to dx because it's the sum over a thin column whose width is dx. You can now add up the contributions from all these towers to get the sum over all the boxes in R. The result is what we called the "outer integral" (10). To summariae, you can add up the contributions from the little boxes one column at a time and then sum over the columns. You get the *nested integral* approach to the double integral, first (9) then (10).

You could add up the contributions by rows instead of columns. For each interval $(y, y + dy)$ there is a row of boxes with x values running from a to c. Adding up the boxes in this row gives

$$\int_{x=a}^{x=c} f(x,y)\, dxdy = \left(\int_{x=a}^{x=c} f(x,y)\, dx \right) dy \; .$$

The integral in parentheses depends on y, and we call it

$$v(y) = \int_{x=a}^{x=c} f(x,y)\, dx \; .$$

Adding up the contributions $v(y)dy$ from all the rows gives the total sum/integral as

$$I = \int_{y=b}^{y=d} v(y)\, dy = \int_{y=b}^{y=d} \left(\int_{x=a}^{x=c} f(x,y)\, dx \right) dy \; .$$

The conclusion is that the double integral (13) may be evaluated either "by columns" or "by rows". The results are the same. The fact that (13), (11), and (12) are equal is sometimes called *changing the order of integration*. It is a useful trick both for theory, as in these notes, and for calculating multiple integrals, which will come up later in this course. The integrals defining u and v may be thought of as line integrals because they are one dimensional integrals in the two dimensional plane.

Green's theorem

We put together the previous two sections to show that the criterion (1) is sufficient for there to be a potential function ψ that satisfies (2). The reasoning is related to what is called *Green's theorem* in vector calculus. For this, we express the fundamental theorem of calculus in different notation. Suppose $f(x,y)$ is any function and

$$u(x,y) = \int_{x'=x_0}^{x'=x} f(x',y)\, dx' \; .$$

The value of the starting point x_0 is irrelevant here. The fundamental theorem of calculus gives

$$\partial_x u(x,y) = f(x,y) \; .$$

Here, y is just a parameter. The integral and derivative are with respect to x. You can do the same thing with x as the parameter and y as the integration and differentiation variable. If

$$v(x,y) = \int_{y'=y_0}^{y'=y} g(x,y')\,dy' \ ,$$

then

$$\partial_y v(x,y) = g(x,y) \ .$$

If there is a potential function ψ that satisfies (2), then we express the value of $\psi(x,y)$ in terms of a *base point*, (a,b), and a line integral. For that, we form the rectangle with corners (a,b), (x,b), (x,y), and (a,y). The integral expression (7) may be written in different notation, and with ψ_f and $\psi_y = g$, as (first integrate along the bottom, then along the right side of the rectangle)

$$\psi(x,y) = \psi(a,b) + \int_{x'=a}^{x'=x} f(x',b)\,dx' + \int_{y'=b}^{y'=y} g(x,y')\,dy' \ . \tag{14}$$

If we take this as the *definition* of ψ, then

$$\partial_y \psi(x,y) = g(x,y) \ .$$

This is because the first two terms on the right of (14) do not depend on y and the the fundamental theorem of calculus applies to the last term.

We could have used the other path and the representation (7) to define a possibly different function

$$\widetilde{\psi}(x,y) = \psi(a,b) + \int_{x'=a}^{x'=x} f(x',y)\,dx' + \int_{y'=b}^{y'=y} g(a,y')\,dy' \ . \tag{15}$$

The first integral on the right is over the top of the rectangle while the second integral (which represents the first part of the path from (a,b) to (x,y)) is over the left side. Since the first and third terms do not depend on x, we see that

$$\partial_x \widetilde{\psi}(x,y) = f(x,y) \ .$$

We will see (next paragraph) that, if the compatibility conditions (1) are satisfied, then (14) and (15) define the same function. This ψ is what we want because it satisfies both $\psi_y = g$ and $\psi_x = f$.

This argument is simple, but the notation could be unnecessarily involved. To make the notation simpler, we go back to (a,b) and (c,d) and drop the primes. We also set $\psi(a,b) = 0$. If $\psi(a,b)$ has some other value, then the function defined would differ only by this value. Thus, showing that $\widetilde{\psi} = \psi$ is the same as showing that functions defined by the two path integrals have the same value at (c,d), which can be any point. We need to show that

$$\int_{x=a}^{x=c} f(x,b)\,dx + \int_{y=b}^{y=d} g(c,y)\,dy = \int_{x=a}^{x=c} f(x,d)\,dx + \int_{y=b}^{y=d} g(a,y)\,dy \ . \tag{16}$$

The x integral on the left has $y = b$, which puts it on the bottom of the rectangle. The x integral on the right has $y = d$, which corresponds to the top of the rectangle. The y integrals are on the right $(x = c)$ and left $(x = a)$ sides. The left and right expressions are equal when $f_y = g_x$, which is part of *Green's theorem* of multivatiate calculus.

The derivatives f_y and g_x get involved when you compare the integrals over the top and bottom (f_y) and over the left and right (g_x). To start,

$$f(x,d) = f(x,b) = \int_{y=b}^{y=d} f_y(x,y)\,dy \ .$$

Thus the difference between the f integrals on the left and right in (16) may be expressed as a double integral

$$\int_{x=a}^{x=c} f(x,d)\,dx - \int_{x=a}^{x=c} f(x,b)\,dx = \int_{x=a}^{x=c} \left(\, f(x,d) - f(x,a)\,\right)\, dx$$
$$= \int_{x=a}^{x=c} \left(\int_{y=b}^{y=d} f_y(x,y)\,dy \right) dx \;.$$

We can change the order of integration in the double integral to get

$$\int_{y=b}^{y=d} \left(\int_{x=a}^{x=c} f_y(x,y)\,dx \right) dy \;.$$

If $f_y = g_x$, this becomes

$$\int_{y=b}^{y=d} \left(\int_{x=a}^{x=c} g_x(x,y)\,dx \right) dy \;.$$

The inner integral is

$$\int_{x=a}^{x=c} g_x(x,y)\,dx = g(c,y) - g(a,y) \;.$$

Therefore

$$\int_{y=b}^{y=d} \left(\int_{x=a}^{x=c} g_x(x,y)\,dx \right) dy = \int_{y=b}^{y=d} \left(\, g(c,y) - g(a,y)\,\right) dy$$
$$= \int_{y=b}^{y=d} g(c,y)\,dy - \int_{y=b}^{y=d} g(a,y)\,dy \;.$$

This is the difference between the left and right g integrals in (16). You can put these together to get the equality in (16). This shows that $\widetilde{\psi} = \psi$.

To summarize, if the compatibility condition (1) is satisfied, then the two ways to define ψ given the same ψ. One way shows that $\psi_y = g$ and the other way shows that $\psi_x = f$. Thus, the two integrals give a constructive proof that the compatibility condition implies that there is a potential function.

A chain rule and differentials

If x and y are both functions of t, the chain rule for this situation is

$$\frac{d}{dt}\psi(x(t),y(t)) = \psi_x\,\dot{x} + \psi_y\,\dot{y} \;.$$

You can write this more formally as

$$\frac{d}{dt}\psi(x(t),y(t)) = \psi_x(x(t),y(t))\,\frac{dx}{dt} + \psi_y(x(t),y(t))\,\frac{dy}{dt} \;.$$

If you multiply both sides by dt you get the differential form

$$d\psi = \psi_x\,dx + \psi_y\,dy \;. \tag{17}$$

This says that if you change x and y by small amounts dx and dy, then the change in ψ is the sum of the change because x changed and the change because y changed, You could think of this as taking a path from (x,y) to $(x+dx, y+dy)$ that first voves in the x direction from (x,y) to $(x+dx,y)$ and then moves in the y direction from $(x + dx, y)$ to $(x + dx, y + dy)$. The x move changes ψ by $\psi_x\, dx$. The y move changes ψ by $\psi_y\, dy$. With both moves, you get the total differential (17). The parts $\psi_x\, dx$ and $\psi_y\, dy$ are "partial differentials", which is why ψ_x and ψ_y are called *partial* derivatives. You might worry that in the path from (x, y) to $(x + dx, y + dy)$, you evaluate ψ_x and ψ_y at slightly different places. The answer is that, for example, the difference between $\psi_y(x, y)\, dy$ and $\psi_y(x + dx, y)\, dx)$ is "tiny" in the sense that the difference is much smaller than the differential $\psi_y\, dy$ itself.

In differential equations we often use this reasoning in reverse. We are given a differential expression

$$f(x, y)\, dx + g(x, y)\, dy \ . \tag{18}$$

The task is to "integrate" this differential expression by finding a function ψ so that

$$d\psi(x, y) = f(x, y)\, dx + g(x, y)\, dy \ .$$

That means finding a function ψ so that

$$\psi_x(x, y) = f(x, y) \ \ , \quad \psi_y(x, y) = g(x, y) \ . \tag{19}$$

The conditions (19) define ψ only "up to a constant of integration". That means that if ψ is a solution, then $\psi + C$ is also a solution. The constant of integration, C would have to be determined from other information about the problem. For example, if you know that $x = a$ when $y = b$, then you know that $\psi(x, y) = \psi(a, b)$ for all x and y.