Differentiation Under the Integral Sign with Weak Derivatives

Steve Cheng

December 2, 2006

Contents

Copyright matters

Preface

In short, in this article we discuss the problem of differentiating under the integral sign.

We assume a passing familiarity of the theory of distributions by Laurent Schwartz, and no more. (Even the author is just a beginner on this theory.) However, even readers who do not care much about the theory may find the section with worked-out computational examples to be useful.

The theory expoused in this article is nothing new: analysts in the nineteenth and twentieth centuries no doubt recognized the issues that crop up with exchanging partial derivatives and integration, and had fixes for them — though perhaps by ad-hoc methods before Schwartz's theory of distributions.

But I hope this article would serve a purpose in explaining the application of the concepts at a more elementary level than the standard works on distribution theory, while not losing rigor for the mathematically-minded audience. Certainly no calculus book nowadays, even non-rigorous ones, dare talk about delta functions, even though they are valuable calculation tools for the engineer and physicist.

I would like to thank Kamyar Hazaveh, who is an engineer, working on the same problem set as I one day; I would not have written this article if not for him posing the question on differentiation under the integral; I would also like to thank Matt Towers ("silverfish") and Raymond Puzio at PlanetMath for their encouraging comments in my investigation of the present problem. (Which only confirms how little people seem to know about differentiation under the integral sign!) And I should also mention that problem set originated from Prof. Sebastian Jaimungal, a former physicist, who also motivated my investigation with his helpful comments on what to do when differentiating integrands that are not smooth.

1 Introduction

Let X be an open subset of $\mathbb{R}^m$, and Ω be a measure space. Given $f\colon X \times \Omega \to \mathbb{R}$, and consider the integral with parameter

$$g(x) = \int_\Omega f(x, \omega)\, d\omega$$

(assuming it is well-defined).

We want to differentiate g, and we hope that

$$\frac{\partial g(x)}{\partial x_i} = \int_\Omega \frac{\partial f(x, \omega)}{\partial x_i}\, d\omega\,,$$

as often asserted in non-rigorous expositions of calculus without qualification.

If f is sufficiently nice — for example, it satisfies the conditions of Theorem A.1 — then the swap of the integration and differentiation can be proven to be valid.

However, when we do enough calculations switching integration and differentiation, we soon begin to realize that the conditions allowing the interchange are probably a lot more general than what the usual mathematical theorems would tell us. For example, in applications, $\partial f(x, \omega)/\partial x_i$ may exist except for a limited number of singularities; if we go ahead and differentiate under the integral sign anyway, experimenting and fudging a little with the Dirac delta function and the like, we seem to always obtain the correct result.

As the involvement of the Dirac delta function suggests, differentiation under the integral sign can be more generally formulated as a problem with *generalized functions* (the *distributions* of Laurent Schwartz) and their integrals and derivatives. In this note, we justify, using generalized derivatives, differentiation under the integral sign, in the cases when f does not satisfy the prerequisite basic conditions, or even when $\partial f(x, \omega)/\partial x_i$ fails to exist for certain values of x and ω.

Although the theory concerns the weak derivatives of generalized functions, in practical calculations, the final results obtained will often be seen to have strong (usual/classical) derivatives that agree with their weak derivatives except at certain singularities. Thus weak derivatives will intervene as a calculation tool, much as complex analysis intermediates results about real functions, and fundamental solutions/Green's functions intermediate results about ordinary solutions of partial differential equations.

2 Theory

For an ordinary function $f \colon X \times \Omega \to \mathbb{R}$, the meaning of the expression $\int_\Omega f(x,\omega)\,d\omega$ is immediate. But to obtain the utmost generality, and to simplify the proofs of the fundamental result, we need to extend the notion of integrating a function with respect to one variable while holding the other fixed, to arbitrary generalized functions f.

This leads us naturally to the following definition.

Definition 2.1. Let $f(x, \omega)$ be a generalized function of $x \in X$, for each fixed $\omega \in \Omega$. Let ϕ denote any function from some suitable class of test functions on X defining the space of generalized functions.

Suppose that the Lebesgue integral

$$\int_\Omega \left(\int_X f(x,\omega)\,\phi(x)\,dx \right) d\omega := \int_\Omega \langle f(\cdot,\omega), \phi \rangle \,d\omega \tag{1}$$

exists as a finite quantity (converges absolutely) for all test functions ϕ. Then we define the generalized function

$$g(x) = \int_\Omega f(x,\omega)\,d\omega$$

by

$$\int_X \left(\int_\Omega f(x,\omega)\,d\omega \right) \phi(x)\,dx := \langle g, \phi \rangle := \int_\Omega \left(\int_X f(x,\omega)\,\phi(x)\,dx \right) d\omega\,. \tag{2}$$

It is critical that, without additional assumptions on f, the linear functional defined by equation (2) may not be a *continuous* linear functional, and hence will not be a genuine *generalized function*.

To fix this problem and ensure that the resulting linear functional (2) will be continuous, we can impose this straightforward hypothesis:

Criterion 2.2 (Continuity)**.** For every convergent[1] sequence $\{\phi_n\}$ of test functions, we stipulate that

$$\int_\Omega \sup_n \Bigl| \langle f(\cdot,\omega), \phi_n \rangle \Bigr| \,d\omega < \infty\,. \tag{3}$$

[1] The convergence is in the topology of the space of test functions for the generalized functions.

Proof that the criterion suffices. For any sequence of test functions $\{\phi_n\}$ converging to ϕ, we have $\langle f(\cdot,\omega),\phi_n\rangle$ converging to $\langle f(\cdot,\omega),\phi\rangle$ as $n\to\infty$, for each $\omega\in\Omega$. Then the bound (3) gives the dominating factor for the Lebesgue Dominated Convergence Theorem that allows us to deduce $\langle g,\phi_n\rangle\to\langle g,\phi\rangle$. ■

Proposition 2.3. *Let $f(x,\omega)$ be a generalized function of $x\in X$, for each $\omega\in\Omega$. If*

$$\int_\Omega f(x,\omega)\,d\omega$$

exists, then so does

$$\int_\Omega \frac{\partial}{\partial x_i} f(x,\omega)\,d\omega\,.$$

Proof. First, we have to show that hypothesis (1) holds for $\partial f/\partial x_i$: that

$$\int_\Omega \langle \frac{\partial f}{\partial x_i},\phi\rangle\,d\omega$$

exists for every test function ϕ. But the integrand is, by the definition of the derivative of a generalized function, equal to $-\langle f(\cdot,\omega),\partial\phi/\partial x_i\rangle$. And the integral over Ω of this quantity exists because $\int_\Omega f(\cdot,\omega)\,d\omega$ exists and $\partial\phi/\partial x_i$ is also a test function.

Criterion 2.2 is verified likewise:

$$\int_\Omega \sup_n \Bigl|\langle\frac{\partial f}{\partial x_i},\phi_n\rangle\Bigr|\,d\omega=\int_\Omega \sup_n\Bigl|-\langle f(\cdot,\omega),\frac{\partial\phi_n}{\partial x_i}\rangle\Bigr|\,d\omega<\infty\,,$$

noting that $\partial\phi_n/\partial x_i\to\partial\phi/\partial x_i$ whenever $\phi_n\to\phi$. ■

Theorem 2.4 (Differentiation under the integral sign)**.** *Let $f(x,\omega)$ be a generalized function of x, for each $\omega\in\Omega$, such that $\int_\Omega f(x,\omega)\,d\omega$ exists. Then*

$$\frac{\partial}{\partial x_i}\int_\Omega f(x,\omega)\,d\omega=\int_\Omega\frac{\partial}{\partial x_i}f(x,\omega)\,d\omega\,.$$

Proof. The integral on the right-hand side of the equation above makes sense in light of Proposition 2.3.

Set $g(x)=\int_\Omega f(x,\omega)\,d\omega$. For every test function ϕ,

$$\langle\frac{\partial g}{\partial x_i},\phi\rangle=-\langle g,\frac{\partial\phi}{\partial x_i}\rangle=-\int_\Omega\langle f(\cdot,\omega),\frac{\partial\phi}{\partial x_i}\rangle\,d\omega=\int_\Omega\langle\frac{\partial f}{\partial x_i},\phi\rangle\,d\omega:=\langle\int_\Omega\frac{\partial f}{\partial x_i}\,d\omega,\phi\rangle\,.\quad ■$$

Needless to say, higher derivatives can also be taken under the integral sign, by applying the theorem multiple times.

Or more directly, and more interestingly, the operator $\partial/\partial x_i$ in the theorem and its proof can in fact be replaced by *any* continuous linear operator T on generalized functions with a dual T':

$$\langle Tg,\phi\rangle=\langle g,T'\phi\rangle=\int_\Omega\langle f(\cdot,\omega),T'\phi\rangle\,d\omega=\int_\Omega\langle Tf,\phi\rangle\,d\omega:=\langle\int_\Omega Tf\,d\omega,\phi\rangle\,.$$

The dual T' of T will be guaranteed to exist if the underlying space of test functions is reflexive, such as that of the compactly-supported smooth functions, $\mathcal{D}=\mathbf{C}_c^\infty(X)$.

Examples of such linear operators T include the higher-order partial differential operators, as well as the Fourier transform.

3 To which functions does the theory apply to

Having obtained some deceptively easy results, now would be a good time to consider for what kind of generalized functions f can the integral $\int_\Omega f(x,\omega)\,d\omega$ be defined.

3.1 Integrable functions

Assumption (1) and Criterion 2.2 in the definition of the integral of a generalized function f may seem to be quite restrictive, but they *are* automatically satisfied by any ordinary function f integrable on $X \times \Omega$, with the usual Lebesgue measure on $X \subseteq \mathbb{R}^m$. The reason is nothing more than that a convergent sequence of test functions must be uniformly bounded on all of X.

If the space of test functions is $\mathcal{D} = \mathbf{C}_c^\infty(X)$ (the infinitely-differentiable functions with compact support), it even suffices that f is only locally integrable in X: that is,

Criterion 3.1 (Local integrability)**.**

$$\int_K \int_\Omega |f(x,\omega)|\,d\omega\,dx < \infty\,, \quad \text{for any compact set } K \subset X.$$

An important case of this situation is when $x \mapsto \int_\Omega |f(x,\omega)|\,d\omega$ is continuous: it is then, of course, locally integrable.

In many applications, these simple results suffice, for most of us are not integrating and differentiating "nasty" Wierstrass-like functions, but only functions that are nice (and smooth) perhaps except at a finite or countable number of points. On the other hand, this does not mean the theory just developed is useless; indeed, as the examples in the next section show, even seemingly minor singularities can turn out to be significant in the final result.

3.2 Relaxing the continuity criterion

If we are desperate about the assumption of Criterion 2.2, we can try to ignore it, with the penalty that we would have to deal with discontinuous functionals. But if the derivative of a discontinuous functional g is defined formally by

$$\langle \frac{\partial g}{\partial x_i}, \phi \rangle = -\langle g, \frac{\partial \phi}{\partial x_i} \rangle \quad \text{for every test function } \phi,$$

just as it is for continuous functionals, then we readily see that the proof of Theorem 2.4 survives.

If a functional g is not known to be continuous, then the consequence is that we would not be able to find a sequence of generalized functions g_n converging to g (in the sense that $\langle g_n, \phi \rangle \to \langle g, \phi \rangle$ for every test function ϕ). However, if we do manage to find such a sequence, then we can deduce that g is in fact continuous (by sequential completeness of the space of distributions).

Alternatively, we might be able to compute $g(x) = \int_\Omega f(x,\omega)\,d\omega$ and discover directly that it is a continuous functional without verifying Criterion 2.2 *a priori*. It even works to first calculate $\partial g/\partial x_i$ formally by differentiation under the integral sign, and discover that it is continuous. *Then* we can also deduce after-the-fact that g must be continuous too, by the theorem of distribution theory that anti-derivatives of any generalized function exist as generalized functions.

3.3 Linear operations other than integration

4 Computational examples

4.1 Fundamental Theorem of Calculus

To warm-up, let us look at the following amusing, if hopelessly roundabout, demonstration of the Fundamental Theorem of Calculus. For f continuous, and $a \leq x \leq b$, we have

$$\frac{d}{dx}\int_a^x f(t)\,dt = \frac{d}{dx}\int_a^b \mathbf{1}(t \leq x) f(t)\,dt = \int_a^b \frac{\partial}{\partial x} f(t)\,\mathbf{1}(x \geq t)\,dt$$
$$= \int_a^b f(t)\,\delta(x-t)\,dt = f(x)\,,$$

where δ, is of course, the (in)famous Dirac delta function, and $\mathbf{1}(\cdot)$ denotes an indicator. (In effect, $\mathbf{1}(x \geq t)$ reduces to $H(x-t)$, where H is the Heaviside step function, so its derivative is $\delta(x-t)$.)

Actually, our derivation is not quite rigorous, though the method does seem to "work". However, rather than obsessing over a trivial calculation, we defer the explanation of the rigorous method to the next examples, where we can make serious blunders when we fail to understand the theory correctly — or at least when we are not careful.

4.2 Smooth integrands

Consider the well-known Gamma function: $\Gamma(x) = \int_0^\infty e^{-t} t^{x-1}\,dt$ for $x > 0$. Since Γ is continuous (and the integrand is non-negative), it satisfies Criterion 3.1, and therefore differentiation under the integral sign is allowed:

$$\frac{d^k}{dx^k}\Gamma(x) = \int_0^\infty e^{-t}\,\frac{\partial^k}{\partial x^k} t^{x-1}\,dt\,.$$

More generally, let $f\colon X \times \Omega \to \mathbb{R}$ be an integrand that is smooth in the parameter, satisfying Criterion 3.1; then we can differentiate (weakly) under the integral sign as many times as we want. However, there is no guarantee that $\partial f/\partial x_i$, say, has to be integrable over Ω in the Lebesgue sense, or that $\int_\Omega f(x,\omega)\,d\omega$ is a smooth function of x.

Such anomalies happen, for example, with the function $f(x,y) = y^{-2}(1-\cos xy)$ for $y > 0$ and $x \in \mathbb{R}$. The integral $\int_0^\infty f(x,y)\,dy$ converges (and is a continuous function of x — see Theorem A.2), but

$$\int_0^\infty \frac{\partial f}{\partial x}\,dy = \text{``}\int_0^\infty \frac{\sin xy}{y}\,dy\text{''} \overset{?}{=} \lim_{s\to\infty}\int_0^s \frac{\sin xy}{y}\,dy = \frac{\pi}{2}\operatorname{sgn}(x)$$

does not exist as a real Lebesgue integral. But it is not hard to see that the interpretation of the divergent integral as the improper limit, as written above, is correct. Here are the gory details:

Let ϕ be a smooth function supported on a compact interval $[a,b]$; then

$$\langle \int_0^\infty \frac{\partial f}{\partial x}\,dy, \phi\rangle := -\int_0^\infty \int_a^b f(x,y)\,\frac{d\phi}{dx}\,dx\,dy = -\lim_{s\to\infty}\int_0^s \int_a^b f(x,y)\,\frac{d\phi}{dx}\,dx\,dy\,,$$

and we can change the outer integral to a limit of integrals since $f \cdot d\phi/dx$ is a $\mathbf{L}^1([a,b] \times (0,\infty))$ function. Next, we perform an integration by parts, then switch the order of integration, which is allowed since $(\partial f/\partial x)\cdot\phi$ is continuous on the compact set $[a,b]\times[0,s]$ (and so has finite integral there):

$$= \lim_{s\to\infty} \int_0^s \int_a^b \frac{\partial f}{\partial x}\, \phi(x)\, dx\, dy = \lim_{s\to\infty} \int_a^b \left(\int_0^s \frac{\partial f}{\partial x} dy \right) \phi(x)\, dx\,.$$

Finally, we can move the limit to inside the first integral since $|\int_0^s \frac{\sin xy}{y}\, dy|$ is uniformly bounded for all $x \in [a,b]$ and $s \in (0,\infty)$:

$$= \int_a^b \left(\lim_{s\to\infty} \int_0^s \frac{\partial f}{\partial x} dy \right) \phi(x)\, dx\,.$$

This verifies the claim that $\int_0^\infty \frac{\partial f}{\partial x}\, dy$ is the locally-integrable function $\lim_{s\to\infty} \int_0^s \frac{\sin xy}{y}\, dy$.

We leave it to the reader to ponder the proper interpretation of this apparent nonsense:

$$\pi\,\delta(x) = \frac{d}{dx}\left(\frac{\pi}{2}\operatorname{sgn}(x)\right) = \frac{d}{dx}\int_0^\infty \frac{\sin xy}{y}\, dy = \int_0^\infty \frac{\partial}{\partial x}\frac{\sin xy}{y}\, dy = \int_0^\infty \cos(xy)\, dy\,.$$

4.3 Smooth integrands: classical differentiability

On the brighter side, we can pose conditions on the integrand f such that its integral has a derivative, in the usual sense, and is given by taking the derivative past the integral:

Theorem 4.1 (Differentiation under the integral sign)**.** *Let $f\colon X\times\Omega \to \mathbb{R}$ be an ordinary function such that the generalized derivatives $\partial f/\partial x_i$ are represented by ordinary functions (for each $\omega \in \Omega$), and both f and $\partial f/\partial x_i$ are locally integrable as in Criterion 3.1. Then almost everywhere on X,*

$$\frac{\partial}{\partial x_i}\int_\Omega f(x,\omega)\, d\omega = \int_\Omega \frac{\partial}{\partial x_i} f(x,\omega)\, d\omega$$

where classical differentiation is used on both sides of the equation.

Of course, if both sides are continuous in x, then equality holds for every $x \in X$.

Proof. Define $g(x) = \int_\Omega f(x,\omega)\, d\omega$. To avoid ambiguity, we temporarily distinguish generalized derivatives from ordinary derivatives by the addition of brackets: $[\partial/\partial x_i]$.

Let ϕ be a test function supported on a compact set K. Then we have

$$\begin{aligned}\left\langle \left[\frac{\partial g}{\partial x_i}\right], \phi \right\rangle &= \int_\Omega \left\langle \left[\frac{\partial f}{\partial x_i}\right], \phi \right\rangle d\omega && \text{(Theorem 2.4)}\\ &= \int_\Omega \int_K \frac{\partial f(x,\omega)}{\partial x_i}\, \phi(x)\, dx\, d\omega && \text{(Theorems C.2, C.4)}\\ &= \int_K \left(\int_\Omega \frac{\partial f(x,\omega)}{\partial x_i}\, d\omega \right) \phi(x)\, dx && \text{(used local integrability).}\end{aligned}$$

Thus

$$\left[\frac{\partial g}{\partial x_i}\right] = h := \int_\Omega \frac{\partial}{\partial x_i} f(\cdot,\omega)\, d\omega\,,$$

in the sense that the locally-integrable function h represents the generalized function $[\partial g/\partial x_i]$ by integration. Now apply Theorem C.4 to g and h. ∎

Compare with Theorem A.1. Loosely speaking, that theorem requires the integral $\int_\Omega |\partial f/\partial x_i|\, d\omega$ to be a locally bounded function of the parameter; here we only demand the integral to be locally integrable, which is a slightly weaker requirement. (For instance, the function $f(x,y) = \sqrt{|x-y|}$ for $x, y \in [0,1]$ fails the first requirement but satisfies the second.)

A nice advantage of our theory though, from the point of the lazy physicist or engineer, is that it allows us to neglect checking the integrability conditions beforehand, and to just go ahead and differentiate under the integral sign formally. If the result of the calculation "makes sense" (i.e. the relevant integrals converge), then the operations will be justified.

4.4 Smooth integrands: another counterexample

(This part may be safely skipped if the reader is getting bored with pathological examples.)

This example comes from [Gelbaum]: let

$$\begin{aligned} f(x,y) &= \frac{x^3}{y^2} e^{-x^2/y}\,, \quad x \in \mathbb{R}, y \in (0,1] \\ \frac{\partial f(x,y)}{\partial x} &= \frac{3x^2}{y^2} e^{-x^2/y} - \frac{2x^4}{y^3} e^{-x^2/y} \end{aligned}$$

Theorem 4.1 applies. Computing directly, we find

$$\frac{d}{dx} \int_0^1 f(x,y)\, dy = \frac{d}{dx}(x e^{-x^2}) = e^{-x^2}(1-2x^2) = \int_0^1 \frac{\partial f(x,y)}{\partial x}\, dy$$

for all $x \neq 0$, but the left- and right-hand sides do not agree at the single point $x = 0$. Of course, there is no contradiction, for Theorem 4.1 only asserts the equality of the derivatives *almost everywhere*.

In the context of Theorem A.1, observe that $|\partial f/\partial x|$ is integrable (with respect to y) when $x \neq 0$, but is not integrable when $x = 0$. Moreover, the dominating factor $\Theta(y)$ can be found for this integrand if x is restricted to a compact set *not including the origin*, but not if x can get arbitrarily close to the origin.

4.5 Integrands that are step functions

Next, we come to a detailed example involving singularities described by delta functions.

Let $\Omega = [0,1]$ (with the elements denoted by the variable y), and $c \geq 0$ be a constant. Define

$$f(x,y) = y\, \mathbf{1}(xy > c)\,, \quad g(x) = \int_0^1 f(x,y)\, dy\,, \quad \text{for } x \in \mathbb{R} \text{ and } y \in [0,1].$$

For comparison, we first compute g and its derivative directly, without differentiation under the integral sign:

$$\begin{aligned} g(x) &= \frac{1}{2}\left(1 - \left(\frac{c}{x}\right)^2\right) \mathbf{1}(x > c)\,, \\ \frac{dg(x)}{dx} &= \frac{1}{x}\left(\frac{c}{x}\right)^2 \mathbf{1}(x > c) \quad \text{(except at } x = c\text{)}. \end{aligned}$$

Next follows a tempting, but unrigorous — *and totally wrong* — computation. (I illustrate these examples because I made the same mistakes before I fully absorbed the proper theory.)

Blunder: We try to apply the chain rule to the Heaviside step function:

$$\frac{dg}{dx} = \int_0^1 \frac{\partial}{\partial x} y\, H(xy - c)\, dy = \int_0^1 y^2\, \delta(xy - c)\, dy = y^2\big|_{xy=c} = \left(\frac{c}{x}\right)^2,$$

for $0 < \frac{c}{x} < 1$.

Even without knowing the true solution for dg/dx, we know that the answer just obtained must be wrong because it has the wrong order in x and fails a dimensional analysis: if y is unitless, but both x and c have units of distance, then g is also unitless, and thus dg/dx must have units of inverse distance.

It turns out that the use of the chain rule is correct, but the substitution

$$\int_0^1 \gamma(y)\, \delta(xy - c)\, dy \to \gamma(y)\Big|_{xy-c=0} = \gamma\left(\frac{c}{x}\right)$$

is incorrect. To understand this intuitively, recall that δ function appeared from partial differentiation *with respect to* x; that means the standard properties of the delta function apply only with respect to the variable x: e.g.

$$\int_{-\infty}^{\infty} \gamma(x)\, \delta(x - a)\, dx = \gamma(a)$$

for continuous functions γ. But we were integrating with respect to the variable y, the integrand $\gamma(y)\, \delta(xy - c)$. Indeed, if we write $\delta(xy - c) \equiv \delta(x - c/y)$, we see that

$$\int_0^1 \delta(xy - c)\, dy = \int_0^1 \delta\Big(x - \frac{c}{y}\Big)\, dy \neq \frac{c}{x},$$

because now y^{-1} in the argument of the delta function causes the integral to accumulate values differently than if the argument had been y itself.

The correct value of the integral is, instead, suggested by a differential change of variables:

$$\begin{aligned}\int_0^1 y\, \delta\Big(x - \frac{c}{y}\Big)\, dy &= \int_c^{\infty} \frac{c}{u} \delta(x - u)\, \frac{c}{u^2} du \quad \text{substitute } u = \frac{c}{y},\ dy = -\frac{c}{u^2} du \\ &= \frac{c^2}{u^3}\bigg|_{u=x} \qquad \text{if } x \geq c, \text{ zero otherwise,}\end{aligned}$$

yielding the desired answer (with the dimensional inconsistency fixed).

Having illustrated this pitfall, we proceed with the rigorous computation. Firstly, f satisfies Criterion 3.1, so differentiation under the integral *is* allowed. Secondly, to dispel any doubts about the validity of $\partial f/\partial x = y\, \delta(x - c/y)$, we compute it formally. For any test function ϕ (with compact support): we have, indeed,

$$\langle \frac{\partial f}{\partial x}, \phi \rangle = -\langle f, \frac{d\phi}{dx} \rangle = -\int_{-\infty}^{\infty} y\, \mathbf{1}(xy > c)\, \frac{d\phi}{dx}\, dx = -y \int_{c/y}^{\infty} \frac{d\phi}{dx}\, dx = y\, \phi\Big(\frac{c}{y}\Big).$$

Now, applying Theorem 2.4 and performing the same differential change of variables $u = c/y$ — it is legitimate now that the integrals are really Lebesgue integrals — we find:

$$\langle \frac{dg}{dx}, \phi \rangle = \int_0^1 \langle \frac{\partial f}{\partial x}, \phi \rangle \, dy = \int_0^1 y \, \phi\Big(\frac{c}{y}\Big) \, dy = \int_c^\infty \frac{c^2}{u^3} \, \phi(u) \, du \, .$$

But we recognize the last integral as the functional, evaluated at ϕ, corresponding to the integrable function $x \mapsto c^2 x^{-3} \, \mathbf{1}(x > c)$.

It is again instructive to examine what Theorem A.1 would say in this situation. Clearly, that theorem cannot possibly apply here, because $\partial f / \partial x = 0$ whenever the derivative exists in the classical sense, and so integrating this naïvely would yield zero identically.

Theorem A.1 asserts that it applies if f is *everywhere* differentiable with respect to x and *almost every* ω, but here we really have f being *almost* everywhere differentiable with respect to x for *every* ω, These two conditions are not the same. In fact, although both X and Ω happened to be subsets of $\mathbb{R}$ in the current example, the roles they play are *not* symmetric. Theorem A.1 uses no measure theory on X-space, but only on Ω-space, so naturally, only in Ω-space do the notions of "almost everywhere" enter into the hypotheses and the conclusions of the theorem.

Even if we arbitrarily redefine the partial derivatives to be zero whenever they do not exist, it is easy to see, in this example, that the difference quotient — see the remarks after Theorem A.1 — cannot be dominated.

Another way to understand this situation, is if we decompose $\partial f / \partial x_i$ into a sum

$$\frac{\partial}{\partial x_i} f(x, \omega) = p(x, \omega) + q(x, \omega) \, ,$$

of an ordinary $\mathbf{L}^1(X \times \Omega)$ function p and a generalized function q, then

$$\frac{\partial}{\partial x_i} \int_\Omega f(x, \omega) \, d\omega = \int_\Omega p(x, \omega) \, d\omega + \int_\Omega q(x, \omega) \, d\omega \, ,$$

where the far right term is exactly the contribution to the derivative that would be missed out when differentiating classically under the integral sign.

4.6 The Black-Scholes pricing formula

Here we present an applied calculation from the field of mathematical finance. (This is one of the problems that I had been working on, leading me to investigate more thoroughly differentiation under the integral.)

Consider the function

$$P(x) = e^{-r\tau} \, \mathbb{E}[\max(xY - K, 0)] \, , \quad x > 0 \, , \, Y = e^{(r - \frac{1}{2}\sigma^2)\tau + \sigma\sqrt{\tau} Z} \, , \, Z \sim \text{Normal}(0, 1) \, .$$

which gives the "arbitrage-free" price of a European call option for a stock, given the current price of the stock x, modelled as a geometric Brownian motion. (The parameter r is the interest rate, σ is the volatility of the stock, τ is the time from the present to the maturity of the European call option, and K is its strike price.)

The expectation can of course be rewritten as an integral over $\mathbb{R}$ of the integrand multiplied by the density of a standard normal distribution, but for the sake of variety, let us stick with the form given above. (Thus, here Ω is taken to be some probability space.)

The function P has the analytical form given by

$$P(x) = x\Phi(d_+) - Ke^{-r\tau}\Phi(d_-)\,, \quad d_\pm(x) = \frac{(r \pm \frac{1}{2}\sigma^2)\tau + \log(x/K)}{\sigma\sqrt{\tau}}\,.$$

Calculating dP/dx is somewhat messy from this formula; we start instead with a differentiation under the expectation sign:

$$\begin{aligned}\frac{dP}{dx} &= e^{-r\tau}\,\mathbb{E}\left[\frac{\partial}{\partial x}\max(xY - K, 0)\right] = e^{-r\tau}\,\mathbb{E}\big[Y\,\mathbf{1}(xY > K)\big] \\ &= e^{-\frac{1}{2}\sigma^2\tau}\,\mathbb{E}\left[e^{-\sigma\sqrt{\tau}Z}\,\mathbf{1}(Z > -d_-)\right] = \Phi(d_+)\,.\end{aligned}$$

The exchange of differentiation and expectation is justified by the following estimate:

$$\int_a^b \mathbb{E}\big[|\max(xY - K, 0)|\big]\,dx \leq \int_a^b \mathbb{E}[xY]\,dx < \infty\,, \quad 0 < a < b\,,$$

and Criterion 3.1 is satisfied[2].

To obtain $d^2P/dx^2 = \Phi(d_+)/(x\sigma\sqrt{\tau})$, it is actually much easier to differentiate dP/dx directly. But to illustrate the techniques, we do a differentiation under the expectation sign again:

$$\begin{aligned}\frac{d^2P}{dx^2} = \frac{d}{dx}\frac{dP}{dx} &= e^{-r\tau}\,\mathbb{E}\left[\frac{\partial}{\partial x}Y\mathbf{1}(xY > K)\right] = e^{-r\tau}\,\mathbb{E}\big[Y\,\delta(x - K/Y)\big] \\ &= e^{-r\tau}\,\mathbb{E}\big[KU^{-1}\,\delta(x - U)\big]\,, \quad U = K/Y\,.\end{aligned}$$

At this point we have no choice but to express the expectation as an integral over the reals, because of the involvement of a delta function. More formally, if we suppose that d^2P/dx^2 is a locally-integrable function h, then we must show that $\int_0^\infty h(x)\psi(x)\,dx = Ke^{-r\tau}\,\mathbb{E}\big[U^{-1}\psi(U)\big]$ for all test functions ψ, and that obviously necessitates expanding the right-hand side as a Lebesgue integral. The calculation is not too hard if we observe that $\log U$ has the distribution $\mathrm{Normal}(\log K - (r - \frac{1}{2}\sigma^2)\tau, \sigma^2\tau)$, so that

$$Ke^{-r\tau}\,\mathbb{E}\left[\frac{\delta(x - U)}{U}\right] = Ke^{-r\tau}\int_0^\infty \frac{\delta(x - u)}{u}\,\overbrace{\frac{\exp(-\frac{1}{2}d_-(u)^2)}{u\sqrt{2\pi}\,\sigma\sqrt{\tau}}}^{\text{density for log-normal}}\,du = \frac{Ke^{-r\tau}}{x^2}\,\Phi'(d_-(x))\,.$$

The latter expression looks different from the direct formula for d^2P/dx^2, but they are in fact equal.

We have played fast-and-loose with delta functions in the last calculation, but clearly it abbreviates the rigorous arguments we have already made. The reader who is still unsure about the method is invited to repeat the explicit calculation with functionals.

[2] We can also verify that the hypotheses of Theorem 4.1 indeed hold for the first derivative, but the conclusions from that theorem, namely that the classical derivative exists and equals the generalized derivative, are already immediate from our calculated results. But Theorem 4.1 does *not* apply for the second derivative, because $\partial Y\mathbf{1}(xY > K)/\partial x$ is *not* an ordinary function.

4.7 Smooth approximations of integrands

A Differentiation under the integral sign with classical derivatives

Theorem A.1 (Differentiation under the integral sign)**.** *Let X be an open subset of $\mathbb{R}^m$, and Ω be a measure space. Suppose that the function $f\colon X \times \Omega \to \mathbb{R}$ satisfies the following conditions:*

1. *$f(x,\omega)$ is a measurable function of ω for each $x \in X$.*
2. *For almost all $\omega \in \Omega$, the derivative $\partial f(x,\omega)/\partial x_i$ exists for all $x \in X$.*
3. *There is an integrable function $\Theta\colon \Omega \to \mathbb{R}$ such that $|\partial f(x,\omega)/\partial x_i| \leq \Theta(\omega)$ for all $x \in X$.*

Then

$$\frac{\partial}{\partial x_i} \int_\Omega f(x,\omega)\, d\omega = \int_\Omega \frac{\partial}{\partial x_i} f(x,\omega)\, d\omega\,.$$

This result is proved by (what else?) the Lebesgue Dominated Convergence Theorem.

A more precise condition, in place of (3) above, that comes out of the proof of the theorem, is that the difference quotient $(f(x+te_i,\omega) - f(x,\omega))/t$ must be dominated by some $\Theta(\omega)$ for all x and t.

For completeness, we also include the following analogous result for continuity:

Theorem A.2 (Continous dependence on integral parameter)**.** *Let X be an open subset of $\mathbb{R}^m$ (or any metric space), and Ω be a measure space. Suppose that the function $f\colon X \times \Omega \to \mathbb{R}$ satisfies the following conditions:*

1. *$f(x,\omega)$ is a measurable function of ω for each $x \in X$.*
2. *For almost all $\omega \in \Omega$, $f(x,\omega)$ is continuous in x.*
3. *There is an integrable function $\Theta\colon \Omega \to \mathbb{R}$ such that $|f(x,\omega)| \leq \Theta(\omega)$ for all $x \in X$.*

Then $\int_\Omega f(x,\omega)\, d\omega$ is a continuous function of x.

B Bibliography

References

[Folland] Gerald B. Folland, *Real Analysis: Modern Techniques and Their Applications*, second ed. Wiley-Interscience, 1999.

[Gelbaum] Bernard R. Gelbaum and John M. H. Olmsted. *Counterexamples in Analysis.* Dover, 2003.

[Hoskins] Roy F. Hoskins and J. Sousa Pinto. *Theories of Generalized Functions.* Horwood Publishing, 2005.

[Jones] D. S. Jones. *The Theory of Generalized Functions*, second ed. Cambridge University Press, 1982.

[Lützen] Jesper Lützen. *The Prehistory of the Theory of Distributions*. Springer-Verlag, 1982.

[Schwartz] Laurent Schwartz. *Théorie des Distributions*, volume I. Hermann, 1957.

[Talvila] Erik Talvila. "Necessary and Sufficient Conditions for Differentiating Under the Integral Sign". *American Mathematical Monthly*, 108 (June-July 2001) 544-548. Found on-line at:

`http://www.math.ualberta.ca/~etalvila/papers/difffinal.pdf`

([Lützen] does not develop any mathematical theory, but I happened upon it researching the current problem. It is an interesting look into the intuition and motivation behind the theory of generalized functions. [Talvila] gives the exact conditions for exchanging integration and classical differentiation, in terms of Henstock integrals. But it seems that those conditions might be difficult to verify in practice, and in any case would fail with step functions whose derivatives are delta functions.)

C List of supporting theorems

Here is a list of some of the theorems we have appealed to implicitly but may not be familiar to the reader. (Actually, I too was not too familiar with them before I started writing this article.)

Theorem C.1 (Fundamental Theorem of Calculus for Lebesgue integrals)**.** *Let f be a Lebesgue-integrable function on a compact interval $[a, b]$. If*

$$G(x) = \int_a^x f(t)\,dt\,, \quad \text{for } x \in [a, b]\,,$$

then G is absolutely continuous, and $G' = f$ almost everywhere. Moreover, if $f = g'$ for an absolutely continous g, then $G = g + c$ for some constant c.

Proof. See, for example, Theorem 3.35 in [Folland].

Theorem C.2 (Integration by Parts)**.** *Let X be an open set in $\mathbb{R}^m$. The usual integration-by-parts formula is valid for locally-integrable functions $f\colon X \to \mathbb{R}$ absolutely continuous*[3] *in x_i, with $\partial f/\partial x_i$ locally integrable, and smooth test functions $\phi\colon X \to \mathbb{R}$ vanishing outside a compact set in X:*

$$\langle \frac{\partial f}{\partial x_i}, \phi \rangle = \int_X \frac{\partial f(x)}{\partial x_i}\,\phi(x)\,dx = -\int_X f(x)\,\frac{\partial \phi(x)}{\partial x_i}\,dx = -\langle f, \frac{\partial \phi}{\partial x_i} \rangle\,.$$

Thus the generalized derivative agrees with the ordinary derivative for such f.

[3] Here and elsewhere, "absolutely continuous in x_i" really means "absolutely continuous in x_i when the other variables x_j, $j \neq i$, are held fixed, after some modification of the function on a set of measure zero.

Proof. This statement appears in [Schwartz], Ch. 2, Sect. 5, Theorem V, part (1).

It can be proven by applying Theorem C.1 and the integration by parts formula for Lebesgue-Stieljes integrals, Theorem 3.36 in [Folland].

Warning. the condition "f is absolutely continuous" cannot be replaced by the weaker condition that "f' exists almost everywhere and is locally integrable". The same sort of restriction enters into the second part of Theorem C.1, where we need to make sure that g is absolutely continuous before concluding $G = g + c$. We do not even have to try hard to find pathological counterexamples to show this necessity: if H is the Heaviside step function, then $H' = 0$ classically, but of course we already know that its generalized derivative is the Dirac delta. In terms of Theorem C.1: H is not absolutely continuous and cannot be expressed as the integral of its derivative.

Theorem C.3. *If f is any generalized function on X, an open set in $\mathbb{R}^m$, then the partial differential equation $\partial u/\partial x_i = f$ has infinitely many generalized solutions, each differing by some generalized function independent of x_i.*

Proof. See [Schwartz], Ch. 2, Sect. 5, Theorem IV.

Theorem C.4. *Let X be an open set in $\mathbb{R}^m$. If a locally-integrable function $f\colon X \to \mathbb{R}$ has a generalized derivative represented by a locally-integrable function $h\colon X \to \mathbb{R}$, then f is absolutely continuous in x_1, and it admits an ordinary derivative of h almost everywhere. (This result is a converse to Theorem C.2.)*

Proof. See [Schwartz], Sect. 5, Theorem V, part (2). We reproduce the proof here:

To avoid ambiguity, we temporarily distinguish generalized derivatives from ordinary derivatives by the addition of brackets: $[\partial/\partial x_i]$.

Define an indefinite integral $G(x) = \int h(x)\,dx_i$. Then $G(x)$ is absolutely continuous in x_i, and $\partial G/\partial x_i = h$ almost everywhere. It is also true that $[\partial G/\partial x_i] = h$.

From the latter statement, we know from Theorem C.3 that $G - f = c$ where c is some generalized function independent of x_i. Since $G - f = c$ is a locally-integrable function, we must have $\partial c/\partial x_i = 0$ *using ordinary derivatives*. It follows that $\partial f/\partial x_i$ exists almost everywhere, and $\partial f/\partial x_i = \partial G/\partial x_i = h$. The function f is absolutely continuous in x_i since $f = G - c$ and both G and c are absolutely continuous in x_i. ■

Theorem C.5. *The differentiation operator $\partial/\partial x_i$ is a continuous mapping from generalized functions to generalized functions.*

Proof. See [Schwartz], Ch. 3, Sect. 5.

Theorem C.6. *The spaces $\mathcal{D}$ and $\mathcal{D}'$ are reflexive.*

Proof. See [Schwartz], Ch. 3, Sect. 3, Theorem XIV.