

4-1-2004

The Wave Equation, Mixed Partial Derivatives, and Fubini's Theorem

Asuman Aksoy
Claremont McKenna College

Mario Martelli
Harvey Mudd College

Recommended Citation

Aksoy, A. Martelli, M., "Wave Equation, Mixed Partial Derivatives and Fubini's Theorem" the American Mathematical Monthly, Vol. 111, 2004, N.4, 2004. doi: 10.2307/4145243

This Article is brought to you for free and open access by the CMC Faculty Scholarship at Scholarship @ Claremont. It has been accepted for inclusion in CMC Faculty Publications and Research by an authorized administrator of Scholarship @ Claremont. For more information, please contact scholarship@cuc.claremont.edu.

NOTES

Edited by William Adkins

The Wave Equation, Mixed Partial Derivatives, and Fubini's Theorem

Asuman Aksoy and Mario Martelli

In a recent paper [1] the two authors of this note have shown that Fubini's theorem on changing the order of integration and Schwarz's lemma on the equality of mixed partial derivatives are equivalent when standard assumptions of continuity and differentiability are made. The proof relies heavily upon the fundamental theorem of integral calculus as usually presented in calculus textbooks. Fubini's theorem is regarded as intuitive and easy to prove. Therefore, the equivalence established in [1] with a straightforward argument provides a simple proof of Schwarz's lemma.

However, there are many cases in which the standard assumptions from which the equality of the mixed partials is derived do not hold. A typical scenario is offered by some cases of the one-dimensional wave equation.

Example 1. In the subset $U = (-\infty, +\infty) \times [0, \infty)$ of \mathbf{R}^2 (= the (x, t) -plane) consider the initial value problem

$$\begin{cases} u_{xx} = u_{tt} & (-\infty < x < \infty, 0 < t < \infty), \\ u(x, 0) = g(x), \\ u_t(x, 0) = 0, \end{cases} \quad (1)$$

where g is the periodic extension to \mathbf{R} of the function $f(x) = x^3 - x$ defined for x in $[-1, 1]$ (see Figure 1). Since $f''(1) = -f''(-1) = 6$, the second derivative of g fails to exist at each point $x_n = 2n + 1$ for $n = 0, \pm 1, \dots$. We neglect, for the time being, any potential problem arising from this lack of differentiability, and we apply D'Alembert's strategy to (1). Thus, we make the substitution $\eta = x + t$, $\xi = x - t$ and invoke the chain rule in conjunction with Schwarz's lemma to transform the differential equation $u_{xx} = u_{tt}$ into $u_{\eta\xi} = 0$.

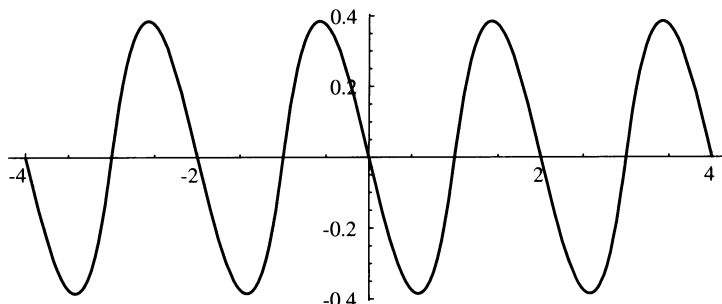


Figure 1. The initial position ($t = 0$) of the wave.

We easily derive the solution and revert to the original variables to obtain

$$u(x, t) = \frac{1}{2}(g(x + t) + g(x - t)). \quad (2)$$

Unfortunately, from the properties of g mentioned before, we see that u_{xt} and u_{tx} do not exist when (x, t) belongs to $\Gamma = \{(x, t) \in U : t > 0, x \pm t = \pm 2n + 1, n = 0, 1, \dots\}$ (see Figure 2). Therefore, one wonders whether D'Alembert's method can in fact be used to solve (1), since formula (2) is obtained assuming the continuity and the equality of the mixed partials of u in U . However, it can be shown that the function $u(x, t)$ given by (2) is the unique solution of (1). Because D'Alembert's method provides it, we are led to the suspicion that Schwarz's lemma can be generalized.

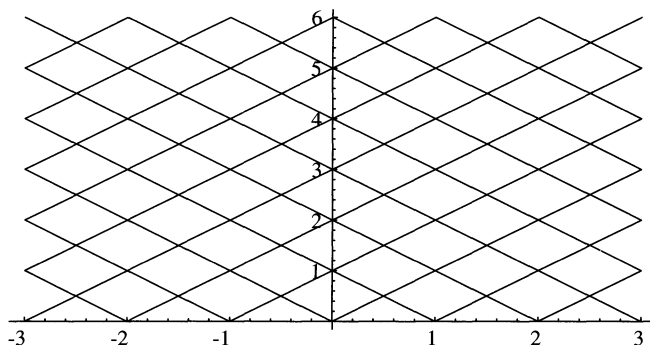


Figure 2. The lines where the second order partial derivatives of $u(x, t)$ do not exist.

This brings us to the purpose of our note. We want to achieve three goals:

- to enlarge the class S of functions for which Schwarz's lemma is valid to include all functions having the same properties of $u(x, t)$ given by (2);
- to enlarge the class F of functions for which Fubini's theorem is valid so that the intuitive character of the theorem is preserved;
- to obtain a simple proof of the equivalence between Fubini's theorem for the class F and Schwarz's lemma for the class S .

All three goals are achieved with Theorem 6. We explicitly state, however, that Theorem 6 proves neither Fubini's theorem for the functions of class F nor Schwarz's lemma for the functions of class S . It simply establishes the equivalence of the two results. The appropriate extension of the fundamental theorem of integral calculus and some of its consequences that are needed in the proof of Theorem 6 are furnished by Theorems 3 and 5.

The astute reader will realize that the three goals impose limitations on membership in S or F . However, we did not list a proof of the most general result on the equivalence between Fubini's theorem and Schwarz's lemma as one of our goals. Moreover, the methods we use are very elementary and the generality we achieve is appropriate for many situations. Hence, the result we obtain, although not the most general, is worth the effort.

To understand the direction we should take, we first recall some facts regarding the basic equality

$$f(x) = f(a) + \int_a^x f'(s) ds. \quad (3)$$

It is well known (see, for example, [4]) that there exist functions $f : [a, b] \rightarrow \mathbf{R}$ that are differentiable at each point of $[a, b]$, with f' bounded, but not Riemann integrable over $[a, b]$. Hence, the Riemann integrability of f' is necessary for the validity of (3). However, the integrability of f' is not sufficient. For example, the Cantor ternary function t is not differentiable at any point of the Cantor middle-third set K in $[0, 1]$, and $t'(x) = 0$ for every x in $[0, 1] \cap K^c$, where K^c denotes the complement of K . Since K is negligible (i.e., has Lebesgue measure zero) we know that t' is Riemann integrable over $[0, 1]$. However,

$$1 = t(1) - t(0) \neq \int_0^1 t'(x) dx.$$

The question naturally arises: In what cases is equality (3) valid? The continuity of f' guarantees (3), but we cannot always count on this property. Theorem 3 provides an answer that, although not as general as possible, is sufficient for our purposes. There are alternatives. For example, one could use Henstock's integral and establish the following result (see [5, Theorem 9.6]):

Theorem 1. *Let $F : [a, b] \rightarrow \mathbf{R}$ be continuous in $[a, b]$. Assume that F is differentiable in $[a, b]$ except possibly on a countable set. Then F' is Henstock integrable and*

$$\int_a^x F'(s) ds = F(x) - F(a).$$

However, as mentioned earlier, we prefer not to follow this path, for it would force us to provide a version of Fubini's theorem in the context of Henstock integration. We feel that this version would not have the intuitive appeal that we seek to preserve. Hence, we use a less general but more elementary approach.

In what follows, a subset A of the real line is termed *admissible* if its closure is at most countable. Recall that the closure of A is the union of A and its boundary. When A is admissible and $f : [a, b] \cap A^c \rightarrow \mathbf{R}$ is continuous, any extension of f to $[a, b]$ is continuous except possibly on a countable set. The following result, whose proof is elementary, can be found in [3].

Theorem 2. *Let A be an admissible set, and let $f : [a, b] \cap A^c \rightarrow \mathbf{R}$ be bounded and continuous. If f_E is a bounded extension of f to the interval $[a, b]$, then f_E is Riemann integrable in $[a, b]$ and the value of the integral is independent of the extension.*

We now obtain a generalization of the fundamental theorem of integral calculus that is suitable for our purposes.

Theorem 3. *Let A be an admissible set, and let $f : [a, b] \cap A^c \rightarrow \mathbf{R}$ be bounded and continuous. Extend f to the entire interval $[a, b]$ so that the extension f_E is bounded, and define $F : [a, b] \rightarrow \mathbf{R}$ by*

$$F(x) = \int_a^x f_E(s) ds.$$

Then F is continuous on $[a, b]$, and $F'(x) = f(x)$ holds for every x in $[a, b] \cap A^c$.

Proof. Theorem 2 ensures the Riemann integrability of f_E and the independence of F from the choice of extension. Standard properties of the Riemann integral imply that

$$|F(x) - F(y)| \leq M|x - y|$$

where $M = \sup\{f_E(x) : x \in [a, b]\}$. Hence F is continuous.

To establish the equality $F'(x) = f(x)$ in $[a, b] \cap A^c$, let x_0 be a point of the set $[a, b] \cap A^c$. Then f is continuous at x_0 . Hence, for each $r > 0$ there exists d with $0 < |d| < r$ such that

$$f(x_0) - r \leq f(x) \leq f(x_0) + r$$

whenever x lies in $(x_0 - |d|, x_0 + |d|)$. Consider the extension of f to $[a, b]$ that has $f(x) = f(x_0)$ for all x in A , and denote it again by f . Integrating the functions involved in the foregoing inequality between x_0 and $x_0 + d$ and dividing by d , we obtain (regardless of whether d is positive or negative)

$$f(x_0) - r \leq \frac{F(x_0 + d) - F(x_0)}{d} = \frac{1}{d} \int_{x_0}^{x_0+d} f(s) ds \leq f(x_0) + r. \quad (4)$$

Letting $r \rightarrow 0$ we see that $F'(x_0) = f(x_0)$. ■

In addition to Theorem 3 we need a suitable extension of the second part of the fundamental theorem of integral calculus, the part that allows one to evaluate a definite integral once an antiderivative of its integrand is found. Recall that in its proof one needs the property (derived from the mean value theorem) that two continuous functions $f, g : [a, b] \rightarrow \mathbf{R}$ such that $f'(x) = g'(x)$ for all x in (a, b) differ by a constant. We now extend this result to functions that may not be differentiable at all points of (a, b) .

Theorem 4. *Let $f, g : [a, b] \rightarrow \mathbf{R}$ be continuous, and let A be an admissible set. If f and g are differentiable with $f'(x) = g'(x)$ for each x in $(a, b) \cap A^c$, then $f - g$ is constant.*

Proof. Let $h = f - g$. Clearly $h([a, b]) = h(A) \cup h([a, b] \cap A^c)$. By known theorems (see, for example, Lemmas 2 and 3 in [6]), neither $h(A)$ nor $h([a, b] \cap A^c)$ has interior points. Hence, h is constant. ■

Let A be admissible, and let $f : [a, b] \cap A^c \rightarrow \mathbf{R}$ be bounded and continuous. We call a continuous function $G : [a, b] \rightarrow \mathbf{R}$ that is differentiable in $(a, b) \cap A^c$ and has the property that $G'(x) = f(x)$ for all x in $(a, b) \cap A^c$ a *generalized antiderivative* of f . By Theorem 3 the set of generalized antiderivatives of f is not empty. Moreover, from Theorems 3 and 4 we derive the following important result:

Theorem 5. *Let A be an admissible set, and let $f : [a, b] \cap A^c \rightarrow \mathbf{R}$ be bounded and continuous. Extend f to the entire interval $[a, b]$ so that the extension is bounded, and denote this extension by f_E . If G is any generalized antiderivative of f , then*

$$\int_a^b f_E(x) dx = G(b) - G(a).$$

Proof. Let

$$F(x) = \int_a^x f_E(s) ds.$$

By Theorem 3, F is a generalized antiderivative of f and

$$\int_a^b f_E(x) dx = F(b).$$

Let G be another generalized antiderivative of f . By Theorem 4, $G - F$ is constant in $[a, b]$. From $F(a) = 0$ we obtain $G(b) - G(a) = F(b)$. ■

We are now ready to obtain the equivalence between Fubini's theorem and Schwarz's lemma in a context where continuity may fail in certain subsets. Introducing some convenient notation will expedite the discussion.

Let U be an open set in \mathbf{R}^2 , and let $\alpha_i : [0, 1] \rightarrow U$ ($i = 1, 2, \dots$) be a countable collection of continuously differentiable functions. Define $\Gamma = \cup_i \Gamma_i$, where Γ_i is the image of α_i . Assume that Γ is closed and locally finite (i.e., each point of Γ has a neighborhood that intersects only finitely many of the sets Γ_i) and that the intersection of Γ with any vertical or horizontal line is at most countable. Denote by $BC(U \cap \Gamma^c, \mathbf{R})$ the vector space of functions $g : U \rightarrow \mathbf{R}$ that are bounded and continuous in each bounded subset of $U \cap \Gamma^c$. Consider the following two statements:

(i) (Fubini's theorem) If $g \in BC(U \cap \Gamma^c, \mathbf{R})$ and $[a, b] \times [c, d] \subset U$, then

$$\int_a^b \int_c^d g(x, y) dy dx = \int_c^d \int_a^b g(x, y) dx dy. \quad (5)$$

(ii) (Schwarz's lemma) If $f \in C(U, \mathbf{R})$, $f_x, f_y, f_{xy} \in BC(U \cap \Gamma^c, \mathbf{R})$, and f_x is continuous with respect to the second variable, then f_y is continuous with respect to the first variable, f_{yx} exists in $U \cap \Gamma^c$, and

$$f_{yx}(x, y) = f_{xy}(x, y)$$

for all (x, y) in $U \cap \Gamma^c$.

Theorem 6. *The statements (i) and (ii) are equivalent.*

Proof. To see that (i) \Rightarrow (ii) consider a point (x, y) in U . Since U is open we can find $r > 0$ such that the open disk D of radius r centered at (x, y) is contained in U . Fix a point (a, c) in D . Since f_{xy} belongs to $BC(U \cap \Gamma^c, \mathbf{R})$ and f_x is continuous with respect to the second variable, we obtain that f_x is a generalized antiderivative of f_{xy} . Hence, a straightforward application of Theorem 5 gives

$$f_x(x, y) - f_x(x, c) = \int_c^y f_{xv}(x, v) dv.$$

Moreover, since f is in $C(U, \mathbf{R})$ and $f_x(x, y)$ belongs to $BC(U \cap \Gamma^c, \mathbf{R})$, the same theorem implies that

$$\int_a^x f_u(u, y) du = f(x, y) - f(a, y)$$

and

$$\int_a^x f_u(u, c) du = f(x, c) - f(a, c).$$

Therefore,

$$\int_a^x \int_c^y f_{uv}(u, v) dv du = f(x, y) - f(a, y) - f(x, c) + f(a, c).$$

By (i) the order of integration can be reversed to give

$$\int_c^y \int_a^x f_{uv}(u, v) du dv = f(x, y) - f(a, y) - f(x, c) + f(a, c). \quad (6)$$

To (6) we now apply Theorem 3 twice. First we differentiate both sides with respect to y to obtain

$$\int_a^x f_{uy}(u, y) du = \frac{\partial}{\partial y} (f(x, y) - f(a, y)).$$

Since f is in $C(U, \mathbf{R})$ and $f_y(x, y)$ belongs to $BC(U \cap \Gamma^c, \mathbf{R})$, the right-hand side gives $f_y(x, y) - df(a, y)/dy$. Hence, f_y is continuous with respect to the first variable. We then differentiate with respect to x . On the left-hand side we obtain $f_{xy}(x, y)$ and on the right-hand side $f_{yx}(x, y)$. Hence, $f_{yx}(x, y)$ exists and

$$f_{yx}(x, y) = f_{xy}(x, y)$$

for all (x, y) in $U \cap \Gamma^c$.

To see that (ii) \Rightarrow (i), let g belong to $BC(U \cap \Gamma^c, \mathbf{R})$. Extend g to U by setting $g(x, y) = 1$ for (x, y) in Γ , and denote the extension by g_E . Given a closed rectangle $Q = [a, b] \times [c, d]$ we can choose bounded open intervals I and J such that $Q \subset V = I \times J \subset U$. Define $h, f : V \rightarrow \mathbf{R}$ by

$$h(x, y) = \int_c^y g_E(x, v) dv, \quad f(x, y) = \int_a^x h(u, y) du.$$

Theorem 3 and the properties of g_E ensure that h is a member of $BC(V \cap \Gamma^c, \mathbf{R})$ and is continuous with respect to the second variable. Moreover, f belongs to $C(V, \mathbf{R})$. In fact,

$$\begin{aligned} |f(x+w, y+z) - f(x, y)| &\leq |f(x+w, y+z) - f(x+w, y)| \\ &\quad + |f(x+w, y) - f(x, y)| \\ &\leq M(|z(x+w-a)| + |w(y-c)|), \end{aligned}$$

where $M = \sup\{|g_E(x, y)| : (x, y) \in V\}$. From Theorem 3 we learn that $f_x = h$ and $h_y = g$ in $V \cap \Gamma^c$. Hence $f_{xy} = h_y$ there, and both f_x and f_{xy} are in $BC(V \cap \Gamma^c, \mathbf{R})$.

Proving that f_y is a member of $BC(V \cap \Gamma^c, \mathbf{R})$ is a bit more delicate. First notice that the intersection of each vertical line with Γ is a closed set that is at most countable. Hence, whenever (x, y) does not lie in Γ , we can find an open neighborhood of (x, y) in which h is continuous and has a continuous partial derivative with respect to y . Thus, we can use the mean value theorem to show that f_y is continuous at every point

of $V \cap \Gamma^c$. Moreover, f_y is bounded in $V \cap \Gamma^c$, since J is a bounded open interval and h_y belongs to $BC(V \cap \Gamma^c, \mathbf{R})$. Therefore f_y is in $BC(V \cap \Gamma^c, \mathbf{R})$.

By (ii) f_y is continuous with respect to the first variable, f_{yx} exists in $V \cap \Gamma^c$, and $f_{yx}(x, y) = f_{xy}(x, y) = g(x, y)$ there. The result now follows easily from Theorem 5. In fact, since f_x is a generalized antiderivative of f_{xy} with respect to y and f is a generalized antiderivative of f_x with respect to x , we obtain

$$\begin{aligned} \int_a^b \int_c^d g(x, y) dy dx &= \int_a^b \int_c^d f_{xy}(x, y) dy dx \\ &= f(b, d) - f(a, d) - f(b, c) + f(a, c). \end{aligned}$$

For similar reasons (namely, that f_y is a generalized antiderivative of f_{yx} with respect to x and f is a generalized antiderivative of f_y with respect to y) we conclude that

$$\begin{aligned} \int_c^d \int_a^b g(x, y) dy dx &= \int_c^d \int_a^b f_{yx}(x, y) dx dy \\ &= f(b, d) - f(b, c) - f(a, d) + f(a, c). \end{aligned}$$

Hence,

$$\int_a^b \int_c^d g(x, y) dy dx = \int_c^d \int_a^b g(x, y) dx dy. \quad \blacksquare$$

We conclude this paper with two important remarks.

Remark 1. Let us return to D'Alembert's method for solving problem (1). First, we observe that Fubini's theorem for the class F of functions described by (i) is intuitive and easy to prove. It is natural to consider solutions of (1) that belong to F . By Theorem 6, Schwarz's lemma is valid for the class S of (ii). We can easily check that the function $u(x, t)$ provided by (2) satisfies all assumptions of (ii). Hence, the application of D'Alembert's method to problem (1) is legitimate.

Remark 2. Theorem 6 can be generalized in many ways. For example, in [2] the authors obtained it from the point of view of Schwartz distributions. However, the most intriguing possibility is probably the one based on the use of the Henstock integral. Many experts today advocate the replacement of the Riemann integral with the Henstock integral in regular calculus courses. One rationale for this is the well-known fact that the Henstock integral is more general than either the Riemann integral or the Lebesgue integral. Another is that Henstock integration does not require any measure theory. A third is the fundamental property expressed by Theorem 1.

In a recent paper E. Talvila [7] gave necessary and sufficient conditions for differentiating under the integral sign using the Henstock integral. The equivalence holds (see [7, Theorem 4]) when for almost all y in $[c, d]$ the function $f(\cdot, y)$ is absolutely continuous in the generalized sense (i.e., belongs to the class ACG_* for almost all y in $[c, d]$). Talvila mentions that the class ACG_* is contained in the class of functions that are differentiable almost everywhere and properly contains the class of functions that are differentiable except possibly on a countable set. In particular, ACG_* encompasses the class S .

We have already explained our reasons for choosing S and for using the Riemann integral. We should add that problems in the solution of the wave equation (as in (1)) to which Theorem 6 can be applied do not require any greater generality than this.

DEDICATION. This paper is dedicated to the memory of our dear friend Barbara Beechler.

REFERENCES

1. A. Aksoy and M. Martelli, Mixed partial derivatives and Fubini's theorem, *College Math. J.* **33** (2002) 126–130.
2. D. D. Ang, K. Schmitt, and L. K. Vy, A multidimensional analogue of the Denjoy-Perron-Henstock-Kurzweil integral, *Bull. Belg. Math. Soc. Simon Stevin* **4** (1997) 355–371.
3. M. W. Botsko, An elementary proof that a bounded a.e. continuous function is Riemann integrable, this MONTHLY **65** (1988) 249–252.
4. C. Goffman, *Real Functions*, Rinehart, New York, 1953.
5. R. A. Gordon, *The Integrals of Lebesgue, Denjoy, Perron, and Henstock*, Graduate Studies in Mathematics, vol. 4, American Mathematical Society, Providence, 1994.
6. M. Martelli and M. Dang, The derivative of a continuous nonconstant function, *Appl. Math. Let.* **7** (1994) 81–84.
7. E. Talvila, Necessary and sufficient conditions for differentiability under the integral sign, this MONTHLY **108** (2001) 544–549.

One Observation behind Two-Envelope Puzzles

Dov Samet, Iddo Samet, and David Schmeidler

1. TWO PUZZLES ON THE THEME “WHICH IS LARGER?” In two famous and popular puzzles a participant is required to compare two numbers of which she is shown only one. Although the puzzles have been discussed and explained extensively, no connection between them has been established in the literature. We show here that there is one simple principle behind these puzzles. In particular, this principle sheds new light on the paradoxical nature of the first puzzle.

According to this principle the ranking of several random variables must depend on at least one of them, except for the trivial case where the ranking is constant. Thus, in the nontrivial case there must be at least one variable the observation of which conveys information about the ranking.

A variant of the first puzzle goes back to the mathematician Littlewood [7], who attributed it to the physicist Schrödinger. See [6], [3], [2] and [1] for more detail on the historical background and for further elaboration on this puzzle. Here is the common version of the puzzle, as first appeared in [5]:

To switch or not to switch? There are two envelopes with money in them. The sum of money in one of the envelopes is twice as large as the other sum. Each of the envelopes is equally likely to hold the larger sum. You are assigned at random one of the envelopes and may take the money inside. However, before you open your envelope you are offered the possibility of switching the envelopes and taking the money inside the other one. It seems obvious that there is no point in switching: the situation is completely symmetric with respect to the two envelopes. The argument for switching is also simple. Suppose you open the envelope and find a sum x . Then, in the other envelope the sum is either $2x$ or $x/2$ with equal probabilities. Thus, the expected sum is