

On the Change of Variables Formula for Multiple Integrals

Shibo Liu^{1,*} and Yashan Zhang²

¹ Department of Mathematics, Xiamen University, Xiamen 361005, P.R. China;

² Department of Mathematics, University of Macau, Macau, P.R. China.

Received January 7, 2017; Accepted (revised) May 17, 2017

Abstract. We develop an elementary proof of the change of variables formula in multiple integrals. Our proof is based on an induction argument. Assuming the formula for $(m-1)$ -integrals, we define the integral over hypersurface in \mathbb{R}^m , establish the divergent theorem and then use the divergent theorem to prove the formula for m -integrals. In addition to its simplicity, an advantage of our approach is that it yields the Brouwer Fixed Point Theorem as a corollary.

AMS subject classifications: 26B15, 26B20

Key words: Change of variables, surface integral, divergent theorem, Cauchy-Binet formula.

1 Introduction

The change of variables formula for multiple integrals is a fundamental theorem in multivariable calculus. It can be stated as follows.

Theorem 1.1. *Let D and Ω be bounded open domains in \mathbb{R}^m with piece-wise C^1 -boundaries, $\varphi \in C^1(\bar{\Omega}, \mathbb{R}^m)$ such that $\varphi: \Omega \rightarrow D$ is a C^1 -diffeomorphism. If $f \in C(\bar{D})$, then*

$$\int_D f(y) dy = \int_{\Omega} f(\varphi(x)) |J_{\varphi}(x)| dx, \quad (1.1)$$

where $J_{\varphi}(x) = \det \varphi'(x)$ is the Jacobian determinant of φ at $x \in \Omega$.

The usual proofs of this theorem that one finds in advanced calculus textbooks involves careful estimates of volumes of images of small cubes under the map φ and numerous annoying details. Therefore several alternative proofs have appeared in recent years. For example, in [5] P. Lax proved the following version of the formula.

*Corresponding author. Email addresses: liusb@xmu.edu.cn (S. Liu), colourful2009@163.com (Y. Zhang)

Theorem 1.2. Let $\varphi: \mathbb{R}^m \rightarrow \mathbb{R}^m$ be a C^1 -map such that $\varphi(x) = x$ for $|x| \geq R$, and $f \in C_0(\mathbb{R}^m)$. Then

$$\int_{\mathbb{R}^m} f(y) dy = \int_{\mathbb{R}^m} f(\varphi(x)) J_\varphi(x) dx.$$

The requirement that φ is an identity map outside a big ball is somewhat restricted. This restriction was also removed by Lax in [6]. Then, Taylor [7] and Ivanov [4] presented different proofs of the above result of Lax [5] using differential forms. See also Báez-Duarte [1] for a proof of a variant of Theorem 1.1 which does not require that $\varphi: \Omega \rightarrow D$ is a diffeomorphism. As pointed out by Taylor [7, Page 380], because the proof relies on integration of differential forms over manifolds and Stokes' theorem, it requires that one knows the change of variables formula as formulated in our Theorem 1.1.

In this paper, we will present a simple elementary proof of Theorem 1.1. Our approach does not involve the language of differential forms. The idea is motivated by Exercise 15 of §1-7 in the famous textbook on classical differential geometry [3] by do Carmo. The exercise deals with the two dimensional case $m = 2$. We will perform an induction argument to generalize the result to the higher dimensional case $m \geq 2$. In our argument, we will apply the Cauchy-Binet formula about the determinant of the product of two matrices. As a byproduct of our approach, we will also obtain the Non-Retraction Lemma (see Corollary 3.2), which implies the Brouwer Fixed Point Theorem.

2 Integral over hypersurface

We will prove Theorem 1.1 by an induction argument. The case $m = 1$ is easily proved in single variable calculus. Suppose we have proven Theorem 1.1 for $(m-1)$ -dimension, where $m \geq 2$. We will define the integral over a hypersurface (of codimension one) in \mathbb{R}^m and establish the divergent theorem in \mathbb{R}^m . Then, in the next two sections we will use the divergent theorem to prove Theorem 1.1 for m -dimension.

Let U be a Jordan measurable bounded closed domain in \mathbb{R}^{m-1} , $x: U \rightarrow \mathbb{R}^m$,

$$(u^1, \dots, u^{m-1}) \mapsto (x^1, \dots, x^m)$$

be a C^1 -map such that the restriction of x in the interior U° is injective, and

$$\text{rank} \left(\frac{\partial x^i}{\partial u^j} \right) = m-1, \quad (2.1)$$

then we say that $x: U \rightarrow \mathbb{R}^m$ is a C^1 -parametrized surface. By definition, two C^1 -parametrized surfaces $x: U \rightarrow \mathbb{R}^m$ and $\tilde{x}: \tilde{U} \rightarrow \mathbb{R}^m$ are equivalent if there is a C^1 -diffeomorphism $\phi: \tilde{U} \rightarrow U$ such that $\tilde{x} = x \circ \phi$. The equivalent class $[x]$ is called a *hypersurface*, and $x: U \rightarrow \mathbb{R}^m$ is called a parametrization of the hypersurface. Since it is easy to see that $x(U) = \tilde{x}(\tilde{U})$ if x and \tilde{x} are equivalent, $[x]$ can be identified as the subset $S = x(U)$.

Let S be a hypersurface with parametrization $x: U \rightarrow \mathbb{R}^m$. By (2.1), for $u \in U$,

$$N(u) = \left(\frac{\partial(x^2, \dots, x^m)}{\partial(u^1, \dots, u^{m-1})}, \dots, (-1)^{m+1} \frac{\partial(x^1, \dots, x^{m-1})}{\partial(u^1, \dots, u^{m-1})} \right) \neq 0, \quad (2.2)$$

where

$$\frac{\partial(x^1, \dots, \hat{x}^i, \dots, x^m)}{\partial(u^1, \dots, u^{m-1})} = \det \begin{pmatrix} \partial_{u^1} x^1 & \cdots & \partial_{u^{m-1}} x^1 \\ \vdots & & \vdots \\ \partial_{u^1} x^{i-1} & \cdots & \partial_{u^{m-1}} x^{i-1} \\ \partial_{u^1} x^{i+1} & \cdots & \partial_{u^{m-1}} x^{i+1} \\ \vdots & & \vdots \\ \partial_{u^1} x^m & \cdots & \partial_{u^{m-1}} x^m \end{pmatrix}.$$

It is well known that $N(u)$ is a normal vector of S at $x(u)$.

Now, we can define the *surface integral* of a continuous function $f: S \rightarrow \mathbb{R}$ by

$$\int_S f d\sigma = \int_U f(x(u)) |N(u)| du. \quad (2.3)$$

By the change of variables formula for $(m-1)$ -integrals, it is not difficult to see that if $\tilde{x}: \tilde{U} \rightarrow \mathbb{R}^m$ is another parametrization of S , then

$$\int_U f(x(u)) |N(u)| du = \int_{\tilde{U}} f(\tilde{x}(v)) |\tilde{N}(v)| dv,$$

where \tilde{N} is defined similar to (2.2). Therefore, our surface integral is well defined.

If $\Sigma = \bigcup_{i=1}^{\ell} S_i$, where $S_i = x_i(U_i)$ are hypersurfaces such that $x_i(U_i^\circ) \cap x_j(U_j^\circ) = \emptyset$ for $i \neq j$, then we call Σ a piece-wise C^1 -hypersurface and define the integral of $f \in C(\Sigma)$ by

$$\int_{\Sigma} f d\sigma = \sum_{i=1}^{\ell} \int_{S_i} f d\sigma.$$

Theorem 2.1 (Divergent Theorem). *Let D be bounded open domain in \mathbb{R}^m with piece-wise C^1 -boundary ∂D , $F: \bar{D} \rightarrow \mathbb{R}^m$ be a C^1 -vector field, n is the unit outer normal vector field on ∂D , then*

$$\int_D \operatorname{div} F dx = \int_{\partial D} F \cdot n d\sigma.$$

Proof. Having defined the surface integral, the proof of the theorem is a standard application of the Fubini Theorem. We include the details here for completeness.

We say that $F = (F^1, \dots, F^m)$ is of i -type if $F^j = 0$ for $j \neq i$. We also say that D is of i -type, if there are a bounded closed domain U in \mathbb{R}^{m-1} with piece-wise C^1 -boundary and $\varphi_{\pm} \in C^1(U)$ such that

$$D = \left\{ x \mid \varphi_-(x') < x^i < \varphi_+(x'), x' \in U^\circ \right\},$$

where $x' = (x^1, \dots, x^{i-1}, x^{i+1}, \dots, x^m)$.

Let $F = (0, \dots, 0, F^m)$ be an m -type vector field. Suppose D is of m -type with U and φ_{\pm} as above. Then ∂D consists of three parts:

$$\Sigma_{\pm} = \{x = (x', \varphi_{\pm}(x')) \mid x' \in U\}$$

and

$$\Sigma_0 = \{x = (x', x^m) \mid \varphi_-(x') \leq x^m \leq \varphi_+(x'), x' \in \partial U\}.$$

On Σ_{\pm} , by (2.2) we have

$$N = (-1)^{m+1}(-\partial_1 \varphi_{\pm}, \dots, -\partial_{m-1} \varphi_{\pm}, 1).$$

Hence $|N| = \sqrt{1 + |\nabla \varphi_{\pm}|^2}$ and

$$n = \pm \frac{1}{\sqrt{1 + |\nabla \varphi_{\pm}|^2}}(-\partial_1 \varphi_{\pm}, \dots, -\partial_{m-1} \varphi_{\pm}, 1).$$

While on Σ_0 , $n = (-, -, 0)$ and $F \cdot n = 0$. Consequently, by (2.3) we obtain

$$\begin{aligned} \int_{\partial D} F \cdot n d\sigma &= \int_{\Sigma_+} F \cdot n d\sigma + \int_{\Sigma_-} F \cdot n d\sigma + \int_{\Sigma_0} F \cdot n d\sigma \\ &= \int_U F^m(x', \varphi_+(x')) dx' - \int_U F^m(x', \varphi_-(x')) dx' \\ &= \int_U dx' \int_{\varphi_-(x')}^{\varphi_+(x')} \partial_m F^m(x', t) dt = \int_D \partial_m F^m(x) dx = \int_D \operatorname{div} F dx. \end{aligned}$$

In a similar maner we can show that the theorem is valid for i -type vector field on i -type domain.

As in most calculus textbooks, we only prove the theorem for the case that D is simultaneously i -type for all $i = 1, \dots, m$. For a general C^1 -vector field $F = (F^1, \dots, F^m)$ on \bar{D} , we set $F_i = (0, \dots, F^i, \dots, 0)$. Since $F = F_1 + \dots + F_m$, and F_i is i -type vector field on i -type domain D , we deduce

$$\int_{\partial D} F \cdot n d\sigma = \sum_{i=1}^m \int_{\partial D} F_i \cdot n d\sigma = \sum_{i=1}^m \int_D \operatorname{div} F_i dx = \int_D \operatorname{div} F dx. \quad \square$$

3 Domains with singly parametrized boundary

In this section, we prove the m -dimensional change of variables formula (1.1) for the case that $\partial\Omega$ can be *singly parametrized*, that is, there exists a C^1 -parametrized surface $x: U \rightarrow \mathbb{R}^m$ such that $\partial\Omega = x(U)$. For example, if Ω is a ball, then $\partial\Omega$ can be singly parametrized by the well known parametrization.

In this case, we only need to require that the transformation φ maps $\partial\Omega$ to ∂D diffeomorphically. We have the following theorem.

Theorem 3.1. Let D and Ω be bounded open domains in \mathbb{R}^m with C^1 -boundaries, $\partial\Omega$ can be singly parametrized. Suppose $\varphi: \bar{\Omega} \rightarrow \bar{D}$ is a C^1 -map so that φ maps $\partial\Omega$ to ∂D diffeomorphically, and $f \in C(\bar{D})$, then

$$\int_D f(y) dy = \pm \int_{\Omega} f(\varphi(x)) J_{\varphi}(x) dx. \quad (3.1)$$

Here, the choice of the signs \pm on the right hand side depends on whether φ preserve the orientation of the boundaries.

Proof. Since $f \in C(\bar{D})$, it can be continuously extended to \mathbb{R}^m . Doing convolution with the mollifiers $\{\eta_{\varepsilon}\}_{\varepsilon>0}$, which are functions $\eta_{\varepsilon} \in C^{\infty}(\mathbb{R}^m)$ such that

$$\int_{\mathbb{R}^m} \eta_{\varepsilon}(y) dy = 1, \quad \text{supp } \eta_{\varepsilon} \subset B_{\varepsilon}(0),$$

we obtain a family of functions $f_{\varepsilon} \in C^{\infty}(\mathbb{R}^m)$ such that as $\varepsilon \rightarrow 0^+$,

$$\sup_{y \in D} |f_{\varepsilon}(y) - f(y)| \rightarrow 0, \quad \sup_{x \in \Omega} |f_{\varepsilon}(\varphi(x)) J_{\varphi}(x) - f(\varphi(x)) J_{\varphi}(x)| \rightarrow 0.$$

It is then easy to see that

$$\int_D f_{\varepsilon}(y) dy \rightarrow \int_D f(y) dy, \quad \int_{\Omega} f_{\varepsilon}(\varphi(x)) J_{\varphi}(x) dx \rightarrow \int_{\Omega} f(\varphi(x)) J_{\varphi}(x) dx.$$

Therefore, we only need to prove (3.1) for $f \in C^{\infty}(\mathbb{R}^m)$. Using a similar approximating argument we may also assume that $\varphi \in C^2(\bar{\Omega}, \mathbb{R}^m)$.

Let $C = (-a, a) \times \cdots \times (-a, a)$ be a cube containing \bar{D} , then define $Q: \bar{D} \rightarrow \mathbb{R}$,

$$Q(y) = \int_{-a}^{y^1} f(t, y^2, \dots, y^m) dt.$$

Then $Q \in C^1(\bar{D})$ and

$$\frac{\partial Q}{\partial y^1} = f \quad \text{in } D.$$

Let $x: U \rightarrow \mathbb{R}^m$ be a parametrization of $\partial\Omega$. Since φ maps $\partial\Omega$ to ∂D diffeomorphically, it follows that $y = \varphi \circ x$ is a parametrization of ∂D . Then

$$N = \left(\frac{\partial(y^2, \dots, y^m)}{\partial(u^1, \dots, u^{m-1})}, \dots, (-1)^{m+1} \frac{\partial(y^1, \dots, y^{m-1})}{\partial(u^1, \dots, u^{m-1})} \right)$$

is a normal vector at $y(u)$ on ∂D and

$$n = \pm N / |N| = (n^1, \dots, n^m) \quad (3.2)$$

is the unit outer normal vector at $y(u)$ on ∂D . By the chain rule we have

$$\begin{pmatrix} \partial_{u^1} y^2 & \cdots & \partial_{u^{m-1}} y^2 \\ \vdots & & \vdots \\ \partial_{u^1} y^m & \cdots & \partial_{u^{m-1}} y^m \end{pmatrix} = \begin{pmatrix} \partial_{x^1} y^2 & \cdots & \partial_{x^m} y^2 \\ \vdots & & \vdots \\ \partial_{x^1} y^m & \cdots & \partial_{x^m} y^m \end{pmatrix} \begin{pmatrix} \partial_{u^1} x^1 & \cdots & \partial_{u^{m-1}} x^1 \\ \vdots & & \vdots \\ \partial_{u^1} x^m & \cdots & \partial_{u^{m-1}} x^m \end{pmatrix}.$$

Applying the Cauchy-Binet formular, we obtain from (3.2) that

$$\begin{aligned} \pm n^1 |N| &= \frac{\partial(y^2, \dots, y^m)}{\partial(u^1, \dots, u^{m-1})} \\ &= \sum_{i=1}^m \frac{\partial(y^2, \dots, y^m)}{\partial(x^1, \dots, \hat{x}^i, \dots, x^m)} \frac{\partial(x^1, \dots, \hat{x}^i, \dots, x^m)}{\partial(u^1, \dots, u^{m-1})} = A \cdot \tilde{N}, \end{aligned} \quad (3.3)$$

where $A = (A_1, \dots, A_m)$, $\tilde{N} = (\tilde{N}^1, \dots, \tilde{N}^m)$, with

$$A_i = (-1)^{i+1} \frac{\partial(y^2, \dots, y^m)}{\partial(x^1, \dots, \hat{x}^i, \dots, x^m)}, \quad \tilde{N}^i = (-1)^{i+1} \frac{\partial(x^1, \dots, \hat{x}^i, \dots, x^m)}{\partial(u^1, \dots, u^{m-1})}.$$

Note that $\tilde{n} = \pm \tilde{N} / |\tilde{N}|$ is the unit outer normal vector at $x(u)$ on $\partial \Omega$. Moreover, A_i is exactly the algebraic cofactor of $\partial_{x^i} y^1$ in the Jacobian

$$J_\varphi(x) = \det \begin{pmatrix} \partial_{x^1} y^1 & \cdots & \partial_{x^m} y^1 \\ \vdots & & \vdots \\ \partial_{x^1} y^m & \cdots & \partial_{x^m} y^m \end{pmatrix}.$$

Thus, since φ is of class C^2 , by the Hadamard identity [2, Page 14] we deduce

$$\operatorname{div} A = \sum_{i=1}^m \frac{\partial A_i}{\partial x^i} = 0. \quad (3.4)$$

Let $\tilde{Q} = Q \circ \varphi$, then $\tilde{Q} \in C^1(\bar{\Omega})$. Using (3.4) we obtain

$$\begin{aligned} \operatorname{div}(\tilde{Q}A) &= \nabla \tilde{Q} \cdot A + \tilde{Q} \operatorname{div} A = \nabla \tilde{Q} \cdot A \\ &= \sum_{i=1}^m \frac{\partial \tilde{Q}}{\partial x^i} A_i = \sum_{i=1}^m \left(\sum_{j=1}^m \frac{\partial Q}{\partial y^j} \Big|_{\varphi(x)} \frac{\partial y^j}{\partial x^i} \right) A_i \\ &= \sum_{j=1}^m \frac{\partial Q}{\partial y^j} \Big|_{\varphi(x)} \left(\sum_{i=1}^m \frac{\partial y^j}{\partial x^i} A_i \right) = \sum_{j=1}^m \frac{\partial Q}{\partial y^j} \Big|_{\varphi(x)} \delta_1^j J_\varphi(x) \\ &= \frac{\partial Q}{\partial y^1} \Big|_{\varphi(x)} J_\varphi(x) = f(\varphi(x)) J_\varphi(x). \end{aligned}$$

Applying Theorem 2.1 and using (3.3), we have

$$\begin{aligned}
 \int_D f(y) dy &= \int_D \frac{\partial Q}{\partial y^1} dy = \int_{\partial D} Q n^1 d\sigma \\
 &= \int_U Q(y(u)) n^1(u) |N(u)| du \\
 &= \pm \int_U \tilde{Q}(x(u)) (A(x(u)) \cdot \tilde{N}(u)) du \\
 &= \pm \int_U (\tilde{Q}(x(u)) A(x(u)) \cdot \tilde{n}(u)) |\tilde{N}(u)| du \\
 &= \pm \int_{\partial \Omega} \tilde{Q} A \cdot \tilde{n} d\sigma = \pm \int_{\Omega} \operatorname{div}(\tilde{Q} A) dx = \pm \int_{\Omega} f(\varphi(x)) J_{\varphi}(x) dx. \quad \square
 \end{aligned}$$

Corollary 3.1. Under the assumption of Theorem 3.1, if $J_{\varphi}(x)$ does not change sign as x varies in Ω , then

$$\int_D f(y) dy = \int_{\Omega} f(\varphi(x)) |J_{\varphi}(x)| dx.$$

Corollary 3.2. (Non-Retraction Lemma.) Let B be the unit closed ball in \mathbb{R}^m , then there does not exist C^1 -map $T: B \rightarrow \mathbb{R}^m$ such that $T(B) \subset \partial B$ and $T|_{\partial B} = 1_{\partial B}$.

Proof. The proof below is essentially a variant form of the argument in [1, Corollarys]. Suppose there is a C^1 -map T with the stated properties. Obviously T map ∂B to itself diffeomorphically. We define a continuous function $f: B \rightarrow \mathbb{R}$,

$$f(y) = \begin{cases} 1 - 4|y|^2, & \text{if } |y| \leq \frac{1}{2}, \\ 0, & \text{if } \frac{1}{2} < |y| \leq 1. \end{cases}$$

Then $f(T(x)) = 0$ for all $x \in B$. By Theorem 3.1,

$$0 < \int_B f(y) dy = \pm \int_B f(T(x)) J_T(x) dx = 0,$$

a contradiction. □

As is well known, the Brouwer Fixed Point Theorem is an easy consequence of Corollary 3.2.

4 General domains

In this section, we prove Theorem 1.1 for the general case that $\partial \Omega$ may not be singly parametrized. Let $f_{\pm} = \max\{\pm f, 0\}$, then $f = f_+ - f_-$. Because f_{\pm} are also continuous on \bar{D} , it follows that we only need to prove the result for nonnegative f . For simplicity, we set

$$\tilde{f}(x) = f(\varphi(x)) |J_{\varphi}(x)|.$$

We want to prove

$$\int_D f(y) dy = \int_\Omega \tilde{f}(x) dx.$$

For any $\varepsilon > 0$, there exist disjoint balls $B_i \subset \Omega$ ($i = 1, \dots, \ell$) such that

$$\int_\Omega \tilde{f}(x) dx \leq \sum_{i=1}^{\ell} \int_{B_i} \tilde{f}(x) dx + \varepsilon. \quad (4.1)$$

Let $U_i = \varphi(B_i)$, then $\varphi: B_i \rightarrow U_i$ is a C^1 -diffeomorphism. Since ∂B_i can be singly parametrized and J_φ is of constant sign in Ω , hence in B_i , by Corollary 3.1 we have

$$\int_{B_i} \tilde{f}(x) dx = \int_{U_i} f(y) dy. \quad (4.2)$$

Because $U_i \cap U_j = \emptyset$ and

$$U = \bigcup_{i=1}^{\ell} U_i \subset D,$$

from (4.1), (4.2), and noting that $f \geq 0$, we deduce that

$$\int_\Omega \tilde{f}(x) dx \leq \sum_{i=1}^{\ell} \int_{U_i} f(y) dy + \varepsilon = \int_U f(y) dy + \varepsilon \leq \int_D f(y) dy + \varepsilon.$$

Let $\varepsilon \rightarrow 0$, we get

$$\int_\Omega \tilde{f}(x) dx \leq \int_D f(y) dy. \quad (4.3)$$

Since $\varphi: \Omega \rightarrow D$ is a diffeomorphism, switching the roles of f and \tilde{f} in the above argument, we obtain

$$\int_D f(y) dy \leq \int_\Omega \tilde{f}(x) dx. \quad (4.4)$$

Now the conclusion of Theorem 1.1 follows from (4.3) and (4.4).

Acknowledgments

Research supported by NSFC (Grant No.11671331) and NSFFJ (Grant No.2014J06002). Part of this work was done while S. Liu is visiting the University of Notre Dame, he is very grateful for the hospitality from the mathematical department and for the Chinese Scholarship Council for financial support. Y. Zhang is grateful to hospitality Professor Huai-Dong Cao for constant encouragement and support. He was supported by the Science and Technology Development Fund (Macao S.A.R.) Grant FDCT/016/2013/A1 and the Project MYRG2015-00235-FST of the University of Macau.

The authors are very grateful to the anonymous referees for their valuable comments and questions.

References

- [1] L. Báez-Duarte. Brouwer's fixed-point theorem and a generalization of the formula for change of variables in multiple integrals. *J. Math. Anal. Appl.*, 1993, 177: 412-414.
- [2] K. Deimling. *Nonlinear Functional Analysis*. Springer-Verlag, Berlin, 1985.
- [3] M. P. do Carmo. *Differential Geometry of Curves and Surfaces*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1976. Translated from the Portuguese.
- [4] N. V. Ivanov. A differential forms perspective on the Lax proof of the change of variables formula. *Amer. Math. Monthly*, 2005, 112: 799-806.
- [5] P. D. Lax. Change of variables in multiple integrals. *Amer. Math. Monthly*, 1999, 106: 497-501.
- [6] P. D. Lax. Change of variables in multiple integrals, II. *Amer. Math. Monthly*, 2001, 108: 115-119.
- [7] M. Taylor. Differential forms and the change of variable formula for multiple integrals. *J. Math. Anal. Appl.*, 2002, 268: 378-383.