Integration with differential forms

Jessica J. Zhang

October 8, 2020

Roughly speaking, a differential form is just "something we can integrate." When we evaluate the integral $\int xdx$, the xdx part is itself a differential form. In a line integral, where we suddenly have $\int_C fdx + gdy$, the entire integrand, namely $fdx + gdy$, is now a differential form. Later on, when working with higher dimensions, we can have more complicated integrands like $fdxdy + gdxdz$, which would actually be more accurately written with wedges, i.e., as $fdx \wedge dy + gdx \wedge dz$.

Now, it isn't immediately obvious why it's useful to talk about the integrands on their own, or why it is useful to suddenly invent a new operation (called a wedge product).

The main motivation comes when we are trying to integrate over spaces more complex than Euclidean space. These spaces are called smooth manifolds, and it is often difficult to integrate over them using the traditional, coordinate-based approach.[1]

But another, equally important, motivation is that the relationship between integration and changes of coordinate is often unnecessarily complicated. In general, even linear changes of coordinate require that we multiply the integrand by some determinant, called the Jacobian. For example, let's try to integrate $x^2 + y^2$ over, say, the unit circle D. In Cartesian coordinates, the integral we would need to evaluate[2] is

$$\iint_D (x^2 + y^2)\ dA = \int_0^1 \int_0^{\sqrt{1-y^2}} (x^2 + y^2)\ dxdy = \frac{\pi}{2}.$$

But in polar coordinates, we can't simply substitute $x = r\cos\theta$ and $y = r\sin\theta$ to find that the integral is

$$\iint_D r^2\ dA = \int_0^{2\pi} \int_0^1 r^2\ drd\theta = \frac{2\pi}{3}.$$

In particular, as we learn in calculus, the dA part also changes upon a change of coordinates. To figure out how it changes, we take the determinant of a matrix determined by the partial derivatives of r and θ with respect to x and y and, basically, get that we need to multiply the whole thing by a factor of r. This gives us that the integral is actually

$$\iint_D r^2\ dA = \int_0^{2\pi} \int_0^1 r^3\ drd\theta = \frac{\pi}{2},$$

which is what we expect.

Finding the Jacobian when working with $\mathbb{R}^2$ isn't the worst thing in the world. But it gets a bit cumbersome in higher dimensions. And, more importantly, it makes it rather difficult to do algebraic manipulations with integrals. Differential forms effectively encode the change without relying on an outside multiplier (e.g., the Jacobian). In particular, we should think of differential forms (i.e., integrands) as a way to assign a function to each point in such a way that the assigned functions change as the coordinates change. This eliminates many of the complications which arise from changes of coordinate.

Differential forms take some buildup to introduce, but they're actually quite familiar things. In fact, the first terms in each of the three equations above, which takes the integral with respect to some vague, amorphous entity "A," show us exactly how we want differential forms to work! In particular, transferring from Cartesian to polar coordinates doesn't mess up dA; instead, it's that the relationship between dA

[1] This basically comes from the fact that manifolds have less structure than $\mathbb{R}^n$. The standard formula for a line integral, for example, involves some notion of a metric—which not all manifolds have and which we do not always want to require. And taking integrals the traditional way basically always requires a choice of coordinates, which often proves unnecessarily cumbersome. Differential forms bypass this all—while retaining much of the traditional notation (helpful for intuition, consistency, and general happiness).

[2] Warning: Do not try to integrate this by hand.

and $dxdy$ is different than that between dA and $drd\theta$. The problem is that dA doesn't *really* encode the information we need, because it doesn't tell us how to change $drd\theta$ to be equal to $dxdy$. Our goal is to figure out how to make the dA have meaning and, in the process, make $dxdy$ and $drd\theta$ behave just as nicely as the dA does.

1 Tensors

As we said, differential forms are, first and foremost, ways to assign functions. In particular, they assign to each point in a given space a specific *type* of function, known as an alternating tensor.

Let V be a vector space. Then a *(covariant) k-tensor on V* is a function

$$\alpha : \underbrace{V \times \cdots \times V}_{k \text{ copies}} \to \mathbb{R}$$

which is *multilinear*. To be multilinear is to be linear with respect to each coordinate. In particular, we must have

$$\alpha(av_1 + a'v_1', v_2, \ldots, v_k) = a\alpha(v_1, v_2, \ldots, v_k) + a'\alpha(v_1', v_2, \ldots, v_k),$$

and similarly for each of the other coordinates.

Example 1.1. Many familiar functions on vector spaces, such as the dot product and the determinant, are real-valued multilinear functions, and therefore are tensors. The cross product in $\mathbb{R}^3$, on the other hand, is multilinear (technically, bilinear), but is vector-valued and thus isn't a tensor. In general, a 0-tensor is just a constant and a 1-tensor is a real-valued linear map $V \to \mathbb{R}$, i.e., an element of the *dual space* V^*.

The main operation on tensors is called the tensor product. In particular, if α is a covariant k-tensor on V and β is a covariant ℓ-tensor on V, then we define their *tensor product* to be the function given by

$$\alpha \otimes \beta : (v_1, \ldots, v_k, w_1, \ldots, w_\ell) \mapsto \alpha(v_1, \ldots, v_k)\beta(w_1, \ldots, w_\ell).$$

One can show that $\alpha \otimes \beta$ is multilinear and is therefore a covariant $(k + \ell)$-tensor on V.

1.1 Alternating tensors

Generally, we shouldn't expect a tensor to behave in any predictable way if we swap two coordinates. Certain tensors, such as the dot product, remain the same if their arguments are swapped. These are called *symmetric tensors*; although they are very useful, they are not central to our discussion of differential forms, and so we now will happily forget about their existence.

More useful for us are tensors that are similar to the determinant. In particular, the determinant is an n-tensor on $\mathbb{R}^n$, and it has the special property that it switches signs whenever two coordinates (i.e., rows/columns of a matrix) are swapped. In general, we say that a k-tensor is *alternating* if swapping two arguments changes the sign. The set of all alternating k-tensors on a vector space V is denoted $\Lambda^k(V^*)$.[3]

Note that it is not particularly surprising that the kind of tensor which proves useful for building a theory of integration is also the kind of tensor which mimics the determinant. After all, the determinant is indeed a clean way to calculate areas; that is, it gives a coordinate-invariant formula to calculate an area, which is precisely what we want differential forms to be able to do.

Example 1.2. All 0-tensors and 1-tensors are alternating. A 2-tensor is alternating if and only if $\alpha(v, w) = -\alpha(w, v)$. For example, if $V = \mathbb{R}^2$, the map

$$\alpha(v, w) = v^1 w^2 - v^2 w^1$$

is alternating. (Note that we index different vectors with subscripts, while indexing the coordinates of each vector with superscripts; we will stick to this convention throughout this handout.)

[3] The reason for using the dual V^*, rather than V, comes from a different, more abstract definition of tensors as elements of the so-called tensor product of *vector spaces*. In this case, a covariant tensor is simply an element of the tensor product $V^* \otimes \cdots \otimes V^*$.

Even if α and β are alternating tensors, there is no guarantee that their tensor product $\alpha \otimes \beta$ is.

Example 1.3. Suppose α is the 2-tensor in Example 1.2 and β is the 1-tensor taking everything to 1. Then

$$(\alpha \otimes \beta)(u, v, w) = \alpha(u, v).$$

However, if we swap u and w, then we find that

$$(\alpha \otimes \beta)(w, v, u) = \alpha(w, v),$$

which does not have any relationship to $\alpha(u, v)$.

There is, however, a projection from the set of all k-tensors to the set of alternating k-tensors. If α is a k-tensor on V, then its *alternation* is given by

$$\operatorname{Alt}(\alpha)(v_1, \dots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma) \alpha(v_{\sigma(1)}, \dots, v_{\sigma(k)}),$$

where S_k is the set of all permutations of k elements. By $\operatorname{sgn} \sigma$, we mean the sign of the permutation; this is either $+1$ or -1.

The actual formula of $\operatorname{Alt}(\alpha)$ is not critical to remember. The main idea is that it is obtained by applying α on every possible permutation of the v_i's, and then taking some alternating sum of the results. When $k = 2$, this gives

$$\operatorname{Alt}(\alpha)(v, w) = \frac{1}{2}(\alpha(v, w) - \alpha(w, v)).$$

Note that $\operatorname{Alt}(\alpha)$ is indeed alternating in this case, because

$$\operatorname{Alt}(\alpha)(w, v) = \frac{1}{2}(\alpha(w, v) - \alpha(v, w)) = -\operatorname{Alt}(\alpha)(v, w).$$

When $k = 3$, we have

$$\operatorname{Alt}(\alpha)(u, v, w) = \frac{1}{6}(\alpha(u, v, w) - \alpha(v, u, w) - \alpha(u, w, v) - \alpha(w, v, u) + \alpha(v, w, u) + \alpha(w, u, v)).$$

It is relatively simple, though tedious, to check in this case, too, that $\operatorname{Alt}(\alpha)$ is alternating. Indeed, we generally have the following proposition.

Proposition 1.4. *The alternation of a k-tensor α is an alternating k-tensor. That is,* $\operatorname{Alt}(\alpha) \in \Lambda^k(V^*)$.

Thus we should think of $\operatorname{Alt}(\alpha)$ as simply an alternating tensor defined by α. The particulars of the definition are not very important to us.

1.2 Wedge products

We already defined the tensor product of two tensors. However, we can define an even more useful operation, known as the *wedge product*, on alternating tensors. If $\alpha \in \Lambda^k(V^*)$ and $\beta \in \Lambda^\ell(V^*)$, then we write

$$\alpha \wedge \beta = \frac{(k + \ell)!}{k!\ell!} \operatorname{Alt}(\alpha \otimes \beta).$$

More explicitly, we have

$$(\alpha \wedge \beta)(v_1, \dots, v_{k+\ell}) = \frac{1}{k!\ell!} \sum_{\sigma \in S_{k+\ell}} (\operatorname{sgn} \sigma) \alpha(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \beta(v_{\sigma(k+1)}, \dots, v_{\sigma(k+\ell)}).$$

Note that $\alpha \otimes \beta$ is a $(k + \ell)$-form, and so its alternation is also a $(k + \ell)$-form.

Proposition 1.5. *The wedge product satisfies is bilinear, associative, and anticommutative, which means that* $\alpha \wedge \beta = (-1)^{k\ell}\beta \wedge \alpha$*. Moreover, suppose* V *(and therefore* V^**) is an* n*-dimensional vector space. Let* V^* *have basis* $\{\varepsilon^1, \dots, \varepsilon^n\}$*. Then* ε^i *is a 1-tensor and*

$$\{\varepsilon^{i_1} \wedge \cdots \wedge \varepsilon^{i_k} : 1 \leq i_1 < \cdots < i_k \leq n\}$$

is a basis for $\Lambda^k(V^*)$*.*

The main takeaway of this proposition is just that the wedge product behaves nicely, is anticommutative, and allows us to construct a natural basis for the vector space of all alternating k-tensors on V. Often, we denote $\varepsilon^{i_1} \wedge \cdots \wedge \varepsilon^{i_k}$ as ε^I, where $I = (i_1, \dots, i_k)$ is an *ascending multi-index*, i.e., where $i_1 < \cdots < i_k$.

An important note

The constant multiple in the definition of the wedge product is technically not necessary, and is not always included, but is helpful because it provides a particularly nice (i.e., constant-free) basis for $\Lambda^k(V^*)$.

Recall that determinants are alternating tensors themselves. In fact, determinants basically *define* our basis elements ε^I. In particular, the typical definition of ε^I is actually as the k-tensor such that

$$\varepsilon^I(v_1, \dots, v_k) = \det \begin{pmatrix} \varepsilon^{i_1}(v_1) & \dots & \varepsilon^{i_1}(v_k) \\ \vdots & \ddots & \vdots \\ \varepsilon^{i_k}(v_1) & \dots & \varepsilon^{i_k}(v_k) \end{pmatrix}.$$

Recall that ε^{i_1} is a 1-tensor, meaning that it takes in a vector and returns some real number, so this definition makes sense.

Generally, we then prove that the collection of all ε^I for ascending multi-indices I is a basis for $\Lambda^k(V^*)$. We finally show that $\varepsilon^{i_1} \wedge \cdots \wedge \varepsilon^{i_k} = \varepsilon^I$, where $I = (i_1, \dots, i_k)$.

The proof of each of these propositions is relatively involved and not particularly enlightening. However, it is worth being aware of this determinant definition of the basis vectors for $\Lambda^k(V^*)$.

In particular, it gives us the following determinant-based interpretation for wedge products of 1-forms.

Proposition 1.6. *Given 1-forms* $\omega^1, \dots, \omega^k$ *and vectors* $v_1, \dots, v_k$*, we have*

$$\omega^1 \wedge \cdots \wedge \omega^k(v_1, \dots, v_k) = \det(\omega^j(v_i)). \tag{1}$$

Proof. When $\omega^i = \varepsilon^i$, this is true by the note above that $\varepsilon^{i_1} \wedge \cdots \wedge \varepsilon^{i_k} = \varepsilon^I$. Now simply observe that both sides are linear in ω^i, from which the proposition follows. □

2 Differential forms

Differential forms basically assign an alternating tensor to each point of the space over which we are performing our desired integration.

To be more specific, let $M \subset \mathbb{R}^n$ be a *domain of integration*. This just means that it is a bounded subset of Euclidean space which doesn't include its boundary.[4] At each point $p \in \mathbb{R}^n$, let $T_p\mathbb{R}^n$ denote the *tangent space*. This basically consists of all the possible directions we can go from p, which we should think of as just a copy of $\mathbb{R}^n$ whose origin is at p. If $\mathbb{R}^n$ has coordinates $x^1, \dots, x^n$, then we write the basis of $T_p\mathbb{R}^n$ as

$$\left.\frac{\partial}{\partial x^1}\right|_p, \dots, \left.\frac{\partial}{\partial x^n}\right|_p.$$

For simplicity, we often identify this space with $\mathbb{R}^n$ by writing

$$v^1 \left.\frac{\partial}{\partial x^1}\right|_p + \cdots + v^n \left.\frac{\partial}{\partial x^n}\right|_p = (v^1, \dots, v^n) \in \mathbb{R}^n.$$

[4]We technically also need that M has *measure zero* boundary. This basically translates to a requirement that the boundary of M (which isn't part of M, but is part of $\mathbb{R}^n$!) has dimension less than n.

Note that $(T_p\mathbb{R}^n)^*$, which we typically just denote as $T_p^*\mathbb{R}^n$, also has dimension n.

Then a *(differential) k-form defined on M* is just a smooth alternating k-tensor field, i.e., a smooth function ω which assigns to each point $p \in M$ an alternating k-tensor $\omega_p = \omega(p) \in \Lambda^k(T_p^*\mathbb{R}^n)$. We denote the set of all differential k-forms on M as $\Omega^k(M)$.

The last part of the definition that we haven't explained yet is the word "smooth." In general, the issue of smoothness only comes into play if we are trying to check if something is or isn't a differential form; we won't worry ourselves about that here, and will just assume that everything that looks like a differential form is one, and so we can safely ignore the smoothness criterion.[5]

Example 2.1. Let $f(x, y) = x^2 + y^2$. This is a smooth function. This gives a smooth 2-form ω taking p to

$$\omega_p = f(x, y)\alpha,$$

where α is the 2-tensor in Example 1.2 taking (v, w) to $v^1w^2 - v^2w^1$ and (x, y) are the Cartesian coordinates of p. The unit disk (without boundary) in $\mathbb{R}^2$ is an example of a domain of integration. These two will be our prototypical examples in the remainder of this paper.

2.1 A basis for $\Omega^k(M)$

In this subsection, we focus on 0-forms, which are just smooth functions, and 1-forms, which are smooth maps taking a point $p \in M$ to a linear functional $\omega_p : T_p\mathbb{R}^n \to \mathbb{R}$. We often refer to 1-forms as *covector fields*, and the linear functionals ω_p as *covectors*.

Our goal is to define and understand the differential of a function; this will allow us in Section 2.2 to define a *differential operator*.

Given a smooth function $f : M \to \mathbb{R}$, we define the *differential of f* to be the map df taking a point $p \in M$ to the map

$$\begin{aligned} df_p : T_p\mathbb{R}^n &\to T_{f(p)}\mathbb{R} \cong \mathbb{R} \\ v &\mapsto \sum_{i=1}^{s} \frac{\partial f}{\partial x^i}\bigg|_p (v), \end{aligned} \tag{2}$$

where $x^1, \ldots, x^n$ are the coordinate axes of $\mathbb{R}^n$. This is a 1-form.

This definition of the differential is analogous to the traditional definition of a total derivative; it effectively provides a map taking the tangent space at p to the tangent space at $f(p)$. As expected, it satisfies several natural properties. In particular, the differential function is linear; satisfies the product (hence quotient) rule, as well as a version of the chain rule; and is zero when f is a constant map.

If $x^j : \mathbb{R}^n \to \mathbb{R}$ denotes the projection of $p = (p^1, \ldots, p^n)$ to its j-th coordinate,then $(dx^j)_p$ is simply the map with

$$(dx^j)_p(v) = \sum_{i=1}^{n} \frac{\partial x^j}{\partial x^i}\bigg|_p (v) = v^j,$$

where $v = (v^1, \ldots, v^n)$. Note that the x^j is a map and x^i is a coordinate, so dx^j is the differential 1-form which assigns to each point p the constant map p^j.

Observe, moreover, that $\{dx^j\}$ is a basis for $T^*\mathbb{R}^n$, which is shorthand for saying that $\{(dx^1)_p, \ldots, (dx^n)_p\}$ is a basis for $T_p^*\mathbb{R}^n$ for every $p \in M$. After all, the basis of $T_p^*\mathbb{R}^n$ can be seen to be equal[6] to the collection of functions $f^1, \ldots, f^n$ such that

$$f^j\left(\frac{\partial}{\partial x^i}\bigg|_p\right) = \delta_{i,j} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{otherwise} \end{cases}.$$

[5] To actually make this definition, we basically need to put a smooth structure on $\coprod_p \Lambda^k(T_p^*\mathbb{R}^n)$. But that's kind of annoying, and not very interesting. As long as you don't *try* to break this definition, it won't break.

[6] That is, the typical proof that the dual space has the same dimension as the original space uses this as its basis.

Yet the definition of dx^j shows that it satisfies this property, and so $\{(dx^j)_p\}$ is indeed a basis for $T_p^*\mathbb{R}^n$. In particular, every 1-form can be written as

$$\omega = \sum_{j=1}^{n} \omega_j dx^j,$$

where

$$\omega_p(v) = \sum_{j=1}^{n} \omega_j(v^1, \ldots, v^n)(dx^j)_p.$$

Note that ω_p denotes the alternating 1-tensor $\omega(p)$, while ω_j denotes the j-th "coordinate" of ω.

In Proposition 1.5, we said that a basis for $\Lambda^k(T_p^*\mathbb{R}^n)$ was given by the wedges of basis elements for $T_p^*\mathbb{R}^n$. Hence, taking all $p \in M$ at once, it follows by definition that a basis for $\Omega^k(M)$ is given by the wedges of basis elements for $\Omega^1(M)$, where everything is taken pointwise. In other words, every $\omega \in \Omega^k(M)$ has the property that there exist unique maps ω_I such that for each $p \in M$ we can write

$$\omega_p(v) = \sum_{I=(i_1,\ldots,i_k)} \omega_I(v^1, \ldots, v^s)(dx^{i_1})_p \wedge \cdots \wedge (dx^{i_k})_p.$$

We are summing over all ascending k-indices, i.e., indices $I = (i_1, \ldots, i_k)$ where $i_1 < \cdots < i_k$. As shorthand, we write

$$\omega = \sum \omega_I dx^I.$$

Here dx^I denotes the wedge of all dx^{i_j}'s, *not* the differential of some function x^I. Note that we are using a slightly different definition of the wedge product, taken between differential forms, rather than tensors: The *wedge product* of two differential forms ω and η to be the differential form $\omega \wedge \eta$ which takes p to $\omega_p \wedge \eta_p$.

Example 2.2. Suppose we have coordinates (x, y) on $\mathbb{R}^2$. (Note that we can do this exercise with any coordinates on $\mathbb{R}^2$, e.g., polar coordinates.) Then dx is the map taking v to its x-coordinate, and similarly for dy. Thus, at every point p, the 2-form $dx \wedge dy$ takes (v, w) to

$$\begin{aligned} 2\,\mathrm{Alt}\Big[(dx)_p \otimes (dy)_p)(v, w)\Big] &= 2 \cdot \frac{1}{2} \cdot \Big[(dx \otimes dy)(v, w) - (dx \otimes dy)(w, v)\Big] \\ &= dx(v)dy(w) - dx(w)dy(v). \end{aligned}$$

This is exactly equal to

$$(dx \wedge dy)(v, w) = v^1 w^2 - w^1 v^2.$$

Hence we can write ω from Example 2.1 in terms of the basis elements:

$$\omega = (x^2 + y^2)\; dx \wedge dy.$$

(Not that, in this case, is only one basis element.)

2.2 Exterior derivatives

This differential map on 0-forms, which is closely related to the definition of a derivative, can be generalized to k-forms, namely to functions $d : \Omega^k(M) \to \Omega^{k+1}(M)$ for all integer $k \geq 0$. We call this natural differential on smooth forms the exterior derivative.

Let $\omega \in \Omega^k(M)$ and write $\omega = \sum \omega_I dx^I$. Then we define the *exterior derivative* of ω to be

$$d\omega = \sum d\omega_I \wedge dx^I \in \Omega^{k+1}(M).$$

Recall that ω_I is just a function, so $d\omega_I$ can be computed as in Equation (2). Moreover, when writing dx^I, the x^I does not denote some *function*; instead, we have $dx^I = dx^{i_1} \wedge \cdots \wedge dx^{i_k}$. To be more explicit, we have

$$d\left(\sum_I \omega_I dx^I\right) = \sum_I \left[\left(\sum_{i=1}^{n} \frac{\partial \omega_I}{\partial x^i} dx^i\right) \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}\right].$$

Note that d really denotes infinitely many derivative functions. If we were being particularly scrupulous, we would have d_0 taking 0-forms to 1-forms, and d_1 taking 1-forms to 2-forms, and so on. But we can ignore the subscripts because which d we are using is implied automatically by the form of which we are taking the exterior derivative (i.e., if $\omega \in \Omega^k(M)$, then the d in $d\omega$ is clearly "d_k," where we borrow our slightly tongue-in-cheek definitions of d_0, d_1, etc.).

For convenience, we collect a few properties of exterior differentiation below.

Proposition 2.3. *The exterior derivative satisfies the following properties:*

- *If $\omega, \eta \in \Omega^k(M)$ and $a, b \in \mathbb{R}$, then $d(a\omega + b\eta) = ad\omega + bd\eta$;*
- *If $\omega \in \Omega^k(M)$ and $\eta \in \Omega^\ell(M)$, then $d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^{k\ell}\omega \wedge d\eta$; and*
- *For any $\omega \in \Omega^k(M)$, we have $(d \circ d)(\omega) = 0 \in \Omega^{k+2}(M)$.*[7]

Example 2.4. Recall our "prototypical example" of $\omega = (x^2 + y^2)\, dx \wedge dy$. Its exterior derivative is

$$d\omega = d(x^2 + y^2) \wedge (dx \wedge dy) = (2x\, dx + 2y\, dy) \wedge (dx \wedge dy).$$

But because the wedge product is anticommutative, this is simply equal to

$$2x - 2y)\, dx \wedge dy.$$

3 Integrating differential forms

Before we are able to fully introduce the theory of integration, we must briefly discuss the idea of an *orientation*. For $\mathbb{R}^n$, this amounts to choosing a positive direction for each of the n coordinate axes. A domain of integration M always has an orientation defined by the orientation of $\mathbb{R}^n$. In three dimensions, for example, the right-hand rule gives the orientation for a lower-dimensional space. There is an opposite orientation of M, achieved by taking the negative of our normal vector, which we denote as $-M$.

The important takeaway from this digression on orientation is that all domains of integration have two orientations. For example, the interval $[0, 1]$ has a positive orientation which moves rightward, and a negative one which corresponds to an arrow pointing to the left. A circle on the xy-plane in $\mathbb{R}^3$ has two possible orientations: The normal can point upward or downward.

Let ω be a differential n-form defined on M. Note that there is only one ascending multi-index in this case—namely $I = (1, 2, \ldots, n)$—and so we can write

$$\omega = f dx^I = f dx^1 \wedge \cdots \wedge dx^n$$

for some smooth function $f : M \to \mathbb{R}$. Then we define the *integral of ω over M* as

$$\int_M \omega = \int_M f dx^1 \wedge \cdots \wedge dx^n = \int_M f dx^1 \ldots dx^n = \int_M f dV.$$

In other words, to calculate the integral of a differential form, we simply "erase the wedges"!

Maybe this is slightly anticlimactic; the integral of a differential form is, after all, basically the exact same as a normal integral. It isn't immediately obvious that smooth maps won't complicate this formula (and it isn't obvious that this formula isn't just a trick of notation). But the following theorem shows why differential forms are in fact the "natural" integrand.

Theorem 3.1. *Suppose M and N are domains of integration. Suppose that $G : M \to N$ is an orientation-preserving or orientation-reversing diffeomorphism. If ω is an n-form on N, then*

$$\int_M G^*\omega = \begin{cases} \displaystyle\int_N \omega & \text{if } G \text{ is orientation-preserving, and} \\ \displaystyle -\int_N \omega & \text{if } G \text{ is orientation-reversing.} \end{cases}$$

The statement of this theorem needs some untangling. A *diffeomorphism* is just a bijective smooth (infinitely differentiable) map whose inverse is also smooth.[8] And G^* is something called the *pullback map*,

[7] This observation actually leads to something called *de Rham cohomology*, which is a cohomology theory based on the spaces $\Omega^k(M)$ and the functions d.

[8] If "bicontinuous" means "continuous with continuous inverse," then can we say that a diffeomorphism is "bismooth"?

which is important enough that we'll spend the next section talking about it.

3.1 Pullbacks of differential forms

Suppose $M, N \subseteq \mathbb{R}^n$ are domains of integration. Let $\omega = f\ dx^1 \wedge \cdots \wedge dx^n$ be a differential n-form on $N \subseteq \mathbb{R}^n$ and let $G : M \to N$ be a smooth map. We define the *pullback* of ω by G as the n-form

$$G^*\omega = (f \circ G)\ d(x^1 \circ G) \wedge \cdots \wedge d(x^n \circ G).$$

In general, if ω is a k-form, then the pullback is defined as

$$G^*\omega = G^*\left(\sum_{I=(i_1,\ldots,i_k)} \omega_I\ dx^I\right) = \sum_{I=(i_1,\ldots,i_k)} (\omega_I \circ G)\ d(x^{i_1} \circ G) \wedge \cdots \wedge d(x^{i_k} \circ G).$$

This is a rather unintuitive definition, but should be thought of as the natural way to pull a differential form on N back into one on M.

Example 3.2. As before, let $\omega = (x^2 + y^2)\ dx \wedge dy$ be a 2-form on $\mathbb{R}^2$. Consider the substitution $x = r\cos\theta$, $y = r\sin\theta$. This amounts to the identity diffeomorphism G on $\mathbb{R}^2$, where we consider the domain under $r\theta$-coordinates and the codomain under xy-coordinates. More specifically, G takes (r, θ) to $(x, y) = (r\cos\theta, r\sin\theta)$. Then we find $G^*(\omega)_p$ is the map taking (v, w) to

$$(1 \circ G)(x^2 + y^2)\ d(x \circ G) \wedge d(y \circ G).$$

We can compute $d(x \circ G)$ as follows:

$$d(x \circ G) = d(r\cos\theta) = \left.\frac{\partial}{\partial r}\right|_p (r\cos\theta) + \left.\frac{\partial}{\partial \theta}\right|_p (r\cos\theta) = \cos\theta dr - r\sin\theta\ d\theta.$$

We can similarly compute $d(y \circ G)$ to find that

$$(1 \circ G)(x^2 + y^2)\ d(x \circ G) \wedge d(y \circ G) = r^2(\cos\theta\ dr - r\sin\theta\ d\theta) \wedge (\sin\theta\ dr + r\cos\theta\ d\theta).$$

Using properties of the exterior derivative—namely, the fact that $d \circ d = 0$ and that the wedge product is a bilinear operation—we can find that this is equal to

$$r^2(r\cos^2\theta\ dr \wedge d\theta - r\sin^2\theta\ d\theta \wedge dr) = r^3\ dr \wedge d\theta.$$

Notice the extra factor of r, just as we'd have in the formula for changing an integral from Cartesian to polar coordinates! (Hint: This isn't a coincidence.)

The pullback helpfully is well-behaved with respect to wedge products and exterior differentiation. In fact, you can try to verify the following proposition.

Proposition 3.3. *Suppose $G : M \to N$ is smooth.*

(a) For every k, $G^* : \Omega^k(N) \to \Omega^k(M)$ is an $\mathbb{R}$-linear map.

(b) Pullbacks distribute over wedge products: $G^*(\omega \wedge \eta) = (G^*\omega) \wedge (G^*\eta)$.

(c) Pullbacks commute with the exterior differentiation operator: $G^*(d\omega) = d(G^*\omega)$.

The most helpful part of the pullback is that for n-forms on $\mathbb{R}^n$, we have the following theorem, which shows that the pullback accounts for the Jacobian, thus encoding a change of variables which usually would need to be a different term in our integrand. This effectively is the heart of integrating with differential forms.

Theorem 3.4. *Suppose $M, N \subseteq \mathbb{R}^n$ are domains of integration. Suppose, moreover, that the $\mathbb{R}^n$ in which M is embedded has coordinates (x^i), while the $\mathbb{R}^n$ in which N is embedded has coordinates (y^i). If $G : M \to N$ is a smooth map, then*

$$G^*(f dx^1 \wedge \cdots \wedge dx^n) = (f \circ G)(\det DG)\, dx^1 \wedge \cdots \wedge dx^n, \tag{3}$$

where DG represents the Jacobian matrix of G in these coordinates, namely

$$DG = \begin{pmatrix} \dfrac{\partial G_1}{\partial x^1} & \cdots & \dfrac{\partial G_1}{\partial x^n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial G_n}{\partial x^1} & \cdots & \dfrac{\partial G_n}{\partial x^n} \end{pmatrix}.$$

Proof. To show that the two sides of the equation are the same, we just need to show that they take any behave the same on the vector $(\partial/\partial x^1|_p, \ldots, \partial/\partial x^n|_p)$ for every p. For simplicity, we omit the subscripts p in the rest of the proof.

The definition of the pullback tells us that

$$G^*(f dx^1 \wedge \cdots \wedge dx^n) = (f \circ G)\, d(x^1 \circ G) \wedge \cdots \wedge d(x^n \circ G).$$

Note that $x^i \circ G$ is, by definition, the i-th coordinate G_i of G under the x-coordinates of M. Hence this is equal to

$$(f \circ G)\, dG^1 \wedge \cdots \wedge dG^n.$$

But Proposition 1.6 implies that

$$(dG^1 \wedge \cdots \wedge dG^n)\left(\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}\right) = \det\left(dG^j\left(\frac{\partial}{\partial x^i}\right)\right) = \det\left(\frac{\partial G^j}{\partial x^i}\right) = \det DG,$$

where the second equality follows from the definition of the differential of a function. This implies that the left side of Equation (3) takes the vector of partial derivative functions to

$$(f \circ G)(\det DG).$$

Basically by definition, the right side of the equation also takes $(\partial/\partial x^1, \ldots, \partial/\partial x^n)$ to $(f \circ G)(\det DG)$. ☐

Again, all this really says is that the pullback takes into account the Jacobian determinant and, as such, it adjusts to changes of coordinate. Compare this to Example 3.2.

3.2 Why differential forms are good

Now recall our main theorem:

Theorem 3.1. *Suppose M and N are domains of integration. Suppose that $G : M \to N$ is an orientation-preserving or orientation-reversing diffeomorphism. If ω is an n-form on N, then*

$$\int_M G^*\omega = \begin{cases} \displaystyle\int_N \omega & \text{if } G \text{ is orientation-preserving, and} \\ \displaystyle -\int_N \omega & \text{if } G \text{ is orientation-reversing.} \end{cases}$$

Proof. Let's use $(x^1, \ldots, x^n)$ for the coordinates on M and $(y^1, \ldots, y^n)$ for those on N. Write $\omega = f\, dy^1 \wedge \cdots \wedge dy^n$. Recall the change of variables formula, which tells us that

$$\int_N \omega = \int_N f\, dV = \int_N (f \circ G)|\det DG|\, dV.$$

If G is orientation-preserving, the Jacobian determinant $\det DG$ is nonnegative. Thus, using Theorem 3.4, we find that this is equal to

$$\int_D (f \circ G)(\det DG)\, dx^1 \wedge \cdots \wedge dx^n = \int_D G^*\omega.$$

Otherwise, if G is orientation-reversing, the determinant $\det DG$ is negative, so taking the absolute value of it swaps the sign of the integral. Other than that, it is the same proof. □

Example 3.5. One final time, let's look at $\omega = dx \wedge dy$. Let M and N be the unit circle with polar and Cartesian coordinates, respectively. Recall from our previous example that $G^*\omega = r\, dr \wedge d\theta$. Now observe that the left side of the equation in Theorem 3.1 is

$$\int_M G^*\omega = \int_M r^3\, dr \wedge d\theta = \int_M r^3\, dr d\theta.$$

We calculated this integral in the very beginning: $\frac{\pi}{2}$. On the right side of the equation, we have

$$\int_N \omega = \int_N (x^2 + y^2)\, dx \wedge dy = \int_N (x^2 + y^2)\, dx dy = \frac{\pi}{2}.$$

Thus the two expressions are indeed equal.

Really, at the heart of this all is the fact that wedges behave like determinants (as mentioned in the "important note" at the end of Section 1). The proof of Theorem 3.1 is just a direct consequence of this. The pullback gives a natural way to encode the Jacobian, which thus allows us to work with changes of variables without worrying about extra factors.