THE HIGH ORDER EXPONENTIALLY FITTED NONEQUIDISTANT EXTRAPOLATION METHODS FOR STIFF SYSTEMS*

XIANG JIA-XIANG (项家祥)
(*Shanghai Teachers University, Shanghai, China*)

Abstract

A class of exponentially fitted nonequidistant extrapolation methods based on an L-stabe linear single-step formula are studied. Theorems about the extrapolation coefficients are given. These methods keep good numerical stability, "quasi-L-stability", while raising the accuracy of the original formula.

§ 1. Introduction

In solving the initial value problem of O. D. E. s

$$\begin{cases} y'=f(y), \\ y(a)=\eta \end{cases}$$

the stiffness will become a serious difficulty when the ratio of the real part of the eigenvalues of the Jacobian $J=\frac{\partial f}{\partial y}$ is very large. Because of the importancy of this class of problems, the research for efficient computing methods is significant.

The technique of extrapolation has been widely used, but most such methods destroy the numerical stability when they raise the accuracy, and the amount of work increases as exponential of 2. J. R. Cash[4] introduced exponential fitted extrapolation methods to improve the stability of local extrapolation, but the eigenvalues of the Jacobian must be calculated, and this costs a lot of machine time.

In this paper, a class of exponentially fitted nonequidistant extrapolation (EFNE) methods are discussed, and the calculation of eigenvalues is no longer needed. With the use of the strategy of nonequidistant extrapolation, the amount of work increases linearly. Another advantage is that the new methods have quasi-L-stability, which is analogous to L-stability.

§ 2. The Construction of EFNE Methods

Consider the linear single-step formula of order 3:

$$y_{n+1}=y_n+\frac{1}{3}h(2f_{n+1}+f_n)-\frac{1}{6}h^2f'_{n+1}. \tag{1}$$

When it is applied to the test equation

* Received November 24, 1984.

$$y' = \lambda y, \quad y(0) = 1 \tag{2}$$

the characteristic function is

$$R_2^1(q) = \frac{1+q/3}{1-2q/3+q^2/6}, \tag{3}$$

where $q=\lambda h$. It is well known that (3) is L-acceptable and therefore (1) is L-stable. Now we try to construct new methods of higher accuracy based on (1). Assume that $y_1, \cdots, y_n$ have been calculated and the $(n+1)$th main step is to compute y_{n+1}. Let h be the step-length of this step, choose real numbers $m_i \geqslant 1$ $(i=1, \cdots, m)$, and calculate y_{n+1} for m times (each calculation denoted as $y_{n+1}^{(i)}$):

$$\begin{aligned} y_{n+\frac{1}{m_i}} &= y_n + \frac{1}{3}\frac{h}{m_i}(2f_{n+\frac{1}{m_i}} + f_n) - \frac{1}{6}\left(\frac{h}{m_i}\right)^2 f'_{n+\frac{1}{m_i}}, \\ y_{n+1}^{(i)} &= y_{n+\frac{1}{m_i}} + \frac{1}{3}\left(\frac{m_i-1}{m_i}h\right)(2f_{n+1}^{(i)} + f_{n+\frac{1}{m_i}}) - \frac{1}{6}\left(\frac{m_i-1}{m_i}h\right)^2 f_{n+1}^{(i)\prime}. \end{aligned} \tag{4}$$

Generally, let $m_1=1$, $m_i \neq m_j$, $i \neq j$. Then take the linear combination

$$y_{n+1} = \sum_{i=1}^{m} u_i y_{n+1}^{(i)}. \tag{5}$$

Connect (4), (5) using (3), and we have the characteristic function of the EFNE methods:

$$R(q) = \sum_{i=1}^{m} u_i R_2^1\left(\frac{q}{m_i}\right) R_2^1\left(\frac{m_i-1}{m_i}q\right). \tag{6}$$

Choose u_i so that $R(q)$ can approximate to e^q as closely as possible, i.e.

$$R(q) = e^q + O(q^{p+1}), \tag{7}$$

where p (positive integer) should be as large as possible.

Definition 1. *An EFNE method is of order p if*

$$R(q) = e^q + c_{p+1}q^{p+1} + O(q^{p+2}), \tag{8}$$

where $c_{p+1} \neq 0$ is the error constant of the method.

The following theorems show how u_i's are chosen and how m is confined.

These theorems are generally true for exponentially fitted nonequidistant extrapolation methods based on formulae different from (1).

Theorem 2. *Let $m_i > 1$, $i=2, \cdots, m$, and $m_i \neq m_j$, $i \neq j$. Denote*

$$a_{ij} = \frac{1+(m_j-1)^{i+2}}{m_j^{i+2}}, \quad i, j = 2, \cdots, m \leqslant 4, \tag{9}$$

$$A = \begin{bmatrix} 1 & 1 \cdots 1 \\ 1 & \\ \vdots & a_{ij} \\ 1 & \end{bmatrix}, \quad e_1 = (1, 0, \cdots, 0)^T$$

Then the coefficients u_i in (6) can be determined by

$$Au = e_1. \tag{10}$$

Proof. Let

$$\begin{aligned} R_j &= R_2^1\left(\frac{q}{m_j}\right) R_2^1\left(\frac{m_j-1}{m_j}q\right) \\ &= \frac{(1+q/3m_j)(1+(m_j-1)q/3m_j)}{(1-2q/3m_j+(q/m_j)^2/6)(1-2(m_j-1)q/3m_j+((m_j-1)q/m_j)^2/6)}, \end{aligned}$$

$$\ln R_j=\sum_{i=1}^{6}\frac{1}{i}[(-1)^{i+1}(q/3m_j)^i+(2q/3m_j-((q/m_j)^2/6)^i$$
$$+(-1)^{i+1}((m_j-1)q/3m_j)^i+(2(m_j-1)q/3m_j-(m_j-1)q/m_j)^2/6)^i]+O(q^7)$$
$$=q-\frac{1}{72}a_{2j}q^4-\frac{1}{270}a_{3j}q^5-\frac{1}{648}a_{4j}q^6+O(q^7),$$

where a_{ij} are defined as in (9).

Using Taylor's expansion of e^x and neglecting the high order terms, we have

$$R(q)=\sum_{j=1}^{m}u_jR_j=\sum_{j=1}^{m}u_j\exp\Big(q-\frac{1}{72}a_{2j}q^4-\frac{1}{270}a_{3j}q^5-\frac{1}{648}a_{4j}q^6\Big)+O(q^7)$$
$$=e^q\sum_{j=1}^{m}u_j\Big(1-\frac{1}{72}a_{2j}q^4-\frac{1}{270}a_{3j}q^5-\frac{1}{648}a_{4j}q^6\Big)+O(q^7)$$
$$=e^q\Big(\sum_{j=1}^{m}u_j-\frac{1}{72}\Big(\sum_{j=1}^{m}u_ja_{2j}\Big)q^4-\frac{1}{270}\Big(\sum_{j=1}^{m}u_ja_{3j}\Big)q^5-\frac{1}{648}\Big(\sum_{j=1}^{m}u_ja_{4j}\Big)q^6+\Big)O(q^7).$$

It is obvious that the EFNE method is of order p $(=m+2)$ if and only if

$$R(q)=e^q(1+d_{p+1}q^{p+1}+O(q^{p+2}))$$

with $d_{p+1}\neq 0$, that is,

$$\sum_{j=1}^{m}u_j=1,\quad \sum_{j=1}^{m}a_{ij}u_j=0,\quad i=2,\cdots,m\leqslant 4.$$

This is just (10). ∎

By using Theorem 2 the coefficients in the EFNE methods of orders 4 to 6 can be calculated easily, and the error constants c_{p+1} can be calculated by (8) (see Table 1).

Table 1

p	m				u				c_{p+1}
4	1	2			−1/7	8/7			−0.00026
5	1	2	3		1/4	24/5	−81/20		0.00006
6	1	2	3	4	−97/60	248/5	−9477/100	3584/75	−0.0000073

In this paper, m_i's are decided previously and u_i's vary along with m_i's. But how we can choose m_i's so that the absolute values of u_i's become small and the numerical stability becomes better is still an open problem.

Unfortunately we cannot construct new methods of order greater than 6 simply using (4) and (5) due to the following theorem:

Theorem 3. *Let A, e_1 be defined as in Theorem 2 with $m>4$. Then* (10) *is inconsistent.*

Proof. Obviously we only need to prove the case $m=5$.

It is easy to testify that the group of polynomials

$$1,\ x^4+(1-x)^4,\ \cdots,\ x^7+(1-x)^7$$

is linearly dependent. In fact, there exists a vector c:

$$(c_1, c_2, c_3, c_4, c_5)=\left(\frac{1}{20}, -\frac{7}{4}, \frac{21}{5}, -\frac{7}{2}, 1\right)$$

such that

$$c_1+\sum_{i=1}^{5} c_i(x^{i+2}+(1-x)^{i+2})\equiv 0.$$

So we can find an elementary matrix (since c_5 cannot be zero)

$$X=\begin{bmatrix} 1 & & \\ & \ddots & \\ c_1 & \cdots & c_5 \end{bmatrix}$$

such that the elements of the last row of XA are all zero. But

$$Xe_1=(1, 0, \cdots, 0, c_1)^T.$$

This verifies the conclusion for $m=5$. ▮

We can get an EFNE method of order greater than 6 by adding a node in the last extrapolation, i.e. choose $m_{51}>m_{52}>1$ and calculate

$$y_{n+\frac{1}{m_{51}}},\ y_{n+\frac{1}{m_{52}}},\ y^{(5)}_{n+1}$$

successively. The theorem for u_i should accordingly be modified, but it is beyond our discussion.

§ 3. Convergency and Stability

Notice that an EFNE method is equivalent to a one-step multiderivative method when both are applied to the test equation (2), that is, they have the same characteristic function. R. Jeltsch[5] has established the theory of multistep-multiderivative methods. We will quote some of his results to simplify our proof.

Theorem 4. *The EFNE methods are convergent is the serse of Dahlquist and Henrici.*

Proof. Now that

$$\Phi(\zeta, q)\equiv\zeta-R(q)=0,$$

so the characteristic function (6) is the unique branch. Since $\sum u_i=1$ we have $\rho_0(\zeta)=\zeta-1$ and therefore the methods are stable (zero-stable).

Since $\zeta(q)\equiv R(q)$ and $R(q)$ is analytic in a neighborhood of the origin (because $R(0)=1$) and $p\geqslant 4$, we have

$$\zeta(q)-e^q\equiv R(q)-e^q=c_{p+1}q^{p+1}+O(q^{p+2}),\quad c_{p+1}\neq 0.$$

By Theorem 2 of [5], the EFNE methods are convergent. ▮

Before the discussion of numerical stability, we first give

Definition 5. *A single-step method is said to be quasi-L-stable if it is stiffly stable and its characteristic function $R(q)$ satisfies*

$$\lim_{\mathrm{Re}\, q\to-\infty}|R(q)|=0.$$

Since (1) is L-stable and for (3) we have

$$\lim_{\operatorname{Re} q\to+\infty}|R_2^1(q)|=\lim_{\operatorname{Re} q\to-\infty}|R_2^1(q)|=0,$$
$$\lim_{\operatorname{Im} q\to+\infty}|R_2^1(q)|=\lim_{\operatorname{Im} q\to-\infty}|R_2^1(q)|=0, \tag{11}$$

(11) holds for (6). This means that if an EFNE method is A_0-stable, it must be quasi-L-stable. We sketch the regions of absolute stability of the EFNE methods of orders 4 to 6 in Fig. 1.

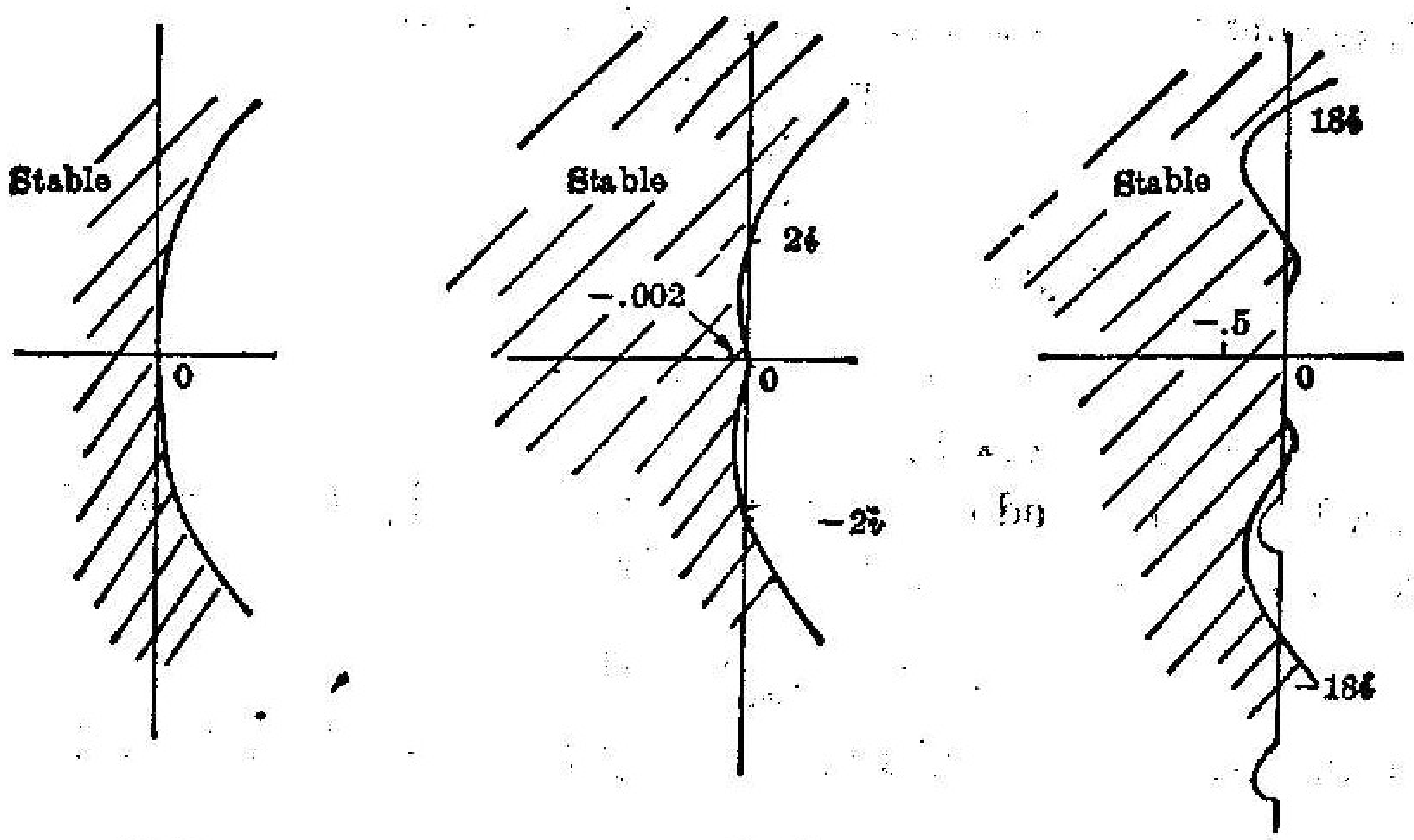


Fig. 1

We can see that the methods of orders 4 and 5 are L-stable and almost L-stable respectively, and the method of order 6 is quasi-L-stable. They are much better than Gear's methods. The comparison is given in Table 2.

Table 2

P	a_{min}		α_{max}	
	EFNE	GEAR	EFNE	GEAR
4	0	0.7	90°	73°
5	0.002	2.4	89.9°	51°
6	0.5	6.1	87.6°	18°

In order to indicate more precisely the numerical stability near the origin, we introduce the following symbols.

Let $R>0$,

$$D=D(R)=\{q\in\overline{\mathbb{C}}:\ |q|\leqslant R\},$$
$$A_s=\{q\in\overline{\mathbb{C}}:\ |R(q)|\leqslant 1\},$$

∂A_s be the boundary of A_s and A_s^c be the complement of A_s,

$$D^+=A_s^c\cap D,\qquad D^-=A_s\cap D,$$
$$H^+=\{q\in\overline{\mathbb{C}}:\ \operatorname{Re}(q)>0\},\qquad H^-=\{q\in\overline{\mathbb{C}}:\ \operatorname{Re}(q)\leqslant 0\}.$$

Theorem 6. *Let $\Phi(\zeta, q) \equiv \zeta - R(q) = 0$ be the characteristic equation of the EFNE methods, and*

$$\lambda = -\rho_1(\zeta(0))/\rho_0'(\zeta(0))$$

be a positive real number. Let

$$R(q) = e^q(1 + d_{p+1}q^{p+1} + d_{p+2}q^{p+2} + \cdots), \quad d_{p+1} \neq 0. \tag{12}$$

Then there exists $R > 0$ such that

$$\Gamma = \partial A_s \cap D$$

is a continuous smooth curve, the imaginary axis at the origin, and devides D into two parts it tangents:

$$D = D^+ \cup D^-.$$

a) *If p is odd and*

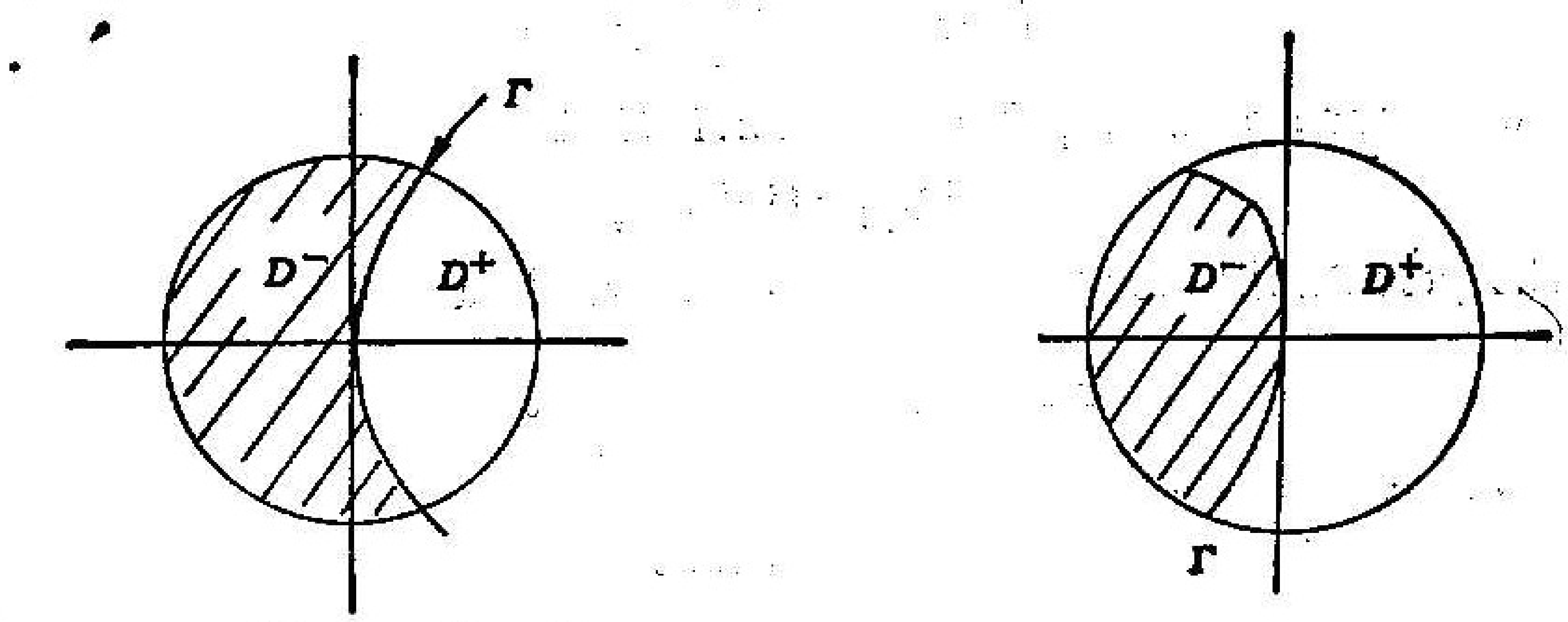
a.1) $d_{p+1}(-1)^{(p+1)/2} < 0$, *then* $D^+ - \{0\} \subset H^+$;

a.2) $d_{p+1}(-1)^{(p+1)/2} > 0$, *then* $D^- \subset H^-$.

b) *If p is even and $d_{p+2j} = 0$, $j = 1, \cdots, s-1$, $s < 2p$, $d_{p+2s} \neq 0$,*

b.1) $d_{p+2s}(-1)^{p/2+s} < 0$, *then* $D^+ - \{0\} \subset H^+$;

b.2) $d_{p+2s}(-1)^{p/2+s} > 0$, *then* $D^- \subset H^-$.



$D^+ - \{0\} \subset H^+$ $\qquad\qquad$ $D^- \subset H^-$

Fig. 2

Proof. The first part of the theorem can be verified directly by Theorem 1 of [6] since the EFNE methods are convergent and consistent. Therefore we only need to prove the second part.

Now we write $q = iy$ where $i^2 = -1$.

a) Assume that p is odd. Then by (12)

$$|R(iy)| = |e^{iy}|\,|1 + d_{p+1}(-1)^{(p+1)/2}y^{p+1} + O(y^{p+2})|$$
$$= \mathrm{sqrt}(1 + 2d_{p+1}(-1)^{(p+1)/2}y^{p+1} + O(y^{p+2})).$$

Obviously, there exists sufficiently small $R > 0$ such that $d_{p+1}(-1)^{(p+1)/2} < 0$ implies $|R(iy)| < 1$, i.e. $D^+ - \{0\} \subset H^+$. Otherwise $D^- \subset H^-$.

b) Let p be even. Now d_{p+2j} may be zero, $j = 1, \cdots$. Assume $d_{p+2s} \neq 0$ $(s < 2p)$. Then

$$|R(iy)| = |e^{iy}|\left|1 + \sum_{j=1}^{\infty} d_{p+2j-1}(iy)^{p+2j-1} + d_{p+2s}(-1)^{(p/2+s)}y^{p+2s} + O(y^{p+2s+2})\right|$$
$$= \mathrm{sqrt}(1 + 2d_{p+2s}(-1)^{p/2+s}y^{p+2s} + O(y^{p+2s+2})).$$

Then b.1), b.2) can be easily established as a) is. ∎

We see that for the EFNE method of order 4, $d_6 \doteq 0.00135$, $d_6(-1)^3<0$. Therefore we have $D^+-\{0\}\subset H^+$ near the origin. (In fact, it is L-stable.) And for the method of order 5, $d_6 \doteq -0.00006$ and $d_6(-1)^3>0$. So $D^-\subset H^-$, that is, it cannot be L-stable. So we can verify for the method of order 6.

§ 4. The Strategy of Step-Length Control

From Table 1 we see that the nodes of lower order method coincide with those of higher order ones. So it is natural to consider a strategy of step-length control like embedded methods.

Let $y(t_{n+1})$ be the true solution at t_{n+1} and y^s_{n+1} the computing value using the method of order s. Then

$$y^p_{n+1}-y(t_{n+1})=C_{p+1}h^{p+1}+O(h^{p+2}),$$
$$y^{p+1}_{n+1}-y(t_{n+1})=C_{p+2}h^{p+2}+O(h^{p+3}).$$

Neglecting the high order terms, we have

$$|h^{p+1}C_{p+1}|=|y^p_{n+1}-y^{p+1}_{n+1}|$$

or

$$|C_{p+1}|=\frac{|y^p_{n+1}-y^{p+1}_{n+1}|}{h^{p+1}}.$$

If we want to choose a new step-length $\bar{h}$ such that

$$|C_{p+1}\bar{h}^{p+1}|<\varepsilon,$$

where ε is the permissible error given previously, then

$$\bar{h}\doteq h\sqrt[p+1]{\frac{\varepsilon}{|y^p_{n+1}-y^{p+1}_{n+1}|}}.$$

We usually take

$$h_{\text{new}}=0.9\bar{h}.$$

§ 5. Numerical Examples

1. Consider the problem:

$$\begin{cases} u'=998u+1998v, \\ v'=-999u-1999v, \end{cases} \quad u(0)=1, \quad v(0)=0.$$

The true solution is

$$\begin{cases} u=2e^{-t}-e^{-1000t}, \\ v=-e^{-t}+e^{-1000t}. \end{cases}$$

The comparison of the accuracy of computing results using the EFNE methods of orders 4 to 6 and the original formula is given in Table 3 with a fixed step-length. One can see at once that the accuracy is improved when extrapolation is used.

Table 3

	order 3	order 4	order 5	order 6	t
max error	6.6E-6	1.8E-7	2.1E-8	8.6E-11	0.5
	3.1E-3	3.3E-5	1.1E-5	1.6E-6	2.0
	5.5E-9	6.4E-11	2.2E-11	2.9E-12	20.0

2. The example for testing the program for stiff systems introduced by Krogh[9]:

$$y' = -By + Uw, \tag{12}$$

$$U = \frac{1}{2}\begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}, \quad z = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix},$$

$$y = Uz, \quad w = (z_1^2\ z_2^2\ z_3^2\ z_4^2)^T,$$

$$B = U\ \mathrm{diag}(\beta_1\ \beta_2\ \beta_3\ \beta_4)U, \quad z_i' = -\beta_i z_i + z_i^2, \quad i = 1, 2, 3, 4.$$

The true solution

$$z_i = \frac{\beta_i}{1 + c_i e^{\beta_i t}}.$$

If $z_i(0) = -1$, then $c_i = -(1+\beta_i)$. The true solution of (12) is given by $y = Uz$. Let $\beta_i = 1000, 800, -10, 0.001$ respectively.

The computing results by the method of order 5 are given in Table 4. The method of order 4 is also used to control the step-length.

Table 4

max error	step number	evaluation	LU	average step-length	time
0.27E-6	11	112	11	0.91E-3	0.0101399
0.37E-5	17	192	17	0.63E-2	0.106844
0.12E-5	34	471	34	0.32E-1	1.09392
0.61E-5	47	653	47	0.21	10.048
0.25E-5	66	877	66	1.53	100.999
0.60E-5	86	1086	86	12.6	1079.00

Table 5 gives the computing results using Gear's method.

Table 5

max error	step number	evaluation	LU	average step-length	time
0.9E-7	70	179	7	0.15E-3	0.0102437
0.26E-5	110	262	12	0.95E-3	0.104887
0.22E-5	168	405	15	0.60E-2	1.012267
0.29E-5	216	523	20	0.46E-1	10.011079
0.30E-5	256	616	25	0.41	102.47713
0.12E-5	283	693	29	3.6	1025.7769

Comparing with Gear's method the EFNE methods have the following advantages:

1. Simple construction and program. It can start working by itself and change its step-length easily.

2. Its numerical stability is better than that of Gear's.

3. High accuracy.

The chief shortcoming, as that of the general extrapolation methods, is that the amount of work is large.

The author is indebted to Professor Kuang Jiao-xun and Wang Guo-rong for their help and guide.

References

[1] J. D. Lambert, Computational methods in O. D. E. s, John Wiley & Sons, 1976.
[2] C. W. Gear, Numerical initial value problems in O. D. E. s, Prentice Hall, Eaglewood Cliffs, 1971.
[3] E. C. Titchmarsh, The theory of functions, Oxford, 2nd ed., 1952.
[4] J. R. Cash, On the design of high order exponentially fitted formulae for the numerical integration of stiff systems, *Numer. Math.*, Vol. 36, 253—266.
[5] R. Jeltsch, On the stability regions of multistep multiderivative methods, Lecture Notes in Mathematics, 631, 63—80.
[6] R. Jeltsch, A necessary condition for A-stability multistep multiderivative methods, *Math. Comp.*, 30, 739.
[7] Li Rui-fu, Cheng Qi-xiang, The theory of complex functions, 1979.