MAT 473 Intermediate Real Analysis II

John Quigg Spring 2009

revised April 29, 2009

Change of variables

In this last section we'll prove a general theorem justifying substitutions in integrals over $\mathbb{R}^n$. It'll be convenient to separate out the special case of linear substitutions:

Lemma 1. *Let T be an invertible linear operator on $\mathbb{R}^n$. If $f \in L^+(\mathbb{R}^n)$ or $L^1(\mathbb{R}^n)$ then so is $f \circ T$, and*

$$\int f = |\det T| \int f \circ T.$$

Proof. By linearity it suffices to prove it for L^+, then then by linearity and the Monotone Convergence Theorem it suffices to prove it for characteristic functions. A little bit of manipulation using the definitions reveals that we must show that if $A \in \mathcal{L}$ then $T(A) \in \mathcal{L}$ and $m(T(A)) = |\det T| m(A)$.

T is a homeomorphism on $\mathbb{R}^n$, so takes open sets to open sets, hence takes Borel sets to Borel sets. Thus we can define $\mu : \mathcal{B}_{\mathbb{R}^n} \to [0, \infty]$ by $\mu(A) = m(T(A))/|\det T|$. It is routine to check that μ is a measure. If we can show that it agrees with Lebesgue measure on boxes, it will coincide with m on $\mathcal{B}_{\mathbb{R}^n}$ by uniqueness, and then a by-now standard technique will allow us to conclude that μ coincides with m on the completed σ-algebra $\mathcal{L}$, as desired.

Now let A be a box. If the conclusion holds for both T_1 and T_2, then it also holds for T_1T_2, since $f \circ T_1 \circ T_2 = f \circ (T_1T_2)$ and $(\det T_1)(\det T_2) = \det T_1T_2$. Thus it is enough to consider T of 3 types:

(1) T switches two coordinates;
(2) T multiplies one coordinate by a nonzero $c \in \mathbb{R}$;
(3) T adds one coordinate to another.

In cases (1)–(2), note that in each case $T(A)$ is a box, with $m(T(A))$ equal to $m(A)$ in (1) and to $|c|m(A)$ in (2). In case (3), we use Tonelli's Theorem and the 1-variable version of translation-invariance: if T adds x_k to x_j (with $j \neq k$), then

$$\begin{aligned}
m(T(A)) &= \int \chi_{T(A)} = \int \chi_A \circ T^{-1} \\
&= \int \cdots \int \chi_A \circ T^{-1}(x_1, \ldots, x_n)\, dx_j\, dx_1 \cdots dx_n && (x_j \text{ first}) \\
&= \int \cdots \int \chi_A(x_1, \ldots, x_j - x_k, \ldots, x_n)\, dx_j\, dx_1 \cdots dx_n \\
&= \int \cdots \int \chi_A(x_1, \ldots, x_j, \ldots, x_n)\, dx_j\, dx_1 \cdots dx_n && (\text{translation-invariance}) \\
&= m(A).
\end{aligned}$$

$$
\begin{aligned}
m(T(A)) &= \int \cdots \int \chi_{T(A)}(x_1, \ldots, x_n)\, dx_1 \cdots dx_n \\
&= \int \cdots \int \chi_A \circ T^{-1}(x_1, \ldots, x_n)\, dx_1 \cdots dx_n \\
&= \int \cdots \int \chi_A(x_1, \ldots, x_j - x_k, \ldots, x_n)\, dx_1 \cdots dx_n \\
&= \int \cdots \int \chi_A(x_1, \ldots, x_j - x_k, \ldots, x_n)\, dx_j\, dx_1 \cdots dx_n \quad (x_j \text{ first}) \\
&= \int \cdots \int \chi_A(x_1, \ldots, x_j, \ldots, x_n)\, dx_j\, dx_1 \cdots dx_n \qquad \text{(translation-invariance)} \\
&= m(A).
\end{aligned}
$$

Since $\det T$ equals -1 in (1), c in (2), and 1 in (3), we are done with boxes, and the proof is complete. $\square$

Theorem 2 (Change of Variables Theorem)**.** *Let $U \subset \mathbb{R}^n$ be open, and let $g : U \to \mathbb{R}^n$ be a C^1 diffeomorphism*[1]*. If f is measurable on $g(U)$, then $f \circ g$ is measurable on U. If also f is nonnegative or integrable, then*

$$\int_{g(U)} f = \int_U f \circ g \,|\det g'|.$$

Proof. This is a hard one, requiring several steps.

Step 1. We begin as in Lemma 1: by linearity it suffices to assume $f \in L^+(g(U))$. However, there is a subtlety: we won't reduce to characteristic functions directly. Rather, we use some trickery: first we'll show that $f \circ g$ is measurable and

$$\int_{g(U)} f \leq \int_U f \circ g |\det g'|.$$

Once we've established this, we can appeal to symmetry: replace g by g^{-1} and f by $(f \circ g)|\det g'|$, and then by the same inequality we get

$$
\begin{aligned}
\int_U f \circ g |\det g'| &= \int_{g^{-1}(g(U))} f \circ g |\det g'| \\
&\leq \int_{g(U)} f \circ g \circ g^{-1} |\det g' \circ g^{-1}| |\det (g^{-1})'| \\
&= \int_{g(U)} f,
\end{aligned}
$$

because for all $y \in g(U)$ we have

$$\left|\det (g^{-1})'(y)\right| = \left|\det g'(g^{-1}(y))^{-1}\right| = \left|\det g'(g^{-1}(y))\right|^{-1}.$$

[1]this means that g is 1-1 and C^1, and also $g^{-1} : g(U) \to U$ is C^1; so in particular $g(U)$ is open by the Inverse Function Theorem, consequently the hypothesis is symmetric in g and g^{-1}

Step 2. As in the proof of Lemma 1, by linearity and the Monotone Convergence Theorem it suffices to consider characteristic functions, and then it suffices to show that if $A \subset U$ and $A \in \mathcal{L}$ then $g(A) \in \mathcal{L}$ and

$$m(g(A)) \leq \int_A |\det g'|,$$

since we would then have $\chi_A \circ g = \chi_{g^{-1}(A)} \in L^+(U)$ because $g^{-1}(A)$ is measurable (since g^{-1} has the same properties as g), and

$$\int_{g(U)} \chi_A = m(A) \leq \int_{g^{-1}(A)} |\det g'| = \int_U \chi_{g^{-1}(A)} |\det g'| = \int_U \chi_A \circ g |\det g'|.$$

Step 3. Similarly to the proof of Lemma 1, it suffices to consider Borel sets, then combine with null sets and monotonicity to handle Lebesgue measurable sets.

Step 4. Suppose we have shown it for closed cubes. Then if A is open we can find nonoverlapping[2] closed cubes $B_1, B_1, \ldots$ such that $A = \bigcup_i B_i$, so that

$$\begin{aligned} m(g(A)) &= m\left(\bigcup_i g(B_i)\right) \leq \sum_i m(g(B_i)) \\ &\leq \sum_i \int_{B_i} |\det g'| = \int_A |\det g'|, \end{aligned}$$

because $\bigcup_i \partial B_i$ is null and the B_i's have disjoint interiors. Then for every Borel set A and every open set $B \supset A$ we have

$$m(g(A)) \leq m(g(B)) \leq \int_B |\det g'|,$$

and taking the inf over B we get $m(g(A)) \leq \int_A |\det g'|$ (because $|\det g'|$ is locally integrable).

Step 5. Now let A be a closed cube. This step alone is quite hard, and will occupy our attention for some time. It will be convenient throughout this step to use the ℓ^∞-norm $\|x\|_\infty = \max_i |x_i|$ on $\mathbb{R}^n$. Note that for this norm the associated operator norm on $L(\mathbb{R}^n)$ is given by the "maximum row sum":

$$\|T\| = \max_i \sum_{j=1}^n |a_{ij}|,$$

where T is identified with an $n \times n$ matrix (a_{ij}).

Let t be the center of the cube A, so that if the side of the cube is $2r$, then

$$A = \overline{B}_r(t) \quad \text{and} \quad m(A) = (2r)^n.$$

Since A is compact and g' is continuous we can put

$$M = \max\{\|g'(y)\| : y \in A\}.$$

[2]this means having disjoint interiors

Let $x \in A$. For each i, since $g_i : A \to \mathbb{R}$ is differentiable, by the Mean Value Theorem there exists $y \in [t, x]$ such that

$$\begin{aligned} |g_i(x) - g_i(t)| &= \big|g_i'(y)(x-t)\big| \\ &\le \|g_i'(y)\|\|x-t\|_\infty \\ &\le \|g'(y)\|r, \end{aligned}$$

because $\|g'(y)\| = \max_i \|g_i'(y)\|$. Hence

$$\|g(x) - g(t)\|_\infty \le Mr.$$

Thus

$$g(A) \subset \overline{B}_{Mr}(g(t)),$$

so

$$m(g(A)) \le m\big(\overline{B}_{Mr}(g(t))\big) = (2Mr)^n = M^n m(A).$$

Now let $T = g'(x)$ with $x \in A$, and apply the above to $T^{-1} \circ g$ instead of g:

$$\begin{aligned} m(g(A)) &= m\big(T \circ T^{-1} \circ g(A)\big) \\ &= |\det T| m\big(T^{-1} \circ g(A)\big) \\ &= |\det T| \left(\max_{y \in A} \big\|(T^{-1} \circ g)'(y)\big\|\right)^n m(A). \end{aligned}$$

Now,

$$\begin{aligned} (T^{-1} \circ g)'(y) &= (T^{-1})'(g(y))g'(y) && \text{Chain Rule} \\ &= T^{-1}g'(y) && \text{because } T^{-1} \text{ is linear} \\ &= g'(x)^{-1}g'(y). \end{aligned}$$

Thus, letting $C = \max_{x,y \in A} \|g'(x)^{-1}g'(y)\|^n$, we have

$$m(g(A)) \le |\det T| C m(A).$$

Define $\phi : A^2 \to \mathbb{R}$ by $\phi(x, y) = \|g'(x)^{-1}g'(y)\|^n$. Then ϕ is uniformly continuous since g is C^1 and A^2 is compact, so given $\varepsilon > 0$ we can choose $\delta > 0$ such that for all $(x, y), (u, v) \in A^2$, if $\|(x, y) - (u, v)\|_\infty < \delta$ then

$$|\phi(x, y) - \phi(u, v)| < \varepsilon,$$

so that, in particular, for all $x, y \in A$, if $\|x - y\|_\infty < \delta$ then

$$\phi(x, y) \le |\phi(x, y) - \phi(x, x)| + \phi(x, x) < 1 + \varepsilon,$$

since

$$\|x - y\|_\infty = \|(x, y) - (x, x)\|_\infty \quad \text{and} \quad \phi(x, x) = 1.$$

Divide A into nonoverlapping closed subcubes $B_1, \ldots, B_k$ with $B_i = \overline{B}_s(t_i)$ and $s < \delta$, and apply the above to each B_i: for all $y \in B_i$ we have $\|t_i - y\| \le s < \delta$, so

$$\begin{aligned} m(g(A)) = m\left(\bigcup_1^k g(B_i)\right) &\le \sum_1^k m(g(B_i)) \\ &\le \sum_1^k |\det g'(t_i)|(1+\varepsilon)m(B_i) \\ &= (1+\varepsilon)\int_A \sum_1^k |\det g'(t_i)|\chi_{B_i}. \end{aligned}$$

Since $|\det g'|$ is continuous, letting $\delta \to 0$ gives

$$m(g(A)) \le (1+\varepsilon)\int_A |\det g'|,$$

because the B_i's have disjoint interiors. Then letting $\varepsilon \to 0$ gives

$$m(g(A)) \le \int_A |\det g'|.$$

This completes the final step, and the proof. □

Now a slight problem arises: many common change-of-variables transformations ("coordinate systems") don't quite satisfy the hypotheses of the Change of Variables Theorem; but these transformations are typically "close enough" in the following general sense: suppose we have measurable sets $A, B \subset \mathbb{R}^n$, open sets $U \subset A$ and $V \subset B$, and a function g such that

- $g(A) = B$;
- g restricts to a C^1 diffeomorphism of U onto V;
- $m(A \setminus U) = m(B \setminus V) = 0$.

Then: $f \circ g$ is measurable on A if f is measurable on B, and in case f is nonnegative or integrable we have $\int_B f = \int_A f \circ g |\det g'|$.

Example 3. *Polar Coordinates.* Define $g : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$g(r, \theta) = (r\cos\theta, r\sin\theta).$$

Then g is C^1, but certainly very far away from being 1-1. In the notation of the above discussion, typically one could take

$$\begin{aligned} A &= [0, \infty) \times [0, 2\pi] \\ B &= \mathbb{R}^2 \\ U &= (0, \infty) \times (0, 2\pi) \\ V &= \{(x, y) \in \mathbb{R}^2 : x < 0 \text{ or } y \ne 0\}. \end{aligned}$$

Then

$$\det g'(r, \theta) = \det \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} = r,$$

so we can use polar coordinates to write

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = \int_0^{2\pi}\int_0^{\infty} f(r\cos\theta, r\sin\theta) r\,dr\,d\theta$$

for $f \in L^+(\mathbb{R}^2)$ or $L^1(\mathbb{R}^2)$.