stichting
mathematisch
centrum

MC

AFDELING NUMERIEKE WISKUNDE
(DEPARTMENT OF NUMERICAL MATHEMATICS)

NW 112/81 OKTOBER

J.G. VERWER

ON THE CONTRACTIVITY OF A COMPLEX RUNGE-KUTTA SCHEME

Preprint

kruislaan 413 1098 SJ amsterdam

Printed at the Mathematical Centre, 413 Kruislaan, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

1980 Mathematics subject classification: 65L05

ACM-Computing Reviews-category: 5.17

On the contractivity of a complex Runge-Kutta scheme *)

by

J.G. Verwer

ABSTRACT

This note reports an attempt to break the order barrier $p \leq 1$ for contractive Runge-Kutta schemes with a special Runge-Kutta scheme of order 2 containing a complex parameter.

KEY WORDS & PHRASES: *Numerical analysis, Runge-Kutta methods, Stiff problems, Contractivity*

---

*) This report will be submitted for publication elsewhere.

Let $\mathbb{K}$ denote the set of real or complex numbers $\mathbb{R}$ or $\mathbb{C}$. Let

$$y' = f(t,y) \ , \ t \geq 0 \ , \ y(0) = y_0 \tag{1}$$

be a differential system in $\mathbb{K}^s$, $s \geq 1$, where f is assumed to be continuous on $[0,\infty) \times \mathbb{K}^s$. Following SPIJKER [3] we define $\bar{F}$ as the class of all continuous functions f for which

$$| \tilde{y}(t_2) - y(t_2) | \leq | \tilde{y}(t_1) - y(t_1) |, \text{ all } t_2 \geq t_1 \geq 0, \tag{2}$$

where $\tilde{y}$ and y denote any two solutions of (1) and $|\cdot|$ stands for some given norm on $\mathbb{K}^s$. Hence, given a norm, $\bar{F}$ is the class of all differential systems (1) with *contractive solutions* with respect to the given norm. A function f belongs to class $\bar{F}$ if for all pairs $(t,\zeta) \in [0,\infty) \times \mathbb{K}^s$ (or an appropriate subspace) we have $\mu[\partial f(t,\zeta)/\partial y] \leq 0$, where μ is the logarithmic norm corresponding to the norm chosen for $\bar{F}$(see e.g. DAHLQUIST [1]).

SPIJKER [3] has investigated the contractivity of numerical solutions $\{u_n\}$ defined by the implicit Runge-Kutta method

$$\begin{aligned} U_i &= u_n + \tau \sum_{j=1}^{m} a_{ij} f(t_n + c_j\tau, U_j) \ , \ 1 \leq j \leq m, \\ u_{n+1} &= u_n + \tau \sum_{i=1}^{m} b_i f(t_n + c_i\tau, U_i), \ n \geq 0. \end{aligned} \tag{3}$$

Here $c_i = a_{i1} + \ldots + a_{im}$ while b_i, a_{ij} are real parameters with $b_1 + \ldots + b_m = 1$. Given a function $f \in \bar{F}$, then method (3) is called *contractive with respect to f*, if, for all $\tau > 0$,

$$| \tilde{u}_{n+1} - u_{n+1} | \leq | \tilde{u}_n - u_n | \tag{4}$$

for any two sequences $\{\tilde{u}_n\}$, $\{u_n\}$ produced by the method.

Spijker's paper now contains a remarkable result, viz., *if method (3) is contractive with respect to all* $f \in \bar{F}$, *where* $|\cdot|$ *may be an arbitrary norm , then its order of accuracy p is not greater than* 1. Hence, if a

Runge-Kutta method is required to be contractive on the whole class F, with respect to any norm, then the *order barrier* $p \leq 1$ is unavoidable (NEVANLINNA & LINIGER [2] proved this result for the one-leg method). An example of a method which is contractive on the whole class F, and for any norm, is Backward Euler. This observation led the author to the following problem.

PROBLEM. Consider the 2-stage integration scheme

$$u_{n+\alpha} = u_n + \alpha\tau\, f(t_n+\alpha\tau, u_{n+\alpha}), \tag{5a}$$

$$u_{n+1} = u_{n+\alpha} + (1-\alpha)\,\tau\, f(t_{n+1}, u_{n+1}). \tag{5b}$$

The computation $(t_n,u_n) \to (t_{n+1},u_{n+1})$ consists of two consecutive backward Euler steps, viz., $(t_n,u_n) \to (t_n+\alpha\tau, u_{n+\alpha})$ and $(t_n+\alpha\tau, u_{n+\alpha}) \to (t_{n+1}, u_{n+1})$. Method (5) may also be considered as a Runge-Kutta method (3) having the Butcher-matrix

$$\begin{array}{c|c} c & A \\ \hline & b^T \end{array} = \begin{array}{c|cc} \alpha & \alpha & 0 \\ 1 & \alpha & 1-\alpha \\ \hline & \alpha & 1-\alpha \end{array}. \tag{6}$$

By choosing α *complex*, namely $\alpha = \frac{1}{2} \pm \frac{1}{2} i$, its order of accuracy $p = 2$ while its stability function R is given by $R(z) = (1-z+\frac{1}{2}z^2)^{-1}$ which is A-stable. Now the question arises whether, for $\alpha = \frac{1}{2} \pm \frac{1}{2}i$, the above complex backward Euler schemes share the contractivity property of the real backward Euler scheme? If so, Spijker's order barrier can be broken with a complex Runge-Kutta method (3).

Unfortunately, the answer to our question turned out to be negative. By taking α complex, the backward Euler schemes (5a),(5b) lose the nice, aforementioned contractivity property. An immediate consequence is that Spijker's order barrier cannot be broken with method (5). We will prove this by means of a counter example.

Consider the function $f : \mathbb{C} \times \mathbb{K}^2 \to \mathbb{K}^2$ defined by

$$(7) \qquad f(t,y) = M(t)y \ , \ M(t) = (m_{ij}(t)) = \begin{pmatrix} 0 & \mathrm{Re}\ \phi(t) \\ 0 & -\mathrm{Re}\ \phi(t) \end{pmatrix},$$

where $\phi : \mathbb{C} \to \mathbb{C}$ is continuous and satisfies Re $\phi(t) \geq 0$, Re$(t) \geq 0$. If we select the ℓ^1-norm in $\mathbb{K}^2$ the logarithmic norm μ_1 is given by

$$(8) \qquad \mu_1[M(t)] = \max_j \left(\mathrm{Re}\ m_{jj} + \sum_{i \neq j} |\, m_{ij}\,|\right) = 0, \ \mathrm{Re}(t) \geq 0.$$

Hence by restricting t to $[0,\infty]$, and by using the ℓ^1-norm, $f \in F$. Observe that this is not true for the ℓ^2-norm and ℓ^∞-norm. Because method (5) takes two internal "complex time steps" we defined ϕ from $\mathbb{C}$ to $\mathbb{C}$. Now introduce the notation $\lambda(t) = \mathrm{Re}\ \phi(t)$. Applying (5) to system (1),(7) then yields $u_{n+\alpha} = B_{n+\alpha}\, u_n, u_{n+1} = B_{n+1}\, u_{n+\alpha}$, where

$$(9) \qquad B_{n+\alpha} = \begin{bmatrix} 1 & \dfrac{\alpha\tau\lambda(t_n+\alpha\tau)}{1+\alpha\tau\lambda(t_n+\alpha\tau)} \\ 0 & \dfrac{1}{1+\alpha\tau\lambda(t_n+\alpha\tau)} \end{bmatrix}, \ B_{n+1} = \begin{bmatrix} 1 & \dfrac{(1-\alpha)\tau\lambda(t_{n+1})}{1+(1-\alpha)\tau\lambda(t_{n+1})} \\ 0 & \dfrac{1}{1+(1-\alpha)\tau\lambda(t_{n+1})} \end{bmatrix}.$$

We first observe that if α is real and $0 < \alpha \leq 1$, then $|B_{n+\alpha}|_1 = |B_{n+1}|_1 = 1$ which implies contractivity. On the other hand, if Im $\alpha \neq 0$, it is immediate that $|B_{n+\alpha}|_1 > 1$ for all $\tau\lambda(t_n+\alpha\tau) > 0$. For the vector $u_n = [0,1]^T$, e.g., it thus follows that $|u_{n+\alpha}|_1 = |B_{n+\alpha}|_1 > 1$ for all $\tau\lambda(t_n+\alpha\tau) > 0$. This means that for α complex the backward Euler scheme (5a) is not contractive on F. Further, by choosing λ in such a way that $\lambda(t_n+\alpha\tau) > 0$ and $\lambda(t_{n+1}) = 0$ it immediately follows that $|u_{n+1}|_1 = |u_{n+\alpha}|_1 > |u_n|_1$. Hence for complex α-values the combined method (5a),(5b) is not contractive either.

We conclude with an inequality which is valid for all members from F:

LEMMA. *Let* m *be a nonnegative real. Let* $f \in F$ *be such that for some given norm* $\mu[\partial f(t,\zeta)/\partial y] \leq -m$ *for all* (t,ζ). *For any two points* $(t,u_0),(t,\tilde{u}_0)$ *the results* $u_\alpha,\tilde{u}_\alpha$ *computed by the backward Euler scheme* (5a), *where Re* $\alpha > 0$, *then satisfy the inequality*

$$(10) \qquad |\, \tilde{u}_\alpha - u_\alpha \,| \leq \frac{\mathrm{Abs}\ \alpha^{-1}}{\tau m + \mathrm{Re}\ \alpha^{-1}}\, |\, \tilde{u}_0 - u_0 \,|, \ \text{all } \tau > 0. \qquad \square$$

This lemma can be proven by making use of known logarithmic norm properties and of some results from DAHLQUIST [1], section 1.3. If we substitute $\alpha = \frac{1}{2} \pm \frac{1}{2} i$, or $1-\alpha$, the constant involved becomes $\sqrt{2}/(1+\tau m)$. Thus we can guarantee contractivity for the second order method (5), if $\tau m \geq \sqrt{2}-1$.

REFERENCES

[1] DAHLQUIST, G., *Stability and error bounds in the numerical integration of ordinary differential equations*, Trans. Royal Inst. of Technology, No. 130, Stockholm, 1959.

[2] NEVANLINNA, O. & W. LINIGER, *Contractive methods for stiff differential equations*, BIT 18, 457-474, 1978 and BIT 19, 53-72, 1979.

[3] SPIJKER, M.N., *Contractivity of Runge-Kutta methods*, in : *Numerical methods for solving stiff initial value problems*, Bericht Nr. 9 der Institut für Geometrie und Praktische Mathematik der RWTH Aachen, eds. G. Dahlquist and R. Jeltsch, 1981.