

A Differential Forms Perspective on the Lax Proof of the Change of Variables Formula

Author(s): Nikolai V. Ivanov

Source: *The American Mathematical Monthly*, Vol. 112, No. 9 (Nov., 2005), pp. 799-806

Published by: Mathematical Association of America

Stable URL: <http://www.jstor.org/stable/30037601>

Accessed: 04-11-2016 08:44 UTC

REFERENCES

Linked references are available on JSTOR for this article:

http://www.jstor.org/stable/30037601?seq=1&cid=pdf-reference#references_tab_contents

You may need to log in to JSTOR to access the linked references.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at

<http://about.jstor.org/terms>



Mathematical Association of America is collaborating with JSTOR to digitize, preserve and extend access to *The American Mathematical Monthly*

A Differential Forms Perspective on the Lax Proof of the Change of Variables Formula

Nikolai V. Ivanov

1. INTRODUCTION. In a beautiful paper [10], Peter Lax presented an elementary proof of a special case of the change of variables theorem. As explained in [10], this special case is sufficient to give a very simple proof of the Brouwer fixed-point theorem. In [11], Lax explained how one can deduce the general case of the change of variables theorem from this special case using some standard tools (e.g., partitions of unity) from the no-man's-land between advanced calculus and the three great "differential" theories (differential topology, differential geometry, ordinary differential equations), to paraphrase Serge Lang (see the foreword of [8]).

The first goal of this article is to present a differential forms version of the Lax proof of (the special case of) the change of variables formula. We have attempted to follow the Lax arguments as closely as possible. Special care was taken to be completely explicit about all results concerned with the integration of differential forms, since the usual expositions assume (or prove in a classical way) the change of variables formula at the very beginning of integration theory for differential forms. One of the exceptions is Lang's book [9], where differential aspects of the theory of differential forms are clearly separated from the integration (because derivatives make sense in infinite dimensions, and among the Lang's objectives is to work in infinite dimensions whenever possible; see [9, chap. 5, sec. 3]).

Our second goal is to present a fairly detailed comparison of our proof with Lax's. Such a comparison is very instructive, for it sheds light on both the efficiency of the differential form theory and the brilliance with which Lax uses classical analysis. To the extent possible, we retain Lax's notation and terminology.

A key role in the Lax proof is played by a fairly mysterious determinantal identity, which surfaces here as (3). It is almost invisible in the differential forms version of the proof. In fact, identity (3) plays a similar role in at least one other proof of the change of variables formula (namely, in the proof by Leinfelder and Simader [12]), and it is ubiquitous in analytical approaches to the Brouwer fixed-point theorem and related topics. With the exception of a 1910 paper of J. Hadamard [4], the original context in which this identity arose is never mentioned. In fact, identity (3) goes back to the Jacobi theory of multipliers for systems of ordinary differential equations [6]. This theory generalizes the well-known theory of integrating factors (due to Euler). While included in such classical treatises of analysis as [2] and [7], this theory apparently fell out favor sometime in the twentieth century.

In this paper we need only a small fragment of the theory of differential forms. All that is needed can be found, for example, in chapter 2 of Warner's textbook [16]. One might also suggest the much more elementary textbook by Edwards [1]. Other textbooks include Guillemin and Pollack [3] and Lang [9] (the last one is more abstract and advanced than the others, partially because it deals with the not-necessarily-finite dimensional situation from the outset). I would like also to recommend to the reader the nice article by H. Samelson in this MONTHLY [13] that summarizes both the theory of differential forms and its history (which contains some surprises).

A version of the Lax proof similar to ours was suggested by Michael Taylor [15]. He also uses differential forms, but his proof differs from Lax's in other respects as well. In contrast with the present article, Taylor aims at more general versions of the change of variables theorem, versions that assume less regularity on the part of the change of variables mapping. Our version is equally amenable to generalization, but the goals of this paper are purely expository.

2. THE LAX SET-UP. The special case of the change of variables formula considered by [10] deals with the following situation. Let $y = \varphi(x)$ be a mapping of n -dimensional x -space into n -dimensional y -space (i.e., let $\varphi : \mathbf{R}^n \rightarrow \mathbf{R}^n$). We make the following two assumptions:

- (i) φ is a C^1 -mapping (i.e., φ has continuous first-order partial derivatives everywhere in \mathbf{R}^n);
- (ii) φ is the identity outside some sphere, say the unit sphere (i.e., $\varphi(x) = x$ when $|x| \geq 1$).

Now we are ready to state the change of the variables theorem in the Lax formulation.

Theorem 1 (Change of Variables). *Let $f : \mathbf{R}^n \rightarrow \mathbf{R}^n$ be a continuous function of compact support. Then*

$$\int f(\varphi(x))J(x) dx = \int f(y) dy,$$

where J denotes the Jacobian determinant of the mapping φ (i.e.,

$$J(x) = \det \left[\frac{\partial \varphi_j}{\partial x_i}(x) \right],$$

in which φ_j is the j th component of φ .)

We start our proof with a simple lemma.

Lemma 1. *If $g : \mathbf{R}^n \rightarrow \mathbf{R}$ is a C^1 -function and $\psi = (\psi_1, \dots, \psi_n) : \mathbf{R}^n \rightarrow \mathbf{R}^n$ is a C^1 -mapping, then*

$$d(g \circ \psi) \wedge d\psi_1 \wedge \dots \wedge \widehat{d\psi_i} \wedge \dots \wedge d\psi_n = (-1)^{i-1} \frac{\partial g}{\partial y_i} \circ \psi \, d\psi_1 \wedge \dots \wedge d\psi_n,$$

for $i = 1, 2, \dots, n$, where the symbol $\widehat{}$ signifies an omitted term.

Proof. Note that, by the chain rule,

$$\begin{aligned} d(g \circ \psi) &= d(\psi^*(g)) = \psi^*(dg) \\ &= \psi^* \left(\sum_{j=1}^n \frac{\partial g}{\partial y_j} dy_j \right) \\ &= \sum_{j=1}^n \left(\frac{\partial g}{\partial y_j} \circ \psi \right) d\psi_j. \end{aligned}$$

(Here $\psi^*(\cdot)$ signifies the pull-back under ψ [16],[3].) Hence

$$\begin{aligned}
d(g \circ \psi) \wedge d\psi_1 \wedge \cdots \wedge \widehat{d\psi_i} \wedge \cdots \wedge d\psi_n \\
&= \left(\sum_{j=1}^n \left(\frac{\partial g}{\partial y_j} \circ \psi \right) d\psi_j \right) \wedge d\psi_1 \wedge \cdots \wedge \widehat{d\psi_i} \wedge \cdots \wedge d\psi_n \\
&= (-1)^{i-1} \left(\frac{\partial g}{\partial y_i} \circ \psi \right) d\psi_1 \wedge \cdots \wedge d\psi_n,
\end{aligned}$$

because $d\psi_j \wedge (d\psi_1 \wedge \cdots \wedge \widehat{d\psi_i} \wedge \cdots \wedge d\psi_n) = 0$ when $j \neq i$. The lemma follows. (Note that the sign $(-1)^{i-1}$ arises from the need to move $d\psi_i$ from the first position to the i th.) ■

Corollary 1. *If $h : \mathbf{R}^n \rightarrow \mathbf{R}$ is a C^1 -function, then*

$$\begin{aligned}
d(h dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_n) &= dh \wedge dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_n \\
&= (-1)^{i-1} \frac{\partial h}{\partial x_i} dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n
\end{aligned}$$

for $i = 1, 2, \dots, n$.

Proof. The first equality is simply the definition of the exterior derivative. In order to prove the second one, just apply the Lemma 1 with $g = h$ and $\psi : \mathbf{R}^n \rightarrow \mathbf{R}^n$ equal to the identity map. ■

We will need the following special case of Stokes's theorem. Let $c > 0$, and let $I = I_c$ be the standard c -cube in n -dimensional x -space (i.e., I is given by the inequalities $|x_i| \leq c$ for $i = 1, 2, \dots, n$).

Theorem 2 (Stokes's Theorem). *Let ω be a smooth $(n - 1)$ -form on \mathbf{R}^n , say*

$$\omega = \sum_{i=1}^n h_i dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_n,$$

in which the h_i are C^1 -functions. Then

$$\int_I d\omega = \int_{\partial I} \omega,$$

where ∂I signifies the boundary of I .

The integrals here can be understood in a naive sense, without any recourse to the general theory of integration of differential forms. First of all, $d\omega$ is an n -form on \mathbf{R}^n , and the integral $\int_I \psi$ of an n -form $\psi = g dx_1 \wedge \cdots \wedge dx_n$ can be defined simply as the usual volume integral $\int_I g$. In order to avoid any discussion of the orientation of I or of the induced orientation of ∂I , we interpret the integral $\int_{\partial I} \omega$ simply as shorthand for

$$\sum_{i=1}^n (-1)^{i-1} \left(\int_{I_i^+} h_i - \int_{I_i^-} h_i \right),$$

where I_i^+ and I_i^- are the faces of the cube I given by the equations $x_i = c$ and $x_i = -c$, respectively. Note that the integral defined in such a way is clearly linear with respect

to the addition of differential forms, since each of terms involved is a standard $(n - 1)$ -dimensional integral.

The possibility of restricting ourselves to such an elementary and limited version of the Stokes theorem is a great advantage of the Lax approach. After this simple case is used to prove Theorem 1 and the change of variables formula is extended to the general case [11], one can return to the theory of integration of differential forms in its full generality (where the strengthened version of the change of variables formula is needed) and easily prove the Stokes theorem in complete generality (say, for manifolds with boundary instead of I , if one already has the notion of a manifold at hand).

Proof of Theorem 2. It is sufficient to deal with the different summands of ω separately. Let $\omega_i = h_i dx_1 \wedge \cdots \wedge \widehat{dx}_i \wedge \cdots \wedge dx_n$. By Corollary 1,

$$d\omega_i = (-1)^{i-1} \frac{\partial h_i}{\partial x_i} dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n.$$

Hence,

$$\begin{aligned} \int_I d\omega &= \int_I (-1)^{i-1} \frac{\partial h_i}{\partial x_i} = (-1)^{i-1} \int \left(\int \frac{\partial h_i}{\partial x_i} dx_i \right) dx_1 \dots \widehat{dx}_i \dots dx_n \\ &= (-1)^{i-1} \left(\int_{I_i^+} h_i dx_1 \dots \widehat{dx}_i \dots dx_n - \int_{I_i^-} h_i dx_1 \dots \widehat{dx}_i \dots dx_n \right) \\ &= (-1)^{i-1} \left(\int_{I_i^+} h_i - \int_{I_i^-} h_i \right) = \int_{\partial I} \omega_i. \end{aligned}$$

The last step uses our definition of the integral $\int_{\partial I}$. Earlier in the computation we used the Fubini theorem (twice) and the fundamental theorem of calculus. (Of course, the Stokes theorem is simply the multidimensional form of the fundamental theorem of calculus.) This proves the theorem. ■

Remark. The foregoing proof is very close to the one found in [9, chap. 17, sec. 1].

Now we are ready for the proof of the change of variables formula.

Proof of Theorem 1. It is sufficient to prove the theorem for C^1 -functions f and for C^2 -mappings φ , since functions and mappings can be approximated in the relevant norms by C^1 -functions and C^2 -mappings, respectively (see [10]). Following Lax, we define $g : \mathbf{R}^n \rightarrow \mathbf{R}$ by

$$g(y_1, y_2, \dots, y_n) = \int_{-\infty}^{y_1} f(z, y_2, \dots, y_n) dz.$$

The integral is well defined because f has compact support and is of class C^1 . Clearly, $\partial g / \partial y_1 = f$. The function g is of class C^1 . Fix $c > 0$ and let $I = [-c, c]^n$. We can choose c to be so large that both the support of f and the unit ball $B = \{y : |y| \leq 1\}$ are contained in I . Then $g(y_1, \dots, y_n) = 0$ when $|y_j| \geq c$ for any $j \neq 1$ and when $y_1 \leq -c$. In addition, since φ agrees with the identity outside of B , $f(\varphi(x))$ vanishes outside I . It follows that we can restrict the integration in the theorem to I .

First, notice that

$$\begin{aligned} d\varphi_1 \wedge \cdots \wedge d\varphi_n &= \sum_{j=1}^n \frac{\partial \varphi_1}{\partial x_j} dx_j \wedge \cdots \wedge \sum_{j=1}^n \frac{\partial \varphi_n}{\partial x_j} dx_j \\ &= \left(\det \left[\frac{\partial \varphi_i}{\partial x_j} \right] \right) dx_1 \wedge \cdots \wedge dx_n = J dx_1 \wedge \cdots \wedge dx_n \end{aligned}$$

in view of the basic relation between determinants and top-dimensional exterior products and the definition of the Jacobian determinant. Therefore, the integrand in the left-hand side of the change of variables formula computes as

$$\begin{aligned} f(\varphi(x))J(x) dx_1 \wedge \cdots \wedge dx_n &= f(\varphi(x)) d\varphi_1 \wedge \cdots \wedge d\varphi_n \\ &= \left(\frac{\partial g}{\partial y_1} \circ \varphi \right) (x) d\varphi_1 \wedge \cdots \wedge d\varphi_n. \end{aligned}$$

By Lemma 1 (applied to the case $i = 1$), the last expression is equal to

$$d(g \circ \varphi) \wedge d\varphi_2 \wedge \cdots \wedge d\varphi_n.$$

Now, obviously,

$$d(g \circ \varphi) \wedge d\varphi_2 \wedge \cdots \wedge d\varphi_n = d(g \circ \varphi) d\varphi_2 \wedge \cdots \wedge d\varphi_n. \quad (1)$$

(Here we have implicitly invoked our assumption that φ is a C^2 -mapping.)

Next, we conclude by integrating that

$$\int_I f(\varphi(x))J(x) dx_1 \wedge \cdots \wedge dx_n = \int_I d(g \circ \varphi) d\varphi_2 \wedge \cdots \wedge d\varphi_n.$$

By Stokes's theorem the last expression is equal to

$$\int_{\partial I} g \circ \varphi d\varphi_2 \wedge \cdots \wedge d\varphi_n.$$

Notice that the boundary ∂I is entirely contained in the domain where φ coincides with the identity mapping. Hence the last integral becomes

$$\begin{aligned} \int_{\partial I} g dy_2 \wedge \cdots \wedge dy_n &= \int_{I_1^+} g dy_2 \cdots dy_n \\ &= \int_{I_1^+} \int_{-\infty}^c f dy_1 dy_2 \cdots dy_n \\ &= \int_{I_1^+} \int_{-c}^c f dy_1 dy_2 \cdots dy_n \\ &= \int_I f dy_1 dy_2 \cdots dy_n = \int f(y) dy. \end{aligned}$$

In the first equality we have exploited the fact that g vanishes on all faces of I with the exception of I_1^+ . This completes the proof of Theorem 1. ■

3. COMPARISON WITH THE LAX PROOF. In this section we compare our proof with Lax's. The reader might want to have the MONTHLY issue with the Lax article [10] at hand, in order to appreciate this discussion fully. At the same time, most of it can be understood independently of [10]. First of all, our proof is based on the same key idea, namely, introducing the function g . We also use the same tool of restricting integration to a large cube I . Lemma 1 in the special case $i = 1$ is the observation of Lax (see [10, p. 489]) expressed in the language of differential forms, with his φ corresponding to our ψ . Note that the standing assumption on φ in [10] (the same as our assumptions on φ) plays no role in this observation (except for the differentiability). We proved Lemma 1 for all i because this is no more difficult than the case $i = 1$, and it provided us with Corollary 1, which turned out to be a key ingredient in our proof of Stokes's theorem.

One may observe that Lax does not use the Stokes theorem, at least not explicitly. Instead he uses integration by parts, which is not used in our proof in an explicit manner, and a determinantal identity (formula (2.10) on page 499 in [10]).

Recall that in the calculus of one variable integration by parts is nothing other than a combination of the Leibniz rule for differentiating a product with the fundamental theorem of calculus. The Leibniz rule naturally generalizes to the calculus of differential forms, where it evolves into the following formula:

$$d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^a \omega_1 \wedge d\omega_2,$$

in which a is the degree of the form ω_1 . The multivariate analogue of the fundamental theorem of calculus is the Stokes theorem, as we have already pointed out. The differential form version of integration by parts in the case of integration over I (the only case we need) is the following computation:

$$\int_I d\omega_1 \wedge \omega_2 + (-1)^a \omega_1 \wedge d\omega_2 = \int_I d(\omega_1 \wedge \omega_2) = \int_{\partial I} \omega_1 \wedge \omega_2.$$

Now, it might appear that we have not used the Leibniz formula either. In fact, it is hidden in formula (1). To be more precise, we can deduce (1) as follows. Observe that

$$d(g \circ \varphi \, d\varphi_2 \wedge \cdots \wedge d\varphi_n) = d(g \circ \varphi) \wedge d\varphi_2 \wedge \cdots \wedge d\varphi_n + g \circ \varphi \, d(d\varphi_2 \wedge \cdots \wedge d\varphi_n)$$

by the Leibniz formula with ω_1 equal to the 0-form $g \circ \varphi$ and ω_2 equal to the form $d\varphi_2 \wedge \cdots \wedge d\varphi_n$. The second summand vanishes because

$$d(d\varphi_2 \wedge \cdots \wedge d\varphi_n) = 0, \tag{2}$$

so (1) follows. From the point of view of differential forms, (2) is obvious. A formal proof follows from the Leibniz rule for an $(n - 1)$ -fold product of forms and the fact that $d(d\varphi_i) = 0$ (here we need φ_i to be of class C^2). A more geometric proof is presented in [5] (see the proof of the lemma therein). Expressed in classical language, formula (2) turns into a not quite trivial and fairly mysterious determinantal identity. Any $(n - 1)$ -form $d\varphi_2 \wedge \cdots \wedge d\varphi_n$ can be written as

$$d\varphi_2 \wedge \cdots \wedge d\varphi_n = \sum_{i=1}^n A_i \, dx_1 \wedge \cdots \wedge \widehat{dx}_i \wedge \cdots \wedge dx_n$$

for suitable functions A_i . (This formula corresponds to expanding the determinant (2.5) in the Lax proof along its first column.) Then

$$d(d\varphi_2 \wedge \cdots \wedge d\varphi_n) = \sum_{i=1}^n (-1)^{i-1} \frac{\partial A_i}{\partial x_i} dx_1 \wedge \cdots \wedge dx_n$$

(by Corollary 1 applied to functions A_i in the role of h). As we saw, the left-hand side of this equation is 0, and this implies the identity

$$\sum_{i=1}^n (-1)^{i-1} \frac{\partial A_i}{\partial x_i} = 0. \quad (3)$$

The A_i are, in fact, determinants of certain matrices of partial derivatives of φ . More precisely, A_i is i th minor (of size $(n-1) \times (n-1)$) of the $(n-1) \times n$ matrix $(\partial\varphi_i/\partial x_j)$, where $2 \leq i \leq n$ and $1 \leq j \leq n$, as the reader can easily check. This means that $(-1)^{i-1}A_i$ is nothing but the cofactor M_i that appears in Lax's paper, and our identity is equivalent to his determinantal identity (2.10):

$$\partial_{x_1} M_1 + \cdots + \partial_{x_n} M_n \equiv 0.$$

To sum up, formula (1) hides within it a special case of the Leibniz rule and the determinantal identity (3). Integrating this formula and applying the Stokes theorem to the result amounts to combining integration by parts with the Lax determinantal identity (2.10).

The remainder of each proof—namely, the computation of

$$\int_{\partial I} g dy_2 \wedge \cdots \wedge dy_n$$

in our proof and the computation of the boundary term in the integration by parts formula in the Lax proof [10]—are exactly the same.

ACKNOWLEDGMENTS. I would like to thank Peter Lax, Hans Samelson, and David Simms for a very interesting correspondence about my note [5], which is concerned with closely related issues in the context of the Dunford-Schwartz proof of the Brouwer fixed-point theorem. This correspondence provided me with a stimulus to look more deeply into related problems. I am grateful to Michael Taylor for sending me the preprint version of [15] before its publication and directing me to his book [14]. I would also like to thank the anonymous referees for many suggestions on how to improve this article. This research was supported in part by NSF Grants DMS-9704817 and DMS-0406946.

REFERENCES

1. H. Edwards, *Advanced Calculus: A Differential Forms Approach*, 2nd ed., Birkhäuser, Boston, 1994.
2. E. Goursat, *Cours d'analyse mathématique, septième édition*, revue et augmentée par J. Favard, *Tome II*, Gauthier-Villars, Paris, 1949.
3. V. Guillemin and A. Pollack, *Differential Topology*, Prentice-Hall, Englewood Cliffs, NJ, 1974.
4. J. Hadamard, Sur quelques applications de l'indice de Kronecker, an appendix to J. Tannery, in *Introduction à la théorie des fonctions d'une variable*, vol. 2, A. Hermann, Paris, 1910, pp. 437-477.
5. N. V. Ivanov, A topologist's view of the Dunford-Schwartz proof of the Brouwer fixed-point theorem, *Math. Intelligencer* **22** (3) (2000) 55-57.
6. C. G. J. Jacobi, Teoria novi multiplicatoris systemati aequationum differentialium vulgarium applicandi, in *Gesammelte Werke*, vol. 4, Chelsea, New York, 1969, pp. 317-509.
7. C. Jordan, *Cours d'analyse*, vol. 3, Gauthier-Villars, Paris, 1887.
8. S. Lang, *Differential Manifolds*, Addison-Wesley, Reading, MA, 1972.
9. ———, *Fundamentals of Differential Geometry*, Graduate Texts in Mathematics, no. 191, Springer-Verlag, New York, 1998.
10. P. D. Lax, Change of variables in multiple integrals, this MONTHLY **106** (1999) 497-501.

11. ———, Change of variables in multiple integrals II, this MONTHLY **108** (2001) 115–119.
12. H. Leinfelder and Ch. G. Simader, The Brouwer fixed point theorem and the transformation rule for multiple integrals via homotopy arguments, *Expositiones Math.* **4** (1983) 349–355.
13. H. Samelson, Differential forms, the early days; or the stories of Deahna's theorem and of Volterra's theorem, this MONTHLY **108** (2001) 522–530.
14. M. E. Taylor, *Partial Differential Equations, I. Basic Theory*, Applied Mathematical Sciences, no. 115, Springer-Verlag, New York, 1996.
15. ———, Differential forms and the change of variable formula for multiple integrals, *J. Math. Anal. Appl.* **268** (2002) 378–383.
16. F. Warner, *Foundations of Differentiable Manifolds and Lie Groups*, 2nd ed., Graduate Texts in Mathematics, no. 94, Springer-Verlag, New York, 1983.

NIKOLAI IVANOV graduated from the Leningrad State University in 1976 and got his Ph.D. from the Steklov Mathematical Institute in 1980, under the direction of V. A. Rokhlin. In 1988 he received the Doctor of Sciences degree (something similar to the Habilitation in West Europe), again from the Steklov Mathematical Institute. From 1979 till 1998 he was a researcher at the Leningrad-branch of the Steklov Mathematical Institute. Since 1991 he has been at Michigan State University, first as a visitor, then after a half year break spent at Duke University, as a permanent faculty member. He kept his position at the Steklov Institute for a while, but, not surprisingly, eventually lost it. His research interests include low-dimensional topology, mapping class groups, Teichmüller spaces, and quasiconformal maps, bridging topology and analysis.
Michigan State University, Department of Mathematics, Wells Hall, East Lansing, MI 48824-1027
ivanov@math.msu.edu

Phyllis explained to him, trying to give of her deeper self, “Don’t you find it so beautiful, math? Like an endless sheet of gold chains, each link locked into the one before it, the theorems and functions, one thing making the next inevitable. It’s music, hanging there in the middle of space, meaning nothing but itself, and so *moving*...”

———John Updike, *Villages*, Alfred A. Knopf, New York, 2004, pp. 93–94;
 submitted by Henry Ricardo, Medgar Evers College (CUNY)