Multivariate Calculus
MANUSCRIPT

William F. Barnes [1]

February 11, 2024

Contents

Chapter 1

Multivariate Calculus

1 Surfaces and Solids

Before getting into heavy jargon, we'll do a brief tour of the extension of the curve $y = f(x)$ into more dimensions.

1.1 Surfaces

The natural extension of a single-input function $f(x)$ is one that takes two arguments. In analogy to $y = f(x)$, we may also write

$$z = f(x, y) \ ,$$

where $f(x, y)$ requires two independent inputs x and y. The domain of f is part (or all) of the Cartesian plane on which x and y occur. The range variable z may be regarded as the 'height above' the plane at (x, y).

If $f(x, y)$ is continuous in both variables, the set of all z-points constitutes a surface. In the same sense that a curve is a continuous arrangement of points in the Cartesian plane, surfaces may are like 'sheets' in a Cartesian volume.

Level Curves

Fixing z constant on a surface restricts the freedom in the xy-plane at height z to a *level curve*. A level curve is the same as a contour line seen on a topographical (not topological) map, or on various weather maps. A small ring surrounds a local minimum or a local maximum. Level curves intersect at a saddle point.

Critical Points

As two-dimensional creatures, surfaces have three kinds of critical points. A maxima in the surface is when both the x- and the y-variables reach a high point simultaneously. The same comment applies to a minima on the surface.

Another kind of critical point is called a *saddle point*, which has one of the variables x, y is a maxima, and the other a minima.

1.2 Solids

Adding another variable into the mix, we can have functions of three variables

$$F = f(x, y, z) \ ,$$

where F is most generally called a *scalar field*. The temperature or air density in a room qualifies as a scalar field. Holding any variable F constant produces a *level surface*, the generalization of the level curve.

We'll stave off the discussion of critical points within scalar fields, or if you see it coming, vector fields, until after a few developments are made.

Topological Remarks

If the field F describes a finite solid, then the notion of minima, maxima, and saddle points on the surface of the solid are given the respective labels 'pits', 'peaks', and 'passes'. It's possible to show using arguments from topology that the following is always true:

$$\text{peaks} - \text{passes} + \text{pits} = 2$$

To understand this, consider an ice cream cone with one scoop of ice cream, supposed spherical. There are two peaks in this situation - the top of the ice cream, and the bottom of the cone, so $2 = 2$ checks out. Press your thumb into the ice cream to introduce a pit and a pass simultaneously. Then you get $2 - 1 + 1 = 2$.

The relationship between critical points has an analogous formula with respect volumes with flat faces and sharp edges. If V is the number of vertices, E is the number of edges, and F is the number of faces, it's possible to prove, much like the above:

$$V - E + F = 2$$

2 Multiple Integration

2.1 Single Integral

The workhorse of integral calculus is the fundamental theorem

$$f(x_1) - f(x_0) = \int_{x_0}^{x_1} f'(t)\, dt\,,$$

where $f'(x)$ is the derivative df/dx.

The integral calculates the area under the curve $y(x) = f'(x)$ between the endpoints x_0, x_1 and hands us the answer in the form $A = f(x_1) - f(x_0)$. A concise way to express the area under such a curve is

$$A = \int_{x_0}^{x_1} y(x)\, dx\,.$$

2.2 Double Integral

The standard area integral calculates the sum of an infinite number of heights $y(x)$ above the x-axis. It stands to reason that $y(x)$ itself could be the result of an integral:

$$y(x) = \int_0^{y(x)} dy$$

The lower limit need not be zero if we take $y(x)$ as the vertical length trapped between two curves $y_0(x)$, $y_1(x)$, or:

$$y(x) = \int_{y_0(x)}^{y_1(x)} dy$$

Inserting this form for $y(x)$ into the one dimensional area integral yields a *double integral*:

$$A = \int_{x_0}^{x_1} \int_{y_0(x)}^{y_1(x)} dy\, dx$$

Notice that the *inner* integration limits are functions of the *outer* integration variable.

Supposing $y_0(x)$, $y_1(x)$ are easily inverted, the same area can be expressed with the integration variables reversed:

$$A = \int_{y_0}^{y_1} \int_{x_0(y)}^{x_1(y)} dx\, dy$$

Order of Integration

Note that in any multiple integral, it's always implied that the order of integration goes from the inner-most to the outer-most, not unlike like simplifying expressions with parentheses.

Integration Region

The curves $y_0(x)$, $y_1(x)$ are called *bounding functions.* The same comment apples to their inverted counterparts $x_0(y)$, $x_1(y)$.

The total information contained in the integration limits, including bounding functions, is called the *integration region*, denoted $\mathcal{D}$. With this we can express the double integral in a less definite form:

$$A = \int\int_{\mathcal{D}} dx\, dy\,.$$

By obscuring the product $dx\, dy$ into an area element dA, the above can be written free of the Cartesian coordinate system:

$$A = \int\int_{\mathcal{D}} dA$$

Volume Integral

The double integral apparatus can be used to calculate the volume trapped between the xy-plane and a given surface $z = f(x, y)$. For this, we simply write

$$V = \int\int_{\mathcal{D}} f(x, y)\, dx\, dy\,.$$

Integrating Functions

Supposing we need to calculate something more abstract than an area, the double integral apparatus takes any reasonable function in its integrand

$$B = \int\int_{\mathcal{D}} f(x, y)\, dx\, dy\ ,$$

where $f(x, y)$ is a generalized surface.

2.3 Polar Integral

To work a specific case, the differential area element in polar coordinates reads

$$dA = r\, dr\, d\theta\ ,$$

which means

$$A = \int\int_{\mathcal{D}} r\, dr\, d\theta$$

Note that the r-variable is usually a function of θ, which means the r-integral should be solved first. In detail,

$$A = \int \left(\int r\, dr \right) d\theta\ ,$$

where the integration limits are left in indefinite form.

Now, even if there were an additional function $f(\theta)$ in the integrand, thus changing the integral to

$$B = \int \left(\int r\, dr \right) f(\theta)\, d\theta\ ,$$

notice how $f(\theta)$ has no r-dependence, and is situated outside the r-integral.

We're thus free to evaluate the inner r-integral, and the above becomes

$$B = \frac{1}{2} \int r^2 f(\theta)\, d\theta\ .$$

This is a nifty result, as if $f(\theta)$ weren't there, then the integral is simply the area under $r(\theta)$ in polar coordinates:

$$A = \frac{1}{2} \int r^2\, d\theta$$

Area of a Circular Arc

Calculating the area of the circular arc in polar coordinates is as easy as

$$A = \int_{\theta_0}^{\theta_1} \int_0^R r\, dr\, d\theta = \frac{1}{2} (\theta_1 - \theta_0) R^2\ ,$$

which is the entire story.

For a bit of self-torture, let us do the same calculation in Cartesian coordinates using double integration. To set up, consider a pair of points (x_0, y_0), (x_1, y_1) with $x_0 > x_1$ that define the endpoints of the arc:

$$\begin{aligned} R\cos(\theta_0) &= x_0 \\ R\sin(\theta_0) &= y_0 \\ R\cos(\theta_1) &= x_1 \\ R\sin(\theta_1) &= y_1 \end{aligned}$$

Note also that the points (x_0, y_0), (x_1, y_1) define straight lines through the origin (no y-intercept), with respective slopes:

$$\begin{aligned} m_0 &= \frac{y_0}{x_0} \\ m_1 &= \frac{y_1}{x_1} \end{aligned}$$

The area of the circular arc will be calculated in two parts, one that resolves to a triangle with all straight edges, and another to handled the curved region. Working this out carefully, find:

$$A = \int_0^{x_1} \int_{m_0 x}^{m_1 x} dy\, dx + \int_{x_1}^{x_0} \int_{m_0 x}^{\sqrt{R^2 - x^2}} dy\, dx$$

The first integral can be evaluated fully with ease, but the second needs to chipped away at. Doing a round of simplifying, reach the intermediate step

$$\begin{aligned} A = &\frac{x_1^2}{2} (m_1 - \cancel{m_0}) \\ &+ \int_{x_0}^{x_1} \sqrt{R^2 - x^2}\, dx \\ &- \frac{m_0}{2} \left(x_0^2 - \cancel{x_1^2} \right)\ . \end{aligned}$$

The remaining integral can be solved with a sine substitution, which has the general solution:

$$\begin{aligned} \int \sqrt{R^2 - x^2}\, dx = &\frac{x\sqrt{R^2 - x^2}}{2} \\ &+ \frac{R^2}{2} \arctan\left(\frac{x}{\sqrt{R^2 - x^2}} \right) + C \end{aligned}$$

For the problem on hand, this means

$$\begin{aligned} \int \sqrt{R^2 - x^2}\, dx = &\frac{m_0 x_0^2 - m_1 x_1^2}{2} \\ &+ \frac{R^2}{2} \left(\arctan\left(\frac{x_0}{y_0} \right) - \arctan\left(\frac{x_1}{y_1} \right) \right)\ . \end{aligned}$$

The x- and m-terms all cancel, and the area reads

$$A = \frac{R^2}{2} \left(\arctan\left(\frac{x_0}{y_0} \right) - \arctan\left(\frac{x_1}{y_1} \right) \right)\ .$$

From trigonometry, note in general that

$$\arctan\left(\cot\left(u\right)\right) = \frac{\pi}{2} - u \ ,$$

and the area becomes

$$A = \frac{R^2}{2}\left(\frac{\pi}{2} - \theta_0 - \frac{\pi}{2} + \theta_1\right) \ ,$$

and finally,

$$A = \frac{1}{2}\left(\theta_1 - \theta_0\right) R^2 \ ,$$

in agreement with the previous answer.

Gaussian Integral

Consider the definite integral

$$I_1 = \int_{-\infty}^{\infty} e^{-ax^2}\, dx \ ,$$

which has no elementary solution. Instead of turning to a numerical approximation, which would ordinarily be the case for such an integral, consider the same exact integral with a swap of variables:

$$I_1 = \int_{-\infty}^{\infty} e^{-ay^2}\, dy$$

If this analysis doesn't seem insane yet, multiply each copy of the integral together to get to what seems like a dead end,

$$I_1^2 = \left(\int_{-\infty}^{\infty} e^{-ax^2}\, dx\right)\left(\int_{-\infty}^{\infty} e^{-ay^2}\, dy\right) \ ,$$

and melt the notation down:

$$I_1^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-a\left(x^2+y^2\right)}\, dx\, dy$$

Now, one must be very careful when doing this, but it just happens that the x, y variables can be regarded as locations on the Cartesian plane, which lends to polar coordinates. Switching to polar, the above integral is

$$I_1^2 = \int_0^{2\pi}\int_0^{\infty} e^{-a\rho^2}\, \rho\, d\rho\, d\phi \ .$$

Observe how the region of integration (the infinite plane) makes the limits on each integral easy to write.

The ϕ-integral is trivial and resolves to 2π. The remaining ρ-integral can is solved straightforwardly by u-substitution. Chugging through each, we find $I_1^2 = \pi/a$, or

$$I_1 = \int_{-\infty}^{\infty} e^{-ax^2}\, dx = \sqrt{\frac{\pi}{a}} \ .$$

This cheat works on several I_1-like problems called *Gaussian integrals*. Let us work through

$$I_2 = \int_{\infty}^{\infty} e^{-ax^2+bx}\, dx \ .$$

Completing the square within the exponential leads to:

$$I_2 = e^{b^2/4a}\frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} e^{-u^2}\, du = \sqrt{\frac{\pi}{a}}\, e^{b^2/4a}$$

Yet another Gaussian integral

$$I_3 = \int_{-\infty}^{\infty} x^2\, e^{-ax^2}\, dx$$

can be solved by taking the derivative of $-I_1$ with respect to a (not x). In detail:

$$\frac{d}{da}\left(-I_1\right) = I_3 = -\frac{d}{da}\left(\sqrt{\frac{\pi}{a}}\right) = \sqrt{\frac{\pi}{4a^3}}$$

2.4 Triple Integral

The apparatus for multiple integration readily generalizes for three and more dimensions. For three dimensions, the idea of a bounding function becomes a *bounding surface*, and we have a *triple integral*:

$$V = \int_{x_0}^{x_1}\int_{y_0(x)}^{y_1(x)}\int_{z_0(x,y)}^{z_1(x,y)} dz\, dy\, dx$$

In the special case $z_0\left(x, y\right) = 0$, the above reduces to a standard volume integral.

The order of integration has greater significance as the dimension of the integral increases. Supposing the required bounding functions and surfaces are easily attained, the triple integral can be written several more ways, for instance:

$$V = \int_{y_0}^{y_1}\int_{z_0(y)}^{z_1(y)}\int_{x_0(y,z)}^{x_1(y,z)} dx\, dz\, dy$$

$$V = \int_{z_0}^{z_1}\int_{x_0(z)}^{x_1(z)}\int_{y_0(z,x)}^{y_1(z,x)} dy\, dx\, dz$$

Or, in terms of an integration region:

$$V = \int\int\int_{\mathcal{D}} dx\, dy\, dz$$

Of course, the triple integral can involve functions in the integrand. For a three-variable function $f\left(x, y, z\right)$, sometimes called a *scalar field*, we can calculate things like

$$B = \int\int\int_{\mathcal{D}} f\left(x, y, z\right) dx\, dy\, dz \ .$$

Non-Cartesian Volume Elements

To go from three-dimensional Cartesian coordinates to a different system, the volume element and specification of the integration need to be changed. For cylindrical coordinates, we may have

$$V = \int_{z_0}^{z_1} \int_{\phi_0(z)}^{\phi_1(z)} \int_{\rho_0(\phi,z)}^{\rho_1(\phi,z)} \rho \, d\rho \, d\phi \, dz \; ,$$

and for spherical coordinates:

$$V = \int_{\phi_0}^{\phi_1} \int_{\theta_0(\phi)}^{\theta_1(\phi)} \int_{r_0(\theta,\phi)}^{r_1(\theta,\phi)} r^2 \sin(\theta) \; dr \, d\theta \, d\phi$$

Like the two-dimensional case, the volume element and integration region can be generalized (a fancy word for 'obscured'):

$$V = \int \int \int_{\mathcal{D}} dV$$

Hurricane Problem

In a simplified model of a hurricane, the velocity of the wind is taken to be purely in the circumferential direction and of magnitude

$$v(\rho, z) = \Omega \, \rho \, e^{-z/h - \rho/a} \; ,$$

where ρ and z are cylindrical coordinates measured from the eye of the hurricane at sea level, and Ω, h, a are positive constants. The density of the atmosphere is approximated by

$$d(z) = d_0 \, e^{-z/h} \; .$$

Find the total kinetic energy of the motion.

As an integral, the kinetic energy is given by

$$T = \int \frac{1}{2} v^2 \, dm \; .$$

This can be converted to a volume integral via the relation

$$\frac{dm}{dV} = d(z) \; ,$$

where dV is the volume element in cylindrical coordinates. The kinetic energy integral becomes

$$T = \int_0^\infty \int_0^{2\pi} \int_0^\infty \frac{1}{2} d(z) \left(v(\rho, z)\right)^2 dV \; ,$$

or

$$T = \frac{d_0 \Omega^2}{2} \int_0^\infty \int_0^{2\pi} \int_0^\infty e^{-z/h} \rho^2 \, e^{-2z/h - 2\rho/a} \rho \, d\rho \, d\phi \, dz \; .$$

The integral can be broken apart into three separate integrals

$$T = \frac{d_0 \Omega^2}{2} \left(\int_0^\infty e^{-3z/h} \, dz \right) \left(\int_0^{2\pi} d\phi \right) \left(\int_0^\infty \rho^3 e^{-2\rho/a} d\rho \right) \; ,$$

each straightforwardly evaluated:

$$T = \frac{d_0 \Omega^2}{2} \left(\frac{h}{3} \right) (2\pi) \left(\frac{3a^4}{8} \right) = \frac{\pi}{8} \Omega^2 d_0 h a^4$$

2.5 Shell Theorem

Newton's law of gravitation tells us that every particle in the universe is trying to pull every other particle toward itself with a force proportional to the masses involved and inversely proportional to the square of the separation, and this is duly used to calculate the force onto planets, moons, satellites, and so on.

Using triple integration and spherical coordinates, something Newton didn't have, we finally address an assumption made early in gravitational analysis, namely *why* we're allowed to represent voluminous objects as single points located at the center of mass. This is called the shell theorem, and entails two important proofs.

Outside a Sphere

Consider a solid sphere of radius R, total mass M, and uniform density λ. Also let there be a test particle of mass m somewhere in space. Without loss of generality, place the test particle on the z-axis at the point $\vec{D} = D \, \hat{z}$. The length D is the distance from the test particle to the center of the sphere.

In order to 'properly' calculate the gravitational attraction between the test mass and the sphere, a volume integral over there entire sphere must be calculated. Choose any element of volume dV inside the sphere at location $\vec{r}$, which is located distance r from the center, at an angle θ from the z-axis.

Let vector $\vec{q}$ denote the line connecting $\vec{D}$ to $\vec{r}$ such that

$$\vec{r} + \vec{q} = D \, \hat{z} \; ,$$

and also let α be the angle between $\hat{z}$ and $\hat{q}$. From the law of cosines, we can say:

$$q^2 = r^2 + D^2 - 2rD\cos(\theta)$$
$$r^2 = q^2 + D^2 - 2qD\cos(\alpha)$$

The total force on the test particle is the vector $\vec{F}$. However, due to the ϕ-symmetry of this picture, only the z-component of the force will have a net effect on the particle. All xy-components cancel equally and oppositely:

$$F = \int_{\mathcal{D}} d\vec{F} \cdot \hat{z} = \int\int\int_{\text{volume}} dF \cos(\alpha)$$

The differential force is

$$dF = \frac{-Gm}{q^2} dm \,,$$

where dm is the mass of the differential volume element influencing the test particle. The mass term can be replaced using the density

$$\frac{dm}{dV} = \frac{M}{4\pi R^3/3} = \lambda \,,$$

where it is appropriate to replace dV with the volume element in spherical coordinates.

The force integral now is

$$F = -Gm\lambda \int_0^{2\pi} \int_0^{\pi} \int_0^{R} \frac{\cos(\alpha)}{q^2} r^2 \sin(\theta)\, dr\, d\theta\, d\phi \,,$$

which, after substituting and simplifying a bit, becomes:

$$F = -Gm\lambda \frac{2\pi}{2D} \int_0^{\pi} \int_0^{R} \left(\frac{1}{q} + \frac{D^2 - r^2}{q^3} \right) r^2 \sin(\theta)\, dr\, d\theta$$

Perform implicit differentiation on the q^2 equation to find, remembering r and θ are independent,

$$q\, dq = rD \sin(\theta)\, d\theta \,,$$

and rewrite the integral with the intent of integrating over r last. Make you you know why the limits are now changed:

$$F = -Gm\lambda \frac{\pi}{D^2} \int_0^{R} \int_{(D-r)}^{(D+r)} \left(1 + \frac{D^2 - r^2}{q^2} \right) r\, dq\, dr$$

The whole q-integral treats r as a constant and resolves to $4r$, so

$$F = -Gm\lambda \frac{\pi}{D^2} \int_0^{R} 4r^2\, dr \,,$$

and the r-integral is elementary. Simplifying everything gives

$$F = -Gm \left(\frac{3M}{4\pi R^3} \right) \frac{\pi}{D^2} \frac{4}{3} R^3 = \frac{-GMm}{D^2} \,.$$

Conveniently, the force acts as if *all* of its mass were concentrated at the center. This result is also true in general, where the notion of 'center' means center of mass, not necessarily the center of the volume.

Inside a Shell

Another interesting question that arises in the course of studying gravity is, what does it feel like inside a hollow uniform shell? To pursue this question, suppose we have a thin spherical shell of radius R and thickness $2a$ that is much less than R, and the test particle is inside anywhere within the shell.

This setup borrows all of the geometry from the previous setup, except this time we have $D < R$, which is the important part. Setting up the same integral and doing the same simplifications, we can jump to

$$F = -Gm\lambda \frac{\pi}{D^2} \int_{R-a}^{R+a} \int_{(r-D)}^{(D+r)} \left(1 + \frac{D^2 - r^2}{q^2} \right) r\, dq\, dr \,.$$

Most notably, the lower integration in the q-integral is swapped to accommodate $D < R$. This causes the q-integral to resolve to zero, and we find

$$F = 0$$

inside the shell.

3 Partial Derivative

Returning to the definition of a function, recall that a function f depends on an input variable x in the function's domain. Given any input value, the output of the function is written $y = f(x)$, and there is only one y for a given x. The set of all y-values constitute the function's range. A 'curve' given by function $y = f(x)$ may exhibit a myriad of features: asymptotic behavior, periodicity, singularities, critical points, inflection points, etc.

Derivative

One star result from the analysis of curves gives the slope of the function at a point x_0, namely

$$\frac{d}{dx} f(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \,.$$

Taylor's Theorem

The so-called derivative turns out to be just the first-order version of something more general we known as Taylor's theorem. Near the point x_0, we have:

$$f\left(x\right)\approx f\left(x_0\right)+\sum_{q=1}^{\infty}\frac{1}{q!}f^{(q)}\left(x_0\right)\left(x-x_0\right)^q$$

3.1 Slope on a Surface

The technical definition of the derivative generalizes to surfaces. For this, we require the surface $z = f\left(x,y\right)$ to be differentiable both in the x-direction and the y-direction, meaning that the slope of the surface at a point $\left(x_0,y_0\right)$ has two answers: a slope along x, and a slope along y.

To express the slope on a surface at the point $\left(x_0,y\right)$, we write the usual slope formula treating x as the 'active' variable with y as a constant:

$$\frac{\partial}{\partial x}f\left(x_0,y\right)=\lim_{x\rightarrow x_0}\frac{f\left(x,y\right)-f\left(x_0,y\right)}{x-x_0} \tag{1.1}$$

Meanwhile, the slope at the point $\left(x,y_0\right)$ allows y to vary while x is constant:

$$\frac{\partial}{\partial y}f\left(x,y_0\right)=\lim_{y\rightarrow y_0}\frac{f\left(x,y\right)-f\left(x,y_0\right)}{y-y_0} \tag{1.2}$$

The familiar d/dx-notation is replaced by $\partial/\partial x$. The symbol ∂ denotes the *partial derivative*.

3.2 Mixed Partial Derivatives

One issue that needs to be settled right away is the idea of *mixed* partial derivatives. For the surface $z = f\left(x,y\right)$, let us find out whether

$$\frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right)=\frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right) \tag{1.3}$$

is true. Using brute force, start with

$$\frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right)=\frac{\partial}{\partial y}\left(\lim_{x\rightarrow x_0}\frac{f\left(x,y\right)-f\left(x_0,y\right)}{x-x_0}\right) ,$$

which becomes

$$\frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right)=\lim_{y\rightarrow y_0}\lim_{x\rightarrow x_0} \frac{f\left(x,y\right)-f\left(x_0,y\right)-f\left(x,y_0\right)+f\left(x_0,y_0\right)}{\left(y-y_0\right)\left(x-x_0\right)} .$$

Now, it takes little to imagine doing a similar calculation with the y-partial derivative first to have

$$\frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right)=\frac{\partial}{\partial x}\left(\lim_{y\rightarrow y_0}\frac{f\left(x,y\right)-f\left(x,y_0\right)}{y-y_0}\right) ,$$

which then simplifies to something nearly identical to the above, save one difference, which that the *order* of the limits is swapped. The task boils down to showing in this context that

$$\lim_{x\rightarrow x_0}\lim_{y\rightarrow y_0}\leftrightarrow\lim_{y\rightarrow y_0}\lim_{x\rightarrow x_0}$$

can be assumed.

To prove this, define two new functions

$$\begin{aligned} X\left(x,y\right)&=f\left(x,y\right)-f\left(x_0,y\right) \\ Y\left(x,y\right)&=f\left(x,y\right)-f\left(x,y_0\right) ,\end{aligned}$$

and notice the following equality:

$$X\left(x,y\right)-X\left(x,y_0\right)=Y\left(x,y\right)-Y\left(x_0,y\right)$$

The left- and right-hand sides of the above each represent the endpoints of a secant line on the surface. By the mean value theorem, each can be replaced by partial derivatives as

$$\left(y-y_0\right)\frac{\partial}{\partial y}X\left(x,b\right)=\left(x-x_0\right)\frac{\partial}{\partial x}Y\left(a,y\right) ,$$

where

$$\begin{aligned} x_0 &< a < x \\ y_0 &< b < y .\end{aligned}$$

Keep simplifying to write

$$\begin{aligned}\left(y-y_0\right)&\left(\frac{\partial}{\partial y}f\left(x,b\right)-\frac{\partial}{\partial y}f\left(x_0,b\right)\right)= \\ &\left(x-x_0\right)\left(\frac{\partial}{\partial x}f\left(a,y\right)-\frac{\partial}{\partial x}f\left(a,y_0\right)\right) ,\end{aligned}$$

and use the mean value theorem a second time on each side to write

$$\begin{aligned}\left(y-y_0\right)\left(x-x_0\right)&\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}f\left(\alpha,b\right)\right)= \\ &\left(x-x_0\right)\left(y-y_0\right)\frac{\partial}{\partial y}\left(\frac{\partial}{\partial x}f\left(a,\beta\right)\right) ,\end{aligned}$$

where

$$\begin{aligned} x_0 &< \alpha < x \\ y_0 &< \beta < y .\end{aligned}$$

Closing the limits tighter, we see that the pair a, α tend to x_0, and also the pair b, β tend to y_0. In the differential limit, the left and right sides are equal and the proof is done.

3.3 Partial Derivative Operator

Like the ordinary derivative operator d/dx, the partial derivative operator is written $\partial/\partial x$. For shorthand, the same operator is often written with one 'partial' symbol and a subscript:

$$\frac{\partial}{\partial x} = \partial_x$$

In this notation, the mixed partial derivative Equation (1.3) is simply written

$$\partial_{yx} f(x, y) = \partial_{xy} f(x, y) \ ,$$

or with just the operators,

$$\partial_{yx} = \partial_{xy} \ .$$

Yet another nomenclature for partial derivatives involves placing a subscript with the function itself:

$$\frac{\partial}{\partial x} f(x, y) = f_x$$

Second Derivative

With the notion of partial derivatives, the idea of the second derivative of a function can go three ways. Each of the following is a second derivative operator

$$\partial_{xx} \qquad \partial_{xy} \qquad \partial_{yy}$$

and each produces, in the general case, a different result.

The partial derivative operator obeys the same algebraic rules as the ordinary derivative operator. Without abusing the notation, we can establish things like:

$$\begin{aligned} (\partial_x + \partial_y)(\partial_x - \partial_y) &= \partial_{xx} - \partial_{xy} + \partial_{yx} - \partial_{yy} \\ &= \partial_{xx} - \partial_{yy} \end{aligned}$$

Third Derivative

The equivalency of the mixed partial derivative extends to any depth. From Equation (1.3), we reason that

$$\partial_{yxx} \qquad \partial_{xyx} \qquad \partial_{xxy}$$

yield the same result. For this reason, it turns out there four *unique* third derivative operations:

$$\partial_{xxx} \qquad \partial_{xxy} \qquad \partial_{yyx} \qquad \partial_{yyy}$$

The notation can be condensed once more by using exponent notation on repeated derivatives:

$$\begin{aligned} \partial_{xxx} &= \partial_{x^3} \\ \partial_{xxy} &= \partial_{x^2 y} \\ \partial_{yyx} &= \partial_{xy^2} \\ \partial_{yyy} &= \partial_{y^3} \end{aligned}$$

3.4 Total Derivative

The notion of 'regular' derivative still survives the jump to more dimensions, and is given the name *total* derivative.

For a function $f(x, y, z)$ the total derivative with respect to a variable t sums across each partial derivative:

$$\frac{d}{dt} f(x, y, z) = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt}$$

The Differential

Stripping away the dt-variable by the chain rule yields the so-called 'differential of' $f(x, y, z)$:

$$df(x, y, z) = \frac{\partial f}{\partial x}\, dx + \frac{\partial f}{\partial y}\, dy + \frac{\partial f}{\partial z}\, dz$$

3.5 Variable Integration Limits

It's possible for an integral to have variable limits, which can make a mess of things like integration by parts. Consider a function $y(x)$ and a two-variable function $f(x, y(x))$. By the fundamental theorem, this setup implies integrals of the form

$$F(x) = \int_{a(x)}^{b(x)} f(x, t)\, dt \ .$$

To attack this, define a helper function

$$G(x, y) = \int_{t_0}^{y} f(x, t)\, dt \ ,$$

so then $F(x)$ reads

$$\begin{aligned} F(x) &= \int_{t_0}^{b(x)} f(x, t)\, dt - \int_{t_0}^{a(x)} f(x, t)\, dt \\ &= G(x, b(x)) - G(x, a(x)) \ . \end{aligned}$$

Take the total derivative of $F(x)$:

$$\begin{aligned} \frac{d}{dx} F(x) = &\frac{\partial}{\partial x} G(x, b(x)) - \frac{\partial}{\partial x} G(x, a(x)) \\ &+ \frac{\partial}{\partial y} G(x, b(x)) \frac{db}{dx} \\ &- \frac{\partial}{\partial y} G(x, a(x)) \frac{da}{dx} \end{aligned}$$

The first two terms combine to make $\partial F/\partial x$. To handle the $\partial G/\partial y$ factors, define Δy such that

$$G(x, y + \Delta y) = \int_{t_0}^{y + \Delta y} f(x, t)\, dt$$

Unpack the right side and divide through by Δy to find

$$\frac{G(x, y+\Delta y)-G(x, y)}{\Delta y}=\frac{1}{\Delta y} \int_{y}^{y+\Delta y} f(x, t)\, dt .$$

In the limit $\Delta y \rightarrow 0$, the left side is $\partial G / \partial x$. The right simplifies to $f(x, y)$. In other words:

$$\frac{\partial}{\partial y} G(x, y)=f(x, y)$$

Putting the whole answer together, we have found

$$\begin{aligned} \frac{d}{d x} \int_{a(x)}^{b(x)} f(x, t)\, d t= & \\ & f(x, b(x)) \frac{d b}{d x}-f(x, a(x)) \frac{d a}{d x} \\ & +\frac{\partial}{\partial x} \int_{a(x)}^{b(x)} f(x, t)\, d t \end{aligned} \tag{1.4}$$

as the chief result, called the *Leibniz integral rule*.

Finally, since the operator ∂_x doesn't touch the y-variable, the last integral obeys:

$$\frac{\partial}{\partial x} \int_{a(x)}^{b(x)} f(x, t)\, d t=\int_{a(x)}^{b(x)} \frac{\partial}{\partial x} f(x, t)\, d t$$

Problem 1

Prove the following:

$$\frac{d}{d x} \int_{0}^{x} e^{x t^{2}}\, d t=e^{x^{3}}+\int_{0}^{x} \frac{\partial}{\partial x} e^{x t^{2}}\, d t$$

3.6 Two-Variable Taylor's Theorem

Taylor's theorem generalizes readily to surfaces. To get started, consider a fixed point (x_0, y_0) in the domain of a surface $z=f(x, y)$. Deviations from the fixed point are tracked by two quantities

$$\begin{aligned} \Delta x & =x-x_{0} \\ \Delta y & =y-y_{0} . \end{aligned}$$

Like the one-dimensional Taylor's theorem, we're allowed to frame the final answer as an infinite sum. To do this, first notice that the one-dimensional case contains every whole number power of the quantity $x-x_0=\Delta x$. Then, for two dimensions, we ought to need every whole number power of $\Delta x \Delta y$.

Of course, the factor $f^{(q)}(x_0)/q!$ that appears in the one-dimensional case can't work for two dimensions. Without knowing what to replace this with, let a set of unknown coefficients $\{C_{jk}\}$ stand in for now.

With all this, the two-dimensional Taylor's theorem looks like:

$$\begin{aligned} f(x, y) \approx & f(x_0, y_0) \\ & +C_{10} \Delta x+C_{01} \Delta y \\ & +C_{20} \Delta x^{2}+C_{11} \Delta x \Delta y+C_{02} \Delta y^{2} \\ & +C_{30} \Delta x^{3}+C_{21} \Delta x^{2} \Delta y+C_{12} \Delta x \Delta y^{2}+C_{03} \Delta y^{3} \\ & +\cdots \end{aligned}$$

Now we have the problem of determining each unknown coefficient C_{jk}. Begin by applying the ∂_x operator across the whole equation, and then evaluate the equation at (x_0, y_0). With almost no effort, we can see that any terms containing Δx^2 or higher power will zero out, and the whole result is

$$\partial_{x} f(x_0, y_0)=C_{10} .$$

Applying the ∂_y instead and doing the exercise again leads to a similar result

$$\partial_{y} f(x_0, y_0)=C_{01} .$$

For the next 'row' of coefficients, apply the ∂_{xx}, ∂_{xy}, ∂_{yy} operators respectively, and evaluate at (x_0, y_0). This saps all but the order-two terms in the equation, from which we find:

$$\begin{aligned} \partial_{xx} f(x_0, y_0) & =2 \cdot C_{20} \\ \partial_{xy} f(x_0, y_0) & =C_{11} \\ \partial_{yy} f(x_0, y_0) & =2 \cdot C_{02} \end{aligned}$$

Solving for the order-three coefficients means using the four operators ∂_{x^3}, $\partial_{x^2 y}$, ∂_{xy^2}, ∂_{y^3} to $f(x, y)$ and evaluate at (x_0, y_0). This gives four new equations:

$$\begin{aligned} \partial_{x^3} f(x_0, y_0) & =3 \cdot 2 \cdot C_{30} \\ \partial_{x^2 y} f(x_0, y_0) & =2 \cdot C_{21} \\ \partial_{xyy} f(x_0, y_0) & =2 \cdot C_{12} \\ \partial_{yyy} f(x_0, y_0) & =3 \cdot 2 \cdot C_{03} \end{aligned}$$

To summarize and condense notation once more, use using the general shorthand

$$z_{x^j y^k}=\partial_{x^j y^k} f(x_0, y_0) ,$$

and we have found

$$\begin{aligned} C_{10} & =z_{x} \\ C_{01} & =z_{y} \end{aligned}$$

$$\begin{aligned} C_{20} & =z_{x^2}/2 \\ C_{11} & =z_{xy} \\ C_{02} & =z_{y^2}/2 \end{aligned}$$

$$C_{30} = z_{x^3}/3!$$
$$C_{21} = z_{x^2y}/2$$
$$C_{12} = z_{xy^2}/2$$
$$C_{03} = z_{y^3}/3!$$

The two-dimensional Taylor's theorem now looks like:

$$\begin{aligned}
&f(x,y) \approx f(x_0,y_0) \\
&+ z_x\Delta x + z_y\Delta y \\
&+ \frac{z_{x^2}\Delta x^2 + 2z_{xy}\Delta x\Delta y + z_{y^2}\Delta y^2}{2} \\
&+ \frac{z_{x^3}\Delta x^3 + 3z_{x^2y}\Delta x^2\Delta y + 3z_{xy^2}\Delta x\Delta y^2 + z_{y^3}\Delta y^3}{3!} \\
&+ \cdots
\end{aligned}$$

Look for a moment at the pattern in the numerical coefficients in the numerator of each term written so far. Jotting these down:

$$\begin{array}{ccccccc}
 & & & 1 & & & \\
 & & 1 & & 1 & & \\
 & 1 & & 2 & & 1 & \\
1 & & 3 & & 3 & & 1
\end{array}$$

The pattern is clearly that of the binomial coefficients, i.e. the entries of Pascal's triangle. This means that the terms in the infinite sum can be regrouped as binomials with the help of the partial derivative operator. For instance, the order-two terms are written

$$\begin{aligned}
&z_{x^2}\Delta x^2 + 2z_{xy}\Delta x\Delta y + z_{y^2}\Delta y^2 \\
&\quad = (\Delta x\,\partial_x + \Delta y\,\partial y)^2 f(x_0,y_0) \ ,
\end{aligned}$$

and similarly for all orders.

Switching to summation notation, we finally have the two-dimensional Taylor's theorem

$$\begin{aligned}
f(x,y) &\approx f(x_0,y_0) \\
&+ \sum_{q=1}^{\infty} \frac{1}{q!}(\Delta x\,\partial_x + \Delta y\,\partial y)^q f(x_0,y_0)
\end{aligned} \tag{1.5}$$

For a sanity check, you can see that if all $y = 0$ then the above reduces to the familiar one-dimensional form.

4 Vectors and Surfaces

4.1 Basis Vectors as Derivatives

In three-dimensional space, there are always three basis vectors from which everything is oriented. In Cartesian coordinates, these are just $\hat{x}$, $\hat{y}$, $\hat{z}$, and are fixed in space. In other systems, such as cylindrical coordinates $\hat{\rho}$, $\hat{\phi}$, $\hat{z}$, and spherical coordinates $\hat{r}$, $\hat{\theta}$, $\hat{\phi}$, each basis vector depends on the coordinates themselves.

In each system mentioned, the respective position vector is:

$$\vec{r} = x\,\hat{x} + y\,\hat{y} + z\,\hat{z}$$
$$\vec{r} = \rho\cos(\phi)\,\hat{x} + \rho\sin(\phi)\,\hat{y} + z\,\hat{z}$$
$$\vec{r} = r\sin(\theta)\left(\cos(\phi)\,\hat{x} + \sin(\phi)\,\hat{y}\right) + r\cos(\theta)\,\hat{z}$$

It's customary using geometry to work out the basis vectors for each system, namely $\hat{\rho}$, $\hat{\phi}$, $\hat{z}$, and also $\hat{r}$, $\hat{\theta}$, $\hat{\phi}$.

Having suffered the tedious derivations once, you're entitled to a secret from the math department. Let q represent any parameter whatsoever - it could be x, or z, or ϕ, etc. It turns out that the basis vector $\hat{q}$ is the normalized q-derivative of the position vector. That is:

$$\hat{q} = \frac{1}{|\partial\vec{r}/\partial q|}\frac{\partial\vec{r}}{\partial q} \tag{1.6}$$

For example, if we want $\hat{\theta}$ from spherical coordinates, write

$$\frac{\partial\vec{r}}{\partial\theta} = r\cos(\theta)\left(\cos(\phi)\,\hat{x} + \sin(\phi)\,\hat{y}\right) - r\sin(\theta)\,\hat{z} \ ,$$

whose magnitude is r. Dividing this out delivers the result promised:

$$\frac{1}{r}\frac{\partial\vec{r}}{\partial\theta} = \hat{\theta}$$

4.2 Surface Tangent Vectors

Parametric Surface Tangents

In the same way that curves $y = f(x)$ can be represented with vectors and parameters, the story is similar for surfaces $z = f(x,y)$. In a generic case, a surface requires two parameters u, v such that

$$\vec{r}(u,v) = x(u,v)\,\hat{x} + y(u,v)\,\hat{y} + z(u,v)\,\hat{z} \ ,$$

which doesn't necessarily need to framed in the Cartesian system.

Choosing any fixed point (u_0,v_0) on a parameterized surface, there exist a pair of embedded tangent vectors we'll call $\vec{u}$, $\vec{v}$ straightforwardly calculated directly from $\vec{r}(u,v)$:

$$\vec{u}(u_0,v_0) = \left(\frac{\partial}{\partial u}\vec{r}(u,v_0)\right)\bigg|_{u_0}$$
$$\vec{v}(u_0,v_0) = \left(\frac{\partial}{\partial v}\vec{r}(u_0,v)\right)\bigg|_{v_0}$$

Like all vectors, the tangents $\vec{u}$, $\vec{v}$ can be converted to normal vectors by dividing out the magnitude:

$$\hat{u} = \vec{u}/u$$
$$\hat{v} = \vec{v}/v$$

Level Curve Tangents

The tangent vectors to a level curve of $z = f(x, y)$ are trickier to determine. To begin, propose choose a point (x_0, y_0) and write the pair of vectors

$$\vec{u}(x_0, y_0) = u_x\,\hat{x} + u_z\,\hat{z}$$
$$\vec{v}(x_0, y_0) = v_y\,\hat{y} + v_z\,\hat{z}\,,$$

where without loss of generality, $\vec{u}$ lacks a y-component and $\vec{v}$ lacks an x-component.

The ratios u_z/u_x, v_z/v_y, receptively, are the partial derivatives in disguise, as

$$\frac{u_z}{u_x} = \left(\frac{\partial}{\partial x} f(x, y_0)\right)\bigg|_{x_0}$$
$$\frac{v_z}{v_y} = \left(\frac{\partial}{\partial y} f(x_0, y)\right)\bigg|_{y_0}\,,$$

which allows the vectors $\vec{u}$, $\vec{v}$ to be written in terms of partial derivatives:

$$\vec{u}(x_0, y_0) = u_x\left(\hat{x} + \left(\frac{\partial}{\partial x} f(x, y_0)\right)\bigg|_{x_0}\hat{z}\right)$$
$$\vec{v}(x_0, y_0) = v_y\left(\hat{y} + \left(\frac{\partial}{\partial y} f(x_0, y)\right)\bigg|_{y_0}\hat{z}\right)$$

For shorthand, denote the fully-evaluated partial derivatives as $f_x(x_0, y_0)$, $f_y(x_0, y_0)$, respectively. Dividing each vector by its own magnitude gives the normalized version of each:

$$\hat{u} = \frac{\hat{x} + f_x\,\hat{z}}{\sqrt{1 + f_x^2}}$$
$$\hat{v} = \frac{\hat{y} + f_y\,\hat{z}}{\sqrt{1 + f_y^2}}$$

4.3 Surface Normal Vector

With a pair of surface tangent vectors $\vec{u}$, $\vec{v}$ in hand for a given point, the cross product of the two yields the vector $\vec{n}$ that is normal to the surface:

$$\vec{n} = \vec{u} \times \vec{v}$$

Parametric Surface Normal

For the parametric surface $\vec{r}(u, v)$, the surface normal is straightforwardly calculated from

$$\vec{n}(u_0, v_0) = \vec{u}(u_0, v_0) \times \vec{v}(u_0, v_0)\ ,$$

which suggests a normalized version

$$\hat{n} = \frac{\vec{u}(u_0, v_0) \times \vec{v}(u_0, v_0)}{|\vec{u}(u_0, v_0) \times \vec{v}(u_0, v_0)|}\ .$$

Of course, there is no need to normalize if we use unit vectors only:

$$\hat{n} = \hat{u} \times \hat{v}$$

Cartesian Surface Normal

The normal vector to the surface $z = f(x, y)$ at a point (x_0, y_0) is the cross product of the tangent vectors $\vec{u}(x_0, y_0)$, $\vec{v}(x_0, y_0)$. Explicitly, this is:

$$\vec{n}(x_0, y_0) = \vec{u} \times \vec{v} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ u_x & 0 & u_x f_x \\ 0 & v_y & v_y f_y \end{vmatrix}\ ,$$

or

$$\vec{n} = u_x v_y\left(-f_x\,\hat{x} - f_y\,\hat{y} + \hat{z}\right)\ .$$

Eliminate the stray coefficients by normalizing:

$$\hat{n} = \frac{-f_x\,\hat{x} - f_y\,\hat{y} + \hat{z}}{\sqrt{1 + f_x^2 + f_y^2}}$$

4.4 Tangent Plane

In either picture, whether it be parametric or Cartesian, the tangent vectors $\vec{u}$, $\vec{v}$ imply the existence of a *tangent plane* to the surface at a given point, much in the same way the slope at a point implies a straight line in the one-dimensional case. The normal vector $\vec{n}$ is always perpendicular to the tangent plane.

If the point (x_0, y_0, z_0) is the base from which the tangent vectors and normal vector are drawn, and (x, y, z) is any other point in space, then the equation of the tangent plane is:

$$\vec{n} \cdot \Delta\vec{x} = 0\,,$$

where

$$\Delta\vec{x} = \langle x, y, z\rangle - \langle x_0, y_0, z_0\rangle\ .$$

From what we know about planes, we can also write

$$ax + by + cz + d = 0$$

to represent the tangent plane. To reconcile this with the vector definition, write out the full dot product:

$$n_x(x - x_0) + n_y(y - y_0) + n_z(z - z_0) = 0\,,$$

or

$$n_x x + n_y y + n_z z + d = 0\,,$$

with

$$d = -n_x x_0 - n_y y_0 - n_z z_0 \; .$$

We can say a bit more about the Cartesian case, as

$$\begin{aligned} n_x &= -f_x \\ n_y &= -f_y \\ n_z &= 1 \end{aligned}$$

would mean

$$-f_x\,(x - x_0) - f_y\,(y - y_0) + (z - z_0) = 0 \; .$$