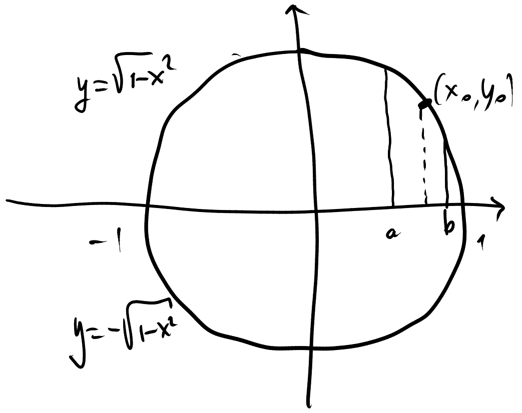
LE "FUNZIONI IMPLICITE"

Questa nota contiene un'introduzione ed alcuni risultati elementari su un classico problema, con importanti risvolti geometrici: esamineremo il caso più semplice. Si consideri $f(x,y) = x^2+y^2-1$ e si osservi che la "curva di livello 0", ossia l'insieme dei punti (x,y) del piano per i quali $f(x,y)=0$, è ben nota: è la circonferenza unitaria centrata nell'origine. Si consideri ora "l'equazione" della curva di livello $x^2+y^2-1=0$, e ci si ponga la domanda se essa sia o no il grafico di una funzione. La risposta è semplicissima: NO! Per ogni $x \in]-1,1[$ esistono due valori di y, $-\sqrt{1-x^2}$ e $\sqrt{1-x^2}$, per i quali $f(x,y)=0$, e dunque non c'è modo di associare univocamente una (sola) y ad ogni x del dominio (che, ragionevolmente, sarà $[-1,1]$). Scegliendo y come variabile indipendente, nulla si conclude: la situazione è identica, e identica resterebbe anche se si operasse una rotazione del sistema di assi. Dunque: gli insiemi di zeri di funzioni di più variabili NON sono, in generale, grafico di UNA funzione. Può capitare che lo siano: $y^3+x^2-1=0$ si può risolvere univocamente rispetto ad y (e non rispetto ad x) ottenendo la "formula risolutiva" $y=(1-x^2)^{1/3}$, definita per ogni $x\in\mathbb{R}$ (se si definisce la radice cubica come la funzione inversa della funzione $t\to t^3$, continua e strettamente crescente su $\mathbb{R}$), i punti del grafico della quale, $(x,(1-x^2)^{1/3})$ sono tutti e sole le soluzioni di $y^3+x^2-1=0$ (infatti: $\left[(1-x^2)^{1/3}\right]^3+x^2-1\equiv 0$ su $\mathbb{R}$ e, se vale $\bar{y}^3+\bar{x}^2-1=0$ allora vale anche $\bar{y}=\sqrt[3]{1-\bar{x}^2}$

In un linguaggio antico, ma ancora utilizzato, la funzione $x\to(1-x^2)^{1/3}$ si dice definita "IMPLICITAMENTE" dalla equazione $y^3+x^2-1=0$

Prima di entrare nel vivo del principale risultato di questa nota, il celebre teorema delle funzioni implicite di Ulisse Dini, occorre approfondire ancora un po' il discorso sull'esempio iniziale. E' vero che non può esistere nessuna funzione il grafico della quale coincide con la circonferenza unitaria ma, se ci si accontenta solo d'una porzione di circonferenza, la situazione muta radicalmente. Infatti, da $x^2+y^2-1=0$ segue $y^2=1-x^2$ che, per ogni $x\in[-1,1]$ ha le due soluzioni (formule risolutive) $y=\pm\sqrt{1-x^2}$ ma è del tutto evidente che, in vicinanza di un qualunque punto della circonferenza (x,y) con $y>0$ si troveranno altre soluzioni appartenenti solo al grafico di $y=\sqrt{1-x^2}$, mentre si DOVRA' scegliere $-\sqrt{1-x^2}$ per le soluzioni (x,y) vicine ad un punto (x_0,y_0) con $y_0<0$



Considerato un intorno di (x_0,y_0), $y_0>0$, tutto contenuto nel semipiano $\{y>0\}$, le uniche soluzioni sono del tipo $(x, \sqrt{1-x^2})$, grafico di $t \to \sqrt{1-t^2}$, e analogamente si potrebbe ragionare in un intorno abbastanza piccolo di (x_0,y_0) se $y_0<0$. Osserviamo invece che, se $y_0=0$, non c'è modo di evitare il doppio segno nella radice, per quanto piccolo si possa scegliere l'intorno: si può invece considerare la y come variabile indipendente e, ad esempio vicino a $(-1,0)$, osservare che $(-\sqrt{1-y^2}, y)$ descrive tutte le soluzioni di $x^2+y^2-1=0$ abbastanza vicino a $(-1,0)$.

Ci sono poi "esplicitazioni" impossibili: in $(0,0)$, l'insieme degli zeri di $f(x,y)=x^2-y^2$ non può in nessun modo essere rappresentato come un grafico, in quanto coincide con le due bisettrici dei quadranti $y=\pm x$, <u>ASSIEME</u>. Nessuna scelta del raggio dell'intorno migliora minimamente la situazione!

Alcune conclusioni preliminari:

- Non c'è motivo di attendersi che l'insieme di <u>tutti</u> gli zeri di una funzione $f(x,y)$ debba essere il grafico di un'<u>unica</u> funzione $y=\varphi(x)$, oppure $x = \psi(y)$

- È possibile che, anche se globalmente il luogo geometrico degli zeri di $f(x,y)$ non sia un grafico, la sua intersezione con un'intorno abbastanza piccolo di una soluzione nota (x_0, y_0) lo sia.
- È possibile che $f(x,y)=0$ sia risolubile in una sola delle forme $y = \varphi(x)$ oppure $x = \psi(y)$, ma non in entrambe (come accade nell'intorno di $(1,0)$, $(0,1)$, $(-1,0)$, $(0,-1)$ per $x^2+y^2-1=0$).
- Sotto quali condizioni su f almeno una delle due formule risolutive $y=\varphi(x)$ o $x=\psi(y)$ è valida, "vicino" a x_0 o a y_0, rispettivamente?

Una risposta soddisfacente è fornita dal teorema di Dini, che proveremo prima in ipotesi meno restrittive, ma di poco agevole verifica, e poi con altre più "pratiche". Chiariamo subito, però, che il teorema di Dini non è un teorema sull'esistenza degli zeri di $f(x,y)$, ma sulla struttura dell'insieme degli zeri vicino ad uno zero (x_0, y_0) già noto, che risulta essere quella di un grafico di una funzione.

TEOREMA (Dini): Siano $f : \Omega \to \mathbb{R}$, $\Omega \subseteq \mathbb{R}^2$, $(x_0, y_0) \in \mathbb{R}^2$ verificanti:

1) (x_0, y_0) è interno ad Ω

2) $f(x_0, y_0) = 0$

3) f continua in Ω

4) $t \to f(x,t)$ è strettamente crescente (in t) per ogni fissato x per cui $(x,t) \in \Omega$.

Allora, esistono $\delta > 0$ e $\varphi : [x_0-\delta, x_0+\delta] \to \mathbb{R}$ tali che:

5) $\varphi(x_0) = y_0$

6) $f(x, \varphi(x)) = 0 \quad \forall x \in [x_0-\delta, x_0+\delta]$

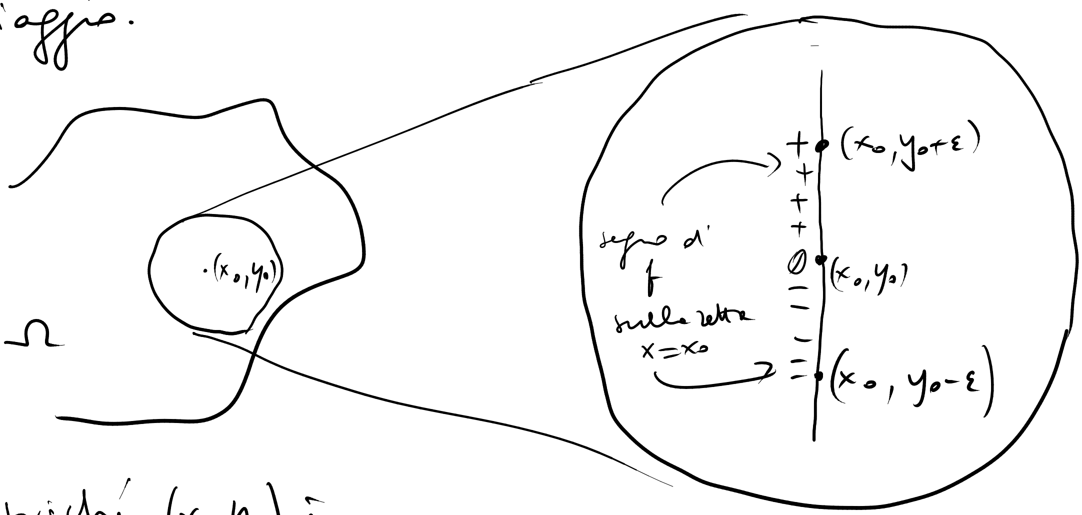
7) φ è continua in $[x_0-\delta, x_0+\delta]$

Prima di dimostrarlo osserviamo che la 2) dice che (x_0, y_0) è lo zero (in $\mathbb{R}^2$) di partenza, mentre 1) garantisce che ci sia spazio attorno ad esso nel dominio, il che è necessario nella prova. La continuità è il "minimo sindacale": ponendo $(x_0, y_0) = (0,0)$, $f(0,0) = 0$ e $f(x,y) = 1$ per $(x,y) \neq (0,0)$

il teorema è certamente falso. L'ultima ipotesi è più specifica: in $(1,0)$, ad esempio $y \to x^2+y^2-1$ non è strettamente crescente, in quanto ha un <u>minimo</u> (ogni punto di $(1,y)$ $y\neq 0$ ha distanza da $(0,0)$ maggiore di 1: cateto - ipotenusa).

La tesi mostra chiaramente il carattere <u>LOCALE</u> delle "formule risolutive" $y=\varphi(x)$, valida solo in $[x_0-\delta, x_0+\delta]$ per δ opportuno, e non scelto da noi a priori, come sarebbe stato in un teorema globale. Che φ sia la "formula risolutiva" dell'equazione $f(x,y)=0$ risulta dalla 6), mentre 5) dice che il punto iniziale (x_0,y_0) è parte del grafico di φ. La 7) sarà molto importante nella prova del teorema per le funzioni di classe C^1.

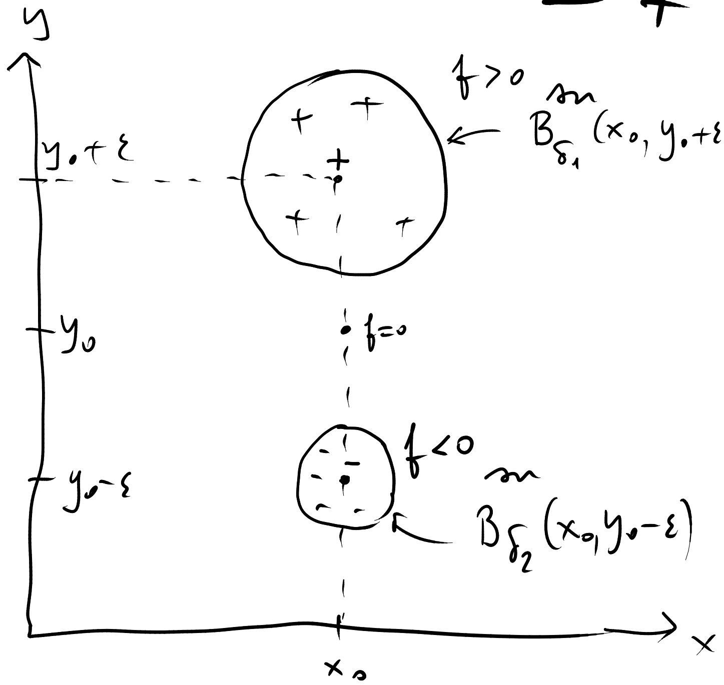
Premettiamo alla vera dimostrazione una specie di piano di viaggio.

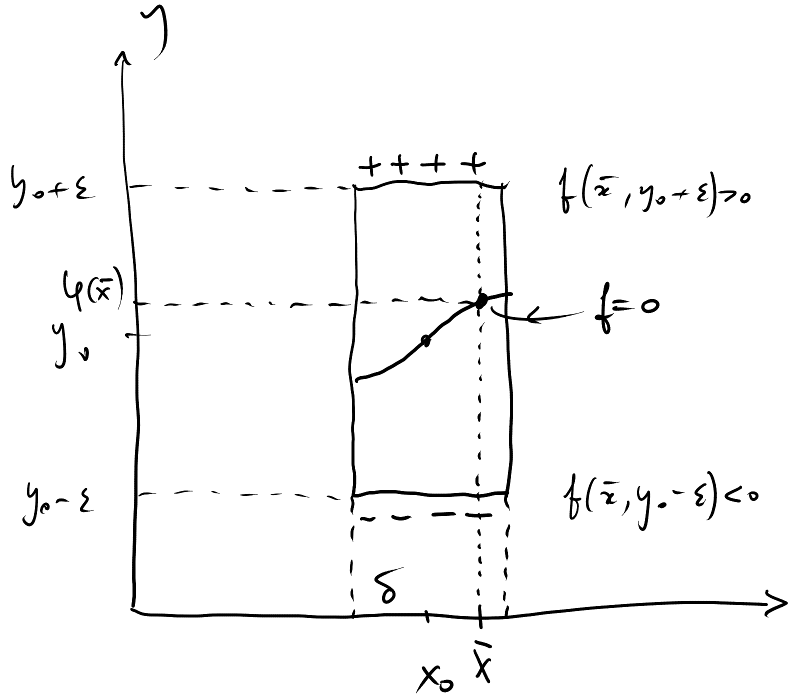


poiché (x_0,y_0) è interno ad Ω

poiché $t \to f(x_0,t)$ è strett. crescente

$$f(x_0,y_0-\varepsilon)<0 \quad e \quad f(x_0,y_0+\varepsilon)>0$$





il segno di f si mantiene costante, per permanenza di segno (continuità), in opportuni intorni di $(x_0, y_0+\varepsilon)$ e $(x_0, y_0-\varepsilon)$ di raggi δ_1 e δ_2.

scelto $\delta < \min(\delta_1, \delta_2)$ per ogni $x \in [x_0-\delta, x_0+\delta]$ si considera $t \to f(x,t)$ che è continua su $[y_0-\varepsilon, y_0+\varepsilon]$, assume valori discordi agli estremi (e quindi ha zeri) ed è strettamente crescente, e quindi lo zero è unico e compreso fra $y_0-\varepsilon$ e $y_0+\varepsilon$. Tale zero è $\varphi(\bar{x})$.

Vedremo che la continuità richiede qualche attenzione supplementare, ma come "road map" può bastare. Passiamo alla prova.

DIM. Poiché (x_0, y_0) è interno a Ω esiste $\rho > 0$ tale che $B_\rho(x_0, y_0) \subseteq \Omega$. Sia $\varepsilon = \frac{\rho}{2}$. Poiché $y \to f(x_0, y)$ è strettamente crescente e vale 0 per $y = y_0$, essa è strettamente positiva per $y > y_0$ e strettamente negativa per $y < y_0$, purché $(x_0, y) \in \Omega = \operatorname{dom} f$. Ne segue che

$$f(x_0, y_0-\varepsilon)<0 \quad \text{e} \quad f(x_0, y_0+\varepsilon)>0$$

Essendo f continua in Ω, e quindi in $(x_0, y_0-\varepsilon)$ e in $(x_0, y_0+\varepsilon)$, esistono $\delta_1, \delta_2>0$ tali che (permanenza di segno):

$$f(x,y)>0 \quad \text{se} \quad (x,y)\in B_{\delta_1}(x_0, y_0+\varepsilon)$$

(ed, in particolare, $f(x, y_0+\varepsilon)>0$ se $|x-x_0|<\delta_1$)

mentre

$$f(x,y)<0 \quad \text{se} \quad (x,y)\in B_{\delta_2}(x_0, y_0-\varepsilon)$$

(e, dunque, $f(x, y_0-\varepsilon)<0$ se $|x-x_0|<\delta_2$)

Si fissi ora $\delta<\min\{\delta_1, \delta_2\}$, e sia $\bar{x}\in[x_0-\delta, x_0+\delta]$.

La funzione $t\to f(\bar{x},t)$ è

- definita in $y_0-\varepsilon$ e in $y_0+\varepsilon$, ed assume in essi valori di segno discorde.
- essendo $(\bar{x}, y_0-\varepsilon)$ e $(\bar{x}, y_0+\varepsilon)$ due punti di $B_\rho(x_0,y_0)\subseteq\Omega$, ed essendo la sfera B_ρ convessa, $t\to f(\bar{x},t)$ è definita sull'intervallo $[y_0-\varepsilon, y_0+\varepsilon]$.
- è continua su tutti i punti per i quali $(\bar{x},t)\in\Omega$, e quindi in particolare in $[y_0-\varepsilon, y_0+\varepsilon]$.

Ne segue che, per il teorema degli zeri di Weierstrass per le funzioni di una variabile, essa ha zeri in $[y_0-\varepsilon, y_0+\varepsilon]$.

Per la stretta monotonia di $t \to f(\bar{x}, t)$ lo zero è unico per ogni $\bar{x} \in [x_0 - \delta, x_0 + \delta]$ prefissato, e definisce dunque univocamente una funzione di $\bar{x}$ che denoteremo con $\varphi(\bar{x})$. Inoltre, per costruzione, lo zero $\varphi(\bar{x})$ così ottenuto è compreso nell'intervallo $[y_0 - \varepsilon, y_0 + \varepsilon]$ ma, dato che negli estremi non si annulla perché assume ivi valori di segno discorde, ne segue

$\varphi(x) \in]y_0 - \varepsilon, y_0 + \varepsilon[$ e cioè $|\varphi(x) - y_0| < \varepsilon$.

Osserviamo ora che per $x = x_0$, essendo $\varphi(x_0)$ l'unico zero di $t \to f(x_0, t)$ per (x_0, t) in Ω, ed essendo per ipotesi $f(x_0, y_0) = 0$, ne segue $\varphi(x_0) = y_0$, che è la 5).

La 6) è immediata dalla costruzione: $\varphi(\bar{x})$ è lo zero (unico) di $f(\bar{x}, t)$, almeno per tutti gli $\bar{x}$ in $[x_0 - \delta, x_0 + \delta]$ e dunque, per essi, $f(\bar{x}, \varphi(\bar{x})) = 0$. La continuità, oggetto della 7) è più delicata. Iniziamo col provarla in x_0. Dalla costruzione risulta che per $|x - x_0| < \delta$, ossia in $]x_0 - \delta, x_0 + \delta[$ si ha $|\varphi(x) - \varphi(x_0)| = |\varphi(x) - y_0| < \varepsilon$ il che sembra chiudere la questione, ma non è così (purtroppo!). Fissare, infatti, un valore di ε' più piccolo del precedente richiede obbligatoriamente di considerare punti $(x_0, y_0 - \varepsilon')$ e $(x_0, y_0 + \varepsilon')$ diversi dai precedenti ai quali applicare la permanenza del segno, ottenendo intorni di raggi δ_1' e δ_2' in generale diversi, il che conduce ad un raggio δ, e ad una funzione φ a priori differenti. Cosa ci garantisce che le due funzioni coincidano sui punti comuni dei rispettivi domini? Semplice: l'unicità! Se $\bar{x}$ appartiene ad entrambi i domini il valore che entrambe le funzioni "φ" associano ad esso non può che essere l'unico zero di $t \to f(\bar{x}, t)$ e dunque esse coincidono in $\bar{x}$.

Analogamente si prova la continuità in un punto $\bar{x}$ di $[x_0 - \delta, x_0 + \delta]$ diverso da x_0: basta riapplicare il teorema fino ad ora dimostrato (continuità in x_0 inclusa) ma scegliendo come punto "iniziale" $(\bar{x}, \varphi(\bar{x}))$ invece di (x_0, y_0). Le due funzioni ottenute dalle due costruzioni diverse coincideranno sull'intersezione dei rispettivi domini, che contiene $\bar{x}$, che è dunque un punto di continuità per la "φ" ottenuta partendo dallo zero "centrale" $(\bar{x}, \varphi(\bar{x}))$, e dunque anche per l'altra, ad essa coincidente ovunque siano entrambe definite. $\boxed{///}$

Il teorema, con minimi aggiustamenti, può essere applicato alle funzioni strettamente decrescenti rispetto alla y, ma anche altrettanto bene alle funzioni strettamente monotone (di ogni tipo) rispetto alla x: in tal caso si otterrà una "formula risolutiva locale" $x = \varphi(y)$, verificante $f(\varphi(y), y) = 0 \quad \forall y \in [y_0 - \delta, y_0 + \delta]$.

Una nota conclusiva "pratica": la verifica diretta della stretta monotonia di una funzione $t \to g(t)$ richiede lo studio della disequazione $g(x) < g(y)$ e la prova che l'insieme delle soluzioni contiene tutte le coppie x, y nel dominio di g verificanti $x < y$, il che è tutt'altro che elementare, in generale.
Dato che la stretta monotonia è più facile da ottenere mediante ipotesi sulla derivata, e dato che il teorema offre comunque un risultato locale, in un intorno il cui raggio non può essere fornito a priori, è conveniente un enunciato di tipo differente, che utilizzi le derivate e la loro regolarità: è il teorema effettivamente dimostrato da Ulisse Dini, oggetto della sezione seguente.