Caltech Division of the Humanities and Social Sciences

Differentiating an Integral: Leibniz' Rule

KC Border
Spring 2002
Revised December 2016
v. 2016.12.25::15.02

Both Theorems 1 and 2 below have been described to me as Leibniz' Rule.

1 The vector case

The following is a reasonably useful condition for differentiating a Riemann integral. The proof may be found in Dieudonné [6, Theorem 8.11.2, p. 177]. One thing you have to realize is that for Dieudonné a partial derivative can be taken with respect to a vector variable. That is, if $f : \boldsymbol{R}^{\mathrm{n}} \times \boldsymbol{R}^{\mathrm{m}}$ where a typical element of $\boldsymbol{R}^{\mathrm{n}} \times \boldsymbol{R}^{\mathrm{m}}$ is denoted (x, z) with $x \in \boldsymbol{R}^{\mathrm{n}}$ and $y \in \boldsymbol{R}^{\mathrm{m}}$. The partial derivative $D_x f$ is a Fréchet derivative with respect to x holding z fixed.

1 Theorem *Let $A \subset \boldsymbol{R}^{\mathrm{n}}$ be open, let $I = [a, b] \subset \boldsymbol{R}$ be a compact interval, and let f be a (jointly) continuous mapping of $A \times I$ into $\boldsymbol{R}$. Then*

$$g(x) = \int_a^b f(x, t)\, dt$$

is continuous in A.

If in addition, the partial derivative $D_x f$ exists and is (jointly) continuous on $A \times I$, then g is continuously differentiable on A and

$$g'(x) = \int_a^b D_x(x, t)\, dt.$$

The next, even more useful, result is listed as an exercise (fortunately with hint) by Dieudonné [6, Problem 8.11.1, p. 177].

2 Leibniz's Rule *Under the hypotheses of Theorem 1, let α and β be two continuously differentiable mappings of A into I. Let*

$$g(x) = \int_{\alpha(x)}^{\beta(x)} f(x, t)\, dt.$$

Then g is continuously differentiable on A and

$$g'(x) = \int_{\alpha(x)}^{\beta(x)} D_x f(x, t)\, dt + f\big(x, \beta(x)\big)\beta'(x) - f\big(x, \alpha(x)\big)\alpha'(x).$$

2 The measure space case

This section is intended for use with expected utility, where instead if integrating with respect to a real parameter t as in Theorem 1, we integrate over an abstract probability space. So let $(\Omega, \mathcal{F}, \mu)$ be a measure space, let $A \subset \boldsymbol{R}^{\mathrm{n}}$ be open. We are interested in the properties of a function $g \colon A \to \boldsymbol{R}$ defined by

$$g(x) = \int_\Omega f(x, \omega)\, d\mu(\omega). \tag{1}$$

We are particularly interested in when g is continuous or continuously differentiable. It seems clear that in order for g to be defined, the function f must be measurable in ω, and in order for g to stand a chance of being continuous, the function f needs to be continuous in x.

3 Definition *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, let A be a topological space. We say that a function $f \colon A \times \Omega \to \boldsymbol{R}$ is a* ***Carathéodory function*** *if for each $x \in A$ the mapping $\omega \mapsto f(x, \omega)$ is $\mathcal{F}$-measurable, and for each $\omega \in \Omega$ the mapping $x \mapsto f(x, \omega)$ is continuous. (Sometimes we say that f is continuous in x and measurable in ω.)*

In order for the function g defined by (1) to be finite-valued we need that for each x, the function $\omega \mapsto f(x, \omega)$ needs to be integrable. But this is not enough for our needs we need the following stronger property.

4 Definition *The function $f \colon A \times \Omega \to \boldsymbol{R}$ is* ***locally uniformly integrably bounded*** *if for every x there is a nonnegative measurable function $h_x \colon \Omega \to \boldsymbol{R}$ such that h_x is integrable, that is, $\int_\Omega h_x(\omega)\, dP(\omega) < \infty$, and there exists a neighborhood U_x of x such that for all*

$$\text{for all } y \in U_x, \quad \bigl|f(y, \omega)\bigr| \leqslant h_x(\omega).$$

Note that since $x \in U_x$, if f is locally uniformly integrably bounded, then we also have that $\omega \mapsto \bigl|f(x, \omega)\bigr|$ is integrable.

Note that if μ is a finite measure, and if f is bounded, then it is also locally uniformly integrably bounded. The next result may be found, for instance, in [2, Theorem 24.5, p. 193], Billingsley [4, Theorem 16.8, pp.181–182], or Cramér [5, ¶ II, p. 67–68].

5 Proposition *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, let $A \subset \boldsymbol{R}^{\mathrm{n}}$ be open, and let the function $f \colon A \times \Omega \to \boldsymbol{R}$ be a Carathéodory function. Assume further that f is locally uniformly integrably bounded. Then the function $g \colon A \to \boldsymbol{R}$ defined by*

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

is continuous.

Suppose further that for each i and each ω, the partial derivative $D_i f(x, \omega)$ with respect to x_i is a continuous function of x and $D_i f$ is locally uniformly integrably bounded. Then g is continuously differentiable and

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

Proof: First we deal with continuity. Since f is locally uniformly integrably bounded, for each x there is a nonnegative integrable function $h_x \colon \Omega \to \boldsymbol{R}$, and a neighborhood U_x of x such that for all $y \in U_x$, we have $\bigl|f(y, \omega)\bigr| \leqslant h_x(\omega)$. Then $\bigl|g(x)\bigr| \leqslant \int_\Omega h_x(\omega)\, d\mu(\omega) < \infty$. Now suppose

$x_n \to x$. Since f is continuous in x, $f(x_n, \omega) \to f(x, \omega)$ for each ω. Eventually x_n belongs to U_x, so for large enough n, $\bigl|f(x_n, \omega)\bigr| \leqslant h_x(\omega)$. Then by the Dominated Convergence Theorem,[1]

$$g(x_n) = \int_\Omega f(x_n, \omega)\, d\mu(\omega) \to \int_\Omega f(x, \omega)\, d\mu(\omega) = g(x).$$

That is, g is continuous.

For continuous differentiability, start by observing that $D_i f(x, \omega)$ is measurable in ω and hence a Carathéodory function. To see this, recall that

$$D_i f(x, \omega) = \lim_{t \to 0} \frac{f(x + te^i, \omega) - f(x, \omega)}{t}.$$

For each t, the difference quotient is a measurable function of ω, so its limit is measurable as well.

Assume that $D_i f(x, \omega)$ is uniformly bounded by the integrable $h_x(\omega)$ on a neighborhood U_x of x. Let e^i denote the i^{th} unit coordinate vector. By the Mean Value Theorem,[2] for each ω and for each nonzero t there is a point $\xi(t, \omega)$ belonging to the interior of the segment joining x and $x + te^i$ with

$$f(x + te^i, \omega) - f(x, \omega) = t D_i f\bigl(\xi(t, \omega), \omega\bigr).$$

Since both functions on the left hand side are measurable, the right-hand side is also a measurable function of ω. [3] For $|t|$ small enough, since $\xi(t, \omega)$ lies between x and $x + te^i$, we must have that $\xi(t, \omega) \in U_x$, so

$$\bigl|t D_i f\bigl(\xi(t, \omega), \omega\bigr)\bigr| \leqslant h_x(\omega).$$

Now

$$g(x + te^i) - g(x) = \int_\Omega f(x + te^i, \omega) - f(x, \omega)\, d\mu(\omega) = \int_\Omega t D_i f\bigl(\xi(t, \omega), \omega\bigr)\, d\mu(\omega).$$

As $t \to 0$, we have $\xi(t, \omega) \to x$, so $D_i\bigl(\xi(t, \omega), \omega\bigr) \to D_i f(x, \omega)$ for each ω. Dividing by t and applying the Dominated Convergence Theorem yields

$$D_i g(x) = \lim_{t \to 0} \frac{g(x + te^i) - g(x)}{t} = \int_\Omega D_i f(x, \omega)\, d\mu(\omega).$$

The proof of continuity of $D_i g$ is the same as the proof of continuity of g. ∎

3 An application to expected utility

The previous section dealt directly with a function f defined on the Cartesian product of a subset of $\boldsymbol{R}^{\mathrm{n}}$ and a measurable space Ω. In practice the dependence on Ω is often via a random vector, which allows for conditions that easier to understand. Here is a common application of these results. See, for instance, Hildreth [9], who refers the reader to Hildreth and Tesfatsion [10] for proofs.

[1] See, for example, Royden [12, Theorem 16, p. 267] or Aliprantis and Border [1, Theorem 11.21, p. 415].

[2] See, for instance, Apostol [3, Theorem 4.5, p. 185]. It is also sometimes known as Darboux's Theorem.

[3] In fact, by the Stochastic Taylor's Theorem 8 below we can show that $\omega \mapsto \xi(t, \omega)$ can be taken to be measurable with respect to ω. But that theorem requires a lot of high-powered machinery for its proof, and contrary to my initial instincts we don't need it for our purposes.

6 Corollary *Let I be an interval of the real line with interior I°, and let $u\colon I \to \boldsymbol{R}$ be strictly increasing, continuous, and concave on I, and twice continuously differentiable on I°. Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let $\boldsymbol{x}, \boldsymbol{y}\colon \Omega \to \boldsymbol{R}$ be measurable functions (random variables). Let A be an open interval of the real line, and assume that for all $\alpha \in A$ and almost all $\omega \in \Omega$ that*

$$\boldsymbol{x}(\omega) + \alpha \boldsymbol{y}(\omega) \in I^\circ.$$

In addition, assume that for each $\alpha \in A$ that

$$\int_\Omega \big| u\big(\boldsymbol{x}(\omega) + \alpha \boldsymbol{y}(\omega)\big)\big| \, dP(\omega) < \infty, \tag{ii}$$

$$\int_\Omega \big| u'\big(\boldsymbol{x}(\omega) + \alpha \boldsymbol{y}(\omega)\big)\boldsymbol{y}(\omega)\big| \, dP(\omega) < \infty, \tag{iii}$$

$$\int_\Omega \big| u''\big(\boldsymbol{x}(\omega) + \alpha \boldsymbol{y}(\omega)\big)\boldsymbol{y}^2(\omega)\big| \, dP(\omega) < \infty. \tag{iv}$$

Define the function

$$g(\alpha) = \int_\Omega u\big(\boldsymbol{x}(\omega) + \alpha \boldsymbol{y}(\omega)\big) \, dP(\omega).$$

Then g is continuously differentiable, and

$$g'(\alpha) = \int_\Omega u'\big(\boldsymbol{x}(\omega) + \alpha \boldsymbol{y}(\omega)\big)\boldsymbol{y}(\omega) \, dP(\omega). \tag{1}$$

If in addition u'' is (weakly) increasing,[4] *then g is twice continuously differentiable and*

$$g''(\alpha) = \int_\Omega u''\big(\boldsymbol{x}(\omega) + \alpha \boldsymbol{y}(\omega)\big)\boldsymbol{y}^2(\omega) \, dP(\omega). \tag{2}$$

Proof: Since u is concave, u' is (weakly) decreasing, and $u'' \leqslant 0$. It also follows that $u' > 0$ on I°.[5] Define $f\colon A \times \Omega \to \boldsymbol{R}$ by

$$f(\alpha, \omega) = u\big(\boldsymbol{x}(\omega) + \alpha \boldsymbol{y}(\omega)\big).$$

Then f is clearly a Carathéodory function. In order to apply Proposition 5, we need to show that f and $D_1 f$ are locally uniformly integrably bounded. So let $\bar\alpha \in A$ and choose $\delta > 0$ so that $A' = [\bar\alpha - \delta, \bar\alpha + \delta] \subset A$. Since u is strictly increasing,

$$\big|f(\alpha, \omega)\big| \leqslant \big|f(\bar\alpha - \delta, \omega)\big| + \big|f(\alpha + \delta, \omega)\big| = h_{\bar\alpha}(\omega)$$

for all $\alpha \in A'$. By (ii), $h_{\bar\alpha}$ is integrable. Thus f is uniformly locally integrably bounded, so g continuous.

Similarly, since u' is decreasing

$$\big|D_1 f(\alpha, \omega)\big| \leqslant \big|D_1 f(\bar\alpha - \delta, \omega)\big| + \big|D_1 f(\alpha + \delta, \omega)\big|$$

for all $\alpha \in A'$, so (iii) implies $D_1 f$ is uniformly locally integrably bounded and the same reasoning implies that g' is continuous and satisfies (1). You can now see how the remainder of the theorem is proven. ∎

[4]This condition is known as **prudence** in the expected utility literature, as it implies a desire to save more in the face of increased risk. For the purposes of twice differentiability of g, we could have assumed that u'' is weakly decreasing, but there is no convincing economic interpretation of that condition.

[5]Since u is strictly increasing, $u' \geqslant 0$ and it cannot attain a maximum on I°. But for concave u, the condition $u' = 0$ implies a maximizer. Thus $u' > 0$.

4 An illustrative (counter)example

To get an idea of what these conditions mean, consider the following example, taken from Gelbaum and Olmsted [7, Example 9.15, p. 123].

7 Example The following example shows what can go wrong when the hypotheses of the previous theorems are violated.

Define $f\colon \boldsymbol{R} \times \boldsymbol{R}_+ \to \boldsymbol{R}$ via

$$f(x,t) = \begin{cases} \dfrac{x^3}{t^2} e^{-x^2/t} & t > 0, \\ 0 & t = 0. \end{cases}$$

First observe that for fixed t the function $x \mapsto f(x,t)$ is continuous at each x, and for each fixed x the function $t \mapsto f(x,t)$ is continuous at each t, including $t = 0$. (This is because the exponential term goes to zero much faster than the polynomial term goes to zero as $t \to 0$.) The function is not jointly continuous though. On the curve $t = x^2$ we have $f(x,t) = e^{-1}/x$, which diverges to ∞ as $x \downarrow 0$ and diverges to $-\infty$ as $x \uparrow 0$. See Figure 1.

Define

$$\begin{aligned} g(x) &= \int_0^1 f(x,t)\,dt \\ &= x^3 \int_0^1 \frac{1}{t^2} e^{-x^2/t}\,dt. \end{aligned}$$

Consulting a table of integrals if necessary, we find the indefinite integral $\int \frac{1}{t^2} e^{-a/t}\,dt = e^{-a/t}/a$. Thus, letting $a = x^2$ we have

$$\boxed{g(x) = xe^{-x^2}}$$

This holds for all $x \in \boldsymbol{R}$. Consequently

$$\boxed{g'(x) = (1 - 2x^2)e^{-x^2}}$$

again for all x.

Now let's compute

$$\int_0^1 D_1 f(x,t)\,dt.$$

For $t = 0$, $f(x,t) = 0$ for all x, so $D_1 f(x,0) = 0$. For $t > 0$, we have

$$\begin{aligned} D_1 f(x,t) &= \frac{3x^2}{t^2} e^{-x^2/t} + \frac{x^3}{t^2} e^{-x^2/t}(-2x/t) \\ &= e^{-x^2/t}\Big(\frac{3x^2}{t^2} - \frac{2x^4}{t^3}\Big). \end{aligned}$$

So

$$\boxed{D_1 f(x,t) = \begin{cases} e^{-x^2/t}\Big(\frac{3x^2}{t^2} - \frac{2x^4}{t^3}\Big) & t > 0 \\ 0 & t = 0. \end{cases}}$$

Note that for fixed x the limit of $D_1 f(x,t)$ as $t \downarrow 0$ is zero, so for each fixed x, $D_1 f(x,t)$ is continuous in t. But again, along the curve $t = x^2$, we have $D_1 f(x,t) = e^{-1}\big(3x^{-2} - 2x^{-2}\big) = -e^{-1}/x^2$ which diverges to ∞ as $x \to 0$. Thus $D_1 f(x,t)$ is not continuous at $(0,0)$. See Figure 2.

The integral

$$I(x) = \int_0^1 e^{-x^2/t}\Big(\frac{3x^2}{t^2} - \frac{2x^4}{t^3}\Big)\,dt$$

satisfies $I(0) = 0$ and for $x > 0$ it can be computed as

$$\begin{aligned}\int_0^1 D_1 f(x,t)\,dt &= \int_0^1 e^{-x^2/t}\Big(\frac{3x^2}{t^2} - \frac{2x^4}{t^3}\Big)\,dt \\ &= 3x^2 \int_0^1 \frac{1}{t^2} e^{-x^2/t}\,dt - 2x^4 \int_0^1 \frac{1}{t^3} e^{-x^2/t}\,dt\end{aligned}$$

so dividing by $x^2 \neq 0$,

$$\begin{aligned}&= 3e^{-x^2/t}\Big|_{t=0}^{t=1} - 2e^{-x^2/t}\Big(1 + \frac{x^2}{t}\Big)\Big|_{t=0}^{t=1} \\ &= (1 - 2x^2)e^{-x^2}\end{aligned}$$

which holds for all $x > 0$.

Thus at $x = 0$, we have

$$g'(0) = (1 - 2\cdot 0^2)e^{-0^2} = 1 \neq 0 = I(0) = \int_0^1 D_1 f(0,t)\,dt.$$

The remarks above show that f and $D_1 f(x,t)$ fail to be continuous at $(0,0)$ so this example does not violate Leibniz' Rule. How does it compare to the hypotheses of Proposition 5?

In this example t plays the role of ω in Proposition 5, so locally uniform integrability requires that for each x there is an integrable function h_x and a neighborhood U_x such that $\sup_{y \in U_x} \big|D_1 f(y,t)\big| \leqslant h_x(t)$. Let's check this for $x = 0$. We need to find a $\delta > 0$ so that $|y| < \delta$ implies $|D_1 f(y,t)| \leqslant h_0(t)$. Now for $t > 0$,

$$D_1 f(y,t) = e^{-y^2/t}\Big(\frac{3y^2}{t^2} - \frac{2y^4}{t^3}\Big).$$

Looking at points of the form $y = \sqrt{t}$, we see that $h_0(t)$ must satisfy

$$h_0(t) \geqslant D_1 f(\sqrt{t}, t) = e^{-1}\Big(\frac{3}{t} - \frac{2}{t}\Big) = e^{-1}/t,$$

which is not integrable over any interval $(0, \varepsilon)$, so the hypotheses of Proposition 5 are also violated by this example. □

5 A Stochastic version of Taylor's Theorem

I used to think the following sort of result was necessary in the proof of Proposition 5, but I was wrong. But I spent a lot of effort figuring out the machinery needed to prove it, so I'm sharing it with you.

8 Stochastic Taylor's Theorem *Let $h\colon [a,b] \to \boldsymbol{R}$ be continuous and possess a continuous n^{th}-order derivative on (a,b). Fix $c \in [a,b]$ and let X be a random variable on the probability space $(S, \mathcal{S}, P)$ such that $c + X \in [a,b]$ almost surely. Then there is a (measurable) random variable ξ satisfying $\xi(s) \in [0, X(s)]$ for all s (where $[0, X(s)]$ is the line segment joining 0 and $X(s)$, regardless of the sign of $X(s)$), and*

$$h\bigl(c + X(s)\bigr) = h(c) + \sum_{k=1}^{n-1} \frac{1}{k!} h^{(k)}(c) X^k(s) + \frac{1}{n!} h^{(n)}\bigl(c + \xi(s)\bigr) X^n(s).$$

Proof: (See [1, Theorem 18.18, p. 603].) Taylor's Theorem without remainder (see, for instance, Landau [11, Theorem 177, p. 120] or Hardy [8, p. 286]) is a generalization of the Mean Value Theorem that asserts that there is such a $\xi(s)$ for each s, the trick is to show that there is a measurable version. To this end define the correspondence $\varphi\colon S \twoheadrightarrow \boldsymbol{R}$ by $\varphi(s) = [0, X(s)]$. It follows from [1, Theorem 18.5, p. 595] that φ is measurable and it clearly has compact values. Set $g(s) = h\bigl(c + X(s)\bigr) - h(c) - \sum_{k=1}^{n-1} \frac{1}{k!} h^{(k)}(c) X^k(s)$, $f(s,x) = \frac{1}{n!} h^{(n)}\bigl(c + x\bigr) X^n(s)$. Then g is measurable and f is a Carathéodory function. (See section 4.10 of [1] for the definition of measurable correspondences.) By Filippov's Implicit Function Theorem [1, Theorem 18.17, p. 603] there is a measurable function ξ such that for all s, $\xi(s) \in \varphi(s)$ and $f\bigl(s, \xi(s)\bigr) = g(s)$, and we are done. ∎

References

[1] C. D. Aliprantis and K. C. Border. 2006. *Infinite dimensional analysis: A hitchhiker's guide*, 3d. ed. Berlin: Springer–Verlag.

[2] C. D. Aliprantis and O. Burkinshaw. 1998. *Principles of real analysis*, 3d. ed. San Diego: Academic Press.

[3] T. M. Apostol. 1967. *Calculus*, 2d. ed., volume 1. Waltham, Massachusetts: Blaisdell.

[4] P. Billingsley. 1979. *Probability and measure*. Wiley Series in Probability and Mathematical Statistics. New York: Wiley.

[5] H. Cramér. 1946. *Mathematical methods of statistics*. Number 34 in Princeton Mathematical Series. Princeton, New Jersey: Princeton University Press. Reprinted 1974.

[6] J. Dieudonné. 1969. *Foundations of modern analysis*. Number 10-I in Pure and Applied Mathematics. New York: Academic Press. Volume 1 of Treatise on Analysis.

[7] B. R. Gelbaum and J. M. Olmsted. 2003. *Counterexamples in analysis*. Minneola, NY: Dover. Slightly corrected reprint of the 1965 printing of the work originally published in 1964 by Holden–Day in San Francisco.

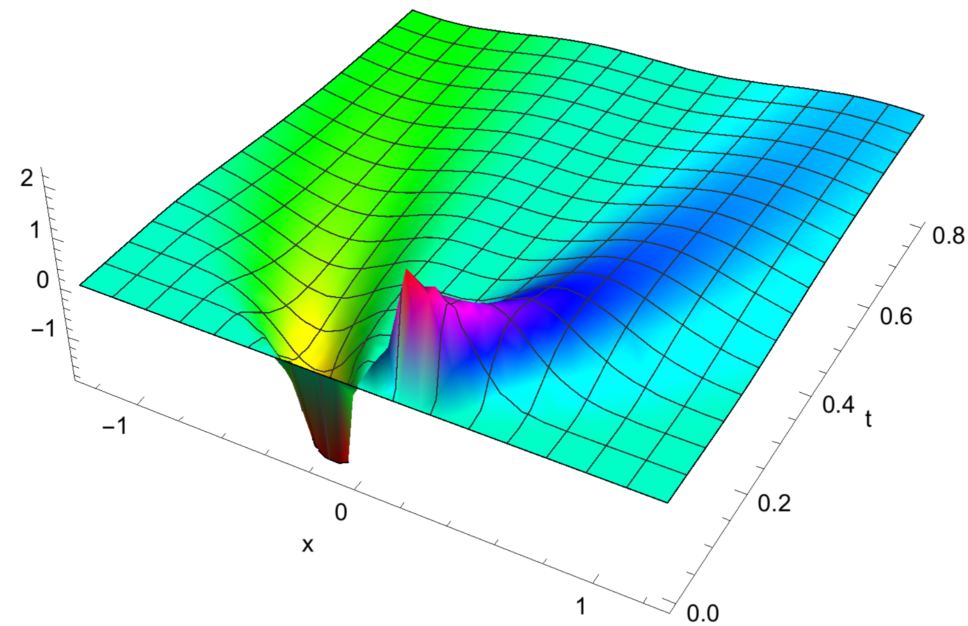
[8] G. H. Hardy. 1952. *A course of pure mathematics*, 10th. ed. Cambridge: Cambridge University Press.

[9] C. Hildreth. 1974. Expected utility of uncertain ventures. *Journal of the American Statistical Association* 69(345):9–17. http://www.jstor.org/stable/2285494.pdf

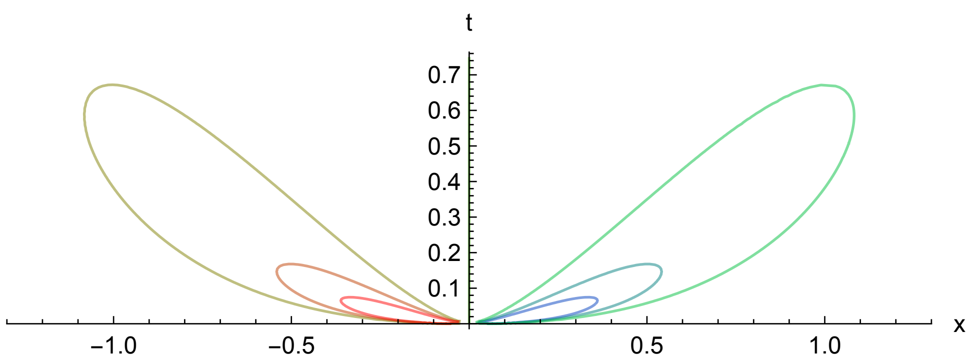
[10] C. Hildreth and L. Tesfatsion. 1974. A model of choice with uncertain initial prospect. Discussion Paper 38, Center for Economic Research, University of Minnesota. http://docs.google.com/a/umn.edu/viewer?a=v&pid=sites&srcid=dW1uLmVkdXxlY29ubGljcmFyeXxneDo3MT

[11] E. Landau. 1960. *Differential and integral calculus*, 2d. ed. New York: Chelsea. Translation of *Einführung in die Differentialrechnung und Integralrechnung*, published in 1934.

[12] H. L. Royden. 1988. *Real analysis*, 3d. ed. New York: Macmillan.

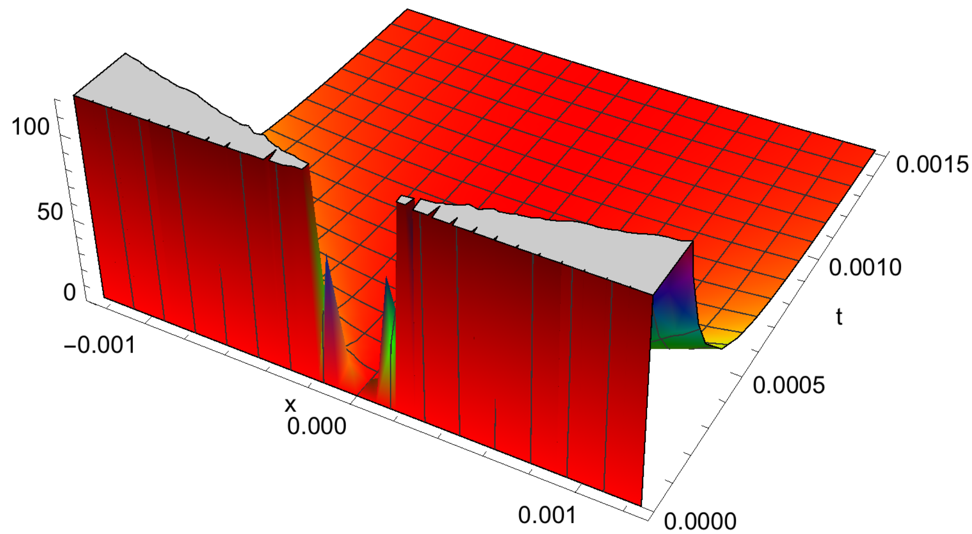


Surface of graph.

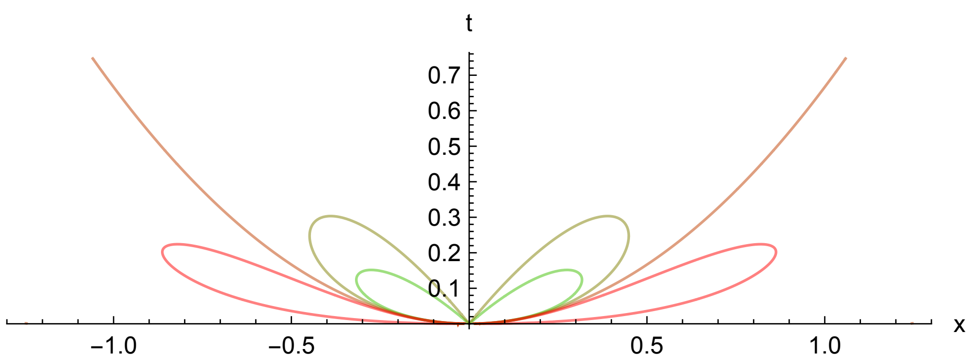


Contours.

Figure 1. Plots of $\dfrac{x^3}{t^2}e^{-x^2/t}$.



Surface of graph.



Contours.

Figure 2. Plots of $e^{-x^2/t}\left(\frac{3x^2}{t^2} - \frac{2x^4}{t^3}\right)$.