Comitato per la Edizione Nazionale delle Opere di

Federigo Enriques

Questioni riguardanti le Matematiche elementari

(III ed.) Zanichelli, Bologna, 1927. (Di Enriques: art. XXVII, Massimi e Minimi nell'Analisi moderna, pp. 311-471)

EDIZIONE NAZIONALE OPERE DI
FEDERIGO ENRIQUES

Il presente testo è stato digitalizzato nell'ambito del progetto "Edizione nazionale delle opere di Federigo Enriques"
promosso dal
Ministero per i Beni e le attività Culturali
Area 4 – Area Archivi e Biblioteche
Direzione Generale per i Beni Librari e gli Istituti Culturali

ARTICOLO VENTISETTESIMO

Massimi e Minimi nell'Analisi moderna di Federigo Enriques a Roma.

INTRODUZIONE

Lo sviluppo dell'Analisi reca un doppio progresso alle questioni concernenti i massimi e minimi:

1) anzitutto un metodo sistematico di ricerca che si basa sugli algoritmi del Calcolo differenziale;

2) in secondo luogo la dimostrazione di generali teoremi d'esistenza e quindi la giustificazione dei metodi fondati su tale presupposto.

Storicamente precede lo sviluppo degli algoritmi, che appartiene alla fase euristica della dottrina; le condizioni differenziali di massimo o minimo, che si presentano dapprima nello studio delle funzioni di una o più variabili, si estendono naturalmente ai problemi concernenti funzioni di linee (Cfr. le note storiche ai §§ 31, 63) manifestando così la potenza e la generalità dei concetti da cui dipendono.

Ma alla scoperta delle condizioni differenziali, succede l'analisi critica volta ad investigare la loro sufficienza per riguardo al massimo o al minimo domandato. In questa fase critica, accanto ai metodi analitici che approfondiscono l'esame delle condizioni suddette (calcolo delle derivate seconde ecc.), sorgono i metodi sintetici basati sui teoremi d'esistenza, la cui importanza tende ad affermarsi prevalente nei recenti sviluppi del Calcolo delle variazioni.

In questo Articolo noi ci siamo proposti di esporre i principali resultati sui massimi e minimi, presentandoli — indipendentemente dall'ordine storico — secondo un ordine logico della dottrina.

La prima parte concerne le funzioni continue di una o più variabili e comincia coi teoremi generali d'esistenza (cap. I). Segue l'analisi infinitesimale dei massimi e minimi basata sul calcolo delle derivate (cap. II): condizioni relative alle funzioni di una variabile, studio del caso di più variabili e particolarmente di due, esame approfondito del così-detto *caso dubbio* (o caso parabolico). Le applicazioni a problemi elementari classici (§§ 13, 27, 30), la legge di reciprocità, il principio di simmetria che resta inesplicato nella trattazione elementare, vengono fatti oggetto di speciali considerazioni.

La seconda parte è consacrata alle funzioni di linee e ai problemi del Calcolo delle variazioni; l'ordine della materia riproduce quello della prima parte. Anche qui il cap. I tratta dei teoremi d'esistenza; esso espone dunque i resultati generali sulle curve limiti d'un insieme di linee, dovuti ad Ascoli, Arzelà, Hilbert, e svolge le questioni — in qualche modo inverse — sull'approssimazione d'una data curva, che sono dominate dal teorema di Weierstrass. La continuità delle elementari funzioni di linee (area, lunghezza) e l'esistenza dei massimi e minimi di queste, formano oggetto di speciale considerazione; a titolo d'esempio ci fermiamo in particolare sopra una questione critica concernente le curve convesse, che viene sollevata dalla teoria di Steiner degl'isoperimetri.

Il cap. II vuol render conto dei principali resultati conseguiti nel Calcolo delle variazioni, avuto riguardo ai recentissimi sviluppi di questo ramo della scienza. A tale scopo abbiamo riferito le nostre considerazioni a due problemi concreti: il problema delle linee di lunghezza minima nel piano e sopra le superficie, e il problema elementare degli isoperimetri. Le dimostrazioni effettive del minimo o del massimo, che costituiscono il punto delicato della teoria, sono state basate sopra i teoremi del cap. I e svolte con una trattazione *in parte nuova* che ci permettiamo di segnalare ai lettori più eruditi (§§ 49, 53, 61), Tuttavia abbiamo voluto render conto, sia pure brevemente e talora anche solo in via storica, di tutti i metodi principali che sono stati elaborati per superare quest'ordine di difficoltà. E stimeremo di non avere speso inutilmente l'opera nostra se essa varrà ad attrarre nuovi studiosi a questo campo di ricerche, e a popolarizzare nel nostro paese le idee sintetiche generali che ne dominano i progressi.

Piacemi terminare ringraziando cordialmente il professor Leonida Tonelli, per le informazioni e i consigli di cui mi è stato largo durante la redazione del presente lavoro.

PARTE PRIMA

I. - Sull'esistenza dei massimi e minimi delle funzioni continue.

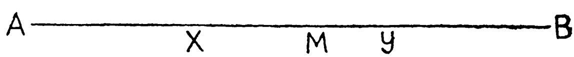
§ 1. **Gruppi di punti sulla retta: punti limiti.** — *Definizione.* Se G è un gruppo di punti sulla retta, si dice *punto limite* di G ogni punto M della retta tale che in qualsiasi segmento contenente M nel suo interno (cioè in qualsiasi *intorno* di M) cadano punti di G, diversi da M.

Teorema. Ogni gruppo, G, costituito da infiniti punti appartenenti ad un segmento, possiede — in questo — almeno un punto limite.

Sia AB il segmento, ordinato — come nella figura — da sinistra a destra. I punti di AB si possono dividere in due classi:

1) punti X tali che fra A e X cada un numero finito (o nullo) di G;

2) punti Y tali che in AY cadano infiniti G.



Ogni punto X trovasi a sinistra di tutti gli Y, il punto A è un X e B è un Y; si può quindi applicare il postulato della continuità (Vol. I, Art. V) in forza del quale esiste un punto M che separa i due gruppi.

Tre casi possono presentarsi:

1) M cade nell'estremo A. Ciò significa che ogni punto a destra di A è un Y, cioè che in ogni intorno di A vi sono punti di G; in altre parole A è punto limite di G.

2) M è interno ad AB.

In questa ipotesi si prendano comunque in AB un punto a sinistra e un punto a destra di M; questi sono risp. un X e un Y; in AX vi è un numero finito di G, in AY infiniti G;

dunque in XY ci sono infiniti G. Ciò significa che M è punto limite di G.

3) M cade in B. Preso comunque un punto interno ad AB questo è un X, per modo che in AX vi è un numero finito di G e quindi in XB infiniti G. Segue che B è punto limite di G.

NOTA. Il punto M che abbiamo costruito è il *primo* punto limite di G nell'ordine del segmento.

Nel caso particolare in cui sia data una *successione* di punti $x_1, x_2, \ldots x_n, \ldots$,

$$x_1 \leq x_2 \leq \ldots \leq x_n \ldots,$$

con

$$x_n \leq l$$

qualunque sia n, esiste un *unico punto limite* del gruppo $(x_1, \ldots x_n, \ldots)$, e questo si dice *il limite* a cui *tende* x_n, o la successione $x_1, \ldots x_n, \ldots$, *per* $n = \infty$. L'esistenza di questo limite equivale direttamente al postulato della continuità della retta.

§ 2. **Limite superiore e inferiore.** — Sia G un gruppo di punti sopra una retta ordinato secondo i valori crescenti delle ascisse x (da sinistra a destra); si dice che un punto (numero) l è *limite superiore* di G, e degli x, se, qualunque sia x:

$$x \leq l,$$

e, per ogni numero positivo ε (piccolo ad arbitrio), esiste qualche x tale che

$$x > l - \varepsilon.$$

Analogamente, scambiando i segni $<$ e $>$ si definisce il *limite inferiore*.

Esistenza e unicità. Un gruppo G può non avere limite superiore (o inferiore); così accade per es. del gruppo

$$x = 0, \quad 1, \quad 2, \quad 3, \ldots,$$

tale che non esiste un

$$l > x.$$

Ma il limite superiore (o inferiore) di G, *se esiste,* è evidentemente *unico.*

Teorema. *Ogni gruppo* G *compreso nella semiretta a destra d'un punto* A *ammette limite inferiore. Similmente ogni* G *compreso nella semiretta a sinistra di* B *ammette limite superiore.*

Riferendoci alla prima ipotesi consideriamo un qualsiasi punto B di G e il gruppo dei G contenuto in AB.

Se questo gruppo è finito si prenda il suo primo punto.

Altrimenti si ha in AB un gruppo G infinito.

Sia il segmento AB ordinato da sinistra a destra. Si costruisca — come al § 1 — il primo punto limite di G: M. Fra A e M può trovarsi o no un qualche punto M' di G; nel secondo caso M è il limite inferiore, nel primo fra A ed M' c'è un numero finito di punti di G e quindi un primo punto di G avanti M.

Definizione. Qualora il limite superiore di un gruppo di punti o numeri appartenga al gruppo, esso dicesi il *massimo*; *minimo* è il limite inferiore che appartenga al gruppo (cfr. la Def. Art. XXV).

Esempî: il gruppo

$$\frac{1}{2},\ \frac{3}{4},\ \frac{7}{8} \ldots\ 1$$

ha per massimo 1 $\left(\text{e minimo } \frac{1}{2}\right)$; il gruppo

$$\frac{1}{2},\ \frac{3}{4},\ \frac{7}{8} \ldots,$$

non ha massimo perchè il suo limite superiore, 1, non appartiene al gruppo.

Per notizie storiche cfr. il § 14.

§ 3. **Funzioni d'una variabile: funzioni limitate.** — Consideriamo i valori (punti) di una variabile indipendente x appartenente ad un *intervallo* definito in uno dei tre modi seguenti:

1) come insieme degli x compresi fra due valori estremi a, b:

$$a \leq x \leq b \qquad \text{(segmento)};$$

2) come insieme degli x per cui

$$a \leq x \quad \text{o} \quad x \geq a \qquad \text{(semiretta)};$$

3) come insieme di tutti i numeri reali (retta).

Definizione. Si dice che y è *funzione di* x, nell'intervallo considerato, se è data una legge di corrispondenza per cui ad ogni x dell'intervallo suddetto corrisponde un valore di y; e si scrive

$$y = f(x) \begin{cases} \text{in } (a, b) \\ \text{o per } x \geq a, \text{ o } x \leq a, \\ \text{o per ogni } x. \end{cases}$$

Definizione. Una funzione $y = f(x)$, definita in un intervallo finito o infinito, si dirà *limitata superiormente*, se esiste un numero l tale che — per ogni x dell'intervallo dato —

$$f(x) < l.$$

La $f(x)$ si dirà *limitata inferiormente* se esiste un l' per cui

$$f(x) > l'.$$

Quando si dice che la $f(x)$ è una *funzione limitata*, senz' altro, s'intende che essa sia in pari tempo limitata superiormente e inferiormente, cioè che esista un numero positivo l per cui, in valore assoluto,

$$|f(x)| < l.$$

Teorema. Ogni funzione $y = f(x)$ limitata superiormente ammette un *limite superiore*, ed ogni funzione limitata inferiormente ammette un *limite inferiore*.

Per ipotesi esiste un l per cui

$$f(x) < l.$$

Si considerino i punti $y = f(x)$ come formanti un gruppo ordinato da sinistra a destra, sulla retta y. Questo gruppo, essendo contenuto nella semiretta a sinistra di l, ammette certo un limite superiore (§ 2).

In modo analogo si dimostra la seconda parte dell'enunciato.

Corollario. Ogni funzione limitata ammette limite superiore e limite inferiore.

§ 4. **Funzioni continue: teorema di Heine sulla continuità uniforme.** — *Definizione.* La funzione $y = f(x)$ si dirà *continua a destra* (o risp. *a sinistra*) nel punto $x = c$, che non sia l'estremo superiore (o risp. inferiore) del dato intervallo di variazione della x, se

$$\lim_{h=+0} f(c+h) = f(c)$$

(o risp.

$$\lim_{h=+0} f(c-h) = f(c));$$

in altre parole se, dato un numero ε, piccolo ad arbitrio, si può determinare un τ tale che, per

$$0 \leq h < \tau$$

si abbia in valore assoluto

$$|f(c+h) - f(c)| < \varepsilon.$$

Se $f(x)$ è *continua a destra e a sinistra* nel punto c, dicesi semplicemente *continua in quel punto*; pertanto la $f(x)$ è continua per $x = c$ se, dato un ε, si può trovare un τ tale che si abbia in valore assoluto

$$|f(c+h) - f(c)| < \varepsilon$$

per

$$|h| < \tau.$$

Definizione. Una funzione continua in tutti i punti d'un intervallo (a, b) dicesi *continua nell'intervallo,* e circa il modo come essa si approssima al limite nei diversi punti dell'intervallo stesso, si può dimostrare un teorema che — pur non essendo necessario per gli sviluppi dei paragrafi seguenti — crediamo di riferire. Il lettore frettoloso può passare oltre.

Teorema di Heine [1]. *Se una funzione* $\mathrm{y} = \mathrm{f(x)}$ *è continua in tutti i punti d'un intervallo* (a, b), *essa è in questo intervallo uniformemente continua,* cioè — *dato* un numero positivo ε piccolo ad arbitrio — si può *trovare un* τ tale che sia in valore assoluto

$$|f(x) - f(x')| \leq \varepsilon,$$

[1] H. E. Heine, *Journal für reine u. angew. Math.*, Bd. 71, p. 361, Bd. 74, p. 188. Cfr. U. Dini, *Fondamenti per la teorica delle funzioni di variabili reali.* Pisa, 1878, p. 48.

per tutti i punti x, x' dell'intervallo per cui

$$|x - x'| \leq \tau.$$

Sia $a < b$ e consideriamo un segmento δ_1 a destra dell'estremo a, tale che sia, per $x - a \leq \delta_1$,

$$|f(x) - f(a)| \leq \frac{\varepsilon}{4}; \tag{1}$$

per tutte le coppie di valori x, x' fra a ed x, si avrà certo

$$|f(x) - f(x')| \leq |f(x) - f(a)| + |f(a) - f(x')| \leq \frac{\varepsilon}{2}. \tag{2}$$

Ora questa diseguaglianza vale finchè δ_1 resta assai piccolo, in guisa che sia soddisfatta la (1). Si considerino tutti gl'intervalli δ_1, a destra di a, per cui la diseguaglianza (2) è soddisfatta; vi sarà certo un limite superiore x_1 di questi intervalli: $x_1 - a = \tau_1$; e finchè

$$x - a \leq \tau_1$$

non potrà aversi

$$|f(x) - f(a)| > \frac{\varepsilon}{4},$$

altrimenti sarebbe

$$|f(x) - f(a)| = \frac{\varepsilon}{4} + \sigma$$

e si potrebbe scegliere un punto x a sinistra di x_1 così vicino ad esso che

$$|f(x) - f(x_1)| < \sigma,$$

$$|f(x) - f(a)| > \frac{\varepsilon}{4}.$$

Si deduce dunque che il punto $x_1 = a + \tau_1$ determinato innanzi è tale che per

$$x - a \leq \tau_1, \quad x' - a \leq \tau_1,$$

$$|f(x) - f(a)| \leq \frac{\varepsilon}{4}$$

e quindi

$$|f(x) - f(x')| \leq \frac{\varepsilon}{2},$$

mentre a destra di x_1, e vicino ad esso quanto si vuole, esistono punti x per cui

$$|f(x) - f(a)| > \frac{\varepsilon}{4}.$$

Ora, procedendo in modo analogo, si determinerà a destra di x_1 un punto

$$x_2 = x_1 + \tau_1,$$

tale che per x, x' fra x_1, x_2

$$|f(x) - f(x_1)| \leq \frac{\varepsilon}{4}$$

$$|f(x) - f(x')| \leq \frac{\varepsilon}{4},$$

mentre a destra di x_2 vi sono punti x vicini quanto si vuole per cui

$$|f(x) - f(x_1)| > \frac{\varepsilon}{4}.$$

Questo procedimento si ripeterà successivamente costruendo una successione di punti $x_1, x_2, \dots x_n$, fino a che non accada di trovare un punto x_n coincidente coll'estremo b dell'intervallo dato. Ma se il processo si esaurisca in tal modo dopo un numero finito di punti $x_1, x_2, \dots x_n$, si potrà determinare un τ uguale al minimo fra i valori di $\tau_1, \tau_2 \dots$, ed allora si proverà facilmente che per

$$|x - x'| \leq \tau$$
$$|f(x) - f(x')| \leq \varepsilon.$$

Infatti x, x' cadranno fra due punti consecutivi x_i, x_{i+1}, oppure si troveranno l'uno fra due punti x_i, x_{i+1}, l'altro fra x_i, x_{i-1}. Nel primo caso si avrà

$$|f(x) - f(x')| \leq \frac{\varepsilon}{2},$$

nel secondo caso

$$|f(x) - f(x')| \leq |f(x) - f(x_i)| + |f(x_i) - f(x')| \leq \varepsilon.$$

Suppongasi invece che i punti $x_1, x_2 \dots x_n \dots$ sieno in numero infinito, e quindi che $\tau_1, \tau_2 \dots \tau_n \dots$ vadano decrescendo oltre ogni limite. In questa ipotesi consideriamo il punto $\bar{x}$, limite superiore di $x_1 x_2 \dots x_n \dots$, il quale cadrà nell' estremo b o a sinistra di b (§ 2). Se si prende un intorno a sinistra di $\bar{x}$, vi saranno certo, in codesto intorno, dei punti x_n, $x_{n+1} \dots$; quindi si potrà determinare un x nell' intorno suddetto per cui

$$|f(x) - f(x_n)| > \frac{\varepsilon}{4}.$$

Questa diseguaglianza si può sempre soddisfare per punti appartenenti ad un intorno sinistro comunque piccolo del punto $\bar{x}$. Ma essa è in contraddizione coll'ipotesi che $f(x)$ sia continua per $x = \bar{x}$ e quindi si potrà trovare un τ tale che per

$$x - \bar{x} < \tau, \quad \bar{x} - x' < \tau$$

$$|f(x) - f(\bar{x})| < \frac{\varepsilon}{8}$$

$$|f(x) - f(x')| \leq |f(x) - f(\bar{x})| + |f(\bar{x}) - f(x') < \frac{\varepsilon}{4}.$$

Dunque quel punto $\bar{x}$ risulta essere un punto di *discontinuità* della funzione $f(x)$, contro le ipotesi fatte.

Segue il teorema che dovevasi dimostrare.

§ 5. **Massimo e minimo.** — Lemma. *Se una funzione continua* $\mathrm{y} = \mathrm{f(x)}$ (in un intervallo finito o infinito) *assume in due punti* x_1, x_2 *i valori* y_1, y_2, *essa assume un valore qualsiasi* l *compreso fra* y_1, y_2, *corrispondentemente a qualche valore di* x *compreso fra* x_1 e x_2.

Sia per es.

$$x_1 < x_2,$$
$$y_1 < l < y_2,$$

I punti del segmento (x_1, x_2) si potranno distribuire in due classi, attribuendo:

1) alla 1ª classe i punti m tali che per $x \leq m$ si abbia $f(x) < l$;

2) alla 2ª classe i punti p tali che esista un $x \leq p$ per cui $f(x) \geq l$.

Si ha così una partizione di DEDEKIND. Il punto di separazione x è tale che nell' intervallo x, $x + h$, qualunque sia $h > 0$, sono compresi punti x per cui $f(x) \geq l$. Per la contiunità di f si deduce

$$f(\bar{x}) = \lim_{h=+0} f(\bar{x}+h) \geq l.$$

D' altra parte

$$f(\bar{x}-h) < l$$

e però

$$f(\bar{x}) = \lim_{h=+0} f(\bar{x}-h) \leq l.$$

Dunque

$$f(\bar{x}) = l.$$

NOTA. Il punto $\bar{x}$ costruito fra x_1 e x_2 è il *primo*, nell' ordine del segmento (x_1, x_2), per cui

$$f(x) = l.$$

Corollario. Ogni *funzione continua* $y = f(x)$ in un *intervallo finito* è *limitata*, e quindi (§ 3) ammette limite superiore ed inferiore.

Dimostriamo l' esistenza del limite superiore, procedendo per assurdo come segue.

Sia $f(x)$ priva di limite superiore. In tal caso preso un numero $l < f(x)$ dovrà trovarsi un x per cui $f(x) > l$ (§ 5). Quindi esisterà un primo punto x_0 nell' ordine (ab) per cui

$$f(x_0) = l.$$

Successivamente si troverà dopo x_0 un primo x_1 per cui

$$f(x_1) = l + 1,$$

e così si costruirà una successione di punti

$$x_0, x_1, x_2, \dots x_n, \dots,$$

procedenti da sinistra a destra, tali che

$$f(x_n) = l + n.$$

La successione nominata tende ad un limite $\bar{x}$ (§ 11) e si dovrebbe avere

$$f(\bar{x}) = \lim_{n=\infty} (l+n),$$

che è assurdo qualunque sia il valore (finito) del numero $f(\bar{x})$.

Teorema. *Ogni funzione continua* y = f(x), *in un intervallo finito,* (a, b), *ammette massimo e minimo.*

Dimostriamo che f ammette massimo.

Sia l il limite superiore degli y, e pongasi che sia

$$l > f(a): \quad l - f(a) = 2e.$$

Allora, procedendo da sinistra a destra a partire da a, esiste certo un primo $x = c$ $(c \geq a)$ per cui

$$f(c) = l - e;$$

questo x si può determinare come nel lemma precedente. Parimente esiste un primo $x = c_1$ per cui

$$f(c_1) = l - \frac{e}{2},$$

un primo $x = c_2$ per cui

$$f(c_2) = l - \frac{e}{2^2},$$

. . . .

un primo $x = c_n$ per cui

$$f(c_n) = l - \frac{e}{2^e}.$$

I punti

$$c, \; c_1, \; c_2, \ldots c_n, \ldots$$

formano una successione crescente che tende ad un limite $\bar{x}(\leq b)$ e si ha

$$f(\bar{x}) = \lim_{n=\infty} f(c_n) = l.$$

Dunque il limite superiore l è un valore assunto da $f(x)$ e perciò è massimo.

§ 6. **Funzioni continue in un intervallo infinito.** — Sia data una funzione continua nell'*intervallo infinito* $x \geq a$, cioè sulla *semiretta* x *a destra* del punto a.

Si operi una *proiezione* della semiretta sopra un segmento; questa operaz one corrisponde analiticamente ad una sostituzione lineare. Per es. la sostituzione

$$x = \frac{x' - a}{x' + 1}$$

trasforma la semiretta a destra di a nel segmento $(0, 1)$, e la sostituzione

$$x = \frac{px' - a}{x' + 1}$$

trasforma la semiretta a destra di a nel segmento $(0, p)$.

Poniamo per semplicità di avere eseguito la prima sostituzione; la funzione $f(x)$ si cambia in una nuova funzione $\bar{f}(x')$ definita nell'intervallo $(0, 1)$ *escluso l'estremo* superiore, e continua in tutti i punti di codesto intervallo, all'infuori dell'estremo suddetto. Sarà possibile definire $\bar{f}(x')$ come funzione continua anche nell'estremo 1, e applicarle quindi i teoremi del paragrafo precedente? Che cosa ne seguirà per $f(x)$?

Occorre distinguere varii casi:

1) Esiste un numero k tale che

$$\lim_{x'=1} \bar{f}(x') = k,$$

e quindi

$$\lim_{x=\infty} f(x) = k.$$

Ciò significa che dato un numero positivo ε, piccolo ad arbitrio, si può determinare un σ tale che

$$1 > x' > 1 - \sigma$$

per

$$|\bar{f}(x') - k| < \varepsilon;$$

e quindi si può determinare un b così grande che

$$x > b$$

per

$$|f(x) - k| < \varepsilon.$$

Nell'ipotesi fatta si può porre per definizione

$$\bar{f}(1) = k,$$

e la $f(x)$ risulta funzione continua in *tutto* l'intervallo $(0, 1)$

(estremo 1 incluso); si dirà quindi che $f(x)$ è *finita e continua anche nel punto* $x = \infty$.

In quest' ipotesi la $\bar{f}(x')$ ammette massimo e minimo (§ 5). La $f(x)$ *ammetterà pure massimo e minimo se alcuno di questi valori non coincide precisamente col valore di*

$$f(\infty) = \lim_{x=\infty} f(x) = k.$$

Se non esiste un numero h per cui

$$\lim_{x'=1} \bar{f}(x') = k,$$

è impossibile definire $\bar{f}(x')$ in $x' = 1$ per modo che essa risulti una funzione continua in *tutto* l' intervallo (0, 1). Ma converrà ancora distinguere varii casi:

2) Sia

$$\lim_{x'=1} \bar{f}(x') = +\infty \quad \text{o} \quad = -\infty,$$

e quindi

$$\lim_{x=\infty} f(x) = +\infty \quad \text{o} \quad = -\infty.$$

Ciò significa che, dato un numero positivo l, grande ad arbitrio, si può trovare un σ così piccolo che

$$1 > x' > 1 - \sigma$$

er

$$\bar{f}(x') > l \quad \text{o risp.} \quad \bar{f}(x') < -l,$$

e quindi si può determinare un b così grande che

$$x > b$$

per

$$f(x) > l \quad \text{o risp.} \quad f(x) < -l.$$

Si dirà nei due casi che $\bar{f}(x')$ è (*continua*) *infinita positiva* o risp. *negativa* per $x' = 1$, e così $f(x)$ per $x = \infty$.

Se ha luogo la prima ipotesi è chiaro che $\bar{f}(x')$, definita fra (0, 1) — estremo 1 escluso — non ha massimo, ma ha invece un minimo, che è il minimo della funzione stessa in un qualsiasi intervallo (0, 1-σ), dove σ sia sufficientemente piccolo.

Parimente *la funzione* f(x) *infinita positiva* per $x = \infty$ ha *minimo* e *non massimo* sulla semiretta $x \geq a$. L' opposto accade se la $f(x)$ è *infinita negativa*.

3) Può accadere che

$$\lim_{x'=1} \bar{f}(x') = \pm \infty$$

e quindi

$$\lim_{x=\infty} f(x) = \pm \infty,$$

cioè la funzione diventi (*continua*) *infinita senza* conservare *segno* costante; questo caso si riduce al precedente considerando il valore assoluto $|f(x)|$, per cui

$$\lim_{x=\infty} |f(x)| = + \infty.$$

La $f(x)$ non avrà ora *nè massimo nè minimo* ma il *suo valore assoluto* avrà sempre un *minimo.*

4) Può accadere che la $\bar{f}(x')$ per $x' = 1$, e quindi $f(x)$ per $x = \infty$, abbia una *discontinuità essenziale,* cioè che $f(x)$ per $x = \infty$ non tenda ad *alcun limite, finito o infinito.*

Ciò accade p. es. se — qualunque sia b — si trovino degli $x > b$ per cui $f(x)$ assume un certo valore k oppure supera in valore assoluto qualsiasi l, e degli $x > b$ per cui la $f(x)$ assume un altro valore $h \neq k$.

Qui nulla si può dire circa l'esistenza del massimo o del minimo di $f(x)$, a meno che non si possa riconoscere che la funzione stessa è *limitata* in un senso, e che le *oscillazioni* della funzione sono *limitate* fra valori inferiori al limite superiore o superiori al limite inferiore.

Così p. es. la funzione

$$y = \operatorname{sen} x - \frac{1}{x} \quad \text{per} \quad x \geq 1$$

ha una discontinuità essenziale nel punto $x = \infty$, e tuttavia

per
$$x > \frac{7\pi}{7}$$

$$y > -\left(1 + \frac{2}{7\pi}\right),$$

mentre per esempio nel punto

$$x = \frac{5\pi}{2},$$

$$y = -\left(-1 + \frac{2}{5\pi}\right);$$

si deduce quindi che *y ha un minimo* che è il minimo della funzione definita nell'intervallo finito $\left(1, \frac{7\pi}{2}\right)$. Ma *manca il massimo,* giacchè il limite superiore 1, non è raggiunto.

Invece la funzione

$$y = \operatorname{sen} x \qquad\qquad \text{per } x \geq 0$$

offre un esempio in cui si ha discontinuità essenziale per $x = \infty$, e *manca massimo* o *minimo.* La funzione suddetta assume, per valori grandi di x, valori positivi e negativi arbitrariamente grandi, ma anche valori nulli per

$$y = \operatorname{sen} x \qquad\qquad (r = 0, 1, 2, ...).$$

Le considerazioni precedenti si estendono subito al caso di una funzione defininita sopra la *semiretta a sinistra* d' un punto a, cioè per $x \leq a$. Esse si estendono pure al caso in cui l' *intervallo* di variazione della x sia *tutta la retta,* potendosi questa riguardare come *somma delle due semirette* a destra e a sinistra d' un punto a. Solo occorre qui considerare il $\lim f(x)$ per $x = \pm \infty$, riguardando il *punto all' infinito* della retta come *unico*; quindi l'esistenza d'un tal limite, finito o infinito, cioè la *continuità* di f all' *infinito* significherà l' *esistenza* e l' *uguaglianza oppure* l' *infinità* dei due limiti: $\lim\limits_{x=+\infty} f(x)$, e $\lim\limits_{x=-\infty} f(x)$.

Pertanto se sia data una *funzione* f(x) *continua su tutta la retta* e *anche nel punto* $x = \infty$, vi sarà luogo a distinguere i seguenti casi:

1) La funzione è *finita* per $x = \infty$:

$$\lim_{x=\pm\infty} f(x) = h.$$

Essa ha *massimo* e *minimo, tranne se* k è il suo *limite superiore* o *inferiore.*

2) La funzione $f(x)$ è *infinita positiva* (oppure *negativa*) per $x = \infty$:

$$\lim_{x=\pm\infty} f(x) = + \infty.$$

In tal caso *non* c' è *massimo, ma* c' è *minimo* (e reciprocamente).

3) La funzione $f(x)$ diviene *infinita senza segno* per $x = \infty$, sia che diventi positiva per $x = + \infty$ e negativa per

$x = -\infty$ (o viceversa), sia che prenda già valori (grandi in valore assoluto) positivi e negativi per x grande positivo (e negativo). Allora *non c' è massimo nè minimo* per $f(x)$.

Ma *in ogni caso il valore assoluto* $|f(x)|$ *di una funzione continua anche all' infinito, ammette almeno un minimo o un massimo.*

§ 7. **Campo a due dimensioni.** — Si considerino i valori x, y (punti) appartenenti ad un *rettangolo* (*abcd*), cioè soddisfacenti alle diseguaglianze

$$a \leq x \leq b$$
$$c \leq x \leq d.$$

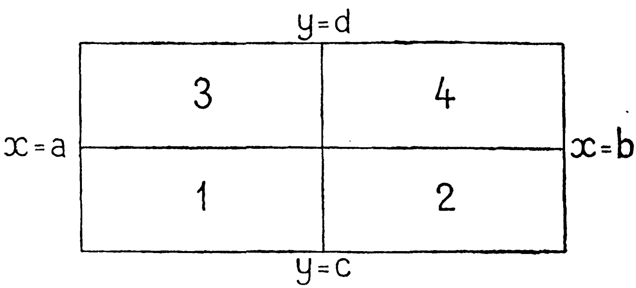
Sia G un gruppo di infiniti punti appartenenti al rettangolo. Si dice che un punto $(\bar{x}, \bar{y})$ è *punto limite* di G se, in ogni rettangolo o quadrato

$$x - h \leq x \leq x + h$$
$$y - h \leq y \leq y + h,$$

per h piccolo quanto si vuole, esistono punti di G diversi da $(\bar{x}, \bar{y})$.

Teorema. Ogni *gruppo infinito* G, dato in un rettangolo (*abcd*), possiede dei *punti limiti.*

Per dimostrare il teorema, immaginiamo di dividere i lati del rettagolo successivamente in 2, 4,.... 2^n, parti uguali; il rettangolo stesso verrà diviso corrispondentemente in 2^2, 4^2.... 2^{2n} rettangoli uguali, e l' insieme dei 2^{2n} rettangoli



così ottenuti si potrà ordinare partendo da quello per cui x, y hanno i valori più piccoli e procedendo per file orizzontali da *destra* a *sinistra* (x crescente) e poi dal *basso* all' *alto* (y *crescente*), Ora in uno almeno R_n fra i sunnominati 2^{2n} rettangoli vi saranno infiniti punti di G. Si scelga p. es. il primo R_n per cui ciò accade, e dentro questo il primo R_{n+1} ove sono pure infiniti punti di G e così via. La successione dei rettangoli $R_n R_{n+1} \ldots$, tende ad un determinato *punto limite* del gruppo G.

Corollario. Se si ha, in $(abcd)$, una *successione* di punti

$$(x_1, y_1) \ldots (x_n, y_n) \ldots,$$

per cui

$$x_1 \leq x_2 \ldots \leq x_{n+1} \ldots,$$

e si assume

$$\bar{x} = \lim_{n=\infty} x_n$$

(cfr. § 1), esiste sulla retta $x = \bar{x}$ qualche punto limite del gruppo $(x_1, y_1) \ldots (x_n, y_n) \ldots$

Un siffatto punto limite $(\bar{x}, \bar{y})$ si ottiene mediante la costruzione precedente, ove si ordinino i rettangoli R_n non gia, come ivi è indicato, *per* x, y *crescenti,* ma all' opposto per *y, x crescenti,* cioè per file verticali da sinistra a destra.

Qualora anche

$$y_1 \leq x_2 \ldots \leq y_n \leq y_{n+1} \ldots,$$

si ha un *unico punto limite* del gruppo $(x_1, y_1) \ldots (x_n, y_n) \ldots$, che dicesi *il limite della successione* o di (x_n, y_n), per $n = \infty$:

$$\lim_{n=\infty} (x_n, y_n).$$

Entro un rettangolo $(abcd)$ si può *delimitare un campo* di variazione per x, y, p. es. ponendo

$$x^2 + y^2 \leq r^2,$$

cioè considerando i *punti interni ad un cerchio.*

In modo generale definiremo come *campo* C *finito e continuo* (x, y), ogni insieme di punti (x, y) che soddisfi alle condizioni seguenti:

1) Sia *finito* cioè contenuto in un rettangolo $(abcd)$:

$$a \leq x \leq b,$$

$$c \leq x \leq d.$$

2) Sia *chiuso*, cioè se un gruppo di punti G appartiene a C, ogni punto limite di G appartiene pure a C.

3) Dati due punti

$$(x_1, y_1), \quad (x_2, y_2) \text{ di } C$$

esista una *linea continua* congiungente i due punti e appartenente a C; cioè si possano determinare due funzioni continue (invertibili) d'una variabile t in un intervallo (t_1, t_2):

$$x = \varphi(t), \quad y = \psi(t)$$

($t =$ funz (x, y)), in modo che per t fra t_1 e t_2 il punto (φ, ψ) appartenga a C, e tali che

$$x_1 = \varphi(t_1), \quad x_2 = \varphi(t_2)$$
$$y_1 = \psi(t_1) \quad y_2 = \psi(t_2).$$

Nota. Un campo C così definito non è necessariamente *a due dimensioni*. Esso può ridursi ad una linea continua o essere *misto*, cioè somma d'una linea e d'una superficie p. es. circolare ecc. Per escludere siffatte possibilità conviene aggiungere alla definizione di C la condizione seguente:

4) Se (x, y) è un punto di C, fra i punti di C per cui in valore assoluto

$$|x - \bar{x}| \leq h, \quad |y - \bar{y}| \leq h,$$

con h positivo assai piccolo, ci sono sempre tutti i punti d'un rettangolo.

Più precisamente: se $(\bar{x}, \bar{y})$ ha un *punto interno* a C, appartengono a C tutti i punti (x, y) per cui

$$|x - \bar{x}| \leq h, \quad |y - \bar{y}| \leq h;$$

e se $(\bar{x}, \bar{y}$ non è interno a C (cioè appartiene al *contorno* di C) esistono due numeri ε_1, ε_2 piccoli quanto si vuole, tali che il punto $(x + \varepsilon_1, y + \varepsilon_2)$ sia interno a C.

§ 8. **Funzioni continue di due variabili.** — *Definizione.* Si dirà che $z = f(x, y)$ è *funzione* di (x, y) in un campo C finito e continuo, se per ogni punto (x, y) di C è determinato *un* valore della variabile z.

La funzione f si dirà *continua* nel punto (p, q) se

$$\lim_{x=p,\ y=q} f(x, y) = f(p, q),$$

cioè dato un numero positivo ε, piccolo ad arbitrio, esiste un h tale che, per

$$p - h < x < p + b$$
$$q - h < y < q + h,$$

cioè per

$$|x - p| < h, \quad |y - q| < h,$$

si ha (in valore assoluto)

$$|f(x, y) - f(p, q)| \leq \varepsilon.$$

E f si dirà *continua nel campo* C se è continua in ogni punto di esso.

Nota. La continuità di $f(x, y)$ come funzione delle due variabili x, y, importa la continuità di $f(x, y)$ come funzione di x per $y =$ cost. e di $f(x, y)$ come funzione di y per $x =$ cost. Ma la funzione $f(x, y)$ potrebbe essere continua *separatamente* rispetto ad x e ad y e *non* essere *continua* rispetto ad x, y *complessivamente*.

Suppongasi sempre $f(x, y)$ continua come funzione di x, y (rispetto alle due variabili complessivamente). Se (x_1, y_1), (x_2, y_2) sono due punti di C si può sempre considerarli come estremi di una linea contenuta in C, rappresentata da due funzioni continue (invertibili)

$$x = \varphi(t), \quad y = \psi(t),$$

in un intervallo (t_1, t_2):

$$\varphi(t_1) = x_1 \quad \varphi(t_2) = x_2,$$
$$\psi(t_1) = y_1 \quad \psi(t_2) = y_2.$$

Allora $z = f(x, y)$ risulta funzione continua di t sopra la linea suddetta.

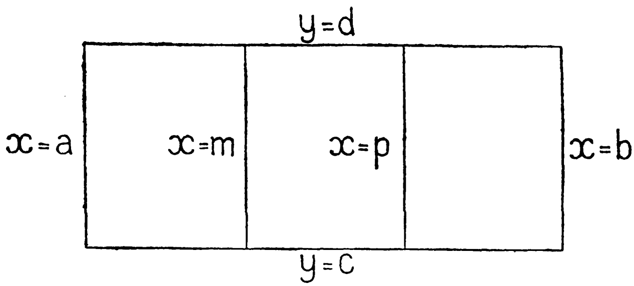
In base a questa osservazione si estende subito alle funzioni di due variabili il lemma fondamentale del § 5.

Per estendere il corollario che ne abbiamo dedotto, cioè per trovare l'esistenza d'un limite superiore e inferiore di $f(x, y)$, si può ripetere lo stesso ragionamento per assurdo.

Poniamo per semplicità di discorso che C sia tutto il rettangolo $(abcd)$. Dato un numero l, comunque grande, si suppone che vi sia qualche punto (x, y) per cui $f(x, y) < l$, e si cerca di determinare un punto (x_0, y_0) per cui

$$f(x_0, y_0) = l.$$

A tale scopo si considerano i segmenti paralleli $x =$ cost, precedenti da sinistra a destra. Se sopra $x = a$ non vi è alcun punto (a, y) per cui $f(a, y) \geq l$, divideremo i segmenti anzidetti in due classi, come segue:



1) un segmento $x = m$ sarà attribuito alla prima classe se, per $x \leq m$,

$$f(x, y) < l;$$

2) un segmento $x = p$ sarà attribuito alla seconda classe se esiste qualche punto (x, y) per cui $x \leq p$, tale che

$$f(x, y) \geq l.$$

Per il postulato di continuità si determina così un x_0, *primo* procedendo da sinistra a destra, tale che esistono punti (x_0, y) per cui

$$f(x_0, y) = l.$$

Fra questi punti ne potremo determinare uno, (x_0, y_0), corrispondente al minimo valore di y, per cui

$$f(x_0, y_0) = l.$$

In modo analogo determineremo *un* punto (x_1, y_1) per cui

$$f(x_1, y_1) = l + 1$$

. . . .

ed un (x_n, y_n) per cui

$$f(x_n, y_n) = l + n;$$

dove

$$x_1 \leq x_2 \leq x_3 \dots$$

Avremo quindi *un* $\bar{x}$ limite della successione $x_1, x_2, x_3 \dots$, e sopra il segmento $x = \bar{x}$ cadrà qualche punto limite del gruppo $(x_1, y_1)(x_2, y_2), \dots (x_n, y_n), \dots$ (§ 7). Scelto uno $(\bar{x}, \bar{y})$ fra tali punti, $\bar{y}$ è punto limite del gruppo $(y_1, y_2 \dots y_n \dots)$ e perciò si può costruire una successione

$$y_{h_1}, y_{h_2} \dots y_{h_n} \dots$$

$(h_1 < h_2 < \dots)$ avente per limite $\bar{y}$. Si ha quindi

$$f(\bar{x}, \bar{y}) = \lim_{n=\infty} f(x_n, y_{h_n}),$$

il che è assurdo, essendo il primo membro finito e il secondo $= \lim\limits_{h_n=\infty} (l + h_n) = \infty$.

Ora si consideri il limite superiore l di $f(x, y)$; vogliamo provare ch' esso è raggiunto in qualche punto del campo C, e però che è massimo

Supposto che sia

$$f(a, c) < l,$$
$$l - f(a, c) = 2e,$$

costruiamo un punto (x_0, y_0) per cui

$$f(x_0, y_0) = l - e,$$

e poi un (x_1, y_1) per cui

$$f(x_1, y_1) = l - \frac{e}{2},$$

. . . .

un (x_n, y_n) per cui

$$f(x_n, y_n) = l - \frac{e}{2},$$

per modo che

$$x_1 < x_2 < x_3 \dots$$

Determiniamo un punto $(\bar{x}, \bar{y})$ fra i punti limiti del gruppo

$(x_1, x_1)(x_2, y_2)\ldots$; esso sarà limite d'una successione di punti

$$(x_{h_1}, y_{h_1}), (x_{h_2}, y_{h_2})\ldots \quad (h_1 < h_2 \ldots)$$

per modo che

$$f(\bar{x}, \bar{y}) = \lim_{x=\infty} f(x_{h_n}, y_{h_n}) = l.$$

Lo stesso ragionamento si può ripetere per il limite inferiore di $f(x, y)$. Concludiamo dunque il

Teorema. *Ogni funzione continua* f(x, y), *in un campo finito e continuo* C, *ammette massimo e minimo, e prende effettivamente tutti i valori compresi fra questi due estremi.*

§ 9. **Campo infinito.** — Come nel § 6 per il caso d'una variabile, possiamo ora estendere le nostre considerazioni alle funzioni definite in *campi continui infiniti,* dove si suppongono verificate le condizioni 2) 3) del paragrafo precedente in un senso esteso, ammettendosi che un gruppo di punti possa avere per limite l'infinito. Ma, per semplicità ci limiteremo ad alcune osservazioni sulle funzioni $f(x, y)$ definite in *tutto il piano,* cioè per *tutti* i valori di x, y.

Consideriamo un cerchio di centro (o, o) e raggio r, che si farà poi crescere oltre ogni limite; dovremo considerare la funzione $f(x, y)$ definita *nel cerchio.*

$$x^2 + y^2 \leq r^2,$$

e la funzione stessa *fuori del cerchio,* cioè per

$$x^2 + y^2 > r^2.$$

1) Se esiste un numero k tale che — preso ε positivo piccolo ad arbitrio — si può determinare un r in modo che

$$x^2 + y^2 < r^2$$

per

$$|f(x, y) - k| < \varepsilon,$$

si dice che

$$\lim_{\infty} f(x, y) = k,$$

e che la funzione si conserva *finita e continua* ancne *all'infinito.*

2) Se dato un numero positivo l, grande ad arbitrio, si

può determinare un r in modo che

$$x^2 + y^2 < r^2$$

per

$$f(x, y) < l \quad (o < 1 - l),$$

si dice che

$$\lim_{\infty} f(x, y) = +\infty \quad (o = -\infty);$$

e che la f è (continua) ***infinita positiva*** (o risp. ***negativa***) all' infinito.

3) Essa dicesi semplicemente ***infinita*** (*senza segno*) se

$$\lim_{\infty} f(x, y) = \pm \infty.$$

4) Infine $f(x, y)$ dicesi essenzialmente ***discontinua*** all' infinito, se non tende ad alcun limite finito o indefinito.

Nel caso 1) ***la*** $f(x, y)$ ammette *massimo e minimo purchè* il $\lim\limits_{\infty} f(x, y) = k$ ***non*** sia il suo *limite superiore* o *inferiore* (nel qual caso può mancare il massimo o il minimo).

Nel caso 2) la f ammette *minimo* ma non *massimo* (o viceversa se è infinita negativa).

Nel caso 3) non c' è *nè massimo nè minimo*. Ma *in tutti questi casi*, in cui la $f(x, y)$ è *continua all' infinito esiste* almeno un *minimo* o un *massimo* del suo *valore assoluto*.

Nel caso 4), in generale, non si può dir nulla (cfr. § 6).

§ 10. **Funzioni di più variabili.** — L' estensione al caso di più variabili è immediata e procede come segue.

Definizione. Dicesi *campo finito e continuo* $(x_1, x_2 \ldots x_n)$ ogni insieme C di *punti* (gruppi di numeri) $(x_1, x_2 \ldots x_n)$ variabili entro limiti assegnati:

1) $$a_1 \leq x_1 \leq a_1', \ldots \quad a_n \leq x_n \leq a_n',$$

in modo che:

2) dati due punti $(x_1^{(1)} \ldots x_n^{(1)})(x_1^{(2)} \ldots x_n^{(2)})$ di C, esista in C una *linea continua* avente per estremi i due punti:

$$x_1 = \varphi_1(t) \ldots x_n = \varphi_n(t),$$

dove le φ_1 sono funzioni continue (invertibili) in (t_1, t_2), tali che

$$\varphi_1(t_1) = x_1^{(1)} \ldots \varphi_n(t_1) = x_n^{(1)}$$
$$\varphi_1(t_2) = x_1^{(2)} \ldots \varphi_n(t_2) = x_n^{(2)};$$

3) dato in C un gruppo G di punti ogni *punto limite* di G appartenga a C; cioè se nell'intorno

$$\bar{x}_1 - h < x_1 < \bar{x} + h, \ldots \quad \bar{x}_n - h < x_n < \bar{x}_n + h,$$

per qualunque h, vi sono punti $(x_1 \ldots x_n)$ di G, il punto $(\bar{x}_1 \ldots \bar{x}_n)$ appartiene a G.

Si dice che è definita una *funzione*

$$z = f(x_1 \ldots x_n)$$

nel campo C, se, per ogni punto $(x_1 \ldots x_n)$ di C, è assegnato un valore di z.

La funzione f si dice *continua* nel campo, se — per ogni punto $(\bar{x}_1 \ldots \bar{x}_n)$ di esso — accade che, dato un numero positivo ε (piccolo ad arbitrio), si può determinare un h per modo che, in valore assoluto:

$$|f(x_1 \ldots x_n) - f(\bar{x}_1 \ldots \bar{x}_n)| < \varepsilon$$

ogni qualvolta

$$|x_1 - \bar{x}_1| < h \ldots |x_n - \bar{x}_n| < h;$$

cioè

$$\lim_{x_i = \bar{x}_i} f(x_1 \ldots x_n) = f(\bar{x}_1 \ldots \bar{x}_n).$$

Ogni funzione continua $f(x_1 \ldots x_n)$ *in un campo* C *finito e continuo ammette un massimo e un minimo e prende effettivamente in punti di* C *tutti i valori compresi fra questi due estremi.*

Una funzione $f(x_1 \ldots x_n)$ può esser definita in un *campo continuo infinito*, p. es. per tutti i valori *positivi* delle x_i, od anche per *tutti* i valori delle x_i o — come si dice — in *tutto lo spazio* $(x_1 \ldots x_n)$.

Questo caso (come nel § 9) conduce a considerare il

$$\lim_{\infty} f(x_1 \ldots x_n).$$

Se questo limite esiste, finito o infinito che sia, la f dicesi *continua anche all'infinito*. In questa ipotesi:

1) Se f è *finito* all'infinito, ammette *massimo* e *minimo purchè* $\lim_{\infty} f$ *non* sia appunto il suo *limite superiore* o *inferiore*.

2) Se la f è *infinita positiva* (o negativa) all'infinito, *non* c'è *massimo ma* c'è *minimo* (o viceversa).

3) Se è *infinita senza segno*, *manca* il *massimo* e il *minimo.*

Ma *in ogni caso* c'è o il *minimo* oppure il *massimo* del *valore assoluto* $|f(x_1, ... x_n)|$. Questo non può mancare che nel caso in cui si abbia all'infinito una *discontinuità* essenziale di f.

§ 11. **Funzioni definite in un gruppo chiuso di punti.** — Nel paragrafo precedente abbiamo enunciato il teorema fondamentale sull'esistenza del massimo e del minimo di una funzione continua di n variabili $f(x_1, x_2 .... x_n)$ definita in un campo C finito e continuo.

La dimostrazione di codesto teorema è stata sviluppata nel caso $n = 2$, da cui si passa — con facile estensione — al caso di n qualunque. Essa si basa sul lemma che « se la funzione continua f prende due valori diversi a e b in due punti di C, la f prende — in qualche punto di C — tutti i valori compresi fra a e b ».

Ma — riprendendo e sviluppando le considerazioni del § 7 sui punti limiti d'un gruppo infinito — si può svolgere la dimostrazione in altro modo, senza appoggiarsi a codesto lemma, e si riesce quindi a stabilire l'esistenza del massimo e del minimo sotto ipotesi più generali, indipendentemente dalla continuità del C (a cui la validità del lemma precedente è strettamente legata). Si abbia un gruppo C di punti $(x_1, x_2 .... x_n)$ che sia

1) *finito*:

$$a_i \leq x_i \leq a_i \qquad (i = 1 .... n);$$

2) *chiuso,* cioè tale che se G è un gruppo di punti appartenenti a C, ogni *punto limite* di G *appartenga* a C.

In queste ipotesi vogliamo dimostrare che *ogni funzione continua* $f(x_1 .... x_n)$ *definita pér i punti del gruppo* C *ammette massimo e minimo.*

Per dimostrare il teorema occorre stabilire due fatti:

Che la funzione f è limitata, cioè (§ 3) ammette un limite superiore l (e un limite inferiore).

Che la f prende effettivamente il valore l in un punto di C.

Pongasi dapprima che la f possa assumere (in punti di C) valori superiori a qualsiasi numero positivo l, grande ad arbitrio. Ci sarà dunque un gruppo G di punti di C per cui

$$f \geq l.$$

Similmente vi sarà un gruppo G_1 di punti per cui

$$f \geq l + 1,$$

. . . .

ed un gruppo G_m di punti per cui

$$f \geq l + m.$$

Il gruppo G contiene $G_1 \dots G_m \dots$; G_1 contiene G_2, $G_3 \dots$; e in generale G_m contiene G_p, per $p > m$.

Ciascuno di questi gruppi contiene infiniti punti.

Dividiamo gli intervalli $a_i' - a_i$, ciascuno in 2 parti uguali, ponendo $h_i = \frac{a_i' - a_i}{2}$. Il campo (parallelipipedo ad n dimensioni) definito dalle diseguaglianze:

$$a_i \leq x_i \leq a_i' \qquad (i = 1, \dots n),$$

verrà diviso in 2^n campi (parallelepipedi), per ciascuno dei quali la x_i varia in una metà dell'intervallo primitivo, cioè fra a_i e $a_i + h_i$ o fra $a_i + h_i$ e a_i'.

I suddetti 2^n parallelepipedi si possono considerare in un ordine determinato; si determinerà un tale ordine confrontando anzitutto l'estremo superiore per x_i e fissando che fra due parallelepipedi P, P' preceda eventualmente quello in cui codesto estremo è minore; nel caso in cui P, P' abbiano lo stesso intervallo di variazione per x_1, si confronteranno analogamente gli estremi superiori per x_2, e così di seguito. Per brevità diremo che la: serie dei parallelepipedi è ordinata *secondo i valori crescenti* di $x_1 \dots x_n$.

Ora, fra i 2^n parallelepipedi suddetti, ve n'è almeno uno che contiene infiniti punti di G_1. Consideriamo il *primo* fra questi parallelepipedi nell'ordine stabilito, e sia P_1.

Dividiamo P_1 in 2^n parallelepipedi, bisecando come innanzi gl'intervalli di variazione delle x_i, e ordiniamo ancora questi parallelepipedi secondo i valori crescenti di $x_1 \dots x_n$. Vi sarà *un* primo parallelepipedo P_2 di questa serie, che contiene infiniti punti di G_2, e così di seguito.

Proseguendo indefinitamente l'operazione precedente otteniamo una seria illimitata di parallelepipedi

$$P_m \quad (m = 1, 2, ...)$$

la quale tende ad un punto limite:

$$x_i \equiv a_i + i_0 h_i + i_1 \frac{h_i}{2} + ,... + i_m \frac{h_i}{2^m} + ...$$

In questa serie i_m designa uno dei due numeri 0, 1, e così appare la convergenza della serie stessa.

Ora il punto anzidetto, punto limite di G_1, G_2,.... G_m.... deve appartenere per ipotesi al nostro campo chiuso C; in esso la f dovrebbe dunque assumere un valore limite di $l + m$, mentre questo limite è infinito.

Si deduce che la f è limitata e perciò (§ 3) ha un limite superiore l.

Si dimostra nello stesso modo che f assume il valore l in un punto di C.

Se in nessun punto $f = l$, vi saranno punti costituenti un gruppo infinito G_1 per cui

$$| f(x_1 .... x_n) - l | < e;$$

quindi infiniti punti costituenti un gruppo G_m per cui

$$| f(x_1 .... x_n) - l | < \frac{e}{2^m} \qquad (m = 1, 2 ....).$$

Il gruppo G conterrà G_1; G_1, G_2 e così di seguito.

Si dividerà come innanzi il parallelepipedo (a_i, a_i') in 2^n parallelepipedi, ordinati secondo i valori crescenti di $x_1 .... x_n$; quindi si considererà il primo parallelepido P_1 contenente infiniti punti di G_1 e così di seguito. La serie dei parallelepipedi $P_m (m = 1, 2, ...)$ ammette un punto limite, appartenente C, in cui

$$f(x_1 .... x_n) = l.$$

§ 12. **Massimi e minimi vincolati.** — Si abbia una funzione continua $f(x_1 .... x_n)$, definita p. es. in tutto lo spazio $(x_1 .... x_n)$ o per i valori $x_i \geq 0$, o in un campo finito e continuo. Sieno

$\varphi_1(x_1 \ldots\ldots x_n) \ldots\ldots \varphi_r(x_1 \ldots\ldots x_n)$, $r < n$ funzioni continue nello stesso campo; e suppongasi che le equazioni

$$\varphi_1(x_1 \ldots\ldots x_n) = 0 \ldots\ldots \varphi_r(x_1 \ldots\ldots x_n) = 0$$

sieno compatibili ed ammettano infinite soluzioni contenute in un *campo finito*. Allora *la funzione* $f(x_1 \ldots\ldots x_n)$ *considerata nel gruppo dei punti* (x_i) *per cui* $\varphi_1 = \ldots\ldots \varphi_r = 0$, *ammette un massimo e un minimo.*

Basta dimostrare che il gruppo dei punti (x_i) suddetti, che si è supposto finito, è chiuso; si applicherà quindi il teorema del precedente paragrafo.

Ora si consideri un gruppo G contenente infiniti punti (x_i) per cui $\varphi_1 = \ldots\ldots = \varphi_r = 0$, e sia $L = (l_i)$ un punto limite di G. Ciò significa che preso un ε piccolo ad arbitrio vi sono degli x_i per cui

$$\varphi_1(x_1 \ldots\ldots x_n) = \ldots\ldots = \varphi_r(x_1 \ldots\ldots x_n) = 0$$

$$|x_i - l_i| < \varepsilon \qquad (i = 1, \ldots\ldots n).$$

Ma, per la continuità di $\varphi_1 \ldots\ldots \varphi_r$, dato un σ piccolo ad arbitrio, si può sempre determinare un ε tale che

$$|x_i - l_i| < \varepsilon$$

per

$$|\varphi(x_1 \ldots\ldots x_n) - \varphi_1(l_1 \ldots\ldots l_n)| < \sigma, \ldots\ldots \quad |\varphi_r(x_1 \ldots\ldots x_n) - \varphi_r(l_1 \ldots\ldots l_r)| < \sigma;$$

si possono scegliere delle x_i per cui

$$\varphi_1(x_i \ldots\ldots x_n) = \ldots\ldots = \varphi_r(x_1 \ldots\ldots x_n) - 0,$$

e si deduce

$$\varphi_1(l_1 \ldots\ldots l_n) = \ldots\ldots = \varphi_r(l_1 \ldots\ldots l_n) = 0. \qquad \text{c. d. d.}$$

§ 13. **Applicazione a problemi elementari.** — I teoremi di esistenza del massimo e del minimo per le funzioni continue, che abbiamo stabiliti nei precedenti paragrafi, vengono a colmare una lacuna che si trova nelle ricerche classiche, giustificando il presupposto accolto fino dai geometri antichi e rendendo quindi legittimi i semplici metodi che vi si fondano.

Anzitutto, a proposito del teorema posto come fondamentale nell'Art. XXV, si ha che il prodotto

$$z = x_1 x_2 \ldots\ldots x_n$$

e la somma

$$x_1 + x_2 + \dots x_n$$

sono funzioni continue delle variabili x_i. Per

$$x_i \geq 0, \quad x_1 + x_2 + \dots x_n = k \qquad (> 0),$$

si ha un campo finito di variazione delle x_i, e quindi *vi sarà un massimo* (e un minimo $z = 0$) *del prodotto* $x_1 x_2 \dots x_n$ $(x_i \geq 0)$ *subordinatamente alla condizione* $\Sigma x_i = k$.

Pertanto, ove si abbia a determinare codesto massimo, siamo ora autorizzati a dedurre ch' esso corrisponde a

$$x_1 = x_2 = \dots = x_n,$$

giacchè se p. es. $x_1 \neq x_2$ il prodotto z può essere aumentato sostituendo ad $x_1 x_2$, due fattori uguali ad $\frac{x_1 + x_2}{2}$ (cfr. Articolo XXV, §§ 9, 10).

Passiamo ora ai teoremi isoperimetrici di Zenodoro e di Cramer, svolti nell'Art. XXVI.

Anche qui possiamo giustificare il presupposto delle dimostrazioni classiche, che:

1) Fra tutti i poligoni di n lati, con un dato perimetro, c' è un poligono d' area massima.

2) Fra tutti i poligoni con n lati dati, c' è parimente un massimo.

Per mostrare che questi resultati sono effettivamente contenuti nel teorema generale del § 12, possiamo procedere come segue.

Si consideri un poligono con n vertici successivi

$$(x_1, y_1), (x_2, y_2), \dots (x_n, y_n),$$

posto nel semipiano positivo $(y_i \geq 0)$. L' area A del poligono è data dal valore assoluto della somma

$$\sum_{i=1}^{i=n} (x_{i+1} - x_i) \frac{y_i + y_{i+1}}{2},$$

dove i singoli termini sono presi ciascuno col proprio segno che è quello della differenza $x_{i+1} - x_i$.

Il lato l_i del contorno, che congiunge i vertici (x_i, y_i) (x_{i+1}, y_{i+1}) vale

$$l_i = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}.$$

Ora la

$$\Sigma_i (x_{i+1} - x_i) \frac{y_{i+1} + y_i}{2}$$

— che per semplicità supporremo positiva e quindi $= A$ — è una funzione continua delle x_i, y_i e così pure le l_i.

È lecito supporre che uno dei vertici del poligono, p. es. (x_i, y_i), abbia — nel piano — una posizione data; allora la condizione

$$\overset{1}{\underset{n}{}} l_i = k \qquad (< 0),$$

delimita un campo finito di variazione per le $x_i y_i$, dovendo essere

$$\sqrt{(x_i - x_1)^2 + (y_i - y_1)^2} \leq l_1 + \dots + l_{i-1} \leq k.$$

È dunque il caso di applicare il teorema generale del § 12, e dedurre che *per un dato perimetro*, o *per dati lati*, l' *area* A del *poligono* di *n lati ammette un massimo.*

Quindi le dimostrazioni classiche fondate su tale presupposto (cfr. Art. XXV), divengono legittime.

Qui è utile osservare che, non soltanto le trattazioni sintetiche di cui si discorre in quell'Articolo, ma anche la trattazione analitica dal teorema di CRAMER, data da LAGRANGE col calcolo delle variazioni [1], acquista valore probativo solo in forza di questo *teorema d' esistenza.* In realtà i calcoli eleganti, ma laboriosi, di LAGRANGE, non provano che il poligono iscritto in un cerchio sia massimo fra quelli aventi gli stessi lati, ma solo che codesto poligono soddisfa alle condizioni necessarie per essere massimo; onde segue che esso è effettivamente il massimo a patto che sia dimostrata l' esistenza d' un massimo.

NOTA. Con un cambiamento di variabili si può far dipendere la funzione A (area del poligono) direttamente dalle lunghezze l_i dei suoi lati e dai suoi angoli ω_i. E si può anche delimitare un campo finito di variazione per le varia-

[1] « Essai d'une nouvelle Méthode pour Déterminer les maxima et les minima... ». Misc. Taur II, 1760-61. App. II. « Oeuvres » Vol. I, p. 357.

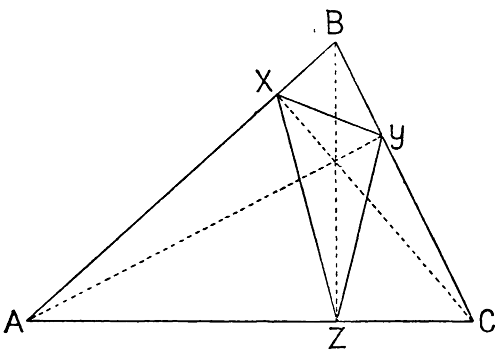
bili l_i, ω_i, in corrispondenza a *poligoni convessi*. Si hanno così per queste variabili delle diseguaglianze a cui si aggiungeranno le *relazioni d'uguaglianza* richieste dall'esistenza del poligono, ed esprimibili coll'annullarsi di funzioni continue. L'area A risulterà funzione continua delle l_i, ω_i; e di essa potrà affermarsi l'esistenza d'un massimo, subordinatamente alle condizioni suddette.

Indichiamo un altro problema classico in cui la nozione dell'esistenza permette facilmente di determinare il minimo.

Si tratta di trovare il *triangolo di perimetro minimo fra quelli che sono iscritti in un triangolo dato* ABC [1].

Si prendano sui lati AB, BC, CA risp. tre punti X, Y, Z. Il perimetro del triangolo XYZ è funzione continua delle variabili AX, BY, CZ (in un campo spaziale) e però ammette un minimo. Ma occorre distinguere due casi:

1) Esiste un minimo *proprio*, cioè per X, Y, Z interni ai rispettivi lati. Allora si ha propriamente un triangolo minimo XYZ. La spezzata $XZ + XY$ deve esser minima fra quelle che hanno per estremi Z, Y e il vertice X sulla retta AB; si conclude quindi (*proprietà della riflessione della luce* cfr. Art. XXVI) che ZX e YX devono essere ugualmente inclinate su AB. Analogamente si dica per le XY, ZY e per le XZ, YZ.



Si deduce facilmente che X, Y, Z sono i *piedi delle altezze del triangolo* che — nelle ipotesi fatte — debbono cadere internamente ai lati. Il triangolo ABC *è acutangolo*.

2) Il *minimo è improprio* cioè due dei punti X, Y, Z, si riuniscono in un vertice del triangolo ABC; p. es. X, Y cadono in A. Ciò accade se l'angolo in A è retto o ottuso,

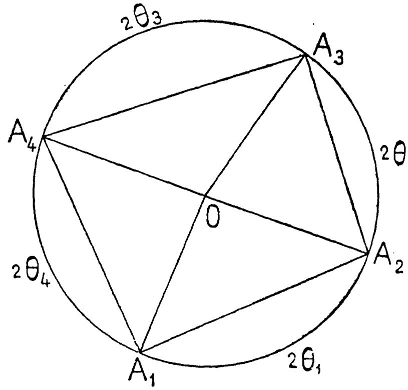
[1] Cfr. STEINER, Werke, Bd II, p. 232. CERTO, Giornale di Mat., voll. 23 e 26. R. STURM, « Maxima und minima in der elementaren Geometrie ». Leipzig. Teubner, 1910.

nel qual caso l'altezza condotta da A su BC, contata due volte, è minore del perimetro di qualsiasi triangolo iscritto in ABC.

Analogo al precedente è il problema di determinare *il poligono* (convesso) *di* n(>2) *lati, di perimetro massimo* (*e di area massima*) *fra quelli che sono iscritti in un cerchio.*

Premettiamo che la costruzione di un poligono $A_1 A_2 \dots A_n$ iscritto in un cerchio dipende dalla divisione dell'angolo al centro in n angoli $2\theta_1, 2\theta_2 \dots 2\theta_n$:

$$\theta_1 + \theta_2 + \dots \theta_n = \pi.$$



Il perimetro del poligono è una funzione continua delle n variabili indipendenti

$$\theta_1 \dots \theta_{n-1},$$

nel campo finito delimitato dalle relazioni

$$\theta_i \geq 0, \quad \theta_1 + \dots + \theta_{n-1} \leq \pi;$$

questa funzione ammetterà dunque un massimo o un minimo.

Il massimo si determina subito in base alla nozione della sua esistenza. Infatti la spezzata $A_1 A_2 A_3$, iscritta nell'arco di cerchio $A_1 A_2$, dovendo godere della proprietà di massimo, si ha che $A_1 A_2 = A_2 A_3$. Analogamente $A_2 A_3 = A_3 A_4$ ecc., cioè *il poligono di* n *lati di perimetro massimo fra quelli che sono iscritti in un cerchio è il poligono regolare.*

(Il minimo si riduce al poligono che ha due lati uguali al diametro e gli altri nulli).

Il poligono regolare anzidetto corrisponde al *massimo dell' area,* come si prova collo stesso ragionamento.

§ 14. **Nota storica.** — La distinzione chiara fra massimo e minimo e limite superiore e inferiore appartiene a Bolzano; essa fu poi svolta da Weierstrass nelle sue lezioni di Berlino. La critica del concetto e del postulato di continuità (Cfr. Vol. I, Art. V) ha condotto Weierstrass alla dimostrazione rigorosa dei teoremi generali d'esistenza dei massimi e minimi per le funzioni continue.

Giova peraltro ricordare che già in alcuni casi particolari importanti si era presentata sostanzialmente la stessa analisi; p. es. la dimostrazione data da GAUSS del teorema fondamentale che un' equazione algebrica

$$f(x) = a_0 x^n + a_1 x^{n-1} + ... + a_n = 0$$

ha (una e quindi) n radici, consiste nel dimostrare che il modulo di f (funzione continua di x, y) ha per limite inferiore 0, e che questo limite è un *minimo effettivamente raggiunto* (Cfr. Vol. II, Art. IX).

Un' esposizione e uno sviluppo della teoria generale delle funzioni continue trovasi in U. DINI « Fondamenti per la teorica delle funzioni di variabile reale » (Pisa, 1878) a cui si riattacca tutta una scuola di critici e di ricercatori.

II. - **Analisi infinitesimale dei massimi e minimi.**

§ 15. **Curve e funzioni.** — L'analisi infinitesimale dei massimi e minimi è suggerita dalla *rappresentazione geometrica* di una funzione continua $y = f(x)$, mediante una *curva.*

A proposito di questa rappresentazione sono da fare le osservazioni seguenti:

1) Il concetto di funzione arbitraria $y = f(x)$, in un intervallo (a, b), è posto in tutta la sua generalità mediante una *definizione logica,* a cui risponde una infinità di enti possibili. La condizione della *continuità* restringe l'arbitrarietà della operazione o legge corrispondente ad f, ma non muta il carattere della definizione logica.

Invece il *concetto geometrico di linea* o *curva* è definito per astrazione come termine generale della classe di tutte le curve intuitivamente immaginabili.

Si tratta dunque di comparare e classificare enti psicologicamente dati, che si presentano al critico definitore come enti ch' egli non pone ad arbitrio ma trova innanzi a sè, quali oggetti proposti alla sua osservazione.

2) Non è detto che sia possibile segnare limiti alla nostra immaginazione costruttrice di curve, così da raggiungere una definizione della curva in generale. Nondimeno l' intuizione

attribuisce alle curve alcuni caratteri che non appartengono sempre al luogo rappresentato da una funzione continua

$$y = f(x);$$

come d' altra parte ci fa riconoscere l' esistenza di curve non rappresentabili in tal modo.

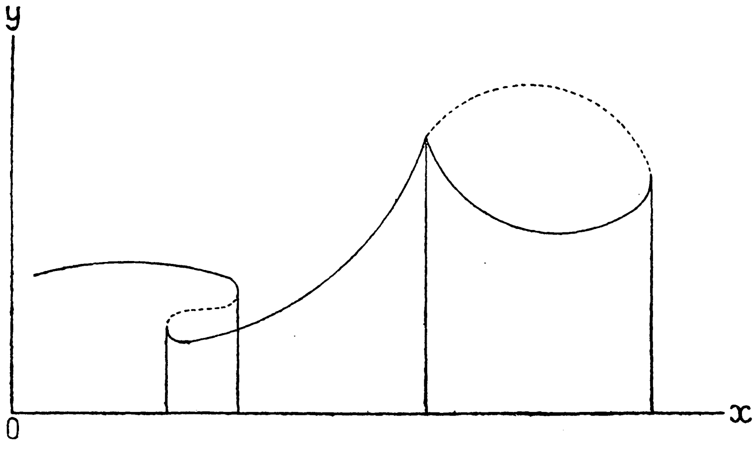
3) Una curva (continua), che immaginiamo rappresentata totalmente nel campo della nostra intuizione, si lascia spezzare in un *numero finito di curve* intersecate in *un* punto dalle rette di una striscia parallela all' asse delle y.

Ciascuna di queste parti è rappresentata da un' equazione

$$y = f(x) \qquad \text{in } (ab),$$

ove f è continua.

4) Una curva continua intuitiva $y = f(x)$ ammette in ogni punto una *tangente a destra* e una *tangente a sinistra*,



il che importa l' esistenza della derivata a destra e a sinistra $f'(+x)$, $f'(-x)$ in ogni punto dell' intervallo di variazione (¹).

(¹) Per definizione:

$$f'(+x) = \lim_{h=+0} \frac{f(x+h) - f(x)}{h};$$

$\frac{f(x+h) - f(x)}{h}$ è la tangente trigonometrica dell' angolo formato dall' asse delle x colla retta che unisce i punti x e $x+h$ della curva; il suo limite (per $h = +0$) corrisponde alla tangente in x all' arco di curva a destra che ha per estremo quel punto.

5) Fatta eccezione da un *numero finito* di punti (*angoli*), la tangente a destra e a sinistra in ogni altro punto della curva coincidono; ciò importa l'*esistenza e l'unicità della derivata* $f'(x)$.

6) L'esercizio ulteriore dell'intuizione sopra le curve immaginabili, conduce ancora a restringere la natura della funzione $f(x)$ che rappresenta una *curva*, nel senso geometrico della parola.

In ispecie si è condotti ad ammettere che se $y=f(x)$ rappresenta una curva intuitiva, *la funzione* f *possegga tutte le derivate successive salvo per un numero finito di punti*, in cui si ha una discontinuità fra le derivate a destra e a sinistra.

Si può anche ammettere di più che la funzione f si riduca ad un *numero finito di funzioni analitiche* o *serie di potenze* (di Taylor).

In modo analogo si può ammettere che una *superficie*, nel senso geometrico-intuitivo della parola, si possa decomporre in un numero finito di superficie rappresentabili con *funzioni analitiche* $x=f(z, y)$, la f essendo *sviluppabile* per x, y, in *serie di potenze convergenti* in un certo campo.

§ 16. Massimi e minimi relativi: annullamento della derivata. — Conformemente al carattere analitico o microscopico dell'analisi infinitesimale, il problema generale di ricercare i massimi e minimi di una funzione $f(x)$, viene ricondotto, dall'intervallo di variazione in cui è data f, ad intervalli convenientemente piccoli, diguisachè si è condotti alla seguente

Definizione. Un punto $\bar{x}$ si dice *punto di massimo* (o *di minimo*) *relativo* per la funzione $y=f(x)$, se in quel punto la funzione assume un valore *maggiore* (o *minore*) *dei valori nei punti vicini*, cioè se si può determinare un h (>0) tale che

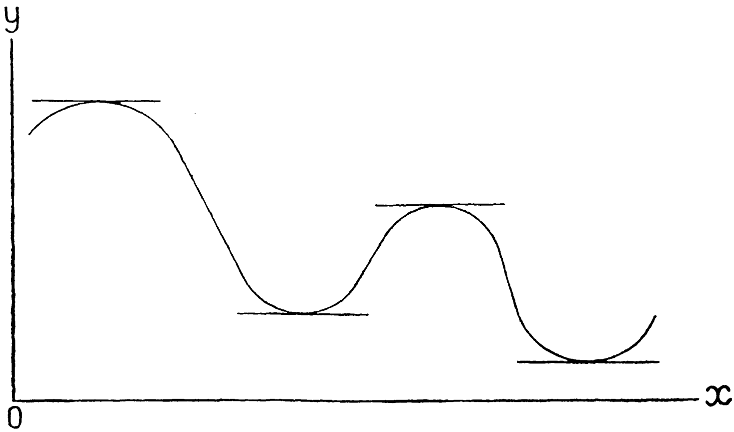
$$f(x) < f(\bar{x})$$

ogni qualvolta

$$|x-\bar{x}| < h.$$

È chiaro che il *massimo* (o il *minimo*) assoluto di f (se esiste) è il massimo (o risp. il minimo) di tutti i massimi (minimi) relativi.

Nell'unita curva $y = f(x)$, sono segnati i massimi e minimi relativi diversi dagli estremi.



Appare tosto che in essi la tangente alla curva è parallela all'asse delle x.

Questa osservazione si traduce nel

Teorema. *Se $f(x)$ è una funzione continua e derivabile, ogni punto di massimo e minimo (relativo) di essa, che non sia un estremo dell'intervallo di variazione di x, è un punto per cui*

$$f'(x) = 0.$$

La dimostrazione è immediata. Si ha infatti per definizione:

$$f'(x) = \lim_{h=0} \frac{f(x+h) - f(x)}{h};$$

quindi se

$$f'(x) > 0$$

si può determinare un numero positivo h tale che per $|h| < h$

$$\frac{f(x+h) - f(x)}{h} (= f'(x) + \varepsilon) < 0,$$

ed allora, per $h > 0$,

$$f(x+h) > f(x) > f(x-h);$$

si deduce che $f(x)$ è sempre *crescente* nell'intervallo fra $x - h$, $x + h$, e non ha massimi nè minimi in x.

Ad analoga conclusione si viene per

$$f'(x) < 0:$$

la f è *decrescente* in x (cioè fra $x-h$, $x+h$).

Dunque un massimo o un minimo di f, fuori degli estremi, è un punto per cui

$$f'(x) = 0.$$

NOTA I. Giova notare subito che questa *condizione è necessaria ma non sufficiente*: p. es. per la funzione

$$y = f(x) = x^3$$

si ha

$$f'(0) = 0,$$

ma il punto 0 non è nè un massimo nè un minimo. Infatti x^3 è sempre crescente da $-\infty$ a $+\infty$.

NOTA II. La conclusione relativa all'annullarsi della derivata $f'(x)$ in un punto di massimo o di minimo di $f(x)$ è essenzialmente fondata sull'*ipotesi* che $f(x)$ ammetta, nel punto, una *derivata finita* e *unica*: derivata a destra = derivata a sinistra.

Così la funzione

$$y = x^{\frac{2}{3}}$$

ha un minimo per $x=0$, mentre

$$y_0' = \left(\frac{2}{3}x^{-\frac{1}{3}}\right)_0 = \infty.$$

(Cfr. il caso della cuspide nel § 21).

§ 17. **Massimi e minimi assoluti.** — Abbiam visto che il teorema precedente esprime una condizione necessaria ma non sufficiente per l'esistenza di massimi e minimi relativi d'una funzione continua e derivabile in un intervallo (ab).

Ma quella condizione basta a risolvere il problema di determinare il massimo e il minimo assoluto della funzione, almeno se — escluso che f rimanga costante in qualche tratto dell'intervallo — si suppone, conformemente all' intuizione, che $f'(x)$ *si annulli* soltanto *un numero finito di volte* fra a e b.

Sieno, in ordine crescente,

$$x_1 < x_2 \ldots . < x_n$$

gli zeri di $f(x)$:

$$a < x_1, \quad b > x_n.$$

Per il teor. prec. la funzione $f(x)$ è sempre crescente o decrescente in ciascuno degli intervalli

$$(ax_1) \quad (x_1x_2) \ldots . \quad (x_nb).$$

Dunque il massimo (minimo) assoluto di f è il massimo (minimo) dei numeri:

$$f(a), \quad f(x_1), \quad f(x_2) \ldots . f(x_n), \quad f(b),$$

§ 18. **Determinazione dei massimi e minimi relativi: metodo sintetico.** — Il metodo sintetico sopra esposto (in cui si considera la funzione f, fra a e b, nella sua interezza) permette anche di risolvere la questione « *se un punto di zero di* $f'(x)$ *corrisponda ad un massimo* o *a un minimo relativo di* $f(x)$ ».

Si consideri p. es. il punto x_2 in cui $f'(x_2) = 0$.

Si devono paragonare i valori di

$$f(x_1), \quad f(x_2), \quad f(x_3);$$

se

$$f(x_1) < f(x_2), \quad f(x_3) < f(x_2),$$

x_2 è un massimo;

se

$$f(x_1) > f(x_2), \quad f(x_3) > f(x_2),$$

x_2 è un minimo;

se

$$f(x_1) < f(x_2) < f(x_3)$$

o

$$f(x_1) > f(x_2) > f(x_3),$$

nel punto x_2 non c'è nè massimo nè minimo.

Nota. L' ipotesi

$$f(x_1) = f(x_2)$$

porterebbe f costante e

$$f' = 0 \quad \text{fra} \quad x_1 \text{ e } x_2,$$

contrariamente al supposto che f' abbia un numero finito di zeri fra a e b.

D' altronde il caso escluso non presenta speciali difficoltà.

Supponendo che la derivata sia ovunque continua e derivabile nell' intervallo (ab), la ricerca delle radici dell' equazione

$$f'(x) = 0,$$

da cui dipende la determinazione dei massimi e minimi di $f(x)$, si può riattaccare alla risoluzione dell' equazione

$$f''(x) = 0,$$

da cui dipende la ricerca dei massimi e minimi di f'. Infatti si supponga che f'' si annulli nei punti

$$x_1', \quad x_2', \ldots, \quad x_m',$$

scritti per ordine crescente. Basterà considerare il segno di f' in a, b e in ciascuno di questi punti: se $f'(x_i')$ e $f'(x'_{i+1})$ hanno lo stesso segno, fra x_i' e x'_{i+1} $f'(x)$ è sempre crescente o decrescente e perciò non si annulla; invece se p. es.

$$f'(x_i') < 0 < f'(x'_{i+1}),$$

fra x_i' e x'_{i+1} cade certo uno zero di $f'(x)$ (cfr. § 5.

L' osservazione precedente contiene un *procedimento ricorrente* per ridurre la ricerca dei massimi e minimi della funzione $f(x)$, a quella delle sue derivate successive. Applicato al caso di una *funzione razionale* questo procedimento conduce ai noti *metodi per la risoluzione numerica* delle *equazioni algebriche.*

§ 19. Condizioni analitiche per i massimi e minimi. — In luogo di considerare sinteticamente la funzione $f(x)$ in tutto l' intervallo (ab) in cui essa è definita, le derivate successive di f permettono di assegnare le *condizioni sufficienti perchè un dato punto di zero di* $f'(x)$ *sia un massimo o un minimo* relativo di $f(x)$.

Rendiamoci conto anzitutto del significato geometrico del problema.

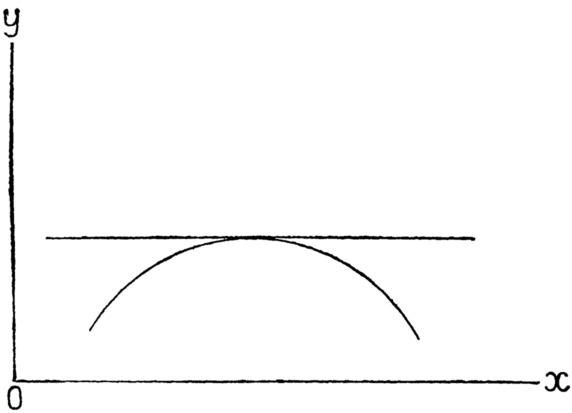
Nei punti per cui

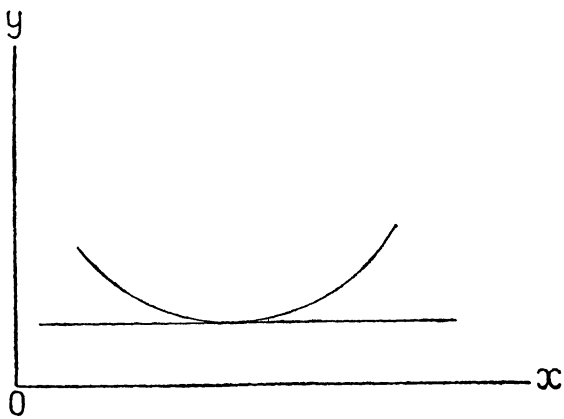
$$f'(x) = 0,$$

la curva $y = f(x)$ ha la tangente parallela all'asse della x.

Ma due casi possono presentarsi:

1) se il punto in questione è un punto generico della curva, la curva stessa resta tutta da una parte (al disotto o

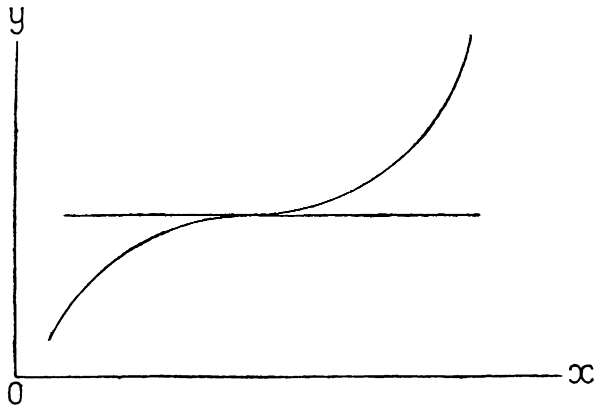




al disopra) della tangente suddetta, almeno per un intorno del punto di contatto; in tale ipotesi si ha un massimo o un minimo di f;

2) ma può accadere invece che il punto sia un *flesso*

della curva e la tangente traversi la curva stessa, in tale ipotesi non vi è massimo nè minimo.



La considerazione della derivata seconda $f''(x)$ illumina la questione.

Pongasi infatti che sia

$$f'(\bar{x}) = 0, \quad f''(\bar{x}) \neq 0.$$

In tale ipotesi si ha un massimo o un minimo di $f(\bar{x})$ secondochè

$$f''(\bar{x}) < 0 \quad \text{o} \quad > 0.$$

Invero sussiste — come è noto — la formula

$$f(\bar{x} + h) = f(\bar{x}) + hf'(\bar{x}) + \frac{h^2}{1 \cdot 2} f''(\bar{x}) + \varepsilon,$$

dove ε diventa infinitesimo con h d'ordine superiore al secondo. Nel nostro caso si ha dunque

$$f(\bar{x} + h) - f(\bar{x}) = \frac{h^2}{1 \cdot 2} f''(x) + \varepsilon,$$

e però — quando h è assai piccolo —

$$f(\bar{x} + h) - f(x) > 0 \quad \text{se} \quad f''(\bar{x}) < 0$$
$$f(\bar{x} + h) - f(x) < 0 \quad \text{se} \quad f''(\bar{x}) > 0.$$

Ma che cosa si può dire se $f''(\bar{x}) = 0$?

In tale ipotesi non è lecito dire senz'altro che si abbia un flesso della curva e manchi massimo o minimo. Tale conclusione è legittima soltanto se

$$f'''(\bar{x}) \neq 0,$$

giacchè

$$f(\bar{x}) + h) - f(\bar{x}) = \frac{h^3}{1 \cdot 2 \cdot 3} f'''(\bar{x}) + \varepsilon,$$

per h piccolo ha il segno di $h^3 f'''(\bar{x})$ che passa — con h — dal positivo al negativo.

Si può ormai enunciare la regola generale, fondata sopra la formula

$$f(\bar{x} + h) = f(\bar{x}) + hf'(\bar{x}) + ... + \frac{h^n}{n!} f^{(n)}(\bar{x}) + \varepsilon,$$

dove ε diventa infinitesimo con h, d'ordine superiore ad h^n;

Teorema. *Se la funzione* $f(\mathrm{x})$ *ammette tutte le derivate successive fino all'ordine* n *in un punto* $\bar{\mathrm{x}}$, *e se tutte le derivate suddette sono nulle meno la* n^{ma} $f^{(\mathrm{n})}(\bar{\mathrm{x}})$, *la* f(x) *non ha massimo nè minimo ove* n *sia dispari; invece se* n *è pari vi è un massimo quando*

$$f^{(n)}(\bar{x}) < 0$$

e un minimo per

$$f^{(n)}(x) > 0.$$

Nota I. Nell'ipotesi precedente la curva ha in x un contatto d'ordine $n - 1$ (n-punto) colla sua tangente. L'analisi precedente prova che la tangente lascia la curva tutta da una parte se n è pari, ed invece la traversa se n è dispari.

Così p. es. la curva

$$y = x^n$$

ha un contatto d'ordine $n - 1$ coll'asse delle x in $x = 0$. La curva sta tutta al disopra dell'asse (si ha un minimo nel punto 0) se n è pari $= 2m$:

$$x^{2m} > 0 \quad \text{per} \quad x \neq 0;$$

invece, per $n = 2m + 1$, la curva passa da sotto a sopra dell'asse mentre x passa da valori < 0 a valori > 0.

Nota II. Il teorema precedente dà un criterio risolutivo per ogni funzione $y = f(x)$ che ammetta le derivate successive, purchè queste non sieno tutte nulle. Ma tale circostanza non si presenta per funzioni — diverse da $y = \text{cost}$ — che rappresentino curve intuitive.

§ 20. **Funzione implicita definita dall'equazione $f(x, y) = 0$: analisi dei punti doppi della curva.** — Sia $f(x, y)$ una funzione algebrica o *analitica*, cioè una serie di potenze di x, y. L'equazione $f(x,y) = 0$ rappresenta (generalmente) una curva, composta di uno o più rami, qualcuno dei quali può estendersi anche all'infinito, ma che — fra due estremi qualsiasi — risponde sempre ai requisiti di una curva intuitiva.

Risolvendo l'equazione $f = 0$ rispetto ad y a partire da una soluzione x_0, y_0, si ha y *funzione implicita* (analitica) della x, generalmente *polidroma*, cioè si hanno più funzioni univoche $y(x)$ che si riattaccano per continuità, prolungandosi l'una nell'altra.

In ciascun punto (x, y) vi è generalmente *una* tangente alla curva $f = 0$, che corrisponde alla derivata

$$y'(x) = -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}}.$$

Ma s'incontra un'eccezione essenziale se

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0.$$

Questo caso corrisponde ad un punto *doppio* (o *multiplo*) della curva $f(x, y) = 0$: ogni retta per il punto ha ivi *due* (o *più*) intersezioni riunite colla curva.

Per esaminare più da vicino il comportamento della curva nell'intorno d'un punto siffatto, portiamo nel punto stesso l'origine 0 delle coordinate; avremo

$$f(x, y) = f_2(x, y) + f_3(x,y) + \dots.$$

dove f_2, $f_3 \dots$ sono *forme* (polinomi omogenei) in x, y, d'ordine crescente:

$$f_2(x, y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2, \text{ ecc.}$$

Anzitutto vi è luogo a considerare il discriminante di f_2

nel punto 0:

$$D = \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 - \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} = a_{12}{}^2 - a_{11}\, a_{22}.$$

I) *Caso iperbolico*:

$$D > 0.$$

Allora qualcuna delle a è $\neq 0$ e perciò 0 è un punto doppio e non multiplo con molteplicità > 2. L'equazione

$$f_2(x,y) = (\lambda_1 x + \mu_1 y)(\lambda_2 x + \mu_2 y) = 0$$

rappresenta due rette distinte, corrispondenti alle due radici $-\frac{\lambda_1}{\mu_1}$, $-\frac{\lambda_2}{\mu_2}$ dell'equazione

$$a_{11} + 2a_{12}\,\frac{y}{x} + a_{22}\left(\frac{y}{x}\right)^2 = 0.$$

Ponendo

$$y = -\frac{\lambda_1}{\mu_1}x$$

in

$$f(x, y) = f_2(x, y) + f_3(x, y) + \ldots.,$$

f_2 si annulla e perciò f si riduce ad un'espressione della forma

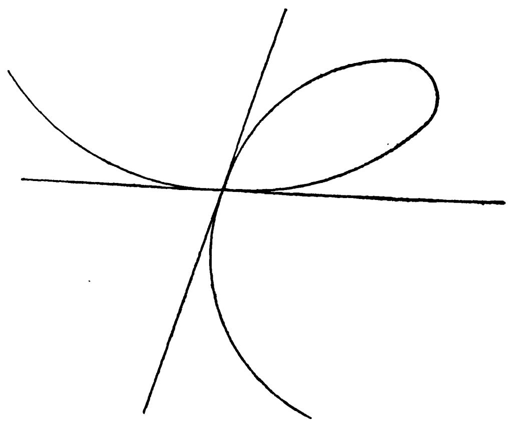
$$Ax^3 + Bx^4 + \ldots.,$$

che ha una radice tripla per $x = 0$.

Si vede così che le due rette

$$\lambda_1 x + \mu_1 y = 0, \quad \lambda_2 x + \mu_2 y = 0$$

sono *osculatrici* alla curva, avendo con essa almeno un contatto tripunto, e toccano *due rami della curva* passanti per 0.



Il punto 0 è un *nodo*.

II) *Caso elittico*:

$$D < 0.$$

Ancora il punto 0 è *doppio* per la curva.

Ma non ci sono rette osculatrici; infatti se la retta

$$\lambda x + \mu y = 0$$

deve avere un contatto tripunto con $f = 0$, bisogna che $\lambda x + \mu y$ sia fattore di f_2, mentre — nell' ipotesi adottata — f_2 si decompone solo in due fattori lineari immaginarii. Si deduce che *non* vi sono *rami della curva* per 0.

Il punto 0 è un *punto isolato.*

III) *Caso parabolico*:

$$D = 0.$$

In questa ipotesi possono nascere varie complicazioni. Anzitutto bisogna distinguere il caso in cui 0 sia soltanto *doppio*, da quello in cui sia *triplo* o *quadruplo* ecc.

Il criterio distintivo è semplice: l' annullarsi delle derivate seconde di f:

$$a_{11} = a_{12} = a_{22} = 0,$$

porta che il punto 0 sia (almeno) triplo ecc.

Pongasi che 0 sia soltanto doppio. Il tipo generale a cui esso appartiene dicesi *cuspide.*

Si avrà

$$f_2(x, y) = (\lambda x + \mu y)^2,$$

e con un cambiamento di coordinate:

$$f_2(x, y) = ay^2 \qquad (a \neq 0).$$

L' asse delle x, $y = 0$, è retta osculatrice ad $f = 0$, e dicesi *tangente cuspidale.*

Poniamo

$$f_3(x, y) = Ax^3 + Bx^2y + Cxy^2 + Ey^3.$$

Nel caso generale si ha

$$A \neq 0;$$

allora per $y = 0$

$$f(x, 0) = Ax^3 + ...,$$

ammette come radice tripla (e non quadrupla) $x = 0$.

Si ha allora la *cuspide ordinaria* o *di 1ª specie* caratterizzata sempre dalla circostanza che la tangente cuspidale ha colla curva un contatto tripunto (e non più elevato),

Se invece

$$A = 0,$$

la *tangente cuspidale* ha con f un *contatto quadripunto,* che diventa quintipunto solo se è nullo il coefficiente di x^4 in f_4.

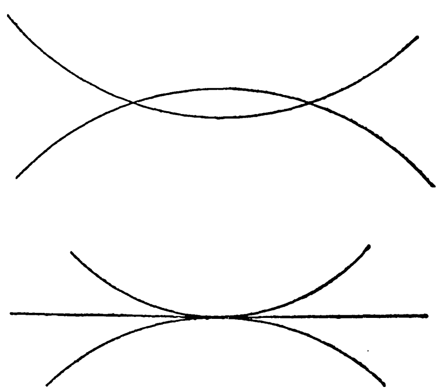
Ora notiamo che si ha in generale, a meno d'infinitesimi d'ordine superiore,

$$\left(\frac{\partial f}{\partial x}\right)_{dx,\,0} = 3\,Ax^2$$

$$\left(\frac{\partial f}{\partial y}\right)_{dx,\,0} = 0.$$

Dunque $A = 0$ si può interpretare come la condizione perchè la curva $f = 0$ abbia un secondo *punto doppio infinitamente vicino al punto* 0 nella direzione dell'asse x.

Più precisamente la curva $f = 0$ si presenta come limite di una curva che consta di due archi secantisi in due punti che si avvicinano indefinitamente.



Della singolarità che essa presenta, porge esempio l'annessa figura. Questa singolarità dicesi *tacnodo* (*iperbolico*), e si può subito immaginare una seconda forma di tacnodo (*ellittico*) che nasce come limite di due punti isolati avvicinantisi e perciò è punto isolato per $f = 0$. Ma è anche possibile una forma di *tacnodo parabolico* che dà luogo a successive distinzioni e complicazioni.

L'analisi della singolarità che la curva

$$f = y^2 + Ax^3 + Bx^2y + \ldots.$$

presenta nella cuspide 0, si può compiere mediante gli *sviluppi di* Puiseux che qui rapidamente accenniamo.

Per x vicino a 0 ci sono due valori di y che tendono a 0, i quali possono designarsi con

$$y = u + \sqrt{v}$$
$$y = u - \sqrt{v}.$$

Sostituendo ad y, $u + \sqrt{v}$ in f, si ha

$$f = u^2 + 2u\sqrt{v} + v + Ax^3 + Bx^2 (u + \sqrt{v}) + \ldots.$$

Questa espressione si riduce alla forma

$$f = P + Q\sqrt{v}.$$

Scrivendo

$$\begin{cases} P(u, v, x) = 0 \\ Q(u, v, x) = 0 \end{cases}$$

si ottiene un sistema d'equazioni che — secondo i teoremi generali sulle funzioni implicite — definisce due funzioni $u(x)$, $v(x)$.

Si possono calcolare le derivate successive di u e v per $x = 0$ e formare così le serie di Taylor:

$$v = - \{ Ax^3 + A_4 x^4 + \ldots + A_s x^s + \ldots \}$$
$$u = - \frac{1}{2} \{ Bx^2 + B_3 x^3 + \ldots B_s x^s + \ldots \},$$

dove si può anche porre, per simmetria di scrittura,

$$A = A_3, \quad B = B_2.$$

Nel caso della cuspide ordinaria

$$A = A_3 \neq 0.$$

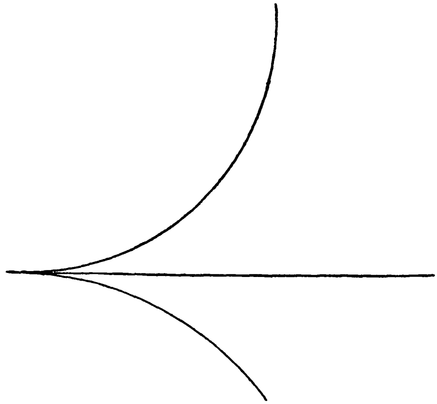
Pongasi per fissare le idee

$$A < 0;$$

allora nell' intorno di $x = 0$

$$y = - \frac{B}{2} x^2 + \ldots \pm \sqrt{- \{ Ax^3 + \ldots}$$

è reale per $x > 0$ ed il suo segno è dato da quello del radicale. Si vede dunque che la curva si compone di due archi tangenti per l'estremo e posti da parte opposta della tangente (v.[i] figura).



Invece sia

$$A = A_3 = 0 \quad A_4 \neq 0,$$

cioè si abbia un *tacnodo ordinario.*

Allora

$$y = -\frac{B}{2}x^2 + \dots \pm \sqrt{-\{A_4 x^4 + \dots\}}$$

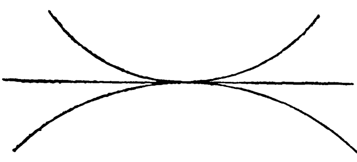
è reale nell'intorno di $x = 0$ soltanto per $A_4 < 0$, ed in questo caso è reale tanto per $x > 0$ che per $x < 0$, ed assume due valori di ugual segno o di segno opposto secondochè

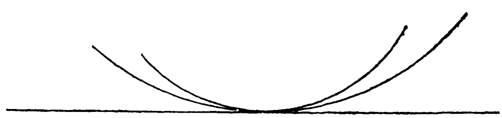
$$\frac{B}{2} > +\sqrt{-A_4} \quad \text{o} \quad \frac{B}{2} < +\sqrt{-A_4}$$

(discussione ulteriore nel caso dell'uguaglianza).

Dunque la curva $f = 0$ in un tacnodo ordinario consta:

1) di due archi tangenti e posti dalla stessa parte o da parte opposta della tangente comune (v.[i] figura): $A_4 < 0$;



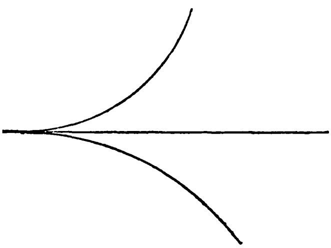


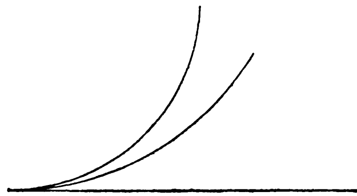
2) oppure di due punti isolati infinitamente vicini ($A_4 > 0$).

Se anche $A_4 = 0$, si deve considerare il primo coefficiente

$$A_5 \neq 0.$$

Se s *è dispari* la curva presenta in sostanza la *forma della cuspide* ordinaria cioè consta di due archi tangenti in un estremo, ma questi archi possono giacere da parti opposte o anche dalla stessa parte della tangente suddetta; e preci-





samente se s'indica con B_r il primo coefficiente B che non si annulla, avverrà il primo caso quando

$$\frac{s}{2} < r,$$

ed il secondo quando

$$\frac{s}{2} > r.$$

Se s *è pari* la curva presenta sostanzialmente la *forma del tacnodo* e precisamente

1) del *tacnodo iperbolico* formato da due archi tangenti se

$$A_s < 0;$$

questi archi si traverseranno o no nel punto di contatto

$$\text{secondochè } \frac{s}{2} \text{ è dispari o pari.}$$

2) invece $f=0$ ha in 0 un punto isolato (riunione di $\frac{s}{2}$ punti isolati infinitamente vicini) se

$$A_s > 0.$$

Merita infine di essere rilevato il *caso particolare* in cui f è un quadrato perfetto, cioè *v è identicamente nullo*:

$$A_s = 0$$

per qualunque s.

In tale ipotesi la curva si riduce ad una *linea passante semplicemente per* 0, *contata due volte.*

§ 21. **Massimi e minimi delle funzioni implicite.** — Ora — essendo f una funzione analitica — proponiamoci di determinare i massimi e minimi della funzione implicita

$$y(x),$$

definita da

$$f(x, y) = 0.$$

Questi potranno cadere: o nei punti dove

$$y'(x) = \frac{\partial f}{\partial x} = 0,$$

e allora si potrà applicare l'analisi del § 19;

oppure nei punti doppi o multipli della curva, per cui

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0.$$

Per esaminare il nuovo caso che qui si presenta, pongasi che si abbia in 0 un punto doppio (non multiplo) per la curva, cioè che non si annullino tutte le derivate seconde di f. Dovremo distinguere i seguenti casi (cfr. paragrafo precedente).

1) Caso iperbolico:

$$D(x, y) > 0.$$

Il punto 0 è un nodo.

Allora evidentemente i due rami della curva traversano da una parte all'altra della paralleta per 0 all'asse x: non c'è massimo nè minimo.

2) Caso ellittico:

$$D(x, y) < 0.$$

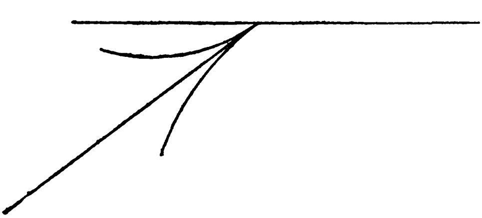
Si ha in 0 un punto isolato di $f = 0$, che può considerarsi a piacere come massimo o minimo relativo.

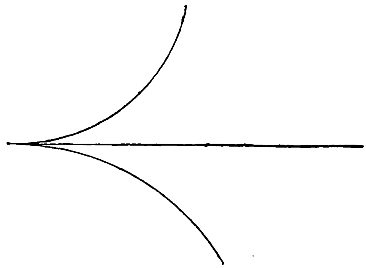
3) Caso parabolico:

$$D(x, y) = 0.$$

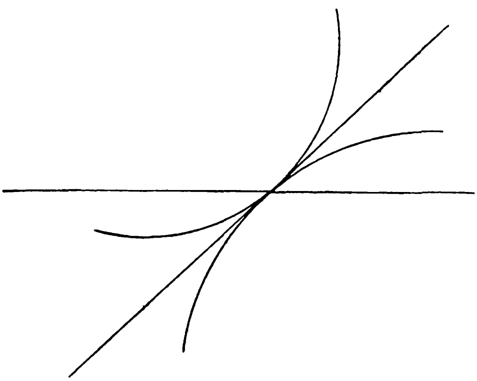
Il punto 0 è una cuspide. Si dà origine a svariatissimi casi. Limitiamoci ai due seguenti:

Se 0 è una cuspide ordinaria (§ 20) si avrà in 0 un mas-



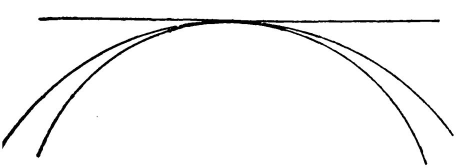


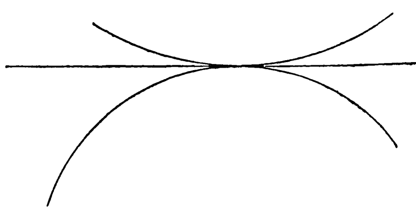
simo o un minimo, semprechè la tangente cuspidale non sia parallela all'asse x; se questa è parallela non c'è massimo nè minimo.



Se 0 è tacnodo iperbolico non si ha certo massimo nè minimo, quando la tangente tacnodale non sia parallela all'asse x.

Se è parallela possono presentarsi i due casi, p. es. quando la curva si approssimi a due archi di cerchio tangenti inter-





namente oppure esternamente.

§ 22. **Metodo della funzione inversa.** — Il problema di determinare i massimi e minimi della funzione implicita definita dall'equazione algebrica $f(x, y) = 0$, può essere trattato anche col *metodo della funzione inversa* che qui giova indicare per l'uso semplice che può farsene nei casi più elementari (cfr. Art. XXV).

Facciamo le seguenti osservazioni fondamentali:

1) Ad ogni punto di massimo o di minimo (relativo) della funzione $y(x)$, che cada in un punto semplice della curva $f(x, y) = 0$, appartiene una tangente parallela all'asse delle x e perpendicolare all'asse delle y.

2) Se codesta tangente viene spostata parallelamente, si ha — da una delle due parti — una perpendicolare all'asse delle y che incontra la curva in due punti, i quali tendono come limite allo stesso punto di contatto.

Si deduce quindi che i punti di massimo e di minimo della funzione implicita $y(x)$, sono punti per cui due rami della funzione inversa $x(y)$ vengono a coincidere. La condizione necessaria perchè ciò avvenga si esprime scrivendo che l'equazione

$$f(x, y) = u_0 + u_1 x + u_2 x^2 + \dots = 0$$
$$(u_i = u_i(y))$$

possiede una radice doppia, cioè annullando il resultante delle equazioni

$$f(x, y) = 0, \ \frac{\partial f}{\partial x} = 0,$$

che è il *discriminante* di

$$u_0 + u_1 x + u_2 x^2 + \dots = 0.$$

Questo metodo assume particolarmente carattere elementare quando il discriminante suddetto, D, è di 2° grado in y, come nei casi considerati nel citato Art. XXV. L'espressione analitica di x contiene allora, razionalmente, la $\sqrt{D}$, e la separazione dei valori per cui $D \geq 0$ e $D < 0$, conduce subito ai massimi e minimi.

§ 23. **Massimi e minimi relativi delle funzioni di più variabili: annullamento delle derivate.** — Sia $z = f(x_1, x_2, \dots x_n)$ una funzione continua e derivabile in un campo continuo C a n dimensioni.

Un punto $(\bar{x}_1, \bar{x}_2, ... \bar{x}_n)$ dicesi punto di massimo o di minimo relativo di f, se si può determinare un numero positivo h tale che nei punti del campo per cui

$$\bar{x}_1 - h < x_1 < \bar{x}_1 + h, ... \bar{x}_n - h < x_n < \bar{x}_n + h,$$

cioè

$$|x_i - \bar{x}_i| < h,$$

si abbia

$$f(x_1, x_2, ... x_n) \leq f(\bar{x}_1, \bar{x}_2, ... \bar{x}_n)$$

o rispett.

$$f(x_1 .... x_n) \geq f(\bar{x}_1 ... \bar{x}_n).$$

Ricordiamo che un punto $(\bar{x}_1, \bar{x}_2, ... \bar{x}_n)$ dicesi *interno* al campo C se si può determinare un numero positivo h, assai piccolo, in modo che ogni punto $(x_1 ... x_n)$ per cui

$$|x_i - \bar{x}_i| < h,$$

appartenga a C.

I massimi e i minimi relativi di $f(x_1 .... x_n)$, *che cadono internamente al campo* C, *sono punti in cui*

$$\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = ... = \frac{\partial f}{\partial x_n} = 0.$$

Si determini un h positivo tale che tutti i punti $(x_1 .... x_n)$ per cui

$$\bar{x}_1 - h < x_1 < \bar{x}_1 + h, ... \quad \bar{x}_n - h < x_n < \bar{x}_n + h,$$

appartengano a C; e si osservi che un massimo o un minimo di f è risp. massimo o minimo della funzione di una variabile

$$f(x_1, \quad \bar{x}_2 .... \bar{x}_n)$$

entro l'intervallo

$$(\bar{x}_1 - h, \quad \bar{x}_1 + h),$$

dove si considerano $\bar{x}_2 .... \bar{x}_n$ come costanti. Si ha dunque la *condizione necessaria*

$$\frac{\partial f}{\partial x_1} = 0,$$

e le altre analoghe.

§. 24. **Analisi dei massimi e minimi delle funzioni di due variabili.** — Nel paragrafo precedente si ha l'estensione immediata, da una a più variabili, del teorema che assegna le condizioni necessarie per l'esistenza di massimi o minimi d'una funzione.

Ma quando si tratta di riconoscere la sufficienza di codeste condizioni, ottenute annullando le derivate, s'incontra una difficoltà essenzialmente nuova.

Per rendersene conto limitiamoci per semplicità alle funzioni $z = f(x, y)$ di due variabili.

Si abbia dunque una funzione $z = f(x, y)$ definita almeno in un rettangolo che contiene il punto $(\bar{x}, \bar{y})$ nel suo interno, ed in questo punto continua e derivabile successivamente; e sia

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0.$$

Distinguiamo tre casi.

I) *Caso ellittico*: *lo hessiano*:

$$D(x, y) = \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 - \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} < 0.$$

Si ha:

$$f(\bar{x} + h, \bar{y} + k) = f(\bar{x}, \bar{y}) + \frac{1}{1 \cdot 2 \cdot}\left\{\frac{\partial^2 f}{\partial x^2} h^2 + 2\frac{\partial^2 f}{\partial x \partial y} hk + \frac{\partial^2 f}{\partial y^2} k^2\right\} + \varepsilon,$$

dove per

$$h = at, \quad k = bt \qquad (a \neq 0, \ b \neq 0,$$

ε diventa infinitesimo con t, d'ordine superiore a t^2.

Per h, k compresi in un rettangolo sufficientemente piccolo, cioè per un intorno conveniente del punto $(\bar{x}, \bar{y})$, il segno della differenza

$$f(\bar{x} + h, \ \bar{y} + k) - f(\bar{x}, \bar{y})$$

è quello della forma

$$\frac{\partial^2 f}{\partial x^2} h^2 + 2\frac{\partial^2 f}{\partial x \partial y} hk + \frac{\partial^2 f}{\partial y^2} k^2,$$

che, essendo $D < 0$, non si annulla mai nell'intorno suddetto e però è una *forma definita*, che conserva sempre il medesimo segno.

Per determinare questo segno, si ponga

$$a_{11} = \frac{\partial^2 f}{\partial x^2}, \quad a_{22} = \frac{\partial^2 f}{\partial y^2}, \quad a_{12} = \frac{\partial^2 f}{\partial x \partial y},$$

e si osservi che, essendo

$$a_{12}^2 - a_{11}a_{22} < 0,$$

a_{11} e a_{22} sono diversi da 0 e hanno lo stessose gno. Se si fa $h = k$ viene

$$(a_{12} + a_{22} + 2a_{12})h^2,$$

e il segno di cui si tratta è quello di

$$a_{11} + a_{22} + 2a_{12}.$$

Ora se

$$a_{11} > 0 \qquad (b > 0)$$

si può provare che

$$a_{11} + a_{22} + 2a_{12} > 0,$$

provando che

$$(a_{11} + a_{22})^2 > 4a_{12}^2,$$

ossia

$$a_{11}^2 + a_{22}^2 + 2a_{11}a_{22} > 4a_{12}^2;$$

ciò risulta evidente perchè

$$a_{11}^2 + a_{22}^2 > 2a_{11}a_{22},$$

e per ipotesi

$$a_{11}a_{22} > a_{12}^2.$$

Dunque il *punto* (x, y) è un *minimo* o un *massimo* secondochè

$$\frac{\partial^2 f}{\partial x^2} > 0 \quad \text{o} \quad \frac{\partial^2 f}{\partial x^2} < 0.$$

Il significato geometrico dell'ipotesi $D < 0$, risulta subito dal § 20.

La superficie

$$z = f(x, y)$$

ha come piano tangente nel punto $(\bar{x}, \bar{y})$ il piano

$$z = k, \quad \text{ove} \quad k = f(\bar{x}, \bar{y}),$$

e giace tutta da una parte di codesto piano nell'intorno del punto del contatto.

Ciò risulta dal fatto che la curva

$$f(x, y) - k = 0,$$

sezione della superficie col piano $z = k$, possiede in $(\bar{x}, \bar{y})$ un punto isolato, cioè che $(\bar{x}, \bar{y})$ è un *punto ellittico* della superficie, punto in cui la curvatura totale è positiva.

Esempii. La funzione

$$z = x^2 + y^2$$

rappresenta un paraboloide elittico di rotazione. Nel punto (0, 0) si ha

$$\frac{\partial z}{\partial x} = 2x = 0$$

$$\frac{\partial z}{\partial y} = 2y = 0$$

$$\frac{\partial^2 z}{\partial x^2} = 2, \quad \frac{\partial^2 z}{\partial y^2} = 2, \quad \frac{\partial^2 z}{\partial x \partial y} = 0,$$

$$D = -4 < 0.$$

C'è dunque un minimo, come è chiaro senz'altro perchè

$$x^2 + y^2 \geq 0,$$

diventa 0 per

$$x = y = 0.$$

Il piano tangente

$$z = 0$$

ha comune col paraboloide soltanto il punto (0, 0).

La funzione

$$z = xyu \quad \text{per} \quad x + y + u = p$$

si riduce alla funzione di due variabili

$$\bar{z} = pxy - (x + y)\, xy.$$

Scriviamo

$$\frac{\partial \bar{z}}{\partial x} = py - 2xy - y^2 = y(p - 2x - y) = 0$$

$$\frac{\partial \bar{z}}{\partial y} = px - 2xy - x^2 = x(p - 2y - x) = 0.$$

Si trovano le soluzioni

$$x = 0 \quad \text{o} \quad y = 0,$$

cioè

$$z = 0,$$

oppure

$$x = y, \quad p = 2x = y,$$

$$x = y = u = \frac{p}{3}.$$

Consideriamo questi ultimi valori che annullano $\dfrac{\partial \bar{z}}{\partial x}, \dfrac{\partial \bar{z}}{\partial y}$; è facile vedere che essi rispondono ad un massimo di $z = \bar{z}$ (cfr. § 13).

Infatti si ha

$$\frac{\partial^2 \bar{z}}{\partial x^2} = -2y = -\frac{2p}{3}, \quad \frac{\partial^2 z}{\partial y^2} = -2x = -\frac{2p}{3},$$

$$\frac{\partial^2 \bar{z}}{\partial x \partial y} = p\, 2(x + y) = -\frac{p}{3},$$

$$D = \frac{p^2}{9} - \frac{4p^2}{9} = -\frac{p^2}{3} < 0.$$

II) *Caso iperbolico*;

$$D(x, y) = \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 - \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} = a_{12}{}^2 - a_{11}a_{22} > 0.$$

In questo caso la forma

$$\frac{\partial^2 f}{\partial x^2} h^2 + 2\frac{\partial^2 f}{\partial x \partial y} hk + \frac{\partial^2 f}{\partial y_z} k^2,$$

relativa al punto $(\bar{x}, \bar{y})$, si annulla per due valori:

$$\frac{h}{k} = \alpha, \frac{h}{k} = \beta,$$

e — attraverso questi valori — cambia di segno. Pertanto la differenza

$$f(\bar{x} + h, \quad \bar{y} + k) - f(\bar{x}, \bar{y})$$

assume valori positivi e negativi in un intorno comunque piccolo di $(\bar{x}, \bar{y})$ e perciò *non vi è massimo nè minimo* della funzione f.

Il significato geometrico dell'ipotesi $D > 0$, è che la superficie $z = f(x, y)$ è attraversata dal piano $z = k$, tangente ad essa in $(\bar{x}, \bar{y})$. Infatti la curva

$$f(x,y) - k = 0,$$

sezione di codesto piano, possiede in $(\bar{x}, \bar{y})$ un nodo, e quindi $(\bar{x}, \bar{y})$ è un *punto iperbolico* della superficie, in cui la curvatura totale è negativa, cioè i due raggi di curvatura principale hanno segno opposto.

Esempii. La funzione

$$z = x^2 - y^2$$

rappresenta un paraboloide iperbolico. Nel punto $x = y = 0$ si ha

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y} = 0, \quad D = +4 > 0,$$

e non vi è massimo nè minimo.

Il piano tangente $z = 0$ sega la superficie secondo le due rette

$$(x + y)\ (x - y) = 0.$$

Anche la funzione

$$z = xy$$

offre un esempio analogo per

$$x = y = 0.$$

La funzione

$$z = pxy - (x + y)xy$$
$$(= xyu \quad \text{per} \quad x + y + u = p)$$

non ha massimo nè minimo nel punto

$$x = y = 0,$$

dove pure si annullano le

$$\frac{\partial z}{\partial x} = py - 2xy - y^2$$

$$\frac{\partial z}{\partial y} = px - 2xy - x^2.$$

Invero (come già vedemmo di sopra)

$$\frac{\partial^2 z}{\partial x^2} = -2y, \quad \frac{\partial^2 z}{\partial y^2} = -2x,$$

$$\frac{\partial^2 z}{\partial x \partial y} = p - 2(x + y);$$

quindi, per $x = y = 0$,

$$D = p^2 > 0.$$

III) *Caso parabolico*:

$$D(x, y) = \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 - \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} = 0.$$

In questo caso la forma

$$\frac{\partial^2 f}{\partial x^2} h + 2 \frac{\partial^2 f}{\partial x \partial y} hk + \frac{\partial^2 f}{\partial y^2} k^2$$

relativa al punto $(\bar{x}, \bar{y})$ è *semidefinita*; essa si riduce ad un quadrato perfetto

$$(\alpha h + \beta k)^2,$$

e però si mantiene costantemente ≥ 0, ma si annulla per

$$k = -\frac{\alpha}{\beta} h.$$

Il segno della differenza

$$f(\bar{x} + h, \quad \bar{y} + k) - f(\bar{x}, \bar{y})$$

non dipende ora solo dall'anzidetta forma quadratica, ma anche dai termini di grado più alto nello sviluppo in serie di potenze di $f(x, y)$. La *conclusione circa l'esistenza* d'un *massimo* o d'un *minimo* di f in $\bar{x}, \bar{y}$ resta *dubbia*.

La considerazione geometrica illumina la natura della difficoltà che qui si presenta.

Posto

$$f(\bar{x}, \bar{y}) = k,$$

si consideri il piano $z = k$ tangente alla superficie in $(\bar{x}, \bar{y})$;

la sezione di questo piano è la curva

$$f(x, y) - k = 0,$$

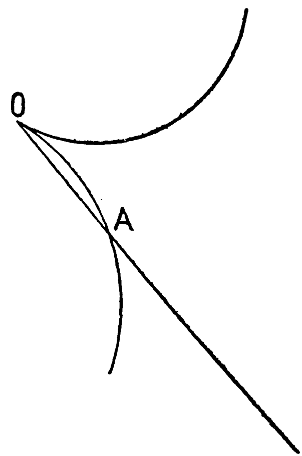
che possiede in $(\bar{x}, \bar{y})$ una cuspide (o un punto di singolarità superiore).

Tre casi occorre distinguere:

1) Se la curva

$$f(x, y) - k = 0$$

possiede un ramo reale nell' intorno di $(\bar{x}, \bar{y})$, che non sia curva di contatto del piano $z = k$ colla superficie (cioè non si riduca ad una curva contata due volte), allora non vi è massimo nè minimo in $O = (\bar{x}, \bar{y})$.

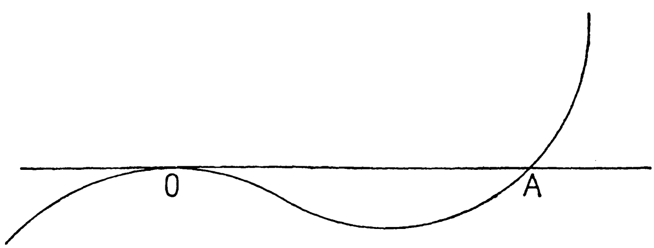


Infatti si prenda sul detto ramo un punto A, la sezione piana verticale per OA tocca OA in O e la traversa in A; siccome A può diventare vicino quanto si vuole ad O, si vede che vi sono, prossimi ad O, punti della superficie dalle due parti del piano tangente.

2) Se la curva

$$f(x, y) - k = 0,$$

non possiede rami reali nell'intorno di $O = (\bar{x}, \bar{y})$, cioè ha ivi



un punto isolato, la superficie stà tutta da una parte del piano tangente nell' intorno del punto: c'è massimo o minimo.

3) Se la curva

$$f(x, y) - k = 0,$$

si riduce ad una curva doppia (non multipla), ossia è una curva parabolica di contatto (semplice) del piano tangente, la superficie resta da una parte del piano tangente, e si ha in $(\bar{x}, \bar{y})$ un massimo o un minimo *improprio*, che fa parte di una linea di massimi o di minimi.

Nel caso d'una curva lungo cui si abbia un contatto superiore al primo, decide la parità di questo ordine di contatto.

Nel § 20 noi abbiamo esaminato il comportamento della curva $f(x, y) = 0$ (o $f(x, y) - k = 0$) in un suo punto doppio (escluso il caso di molteplicità superiore). In base a questo esame possiamo affermare che *per* $D = 0$ *non si ha in generale massimo nè minimo*, la sezione del piano tangente presentando una cuspide ordinaria; ma quando il punto di contatto di questa sezione diventa un tacnodo si avrà *massimo o minimo se* trattasi d'un *tacnodo ellittico* (punto isolato), e *non* si avrà nel caso del *tacnodo iperbolico*. Una particolarità ulteriore riporta il dubbio ecc.

I resultati principali si possono desumere dal seguente quadro che esprime le circostanze relative alla curva sezione $f(x, y) - k = 0$.

Si suppone che qualcuna delle derivate seconde di f *sia diversa da* 0.

$D < 0$: Caso ellittico (punto isolato)

$D > 0$: Caso iperbolico (nodo)

In particolare:

$$D = 0:\ \text{Caso parabolico}\left\{\begin{array}{c}\text{cuspide}\\ \text{o in}\\ \text{particolare}\end{array}\right.\left\{\begin{array}{l}\text{tacnodo}\left\{\begin{array}{l}\text{ellittico}\\ \text{iperbolico}\end{array}\right.\\ \begin{array}{l}\text{o in particolare}\\ \text{parabolico}\end{array}\left\{\begin{array}{l}\text{cuspide}\\ \text{o in particolare}\end{array}\right.\left\{\ \dots\right.\end{array}\right.$$

La funzione $f(x, y)$ avrà un massimo o un minimo solo quando ci si trovi nel caso ellittico, o parabolico-ellittico, o parabolico-parabolico-ellittico ecc., oppure nel caso di una *infinita parabolicità* che corrisponde ad una curva di contatto e ad un massimo o minimo improprio (cfr. il precedente caso 3).

Ad illustrazione e complemento delle cose dette facciamo seguire alcuni:

Esempii. Le seguenti funzioni hanno le derivate $\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y} = 0$ nel punto $O = (0, 0)$ e danno luogo a varie circostanze:

1) $$z = y^2 + x^3$$

corrisponde al caso parabolico generale (la sezione col piano $z = 0$ ha una cuspide in O); non vi è massimo nè minimo.

2) $$z = y^2 + x^4,$$

ha un minimo per $x = y = 0$; la curva $y^2 + x^4 = 0$ ha in O un tacnodo ellittico (punto isolato).

3) $$z = y^2 + x^2y^2,$$
$$z = y^2 + x^4 - y^4,$$

corrispondono al medesimo caso precedente ed hanno sempre un minimo (relativo) per $x = y = 0$.

4) $$z = y^2 - x^4,$$
$$z = y^2 + y^4 - x^4,$$

non hanno in O nè massimo nè minimo; la sezione col piano tangente ha un tacnodo iperbolico.

5) $$z = y^2 + x^5$$

corrisponde al caso parabolico-parabolico generale; la sezione del piano tangente $y^2 + x^5 = 0$ presenta il caso particolare del tacnodo in cui la curva riassume forma di cuspide; non c'è massimo nè minimo.

6) $$z = (x + y)^2$$

presenta un minimo in O; la relativa superficie tocca il piano tangente secondo la curva parabolica

$$x + y = 0.$$

7) $$z = x^3 + y^3,$$
$$z = xy^2,$$

non hanno massimo nè minimo per $x = y = 0$.

Queste funzioni hanno *nulle* in O *tutte le derivate seconde* ma *non le derivate terze*; quindi le relative superficie sono segate dal piano tangente $z = 0$ secondo curve che hanno in O un punto triplo, e però posseggono certo qualche ramo reale.

8) $$z = x^4 + y^4$$

ha un minimo per $x = y = 0$.

Sono nulle in O, oltre le derivate prime e seconde, anche le terze. La superficie è segata dal piano tangente secondo una curva con un *punto quadruplo isolato.*

9) $$z = x^4 - y^4,$$

non ha massimo nè minimo per $x = y = 0$.

Tutte le circostanze sono analoghe a quelle dell'es. prec.te, ma la sezione della superficie col piano tangente in O, cioè la curva

$$x^4 - y^3 = 0,$$

passa con *due rami reali per il punto quadruplo* 0.

§ 25. **Cenno sulle funzioni di tre variabili.** Si abbia

$$u = f(x, y, z),$$

funzione continua e derivabile (successivamente), in un campo continuo a tre dimensioni.

Sia p. es.

$$x = y = z = 0$$

un punto, interno al campo, per cui

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = \frac{\partial f}{\partial z} = 0.$$

Si avrà

$$f(x, y, z) - f(0, 0, 0) = a_{11}x^2 + a_{22}y^2 + a_{33}z^2 +$$
$$+ 2a_{12}xy + 2a_{13}xz + 2a_{23}yz + \varepsilon,$$

ed il segno della differenza che è nel primo membro, quando x, y, z sono piccoli, dipenderà dal segno della forma quadratica $a_{11}x^2 + \ldots$.

Supponendo il descriminante non nullo, questa si può ridurre alla forma canonica

$$Dx^2 + D'y^2 + D''z^2,$$

ed allora si vede che è *definita* quando

$$D, \ D', \ D''$$

hanno il medesimo segno; onde in tal caso f ha in $O = (0, 0, 0)$ un massimo o un minimo.

Se due delle D hanno segno opposto, p. es.

$$D > 0, \quad D' < 0,$$

si deduce che f non ha massimo nè minimo.

Posto

$$k = f(0, 0, 0),$$

la superficie

$$f(x, y, z) - k = 0$$

ha generalmente nel punto $O = (0, 0, 0)$ un punto doppio conico; l'annullarsi della forma quadratica rappresenta il cono delle rette osculatrici in O. Questo deve essere immaginario, ossia O deve essere un *punto conico isolato*, perche vi sia massimo o minimo.

Il caso del discriminante nullo, che sopra abbiamo escluso è un *un caso dubbio*. Il cono osculatore in O si riduce ad una coppia di piani osculatori o ad un piano contato due volte; oppure diventa indeterminato, ed O ha una molteplicità > 2 per la superficie (annullamento di tutte le derivate seconde. Che cosa si può dire nel caso in cui O sia doppio per la superficie?

Osserviamo che l'esistenza d'una retta per O

$$(y = rx, \ z = sx, \ u = k)$$

avente un contatto tripunto, esclude il massimo o il minimo, giacchè nel piano delle coordinate x, u, la curva $u = f(x, rx, sx)$ traversa la retta medesima: $u = k$. Dunque se il cono osculatore in O è costituito di due piani, anche se questi sieno immaginari, non vi è *in generale* nè massimo nè minimo, a meno che la retta intersezione dei due piani osculatori non abbia colla superficie un contatto quadripunto. Anche quando il cono osculatore in O si riduca a un piano contato due volte non vi è in generale massimo o minimo, le rette per O in questo piano avendo colla superficie un contatto tripunto: ma in particolare può darsi che le rette osculatrici per O abbiano un contatto quadripunto colla superficie; allora O è in generale un tacnodo; si ha massimo o minimo se è tacnodo ellittico (punto isolato), non si ha se è tacnodo iperbolico ecc.

Infine l'annullamento delle derivate seconde esclude il massimo e il minimo se c'è qualche derivata terza diversa da zero, ossia se 0 è *triplo* per la superficie $f(x, y, z) - k = 0$; se invece O diventa *quadruplo*, occorre esaminare se è punto isolato o no per la superficie ecc.

§ 26. **Massimi e minimi assoluti.** — Indipendentemente dall'analisi più delicata dei §§ precedenti, la condizione necessaria del § 23, permette sotto larghe condizioni di determinare i massimi e minimi assoluti d'una funzione continua e derivabile $f(x_1, x_2, ... x_n)$, definita entro un *campo continuo finito*:

$$\varphi(x_1, x_2, ... x_n) \leq 0 ,$$

dove anche la funzione φ si supponga derivabile.

I punti in cui si raggiunge il massimo o il minimo di f sono punti *interni* al campo per cui

$$\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = ... = \frac{\partial f}{\partial x_n} = 0,$$

o punti *del contorno* per cui

$$\varphi(x_1, x_2, ... x_n) = 0.$$

Cerchiamo i punti del contorno dove f può diventare massima o minima.

Osserviamo anzitutto che il campo — per ipotesi finito e continuo — determinato dall'equazione

$$\varphi(x_1, x_2, ... x_n) = 0,$$

può essere considerato come un *campo ad* n — 1 *dimensioni senza contorno*, costituito dalla riunione di due o più campi contornati per cui

$$x_n = \theta(x_1 .... x_{n-1})$$

è funzione implicita di $x_1 .... x_{n-1}$.

Se p. es. $(n=3)$:

$$\varphi(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - r^2,$$

il campo $\varphi = 0$ è una superficie sferica, che consta della riunione di due superficie (x_1, x_2) limitate da $x_1^2 + x_2^2 \leq r^2$; per la prima di queste superficie si ha

$$x_3 = +\sqrt{r^2 - x_1^2 - x_2^2},$$

e per la seconda

$$x_3 = -\sqrt{r^2 - x_1^2 - x_2^2},$$

Ciò posto, per avere i punti dove $\varphi = 0$ per cui f diventa massima o minima, bisogna annullare le derivate di

$$f(x_1, x_2, \ldots. x_{n-1})$$

dove

$$x_n = \theta(x_1 \ldots. x_{n-1}).$$

Differenziando l'identità:

$$\varphi(x_1, x_2, \ldots. x_{n-1}, \theta(x_1, \ldots. x_{n-1})) = 0$$

si avrà identicamente

$$\frac{\partial\varphi}{\partial x_1} + \frac{\partial\varphi}{\partial\theta_1}\frac{\partial\theta}{\partial x_1} = 0$$
$$\ldots\ldots\ldots\ldots$$

cioè

$$\frac{\partial x_n}{\partial x_1} = \frac{-\dfrac{\partial\varphi}{\partial x_1}}{\dfrac{\partial\varphi}{\partial x_n}},$$
$$\ldots\ldots\ldots;$$

quindi le equazioni da cui dipende il nostro problema sono

$$\frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial x_1} = \frac{\partial f}{\partial x_1}\frac{\partial\varphi}{\partial x_n} - \frac{\partial f}{\partial x_n}\frac{\partial\varphi}{\partial x_1} = 0,$$

Consideriamo i due sistemi di equazioni:

$$(1) \qquad \frac{\partial f}{\partial x_i} = 0 \qquad (i = 1, 2 \ldots. n)$$

$$(2) \qquad \begin{cases} \dfrac{\partial f}{\partial x_i}\dfrac{\partial\varphi}{\partial x_n} - \dfrac{\partial f}{\partial x_n}\dfrac{\partial\varphi}{\partial x_i} = 0 \qquad (i = 1, 2 \ldots. n-1), \\ \varphi(x_1, x_2, \ldots. x_n) = 0, \end{cases}$$

e poniamo che ciascuno di essi abbia un *numero finito di soluzioni*, oppure infinite soluzioni a cui corrisponda un numero finito di valori per z. Basterà confrontare i valori di f in corrispondenza a tali soluzioni; *il maggiore* di questi *sarà il massimo, il minore sarà il minimo* di f.

§ 27. **Applicazioni elementari.** — Applichiamo il procedimento indicato, al seguente esempio (cfr. § 13). Trattisi di determinare il massimo assoluto di

$$z = x_1 x_2 \dots x_n,$$

per

$$0 \leq x_i, \quad \sum_1^n x_i \leq p. \qquad (p > 0).$$

Si dovranno considerare i punti, interni al campo di variazione delle x_i, per cui

$$(1) \qquad \frac{\partial z}{\partial x_i} = x_1 \dots x_{i-1} x_{i+1} \dots x_n = 0$$

$$(i = 1, 2 \dots n),$$

e i punti del contorno per cui

$$(2) \qquad \begin{cases} \dfrac{\partial z}{\partial x_i} - \dfrac{\partial z}{\partial x_n} = 0 & (i = 1, 2, \dots n) \\ \Sigma x_i = p. \end{cases}$$

Ora le equazioni (1) danno almeno una $x_i = 0$ e quindi $z = 0$ Le (2), nell'ipotesi $x_i \neq 0$, danno

$$x_1 = x_2 = \dots = x_n = \frac{p}{n}$$

$$z = \frac{p^n}{n^n} > 0.$$

In alcuni casi la ricerca del massimo o minimo assoluto può essere semplificata, quando p. es. si tratti di una funzione definita in un campo infinito che nel punto all'infinito tenda a divenire infinita positiva o negativa (cfr. §§ 9, 10); nel qual caso si ha *a priori* la certezza che il minimo (o il massimo) cercato cade necessariamente in un punto interno al campo.

Valga ad esempio il seguente:

Problema del baricentro. Si abbiano nello spazio n punti materiali $P_1, P_2, ... P_n$ di pesi rispettivi $m_1, m_2, ... m_n$; si tratta di determinare un punto X per cui la somma

$$f = m_1 \overline{XP_1}^2 + m_2 \overline{XP_2}^2 + ... + m_n \overline{XP_n}^2,$$

riesca minima.

Designando con x, y, z le coordinate di X e con x_i, y_i, z_i quelle di P_t, si avrà

$$f(x, y, z) = \Sigma m_i \{ (x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2 \},$$

e perciò

$$\frac{\partial f}{\partial x} = 2\Sigma m_i (x - x_i)$$
$$\frac{\partial f}{\partial y} = 2\Sigma m_i (y - y_i)$$
$$\frac{\partial f}{\partial z} = 2\Sigma m_i (z - z_i).$$

Il punto in cui s'annullano queste derivate:

$$\begin{cases} x = \dfrac{\Sigma m_i x_i}{\Sigma m_i} \\[2ex] y = \dfrac{\Sigma m_i y_i}{\Sigma m_i} \\[2ex] z = \dfrac{\Sigma m_i z_i}{\Sigma m_i} \end{cases}$$

è unico; esso prende il nome di *baricentro degli* n *punti* P_i *di peso* rispettivo m_i. Questo baricentro corrispondere al minimo valore della funzione continua f, poichè questa prende un valore grande quanto si vuole nei punti esterni ad una sfera di raggio assai grande col centro in uno dei punti P_i e perciò ha il suo minimo nell'interno della sfera stessa.

La proprietà di minimo del baricentro X si dimostra anche direttamente, osservando che per ogni altro punto P il valore di f è

$$f(P) = f(X) + \overline{PX^2};$$

il luogo dei punti per cui f ha un valore costante è una sfera col centro nel baricentro (cfr. APOLLONIO t. 2, p. 116).

NOTA. Nell'applicare il procedimento sopra indicato devesi sempre avere riguardo all'ipotesi che le derivate della funzione f, da rendere massima o minima, si conservino determinate e finite. Altrimenti occorre anche prendere in considerazione i punti singolari per le derivate. Così accade nel problema classico che ha per oggetto di determinare *il punto per cui riesce minima la somma delle distanze da tre punti dati.*

Sieno

$$A_1 \equiv (x_1, y_1), \quad A_2 \equiv (x_2, y_2), \quad A_3 \equiv (x_3, y_3)$$

tre punti, vertici d'un triangolo. La distanza d'un punto (x, y) da (x_i, y_i) $(i = 1, 2, 3)$ si designi con

$$z_i = \sqrt{(x_i - x_i) + (y - y_i)^2}.$$

Si tratta di render minima la funzione

$$z = z_1 + z_2 + z_3.$$

Avremo

$$\frac{\partial z}{\partial x} = \frac{x - x_1}{z_1} + \frac{x - x_2}{z_2} + \frac{x - x_3}{z_3}$$

$$\frac{\partial z}{\partial y} = \frac{y - y_1}{z_1} + \frac{y - y_2}{z_2} + \frac{y - y_3}{z_3}.$$

Ora queste derivate diventano indeterminate per

$$z_1 = 0, \quad z_2 = 0, \quad z_3 = 0,$$

cioè nei tre punti dati A_1, A_2, A_3. Il minimo assoluto di z può cadere sia in un punto per cui $\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y} = 0$, sia in uno dei tre punti suddetti.

L'analisi geometrica del problema mostra assai facilmente che il punto di minimo si trova di fatto nell'interno del triangolo $A_1A_2A_3$ (in un punto per cui $\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y} = 0$) se il triangolo ha i tre angoli $< 120°$; in questo caso il *punto di minimo* è il *punto da cui si vedono i tre lati sotto angoli di 120°*, e così resta definito come intersezione di tre cerchi

Ma *se il triangolo* $A_1A_2A_3$ *possiede un angolo* $\geq 120°$, il *punto di minimo cade nel vertice corrispondente.*

Questo problema, proposto da FERMAT, fu risoluto (pel caso $\widehat{A_i} < 120°$) da TORRICELLI e da CAVALIERI ([1]). Esso dette luogo ad una corrispondenza fra TORRICELLI, FERMAT e ROBERVAL, su cui riferisce il VIVIANI ([2]). In tempi più recenti fu ripreso da STEINER ([3]).

§ 28. **Massimi e minimi vincolati.** — Nel § 26 si è presentata la ricerca dei massimi e minimi di una funzione $f(x_1, x_2, \dots x_n)$ pei valori di $x_1, x_2, \dots x_n$ soddisfacenti ad una equazione data

$$\varphi(x_1, x_2, \dots x_n) = 0.$$

Le equazioni a cui siamo pervenuti, corrispondono ugualmente alla condizione

$$\varphi(x_1, \dots x_n) = \text{cost.},$$

e sono della forma

$$\frac{\partial f}{\partial x_i}\frac{\partial \varphi}{\partial x_i} - \frac{\partial \varphi}{\partial x_i}\frac{\partial f}{\partial x_r} = 0;$$

esse sussistono nei punti di massimo e minimo per tutte le coppie di numeri i, r fra 1 e n.

Ora è notevole il fatto che a codeste equazioni si arriva ugualmente ove si proponga il problema di trovare i massimi e minimi della funzione

$$f(x_1, \dots x_n) + \lambda\varphi(x_1, \dots x_n),$$

ove λ designa una costante il cui valore risulterà poi determinato. Si avranno infatti le equazioni

$$\frac{\partial f}{\partial x_i} + \lambda\frac{\partial \varphi}{\partial x_i} = 0,$$

([1]) « Exercitationes geometricae sex », Bononiae, 1647, (p. 504).

([2]) VINCENTIUS VIVIANUS, « De maximis et minimis ». Firenze, 1659. Appendix, p. 149.

([3]) WERKE, Bd II, p. 6, 93, 729. Cfr. R. STURM « Maxima und minima in der elementaren Geometrie ». Teubner, Lipsia, 1910, (§ 8, p. 55).

le quali dovendo sussistere per tutti i valori di $i = 1, 2 .... n$, e definire lo stesso valore di λ, danno appunto le equazioni resultanti

$$\frac{\partial f}{\partial x_i}\frac{\partial \varphi}{\partial x_r} - \frac{\partial \varphi}{\partial x_i}\frac{\partial f}{\partial x_r} = 0.$$

Se invece di una sola condizione data $\varphi = \text{cost.}$ si hanno più relazioni

$$\varphi = \text{cost}, \quad \psi = \text{cost}, ... ,$$

vale ancora la stessa regola. *Le equazioni da cui dipende la ricerca dei massimi o minimi d'una funzione* $f(x_1 .... x_n)$ *pei valori delle* x_i *soggetti a più condizioni*

$$\varphi(x_1 .... x_n) = \text{cost}, \quad \psi(x_1 .... x_n) = \text{cost}, ... ,$$

si ottengono semplicemente scrivendo le condizioni per il massimo e il minimo della funzione

$$f + \lambda\varphi + \mu\psi + ... ,$$

dove λ, μ .... *designano delle costanti* e *le* $x_1, x_2 .... x_n$ *si assumono liberamente variabili.*

Eliminando λ, μ .... dalle equazioni così ottenute, si deducono le equazioni seguenti:

$$\begin{vmatrix} \frac{\partial f}{\partial x_i} & \frac{\partial \varphi}{\partial x_i} & \frac{\partial \psi}{\partial x_i} & .... \\ \frac{\partial f}{\partial x_r} & \frac{\partial \varphi}{\partial x_r} & \frac{\partial \psi}{\partial x_r} & .... \\ \frac{\partial f}{\partial x_s} & \frac{\partial \varphi}{\partial x_s} & \frac{\partial \psi}{\partial x_s} & .... \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix} = 0;$$

le quali contengono simmetricameute le funzioni f, φ, ψ .....

§ 29. **Legge di reciprocità.** — La simmetria delle equazioni dedotte nel precedente paragrafo, rispetto alle funzioni

f, φ, ψ, ..., risponde ad una legge di reciprocità che si può rendere manifesta con semplici considerazioni geometriche [1].

Sieno $f(x_1 \dots x_n)$, $\varphi(x_1 \dots x_n)$, $\psi(x_1 \dots x_n)$ ...., funzioni continue in dato campo.

Ad ogni massimo o minimo relativo della funzione* f *per* $\varphi = \text{cost}$, $\psi = \text{cost}$.... *risponde in generale un massimo o un minimo di* φ *per f $= \text{cost}$, $\psi = \text{cost}$.... ecc.

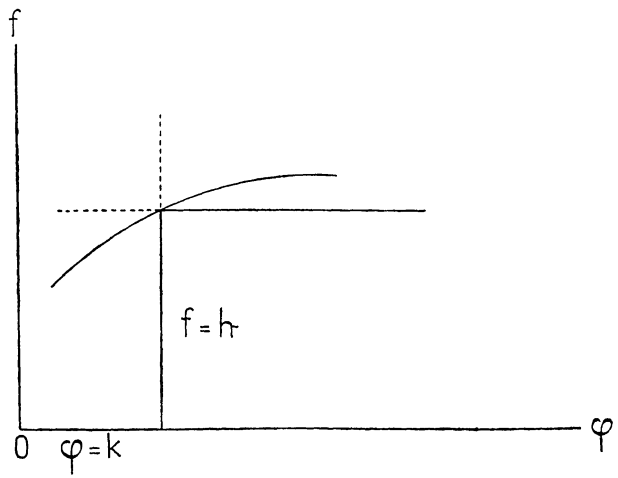
Più precisamente si considerino i valori di f, φ, corrispondenti alle $x_1 \dots x_n$ di un campo convenientemente limitato, come coordinate d' un punto (f, φ) nel piano. I punti (f, φ) copriranno una superficie che sarà, in generale, ***limitata superiormente*** (e così pure inferiormente) da una linea l:

$$f = f(\varphi).$$

Posto

$$\varphi = k$$

si avrà una retta parallela all' asse f; se questa incontra la linea anzidetta in un punto, $f = h$, c' è corrispondentemente

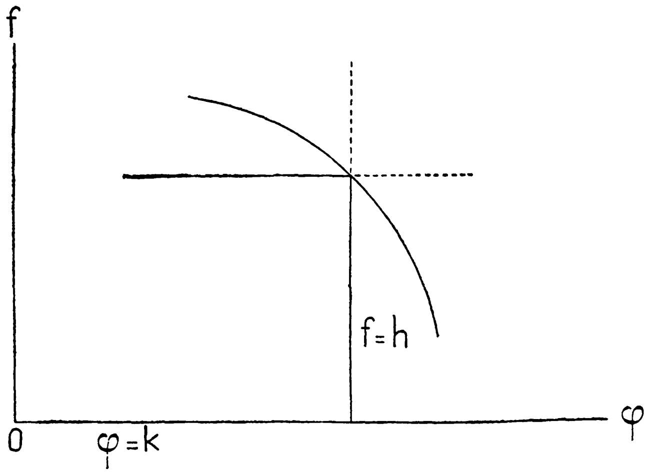


un massimo (relativo) di f *per* $\varphi = k$.

Ora debbonsi distinguere varii casi:

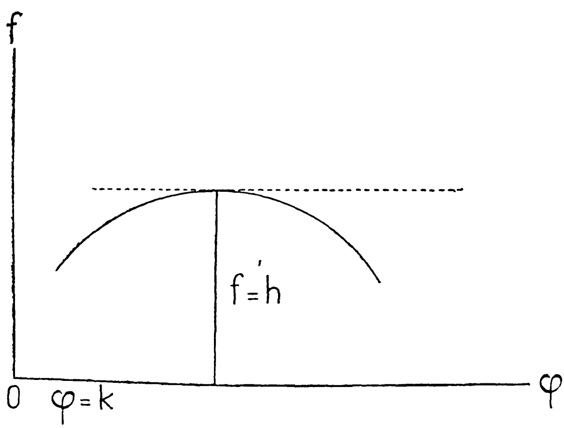
(1) Una trattazione simile trovasi in J. HADAMARD, « Leçons sur le calcul des variations ». Paris, 1910, p. 9.

1) La funzione $f(\varphi)$ (il massimo di f per un dato φ) è *crescente* nel punto $\varphi = h$; allora per $f = h$ si ha ivi un *mi-*



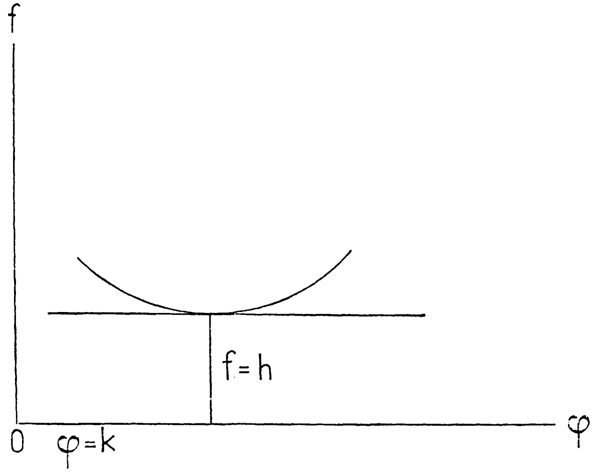
nimo di φ. Infatti la linea *l limita a sinistra* la regione di piano coperta dai punti $(f\varphi)$.

2) La $f(\varphi)$ è *decrescente* nel punto $\varphi = k$. Allora per $f = h$ si ha ivi un *massimo* di f.



3) Il punto $\varphi = k$ corrisponde ad un *massimo di* f(φ), cioè $f = h$ è un *massimo di* f *indipendente dalla condizione* $\varphi = k$. Per $f = h$, si ha, generalmente, *un solo valore di* φ, che è $\varphi = k$, e può considerarsi ad arbitrio come massimo o minimo.

4) Il punto $\varphi = k$ corrisponde ad un *minimo di* f(φ).



Allora per $f = k$ si hanno valori di $\varphi > k$ e $< k$; ***non c'è massimo nè minimo.***

5) Un sotto-caso che rientra come ***caso particolare*** in 4) è quello in cui la curva l riesce indipendente da φ, cioè si riduce ad un segmento di retta $f = h$; ed è chiaro che allora *nulla* si può dire a priori *circa il massimo o il minimo* di φ in un punto $\varphi = k$ su $f = h$.

L'analisi precedente si ripete ugualmente pel caso in cui $f = h$ sia *un minimo di* f per $\varphi = k$.

Appare così che la legge di reciprocità comporta *una sola eccezione essenziale*, cioè quando il massimo di f per $\varphi =$ cost è minimo per la curva dei massimi $f(\varphi)$, oppure il minimo di f per $\varphi =$ cost è massimo per la curva dei minimi.

Diamo un esempio relativo a questa eccezione.

Si assuma

$$\psi(x, y, z) = y^2 + z^2 - x^2 = r^2,$$
$$\varphi = x,$$
$$f = x.$$

La prima equazione rappresenta un iperboloide di rotazione; i punti più alti di z per $x =$ cost si trovano sopra un'iperbole $z^2 - x^2 = r^2$ nel piano $y = 0$; ciascuno di questi punti, a destra o a sinistra del piano $x = 0$, è un vertice dell'iperbole sezione d'un piano $x = \text{cost}$; e un punto che stia a destra del suddetto piano $(x > 0)$ corrisponde dunque ad un minimo di x per $z = \text{cost}$. *Fa eccezione il punto*

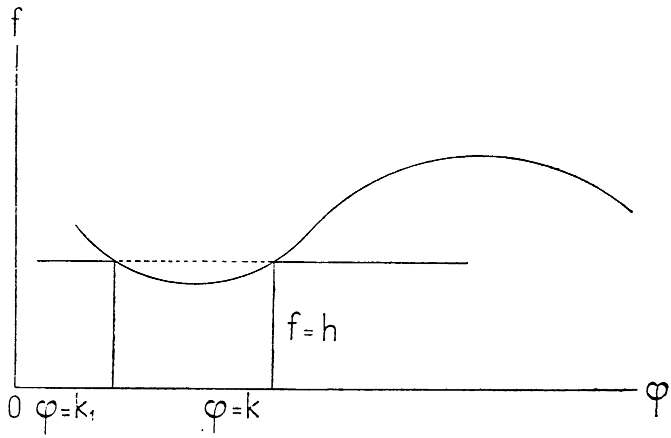
$$x = y = 0, \quad z = + r,$$

Infatti il piano $z = + r$, è tangente all'iperboloide: la sezione è costituita dalla dalla coppia di rette

$$(y + x)(y - x) = 0,$$

e *non c'è massimo nè minimo per* x.

Nota. Il ragionamento con cui si è giustificata la legge di reciprocità può tradursi — ove piaccia — sotto forma



algebrica. Esso non differisce sostanzialmente da quello che è svolto nell'Art. XXV, per i massimi e minimi assoluti. L'ipotesi — ivi considerata — di una corrispondenza *concorde* fra il valore dato $\varphi = \text{cost}$ e il massimo di f, significa appunto che la funzione $f(\varphi)$ è crescente (caso 1)); il caso *discorde* corrisponde ad $f(\varphi)$ decrescente (caso 2)). Ma perchè la *legge di reciprocità* possa stabilirsi in tal modo *per i massimi e minimi assoluti*, bisogna supporre *che* f(φ) *sia sempre crescente* o *sempre decrescente*, il che — in generale — non avverrà.

Se la curva dei massimi $f(\varphi)$ si presenta come nell'annessa figura, per $f = h$ si ha un minimo relativo $\varphi = k$, ma allargando il campo di variazione delle variabili $x_1 \ldots . x_n$ si trova ancora un massimo relativo $\varphi = k_1$, dove $k_1 < k$.

§ 30. **Principio di simmetria.** — Nelle ricerche sui massimi e minimi, specialmente nel campo geometrico, STEINER ha fatto valere un principio di simmetria, secondo cui la figura che realizza il massimo o il minimo è quella che gode la più perfetta simmetria rispetto ai suoi elementi. Questo principio ha d'altronde un campo limitato di validità; esso si applica soltanto al caso in cui la funzione da rendere massima o minima sia una funzione simmetrica rispetto ai nominati elementi (variabili). Ma — anche così limitato — il principio di simmetria non costituisce una legge generale, bensì — piuttosto — un principio di scoperta.

Il suo esatto valore si desume dalle seguenti considerazioni che — per brevità — limiteremo ai casi più semplici.

1) Si abbia da rendere massima o minima la *funzione simmetrica*:

$$z = f(x_1, x_2, \dots x_n).$$

Le condizioni differenziali del 1° ordine sono, secondo il § 23:

$$\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = \dots = \frac{\partial f}{\partial x_n} = 0.$$

Ora, per ipotesi, la f contenendo simmetricamente le $x_1 \dots x_n$, rimane inalterata per uno scambio qualsiasi di variabili (x_i, x_h). Un tale mutamento scambia invece le due derivate parziali $\frac{\partial f}{\partial x_i}, \frac{\partial f}{\partial x_h}$. Si deduce che per $x_i = x_h$, $\frac{\partial f}{\partial x_i} = \frac{\partial f}{\partial x_h}$.

Quindi per

$$x_1 = x_2 = \dots = x_n = x,$$

$$\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = \dots = \frac{\partial f}{\partial x_n};$$

e ponendo

$$f(x, x, \dots x) = F(x),$$

$$\frac{\partial F}{\partial x} = n \frac{\partial f}{\partial x_1}.$$

Pertando *le condizioni differenziali del 1° ordine per l'esistenza d'un massimo o d'un minimo della funzione simmetrica* f *sono soddisfatte nel punto*

$$x_1 = x_2 = \dots = x_n = x,$$

per cui

$$\frac{dF}{dx} = 0.$$

2) Si vogliano i massimi o minimi della *funzione simmetrica*

$$z = f(x_1, x_2, ... x_n),$$

subordinatamente alla *condizione simmetrica*

$$\varphi(x_1, x_2, ... x_n) = 0.$$

Le condizioni differenziali del 1° ordine (§ 28) sono:

$$\frac{\partial f}{\partial x_i}\frac{\partial \varphi}{\partial x_r} - \frac{\partial f}{\partial x_i}\frac{\partial \varphi}{\partial x_r} = 0;$$

esse *sono soddisfatte per*

$$x_1 = x_2 = ... = x_n.$$

Infatti, ponendo

$$x_1 = x_2 = ... = x_n = x,$$

dove x resta determinato dalla condizione

$$\varphi(x, x, ... x) = \Phi(x) = 0,$$

si ottiene

$$\frac{\partial f}{\partial x_i} = \frac{\partial f}{\partial x_r}, \quad \frac{\partial \varphi}{\partial x_i} = \frac{\partial \varphi}{\partial x_r}.$$

3) Si abbiano da determinare i massimi e minimi della funzione simmetrica

$$z = f(x, y),$$

subordinatamente alla condizione simmetrica

$$\varphi(x, y) = 0.$$

In questo *problema simmetrico vincolato, che dipende da una sola variabile indipendente, si ha per*

$$x = y$$

un massimo o un minimo.

Infatti le due superficie

$$z = f(x, y), \quad \varphi(x, y) = 0$$

s'intersecano secondo una linea che ha come piano di simmetria il piano $x = y$; una linea siffatta ha un punto di massimo o di minimo (relativo) nel punto in cui interseca il piano $x = y$.

4) Nel problema libero, dove si tratta di rendere massima o minima la funzione $z = f(x, y)$ (oppure dove si tratti di render massima o minima una funzione simmetrica di tre variabili subordinamente ad *una* condizione simmetrica) *non si ha sempre un massimo o un minimo per* $\mathrm{x} = \mathrm{y}$.

Si assuma per esempio

$$z = + \sqrt{xy + r^2}.$$

Per $x = y$ il minimo di z si ha facendo

$$x = y = 0, \quad z = r,$$

Ebbene, questo non è un minimo per la funzione di due variabili, giacchè per $x > 0$, $y > 0$ o $x < 0$, $y < 0$, si hanno valori più grandi di z, ma per $x > 0$, $y < 0$, si hanno valori più piccoli.

Se si pone

$$x = u_1 + u_2, \quad y = u_1 - u_2,$$

si ottiene

$$z = + \sqrt{u_1^2 - u_2^2 + r^2},$$

che è l'equazione d'un iperboloide ad una falda riferito ai piani coordinati u_1, u_2, z; il punto $u_1 = u_2 = 0$ (corrispondente ad $x = y = 0$) è un punto iperbolico.

5) Il principio di simmetria acquista valore trattandosi di ricercare il *massimo* o *il minimo assoluto d'una funzione simmetrica* $\mathrm{f}(\mathrm{x}_1 \ldots . \mathrm{x}_n)$, eventualmente con date condizioni simmetriche, quando si possa stabilire insieme *l'esistenza* e *l'unicità del massimo o del minimo* richiesto.

In questo caso infatti si avrà senz'altro che questo corrisponde ad

$$x_1 = x_2 = \ldots . = x_n.$$

Ci riferiremo come esempio a due problemi classici già innanzi accennati.

Minimo della somma

$$z = x_1 + x_2 + ... + x_n,$$

per

$$x_1 x_2 .... x_n = p$$

(problema reciproco del massimo del prodotto).

Il minimo esiste (§ 13).

Il minimo *è unico*. Infatti se si ha

$$a_1 + a_2 + ... + a_n = b_1 + b_2 .... + b_n = s$$
$$a_1 a_2 .... a_n = b_1 b_2 .... b_n = p,$$

si ha pure

$$\sqrt{a_1 b_1} < \frac{a_1 + b_1}{2},$$

quindi

$$\sqrt{a_1 b_1} + \sqrt{a_2 b_2} + ... + \sqrt{a_n b_n} < s,$$

con

$$\sqrt{a_1 b_1} \cdot \sqrt{a_2 b_2} .... \sqrt{a_n b_n} = p.$$

Dunque il minimo si ha per

$$x_1 = x_2 = ... = x_n = \sqrt[n]{p}.$$

Poligono di n *lati, iscritto in un cerchio, di perimetro massimo.*

Indicando con $\theta_1 .... \theta_n$ le metà degli angoli al centro corrispondenti ai lati, questi sono

$$2 \operatorname{sen} \theta_1, ... \quad 2 \operatorname{sen} \theta_n.$$

Si tratta dunque di render massima la funzione

$$z = 2 \operatorname{sen} \theta_1 + 2 \operatorname{sen} \theta_2 .... + 2 \operatorname{sen} \theta_n$$

con

$$\theta_1 + \theta_2 + ... + \theta_n = \pi.$$

Il massimo esiste (§ 13). Il minimo pure esiste e corrisponde all'annullarsi o al divenire $= \pi$ di qualcuna delle θ.

Le condizioni differenziali corrispondono al massimo della funzione

$$Z = z + \lambda \sum_1^n \theta_i$$

(§ 28). Si ha dunque

$$\frac{\partial Z}{\partial \theta_1} = 2 \cos \theta_1 + \lambda, \ldots ,$$

$$\cos \theta_1 = \cos \theta_2 = \ldots = \cos \theta_n$$

con

$$0 < \theta_i < \pi.$$

Queste equazioni ammettono la *soluzione unica*

$$\theta_1 = \theta_2 = \ldots = \theta_n,$$

che corrisponde al massimo cercato. Il *massimo* è dunque il *poligono regolare.*

NOTA. L'area del poligono iscritto nel cerchio è

$$\frac{1}{2} \{ \operatorname{sen} 2\theta_1 + \ldots + \operatorname{sen} 2\theta_n \},$$

$$(\theta_1 + \ldots + \theta_n = \pi);$$

quindi essa diventa pure massima per il poligono regolare:

$$\cos 2\theta_1 = \cos 2\theta_2 = \ldots = \cos 2\theta_n,$$

$$0 < \theta_i < \pi,$$

$$\theta_1 = \theta_2 = \ldots = \theta_n.$$

§ 31. **Nota storica** [1]. — Il problema dei massimi e minimi, e il problema della tangente ad una curva, della velocità o della derivata debbono riguardarsi storicamente come varii aspetti di un medesimo problema, attraverso a cui si svolgono storicamente i concetti del calcolo differenziale.

[1] Cfr. tre note storiche di G. VACCA nel Formulario edito da G. Peano (t. IV e V) e l'articolo di A. Voss nella « Encyclopädie der Mathematischen Wissenschaften » (II A 2).

Già negli Elementi d'EUCLIDE (libro III) si considera il contatto di rette e cerchi e si risolve il problema di determinare la minima distanza di un punto da una circonferenza.

Problemi analoghi sono trattati da APOLLONIO per le coniche (Libro IV e Frammenti I. VII ed HEIBERG, vol. II - cfr. PAPPO, libro VII ed HULTSCH, III, p. 339, n. VI e seg.).

La tangente al cerchio è considerata in DESCARTES come retta che ha un punto comune colla linea e lascia la linea tutta da una parte. Il medesimo concetto serve ad ARCHIMEDE per definire la tangente alla *spirale*.

Si può dire che la considerazione sopra accennata della tangente, contiene implicitamente il rapporto che si ha in generale fra il problema del massimo e minimo e quello della tangente.

Il concetto generale della tangente che appartiene alle Matematiche moderne, si trova in DESCARTES (« Géométrie », 1637, Oeuvres t. VI, p. 418); questi chiama tangente una retta che incontra la curva in due punti « riuniti in uno solo ». Tale definizione si riduce al concetto moderno se s' interpreta dicondo « tangente è il limite d'una retta secante quando uno dei punti d'incontro s' avvicina indefinitamente all' altro ». Il procedimento analitico indicato da DESCARTES per determinare la tangente ad una curva $f(x, y) = 0$, ha condotto più tardi HUDDE, (Geometria a Renato Des Cartes...., 1659) a quello che oggi si chiamerebbe calcolo delle derivate delle funzioni implicite.

Contemporaneamente allo sviluppo del pensiero di DESCARTES devesi menzionare quello di ROBERVAL e TORRICELLI che considerarono la tangente ad una curva come *velocità* d'un punto mobile che la descrive, sebbene non pervenissero veramente a trasformare quest' idea in un generale procedimento analitico. Ma soprattutto c' interessano qui le ricerche di P. FERMAT, (« Méthode pour la recherche du Maximum et du Minumum », Oeuvres trad. TANNERY, t. III, p. 121), che si riferiscono specialmente ai problemi di massimo e di minimo e sono da riguardare come il primo passo verso la costituzione del calcolo differenziale.

FERMAT prese le mosse da un' osservazione già fatta in PAPPO, cioè che nel punto di massimo o di minimo della funzione $y = f(x)$ ci sono — per x — due radici uguali. Quest' osservazione lo conduceva in particolare alla regola (§ 22) che consiste nell'annullare il discriminante della funzione inversa (l. c. e § 22), ripresa poi da A. DE MONFORTE (1).

Ma il metodo principale proposto da FERMAT si può riassumere come segue:

Per determinare il massimo o il minimo della funzione $f(x)$ si deve dare ad x un incremento h e porre l'equazione

$$f(x)=f(x+h).$$

Divisi i due membri dell'equazione per h e posto quindi $h=0$, si trova la condizione di massimo o minimo.

Per es. se si tratti di trovare il massimo di $x(a-x)$ si scrive:

$$a(x+h)-(x+h)^2=ax-x^2$$
$$ah-2xh-h^2=0;$$

dividendo per h si ha

$$a-2x-h=0,$$

e facendo $h=0$ viene cioè

$$x=\frac{a}{2}.$$

LAGRANGE nelle sue « Leçons sur le calcul des fonctions » (Oeuvres, t. V, p. 294) osserva che la regola di FERMAT coincide colla regola del Calcolo differenziale, dove si trascurano come infinitesimi d'ordine superiore i termini che sono considerati da FERMAT come nulli; egli pure fa rilevare la stretta analogia del procedimento di FERMAT con quello di LEIBNIZ, « Nova methodus pro maximis el minimis... » (Acta eruditorum, 1684). Tuttavia i geometri contemporanei non compresero lo spirito del metodo di FERMAT riguardandolo come un puro artificio, come risulta dalla lunga discussione a cui quel metodo dette origine (Cfr. FERMAT, Oeuvres trad. TANNERY, t. IV, p. 25). Perciò l'invenzione di FERMAT ebbe poca influenza. Fa eccezione la regola di SLUZE per trovare le tangenti, che sembra ispirata appunto da FERMAT.

(1) Cfr. F. AMODEO, « Periodico di Matematica », 1909.

Barrow nel 1674 espose il suo metodo delle tangenti, in cui le quantità supposte nulle da Fermat vengono sostituite con quantità evanescenti. L'algoritmo del calcolo differenziale si costituisce quindi con Newton « Principia » (1687) e Leibniz (l. c. cfr. Mathematische Schriften ed. Gerhardt, 1849-63).

La regola per riconoscere se effettivamente un punto di zero della derivata $f'(x)$ corrisponda ad un massimo della funzione $f(x)$, fondata sull' uso delle derivate successive, apparsiene a Maclaurin (cfr. Lagrange, Oeuvres 1, p. 4).

La determinazione dei massimi e minimi delle funzioni di più variabili è già contemplata da Fermat (l. c.).

Il criterio sufficiente per l' esistenza del massimo o minino di una funzione di due o più variabili, che la differenza

$$f(x_1 + h_1, \ldots x_n + h_n) - f(x_1 \ldots x_n)$$

si esprima a meno d' ordine superiore con una forma quadratica *definita*

$$\frac{1}{2} \Sigma \frac{\partial^2 f}{\partial x_i \partial x_h} h_i h_h,$$

trovasi in Lagrange, 1759 (Oeuvres 1, p. 1 e segg.). Il caso della forma *semidefinita* deve essersi presentato nelle ricerche sulle superficie di Monge e di Gauss (punto parabolico). Esso trovasi notato espressamente da Gergonne nel 1830 (Annales de Gergonne, t. 20). Su questo caso regnò fino ai tempi più recenti l' erronea opinione che si avesse massimo o minimo quando ciò accade pei valori di h_i che annullano la forma quadratica. Nel « Calcolo differenziale » di Genocchi e Peano trovasi corretto questo errore con l' esempio della funzione

$$f(x, y) = (x^2 - 2py)(x^2 - 2qy)$$
$$p > 0, \quad q > 0.$$

Lo studio approfondito del caso in questione è stato sviluppato da L. Scheeffer (Math. Annalen, Bd 35, 1885).

Per l' estensione a più di due variabili cfr. O. Stolz, Wien, Ber. (1890-93).

Una chiara esposizione delle anzidette teorie trovasi in E. Goursat, « Cours d' Analyse mathématique », II ed. Paris, 1910, t. I.

PARTE SECONDA

I. MASSIMI E MINIMI DELLE FUNZIONI DI LINEE E IL CALCOLO DELLE VARIAZIONI

I. - Sull' esistenza dei massimi e minimi delle funzioni di linee.

§ 32. **Introduzione.** — La Geometria elementare ci offre esempio di problemi in cui si tratta di trovare il massimo e il minimo di qualche dato che varia non già come funzione di un numero finito di variabili, ma piuttosto in dipendenza della *forma di una linea* cioè come *funzione di una linea.*

Basta infatti richiamare i problemi seguenti:

1) Determinare la linea di lunghezza minima fra quelle che congiungono due punti *dati* A, B.

2) Determinare la linea piana di area massima fra quelle che hanno una data lunghezza (isoperimetre).

Il primo problema trovasi risolto implicitamente in EUCLIDE, Prop. XX, ove dimostrasi che in ogni triangolo un lato è minore della somma degli altri due. Quando si definisca la lunghezza d'una linea (rettificabile) come limite delle poligonali iscritte, si deduce che « *la linea di lunghezza minima è la retta* » (cfr. § 39).

Quanto al 2° problema, esso ha formato oggetto di trattazione speciale nell' Art. XXVI.

Ora, prendendo le mosse da codesti problemi elementari si è condotti a sviluppare alcuni principii generali che forniscono *criteri per l'esistenza e per la determinazione dei massimi e minimi delle funzioni di linee.*

33. **Insieme di linee: linee limiti.** — In ciò che segue il concetto di *linea* o *curva continua* è posto nella sua maggiore generalità, conforme al punto di vista *analitico*; dicesi « linea » il luogo rappresentato da una funzione continua $y = \varphi(x)$, o più generalmente da due funzioni continue invertibili

$$x = \varphi_1(t), \quad y = \varphi_2(t),$$

in un certo intervallo.

Si consideri una famiglia di linee continue

$$y = \varphi(x),$$

definite fra due estremi a, b ($a < b$). E pongasi che le funzioni φ variino esse stesse in un *intervallo funzionale* o *striscia*, fra due funzioni estreme:

$$\begin{matrix} \psi(x) \leq \varphi(x) \leq \theta(x) \\ \psi(x) \leq \theta(x) \end{matrix} \qquad \text{(per ogni } x \text{ fra } a, b).$$

Def. Si abbia un insieme di linee $\varphi(x)$ appartenenti alla famiglia considerata. Si dirà che la linea $\bar{\varphi}$ è una *linea limite dell'insieme*, se, per ogni numero positivo ε piccolo ad arbitrio, esiste una φ, dell'insieme suddetto, tale che — in valore assoluto —

$$| \varphi(x) - \bar{\varphi}(x) | < \varepsilon \quad \text{(per ogni } x \text{ fra } a, b).$$

In *modo più generale* si può considerare una famiglia di linee rappresentate da equazioni parametriche del tipo

$$(1) \qquad \begin{cases} x = \varphi_1(t) \\ y = \varphi_2(t) \end{cases} \qquad \text{(per } t \text{ fra } a \text{ e } b)$$

dove φ_1 e φ_2 sono funzioni continue e invertibili ($t =$ funz. (x, y)) comprese in un certo intervallo funzionale:

$$\begin{matrix} \psi_1(t) \leq \varphi_1(t) \leq \theta_1(t) \\ \psi_2(t) \leq \varphi_2(t) \leq \theta_2(t) \\ \psi_1 \leq \theta_1, \quad \psi_2 \leq \theta_2. \end{matrix}$$

Si dirà che la linea

$$\begin{cases} x = \bar{\varphi}_1(t) \\ y = \bar{\varphi}_2(t) \end{cases}$$

è *linea limite* di un insieme di linee (1), se — per ogni numero positivo ε piccolo ad arbitrio — esiste una linea $\begin{cases} x = \varphi_1(t) \\ y = \varphi_2(t) \end{cases}$ dell'insieme considerato, tale che, in valore assoluto,

$$\begin{matrix} | \varphi_1(t) - \bar{\varphi}_1(t) | < \varepsilon \\ | \varphi_2(t) - \bar{\varphi}_2(t) | < \varepsilon \end{matrix} \qquad \text{(per ogni } t \text{ fra } a \text{ e } b).$$

Esempio. Sieno $\varphi(x)$, $\omega(x)$, $\psi(x)$, $\theta(x)$, funzioni continue in un intervallo (a, b), e sia — per t compreso fra 0 e $\bar{t}$ (> 0) —

$$\psi(x) \leq \varphi(x) + t\omega(x) \leq \theta(x).$$

La linea

$$y = \varphi(x) + t_0\omega(x)$$

$$(0 \leq t_0 \leq \bar{t})$$

è linea limite della famiglia

$$y = \varphi(x) + t\omega(x).$$

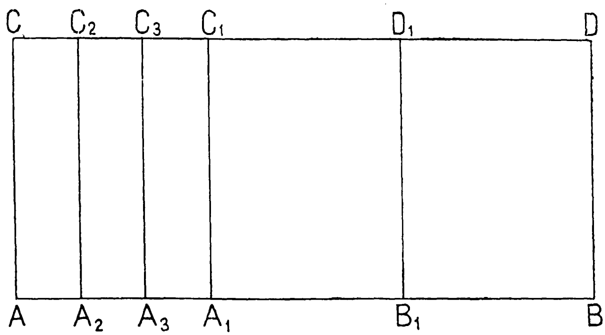
Infatti

$$y = \varphi(x) + t\omega(x)$$

è — nelle ipotesi fatte — una funzione continua delle due variabili x, t, che — per $t = t_0$ — tende alla

$$y = \varphi(x) + t_0\omega(x).$$

Le definizioni precedenti sono analoghe a quella di punto limite d'un gruppo di punti sulla retta o nel piano. Ma



— mentre un gruppo infinito di punti, in un intervallo o campo finito ha sempre qualche punto limite — qui ci troviamo di fronte al caso nuovo:

Un insieme di linee continue — appartenenti ad un intervallo funzionale — non ha sempre qualche linea limite.

Può darsi infatti che i punti limiti di $\varphi(x)$ per $x = \text{cost}$, non descrivano al variare di x delle linee, ma invadano addirittura delle superficie.

Esempio. Si abbia il rettangolo $ABCD$ (v.[i] figura).

Si dividano i lati opposti AB, CD in tre parti uguali mediante i punti A_1, B_1, C_1, D_1, e si consideri quindi la linea $ACC_1A_1B_1D_1DB$; si dividano quindi i lati opposti AA_1, CC_1 del rettangolo AA_1CC_1 in tre parti uguali mediante i punti A_2A_3, C_2C_3 e si costruisca la linea poligonale

$$ACC_2A_2A_3C_3C_1A_1\,.$$

In modo analogo si operi sopra i rettangoli $C_1D_1A_1B_1$, B_1BD_1D e poi su ciascuno dei rettangoli AA_2CC_2 ecc. Questo procedimento ripetuto dà luogo ad una *successione di poligonali che ha per limite l'intera superficie del rettangolo* ABCD.

Un altro semplice esempio è costituito dall'insieme di linee

$$y = \operatorname{sen} nx,$$

per x compreso fra 0 e π e per $n = 1, 2, \ldots$ —. Per n crescente queste linee fanno un numero d'oscillazioni crescente e tendono — per $n = \infty$ — all'intero rettangolo limitato dalle rette

$$x = 0, \quad x = \pi, \quad y = 1, \quad y = -1.$$

L'esistenza di linee-limiti per un insieme di linee può essere dimostrata in condizioni assai generali, quali sono espresse dai teoremi dei §§ seguenti.

§ 34. **Uguale continuità: teorema di Ascoli.** — Un criterio fondamentale per l'esistenza di una linea limite d'un insieme di linee, si basa sul concetto della *uguale continuità* posto da G. Ascoli [1].

Def. Un insieme di funzioni (linee) $y = \varphi(x)$, definite entro un certo intervallo (a, b), si dice costituito da *funzioni* (linee) *ugualmente continue* se — dato un numero positivo ε, piccolo ad arbitrio — si può determinare un δ per modo che sia in valore assoluto,

$$|\,\varphi(x) - \varphi(x')\,| < \varepsilon,$$

contemporaneamente per tutte le funzioni φ e per ogni coppia

[1] « Le curve limiti di una varietà data di curve ». Mem. Acc. Lincei, 1884.

di valori x, x', per cui

$$|x - x'| < \delta.$$

Ora sussiste il

Teorema di ASCOLI [1]: Un insieme di infinite linee,

$$y = \varphi(x)$$

ugualmente continue in un intervallo (a, b), e comprese in un intervallo funzionale ψ, θ, possiede almeno una linea limite.

Per costruire una linea limite si procede come segue: 1) si definisce un punto limite dei punti $y = \varphi(x)$ per ogni valore razionale di x, e si determina quel punto $f(x)$ in modo che le medesime curve φ si accostino contemporaneamente ai rispettivi punti limiti; 2) si stabilisce la continuità della funzione f sul gruppo dei punti x razionali e quindi si estende — come funzione continua — per tutti i punti dell'intervallo (a, b); 3) si dimostra che $y = f(x)$ è linea limite delle φ.

Si considerino dunque i punti x razionali fra a e b; essi formano un gruppo numerabile (Vol. I, Art. VI), le cui ascisse verranno prese in certo ordine

$$x_1, x_2, \ldots\, x_n, \ldots$$

Tutte le funzioni φ, per $x = x_1$, danno luogo ad un gruppo di infiniti punti $\varphi(x_1)$, computandosi eventualmente come punti sovrapposti quelli che rispondessero a φ diverse prendenti in x_1 lo stesso valore. Ora designando con k il minimo di $\psi(x)$ e con k' il massimo di $\theta(x)$, tutte le $\varphi(x)$ — e quindi in particolare i punti $\varphi(x_1)$ — si trovano fra

$$k \quad \text{e} \quad k',$$

e perciò quel gruppo di punti ammette certo qualche punto limite. Fra tali punti sceglieremo *il più basso* $f(x_1)$, notando

[1] L. c. Nella dimostrazione originale ed anchè nelle successive esposizioni o modificazioni di essa, si suppone sempre — esplicitamente o meno — che, dall'insieme dato di curve, si possa estrarre una *successione* $C_1, C_2, \ldots\, C_n, \ldots$. L. TONELLI nella sua nota « Sul valore d'un certo ragionamento » (Atti Acc. delle Scienze di Torino, 1913) ha reso la dimostrazione del teorema indipendente da questa ipotesi. Il testo riproduce il ragionamento di Tonelli.

che il limite inferiore del gruppo dei punti limiti è certo un punto limite.

Assumasi ora una successione di numeri positivi decrescenti e tendenti a zero:

$$\varepsilon_1, \varepsilon_2, \ldots \varepsilon_m, \ldots$$

Fra le funzioni φ si può distaccare un insieme che si approssima al punto limite $f(x_1)$ più di ε_m; consideriamo infatti le infinite funzioni φ per cui è soddisfatta (in valore assoluto) la diseguaglianza

$$|\varphi(x_1) - f(x_1)| \leq \varepsilon_m;$$

queste funzioni $\varphi_m^{(1)}$ formeranno un insieme $I_m^{(1)}$. Per $m = 1, 2, \ldots$ si ha così una successione d'insiemi ciascuno contenuto nel precedente.

Ora si consideri il gruppo d'infiniti punti $\varphi_m^{(1)}(x_2)$; questo avrà un gruppo di punti limiti $G_m^{(2)}$. Per $m = 1, 2, \ldots$, si ottiene una successione di gruppi infiniti

$$G_1^{(2)} G_2^{(2)} \ldots,$$

ciascuno contenuto nel precedente.

In forza del postulato di continuità si conclude facilmente che i gruppi chiusi $G_m^{(2)}$ — per $m = 1, 2 \ldots$ — contengono un gruppo comune, costituito almeno di un punto, $G_\infty^{(2)}$. Si può considerare il punto più basso di $G_\infty^{(2)}$ e chiamarlo $f(x_2)$.

Si ripeterà quindi il procedimento innanzi descritto, passando da x_2 ad x_3 come si è passati da x_1 ad x_2, e così di seguito. Resterà così definita per tutti i punti razionali la funzione $f(x_n)$, e si avranno infinite successioni d'insiemi di funzioni φ:

$$\begin{array}{l} I_1^{(1)}\, I_2^{(1)} \ldots I_m^{(1)} \ldots, \\ I_1^{(2)}\, I_2^{(2)} \ldots I_m^{(2)} \ldots \\ \ldots\ldots\ldots\ldots \end{array}$$

Si tratta di provare che le funzioni $\varphi_m^{(m)}$ di $I_m^{(m)}$, al crescere di m, convergono verso una funzione limite, che per i punti razionali si riduce ad $f(x_n)$.

Dimostriamo anzitutto che $f(x_n)$ è continua rispetto all'insieme dei punti razionali fra a e b.

Intanto le φ sono ugualmente continue, cioè dato un numero positivo ε, piccolo ad arbitrio, si può trovare un δ tale che per

$$|x - x'| < \delta$$

sia

$$|\varphi(x) - \varphi(x')| < \varepsilon,$$

per ogni φ e per ogni x fra a e b.

Ciò posto si consideri un punto razionale x_r e sia x_s un altro punto razionale assai vicino, cioè tale che

$$|x_r - x_s| < \delta.$$

Si avrà — qualunque sia m —

$$|\varphi_m^{(m)}(x_r) - \varphi_m^{(m)}(x_s)| < \varepsilon.$$

Ma per la definizione delle funzioni $\varphi_m^{(m)}$ di $I_m^{(m)}$, purchè sia

$$r < m, \quad s < m,$$

si avrà

$$\begin{cases} |\varphi_m^{(m)}(x_r) - f(x_r)| \leq \varepsilon_m \\ |\varphi_m^{(m)}(x_s) - f(x_s)| \leq \varepsilon_m, \end{cases}$$

e quindi

$$|f(x_r) - f(x_s)| \leq \varepsilon + 2\varepsilon_m.$$

Questa diseguaglianza deve sussistere per m sufficientemente grande, quanlunque sia m, e perciò (essendo $\lim_{m=\infty} \varepsilon_m = 0$):

$$|f(x_r) - f(x_s)| \leq \varepsilon,$$

ogni qualvolta

$$|x_r - x_s| < \delta.$$

Ciò significa appunto che la funzione $f(x_r)$ è continua nell'insieme dei punti razionali x_r, e perciò resta definita *come funzione continua* f(x) *per tutti i valori di* x fra a e b.

Dico che $y = f(x)$ è curva limite delle $y = \varphi_m^{(m)}(x)$, alla quale si avvicinano indefinitamente delle curve $\varphi_m^{(m)}$ per m alto quanto si vuole.

A tale scopo scegliamo fra i numerosi razionali compresi nell'intervallo (a, b) una successione di numeri crescenti

$$x_{r_1}, x_{r_2}, \dots x_{r_i}, \dots x_{r_p},$$

tali che:

$$a < x_{r_1} < x_{r_2} \dots < x_{r_p} < b,$$
$$x_{r_1} - a < \delta, \quad x_{r_2} - x_{r_1} < \delta, \dots, \quad b - x_{r_p} < \delta$$

Si consideri un punto qualsiasi x fra a, b, e sia p. es.

$$x_{r_i} \leq x \leq x_{r_{i+1}}.$$

Avremo

$$\begin{aligned} | \varphi_m{}^{(m)}(x) - f(x) | \leq & \, | \varphi_m{}^{(m)}(x) - \varphi_m{}^{(m)}(x_{r_i}) | \\ & + | \varphi_m{}^{(m)}(x_{r_i}) - f(x_{r_i}) | \\ & + | f(x_{r_i}) - f(x) | . \end{aligned}$$

Ora — per la uguale continuità delle φ —

$$| \varphi_m{}^{(m)}(x) - \varphi_m{}^{(m)}(x_{r_i}) | < \varepsilon$$

essendo

$$| x - x_{r_i} | < \delta;$$

inoltre — per la definizione delle $\varphi_m{}^{(m)}$ e di f, supposto $m > r_i$ —

$$| \varphi_m{}^{(m)}(x_{r_i}) - f(x_{r_i}) | < \varepsilon_m;$$

finalmente

$$| f(x_{r_i}) - f(x) | \leq \varepsilon,$$

giacchè tale diseguaglianza dimostrata per tutti i punti razionali di un intervallo d'ampiezza δ si estende anche ai punti irrazionali d'un tale intervallo, in forza della continuità di f,

Si conclude dunque

$$| \varphi_m{}^{(m)}(x) - f(x) | \leq 2\varepsilon + \varepsilon_m,$$

In questa diseguaglianza si deve supporre m superiore a tutti gl'indici $r_1, r_2, \dots r_p$. Si può anche supporre m così grande che sia

$$\varepsilon_m < \varepsilon,$$

e si avrà dunque — per m assai grande —

$$| \varphi_m{}^{(m)}(x) - f(x) | < 3\varepsilon,$$

qualunque sia x fra a e b.

Ciò significa appunto che $y = f(x)$ è curva limite delle $\varphi_m^{(m)}$ c. d. d.

Nota. Risulta dalla dimostrazione che precede anche il teorema seguente:

Se è data una successione d'insiemi di linee (funzioni)

$$I_1, I_2, \dots I_m \dots,$$

ciascuno contenuto nel precedente, e se le nominate funzioni (linee) — comprese in un intervallo funzionale — sono ugualmente continue, si può determinare una linea limite a cui si approssimano indefinitamente linee dell'insieme I_m, per m grande quanto si vuole.

§ 35. **Teorema di Arzelà.** [1] — Sia $y = \varphi(x)$ un insieme d'infinite funzioni, continue e derivabili, comprese in un intervallo funzionale:

$$\psi(x) \leq \varphi(x) \leq \theta(x) \qquad (a \leq x \leq b),$$

e sia soddisfatta — per tutte le φ e per qualsiasi x fra a e b — la *condizione di* Lipschitz, cioè la derivata $\varphi(x)$ sia inferiore in valore assoluto ad un numero positivo l:

$$|\varphi'_{(x)}| < l;$$

allora *esiste una linea limite* delle $y = \varphi(x)$.

Si prova infatti che le $\varphi(x)$ godono della uguale continuità.

Sia dato dunque un numero positivo ε piccolo ad arbitrio, si vuol determinare un δ tale che per

$$|x - x'| < \delta$$

sia sempre

$$|\varphi(x) - \varphi(x')| < \varepsilon,$$

qualunque sia la funzione φ e la coppia x, x' che si considera.

A tale scopo basta invocare la nota eguaglianza

$$\varphi(x) - \varphi(x') = (x - x')\varphi'(\bar{x}),$$

[1] Cfr. C. Arzelà, «Sulle serie di funzioni» Mem. Accad. di Bologna, 1889-1900.

ove $\bar{x}$ è un punto compreso fra x e x'. Essendo

$$|x - x'| < \delta$$
$$|\varphi'(\bar{x})| < l,$$

si deduce

$$|\varphi(x) - \varphi(x')| < \delta l;$$

e perciò la differenza del primo membro si rende $< \varepsilon$, prendendo

$$\delta < \frac{\varepsilon}{l}.$$

§ 39. **Teorema di Hilbert** [1]. — Si abbia un insieme di linee rettificabili, aventi gli estremi fissi a, b, e appartenenti a un dato intervallo funzionale, e suppongasi che la lunghezza di tutte le linee sia inferiore ad un dato limite l. Allora esiste certo una linea limite dell'insieme.

Si assumano le linee dell'insieme dato rappresentate parametricamente mediante le funzioni *continue e derivabili*:

$$(1) \qquad \begin{cases} x = \varphi_1(t) \\ y = \varphi_2(t) \end{cases}$$

ove il parametro t si esprime mediante l'arco s colla formula

$$t = s\frac{l}{k},$$

k essendo la lunghezza della linea (1) fra gli estremi dati.

Si avrà

$$\varphi_1'(t) = \frac{k}{l}\varphi_1'(s),$$

ma

$$\frac{k}{l} \leq 1,$$

$$|\varphi_1'(s)| = \left|\frac{dx}{ds}\right| = \left|\frac{dx}{\sqrt{dx^2 + dy^2}}\right| < 1,$$

[1] D. Hilbert, « Ueber das Dirichlet'sche Prinzip » Jahresbericht der Deutschen Math. Vereinigung, 1900, ha dato questo criterio applicandolo anche al caso delle geodetiche, Cfr. Math. Annalen, Bd LIX, p. 161.

e perciò

$$|\varphi_1'(t)| < 1.$$

Si applica dunque alle funzioni $x = \varphi_1(t)$, il teorema di ARZELÀ. Lo stesso ragionamento si ripete per le funzioni $y = \varphi_2(t)$. Si designino ora con

$$x = f_1(t), \quad y = f_2(t),$$

le rispettive curve limiti di φ_1, φ_2, costruite secondo il procedimento già indicato; la curva

$$\begin{cases} x = f_1(t) \\ y = f_2(t), \end{cases}$$

sarà curva limite delle

$$\begin{cases} x = \varphi_1(t) \\ y = \varphi_2(t). \end{cases}$$

E pertanto il teorema è dimostrato.

§ 37. **Approssimazione di una curva continua con una curva algebrica.** — I precedenti teoremi di ASCOLI, ARZELÀ ed HILBERT, si riferiscono al caso in cui sia data una famiglia di curve continue, e si cerchi una curva limite. Una questione inversa è quella di « approssimare una curva data con curve di specie determinata », cioè costruire una famiglia di curve, della specie suddetta, che abbia la curva data come curva limite.

Accenniamo brevemente ad alcuni resultati in quest'ordine d'idee.

Anzitutto si ha:

***Data una curva continua* y $= \varphi(x)$, *per* x *fra* a *e* b, *si può costruire una poligonale che si approssimi quanto si vuole alla curva* φ.**

Invero per il teorema di HEINE (§ 4), dato un numero positivo ε piccolo ad arbitrio, si può determinare un δ per modo che — dividendo l'intervallo (a, b) in tratti successivi (a, x_1), $(x_1, x_2), \ldots. (x_n, b)$, di ampiezza δ —, si abbia ogni qualvolta

$$x_i \leq x \leq x_{i+1}$$

$$|\varphi(x) - \varphi(x_i)| < \frac{\varepsilon}{2}.$$

Congiungendo i punti

$$\varphi(a),\ \varphi(x_1),\quad \varphi(x_2)\ \ldots\ \varphi(x_n),\quad \varphi(b),$$

si ha appunto una poligonale $y = \psi(x)$ che approssima alla curva $y = \varphi(x)$ più di ε.

Infatti da

$$|\psi(x) - \psi(x_i)| < \frac{\varepsilon}{2}$$

$$|\varphi(x) - \varphi(x_i)| < \frac{\varepsilon}{2}$$

segue

$$|\psi(x) - \varphi(x)| \leq |\psi(x) - \psi(x_i)| + |\varphi(x_i) - \varphi(x)| < \varepsilon.$$

NOTA. Lo stesso teorema vale anche, e si stabilisce facilmente, per una curva continua rappresentata da equazioni parametriche

$$(1) \qquad \begin{cases} x = \varphi_1(t) \\ y = \varphi_2(t), \end{cases}$$

dove si suppone inversamente ψ funziona continua di x, y.

Il procedimento precedente, qui applicato, conduce in generale a poligonali con punti doppi; ma si trova una poligonale sciolta vicina quanto si vuole ad una poligonale con punti doppi.

Ora dimostriamo il

Lemma. *Una poligonale sciolta di* n *lati, si può approssimare quanto si vuole con un arco continuo reale appartenente ad una curva algebrica d'ordine* n,

$$f(x, y) = 0,$$

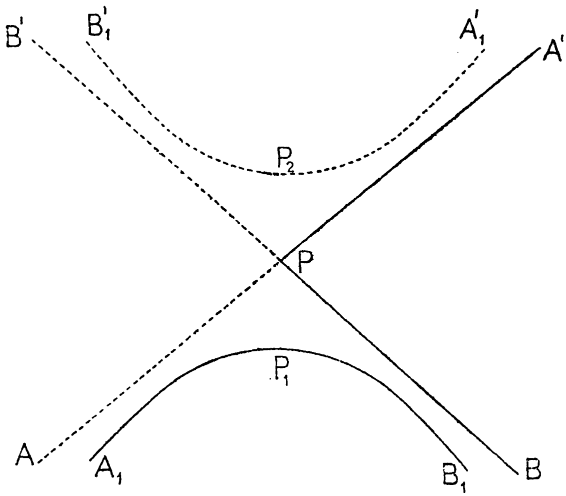
senza punti doppi.

Anzitutto gli n lati della poligonale appartengono ad una curva algebrica riducibile, d'ordine n,

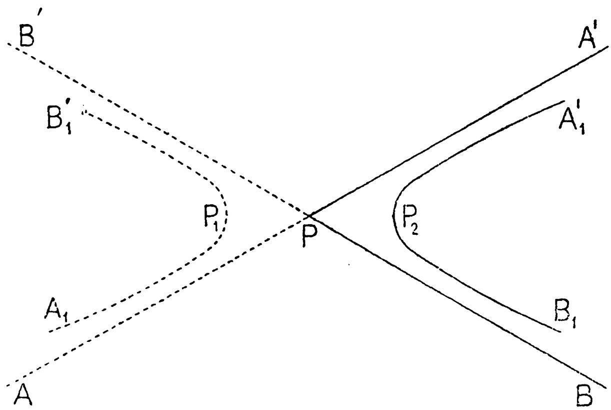
$$\varphi(x, y) = 0.$$

Dando una piccola variazione ai coefficienti di φ, in guisa che il discriminante dell'equazione divenga diverso da 0, si ottengono curve algebriche $f(x, y) = 0$, senza punti doppi, vicine quanto si vuole alla curva φ. Una curva f così determinata consterà di *più rami reali.*

Si designi con P un vertice della data poligonale, con A, B due punti appartenenti ai lati adiacenti a P, e con A', B' due punti appartenenti ai prolungamenti di AP, BP, dal-



l'altra parte di P. Se s'immagina che la curva f varii in modo determinato avvicinandosi alla curva limite φ, vi sa-



ranno su f dei punti A_1, B_1, A_1', B_1', che variano avvicinandosi rispettivamente ad A, B, A', B'; ma il punto P apparirà come limite di due punti P_1, P_2, appartenenti a due rami di f.

Due casi potranno presentarsi, cioè:

1) vi sono due rami di f ($A_1P_1B_1$, $A_1'P_2B_1'$) che si approssimano rispettivamente alle poligonali APB, $A'PB'$;

2) oppure vi sono due rami di f ($A_1P_1B_1'$, $B_1P_2A_1'$) che, si avvicinano rispettivamente ad APB', BPA'.

Nel primo caso l'iperbole osculatrice ad f nel punto P_1 ha un ramo vicino ad APB; nel secondo caso codesto ramo d'iperbole è vicino ad APB'.

Ora esistono iperbole vicine alla coppia di rette AP, BP, e di cui i rami si approssimano ad APB, $A'PB'$ come nel caso (1), o invece ad APB' BPA', come nel caso (2). Se ne deduce che esistono certe curve f tendenti a φ che si approssimano con un ramo ad APB come nel caso (1). Si può anzi supporre che questa condizione sia soddisfatta da f per riguardo a ciascun vertice della nostra poligonale di n lati, n essendo abbastanza grande: infatti la determinazione di una f dipende da $\frac{n(n+3)}{2}$ costanti arbitrarie, e l'assegnazione di n iperbole osculatrici in n punti, importa l'uso di sole $5n$ costanti, dove

$$5n < \frac{n(n+3)}{2}$$

se

$$n > 7.$$

Ciò posto si avrà una curva f, un ramo della quale (costituente un arco di curva continua) si approssima alla poligonale data. c· d. d.

Di qui segue il

Teor. *Una curva continua nell'intervallo* (a, b), *può essere approssimata quanto si vuole da una curva continua algebrica irriducibile.*

§ 38. **Teorema di Weierstrass.** — Sussiste un teorema più significativo, dovuto a WEIERSTRASS (1885): *Una funzione continua* y = f(x), *in un intervallo* (a, b), *può essere approssimata quanto si vuole da un polinomio.*

A prima vista parrebbe facile dedurre questo teorema da quello del paragrafo precedente. Invero se la curva continua, di cui ivi si discorre, è rappresentata da un'equazione della forma $y = \varphi(x)$, si può supporre che la curva algebrica che le si approssima sia pure incontrata in un sol punto dalle rette $x =$ cost appartenenti alla striscia $a \geq x \geq b$, escludendosi pure che le rette $x = a$, $x = b$ sieno tangenti. In tal caso dunque questa curva sarà rappresentata da una funzione algebrica

$$y = y(x)$$

che avrà la derivata y sempre finita fra a e b; e perciò la $y(x)$ sarà sviluppabile in serie di TAYLOR, procedente per le potenze di $(x - x_0)$, a partire da un qualsiasi punto x_0 fra a e b.

Se lo sviluppo anzidetto convergerà per tutti i valori di x fra a e b, arrestandosi al termine d'ordine n, per n assai alto, si otterrà un polinomio $y = f(x)$ prossimo quanto si vuole alla $y(x)$ e perciò alla funzione continua $y = \varphi(x)$, fra a e b.

Ma in generale la curva algebrica, che abbiamo costruita, non verrà rappresentata da un solo sviluppo di TAYLOR, bensì da un numero finito di serie di TAYLOR nell'intervallo (a, b). E si otterrà allora, non una sola parabola $y = f_n(x)$, approssimata alla curva data, ma una curva (approssimata) composta di un numero finito di parabole. Sarà possibile rimpiazzare questa curva composta, con una parabola unica, vicina ad essa quanto si vuole?

La dimostrazione di questo teorema, con considerazioni sintetiche di carattere algebrico non sembra facile [1].

Conviene perciò ricorrere a considerazioni analitiche di altra natura, riferendosi direttamente al fatto che una funzione continua qualsiasi $y = \varphi(x)$ può essere approssimata quanto si vuole da una poligonale (§ 37). La prima dimostra-

[1] La questione si riduce ad approssimare con una parabola unica d'ordine n, alto quanto si vuole, $y = f_n(x)$, la curva composta di due parabole d'un certo ordine m: $y = \varphi_m(x)$, $y = \psi_m(x)$, dove φ_m si considera per x fra 0 e 1, ψ_m per x fra 1 e p (> 1). e dove $\varphi_m(1) = \psi_m(1)$. Si presenta naturale l'idea di definire f_{2n+1}, imponendo un contatto d'ordine n di essa con φ_m e ψ_m, rispett. nei punti 0 e p. Ma il polinomio f_{2n+1} così definito diverge per $n = \infty$ come ho potuto verificare col calcolo di esso.

Invece si riesce ad approssimare la nostra curva composta ponendo

$$y = \left(\frac{\varphi + \psi}{2}\right) + \mu(x)\left(\frac{\varphi - \psi}{2}\right)$$

dove

$$\mu(x) = -1 \quad \text{per} \quad x < 1$$
$$\mu(x) = +1 \quad \text{per} \quad x > 1$$

($\mu(x)$ finita per $x = 1$); ma la ricerca di approssimare una tale funzione $\mu(x)$ ci riporta a quella per la funzione $\lambda(x)$ adoperata nel testo.

zione, data da WEIERSTRASS, si fonda sul calcolo dell'integrale

$$\int_{-\infty}^{+\infty} e^{-\psi^2}\, dt = \sqrt{\pi}.$$

LERCH e VOLTERRA hanno dedotto il teorema di WEIERSTRASS dallo sviluppo d'una funzione periodica in serie trigonometrica di FOURIER.

Infine, seguendo una via indicata da RUNGE, si è pervenuti a una dimostrazione più elementare che può svolgersi in diversi modi con MITTAG-LEFFLER e LEBESGUE [1].

Il punto di partenza di questa dimostrazione è la rappresentazione di una poligonale nel modo che segue.

Sieno (x_i, y_i) per $i = 0, 1, 2 \ldots.\ n+1$, $n+2$ punti del piano, vertici di una poligonale, e suppongasi

$$x_0 < x_1 < x_2 < \ldots. < x_n < x_{n+1}.$$

Il lato della poligonale che unisce i vertici (x_{i-1}, y_{i-1}), (x_i, y_i) ha per equazione

$$y = f_i(x) = y_{i-1} + \frac{x - x_{i-1}}{x - x_i}(y_i - y_{i-1}).$$

Ciò posto la poligonale viene rappresentata dalla funzione continua

$$y = f_1(x) + \sum_{i=1}^{i=n} (f_{i+1}(x) - f_1(x))\lambda(x - x_i)$$

dove $\lambda(x)$ designa una funzione tale che

$$\begin{aligned} \lambda(x) &= 0 \quad \text{per} \quad x < 0 \\ \lambda(x) &= 1 \quad \text{per} \quad x > 0 \end{aligned}$$

($\lambda(0)$ avendo un valore finito).

Secondo MITTAG-LEFFLER la funzione $\lambda(x)$ può essere definita ponendo

$$\lambda(x) = \lim_{n=\infty} \{ 1 - 2^{-(1+x)n} \}.$$

[1] Per un resoconto su queste varie dimostrazioni e per considerazioni critiche concernenti il teorema cfr. E. BOREL, « Leçons sur les fonctions de variables réelles », Paris, Gauthier-Villars, 1905, Cap. IV.

qualora però si sia preventivamente ridotto l'intervallo (x_0, x_{n+1}) ad essere in lunghezza minore di 1.

Ma poichè l'esponenziale è sviluppabile in serie di potenze per qualsiasi valore dell'esponente, si può rimpiazzare $\lambda(x)$ con un polinomio approssimato ad essa quanto si vuole. Basterà infatti prendere n così grande che sia $1-2^{-(1+x)^n}$ approssimato a $\lambda(x)$, più di $\frac{\varepsilon}{2}$ entro il tratto in cui occorre considerare questa funzione; quindi prendere tanti termini dello sviluppo di $M_n(x)=1-2^{-(1+x)^n}$ in modo da avere un polinomio $P^n(x)$ approssimato a $M_n(x)$, anch'esso più di $\frac{\varepsilon}{2}$; il poligono suddetto sarà approssimato a $\lambda(x)$ più di ε.

Segue di qui il teorema di WEIERSTRASS enunciato in principio.

NOTA. Si può definire — in diversi modi — una successione di polinomi $P_n(x)$ d'ordine $n=1, 2, 3 \ldots.$, che si approssimino ad una funzione continua data $y=\varphi(x)$, e rappresentare quindi $\varphi(x)$ colla serie convergente:

$$\varphi(x)=P_1+(P_2-P_1)+\ldots.+(P_n-P_{n-1})+\ldots..$$

La più rapida convergenza della serie si ottiene coi polinomi di TCHEBICHEFF [1]. Per ogni polinomio d'ordine n, $x=f_n(x)$, il valore assoluto della differenza

$$|\varphi(\varphi)-f_n(x)|$$

ammette un massimo

$$M=M(a_1 a_2 \ldots.)$$

funzione dei coefficienti di f_n (cfr. § 10). Ora per ogni n esiste *un poligono determinato* che rende minimo M; questo è il polinomio P_n più approssimato a $\varphi(x)$, la cui unicità fu stabilita appunto da TCHEBICHEFF.

§ 32. **Funzioni di linee: area, lunghezza.** — *Def.* Si abbia una famiglia di linee

$$y=\varphi(x)$$

[1] Bollettino della Società fisico-matematica dell'Accademia di Pietroburgo, 1858. Cfr. BOREL, l. c., p. 82.

(x variabile fra a, b) comprese in una striscia (ψ, θ):

$$\psi(x) \leq \varphi(x) \leq \theta(x);$$

si dirà che f è *funzione delle linee* φ, appartenenti alla striscia suddetta, se — per ogni φ — è definito un valore di f:

$$f = f \left\{ \overset{b}{\underset{a}{\varphi(x)}} \right\}.$$

Diamo alcuni esempi: *L' area* racchiusa fra la linea $\varphi(x)$, le perpendicolari nei suoi estremi (a, b) all'asse x e questo asse, è una funzione della linea:

$$f = \int_a^b \varphi(x)dx.$$

L'area suddetta è notoriamente definita per tutte le funzioni continue φ, ed anche per funzioni *generalmente continue*; infatti adottando la definizione dell'integrale di RIEMANN, l'integrale suddetto riesce determinato per le funzioni aventi punti di discontinuità che costituiscono un insieme di *misura nulla*, cioè racchiudibile con un numero finito o infinito di tratti la cui somma converga a zero. (Cfr. LEBESGUE « Leçons sur l'intégration » Paris, 1904, p. 29).

Inoltre LEBESGUE (« Integrale, longueur, aire » Annali di Mat. 1902) ha generalizzato le nozione d'integrale e di area, in modo che f risulta definito per ogni funzione φ limitata e tale che si possa definire la *misura* della somma degl'intervalli entro cui φ ha una data oscillazione ed anche la misura del gruppo di punti in cui essa prende un valore qualsiasi assegnato.

La *lunghezza* della linea φ fra gli estremi a, b si esprime coll'integrale

$$f = \int_a^b \sqrt{1 + \varphi'(x)^2}\, dx,$$

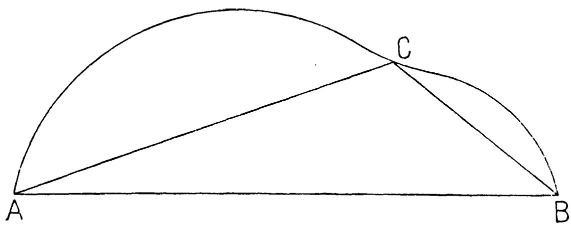
il quale risulta definito nell'ipotesi che esista una derivata $\varphi'(x)$ generalmente continua ed anche in altri casi (1). La

(1) Cfr. L. TONELLI « Sulla rettificazione delle curve ». Atti Accad., Torino, 1908, (v.i anche ibidem 1912).

lunghezza f si può dunque considerare come funzione delle linee φ soddisfacenti alla condizione accennata.

Ma *il concetto di lunghezza d'una linea continua si può porre per mezzo di una definizione geometrica più generale* (1).

Si divida l'intervallo (a, b) in un numero finito n di tratti successivi $\varepsilon_1, \varepsilon_2, \dots \varepsilon_n$, tendenti a zero col crescere di n (teorema



di Heine, § 4) e si consideri la poligonale iscritta nella linea $y = \varphi(x)$, i cui lati congiungono i punti successivi della serie

$$a, \quad a + \varepsilon_1, \quad a + \varepsilon_1 + \varepsilon_2, \dots, \quad b - \varepsilon_n, \quad b.$$

Se la lunghezza della poligonale tende ad un limite finito per $n = \infty$ $(\varepsilon_i = 0)$, questo si dice *lunghezza della curva* φ fra a e b.

Qui si può ricordare il *teorema di* L. Scheeffer (2):

Se le lunghezze delle *poligonali* iscritte nella linea continua φ rimangono tutte *inferiori ad un numero fisso,* esse tendono ad uno stesso limite finito, cioè *la curva è rettificabile.* Se una poligonale iscritta in φ, al crescere del numero dei lati, può superare un numero grande ad arbitrio, ogni altra poligonale iscritta i cui lati tendono a zero, diventa pure superiore ad ogni limite.

In altre parole le *curve continue non rettificabili* sono da considerare come linee di *lunghezza infinita.*

Noteremo infine l'importante *teorema* di Lebesgue (3).

Ogni funzione continua (curva) *rettificabile ammette la*

(1) Le condizioni di *rettificabilità* per φ sono date da Jordan, « Cours d'Analyse ». Paris, 1893. Vol. I, nn. 105-111.

(2) « Allgemeine Untersuchungen über Rectification der Curven ». Acta Mathem. 5, (1884-85).

(3) « Leçons sur l'intégration ». Paris, 1904.

derivata in tutti i punti, eccezion fatta, al più, da un insieme di punti di misura nulla.

NOTA. Data la precedente definizione generale della lunghezza d'una linea, si ha senz'altro che la lunghezza d'una linea AB vale

$$l \geq AB.$$

Ma se c'è un punto C della linea fuori del segmento AB segue

$$l \geq AC + CB > AB.$$

Così si conclude che *la retta è linea di lunghezza minima.*

§ 40. **Continuità.** — Una funzione delle linee φ — in un dato intervallo — dicesi continua per $\bar{\varphi} = \varphi$, se — essendo $\bar{\varphi}$ linea limite delle φ — per ogni numero positivo ε piccolo ad arbitrio si può determinare un δ tale che, per tutte le φ soddisfacenti alla diseguaglianza

$$|\varphi - \bar{\varphi}| < \delta,$$

si abbia

$$|f(\varphi) - f(\bar{\varphi})| < \varepsilon,$$

cioè

$$\lim_{\varphi = \bar{\psi}} f(\varphi) = f(\bar{\varphi}).$$

Teor. L' *area*

$$\int_a^b \varphi(x) dx$$

è *funzione continua* della linea φ.

Infatti se si ha — in valore assoluto —

$$|\varphi - \bar{\varphi}| < \varepsilon,$$

segue

$$|f\{\varphi\} - f\{\bar{\varphi}\}| < \int_b^a \varepsilon dx$$

cioè

$$|f\{\varphi\} - f\{\bar{\varphi}\}| < \varepsilon(b - a).$$

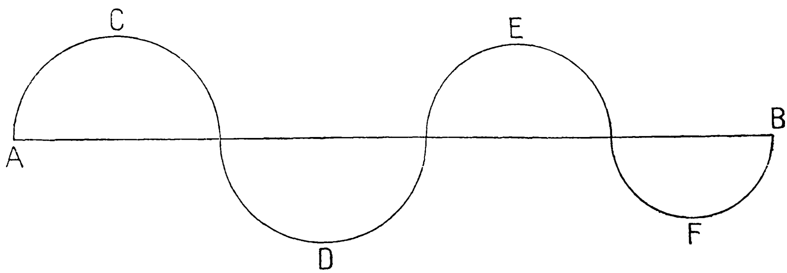
Teor. La lunghezza della linea φ non è funzione continua della linea stessa.

Infatti si consideri il segmento rettilineo AB e la linea $ACDEFB$ composta di n semicerchi alternativamente al di sopra e al disotto della retta AB; la lunghezza di questa linea — qualunque sia n — è πAB; ma per $n = \infty$ la linea suddetta ha per limite il segmento AB.

Anche dall'espressione analitica della lunghezza

$$f = \int_a^b \sqrt{1 + \varphi'^2} . dx ,$$

risulta la non continuità di φ, perchè al variare di φ approssimantesi alla curva limite $\bar{\varphi}$, non accade in generale che la



derivata φ' tenda a $\bar{\varphi}'$ (anche a prescindere dai casi in cui la derivata venga a mancare).

Nota. Se si suppone data la linea continua rettificabile $y = \varphi(x)$, essa può essere approssimata con una poligonale la cui lunghezza tende — per definizione — alla lunghezza l di φ. Ora, anche una curva continua algebrica $y = \psi(x)$ prossima alla poligonale — costruita come al § 37 — ha una lunghezza tendente ad l. Approssimiamo la curva algebrica

$$y = \psi(x)$$

con un polinomio

$$y = f_n(x) \qquad (\S\ 38);$$

allora il *polinomio*

$$F_{n+1}(x) = \int_a^x f_n(x) + \psi(a)$$

si *approssimerà* a $\psi(x)$ e quindi *a* $\varphi(x)$, *in modo che la lunghezza della parabola* F_{n+1} *tenderà alla lunghezza di* φ (osservazione di Painlevè).

Si può anche dimostrare che la convergenza della lunghezza alla lunghezza, sussiste per i polinomi $y = f_n(x)$ approssimanti una poligonale, quando essi vengano costruiti col metodo di MITTAG-LEFFLER (§ 38). Basterà dimostrare che, in ogni intervallo fra due vertici della poligonale, la derivata $f_n'(x)$ tende ad un valore costante, uguale al coefficiente angolare del lato della poligonale stessa; invero da ciò risulta che la lunghezza dell'arco di parabola fra i due estremi corrispondenti ai vertici sopra nominati, tende alla lunghezza del lato.

Ora la dimostrazione richiesta si riduce alla verifica che — per n assai elevato — mentre la funzione

$$M_n(x) = 1 - 2^{-(1+x)^n}$$

si approssima a

$$\lambda(x) \begin{cases} = 0 & \text{per} \quad x < 0 \\ = 1 & \text{per} \quad x > 0, \end{cases}$$

anche la derivata

$$M_n'(x)$$

si approssima a

$$\lambda'(x) = 0,$$

e quindi la derivata del polinomio $P_n(x)$, costituito da un numero assai grande di termini dello sviluppo di TAYLOR di $M_n(x)$, si approssima pure a 0.

Calcoliamo dunque

$$M_n'(x) = n \log_e 2 \cdot (1 + x)^{n-1} 2^{-(1+x)^n}$$

appare tosto che per $x > 0$

$$\lim_{n=\infty} M_n'(x) = 0$$

crescendo $2^{(1+x)^n}$ più rapidamente che $n(1 + x)^{n-1}$; parimente per $x < 0$ si ha

$$\lim_{n=\infty} M_n'(x) = \log_e 2 \lim_{n=\infty} n(1 + x)^{n-1} = 0,$$

per essere (v. § 38)

$$|x| < 1.$$

§ 41. **Semicontinuità.** — Dove manca la continuità supplisce spesso — nella teoria delle funzioni di linee — la semicontinuità, di cui diamo la seguente

Def. Una funzione f delle linee φ, in un dato intervallo, si dice godere per $\varphi = \bar{\varphi}$ della *semicontinuità inferiore* (o *superiore*) secondo la denominazione di Baire (« Leçons sur les fonctions discontinues ») se — essendo $\bar{\varphi}$ linea limite delle φ — per ogni numero positivo ε piccolo ad arbitrio, si può determinare un δ tale che, per tutte le φ soddisfacenti alla diseguaglianza

$$|\varphi - \bar{\varphi}| < \delta,$$

si abbia

$$f\{\varphi\} \geq f\{\bar{\varphi}\} - \varepsilon$$

(o rispettivamente

$$f\{\varphi\} \leq f\{\bar{\varphi}\} + \varepsilon).$$

Esempio. Se si ha un insieme di linee rettificabili avente una linea limite $\bar{\varphi}$ di lunghezza l, *la lunghezza di φ gode della semicontinuità inferiore,* cioè dato un ε piccolo ad arbitrio, essa si mantiene sempre superiore ad $l - \varepsilon$ per tutte le φ assai vicine a $\bar{\varphi}$.

Per dimostrare questo teorema procederemo come segue.

Essendo l la lunghezza di $\bar{\varphi}$, si può costruire una poligonale iscritta nella linea ed avente una lunghezza $l - \varepsilon$ dove ε è piccolo ad arbitrio (§ 39). La poligonale suddetta avrà n lati tutti maggiori di un segmento k; i vertici della poligonale corrisponderanno a punti che designeremo, secondo i valori crescenti delle ascisse, con $x_1, x_2, \ldots.$ Ora si determinerà una curva φ assai vicina alla curva limite $\bar{\varphi}$, in guisa che si abbia, per tutti i punti x dell'intervallo che si considera, in valore assoluto

$$|\varphi(x) - \bar{\varphi}(x)| < \delta;$$

in particolare sarà

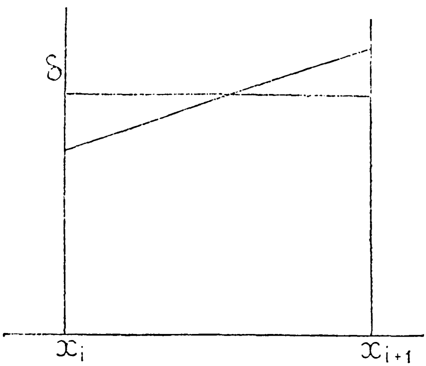
$$|\varphi(x_i) - \bar{\varphi}(x_i)| < \delta.$$

I punti $\varphi(x_i)$ sono vertici d'una poligonale iscritta nella linea φ. La lunghezza del lato che congiunge i punti i^{mo} e $(i+1)^{mo}$ si potrà considerare come lato di un triangolo che ha gli altri due lati rispett. $< k$ e $< 2\delta$; pertanto quella lunghezza sarà

$$< k + 2\delta,$$

ove si può supporre che δ sia piccolo quanto si vuole rispetto a k.

La possibilità di rendere δ piccolo ad arbitrio rispetto a k, per es. $\delta < k^2$, porta la conseguenza che esiste una poligonale, iscritta nella linea φ, la cui lunghezza si avvicina ad $l - \varepsilon$ e quindi ad l quanto si vuole.



La lunghezza di φ è certo maggiore di quella della poligonale suddetta e così resta dimostrato il teorema della semi-continuità.

NOTA. La dimostrazione si estende facilmente al caso in cui si abbiano linee, rappresentate da equazioni parametriche, che vengano incontrate in più d'un punto dalle rette $x =$ cost.

§ 42. **Esistenza delle linee di lunghezza minima.** — Come nella teoria delle funzioni, dipendenti da una o più variabili, anche per le funzioni di linee si hanno alcuni teoremi generali sull'esistenza dei massimi e minimi. Ci riferiremo qui a due casi particolari che sono in qualche modo tipici per più vaste classi di problemi, cioè al *problema delle linee di lunghezza minima* e al *problema degli isoperimetri*.

Poniamo — per semplicità — di essere nel piano. Il problema delle linee di lunghezza minima consiste in questo:

Dati due punti A, B, appartenenti ad una superficie piana — che costituisca un *campo finito* e *continuo* (§ 7) — si considerino tutte le linee rettificabili di estremi A, B, entro la superficie suddetta; si tratta di provare che fra queste *esiste una linea di lunghezza minima.*

La dimostrazione del teorema si svolge come segue:

Anzitutto le lunghezze di tutte le linee AB comportano un limite inferiore l; si tratta di provare che questo è un minimo effettivamente raggiunto. A tale scopo si considerino tutte le linee AB la cui lunghezza non supera $l+\varepsilon$, ε essendo un numero positivo piccolo ad arbitrio; si considerino poi le linee AB la cui lunghezza non supera

$$l+\frac{\varepsilon}{2},\quad l+\frac{\varepsilon}{3},\ldots .$$

Si avrà in tal modo una successione d'insiemi di linee

$$I_1, I_2, I_3, \ldots,$$

ciascuno contenuto nel precedente. Vi sarà quindi (§ 36) una linea limite a cui si approssimano linee di I_n, per n grande quanto si vuole. Il ragionamento stesso che prova la semicontinuità della lunghezza, vale a provare che codesta linea limite è rettificabile ed ha una lunghezza $< l+\frac{\varepsilon}{n}$, cioè la lunghezza minima l.

Questa dimostrazione si estende dal piano alle superficie; si ha pertanto il seguente:

Teorema d'esistenza di HILBERT:

Sia

$$x=x(u,v) \quad y=y(u,v) \quad z=z(u,v)$$

una superficie rappresentata da funzioni continue e derivabili in un dato campo continuo; fra tutte le linee che congiungono due punti della superficie, esiste sempre una linea di lunghezza minima.

Giova riportare in breve la semplicissima

Costruzione di HILBERT (¹). Essa si basa sul presupposto che si possa determinare una successione di linee $C_1, C_2, \ldots$, cogli estremi fissi PP', le cui lunghezze $l_1, l_2, \ldots$, tendono al limite inferiore l.

Si supponga dunque data una tale successione.

A partire da P si stacchi sopra C_i il punto P_i tale che l'arco

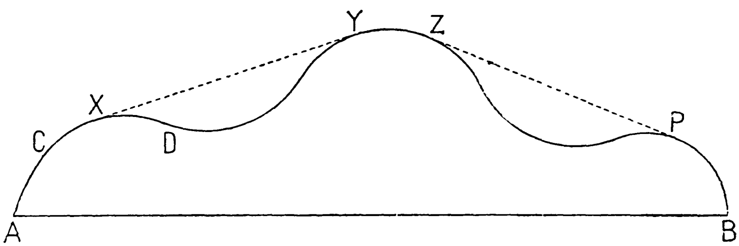
$$PP_i=\frac{m}{n}\, l_i.$$

(¹) L. c. Cfr. CARATHEODORY, « Dissertation », Göttingen, 1904. Göttinger Nachrichten, 1905.

Si dimostra che i punti P_i — per $i = 1, 2 \ldots.$ — tendono ad un punto limite determinato $P^{\left(\frac{m}{n}\right)}$. I punti limiti $P^{\left(\frac{m}{n}\right)}$ appartengono ad una determinata curva continua, limite delle C_i, che è rettificabile e di lunghezza l.

Nota. La dimostrazione, da noi esposta, suppone che le curve che si accostano alla lunghezza minima sieno rappresentate da funzioni derivabili. Non è questa una restrizione essenziale, giacchè in ogni caso vi sono prossime quanto si vuole a una funzione data, funzioni algebriche derivabili (§ 37).

§ 43. **Curva convessa minima comprendente una data superficie piana.** — Possiamo fare un' elegante applicazione del



teorema precedente, nel caso di una superficie piana, dimostrando rigorosamente un principio di cui si vale Steiner nelle sue ricerche sugl' isoperimetri (Art. XXVI).

Si abbia una curva continua $y = f(x)$, (in un intervallo finito) giacente nel semipiano superiore all' asse x ed avente i suoi estremi sull' asse anzidetto. Questa curva divide il semipiano in due parti (inferiore e superiore) costituite rispett. dai punti per cui

$$y \leq f(x) \quad \text{e} \quad y \geq f(x);$$

la prima parte è la superficie compresa dalla curva f (insieme alla retta AB).

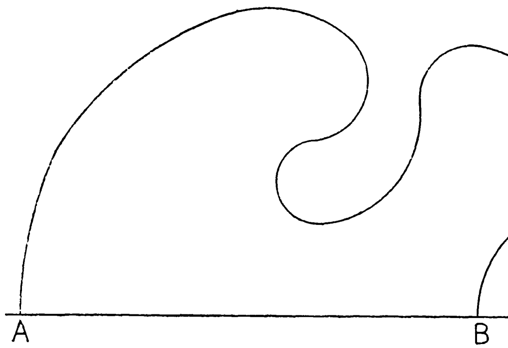
Un arco CD della curva si dice *convesso* se tutti i segmenti congiungenti due punti dell' arco appartengono alla superficie compresa dalla curva.

Ora:

Esiste una linea convessa di lunghezza minima che ha per estremi A, B, *e comprende la superficie compresa da* f.

La linea convessa anzidetta (composta di archi della curva f e di tratti rettilinei) è la linea di lunghezza minima che congiunge A, B, *entro la parte di piano superiore ad* f; perciò la sua esistenza segue dal teorema del § precedente.

Nota. Il principio qui stabilito vale indipendentemente dall'ipotesi che la curva continua cogli estremi A, B, gia-



cente nel semipiano superiore alla retta AB, sia intersecata in un sol punto dalle perpendicolari ad AB e perciò possa rappresentarsi con un'equazione unica

$$y = f(x).$$

Infatti la dimostrazione del principio anzidetto si basa sulla *proprietà della curva continua chiusa senza punti doppi*, formata dalla curva f e dal segmento AB, *di delimitare una superficie*, e Jordan [1] ha dimostrato che questa proprietà appartiene in generale alle curve chiuse

$$(1) \qquad \begin{cases} x = \varphi_1(t) \\ y = \varphi_2(t) \end{cases}$$

dove φ_1 e φ_2 sono funzioni continue e le equazioni (1) sono invertibili:

$$t = \text{funz.}\,(x, y).$$

§ 44. **Esistenza d'una curva d'area massima in una famiglia d'isoperimetre.** — Si considerino ora nel piano, e precisamente nel semipiano superiore ad una retta data x,

[1] Cfr. C. Jordan, « Cours d'Analyse ». Gauthier-Villars. Paris, 1893, (t. I, n. 95); cfr. Art. V, nel Vol. I (§ 9).

tutte le linee rettificabili, di lunghezza $l(>b-a)$, aventi gli estremi fissi a, b; *fra queste linee isoperimetre ne esiste una che racchiude area massima coll'asse* x.

Il teorema si dimostra come segue:

Anzitutto vi è un limite superiore A per le aree racchiuse dalle linee isoperimetre suddette, giacchè tutte codeste aree sono certo inferiori a quella del cerchio di raggio l.

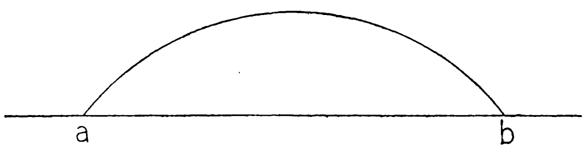
Ora vi saranno certo, nella famiglia suddetta, linee per cui l'area racchiusa risulta

$$\geq A-\varepsilon, \quad \geq A-\frac{\varepsilon}{2}, \quad \geq A-\frac{\varepsilon}{3}, \ldots,$$

designando ε un numero positivo dato. Si otterrà quindi una successione d'insiemi di linee

$$I_1, I_2, \ldots,$$

ciascuno contenuto nel precedente In forza del teorema di Hilbert vi sarà una linea limite a cui si approssimano le



linee di I_n, per n grande quanto si vuole. Si vede facilmente che codesta linea limite è rettificabile ed ha la lunghezza l e che l'area da essa racchiusa coll'asse x è $> A-\frac{\varepsilon}{n}$ dove n è grande quanto si vuole, e perciò ha precisamente il valore massimo A.

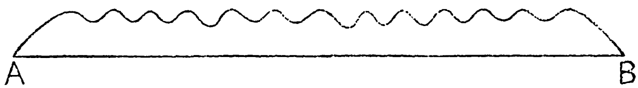
Nota. La dimostrazione precedente — per ciò che riguarda la rettificabilità della linea limite — si riduce semplicissima se, approfittando dell'osservazione di Steiner giustificata nel paragrafo precedente, si opera esclusivamente sopra linee convesse, convessa resultando quindi, *a priori*, anche la linea limite.

§ 45. **Funzioni di linee prive di massimo o minimo.** — L'esistenza del massimo o del minimo d'una funzione di linee, a cui si riferiscono i teoremi dei numeri precedenti, fu ritenuta in generale evidente, e come tale presupposta nelle dimostrazioni, fino dai geometri antichi (cfr. Art. XXVI). Nei tempi moderni STEINER ha edificato su tale presupposto la sua bella trattazione degl'isoperimetri (ibidem); RIEMANN ha stabilito sulla stessa base la dimostrazione del cosidetto *principio di* DIRICHLET, fondamento della sua grandiosa teoria delle funzioni abeliane.

Pare (1) che per primo DIRICHLET avvertisse la lacuna delle dimostrazioni di STEINER. Per lo stesso motivo WEIERSTRASS oppugnò la dimostrazione accennata del principio di DIRICHLET, portando l'esempio (2) dell'integrale

$$I = \int_{-1}^{+1} x^2 y'^2 dx,$$

che non può rendersi minimo in corrispondenza ad una curva continua $y = y(x)$ passante per i punti $(-1, a)$ $(+1, b)$. dove



$a \neq b$. Si dimostra infatti che il *limite inferiore* di I è $= 0$; questo limite si raggiunge solo per $y' = 0$, $y = \text{cost}$, ed è quindi irraggiungibile da una funzione continua che soddisfi alle date condizioni ai limiti.

Un *esempio elementare in cui manca il minimo* può darsi già nella teoria degli isoperimetri.

Si cerchi di determinare una curva continua di data lunghezza $l \geq AB$ e di estremi A e B, che giaccia nel semipiano superiore alla retta AB, e renda minima l'area compresa con questa retta. *Se* $l > AB$, *la curva che dà il minimo non esiste;* l'area si può render piccola quanto si vuole con una curva che si approssima alla retta AB (§ 41), ma non può mai essere nulla fino a che si conservi la definizione data di linea, che *esclude* la possibilità di *tratti di curva contati più volte.*

È chiaro come il precedente problema si trasformi facilmente in un problema di massimo. Si tratti p. es., di trovare

(1) Cfr. LAMPE, Bibliotheca Math. (3) 1, p. 134.
(2) WERKE, Bd II, p. 49.

una curva di data lunghezza l, avente gli estremi A, B e giacente in un quadrato di lato AB, la quale debba render massima l'area compresa con AB.

Il *massimo non esiste per*

$$l > 3AB.$$

Nelle condizioni in cui si pone STEINER, il presupposto del massimo è giustificato dal teorema del § 44. E così — colmata la lacuna — *la teoria di* STEINER *ha la sua piena giustificazione* (cfr. Art. XXVI) (1).

II.

Su alcuni problemi del calcolo delle variazioni.

§ 46. **Introduzione.** — La ricerca delle linee (superficie ecc.) per cui una data funzione riesce massima o minima, si fa per mezzo del *calcolo delle variazioni* (cfr. Nota storica, § 63).

La funzione della linea $y = f(x)$ essendo espressa da un integrale, il calcolo delle variazioni insegna a dedurre un'equazione differenziale a cui la curva $y = f(x)$ deve soddisfare, affinchè essa possa rispondere al massimo o al minimo richiesto. Quest'equazione differenziale, che porta il nome di EULERO, si ottiene nel modo più spedito col procedimento di LAGRANGE.

Ma la *condizione d'* EULERO costituisce soltanto una *condizione necessaria*, non sufficiente, per il massimo o il minimo. Le linee che vi soddisfano (indipendentemente dal rispondere o meno al massimo o al minimo) si chiamano (con KNESER) *estremali* del problema proposto; diversi criterii permettono di riconoscere *se* — e in quali limiti — *le estremali* godano *effettivamente* la proprietà di *massimo* o *minimo*.

Noi prenderemo le mosse dal problema elementare di « determinare, nel piano, la linea più breve fra due punti ». E tratteremo quindi il problema delle geodetiche sopra una

(1) Sull'esistenza dei massimi e minimi delle funzioni semicontinue si ha un teorema generale di L. TONELLI, « Sul caso regolare nel calcolo delle variazioni ». Circolo Mat. di Palermo, 1913.

superficie, che costituisce insomma il tipo d'una classe generale di problemi.

Quindi daremo un'idea dei problemi detti isoperimetrici, trattando appunto il problema elementare che ha dato nome alla classe: « determinare la linea di data lunghezza e con dati estremi, che racchiude l'area massima colla retta congiungente i due punti ».

§ 47. **Linee di lunghezza minima: equazione d'Eulero.** — Si tratti di determinare la linea di lunghezza minima fra quelle che congiungono — nel piano — due punti A, B.

Supponiamo che esista una linea di lunghezza minima e che questa sia decomponibile in un numero finito di archi intersecati in un sol punto dalle parallele all'asse y, comprese in una certa striscia. In questa ipotesi si deve dunque determinare la funzione $y=f(x)$ che prende dati valori y_a, y_b per $x=a$, $x=b$, e tale che la lunghezza

$$\int_a^b \sqrt{1+f'^2}\,dx$$

riesca minima.

Si consideri un fascio di linee

$$y=f(x)+t\omega(x),$$

a cui appartiene la f.

Si avrà

$$L(t)=\int_a^b \sqrt{1+(f+t\omega)'^2}\,dx,$$

e poichè $L(t)$ ha un minimo per $t=0$:

$$\frac{dL}{dt}=0.$$

Il differenziale di L, riguardato come funzione del parametro t, è appunto ciò che si chiama la *variazione* δL, sicchè può dirsi che il minimo richiesto corrisponde necessariamente all'annullarsi della variazione

$$\delta L=\left(\frac{dL}{dt}\right)_0 dt=0.$$

Ora si ha:

$$\frac{dL}{dt} = \int_a^b \frac{d}{dt} \{ \sqrt{1 + (f + t\omega)'^2} \} \, dx = \int_a^b \frac{(f' + t\omega')\omega'}{\sqrt{1 + (f' + t\omega')^2}} . \, dx,$$

e qui viene

$$\left(\frac{dL}{dt}\right)_{t=0} = \int_a^b \frac{f' \, \omega'}{\sqrt{1 + f'^2}} . \, dx = 0.$$

Integrando per parti, ove si assuma come fattor differenziale $\omega' dx$, si ottiene

$$\left(\frac{dL}{dt}\right)_{t=0} = \left[\frac{f' . \, \omega}{\sqrt{1 + f'^2}}\right]_a^b = \int_a^b \omega \frac{f'' \sqrt{1 + f'^2} - \dfrac{f'^2 f''}{\sqrt{1 + f'^2}}}{(1 + f'^2)} dx,$$

ed essendo, per ipotesi,

$$\omega(a) = \omega(b) = 0,$$

$$\left(\frac{dL}{dt}\right)_{t=0} = - \int_a^b \omega \frac{f''}{(1 + f'^2)^{\frac{3}{2}}} \, dx = 0.$$

Questa equazione dovendo sussistere *qualunque sia* $\omega(x)$, segue

$$\frac{f''}{(1 + f'^2)^{\frac{3}{2}}} = 0,$$

cioè

$$f''(x) = 0,$$

che è l'*equazione differenziale delle rette*

$$y = px + q.$$

Le due costanti arbitrarie si determinano in modo unico dalla condizione

$$y_a = pa + q$$
$$y_b = pb + q.$$

Nota I. Facendo uso del simbolo δ di Lagrange, il calcolo precedente si può ripetere senza designare esplicitamente la funzione $\omega(x)$, trattando il δ come differenziale rispetto ad una variabile ausiliaria t, che pure non viene designata.

Si avrà

$$\delta L = \int_a^b \delta\sqrt{1+f'^2}.\,dx = \int_a^b \frac{1}{2}\,\frac{\delta(1+f'^2)}{\sqrt{1+f'^2}}\,dx = \int_a^b \frac{f'\delta f'\,dx}{\sqrt{1+f'^2}},$$

e, posto

$$\delta f' dx = \delta df = d\delta f,$$

integrando per parti,

$$\delta L = \left[\frac{f'\delta f}{\sqrt{1+f'^2}}\right]_a^b - \int_a^b \delta f\,\frac{d}{dx}\left(\frac{f'}{\sqrt{1+f'^2}}\right),$$

L'espressione fra parentesi è nulla identicamente perchè

$$(\delta f)_a = (\delta f)_b = 0.$$

L'annullarsi del 2° termine *qualunque sia* δf, porta

$$\frac{d}{dx}\,\frac{f'}{\sqrt{1+f'^2}} = 0,$$

cioè l'equazione già ottenuta

$$f'' = 0.$$

NOTA II. Il procedimento svolto innanzi si può ripetere, indipendentemente dal presupposto che la linea da determinare sia intersecata in un punto dalle rette $x =$ cost, riferendosi alle equazioni parametriche

$$\begin{cases} x = \varphi_1(t) \\ y = \varphi_2(t), \end{cases}$$

dove il parametro t può essere scelto poi convenientemente, ad es. come lunghezza dell'arco di linea corrispondente al punto (x, y). Si trovano allora (cfr. § 51) le equazioni differenziali

$$\varphi_1'' = 0, \quad \varphi_2'' = 0,$$

che danno

$$x = p_1 t + q_1$$
$$y = p_2 t + q_2;$$

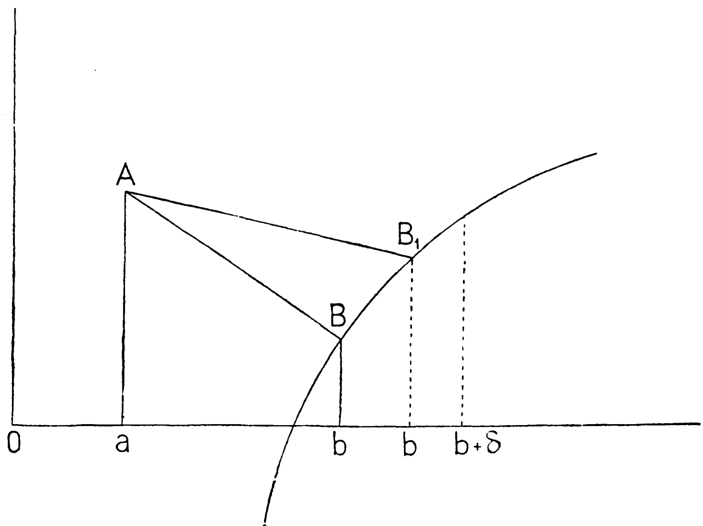
e così risulta che la *linea estremale,* per cui la variazione

$$\delta L = 0,$$

può essere rappresentata in generale (cioè escluso il caso $p_1 = 0$, $x =$ cost) coll'equazione precedente del tipo

$$y = px + q.$$

Nota III. Per la deduzione precedente è essenziale che il campo di variazione di x, y, sia tale che la *linea* minima



$y = f(x)$ *non faccia parte del contorno* o non contenga tratti del contorno; nel qual caso la linea variata,

$$y = f(x) + t\omega(x),$$

subirebbe limitazioni che non permetterebbero più di dedurre $f'' = 0$. A tale proposito vedasi la costruzione della minima linea comprendente una linea non convessa (§ 43).

§ 48. **Condizioni ai limiti.** — Possiamo generalizzare il problema trattato innanzi, proponendoci di « determinare la linea di lunghezza minima fra quelle che hanno un estremo fisso A, e l'altro estremo B variabile sopra una linea $y = \varphi(x)$ ».

A tale scopo occorre riprendere il calcolo precedente, tenendo conto della circostanza che il limite superiore dell'integrale L non è più fisso.

Riferendoci alla figura, prolunghiamo $f(x)$ e $\omega(x)$ al di là di b fino a $b+\delta$, in modo che sieno continue insieme alle loro derivate e che $\omega(x)$ sia diversa da 0 in $(b, b+\delta)$.

Consideriamo una curva variata

$$y=f(x)+t\omega(x)$$

che va da A a B_1. L'ascissa b_1 di B_1 è data dall'equazione

$$f(b_1)+t\omega(b_1)=\varphi(b_1);$$

essa varia al variare di t, e può quindi riguardarsi come funzione di questo parametro in un piccolo intervallo $(b-\delta, b+\delta)$.

La funzione inversa della $b_1(t)$ è

$$t=\frac{\varphi(b_1)-f(b_1)}{\omega(b_1)}$$

e perciò

$$\frac{db_1}{dt}=\frac{1}{\frac{dt}{db_1}}=\frac{\omega^2(b_1)}{\{\varphi'(b_1)-f'(b_1)\}\,\omega(b_1)-\{\varphi(b_1)-f(b_1)\}\,\omega'(b_1)}$$

Ciò posto calcoliamo la derivata

$$\frac{dL}{dt},$$

essendo

$$L(t)=\int_a^{b_1(t)}\sqrt{1+(f'+t\omega')^2}\,.\,dx.$$

Si ha

$$L(t)=\int_a^b\sqrt{1+(f'+t\omega')^2}\,.\,dx+\int_b^{b_1(t)}\sqrt{1+(f'+t\omega')^2}\,.\,dx,$$

e quindi

$$\frac{dL}{dt} = \int_a^b \frac{(f' + t\omega')\omega'}{\sqrt{1 + (f' + t\omega')^2}} \cdot dx +$$

$$+ \left\{ \int_b^{b_1(t)} \frac{(f' + t\omega')\omega'}{\sqrt{1+(f'+t\omega')^2}} dx + \sqrt{1+\{f'(b_1(t)) + t\,\omega'[b_1(t)]\}^2} \frac{db_1}{dt} \right\}.$$

Ora per $t = 0$ è

$$b_1(0) = b$$

e quindi sparisce lo $\int_b^{b_1}$, corrispondente ad un intervallo d'integrazione nullo.

Sostituiamo a $\frac{db_1}{dt}$ il suo valore, osservando che

$$\{ \varphi(b_1) - f(b_1) \} = 0$$

per $b_1 = b_1(0)$, avremo

$$\left(\frac{dL}{dt}\right)_0 = \int_a^b \frac{f'\,\omega'}{\sqrt{1+f'^2}} dx + \sqrt{1 + f'(b)^2} \cdot \frac{\omega(b)}{\varphi'(b) - f'(b)}\,.$$

Integrando per parti, come innanzi, si ha

$$\int_a^b \frac{f'\,\omega'}{\sqrt{1+f'^2}} dx = \left[\frac{f'\omega}{\sqrt{1+f'^2}}\right]_a^b - \int_a^b \omega \frac{f''}{(1+f'^2)^{\frac{3}{2}}} dx.$$

Affinchè sia

$$\left(\frac{dL}{dt}\right)_0 = 0$$

qualunque sia ω, dovrà essere intanto — come nel caso degli estremi fissi a, b —

$$\frac{f''}{(1+f'^2)^{\frac{3}{2}}} = 0,$$

cioè

$$f'' = 0.$$

Inoltre si ha l'equazione in b

$$\frac{f'(b)\omega(b)}{\sqrt{1+f'^2(b)}} + \frac{\sqrt{1+f'^2(b)}}{\varphi'(b) - f'(b)} . \omega(b) = 0,$$

cioè — essendo $\omega(b) \neq 0$ —

$$\{ \varphi'(b) - f'(b) \} f'(b) + (1 + f'(b)^2) = 0,$$

ossia

$$\varphi'(b) f'(b) + 1 = 0.$$

In conclusione si ha che la linea di lunghezza minima richiesta (nelle ipotesi adottate di continuità e derivabilità) deve soddisfare alle due condizioni seguenti:

1) deve essere linea minima fra i suoi due estremi considerati come fissi, cioè deve esser retta (§§ 39, 49);

2) deve essere *ortogonale* alla linea $y = \varphi(x)$ su cui l'estremo b può variare (non cadendo b in un estremo di φ); la condizione d'ortogonalità viene appunto significata dall'equazione

$$\varphi'(b) f'(b) + 1 = 0.$$

Nota. Giova osservare che in sostanza, nell'ipotesi di un estremo variabile, la ricerca del minimo si riduce a due problemi analoghi: uno dei quali si riferisce al caso di estremi fissi e l'altro è un problema di minimo per una funzione di una variabile. Questa riduzione è evidente *a priori*, perchè la linea minima deve esser tale anche se si aggiunge il legame che gli estremi restino fissi, e una volta che si supponga determinata questa linea minima, la sua lunghezza diventa una funzione dell'estremo variabile, la quale funzione appunto si vuole render minima.

§ 49. Dimostrazione che la retta è linea di lunghezza minima. — Ritorniamo al problema fondamentale di Geometria piana in cui si ricerca la linea di lunghezza minima, congiungente due punti fissi.

Lo sviluppo del § 47 prova che « se una linea minima esiste, e se essa può essere rappresentata da un'equazione

$$y = f(x)$$

dove f è continua e derivabile, deve essere soddisfatta l'equazione differenziale (delle rette)

$$f'' = 0 \text{ ».}$$

Le linee, *estremali*, che soddisfano a questa equazione differenziale, cioè le rette, porgeranno effettivamente il minimo richiesto?

La dimostrazione di questa proprietà non risulta affatto dal procedimento seguito, il quale — ove non venga completato con altre considerazioni — ha soltanto il valore di un *metodo di scoperta,* che ci addita la retta come possibile linea di lunghezza minima, e non di un *metodo di dimostrazione.* All'opposto il ragionamento elementare della Prop. XX di EUCLIDE, porge una semplicissima dimostrazione della proprietà di minimo della retta (§ 39), ma non ci dice in qual modo essa potrebbe essere scoperta e non ci addita quindi la via da tenere in casi analoghi.

Ora, in vista di successive generalizzazioni, giova prendere in esame la questione « se il procedimento del calcolo delle variazioni sia suscettibile di condurre ad una dimostrazione effettiva della proprietà di minimo, per cui fu posto in opera ».

A tale scopo giova richiamare il teorema d'esistenza delle linee di lunghezza minima (§ 42). Sapendo che queste linee esistono e sapendo che esse debbono soddisfare alla condizione necessaria d'EULERO (equazione differenziale) che caratterizza le estremali; sapendo inoltre che due punti determinano un'estremale (retta) che li congiunge, sembra si possa affermare che questa estremale (unica) porge effettivamente il minimo richiesto.

Ma tale conclusione non è ancora legittima; essa urta contro la seguente obiezione:

Le linee che dànno il minimo richiesto potrebbero essere rappresentate da *funzioni continue*

$$y = f(x),$$

che non ammettono derivata.

A prima vista può sembrare che si tratti qui di un dubbio pedantesco. Ma se non si vuol postulare *a priori* l'esistenza d'una linea minima dotata di tutti i caretteri intuitivi, e perciò appunto si fa appello al teorema del § 42, non si può dimenticare che la linea minima viene ivi definita come *linea limite* d'un insieme di linee, ed è facile persuadersi che una siffatta definizione può condurre a linee non soddisfacenti alle condizioni di derivabilità di cui si tratta.

È dunque essenziale rimuovere l'obiezione « provando che effettivamente la linea di lunghezza minima che congiunge due punti coincide coll'estremale (retta) da essi determinata ».

Tale scopo può essere raggiunto nel modo più semplice ricorrendo al teorema di Weierstrass sulla rappresentazione approssimata delle funzioni continue.

La linea minima che congiunge i punti A, B, suppongasi rappresentata dall'equazione

$$y = f(x),$$

dove f è una funzione continua fra a e b. Allora si possono costruire dei polinomi di grado n assai alto,

$$y = f_n(x),$$

che si approssimino quanto si vuole alla funzione (§ 38), in modo che le lunghezze delle linee $y = f_n$ convergano a quella di f (§ 38). Cerchiamo fra tutti i polinomi f_n quello $\bar{f}_n$ che corrisponde alla linea più breve entro la famiglia delle parabole $y = f_n(x)$. Una linea limite delle $\bar{f}_n$ è necessariamente minore o uguale a quella di tutte le f_n e perciò linea di lunghezza minima. Se dunque si può provare che le $\bar{f}_n$, per $n = 1, 2, 3....$, si approssimano indefinitamente all'estremale che congiunge AB, resta anche provato che questa è linea di lunghezza minima.

Ma questa prova si ottiene senza difficoltà, poichè tutte le $\bar{f}_n$ si riducono proprio alla retta $y = px + q$ che congiunge i due punti; infatti basta verificare che questa retta dà la minima lunghezza in ogni famiglia di parabole

$$y = f_n(x) = px + q + \varphi_n(x),$$
$$\varphi_n = h_0 + h_1 x + ... + h_n x^n.$$

E non vi è luogo ad indugiarvici.

NOTA. Nel ragionamento precedente si è supposto *a priori*, per semplicità di discorso, che la linea di lunghezza minima congiungente A, B, sia incontrata in un sol punto dalle perpendicolari all'asse x e perciò venga rappresentata da una sola equazione

$$y = f(x).$$

Se si vuole eliminare la restrizione apparente che così viene introdotta, si ricorrerà alle equazioni parametriche

$$x = f_1(t) \quad y = f_2(t);$$

si dimostrerà quindi, con ragionamento del tutto analogo, che la linea di lunghezza minima coincide colla estremale rappresentata dalle equazioni parametriche

$$\begin{cases} x = p_1 t + q_1 \\ x = p_2 t + q_2 \end{cases}$$

(cfr. § 47, Nota II).

§ 50. **Linee di lunghezza minima sul cilindro.** — Estendiamo ora la ricerca delle linee di lunghezza minima, dal piano al cilindro.

Assumasi il cilindro parallelo all'asse z e nel piano xy un sistema di coordinate polari ρ, θ.

Designando r il raggio del cilindro, i punti della superficie corrispondono alle coordinate

$$z, \quad \theta. \qquad (\rho = r).$$

Il differenziale della lunghezza d'una linea tracciata sul cilindro è

$$ds = \sqrt{dz^2 + r^2 d\theta^2}.$$

Poichè r è costante, il calcolo procede come al numero precedente e si trova l'*equazione differenziale delle geodetiche del cilindro*:

$$f''(\theta) = 0,$$

che — integrata — dà

$$z = p\theta + q.$$

Questa è l'equazione generale delle *eliche* tracciate sul cilindro, cioè delle linee che vengono descritte da un movimento resultante di due moti uniformi: una traslazione parallela all'asse e una rotazione intorno al medesimo.

L'identità del procedimento analitico, nei due casi del piano e del cilindro, proviene da ciò che: *il cilindro si può sviluppare sul piano* (rotolando su questo) *in guisa che i punti* (z, θ) *del cilindro vadano a coprire i punti* (x, y) *del piano, entro la striscia*

$$y = 0 \quad y = 2\pi r.$$

Fino a che si considera una *striscia* del cilindro limitata fra due generatrici, c'è *identità perfetta fra la geometria* sopra questa *superficie* e la *geometria sulla superficie piana sviluppata.* Infatti per due punti del cilindro tali che

$$0 \leq \theta < 2\pi,$$

passa un'elica ed una sola, che è la linea di lunghezza minima fra quelle congiungenti i due punti tracciate sulla striscia di superficie limitata dall'ineguaglianza posta per θ.

Ma se — in luogo di considerare una striscia di cilindro limitata da due generatrici distinte — si considera il cilindro intero, e perciò si consente a θ di uguagliare e superare 2π o di assumere valori negativi, i punti del cilindro non corrispondono più biunivocamente alle coordinate z, θ, avendosi

$$(z,\ \theta) \equiv (z,\ \theta + 2n\pi),$$

per

$$n = 1,\ 2, \ldots.$$

o

$$n = -1, -2 \ldots.$$

Questa circostanza costituisce una fondamentale differenza fra il cilindro considerato nella sua interezza e il piano. Ne consegue che *per due punti del cilindro passano infinite eliche,* potendosi trovare fra queste un'elica che faccia un numero comunque grande di giri, così in un verso come nell'altro.

Una di queste eliche, e precisamente quella che fa il minimo giro, porge la linea di minima lunghezza. Ci sono — in generale — due eliche di verso opposto che hanno per estremi i due punti, una delle quali corrisponde a un intervallo per θ minore di mezzo giro; e questa è appunto la minima.

Si conclude pertanto che:

Ogni elica tracciata sul cilindro è linea di lunghezza minima per due punti presi su di essa, i quali determinino un arco minore o uguale alla metà di un giro (cioè dell'arco determinato da due intersezioni consecutive dell'elica con una generatrice).

Nel caso dell'arco di mezzo giro vi sono per i due punti due eliche di verso opposto che rispondono alla stessa lunghezza minima.

§ 51. Linee di lunghezza minima sopra una superficie: equazioni delle geodetiche. — Il problema trattato per il piano e per il cilindro (superficie sviluppabile) si può estendere ad una qualsiasi ***superficie elementare, senza punti singolari,*** cioè ad una superficie rappresentata da tre funzioni continue e derivabili (ed anche — se si vuole — analitiche) in un campo continuo:

$$x = x(u, v)$$
$$y = y(u, v)$$
$$z = z(u, v).$$

Sia

$$\begin{cases} u = u(t) \\ v = v(t) \end{cases}$$

una linea di lunghezza minima, congiungente due punti sulla superficie. Dovrà esser minimo l'integrale

$$L = \int_{t_1}^{t_2} ds = \int_{t_1}^{t_2} \sqrt{2T} dt,$$

essendo l'elemento lineare

$$ds = \sqrt{E du^2 + 2F du dv + G dv^2} = \sqrt{2T} dt,$$

cioè

$$2T = Eu'^2 + 2Fu'v' + Gv'^2,$$

dove

$$E = \Sigma \left(\frac{\partial x}{\partial u}\right)^2, \quad F = \Sigma \frac{\partial x}{\partial u} \frac{\partial x}{\partial v}, \quad G = \Sigma \left(\frac{\partial x}{\partial v}\right)^2.$$

Supponiamo che la linea minima giaccia tutta *entro la superficie* data (cioè non contenga tratti del contorno).

Facciamo subire ad essa una variazione di forma dipendente da un parametro τ, in guisa che le funzioni $u(t)$, $v(t)$ si cambino in

$$u(t)+\tau\omega_1(t),$$

$$v(t)+\tau\omega_2(t),$$

e annulliamo la variazione

$$\delta L=\frac{dL}{d\tau}d\tau.$$

Si avrà

$$\delta L=\delta\int_{t_1}^{t_2}\sqrt{2T}.\,dt=\int_{t_1}^{t_2}\delta\sqrt{2T}.dt=\int_{t_1}^{t_2}\frac{\delta T}{\sqrt{2T}}\,dt.$$

Consideriamo $2T$ come una funzione di u, v, u', v', e scriviamo quindi

$$\delta T=\frac{\partial T}{\partial u}\delta u+\frac{\partial T}{\partial v}\delta v+\frac{\partial T}{\partial u'}\delta u'+\frac{\partial T}{\partial v'}\delta v'.$$

Verrà

$$\delta L=\int_{t_1}^{t_2}\frac{\left(\frac{\partial T}{\partial u}\delta u+\frac{\partial T}{\partial v}\delta v+\frac{\partial T}{\partial u'}\delta u'+\frac{\partial T}{\partial v'}dv'\right)}{\sqrt{2T}}dt=$$

$$=\int_{t_1}^{t_2}\frac{\left(\frac{\partial T}{\partial u}.\delta u+\frac{\partial T}{\partial u'}\delta u'\right)}{\sqrt{2T}}dt+\int_{t_1}^{t_2}\frac{\left(\frac{\partial T}{\partial v}\delta v+\frac{\partial T}{\partial v'}\delta v'\right)}{\sqrt{2T}}dt.$$

Integriamo per parti lo

$$\int_{t_1}^{t_2}\frac{\frac{\partial T}{\partial u'}\delta u'}{\sqrt{2T}}dt,$$

prendendo $\frac{1}{\sqrt{2T}}\frac{\partial T}{\partial u'}$ come fattor finito e

$$\delta u'dt=\delta du=d\delta u$$

come fattor differenziale; si otterrà

$$\left[\frac{1}{\sqrt{2T}}\frac{\partial T}{\partial u'}\delta u\right]_{t_1}^{t_2} - \int_{t_1}^{t_2} \delta u \, \frac{d\,\frac{1}{\sqrt{2T}}\frac{\partial T}{\partial u'}}{dt} \cdot dt.$$

Qui il primo termine è nullo perchè gli estremi della linea data si suppongono fissi. Pertanto, operando le trasformazioni analoghe sul secondo termine, si avrà

$$\delta L = \int_{t_1}^{t_2} \cdot \left\{ \frac{\frac{\partial T}{\partial u}}{\sqrt{2T}} - \frac{d\,\frac{1}{\sqrt{2T}}\frac{\partial T}{\partial u'}}{dt} \right\} \delta u \cdot dt +$$

$$+ \int_{t_1}^{t_2} \left\{ \frac{\frac{\partial T}{\partial v}}{\sqrt{2T}} - \frac{d\,\frac{1}{\sqrt{2T}}\frac{\partial T}{dv'}}{dt} \right\} \delta v \cdot dt.$$

La somma di questi due integrali deve annullarsi identicamente qualunque sieno δu, δv $\left(\delta u = \frac{d(u + \tau\omega_1)}{d\tau} d\tau = \omega_1(x) d\tau,\right.$ $\left.\delta v = \omega_2(x) d\tau\right)$; si conclude che ciascuno dei due integrali deve pure annullarsi identicamente e perciò

$$\left\{ \begin{aligned} &\frac{\frac{\partial T}{\partial u}}{\sqrt{2T}} - \frac{d\,\frac{1}{\sqrt{2T}}\frac{\partial T}{\partial u'}}{dt} = 0 \\ &\frac{\frac{\partial T}{\partial v}}{\sqrt{2T}} - \frac{d\,\frac{1}{\sqrt{2T}}\frac{\partial T}{\partial v'}}{dt} = 0, \end{aligned} \right.$$

ossia, eseguendo le derivazioni rispetto a t, e moltiplicando per $\sqrt{2T}$,

$$\left\{ \begin{aligned} &\frac{\partial T}{\partial u} - \frac{d}{dt}\frac{\partial T}{\partial u'} + \frac{1}{2T}\frac{dT}{dt}\frac{\partial T}{\partial u'} = 0 \\ &\frac{\partial T}{\partial v} - \frac{d}{dt}\frac{\partial T}{\partial v'} + \frac{1}{2T}\frac{dT}{dt}\frac{\partial T}{\partial v'} = 0. \end{aligned} \right.$$

Queste sono le equazioni differenziali delle *geodetiche*, linee estremali del problema proposto. Le possiamo semplificare scegliendo, come parametro t, l'arco s della linea, per cui

$$ds^2 = Edu^2 + 2Fdudv + Gdv^2,$$
$$2T = Eu'^2 + 2Fu'v' + Gv'^2 = 1,$$

e perciò

$$\frac{dT}{ds} = 0.$$

Allora le equazioni predette si riducono alla forma semplice

$$(1) \qquad \begin{cases} \dfrac{\partial T}{\partial u} - \dfrac{d}{ds}\dfrac{\partial T}{\partial u'} = 0 \\[2mm] \dfrac{\partial T}{\partial v} - \dfrac{d}{ds}\dfrac{\partial T}{\partial v'} = 0. \end{cases}$$

Calcolando le derivate parziali della forma T, queste equazioni divengono:

$$(1) \qquad \begin{cases} \alpha = \dfrac{\partial E}{\partial u}\left(\dfrac{du}{ds}\right)^2 + 2\dfrac{\partial F}{\partial u}\dfrac{du}{ds}\dfrac{dv}{ds} + \dfrac{\partial G}{\partial u}\left(\dfrac{dv}{ds}\right)^2 - 2\dfrac{d}{ds}\left(E\dfrac{du}{ds} + F\dfrac{dv}{ds}\right) = 0 \\[2mm] \beta + \dfrac{\partial E}{\partial v}\left(\dfrac{du}{ds}\right)^2 + 2\dfrac{\partial F}{\partial v}\dfrac{du}{ds}\dfrac{dv}{ds} + \dfrac{\partial G}{\partial s}\left(\dfrac{dv}{ds}\right)^2 - 2\dfrac{d}{ds}\left(F\dfrac{du}{ds} + G\dfrac{dv}{ds}\right) = 0. \end{cases}$$

Giova notare che queste equazioni si deducono l'una dall'altra in virtù della condizione

$$E\left(\frac{du}{ds}\right)^2 + 2F\frac{du}{ds}\frac{dv}{ds} + G\left(\frac{dv}{ds}\right)^2 = 1,$$

che, derivata rispetto ad s, fornisce l'identità

$$\alpha\frac{du}{ds} + \beta\frac{dv}{ds} = 0.$$

Se invece di fare $t = s$ nelle equazioni precedenti, si prende

$$t = u,$$

si trova l'equazione differenziale cui deve soddisfare la funzione

$$v = v(u)$$

perchè rappresenti una linea di lunghezza minima. Si arriva direttamente a questa equazione annullando la variazione

$$\delta L = \delta \int_a^b \sqrt{E + 2Fv' + Gv'^2} \cdot du,$$

da cui segue

$$\frac{d}{du} \frac{Gv' + F}{\sqrt{E + 2Fv' + Gv'^2}} - \frac{\frac{\partial E}{\partial v} + 2 \frac{\partial F}{\partial v} v' + \frac{\partial G}{\partial v} v'^2}{2\sqrt{E + 2Fv' + Gv'^2}}$$

e — sviluppando i calcoli —

$$(2) \quad \left\{ \begin{aligned} 2(EG - F^2)v'' = & \left(E \frac{\partial E}{\partial v} + F \frac{\partial E}{\partial u} - 2E \frac{\partial F}{\partial u}\right) \\ & + \left(3F \frac{\partial E}{\partial v} + G \frac{\partial E}{\partial u} - 2F \frac{\partial F}{\partial u} - 2E \frac{\partial G}{\partial u}\right) v' \\ & - \left(3F \frac{\partial G}{\partial u} + \frac{\partial G}{\partial v} - 2F \frac{\partial F}{\partial v} - 2G \frac{\partial E}{\partial v}\right) v'^2 \\ & - \left(G \frac{\partial G}{\partial u} + E \frac{\partial G}{\partial v} - 2G \frac{\partial F}{\partial v}\right) v'^3 \end{aligned} \right.$$

dove (com'è noto) $EG - F^2 > 0$.

Il significato geometrico delle equazioni (1) e della (2) è la *proprietà caretteristica delle linee geodetiche* (scoperta da Bernouilli): *il piano osculatore* alla linea è in ogni punto *normale alla superficie* ([1]).

Tale proprietà ha anche un evidente *significato meccanico:* un punto mobile sopra una superficie, su cui non agiscano forze, descrive una geodetica, percorrendo archi uguali in tempi uguali. Questo è il *principio d'inerzia generalizzato* di Hamilton che si traduce nelle equazioni (1), alle quali si perviene — com'è noto — applicando il *principio di* D'Alembert (s è preso come misura del tempo, $T = \frac{1}{2}(x'^2 + y'^2 + z'^2)$ designa qui la *forza viva* del punto mobile).

([1]) Cf. p. es. L. Bianchi, «Lezioni di Geometria differenziale». Bologna, 1927, Cap. VI.

NOTA I. Se nelle precedenti equazioni (1′) si pone

$$u = x, \quad v = y, \quad E = G = 1, \quad F = 0$$

cioè

$$ds^2 = dx^2 + dy^2,$$

si ottengono le equazioni parametriche delle geodetiche sul piano (o sopra una superficie sviluppabile):

$$\frac{d^2x}{ds^2} = \frac{d^2y}{ds^2} = 0,$$

da cui

$$x = p_1 s + q_1$$
$$y = p_2 s + q_2.$$

NOTA II. Il calcolo che ci ha condotto all'equazione (2) si può applicare in generale al problema in cui si tratta di render massimo o minimo l'integrale

$$I = \int_a^b F(x, y, y')dx = 0 .$$

La condizione

$$\delta I = 0,$$

eseguendo l'integrazione per parti di LAGRANGE, dà l'equazione d'EULERO

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = 0 .$$

§ 52. **Esistenza e determinazione delle geodetiche.** — Il metodo del calcolo delle variazioni ci ha condotto, nel paragrafo precedente, a trasformare la condizione di minimo della lunghezza d'una linea tracciata sopra una superficie in un'equazione differenziale.

Studiamo per un momento le condizioni d'*esistenza* delle linee *estremali* che soddisfano a questa equazione, le quali prendono il nome di *geodetiche*. Le geodetiche vengono dunque considerate come linee che hanno in ciascun punto il piano osculatore normale alla superficie, senza che ci domandiamo ora se esse sieno effettivamente, ed in quali limiti, linee di lunghezza minima.

Anzitutto il teorema fondamentale di CAUCHY relativo all'esistenza degl'integrali delle equazioni differenziali ci permette di affermare che:

Esiste una geodetica passante per un punto della superficie ed avente in esso una data tangente.

Se invece si cercano le geodetiche passanti per due punti, può accadere che si abbiano sempre infinite geodetiche, come abbiamo veduto per il cilindro (§ 50).

Ma si dimostra con DARBOUX [1] il seguente

Teorema. *Si può determinare un numero* l *tale che se si prende sopra ogni geodetica passante per un punto* P *un arco* $PP' \leq l$ *non vi è alcun' altra geodetica di lunghezza* $\leq l$ *passante per* P, P'.

Per la dimostrazione riprendiamo le equazioni (1') del paragrafo precedente, scrivendole sotto la forma

$$(1'')\quad \left\{\begin{aligned} \frac{d^2u}{ds^2} &= a\left(\frac{du}{ds}\right)^2 + 2b\frac{du}{ds}\frac{dv}{ds} + c\left(\frac{dv}{ds}\right)^2 \\ \frac{d^2v}{ds^2} &= a_1\left(\frac{du}{ds}\right)^2 + 2b_1\frac{du}{ds}\frac{dv}{ds} + c_1\left(\frac{dv}{ds}\right)^2; \end{aligned}\right.$$

ove a, b, c, a_1, b_1, c_1, designano qui date funzioni di u, v.

Sia $P = (u_0, v_0)$; la geodetica uscente da P e tangente in P a una retta data, sarà rappresentata dalle funzioni

$$\begin{aligned} u &= u\left\{ s, u_0, v_0, \left(\frac{du}{ds}\right)_0, \left(\frac{dv}{ds}\right)_0 \right\} \\ v &= v\left\{ s, u_0, v_0, \left(\frac{dv}{ds}\right)_0, \left(\frac{dv}{ds}\right)_0 \right\}, \end{aligned}$$

dove l'arco s si supporrà contato a partire da P. Ora le equazioni (1') non cambiano di forma se si cambia s in ks, dove k designa una costante. Bisognerà quindi che i valori di u, v, non cambino se si sostituisca ad

$$s, \quad \left(\frac{du}{ds}\right)_0, \quad \left(\frac{dv}{ds}\right)_0$$

rispettivamente

$$ks, \quad \frac{1}{k}\left(\frac{du}{ds}\right)_0, \quad \frac{1}{k}\left(\frac{dv}{ds}\right)_0.$$

[1] « Leçons sur la théorie des surfaces ». II Partie, 1889, p. 408.

Ciò porta che u, v dipendano, oltrechè da u_0, v_0, dalle sole variabili

$$u_1 = s\left(\frac{du}{ds}\right)_0 \quad v_1 = s\left(\frac{dv}{ds}\right)_0;$$

si avrà dunque

$$\begin{cases} u = \varphi_1(u_1, v_1, u_0, v_0) \\ v = \varphi_2(u_1, v_1, u_0, v_0). \end{cases}$$

Se le funzioni E, F, G e quindi a, b, c, a_1, b_1, c_1, sono sviluppabili in serie di TAYLOR per $(u - u_0)$, $(v - v_0)$, le funzioni φ_1, φ_2 saranno sviluppabili per le potenze di u_1, v_1 e si avrà:

$$u - u_0 = u_1 + \alpha u_1^2 + 2\alpha_1 u_1 v_1 + \alpha_2 u^2 + \ldots.$$
$$v - v_0 = v_1 + \beta u_1^2 + 2\beta_1 u_1 v_1 + \beta_2 v_1^2 + \ldots.,$$

entro un certo cerchio di convergenza delle serie indicate.

Ora il determinante jacobiano

$$\left|\begin{matrix} \frac{\partial u}{\partial u_1} & \frac{\partial v}{\partial u_1} \\ \frac{\partial u}{\partial v_1} & \frac{\partial v}{\partial v_1} \end{matrix}\right|_{u_1 = v_1 = 0} = 1,$$

quindi le equazioni precedenti potranno essere risolute rispetto ad u, v, e daranno serie procedenti per le potenze di $u - u_0$, $v - v_0$, convergenti in un certo campo, finchè $u - u_0$, $v - v_0$ restino inferiori a un dato limite. In altri termini per il punto P e per un punto P' assai vicino ad essi non passerà che una sola linea geodetica per cui u_1 e v_1, cioè la lunghezza dell'arco contato da P, non supera un certo limite. c. d. d.

Il BIANCHI chiama *cerchio geodetico* la linea luogo degli estremi degli archi geodetici di ugual lunghezza (*raggio*) uscenti da un punto (*centro*) della superficie, linea che è traiettoria ortogonale delle geodetiche uscenti dal centro; si può denominare collo stesso nome la superficie limitata da quella linea. Adottando questa definizione del cerchio (che il DARBOUX riserva invece alle linee di curvatura geodetica costante) si può dire che:

Data una geodetica, per ogni punto di essa si può determinare un cerchio geodetico tale che per tutti i punti interni ad esso e per il centro passi una geodetica ed una sola.

§ 53. **Dimostrazione della proprietà di minimo delle geodetiche.** — Ciò posto possiamo dimostrare che alle geodetiche compete *in piccolo*, cioè per archi sufficientemente limitati, la proprietà di minimo.

Si abbia una geodetica AB tracciata sopra una superficie, e a partire da un punto A di essa come centro, si costruisca un cerchio geodetico in guisa da soddisfare alle condizioni del paragrafo precedente.

Preso nel cerchio un punto B, per es. sulla nostra geodetica, avremo:

1) per i punti A, B passa una sola geodetica, cioè una sola linea estremale dell'equazione

$$\delta L = 0;$$

2) per i punti A, B passa una linea di lunghezza minima (§ 42).

Se questa linea minima non giacesse tutta internamente al cerchio (cioè contenesse degli archi del cerchio geodetico che delimita la nostra regione superficiale) si troverebbe in ogni caso un limite inferiore — ed anche un minimo — delle *distanze* di A dai punti del cerchio geodetico; si può supporre in ogni caso che l'arco geodetico AB sia stato preso inferiore a codesto minimo.

Con tale avvertenza ci assicuriamo *a priori* che la linea minima AB (inferiore alla distanza di A dal contorno) dovrà giacere tutta internamente al cerchio.

Qualora poi sia provato che la geodetica AB è minima entro il cerchio di raggio $> AB$, risulterà minima anche nel cerchio di raggio AB.

Ora si potrà dedurre che la geodetica $v = f(u)$ per A, B, e la linea minima $v = v(u)$ che congiunge i medesimi punti, coincidono, a patto che $v(u)$ *sia una funzione derivabile* (condizione richiesta per dimostrare l'equazione delle geodetiche). Ma ciò non può essere affermato *a priori*.

Per superare la difficoltà, ricorriamo al teorema di Weierstrass sulla rappresentazione approssimata d'una funzione continua. La funzione $v(u) - f(u)$, può essere approssimata con polinomi $f_n(u) = h_0 + h_1 u + \ldots + h_n u^n$ di grado $n = 1, 2, \ldots$, e quindi $v(u)$ può essere approssimata con funzioni del tipo

$$f(u) + f_n(u),$$

per $n = 1, 2, \ldots$ Anzi (§ 39) l'approssimazione può farsi in modo che la lunghezza della curva

$$v = f(u) + f_n(u)$$

si approssimi alla lunghezza (minima) di

$$v = v(u).$$

Si dedurrà quindi che $v = f(u)$ è la curva di minima lunghezza che congiunga A, B, se si possa dimostrare che essa è minore di tutte le curve

$$v = f(u) + f_n(u),$$

per

$$n = 1, 2, \ldots .$$

Ora si cerchi la curva di lunghezza minima, congiungente AB, entro la famiglia

$$v = f(u) + f_n(u),$$

per un dato n.

Questa curva, che rende minima la lunghezza $L(h_0, h_1, \ldots h_n)$, esiste certo (§ 10) e soddisfa alle condizioni necessarie

$$\frac{\partial L}{\partial h_i} = 0 \qquad (i = 0, 1, 2, \ldots u),$$

le quali, prese unitamente alle condizioni ai limiti, sono soddisfatte soltanto dalla funzione che annulla la variazione δL, cioè per

$$h_0 = h_1 = \ldots = h_n = 0.$$

Segue di qui che la curva $f(u)$ è di minima lunghezza nella famiglia

$$v = f(u) + f_n(u),$$

qualunque sia il valore di

$$n = 1, 2, 3 \ldots,$$

e perciò essa è assolutamente di lunghezza minima, rispetto a qualunque altra linea congiungente A, B, entro la superficie considerata. c. d. d.

La dimostrazione precedente fa vedere che la geodetica AB è minima entro la superficie d'un cerchio geodetico, il cui contorno disti da A più che la lunghezza l di AB. Ne consegue che la geodetica AB è certo minore di qualsiasi linea che passi per AB e giaccia entro una superficie contenente il cerchio predetto, giacchè se una linea ACB contenga un punto C fuori del cerchio, sarà $ACB > AC > l$.

Dunque la AB è minima in senso assoluto sopra la superficie data.

Concludiamo pertanto: Sopra una superficie senza punti singolari sia tracciata una linea geodetica non avente punti sul contorno; a partire da un qualsiasi punto A della linea e in un verso di essa (a destra), esisteranno sempre dei punti B per cui la geodetica AB è linea di lunghezza minima fra tutte quelle tracciate sopra la superficie per A, B; esisteranno in generale anche dei punti B' per cui la geodetica AB' non gode più la proprietà di minimo, e i punti B saranno separati dai B' per mezzo di un *punto* A' che sarà *il punto più lontano della linea geodetica a destra di* A *per cui vale la proprietà di minimo;* la geodetica AA' è minore o uguale di ogni altra linea cogli stessi estremi.

Sopra una linea geodetica aperta, i cui estremi non sieno sul contorno della superficie, od anche sopra una geodetica chiusa, a partire da ogni punto A, determiniamo a destra di esso il più lontano punto A' per cui l'arco geodetico AA' sia minimo, e analogamente determiniamo a sinistra di A il punto A'' per cui AA'' sia minimo. Col ragionamento del § 4, concluderemo che:

Esiste un numero positivo ℓ *tale che ogni arco della linea geodetica di lunghezza* $\leq \ell$ *gode della proprietà di minimo.*

Questa proprietà si estende al caso delle *superficie chiuse* senza punti singolari.

Esempii. Abbiamo già verificato le proprietà sopra enunciate per le eliche, geodetiche del cilindro (§ 50).

Nel caso della sfera le geodetiche sono archi di circolo massimo; infatti il piano del circolo (osculatore in ogni suo punto) passa pel centro e perciò è normale alla superficie in tutti i punti della linea.

Per due punti che non sieno opposti sulla sfera, passa un circolo massimo; l'arco minore dà la linea di lunghezza minima fra i due punti. Ogni circolo massimo gode dunque effettivamente della proprietà di minima lunghezza per due punti che sieno estremi d'un arco minore o uguale alla semicirconferenza. Ma per la semicirconferenza tale proprietà è soddisfatta in senso largo: per due punti opposti passano infiniti semicerchi geodetici di ugual lunghezza.

§ 54. Verifica diretta della proprietà di minimo delle geodetiche. — La proprietà di minimo delle geodetiche entro una superficie convenientemente limitata, è stata dimostrata da noi sulla base dei teoremi d'esistenza del minimo e di unicità della geodetica determinata da due punti.

Ora in queste condizioni la stessa proprietà si può verificare direttamente con una trasformazione di coordinate, che conduce ad una forma dell'elemento lineare messa in luce da Gauss.

Questo procedimento è stato indicato da Darboux nelle sue « Leçons sur la théorie des surfaces », II Partie, 1889, cap. IV, p. 402.

La dimostrazione (cfr. anche Bianchi, op. cit., § 105) si svolge semplicemente come segue:

Si considerino le geodetiche tracciate sulla superficie, per un punto P; e si stacchi su ciascuna di esse a partire da P un arco di lunghezza costante PP', inferiore ad un certo limite l.

Il luogo del punto P' sarà una curva traiettoria ortogonale delle geodetiche per P, che abbiamo denominato (con Bianchi) cerchio geodetico di centro P.

Ora finchè l è assai piccolo vi è una sola geodetica passante per P e P' (§ 52). Quindi nel cerchio geodetico anzidetto si possono assumere come linee coordinate:

$v = \text{cost}$ le geodetiche per P,

e come

$u = \text{cost}$ le traiettorie ortogonali.

Si può anche prendere come parametro u l'arco delle geodetiche per P.

Così facendo — e tenuto conto del modo come viene trasformata l'equazione delle geodetiche — si riconosce (con

Gauss) che l'elemento lineare assume la forma

$$ds = + \sqrt{du^2 + G dv^2}.$$

Questa forma si presta alla verifica diretta che le linee $v = \text{cost}$ sono di lunghezza minima. Infatti la lunghezza della linea u che congiunge due punti (u_1, k), (u_2, k), sarà

$$\int_{u_1}^{u_2} du = u_2 - u_1 \qquad (u_2 > u_1).$$

Invece la lunghezza d'una linea

$$v = \varphi(u),$$

congiungente i medesimi punti, sarà in generale

$$\int_{u_1}^{u_2} \sqrt{1 + \varphi'^2}\, du;$$

ed il secondo integrale — ad elementi positivi — supera certo il primo, ove non sia $\varphi' = 0$.

§ 55. **Proprietà di minimo delle geodetiche in grande.** — Dopo che si è dimostrata la proprietà di minimo delle geodetiche *in piccolo* (§§ 53, 54), che cosa aggiunge ancora il teorema esistenziale di Hilbert?

Supponiamo data una *superficie illimitata*, sia p. es. *chiusa* o anche *infinita*, *senza punti singolari*, la quale si supponga rappresentata da un'equazione analitica $f(x, y, z) = 0$ risolubile per mezzo di due parametri nell'intorno d'ogni punto.

Dati sulla superficie due punti A e B, sappiamo che esiste una linea di lunghezza minima che li congiunge. Ora questa linea deve godere la proprietà di minimo per ogni suo arco piccolo quanto si vuole, e perciò anche dentro un cerchio geodetico determinato come nei paragrafi precedenti. Segue che la linea minima soddisfa in ogni suo punto all'equazione differenziale delle geodetiche, cioè è una geodetica.

Dunque: *Dati due punti comunque lontani sopra la superficie, esiste sempre una geodetica che li congiunge e che è linea di lunghezza minima fra di essi.*

Questa può dirsi la proprietà di minimo delle geodetiche *in grande*; essa esprime un teorema d'esistenza *esteso*, rispetto all'equazione differenziale delle geodetiche.

NOTA. In una superficie limitata da contorno, la linea minima che congiunge due punti AB, può contenere dei tratti del contorno. Si avrà ancora che « ogni linea minima è geodetica » quando il contorno sia determinato in guisa da soddisfare ad una *condizione di convessità* rispetto alle geodetiche minime.

§ 56. Il lemma di Du Bois Reymond: la dimostrazione del minimo resa indipendente dall'unicità dell'estremale per due punti in un campo. — In una Memoria « Erlaüterungen zur den Anfangsgründen der Variationsrechnung » (1) che risale al 1879, DU BOIS REYMOND si propose di esaminare criticamente fino a che punto il calcolo delle variazioni permettesse di dimostrare che la retta è linea di lunghezza minima. Egli accolse come evidenti i presupposti che:

1) fra le linee continue rettificabili congiungenti due punti debba esserci una linea continua minima, avente una certa lunghezza L;

2) che vi sia una linea rettificabile di lunghezza assolutamente minima $L_1 \leq L$, la quale possa essere dotata di un certo numero di punti di discontinuità.

Dimostrò quindi che, se fosse $L_1 < L$, si potrebbe approssimare la linea discontinua con una linea continua e perciò L non sarebbe più il minimo per le linee continue.

DU BOIS REYMOND credette aver dimostrato in tal modo « l'esistenza d'una linea continua minima », mentre a tale scopo occorre giustificare il presupposto 1) (cfr. § 42). Superata oggi questa difficoltà col teorema di HILBERT, vediamo quale serie di considerazioni svolge il DU BOIS per dedurre che la linea minima è una retta, estremale dell'equazione d'EULERO.

Ammesso che vi sia una linea minima continua, il nostro autore ammette pure, in conformità coll'ipotesi della rettificabilità (intesa come condizione analitica per l'esistenza del-

(1) Math. Annalen, Bd 15.

l'integrale di RIEMANN $\int_a^b \sqrt{1+f'^2} \cdot dx$)) che essa sia generalmente derivabile, cioè che *esista* — tranne in punti eccezionali — una *derivata continua* f', ma osserva che il metodo di LAGRANGE — deducendo l'equazione di minimo

$$\frac{f''}{\sqrt{1+f'^2}^3}=0$$

da un'integrazione per parti — si basa sull'*esistenza della derivata seconda* f''. Sorge quindi il dubbio che la linea minima $y=f(x)$ corrisponda ad una funzione priva di derivata seconda, la quale non soddisfi alla condizione precedente, cioè non sia una retta.

Per eliminare questo dubbio DU BOIS ha ripreso in esame la condizione d'annullamento della variazione δL cfr. § 47):

$$\int_a^b \frac{f'(x)}{\sqrt{1+f'(x)^2}}\,\omega'(x)dx=0,$$

che deve essere soddisfatta per qualsiasi funzione $\omega(x)$, e ne ha dedotto *direttamente* la condizione

$$\frac{f'}{\sqrt{1+f'^2}}=\text{cost},$$

che è un'integrale dell'equazione d'EULERO.

Il *metodo* di DU BOIS REYMOND ha un *valore generale*. Quando si voglia annullare la variazione dell'integrale

$$\int_a^b F(x, y, y')dx,$$

si può modificare l'integrazione per parti di LAGRANGE, pervenendo così all'equazione

$$\int_a^b \left\{\frac{\partial F}{\partial y'}-\int_a^x \frac{\partial F}{\partial y}dx\right\}\omega'(x)\cdot dx=0.$$

Si tratta ora di dimostrare il

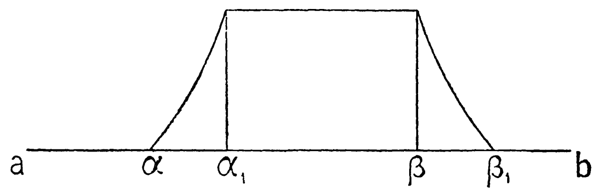
Lemma di Du Bois Reymond [1].

Se l'equazione

$$\int_a^b M(x)\omega'(x)dx = 0$$

è soddisfatta per una funzione continua $M(x)$, qualunque sia la funzione $\omega(x)$ di cui figura la derivata, si ha

$$M(x) = \text{cost.}$$



Si considerino fra a e b i punti successivi α, α_1, β, β_1 e si assuma

$$\omega(x) = 0 \quad \text{per} \quad a \leq x \leq \alpha_1,$$

e

$$\text{per} \quad \beta_1 \leq x \leq b,$$

$$\omega(x) = 1 \quad \text{per} \quad \alpha_1 \leq x \leq \beta,$$

ω crescente fra α e α_1 e decrescente fra β e β_1.

Sarà

$$\int_a^b M\omega' dx = \int_\alpha^{\alpha_1} M\omega' dx + \int_\beta^{\beta_1} M\omega' dx = 0.$$

Ora per il primo teorema della media, essendo $\omega' > 0$ fra α e α_1, si avrà

$$\int_\alpha^{\alpha_1} M\omega' dx = M(x_1)\int_\alpha^{\alpha_1} \omega' dx = M(x_1)\{\omega(\alpha_1) - \omega(\alpha)\} = M(x_1),$$

[1] Esponiamo il ragionamento semplificato di Hilbert. Cfr. Whittemore, Annals of Mathematics (2), Bd II, (1901). — Bolza, op. c., p. 28. Sul lemma di Du Bois Reymond cfr. anche Hadamard, op. c., p. 70, 484.

e così

$$\int_{\beta}^{\beta_1} M\omega' dx = -M(x_2),$$

designando x_1 un punto fra α e α_1 ed x_2 un punto fra β e β_1.

Si conclude

$$M(x_1) = M(x_2),$$

e poichè impiccolendo i tratti $\alpha\alpha_1$, $\beta\beta_1$, si ha $x_1 = \lim \alpha$, $x_2 = \lim \beta$, i punti x_1, x_2 sono in sostanza due punti arbitrari di (a, b). Dunque

$$M(x) = \text{cost} \qquad \text{c. d. d.}$$

Ragionando sul caso particolare della linea di lunghezza minima si ottiene in tal modo la *dimostrazione diretta che la linea minima — di cui si conosce l'esistenza — è un'estremale, senza appoggiarsi* all'*unicità* dell'*estremale* (retta nel piano, geodetica sopra una superficie) passante *per due punti.*

Questa conclusione sussiste nell'ipotesi che $M(x)$ sia *continua.* Ma si estende facilmente al caso in cui vi sia un numero finito di discontinuità.

Più in generale se la $M(x)$ è *generalmente continua, fatta eccezione* da un *gruppo* di punti G di *misura nulla,* immaginando surrogato l'integrale precedente dall'*integrale di* LEBESGUE, si deduce $M = 0$ tranne tutt'al più pei punti di G.

Così completato, il lemma di DU BOIS REYMOND acquista valore per il teorema di LEBESGUE che « ogni funzione continua rettificabile ammette derivata in *quasi* tutti i punti, cioè fatta eccezione pei punti d'un gruppo di misura nulla » (cfr. § 39).

Riferiamoci al caso della linea piana minima e si supponga che questa sia rappresentata dall'equazione

$$y = f(x),$$

dove f sia continua e derivabile tranne — tutt'al più — pei punti d'un gruppo G di misura nulla, e che la sua lunghezza sia espressa dall'integrale di LEBESGUE

$$L = \int_a^b \sqrt{1 + f'(x)^2} \cdot dx,$$

dove i punti di G possono trascurarsi nel calcolo dell'integrale. Annullando la variazione δL avremo,

$$\int_a^b \frac{f'(x)}{\sqrt{1+f'(x)^2}}\,\omega'(x)=0$$

e quindi — per tutti i punti fuori di G —

$$\frac{f'(x)}{\sqrt{1+f'(x)^2}}=\text{cost},\quad \text{o}\quad f'(x)=\text{cost}.$$

Ora si riconosce che, la linea $y=f(x)$ avendo lunghezza finita, la $f(x)$ è *assolutamente continua* in guisa che essa è l'integrale di $f'(x)$ anche se questa f' non è definita in un insieme di punti di misura nulla. Si avrà dunque

$$f(x)=\int_a^x f'(x)dx+c_1,$$

$$f'(x)=c_2,$$

e perciò

$$f(x)=c_2(x-a)+c_1 \text{ [1]}.$$

§ 57. **Criterii per il minimo desunti dalla seconda variazione: condizione di Legendre.** — Nei §§ 53, 54, 56, riferendoci al problema tipico delle geodetiche, abbiamo svolto sostanzialmente i principii del calcolo delle variazioni, fondando il riconoscimento della proprietà di minimo (o massimo) delle estremali sul teorema d'esistenza del minimo (di HILBERT) e sulla determinazione dell'estremale per due punti in un campo convenientemente limitato; abbiamo anche visto che

1) il teorema di determinazione delle estremali permette di dedurre la proprietà di minimo in piccolo, mediante una verifica diretta (§ 54);

[1] Un uso più largo del lemma di DU BOIS REYMOND, in rapporto all'integrale di LEBESGUE, è fatto da TONELLI allo scopo di stabilire la dimostrazione diretta del minimo per una vasta classe di problemi isoperimetrici, estendendo un teorema di HADAMARD. Cfr. la nota sul problema degl' isoperimetri, Rendic. Accad, dei Lincei, 6 aprile 1913.

2) il teorema d'esistenza basta da solo a dedurre la proprietà di minimo, seguendo il metodo di Du Bois Reymond.

Ora importa ricordare un'altra teoria indipendente dai teoremi d'esistenza, che si basa sulla *seconda variazione*; teoria che risale a Legendre (1786) e fu elaborata da Iacobi (1837), ma che ha ricevuto assetto rigoroso soltanto per opera di Weierstrass-Scheeffer, e *dopo* che il resultato essenziale era già stato ottenuto da Weierstrass mediante la considerazione d'un campo d'estremali (§ 59). Cerchiamo di riassumere in breve questa classica teoria.

Già notammo che il problema della geodetiche si può ridurre al tipo generale seguente: determinare la funzione $y = f(x)$ che rende massimo o minimo un integrale della forma

$$L = \int_a^b F(x, y, y')dx$$

dove F è una funzione continua e derivabile degli argomenti

$$x, \quad y = f(x), \quad y' = f'(x).$$

L'annullamento della variazione

$$\delta L = 0,$$

conduce ad un'equazione differenziale

(1) $$\varphi(x, y, y', y'') = 0$$

che è *condizione necessaria* pel massimo o minimo, presupposta la derivabilità di f.

Questa è l'*equazione* d'Eulero, i cui integrali abbiamo chiamato *estremali* del problema proposto. È chiaro che tale condizione non può essere *a priori* sufficiente. Considerando due fasci di linee

$$y = f(x) + t\omega_1(x)$$

$$y = f(x) + t\omega_2(x),$$

potrà avvenire p. es. che f, determinata in guisa che sia $\dfrac{dL}{dt} = 0$, riesca un massimo per il primo fascio e un minimo per il secondo; in tal caso l'estremale f non corrisponde certo ad un massimo o ad un minimo per riguardo e tutte le linee fra a e b, comunque queste vengano limitate entro un conniente intervallo funzionale.

Per riconoscere se l'integrale f dell'equazione d'Eulero, $\varphi = 0$, corrisponde effettivamente al massimo o al minimo richiesto, si presenta naturale (in rapporto al § 19) di considerare la *variazione seconda*

$$\delta^2 L = \frac{d^2 L}{dt^2} dt^2.$$

Affinchè la linea estremale f corrisponda ad un minimo rispetto a qualsiasi famiglia di curve $y = f(x) + t\omega(x)$, deve essere

$$\delta^2 L \geq 0$$

per tutti i punti della linea; per il massimo si ha

$$\delta^2 L \leq 0.$$

Ora Legendre [1] ha insegnato a trasformare la variazione seconda riducendo

$$(2) \qquad \delta^2 L = dt^2 \int_a^b \frac{\partial^2 F}{\partial y'^2} \left(\omega'(x) + \frac{\dfrac{\partial^2 F}{\partial y \partial y'} + w(x)}{\dfrac{\partial^2 F}{\partial y'^2}} \, \omega(x) \right)^2 dx,$$

dove $w(x)$ è una funzione che si suppone determinata in guisa da soddisfare — per x fra a e b — all'equazione di Riccati

$$(3) \qquad \left(\frac{\partial^2 F}{\partial x \partial x'} + w(x) \right)^2 - \frac{\partial^2 F}{\partial y \partial y'} \left(\frac{\partial^2 F}{\partial y^2} + w'(x) \right) = 0 .$$

Da questa trasformazione Legendre deduce che la condizione

$$(4) \qquad \frac{\partial^2 F}{\partial y'^2} > 0 \quad (\text{o} < 0),$$

[1] « Mémoire sur la manière de distinguer les maxima des minima dans le calcul des variations ». Mémoires de l'Acad. des Sciences, 1786,

Cfr. O. Bolza, « Vorlesungen über Variationsrechnung ». Teubner, Lipsia, 1909, p. 65.

ove sia *soddisfatta per tutti i punti dell'estremale,* porta $\delta^2 L > 0$ (o < 0), e quindi che f corrisponda effettivamente ad un minimo (o ad un massimo), *almeno* per tutte le linee comprese *in un certo intervallo funzionale* o *striscia* contenente la linea data.

Vedremo fra un momento che la deduzione precedente è illegittima.

Intanto notiamo che la condizione di LEGENDRE è soddisfatta nel piano sopra una retta qualsiasi, e sopra una superficie pei punti di una qualsiasi geodetica (ed anche d'una linea qualsiasi).

Infatti la lunghezza d'una linea piana $y = f(x)$ vale

$$L = \int_a^b \sqrt{1 + f'^2}\, dx;$$

qui

$$F(x, y, y') = \sqrt{1 + y'^2},$$

$$\frac{\partial^2 F}{\partial y'^2} = \frac{1}{(1 + y'^2)^{\frac{3}{2}}},$$

e posto $y' = k$ si ha sempre

$$\frac{\partial^2 F}{\partial y'^2} > 0.$$

Si consideri invece la lunghezza d'una linea $v = v(u)$ sopra una superficie (cfr. § 51). Si ha

$$L = \int_a^b \sqrt{E + 2Fv' + Gv'^2} \cdot du,$$

$$F(u, v, v') = \sqrt{E + 2Fv' + Gv'^2},$$

e si ottiene quindi facilmente

$$\frac{\partial^2 F}{\partial v'^2} = \frac{EG - F^2}{(E + 2Fv' + Gv'^2)^{\frac{3}{2}}}$$

che è > 0 qualunque siano i valori di u, v, v'.

§ 58. **Condizione di Iacobi.** — Ma l'anzidetta condizione di LEGENDRE pei punti dell'estremale non basta ad accertare che questa corrisponda ad un minimo (o ad un massimo).

Anzitutto per la legittimità della trasformazione di LEGENDRE occorre che *esista in tutto l'intervallo* (a, b) *l'integrale dell'equazione* (1) del paragrafo precedente.

IACOBI [1] investigando il significato di questa condizione è pervenuto a trasformarla come segue:

Si considerino le linee estremali uscenti dal punto $a, y = f(a)$, e si determini il loro inviluppo, il punto di contatto di questo inviluppo con $y = f(x)$ verrà designato come *fuoco coniugato* di *a*. *La condizione di* IACOBI *afferma che il fuoco coniugato di* a *sulla linea* y = f(x) *non appartiene all'arco* ab *della linea, ma al suo prolungamento.* Questa condizione deve in particolare ritenersi soddisfatta se le estremali uscenti da *a* non hanno altri punti comuni e manca quindi l'inviluppo anzidetto.

Esempi. Sul cilindro un'elica passante per un punto P non ha altre intersezioni coll'elica infinitamente vicina: il fuoco coniugato di P può dirsi il punto all'infinito dell'elica.

Sulla sfera un cerchio massimo per il punto P incontra il cerchio infinitamente vicino per P nel punto P' opposto a P: P' è il fuoco coniugato di P.

Qui appare la *necessità* della condizione di IACOBI. Infatti un arco PQ di cerchio massimo maggiore alla semicirconferenza non è linea minima, neppure in senso *relativo*, per riguardo ad una striscia comunque piccola che comprenda la linea.

§ 59. **Condizione sufficiente di Weierstrass.** — Suppongasi che l'estremale del problema proposto congiungente i punti *a, b,* soddisfi alle condizioni di LEGENDRE e di IACOBI. Potrà ora affermarsi che codesta linea risponde effettivamente ad un minimo (o ad un massimo)?

WEIERSTRASS ha mosso al riguardo l'obiezione seguente:

[1] « Zur Theorie der Variationsrechnung und der Differentialgleichung ». Journal für Math., Bd XVII, (1837).
Cfr. BOLZA, op. c., p. 59-79.

La linea $y = f(x)$ per cui

$$\delta L = 0$$

e su cui

$$\delta^2 L > 0$$

risponde certo ad un *minimo relativo* entro ciascuna famiglia di linee

$$y = f(x) + t\omega(x),$$

qualunque sia $\omega(x)$; ma questo minimo vale per

$$|t| < \varepsilon_\omega,$$

dove ε_ω dipende dalla famiglia di linee considerata. Per affermare che f sia una *linea minima, occorre* che vi sia un *minimo di tutti gl' intervalli* ε_ω!

L'obiezione precedente si può anche presentare sotto un altro aspetto.

Se si ha, sulla linea $y = f(x)$,

$$\frac{\partial^2 F}{\partial y'^2} > 0, \qquad \delta'^2 L > 0,$$

non è lecito affermare che sia $\frac{\partial^2 F}{\partial y'^2} > 0$ *su tutte le linee d'una striscia contenente* f, giacchè in una tale striscia vi sono linee $y = \varphi(x)$ che si approssimano ad f senza che $y' = \varphi'$ si approssimi ad f'.

Le condizioni di Legendre e *di* Iacobi portano soltanto che l'*estremale* y = f(x) su cui si suppongono soddisfatte sia un *minimo* (*o un massimo*) *debole* (¹), cioè un minimo relativo a tutte le linee φ che sono prossime ad f in guisa che φ' sia prossimo ad f'.

Affinchè l'estremale corrisponda a un *minimo* (o a un *massimo*) *forte*, cioè ad un *minimo* (o *massimo*) *relativo a tutte le linee d'una striscia* convenientemente delimitata intorno ad essa, è *sufficiente* che, *oltre la condizione di* Iacobi, sia verificata (non solo sull'estremale come voleva Legendre ma)

(¹) Cfr. Scheeffer, Math. Annalen, Bd XXVI, (1886); Bolza, op. c., p. 88.

entro la striscia la *condizione di* WEIERSTRASS

$$\frac{\partial^2 F}{\partial y'^2} > 0, \qquad (0 > 0)$$

dove al posto di y′ si metta un *numero reale qualsiasi* fra $-\infty$ e $+\infty$ (1).

Appunto questa condizione è soddisfatta se per F si pone l'espressione che figura sotto l'integrale della lunghezza d'una linea, nel piano o sopra una superficie.

WEIERSTRASS ha dato in tal modo la sistemazione rigorosa del calcolo delle variazioni nelle sue *Lezioni di Berlino* del 1879, ormai conosciute universalmente per le pubblicazioni di numerosi scolari (2).

La via con cui WEIERSTRASS è giunto al resultato fondamentale si può in breve riassumere così:

1) in base alla condizione

$$\frac{\partial^2 F}{\partial y'^2} > 0,$$

e colla limitazione di IACOBI, si determina un campo d'estremali con un estremo comune, in guisa che per due punti passi una sola estremale;

2) si verifica che l'integrale esteso all'estremale è minore di quello esteso ad un'altra linea qualunque.

Questo procedimento non differisce dunque sostanzialmente da quello a cui giunse indipendentemente DARBOUX per il caso delle geodetiche, e che abbiamo esposto nel § 54.

Ciò che giuoca è in ogni caso il teorema d'esistenza e di determinazione per gl'integrali delle equazioni differenziali (3), su cui pure abbiamo fondato il nostro metodo del § 53.

(1) Cfr. O. BOLZA, op. c., p. 123.

(2) Cfr. O. BOLZA, op. c.

(3) Cfr. G. A. BLISS, An esistences theorem for a differential equation of the second order, with an application to the calculus of variations. Transactions of the American, Math. Soc., Vol. V. (1904).

Per lo sviluppo generale della teoria di DARBOUX vedi KNESER « Lehrbuch der Variationsrechnung ». E per il confronto colla teoria di WEIERSTRASS cfr. HADAMARD, « Leçons sur le calcul des variations ». Paris, Hermann, 1910, p. 381.

La determinazione dell'estremale per due punti figura d'altronde nella condizione di IACOBI e però è implicita anche nello sviluppo della teoria di WEIERSTRASS-SCHEEFFER, innanzi abbozzato.

NOTA SULLA DISTINZIONE FRA MINIMI ASSOLUTI E RELATIVI.

Ora importa osservare che: le condizioni di IACOBI e di WEIERSTRASS assicurano soltanto che l'estremale goda della proprietà di *minimo relativo entro una striscia* che comprende la linea, e *non* che sia un *minimo assoluto.*

Così nell'esempio delle geodetiche sul cilindro: ogni arco di elica — per quanto grande — è linea di lunghezza minima entro una piccola striscia (il fuoco coniugato d'un punto essendo all'infinito § 58); ma soltanto un arco che non superi mezzo giro è linea di lunghezza minima su tutta la superficie (§ 50) (1).

§ 60. Il problema elementare degli isoperimetri. — Il problema elementare della teoria degl'isoperimetri trattato nell'Art. XXVI, si può ridurre al seguente:

Dati sopra l'asse delle x due punti a, b, e data una lunghezza $l > b - a$, determinare una linea di lunghezza l e di estremi a, b, che racchiuda un'area massima coll'asse x.

Designando la linea incognita con $y = f(x)$ e l'area suindicata con A, si tratta di render massimo l'integrale

$$A = \int_a^b f(x)dx$$

dove

$$l = \int_a^b \sqrt{1 + f'^2} \cdot dx$$

$$f(a) = f(b) = 0.$$

Inoltre, per l'*ipotesi*

$$l > b - a,$$

la curva $y = f(x)$ che rende *minimo* A *non* è un'*estremale*

(1) Per questa distinzione cfr. DARBOUX, « Leçons, Vol. III, Cap. V. p. 86.

per l'integrale $L = \int_a^b \sqrt{1 + f'^2} \cdot dx$, diguisachè la variazione di questo integrale è diversa da 0.

Esisterà dunque una curva

$$y = f(x) + \tau\omega_1(x)$$

dove

$$\omega_1(a) = \omega_1(b) = 0,$$

tale che

$$\left[\frac{dL(\tau)}{d\tau}\right]_0 \neq 0.$$

Ora poniamo

(1) $$y = f(x) + t\omega(x) + \tau\omega_1(x),$$

dove anche

$$\omega(a) = \omega(b) = 0.$$

La lunghezza L di questa linea, dipendente dai parametri t e τ, è funzione di t, τ; se la linea stessa deve appartenere alla classe di linee date, dovrà aversi

(2) $$L(t, \tau) = l.$$

Essendo qui

$$\left(\frac{\partial L(t, \tau)}{d\tau}\right)_{0,\,0} = \left(\frac{dL(0, \tau)}{d\tau}\right)_0 \neq 0,$$

l'equazione (2) definisce una funzione implicita $\tau(t)$ che si annulla nel punto $t = 0$. E la curva

$$y = f(x) + t\omega(x) + \tau(t)\omega_1(x),$$

considerata per tutti i valori di t in un certo intorno del punto 0, ha costantemente la lunghezza

$$L = l.$$

Poniamo

$$y = f(x) + t\omega(x) + \tau(t) \cdot \omega_1(x)$$

entro l'integrale A; avremo

$$A(y) = A(f + t\omega + \tau(t) \cdot \omega_1),$$

$$\frac{dA}{dt} = \frac{\partial A}{\partial t} + \frac{\partial A}{\partial \tau}\frac{d\tau}{dt}$$

e per la (2)

$$\frac{dA}{dt} = \frac{\partial A}{\partial t} - \frac{\partial A}{\partial \tau} \frac{\frac{\partial L}{\partial t}}{\frac{\partial L}{\partial \tau}}$$

Affinchè $y = f(x)$ dia un minimo di A dovrà essere

$$\left(\frac{dA}{dt}\right)_{t=0} = \left(\frac{\partial A}{\partial t}\right)_{t=0} - \left(\frac{\partial A}{\partial \tau}\right)_{t=0} \left[\frac{\frac{\partial L}{\partial t}}{\frac{\partial L}{\partial \tau}}\right]_{t=0} = 0.$$

Ora $\left(\frac{\partial A}{dt}\right)_{t=0}$ e $\left(\frac{\partial L}{\partial t}\right)_{t=0}$ sono indipendenti da ω_1, essendo

$$A = A(f + t\omega + \tau(t) \cdot \omega_1) \text{ e } \tau(0) = 0,$$

e così $\left(\frac{\partial A}{\partial \tau}\right)_{t=0}$, $\left(\frac{\partial L}{\partial \tau}\right)_{t=0}$ sono indipendenti da ω; il rapporto $-\lambda = \left(\frac{\frac{\partial A}{\partial \tau}}{\frac{\partial L}{\partial \tau}}\right)_0$ risulta perciò indipendente da ω. Quindi la equazione

$$\left(\frac{\partial A}{\partial t}\right)_0 + \lambda \left(\frac{\partial L}{\partial t}\right) = \left(\frac{\partial(A + \lambda L)}{\partial t}\right)_0 = 0,$$

dove λ figura come una costante da determinare, si deve ritenere soddisfatta per tutte le linee

$$y = f(x) + t\omega(x),$$

qualunque sia ω, cioè per qualsiasi *variazione libera* della linea $y = f(x)$ che si suppone racchiudere l'area minima.

Dunque, nelle ipotesi fatte circa la derivabilità, *la curva di lunghezza costante* L = 1 *che dà il minimo dell'area* A, *è un'estremale per l'integrale*

$$A + \lambda L.$$

Eseguendo effettivamente le derivazioni rispetto a t troviamo

$$\delta(A+\lambda L)=\frac{d(A+\lambda L)}{dt}dt=\delta\int_a^b\left\{f+\lambda\sqrt{1+f'^2}\right\}dx=$$
$$=\delta\int_a^b f dx+\lambda\delta\int_a^b\sqrt{1+f'^2}\,.\,dx.$$

Ma (§ 47):

$$\delta\int_a^b\sqrt{1+f'^2}\,dx=\int_a^b\omega\cdot\frac{f''}{(1+f'^2)^{\frac{3}{2}}}dx,$$

mentre, per definizione,

$$\delta f(x)=\omega(x)=\delta x,$$
$$\delta\int_a^b f(x)\cdot dx=\int_a^b\omega(x)dx.$$

Si trova dunque l'equazione

$$\int_a^b\omega\left\{1+\lambda\frac{f''}{(1+f'^2)^{\frac{3}{2}}}\right\}dx=0.$$

Questa equazione dovendo sussistere qualunque sia δx, si ottiene

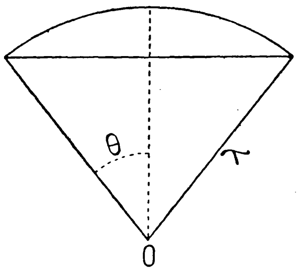
$$\frac{f''}{(1+f'^2)^{\frac{3}{2}}}=-\frac{1}{\lambda}.$$

Il primo membro è l'espressione della *curvatura* della linea $y=f(x)$. La *linea* di *minimo* cercata deve dunque avere curvatura costante $=-\frac{1}{\lambda}$ cioè deve essere un *arco di cerchio di raggio* $|\lambda|$.

La costante λ (detta *costante isoperimetrica*) viene determinata mediante la condizione data:

$$L=l.$$

A tale scopo osserviamo che nel cerchio di raggio $|\lambda|$, l'arco corrispondente all'angolo al centro 2θ ha per lunghezza $2\theta\,|\lambda|$ e che la lunghezza della sua corda è $2\,|\lambda|\,\text{sen}\,\theta$. Si avrà dunque



$$2\,|\lambda|\,\text{sen}\,\theta = b - a$$

$$2\theta\,|\lambda| = l,$$

e perciò

$$|\lambda| = \frac{l}{2\theta}$$

dove θ è definito dall'equazione

$$\frac{\text{sen}\,\theta}{\theta} = \frac{b-a}{l}.$$

Nota. Presupposto che la linea f che dà l'area massima:

1) esista (cfr. § 44),

2) sia determinata, in modo unico, dai suoi estremi a, b, come dimostra Steiner (cfr. Art. XXVI);

si deduce che questa curva deve essere simmetrica rispetto all'asse del segmento che ne congiunge gli estremi.

Ora la *simmetria* della curva di massimo rispetto all'asse del segmento che ne congiunge gli estremi, deve valere *per ogni arco della curva stessa,* giacchè un tale arco deve pur godere della proprietà di massimo (principio di Steiner).

Questa conclusione, applicata all'arco infinitesimo della curva, dà l'equazione differenziale trovata innanzi; infatti la simmetria porta che il *cerchio osculatore* in un punto (vertice d'un arco simmetrico) abbia un *contatto quadripunto,* colla curva, e quindi che il cerchio osculatore suddetto abbia raggio costante. Appare così che l'equazione differenziale, a cui deve soddisfare la curva di massima area, si può riguardare come l'interpretazione analitica della simmetria che caratterizza ogni parte della curva stessa.

§ 61. Dimostrazione della proprietà isoperimetrica del cerchio. — Resta a dimostrare che l'arco di cerchio di raggio $|\lambda|$ e di lunghezza l, che congiunge i punti a, b, racchiude effettivamente coll'asse x l'area massima fra le linee di ugual lunghezza terminate ai medesimi estremi.

La dimostrazione si compie facilmente col metodo già spiegato innanzi (§§ 49, 53) basandosi sul teorema d'*esistenza* della *linea* che dà l'*area massima* (§ 44), e sulla *unicità* dell'*arco* di *cerchio* di raggio $|\lambda|$ cogli estremi a, b, entro il semipiano superiore alla retta congiungente i due punti.

Si scriva l'equazione dell'arco di cerchio

$$y = f(x),$$

e si considerino le linee

$$y = f(x) + f_n(x)$$

dove f_n indica un polinomio di grado $n = 1, 2....$:

$$f_n = a_0 + a_1 x + ... + a_n x^n.$$

La linea di lunghezza l che dà l'area massima sia

$$y = y(x).$$

La funzione $y(x)$ si può approssimare quanto si vuole con funzioni del tipo

$$y = f(x) + f_n(x),$$

e basta quindi dimostrare che la linea

$$y = f(x)$$

gode della proprietà di massimo per riguardo alle linee

$$y = f(x) + f_n(x)$$
$$(n = 1, 2, ...).$$

Si considerino dunque le linee suddette per un certo valore di n.

Avremo l'area A e la lunghezza L funzioni continue di $a_0, a_1, ... a_n$, e il massimo di A per $L = l$ si otterrà mediante le condizioni

$$\frac{\partial(A + \lambda L)}{\partial a_i} = 0 \qquad (i = 0, 1, ... n)$$

(cfr. § 28) equivalenti a $\delta(A + \lambda L) = 0$.

Queste equazioni, tenuto conto delle condizioni ai limiti

$$y_a = y_b = 0,$$

portano

$$a_i = 0 \qquad (i = 0, 1, \ldots. n).$$

Ciò significa che la linea di lunghezza l che dà l'area **massima**, fra quelle appartenenti alla famiglia

$$y = f(x) + f_n(x),$$

soddisfa alle condizioni necessarie

$$a_i = 0;$$

e perciò tale linea, necessariamente esistente (§ 12), coincide colla

$$y = f(x),$$

onde segue il teorema da dimostrare.

§ 62. Cenno sui problemi isoperimetrici in generale. — Il problema elementare trattato nei paragrafi precedenti, è il tipo di una famiglia generale di problemi, detti per estensione *problemi isoperimetrici*.

Si tratta di:

Determinare una linea

$$y = f(x)$$

che renda massimo o minimo un integrale del tipo

$$I = \int_a^b F(x, y, y') = 0$$

subordinatamente alla condizione che un altro integrale

$$I_1 = \int_a^b F_1(x, y, y')dx$$

assuma un valore assegnato:

$$I_1 = k.$$

Se la funzione f *non è un'estremale dell'integrale* I_1, *le condizioni differenziali a cui* f *deve soddisfare si ottengono* (come al § 60) *annullando la variazione*

$$\delta(I + \lambda I_1),$$

dove λ *è una costante.*

Questa è la ***regola isoperimetrica*** d'EULERO [1], suscettibile di estensione in varii sensi.

Quanto ai criterii per riconoscere se le linee estremali f, soddisfacenti alla condizione

$$\delta(I + \lambda I_1) = 0,$$

rispondano effettivamente ad un massimo o ad un minimo, osserveremo che:

1) Per il problema *vincolato* (massimo o minimo di I per $I_1 = k$) si hanno per il massimo o il minimo le stesse condizioni di LEGENDRE e di WEIERSTRASS relative al problema *libero* (massimo o minimo di $I + \lambda I_1$).

2) Ma *la condizione analoga alla condizione di* IACOBI *per il problema vincolato è meno restrittiva di quella che appartiene al corrispondente problema libero* [2], cioè il fuoco coniugato ad un punto P su un'estremale per $I_1 = k_1$ si trova generalmente dopo — e non mai prima — di quello relativo alla curva stessa considerata come estremale di $I + \lambda I_1$.

Questa circostanza si verifica già nel problema elementare del paragrafo precedente.

Se si considerano i cerchi di raggio $|\lambda|$ passanti per un punto P, il fuoco coniugato a P, definito come intersezione di un cerchio C col cerchio infinitamente vicino (problema libero), è il punto P' di C estremo del diametro che passa per P; ma per il problema vincolato l'unicità dell'arco di cerchio di lunghezza l con dati estremi (entro una certa striscia o addirittura nel semipiano limitato dalla corda) non cessa di valere anche per archi maggiori del semicerchio.

(1) «Methodus inveniendi....» (1744), Cap. V; BERTRAND (Journal de Math., t. VII, (1842); DU BOIS REYMOND, (Math. Annalen, Bd XV, (1879) e MAYER (ibidem Bd XVI) si occuparono di dimostrare questa regola. Ma la sua dimostrazione rigorosa fu data soltanto da WEIERSTRASS nelle sue Lezioni del 1877, avendo WEIERSTRASS messo in luce la condizione che f non sia un'estremale per l'integrale I_1.

(2) Cfr. LUNDSTRÖM, «Distinction des maxima et des minima dans un problème isopérimétrique». Nova acta reg. soc. sc. Upsaliensis § 3 t. VII, (1869). MAYER, Math. Annalen XIII, (1878).

3) Le condizioni di LEGENDRE e di WEIERSTRASS $\left(\frac{\partial^2 F}{\partial y'^2} > 0 \text{ o } < 0\right.$, sull'estremale e in una striscia intorno ad essa) unitamente alla condizione di IACOBI, nella forma meno restrittiva che appartiene al problema vincolato, sono *sufficienti* perchè l'estremale corrisponda ad un *minimo* o ad un *massimo.* La dimostrazione di tale sufficienza appartiene a LINDEBERG (1909). (Cfr. BOLZA, op. cit., Cap. X).

NOTA. Si estendono facilmente ai problemi isoperimetrici le considerazioni del § 29 su cui si basa la *legge di reciprocità*: ad un massimo o minimo relativo dell'integrale I per $I_1 = \text{cost}$ corrisponde in generale un massimo o minimo di I_1 per $I = \text{cost}$. Nulla vi è da modificare o da aggiungere a tale proposito se I, I_1, sono funzioni continue della linea d'integrazione; se godono della semicontinuità si arriva alla stessa conclusione purchè questa valga nel senso che è richiesto dai teoremi d'esistenza del massimo o del minimo.

§ 63. **Nota storica sui problemi del calcolo delle variazioni.** — La ricerca di un metodo proprio per determinare le linee (superficie ecc.) che rendono minima o massima una certa funzione, si affaccia come un naturale prolungamento del calcolo differenziale.

Già NEWTON nei « Principia mathematica.... » (Londra, 1686), dà l'equazione differenziale della curva meridiana d'un solido rotondo che corrisponde al problema di render minima la resistenza al movimento in un fluido. La dimostrazione inedita è stata ritrovata fra le sue carte (cfr. O. BOLZA « Bibliotheca Math. » maggio, 1913, p. 146).

Ma il problema a cui storicamente si riattaccano le principali ricerche sul calcolo delle variazioni è quello della *brachistocrona,* proposto da GIOVANNI BERNOUILLI (Acta Eruditorum, Lipsia, 1686):

Determinare la traiettoria di un punto mobile cadente per effetto della gravità, che debba portarsi — nel minimo tempo possibile — da un punto dato ad un altro punto più basso, parimente dato.

La curva che risponde al problema è una *cicloide* colla base orizzontale.

La soluzione del problema, data dallo stesso GIOVANNI BERNOUILLI nel 1697, e le ricerche sullo stesso soggetto pubblicate da GIACOMO BERNOUILLI, furono il punto di partezza dello studio generale dei *problemi isoperimetrici*, in cui si tratta di determinare una linea che renda massima o minima una certa funzione, cioè un integrale dipendente dalla forma della linea stessa, mentre uno o più integrali, dipendenti dalla linea, assumono valori assegnati (§ 62). In questa classe di problemi rientrano appunto le ricerche degli antichi sulla linea di data lunghezza che racchiude l'area massima (cfr. § 60).

Un altro esempio è porto dal problema di « trovare una linea piana di lunghezza data che, ruotando intorno ad un asse del suo piano, generi una superficie di rotazione massima o minima ».

La curva che risolve il problema è la *catenaria* (EULERO, 1744).

Dal punto di vista del metodo si deve notare che i BERNOUILLI rimpiazzano la curva di cui si cerca il massimo o il minimo con una poligonale di cui si fanno variare gli elementi. GIOVANNI BERNOUILLI aveva creduto sufficiente, per ottenere l'equazione differenziale della curva richiesta, di far variare due lati consecutivi della poligonale, ma GIACOMO BERNOUILLI nella memoria « Analysis magni problematis isoperimetrici » (Basilea e Lipsia 1701) osservò essere necessario tener conto di tre lati; altrimenti l'equazione proposta si riduce a un'identità.

Oltre alle accennate ricerche, citiamo la scoperta della proprietà fondamentale delle *geodetiche* d'una superficie, fatta da GIOVANNI BERNOUILLI, e gli studii di TAYLOR (« Methodus incrementorum ») e di EULERO nei t. VI, VII dei Commentarii Petropolitani. Finalmente la soluzione più generale e completa pei problemi isoperimetrici si trova nella memoria di EULERO « Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes.... » (1744). Ivi in particolare è data la *regola isoperimetrica* che riconduce la ricerca del massimo o minimo di linee, vincolate a date condizioni, ad un problema analogo relativo a linee liberamente variabili (§ 60).

Il procedimento d'EULERO, limitandoci alla ricerca del massimo o minimo dell'integrale

$$\int_a^b F(x, y, y')dx = 0,$$

consiste nel sostituire alla linea una poligonale approssimata i cui vertici successivi rispondono a una differenza costante Δx; cambiando y in $y + \varepsilon$ per un vertice della poligonale, l'integrale subisce allora l'incremento:

$$\varepsilon \Delta x \left\{ \frac{\partial F}{\partial y} - \frac{d}{dx}\frac{\partial F}{\partial y'} \right\},$$

e, se deve aversi massimo o minimo, si trova la condizione necessaria

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\frac{\partial F}{\partial y'} = 0.$$

LAGRANGE « Memorie dell'Accad. delle Scienze di Torino » t. II, 1762 (cfr. Oeuvres, t. I, p. 335, II, p. 37) si è liberato dalla considerazione della poligonale approssimata alla linea, riguardando l'integrale come una funzione *della linea*, e calcolando la variazione — da lui designata col simbolo δ — come differenziale relativo ad un parametro ausiliario, da cui si fa dipendere il cambiamento della linea in una certa famiglia. Una semplice integrazione per parti conduce allora all'equazione d'EULERO. Inoltre LAGRANGE ha approfondito la questione delle *equazioni ai limiti* nel caso in cui gli estremi della linea richiesta non sieno fissi (cfr. § 48).

Quest'ordine d'idee è particolarmente importante nei problemi relativi a *funzioni di due variabili*. Ad esso si collega la determinazione delle *superficie ad area minima* aventi un dato contorno, problema risoluto da LAGRANGE, di cui s'incontra solo un caso particolare in EULERO. L'equazione alle derivate parziali corrispondente è stata interpretata da MEUSNIER (1776) come condizione che la *curvatura media* sia *costante*, per i punti della superficie.

Colle ricerche di LAGRANGE, (il cui metodo fu adottato quindi da EULERO) e cogli studii di GAUSS relativi ad integrali multipli a limiti variabili (Gottinga 1833), si può dire che l'algoritmo del Calcolo delle variazioni abbia toccato la più grande generalità e perfezione. Grazie a questo algoritmo diventa facile in ogni caso di *scoprire* le linee, superficie ecc. rispondenti a condizioni di massimo o minimo, assegnandone le equazioni differenziali. Ma la questione di riconoscere se le estremali che integrano codeste equazioni diano luogo effettivamente al massimo o al minimo proposto, non ricevette ugualmente una soluzione soddisfacente. Nonostante le ricerche di LAGRANGE e di IACOBI (§§ 57, 58), su questo argomento ha regnato la più grande oscurità ed incertezza, quasi fino alla fine del secolo scorso.

Una risposta rigorosa al problema fu data — come dicemmo (§ 59) — da WEIERSTRASS (proseguito da SCHEEFFER) e indipendentemente da DARBOUX (proseguito da KNESER). Finalmente una nuova epoca per questi problemi fu aperta dallo sviluppo di una *teoria generale delle funzioni di linee* per opera di VOLTERRA, ASCOLI, ARZELÀ ecc. Su questa base naturale veniva posto il calcolo delle variazioni dal teorema esistenziale di HILBERT del 1900, a cui seguirono analoghi e più vasti teoremi d'esistenza per opera di LEBESGUE, CARATHEODORY, HADAMARD, TONELLI ecc. Una trattazione sistematica del soggetto, in tal senso, si può ora trovare in L. TONELLI « Fondamenti di Calcolo delle Variazioni », 2 volumi, Bologna, Zanichelli, 1921 e 1923.