

APROXIMATE SOLUTIONS OF THE OPERATOR LINEAR DIFFERENTIAL EQUATION II

D. Herceg and B. Stanković

(Received January 19, 1977)

1. Introduction

In the first part [4] we constructed a theory which gave the approximate solutions of the linear differential equation

$$(1) \quad \sum_{k=0}^n a_k(s) x^{(k)}(\lambda) = 0, \quad \lambda_1 \leq \lambda \leq \lambda_2,$$

in the field of Mikusiński operators; where

$$(2) \quad a_k(s) = \sum_{\nu=0}^{d_k} \alpha_{k,\nu} s^\nu, \quad d_k \leq m, \quad k = 0, 1, \dots, n,$$

s is the differential operator.

In this second part we shall give the numerical side of the mentioned theory especially when we apply it to partial differential equations using a computer; at the end we shall illustrate it on a concrete example.

2. Character of the methode

To the partial differential equation with constant coefficients

$$(3) \quad \sum_{k=0}^n \sum_{\nu=0}^{d_k} \alpha_{k,\nu} x_{\lambda k, t^\nu}(\lambda, t) = e(\lambda, t) \quad \begin{array}{l} \lambda_1 \leq \lambda \leq \lambda_2, \\ 0 \leq t \leq \infty, \end{array}$$

corresponds, in the field of Mikusiński, the differential equation

$$(4) \quad \sum_{k=0}^n a_k(s) x^{(k)}(\lambda) = f(\lambda)$$

where $a_k(s)$ are given by (2).

The characteristics of our results are:

— The construction of a unique program for a computer to realize the approximation of the linearly independent solutions of the equation (1) in the interval $[0, T]$. These solutions are used to construct the solutions of the equation (3) restricted with some conditions (initial, boundary, ...); they have not to be solutions of the equations (3) in classical sense.

— All the approximations are given by a unique class of functions — Wright's functions.

— The upper bound for the measure of the approximation and in case of a function also the error is given and a program to calculate this bound by a computer.

— For the Wright's function we know an approximation by polynomials and the error for such an approximation.

3. Calculation of the approximate solutions

We have seen that the linearly independent solutions of the equations (1) are of the form

$$(5) \quad x(\lambda) = \lambda^i e^{w\lambda}, \quad i=0, 1, \dots, k-1$$

where w is a k -tuple zero of the polynomial equation

$$(6) \quad F(l, w) \equiv \sum_{k=0}^n \sum_{v=0}^{d_k} \alpha_{k,v} l^{m-v} w^k = 0, \quad d_k \leq m, \quad k=0, 1, \dots, n$$

and they have the form

$$(7) \quad w = l^{-q/p} \sum_{i \geq 0} a_i l^{i/p}, \quad q=0 \text{ if } d_n=0.$$

We know that the value of q/p and a_0 predetermine the existence of the solution (5). The values for q/p can be determined from the inequalities:

$$(8) \quad d_i + \frac{q}{p} i = d_j + \frac{q}{p} j \geq d_k + \frac{q}{p} k, \quad k=0, 1, \dots, n.$$

If $a_k(s)=0$ for one of the value $k=k_0$ then we take as d_{k_0} the value -1 .

All the numbers q and p for which q/p has different values, we count up by the subprogram QIP and NZB (see the schedule which follows).

The corresponding coefficient a_0 can be found as a nonzero solution of the equation

$$(9) \quad Q(a_0) \equiv \alpha_{i, d_i} a_0^i + \alpha_{j, d_j} a_0^j + \alpha_{k_1, d_{k_1}} a_0^{k_1} + \dots + \\ + \alpha_{k_t, d_{k_t}} a_0^{k_t} = 0$$

where k_1, \dots, k_k are these values of k in (8) for which we have equality.

Let d_j be chosen in such a way that $d_j - d_k \geq 0$, $k = i, j, k_1, \dots, k_t$; using the relation (8) we can transform the equation (9) to:

$$(10) \quad \alpha_{j, d_j} + \alpha_{i, d_i} (a_0^{p/q})^{d_j - d_i} + \alpha_{k_1, d_{k_1}} (a_0^{p/q})^{d_j - d_{k_1}} + \dots + \\ + \alpha_{k_t, d_{k_t}} (a_0^{p/q})^{d_j - d_{k_t}} = 0.$$

Now, the coefficient a_0 can be counted up by a subprogram ANULA using also a subprogram ROOT in which by the method Lin-Bairstow we find all the solutions of the equation (9) or (10). In the following we use one value for a_0 but real.

Let us suppose that we fixed a q/p and a real a_0 . In the polynomial $F(l, w)$ (see relation (6)) we introduce new variables: $\omega = l^{q/p} w$ and $u = l^{1/p}$ so that

$$F(l, w) = P(u, \omega) \equiv \sum_{k=0}^n \sum_{v=0}^{d_k} \alpha_{k,v} u^{pm-pv-qk} \omega^k.$$

$P(u, \omega)$ can be written in the form

$$(11) \quad P(u, \omega) = \sum_{k=0}^n \sum_{v=0}^{\mu} \frac{1}{k! v!} P_{u^v, \omega^k}(0, a_0) u^v (\omega - a_0)^k$$

where $\mu = \max p \{m - v\} - qk$, $k = 0, 1, \dots, n$; $v = 0, 1, \dots, d_k$.

It is easy to see that $P_{u^0, \omega^0}(0, a_0) = Q(a_0) = 0$ and $P_{u^0, \omega^0}(0, a_0) = Q'(a_0)$. We can suppose that $Q'(a_0) \neq 0$ because our supposition is that the polynomial $F(l, w)$ is irreducible.

Let us denote by

$$(12) \quad A_{v,k} = \frac{-P_{u^v, \omega^k}(0, a_0)}{v! k! Q'(a_0)}$$

$A_{v,k}$ can be counted up by the programs PARCIZ and POLIZ.

From the equation $P(u, \omega) = 0$ we have

$$(13) \quad \omega - a_0 = \sum_{k=0}^n \sum_{v=0}^{\mu} A_{v,k} u^v (\omega - a_0)^k$$

where $A_{0,0} = A_{0,1} = 0$.

Taking $\omega - a_0 = \sum_{i \geq 1} a_i u^i$ and $(\sum_{i \geq 1} a_i u^i)^k = \sum_{i \geq k} \beta_{i,k} u^i$,

$$\beta_{i,1} = a_i, i \geq 1, \beta_{i,k} = \sum_{m=k-1}^{i-1} a_{i-m} \beta_{m,k-1}, i \geq k,$$

we have

$$(14) \quad \sum_{i \geq 1} a_i u^i = \sum_{k=0}^n \sum_{v=0}^{\mu} A_{v,k} u^v \beta_{i,k} u^i.$$

Let us denote by

$$\gamma_{j,k} = \sum_{i=0}^{P_k} A_{i,k} \beta_{j-i,k}, \quad P_k = \min(\mu, j-k), \quad N = \min(j, n),$$

$$H = \begin{cases} 1, & j \leq \mu \\ 0, & j > \mu \end{cases},$$

we have

$$(15) \quad a_j = HA_{j,0} + \sum_{k=1}^N \gamma_{j,k}, \quad j = 1, 2, \dots$$

In special cases:

$$a_1 = A_{1,0}$$

for $n = 1$

$$(16) \quad a_j = H \cdot A_{j,0} + \sum_{i=1}^{P_1} A_{i,1} a_{j-i}, \quad j \geq 2$$

for $n \geq 2$

$$(17) \quad a_j = H \cdot A_{j,0} + \sum_{i=1}^{P_1} A_{i,1} a_{j-i} + \sum_{k=2}^N \sum_{i=0}^{P_k} A_{i,k} \beta_{j-i,k}, \quad j \geq 2$$

The needed calculations are given by the following schedule:

1. $m = \max \{d_k\}$;
2. q and p ; subprograms QIP and NZB;
3. One chooses a fixed value for q and p ;
4. a_0 ; subprograms ANULA and ROOT;
5. One chooses only one nonzero and real value for a_0 ;
6. $Q'(a_0)$; subprogram POLIZ;
7. $\mu = \max \{p(m-v) - qk\}$, $k = 0, 1, \dots, n$; $v = 0, 1, \dots, d_k$;
8. $A_{v,k}$; subprograms POLIZ, PARCIZ, FACT;
9. a_i , $i = 1, 2, \dots$; subprogram COEF.

4. Measure of the approximation

Let us suppose that we have computed the first i_0 coefficients a_i . As the approximate solution of equation (1) we take

$$(18) \quad \begin{aligned} \tilde{x}(\lambda) &= \exp \left(\lambda \sum_{i=0}^{i_0} a_i l^{(i-q)/p} \right) = \\ &= \exp \left(\lambda \cdot \sum_{i=0}^q a_i l^{(i-q)/p} \right) \exp \left(\lambda \cdot \sum_{i=q+1}^{i_0} a_i l^{(i-q)/p} \right). \end{aligned}$$

About the character of this approximate solution and how it can be expressed by Wright's functions see part I, 2. 3.

The difference from the exact solution $x(\lambda)$ is:

$$x(\lambda) - \tilde{x}(\lambda) = \tilde{x}(\lambda) \left(\exp \left(\lambda \cdot \sum_{i \geq i_0+1} a_i l^{(i-q)p} \right) - I \right).$$

The expression

$$(19) \quad \exp \left(\lambda \cdot \sum_{i \geq i_0+1} a_i l^{(i-q)p} \right) - I$$

gives the measure of the approximation. (We suppose that $i_0+1 \geq q$ so that it is a function from L .)

We know how to find M and r in such a way that $|a_i| \leq Mr^i$ (part I 3.). With these notations we have:

$$(20) \quad \left| \exp \left(\lambda \cdot \sum_{i \geq i_0+1} a_i l^{(i-q)p} \right) - I \right| \leq \{t^{-1} \Phi(0, \delta, \nu t^\delta)\}_T$$

where $\delta = (i_0 - q + 1)/p$ and (see part I, 4.2 and 4.3)

$$(21) \quad \nu \geq Mr^{i_0+1} \sum_{j \geq 0} r^j \frac{T^{j/p}}{\Gamma(j/p + 1)}.$$

4.1. A bound for the parameter

To find a bound for ν we can use the inequality $\Gamma(x) \geq \sqrt{2\pi} x^x e^{-x} x^{1/2}$, $x > 0$, [2] and in this case

$$(22) \quad \nu = Mr^{i_0} \sqrt{\frac{p}{2\pi}} \left[\sum_{j=1}^{j_0-1} \omega^j + \omega^{j_0} \frac{1}{1-\omega} \right] l^{(i_0-q)p},$$

where $\omega = (eTp/j_0)^{1/p} r$; j_0 is chosen in such a way that $\omega < 1$.

An other bound for ν we can realize if we start from the relation

$$(23) \quad \frac{1}{\Gamma(z)} = \frac{1}{2\pi i} \int_C e^u u^{-z} du$$

where C is given by the fig. 1.

We have for ν

$$(24) \quad \nu = Mr^{i_0+1} T^{1/p} \int_C e^u \frac{1}{u} \frac{du}{u^{1/p} - (rT^{1/p})}, \quad |rT^{1/p}| < u^{1/p}.$$

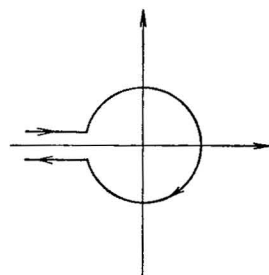


fig. 1.

4.2. An upper bound for the function $t^{-1} \Phi(0, \delta; \nu t^\delta)$

The relation (20) says that the measure of the approximation is given by the function $t^{-1} \Phi(0, \delta; \nu t^\delta)$. To have an upper bound of this function we can use the integral expression

$$(25) \quad t^{-1} \Phi(0, \delta; \nu t^\delta) = \frac{1}{2\pi i} \int_C \exp\left(u + \frac{\nu t^\delta}{u^\delta}\right) du$$

where C is given by the fig. 1, or the Taylor-series for the function

$$(26) \quad t^{-1} \Phi(0, \delta; \nu t^\delta) = t^{-1} \sum_{k=1}^{\infty} \frac{(\nu t^\delta)^k}{\Gamma(k+1) \Gamma(k\delta)}.$$

We shall decide on the relation (26);

$$(27) \quad \begin{aligned} \{t^{-1} \Phi(0, \delta; \nu t^\delta)\} &= \left\{ \frac{t^{\delta-1}}{\Gamma(\delta)} \nu \Gamma(\delta) \sum_{k=1}^{\infty} \frac{(\nu t^\delta)^{k-1}}{\Gamma(k+1) \Gamma(k\delta)} \right\} \\ &\leq \nu \Gamma(\delta) \sum_{k=0}^{\infty} \frac{(\nu T^\delta)^k}{\Gamma(k+1) \Gamma((k+1)\delta)} T^\delta \\ &\leq \nu \Omega_{\delta, \nu} T^\delta \end{aligned}$$

where

$$(28) \quad \Omega_{\delta, \nu} = \nu \Gamma(\delta) \sum_{k=0}^{\infty} \frac{(\nu T^\delta)^k}{\Gamma(k+2) \Gamma[(k+1)\delta]}.$$

It remains to find a bound for $\Omega_{\delta, \nu}$; we have to distinguish three cases:

1. $0 < \delta < 1$. We know that the function $\Gamma(x)$ is positive for $x > 0$ and it has a minimum 0.88560... in the point $x = 1.46163...$ For this reason

$$(29) \quad \begin{aligned} \Omega_{\delta, \nu} &< \frac{\nu \Gamma(\delta)}{0.88} \sum_{k=0}^{\infty} \frac{(\nu T^\delta)^k}{\Gamma(k+1)} \\ &< \frac{\nu \Gamma(\delta)}{0.88} \exp(\nu T^\delta). \end{aligned}$$

2. $\delta \geq 1$. In this case $\Gamma(k\delta) > \Gamma(\delta)$, $k = 1, 2, \dots$ because $k\delta \geq \delta$, $\Gamma(1) = \Gamma(2) = 1$ and $\Gamma(x)$ is monotone increasing for $x > 1.46163...$ Now

$$(30) \quad \Omega_{\delta, \nu} \leq \nu \sum_{k=0}^{\infty} \frac{(\nu T^\delta)^k}{\Gamma(k+1)} = \nu \exp(\nu T^\delta).$$

3. $\delta \geq 3/2$. For this case we shall use the following relation [5]:

$$(31) \quad \Gamma(k\delta) = \left(\frac{1}{\sqrt{2\pi}} \right)^{k-1} k^{k\delta-1/2} \prod_{r=0}^{k-1} \Gamma(\delta + r/k)$$

from which we have:

$$(32) \quad \Gamma(k\delta) > k^{-1/2} \Gamma^k(\delta).$$

$$\begin{aligned} \text{Now} \quad \Omega_{\delta, \nu} &\leq \nu \Gamma(\delta) \sum_{k=0}^{\infty} \frac{(\nu T^{\delta})^k (k+1)^{1/2}}{(k+1) \Gamma(k+1) \Gamma^{k+1}(\delta)} \\ &\leq \nu \sum_{k=0}^{\infty} \left(\frac{\nu T^{\delta}}{\Gamma(\delta)} \right)^k \frac{1}{\Gamma(k+1)} \\ &\leq \nu \exp \left(\frac{\nu T^{\delta}}{\Gamma(\delta)} \right). \end{aligned}$$

At last we can give a bound of the error if the approximate solution $\tilde{x}(\lambda)$ is a function:

$$\begin{aligned} &\left| x(\lambda) - \tilde{x}(\lambda) \right| \leq_T \left| \tilde{x}(\lambda) \right| \left| \exp \left(\lambda \cdot \sum_{i \geq i_0+1} a_i l^{\frac{i-q}{p}} \right) - I \right| \\ &\leq_T \left| \exp \left(\lambda \cdot \sum_{i=0}^q a_i l^{\frac{i-q}{p}} \right) \right| \left| \prod_{i=q+1}^{i_0} I + \left\{ t^{-1} \Phi \left(0, \frac{i-q}{p}; a_i + \frac{i-q}{p} \right) \right\} \right| \\ &\quad \times \left| \left\{ t^{-1} \Phi \left(0, \frac{i_0-q+1}{p}; \nu t^{\frac{i_0-q+1}{p}} \right) \right\} \right| \\ &\leq_T \left| \exp \left(\lambda \cdot \sum_{i=0}^q a_i l^{\frac{i-q}{p}} \right) \right| \left| \prod_{i=q+1}^{i_0} I + \Omega_{\frac{i-q}{p}, \lambda a_i} l^{\frac{i-q}{p}} \right| \times \\ &\quad \times \Omega_{\frac{i_0-q+1}{p}, \nu} l^{\frac{i_0-q+1}{p}}. \end{aligned}$$

A bound of the first part of this product is given in I 4. proposition A.

5. Application to a special operator

In [1] L. Berg applies his method of finding an approximate inversion of the Laplace transform to the function

$$F(p) = \exp \left(-\sqrt{p+a+b-b^2(cp+b)^{-1}} \right)$$

$a, b > 0, c > 1$ which appears by finding of the dispersion's coefficients [3].

We shall apply our theory to the corresponding operator in:

$$(34) \quad F(s) = \exp \left(-(\sqrt{s+d-b^2(cs+b)^{-1}}) \right)$$

$b > 0, c > 1, d > b$, s -differential operator.

$F(s)$ is of the form (5) where $\lambda = 1$ and w satisfies the equation

$$(cs+b)w^2 - ((s+d)(cs+b) - b^2) = 0$$

or

$$(c+bl)lw^2 - ((I+dl)(c+bl) - b^2I^2) = 0$$

which corresponds to the equation (6). The last equation can be written in the form:

$$P(l, \omega) = (c + bl)\omega^2 - ((I + dl)(c + bl) - b^2 l^2) = 0$$

where $\omega = l^{1/2} w$ and we see that $\frac{q}{p} = \frac{1}{2}$ and for a_0 we have $P(0, a_0) = a_0^2 = 1$,

whence $a_0 = \pm 1$. The function $F(s)$ shows that we have to take $a_0 = -1$.

Our w is of the form

$$(35) \quad w = l^{-1/2} \sum_{i \geq 0} a_i l^i \quad a_0 = -1 \text{ or } \omega = \sum_{i \geq 0} a_i l^i.$$

To find the coefficients a_i , $i \geq 1$ we have

$$\omega + 1 = \frac{1}{2c} \left[bl(\omega + 1)^2 + (b^2 - db)l^2 - 2bl(\omega + 1) + c(\omega + 1)^2 - dcl \right]$$

which corresponds to equation (13). Now

$$a_1 = -\frac{d}{2}, \quad a_2 = \frac{d^2}{8} + \frac{b^2}{2c},$$

$$a_i = \frac{1}{2c} \sum_{k=1}^{i-1} a_k (ba_{i-1-k} + ca_{i-k}) - ba_{i-1}, \quad i \geq 3.$$

To determine M and ρ (see I 4.) we have $S_1 = \left\{ z_0 = \frac{-c}{b} \right\}$, $S_2 = S_3 =$
 $= \left\{ z_{1,2} = \frac{-(cd+b) \pm \sqrt{(cd-b)^2 + 4cb^2}}{2(db-b^2)} \right\}$. So $\rho \leq \min \{ |z_0|, |z_1|, |z_2| \}$ and $M_\rho =$
 $= \sqrt{\frac{||c| + |cd+b|\rho + |db-b^2|\rho^2}{||c| - |b|\rho|}}$.

For the other two parameters ν and ρ we have

$$\left| \sum_{i \geq i_0+1} a_i l^{i-1/2} \right| \leq_T l^{1/2} \left\{ \sum_{i \geq i_0+1} |a_i| \frac{t^{i-2}}{(i-2)!} \right\} \\ \leq_T \frac{M}{\rho^{i_0+1}} eT/\rho l^{i_0+1/2},$$

whence

$$\nu = \frac{M}{\rho^{i_0+1}} eT/\rho \quad \text{and} \quad \delta = i_0 + 1/2.$$

One approximation of our function $F(s)$ is:

$$\tilde{x} = e^{-\sqrt{s}} \exp \left(\sum_{i=1}^{i_0} a_i l^{i-1/2} \right) \\ \tilde{x} = \{t^{-1} \Phi(0, -1/2, -t^{-1/2})\} \prod_{i=1}^{i_0} [(I + \{t^{-1} \Phi(0, i-1/2; a_i t^{i-1/2})\})].$$

We know that

$$|t^{-1} \Phi(0, -1/2; -t^{-1/2})| \leq \frac{3\sqrt{6}}{\sqrt{\pi}} e^{-3/2}, \quad t \geq 0$$

whence:

$$|F(s) - \tilde{x}| \leq \tau \frac{3\sqrt{6}}{\sqrt{\pi}} e^{-3/2} l \prod_{i=1}^{i_0} (I + \Omega_{i-1/2, a_i} l^{i-1/2}) \times \\ \times \Omega_{i_0+1/2, v} l^{i_0+1/2}.$$

For the special case $b=1$, $c=2$, $d=2$, $i_0=30$ we obtained:

$$M=1,6756229 \quad \text{and} \quad \rho=0,4384379,$$

i	a_i
1	-.1000000 E 01
2	.7500000 E 00
3	-.8750000 E 00
4	.1218750 E 01
5	-.1906250 E 01
6	.3218750 E 01
7	-.5722656 E 01
8	.1055127 E 02
9	-.1998486 E 02
10	.3864636 E 02
11	-.7597809 E 02
12	.1513983 E 03
13	-.3050873 E 03
14	.6206497 E 03
15	-.1272915 E 04
16	.2629128 E 04
17	-.5463891 E 04
18	.1141710 E 05
19	-.2397240 E 05
20	.5055314 E 05
21	-.1070232 E 06
22	.2273742 E 06
23	-.4846165 E 06
24	.1035925 E 07
25	-.2220371 E 07
26	.4770868 E 07
27	-.1027452 E 08
28	.2217400 E 08
29	-.4794906 E 08
30	.1038749 E 09

T	i_0	$ F(s) - \tilde{x} $
.125	1	.3740859 E 00
	2	.8288866 E-01
	3	.9763155 E-02
	4	.7982501 E-03
	5	.5058944 E-04
	6	.2622452 E-05
	7	.1150262 E-06
	8	.4372575 E-08
	9	.1466628 E-09
	10	.4401477 E-11
	11	.1195119 E-12
	12	.2962890 E-14
.250	1	.2853854 E 01
	2	.1664468 E 01
	3	.4604768 E 00
	4	.7854897 E-01
	5	.1002882 E-01
	6	.1040727 E-02
	7	.9130588 E-04
	8	.6941813 E-05
	9	.4656787 E-06
	10	.2795083 E-07
	11	.1517879 E-08
	12	.7526121 E-10
.500	1	.1439097 E 05
	2	.3212855 E 05
	3	.3878998 E 05
	4	.2127287 E 05
	5	.6519686 E 04
	6	.1421877 E 04
	7	.2521281 E 03
	8	.3840903 E 02
	9	.5154650 E 01
	10	.6188052 E 00
	11	.6720921 E-01
	12	.6664889 E-02

The first 30 coefficients a_i are computed and given in the first scheme. In the second scheme one can find the estimation of the difference $|F(s) - \tilde{x}|$ for three values of T and twelve of i_0 .

REFERENCES

- [1] Berg L., *Zur approximativen Umkehrung der Laplace Transformation*, ZAMM 55, 346—348 (1975)
- [2] Favard J., *Conrs d'analyse de l'Ecole polytechnique*, Paris T. II, p. 21
- [3] Roethe, A., Roethe K.—P., und D. Gelbin, *Bestimmung axialer Dispersionskoeffizienten in Gas-fest-Systemen*, Chemische Technik 23, 88—92, (1971).
- [4] Stanković B., *Approximate solutions of the operator linear differential equation I*, Publ. Inst. Math. 21 (35) 1977, 185—196.
- [5] *Higher Transcendental Functions*, Volume 1, McGraw-Hill Book 1953, p. 4