Bollettino
Unione Matematica Italiana

Roberto Infantino

Sulla convergenza di un noto metodo iterativo per la risoluzione numerica delle equazioni differenziali di tipo ellittico e di tipo parabolico.

Bollettino dell'Unione Matematica Italiana, Serie 3, Vol. **19** (1964), n.4, p. 391–399.
Zanichelli
<http://www.bdim.eu/item?id=BUMI_1964_3_19_4_391_0>

Bollettino dell'Unione Matematica Italiana, Zanichelli, 1964.

SEZIONE SCIENTIFICA

BREVI NOTE

Sulla convergenza di un noto metodo iterativo per la risoluzione numerica delle equazioni differenziali di tipo ellittico e di tipo parabolico

Nota di ROBERTO INFANTINO (a Napoli) (*) (**)

Sunto. - *Viene data una dimostrazione, basata sulla teoria delle «contrazioni», della convergenza del metodo di iterazione per la risoluzione del sistema di equazioni lineari, che si ottiene applicando il metodo delle differenze finite al secondo problema al contorno per le equazioni di tipo ellittico, e al primo e secondo problema al contorno per le equazioni di tipo parabolico.*

Dato il dominio D limitato da una curva semplice e chiusa Γ, consideriamo il problema:

$$au_{xx} + cu_{yy} + du_x + eu_y - gu = r \tag{1}$$

$$A_1(B)u(B) + A_2(B)\frac{\partial u(B)}{\partial n} = f(B), \qquad B \in \Gamma, \tag{2}$$

dove $\frac{\partial u(B)}{\partial n}$ indica la derivata secondo la normale interna, $f(B)$, $A_1(B)$, $A_2(B)$ sono funzioni note su Γ, e a, c, d, e, g, r sono funzioni assegnate, definite nel dominio D. Inoltre sia:

$$a > 0, \qquad c > 0, \qquad g \geqq 0. \tag{3}$$

(*) Pervenuta alla Segreteria dell'U.M.I. il 9 ottobre 1964.

(**) Lavoro eseguito nell'ambito dell'attività del Gruppo di Ricerca n. 13 del Comitato Nazionale per la Matematica del C.N.R. per l'anno 1963-64.

M. Iannuzzi [1], nella ipotesi $A_2 \equiv 0$, ha dato una nuova dimostrazione della convergenza del metodo di iterazione per la risoluzione del sistema di equazioni lineari, che si ottiene applicando il metodo delle differenze finite al problema (1), (2).

In questo lavoro dimostro che il procedimento adoperato da M. Iannuzzi *si può applicare al problema* (1), (2) *anche nel caso in cui sia sempre* $A_2 \neq 0$ *e inoltre per* h *positivo e abbastanza piccolo risulti:*

$$L \equiv | A_2 - mhA_1 | \geqq | A_2 |, \qquad 1 \leqq m \leqq \sqrt{2}. \tag{4}$$

Inoltre, detto G il dominio del piano xy che ha per frontiera il segmento AB, parallelo all'asse x, e la curva regolare e semplice Γ_1, di punti estremi A, B tutta al disotto della retta AB, consideriamo il problema:

$$au_{xx} + bu_x + cu_y + du = e \tag{5}$$

$$A_1(B)u(B) + A_2(B)\frac{\partial u(B)}{\partial n} = f(B), \qquad B \in \Gamma_1, \tag{6}$$

dove $f(B)$, $A_1(B)$, $A_2(B)$ sono funzioni note su Γ_1 e a, b, c, d, e funzioni assegnate definite in G, con $a > 0$, $c < 0$, $d \leqq 0$.

In questo lavoro mostro anche che il detto procedimento *si può applicare al problema* (5), (6) *nei due casi seguenti:*

a) $A_2(B) \equiv 0$, $| A_1(B) | \geqq 1$ per ogni $B \in \Gamma_1$;

b) $A_2 \neq 0$ per ogni $B \in \Gamma_1$, vale la (4) su Γ_1, con una opportuna ipotesi sulla Γ_1 che preciseremo in seguito.

1. Fissiamo nel piano un punto (x_0, y_0) e consideriamo i punti di coordinate:

$$x_i = x_0 + ih,$$

$$y_k = y_0 + kh, \qquad (i, k = 0, \pm 1, \pm 2, \ldots), \qquad h > 0;$$

(1) M. Iannuzzi, *Una nuova dimostrazione della convergenza di un noto metodo iterativo per la risoluzione numerica delle equazioni differenziali alle derivate parziali di tipo ellittico*, « Bollettino U.M.I. », Serie III, Anno XVII, N. 1 (1962), pagg. 15-19.

i quali possono considerarsi come i nodi di un reticolato a maglie quadrate di lato h.

Due nodi li diremo *vicini* se appartengono alla stessa retta del reticolato e sono consecutivi.

Indichiamo con I la totalità dei nodi che appartengono a D e con $I_1, I_2, \dots, I_n$ n sottoinsiemi di I, a due a due disgiunti e tali che $\bigcup_{k=1}^{n} I_k = I$ e così definiti:

$X \in I_1$ se $X \in I$ ed esiste un suo *vicino* $Y \notin I$;

$X \in I_k$, $k = 2, 3, \dots, n$, se $X \in I$, $X \notin I_{k-1}$ ed esiste un suo *vicino* $Y \in I_{k-1}$.

In ogni nodo $X \equiv (x_i, y_k) \in I - I_1$ scriviamo l'equazione alle differenze che corrisponde alla (1):

$$\begin{aligned} &a_{i,k} \frac{U_{i+1,k} - 2U_{i,k} + U_{i-1,k}}{h^2} + c_{i,k} \frac{U_{i,k+1} - 2U_{i,k} + U_{i,k-1}}{h^2} + \\ (7) \qquad &+ d_{i,k} \frac{U_{i+1,k} - U_{i-1,k}}{2h} + e_{i,k} \frac{U_{i,k+1} - U_{i,k-1}}{2h} - g_{i,k} U_{i,k} = r_{i,k} \,. \end{aligned}$$

Se $X \in I_1$, detto B_X il punto di Γ che ha la minima distanza da X, la retta $B_X X$ risulta normale alla Γ in B_X; orientiamo questa nel verso che va da B_X a X e sia C_X il primo punto, dopo X, che tale normale ha in comune con il reticolato ed il dominio D.

Denotiamo con D_X, E_X i due nodi *vicini* tra i quali si trova C_X e supponiamo h abbastanza piccolo in modo che risulti $\{D_X\} \cup \{E_X\} \not\subset I_1$ (2).

(2) Nel caso particolare in cui C_X coincida con uno dei punti D_X, E_X, cioè $C_X \in I$, supporremo che h sia tale che $C_X \notin I_1$. Mostriamo brevemente che tali ipotesi su h sono possibili. Siano $x = x(s)$, $y = y(s)$, con $s \in [a, b]$ ascissa curvilinea, le equazioni parametriche della curva Γ. È facile vedere che comunque si fissi un numero positivo δ esiste un numero positivo δ' tale che per ogni $P \in \Gamma$ $C_{\delta'}(P)$, il cerchio di centro P e raggio δ', contenga un solo arco continuo di Γ di lunghezza minore di δ. Essendo Γ un insieme compatto, ne viene che si può ricoprire Γ con un numero finito di cerchi $C_{\bar\delta_0}(P_0)$, $C_{\bar\delta_1}(P_1), \dots$, $C_{\bar\delta_n}(P_n)$ contenenti rispettivamente gli archi $\Gamma_0, \Gamma_1, \dots, \Gamma_n$ tali che $\bigcup_{i=0}^{n} \Gamma_i = \Gamma$ e per ogni $P \in \Gamma_i$ $C_{\bar\delta_i}(P)$ contiene

Posto:

$$\overline{XC_X} = mh, \qquad \overline{D_X C_X} = m'h, \qquad \overline{E_X C_X} = (1 - m')h,$$

dove $1 \leqq m \leqq \sqrt{2}$, $0 \leqq m' \leqq 1$, sostituiamo la condizione (2) con la seguente [3]:

$$(8) \qquad A_1(B_X)U(X) + A_2(B_X)\frac{(1-m')U(D_X) + m'U(E_X) - U(X)}{mh} = f(B_X),$$

$$X \in I_1, \qquad B_X \in \Gamma.$$

Per risolvere il sistema (7), (8) col metodo di iterazione si procede nel modo seguente: si danno valori arbitrari alle $U_{i,k}^{(0)}$ e $U^{(0)}(X)$, quindi, per ogni coppia di valori i, k, che corrisponde al punto $(x_i, y_k) \in I - I_1$, e per ogni $X \in I_1$, si costruiscono le successioni:

$$U_{i,k}^{(1)}, \quad U_{i,k}^{(2)}, \ldots, \quad U_{i,k}^{(\nu)}, \ldots$$

$$U^{(1)}(X), \quad U^{(2)}(X), \ldots, \quad U^{(\nu)}(X), \ldots,$$

un solo arco continuo di Γ di lunghezza minore di δ. Pertanto, posto $\bar{\delta} = \min\{\bar{\delta}_0, \bar{\delta}_1, \ldots, \bar{\delta}_n\}$, per ogni $P \in \Gamma$ $C_{\bar{\delta}}(P)$ contiene un solo arco continuo di Γ di lunghezza minore di δ. Per ogni coppia di punti P', P'' indichiamo con $\alpha(P', P'')$ l'angolo delle normali interne in P', P'' alla Γ. Poichè le funzioni $x'(s)$, $y'(s)$ sono uniformemente continue in $[a, b]$, fissato $\varepsilon > 0$, è possibile determinare un numero $\rho > 0$ tale che per $\widehat{P'P''} < \rho$ risulti $\alpha(P', P'') < \varepsilon$.

Allora, per quanto dimostrato sopra, è possibile determinare un numero positivo η tale che per ogni $P \in \Gamma$ $C_\eta(P)$ contenga un solo arco continuo Γ_P di Γ di lunghezza minore di ρ e quindi tale che per ogni $P', P'' \in \Gamma_P$ risulti $\alpha(P', P'') < \varepsilon$.

Ciò premesso, poniamo $\overline{B_X X} = d_X$ e sia θ l'angolo che ha per bisettrice la normale interna $\overline{B_X X}$ e ampiezza $\pi - 2\varepsilon$. Nel caso $C_X \in I$, si vede facilmente che l'arco di Γ contenuto in $C_\eta(B_X)$ è esterno all'angolo θ. Pertanto, essendo $0 \leqq d_X < h$, se è $\varepsilon \leqq \pi/4$, $h \leqq \eta/3$, nessuno dei quattro *vicini* di C_X è esterno a Γ, cioè $C_X \notin I_1$.

Nel caso generale, con ragionamenti analoghi si vede che almeno uno dei punti D_X, E_X non appartiene a I_1, se è $\varepsilon \leqq \pi/8$, $h \leqq \eta/3$. Perciò, fissato $\varepsilon \leqq \pi/8$ e preso $h \leqq \eta/3$, risulta $\{D_X\} \cup \{E_X\} \not\subset I_1$.

(3) Cfr. L. Collatz, *The numerical treatment of differential equations*, Springer Verlag, Berlin 1960, pag. 346.

mediante le formule ricorrenti:

$$
(9)\quad \left\{
\begin{aligned}
&U_{i,k}^{(\nu+1)} = \frac{a_{i,k}+\frac{h}{2}d_{i,k}}{2a_{i,k}+2c_{i,k}+h^2g_{i,k}}\,U_{i+1,k}^{(\nu)} + \frac{a_{i,k}-\frac{h}{2}d_{i,k}}{2a_{i,k}+2c_{i,k}+h^2g_{i,k}}\,U_{i-1,k}^{(\nu)} + \\
&+ \frac{c_{i,k}+\frac{h}{2}e_{i,k}}{2a_{i,k}+2c_{i,k}+h^2g_{i,k}}\,U_{i,k+1}^{(\nu)} + \frac{c_{i,k}-\frac{h}{2}e_{i,k}}{2a_{i,k}+2c_{i,k}+h^2g_{i,k}}\,U_{i,k-1}^{(\nu)} + \\
&- \frac{h^2 r_{i,k}}{2a_{i,k}+2c_{i,k}+h^2g_{i,k}} \\
&U^{(\nu+1)}(X) = \frac{A_2(B_X)}{L}\{m'U^{(\nu)}(E_X)+(1-m')U^{(\nu)}(D_X)\} - mhf(B_X), \\
&X \in I_1 .
\end{aligned}
\right.
$$

Le quali, con ovvio significato dei simboli, possono scriversi:

$$
(9')\quad \left\{
\begin{aligned}
&U_{i,k}^{(\nu+1)} = A_{i,k}U_{i+1,k}^{(\nu)} + B_{i,k}U_{i-1,k}^{(\nu)} + C_{i,k}U_{i,k+1}^{(\nu)} + D_{i,k}U_{i,k-1}^{(\nu)} + E_{i,k} \\
&U^{(\nu+1)}(X) = \frac{A_2(B_X)}{L}\{m'U^{(\nu)}(E_X)+(1-m')U^{(\nu)}(D_X)\} - mhf(B_X), \\
&X \in I_1 .
\end{aligned}
\right.
$$

Possiamo supporre h tale che:

$$a_{i,k} \pm \frac{h}{2} d_{i,k} > 0, \qquad e_{i,k} \pm \frac{h}{2} e_{i,k} > 0.$$

Ne viene:

$$
(10)\quad \left\{
\begin{aligned}
&A_{i,k} > 0, \quad B_{i,k} > 0, \quad C_{i,k} > 0, \quad D_{i,k} > 0 \\
&A_{i,k} + B_{i,k} + C_{i,k} + D_{i,k} = \frac{2a_{i,k}+2c_{i,k}}{2a_{i,k}+2c_{i,k}+h^2g_{i,k}} \leqq 1.
\end{aligned}
\right.
$$

Sia S l'insieme delle funzioni reali definite in I. Posto:

$$\alpha = \min_{(x_i,\, y_k)\in I} \{A_{i,k},\ B_{i,k},\ C_{i,k},\ D_{i,k}\},$$

$$\eta = \frac{1-\alpha}{\alpha}, \qquad \beta = \max(1,\ \eta),$$

esistono due convenienti numeri σ e θ appartenenti all'intervallo aperto (0, 1) tali che, posto ancora:

$$z_k = \sigma(\eta^{k-1} + \eta^{k-2} + \dots + 1), \qquad (k = 1, 2, \dots, n-1),$$

$$w_1 = 1, \qquad w_i = 1 - \sum_{k=1}^{i-1} z_k, \qquad (i = 2, 3, \dots, n-1, n),$$

risulti: (4)

$$(11) \qquad \begin{cases} w_{i+1} < w_i, & (i = 1, 2, \dots, n-1), \\ w_i > 0 \\ \alpha w_{i+1} + (1-\alpha) w_{i-1} \leq \theta w_i, & (i = 2, 3, \dots, n-1). \end{cases}$$

Assumendo come distanza di due funzioni u', u'' di S la quantità:

$$\| u' - u'' \| = \max_{k \in \{1, 2, \dots, n\}} \frac{\max_{X \in I_k} | u'(X) - u''(X) |}{w_k},$$

S diviene uno spazio metrico e le (9') definiscono una trasformazione T di S in sé. In forza di un noto teorema di Analisi funzionale (5), per dimostrare che il sistema (9') ha una sola soluzione ci basta far vedere che per ogni coppia di funzioni u', u'' di S risulta:

$$(12) \qquad \| Tu' - Tu'' \| \leqq \lambda \| u' - u'' \|, \qquad 0 < \lambda < 1.$$

Ora, tenuto conto delle (10) e delle (11), se $X \equiv (x_i, y_k) \in I_s$, $s = 2, 3, \dots, n$), é:

$$(13) \qquad \frac{| Tu'(X) - Tu''(X) |}{w_s} \leqq \frac{\alpha w_{s+1} + (1-\alpha) w_{s-1}}{w_s} \| u' - u'' \| \leqq \theta \| u' - u'' \|.$$

Se $X \in I_1$, $E_X \in I_{k_X}$, $D_X \in I_{k'_X}$, risulta:

$$\frac{| Tu'(X) - Tu''(X) |}{w_1} \leqq \frac{| A_2(B_X) |}{L} \frac{m'}{w_1} \max_{I_{k_X}} | u' - u'' | + $$
$$+ \frac{1 - m'}{w_1} \max_{I_{k'_X}} | u' - u'' |,$$

(4) Basta prendere $0 < \sigma < 2[n(n-1)\beta^{n-2}]^{-1}$,

$$\theta = \max_{i \in \{2, 3, \dots, n\}} \frac{w_i - \alpha(z_{i-1} + z_i)}{w_i}.$$

(5) Si veda, per esempio, L. COLLATZ, op. cit., pagg. 36-38.

e se h é tale che $\frac{|A_2(B_X)|}{L} \leq 1$, si ha:

$$\frac{|Tu'(X) - Tu''(X)|}{w_1} \leq \frac{m'w_{k_X} + (1-m')w_{k'_X}}{w_1} \| u' - u'' \|, \quad X \in I_1, \tag{14}$$

Dato che si é supposto h tale che $\{D_X\} \bigcup \{E_X\} \not\subset I_1$, a motivo della prima delle (11), risulta:

$$\mu = \max_{X \in I_1} \frac{m'w_{k_X} + (1-m')w_{k'_X}}{w_1} < 1\,;$$

allora, posto $\lambda = \max\{\theta, \mu\}$, si ha: $0 < \lambda < 1$ e, in forza delle (13) e (14), la (12).

2. Consideriamo ora il problema (5), (6). Prendiamo $(x_0, y_0) \in AB$; allora il reticolato, definito nel n. 1, avrà una retta coincidente con la retta AB.

Indichiamo con I la totalità dei nodi di G che non appartengono al segmento AB e con I_1 l'insieme dei nodi $X \in I$ tali che almeno una delle due rette del reticolato per X intersechi Γ_1 in un punto Z avente distanza da X minore di h. Definiamo poi gli insiemi $I_2, I_3, \ldots, I_n$ come nel n. 1.

In ogni $X \equiv (x_i, y_k) \in I - I_1$ scriviamo l'equazione alle differenze che corrisponde alla (5):

$$a_{i,k} \frac{U_{i+1,k} - 2U_{i,k} + U_{i-1,k}}{h^2} + b_{i,k} \frac{U_{i,k} - U_{i-1,k}}{h} +$$

$$+ c_{i,k} \frac{U_{i,k} - U_{i,k-1}}{h} + d_{i,k} U_{i,k} = e_{i,k}\,.$$

Distinguiamo ora i due casi *a*) e *b*).

Nel caso *a*) ad ogni punto $X \in I_1$ associamo un punto Z, intersezione di Γ_1 con una delle due rette del reticolato per X, tale che la distanza $\delta_X \geq 0$ di X da Z sia minore di h.

Supponiamo h tanto piccolo in modo che il nodo C_X, il *vicino* di X più lontano da Z, appartenga ad I e non si trovi su Γ_1. Per ogni $X \in I_1$ scriviamo l'equazione che corrisponde alla (6) nel caso *a*):

$$U(X) = \frac{1}{A_1} \frac{\delta_X U(C_X) + hf(Z)}{\delta_X + h}\,. \tag{16}$$

Facciamo vedere che il sistema (15), (16) si può risolvere col metodo di iterazione descritto nel n. 1, mediante le formule ricorrenti:

$$U_{i,k}^{(\nu+1)} = \frac{a_{i,k} U_{i+1,k}^{(\nu)} + (a_{i,k} - hb_{i,k}) U_{i-1,k}^{(\nu)} - hc_{i,k} U_{i,k-1}^{(\nu)} - e_{i,k} h^2}{2a_{i,k} - h(b_{i,k} + c_{i,k}) - h^2 d_{i,k}}$$

$$U^{(\nu+1)}(X) = \frac{1}{A_1} \frac{\delta_X U^{(\nu)}(C_X) + hf(Z)}{\delta_X + h}.$$

Essendo $a > 0$, per ipotesi, per h abbastanza piccolo risulta:

$$a_{i,k} - hb_{i,k} \geqq 0.$$

Inoltre essendo, ancora per ipotesi, $c < 0$, $d \leqq 0$ si ha:

$$2a_{i,k} - h(b_{i,k} + c_{i,k}) - h^2 d_{i,k} > 0,$$

e il sistema (17), con ovvio significato dei simboli, può scriversi:

$$(17') \qquad \begin{cases} U_{i,k}^{(\nu+1)} = A_{i,k} U_{i+1,k}^{(\nu)} + B_{i,k} U_{i-1,k}^{(\nu)} + C_{i,k} U_{i,k-1}^{(\nu)} + E_{i,k} \\ U^{(\nu+1)}(X) = \dfrac{1}{A_1} \dfrac{\delta_X U^{(\nu)}(C_X) + hf(Z)}{\delta_X + h}, \end{cases}$$

con:

$$(18) \qquad \begin{cases} A_{i,k} > 0, \qquad B_{i,k} \geqq 0, \qquad C_{i,k} > 0, \\ A_{i,k} + B_{i,k} + C_{i,k} \leqq 1, \\ \dfrac{1}{A_1} \dfrac{\delta_X}{\delta_X + h} < 1. \end{cases}$$

Sia S l'insieme delle funzioni reali definite in I e diciamo T la trasformazione definita dalle (17').

Posto:

$$(19) \qquad \alpha = \max_{(x_i, y_k) \in I} \{ A_{i,k}, \ B_{i,k}, \ C_{i,k} \},$$

risulta $\alpha < 1$.

Definendo in S una distanza nel modo indicato nel numero 1, dove però α ora é il valore dato dalla (19), procedendo come nel

n. 1, per ogni $X \in I_s$, $(s = 2, 3, \dots, n)$, si ha:

$$\frac{|\,Tu'(X) - Tu''(X)\,|}{w_s} \leqq \frac{\alpha w_{s+1} + (1-\alpha) w_{s-1}}{w_s} \|\, u' - u'' \,\|. \tag{20}$$

Per ogni $X \in I_1$ si ha poi:

$$\frac{|\,Tu'(X) - Tu''(X)\,|}{w_1} \leqq \frac{\delta_X}{\delta_X + h} \|\, u' - u'' \,\|. \tag{21}$$

Allora, posto

$$\lambda = \max_{X \in I_1} \left\{ \theta, \; \frac{\delta_X}{\delta_X + h} \right\},$$

riesce: $0 < \lambda < 1$, e, in forza delle (20), (21),

$$\|\, Tu' - Tu'' \,\| \leqq \lambda \,\|\, u' - u'' \,\|.$$

Nel caso b), se $X \in I_1$, consideriamo il punto B_X di Γ_1 di minima distanza da X. Dato che il reticolato é stato preso in modo che una delle rette di esso coincida con la retta AB, la retta $B_X X$ risulta normale a Γ_1 ([6]). Orientiamo questa nel verso che va da B_X a X e sia C_X il primo punto, dopo X, che tale normale ha in comune con il reticolato e il dominio G.

Se non si fa alcuna ipotesi supplementare su Γ_1 può capitare che C_X appartenga al segmento AB. Per evitare ciò, orientata la Γ_1 nel verso che va da A a B, supponiamo che le tangenti in A e in B siano tali che, orientate secondo il verso positivo sulla Γ_1, le loro direzioni positive formino angoli ottusi con l'asse $\overrightarrow{AB}$.

Consideriamo le intersezioni con Γ_1 della retta r di equazione $y = y_0 - 2h$; per h abbastanza piccolo queste sono soltanto due. Sia A' quella più vicina ad A e B' quella più vicina a B. Sotto l'ulteriore ipotesi posta per la Γ_1, per h convenientemente piccolo, l'arco AA' é tale che in ogni suo punto la tangente gode delle proprietà della tangente in A e l'arco BB' é tale che in ogni suo punto la tangente gode della proprietà della tangente in B.

Ne viene che C_X non può appartenere ad AB perché B_X deve appartenere ad $\widehat{AA'} \cup \widehat{BB'}$, quando X appartiene alla retta $y = y_0 - h$, in quanto, in tal caso, ogni altro punto non appartenente ad $\widehat{AA'} \cup \widehat{BB'}$ ha distanza da X maggiore di h.

Dopo ciò tutto procede come nel n. 1 a partire dal secondo capoverso dopo la (7).

([6]) Infatti non può risultare $B_X \in \{A\} \cup \{B\}$ perchè $X \notin \{A\} \cup \{B\}$.