

Sulla teoria della convergenza degli algoritmi di iterazione.

(Di ONORATO NICOLETTI, a Pisa.)

In una Memoria: *Sulla teoria dell'iterazione* (*), che ha avuto l'onore di essere pubblicata tra le *Memorie della Società Italiana delle Scienze*, ho sviluppato una teoria della *convergenza degli algoritmi di iterazione in n variabili*, assegnando dei criterî di convergenza, i quali permettono, nella maggior parte dei casi, di riconoscere, *mediante operazioni razionali*, se un determinato algoritmo di iterazione converge a limiti determinati e finiti. Per la teoria stessa, gli algoritmi convergenti possono classificarsi, in ordine alla loro convergenza, introducendo la nozione di *ordine* e *grado* di convergenza. Vi è però, a questo punto, una grave lacuna; in quanto, mentre è agevole riconoscere se un algoritmo ha, o meno, convergenza di un ordine e grado determinati, quando invece sia dato soltanto l'algoritmo, non è possibile, per i criterî stessi, decidere con un numero finito di operazioni se l'algoritmo soddisfi, o meno, ad uno qualsiasi di essi criterî. Questa lacuna mi è ora riuscito colmare, in guisa molto semplice, ed espongo, nelle pagine che seguono, il risultato cui sono pervenuto.

Per maggior chiarezza, credo opportuno richiamare prima alcune definizioni e risultati della Memoria sopra ricordata.

(*) *Memorie della Società Italiana delle Scienze (detta dei XL)*, Serie III, Tomo XIV, Roma, 1906. Dovrò nel seguito riferirmi continuamente a questa Memoria; la indicherò sempre colla lettera A.

I. NOTAZIONI E DEFINIZIONI.

1. Si abbia (*), per considerare il caso più semplice, un sistema di n equazioni in due serie di n variabili $x_1 x_2 \dots x_n, x'_1 x'_2 \dots x'_n$

$$\Phi_i(x, x') = 0, \quad (i = 1, 2, \dots, n), \quad (1)$$

dove le $\Phi_i(x, x')$, quando le x e le x' siano variabili complesse, si suppongono analitiche in queste variabili; quando le x ed x' siano invece reali, sono loro funzioni reali, finite e continue nel campo che si considera con tutte quelle loro derivate che ci occorrerà di considerare. Facendo nelle $\Phi_i(x, x')$ $x'_k = x_k$ ($k = 1, 2, \dots, n$), otteniamo n funzioni delle $x_1 x_2 \dots x_n$

$$\varphi_i(x_1 x_2 \dots x_n) = \varphi_i(x), \quad (i = 1, 2, \dots, n) \quad (2)$$

tali che si ha per esse, coi simboli di KRONECKER (**):

$$\varphi_i(x, x') \equiv \varphi_i(x) \pmod{x'_k - x_k}, \quad (i, k = 1, 2, \dots, n). \quad (2^*)$$

Sia ora $M \equiv (g_1, g_2, \dots, g_p)$, con $g_\mu = g_\mu(x) = g_\mu(x_1 x_2 \dots x_n)$ ($\mu = 1, 2, \dots, p$), un divisore di rango p del sistema di funzioni $\varphi_1, \varphi_2, \dots, \varphi_n$, tale cioè che la matrice funzionale

$$\frac{d(g_1 g_2 \dots g_p)}{d(x_1 x_2 \dots x_n)}$$

abbia la caratteristica p . Abbiamo allora

$$\varphi_i(x) \equiv 0 \pmod{g_1, g_2, \dots, g_p}, \quad (i = 1, 2, \dots, n); \quad (3)$$

di più le equazioni

$$g_1(x) = 0, g_2(x) = 0, \dots, g_p(x) = 0 \quad (4)$$

definiscono nell' S_n di coordinate $(x_1 x_2 \dots x_n)$ una varietà M ad $n - p$ dimensioni.

Ammettiamo ora che il determinante funzionale

$$\frac{d(\Phi_1 \Phi_2 \dots \Phi_n)}{d(x'_1 x'_2 \dots x'_n)}$$

(*) Cf. A, n.° 1-3.

(**) Sui metodi di KRONECKER, cf. ad es.: KÖNIG, *Algebraischen Grössen* (Lipsia-Teubner, 1903) Cf. anche A, n.° 2.

sia diverso da zero, quando in esso si faccia $x'_k = x_k$, $g_\mu(x) = 0$ ($k = 1, 2, \dots, n$; $\mu = 1, 2, \dots, p$). In questo caso le equazioni (1) definiscono, in un intorno (di S_n) convenientemente piccolo della varietà M , le x' come funzioni determinate, finite, continue, derivabili (analitiche nel caso complesso) delle x . Facciamo infine l'ipotesi che detto C l'intorno della varietà M , entro il quale va preso il punto x perchè sian soddisfatte le condizioni superiori, anche il punto x' , definito dalle (1), cada nello stesso intorno.

2. Ciò posto, sia $x^{(0)} \equiv (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})$ un punto qualunque di questo intorno, e consideriamo la successione infinita dei sistemi di equazioni:

$$\Phi_i(x^{(0)}, x^{(1)}) = 0; \Phi_i(x^{(1)}, x^{(2)}) = 0; \dots; \Phi_i(x^{(p-1)}, x^{(p)}) = 0, \dots, (i = 1, 2, \dots, n).$$

Nelle ipotesi fatte, esse equazioni definiscono una successione di punti in S_n

$$x^{(0)}, x^{(1)}, x^{(2)}, \dots, x^{(p)}, \dots; \quad (5)$$

ed ove questa successione sia convergente ad un punto limite $X \equiv (X_1, X_2, \dots, X_n)$, diremo che le equazioni (1) definiscono un algoritmo K di iterazione (*). Per la continuità delle $\Phi_i(x, x')$ e per le (2*) il punto limite X soddisferà alle equazioni

$$\varphi_i(X) = 0, \quad (i = 1, 2, \dots, n); \quad (6)$$

se in particolare, qualunque sia il punto iniziale $x^{(0)}$ scelto nell'intorno C della varietà M , si avrà

$$g_\mu(X) = 0, \quad (\mu = 1, 2, \dots, p) \quad (7)$$

(con che le (6) sono manifestamente soddisfatte), cioè se il punto limite X appartiene alla varietà M , diremo anche che l'algoritmo K appartiene alla varietà M definita dalle equazioni (4).

II. CONDIZIONI DI CONVERGENZA. — ORDINE E GRADO DI CONVERGENZA.

3. Nelle ipotesi del n.º 1, quando il punto x si prenda in un intorno sufficientemente piccolo di M , dalle equazioni (1) si ottengono delle relazioni della forma (**):

$$x'_k - x_k = \sum_1^p \beta_{k\mu} g_\mu(x), \quad (k = 1, 2, \dots, n) \quad (8)$$

(*) Cf. A, n.º 4.

(**) Cf. A, n.º 5.

dove le $\beta_{k\mu}$ ($k = 1, 2, \dots, n$; $\mu = 1, 2, \dots, p$) sono funzioni determinate e finite delle $x_1 x_2 \dots x_n$. Da queste si traggono p identità della forma:

$$g_\mu(x') = \sum_1^p \gamma_{\mu\nu} g_\nu(x), \quad (\mu = 1, 2, \dots, p) \quad (9)$$

essendo le $\gamma_{\mu\nu}$ funzioni determinate delle x ; cioè: *nel passare dal punto x al punto successivo x' (al suo conseguente, secondo la denominazione di POINCARÉ) le p funzioni g_1, g_2, \dots, g_p subiscono una sostituzione lineare.*

Posto, per brevità di scrittura, per qualunque ρ

$$g_\mu^{(\rho)} = g_\mu(x^{(\rho)}) = g_\mu(x_1^{(\rho)} x_2^{(\rho)} \dots x_n^{(\rho)}),$$

dalle (8) si ha immediatamente che quando le p serie

$$\sum_1^\infty g_\mu^{(\rho)} \quad (\mu = 1, 2, \dots, p) \quad (10)$$

convergono assolutamente (ed uniformemente) nel campo dato, altrettanto è delle altre

$$x_k^{(\rho)} + \sum_1^\infty (x_k^{(\rho)} - x_k^{(\rho-1)}) \quad (k = 1, 2, \dots, n) \quad (10^a)$$

e quindi l'algoritmo K converge (uniformemente) ad un punto della varietà M , appartiene cioè alla varietà M (*).

(*) È chiaro che la convergenza (assoluta ed uniforme) delle serie (10) è condizione soltanto sufficiente perchè l'algoritmo converga ed appartenga al divisore $(g_1 g_2 \dots g_p)$ (cf. anche A, n.º 18). Però, se si vuole che l'algoritmo tenda al suo limite, *soddisfacendo a particolari condizioni*, si è condotti a porre la convergenza delle serie stesse.

Ammettiamo infatti che, quando il punto iniziale $x^{(0)}$ sia scelto in un intorno conveniente della varietà M , l'algoritmo sia convergente, sia $X_k = \lim_{\rho \rightarrow \infty} x_k^{(\rho)}$ e valgano insieme le (7). Consideriamo i moduli delle n differenze $X_k - x_k^{(\rho)}$ e sia δ_ρ il limite superiore di questi moduli, mentre il punto iniziale $x^{(0)}$ varia nell'intorno considerato; δ_ρ potrà definirsi come l'errore che si commette quando al punto limite (X) si sostituisca il punto ρ -iterato ($x^{(\rho)}$). Per la convergenza dell'algoritmo si ha evidentemente

$$\lim_{\rho \rightarrow \infty} \delta_\rho = 0.$$

Ammettiamo ora che, essendo a una costante positiva minore dell'unità, da un certo valore ρ_0 di ρ in poi si abbia

$$\delta_{\rho+1} \leq a \delta_\rho, \quad (\rho \geq \rho_0),$$

4. Siamo in tal guisa condotti a studiare le p serie (10) ed a cercare delle condizioni, sotto le quali esse convergono assolutamente (ed uniformemente). Queste condizioni sono espresse dal teorema seguente (*):

Poniamo:

$$\Gamma_\mu \equiv \sum_1^p |\gamma_{\mu\nu}| \quad (\text{mod } g_1, g_2, \dots, g_p). \quad (11)$$

Se su tutta la varietà M si ha

$$\Gamma_\mu \leq a_0, \quad (12)$$

oppure, essendo L una quantità finita:

$$\delta_\rho < L a^\rho, \quad (\rho \geq \rho_0).$$

In una di queste ipotesi convergeranno assolutamente ed uniformemente le n serie

$$\sum_0^\infty (X_k - x_k^{(\rho)}) \quad (k=1, 2, \dots, n)$$

e con esse le altre

$$\sum_0^\infty g_\mu^{(\rho)} = \sum_0^\infty (g_\mu(x^{(\rho)}) - g_\mu(X)) = \sum_0^\infty \left\{ \sum_1^n \bar{g}_{\mu k} (x_k^{(\rho)} - X_k) \right\} \quad (\nu=1, 2, \dots, p)$$

(dove le $\bar{g}_{\mu k}$ sono le derivate $\frac{\partial g_\mu}{\partial x_k}$ in punti intermedi) cioè le serie (10). Se dunque si vuole che gli errori successivi decrescano (almeno) come una progressione geometrica decrescente, ne segue necessariamente la convergenza (assoluta ed uniforme) delle serie (10).

Reciprocamente, quando le serie (10) convergono ed è per esse soddisfatta la (12), indicando con a un qualunque numero compreso tra a_0 ed 1 (gli estremi escl.), il punto iniziale $x^{(0)}$ può prendersi in un intorno determinato della varietà M , in guisa che per il resto $R_n(g_\mu)$ di una qualunque delle serie (10) si abbia

$$(R_n(g_\mu)) < G_0 \frac{a^n}{1-a}, \quad (\mu=1, 2, \dots, p)$$

dove G_0 è il massimo modulo delle $g_i^{(0)}$ nell'intorno considerato. Una relazione del tutto analoga varrà allora per il resto R_n di una qualunque delle serie (10^a) e quindi anche per l'errore δ_n : si potrà cioè determinare un numero ρ_0 tale che per $\rho \geq \rho_0$ sarà sempre, indicando con L un numero finito

$$\delta_\rho \leq L a^\rho, \quad (\rho \geq \rho_0);$$

cioè gli errori successivi saranno ancora confrontabili coi termini di una progressione geometrica decrescente.

(*) Cf. A, n. 5 e 6.

dove a_0 è un numero positivo determinato, **minore dell'unità**, ed il punto iniziale $x^{(0)}$ si prende in un intorno convenientemente piccolo della varietà M , le serie (10) convergono assolutamente ed uniformemente.

Valgano, sulla varietà M , le disuguaglianze (12). Due casi sono allora possibili (*).

a) Le Γ_μ non sono tutte nulle in ogni punto di M . In questo caso detto a_0 il limite superiore dei valori delle Γ_μ sulla M , indicando con a un qualunque numero positivo compreso tra a_0 ed 1, le serie (10) sono confrontabili (i moduli dei loro termini sono cioè minori dei termini corrispondenti moltiplicati per un numero finito) colla progressione geometrica $\sum_0^\infty a^\rho$, ma non si può dire ugualmente per una progressione analoga $\sum_0^\infty a'^\rho$, per la quale sia invece $0 < a' < a_0$. Diremo in questo caso che l'algoritmo K ha *convergenza lineare o di 1.° grado (e del 1.° ordine)* e diremo anche che *la convergenza è misurata dalla costante a_0* .

b) Tutte le Γ_μ e quindi anche le $\gamma_{\mu\nu}$ ($\mu, \nu = 1, 2, \dots, p$) sono nulle sulla varietà M . Può allora (in generale) determinarsi un intero $s \geq 2$, tale che in un intorno convenientemente piccolo della varietà M si hanno delle relazioni della forma:

$$g_\mu(x) = \sum_1^p \gamma_{\mu, k_1 k_2 \dots k_s} \gamma_{\mu, k_1 k_2 \dots k_s} g_{k_1}(x) g_{k_2}(x) \dots g_{k_s}(x), \quad (\mu = 1, 2, \dots, p) \quad (13)$$

dove le $\gamma_{\mu, k_1 k_2 \dots k_s}$ sono funzioni determinate delle x , che sulla varietà M rimangono inferiori in modulo ad un numero finito e *non sono tutte nulle*. In questo caso le serie (10) convergono assolutamente ed uniformemente in un intorno conveniente di M ; esse sono inoltre confrontabili colla serie:

$$1 + \sum_1^\infty \theta^{i^s}$$

dove θ è ancora un numero positivo determinato, minore dell'unità, e che ha quindi una convergenza estremamente rapida. Diciamo in questo caso che l'algoritmo K ha *convergenza (del 1.° ordine e) di grado s* .

5. Le considerazioni precedenti sono suscettibili di una importante generalizzazione (**). Sia r un intero qualunque, maggiore dell'unità e nelle (9)

(*) Cf. A, n.° 7.

(**) Cf. A, n.° 8.

cambiamo successivamente le x nelle $x^{(1)}, x^{(2)}, x^{(3)}, \dots, x^{(r-1)}$; eliminando poi le $g^{(1)}, g^{(2)}, \dots, g^{(r-1)}$, otteniamo delle relazioni affatto analoghe alle (9) stesse, che scriviamo:

$$g_{\mu}^{(r)} = g_{\nu} (x^{(r)}) = \sum_{\nu=1}^p \gamma_{\mu\nu}^{(r)} \cdot g_{\nu} (x), \quad (\mu = 1, 2, \dots, p) \quad (9)_r$$

e le $\gamma_{\mu\nu}^{(r)}$ possono ancora riguardarsi, mediante le equazioni dell'algoritmo, come funzioni determinate delle x . Poniamo ancora:

$$\Gamma_{\mu}^{(r)} \equiv \sum_{\nu=1}^p |\gamma_{\mu\nu}^{(r)}|, \quad (\text{modd } g_1, g_2, \dots, g_p); \quad (11)_r$$

abbiamo allora il teorema:

Se su tutta la varietà M si ha

$$\Gamma_{\mu}^{(r)} \leq a_0, \quad (\mu = 1, 2, \dots, p), \quad 0 \leq a_0 < 1, \quad (12)_r$$

le serie (10) convergono ancora assolutamente ed uniformemente in un intorno conveniente della varietà M .

In questo caso diremo che l'algoritmo K ha convergenza *dell'ordine* r e la convergenza sarà *lineare*, quando non tutte le $\Gamma_{\mu}^{(r)}$ siano nulle su M ; quando invece le $\Gamma_{\mu}^{(r)}$ e quindi anche le $\gamma_{\mu\nu}^{(r)}$ ($\mu, \nu = 1, 2, \dots, p$) sian tutte nulle su M , la convergenza sarà *di grado* $s \geq 2$, se in un intorno di M varranno delle relazioni, analoghe alle (13)

$$g_{\mu} (x^{(r)}) = \sum_{k_1, k_2, \dots, k_s} \gamma_{\mu; k_1, \dots, k_s}^{(r)} g_{k_1} (x) g_{k_2} (x) \dots g_{k_s} (x), \quad (13)_r$$

dove le $\gamma_{\mu; k_1, k_2, \dots, k_s}^{(r)}$ sono ancora funzioni finite e determinate delle x , *non tutte nulle su M* . E nel caso della convergenza lineare le serie (10) sono confrontabili colla serie

$$\sum_{\nu=0}^{\infty} a \left[\frac{\rho}{r} \right]^{\nu}$$

(dove a è un qualunque numero compreso tra a_0 ed 1 e $\left[\frac{\rho}{r} \right]$ indica il massimo intero contenuto in $\frac{\rho}{r}$), la cui convergenza è simile a quella della progressione geometrica; nel caso della convergenza di grado s sono invece

confrontabili colla serie :

$$\sum_0^{\infty} \theta^s \left[\frac{\theta}{r} \right], \text{ con } 0 < \theta < 1$$

la quale è ancora rapidissimamente convergente.

Notiamo infine che, se $p = 1$, non è il caso di occuparsi della convergenza di ordine superiore al primo; se l'algoritmo ha convergenza di un ordine $r > 1$, l'ha ancora del 1.° ordine. Non altrettanto può dirsi per $p > 1$.

6. Sull'ordine ed il grado di convergenza di un algoritmo K abbiamo i teoremi seguenti (*):

a) *Se l'algoritmo K ha (al divisore M) convergenza di ordine r e di grado $s \geq 2$, avrà anche, in convenienti intorno di M , convergenza di qualunque ordine $r' > r$ e di grado non minore di s .*

b) *Se l'algoritmo K ha convergenza dell'ordine r e grado $s \geq 2$, ed insieme dell'ordine r' e del grado $s' \geq 2$, ha anche convergenza dell'ordine $r + r'$ e di grado non minore di $s s'$.*

In particolare, ne segue:

c) *Se l'algoritmo K ha convergenza dell'ordine r e grado $s \geq 2$, ha anche convergenza di qualunque ordine $r q$, multiplo di r , e di grado non minore di s^q .*

Nel caso della convergenza lineare si ha poi:

d) *Se l'algoritmo K ha convergenza lineare degli ordini r, r' , misurata rispettivamente dalle costanti a, a' , ha ancora convergenza (almeno) lineare dell'ordine $r + r'$, misurata da una costante non maggiore di $a a'$.*

In particolare:

e) *Se l'algoritmo K ha convergenza lineare dell'ordine r , misurata dalla costante a , ha anche convergenza (almeno) lineare di ogni ordine $r q$, multiplo di r , e misurata da una costante non maggiore di a^q (**).*

È opportuno fare anche l'osservazione seguente (**).

L'algoritmo K non cambia in modo essenziale, quando alle equazioni (1) che lo definiscono si sostituisca un sistema equivalente, o una trasformazione analoga si eseguisca sulle equazioni (4) del divisore M , o quando infine si faccia nell' $S_n \equiv (x_1 \dots x_n)$ una qualunque trasformazione di variabili.

(*) Cf. A, n.° 9.

(**) Si può ancora aggiungere che: l'algoritmo avrà anche convergenza lineare di qualunque ordine maggiore di un numero r_0 sufficientemente grande (cf. più oltre n.° 19, a)).

(***) Cf. A, n.° 25.

Una qualunque di queste trasformazioni lascia invariati l'ordine r ed il grado s di convergenza dell'algoritmo, quando sia $s > 1$; nel caso della convergenza lineare (di un ordine qualunque) la 1^a e la 3^a trasformazione lasciano invariati l'ordine della convergenza e la costante che la misura; cambiando invece le equazioni (4) che definiscono il divisore M può (se $p > 1$) mutare il minimo ordine di convergenza delle relative serie; quel che rimane invariato è (in una certa misura) il rapporto delle costanti che misurano la convergenza dei successivi ordini.

7. Quali sono le condizioni necessarie e sufficienti, perchè l'algoritmo K definito dalle equazioni (1) abbia, al divisore M , convergenza di ordine r e grado s (e non maggiore) ? (*).

Sia dapprima $s \geq 2$. Si hanno allora come *condizioni necessarie e sufficienti* le $p \left\{ \binom{n+s-1}{n} - 1 \right\}$ congruenze (non tutte però indipendenti)

$$\Gamma_{\mu, t_1, t_2, \dots, t_p}^{(r)} = \frac{\partial^s g_\mu(x^{(v)})}{\partial x_{t_1} \partial x_{t_2} \dots \partial x_{t_p}} \equiv 0 \pmod{g_1, g_2, \dots, g_p} \quad \left(\begin{array}{l} \rho = 1, 2, \dots, s-1 \\ \mu = 1, 2, \dots, p \\ t_1, t_2, \dots, t_p = 1, 2, \dots, n \end{array} \right) \quad (1)$$

insieme colla disuguaglianza

$$\sum_1^p \sum_1^n \sum_{t_1, t_2, \dots, t_p} \left| \frac{\partial^s g_\mu(x^{(v)})}{\partial x_{t_1} \partial x_{t_2} \dots \partial x_{t_p}} \right| \equiv \equiv 0 \pmod{g_1, g_2, \dots, g_p}. \quad (15)_r$$

Si osservi ora, che, per i teoremi fondamentali sulle funzioni implicite, ogni $\Gamma_{\mu, t_1, \dots, t_p}^{(r)}$ (a meno di un denominatore che su M è diverso da zero e finito) si può esprimere come una funzione razionale intera delle derivate delle $g_\mu(x)$ e delle $\Phi_i(x, x')$ di ordine non maggiore di ρ , e nella quale figurano in modo essenziale le derivate delle g_μ e delle $\Phi_i(x, x')$ di ordine ρ . Ne segue che:

Operando sulle g_μ e sulle $\Phi_i(x, x')$ con sole operazioni razionali e di derivazione, è possibile riconoscere se l'algoritmo K ha al divisore M convergenza di un determinato ordine e grado.

Per $s = 2$ i calcoli si eseguono con tutta facilità e danno il risultato seguente. Poniamo

$$\frac{d(\Phi_1 \Phi_2 \dots \Phi_n)}{d(x'_1 \dots x'_{t-1} x_k x'_{t+1} \dots x'_n)} \equiv \Delta_{tk} \pmod{g_1, g_2, \dots, g_p} \quad \left. \begin{array}{l} (t, k = 1, 2, \dots, n) \\ \Delta_{tk}^{(r)} = \sum_1^n \sum_{k_1 k_2 \dots k_{r-1}} \Delta_{tk_1} \Delta_{k_1 k_2} \dots \Delta_{k_{r-2} k_{r-1}} \Delta_{k_{r-1} k} \end{array} \right\} (16)_r$$

(*) Cf. A, n. 11 e 12.

le condizioni necessarie e sufficienti per la convergenza quadratica e di ordine r sono espresse dalle np congruenze:

$$\sum_1^n g_{\mu k} \Delta_{ik}^{(r)} \equiv 0 \pmod{g_1, g_2, \dots, g_p}, \quad (\mu = 1, 2, \dots, p; k = 1, 2, \dots, n). \quad (17),$$

In particolare, per $r = 1$, le condizioni necessarie e sufficienti per la convergenza quadratica del 1.º ordine si scrivono

$$\sum_1^n g_{\mu k} \frac{d(\Phi_1 \Phi_2 \dots \Phi_n)}{d(x'_1 \dots x'_{i-1} x_k x'_{i+1} \dots x'_n)} \equiv 0 \pmod{g_1, g_2, \dots, g_p}. \quad (17)_1$$

Per $s = 1$, quando cioè la convergenza è lineare, si ottengono p disuguaglianze (*), che per brevità omettiamo di scrivere, non occorrendoci nel seguito.

8. Ammettiamo (**) che anche il determinante funzionale $\frac{d(\Phi_1 \Phi_2 \dots \Phi_n)}{d(x_1 x_2 \dots x_n)}$ sia diverso da zero. Allora, sotto condizioni analoghe a quelle poste al n.º 1, le equazioni $\Phi_i(x, x') = 0$ possono risolversi anche rispetto alle $x_1 x_2 \dots x_n$ e definiscono un altro algoritmo, che diciamo *inverso* del primitivo K e indichiamo col simbolo K^{-1} .

Sia ora $p = 1$; se per l'algoritmo K , nell'intorno di M , si ha

$$g(x') = \gamma(x) \cdot g(x),$$

per l'inverso K^{-1} sarà invece:

$$g(x) = \frac{1}{\gamma(x)} \cdot g(x').$$

Consideriamo allora le due equazioni $g(x) = 0$, $|\gamma(x)| = 1$, la seconda delle quali escludiamo sia una conseguenza della prima. Esse definiscono una varietà M_1 ad $n - 2$ dimensioni, che divide la varietà M in due parti M' , M'' (una delle quali può anche mancare), tali che in una di esse, ad es. M' , si ha $|\gamma(x)| < 1$, nell'altra M'' si ha invece $|\gamma(x)| > 1$. In un conveniente intorno di M' è chiaro che converge l'algoritmo K (al divisore g), nè può questo algoritmo convergere ad un punto di M fuori di M' , in quanto

(*) Cf. A, n.º 13, c).

(**) Cf. A, n.º 14.

nella serie $\sum_0^{\infty} g^{(p)}$, quando sia $|\gamma| \geq 1$, i termini rimangono in modulo superiori ad un numero finito, il punto $x^{(p)}$ non può quindi tendere ad un punto di M , nel quale è $g = 0$. Analogamente l'algoritmo K^{-1} converge nell'intorno di M'' e non può convergere ad altri punti di M ; di guisa che, fatta eccezione della varietà M_1 ad $n-2$ dimensioni, su tutta la varietà M (in un suo conveniente intorno) converge uno ed uno solo dei due algoritmi K, K^{-1} .

Sia invece $p > 1$. *I due algoritmi K, K^{-1} non possono aver convergenza (al divisore M) nè dello stesso ordine, nè di ordini distinti.* Questo però non dimostra, a tutto rigore, (come per $p = 1$) che, quando l'algoritmo K converge nell'intorno di M , non possa convergere anche l'algoritmo K^{-1} ; possiamo affermare soltanto che esso *non può soddisfare ad alcuno dei criteri di convergenza da noi assegnati. I due algoritmi K, K^{-1} possono però convergere ambedue quando appartengano a divisori distinti delle funzioni $\varphi_1, \varphi_2, \dots, \varphi_n$ (*)*.

9. È chiaro, da quanto precede, come il caso della convergenza di grado superiore al primo conduca a risultati molto più precisi e notevoli che non quello della convergenza lineare (cf. anche A, n.º 20 e 26). Si presenta quindi naturalmente la questione: *Riconoscere, mediante un numero finito di operazioni, se l'algoritmo K definito dalle equazioni (1) ha ad un divisore $(g_1 g_2 \dots g_p)$ di rango p delle funzioni (2) convergenza (di un qualche ordine) e di grado superiore al primo.*

Per la risoluzione della questione stessa non bastano i criteri di convergenza dati ai n.º 4 e 5; volendo infatti applicare questi criteri, dovremmo cominciare dal verificare se l'algoritmo K ha al divisore $(g_1 \dots g_p)$ convergenza (almeno) quadratica dei successivi ordini; ed è chiaro che un numero qualunque di risultati negativi non autorizza, a priori, alcuna conclusione circa la convergenza o meno dell'algoritmo stesso.

La questione si può però risolvere e, come vedremo, in modo semplice ed elegante. Dimostreremo infatti, nelle pagine seguenti, un teorema, il quale permette di riconoscere, *con un numero finito di operazioni razionali*, se l'algoritmo K ha al divisore $(g_1 g_2 \dots g_p)$ convergenza di grado superiore al primo (e di un ordine qualunque) ed inoltre, quando questo accada, assegna l'ordine *minimo* di convergenza (almeno) quadratica che l'algoritmo possiede; di guisa che quando (come nella maggior parte dei casi si farà per evitare dei calcoli molto lunghi) ci si limiti alla considerazione della sola convergenza quadratica, *nessun altro calcolo è più necessario.*

(*) Cf. anche A, n.º 15-17.

Infine, quando si sappia, per lo stesso teorema, che l'algoritmo non ha al divisore $(g_1 \dots g_x)$ convergenza di grado superiore al primo, le considerazioni stesse conducono ad un notevole criterio (naturalmente soltanto sufficiente) per la convergenza lineare.

III. IL DETERMINANTE CARATTERISTICO DI UN ALGORITMO.

10. Diremo *determinante caratteristico* dell'algoritmo K , definito dalle equazioni (1), il determinante del grado n in ω

$$E(\omega) = \left| \frac{\partial \Phi_i}{\partial x_k} + \omega \frac{\partial \Phi_i}{\partial x'_k} \right|^{(*)}, \quad (i, k = 1, 2, \dots, n). \quad (18)$$

Questa denominazione si giustifica coll'osservare che:

a) Se alle equazioni (1) che definiscono l'algoritmo si sostituisce un sistema equivalente

$$\Psi_i(x, x') = 0 \quad (i = 1, 2, \dots, n) \quad (1')$$

si hanno delle relazioni della forma

$$\Psi_i(x, x') = \sum_1^n \alpha_{ii} \Phi_i(x, x'), \quad (19)$$

dove le α_{ii} sono funzioni determinate delle x ed x' , il cui determinante è $\equiv 0 \pmod{\Phi_1, \Phi_2, \dots, \Phi_n}$. Derivando le (19) otteniamo:

$$\frac{\partial \Psi_i}{\partial x_k} + \omega \frac{\partial \Psi_i}{\partial x'_k} \equiv \sum_1^n \alpha_{ii} \left(\frac{\partial \Phi_i}{\partial x_k} + \omega \frac{\partial \Phi_i}{\partial x'_k} \right) \pmod{\Phi_1, \Phi_2, \dots, \Phi_n}, \quad (i = 1, 2, \dots, n)$$

e quindi anche

$$\left| \frac{\partial \Psi_i}{\partial x_k} + \omega \frac{\partial \Psi_i}{\partial x'_k} \right| \equiv |\alpha_{ii}| \cdot \left| \frac{\partial \Phi_i}{\partial x_k} + \omega \frac{\partial \Phi_i}{\partial x'_k} \right| \pmod{\Phi_1, \Phi_2, \dots, \Phi_n}$$

$$(i, l, k = 1, 2, \dots, n)$$

(*) Il determinante caratteristico di un algoritmo si trova per la prima volta, io credo, nella Memoria di POINCARÉ: *Sur une classe nouvelle de transcendentes uniformes* (pagg. 342 e segg.); fu quindi nuovamente considerato dal sig. LEAU (*Sur les équations fonctionnelles à une ou à plusieurs variables*, Annales de Toulouse, 1897, E, pag. 71 e segg.) e poi dal LEVICIVITA e dal LATTÈS (rispettivamente a pag. 231 e alle pagg. 10 e 38 delle Memorie citate in ultimo).

cioè: il determinante $E(\omega)$ si moltiplica (modd $\Phi_1, \Phi_2, \dots, \Phi_n$) per un fattore diverso da zero ed indipendente da ω , quando alle equazioni di definizione dell'algoritmo si sostituisca un sistema equivalente.

b) Eseguiamo un cambiamento di variabili, mediante le formule:

$$x_i = f_i(\xi_1, \xi_2, \dots, \xi_n), \quad x'_i = f'_i(\xi'_1, \xi'_2, \dots, \xi'_n) \quad (i = 1, 2, \dots, n),$$

essendo i due determinanti funzionali

$$\frac{d(x_1, x_2, \dots, x_n)}{d(\xi_1, \xi_2, \dots, \xi_n)}, \quad \frac{d(x'_1, x'_2, \dots, x'_n)}{d(\xi'_1, \xi'_2, \dots, \xi'_n)}$$

diversi da zero. Dette $\bar{\Phi}_i(\xi, \xi') = 0$ le nuove equazioni dell'algoritmo, sarà

$$\frac{\partial \bar{\Phi}_i}{\partial \xi_k} = \sum_1^n \frac{\partial \Phi_i}{\partial x_l} \frac{\partial x_l}{\partial \xi_k}; \quad \frac{\partial \bar{\Phi}_i}{\partial \xi'_k} = \sum_1^n \frac{\partial \Phi_i}{\partial x'_l} \frac{\partial x'_l}{\partial \xi'_k}, \quad (i, k = 1, 2, \dots, n).$$

Sia ora $M \equiv (g_1, g_2, \dots, g_p)$ un divisore di rango p delle funzioni $\varphi_1, \varphi_2, \dots, \varphi_n$; si avrà allora

$$x'_i \equiv x_i; \quad \xi'_i \equiv \xi_i \pmod{g_1, g_2, \dots, g_p}, \quad (i = 1, 2, \dots, n)$$

e quindi anche

$$\frac{\partial x'_l}{\partial \xi'_k} \equiv \frac{\partial x_l}{\partial \xi_k} \pmod{g_1, g_2, \dots, g_p}, \quad (k, l = 1, 2, \dots, n)$$

e perciò

$$\left| \frac{\partial \bar{\Phi}_i}{\partial \xi_k} + \omega \frac{\partial \bar{\Phi}_i}{\partial \xi'_k} \right| \equiv \left| \frac{\partial \Phi_i}{\partial x_l} + \omega \frac{\partial \Phi_i}{\partial x'_l} \right| \cdot \left| \frac{\partial x_l}{\partial \xi_k} \right| \pmod{g_1, g_2, \dots, g_p};$$

cioè: Mutando in un modo qualunque le variabili, il determinante $E(\omega)$ si moltiplica, rispetto ad un qualunque divisore delle funzioni (2), per il determinante funzionale delle antiche variabili rispetto alle nuove.

In ambedue i casi i divisori elementari del determinante caratteristico rimangono invariati.

11. Sia ancora $M \equiv (g_1, g_2, \dots, g_p)$ un divisore di rango p delle funzioni (2) e proponiamoci di vedere se l'algoritmo K definito dalle (1) converge o meno al divisore M . Poichè, come abbiamo visto, il determinante $E(\omega)$ si moltiplica per un fattore indipendente da ω e diverso da zero, quando alle (1) si sostituisca un sistema equivalente, possiamo in particolare sup-

porre che le equazioni di definizione dell'algoritmo abbiano la forma seguente (*):

$$\left. \begin{aligned} \Phi_\nu(x, x') &= -g_\nu(x') + \sum_1^p \gamma_{\mu\nu}(x) g_\mu(x) = 0, \quad (\nu = 1, 2, \dots, p) \\ \Phi_t(x, x') &= \sum_1^p h_{t\nu} g_\nu(x) + \sum_{k=1}^n \varphi_{tk}(x, x')(x'_k - x_k) = 0, \quad (t = p+1, \dots, n) \end{aligned} \right\} \quad (20)$$

dove le $\gamma_{\mu\nu}(x)$ sono le funzioni già definite al n.º 3 e le $h_{t\nu}$, φ_{tk} sono funzioni determinate dei loro argomenti.

Si ha in questa ipotesi:

$$E(\omega) \equiv \begin{vmatrix} \sum_1^p \gamma_{\mu\nu} g_\nu & -\omega g_{\rho k} \\ \sum_1^p h_{t\nu} g_\nu & +(\omega-1)\varphi_{tk} \end{vmatrix} \pmod{g_1 g_2 \dots g_p}, \quad \begin{matrix} (\nu = 1, 2, \dots, p) \\ (t = p+1, \dots, n) \\ (k = 1, 2, \dots, n) \end{matrix}$$

dove abbiám posto

$$g_{\nu k} \equiv \frac{\partial g_\nu(x)}{\partial x_k} \pmod{g_1, g_2, \dots, g_p}; \quad (\nu = 1, 2, \dots, p; k = 1, 2, \dots, n);$$

si avrà quindi anche

$$E(\omega) \equiv \begin{vmatrix} \gamma_{\mu\nu} - \varepsilon_{\mu\nu} \omega & 0 \\ h_{t\nu} & (\omega-1)\varepsilon_{tk} \end{vmatrix} \cdot \begin{vmatrix} g_{\rho k} \\ \varphi_{tk} \end{vmatrix}, \quad \begin{matrix} (\mu, \nu, \rho = 1, 2, \dots, p) \\ (t, u, v = p+1, \dots, n) \\ (k = 1, 2, \dots, n) \end{matrix} \quad (21)$$

(con $\varepsilon_{ik} = 1$ per $i = k$, $= 0$ per $i \neq k$. Ma avendosi

$$\begin{vmatrix} g_{\rho k} \\ \varphi_{tk} \end{vmatrix} \equiv (-1)^p \frac{d(\Phi_1 \Phi_2 \dots \Phi_n)}{d(x'_1 x'_2 \dots x'_n)} \pmod{g_1, g_2, \dots, g_p},$$

sarà il determinante stesso diverso da zero e quindi (**) i divisori elementari del determinante $E(\omega)$ sono gli stessi del determinante

$$\begin{vmatrix} \gamma_{\mu\nu} - \varepsilon_{\mu\nu} \omega & 0 \\ h_{t\nu} & (\omega-1)\varepsilon_{tk} \end{vmatrix}, \quad \begin{matrix} (\mu, \nu = 1, 2, \dots, p) \\ (t, u = p+1, \dots, n) \end{matrix} \quad (21')$$

(*) Cf. A, n.º 7, formula (30.*).

(**) Quando due matrici simili A , A' si deducono l'una dall'altra componendole con un determinante B diverso da zero, i minori di un ordine qualunque di una delle due matrici sono funzioni lineari omogenee dei minori dello stesso ordine dell'altra. Ne segue subito l'asserzione del testo.

Per risparmiar delle considerazioni troppo minute, e che, almeno per la questione che ci occupa, non avrebbero alcun interesse (cf. n.º 20) supporremo che il determinante di ordine p

$$F(\omega) = |\gamma_{\mu\nu} - \varepsilon_{\mu\nu} \omega|, \quad (\mu, \nu = 1, 2, \dots, p) \quad (22)$$

non si annulli (mod $g_1 \dots g_p$) per $\omega = 1$.

Dalle (21), (21') abbiamo allora il teorema:

Se $M \equiv (g_1 g_2 \dots g_p)$ è un divisore di rango p delle funzioni $\varphi_1, \varphi_2, \dots, \varphi_n$, il determinante $E(\omega)$ caratteristico dell'algoritmo ha (mod $g_1 g_2 \dots g_p$) $n - p$ divisori elementari uguali ad $\omega - 1$ ed inoltre tutti i divisori elementari del determinante caratteristico della sostituzione lineare (9) (che supponiamo diversi da una potenza di $\omega - 1$)

$$F(\omega) = |\gamma_{\mu\nu} - \varepsilon_{\mu\nu} \omega|, \quad (\mu, \nu = 1, 2, \dots, p).$$

Si ha insieme identicamente

$$E(\omega) \equiv (-1)^n \frac{d(\Phi_1 \Phi_2 \dots \Phi_n)}{d(x'_1 x'_2 \dots x'_n)} (\omega - 1)^{n-p} F(\omega), \quad (\text{mod } g_1, g_2, \dots, g_p). \quad (21'')$$

IV. EQUAZIONI NORMALI DEL DIVISORE $M \equiv (g_1 g_2 \dots g_p)$.

12. Abbiamo osservato al n.º 3 che nel passare, mediante l'algoritmo K , da un punto x al suo conseguente x' , le p funzioni $g_1 g_2 \dots g_p$ subiscono una sostituzione lineare data dalla formula

$$g_\mu(x') \equiv \sum_{\nu=1}^n \gamma_{\mu\nu} g_\nu(x). \quad (9)$$

Consideriamo la sostituzione

$$\Gamma \equiv (\gamma_{\mu\nu}) \quad (\text{mod } g_1 g_2 \dots g_p), \quad (\mu, \nu = 1, 2, \dots, p).$$

Il suo *determinante caratteristico* è il determinante (22), e se questo si decompone nei suoi divisori elementari al modo seguente

$$F(\omega) \equiv (-1)^n (\omega - \omega_1)^{p_1} (\omega - \omega_2)^{p_2} \dots (\omega - \omega_r)^{p_r}, \quad (p_1 + p_2 + \dots + p_r = p) \quad (22')$$

è noto (*), che si può determinare una conveniente trasformata $\Gamma' \equiv T^{-1} \Gamma T$ della Γ (mediante una trasformazione lineare T non degenere) la quale abbia la forma

$$\Gamma' \equiv \Gamma_1 \cdot \Gamma_2 \dots \Gamma_r \pmod{g_1 g_2 \dots g_p}, \quad (23)$$

dove le $\Gamma_1, \Gamma_2, \dots, \Gamma_r$ sono sostituzioni (tra loro permutabili) che operano su variabili diverse, e delle quali la Γ_i opera su p_i variabili ed ha $\pmod{g_1 \dots g_p}$ la forma normale

$$\Gamma_i \equiv \begin{pmatrix} \omega_i & 0 & 0 & \dots & 0 & 0 \\ 1 & \omega_i & 0 & \dots & 0 & 0 \\ 0 & 1 & \omega_i & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & \omega_i \end{pmatrix}^{(p_i)}. \quad (24)$$

Vogliamo ora dimostrare che si può, ed in infiniti modi, determinare un sistema $(G_1 G_2 \dots G_p)$ di funzioni, *equivalente* al sistema $(g_1 g_2 \dots g_p)$, il quale, nel passaggio dal punto x al conseguente x' (mediante l'algoritmo K), subisce una sostituzione lineare Γ' che ha $\pmod{g_1 g_2 \dots g_p}$ la forma normale espressa dalle (23) e (24). Diremo per brevità che in tal caso le $(G_1 G_2 \dots G_p)$ costituiscono un *sistema normale* di equazioni del divisore $M \equiv (g_1 \dots g_p)$.

13. È evidente, per la sua definizione, che un sistema normale $(G_1 G_2 \dots G_p)$ deve essere indipendente dal particolare sistema $(g_1 g_2 \dots g_p)$ di equazioni del divisore M che si considera. È quindi naturale pensare (e così è veramente) che un tal sistema possa definirsi in guisa affatto indipendente dal sistema $(g_1 g_2 \dots g_p)$.

Sia per questo

$$G = G(x) = \sum_1^n \xi_i(x) \cdot \varphi_i(x) \quad (25)$$

una qualunque combinazione lineare delle funzioni $\varphi_1, \varphi_2, \dots, \varphi_n$ definite dalle (2*) del n.º 1, essendo le $\xi_i(x)$ funzioni, per ora indeterminate, delle x , e indichiamo con

$$G' = G(x') = \sum_1^n \xi_i(x') \varphi_i(x') \quad (25')$$

la funzione *conseguente* della G mediante l'algoritmo K .

(*) Cf. ad es.: HAMBURGER, *Giornale di Crelle*, Vol. 76, pagg. 113-125 ed anche: MUTH, *Theorie und Anwendung der Elementarteiler* (Lipsia-Teubner, 1899, pag. 82).

Si ha ora identicamente, per il teorema del valor medio (*):

$$\Phi_i(x, x') = \varphi_i(x) + \sum_1^n \frac{\partial \bar{\Phi}_i}{\partial x'_k} (x'_k - x_k) = \varphi_i(x') + \sum_1^n \frac{\partial \bar{\Phi}_i}{\partial x_k} (x_k - x'_k) = 0,$$

dove le derivate $\frac{\partial \bar{\Phi}_i}{\partial x'_k}$, (e le $\frac{\partial \bar{\Phi}_i}{\partial x_k}$) sono prese in un conveniente punto intermedio tra i punti (x, x') , (x, x) (e (x, x') , (x', x')); si avrà quindi

$$G = - \sum_1^n \mathfrak{S}_i(x'_k - x_k) \frac{\partial \bar{\Phi}_i}{\partial x'_k}, \quad G' = \sum_1^n \mathfrak{S}'_i(x'_k - x_k) \frac{\partial \bar{\Phi}_i}{\partial x_k},$$

od anche, poichè

$$\frac{\partial \bar{\Phi}_i}{\partial x'_k} \equiv \frac{\partial \Phi_i}{\partial x'_k}, \quad \frac{\partial \bar{\Phi}_i}{\partial x_k} \equiv \frac{\partial \Phi_i}{\partial x_k}, \quad \mathfrak{S}'_i \equiv \mathfrak{S}_i \pmod{g_1 g_2 \dots g_p},$$

e per le (8) del n.º 3 sarà:

$$\left. \begin{aligned} G &\equiv - \sum_1^n \mathfrak{S}_i(x'_k - x_k) \frac{\partial \Phi_i}{\partial x'_k}, \\ G' &\equiv \sum_1^n \mathfrak{S}_i(x'_k - x_k) \frac{\partial \Phi_i}{\partial x_k}. \end{aligned} \right\} \pmod{(g_1 g_2 \dots g_p)^2} (**) \quad (26)$$

Supponiamo ora che, *indicando con ω una funzione determinata dei punti della varietà M* , si abbia

$$G' \equiv \omega G \pmod{(g_1 g_2 \dots g_p)^2}; \quad (27)$$

per le (26) questa diventa:

$$\sum_1^n \mathfrak{S}_i(x'_k - x_k) \left\{ \frac{\partial \Phi_i}{\partial x_k} + \omega \frac{\partial \Phi_i}{\partial x'_k} \right\} \equiv 0 \pmod{(g_1 g_2 \dots g_p)^2}. \quad (27^*)$$

Quando inversamente questa congruenza sia soddisfatta, per le (26), varrà ancora la (27), cioè la funzione $G = \sum \mathfrak{S}_i \varphi_i$, *ove non sia* $\equiv 0 \pmod{(g_1 g_2 \dots g_p)^2}$, si riprodurrà, per l'algoritmo K , moltiplicata per $\omega \pmod{(g_1 g_2 \dots g_p)^2}$.

(*) Cf. A, n.º 5.

(**) Col simbolo M^2 o $(g_1 g_2 \dots g_p)^2$ indichiamo il modulo *quadrato* di $(g_1 g_2 \dots g_p)$, ottenuto cioè *componendo* due volte il modulo $(g_1 g_2 \dots g_p)$. (Cf. anche A, n.º 2).

14. Poniamo per brevità

$$\Lambda_k \equiv \sum_1^n \mathfrak{S}_i \left\{ \frac{\partial \Phi_i}{\partial x_k} + \omega \frac{\partial \Phi_i}{\partial x'_k} \right\} \pmod{g_1 g_2 \dots g_p}, \quad (k = 1, 2, \dots, n); \quad (28)$$

la (27*) si scrive allora

$$\sum_1^n \Lambda_k (x'_k - x_k) \equiv 0 \pmod{(g_1 g_2 \dots g_p)^2} \quad (27^{**})$$

e, per le (8) del n.º 3, sarà soddisfatta quando si abbia

$$\Lambda_k \equiv 0 \pmod{g_1 g_2 \dots g_p}, \quad (k = 1, 2, \dots, n). \quad (29)$$

Dalle (28) poi è chiaro, che si potrà soddisfare alle (29) con valori delle \mathfrak{S}_i non tutti nulli $\pmod{g_1 g_2 \dots g_p}$ allora ed allora soltanto che ω sia una radice della equazione $E(\omega) \equiv 0 \pmod{g_1 g_2 \dots g_p}$, che si ha annullando $\pmod{g_1 g_2 \dots g_p}$ il determinante caratteristico dell'algoritmo K .

Sia inversamente ω una radice della equazione $E(\omega) \equiv 0 \pmod{g_1 g_2 \dots g_p}$, e si determinino le \mathfrak{S}_i (non tutte nulle $\pmod{g_1 \dots g_p}$), in guisa da soddisfare alle congruenze (29); varrà allora anche la (27*) e quindi la (27), purchè non sia $G \equiv 0 \pmod{(g_1 g_2 \dots g_p)^2}$.

Vediamo allora se può essere per la funzione così determinata

$$G = \sum \mathfrak{S}_i \Phi_i \equiv 0 \pmod{(g_1 g_2 \dots g_p)^2}.$$

In questa ipotesi, derivando rispetto ad una qualunque x , si avrà:

$$\frac{\partial G}{\partial x_k} \equiv \sum \mathfrak{S}_i \frac{\partial \Phi_i}{\partial x_k} \equiv 0 \pmod{g_1 g_2 \dots g_p}, \quad (k = 1, 2, \dots, n).$$

Ma dalle congruenze (2*) del n.º 1 si ha subito:

$$\frac{\partial \Phi_i}{\partial x_k} \equiv \frac{\partial \Phi_i}{\partial x_k} + \frac{\partial \Phi_i}{\partial x'_k} \pmod{g_1 g_2 \dots g_p}, \quad (k = 1, 2, \dots, n)$$

e sostituendo nelle precedenti:

$$\frac{\partial G}{\partial x_k} \equiv \sum_1^n \mathfrak{S}_i \left(\frac{\partial \Phi_i}{\partial x_k} + \frac{\partial \Phi_i}{\partial x'_k} \right) \equiv 0 \pmod{g_1 g_2 \dots g_p}, \quad (k = 1, 2, \dots, n).$$

Sottraendo queste congruenze dalle (29), otteniamo infine

$$(\omega - 1) \sum_1^n \mathfrak{S}_i \frac{\partial \Phi_i}{\partial x'_k} \equiv 0 \pmod{g_1 g_2 \dots g_p}, \quad (k = 1, 2, \dots, n);$$

se dunque supponiamo che ω sia diversa dall'unità, soddisfi cioè, oltrechè alla equazione $E(\omega) \equiv 0$, all'altra $F(\omega) \equiv 0$, potrà essere $G \equiv 0 \pmod{(g_1 g_2 \dots g_p)^2}$, solo quando si abbia

$$\sum_{i=1}^n \mathfrak{S}_i \frac{\partial \Phi_i}{\partial x'_k} \equiv 0 \pmod{g_1 g_2 \dots g_p}, \quad (k = 1, 2, \dots, n)$$

cioè, poichè il determinante $\frac{d(\Phi_1 \Phi_2 \dots \Phi_n)}{d(x'_1 x'_2 \dots x'_n)}$ ha $\pmod{g_1 g_2 \dots g_p}$ la caratteristica n , quando sia

$$\mathfrak{S}_i \equiv 0 \pmod{g_1 g_2 \dots g_p}, \quad (i = 1, 2, \dots, n),$$

contro la nostra ipotesi, che non tutte le \mathfrak{S}_i sian nulle $\pmod{g_1 g_2 \dots g_p}$.

Adunque:

Ad ogni radice ω , diversa dall'unità, dell'equazione

$$E(\omega) \equiv 0 \pmod{g_1 g_2 \dots g_p}$$

che si ha annullando $\pmod{g_1 \dots g_p}$ il determinante caratteristico dell'algoritmo corrispondono delle funzioni G , che per l'algoritmo si riproducono moltiplicate per ω .

15. Possiamo precisare e completare il risultato superiore, ricorrendo alla teoria delle sostituzioni lineari.

Sia per questo ω_1 una radice (diversa da 1) della equazione

$$E(\omega) \equiv 0 \pmod{g_1 g_2 \dots g_p},$$

e siano

$$(\omega - \omega_1)^{e_1}, \quad (\omega - \omega_1)^{e_2}, \dots, \quad (\omega - \omega_1)^{e_n}, \quad (e_1 \geq e_2 \geq \dots \geq e_n) \quad (30)$$

i divisori elementari del determinante $E(\omega)$ corrispondenti alla radice ω_1 ; diciamo h_1 il numero degli esponenti e uguali ad 1, h_2 il numero di quelli uguali a 2, ..., h_{e_1} il numero di quelli tra essi esponenti che hanno il valore massimo e_1 (*).

Poniamo inoltre, per brevità

$$a_{ik} \equiv \frac{\partial \Phi_i}{\partial x_k} + \omega_1 \frac{\partial \Phi_i}{\partial x'_k}, \quad b_{ik} \equiv \frac{\partial \Phi_i}{\partial x'_k} \pmod{g_1 g_2 \dots g_p} \quad (31)$$

(*) Quando non vi siano esponenti e uguali ad un valore t , porremo $h_t = 0$.

e consideriamo i sistemi di equazioni lineari omogenee

$$(a) \left. \begin{aligned} \sum_1^n a_{ik} \mathfrak{S}_i &\equiv 0, & (b_u) \sum a_{ik} \mathfrak{S}_i^{(u)} + \sum_1^n b_{ik} \mathfrak{S}_i^{(u-1)} &\equiv 0 \pmod{g_1 g_2 \dots g_p} \\ (k = 1, 2, \dots, n; & u = 1, 2, \dots, e_1 - 1, \mathfrak{S}_i^{(0)} = \mathfrak{S}_i). \end{aligned} \right\} (32)$$

Diremo che una soluzione $(\mathfrak{S}_1, \mathfrak{S}_2, \dots, \mathfrak{S}_n)$ delle (32 a) è di *rango* t , quando essa renda compatibili i primi $t-1$ sistemi (32, b_u), ma non i primi t , quando cioè è possibile determinare le $\mathfrak{S}_i^{(u)}$ ($u = 1, 2, \dots, t-1, i = 1, 2, \dots, n$) in guisa che insieme alle (32 a) sian soddisfatti anche i primi $t-1$ sistemi (32, b_u), ma non i primi t . Si può allora dimostrare che tra le ∞^h soluzioni del sistema (32 a) possono determinarsene h_1 (e non più) linearmente indipendenti di 1.° rango, h_2 di 2.° rango, ..., h_{e_1} di rango e_1 ; e queste $h_1 + \dots + h_{e_1} = h$ soluzioni formano tutte insieme un sistema *fondamentale* di soluzioni delle (32) stesse.

Se $(\mathfrak{S}_1, \mathfrak{S}_2, \dots, \mathfrak{S}_n)$ è una soluzione di rango t delle (32 a) si soddisfa ai primi $t-1$ sistemi (32, b_u) ponendo:

$$\mathfrak{S}_i^{(u)} = \left(\frac{\partial^u}{\partial \omega^u} \mathfrak{S}_i \right)_{\omega=\omega_1}, \quad (i = 1, 2, \dots, n), \quad (u = 1, 2, \dots, t-1), \quad (33)$$

dove s'intende che una qualunque \mathfrak{S}_i va pensata come funzione di ω , come variabile, e va quindi fatto $\omega = \omega_1$ dopo la derivazione.

Sia ancora $(\mathfrak{S}_1, \mathfrak{S}_2, \dots, \mathfrak{S}_n)$ una soluzione di rango t delle (32 a); e si ponga

$$G = \sum_1^n \mathfrak{S}_i \varphi_i, \quad G^{(u)} = \sum_1^n \mathfrak{S}_i^{(u)} \varphi_i = \left(\frac{\partial^u G}{\partial \omega^u} \right)_{\omega=\omega_1}, \quad (u = 1, 2, \dots, t-1); \quad (34)$$

la $G(x)$ e quindi anche le $G^{(u)}$ non saranno nulle $\pmod{(g_1 g_2 \dots g_p)^2}$ e costituiscono un sistema di t funzioni linearmente indipendenti $\pmod{g_1 \dots g_p}$, per le quali si ha identicamente

$$\left. \begin{aligned} G(x') &\equiv \omega_1 G(x), & G^{(u)}(x') &\equiv G^{(u-1)}(x) + \omega_1 G^{(u)}(x) \pmod{(g_1 g_2 \dots g_p)^2} \\ (u = 1, 2, \dots, t-1). \end{aligned} \right\} (35)$$

Operando in tal guisa, per ogni valore di t , su ciascuna delle h_t soluzioni indipendenti di rango t delle (32) stesse, otteniamo $h_1 + 2h_2 + \dots + e_1 h_{e_1} = \nu_1$ (dove ν_1 è la molteplicità della radice ω_1 per la $E(\omega) \equiv 0$) funzioni lineari omogenee (indipendenti) delle $\varphi_1, \varphi_2, \dots, \varphi_n$, le quali per l'algoritmo K subiscono $\pmod{g_1 g_2 \dots g_p}$ la sostituzione lineare corrispondente alle formule (35).

Siano ora $\omega_1, \omega_2, \dots, \omega_\lambda$ le radici della equazione $E(\omega) \equiv 0 \pmod{g_1 g_2 \dots g_p}$ diverse dall'unità (cioè le radici diverse della equazione

$$F(\omega) \equiv 0 \pmod{g_1 g_2 \dots g_p},$$

siano $\nu_1, \nu_2, \dots, \nu_\lambda$ le loro molteplicità, e si determinino nel modo sopra descritto $\nu_1, \nu_2, \dots, \nu_\lambda$ funzioni lineari omogenee delle $\varphi_1 \varphi_2 \dots \varphi_n$. Queste $\nu_1 + \nu_2 + \dots + \nu_\lambda = p$ funzioni G delle $g_1 g_2 \dots g_p$, si dimostra facilmente, sono linearmente indipendenti e, per le (35) e analoghe, costituiscono un sistema normale di equazioni del divisore $M \equiv (g_1 g_2 \dots g_p)$.

La nostra asserzione è così dimostrata.

V. CONDIZIONI NECESSARIE E SUFFICIENTI PER LA CONVERGENZA
DI GRADO SUPERIORE AL PRIMO.

16. Perchè l'algoritmo K definito dalle (1) abbia al divisore $M \equiv (g_1 g_2 \dots g_p)$ convergenza (di un qualunque ordine) di grado superiore al primo, è necessario e sufficiente che esso abbia convergenza quadratica di un certo ordine (n.º 9). Converrà quindi vedere quando questo accade.

Osserviamo perciò, che, come nel passare dal punto x al conseguente x' le g_μ subiscono la sostituzione lineare $\Gamma \equiv (\gamma_{\mu\nu})$, ($\mu, \nu = 1, 2, \dots, p$) data dalle (9) del n.º 3, in guisa del tutto simile, nel passare dal punto x al suo ρ -conseguente $x^{(\rho)}$, (con ρ qualunque) le g_μ subiscono una sostituzione lineare $\Gamma^{(\rho)} \equiv (\gamma_{\mu\nu}^{(\rho)})$, data dalle (9) $_\rho$ del n.º 5; e l'algoritmo K avrà, al divisore $(g_1 g_2 \dots g_p)$ convergenza quadratica dell'ordine ρ allora ed allora soltanto che la sostituzione $\Gamma^{(\rho)}$ sia identicamente nulla, abbia cioè nulli tutti i suoi elementi $\pmod{g_1 g_2 \dots g_p}$. Ma, per la definizione stessa della sostituzione $\Gamma^{(\rho)}$, si ha identicamente (*):

$$\Gamma^{(\rho)} \equiv \Gamma^\rho \pmod{g_1 g_2 \dots g_p};$$

quindi perchè l'algoritmo K abbia al divisore $(g_1 g_2 \dots g_p)$ convergenza quadratica (di un qualche ordine), è necessario e sufficiente che una conveniente potenza della sostituzione Γ sia identicamente nulla $\pmod{g_1 g_2 \dots g_p}$; ed il minimo esponente ρ pel quale si avrà, in questa ipotesi, $\Gamma^\rho \equiv 0 \pmod{g_1 g_2 \dots g_p}$

(*) Cf. A, n.º 8.

sarà uguale al minimo ordine, per il quale l'algoritmo K ha al divisore $(g_1 g_2 \dots g_p)$ convergenza almeno quadratica.

Consideriamo in particolare un sistema normale $(G_1 G_2 \dots G_p)$ di equazioni del divisore M . Per esso la Γ ha (modd $g_1 g_2 \dots g_p$) la forma normale data dalle (23) e (24) del n.° 12, e si avrà, per qualunque ρ :

$$\Gamma^\rho \equiv \Gamma_1^\rho \cdot \Gamma_2^\rho \dots \Gamma_r^\rho \quad (23)_\rho$$

ed insieme, come si ha subito per induzione dalle (24):

$$\Gamma_i^\rho \equiv \left(\binom{\rho}{s-t} \omega_i^{\rho-s+t} \right) \quad (s, t = 1, 2, \dots, p_i; i = 1, 2, \dots, r) \quad (24)_\rho$$

(dove si debbono fare uguali allo zero i coefficienti binomiali con indice inferiore negativo); e potrà essere $\Gamma^\rho \equiv 0 \pmod{g_1 g_2 \dots g_p}$ soltanto quando, per tutti i valori di i , sia $\Gamma_i^\rho \equiv 0 \pmod{g_1 \dots g_p}$. Ma, per le (24) $_\rho$, è per questo necessario e sufficiente che si abbia $\omega_i \equiv 0 \pmod{g_1 g_2 \dots g_p}$ e quando questa condizione sia soddisfatta, è $\Gamma_i^\rho \equiv 0 \pmod{g_1 g_2 \dots g_p}$, per ogni valore di ρ maggiore od uguale a p_i .

Ricordando la (22'), abbiamo il teorema:

a) *Condizione necessaria e sufficiente perchè l'algoritmo K definito dalle equazioni (1) appartenga al divisore, di rango p , $M \equiv (g_1 g_2 \dots g_p)$ ed abbia a questo divisore convergenza di un grado superiore al primo (di un ordine qualunque) è che tutti i divisori elementari del determinante $F(\omega)$ siano (modd $g_1 \dots g_p$) uguali ad una potenza di ω .*

In questa ipotesi, se il determinante $F(\omega)$ decomposto nei suoi divisori elementari è

$$\left. \begin{aligned} F(\omega) &\equiv (-1)^p \omega^{p_1} \cdot \omega^{p_2} \dots \omega^{p_r} \pmod{g_1 \dots g_p} \\ (p_1 \geq p_2 \geq \dots \geq p_r \geq 1; p_1 + p_2 + \dots + p_r = p) \end{aligned} \right\} \quad (35)$$

l'algoritmo ha al divisore $(g_1 g_2 \dots g_p)$ convergenza quadratica dell'ordine p_1 e non di un ordine minore.

Per la relazione, dimostrata al n.° 11, tra i due determinanti $E(\omega)$, $F(\omega)$ il teorema superiore può enunciarsi al modo seguente:

b) *Condizione necessaria e sufficiente perchè l'algoritmo K definito dalle equazioni (1) appartenga al divisore di rango p $(g_1 g_2 \dots g_p)$ ed abbia a questo divisore convergenza di grado superiore al primo (di un ordine qualunque) è che il determinante $E(\omega)$ caratteristico dell'algoritmo (il quale ha (modd $g_1 g_2 \dots g_p$)*

$n - p$ divisori elementari uguali ad $\omega - 1$ abbia (modd $g_1 g_2 \dots g_p$) gli altri divisori elementari tutti uguali ad una potenza di ω .

In questa ipotesi, il massimo esponente di un divisore elementare del determinante $E(\omega)$ (modd $g_1 g_2 \dots g_p$) uguale ad una potenza di ω dà l'ordine minimo, per il quale l'algoritmo K ha convergenza quadratica (ed eventualmente di grado maggiore) al divisore (g_1, g_2, \dots, g_p) .

Possiamo anche dire evidentemente:

b') Condizione necessaria e sufficiente perchè l'algoritmo K definito dalle (1) abbia al divisore $(g_1 g_2 \dots g_p)$ convergenza di grado superiore al primo (di un ordine qualunque) è che l'equazione caratteristica dell'algoritmo $E(\omega) \equiv 0$ (modd $g_1 g_2 \dots g_p$) abbia (oltre la radice 1 multipla di ordine $n - p$) la radice zero multipla dell'ordine p .

In questa ipotesi se ω^{p-p_1} (con $1 \leq p_1 \leq p$) è la più alta potenza di ω che divide tutti i minori di ordine $n - 1$ del determinante $E(\omega)$ (modd $g_1 g_2 \dots g_p$), sarà p_1 l'ordine minimo, per il quale l'algoritmo K ha convergenza quadratica (ed eventualmente in grado maggiore) al divisore $(g_1 g_2 \dots g_p)$.

17. Il teorema b) o b') permette, come abbiamo affermato al n.º 9, di riconoscere con un numero finito di derivazioni e di operazioni razionali, se l'algoritmo (1) ha o no al divisore $(g_1 g_2 \dots g_p)$ convergenza (di qualunque ordine e) di grado superiore al primo. Si deducono inoltre da esso corollari notevoli.

a) L'esponente di un divisore elementare del determinante $F(\omega)$ non può evidentemente superare p . Ne segue:

Qualunque algoritmo, il quale appartenga al divisore $(g_1 g_2 \dots g_p)$ delle (2) ed abbia a questo divisore convergenza (di qualunque ordine) di grado superiore al primo, ha certamente convergenza quadratica di ordine non maggiore di p .

Questo limite massimo non può abbassarsi; se infatti le equazioni dell'algoritmo sono

$$\begin{cases} g_1(x') \equiv 0; & g_{i+1}(x') \equiv g_i(x) \pmod{(g_1 g_2 \dots g_p)^2} & (i = 1, 2, \dots, p-1) \\ \sum_1^p h_{t_1}(x) g_1(x) + \sum_1^n \varphi_{t_k}(x, x') (x'_k - x_k) = 0 & & (t = p+1, \dots, n) \end{cases}$$

l'algoritmo ha convergenza quadratica dell'ordine p e non minore.

b) Consideriamo più particolarmente il caso che l'algoritmo abbia al divisore M convergenza quadratica del 1.º ordine. Perchè questo accada, è necessario e sufficiente che il determinante caratteristico $E(\omega)$ abbia

(modd $g_1 g_2 \dots g_p$) p divisori elementari uguali ad ω . Ora, per teoremi noti della teoria dei divisori elementari (*), è per questo necessario e sufficiente che tutti i minori di ordine $n - p + 1$ del determinante stesso contengano il fattore ω linearmente, o, in altre parole, che tutti i minori di ordine $n - p + 1$ del determinante $\frac{d(\Phi_1 \Phi_2 \dots \Phi_n)}{d(x_1 x_2 \dots x_n)}$ siano $\equiv 0 \pmod{g_1, g_2, \dots, g_p}$. I minori di ordine $n - p$ di questo determinante funzionale non possono invece esser tutti nulli (modd $g_1 \dots g_p$) (altrimenti i minori di ordine $n - p + 1$ sarebbero divisibili per ω^2 almeno). Il determinante $\frac{d(\Phi_1 \Phi_2 \dots \Phi_n)}{d(x_1 x_2 \dots x_n)}$ avrà quindi

(modd $g_1 g_2 \dots g_p$) la caratteristica $n - p$. Abbiamo così il teorema notevole:

Condizione necessaria e sufficiente perchè l'algoritmo definito dalle (1) abbia al divisore ($g_1 g_2 \dots g_p$) convergenza (almeno) quadratica del 1.° ordine, è che il determinante funzionale

$$\frac{d(\Phi_1 \Phi_2 \dots \Phi_n)}{d(x_1 x_2 \dots x_n)}$$

abbia (modd $g_1 g_2 \dots g_p$) caratteristica $n - p$.

Per un noto teorema di KRONECKER, questo importa p^2 condizioni indipendenti, che debbono naturalmente equivalere alle congruenze (17)₁ del n.° 7 (**).

c) Per un altro teorema sui divisori elementari (***) abbiamo anche:

Condizione necessaria e sufficiente perchè l'algoritmo definito dalle (1) abbia al divisore ($g_1 g_2 \dots g_p$) convergenza quadratica del 1.° ordine, è che il determinante caratteristico $E(\omega)$ ammetta (modd $g_1 g_2 \dots g_p$) il fattore ω^p ed i suoi minori di ordine $n - 1$ ammettano il fattore ω^{p-1} .

(*) Cf. МУТН, loc. cit., pag. 4.

(**) Questa equivalenza risulta immediatamente dalle (43) di A, n. 13, facendo in esse $r = 1$. Un'altra dimostrazione, diretta, del teorema superiore si ha dalla (21) del n.° 11. Facendo in questa $\omega = 0$ si ottiene

$$\frac{d(\Phi_1 \Phi_2 \dots \Phi_n)}{d(x_1 x_2 \dots x_n)} \equiv (-1)^n \frac{d(\Phi_1 \Phi_2 \dots \Phi_n)}{d(x'_1 x'_2 \dots x'_n)} \left| \begin{array}{cc} \gamma_{\mu\nu} & 0 \\ h_{t\nu} & \varepsilon_{tn} \end{array} \right| \pmod{g_1 g_2 \dots g_p} \begin{pmatrix} i, k = 1, 2, \dots, n \\ \mu, \nu = 1, 2, \dots, p \\ t, u = p + 1, \dots, n \end{pmatrix},$$

e quindi la caratteristica del determinante $\frac{d(\Phi_1 \Phi_2 \dots \Phi_n)}{d(x_1 x_2 \dots x_n)} \pmod{g_1 \dots g_p}$ supera di $n - p$ unità quella del determinante $|\gamma_{\mu\nu}|$. Se questo si suppone identicamente nullo, si ha il teorema superiore.

(***) Cf. МУТН, loc. cit., pag. 12.

d) Se in b) supponiamo sia $p = 1$, abbiamo :

Perchè un algoritmo abbia convergenza (almeno) quadratica (*) ad un divisore g di rango uno, è necessario e sufficiente che si abbia

$$\frac{d(\Phi_1 \Phi_2 \dots \Phi_n)}{d(x_1 x_2 \dots x_n)} \equiv 0 \pmod{g}.$$

e) Supponiamo invece che sia $p = n$, e quindi si tratti di approssimare, iterando, una radice semplice (poichè $\frac{d(g_1 g_2 \dots g_n)}{d(x_1 x_2 \dots x_n)} \neq 0$) di un sistema di n equazioni in n incognite: $g_1 = 0, g_2 = 0, \dots, g_n = 0$. In questo caso abbiamo :

Perchè l'algoritmo abbia convergenza quadratica (del 1.° ordine) ad un divisore di rango n ($g_1 g_2 \dots g_n$) è necessario e sufficiente che si abbiano le n^2 congruenze

$$\frac{\partial \Phi_i}{\partial x_k} \equiv 0 \pmod{g_1 g_2 \dots g_n}, \quad (i, k = 1, 2, \dots, n).$$

Se ancora più particolarmente supponiamo che le equazioni che definiscono l'algoritmo abbiano la forma

$$\Phi_i = -x'_i + x_i + \sum_1^n h_{iv}(x) g_v(x) = 0 \quad (1^*)$$

(cioè l'algoritmo sia esplicito, e sia posto in evidenza il divisore cui l'algoritmo appartiene), le congruenze superiori diventano :

$$\varepsilon_{ik} + \sum_1^n h_{iv}(x) g_{vk}(x) \equiv 0 \pmod{g_1 g_2 \dots g_n};$$

donde, posto :

$$\frac{d(g_1 g_2 \dots g_n)}{d(x_1 x_2 \dots x_n)} \equiv G, \quad G_{ik} \equiv \frac{\partial \log G}{\partial g_{ik}} \pmod{g_1 g_2 \dots g_n}, \quad (i, k = 1, 2, \dots, n)$$

si trae immediatamente che l'algoritmo ha la forma :

$$x'_i \equiv x_i - \sum_1^n G_{vi} g_v \pmod{(g_1 g_2 \dots g_n)^2}, \quad (i = 1, 2, \dots, n). \quad (37)$$

Facendo in particolare

$$x'_i = x_i - \sum_1^n G_{vi} g_v, \quad (i = 1, 2, \dots, n) \quad (37^*)$$

(*) Naturalmente del 1.° ordine (cfr. il n.° 5).

si hanno le formule dell'algoritmo di NEWTON, generalizzato ad un sistema di n equazioni in n incognite. Questo algoritmo ha dunque convergenza quadratica per ogni radice semplice del sistema di equazioni considerato; di più qualunque algoritmo che abbia la forma esplicita (1*) e convergenza quadratica di 1.° ordine ad una radice (semplice) del sistema stesso, differisce da quello di NEWTON per termini che hanno nelle g un grado maggiore od uguale al secondo, ma sono del resto affatto arbitrari (*).

f) Poniamo in b) $p = n - 1$; il determinante funzionale $\frac{d(\Phi_1 \Phi_2 \dots \Phi_n)}{d(x_1 x_2 \dots x_n)}$ dovrà avere (modd $g_1 g_2 \dots g_{n-1}$) la caratteristica 1. Supponiamo in particolare che si abbia $g_i(x) = x_i - x_{i+1}$ ($i = 1, 2, \dots, n - 1$) e l'algoritmo sia simmetrico, cioè le $\Phi_i(x, x')$ siano simmetriche nelle $x_1 x_2 \dots x_n$ e per $x_i = x'_i = X$ (con X qualunque) si annullino tutte identicamente. Per la simmetria delle Φ_i nelle $x_1 x_2 \dots x_n$ si ha allora:

$$\frac{\partial \Phi_i}{\partial x_h} \equiv \frac{\partial \Phi_i}{\partial x_k} \pmod{x_i - x_{i+1}, x'_r - x_r}$$

$$(i, h, k, r = 1, 2, \dots, n; l = 1, 2, \dots, n - 1)$$

e quindi il determinante $E(\omega)$ ha effettivamente la caratteristica 1. Ne segue il risultato notevole:

Qualunque algoritmo simmetrico in n variabili, che appartenga al divisore di rango $n - 1$: $g_i = x_i - x_{i+1}$ ($i = 1, 2, \dots, n - 1$) ha convergenza quadratica del 1.° ordine.

Tali sono ad es.: l'algoritmo della media aritmetico-geometrica di LAGRANGE e GAUSS e le sue generalizzazioni di BORCHARDT e SCHAPIRA, l'algoritmo della media aritmetico-armonica, quello più generale di STIELTJES per la radice n^{ma} , etc.

18. L'algoritmo K abbia al divisore $(g_1 g_2 \dots g_p)$ convergenza (almeno) quadratica dell'ordine p_1 e non di un ordine minore. Le congruenze (14)_r e la disuguaglianza (15)_r del n.° 7 permettono allora (per quanto i calcoli relativi siano un po' lunghi) di determinare il grado massimo $s_1 \geq 2$ di convergenza dell'ordine p_1 che l'algoritmo K ha al divisore $(g_1 g_2 \dots g_p)$. Se allora r è un intero qualunque maggiore di p_1 , per il teorema c) del n.° 6, l'algoritmo K avrà al divisore $(g_1 g_2 \dots g_p)$ convergenza dell'ordine r e di un

(*) Cf. A, n.° 13, d).

grado s_r non minore di $s_1^{\lceil \frac{r}{p_1} \rceil}$ (dove, come al solito, $\lceil \frac{r}{p_1} \rceil$ indica il massimo intero contenuto in $\frac{r}{p_1}$). Facciamo ora percorrere ad r i successivi valori interi maggiori di p_1 ; due casi sono allora logicamente possibili. O, per qualunque valore di r , il grado massimo s_r di convergenza che l'algoritmo ha al divisore $(g_1 g_2 \dots g_p)$ è uguale ad $s_1^{\lceil \frac{r}{p_1} \rceil}$; oppure esistono dei valori di r maggiori di p_1 , per i quali il grado massimo s_r relativo di convergenza è maggiore di $s_1^{\lceil \frac{r}{p_1} \rceil}$. Nel primo caso la convergenza dell'algoritmo potrebbe opportunamente dirsi *pura* o di *primo rango*, nel secondo *mista* o di *rango superiore al primo*. E con considerazioni del tutto simili alle superiori, potremmo introdurre la nozione di convergenza del 2.°, 3.°, 4.°, ... rango, di un rango qualunque. Ma non mi è riuscito finora di trovare un criterio che permetta, con un numero finito di operazioni, di distinguere la convergenza pura da quella mista o di rango superiore al primo; d'altra parte coll'aumentare dell'ordine, aumentano sempre più i calcoli necessari alle relative verifiche di convergenza; nè infine può affermarsi a priori che una tale ricerca possa avere una grande importanza per la teoria generale dell'iterazione; crediamo perciò opportuno limitarci alle osservazioni già fatte.

Aggiungiamo soltanto l'osservazione seguente.

Supponiamo che l'algoritmo K abbia al divisore $(g_1 g_2 \dots g_p)$ convergenza di un certo ordine $r > p_1$ e di un grado s_r maggiore di $s_1^{\lceil \frac{r}{p_1} \rceil}$. Quando il grado s_r sia dato, si può assegnare per r un limite superiore; se è infatti $s_1^{q-1} < s_r < s_1^q$, da quanto precede è chiaro che l'ordine relativo r di convergenza deve essere minore di $p_1 q$. Ne segue che: è possibile, con un numero finito di operazioni, riconoscere se l'algoritmo K possiede al divisore $(g_1 g_2 \dots g_p)$ (oltre la convergenza di ordine p_1 e grado s_1) convergenza (di un qualche ordine e) di un grado s **assegnato**.

VI. CONDIZIONI PER LA CONVERGENZA LINEARE.

19. Se il determinante caratteristico $E(\omega)$ dell'algoritmo K non ha (modd $g_1 \dots g_p$), oltre gli $n - p$ divisori elementari uguali ad $\omega - 1$, tutti gli altri divisori elementari uguali ad una potenza di ω , l'algoritmo non ha, al divisore stesso, convergenza di grado superiore al primo. La considerazione del determinante caratteristico permette in questo caso di riconoscere se l'algoritmo ha, allo stesso divisore, convergenza lineare di un ordine qualunque. Si hanno infatti i teoremi:

a) Se l'equazione $E(\omega) \equiv 0 \pmod{g_1 g_2 \dots g_p}$ ha (oltre la radice 1 multipla di ordine $n - p$) tutte radici in modulo minori dell'unità, l'algoritmo ha, al divisore $(g_1 \dots g_p)$, convergenza lineare (di un ordine sufficientemente elevato).

Infatti, in questa ipotesi, tutte le radici della equazione

$$F(\omega) \equiv 0 \pmod{g_1 g_2 \dots g_p}$$

si manterranno su tutta la varietà M minori in modulo dell'unità; essendo quindi funzioni continue dei punti di M , potrà determinarsi un numero positivo Ω , minore dell'unità, che sarà maggiore ed uguale al modulo di ogni radice della equazione stessa su tutta la varietà M .

Sia ora $(G_1 G_2 \dots G_p)$ un sistema normale, pel quale la sostituzione Γ indotta dall'algoritmo ha la forma data dalle (23) e (24). Varranno allora, per qualunque ρ , le $(23)_\rho$, $(24)_\rho$; quindi se, ad es.: p_1 è il massimo degli esponenti $p_1 \dots p_n$, ogni elemento $\gamma_{\mu\nu}^{(\rho)}$ della $\Gamma^{(\rho)}$ ha, sulla varietà M , un modulo sempre minore di $\rho^{p_1} \cdot \Omega^{\rho - p_1}$; si avrà quindi anche, per la sostituzione stessa

$$\Gamma_{\mu}^{(\rho)} = \sum_{\nu} |\gamma_{\mu\nu}^{(\rho)}|_M < p_1 \rho^{p_1} \Omega^{\rho - p_1} = A \rho^{p_1} \Omega^\rho, \quad (\nu = 1, 2, \dots, p)$$

dove $A = p_1 \Omega^{-p_1}$ è un numero finito.

Si osservi ora che, poichè $\Omega < 1$, è $\lim_{\rho \rightarrow \infty} \rho^{p_1} \Omega^\rho = 0$; assegnato quindi un numero α , positivo e minore dell'unità, si può determinare un numero ρ_0 tale che per $\rho \geq \rho_0$ si abbia

$$\rho^{p_1} \Omega^\rho < \frac{\alpha}{A};$$

per questi valori di ρ si avrà allora

$$\Gamma_{\mu}^{(q)} < a, \quad (\mu = 1, 2, \dots, p)$$

e quindi l'algoritmo K avrà, al divisore $(G_1 G_2 \dots G_p)$, convergenza lineare di ogni ordine $\rho \geq \rho_0$, misurata da una costante non maggiore di $\alpha_\rho = A \rho^{p_1} \Omega^\rho$ (*).

Ritornando al primitivo sistema $g_1 g_2 \dots g_p$, abbiamo (cf. n.° 6 ed anche A , n.° 25, b)) che l'algoritmo avrà, al divisore $(g_1 g_2 \dots g_p)$, ancora convergenza lineare di ogni ordine ρ maggiore di un conveniente numero $\rho_1 \geq \rho_0$.

Il teorema enunciato è così dimostrato.

b) Se l'equazione $E(\omega) \equiv 0 \pmod{g_1 g_2 \dots g_p}$ ha qualche radice che abbia un modulo maggiore dell'unità, l'algoritmo K non può convergere al divisore $(g_1 \dots g_p)$.

Sia infatti ω_1 una radice della $E(\omega) \equiv 0$ che abbia un modulo maggiore dell'unità; vi sarà una funzione del sistema normale, ad es.: la G_1 , per la quale si ha:

$$G'_1 \equiv \omega_1 G_1 \pmod{(g_1 g_2 \dots g_p)^2};$$

se allora θ è un numero maggiore dell'unità, ma minore del modulo della ω_1 , sarà in un intorno conveniente di M

$$|G'_1| > \theta \cdot |G_1|,$$

e quindi anche, per ρ qualunque:

$$|G_1^{(q)}| > \theta^q \cdot |G_1|.$$

Se quindi G_1 è inizialmente diverso da zero, quando si prenda ρ sufficientemente grande, sarà $|G_1^{(q)}|$ maggiore di un numero grande a piacere. L'algoritmo non può allora convergere al divisore $(G_1 G_2 \dots G_p)$; in questo caso infatti dovrebbe aversi

$$\lim_{q \rightarrow \infty} G_1^{(q)} = 0.$$

Ne segue il teorema enunciato. Si noti che condizione essenziale della dimostrazione è che inizialmente sia $G_1 \neq 0$.

c) Se (oltre la radice 1 multipla di ordine $n - p$) tutte le radici della $E(\omega) \equiv 0 \pmod{g_1 \dots g_p}$ hanno un modulo maggiore dell'unità, l'algoritmo K^{-1} inverso di quello dato avrà al divisore $(g_1 g_2 \dots g_p)$ convergenza lineare.

(*) Notiamo, di passaggio, che è $\lim_{q \rightarrow \infty} \frac{\alpha_{q+1}}{\alpha_q} = \Omega$.

Infatti il determinante $E^{-1}(\omega)$, caratteristico dell'algoritmo inverso K^{-1} , si ha da $E(\omega)$, mutandovi ω in $\frac{1}{\omega}$ e moltiplicando poi per ω^n .

I teoremi *a), b), c)* furono enunciati da POINCARÉ per $p = n$, e dimostrati dal LATTÈS per $n = 2, 3$ e p qualunque nelle Memorie citate.

20. È chiaro da quanto precede che la considerazione del determinante caratteristico $E(\omega)$ non serve a decidere della convergenza (o meno) dell'algoritmo K al divisore $(g_1 g_2 \dots g_p)$, nel solo caso che l'equazione

$$E(\omega) \equiv 0 \pmod{g_1 g_2 \dots g_p}$$

(liberata dalla radice 1 multipla di ordine $n - p$) abbia radici che hanno tutte un modulo non maggiore dell'unità e vi sia qualche radice di modulo 1.

È interessante osservare che: in questo caso le condizioni di convergenza espresse dalle disuguaglianze (12), del n.° 5 non sono soddisfatte per nessun valore di r .

Ammettiamo infatti, più generalmente, che l'equazione

$$E(\omega) \equiv 0 \pmod{g_1 \dots g_p}$$

abbia (oltre la radice 1 multipla di ordine $n - p$) delle radici di modulo maggiore od uguale ad uno, e siano $\omega_1, \omega_2, \dots, \omega_k$ (con $k \leq p$) le radici, uguali o diverse, che hanno il (medesimo) modulo massimo (e potrà anche k variare da regione a regione di M) (*). Prendiamo allora l'equazione caratteristica della sostituzione lineare $\Gamma^{(\rho)}$:

$$F^{(\rho)}(\omega) = |\gamma_{\mu\nu}^{(\rho)} - \varepsilon_{\mu\nu} \omega| \equiv 0 \pmod{g_1 g_2 \dots g_p}, \quad (\mu, \nu = 1, 2, \dots, p; \rho = 1, 2, \dots);$$

essa avrà come radici di ugual modulo massimo $\omega_1^{(\rho)}, \omega_2^{(\rho)}, \dots, \omega_k^{(\rho)}$ (**). e considerando la funzione simmetrica elementare di grado k delle radici di questa equazione, si avrà per essa

$$\Sigma \omega_1^{(\rho)} \omega_2^{(\rho)} \dots \omega_k^{(\rho)} \equiv \Sigma \begin{vmatrix} \gamma_{11}^{(\rho)} & \gamma_{12}^{(\rho)} & \dots & \gamma_{1k}^{(\rho)} \\ \cdot & \cdot & \cdot & \cdot \\ \gamma_{k1}^{(\rho)} & \gamma_{k2}^{(\rho)} & \dots & \gamma_{kk}^{(\rho)} \end{vmatrix} \pmod{g_1 g_2 \dots g_p} \quad (40)$$

(*) Considereremo allora separatamente le diverse regioni di M .

(**) Cf. ad es.: MUTH, l. c., p. 33. L'asserzione del testo si dimostra subito in modo diretto sulla forma normale (23)_Q, (24)_Q che può darsi alla $F^{(\rho)}$.

dove il secondo membro è l'invariante *ortogonale* di grado K della sostituzione lineare $\Gamma^{(p)}$.

Poichè le altre radici $\omega_{k+1} \dots \omega_p$ hanno un modulo minore di $\omega_1, \omega_2, \dots, \omega_k$, si potrà determinare un numero ρ_0 sufficientemente grande tale che per qualunque valore di $\rho \geq \rho_0$ nella funzione simmetrica del primo membro il primo termine sia preponderante e quindi *il modulo della stessa funzione*, (essendo $|\omega_1| \geq 1, \dots, |\omega_k| \geq 1$) *sia sempre maggiore, sulla varietà M , di un numero determinato n positivo minore dell'unità.*

Ammettiamo ora, se è possibile, che l'algoritmo K abbia, al divisore $(g_1 g_2 \dots g_p)$ convergenza lineare di un certo ordine r , misurata dalla costante positiva $a < 1$; per il teorema *e*) del n.º 6 esso avrà allora anche convergenza lineare di ogni ordine $r q$, multiplo di r , misurata da una costante non maggiore di a^q ; sarà quindi anche per q qualunque

$$|\gamma_{\mu\nu}^{(rq)}| < a^q, \quad (\mu, \nu = 1, 2, \dots, p)$$

ed insieme

$$\left| \Sigma \begin{vmatrix} \gamma_{11}^{(rq)} & \gamma_{12}^{(rq)} & \dots & \gamma_{1k}^{(rq)} \\ \dots & \dots & \dots & \dots \\ \gamma_{k1}^{(rq)} & \gamma_{k2}^{(rq)} & \dots & \gamma_{kk}^{(rq)} \end{vmatrix} \right| < \binom{p}{k} k! a^{qk} < p! a^q. \quad (41)$$

Facciamo allora nella (40) $\rho = r q$; per la (41) avremo, sulla varietà M

$$|\Sigma \omega_1^{r^q} \omega_2^{r^q} \dots \omega_k^{r^q}| < p! a^q,$$

e può quindi rendersi piccolo a piacere, prendendo q sufficientemente grande. Ma ciò è assurdo; l'algoritmo K non può quindi avere, al divisore $(g_1 g_2 \dots g_p)$ convergenza lineare di nessun ordine (*).

Ricordando il teorema *a*) del n.º 19, ne segue evidentemente che: *le condizioni per la convergenza lineare di un ordine qualunque saranno soddisfatte allora ed allora soltanto che l'equazione $E(\omega) \equiv 0 \pmod{g_1 g_2 \dots g_p}$ abbia (oltre la radice 1 multipla dell'ordine $n - p$) tutte le radici inferiori in modulo all'unità: sicchè questa è, in ultima analisi, la condizione necessaria e sufficiente perchè l'algoritmo K abbia al divisore $(g_1 g_2 \dots g_p)$ convergenza lineare di un ordine qualunque.*

Pisa, Gennaio 1907.

(*) Tutto questo non dimostra però che l'algoritmo *non* converge.

Riporto, qui sotto, i titoli di alcuni lavori che trattano di questioni di iterazione in una o più variabili, che, per dimenticanza o perchè pubblicati contemporaneamente, furono omissi nell'indice bibliografico che è in fine alla Memoria ricordata dell'Accademia dei XL.

1875. C. FORMENTI, *Su alcuni problemi di Abel* (Rendiconti Istituto Lombardo, serie 2.^a, vol. 8.^o, pag. 276).
1890. H. POINCARÉ, *Sur une classe nouvelle de transcendentes uniformes* (Journal de Liouville, 4.^{me} Serie, Tom. 6.^{me}, pag. 313).
1897. F. PODETTI, *Sulle sostituzioni uniformi* (Giornale di Battaglini, Vol. 35, pag. 264).
1898. C. BOURLET, *Sur le problème de l'itération* (Annales de la Faculté des Sciences de Toulouse, Tom. XII, C. I).
1901. J. HADAMARD, *Sur l'itération et les solutions asymptotiques des équations différentielles* (Bulletin de la Société Math. de France, Tome XXIX, pag. 244).
1901. T. LEVI-CIVITA, *Sopra alcuni criteri di instabilità* (Annali di Matematica, Serie 3.^a, Tomo V^o, pag. 221).
1902. O. SPIESS, *Die Grundbegriffe der Iterationsrechnung* (Dissertation Basel, 1902).
1904. CIGALA, *Sopra un criterio di instabilità* (Annali di Matematica, Serie 3.^a, Tomo XI, pag. 67).
1906. S. LATFÈS, *Sur les équations fonctionnelles qui définissent une courbe ou une surface invariante par une transformation* (Annali di Matematica, Serie 3.^a, Tomo XIII, pag. 1).
1906. SPIESS, *Theorie der linearen Itergleichung mit konstanten Koeffizienten* (Mathematische Annalen, Bd. 62, S. 226).

Pisa, li 18 Marzo 1907.
