Il teorema di Vitali-Lebesgue

Gianluca Gorni
Università di Udine

8 gennaio 2013

Nel 1907 Giuseppe Vitali e Henri Lebesgue, indipendentemente uno dall'altro, trovarono che si possono caratterizzare in modo elegante le funzioni integrabili secondo Riemann in termini della misura di Lebesgue. Grosso modo, *le funzioni integrabili secondo Riemann sono quelle i cui punti di discontinuità formano un insieme di misura nulla secondo Lebesgue*, e per loro l'integrale secondo Riemann e secondo Lebesgue coincidono. L'enunciato preciso è più avanti.

Faremo la trattazione per funzioni reali definite su tutto $\mathbb{R}^N$ e da integrare su tutto $\mathbb{R}^N$. Questo non è restrittivo, perché un integrale del tipo $\int_E f(x)\,dx$ si può sempre riscrivere come $\int_{\mathbb{R}^N} f(x)\chi_E(x)\,dx$, dove χ_E è la funzione caratteristica di E in $\mathbb{R}^N$, cioè $\chi_E(x) = 1$ se $x \in E$, $\chi_E(x) = 0$ se $x \in \mathbb{R}^N \setminus E$.

La trattazione che faremo qui segue quella del libro di Giuseppe De Marco, Analisi Due, Decibel-Zanichelli, prima edizione 1993, vol. 2, cap. VII, sez. 5.

1 Funzioni a gradino e integrale di Riemann

Definizione 1.1. Chiameremo *rettangolo* (o anche pluriintervallo, o cuboide, o parallelepipedo, o scatola, in inglese *box*) in $\mathbb{R}^N$ qualsiasi prodotto cartesiano di N intervalli di $\mathbb{R}$.

Definizione 1.2. Chiameremo *funzione a gradino* in $\mathbb{R}^N$ una qualsiasi combinazione lineare finita di funzioni caratteristiche di rettangoli limitati di $\mathbb{R}^N$:

$$\sum_{k=1}^{n} a_k \chi_{R_k}\,, \tag{1}$$

con $a_k \in \mathbb{R}$ e R_k rettangolo limitato.

Si vedano le Figure 1 e 2 per dei grafici di funzioni a gradino di una e di due variabili.

Una funzione a gradino non individua univocamente i coefficienti a_k e i rettangoli R_k. Comunque, sia M l'unione delle frontiere dei rettangoli R_k. Questo M è l'unione di un numero finito di rettangoli che hanno almeno un lato di lunghezza 0, e quindi ha misura di Lebesgue N-dimensionale nulla. Se $x_0 \notin M$ esiste un intorno di x_0 su cui la funzione è costante. Potremmo dire che una funzione a gradino è quasi ovunque localmente costante.

Una funzione a gradino è in particolare una funzione semplice e misurabile nel senso di Lebesgue, ed è anche integrabile:

$$\int_{\mathbb{R}^N} \sum_{k=1}^{n} a_k \chi_{R_k}\, d\lambda = \sum_{k=1}^{n} a_k \operatorname{vol} R_k\,,$$

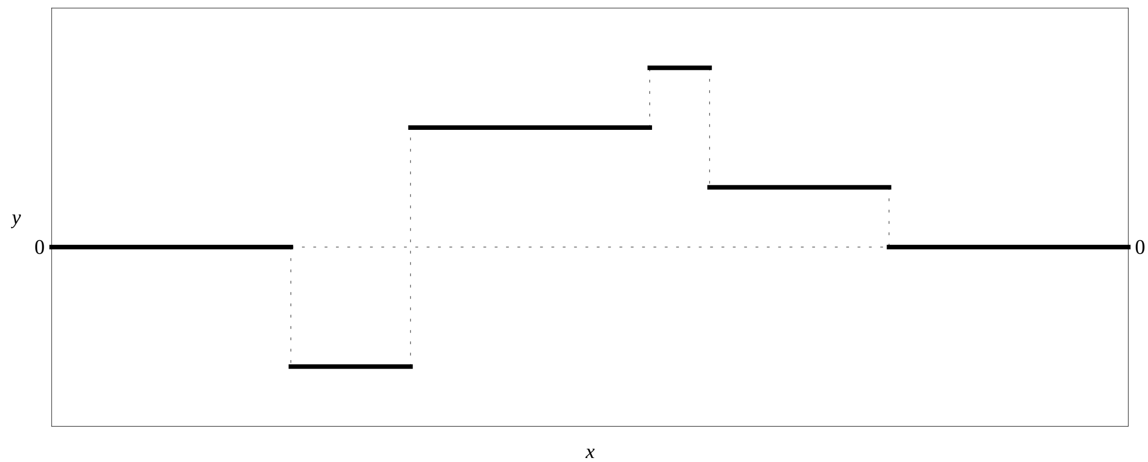

Figura 1: Una funzione a gradino di una variabile

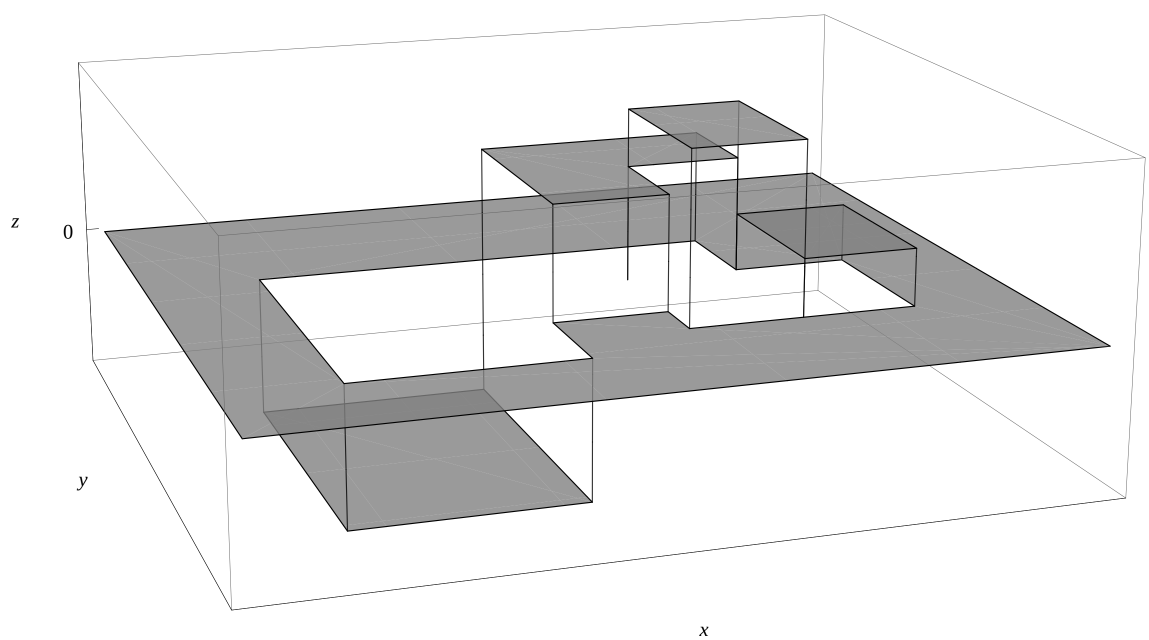

Figura 2: Una funzione a gradino di due variabili

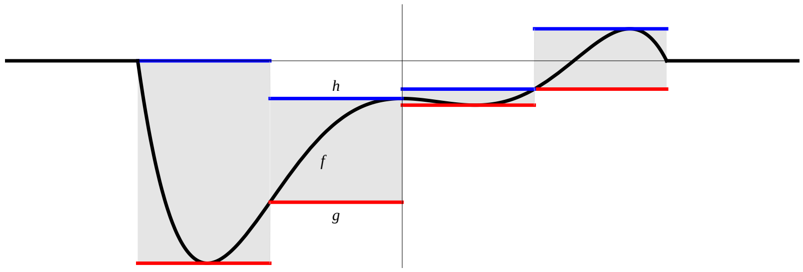

Figura 3: Una funzione f compresa fra le funzioni a gradino g (in rosso) ed h (in blu)

dove λ è la misura di Lebesgue in $\mathbb{R}^N$ e indichiamo con $\operatorname{vol} R$ il volume del rettangolo R. Inoltre ogni funzione a gradino è limitata ed è nulla al di fuori di un insieme limitato di $\mathbb{R}^N$. Qui useremo l'integrale di Lebesgue per le funzioni a gradino, anche se si potrebbe fare per loro una teoria elementare ad hoc.

Definizione 1.3. Data una funzione $f\colon \mathbb{R}^N \to \mathbb{R}$, chiameremo *integrale inferiore* e *integrale superiore* secondo Riemann della funzione f le seguenti quantità rispettivamente:

$$\operatorname*{int\,inf}_{\mathbb{R}^N} f := \sup\Big\{ \int_{\mathbb{R}^N} g\, d\lambda \;\Big|\; g \le f,\ g \text{ a gradino}\Big\} \tag{2}$$

$$\operatorname*{int\,sup}_{\mathbb{R}^N} f := \inf\Big\{ \int_{\mathbb{R}^N} h\, d\lambda \;\Big|\; h \ge f,\ h \text{ a gradino}\Big\} \tag{3}$$

Quando le due quantità sono finite e coincidenti, diremo che la funzione è *integrabile secondo Riemann* e il valore sarà l'integrale secondo Riemann di f.

Perché l'integrale inferiore non sia $-\infty$ occorre che esista almeno una funzione a gradino che sta sotto la f. Analogamente, perché l'integrale superiore non sia $+\infty$ occorre che esista almeno una funzione a gradino che sta sopra la f. Quindi affinché la f sia integrabile secondo Riemann è necessario che esistano due funzioni a gradino g, h tali che $g \le f \le h$. In particolare, *la f deve essere limitata e nulla al di fuori di un limitato.* Questa ipotesi preliminare ricorrerà spesso nel seguito.

In figura 3 è raffigurata una funzione f in nero, ingabbiata fra due funzioni a gradino g ed h in rosso e blu. L'area della regione in grigio è $\int_{\mathbb{R}} (h-g)d\lambda$. Per il caso di due variabili si vedano le figure 9 e 10 rispettivamente a pag. 10 e 11.

Proposizione 1.4. *Sia $f\colon \mathbb{R}^N \to \mathbb{R}$ limitata e nulla al di fuori di un limitato. Allora*

$$-\infty < \operatorname*{int\,inf}_{\mathbb{R}^N} f \le \operatorname*{int\,sup}_{\mathbb{R}^N} f < +\infty\,. \tag{4}$$

Inoltre f è integrabile secondo Riemann se e solo se per ogni $\varepsilon > 0$ esistono $g_\varepsilon, h_\varepsilon$ a gradino tali che

$$g_\varepsilon \le f \le h_\varepsilon \qquad e \qquad \int_{\mathbb{R}^N} (h_\varepsilon - g_\varepsilon)\, d\lambda < \varepsilon\,. \tag{5}$$

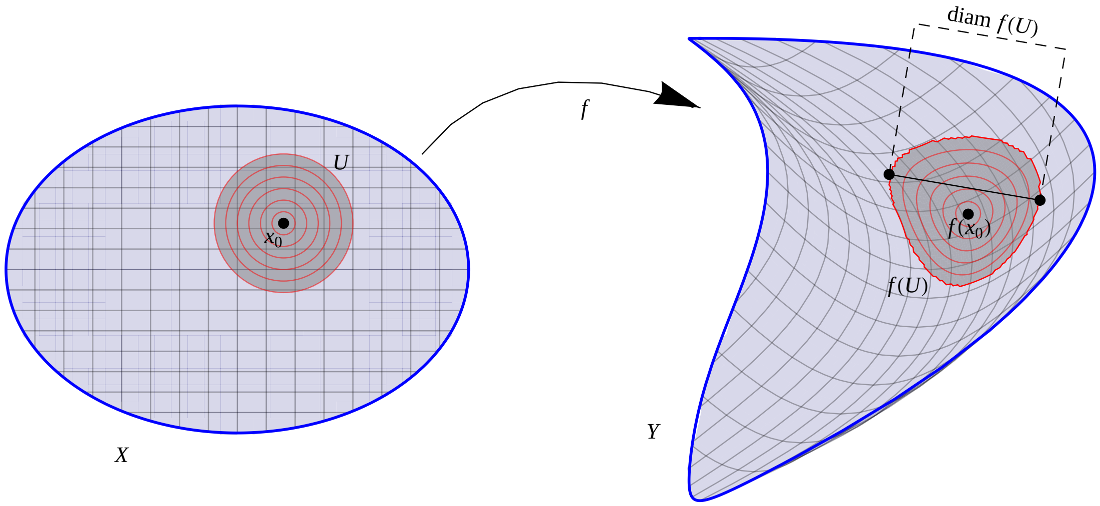

Figura 4: Come si definisce l'oscillazione di una funzione in un punto

2 Oscillazione di una funzione

Definizione 2.1. Sia Y uno spazio metrico con metrica d ed $A \subseteq Y$. Chiameremo *diametro* di A la quantità

$$\operatorname{diam} A := \sup\{d(x,y) \mid x, y \in A\}. \tag{6}$$

Se $A \neq \varnothing$ allora $0 \leq \operatorname{diam} A \leq +\infty$. Il diametro è finito se e solo se A è limitato (e non vuoto). Se $A \subseteq B$ allora $\operatorname{diam} A \leq \operatorname{diam} B$.

Esercizio 2.2. Siano $Y = \mathbb{R}^2$ con la metrica euclidea, $x_0, y_0 \in A \subset \mathbb{R}^2$ e $d(x_0, y_0) = \operatorname{diam} A$. Vero o falso: A è contenuto nel disco chiuso avente per diametro il segmento di estremi x_0 e y_0.

Definizione 2.3. Sia X uno spazio topologico, Y uno spazio metrico, $f\colon X \to Y$ una funzione e $x_0 \in X$. Chiameremo *oscillazione di f in x_0* la quantità

$$\operatorname{osc}(f, x_0) := \inf\{\operatorname{diam} f(U) \mid U \text{ intorno di } x_0\}. \tag{7}$$

L'oscillazione è un numero in $[0, +\infty]$ (se X è non vuoto).

Proposizione 2.4. *La f è continua in x_0 se e solo se* $\operatorname{osc}(f, x_0) = 0$.

Dimostrazione. Supponiamo che f sia continua in x_0 e sia $\varepsilon > 0$. Esiste un intorno U_ε di x_0 tale che $x \in U_\varepsilon \Rightarrow d(f(x), f(x_0)) < \varepsilon$. Ma allora $f(U_\varepsilon)$ è contenuta nella palla $B_Y(x_0, \varepsilon)$ di centro $f(x_0)$ e raggio ε, il cui diametro è $\leq 2\varepsilon$. Quindi

$$0 \leq \operatorname{osc}(f, x_0) \leq \operatorname{diam} f(U_\varepsilon) \leq \operatorname{diam} B_Y\big(f(x_0), \varepsilon\big) \leq 2\varepsilon.$$

Valendo questo per ogni $\varepsilon > 0$ deduciamo che $\operatorname{osc}(f, x_0) = 0$.

Viceversa, supponiamo che $\operatorname{osc}(f, x_0) = 0$ e sia di nuovo $\varepsilon > 0$. Per definizione di oscillazione, esiste un intorno U_ε di x_0 tale che $\operatorname{diam} f(U_\varepsilon) < \varepsilon$. Se $x \in U_\varepsilon$ abbiamo che $d(f(x), f(x_0)) \leq \operatorname{diam} f(U_\varepsilon) < \varepsilon$. Concludiamo che f è continua in x_0. □

Proposizione 2.5. *Dato $\alpha > 0$, l'insieme $\{x \in X \mid \operatorname{osc}(f, x) < \alpha\}$ è aperto in X.*

Dimostrazione. Sia x_0 un punto dell'insieme, cioè tale che $\operatorname{osc}(f, x_0) < \alpha$. Per definizione di oscillazione, esiste un intorno U_{x_0} di x_0 tale che $\operatorname{diam} f(U_{x_0}) < \alpha$. Sia V la parte interna di U_{x_0}. Se $x \in V$ allora U_{x_0} è un intorno anche di x, e quindi

$$\operatorname{osc}(f, x) = \inf\bigl\{\operatorname{diam} f(U) \mid U \text{ intorno di } x\bigr\} \leq \operatorname{diam} f(U_{x_0}) < \alpha\,.$$

Quindi $V \subseteq \{x \in X \mid \operatorname{osc}(f, x) < \alpha\}$. Ogniqualvolta l'insieme $\{x \in X \mid \operatorname{osc}(f, x) < \alpha\}$ contiene un punto x_0, contiene anche tutto un intorno V del punto. Concludiamo che l'insieme è aperto. □

Esercizio 2.6. Indagare se è vero o falso che $\operatorname{osc}(f, x_0) = \lim_{r\to 0^+} \operatorname{diam} f(B_r(x_0))$, dove $B_r(x_0)$ è la palla aperta di centro x_0 e raggio r.

Esercizio 2.7. Quando $Y = \mathbb{R}$ c'è una semplice relazione fra le quantità

$$\operatorname{osc}(f, x_0), \qquad f(x_0)\,, \qquad \min_{x\to x_0}\lim f(x)\,, \qquad \max_{x\to x_0}\lim f(x)\,.$$

3 Il teorema

Teorema 3.1. *Sia $f\colon \mathbb{R}^N \to \mathbb{R}$ una funzione limitata e nulla al di fuori di un limitato. Allora si equivalgono le condizioni seguenti:*

- *f è integrabile secondo Riemann;*
- *l'insieme dei punti di discontinuità di f è trascurabile per la misura di Lebesgue.*

Se valgono le condizioni, allora f è misurabile e integrabile anche secondo Lebesgue e gli integrali secondo Riemann e secondo Lebesgue coincidono.

Dimostrazione. **Prima parte.** Cominciamo supponendo che f sia integrabile secondo Riemann. Dobbiamo dimostrare che l'insieme

$$\{x \in \mathbb{R}^N \mid f \text{ è discontinua in } x\} = \bigl\{x \in \mathbb{R}^N \bigm| \operatorname{osc}(f, x) > 0\bigr\}$$

è trascurabile secondo Lebesgue. Poiché

$$\begin{aligned}\bigl\{x \in \mathbb{R}^N \bigm| \operatorname{osc}(f, x) > 0\bigr\} &= \Bigl\{x \in \mathbb{R}^N \Bigm| \exists k \in \mathbb{N} \text{ t.c. } \operatorname{osc}(f, x) \geq \frac{1}{k}\Bigr\} = \\ &= \bigcup_{k\in\mathbb{N}} \Bigl\{x \in \mathbb{R}^N \Bigm| \operatorname{osc}(f, x) \geq \frac{1}{k}\Bigr\}\,,\end{aligned}$$

basterà dimostrare che fissato $\varepsilon > 0$ l'insieme

$$\bigl\{x \in \mathbb{R}^N \bigm| \operatorname{osc}(f, x) \geq \varepsilon\bigr\} \tag{8}$$

è trascurabile. Notare che si tratta di un insieme chiuso, e quindi misurabile secondo Lebesgue. Ora, per definizione di integrabilità secondo Riemann, per ogni $n > 0$ esistono funzioni a gradino g_n, h_n tali che

$$g_n \leq f \leq h_n \qquad \text{e} \qquad \int_{\mathbb{R}^N} (h_n - g_n)\, d\lambda < \frac{1}{n}\,. \tag{9}$$

Sia M l'unione delle frontiere dei rettangoli che definiscono le funzioni a gradino g_n, h_n. Essendo l'unione di una quantità numerabile di rettangoli con almeno un lato di lunghezza nulla, M è trascurabile secondo Lebesgue. Possiamo scrivere

$$\big\{x \in \mathbb{R}^N \;\big|\; \mathrm{osc}(f,x) \geq \varepsilon\big\} \subseteq M \cup \big\{x \in \mathbb{R}^N \setminus M \;\big|\; \mathrm{osc}(f,x) \geq \varepsilon\big\} \tag{10}$$

Basterà dimostrare che il nuovo insieme

$$\big\{x \in \mathbb{R}^N \setminus M \;\big|\; \mathrm{osc}(f,x) \geq \varepsilon\big\} \tag{11}$$

è trascurabile. Prendiamo un x_0 in questo insieme, cioè tale che $x_0 \notin M$ e $\mathrm{osc}(f, x_0) \geq \varepsilon$. Ognuna delle funzioni a gradino g_n, h_n è localmente costante in x_0 (Figura 5). Quindi, fissato n, esiste un intorno U_n di x_0 tale che g_n e h_n sono costanti su U_n. Poiché $g_n \leq f \leq h_n$, abbiamo che

$$g_n(x_0) = g_n(x) \leq f(x) \leq h_n(x) = h_n(x_0) \qquad \forall x \in U_n\,, \tag{12}$$

per cui

$$f(U_n) \subseteq [g_n(x_0), h_n(x_0)] \tag{13}$$

e quindi

$$\varepsilon \leq \mathrm{osc}(f, x_0) \leq \mathrm{diam}\, f(U_n) \leq \mathrm{diam}[g_n(x_0), h_n(x_0)] = h_n(x_0) - g_n(x_0)\,. \tag{14}$$

Abbiamo stabilito che

$$\big\{x \in \mathbb{R}^N \setminus M \;\big|\; \mathrm{osc}(f,x) \geq \varepsilon\big\} \subseteq \big\{x \in \mathbb{R}^N \mid h_n(x) - g_n(x) \geq \varepsilon\big\} = \{h_n - g_n \geq \varepsilon\}\,. \tag{15}$$

Stimiamo ora la misura di quest'ultimo insieme, con la tecnica della disuguaglianza di Čebyšëv:

$$h_n(x) - g_n(x) \geq \varepsilon \qquad \Longleftrightarrow \qquad 1 \leq \frac{h_n(x) - g_n(x)}{\varepsilon}\,,$$

da cui

$$\chi_{\{(h_n - g_n)/\varepsilon \geq 1\}}(x) \leq \frac{h_n(x) - g_n(x)}{\varepsilon} \qquad \forall x\,. \tag{16}$$

Integrando

$$\begin{aligned}\lambda\big(\big\{x \in \mathbb{R}^N \setminus M \;\big|\; \mathrm{osc}(f,x) \geq \varepsilon\big\}\big) \leq \lambda\big(\{h_n - g_n \geq \varepsilon\}\big) = \int_{\mathbb{R}^N} \chi_{\{h_n - g_n \geq \varepsilon\}}\, d\lambda \leq \\ \leq \int_{\mathbb{R}^N} \frac{h_n - g_n}{\varepsilon}\, d\lambda = \frac{1}{\varepsilon} \int_{\mathbb{R}^N} (h_n - g_n)\, d\lambda < \frac{\varepsilon}{n}\,.\end{aligned}$$

Mandando $n \to +\infty$ otteniamo

$$\lambda\big(\big\{x \in \mathbb{R}^N \setminus M \;\big|\; \mathrm{osc}(f,x) \geq \varepsilon\big\}\big) = 0\,, \tag{17}$$

che era quanto serviva.

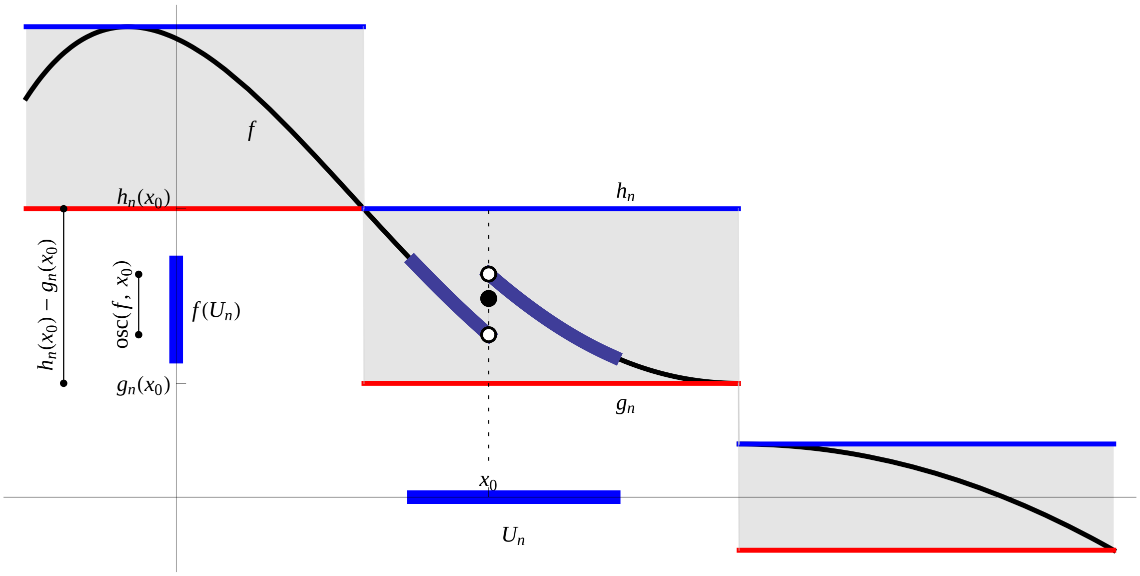

Figura 5: Oscillazione in x_0 e approssimanti a gradino

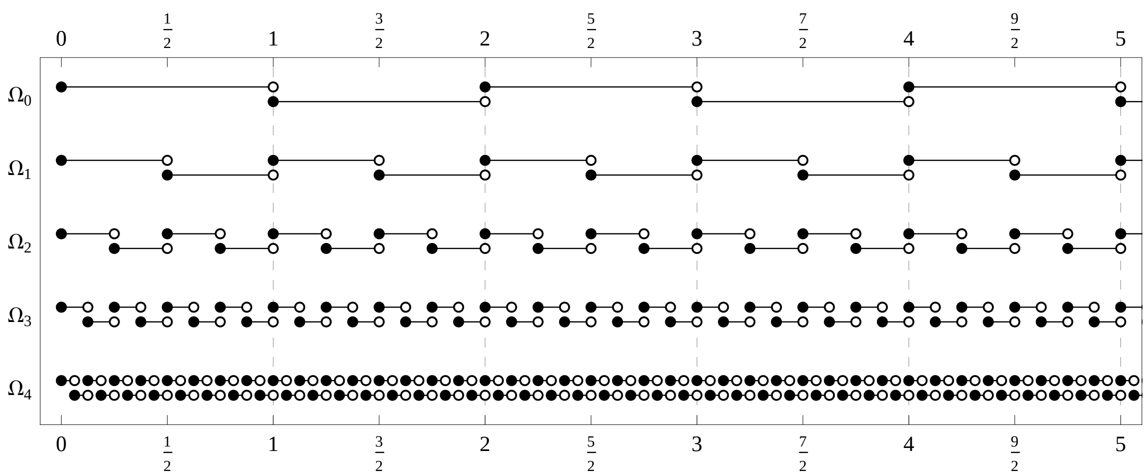

Figura 6: Cubetti standard in dimensione 1 secondo la formula (18), alternativamente sfalsati in modo da renderli distinguibili

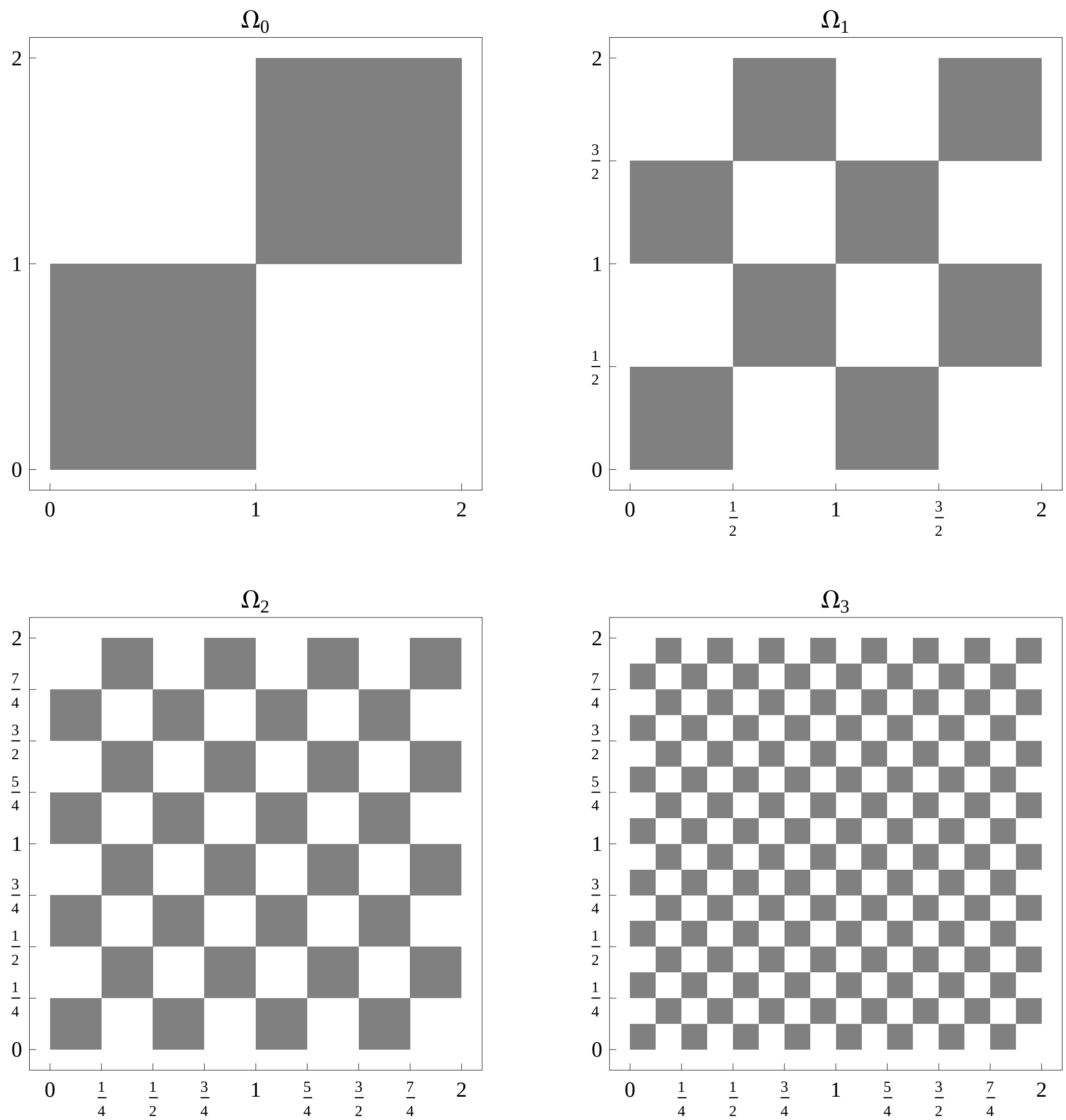

Figura 7: Cubetti standard in dimensione 2 secondo la formula (18)

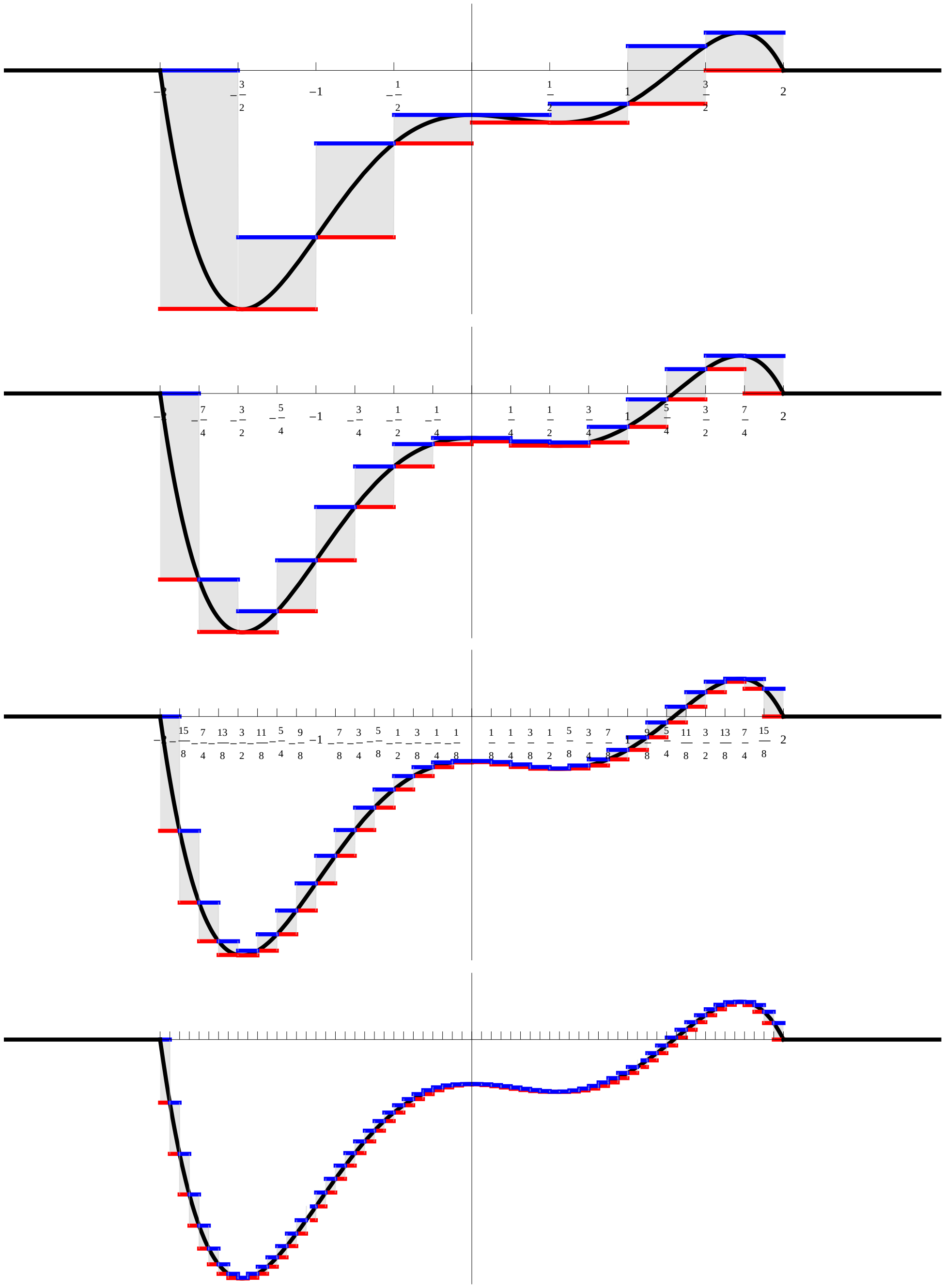

Figura 8: Gli approssimanti successivi g_n, h_n della funzione in nero secondo la formula (19) in dimensione 1

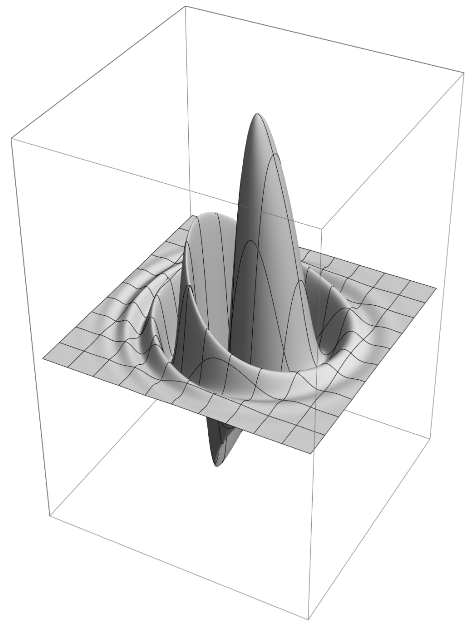

Figura 9: Il grafico di una funzione di due variabili, le cui approssimanti sono in figura 10

Seconda parte. Supponiamo ora che f sia continua in quasi ogni punto. Per $n \in \mathbb{N}$, sia Ω_n l'insieme dei cubetti standard di lato $1/2^n$:

$$\Omega_n := \left\{ x + [0, 2^{-n}[^N \;\middle|\; x \in 2^{-n}\mathbb{Z} \right\}. \tag{18}$$

Vedi le figure 6 e 7 rispettivamente a pag. 7 e pag. 8. Lo spazio $\mathbb{R}^N$ è l'unione disgiunta dei cubetti di Ω_n. Inoltre dati due cubetti di lati diversi $Q_1 \in \Omega_n$, $Q_2 \in \Omega_m$, o sono disgiunti o quello di lato minore è contenuto nell'altro. Come candidate approssimanti scegliamo

$$g_n := \sum_{Q \in \Omega_n} \Bigl(\inf_Q f\Bigr)\chi_Q\,, \qquad h_n := \sum_{Q \in \Omega_n} \Bigl(\sup_Q f\Bigr)\chi_Q\,. \tag{19}$$

Vedi la figura 8 per il caso di una variabile, e le figure 9 e 10 per due variabili. Gli estremi inferiori e superiori sono finiti perché f è limitata. Le due somme sono numerabili, ma soltanto un numero finito di addendi può essere non nullo, perché soltanto un numero finito di cubetti di Ω_n interseca la regione limitata in cui f è non nulla. Quindi g_n ed h_n sono funzioni *a gradino*. Inoltre

$$x \in Q \in \Omega_n \quad \Rightarrow \quad g_n(x) = \inf_Q f \quad \text{e} \quad h_n(x) = \sup_Q f\,. \tag{20}$$

per cui in particolare

$$g_n \le g_{n+1} \le f \le h_{n+1} \le h_n\,. \tag{21}$$

In figura 8 si vedono g_n, h_n (rosso e blu) per $n = 1, \ldots, 4$. Vogliamo dimostrare che

$$\int_{\mathbb{R}^N} (h_n - g_n)\, d\lambda \to 0 \quad \text{per } n \to +\infty\,, \tag{22}$$

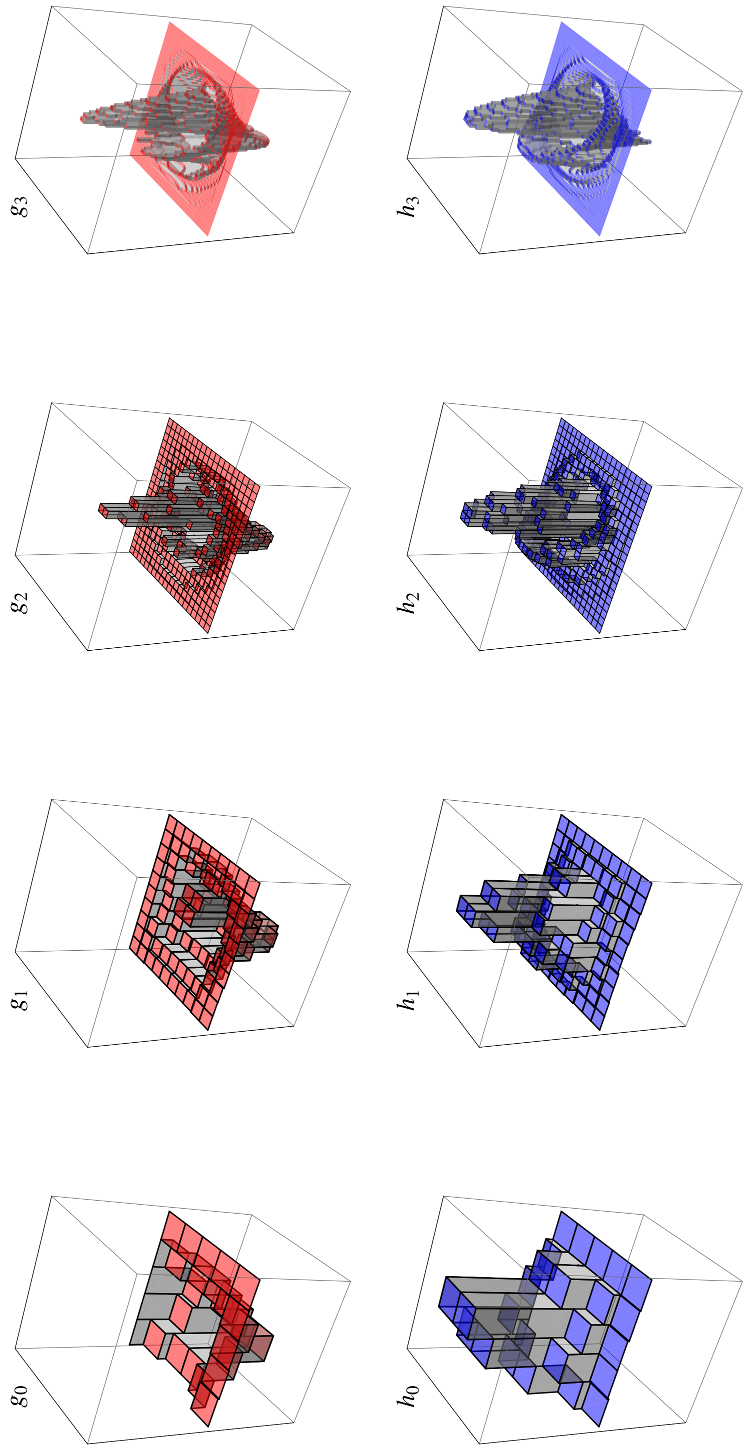

Figura 10: Gli approssimanti successivi g_n, h_n della funzione di due variabili di figura 9 secondo la formula (19)

e lo faremo usando il teorema della convergenza dominata. Una dominazione è la seguente:

$$0 \le h_n - g_n \le h_0 - g_0 \in L^1(\mathbb{R}^N)\,.$$

Sia ora x_0 un punto in cui f è continua e sia $\varepsilon > 0$. Esiste $\delta_\varepsilon > 0$ tale che

$$\|x - x_0\| < \delta_\varepsilon \quad \Rightarrow \quad f(x_0) - \varepsilon < f(x) < f(x_0) + \varepsilon\,.$$

Sia $n_\varepsilon \in \mathbb{N}$ tale che il diametro dei cubetti di lato $1/2^{n_\varepsilon}$ sia minore di δ. Se $n \ge n_\varepsilon$

$$x_0, x \in Q \in \Omega_n \quad \Rightarrow \quad \|x - x_0\| < \delta_\varepsilon \quad \Rightarrow \quad f(x_0) - \varepsilon < f(x) < f(x_0) + \varepsilon\,.$$

Quindi se $n \ge n_\varepsilon$

$$x_0 \in Q \in \Omega_n \quad \Rightarrow \quad f(x_0) - \varepsilon \le \inf_Q f \le \sup_Q f \le f(x_0) + \varepsilon$$

cioè

$$n \ge n_\varepsilon \quad \Rightarrow \quad f(x_0) - \varepsilon \le g_n(x_0) \le h_n(x_0) \le f(x_0) + \varepsilon\,.$$

Insomma, se f è continua in x_0 allora $g_n(x_0)$ e $h_n(x_0)$ tendono entrambe a $f(x_0)$ per $n \to +\infty$. Poiché f è continua in quasi ogni punto, la successione di funzioni $h_n - g_n$ tende a zero quasi ovunque. Per la convergenza dominata possiamo concludere che effettivamente vale la relazione (22). Concludiamo che f è integrabile secondo Riemann.

Terza parte. Riprendendo la formula (21), vediamo che la successione di funzioni g_n converge puntualmente ovunque a una funzione h_∞, essendo monotona debolmente crescente rispetto a n e limitata superiormente da h_0. Questo limite puntuale è misurabile secondo Lebesgue (anzi, è boreliano). Abbiamo visto che quasi ovunque questo limite coincide con f. Quindi f è misurabile perché coincide quasi ovunque con una funzione misurabile (lo spazio di misura di Lebesgue è completo). È anche sommabile perché compresa fra le funzioni g_0 ed h_0. Analogamente h_n tende puntualmente decrescendo a una funzione h_∞, che coincide quasi ovunque con g_∞ e con f. Per convergenza dominata gli integrali di g_n e di h_n tendono all'integrale (secondo Lebesgue) di g_∞. Poiché

$$\int_{\mathbb{R}^N} g_n \, d\lambda \le \int_{\mathbb{R}^N} g_\infty \, d\lambda = \int_{\mathbb{R}^N} f \, d\lambda = \int_{\mathbb{R}^N} h_\infty \, d\lambda \le \int_{\mathbb{R}^N} h_n \, d\lambda$$

e dato che il primo e l'ultimo membro tendono entrambi a $\int_{\mathbb{R}^N} f \, d\lambda$, per definizione di integrale inferiore e superiore secondo Riemann deduciamo che

$$\operatorname*{int\,inf}_{\mathbb{R}^N} f = \int_{\mathbb{R}^N} f \, d\lambda = \operatorname*{int\,sup}_{\mathbb{R}^N} f\,,$$

cioè che l'integrale di Riemann e di Lebesgue di f coincidono. □