Kiril Datchev
MA 562
Fall 2022

This note presents the change of variables theorem, following Lax's 1998 paper *Change of Variables in Multiple Integrals* and Section 4.5 of Taylor's book *Introduction to Analysis in Several Variables.* Questions, comments, and corrections are gratefully received at kdatchev@purdue.edu.

The first section covers the theorem itself, and the second uses it to prove Brouwer's fixed point theorem.

1. The Change of Variables Theorem

The change of variables theorem in one dimension says that if $f\colon \mathbb{R} \to \mathbb{R}$ is continuous and $\varphi\colon \mathbb{R} \to \mathbb{R}$ is C^∞, and if $\varphi(a) = A$ and $\varphi(b) = B$, then

$$\int_a^b f(\varphi(x))\varphi'(x)\,dx = \int_A^B f(y)\,dy.$$

One can prove this using the fundamental theorem of calculus: let $g(y) = \int_A^y f(z)\,dz$. Then $g'(y) = f(y)$ and the right hand side becomes $g(B) - g(A)$ while the left hand side becomes

$$\int_a^b g'(\varphi(x))\varphi'(x)\,dx = \int_a^b \frac{d}{dx} g(\varphi(x))\,dx = g(\varphi(b)) - g(\varphi(a)) = g(B) - g(A).$$

Note that writing it this way gives a result which does not require φ to be invertible. Our plan is to mimic this proof in higher dimensions.

Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a continuous function of compact support, and let $\varphi\colon \mathbb{R}^n \to \mathbb{R}^n$ be C^∞. We ask, under what conditions on φ can we prove that

$$\int f(\varphi(x))J(x)\,dx = \int f(y)\,dy, \tag{1}$$

where $J(x)$ is the Jacobian determinant $\det(\partial_{x^j}\varphi^k(x))$? By the chain rule and the multiplicativity of the determinant, the class of functions φ for which (1) holds is closed under compositions. It clearly includes translations.

Case 1: φ is linear, $\varphi(x) = Ax$, with $J(x) = \det A > 0$.

Approximating $\int f$ by Riemann sums and using linearity, we see that it is enough to prove (1) in Case 1 when f is the characteristic function of a n-cube C. Then we must show $\det A \operatorname{vol}(\varphi^{-1}(C)) = \operatorname{vol}(C)$. This is true if and only if $\det A > 0$.

Case 2: $\varphi(x) = x$ for $|x|$ large enough.

Let

$$g(y^1, \dots, y^n) = \int_{-\infty}^{y^1} f(z, y^2, \dots, y^n)\,dz, \qquad \partial_{y^1} g(y) = f(y).$$

Then we rewrite the integrand on the left side of (1) as

$$(\partial_{y^1} g)(\varphi(x))J(x)\,dx^1 \wedge \cdots \wedge dx^n = \varphi^*(\partial_{y^1} g)\varphi^*(dy^1 \wedge \cdots \wedge dy^n) = \varphi^*(dg \wedge dy^2 \wedge \cdots \wedge dy^n)$$
$$= \tilde{J}(x)\,dx^1 \wedge \cdots \wedge dx^n,$$

where in the second equality we used $dg = \sum_{k=1}^n \partial_{y_k} g\, dy^k$, and where $\tilde{J}(x)$ is the Jacobian of the transformation $(g \circ \varphi, \varphi^2, \ldots, \varphi^n)$, i.e.

$$f(\varphi(x))J(x) = \sum_\sigma \operatorname{sgn}(\sigma)\partial_{\sigma(1)}(g \circ \varphi)\partial_{\sigma(2)}\varphi^2 \cdots \partial_{\sigma(n)}\varphi^n,$$

where the sum is over permutations of $\{1, \ldots, n\}$, and the partials on the right are with respect to the x variables and everything is evaluated at x. We now plug this into the left side of (1) and integrate by parts over a cube $(-c, c)^n$ large enough to encompass the supports of f and $\varphi - I$, taking the $\partial_{\sigma(1)}$ derivative off $\partial_{\sigma(1)}(g \circ \varphi)$ and putting it on $\partial_{\sigma(2)}\varphi^2 \cdots \partial_{\sigma(n)}\varphi^n$, to get

$$\int f(\varphi(x))J(x)\,dx = \quad \textit{boundary terms} \quad -$$
$$\int \sum_\sigma \operatorname{sgn}(\sigma)(g \circ \varphi) \sum_{k=2}^n \partial_{\sigma(2)}\varphi^2 \cdots \partial_{\sigma(k-1)}\varphi^{k-1}\partial_{\sigma(1)}\partial_{\sigma(k)}\varphi^k \partial_{\sigma(k+1)}\varphi^{k+1} \cdots \partial_{\sigma(n)}\varphi^n.$$

But for every k, we have $\sum_\sigma \operatorname{sgn}(\sigma)\partial_{\sigma(2)}\varphi^2 \cdots \partial_{\sigma(k-1)}\varphi^{k-1}\partial_{\sigma(1)}\partial_{\sigma(k)}\varphi^k\partial_{\sigma(k+1)}\varphi^{k+1} \cdots \partial_{\sigma(n)}\varphi^n = 0$ because for smooth φ we have $\partial_{\sigma(1)}\partial_{\sigma(k)}\varphi^k = \partial_{\sigma(k)}\partial_{\sigma(1)}\varphi^k$. This leaves the boundary terms, and to simplify these we use the fact that $\varphi = I$ on the boundary and $g = 0$ on all boundary faces except $\{c\} \times [-c, c]^{n-1}$, giving

$$\int f(\varphi(x))J(x)\,dx = \int g(c, y^2, \ldots, y^n)\,dy^2 \cdots dy^n = \int f\,dy,$$

where we inserted the definition of g and used the fact that f is supported in $(-c, c)^n$. This completes the proof that (1) holds in the case where $\varphi(x) = x$ for $|x|$ large enough.

Case 3: There is a neighborhood U of the support of f such that φ is injective[1] on $\varphi^{-1}(U)$ and $J > 0$ on $\varphi^{-1}(U)$.

By a partition of unity argument, it is enough to show that every p in the support of f has a neighborhood U_p such that (1) holds with f replaced by f_p, where f_p is a continuous function supported in U_p. By composing with a linear function and a translation we may assume that $p = 0$ and that $D\varphi(0) = I$. By Case 2, and using the injectivity of φ, it is enough to prove that there exists a C^∞ injective map $\Phi\colon \mathbb{R}^n \to \mathbb{R}^n$ such that $\Phi = \varphi$ near 0 and $\Phi = I$ off a compact set.

Let

$$\Phi(x) = \varphi(x)b(x/\varepsilon) + x(1 - b(x/\varepsilon)),$$

where $b \in C^\infty(\mathbb{R}^n)$ takes values in $[0, 1]$, is 1 on $B_{1/2}$, and vanishes outside of B_1, where B_r is the open ball of radius r centered at the origin, and $\varepsilon > 0$ is to be determined. It is clear that $\Phi = \varphi$ in $B_{\varepsilon/2}$ and $\Phi = I$ off B_ε, so it remains to choose ε small enough that Φ is injective. But

$$|x - \tilde{x}| \le |\Psi(x) - \Psi(\tilde{x})| + |\Phi(x) - \Phi(\tilde{x})|,$$

[1] Note that by the inverse function theorem replacing 'injective on $\varphi^{-1}(U)$' with 'is a diffeomorphism from $\varphi^{-1}(U)$ to U' results in an equivalent condition: see Corollary II.6.7 of Boothby's *Introduction to Differentiable Manifolds and Riemannian Geometry.*

where $\Psi(x) = \Phi(x) - x$, and using the mean value bound we have

$$|\Psi(x) - \Psi(\tilde{x})| = |b(x/\varepsilon)(\varphi(x) - x) - b(\tilde{x}/\varepsilon)(\varphi(\tilde{x}) - \tilde{x})|$$
$$\leq \Big(\max|b(z/\varepsilon)(\partial_j\varphi^k(z) - \delta_j^k) - \partial_j b(z/\varepsilon)(\varphi^k(z) - z^k)/\varepsilon|\Big) n|x - \tilde{x}|,$$

where the max is taken over $z \in B_\varepsilon$ and $j, k \in \{1, \ldots, n\}$. It is enough to show that this max tends to 0 as $\varepsilon \to 0$, because then for ε small enough we have $|\Psi(x) - \Psi(\tilde{x})| \leq |x - \tilde{x}|/2$, and thus $|x - \tilde{x}| \leq 2|\Phi(x) - \Phi(\tilde{x})|$ which implies Φ is injective.

To show the max tends to 0 as $\varepsilon \to 0$, use the mean value bound again to write

$$\max|b(z/\varepsilon)(\partial_j\varphi^k(z) - \delta_j^k)| \leq \max|\partial_j\varphi^k(z) - \delta_j^k| \leq \Big(\max|\partial_i\partial_j\varphi^k|\Big) n^{1/2}\varepsilon,$$

and

$$\max|\partial_j b(z/\varepsilon)(\varphi^k(z) - z^k)/\varepsilon| \leq \max|\partial_j b| \max|\varphi^k(z) - z^k|/\varepsilon$$
$$\leq \max|\partial_j b| \max|\partial_j\varphi^k(z) - \delta_j^k| n^{1/2} \leq \max|\partial_j b| \max|\partial_i\partial_j\varphi^k| n\varepsilon.$$

More general cases:

We can treat the case $J(x) < 0$ by switching two rows or columns of φ. That leads to the statement that

$$\int f(\varphi(x))|J(x)|dx = \int f(y)\,dy, \tag{2}$$

provided either $J(x)$ never vanishes on the support of f or $\varphi(x) = \pm x$ outside of a compact set.

We can replace the right side of (2) by $\int_U f(y)\,dy$ for an open set U by using a partition of unity to write $f \in C(U)$ as $f = \sum_{j=1}^{\infty} f_j$, where each f_j is continuous and compactly supported. Then the left side of (2) becomes $\int_{\varphi(U)} f(\varphi(x))|J(x)|dx$

Using Lebesgue theory, we can allow $f \in L^1(\mathbb{R}^n)$ by using a partition of unity to reduce to the case where f is compactly supported, then linearity to reduce to the case where f is nonnegative, then a sequence of truncations $f_n(x) = \min\{n, f(x)\}$ and the monotone convergence theorem to reduce to the case where f is bounded, and finally the fact that continuous functions of compact support are dense in bounded compactly supported integrable functions.

2. The Brouwer Fixed Point Theorem

The Brouwer Fixed Point Theorem. Let $F \colon \overline{B_1} \to \overline{B_1}$ be continuous. Then there is $x \in \overline{B_1}$ such that $F(x) = x$.

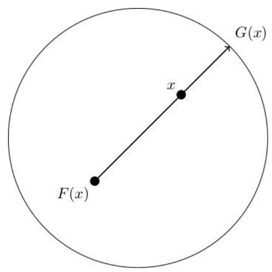
Here and below, $B_a = \{x \in \mathbb{R}^n \colon |x| < a\}$, and $\overline{B_a}$ and ∂B_a are its closure and boundary.

A fun example is that if you take a map of the room you are in (or of the city, or of the country, etc.) and crumple it, there is always a point on the map which lies directly over its corresponding point on the ground. Here we are using the fact that there is a homeomorphism from $\overline{B_1} \subset \mathbb{R}^2$ to the shape of this room.

The Brouwer fixed point theorem follows from a higher dimensional version of the familiar intermediate value theorem, which we simply call

The Intermediate Value Theorem. Let $G\colon \overline{B_1} \to \overline{B_1}$ be continuous. If $G|_{\partial B_1}$ is the identity, i.e. if $G(x) = x$ whenever $|x| = 1$, then G is surjective.

Proof that IVT implies BFPT. We argue by contradiction. Suppose $F\colon \overline{B_1} \to \overline{B_1}$ is continuous and $F(x) \neq x$ for all $x \in \overline{B_1}$. Then define $G\colon \overline{B_1} \to \partial B_1$ by taking each point $x \in \overline{B_1}$ to the point on ∂B_1 obtained by following the ray from $F(x)$ to x until it meets ∂B_1, as in the picture below



Then G satisfies all the hypotheses of the Intermediate Value Theorem but not the conclusion, so we have a contradiction. $\square$

We will prove IVT using the change of variables formula in this form: if $\varphi\colon \mathbb{R}^n \to \mathbb{R}^n$ is C^∞ and there is R such that $\varphi(x) = x$ whenever $|x| \geq R$, then

$$\int f(\varphi(x))J(x)\,dx = \int f(y)\,dy, \tag{3}$$

for any continuous and compactly supported function f. It is significant that we did not have to assume that φ is injective or surjective.

Lemma. If $\varphi\colon \mathbb{R}^n \to \mathbb{R}^n$ is C^∞ and there is R such that $\varphi(x) = x$ whenever $|x| \geq R$, then φ is surjective.

Proof. We again argue by contradiction. Suppose φ is not surjective. Since φ is the identity off of B_R, there must exist $y_0 \in B_R$ such that $\varphi(x) \neq y_0$ for all $x \in \overline{B_R}$. But $\varphi(\overline{B_R})$ is compact, so there is a neighborhood U of y_0 such that $\varphi(x) \notin U$ for all $x \in \mathbb{R}^n$.

Let f be a smooth function supported in U such that $\int f = 1$. Then $f \circ \varphi$ vanishes identically, and we obtain the contradiction that the right side of (3) is 1 and the left side is 0. $\square$

Proof of IVT. Extend G to a continuous function $\mathbb{R}^n \to \mathbb{R}^n$ by putting $G(x) = x$ for all $x \notin B_1$. If G happens to be C^∞ then we may apply the Lemma with $\varphi = G$ and we are done.

If not, take a sequence of C^∞ functions $\varphi_1, \varphi_2, \ldots$ converging to G such that for all x we have $\varphi_j(x) = x$ for when $|x| \geq 2$ and such that

$$\max_{x\in\mathbb{R}^n} |G(x) - \varphi_j(x)| \to 0 \qquad \text{as } j \to \infty. \tag{4}$$

Let $y \in B_1$ be given, let $x_1,\ x_2,\ \cdots \in B_1$ obey $\varphi_j(x_j) = y$. Let $x_{j_1}, x_{j_2}, \ldots$ be a subsequence converging to some limit $x^* \in \overline{B_1}$. Then

$$|G(x^*) - y| \leq |G(x^*) - G(x_{j_k})| + |G(x_{j_k}) - \varphi_{j_k}(x_{j_k})|,$$

which tends to zero as $k \to \infty$ by (4) and the continuity of G. This proves $G(x^*) = y$ and thus that G is surjective.

To construct such a sequence $\varphi_1, \varphi_2, \ldots$, let ψ be a smooth function supported in B_1 such that $\int \psi = 1$ (as in the proof of the Lemma above), put $H(y) = G(y) - y$ and put

$$\varphi_j(x) = x + j^n \int \psi(j(x-y))H(y)dy.$$

To see where this comes from, note that $\psi(j(x-y))$ is nonzero only when $|x-y| < 1/j$ and hence roughly speaking for j large we have $H(y) \approx H(x)$ in the integrand, after which using $j^n \int \psi(j(x-y))dy = \int \psi = 1$ gives $\varphi_j(x) \approx x + H(x) = G(x)$.

Now observe that $\varphi_j \in C^\infty$ by differentiating under the integral sign, using the fact that the integrands are continuous and compactly supported in y, even after any number of differentiations with respect to x. To see that $\varphi_j(x) = x$ when $|x| \geq 2$, observe that $H(y)$ is nonzero only when $|y| < 1$ and $\psi(j(x-y))$ is nonzero only when $|x-y| < 1/j$, and hence the integrand can only be nonzero when $|x| < 1 + 1/j$.

It remains to check (4). We make precise the reasoning above using '$\approx$' by writing

$$G(x) - \varphi_j(x) = j^n \int \psi(j(x-y))(H(x) - H(y))dy = \int \psi(z)(H(x) - H(x + zj^{-1}))dz,$$

where for the first equality we used $j^n \int \psi(j(x-y))dy = \int \psi = 1$, and for the second the change of variables $z = j(y-x)$, $dz = j^n dy$. Since a continuous function on a compact set is uniformly continuous, for every $\varepsilon > 0$ there is N such that $|H(x) - H(x + zj^{-1})| < \varepsilon$ when $j \geq N$ and $|z| < 1$, regardless of x. That gives $|G(x) - \varphi_j(x)| < \varepsilon \int |\psi|$, and hence (4). $\square$

Later in life Brouwer became a renowned enemy of proof by contradiction, so it is always a pleasure to use contradiction to prove his most famous theorem. Actually a direct proof is easy in the one-dimensional case, because one can prove the intermediate value theorem by repeated bisection (a binary search). In the higher dimensional cases I'm told you can do it using Sperner's lemma from combinatorics.