The Wonders of Integration
Northwestern University, Spring 2014

Here we outline two basic ideas which every calculus student should know: all integrals we have ever considered are the "same" and all integration theorems we have ever considered are the "same". Now, by "same" we don't mean literally the same, but rather in the sense that they all reflect the same basic idea.

All Integrals Are The Same

Consider the following diagram:

$$\begin{array}{ccccc}
\int_{[a,b]} f(x)\,dx & \longrightarrow & \iint_D f(x,y)\,dA & \longrightarrow & \iiint_E f(x,y,z)\,dV \\
\downarrow & & \downarrow & & \downarrow \\
\int_C f(x,y)\,ds & \longrightarrow & \iint_S f(x,y,z)\,dS & \longrightarrow & ? \\
\downarrow & & \downarrow & & \downarrow \\
\int_C f(x,y,z)\,ds & \longrightarrow & ? & \longrightarrow & ?
\end{array}$$

which contains all types of integrals we've seen: an ordinary single-variable integral, a double integral, a triple integral, (scalar) line integrals, and a (scalar) surface integral. Ignore the question marks for now.

Two things to note: first, usually you would write a single-variable integral as $\int_a^b f(x)\,dx$, but here we write it in a way so as to emphasize the idea that it is an integral "over" the interval $[a,b]$ to be consistent with the other integral notations; and second, vector line and surface integrals are just special cases of scalar line and surface integrals (where the "function" you integrate comes from vector field dot tangent vector in the line integral case and vector field dot normal vector in the surface integral case), so even these types of integrals are covered by this diagram.

Here is the idea to recognize:

> All of these types of integrals amount to "adding" up the values of a function as we vary through all points of some geometric object.

The only difference between these integrals comes from the types of functions we consider (depending on the number of variables) and the types of geometric objects we integrate over.

We use the term *ambient space* to mean the "larger" space which a geometric object sits inside of. For instance, for a circle in $\mathbb{R}^2$ the ambient space is $\mathbb{R}^2$ but for a circle in $\mathbb{R}^3$ the ambient space is $\mathbb{R}^3$. The ambient space is what determines the types of functions we consider, where the number variables corresponds to the dimension of the ambient space. So, in the diagram, the very first integral in the upper left is one where the ambient space is $\mathbb{R}$, reflecting the fact that we are integrating a single-variable function. Going "one diagonal" down gives the (two-variable) line integral in the second row and double integral in the first, where in both of these the ambient space is $\mathbb{R}^2$. Going one diagonal down further gives integrals where the ambient space is $\mathbb{R}^3$: a (three-variable) line integral in the lower left, a surface integral in the middle, and a triple integral in the upper right. Thus:

When moving from an entry in one row to the corresponding entry in the next row, the dimension of the ambient space involved increases by 1. In other words, the vertical arrows in the diagram indicate moving to a larger-dimensional ambient space.

In particular, looking at the integrals in the first column, we move from ambient space $\mathbb{R}$, to ambient space $\mathbb{R}^2$, to ambient space $\mathbb{R}^3$, while in the integrals in the second column we move from ambient space $\mathbb{R}^2$ to ambient space $\mathbb{R}^3$.

Now, every integral in the first column involves a 1-dimensional region of integration: an interval in the first and curves in the second and third entries. If we view an interval as a "curve" in the ambient space $\mathbb{R}$, as we should, then all integrals in the first column are taken over curves, with the only difference between the type of ambient space that these curves sit inside of. Similarly, the integrals in the second column involve 2-dimensional regions of integration: a region in the xy-plane for the double integral and a surface in the surface integral. Of course, we should view a region in the xy-plane as a "flat" surface, so that both integrals in the second column take place over surfaces, with the difference between the type of ambient space that the surface sits inside of: a surface in $\mathbb{R}^2$ for the double integral vs a surface in $\mathbb{R}^3$ for the surface integral. Thus:

When moving from one column to the next, the dimension of the region of integration involved increases by 1: from curves, to surfaces, to 3-dimensional solids. In other words, the horizontal arrows in the diagram indicate moving to a larger-dimensional region of integration.

And that's the point: these are all essentially the same type of integral, only taking place over regions of varying dimension which sit inside of ambient spaces of varying dimension. It should be clear that if we wanted to we could continue this process indefinitely, filling in the questions marks with other types of integrals as well. For instance, the third entry in the second row would involve integrating a four-variable function over a 3-dimensional solid inside $\mathbb{R}^4$, whereas the missing terms in the third row would involve integrating a four-variable function over a 2-dimensional surface inside $\mathbb{R}^4$ and integrating a five-variable function over a 3-dimensional solid inside $\mathbb{R}^5$ respectively. (Note that in all integrals in the first row, the region of integration and ambient space have the same dimensions, whereas in the second the difference in these dimensions is 1 and in the final row it is 2.) How would you compute all of these fancy types of integrals? By using parametric equations of course!

One final thing to note: the integrals in the first row are taken with respect to dx, $dx\,dy$, and $dx\,dy\,dz$ respectively, whereas once we hit line and surface integrals we have to throw in some "Jacobian factors", either $\|\mathbf{x}'(t)\|$ in the line integral case or $\|\mathbf{X}_s \times \mathbf{X}_t\|$ in the surface integral case. This is due to the fact that all regions of integration in the first row are "flat" (to be precise, the dimensions of the region and of the ambient space are the same), whereas everywhere else this is no longer necessarily the case. Moving from flat to non-flat things requires including some Jacobian term which tells us how lengths, areas, volumes, etc. are affected by such "changes of ambient dimensions".

All Integration Theorems Are The Same

Recall the main integration theorems we've seen when looking at vector calculus: the Fundamental Theorem of Line Integrals, Stokes' Theorem, and Gauss's Theorem. Now, we also saw Green's Theorem, but of course this is just a special case of Stokes' Theorem (when we have a flat surface in the xy-plane), so Green's Theorem is also covered by the three main theorems above. Also, note that the usual single-variable Fundamental Theorem of Calculus is itself a special case of the

Fundamental Theorem of Line Integrals (applied to the conservative field $\mathbf{F} = f'(x)\mathbf{i}$ with potential $f(x)$, when integrating over a curve which is an interval $[a, b]$ on the x-axis), so even this theorem is covered by the theorems above.

The formulas for Stokes' Theorem and Gauss's Theorem explicitly involve integrating over a boundary on one side, and indeed so does the Fundamental Theorem of Line Integrals if we consider

$$f(\text{end point}) - f(\text{start point})$$

to be an "integral" of f over the "space" consisting of only the two points "end point of C" and "start point of C". So, viewing these two points as forming the boundary ∂C of C, we will use $\int_{\partial C} f$ to mean $f(\text{end point}) - f(\text{start point})$. (This makes sense, since after all integrating f over ∂C amounts to adding up the values of f among all points of ∂C, and in this case there are only two such points. The fact that the second value is subtracted instead of added has to do with having an "orientation" on ∂C, the details of which here we'll just ignore.)

With the above notation in mind, here then are the three main theorems:

$$\underbrace{\int_{\partial C} f = \int_C \nabla f \cdot d\mathbf{s}}_{\text{Fundamental Theorem}} \qquad \underbrace{\int_{\partial S} \mathbf{F} \cdot d\mathbf{s} = \iint_S \operatorname{curl} \mathbf{F} \cdot d\mathbf{S}}_{\text{Stokes' Theorem}} \qquad \underbrace{\iint_{\partial E} \mathbf{F} \cdot d\mathbf{S} = \iiint_E \operatorname{div} \mathbf{F}\, dV}_{\text{Gauss's Theorem}}.$$

Notice that all three theorems thus involving integrating over a boundary on one side and on the other side integrating something made out of derivatives, whether it be a gradient, a curl, or a divergence. Thus all theorems have the form:

$$\int_{\text{boundary (object)}} \text{something} = \int_{\text{object}} \text{derivative (something)}.$$

Derivatives measure how something "changes", so the point is:

> All integration theorems express the idea that integrating something over the boundary of a geometric object is the same as integrating a "derivative" of that something over the full object. That is: how something behaves along a boundary is directly related to how it "changes" away from that boundary.

The only difference in the three theorems comes from the types of derivatives considered and the types of regions we integrate over: the Fundamental Theorem integrates a gradient over a curve, Stokes' Theorem a curl over a surface, and Gauss's Theorem a divergence over a 3-dimensional solid.

Consider now the following diagram:

$$\begin{tikzcd}
\text{function} \arrow[r, "\text{grad}"] \arrow[d] \arrow[dr, phantom, "\text{Fundamental}"] & \text{vector field} \arrow[r, "\text{curl}"] \arrow[d] \arrow[dr, phantom, "\text{Stokes'}"] & \text{vector field} \arrow[r, "\text{div}"] \arrow[d] \arrow[dr, phantom, "\text{Gauss's}"] & \text{function} \arrow[d] \\
\text{points} & \text{curve} \arrow[l, "\text{boundary}"'] & \text{surface} \arrow[l, "\text{boundary}"'] & \text{solid} \arrow[l, "\text{boundary}"']
\end{tikzcd}$$

The first row contains the types of things we integrate and each arrow indicates a type of derivative we can take to get from one thing to another. The second row contains the types of geometric objects we integrate over, listed with increasing dimension; the operation of taking a boundary moves us from one object to another of dimension one less. Finally, the vertical arrows indicate a type of integral—namely, the one obtained by integrating the thing in the first row over the corresponding region in the second row. To be precise, from left to right we have the "integral" which evaluates a function at two points and subtracts the values, then a line integral, then a surface integral, and finally a triple integral.

The upshot is that each square in the diagram essentially characterizes one of our main theorems. The first square gives the statement of the Fundamental Theorem: start with a function f in the upper left, apply the horizontal arrow to get ∇f, then apply the vertical arrow to get $\int_C \nabla\mathbf{F} \cdot d\mathbf{s}$ where C is some curve, and you get the same thing as taking f and applying the first vertical arrow to get $f(\text{end point}) - f(\text{start point})$, which is the "integral" of f over the object obtained by applying the horizontal "boundary" arrow to the curve C. The second square gives Stokes' Theorem: start with a vector field $\mathbf{F}$ in the second entry of the first row, apply the horizontal arrow to get $\operatorname{curl}\mathbf{F}$, then apply the vertical arrow to get $\iint_S \operatorname{curl}\mathbf{F} \cdot d\mathbf{S}$ for some surface S, and you get the same thing as taking $\mathbf{F}$ and applying the second vertical arrow to get $\int_{\partial S} \mathbf{F} \cdot d\mathbf{s}$ where ∂S is the object obtained by applying the second horizontal arrow in the second row to S. Finally, the last square is Gauss's Theorem: take $\mathbf{F}$ in the third entry of the first arrow, apply the horizontal arrow to get $\operatorname{div}\mathbf{F}$, then apply the vertical arrow to get $\iiint_E \operatorname{div}\mathbf{F}\, dV$ for some solid E, and you get the same thing as taking $\mathbf{F}$ and applying the third vertical arrow to get $\iint_{\partial E} \mathbf{F} \cdot d\mathbf{S}$ where ∂E is the object obtained by applying the final horizontal arrow in the second row to E.

> Thus, all of these theorems are the same, with the only difference being the parts of the diagram above they relate to one another.

One last thing to note: if you follow two consecutive horizontal arrows in the first row of the diagram you always get zero! Indeed, this just expresses the facts that

$$\operatorname{curl}(\nabla f) = \mathbf{0} \text{ and } \operatorname{div}(\operatorname{curl}\mathbf{F}) = 0.$$

Now, notice that if you follow two consequence horizontal arrows in the second row a similar thing happens: starting with a solid, taking its boundary gives a closed surface, and taking the boundary of that gives nothing since closed surfaces have no boundary; and starting with a surface, taking its boundary gives a closed curve, and taking the boundary of that gives nothing since closed curves have no boundary. That is,

$$(\text{boundary}) \circ (\text{boundary}) = \text{nothing},$$

where $\circ$ denotes a composition of operations, whereas in the first row we have

$$(\text{derivative}) \circ (\text{derivative}) = \text{zero}$$

when we interpret "derivative" correctly. In some sense, the operation of taking a boundary is some type of geometric "analog" of the operation of taking a derivative, but explaining this further would require a whole different course, so we'll just have to leave it at that.