Integration on curves

Differential 1-forms (that is differential forms without the wedge product that we will get to later) can be integrated along curves. To a large extent, that is what they are for. Since differential forms are made of differentials and the definition of the differential of an expression (at least the one that I gave in the handout from January 18) is ultimately about curves, this is a very natural operation.

The definition

Like the textbook does for one-variable Calculus, I'll define the Riemann integral as a limit of Riemann sums, although there are more general notions of integration that can handle more expressions. The Riemann integral will be sufficient for *piecewise continuous* differential forms (those defined in one or more pieces using continuous operations applied to continuous quantities and the differentials of continuously differentiable quantities) along *piecewise continuously differentiable curves* (those with parametrizations defined in one more pieces using continuously differentiable operations applied to the parameter).

So, suppose that we have a differential form α written using the variables $P = (x, y, \ldots)$ and their differentials, and a curve in the same number of dimensions, given by some parametrization function C whose domain is a closed interval $[a, b]$. Then we can try to integrate α along the curve where $P = C(t)$, by defining the integral

$$\int_{P=C(t)} \alpha.$$

Given any way of dividing the interval $[a, b]$ into a partition $a = t_0 \leq t_1 \leq \cdots \leq t_{n-1} \leq t_n = b$ (with n subintervals) and tagging this partition with n values c_k with $t_{k-1} \leq c_k \leq t_k$ for k from 1 to n (this is exactly the kind of partition considered in one-variable Calculus, as on pages 297–299 of the textbook), there is a **Riemann sum**

$$\sum_{k=1}^{n} \alpha\big|_{\substack{P=C(c_k),\\ \mathrm{d}P=C(t_k)-C(t_{k-1})}}.$$

That is, on the kth subinterval, we evaluate the form α at the point through which the curve passes at time c_k within that subinterval along the vector from where the curve is at the beginning of the subinterval to where it is at the end of the subinterval. If we require that the magnitude of this vector be less than δ and take the limit as $\delta \to 0^+$, then this limit (if it exists) is the value of the integral. And there is a theorem that it does exist, at least if α is piecewise continuous and C is piecewise continuously differentiable (and sometimes otherwise); I don't know a nice proof of this directly, but you can prove that it exists because the practical calculation method on page 2 works.

There is now another nice theorem, that the value of this integral does not depend on the parametrization of the curve, at least not very much. That is, if ϕ is a function in the ordinary sense (a real-valued function of one real variable), then $C \circ \phi$ is another parametrized curve; if ϕ is one-to-one and increasing (so that we travel along the curve in the same direction without repetition) and its range lies entirely within the domain of C (so that we cover the entire curve), then the theorem is that $\int_{P=C(t)} \alpha = \int_{P=(C\circ\phi)(t)} \alpha$. The proof is that any Riemann sum for C is also a Riemann sum for $C \circ \phi$; the same points $C(t_k)$ and $C(c_k)$ occur in the same order, just at different values of the parameter. So the Riemann integrals, which are the limits of these Riemann sums, must also be the same.

For this reason, we usually don't specify a parametrized curve in the notation at all. Instead, we specify an **oriented curve**, which is anything that *could* be given as a parametrized curve, keeping track of which direction we travel along the curve (this is the **orientation** of the curve) but otherwise ignoring the parametrization.

Evaluating integrals along curves

The practical method of evaluating integrals along curves is to pick any convenient parametrization (preferably one that is continuously differentiable) and put everything in terms of that parameter. For example, to integrate $2x\,\mathrm{d}x + 3xy\,\mathrm{d}y$ along the top half of the circle $x^2 + y^2 = 4$, oriented counterclockwise, try the parametrization where $x = 2\cos t$, $y = 2\sin t$, and $0 \leq t \leq \pi$. Then $\mathrm{d}x = -2\sin t\,\mathrm{d}t$ and $\mathrm{d}y = 2\cos t\,\mathrm{d}t$, so the value of the integral is

$$\begin{aligned}\int_{\substack{x^2+y^2=4,\ y\geq 0\\ \mathrm{d}x\leq 0}} (2x\,\mathrm{d}x + 3xy\,\mathrm{d}y) &= \int_{t=0}^{\pi} \big(2(2\cos t)(-2\sin t\,\mathrm{d}t) + 3(2\cos t)(2\sin t)(2\cos t\,\mathrm{d}t)\big) \\ &= \int_{t=0}^{\pi} (-8\sin t\cos t + 24\sin t\cos^2 t)\,\mathrm{d}t = 16.\end{aligned}$$

(You can do this last integral with the substitution $u = \cos t$.) I've described the curve of integration with an equation (of a circle) and an inequality (to get the top half only) and oriented it by saying that x is always decreasing (so that $\mathrm{d}x$ is always negative), but usually people write that all out to the side somewhere, call the resulting oriented curve C (for example), and write $\int_C (2x\,\mathrm{d}x + 3xy\,\mathrm{d}y)$.

The reason why this gives the correct result is that any Riemann sum for the integral involving t involves almost the same calculations as a Riemann sum for the integral along the curve. The only difference is that the integral involving t looks at the point from within of each subinterval to handle the differentials, whereas as the integral of the curve looks at the points on each end of the subinterval. But in the limit, all of these points approach each other, and the result is the same. (There is another slight complication because the integral involving t takes a limit as the change in t goes to 0, while the integral along the curve takes a limit as the magnitude of the change in position goes to 0. However, these are the same because the parametrization is continuous. If you can calculate $\mathrm{d}x$ and $\mathrm{d}y$ at all, then the parametrization must be differentiable and so definitely continuous.)

You should be able to visualize this example geometrically well enough to see that the answer would have to be positive. The term $2x\,\mathrm{d}x$ should completely cancel, because the right half of the curve exactly mirrors the left half, with $\mathrm{d}x$ the same on both halves (always negative because of movement to the left) but x being the opposite on the two halves (first positive, then negative). On the other hand, the term $3xy\,\mathrm{d}y$ will be negative on both sides; while y is always positive (above the horizontal axis), x and $\mathrm{d}y$ are both positive on the right half (right of the vertical axis and moving upwards) and both negative on the left half (left of the axis and moving downwards), making for a positive product everywhere.

If you are asked to integrate a vector field $\mathbf{F}$ along an oriented curve, then they really want you to integrate the differential form $\mathbf{F}(x,y) \cdot \langle \mathrm{d}x, \mathrm{d}y \rangle$, or more generally $\mathbf{F}(P) \cdot \mathrm{d}P$, where P is (x,y) or (x,y,z). If you write $\mathbf{r}$ for the vector $P - \mathrm{O}$ (where O is the origin $(0,0)$ or $(0,0,0)$), then $\mathrm{d}P = \mathrm{d}\mathbf{r}$, and this is the reason for the traditional notation $\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r}$, which is used in the textbook. (You may also see $\int_C \mathbf{F} \cdot \mathbf{T}\,\mathrm{d}s$, where $\mathrm{d}s$ is the $\mathrm{d}s$ that appears at the very bottom of this page and $\mathbf{T}$ is defined to be $\mathrm{d}\mathbf{r}/\mathrm{d}s$. This is usually completely pointless; if you see $\mathbf{T}\,\mathrm{d}s$, just think of it as $\mathrm{d}\mathbf{r}$.)

For example, to integrate $\langle 2x, 3xy \rangle$ along the same semicircle as in the previous example (with the same orientation), you do exactly the same integral as in the previous example. This is because

$$\langle 2x, 3xy \rangle \cdot \langle \mathrm{d}x, \mathrm{d}y \rangle = 2x\,\mathrm{d}x + 3xy\,\mathrm{d}y,$$

so

$$\int_C \langle 2x, 3xy \rangle \cdot \mathrm{d}\mathbf{r} = \int_C (2x\,\mathrm{d}x + 3xy\,\mathrm{d}y) = 16$$

as before. Since the vector $\langle 2x, 3xy \rangle$ points to the right on the right side and to the left on the left side, while we move along the curve consistently to the left, this suggests that the horizontal component should cancel. However, since this vector points upwards where we move upwards along the curve (on the right side) and points downwards where we move downards along the curve (on the left side), this suggests a positive contribution from the vertical component. So as in the first example, you should expect a positive result even before doing the calculation.

If you are asked to integrate a function f along a curve, then they really want you to integrate the differential form $f(x,y)\sqrt{\mathrm{d}x^2 + \mathrm{d}y^2}$, or more generally $f(P)\,|\mathrm{d}P|$. It's traditional to write $\mathrm{d}s$ for $|\mathrm{d}P|$ (or

$|\mathrm{d}\mathbf{r}|$, which is the same), but it's important that there is no quantity s defined everywhere on the coordinate plane that $\mathrm{d}s$ is the differential of. To emphasize this, you can write $\text{đ}s$; 'đ' is a symbol that some people use when something is traditionally written with 'd' but is not really a differential.

As long as the differentials $\mathrm{d}x$ etc appear only in $\text{đ}s$, then the result of the integral is independent of orientation, because replacing $\mathrm{d}x$ with $-\mathrm{d}x$ (as would happen upon reversing the orientation) doesn't change $\text{đ}s$. For this reason, you can integrate a function on an *unoriented* curve. When parametrizing, everything will come out using $|\mathrm{d}t|$ instead of $\mathrm{d}t$, but as long as the integral involving t has its bounds set up so that t is increasing, then $\mathrm{d}t$ is positive and so $|\mathrm{d}t| = \mathrm{d}t$, after which you can integrate normally.

For example, to integrate $f(x,y) = 6x^2y$ on the same semicircle as in the previous examples, you get

$$\text{đ}s = \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2} = \sqrt{(-2\sin t\,\mathrm{d}t)^2 + (2\cos t\,\mathrm{d}t)^2} = \sqrt{(4\sin^2 t + 4\cos^2 t)\,\mathrm{d}t^2} = \sqrt{4}\sqrt{\mathrm{d}t^2} = 2\,|\mathrm{d}t|.$$

Thus, the integral is

$$\int_{x^2+y^2=4,\ y\geq 0} 6x^2y\,\text{đ}s = \int_{t=0}^{\pi} 6(\cos t)^2(\sin t)(2\,|\mathrm{d}t|) = \int_{t=0}^{\pi} 12\sin t\cos^2 t\,\mathrm{d}t = 8.$$

Since x^2y is positive everywhere on this curve, you should have expected a positive result.

If for some reason you set the integral up backward, then $\mathrm{d}t$ would be negative and so $|\mathrm{d}t|$ would be $-\mathrm{d}t$, and the result would be the same in the end:

$$\int_C \text{đ}s = \int_{t=\pi}^{0} 12\sin t\cos^2 t\,|\mathrm{d}t| = \int_{t=\pi}^{0} 12\sin t\cos^2 t(-\mathrm{d}t) = -\int_{t=\pi}^{0} 12\sin t\cos^2 t\,\mathrm{d}t = -(-8) = 8.$$

(But it's simpler to always set things up so that the parameter is increasing.)

Pseudooriented curves

In 2 dimensions, you'll sometimes be asked to integrate a vector field *across* a curve rather than *along* it as usual. Although there is no standard notation for this, you can write it as as $\mathbf{F} \times \mathrm{d}\mathbf{r}$ in analogy with the usual $\mathbf{F} \cdot \mathrm{d}\mathbf{r}$. The book sometimes writes $\mathbf{F} \cdot \mathbf{n}\,\mathrm{d}s$, where $\mathbf{n} = \times\mathbf{T}$ and $\mathrm{d}\mathbf{r} = \mathbf{T}\,\mathrm{d}s$, but this just results in $\mathbf{F} \cdot \times\mathrm{d}\mathbf{r} = \mathbf{F} \times \mathrm{d}\mathbf{r}$.

This is the 2-dimensional cross product, so the result is still a scalar. Technically, however, it is actually a **pseudoscalar**, because its sign depends on how you orient the plane (counterclockwise as is the convention, or clockwise instead). Similarly, specifying a direction across a curve really gives the curve a **pseudoorientation**, because it only defines a direction along the curve (an orientation) by picking a convention about how these directions correspond. In practice, we orient the plane counterclockwise, meaning that counterclockwise cross products are positive, the rotation $\times\mathbf{v}$ of a vector $\mathbf{v}$ is obtained by rotating it clockwise, a direction across a curve turns into a direction along it by rotation counterclockwise, and a direction along a curve turns into a direction across it by rotating clockwise. But if you consistently did all of these the other way, then the results of all integrals would be the same.

For example, to integrate $\langle 2x, 3xy\rangle$ across our semicircle, now pseudooriented upwards, integrate

$$\langle 2x, 3xy\rangle \times \langle \mathrm{d}x, \mathrm{d}y\rangle = 2x\,\mathrm{d}y - 3xy\,\mathrm{d}x,$$

and use the orientation counterclockwise from upwards, which is leftwards (the same as in first example):

$$\begin{aligned}
\int_{\substack{x^2+y^2=4,\ y\geq 0\\ \mathrm{d}y\geq 0}} \langle 2x, 3xy\rangle \times \mathrm{d}\mathbf{r} &= \int_{\substack{x^2+y^2=4,\ y\geq 0\\ \mathrm{d}x\leq 0}} (2x\,\mathrm{d}y - 3xy\,\mathrm{d}x) \\
&= \int_{t=0}^{\pi} \Big(\big(2(2\cos t)(2\cos t\,\mathrm{d}t)\big) - 3(2\cos t)(2\sin t)(-2\sin t\,\mathrm{d}t)\Big) \\
&= \int_{t=0}^{\pi} (8\cos^2 t + 24\sin^2 t\cos t)\,\mathrm{d}t = 4\pi.
\end{aligned}$$

Since the vector $\langle 2x, 3xy\rangle$ points to the right where we cross the curve to the right (on the right side) and points to the left where we cross to the left, this suggests that the horizontal component should give a positive result. However, since this vector points upwards on the right side and downwards on the left side, while we cross the curve consistently upwards, this suggests that the vertical component should cancel. So you should again expect a positive result before doing the calculation.

The Fundamental Theorem of Calculus

In one-variable Calculus, the second Fundamental Theorem states that

$$\int_{x=a}^{b} f'(x)\,\mathrm{d}x = f(b) - f(a).$$

If we write u for the quantity $f(x)$, then its differential $\mathrm{d}u$ is precisely the integrand $f'(x)\,\mathrm{d}x$, so the Fundamental Theorem can also be written as

$$\int_{a}^{b} \mathrm{d}u = u|_{a}^{b}.$$

This works just as well when there are several independent variables as when there is just one. Now if $u = f(P)$, then $\mathrm{d}u$ is $\nabla f(P) \cdot \mathrm{d}\mathbf{r}$, so

$$\int_{P=a}^{b} \nabla f(P) \cdot \mathrm{d}\mathbf{r} = f(b) - f(a).$$

Although this is now a theorem about integrating a gradient along a curve, in essence it is still just the FTC, a theorem about integrating differentials. This has a massive generalization to higher-rank differential forms, called the *Stokes Theorem*, which we'll get to later.

A differential form is called **exact** if there exists a quantity u such that $\alpha = \mathrm{d}u$. Similarly, a vector field $\mathbf{F}$ is called **conservative** if there is a scalar field f such that $\mathbf{F} = \nabla f$. The connection between these is that $\mathbf{F}$ is conservative if and only if $\mathbf{F}(P) \cdot \mathrm{d}\mathbf{r}$ is exact. (After all, if $\mathbf{F} = \nabla f$, then $\mathbf{F}(P) \cdot \mathrm{d}\mathbf{r} = \mathrm{d}\big(f(P)\big)$.) An oriented curve is called **closed** if its beginning and ending points are the same; one sometimes emphasizes that an integral is along a closed curve by writing $\oint$ in place of $\int$. Then the integral of an exact differential form or a conservative vector field along a closed curve is zero, because

$$\oint_C \alpha = \int_a^a \mathrm{d}u = u|_a^a = u\big|_a - u\big|_a = 0.$$

Similarly, the integral of a conservative vector field along a closed curve is zero. In this case, we can use notation more like that of a definite integral in one variable:

$$\int_{P=P_1}^{P_2} \alpha$$

means the integral of α along *any* curve from P_1 to P_2. It doesn't matter which curve you use; if C_1 and C_2 are both curves like this, then these combine into a closed curve $C_1 - C_2$, in which you start at P_1, follow C_1 to P_2, then follow C_2 backwards (hence the minus sign) back to P_1. Then

$$\int_{C_1} \alpha - \int_{C_2} \alpha = \oint_{C_1 - C_2} \alpha = 0,$$

so $\int_{C_1} \alpha = \int_{C_2} \alpha$. (This is still undefined if there is *no* curve from P_1 to P_2 through the domain of α. This is analogous to the case in one dimension of an integral $\int_{x=a}^{b} f(x)\,\mathrm{d}x$ where f is undefined somewhere between a and b.)

Conversely, if the integral of a differential form or of a vector field is zero along *every* closed curve, then that differential form must be exact or that vector field must be conservative. The reason is that in this case (and only in this case) we can pick a point P_0 to start from and define a semidefinite integral

$$u = \int_{P=P_0} \alpha = \int_{P_0}^{P} \alpha.$$

Because α is exact, you get the same result no matter which path you use from P_0 to P. (Ideally, the domain of α should be *path-connected*, meaning that there exists a curve between any two points. If not, then you must split the domain into various path-connected components and pick a point in each.) That $\mathrm{d}u = \alpha$ in this case is essentially the multivariable version of the *first* Fundamental Theorem of Calculus.

Given a differential form α, finding such an expression u is a form of *indefinite* integration. It's not practical to check every possible curve, of course, so we need other methods to decide if α is exact, and this can also help us to find u. There are actually several methods; one is given in the textbook, essentially reversing the process of partial differentiation with a kind of partial integration. (If you try this method when α is not exact, then it will fail.)

If the domain of α is reasonably simple, then it's possible to pick a point P_0 and write down a general formula for a parametrized curve from P_0 to any point P. (For example, you could always use a straight line segment, as long as these line segments always lie entirely within the domain.) If you try this method when α is not exact, then you may get a result; but when you check it, then you'll find that it's wrong when α is not exact.

It's often possible to tell ahead of time whether α is exact. To really explain what's going on here, I'll need to talk about the *exterior differential*, which is a topic that we'll get to in a couple of weeks. For now, I'll describe it in terms of partial derivatives. So, if $\alpha = \mathrm{d}u$, then

$$\alpha = \frac{\partial u}{\partial x}\,\mathrm{d}x + \frac{\partial u}{\partial y}\,\mathrm{d}y + \cdots.$$

(The dots are meant to indicate that more terms may appear if there are more than two variables.) Assuming that u is twice differentiable, then mixed second partial derivatives are equal:

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

So if you start with an arbitrary linear differential 1-form

$$\alpha = \alpha_x\,\mathrm{d}x + \alpha_y\,\mathrm{d}y + \cdots,$$

then it could only be exact if it is **closed**, meaning that

$$\frac{\partial \alpha_x}{\partial y} = \frac{\partial \alpha_y}{\partial x}$$

(and similarly for other mixtures of derivatives if there are more than two variables), assuming that it's differentiable in the first place. Similarly, a vector field

$$\mathbf{F}(x, y, \ldots) = \mathbf{F}_1(x, y, \ldots)\mathbf{i} + \mathbf{F}_2(x, y, \ldots)\mathbf{j} + \cdots$$

can only be conservative if it is **irrotational**, meaning that

$$D_2\mathbf{F}_1 = D_1\mathbf{F}_2$$

(and similarly for other mixtures of derivatives if there are more than two variables), assuming that it's differentiable in the first place.

Conversely, a closed differential form or an irrotational vector field must be exact or conservative (respectively) if its domain is **precisely-simply connected**, which means that any simple closed curve (one that doesn't intersect itself except where its two endpoints are equal) in the domain of the differential form or the vector field is the boundary of a region that lies entirely within that domain. (The domain is *simply connected* if it is both path-connected and precisely-simply connected. Conversely, it is precisely-simply connected if each of its path-connected components is simply connected. If you take a class in Topology such as MATH 471 at UNL, then you'll learn a hundred specific terms like these.) But a full discussion of the reasons for this must wait until we've covered higher-order differential forms.