Averages. The word “average” probably suggests to you the process of adding a list of numbers and dividing by the number of terms in the sum. In particular, the average of two numbers, a and b is

$$\frac{a+b}{2}.$$

This seems so familiar that you may be surprised to see a whole section in the calculus book devoted to averages. In fact, averages provide a very useful viewpoint on the integral calculus, both in theory and in practice.

Weighted averages. Suppose we seek to average a list of numbers, and we notice that every element is either a or b, but there are many of each. To continue, suppose we count the numbers of each type and find that m of then are a and n of them are b. Then the sum of the list will be $ma + nb$ and the number of elements in the list is $m + n$. We can easily find the average from this, but I want to call attention to an equivalent form. The average is given by

$$\frac{ma+nb}{m+n} = \frac{m}{m+n}a + \frac{n}{m+n}b.$$

This shows that the average is of the form $\lambda a + \mu b$ with

$$\lambda \geq 0$$

$$\mu \geq 0$$

$$\lambda + \mu = 1.$$

Such an expression is called a **weighted average**, even if λ and μ are not rational.

Why average? It follows from the properties of λ and μ in our expression for an average that all averages of a and b are between a and b. In this sense, the average value of a list is *typical* of the list. Indeed, it would be hard to imagine any other **single number** that does as good a job of describing the values in the list.

Averaging many values. The same results hold if our list contains many different values. The average value is always between the smallest and largest values in the list.

Connections with Riemann sums. Suppose that the list of values is generated by taking the value of a function at certain points. For example, suppose you record the temperature at a particular location every hour on the hour. Adding this list of numbers and dividing by the number of observations is a reasonable interpretation of the “average temperature at that location” on the day that you perform this experiment.

The underlying idea behind this should seem familiar. The domain of possible times is divided into a large number of short intervals. For convenience, we make those intervals of equal length. Then, for each of those intervals, we choose a time t inside that interval (in a systematic way) at which we evaluate the temperature function. This sounds very much like a Riemann sum. Since the time intervals Δt are all of equal length, each has length equal to the total length divided by the number of measurements. A slight rearrangement of these factors shows that the Riemann sum is the length of the interval multiplied by the average value of our observations.

The meaning of convergence. The statement that finer and finer Riemann sums converge to a definite value is equivalent, after factoring out the factor giving the length of the domain, to the statement that all averages of a large number of values of the function at equally spaced points are close to one another. Since these values are averages, they are always trapped between the smallest and largest values of the function.

This limiting average will be taken as the definition of the “average value of the function between a and b”. That is,

$$f_{\text{ave}} = \frac{1}{b-a}\int_a^b f(x)\,dx.$$

Another way to describe f_{ave} is to say that it is the **constant** that has the same integral as the function f. If $f(x) \geq 0$ for all $a \leq x \leq b$, so that we can use the interpretation of the integral as area, then the average is the height of a rectangle on the base $[a, b]$ that has area equal to the area under the graph of $y = f(x)$ on $[a, b]$.

Since average values are related to values of the function being averaged, it is easy to notice some errors in the computation of an average. In particular, using the result of an integration to find an average is a way to see if the answer “looks right”. If the function f being averaged is continuous, there will be a value of $x \in [a, b]$ such that $f(x)$ is any particular values in interval between the extreme values of f on $[a, b]$. Since the average is such a value, there will always be someplace in $[a, b]$ such that the average is found by evaluating f there. This doesn’t add much to our understanding of averages, but it allows *averages of values* of f to be thought of as *average values* of f.

Numerical integration. Riemann sums turn around the idea of averages being typical values of the function, by using a value of the function to represent its average on a small interval. Although Riemann sums are general enough for all theoretical purposes, the midpoint rule is the only simple formula using a single point x_i^* in the i^{th} interval that is close to being useful. The use of a finite weighted average to approximate the average value of the function opens new possibilities. One example is the *trapezoidal rule* in which the contribution of the i^{th} is

$$\frac{f(x_{i-1}) + f(x_i)}{2}(x_i - x_{i-1}).$$

Simpson’s rule isn’t much more complicated, but it turns out to be much better for many functions. The average of function values used is

$$\frac{f(x_{i-1}) + 4f\left(\frac{x_{i-1}+x_i}{2}\right) + f(x_i)}{6}.$$

The statements of these rules that one usually sees combines the contribution that $f(x_i)$ makes as the right endpoint of the $(i-1)^{\text{st}}$ interval and the left endpoint of the i^{th} interval. This speeds up computation at the expense of a slight complication in the formula. We will return to this topic in Section 7.8.

Examples. Exercises 7 through 10 of Section 6.5 have the general instructions to: (a) find the average; (b) find a c such that $f(c)$ is the average; (c) sketch to show equality of area. Part (b) is somewhat of a distraction from what is important here. While the average itself is useful, we only need to know that it is attained *somewhere*. Solving $f(x) = c$ may be a lot of work for very little benefit; we do it only because our calculators make it available at the push of a button. Two of these exercises are

$$f(x) = 4 - x^2 \text{ on } [0, 2] \qquad (\#7)$$

$$f(x) = x^3 - x + 1 \text{ on } [0, 2] \qquad (\#9)$$

Integration by parts. Suppose, as in Exercise 1 of Section 7.1 that you want to find

$$\int xe^{2x}\,dx.$$

This can't be too hard. All we need to do is to find something whose derivative is xe^{2x}. We probably got that answer to some differentiation problem — if only we could remember. One approach is to this would be to differentiate some reasonable guesses at the answer and use what we learn to come up with better guesses. This is actually more efficient than it sounds because of the **linearity of the integral**. For example, if you

get within a *constant factor* of the correct answer, you just multiply F by a constant to multiply F' by the same constant to make it equal to the given f. This can be extended to the case in which the derivatives of F and G share a common term; then you add a multiple of G to F to get a function whose derivative does not contain that term.

There is another reason for approaching integration problems this way. Although the theory assures us that the integral must exist, it does show by proving that we can approximate the value of any definite integral as well as we like. There is no guarantee that there will be a nice formula for an indefinite integral. Indeed, one way of convincing ourselves that exponentials, logarithms, trig functions and their inverses really exist is to define the logarithm and arctangent as integrals of rational functions. Even at the level of familiar functions, integrals are often much more *interesting* than the integrands. By contrast, whenever we accept a function into our family of everyday functions, we calculate its derivative. The first time we meet the derivative, we may need limits to find the derivative, but all later work will use the general methods of calculus to differentiate other expressions in which our new friend appears. In short, formal differentiation is a mechanical process, i.e., a *Calculus*. Although we talk about an Integral Calculus, there is much more of an art to finding symbolic indefinite integrals.

To get back to Exercise 1, I will write the derivatives suggested by the problem. The table below will just contain pairs F, F', but I will try to explain the motivation for selecting each F

F	F'
e^{2x}	$2e^{2x}$
xe^{2x}	$2xe^{2x} + e^{2x}$
$2xe^{2x} - e^{2x}$	$4xe^{2x}$
$\frac{1}{4}\left(2xe^{2x} - e^{2x}\right)$	xe^{2x}

The last line is the answer.

The second line of the table uses the product rule. It plays an important role in finding the answer because

the product rule gives us two terms: one is (a constant multiple of) the one we want; the other is simpler. In this example, we began with something that allowed us to integrate the simpler term, because it was a formula that was needed to determine the differentiation formula on the second line. If we could count on the product rule to always give us one term that is easier than the other, then its use in finding integrals could be described by rearranging the result of the product rule as

$$\int u\,dv = uv - \int v\,du. \tag{P}$$

Formula (P) is called "integration by parts". The idea is that you break the given integral up as a product of one part u which you expect to be simplified by differentiation, and a second factor dv (which includes the dx) which is not made worse by integrating. As applied to Exercise 1, $u = x$ and $dv = e^{2x}\,dx$. Differentiating the first gives $du = dx$, and integrating the second gives $v = \left(\frac{1}{2}\right)e^{2x}$. Formula (P) then says

$$\int xe^{2x}\,dx = x\left(\tfrac{1}{2}\right)e^{2x} - \int \left(\tfrac{1}{2}\right)e^{2x}\,dx$$

The integral at the end of the line is easily found, leading to the same formula obtained in the last line of the table.

Unfortunately, there are some integrals that are found by using the product rule for differentiation in combination with algebraic manipulation that rewrites the derivative. Although this algebra is natural in the context of finding the derivative, it looks artificial when it must be introduced into (P). Examples will be given later.

Easy cases. First let us look at some examples where (P) does everything that is claimed for it. The main example of functions that are simplified by differentiation are polynomials. Differentiation always lowers the degree, so if you can differentiate enough times, you can make the polynomial go away. Linearity allows us to break the polynomial up as a combination of powers, so most of the examples of this type that we consider will have $u = x^n$.

The main types of functions that are not made worse by integrating are exponential or trigonometric functions. We have seen on example in Exercise 1. Now

look at Exercise 3.

$$\int x \cdot \sin 4x \, dx \,,$$

where I have separated $u = x$ from $dv = \sin 4x \, dx$ by a dot. Differentiating the first gives $du = dx$, and integrating the second gives $v = -(1/4) \cos 4x$. Note that we are content to find *one* possible v. The $+\,C$ that appears in the general integration formula would complicate all the intermediate steps without doing anything for the final answer. It still remains true that indefinite integration require finding *one* expression with the correct derivative and adding a constant to it to find all function with that derivative. Formula (P) says

$$\int x \cdot \sin 4x \, dx = x \cdot \big(-(1/4) \cos x \big) \\ - \int \big(-(1/4) \cos 4x \big) \cdot dx \,.$$

Thus

$$\int x \cdot \sin 4x \, dx = \frac{-4x \cos 4x + \sin 4x}{16} + C.$$

We can now do exercises like

$$\int x^2 \cos 3x \, dx \tag{\#5}$$

$$\int_0^1 te^{-t} \, dt \tag{\#17}$$

$$\int (2x+3)e^x \, dx \tag{\#25}$$

Definite integrals. When computing a definite integral, it is not necessary to complete the calculation of the indefinite integral before substituting the limits of integration for x. The fundamental theorem of calculus tells us that the expression $\int v \, du$ on the right side of (P) can be taken to be the definition of a function. The difference of this expression, evaluated at $x = a$ and $x = b$, is exactly the definite integral between the same limits as in the given $\int u \, dv$. The contribution of the evaluated part is simply the difference of this function at the given limits of integration.

Slightly harder examples. If you can take $u = \ln x$ or $u = \arctan x$, then differentiation will make du

so much simpler that dv can be allowed to become *a little* more complicated when it is integrated. The basic examples are those with $dv = dx$. (These are examples 2 and 5 in the text.) Once you have done these, you can extend to the case with $dv = x^n \, dx$

Back to the tables. Now for the examples where the **only** clear method is to simplify the result of the product rule before applying it to the integration problem. Consider the differentiation formulas

F	F'
$e^x \cos x$	$e^x \cos x - e^x \sin x$
$e^x \sin x$	$e^x \cos x + e^x \sin x$
$e^x \cos x + e^x \sin x$	$2e^x \cos x$
$e^x \sin x - e^x \cos x$	$2e^x \sin x$

These formulas show how to compute $\int e^x \sin x \, dx$ (compare to Example 4). I find this approach much clearer than the usual approach that insists on using formula (P).

Also consider

F	F'
$\cos x \sin^n x$	$\cos x \cdot n \sin^{n-1} x \cdot \cos x +$
	$+ (-\sin x) \sin^n x$
	$n \cos^2 x \sin^{n-1} x - \sin^{n+1} x$
	$n(1 - \sin^2 x) \sin^{n-1} x -$
	$- \sin^{n+1} x$
	$n \sin^{n-1} x - (n+1) \sin^{n+1} x$

Rearranging the terms on the last line gives

$$\begin{aligned} \int \sin^{n+1} x \, dx = &-\frac{1}{n+1} \cos x \sin^n x \\ &+ \frac{n}{n+1} \int \sin^{n-1} x \, dx, \end{aligned}$$

which is similar to the formula found in Example 6.

An interesting special case involves the definite integral from 0 to π. Here the integrated part (i.e., what would be the uv part if we were using formula (P)) (when $n > 0$) is a function that is zero at both endpoints. Thus, this term contributes nothing to the definite integral and one gets only a multiple (depending on n) of a similar integral in which the exponent has been decreased by 2. The special cases $n = 0$ and $n = -1$ allow all powers of $\sin x$ to be integrated over this interval.