Mathematics 421 Essay 5
Vector Calculus
Spring 2010

0. Introduction

In our multivariable calculus course, there is a chapter on vector calculus covering the **gradient** of a function of several variables, the **curl** and **divergence** of vector fields in three dimensions, and some integrals that invert these differentiation operators. The different versions of the **fundamental theorem of calculus** are treated separately, which emphasizes the technical differences instead of the relation to the fundamental theorem. Chapter 9 of the present text also gives many details. Another source of information about Vector Calculus is "div, grad, curl, and all that" by h. m. schey (the title page really is written entirely in lower case), fourth edition published by W. W. Norton, New York, 2005 (ISBN 0-393-92516-1). Another complete exposition is not needed; these notes only aim to introduce a collection of exercises illustrating what we need to express the **Laplacian** in order to work with the classical partial differential equations in coordinate systems other that the usual **rectangular** or **Cartesian** coordinates. The appearance of the Laplacian in these equations is based on geometric considerations, with the role of coordinates being only to allow it to be evaluated. Thus, it **must** change its appearance in different coordinate systems.

1. The del operator

In two or three dimensions, formal properties of the differentiation operators are expressed using

$$\nabla = \left\langle \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right\rangle$$

which operates on the function written to its right. When applied to a scalar function f, it gives a vector ∇f, called the **gradient** of f. If $u = f(x, y, z)$, then

$$\nabla f = \left\langle \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z} \right\rangle$$

Such vector functions are often call **vector fields**. In addition, the rules used to define vector products give vector operators that have suggestive notation. For a vector field

$$\mathbf{F}(x, y, z) = \langle P(x, y, z), Q(x, y, z), R(x, y, z) \rangle,$$

the **curl** of $\mathbf{F}$ is the vector field

$$\nabla \times \mathbf{F} = \langle R_y - Q_z, P_z - R_x, Q_x - P_y \rangle,$$

with subscripts used to denote partial derivatives. The **divergence** of $\mathbf{F}$ is the scalar function

$$\nabla \cdot \mathbf{F} = P_x + Q_y + R_z,$$

again with subscripts used to denote partial derivatives.

It is a simple consequence of the definitions that we have the analogs of the product rule:

$$\begin{aligned}
\nabla(fg) &= f\nabla g + (\nabla f)g \\
\nabla \cdot (f\mathbf{G}) &= f\nabla \cdot \mathbf{G} + (\nabla f) \cdot \mathbf{G} \\
\nabla \times (f\mathbf{G}) &= f\nabla \times \mathbf{G} + (\nabla f) \times \mathbf{G} \\
\nabla \cdot (\mathbf{F} \times \mathbf{G}) &= (\nabla \times \mathbf{F}) \cdot \mathbf{G} - \mathbf{F} \cdot (\nabla \times \mathbf{G})
\end{aligned}$$

A **great observation** is that the fundamental rule

$$\frac{\partial^2 u}{\partial x \partial y} = \frac{\partial^2 u}{\partial y \partial x}$$

tells us that $\nabla \times \big(\nabla f\big) = 0$ and $\nabla \cdot \big(\nabla \times \mathbf{F}\big) = 0$.

In two dimensions, a formula defining $f(x, y)$ can be thought of as defining a function of x, y and z, that just happens to be independent of z. The gradient of f has third coordinate zero and first two coordinates giving a vector field in the plane. The curl of a vector field that is independent of z with third coordinate zero has first two coordinates zero, so that its third coordinate can be considered as a scalar derivative of a two dimensional vector field. It can be related to the divergence by noting that, when applied to $\mathbf{F}$, it is the divergence of a vector field obtained by rotating $\mathbf{F}$ through an angle of $\pi/2$ (clockwise).

In higher dimensions, the role of the cross product must be replaced by operators involving things that might be called multivectors. This is mentioned only as a reminder that special techniques used in low dimensions may need to be modified for problems in higher dimensions.

2. Line integrals

If a **force** is represented by a **vector field** (in any number of dimensions), the **work** done by that force in moving an object along a path $\mathcal{P}$ is computed by an integral

$$\int_{\mathcal{P}} \mathbf{F} \cdot \mathbf{ds} = \int_{\mathcal{P}} \mathbf{F} \cdot \mathbf{T}\, ds$$

computed by setting $\mathbf{ds} = \mathbf{T}\, ds = \langle dx, dy, \ldots \rangle$ on the curve $\mathcal{P}$. If $\mathcal{P}$ is given by expressing $x, y, \ldots$ as functions of a **parameter** t, then $dx = (dx/dt)dt$ as usual to give an integral in terms of t. The **chain rule** shows that **any** parameterization may be used to compute the integral, since all will give the same answer.

If $\mathbf{F} = \nabla f$, the **multivariable** chain rule shows that the integral is the difference of the values of f at the endpoints of $\mathcal{P}$, so it is “independent of path”. In particular, the integral is zero if $\mathcal{P}$ is a **closed curve**.

Some convenient parameterizations that can be used to evaluate line integrals are:

(i) **line segments**. The line segment from (x_0, y_0) to (x_1, y_1) is given by $x = x_1 t + x_0(1-t)$, $y = y_1 t + y_0(1-t)$ for $0 \le t \le 1$;

(ii) the **ellipse** $(x/a)^2 + (y/b)^2 = 1$ is given by $x = a\cos t$, $y = b \sin t$ for $0 \le t \le 2\pi$.

A **quantitative** form of the independence of path for ∇f is given by **Green’s theorem**:

$$\oint_{\mathcal{C}} P\, dx + Q\, dy = \iint_{\mathcal{D}} Q_x - P_y\, dA$$

where $\mathcal{C}$ is a curve tracing the boundary of the region $\mathcal{D}$ in the counterclockwise sense, and dA signifies integration with respect to **area**. In rectangular coordinates, $dA = dx\, dy$. For simple regions the proof consists of noting that integrating Q_x with respect to x gives the difference of the values of Q at the endpoints and integrating this result with respect to y is exactly how one computes the path integral of $Q\, dy$, and similarly for the other terms on each side of this formula. Less simple regions need to be cut into simpler pieces. If a region is cut by a line, the integral over the whole region is the sum of the integrals over the pieces. To get closed curves, the pieces of the path integral around the boundary need to be completed with integrals along the cut, but one adds an integral in one direction to one piece and its reverse to the other piece. The sum of these new terms is zero.

It is important to note that this integral is **oriented**. Reversing the direction of motion on the boundary changes the value of the integral to its negative. The two-dimensional integral with respect to area must also have an orientation. This appears in the proof of Green’s theorem in the consideration of the direction followed by the top and bottom as x increases. Reversing the roles of x and y reverses the orientation of the area integral. In elementary work, this is hidden by using a geometric description of the domain of integration that leads directly to the appropriate iterated integral.

Green's theorem can be used to **define** the integral over the region $\mathcal{D}$ with respect to area in terms of the conceptually simpler path integral around the boundary. One application is a change-of-variables formula for double integrals. It can also be used in the other direction to evaluate path integrals as double integrals with an integrand in which a great deal of redundancy has been removed. In particular, we know that the integral of 1 with respect to area is just the area of the region, and the integral of x is the moment with respect to the y axis, which is the product of the area with the x coordinate of the centroid. For simple regions, these are all known on the basis of computations done many times in earlier courses.

In a first course on Multivariable Calculus, exercises usually give one of the integrals related by the theorem and ask to compute it using the **other** integral. Usually, the double integral is used for the computation because many different line integrals correspond to a single double integral and some effort has already been made to find efficient evaluation of double integrals. This is misleading: the direct evaluation of double integrals as iterated integrals is only the construction of a particular line integral; changing variables is proved by consideration of the line integrals. In reality, these exercises only replace one line integral by another one that corresponds to the same double integral. The use of physical properties described in the last paragraph to recognize integrals that had previously been computed is kept secret. An **advanced** study of calculus should include reliable computations to replace error-prone computations based on the basic techniques of calculus.

3. Changes of variables

Suppose that x and y are functions of another pair of variable u and v. If z is a function of x and y, a multivariable chain rule gives the derivatives of z with respect to the new variables by a formula that can be written in matrix form as

$$\begin{bmatrix} z_u \\ z_v \end{bmatrix} = \begin{bmatrix} x_u & y_u \\ x_v & y_v \end{bmatrix} \begin{bmatrix} z_x \\ z_y \end{bmatrix}$$

This expression can be **inverted** to give

$$\begin{bmatrix} z_x \\ z_y \end{bmatrix} = \frac{1}{x_u y_v - x_v y_u} \begin{bmatrix} y_v & -y_u \\ -x_v & x_u \end{bmatrix} \begin{bmatrix} z_u \\ z_v \end{bmatrix}$$

which tells how to directly construct the expressions obtained by evaluating partial derivatives with respect to the old variables, or combining them into a gradient, and substituting expressions in terms of the new variables.

A benefit of the use of matrix notation is that the same formula may be interpreted in different ways: the product of a matrix M and a vector v may be considered as a vector of linear combinations of elements of v whose coefficients are the entries of M or a linear combination of the columns of M whose coefficients are the entries of v. The latter interpretation leads to a chain rule:

$$\nabla f = \nabla u \frac{\partial z}{\partial u} + \nabla v \frac{\partial z}{\partial v}$$

where $z = f(x, y)$ that is expressed in terms of the new variables u and v by the change of variables formula. This is particularly useful where the function f itself has not given, but one has the result of expressing its result in terms of the new variables. The expressions for ∇u and ∇v in terms of u and v are the columns of the inverse of the change of variables matrix that we constructed.

An important example is given by **polar coordinates**, in which the new variables are usually denoted r and θ with $x = r\cos\theta$ and $y = r\sin\theta$. Then, $x_r = \cos\theta$, $y_r = \sin\theta$, $x_\theta = -r\sin\theta$, and $y_\theta = r\cos\theta$. Then $x_r y_\theta - x_\theta y_r = r$ and

$$\nabla r = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \text{ and } \nabla\theta = \frac{1}{r}\begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$$

This formula is remarkable! Although $r = \sqrt{x^2 + y^2}$, there is no **single formula** for θ. The usual formula , $\theta = \arctan y/x$ fails to be defined when $x = 0$. Indeed, there can be no nice function of x and y that gives θ because the value of θ should change by 2π when it is extrapolated along a curve that wraps around the origin.

Given a line integral around a closed curve in the xy-plane

$$I = \oint P\,dx + Q\,dy.$$

and expressions for x and y as functions of new variables u and v so that there is a figure in the uv-plane that is mapped **exactly** to the given curve **and** the region bounded by the curve, we can change variables in the integral to obtain

$$I = \oint (Px_u + Qy_u)\,du + (Px_v + Qy_v)\,dv.$$

Our assumptions allow us to apply Green's Theorem in th uv-plane to get the double integral with respect to u and v of

$$\begin{aligned}
(Px_v + Qy_v)_u - (Px_u + Qy_u)_v = \\
Px_{vu} + Qy_{vu} + (P_x x_u + P_y y_u)x_v \\
+(Q_x x_u + Q_y y_u)y_v \\
-Px_{uv} - Qy_{uv} - (P_x x_v + P_y y_v)x_u \\
-(Q_x x_v + Q_y y_v)y_u \\
= (Q_x - P_y)(x_u y_v - x_v y_u)
\end{aligned}$$

Here the first part is the integrand of the original double integral in the xy plane composed with the definition of x and y in terms of u and v, and the second factor is the local ratio of areas in the two planes. This factor is called the **Jacobian**. A negative Jacobian corresponds to a map that reverses orientation.

4. Stokes' Theorem

A surface in 3-dimensional space can be described by giving x, y and z in terms of two parameters u and v. A region in this surface corresponds to a region in the uv-plane, and the boundary of the region on the surface corresponds to the boundary of region in the uv-plane. To parameterize the boundary in the surface, one takes a parameterization of the curve in the uv-plane and applies the function parameterizing the surface. A simple example will be good enough to illustrate the nature of the proof: let the surface be the graph of a function, $z = g(x, y)$, and take $x = u$ and $y = v$. Then the **projection** of the region on the surface and its boundary are the region and curve in the parameter space.

One side of the Stokes Theorem equation is the **integral of a vector field** $\mathbf{F}$ around a curve $\mathcal{C}$. The projection of this curve is parameterized by ignoring the third coordinate z. The integral of a vector field is given by

$$\begin{aligned}
\int_a^b & P\big(x(t), y(t), z(t)\big)\,x'(t) + \\
& Q\big(x(t), y(t), z(t)\big)\,y'(t) + \\
& R\big(x(t), y(t), z(t)\big)\,z'(t) \quad dt,
\end{aligned}$$

where the values $t = a$ and $t = b$ correspond to **going once around** $\mathscr{C}$ in the counterclockwise direction when viewed from above. Substituting $z = g(x, y)$ in this gives

$$\oint \big(P\big(x, y, g(x, y)\big) + R\big(x, y, g(x, y)\big)\, g_x(x, y)\big)\, dx + \\ \big(Q\big(x, y, g(x, y)\big) + R\big(x, y, g(x, y)\big)\, g_y(x, y)\big)\, dy.$$

Applying Green's theorem requires calculating the difference of

$$\begin{aligned}\frac{\partial}{\partial x}\big(Q\big(x, y, g(x, y)\big) + R\big(x, y, g(x, y)\big)g_2(x, y)\big) = \\ Q_1\big(x, y, g(x, y)\big) + Q_3\big(x, y, g(x, y)\big)g_1(x, y) + \\ R\big(x, y, g(x, y)\big)g_{21}(x, y) + R_1\big(x, y, g(x, y)\big)g_2(x, y) + \\ R_3\big(x, y, g(x, y)\big)g_1(x, y)g_2(x, y)\end{aligned}$$

and

$$\begin{aligned}\frac{\partial}{\partial y}\big(P\big(x, y, g(x, y)\big) + R\big(x, y, g(x, y)\big)g_1(x, y)\big) = \\ P_2\big(x, y, g(x, y)\big) + P_3\big(x, y, g(x, y)\big)g_2(x, y) + \\ R\big(x, y, g(x, y)\big)g_{12}(x, y) + R_2\big(x, y, g(x, y)\big)g_1(x, y) + \\ R_3\big(x, y, g(x, y)\big)g_1(x, y)g_2(x, y)\end{aligned}$$

This simplifies to

$$(R_2 - Q_3)(-g_1) + (P_3 - R_1)(-g_2) + (Q_1 - P_2).$$

The second factors in these terms are the components of $\langle -g_1, -g_2, 1\rangle$ which is **perpendicular to** $\mathscr{S}$. The first factor must then be the components of a **vector field** being integrated over the surface. Note that the normalization of our normal vector to have third coordinate 1 signifies upward orientation and integration with respect to $dx\, dy$. These surface integrals are then the integral of a vector field with respect to a **vector element of surface area** since the area in the tangent plane to the surface over a small rectangle in the xy-plane has area equal to the length of the normal vector we found times the area of the rectangle in the xy-plane.

5. The divergence theorem

If a vector field $\mathbf{G}$ is the curl of another vector field $\mathbf{F}$, then the integral of $\mathbf{G}$ over a piece of a surface depends only on the boundary , so the integral over a closed surface will be zero. We have seen that $\nabla \cdot \mathbf{G} = 0$ in this case. The divergence theorem makes this quantitative by showing that for **all** vector fields $\mathbf{G}$ the integral of $\mathbf{G}$ over a closed surface can be expressed in terms of $\nabla \cdot \mathbf{G}$. More precisely, the surface integral of $\mathbf{G}$ is equal to the integral of $\nabla \cdot \mathbf{G}$ over the region bounded by the surface. The proof is essentially the same as the proof we gave for Green's theorem.

In the plane, an integral **through** a curve, called a **flux** integral can be viewed as the integral of the dot product of the vector field with a unit normal vector with respect to arc length. This interpretation is similar to considering the usual line integral, a **flow** integral, as the integral of the dot product of a vector field with the unit tangent vector with respect to arc length. Thus the flow integral of $\mathbf{F}$ is always equal to the flux integral of the vector field obtained by rotating $\mathbf{F}$ through and angle of $\pi/2$ clockwise. In particular, rotatimg a graient vector field through π/s always gives a vector fireld whose divergence is zero. This can be used to describe the vector calculus of objects expressed in terms of coordiates systems other than the Cartesian system.

6. Exercises

A. Given $f(x, y) = \ln(x^2 + 3y^2)$, find ∇f. Note: $f(x, y)$ is undefined when $(x, y) = (0, 0)$.

B. Given $A\colon (0, 0)$, B :: (3.0), $C\colon (3, 1)$, the boundary of the triangle ABC is traced in the counterclockwise direction by the three intervals AB, BC, CA.

(a) use this description to find

$$\oint_{\mathcal{C}} (2xy - 4y)\,dx + (2x + 1)\,dy$$

around the boundary $\mathcal{C}$ of triangle ABC.

(b) Let $\mathcal{T}$ denote triangle ABC, and apply Green's theorem to write this integral as a double integral with respect to area over $\mathcal{T}$.

(c) Since we have formulas for the area of a triangle and the location of its centroid, we know

$$\iint_{\mathcal{T}} dA = \frac{3}{2}, \quad \iint_{\mathcal{T}} x\,dA = 3, \quad \iint_{\mathcal{T}} y\,dA = \frac{1}{2}.$$

Use this to evaluate the integral in (b). If this does not agree with the value in (a), there is a mistake somewhere. Correct any such mistake before considering the exercise complete.

C. Green's theorem shows that the area of a region with boundary $\mathcal{B}$ is given by

$$\frac{1}{2}\oint_{\mathcal{B}} x\,dy - y\,dx$$

Suppose that $\mathcal{B}$ is given in polar coordinates by an equation $r = g(\theta)$ for $0 \le \theta \le 2\pi$. Determine the result of using this to express the area as an integral with respect to θ. The result should look familiar.

end of supplement