La derivata seconda di ϕ in x_0 può essere calcolata usando la regola di derivazione della funzione composta, così come abbiamo fatto vedere per la derivata prima nell'Osservazione 3.28. Sia infatti $h\colon I = (x_0 - h, x_0 + h) \to \mathbb{R}$ la funzione composta definita in (3.13). Abbiamo già osservato che, essendo h costante nell'intervallo I, la sua derivata, che abbiamo calcolato in (3.14) utilizzando la regola di derivazione della funzione composta, è nulla in tale intervallo.
Supponiamo ora che f sia di classe C^2, in maniera tale che ϕ sia anch'essa di classe C^2 in I. Poiché $h'(x) = 0$ per ogni $x \in I$, avremo anche che $h''(x) = 0$ per ogni $x \in I$. D'altra parte, $h''(x)$ può essere calcolata a partire da (3.14) usando la regola di derivazione della funzione composta;

$$h''(x) = f_{xx} + f_{xy}\,\phi' + (f_{yx} + f_{yy}\phi')\phi' + f_y\,\phi'' = 0, \qquad \forall x \in I; \tag{3.16}$$

per semplificare le notazioni, abbiamo tralasciato l'argomento $(x, \phi(x))$ in tutte le derivate parziali di f e l'argomento (x) nelle derivate di ϕ. Da questa equazione è possibile ricavare ϕ''; in particolare, nel punto $x = x_0$ si ha

$$\phi''(x_0) = -\frac{f_{xx}(x_0, y_0) + 2f_{xy}(x_0, y_0)\,\phi'(x_0) + f_{yy}(x_0, y_0)\,[\phi'(x_0)]^2}{f_y(x_0, y_0)}\,. \tag{3.17}$$

✂ 3.5 Metodo dei moltiplicatori di Lagrange

In questo paragrafo faremo vedere come il Teorema 3.27 della funzione implicita possa essere applicato all'analisi dei problemi di ottimizzazione vincolata. Abbiamo già visto, nel Paragrafo 3.3, come si possono trattare i problemi di ottimizzazione vincolata nel caso in cui il vincolo Γ sia il sostegno di una assegnata curva regolare γ, oppure il grafico di una funzione $y = \phi(x)$ (o anche $x = \psi(y)$). Esporremo ora un altro modo, noto come **metodo dei moltiplicatori di Lagrange**, per l'analisi dei problemi di ottimizzazione vincolata nel caso in cui il vincolo Γ sia definito in forma implicita.

Esempio 3.32. Riprendiamo l'Esempio 3.24, nel quale abbiamo determinato i punti di estremo assoluto della funzione $f(x, y) = x^3 + 3y$ nel cerchio unitario K centrato nell'origine. Abbiamo già osservato che tali punti di estremo, che esistono per il teorema di Weierstrass, devono necessariamente stare sulla frontiera Γ di K, dal momento che f non ha punti stazionari interni a K. La frontiera Γ di K è la circonferenza unitaria centrata nell'origine, che può essere rappresentata come l'insieme di livello zero della funzione

$$g(x, y) = x^2 + y^2 - 1\,.$$

Ricordiamo che l'unico punto di massimo assoluto di f in Γ è $(0,1)$, mentre l'unico punto di minimo assoluto è $(0,-1)$. Come abbiamo già detto, il gradiente di f in questi due punti non si annulla, e si ha $\nabla f(0,1) = \nabla f(0,-1) = (0,3)$. Osserviamo anche che $\nabla g(x,y) = (2x, 2y)$ e dunque $\nabla g(0,1) = (0,2)$, $\nabla g(0,-1) = (0,-2)$. In entrambi i punti il gradiente di f è quindi parallelo a quello di g. Come dimostreremo fra breve nel Teorema 3.33, questo non è un caso.

Sia data una funzione $f\colon A \to \mathbb{R}$ di classe C^1 in un insieme aperto $A \subseteq \mathbb{R}^2$, di cui vogliamo determinare i punti di estremo (relativo o assoluto) su un vincolo Γ che è dato in forma implicita, cioè del tipo

$$\Gamma = \{(x,y) \in A;\ g(x,y) = c\}\,,$$

dove $g\colon A \to \mathbb{R}$ è una funzione di classe C^1. Come abbiamo accennato nell'Esempio 3.30, anche quando Γ è rappresentabile localmente come sostegno di una curva, non è detto che si possa scrivere in maniera esplicita questa parametrizzazione. Nell'Esempio 3.24 questo è stato possibile, poiché il vincolo era abbastanza semplice. Per un vincolo del tipo

$$(1+x)^6 y + \log(1 + x^2 + y^2) = 0$$

risulta invece praticamente impossibile scrivere una parametrizzazione esplicita anche solo in un intorno dell'origine.

Il metodo dei moltiplicatori di Lagrange fornisce una condizione necessaria affinché un punto sia di estremo vincolato, evitando di scrivere la parametrizzazione esplicita del vincolo.

Teorema 3.33 (Moltiplicatori di Lagrange). *Siano $f, g\colon A \to \mathbb{R}$ due funzioni di classe C^1 nell'aperto $A \subset \mathbb{R}^2$ e sia*

$$\Gamma = \{(x,y) \in A;\ g(x,y) = c\}\,.$$

Sia $(x_0, y_0) \in \Gamma$ un punto di estremo relativo vincolato di f su Γ. Allora il determinante della matrice 2×2 che ha come righe i gradienti delle funzioni f e g in (x_0, y_0) è nullo, cioè

$$f_x(x_0,y_0)\, g_y(x_0,y_0) - f_y(x_0,y_0)\, g_x(x_0,y_0) = 0\,. \tag{3.18}$$

In altri termini, i vettori $\nabla f(x_0,y_0)$ e $\nabla g(x_0,y_0)$ sono paralleli.

Dimostrazione. Se $\nabla g(x_0, y_0) = (0,0)$ allora la condizione (3.18) è evidentemente verificata. Supponiamo quindi $\nabla g(x_0, y_0) \neq (0,0)$; per fissare le idee, assumiamo ad esempio che $g_y(x_0, y_0) \neq 0$ (l'altro caso si tratta in maniera analoga). Possiamo applicare il Teorema della Funzione Implicita: esistono quindi un rettangolo $Q = (x_0 - h, x_0 + h) \times (y_0 - k, y_0 + k)$ contenuto in A e una funzione $\phi \colon (x_0 - h, x_0 + h) \to (y_0 - k, y_0 + k)$ di classe C^1, tale che il grafico di ϕ coincide con $\Gamma \cap Q$. Inoltre sappiamo che

$$\phi(x_0) = y_0, \quad \phi'(x_0) = -\frac{g_x(x_0, y_0)}{g_y(x_0, y_0)} \,. \tag{3.19}$$

Consideriamo dunque la restrizione

$$h(x) = f(x, \phi(x)), \qquad x \in (x_0 - h, x_0 + h)\,. \tag{3.20}$$

Per il Teorema 2.57 di derivazione della funzione composta, la derivata di h si può scrivere come

$$h'(x) = f_x(x, \phi(x)) + f_y(x, \phi(x))\, \phi'(x), \qquad x \in (x_0 - h, x_0 + h)\,. \tag{3.21}$$

Poiché (x_0, y_0) è, per ipotesi, un punto di estremo vincolato di f su Γ, avremo che x_0 è un punto di estremo della funzione di una variabile h. Per il Teorema di Fermat (si veda il Teorema 4.38 del primo volume), abbiamo che $h'(x_0) = 0$; utilizzando l'espressione (3.21) della derivata di h e la formula (3.19) si ha dunque

$$0 = h'(x_0) = f_x(x_0, y_0) + f_y(x_0, y_0)\, \phi'(x_0) = f_x(x_0, y_0) - f_y(x_0, y_0)\, \frac{g_x(x_0, y_0)}{g_y(x_0, y_0)} \,,$$

cioè $f_x(x_0, y_0)\, g_y(x_0, y_0) - f_y(x_0, y_0)\, g_x(x_0, y_0) = 0$, che è quanto si voleva dimostrare. $\square$

Osservazione 3.34. Ricordiamo che due vettori sono paralleli se e solo se sono proporzionali. Di conseguenza, il Teorema 3.33 afferma che in un punto $P = (x_0, y_0) \in \Gamma$ di estremo relativo vincolato di f su Γ tale che $\nabla g(P) \neq \mathbf{0}$, si ha $\nabla f(P) = \lambda \nabla g(P)$ per un opportuno $\lambda \in \mathbb{R}$. Il numero λ prende il nome di moltiplicatore di Lagrange. Se anche $\nabla f(P) \neq \mathbf{0}$, allora per il Teorema 3.27 della funzione implicita abbiamo che, in un intorno di P, sia l'insieme di livello l_f di f a quota $f(P)$ che il vincolo Γ sono localmente il supporto di curve regolari. Inoltre, $\nabla f(P)$ e $\nabla g(P)$ sono ortogonali rispettivamente a l_f e Γ nel punto P. Di conseguenza, la condizione di parallelismo fra i due gradienti si traduce in una condizione di tangenza fra l_f e Γ nel punto P (cfr. Osservazione 3.29), come è mostrato in Figura 3.11.

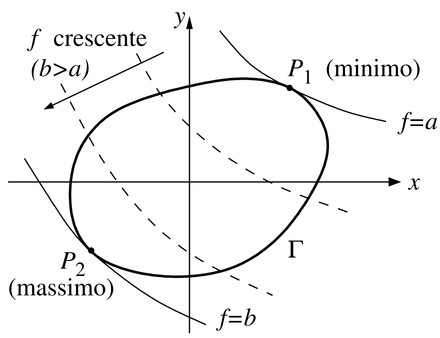

Figura 3.11: Significato geometrico del Teorema 3.33

Osservazione 3.35. I punti $(x_0, y_0) \in \Gamma$ soddisfacenti la relazione (3.18) e tali che $\nabla g(x_0, y_0) \neq (0,0)$ sono punti stazionari vincolati di f su Γ. Infatti, in tal caso la restrizione $h(x)$ di f a Γ (in un intorno di (x_0, y_0)), definita in (3.20), ha derivata prima che si annulla in x_0. Il Teorema 3.33 dice quindi che i punti di estremo relativo vincolato sono punti stazionari vincolati. Osserviamo esplicitamente che tali punti stazionari vincolati sono dunque da ricercare fra le soluzioni del sistema

$$\begin{cases} g(x_0, y_0) = c\,, \\ f_x(x_0, y_0)\, g_y(x_0, y_0) - f_y(x_0, y_0)\, g_x(x_0, y_0) = 0\,, \end{cases} \tag{3.22}$$

dove la prima equazione è esattamente la condizione di appartenenza del punto (x_0, y_0) al vincolo Γ, mentre la seconda è la condizione di stazionarietà.

Esempio 3.36. Vogliamo determinare i punti di estremo assoluto della funzione $f(x,y) = x^2 + 4y^2 - 4x - 16y$ sull'ellisse

$$\Gamma = \{(x,y) \in \mathbb{R}^2;\ x^2 + 4y^2 = 1\}\,.$$

La funzione f è continua sull'insieme chiuso e limitato Γ. Per il Teorema 3.7 di Weierstrass essa ha quindi punti di massimo e minimo assoluto su Γ. Poiché f è di classe C^1 in $\mathbb{R}^2$ e Γ è l'insieme di livello zero della funzione di classe C^1

$$g(x,y) = x^2 + 4y^2 - 1\,, \qquad (x,y) \in \mathbb{R}^2,$$

possiamo usare il metodo dei moltiplicatori di Lagrange per identificare i punti stazionari vincolati. Abbiamo che $f_x = 2x - 4$, $f_y = 8y - 16$, $g_x = 2x$, $g_y = 8y$, quindi il sistema (3.22) diviene

$$\begin{cases} x^2 + 4y^2 - 1 = 0, \\ 8y(2x-4) - 2x(8y-16) = 0, \end{cases} \iff \begin{cases} x^2 + 4y^2 - 1 = 0, \\ y = x. \end{cases}$$

Sostituendo $y = x$ nella prima equazione otteniamo $5x^2 = 1$, cioè $x = \pm 1/\sqrt{5}$. Otteniamo i due punti stazionari

$$P_1 = \left(\frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right), \qquad P_2 = \left(-\frac{1}{\sqrt{5}}, -\frac{1}{\sqrt{5}}\right).$$

Calcoliamo f su questi due punti:

$$f(P_1) = 1 - 4\sqrt{5}, \qquad f(P_2) = 1 + 4\sqrt{5},$$

da cui possiamo concludere che P_1 è l'unico punto di minimo assoluto di f su Γ, mentre P_2 è l'unico punto di massimo assoluto di f su Γ.

Riprendiamo ora il problema di ottimizzazione generale, già esposto nel Paragrafo 3.3. Supponiamo cioè di essere nella situazione $f\colon A \to \mathbb{R}$ di classe C^1 in un aperto $A \subset \mathbb{R}^2$, e $K \subset A$ chiuso e limitato. Per il Teorema 3.7 di Weierstrass, f ammette punti di estremo assoluto in K. Tali punti di estremo assoluto vanno ricercati fra:

1. i punti stazionari di f interni a K;
2. i punti regolari della frontiera Γ di K, che siano punti stazionari vincolati: questi punti possono essere determinati o parametrizzando esplicitamente il vincolo, oppure utilizzando il metodo dei moltiplicatori di Lagrange;
3. i punti in cui Γ non è regolare: se Γ è l'unione di un numero finito di curve regolari, i punti in cui Γ non è regolare sono tipicamente gli estremi di tali curve regolari.

Esempio 3.37. Riprendiamo l'Esempio 3.24 e applichiamo il metodo dei moltiplicatori di Lagrange per lo studio dei punti di estremo vincolati. Il vincolo e alcuni insiemi di livello di f sono mostrati in Figura 3.12 (l'insieme di livello c della funzione f è il grafico della funzione $y = (c - x^3)/3$).
Cerchiamo i punti stazionari vincolati risolvendo il sistema (3.22):

$$\begin{cases} x^2 + y^2 = 1, \\ 6x^2y - 6x = 0, \end{cases} \qquad \Longleftrightarrow \qquad \begin{cases} x^2 + y^2 = 1, \\ x(xy - 1) = 0. \end{cases}$$

La seconda equazione è soddisfatta in due casi:

(a) $x = 0$; dalla prima equazione ricaviamo $y^2 = 1$, cioè $y = \pm 1$. Si ottengono quindi due punti stazionari vincolati, $P_1 = (0, 1)$ e $P_2 = (0, -1)$.

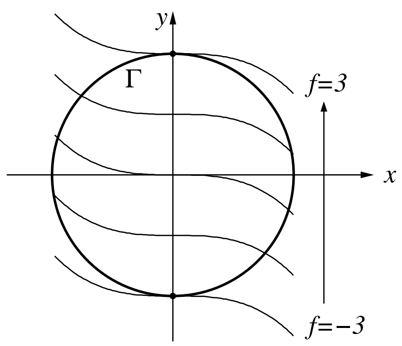

Figura 3.12: Insiemi di livello e vincolo dell'Esempio 3.24

(b) $xy - 1 = 0$; possiamo ricavare, ad esempio, $y = 1/x$, con $x \neq 0$. Sostituendo nella prima equazione del sistema si ottiene $x^2 + 1/x^2 = 1$, cioè $x^4 - x^2 + 1 = 0$. Si verifica facilmente che quest'ultima equazione biquadratica non ammette soluzioni reali.

Abbiamo già osservato che f ammette sicuramente, per il Teorema di Weierstrass, punti di massimo e minimo assoluto in K e che questi devono trovarsi su Γ, cioè vanno ricercati fra i punti stazionari vincolati. Valutando la funzione f su tutti i punti stazionari vincolati prima determinati si ha

$$f(P_1) = 3, \qquad f(P_2) = -3,$$

da cui concludiamo che $(0, 1)$ e $(0, -1)$ sono rispettivamente punto di massimo e minimo assoluto di f in K.

Esempio 3.38. Vogliamo determinare i punti di estremo assoluto della funzione $f(x, y) = x^2 + y^2 - 2y$ nell'insieme

$$K = \{(x, y) \in \mathbb{R}^2;\ x^2 + y^2 \leq 4,\ y \geq 0\}.$$

La funzione f è di classe C^1 su tutto $\mathbb{R}^2$; l'insieme K è un semicerchio chiuso (si veda la Figura 3.13). Per il Teorema 3.7 di Weierstrass la funzione f ammette punti di massimo e minimo assoluto in K, e tali punti vanno ricercati fra:

(1) I punti interni di K sui quali si annulli il gradiente di f: poiché abbiamo $\nabla f(x, y) = (2x, 2y - 2)$, l'unico punto dove si annulla il gradiente di f è $P_1 = (0, 1)$, che è interno a K.

(2) I punti regolari della frontiera Γ di K, che siano punti stazionari vincolati: Γ è l'unione di due archi regolari (la semicirconferenza superiore e il segmento

$(-2,2)$ sull'asse x). Per cercare i punti stazionari sulla semicirconferenza possiamo usare il metodo dei moltiplicatori di Lagrange; posto $g(x,y) = x^2+y^2-4$, determiniamo quindi le eventuali soluzioni del sistema (3.22), aventi $y > 0$:

$$\begin{cases} g = 0\,, \\ f_x\, g_y - f_y\, g_x = 0\,, \end{cases} \qquad \Longleftrightarrow \qquad \begin{cases} x^2 + y^2 - 4 = 0\,, \\ 4xy - 4x(y-1) = 0\,. \end{cases}$$

La seconda equazione è soddisfatta solo per $x = 0$; sostituendo nella prima si ottiene $y = \pm 2$. Ricordando che sulla semicirconferenza deve essere $y > 0$, otteniamo il punto stazionario vincolato $P_2 = (0,2)$.
Cerchiamo ora gli eventuali punti stazionari vincolati sul segmento che congiunge i punti $(-2,0)$ e $(2,0)$. Possiamo ancora applicare il metodo dei moltiplicatori di Lagrange, o più semplicemente considerare la restrizione

$$h(x) = f(x,0) = x^2, \qquad x \in (-2,2)\,.$$

Tale restrizione ha un punto di minimo per $x = 0$; otteniamo quindi il punto stazionario vincolato $P_3 = (0,0)$.
(3) I punti non regolari di Γ: essi sono $P_4 = (-2,0)$ e $P_5 = (2,0)$.

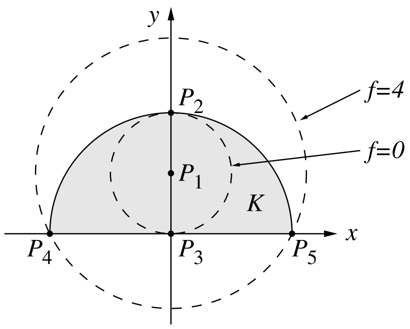

Figura 3.13: Esempio 3.38

Abbiamo dunque ottenuto cinque candidati a punti di estremo assoluto; calcoliamo f in questi punti:

$$f(P_1) = -1,\ f(P_2) = 0,\ f(P_3) = 0,\ f(P_4) = 4,\ f(P_5) = 4.$$

Di conseguenza, P_1 è l'unico punto di minimo assoluto di f, mentre P_4 e P_5 sono i due punti di massimo assoluto di f in K.

Osserviamo che l'esercizio poteva essere svolto anche in un modo più geometrico. Infatti la funzione f può essere scritta come

$$f(x,y) = x^2 + (y-1)^2 - 1 = d((x,y),P_1)^2 - 1\,,$$

dove $d((x,y),P_1)$ indica la distanza del punto (x,y) dal punto $P_1=(0,1)$. Di conseguenza, f assumerà massimo e minimo dove la distanza da P_1 assume massimo e minimo. È facile verificare, geometricamente, che i punti P_4 e P_5 sono i punti in K di massima distanza da P_1, e dunque sono i punti di massimo assoluto di f in K. Ovviamente il punto P_1 rappresenta invece il minimo della distanza, e quindi il minimo assoluto di f in K.

✂ 3.6 Estremi liberi in dimensione maggiore di due

Nel Paragrafo 3.2 abbiamo visto come si possono determinare i punti di estremo relativo di una funzione di due variabili reali. In questo paragrafo daremo alcuni cenni di come si può procedere quando si ha a che fare con una funzione $f\colon A\to\mathbb{R}$, dove A è un sottoinsieme di $\mathbb{R}^n$ con $n\geq 2$. Se f è derivabile parzialmente, il gradiente di f è un vettore n-dimensionale, le cui componenti sono le n derivate parziali di f, che indicheremo con $f_{x_1},\dots,f_{x_n}$. Analogamente, se f è di classe C^2, la matrice hessiana di f è una matrice quadrata simmetrica di ordine n avente per elementi le derivate parziali seconde $f_{x_ix_j}$, $i,j=1,\dots,n$.

Il Teorema 3.10 di Fermat vale anche in questo caso generale e la dimostrazione è identica al caso bidimensionale.

Teorema 3.39 (di Fermat). *Sia $f\colon A\to\mathbb{R}$ una funzione definita nell'insieme $A\subset\mathbb{R}^n$. Sia $\mathbf{x_0}$ un punto di estremo relativo di f in A; se $\mathbf{x_0}$ è un punto interno di A in cui f è derivabile parzialmente, allora $\nabla f(\mathbf{x_0})=\mathbf{0}$.*

Il Teorema 3.17 di classificazione dei punti stazionari è un po' più complicato da enunciare in dimensione maggiore di due. Come per il caso bidimensionale, anche in $\mathbb{R}^n$ possiamo definire le forme quadratiche, che sono polinomi omogenei di grado due nelle n variabili $x_1,\dots,x_n$; con notazione matriciale, una generica forma quadratica in $\mathbb{R}^n$ può essere scritta come

$$Q(\mathbf{h})=\mathbf{h}A\mathbf{h}^t=\sum_{i,j=1}^{n}a_{ij}h_ih_j,\qquad \mathbf{h}=(h_1,\dots,h_n)\in\mathbb{R}^n,$$

dove $A=(a_{ij})$, $i,j=1,\dots,n$, è una matrice quadrata di ordine n simmetrica. (Qui stiamo considerando il vettore $\mathbf{h}$ come una matrice $1\times n$ e con il simbolo $\mathbf{h}^t$ indichiamo la sua trasposta; in questo modo il prodotto righe per colonne $\mathbf{h}A\mathbf{h}^t$ è ben definito.) Anche in questo caso diremo che Q è una forma quadratica definita positiva [risp. negativa] se $Q(\mathbf{h})>0$ [risp. <0] per ogni $\mathbf{h}\neq\mathbf{0}$,