5 Il teorema del cambio di variabili

Teorema 1 *Sia $A \subset \mathbb{R}^n$ un aperto misurabile e $\phi \in C^1(\overline{A}, \mathbb{R}^n)$ tale che ϕ sia iniettiva su A e*

$$\det \frac{\partial \phi}{\partial x} \neq 0 \ , \qquad \forall \ x \in A \ . \tag{1}$$

Allora $B := \phi(A)$ è un aperto misurabile di $\mathbb{R}^n$ e se f è una funzione integrabile su[1] *B allora $f \circ \phi$ è integrabile su A e si ha che*

$$\int_B f(y)dy = \int_A f \circ \phi(x) \Big| \det \frac{\partial \phi}{\partial x}(x) \Big| dx \ . \tag{4}$$

In particolare, vale

$$\mathrm{mis}(\phi(A)) = \int_A \Big| \det \frac{\partial \phi}{\partial x}(x) \Big| dx \ . \tag{5}$$

Nella dimostrazione useremo il seguente

Lemma 2 *Sia A un aperto limitato di $\mathbb{R}^n$ ed $F \in C(\overline{A}, \mathbb{R}^m)$.*
(i) *Se $F(A)$ è aperto in $\mathbb{R}^m$ allora $\partial(F(A)) \subset F(\partial A)$.*
(ii) *Se F è iniettiva su $\overline{A}$ allora $F(\partial A) \subset \partial(F(A))$.*

Dimostrazione (i) Sia $B := F(A)$ e $y \in \partial B$, cioè (essendo B aperto per ipotesi) $\exists$ $y_n \in B$ tali che $y_n \to y \notin B$. Siano $x_n \in A$ tali che $F(x_n) = y_n$. Poiché $\overline{A}$ è compatto, esiste una sottosuccesione x_{n_k} convergente in $\overline{A}$, e sia $x = \lim_{k\to\infty} x_{n_k}$. Per la continuità

[1] Ricordiamo che f è integrabile secondo Riemann su $B \subset \mathbb{R}^n$ se B è misurabile secondo Peano-Jordan e se, dato un qualunque rettangolo, limitato e non degenere E che contenga B, la funzione $\bar{f}_B$ (cioè la funzione che coincide con f su B e vale 0 su B^c) è integrabile secondo Riemann su E. E' facile vedere che (**esercizio**) f *è integrabile secondo Riemann su B se e solo se per ogni $\varepsilon > 0$ esiste una partizione $\mathcal{R}$ di E tale che*

$$\sum_{R \in \mathcal{R}} \Big(\sup_{R\cap B} f - \inf_{R \cap B} f \Big) \,\mathrm{mis}(R \cap B) \leq \varepsilon \tag{2}$$

ed in tal caso

$$\sup_{\{\mathcal{R} \text{ partizione di } E\}} \ \sum_{R \in \mathcal{R}} (\inf_{R\cap B} f) \,\mathrm{mis}(R \cap B) = \int_B f = \inf_{\{\mathcal{R} \text{ partizione di } E\}} \ \sum_{R \in \mathcal{R}} (\sup_{R\cap B} f) \,\mathrm{mis}(R \cap B) \ . \tag{3}$$

di F, $F(x_{n_k}) = y_{n_k} \to F(x)$, cioè $y = F(x)$. Ma allora $x \in \partial A$: se fosse $x \in A$, $y = F(x)$ apparterrebbe a B, contrariamente alla nostra assunzione. Quindi $\partial(F(A)) \subset F(\partial A)$.
(ii) Sia $x \in \partial A$ e $y = F(x)$. Poiché A è aperto, x non appartiene ad A ed esistono $x_n \in A$ tali che $x_n \to x$. Per la continuità di F, $y_n := F(x_n) \to F(x) = y$ e naturalmente $y_n \in B$. D'altra parte, $y \notin B$ e quindi $y \in \partial B$: se fosse $y \in B$, per l'iniettività di F, $x = F^{-1}(y)$ e quindi x apparterrebbe ad A contrariamente alla nostra assunzione. Dunque $F(\partial A) \subset \partial(F(A))$. ∎

Osservazione 3 (i) Per $n = 1$, la (4) non è altro che la nota formula del cambio di variabile in integrali semplici: in tal caso ϕ è una funzione scalare di variabile reale $x \in \mathbb{R}$ ed il suo jacobiano è la derivata di ϕ rispetto ad x. Si noti che l'apparire del modulo è dovuto al fatto che se $\phi' < 0$ la funzione ϕ non conserva l'ordine: supponiamo infatti che $A = [a, b]$ e $\phi' < 0$ su (a, b) (cioè ϕ è strettamente decrescente su $[a, b]$), allora $\phi([a, b]) = [\phi(b), \phi(a)]$. Quindi, per la formula del cambio di variabile in integrali semplici (ponendo $y = \phi(x)$), si ha

$$\begin{aligned}\int_{\phi(b)}^{\phi(a)} f(y)dy &= \int_b^a f \circ \phi(x)\phi' dx \\ &= -\int_a^b f \circ \phi(x)\phi' dx = \int_{[a,b]} f \circ \phi(x)|\phi'| dx \ .\end{aligned}$$

Nel resto di questa sezione considereremo $n \geq 2$.
(ii) La formula (5) si ottiene da (4) prendendo $f \equiv 1$.
(iii) Il fatto che $B = \phi(A)$ sia un insieme aperto è conseguenza del teorema della funzione inversa (e dell'assunzione che $\det \phi' \neq 0$ su A).
(iv) Per il Lemma 2, punto (i), essendo $B = \phi(A)$ un insieme aperto, $\partial\phi(A) \subset \phi(\partial A)$. Inoltre ∂A è un insieme di misura nulla (essendo A misurabile). Quindi poiché ϕ (essendo $C^1(\overline{A})$) è uniformemente lipschitziana su $\overline{A}$, segue che $\phi(\partial A)$ è un insieme di misura nulla e quindi lo è anche il suo sottoinsieme $\partial(\phi(A))$. Dunque, per il teorema di Riemann-Lebesgue, B è un insieme misurabile.
(v) Poiché ϕ è un diffeomeorfismo, $y_0 \in B$ è un punto di continuità per f se e solo se $x_0 = \phi^{-1}(y_0)$ è un punto di continuità per $f \circ \phi$. Quindi, ancora, per il teorema di Riemann-Lebesgue, f è integrabile su B se e solo se $f \circ \phi$ è integrabile su A ed essendo le funzioni integrabili secondo Riemann un'algebra, $f \circ \phi \, \Big|\det \phi'\Big|$ è integrabile su A.
(vi) Sia $\phi(x) = Tx$ con $T \in \mathrm{Mat}(n \times n)$ matrice invertibile ed A il cubo unitario n–dimensionale aperto $(0, 1)^n$ e $K_1 = \overline{A}$. In tal caso la matrice jacobiana $\frac{\partial\phi}{\partial x}$ coincide con T stessa e quindi le ipotesi del teorema sono chiaramente soddisfatte. La formula (5) diviene, allora,

$$\mathrm{mis}(TK_1) = \mathrm{mis}(TA) = |\det T| \, \mathrm{mis}(A) = |\det T| \ . \tag{6}$$

In realtà la parte più significativa del teorema è proprio l'affermazione (6): il caso generale, infatti, si ricondurrà a (6) approssimando, localmente, $\phi(x)$ con la formula di Taylor al prim'ordine.

Cominciamo dunque con il discutere (6). Si ricordi che il determinante, visto come funzione delle colonne, è caratterizzato dalle seguenti tre proprietà: 1) scambiando di posto a due colonne, il valore del determinante cambia segno, 2) il determinante è una funzione lineare della prima colonna[2], 3) il determinante di $(e^{(1)},...,e^{(n)})$, dove $\{e^{(i)}\}$ è la base standard di $\mathbb{R}^n$, vale 1. Tale caratterizzazione significa che se Δ è una funzione a valori reali e definita su n–nuple di vettori in $\mathbb{R}^n$ che verifica

$$\begin{aligned}
&1)\quad \Delta(...,v^{(i)},...,v^{(j)},...) = -\Delta(...,v^{(j)},...,v^{(i)},...)\ ,\quad i\neq j\ ,\\
&2)\quad \Delta(av^{(1)}+bw^{(1)},v^{(2)},...,v^{(n)}) =\\
&\qquad\qquad = a\Delta(v^{(1)},v^{(2)},...,v^{(n)}) + b\Delta(w^{(1)},v^{(2)},...,v^{(n)})\ ,\\
&3)\quad \Delta(e^{(1)},...,e^{(n)}) = 1\ ,
\end{aligned}\tag{7}$$

(dove la scrittura simbolica nel punto 1) sta a significare che scambiando la i–esima colonna con la j–esima il valore di Δ cambia segno) allora Δ coincide con il determinante, cioè

$$\Delta(v^{(1)},...,v^{(n)}) = \det[v^{(1)},...,v^{(n)}] = \sum_\sigma \varepsilon_\sigma v^{(1)}_{\sigma_1}\cdots v^{(n)}_{\sigma_n}\ ,$$

dove la sommatoria è estesa a tutte le permutazioni σ dell'insieme $\{1,...,n\}$ e ε_σ denota il[3] "segno di σ".

La relazione tra determinanti e volumi si basa sulla seguente definizione: siano $v^{(1)}$,..., $v^{(n)}$, n vettori in $\mathbb{R}^n$ e sia $T := [v^{(1)},...,v^{(n)}]$ la matrice che ha come j–esima colonna le n componenti, $v^{(j)}_i$, di $v^{(j)}$.

Definizione 4 *Si chiama "parallelepipedo con lati $v^{(1)},...,v^{(n)}$" (o anche "parallelepipedo generato da $v^{(1)},...,v^{(n)}$") l'insieme*

$$\Pi(v^{(1)},...,v^{(n)}) := \{y\in\mathbb{R}^n:\ y=\sum_{j=1}^n x_j v^{(j)}\ \text{ con }\ 0\le x_j\le 1\ ,\forall 1\le j\le n\}\ .\tag{8}$$

Se $x\in\mathbb{R}^n$, $Tx=\sum_{j=1}^n x_j v^{(j)}$, dunque da tale definizione segue immediatamente che

$$\Pi(v^{(1)},...,v^{(n)}) = TK_1\ ,\qquad \Big(T:=[v^{(1)},...,v^{(n)}]\ ,K_1:=[0,1]^n\Big)\ .\tag{9}$$

[2]E dunque, per 1), il determinante è una funzione lineare della j–esima colonna, con j qualunque.

[3]Una permutazione σ di $\{1,...,n\}$ è una mappa uno–uno $\sigma: j\in\{1,...,n\}\to\sigma_j\in\{1,...,n\}$ e ε_σ è il segno di σ (cioè $(-1)^p$ dove p è il numero di scambi che bisogna fare per ordinare la n–nupla $(\sigma_1,...,\sigma_n)$ o, più analiticamente, $\varepsilon_\sigma=\prod_{i<j}$ segno$(\sigma_j-\sigma_i)$).

Dimostrazione di (6). Identificando una n-pla di vettori[4] $(v^{(1)},...,v^{(n)})$ con la matrice $T := [v^{(1)},...,v^{(n)}]$, definiamo

$$\Delta_0(v^{(1)},...,v^{(n)}) := \mathrm{mis}(TK_1) = \mathrm{mis}\left(\Pi(v^{(1)},...,v^{(n)}\right) , \tag{10}$$

e[5]

$$\Delta(T) := \ \mathrm{segno}\,(\det T)\ \Delta_0(T) \ . \tag{11}$$

Si noti che se $\det T = 0$, $\Pi(T)$ è contenuto in uno spazio vettoriale di dimensione $m < n$ e dunque , in tal caso, $\mathrm{mis}(\Pi(T)) = 0$ e $\Delta(T) = 0$. Per dimostrare (6) basterà dimostrare che la funzione Δ verifica 1), 2), 3) di (7) cosicché si avrà $\Delta(T) = \det T$ e prendendo il modulo di tale relazione si otterrà (6). Dalla definizione di $\Pi(v^{(1)},...,v^{(n)})$ segue immediatamente che

$$\Pi(...,v^{(i)},...,v^{(j)},...) = \Pi(...,v^{(j)},...,v^{(i)},...) \tag{12}$$

e da (9) e da (11) segue immediatamente che, scambiando due vettori di posto, Δ cambia segno: la proprietà 1) è verificata. E' anche chiaro che

$$\Pi(e^{(1)},...,e^{(n)}) = K_1 \tag{13}$$

e dunque $\Delta(e^{(1)},...,e^{(n)}) = 1$ e cioè vale la proprietà 3). Più delicata è la verifica della linearità. Cominciamo con l'osservare che se p è un intero positivo, notando che $[0,p] = \{x_1 p : 0 \le x_1 \le 1\} = \cup_{i=1}^{p}(i-1) + [0,1]$,

$$\begin{aligned}\Pi(pv^{(1)},v^{(2)},...,v^{(n)}) \ &:= \{x_1 pv^{(1)} + \sum_{j=2}^{n} x_j v^{(j)} :\ 0 \le x_j \le 1\} \\ &= \bigcup_{i=1}^{p} \{(i-1+x_1)v^{(1)} + \sum_{j=2}^{n} x_j v^{(j)} :\ 0 \le x_j \le 1\} \\ &= \bigcup_{i=1}^{p} \Big((i-1)v^{(1)} + \Pi(v^{(1)},...,v^{(n)})\Big) \ .\end{aligned} \tag{14}$$

Poiché l'insieme nell'ultima riga di (14) è formato dall'unione di p insiemi di eguale misura (essendo la misura invariante per traslazioni) aventi in comune solo insiemi di misura nulla ("facce" dei parallelepipedi), si ha che la misura di $\Pi(pv^{(1)},v^{(2)},...)$ è p volte la misura di $\Pi(v^{(1)},v^{(2)},...)$; cioè

$$\Delta_0(av^{(1)},...,v^{(n)}) = a\Delta_0(v^{(1)},...,v^{(n)}) \tag{15}$$

[4]Si ricordi che $n \ge 2$.

[5]Useremo qui la convenzione che $\mathrm{segno}\,(0) = 0$.

vale con $a = p$ intero positivo. Ora se $\det T > 0$ da (15) segue

$$\Delta(av^{(1)}, ..., v^{(n)}) = a\Delta(v^{(1)}, ..., v^{(n)}) \ , \tag{16}$$

e se $\det T < 0$, per (15) si ha che

$$\begin{aligned} \Delta(av^{(1)}, ..., v^{(n)}) &= -\Delta_0(av^{(1)}, ..., v^{(n)}) \\ &= -a\Delta_0(v^{(1)}, ..., v^{(n)}) \\ &= a\Delta(v^{(1)}, ..., v^{(n)}) \ , \end{aligned} \tag{17}$$

il che dimostra la (16) per ogni T. Ora,

$$\Delta(v^{(1)}, ..., v^{(n)}) = \Delta(p\ \frac{1}{p}v^{(1)}, ..., v^{(n)}) = p\Delta(\frac{1}{p}v^{(1)}, ..., v^{(n)}) \tag{18}$$

e quindi (dividendo la (18) per p, si vede che) (16) vale anche per $a = \frac{1}{p}$ con p intero positivo. Combinando questi due fatti si ottiene subito che (16) vale per a numero razionale positivo. Dopodiché osservando che, se $0 < \alpha < \beta$

$$\Pi(\alpha v^{(1)}, ..., v^{(n)}) \subset \Pi(\beta v^{(1)}, ..., v^{(n)}) \tag{19}$$

si ottiene facilmente (15) per $a \in (0, \infty)$: basta infatti considerare due successioni monotone di numeri razionali positivi, $a_i < a < a_i'$ che tendano, rispettivamente da sinistra e da destra al numero reale a, ed usare (16) per razionali positivi che insieme alla relazione (19) implica

$$a_i\Delta_0(v^{(1)}, ..., v^{(n)}) \leq \Delta_0(av^{(1)}, .., v^{(n)}) \leq a_i'\Delta_0(v^{(1)}, ..., v^{(n)})$$

e prendendo il limite per $i \to \infty$ si ottiene (15) con a numero reale positivo. Da (15) segue la (16) quando $\det T > 0$ e, nel caso $\det T < 0$, la (16) per $a > 0$ segue da (17). Per $a = 0$, (16) deriva immediatamente dalla definizione, essendo $\det[0, v^{(2)}, .., v^{(n)}] = 0$. Infine si noti che

$$\begin{aligned} v^{(1)} + \Pi(-v^{(1)}, ..., v^{(n)}) &= \{(1 - x_1)v^{(1)} + \sum_{j=2}^{n} x_j v^{(j)} :\ 0 \leq x_j \leq 1\} \\ &= \Pi(v^{(1)}, ..., v^{(n)}) \end{aligned} \tag{20}$$

e dunque, essendo il segno di $\det(-v^{(1)}, ..., v^{(n)})$ opposto a quello di $\det(v^{(1)}, ..., v^{(n)})$, si ottiene (16) con $a = -1$. Infine se a è un numero reale negativo, $a = -|a|$, si ottiene

$$\Delta(av^{(1)}, ..., v^{(n)}) = -\Delta(|a|v^{(1)}, ..., v^{(n)}) = a\Delta(v^{(1)}, ..., v^{(n)}) \ ,$$

quindi (16) è vera per ogni $a \in \mathbb{R}$. Usando la già dimostrata proprietà 1) si ottiene immediatamente che

$$\Delta(v^{(1)}, ..., av^{(j)}, ..., v^{(n)}) = a\Delta(v^{(1)}, ..., v^{(n)}) \ , \tag{21}$$

per qualunque $1 \le j \le n$. Resta ora da dimostrare

$$\Delta(v + w, v^{(2)}, ..., v^{(n)}) = \Delta(v, v^{(2)}, ..., v^{(n)}) + \Delta(w, v^{(2)}, ..., v^{(n)}) \ . \tag{22}$$

Cominciamo col discutere un caso particolare di (22) e cioè $v = v^{(1)}$, $w = v^{(2)}$ (in qual caso il secondo addendo a destra di (22) è nullo per definizione di Δ):

$$\Delta(v^{(1)} + v^{(2)}, v^{(2)}, ..., v^{(n)}) = \Delta(v^{(1)}, v^{(2)}, ..., v^{(n)}) \ . \tag{23}$$

Se $\det T = \det[v^{(1)}, ..., v^{(n)}] = 0$, la (23) è chiaramente vera. Assumiamo che $\det T > 0$ e introduciamo i seguenti "prismi"

$$D_1 := \{x \in K_1 : x_2 \le x_1\} \ , \quad D_2 := \{x \in K_1 : x_2 \ge x_1\} \ , \quad D_3 := e^{(2)} + D_1 \tag{24}$$

e si noti che $\text{mis}(D_1 \cap D_2) = 0$, $\text{mis}(D_2 \cap D_3) = 0$; che $\Pi(e^{(1)}, ..., e^{(n)}) = K_1$ e che

$$\begin{aligned} \Pi(e^{(1)} + e^{(2)}, e^{(2)}, ..., e^{(n)}) &= \{x_1 e^{(1)} + (x_1 + x_2)e^{(2)} + \cdots \ , \ 0 \le x_i \le 1\} \\ &= \{x_1 e^{(1)} + y e^{(2)} + \cdots \ , \ 0 \le x_1 \le 1 \ , \ x_1 \le y \le 1 + x_1 \ , ...\} \\ &= \{x_1 e^{(1)} + y e^{(2)} + \cdots \ , \ 0 \le x_1 \le 1 \ , \ x_1 \le y \le 1 \ , ...\} \cup \\ &\quad\ \{x_1 e^{(1)} + (1 + y)e^{(2)} + \cdots \ , \ 0 \le x_1 \le 1 \ , \ 0 \le y \le x_1 \ , ...\} \\ &= D_2 \cup \Big(e^{(2)} + D_1\Big) =: D_2 \cup D_3 \end{aligned} \tag{25}$$

Si ricordi anche che, se $T = [v^{(1)}, ..., v^{(n)}]$, $Te^{(i)} = v^{(i)}$ e che, per ogni $A \in \text{Mat}(n \times n)$, $AT = [Av^{(1)}, ..., Av^{(n)}]$ e $A\Pi(T) = \Pi(AT)$. Quindi,

$$\begin{aligned} \Pi(v^{(1)}, ..., v^{(n)}) &= TK_1 = TD_1 \cup TD_2 \ , \\ \Pi(v^{(1)} + v^{(2)}, v^{(2)}, ..., v^{(n)}) &= \Pi\Big(Te^{(1)} + Te^{(2)}, Te^{(2)}, ..., Te^{(n)}\Big) \\ &= \Pi\Big(T(e^{(1)} + e^{(2)}, e^{(2)}, ..., e^{(n)})\Big) \\ &= T\Pi(e^{(1)} + e^{(2)}, e^{(2)}, ..., e^{(n)}) \\ &= TD_2 \cup TD_3 \\ &= TD_2 \cup (v^{(2)} + TD_1) \ . \end{aligned}$$

Quindi, poiché $\text{mis}(TD_1 \cap TD_2) = 0$ e $\text{mis}(TD_2 \cap TD_3) = 0$ (essendo $x \to Tx$ derivabile) e poiché la misura è invariante per traslazioni, si ottiene

$$\begin{aligned} \text{mis}\left(\Pi(v^{(1)} + v^{(2)}, v^{(2)}, ..., v^{(n)})\right) &= \text{mis}(TD_2) + \text{mis}(TD_3) \\ &= \text{mis}(TD_2) + \text{mis}(TD_1) \\ &= \text{mis}(TD_1 \cup TD_2) \\ &= \text{mis}\left(\Pi(v^{(1)}, v^{(2)}, ..., v^{(n)}\right) \ , \end{aligned}$$

il che dimostra (23) nel caso $\det T > 0$; poiché $\det[v^{(1)} + v^{(2)}, v^{(2)}, ...] = \det[v^{(1)}, v^{(2)}, ...]$, il caso $\det T < 0$ segue immediatamente. Combinando (23) con la proprietà 1) si ottiene

$$\Delta(v^{(1)} + v^{(j)}, v^{(2)}, ..., v^{(n)}) = \Delta(v^{(1)}, v^{(2)}, ..., v^{(n)}) \tag{26}$$

per ogni $2 \leq j \leq n$. Sia ora a un numero diverso da 0. Allora da (21) (con $j = 2$) e da (23), si ottiene

$$\begin{aligned} \Delta(v^{(1)} + av^{(2)}, v^{(2)}, ..., v^{(n)}) &= \frac{1}{a}\Delta(v^{(1)} + av^{(2)}, av^{(2)}, ..., v^{(n)}) \\ &= \frac{1}{a}\Delta(v^{(1)}, av^{(2)}, ..., v^{(n)}) \\ &= \Delta(v^{(1)}, v^{(2)}, ..., v^{(n)}) \ . \end{aligned} \tag{27}$$

Naturalmente per $a = 0$, l'uguaglianza tra il primo e l'ultimo membro di (27) è banalmente vera. Usando ancora la proprietà 1), otteniamo, per ogni $2 \leq j \leq n$, e per ogni $a \in \mathbb{R}$,

$$\Delta(v^{(1)} + av^{(j)}, v^{(2)}, ..., v^{(n)}) = \Delta(v^{(1)}, v^{(2)}, ..., v^{(n)}) \ ; \tag{28}$$

da tale relazione segue facilmente che, per ogni a_j,

$$\Delta(v^{(1)} + \sum_{j=2}^{n} a_j v^{(j)}, v^{(2)}, ..., v^{(n)}) = \Delta(v^{(1)}, v^{(2)}, ..., v^{(n)}) \ , \tag{29}$$

che equivale alla (22) nel caso in cui w appartenga allo spazio generato da $\{v^{(j)}, j \geq 2\}$. Naturalmente se $v^{(2)}, ..., v^{(n)}$ sono linearmente dipendenti (22) è banalmente vera, essendo entrambi i membri a sinistra e destra dell'uguaglianza nulli. Assumiamo dunque $w =: v^{(1)}$, $v^{(2)}, ..., v^{(n)}$ linearmente indipendenti, cosicché $v^{(1)}, ..., v^{(n)}$ forma una base in $\mathbb{R}^n$. Allora esistono n costanti a_i tali che $v = \sum_{i=1}^{n} a_i v^{(i)}$ ed usando (28) ed (16),

$$\begin{aligned} &\Delta(v + w, v^{(2)}, ..., v^{(n)}) \\ &\quad = \Delta(\sum_{i=1}^{n} a_i v^{(i)} + v^{(1)}, v^{(2)}, ..., v^{(n)}) \\ &\quad = \Delta((1 + a_1)v^{(1)} + \sum_{i=2}^{n} a_i v^{(i)}, v^{(2)}, ..., v^{(n)}) \\ &\quad = \Delta((1 + a_1)v^{(1)}, v^{(2)}, ..., v^{(n)}) \\ &\quad = (1 + a_1)\Delta(v^{(1)}, v^{(2)}, ..., v^{(n)}) \\ &\quad = \Delta(v^{(1)}, v^{(2)}, ..., v^{(n)}) + \Delta(a_1 v^{(1)}, v^{(2)}, ..., v^{(n)}) \\ &\quad = \Delta(v^{(1)}, v^{(2)}, ..., v^{(n)}) + \Delta(a_1 v^{(1)} + \sum_{i=2}^{n} a_i v^{(i)}, v^{(2)}, ..., v^{(n)})) \\ &\quad = \Delta(w, v^{(2)}, ..., v^{(n)}) + \Delta(v, v^{(2)}, ..., v^{(n)}) \ . \end{aligned}$$

Questo completa la dimostrazione della proprietà 2) e quindi per quanto discusso sopra, vale (6). $\blacksquare$

Osservazione 5 Si noti che, infatti, abbiamo dimostrato (6) per *ogni matrice* T (cioè non solo per matrici T invertibili).

Sia ora R il rettangolo $[0,b_1]\times\cdots\times[0,b_n]$ con $b_n>0$. Allora $R=\Lambda K_1$ dove $\Lambda=\mathrm{diag}\,(b_1,...,b_n)$ e quindi

$$\begin{aligned}\mathrm{mis}(TR) &= \mathrm{mis}\,(T(\Lambda K_1))=\mathrm{mis}\,((T\Lambda)K_1)=|\det(T\Lambda)|\\ &=|\det T|\;|\det\Lambda|=|\det T|(b_1b_2\cdots b_n)=|\det T|\,\mathrm{mis}(R)\ .\end{aligned}$$

E poiché, se $a=(a_1,...,a_n)$,

$$[a_1,b_1]\times\cdots\times[a_n,b_n]=a+[0,b_1-a_1]\times\cdots\times[0,b_n-a_n]$$

dall'invarianza della misura per traslazione segue che

$$\mathrm{mis}(TR)=|\det T|\,\mathrm{mis}(R) \tag{30}$$

per ogni rettangolo in $\mathbb{R}^n$.

Siamo ora pronti per la

Dimostrazione di (5). Sia $\Delta_\phi(x):=|\det\frac{\partial\phi}{\partial x}(x)|$. Dividiamo la dimostrazione di (5) in tre passi: (1) dimostriamo prima che per ogni cubo chiuso $R\subset A$, vale

$$\mathrm{mis}(\phi(R))=\int_R|\Delta_\phi(x)|dx\ ; \tag{31}$$

(2) vediamo poi che (31) vale anche se R è un qualunque rettangolo chiuso contenuto in A e, (3), dimostriamo (5).

Nel corso della dimostrazione denoteremo con $|x|$ la norma del massimo in $\mathbb{R}^n$ ($|x|=\max_i|x_i|$) e con $\|T\|$ la relativa norma matriciale ($\|T\|=\sup_{|x|=1}|Tx|$).

(1) Sia R un cubo chiuso contenuto in A. Possiamo scrivere per ogni $x\in R$ ed ogni h tale che $x_0+h\in R$

$$\phi(x_0+h)=\phi(x_0)+\frac{\partial\phi}{\partial x}(x_0)h+\theta(h,x_0) \tag{32}$$

con

$$\theta(h,x_0):=\Big[\int_0^1\Big(\frac{\partial\phi}{\partial x}(x_0+th)-\frac{\partial\phi}{\partial x}(x_0)\Big)dt\Big]h \tag{33}$$

(si noti che R è un insieme convesso e quindi il segmento $\{x_0+th : t \in [0,1]\}$ è interamente contenuto in R). Fissiamo $\varepsilon > 0$ e sia $\delta > 0$ tale che

$$|\Delta_\phi(x) - \Delta_\phi(y)| \le \varepsilon \ , \quad \Big\|\frac{\partial\phi}{\partial x}(x) - \frac{\partial\phi}{\partial x}(y)\Big\| \le \varepsilon \tag{34}$$

per ogni $x, y \in R$ tali che $|x-y| \le \delta$. Da (34) e (33) segue che, per ogni $x \in R$, $x+h \in R$ con $|h| \le \delta$

$$|\theta(h,x)| \le |h|\varepsilon \ . \tag{35}$$

Sia $\mathcal{R}$ una partizione di R in cubi di lato r con $r \le \delta$. Introduciamo la seguente notazione: indichiamo con $K_\rho(x)$ il cubo chiuso di lato ρ centrato in x ovvero[6]:

$$K_\rho(x) := \{y : \ |y-x| \le \frac{\rho}{2}\} \ . \tag{36}$$

Dimostriamo che esiste una costante $c_1 > 0$ indipendente da ε tale che, per ogni cubo $K = K_r(x_0)$ della partizione $\mathcal{R}$,

$$TK_{r(1-c_1\varepsilon)}(y_0) \subset \phi(K) \subset TK_{r(1+c_1\varepsilon)}(y_0) \ , \tag{37}$$

dove $T := \phi'(x_0)$ e $y_0 = T^{-1}\phi(x_0)$. Riscriviamo la (37) come

$$K_{r(1-c_1\varepsilon)}(y_0) \subset T^{-1}\phi(K) \subset K_{r(1+c_1\varepsilon)}(y_0) \ . \tag{38}$$

Poiché la mappa $x \to T^{-1}\phi(x)$, su $R = \overline{R} \subset A$, è continua, invertibile ed aperta[7], per il lemma 2 si ha che

$$\partial\Big(T^{-1}\phi(K)\Big) = T^{-1}\phi(\partial K) \ , \tag{39}$$

e quindi, se $x \in \partial K$ ovvero se $x = x_0 + h$ con $|h| = \frac{r}{2}$, usando (32) (moltiplicata per T^{-1}), si ottiene

$$\begin{aligned} |T^{-1}\phi(x) - y_0| \quad &= |h + T^{-1}\theta| \\ &\ge |h| - |T^{-1}\theta| = \frac{r}{2} - |T^{-1}\theta| \\ &\ge \frac{r}{2}(1 - \|T^{-1}\|\,\varepsilon) \end{aligned} \tag{40}$$

dove, nell'ultima disuguaglianza abbiamo usato (35). Se poniamo

$$c_1 = \sup_A \Big\|(\frac{\partial\phi}{\partial x})^{-1}\Big\| \tag{41}$$

[6]Si ricordi che stiamo denotando con $|\cdot|$ la norma del massimo.

[7]Si ricorda che una funzione F si dice "aperta" se l'immagine di ogni insieme aperto è un aperto ovvero se $F(A)$ è un insieme aperto, per ogni insieme aperto A.

otteniamo che un punto sulla frontiera di $T^{-1}\phi(K)$ ha distanza (in norma del massimo $|\cdot|$) *almeno* $\frac{r}{2}(1-c_1\varepsilon)$ da y_0 e quindi

$$K_{r(1-c_1\varepsilon)}(y_0) \subset T^{-1}\phi(K) \ .$$

Analogamente, se x è un punto di K, e quindi $x = x_0 + h$ con $|h| \le \frac{r}{2}$,

$$\begin{aligned} |T^{-1}\phi(x) - y_0| &= |h + T^{-1}\theta| \\ &\le |h|(1 + c_1\varepsilon) \le \frac{r}{2}(1 + c_1\varepsilon) \ , \end{aligned} \tag{42}$$

ovvero vale anche la seconda inclusione di (38).
Sia $\varepsilon > 0$ e δ come in (34). Dimostriamo che esiste una costante c_2 (indipendente da ε) tale che, per ogni cubo $K_r(x_0) \subset A$ con $r \le \delta$, si ha

$$\begin{aligned} &\Big|\Delta_\phi(x_0)\,\mathrm{mis}\,(K_r(x_0)) - \mathrm{mis}\,(\phi(K_r(x_0))\Big| \\ = &\Big|\,\mathrm{mis}\,(\phi'(x_0)\,K_r(x_0)) \ - \ \mathrm{mis}\,(\phi(K_r(x_0))\Big| \le c_2\,\varepsilon\ \mathrm{mis}(K_r(x_0)) \ . \end{aligned} \tag{43}$$

Infatti, assumendo che $\varepsilon < 1$ e $c_1\varepsilon < 1$ ed osservando che

$$\begin{aligned} &1 = (1 - c_1\varepsilon + c_1\varepsilon)^n \le (1 - c_1\varepsilon)^n + \bar{c}\varepsilon \ , \qquad \bar{c} := c_1 \sum_{j=1}^{n} \binom{n}{j} \ , \\ &1 = (1 + c_1\varepsilon - c_1\varepsilon)^n \ge (1 + c_1\varepsilon)^n - \bar{c}\varepsilon \ , \end{aligned}$$

si ottiene che

$$\begin{aligned} \mathrm{mis}\,(\phi'(x_0)\,K_r(x_0)) &= |\det\phi'(x_0)|\,r^n \\ &\le |\det\phi'(x_0)|\,\Big(r(1 - c_1\varepsilon)\Big)^n + c_2\,\varepsilon\,r^n \\ &= |\det\phi'(x_0)|\ \mathrm{mis}\,(K_{r(1-c_1\varepsilon)}(y_0)) + c_2\,\varepsilon\ \mathrm{mis}(K_r(x_0)) \\ &= \mathrm{mis}\,(\phi'(x_0)\,K_{r(1-c_1\varepsilon)}(y_0)) + c_2\,\varepsilon\ \mathrm{mis}(K_r(x_0)) \\ &\le \mathrm{mis}\,\Big(\phi(K_r(x_0))\Big) + c_2\,\varepsilon\ \mathrm{mis}(K_r(x_0)) \end{aligned}$$

dove $y_0 := (\phi'(x_0))^{-1}\phi(x_0)$ e $c_2 := \bar{c}\,\sup_A|\det\phi'(x)|$. In modo del tutto analogo si ottiene

$$\mathrm{mis}\,(\phi'(x_0)\,K_r(x_0)) \ge \mathrm{mis}\,\Big(\phi(K_r(x_0))\Big) - c_2\,\varepsilon\ \mathrm{mis}(K_r(x_0))$$

e, dunque, la validità di (43).

Ora, sia $\varepsilon > 0$ arbitrario e δ come sopra e sia $\{R_j\}$, $1 \le j \le N$, una partizione del cubo R in cubi di lato $r \le \delta$ e centro $x^{(j)}$. Allora, per (34) e (43), si ha

$$\begin{aligned}
\Big|\int_R \Delta_\phi - \mathrm{mis}(\phi(R))\Big| &= \Big|\sum_j \int_{R_j} \Delta_\phi - \sum_j \mathrm{mis}(\phi(R_j))\Big| \\
&\le \sum_j \Big|\int_{R_j} \Delta_\phi \, dx - \Delta_\phi(x^{(j)}) \, \mathrm{mis}(R_j)\Big| \\
&\qquad + \sum_j \Big|\Delta_\phi(x^{(j)}) \, \mathrm{mis}(R_j) - \mathrm{mis}(\phi(R_j))\Big| \\
&\le \varepsilon \sum_j \mathrm{mis}(R_j) + c_2 \varepsilon \sum_j \mathrm{mis}(R_j) \\
&\le \mathrm{mis}(A)(1 + c_2) \, \varepsilon \ .
\end{aligned}$$

Dall'arbitrarietà di ε segue la (31) con R cubo chiuso in A.

(2) Ricordiamo i seguenti fatti:

(i) Dato un rettangolo chiuso $R \subset \mathbb{R}^n$ e dato $\varepsilon > 0$, possiamo trovare N cubi a due a due disgiunti $K_i \subset R$ tali che, se

$$R_1 := \bigcup_{i=1}^N K_i \ , \qquad R_2 := R \backslash R_1$$

si abbia

$$\mathrm{mis}(R_2) \le \varepsilon \ .$$

In altre parole, per ogni $\varepsilon > 0$, è possibile decomporre un rettangolo nell'unione di N cubi più un insieme elementare di misura non superiore ad ε.

(ii) Se ϕ è lipschitziana con costante di Lipschitz L (rispetto alla norma del massimo) su E e se K_r è un cubo tale che $K_r \cap E \neq \emptyset$, allora per ogni $x_0 \in K_r \cap E$, si ha che

$$\phi(K_r \cap E) \subset K_{2Lr}(\phi(x_0)) \ . \tag{44}$$

Quindi, se $R \subset E$ è un insieme elementare, allora

$$\mathrm{mis}(\phi(R \cap E)) \le (2L)^n \, \mathrm{mis}(R) \ . \tag{45}$$

Sia, dunque, R un rettangolo chiuso contenuto in A e sia $\varepsilon > 0$. Siano R_1 e R_2 come al punto (i) della precedente osservazione. Allora, per il passo (1) e le osservazioni (i) e (ii) qui sopra si ha che

$$\int_{R_1} \Delta_\phi = \sum_{i=1}^N \int_{K_i} \Delta_\phi = \sum_{i=1}^N \mathrm{mis}(\phi(K_i)) = \mathrm{mis}(\phi(R))$$

da cui

$$\begin{aligned}\Big|\int_R \Delta_\phi - \mathrm{mis}(\phi(R))\Big| &= \Big|\int_{R_2} \Delta_\phi - \mathrm{mis}(\phi(R_2))\Big| \\ &\le \Big(\sup_A \Delta_\phi + (2L)^n + 1\Big)\,\varepsilon\ .\end{aligned}$$

Dall'arbitrarietà di ε segue la (31) vale per ogni rettangolo chiuso $R \subset A$.

(3) Poiché A è misurabile, la sua frontiera è misurabile ed ha misura nulla. Quindi, fissato $\varepsilon > 0$, esiste un insieme elementare $R_2 \supset \partial A$ di misura minore di ε. A questo punto definenedo $R_1 := A \backslash R_2$ (che è un sottoinsieme elementare di A per cui vale la (31)) possiamo ripetere l'argomento usato al passo (2) e concludere la validità di (5). ∎

Dimostrazione di (4). Sia $\mathcal{R}$ una partizione di $E \supset B = \phi(A)$. Allora

$$\begin{aligned}\int_A f\circ\phi\ \Delta_\phi = &\sum_{R\in\mathcal{R}} \int_{\phi^{-1}(R\cap B)} f\circ\phi\ \Delta_\phi \\ &\le \sum_{R\in\mathcal{R}} \big(\sup_{\phi^{-1}(R\cap B)} f\circ\phi\big)\int_{\phi^{-1}(R\cap B)} \Delta_\phi \\ &= \sum_{R\in\mathcal{R}} \big(\sup_{R\cap B} f\big)\,\mathrm{mis}(R\cap B)\end{aligned}$$

e quindi da (3), prendendo l'estremo inferiore su tutte le partizioni $\mathcal{R}$ di E, segue che

$$\int_A f\circ\phi\ \Delta_\phi \le \int_B f\ .$$

La disuguaglianza inversa si dimostra in maniera del tutto analoga. ∎