Integration by parts

The usual rule of integration by parts taught in High School say that

$$\int_a^b F\,G'\,dt = F\,G\Big|_a^b - \int_a^b G\,F'\,dt\,,$$

for functions $F, G : [a, b] \to \mathbb{R}$ which admit derivative.

This is an straight-forward consequence of the rule of derivative of a product: $(F\,G)' = F'\,G + F\,G'$, and the Barrow rule:

$$F\,G\Big|_a^b = \int_a^b (F\,G)'\,dt = \int_a^b (F'\,G + F\,G')\,dt\,.$$

What is curious is that the integration by parts does not need derivability. Moreover, the proof is simple and accesible to undergraduates.

Suppose that we have

$$\begin{cases} F(x) = c_1 + \int_a^x f(t)\,dt \\ G(x) = c_2 + \int_a^x g(t)\,dt \end{cases}$$

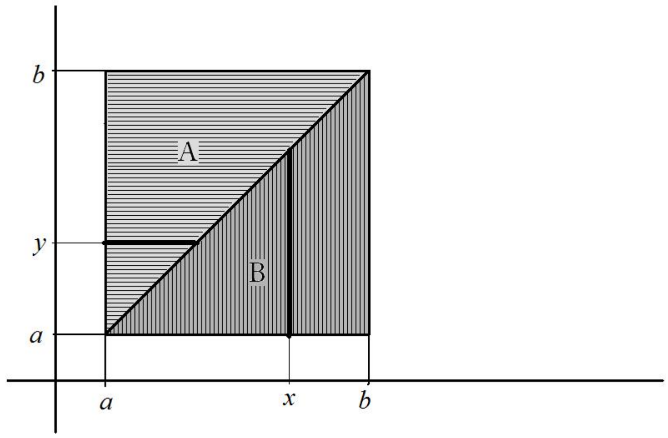
where $f, g : [a, b] \to \mathbb{R}$ are integrable functions and c_1, c_2 are constants. When f, g are continuous, then F, G have derivatives $F' = f$, $G' = g$. However, it is usual that F, G of the above form are not derivable.

Let us see that the integration by parts formula holds in this case, in two steps.

First, assume that $c_1 = c_2 = 0$. Then

$$\begin{aligned}
\int_a^b (F\,g + G\,f)\,dt &= \int_a^b g(t)\left(\int_a^t f(s)\,ds\right)dt + \int_a^b f(t)\left(\int_a^t g(s)\,ds\right)dt \\
&= \int_a^b\int_a^y f(x)g(y)\,dx\,dy + \int_a^b\int_a^x f(x)g(y)\,dy\,dx \\
&= \int_a^b\int_a^b f(x)g(y)\,dx\,dy \\
&= \left(\int_a^b f(x)\,dx\right)\left(\int_a^b g(y)\,dy\right) \\
&= F(b)\,G(b) = F\,G\Big|_a^b .
\end{aligned}$$

The first equality is by definition; in the second equality we have set $x = s, y = t$ in the first summand, and $x = t, y = s$ in the second summand; in the third equality we have added up the integration regions in the following figure; the other equalities are clear.



Now, let c_1, c_2 be any real numbers. Then let $F = \tilde{F} + c_1$, $G = \tilde{G} + c_2$. With the previous case at hand, we have:

$$\begin{aligned}
\int_a^b (F\,g + G\,f)\,dt &= \int_a^b ((\tilde{F} + c_1)\,g + (\tilde{G} + c_2)\,f))\,dt \\
&= \int_a^b (\tilde{F}\,g + \tilde{G}\,f)\,dt + c_1 \int_a^b g + c_2 \int_a^b f \\
&= \tilde{F}\,\tilde{G}\Big|_a^b + c_1(G(b) - G(a)) + c_2(F(b) - F(a)) \\
&= (F(b) - F(a))(G(b) - G(a)) + F(a)(G(b) - G(a)) + G(a)(F(b) - F(a)) \\
&= F(b)G(b) - F(a)G(a) = F\,G\Big|_a^b .
\end{aligned}$$

The fourth equality uses $\tilde{F}(b) = F(b) - c_1 = F(b) - F(a)$, $\tilde{G}(b) = G(b) - c_2 = G(b) - G(a)$, and $c_1 = F(a)$, $c_2 = G(a)$.

Notes for helping the referee.

1. The interest of the result lies in the fact that it can be used for many examples of functions f, g which are not continuous. For instance, let f, g be piecewise continuous, with finite lateral limits at the discontinuity points. Then F, G are continuous and piecewise C^1, and the result is valid in this case.

Let us do an alternative proof for this specific case: assume that $a = t_0 < t_1 < \ldots < t_{k-1} < b = t_k$, and f, g are continuous at each interval $[t_{i-1}, t_i]$. Then

$$F\,G|_a^b = \sum_{i=1}^k F\,G|_{t_{i-1}}^{t_i} = \sum_{i=1}^k \int_{t_{i-1}}^{t_i} (f\,G + F\,g) = \int_a^b (f\,G + F\,g)\,.$$

(Our previous proof however is more direct; in this one, one has to consider first the case C^1 and then the case piecewise C^1.) Here F, G are not derivable at the points t_i, though they are derivable everywhere else.

There are functions f with a numerable set of un-removable discontinuity points, even a dense set. An easy example is:

$$f : [0,1] \to \mathbb{R},\ f(0) = 0, f(x) = 1/n, x \in [\frac{1}{n+1}, \frac{1}{n}], n \geq 1\,.$$

Then F is not derivable at the points $x = 1/n$, $n \geq 2$.

A more sophisticated example is as follows: enumerate the rationals in $[0,1]$ as $\{x_n\}$; consider $\epsilon_n = \frac{1}{2^n}$. Define

$$f(x) = \sum_{k\,|\,x_k < x} \epsilon_k\,.$$

This function is discontinuous at $(0,1) \cap \mathbb{Q}$. The function F is not derivable at $(0,1) \cap \mathbb{Q}$, a dense set.

2. The hypothesis of our result covers functions which are Lebesgue integrable in $[a,b]$. In particular, it covers Riemann integrable functions. (A function is Riemann integrable if and only if it is bounded and the set of discontinuity points is of zero measure [T. Apostol, (1974), Mathematical Analysis, Addison-Wesley, pp. 169–172].)

The integrability hypothesis is used for the double integral: if $f(x)$, $g(y)$ are Lebesgue integrable, then $f(x)g(y)$ is Lebesgue integrable in $[a,b] \times [a,b]$, and there is a Fubini rule that says that the double integral can be done first in the variable x, then in variable y (see G. Fubini, Sugli integrali multipli, Opere scelte, 2, Cremonese (1958) pp. 243–249; or http://eom.springer.de/F/f041870.htm)

3. The result is probably known to specialists, understood from other point of view. But I am not aware of a reference of it.

If F, G are continuous functions (they are always the case here), they admit derivative **as distributions** (see http://en.wikipedia.org/wiki/Distribution_(28mathematics) for distributions). This means that the derivatives $D(F), D(G)$ are distributions satisfying

$$\int_a^b D(F)h = F\,h|_a^b - \int_a^b F(t)h'(t)dt \tag{1}$$

for any C^∞ function $h : [a,b] \to \mathbb{R}$. In our situation, it is clear that $D(F) = f$ almost everywhere (this follows from the integration by parts rule that we have proved, applied to f, h', and from the defining equation (1)).

Note that it is possible that F is not derivable (as function) at many points (by item **1.**).

Our result should follow from the theory of distributions. First there is a Barrow rule

$$\int_a^b D(F) = F|_a^b,$$

which is (1) with $h = 1$. If there is a Leibnitz rule for derivatives of distributions

$$D(F\,G) = D(F)\,G + F\,D(G)\,, \tag{2}$$

then our result follows by integrating $\int_a^b$ (2).

However, I do not know of a proof of (2). Note that in general, the issue of multiplying distributions is a delicate one.

I can provide a proof of (2) **using the results of this note.** It goes as follows: First, consider $h : [a, b] \to \mathbb{R}$, a C^∞ function. Then $D(Gh) = gh + Gh'$. This is proved by taking any $k : [a, b] \to \mathbb{R}$, any C^∞ function,

$$\begin{aligned} \int_a^b D(Gh)k &= Ghk|_a^b - \int_a^b Ghk' = \int_a^b (g\,hk + G(hk)') - \int_a^b Ghk' \\ &= \int_a^b (ghk + Gh'k + Ghk' - Ghk') = \int_a^b (gh + Gh')k\,, \end{aligned}$$

using our result, in the second equality, for g and $(hk)'$ (with "integrals" G and hk).

Now use again our result for f and $gh + Gh'$ (with "integrals" F and Gh)

$$\int_a^b fGh = FGh|_a^b - \int_a^b F(gh + Gh')$$

which is rewritten as

$$\int_a^b (fG + Fg)h = FGh|_a^b - \int_a^b FGh'\,.$$

This means that $D(FG) = fG + Fg$.

Note that such Leibnitz rule holds for our class of functions F, G, where f, g are integrable.

4. The submitted proof is of interest because it is very simple, and it does not mention distributions. It can be taught in an undergraduate course when Calculus of several variables is available, but not the theory of distributions. I used it myself to solve elementary differential equations in a Calculus course, at a point where I needed integration by parts for non-smooth functions (e.g. propagation of singularities in the wave equation).

It is also interesting to see that a result like the integration by parts can be proved **avoiding** the use of derivatives.