

ΧΑΡΑΛΑΜΠΟΥ ΓΡΗΓ. ΤΣΙΤΟΥΡΑ

ΠΤΥΧΙΟΥΧΟΥ ΜΑΘΗΜΑΤΙΚΩΝ ΠΑΝ.ΑΘΗΝΩΝ

ΠΤΥΧΙΟΥΧΟΥ Ε.Ε.Π.Ε.Ε. ΠΑΝ.ΑΘΗΝΩΝ

**ΣΥΝΕΧΕΙΣ ΕΠΕΚΤΑΣΕΙΣ ΜΕΘΟΔΩΝ
RUNGE-KUTTA ΓΙΑ ΠΡΟΒΛΗΜΑΤΑ
ΑΡΧΙΚΩΝ ΤΙΜΩΝ**

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

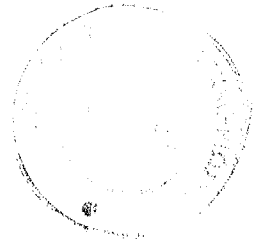
ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΑΘΗΝΑ 1990

ΧΑΡΑΛΑΜΠΟΥ ΓΡΗΓ. ΤΣΙΤΟΥΡΑ

ΠΤΥΧΙΟΥΧΟΥ ΜΑΘΗΜΑΤΙΚΩΝ ΠΑΝ.ΑΘΗΝΩΝ

ΠΤΥΧΙΟΥΧΟΥ Ε.Ε.Π.Ε.Ε. ΠΑΝ.ΑΘΗΝΩΝ



ΣΥΝΕΧΕΙΣ ΕΠΕΚΤΑΣΕΙΣ ΜΕΘΟΔΩΝ
RUNGE-KUTTA ΓΙΑ ΠΡΟΒΛΗΜΑΤΑ
ΑΡΧΙΚΩΝ ΤΙΜΩΝ

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΑΘΗΝΑ 1990

*Στή μνήμη του
Πατέρα μου*

Π Ρ Ο Λ Ο Γ Ο Σ

Η Παρούσα Διδακτορική Διατριβή εκπονήθηκε στο Εθνικό Μετσόβιο Πολυτεχνείο από τον Ιούνιο του 1987 ως τον Σεπτέμβριο του 1990. Για την επιτυχή διεκπεραίωσή της θα ήθελα να ευχαριστήσω τους

--Γιώργο Παπαγεωργίου, αφ' ενός μεν για την πάρα πολύ καλή επιστημονική του καθοδήγηση, η οποία συνίσταται, μεταξύ άλλων, στην άριστη επιλογή του θέματος, τη πλήρη βιβλιογραφική ενημέρωση που μου παρέσχε και στη σωστή κατεύθυνση που έδωσε στην έρευνα που διεξήχθη, αφ' ετέρου δε, διότι υπήρξε φίλος και με βοηθήσε σε όλες τις δυσκολίες, επιστημονικές και μη.

- Α.Ράπτη και Ι.Χρυσοβέργη, μέλη της τριμελούς συμβουλευτικής μου επιτροπής, για την συμπαράστασή τους στη προσπάθειά μου.

- Τομέα Μαθηματικών, Γενικού Τμήματος του Ε.Μ.Π., για την υποτροφία που μου χορήγησαν και τη δυνατότητα που μου έδωσαν να χρησιμοποιήσω τον υλικό εξοπλισμό τους.

- την Μητέρα μου για την ανοχή και την φροντίδα της.

- την Μνηστή μου Κ.Βακαλοπούλου, για την υπομονή και την κατανόηση που έδειξε καθώς και την αμέριστη συμπαράσταση της.

Π Ε Ρ Ι Ε Χ Ο Μ Ε Ν Α

ΚΕΦΑΛΑΙΟ & Τίτλος	σελ.
ΕΙΣΑΓΩΓΗ	1
ΚΕΦ. 1 ^{ΟΥ} ΒΑΣΙΚΗ ΘΕΩΡΙΑ ΤΩΝ ΑΡΙΘΜΗΤΙΚΩΝ ΜΕΘΟΔΩΝ	7
1.1 Εισαγωγή	7
1.2 Μέθοδος απλού βήματος. Βασικές έννοιες	10
1.3 Η μέθοδος του Euler	14
ΚΕΦ. 2 ^{ΟΥ} ΜΕΘΟΔΟΙ RUNGE-KUTTA ΚΑΙ RUNGE-KUTTA-NYSTROM	18
2.1 Εισαγωγή	19
2.2 Μέθοδοι Runge-Kutta (RK)	20
2.3 Εξισώσεις προσδιορισμού των παραμέτρων RK	24
2.4 Εξισώσεις συνθηκών RK μέχρι 6 ^{ης} τάξης	29
2.5 Embedded Runge-Kutta	34
2.6 Μεθοδοι Runge-Kutta-Nystrom (RKN)-Γενική θεωρία	38
2.7 Απόλυτη ευστάθεια	47
ΚΕΦ. 3 ^{ΟΥ} SCALED RUNGE-KUTTA ΜΕΘΟΔΟΙ	49
3.1 Γενικά περι scaling	49
3.2 Scaling - Γενική θεωρία	50
3.3 Scaling 4 ^{ης} τάξης	55
3.4 Κατασκευή μεθόδων 4 ^{ης} τάξης	60
3.5 Μέθοδοι 5 ^{ης} τάξης	66
3.6 Αριθμητικά αποτελέσματα	72
ΚΕΦ. 4 ^{ΟΥ} ΠΑΡΕΜΒΟΛΙΚΑ ΠΟΛΥΩΝΥΜΑ ΓΙΑ ΜΕΘΟΔΟΥΣ RUNGE-KUTTA	77
4.1 Εισαγωγή	77
4.2 Γενική θεωρία	79
4.3 Παρεμβολικά πολυώνυμα δύο βημάτων για μεθόδους 5 ^{ης} τάξης	83

4.4	Ανάλυση σφάλματος για μεθόδους 5 ^{ης} τάξης	87
4.5	Μέθοδοι ανώτερης τάξης	92
4.6	Αποτελέσματα - Συμπεράσματα	98
4.7	Μονοβηματικό interpolant	102
ΚΕΦ. 5 ^ο	SCALED RUNGE-KUTTA-NYSTROM ΜΕΘΟΔΟΙ	107
5.1	Εισαγωγή	107
5.2	Scaled Runge-Kutta-Nystrom μέθοδοι. Γενική θεωρία	108
5.3	Scaled Runge-Kutta-Nystrom χαμηλής τάξης	115
5.4	SRKN μέθοδοι υψηλής τάξης	122
5.5	Αριθμητικά αποτελέσματα	128
ΚΕΦ. 6 ^ο		
6.1	C ² SRKN αλγόριθμοι	133
6.2	Κατασκευή SRKN-Fehlberg 45 μεθόδου	137
6.3	Παρεμβολή για RKN μεθόδους	137
ΕΠΙΛΟΓΟΣ	ΣΥΜΠΕΡΑΣΜΑΤΑ - ΠΡΟΟΠΤΙΚΕΣ	145
ΒΙΒΛΙΟΓΡΑΦΙΑ		151

Π Ε Ρ Ι Λ Η Ψ Η

Οι συνήθεις διαφορικές εξισώσεις 1ης και 2ης τάξης εμφανίζονται σε πολλές περιοχές των Εφαρμοσμένων Μαθηματικών και της Τεχνολογίας.

Με την ραγδαία ανάπτυξη των Ηλεκτρονικών Υπολογιστών υπάρχει και παράλληλη ανάπτυξη ικανών αριθμητικών μεθόδων διάφορων τύπων για την προσέγγιση της λύσης των προβλημάτων αρχικών τιμών 1ης και 2ης τάξης, της μορφής :

$$y' = f(x,y) , y(x_0) = y_0 , x \geq x_0 \quad (1)$$

και

$$y'' = f(x,y) , y(x_0) = y_0 , y'(x_0) = y'_0 , x \geq x_0 \quad (2)$$

αντίστοιχα.

Οι μέθοδοι *Runge-Kutta* για διαφορικές εξισώσεις 1ης τάξης και *Runge-Kutta-Nystrom* για διαφορικές εξισώσεις 2ης τάξης ανήκουν σήμερα στις πύο δημοφιλείς κατηγορίες μεθόδων για την προσέγγιση της λύσης πολλών διαφορικών εξισώσεων. Οι παραπάνω μέθοδοι βρίσκονται συνεχώς στην ερευνητική επικαιρότητα, βελτιώνοντας αφ' ενός την μεθοδολογία κατασκευής τους και τα ποιοτικά τους χαρακτηριστικά, και αφ' ετέρου δημιουργώντας νέες ιδιότητες και προοπτικές γι' αυτές.

Η πλειοψηφία των μεθόδων τύπου *Runge-Kutta* έχουν σχεδιαστεί να παράγουν προσεγγίσεις $y_n, n=0,1,\dots,N$ σε διακεκριμένα σημεία $x_n, n=0,1,2,\dots,N$ του διαστήματος ολοκλήρωσης $[x_0, x_{\text{end}}]$. Ο αλγόριθμος προχωρεί βήμα προς βήμα επιλέγοντας αυτόματα το βήμα ολοκλήρωσης $h_n = x_{n+1} - x_n$, έτσι ώστε η εκτίμηση του τοπικού σφάλματος αποκοπής να είναι μικρότερη από μία προκαθορισμένη ανοχή TOL. Ομως σε πολλές εφαρμογές, π.χ. πυκνή έξοδος

προσεγγίσεων, προβλήματα με ασυνέχειες κ.λ.π. απαιτείται μια συνεχής προσέγγιση σ' ολόκληρο το διάστημα ενός βήματος. Μέχρι πολύ πρόσφατα οι μέθοδοι τύπου *Runge-Kutta* δεν είχαν αυτή τη δυνατότητα χωρίς την ουσιαστική αύξηση του υπολογιστικού κόστους.

Στην παρούσα διατριβή ασχολούμεθα με την αριθμητική επίλυση των διαφορικών εξισώσεων της μορφής (1) και (2), και συγκεκριμένα με την μελέτη και κατασκευή μεθόδων τύπου *Runge-Kutta* που έχουν παρεμβολικές ιδιότητες. Με τις νέες μεθόδους επεκτείνεται η ήδη υπάρχουσα λύση στα διακριτά σημεία και γίνεται συνεχής. Επίσης μελετώνται οι νέες μέθοδοι, έτσι ώστε, η συνεχής αυτή προσέγγιση της λύσης να διατηρεί την ακρίβεια της βασικής μεθόδου με το ελάχιστο κατά το δυνατόν πρόσθετο υπολογιστικό κόστος, καθώς επίσης καταβάλλεται προσπάθεια οι συναρτήσεις-πολυώνυμα της συνεχούς επέκτασης που κατασκευάζονται να είναι λείες στα κομβικά σημεία, δηλαδή C^1 ή C^2 .

CONTINUOUS EXTENSIONS OF RUNGE-KUTTA TYPE METHODS FOR INITIAL VALUE PROBLEMS

by

Charalampos Tsitouras

National Technical University of Athens

Doctoral Dissertation

SUMMARY

The numerical solution of first and second order ordinary differential equations is one of the most important application areas for initial value methods.

Runge-Kutta methods for the problem

$$y' = f(x,y), y(x_0)=y_0, x \geq x_0 \quad (1)$$

and Runge-Kutta-Nystrom methods for the problem

$$y'' = f(x,y), y(x_0)=y_0, y'(x_0)=y'_0, x \geq x_0 \quad (2)$$

are known to be efficient for the treatment of non-stiff differential equation of the form (1) and (2). It is well known that most explicit (RK) codes for the numerical solution of (1) and (RKN) codes for the numerical solution of (2) are designed to produce approximations at discrete points in the integration interval, selecting automatically the step size. However, in many applications (e.g. problems with dense output, discontinuous IVP's, Delay differential equations) a continuous approximation in the entire interval, or some part of it is required. The main disadvantage, of (RK) and (RKN) methods, is the lack of capability to obtain continuous approximation in an inexpensive way.

The subject of this dissertation is to develop continuous extensions of Runge-Kutta and Runge-Kutta-Nystrom methods for

initial value problems of first and second order with the minimum possible additional computational cost. Especially, we first construct, Scaled Runge-Kutta (SRK) and Scaled Runge-Kutta-Nystrom (SRKN) methods for most of the well known methods of several orders from these two categories.

Second, a general procedure for the construction of interpolants for (RK) formulas, based on values from two successive steps is presented. As an application interpolants for specific methods of several orders are constructed. Finally, C^2 interpolants for (RKN) methods of low and high order are constructed. All these interpolants are very efficient and more inexpensive comparing with other well known interpolants.

ΕΙΣΑΓΩΓΗ

Στην παρούσα διατριβή ασχολούμεθα με την μελέτη και κατασκευή μεθόδων οι οποίες δίνουν συνεχείς προσεγγίσεις της λύσης των συνήθων διαφορικών εξισώσεων 1^{ης} και 2^{ης} τάξης με αρχική ή αρχικές τιμές αντίστοιχα.

Η γενική μορφή του προβλήματος αρχικών τιμών για συνήθεις διαφορικές εξισώσεις 1^{ης} τάξης μπορεί να γραφτεί:

$$y'(x) = f(x, y(x)), \quad x \in [x_0, x_{\text{end}}]$$

0-1

$$y(x_0) = y_0$$

όπου γενικά $y: \mathbb{R} \rightarrow \mathbb{R}^m$ και $f: \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$.

Η γενική μορφή του προβλήματος αρχικών τιμών για συνήθεις διαφορικές εξισώσεις 2^{ης} τάξης γράφεται:

$$y''(x) = f(x, y(x), y'(x)), \quad x \in [x_0, x_{\text{end}}]$$

0-2

$$y(x_0) = y_0, \quad y'(x_0) = y'_0$$

όπου επίσης $y: \mathbb{R} \rightarrow \mathbb{R}^m$, $y': \mathbb{R} \rightarrow \mathbb{R}^m$ και $f: \mathbb{R} \times \mathbb{R}^{2m} \rightarrow \mathbb{R}^m$.

Σημαντικό ενδιαφέρον παρουσιάζουν οι διαφορικές εξισώσεις 2^{ης} τάξης ειδικής μορφής. Το πρόβλημα αρχικών τιμών γι' αυτή την κατηγορία είναι:

$$y'' = f(x, y(x)) , x \in [x_0, x_{\text{end}}]$$

0-3

$$y(x_0) = y_0 , y'(x_0) = y'_0$$

όπου $y : \mathbb{R} \rightarrow \mathbb{R}^m$, $y' : \mathbb{R} \rightarrow \mathbb{R}^m$ και $f : \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$.

Τα τελευταία χρόνια με την ανάπτυξη των υπολογιστών έχουν αναπτυχθεί διάφοροι τύποι αριθμητικών μεθόδων για την προσέγγιση της λύσης $y(x)$ ενός προβλήματος αρχικών τιμών. Αυτές οι μέθοδοι μπορούν να χωριστούν σε δύο κατηγορίες. Τις συνεχείς και τις διακριτές. Η πρώτη περιλαμβάνει τις collocation μεθόδους ενώ η δεύτερη περιλαμβάνει τις Extrapolation (μεθόδους παρεκβολής), τις μεθόδους Runge-Kutta καθώς και τις γραμμικές πολυβηματικές μεθόδους. Αυτές οι δύο κατηγορίες δεν είναι τελείως ανεξάρτητες. Γενικά κάθε μέθοδος της πρώτης κατηγορίας μπορεί να χρησιμοποιηθεί για να παράγουμε προσεγγίσεις σε διακριτά σημεία. Αντίθετα, αυτό ισχύει μόνο για μερικές μεθόδους της δεύτερης κατηγορίας. Οι περισσότερες όμως διακριτές μέθοδοι όπως οι Runge-Kutta και Extrapolation ,δεν μπορούν να παράγουν συνεχείς λύσεις για τα προβλήματα αρχικών τιμών. Αυτό έχει σαν αποτέλεσμα οι μέθοδοι αυτές να είναι ανεπαρκείς για προβλήματα τα οποία απαιτούν ολικά (globally) συνεχή λύση ή πυκνή έξοδο, όπως απαιτεί για παράδειγμα το γράφημα της λύσης.

Το πρόβλημα της επέκτασης μιας διακριτής αριθμητικής μεθόδου σε συνεχή επέκταση έχει απασχολήσει τα τελευταία χρόνια διάφορους ερευνητές. Ειδικά στην κατηγορία των

Runge-Kutta μεθόδων που υπάρχει η μεγαλύτερη δυσκολία έχει γίνει μεγάλη πρόοδος. Πρώτη η Horn [1983] εφάρμοσε την τεχνική του scaling για να επεκτείνει μια μέθοδο Runge-Kutta σε συνεχή, χωρίς τα αποτελέσματά της να είναι τα καλύτερα δυνατά. Αργότερα πολλοί άλλοι έχουν ασχοληθεί με το πρόβλημα της συνεχούς επέκτασης μιας Runge-Kutta μεθόδου. Shampine [1985],[1986], Enright et.al [1986], Calvo et.al. [1989],[1990], Zennaro [1986], Owren & Zennaro [1990], Bellen-Zennaro [1988], Verner [1990a] ,[1990b], Dormand & Prince [1986],[1987] ,Papageorgiou et.al.[1987],[1988],[1989] ,Tsitouras & Papageorgiou [1990] .

Ο βασικός σκοπός αυτής της διατριβής είναι να αναπτύξουμε και να μελετήσουμε νέες μεθόδους οι οποίες καθιστούν τις ήδη υπάρχουσες μεθόδους Runge-Kutta και Runge-Kutta-Nystrom ικανές να παράγουν συνεχείς λύσεις για τα προβλήματα αρχικών τιμών 0-1 και 0-3 αντίστοιχα, με όσο το δυνατόν μικρότερο πρόσθετο κόστος και με ακρίβεια της ίδιας τάξης με την ακρίβεια των αρχικών μεθόδων. Για να το επιτύχουμε αυτό πρέπει να καταφύγουμε σε τεχνικές που λαμβάνουν υπ' όψη την βασική φιλοσοφία κατασκευής των μεθόδων Runge-Kutta και Runge-Kutta-Nystrom .Αυτό σημαίνει ότι πρέπει να λύσουμε τις εξισώσεις που ορίζουν τους συντελεστές αυτών των μεθόδων θεωρώντας ότι εξαρτώνται από μια παράμετρο σ ώστε να έχουμε λύση στη θέση $x_n + \sigma(x_{n+1} - x_n) = x_n + \sigma h_n$.

Στα πλαίσια της διατριβής αυτής και συγκεκριμένα στα δύο πρώτα κεφάλαια αναπτύσσεται η βασική θεωρία των αριθμητικών

μεθόδων και των μεθόδων Runge-Kutta και Runge-Kutta-Nystrom. Εν συνεχεία στο 3^ο κεφάλαιο δόθηκαν τεχνικές αποφυγής των περιττών υπολογισμών με ταυτόχρονη ομαλοποίηση της λύσης σε C^1 συνάρτηση. Αυτό επιτεύχθηκε με την εισαγωγή της έννοιας του *scaling* για την παράγωγο στις Σ.Δ.Ε 1^{ης} τάξης και λύση των σχετικών εξισώσεων με άλλους τρόπους. Για τις πιο δημοφιλείς μεθόδους δόθηκαν νέες συνεχείς επεκτάσεις που εκτός των παραπάνω έχουν ελαχιστοποιημένα τα σφαλματά τους και βελτιωμένα τα ποιοτικά χαρακτηριστικά τους.

Για τις μεθόδους του Fehlberg 4(5) [1969] και των Dormand και Prince (4)5 [1980] κατασκευάστηκαν scaled μέθοδοι 4^{ης} τάξης με τα παραπάνω χαρακτηριστικά. Επίσης προτείνεται νέος τρόπος λύσης των εξισώσεων 5^{ης} τάξης που μας επιτρέπει την προσέγγιση λύσης σε ένα προκαθορισμένο σημείο με ένα υπολογισμό της f λιγότερο, από τις μέχρι τώρα γνωστές μεθόδους, σε κάθε βήμα της ολοκλήρωσης. Τέλος δίνουμε αριθμητικά tests σε παραδείγματα από το DETEST, Hull et.al. [1972], το διεθνώς αποδεκτό πρόγραμμα ελέγχου των αριθμητικών μεθόδων για λύση Σ.Δ.Ε. Περισσότερα αποτελέσματα και παραδείγματα μεθόδων βρίσκονται στην εργασία Parageorgiou et.al [1989]

Στο 4^ο κεφάλαιο εξετάζουμε την κατασκευή παρεμβολικών πολυωνύμων τα οποία συνοδεύουν μια μέθοδο Runge-Kutta για την προσέγγιση του προβλήματος 0-1 ώστε να δίνει συνεχείς προσεγγίσεις. Αυτά βασίζονται σε έναν επαρκή αριθμό πληροφοριών από ένα βήμα ολοκλήρωσης $[x_n, x_{n+1}]$. Για μεθόδους

Runge-Kutta υψηλής τάξης απαιτούνται πρόσθετες τιμές της λύσης και της παραγώγου, οι οποίες μπορούν να υπολογιστούν εφαρμόζοντας σε συνδυασμό την τεχνική του scaling. Οι Enright et.al [1986] και ο Shampine [1985],[1986] ασχολήθηκαν με την περίπτωση αυτή δίνοντας ώθηση στην έρευνα προς την κατεύθυνση αυτή. Στο κεφάλαιο αυτό προτείνουμε αρχικά νέες επεκτάσεις Runge-Kutta μεθόδων οι οποίες έχουν μικρότερο κόστος από αντίστοιχες που έχουν προταθεί και καλύτερη ακρίβεια. Εν συνεχεία αποδεικνύουμε ότι με κατάλληλη επιλογή σημείων έξω από το τρέχον διάστημα $[x_n, x_{n+1}]$ μπορούμε να επιτύχουμε καλλίτερες μεθόδους παρεμβολής χωρίς κόστος. Μερικά από τα αποτελέσματα αυτά έχουν δημοσιευθεί στην εργασία Tsitouras & Papageorgiou [1990].

Στο πέμπτο κεφάλαιο εφαρμόζουμε την τεχνική του scaling για την επέκταση μεθόδων τύπου Runge-Kutta-Nystrom, που προσεγγίζουν την λύση του προβλήματος 0-3, σε συνεχή. Εδώ για πρώτη φορά αναλύσαμε τις ιδιαιτερότητες που μπορούν να εμφανιστούν από την εφαρμογή του scaling σ' αυτή την κατηγορία μεθόδων. Η γενική θεωρία εφαρμόζεται στις πιο γνωστές μεθόδους αυτής της κατηγορίας για τις οποίες παράγονται οι αντίστοιχες συνεχείς επεκτάσεις των οποίων αναλύονται τα διάφορα χαρακτηριστικά και για τις οποίες γίνονται αριθμητικά tests. Μερικά από τα αποτελέσματα δίνονται στις εργασίες Papageorgiou & Tsitouras [1987],[1989]. Στο έκτο κεφάλαιο αναλύεται η δυνατότητα του να κατασκευάσει κανείς C^2 scaled μεθόδους για τις Runge-Kutta-Nystrom

μεθόδους. Μια μέθοδος τετάρτης τάξης κατασκευάζεται που ανήκει στην κλάση C^2 . Επίσης για πρώτη φορά κατασκευάζονται 8ης τάξης interpolants (παρεμβολικά πολυώνυμα) με ελαχιστοποιημένα τα τοπικά σφάλματα αποκοπής.

Τέλος στο έβδομο κεφάλαιο αναφέρονται οι προοπτικές του θέματος με το οποίο ασχοληθήκαμε. Εξετάζονται οι δυνατότητες για κατασκευή πολυωνύμων παρεμβολής υψηλής τάξης. Αναλύονται τεχνικές για περαιτέρω μείωση του υπολογιστικού κόστους. Η πιθανότητα παρεμβολής σε περιοδικά προβλήματα επίσης εξετάζεται. Ακόμα οι εφαρμογές που μπορεί να έχουν τα αποτελέσματα σε παρεμφερείς περιοχές δίνονται. Ειδική αναφορά γίνεται στις delay - διαφορικές εξισώσεις, στη τεχνική του defect correction και στα προβλήματα με ασυνέχειες που χρησιμοποιούν τεχνικές παρεμβολής.

Κ Ε Φ Α Λ Α Ι Ο 1^{ΟΝ}

ΒΑΣΙΚΗ ΘΕΩΡΙΑ ΤΩΝ ΑΡΙΘΜΗΤΙΚΩΝ ΜΕΘΟΔΩΝ

1.1 Εισαγωγή

Πολλά προβλήματα της Επιστήμης και της Τεχνολογίας καταλήγουν σε Μαθηματικά προβλήματα της μορφής 0-1 και 0-3. Για να εξασφαλίσει κανείς ύπαρξη μοναδικής λύσης πρέπει για την f να ισχύει το επόμενο θεώρημα

Θεώρημα 1.1

Εστω $f(x,y)$ μια συνάρτηση ορισμένη και συνεχής σε όλα τα σημεία (x,y) του χωρίου που ορίζεται από $x_0 \leq x \leq x_{end}$, $y \in \mathbb{R}^m$. Υποθέτουμε ότι υπάρχει μια σταθερά L , έτσι ώστε

$$\| f(x,y) - f(x,y^*) \| \leq L \cdot \| y - y^* \| \quad 1-1-1$$

για κάθε (x,y) στο χωρίο. Τότε αν y_0 είναι η αρχική συνθήκη του Π.Α.Τ, υπάρχει μοναδική λύση του $y(x)$ που είναι συνεχής και διαφορίσιμη. Για απόδειξη βλέπε Henrici [1962].

Η συνθήκη 1-1-1 είναι γνωστή ως συνθήκη Lipschitz και η L ως σταθερά Lipschitz. Αν D το χωρίο στο οποίο βρίσκονται οι τιμές (x,y) τότε

$$L = \sup_{(x,y) \in D} \left\| \frac{\partial f}{\partial y} \right\| \quad 1-1-2$$

όπου $\partial f/\partial y$ η Ιακωβιανή της f ως προς y .

Επειδή η f εν γένει είναι διάνυσματική συνάρτηση είναι φανερό ότι αν έχουμε διαφορική εξίσωση ανώτερης τάξης, με μετατροπή των παραγώγων σε συντεταγμένες της f προκύπτει σύστημα διαφορικών εξισώσεων πρώτης τάξης και τα παραπάνω θα ισχύουν όσον αφορά την ύπαρξη και μοναδικότητα της λύσης.

Η συνθήκη 1-1-1 είναι ικανή για να έχει κανείς μοναδική λύση αλλά αρκετά περιοριστική. Η απλή συνάρτηση π.χ. y^2 δεν την ικανοποιεί. Συνήθως χρησιμοποιούμε κάποιον περιορισμό της. Για παραπέρα ανάπτυξη όμως του θέματος μπορεί να δει κανείς το κλασικό βιβλίο του Henrici [1961], καθώς και τα Ελληνικά κείμενα επί του θέματος του Παπαγεωργίου [1985] και του Δουγαλή [1987].

Πολύ σημαντικά είναι τα βιβλία ανθρώπων που ασχολούνται με την αριθμητική επίλυση συνήθων διαφορικών εξισώσεων και πρώτα απ' όλα του J.C. Butcher [1987]. Επίσης πολλοί άλλοι ερευνητές πρέπει να αναφερθούν, επειδή ασχολήθηκαν με τη συγγραφή μονογραφιών πάνω στο αντικείμενό μας όπως οι Gear [1971], Lambert [1973], Hairer, Norsett, Wanner [1987], Fatunla [1988].

Σήμερα με την ανάπτυξη των υπολογιστών οι αριθμητικές μέθοδοι είναι πολύ αποδοτικές και δίνουν προσεγγίσεις των λύσεων που ζητάμε με σημαντική ακρίβεια σε σύντομο χρονικό διάστημα. Συνήθως αφού δοθεί ένα πρόβλημα αρχικών τιμών και εξασφαλιστεί ότι ισχύει το θεώρημα 1.1, θεωρούμε μια

ακολουθία σημείων $\{x_n\}$ που δεν είναι απαραίτητο να ισαπέχουν. Κατόπιν ζητάμε προσέγγιση της λύσης όχι στο συνεχές διάστημα ορισμού της f αλλά στο σύνολο των διακεκριμένων σημείων $\{x_n\}$.

Οι κυριότερες κατηγορίες διακριτών μεθόδων για την αριθμητική επίλυση του προβλήματος αρχικών τιμών είναι

(α) Μέθοδοι απλού βήματος με κύριο εκπρόσωπο τις μεθόδους τύπου Runge-Kutta

(β) Μέθοδοι πολλαπλού βήματος που συνήθως προκύπτουν από παρεμβολικά πολυώνυμα

Συνδιασμοί των μεθόδων (α) και (β) ή γενικεύσεις αυτών δημιουργούν άλλες κατηγορίες από τις οποίες πιά γνωστές είναι.

(i) Hybrid μέθοδοι που είναι πολυβηματικές Runge-Kutta μέθοδοι

(ii) Extrapolation μέθοδοι

(iii) Μη γραμμικές πολυβηματικές μέθοδοι

Αν τώρα y_n μια προσέγγιση της θεωρητικής λύσης στο σημείο x_n , δηλαδή της $y(x_n)$, τότε οι μέθοδοι αυτές μπορεί να παρασταθούν από το γενικό τύπο.

$$\sum_{i=0}^k a_i \cdot y_{n+i} = h \cdot \Phi_f (x_n ; y_{n+k}, y_{n+k-1}, \dots, y_n, h) \quad 1-1-3$$

για $0 \leq n \leq N-k$, N το πλήθος των σημείων της διαμέρισης. Τα $\{a_i\}$ $i=0,1,\dots,k$ είναι σταθερές, $a_k \neq 0$, και οι αρχικές

τιμές y_0, y_1, \dots, y_{k-1} θεωρούνται γνωστές. Αν η συνάρτηση Φ_f είναι ανεξάρτητη της y_{n+k} ή των παραγώγων της, $y_{n+k}^{[n]}$, η μέθοδος λέγεται άμεση διαφορετικά λέγεται έμμεση. Επίσης αν $k=1$ έχουμε μέθοδο απλού βήματος ενώ αν $k > 1$ έχουμε πολυβηματική μέθοδο. Στην συνέχεια θα ασχοληθούμε ειδικότερα με τις μεθόδους απλού βήματος στις οποίες βασίζεται και το θέμα της διατριβής αυτής.

1.2 Μέθοδος απλού βήματος. Βασικές έννοιες.

Θεωρούμε το πρόβλημα αρχικών τιμών όπως ορίζεται από τις (0-1). Ο συνήθης τρόπος έκφρασης των μονοβηματικών μεθόδων είναι :

$$y_{n+1} = y_n + h_n \cdot \Phi(x_n, y_n, h_n) \quad 1-2-1$$

όπου $\Phi(x_n, y_n, h_n)$ είναι η συνάρτηση αύξησης και $h_n = x_{n+1} - x_n$. Για ευκολία μερικές φορές το h_n μπορεί να λαμβάνεται σταθερό ίσο με h .

Ορισμός 1-1 Ο τύπος ολοκλήρωσης 1-2-1 θα λέγεται συνεπής με το πρόβλημα αρχικών τιμών (Π.Α.Τ) 0-1 όταν η συνάρτηση $\Phi(x, y, h)$ ικανοποιεί την ακόλουθη σχέση:

$$\Phi(x, y, 0) = f(x, y) \quad 1-2-2$$

Η συνέπεια για μια μέθοδο 1ης τάξης εξασφαλίζει ότι αυτή

είναι τουλάχιστον πρώτης τάξης και προσεγγίζει κατά κάποιο τρόπο την λύση.

Ορισμός 1-2 Το τοπικό σφάλμα διακριτοποίησης δ_{n+1} (Τ.Σ.Δ) της μεθόδου 1-2-1 δίνεται από τη σχέση.

$$\delta_{n+1} = y(x_{n+1}) - y(x_n) - h \cdot \Phi(x_n, y(x_n), h), \quad 1-2-3$$

όπου $y(x)$ η αληθής λύση του (Π.Α.Τ).

Το Τ.Σ.Δ ορίζει απλά το πόσο η πραγματική λύση του Π.Α.Τ αποτυγχάνει να ικανοποιήσει την πρώτη τάξης εξίσωση διαφορών 1-2-1 κάτω από την απλοποιητική παραδοχή ότι οι προηγούμενες λύσεις ήταν ακριβείς (δηλ. $y_n = y(x_n)$). Ομως μετά η βήματα έχουμε να λύσουμε το τοπικό πρόβλημα

$$u' = f(x, u), \quad u(x_n) = y_n \quad 1-2-3a$$

με τοπική λύση την $u(x)$.

Ορισμός 1-2-α Το Τοπικό σφάλμα αποκοπής (Τ.Σ.Α) για το τοπικό πρόβλημα 1-2-3a είναι,

$$t_{n+1} = y_{n+1} - u(x_{n+1}) \quad 1-2-3b$$

Υποθέτουμε ότι η θεωρητική λύση ανήκει στην κλάση $C^\infty [a, b]$

Ορισμός 1-3 Η μέθοδος 1-2-1 λέμε ότι έχει τάξη p αν p είναι ο μεγαλύτερος θετικός ακέραιος έτσι ώστε το Τ.Σ.Α να

ικανοποιεί τη σχέση.

$$t_{n+1} = O(h^{p+1}) \quad 1-2-4$$

όπου $O(h^{p+1})$ σημαίνει ότι υπάρχουν δύο πεπερασμένες σταθερές C & $h_0 > 0$ έτσι ώστε

$$t_{n+1} \leq C \cdot h^{p+1}$$

για κάθε $h \leq h_0$.

Ορισμός 1-4 Το ολικό σφάλμα e_{n+1} της μεθόδου 1-2-1 ορίζεται από την σχέση,

$$e_{n+1} = y_{n+1} - y(x_{n+1}) \quad 1-2-5$$

Ορισμός 1-5 Η μέθοδος 1-1-1 λέμε ότι συγκλίνει για κάποια αρχική συνθήκη y_0 και ένα αυθαίρετο σημείο στο διάστημα $[x_0, x_{end}]$ αν

$$\lim_{h \rightarrow 0} \max_n e_n \rightarrow 0 \quad 1-2-6$$

Ορισμός 1-6 Η μέθοδος 1-2-1 είναι ευσταθής αν για κάθε αρχικό σφάλμα e_0 , υπάρχουν σταθερές K και $h_0 > 0$ έτσι ώστε, όταν η 1-2-1 εφαρμόζεται στο Π.Α.Τ με βήμα $h \in (0, h_0]$, το ολικό σφάλμα e_n ικανοποιεί την σχέση,

$$e_n \leq K \cdot e_0, K < \infty \quad 1-2-7$$

Μια ενδιαφέρουσα έννοια ευστάθειας είναι και η απόλυτη ευστάθεια η οποία έχει νόημα όταν έχουμε Π.Α.Τ με τις ιδιοτιμές των Ιακωβιανών τους να έχουν αρνητικά πραγματικά μέρη. Οι ιδιότητες της απόλυτης ευστάθειας μπορούν να εξετασθούν εφαρμόζοντας την 1-2-1 στο πρόβλημα μοντέλο

$$y' = \lambda y \quad 1-2-8$$

οπότε δημιουργείται η παρακάτω εξίσωση διαφορών

$$y_{n+1} = \mu(z) y_n, z = \lambda h \quad 1-2-9$$

όπου $\mu(z)$ πολυώνυμο του z .

Ορισμός 1-7 Η μονοβηματική μεθοδος 1-2-1 είναι απόλυτα ευσταθής σε ένα χωρίο του μιγαδικού επιπέδου αν η $\mu(z)$ ικανοποιεί την

$$|\mu(z)| < 1 \quad 1-2-10$$

καί τότε το χωρίο απόλυτης ευστάθειας R ορίζεται ως

$$R = \{ z / \text{ισχύει η 1-2-10} \} \quad 1-2-11$$

Ορίσμος 1-8 Η μέθοδος 1-2-1 λέγεται A-ευσταθής όταν το χωρίο απόλυτης ευστάθειας της περιλαμβάνει όλο το αριστερό μέρος του μιγαδικού επιπέδου.

Στη συνέχεια θεωρώντας την βασική μέθοδο του Euler αναλύουμε βασικά θεωρητικά αποτελέσματα τα οποία μπορούν να θεωρηθούν σαν βάση για γενικότερες μονοβηματικές μεθόδους.

1.3 Η μέθοδος του Euler.

Ο Euler [1768] πρότεινε την απλούστερη και τη περισσότερο αναλυμένη μέθοδο επίλυσης του Π.Α.Τ. Προκειται για μια μονοβηματική μέθοδο στην οποία σύμφωνα με τα παραπάνω έχουμε τη Φ να ορίζεται ως εξής:

$$\Phi(x, y, h) = f(x, y) \quad 1-3-1$$

Δηλαδή

$$y_{n+1} = y_n + h f(x_n, y_n) \quad 1-3-2$$

Παρακάτω αποδεικνύεται η σύγκλιση της μεθόδου του Euler

Λήμμα 1-1 Εστω $\{e_j, j=0(1),n\}$ ένα σύνολο πραγματικών αριθμών. Αν υπάρχουν πεπερασμένοι R και S έτσι ώστε

$$|e_{j+1}| \leq R \cdot |e_j| + S, \quad j=0(1)n-1 \quad 1-3-3$$

τότε

$$|e_j| \leq \frac{R^j - 1}{R - 1} S + R^j |e_0| \quad 1-3-4$$

Θεώρημα 1-2 Εστω ότι η $f(x,y)$ ικανοποιεί την συνθήκη ύπαρξης τότε η μέθοδος του Euler συγκλίνει.

Απόδειξη. Εστω t_{n+1} και e_{n+1} το Τ.Σ.Α και το ολικό σφάλμα αντίστοιχα της μεθόδου Euler τότε

$$y(x_{n+1}) = y(x_n) + hf(x_n, y(x_n)) + \delta_{n+1} \Rightarrow \quad 1-3-5$$

$$e_{n+1} = e_n + h(f(x_n, y(x_n)) - f(x_n, y_n)) + \delta_{n+1} \quad 1-3-6$$

Αν

$$\max_{x_0 \leq x_n \leq x_{end}} |t_n| = T \quad 1-3-7$$

και K, L οι Lipschitz σταθερές της $f(x,y)$ για το y και x , αντίστοιχα τότε από 1-2-5 έχουμε

$$|e_{n+1}| \leq (1+hL) |e_n| + T, \quad n=0,1,\dots \quad 1-3-8$$

Απο το λήμα 1-1 με $R=1+hL$ και $S=T$ έχουμε

$$|e_n| \leq \frac{(1+hL)^n - 1}{hL} T + (1+hL)^n |e_0|$$

Αφού $(1+hL)^n \leq e^{nhL} = e^{L(x_{\text{end}}-x_0)} \Rightarrow e^{L(x_{\text{end}}-x_0)} \geq e^{L(x_n-x_0)} \Rightarrow$

$$|e_n| \leq \frac{e^{L(x_{\text{end}}-x_0)} - 1}{hL} T + e^{L(x_{\text{end}}-x_0)} |e_0| \quad 1-3-9$$

Απο το Θεώρημα μέσης τιμής \Rightarrow

$$t_{n+1} = h(f(x_n + \theta h, y(x_n + \theta h)) - f(x_n, y(x_n))) \quad , 0 \leq \theta \leq 1 \Rightarrow$$

$$|t_{n+1}| \leq h |f(x_n + \theta h, y(x_n + \theta h)) - f(x_n + \theta h, y(x_n))| +$$

$$+ h |f(x_n + \theta h, y(x_n)) - f(x_n, y(x_n))| \leq hL |y(x_n + \theta h) - y(x_n)| + Kh^2 \theta =$$

$$Lh^2 \theta |y'(\xi)| + Kh^2 \quad , x_n \leq \xi \leq x_{n+1}$$

Εστω Y το ανω φράγμα της $y'(x)$ στο $[x_0, x_{\text{end}}]$, δηλ.

$$Y = \max_{x_0 \leq x \leq x_{\text{end}}} |y'(x)| \quad 1-3-10$$

Ετσι $|t_{n+1}| \leq h^2 \theta (LY+K) \leq h^2 (LY+K) = T$ και με $T = h^2 (LY+K)$ η 1-2-9 γίνεται

$$|e_n| \leq h \frac{e^{L(x_{\text{end}}-x_0)} - 1}{L} (LY+K) + e^{L(x_{\text{end}}-x_0)} |e_0| \quad 1-3-11$$

οπότε οριακά έχουμε

$$\lim_{\substack{h \rightarrow 0 \\ |e_0| \rightarrow 0}} e_n = 0 \quad \text{και} \quad \lim_{\substack{h \rightarrow 0 \\ |e_0| \rightarrow 0}} y_n = y(x_n)$$

που συμπληρώνει τήν απόδειξη.

Θεώρημα 1-3 Εστω ότι η $\Phi(x,y,h)$ είναι συνεχής για κάθε $(x,y) \in \mathbb{R}^{m+1}$ (όπου m η διασταση του Π.Α.Τ) και $0 < h < h_0$ και η Φ ικανοποιεί την συνθήκη Lipschitz στο y . Τότε η 1-1-1 συγκλίνει εάν και μόνο εάν είναι συνεπής. (Henrici [1962])

Κ Ε Φ Α Λ Α Ι Ο 2^{ΟΝ}

ΜΕΘΟΔΟΙ RUNGE-KUTTA ΚΑΙ RUNGE-KUTTA-NYSTROM

2.1 Εισαγωγή.

Την ιδέα επέκτασης της μεθόδου Euler για την προσέγγιση του Π.Α.Τ 1^{ης} τάξης, η οποία επιτρέπει την πολλαπλή χρήση υπολογισμών της f μέσα σε κάθε βήμα ολοκλήρωσης πρότείνει ο Runge [1895]. Εν συνεχεία ο Heun [1900] και ο Kutta [1901] εφάρμοσαν αυτή την ιδέα και πρότειναν συγκεκριμένες μεθόδους οι οποίες σήμερα αναφέρονται σαν μέθοδοι τύπου Runge-Kutta. Ειδικότερα, ο Kutta χαρακτήρισε πλήρως το σύνολο των μεθόδων Runge-Kutta 4^{ης} τάξης και πρότείνει την πρώτη μέθοδο 5^{ης} τάξης.

Αργότερα ο Nystrom [1925], πρότείνει ειδικές μεθόδους για δεύτερης τάξης διαφορικές εξισώσεις, και είχε επίσης την δική του προσφορά στην ανάπτυξη μεθόδων για το Π.Α.Τ 1^{ης} τάξης. Τέλος ο Huta [1957] ήταν αυτός που πρωτοασχολήθηκε με μεθόδους 6^{ης} τάξης.

Με την εμφάνιση και ανάπτυξη των Ηλεκτρονικών Υπολογιστών, ανανεώθηκε το ενδιαφέρον για τις μεθόδους Runge-Kutta. Ένας μεγάλος αριθμός ερευνητών συνέβαλε στις πρόσφατες επεκτάσεις της θεωρίας καθώς και στην ανάπτυξη συγκεκριμένων μεθόδων. Ο Butcher [1963],[1965],[1966],[1972] βασιζόμενος στην δουλειά

του Gill [1951] και κυρίως του πρωτοποριακού Merson [1957] συστηματικοποίησε την ανάλυση και θεωρία των μεθόδων Runge-Kutta και πρότεινε (Butcher [1964]), νέες μεθόδους υψηλής τάξης.

Αργότερα ο Fehlberg [1968],[1969],[1970] εισήγαγε την ιδέα των embedded (εμφυτευμένων) μεθόδων Runge-Kutta. Μια embedded Runge-Kutta μέθοδος δίνει προσεγγίσεις της λύσης τάξης p και $p+1$ χρησιμοποιώντας τους ίδιους υπολογισμούς, οπότε αν τις συνδυάσουμε μπορούμε να έχουμε μια εκτίμηση του σφάλματος της μεθόδου τάξης p . Μετά τον Fehlberg ένας μεγάλος αριθμός ερευνητών πρότεινε embedded Runge-Kutta μεθόδους όπως οι Verner [1978a],[1978b], Dormand & Prince [1980], Zonneveld [1964], Shintani [1966]. Επίσης και πολλές Runge-Kutta-Nystrom μέθοδοι διαφόρων τάξεων έχουν προταθεί από τους Fehlberg [1972],[1974], Hairer [1982], Bettis [1973], Dormand et.al [1986].

2.2 Μέθοδοι Runge-Kutta

Ας θεωρήσουμε το πρόβλημα 0-1 και την μέθοδο απλού βήματος 1-2-1. Τότε η συνάρτηση αύξησης $\Phi(x_n, y_n, h)$ γράφεται.

$$\Phi(x_n, y_n, h) = \sum_{r=0}^{p-1} \frac{h^r}{(r+1)!} f^{[r]}(x_n, y_n) \quad 2-2-1$$

όπου $f^{[r]}(x_n, y_n)$, $r=0(1)p-1$ είναι η r -παράγωγος της $f(x, y(x))$.

$$\text{Δηλ. } f^{[r]}(x, y(x)) = y^{[r+1]}(x) = \left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y} \right)^r f(x, y(x)) \quad 2-2-2$$

Το Τ.Σ.Α της 2-2-1 είναι

$$t_{n+1} = \frac{h^{p+1}}{(p+1)!} f^{[p]}(x_n, y_n) + O(h^{p+2}) \quad 2-2-3$$

Ορισμός 2-1. Εστω s ακέραιος αριθμός (ο αριθμός βαθμίδων) και $a_{21}, a_{31}, a_{32}, \dots, a_{s,s-1}, b_1, \dots, b_s, c_1, \dots, c_s$ πραγματικοί αριθμοί. Τότε η μέθοδος

$$f^1 = f(x_n, y_n)$$

$$f^2 = f(x_n + c_2 h, y_n + h a_{21} f^1)$$

...

2-2-4a

$$f^s = f(x_n + c_s h, y_n + h(a_{s1} f^1 + \dots + a_{s,s-1} f^{s-1}))$$

$$y_{n+1} = y_n + h(b_1 f^1 + b_2 f^2 + \dots + b_s f^s) = y_n + \Phi_{RK}(x_n, y_n, h)$$

της οποίας οι παράμετροι σε μορφή πινάκων μπορούν να γραφτούν.

$$c \begin{array}{|c} A \\ \hline b^T \end{array}$$

2-2-4b

όπου $b, c \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times m}$

καλείται μια s -βαθμίδων άμεση *Runge-Kutta* μέθοδος. Η αλγεβρική τάξη της 2-2-4 σύμφωνα με τον ορισμό 1-3 και την παρατήρηση 2-2-3 είναι p . Για να το επιτύχουμε αυτό πρέπει να βρούμε παραμετρους a_{ij}, b_i, c_i έτσι ώστε το ανάπτυγμα της συνάρτησης $\Phi_{RK}(x_n, y_n, h)$ σε δυνάμεις του h , εφαρμόζοντας θεώρημα Taylor στο σημείο (x_n, y_n) , να διαφέρει από το ανάπτυγμα της $\Phi(x_n, y_n, h)$ στην 2-2-1, μόνο στην p -στη ή μεγαλύτερες δυνάμεις του h . Είναι φανερό ότι για μεγάλο p απαιτούνται πολύπλοκες αλγεβρικές πράξεις ώστε να προσδιορισθούν οι σταθερές.

Οι μέθοδοι αυτές με βάση τον ορισμό 1-5 συγκλίνουν, δηλαδή παράγουν λύση που πράγματι προσεγγίζει την αληθινή λύση. Ακολουθεί σε γενικές γραμμές η απόδειξη αυτού του θεωρήματος. Για πιο λεπτομερή απόδειξη βλέπε Henrici [1962].

Λήμμα 2-2. Αν το Π.Α.Τ έχει μοναδική λύση τότε η Φ_{RK} ικανοποιεί μια συνθήκη Lipschitz για την εξαρτημένη μεταβλητή y .

Απόδειξη. Εστω L η σταθερά Lipschitz της $f(x, y)$ για την y και

$$Y_1(y_n) = f(x_n, y_n), \quad Y_1(z_n) = f(x_n, z_n)$$

τότε

$$|Y_1(y_n) - Y_1(z_n)| = |f(x_n, y_n) - f(x_n, z_n)| \leq L|y_n - z_n| \quad 2-2-5$$

Ομοίως

$$|Y_2(y_n) - Y_2(z_n)| = |f(x_n + c_2 h, y_n + h a_{21} Y_1(y_n)) - f(x_n + c_2 h, y_n + h a_{21} Y_1(z_n))| \leq$$

$$\leq L\{|y_n - z_n| + h|a_{21}| |Y_1(y_n) - Y_1(z_n)|\} \leq L|y_n - z_n| \{1 + hL|a_{21}|\}. \quad 2-2-6$$

οπότε για μεθοδο *Runge-Kutta* δυο βαθμίδων έχουμε

$$\begin{aligned} & |\Phi_{RK2}(x_n, y_n, h) - \Phi_{RK2}(x_n, z_n, h)| = \\ & = |b_1 Y_1(y_n) + b_2 Y_2(y_n) - (b_1 Y_1(z_n) + b_2 Y_2(z_n))| \leq \\ & \leq |b_1| |Y_1(y_n) - Y_1(z_n)| + |b_2| |Y_2(y_n) - Y_2(z_n)|. \end{aligned}$$

Απο 2-2-5 & 2-2-6 $\Rightarrow |\Phi_{RK2}(x_n, y_n, h) - \Phi_{RK2}(x_n, z_n, h)| \leq L^* |y_n - z_n|$,

οπου $L^* = L\{|b_1| + |b_2|(1 + hL|a_{21}|)\}$

Το προηγούμενο λήμμα ισχύει με τελείως προφανείς επεκτάσεις και για μεθόδους με περισσότερες βαθμίδες, δηλαδή για οποιοδήποτε Φ και είναι χρήσιμο στο επόμενο θεώρημα που εξασφαλίζει την ευστάθεια σύμφωνα με τον ορισμό 1-6.

Fatunla [1988]

Θεώρημα 2-2 Αν το Π.Α.Τ έχει μοναδική λύση τότε η μέθοδος 2-2-4 είναι ευσταθής.

Απόδειξη. Εστω $\{y_n\}$, $\{z_n\}$ δύο σύνολα από λύσεις που κατασκευάστηκαν από μια *Runge-Kutta* μέθοδο με αρχικές συνθήκες $y(x_0) = y_0$, $z(x_0) = z_0$

$$|y_0 - z_0| = \delta'_0$$

Εστω

$$\delta'_n = y_n - z_n, \quad n \geq 0$$

και

$$y_{n+1} = y_n + h\Phi(x_n, y_n, h), \quad 2-2-7$$

$$z_{n+1} = z_n + h\Phi(x_n, z_n, h) \quad 2-2-8$$

$$\text{Απο τριγωνική ανισότητα} \Rightarrow |\delta'_{n+1}| \leq (1 + hL^*) |\delta'_n|, \quad n \geq 0 \quad 2-2-9$$

Απο το λήμμα 1-1 και την 2-2-9 με $R=1+hL^*$, $S=0 \Rightarrow$

$$|\delta'_n| \leq K |\delta'_0| \quad \text{με} \quad K = e^{L^*(x_{\text{end}} - x_0)} \quad \text{Ο.Ε.Δ.}$$

Το θεώρημα 1-3 ολοκληρώνει την απόδειξη της σύγκλισης.

2.3 Εξισώσεις προσδιορισμού των παραμέτρων Runge-Kutta

Για την εύρεση των παραμέτρων της μεθόδου *Runge-Kutta* ακολουθείται μια διαδικασία επίπονη. Κατ' αρχάς θα υποθέσουμε ότι έχουμε ένα σύστημα διαφορικών εξισώσεων και μάλιστα αυτόνομο, προσθέτοντας την ανεξάρτητη μεταβλητή x στις εξαρτημένες μεταβλητές.

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ f(x, y) \end{bmatrix} \quad 2-3-1$$

με y m -διάστατο διάνυσμα όπως και το y' ενώ η $f(x,y): \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ είναι μια συνάρτηση με τις γνωστές ιδιότητες. Από τον ορισμό της γενικής μεθόδου Runge-Kutta έχουμε

$$f_j^i = f(y_n + h \sum_{k=1}^{i-1} a_{ik} f_j^k) \Rightarrow f_1^i = f(x_n + h \sum_{k=1}^{i-1} a_{ik}) \quad 2-3-2$$

όπου οι εκθέτες δείχνουν τη βαθμίδα της μεθόδου Runge-Kutta, και οι δείκτες τη συντεταγμένη της διαφορικής εξίσωσης. Δηλαδή από τη πρώτη συντεταγμένη που πλέον είναι η ανεξάρτητη μεταβλητή έχουμε με βάση την 2-3-2

$$\sum_{j=1}^{s-1} a_{ij} = c_i \quad i=2,3,\dots,s \quad 2-3-3$$

Η σχέση αυτή οδηγεί μεν στην απλοποίηση 2-3-1 αλλά είναι και χρήσιμη στην απαλοιφή αργότερα εξισώσεων που προκύπτουν για τις παραμέτρους της μεθόδου Runge-Kutta. Επίσης εξασφαλίζει ότι τα f^i είναι πρώτης τάξης προσεγγίσεις της $y'(x_n + c_i h)$ και είναι ενδεικτική τυχόν άλλων παραδοχών που θα μπορούσε να κάνει κανείς για περαιτέρω απλοποιήσεις των εξισώσεων. Γενικεύοντας την 2-2-7 και θεωρώντας το διαφορικό

$$Df = \sum_{i=1}^{\infty} \frac{\partial f}{\partial y_i} f_i \quad 2-3-4$$

έχουμε

$$f^{[r]}(x, y(x)) = y^{[r+1]}(x) = \left(\sum_{i=1}^{\infty} \frac{\partial f}{\partial y_i} f_i \right)^r \quad 2-3-5$$

όπου f_i οι συντεταγμένες της f και r η τάξη του διαφορικού. Έτσι θεωρώντας για παράδειγμα ότι η f είναι m -διάστατη θα έχουμε τις σχέσεις

$$f'(x, y(x)) = Df = \sum_{i=1}^m \frac{\partial f}{\partial y_i} f_i \quad 2-3-6$$

$$f''(x, y(x)) = D^2f = \sum_{i=1}^m \sum_{j=1}^m \frac{\partial^2 f}{\partial y_i \partial y_j} f_i f_j + \sum_{i=1}^m \sum_{j=1}^m \frac{\partial f}{\partial y_i} \frac{\partial f_i}{\partial y_j} f_j$$

από τις οποίες καθορίζεται πλήρως η 2-2-1. Η συνάρτηση Φ_{RK} για να καθορισθεί πρέπει να αναπτυχθούν τα m -διάστατα διανύσματα f^1, f^2, \dots, f^s σε ανάπτυγμα Taylor γύρω από το (x_n, y_n) που για λόγους ευκολίας θα σημειώνω (x, y) . Πράγματι έχουμε από το ανάπτυγμα Taylor για m -μεταβλητές, όσες δηλαδή και οι συντεταγμένες της f ,

$$f^1 = f = f(x, y)$$

$$f^2 = f(y + a_{21} h f^1) = f + a_{21} h \sum_{i=1}^m \frac{\partial f}{\partial y_i} f_i^1 + \frac{1}{2} h^2 a_{21} \sum_{j=1}^m \sum_{j=1}^m f_i^1 f_j^1$$

$$f^3 = f(y + h(a_{31}f^1 + a_{32}f^2)) = f + h \sum_{i=1}^m \frac{\partial f}{\partial y_i} (a_{31}f_i^1 + a_{32}f_i^2) + \quad 2-3-7$$

$$+ \frac{1}{2}h^2 a_{21} \sum_{i=1}^m \sum_{j=1}^m \frac{\partial^2 f}{\partial y_i \partial y_j} (a_{31}f_i^1 + a_{32}f_i^2) (a_{31}f_j^1 + a_{32}f_j^2)$$

$$f^4 = \dots$$

:

Κ.Ο.Κ

Τελικά από 2-2-1 και 2-3-5 έχω

$$h\Phi(x_n, y_n, h) = hf + \frac{1}{2}h^2 \sum_{i=1}^{\infty} \frac{\partial f}{\partial y_i} f_i +$$

2-3-8

$$\frac{1}{6}h^3 \left(\sum_{i=1}^m \sum_{j=1}^m \frac{\partial^2 f}{\partial y_i \partial y_j} f_i f_j + \sum_{i=1}^m \sum_{j=1}^m \frac{\partial f}{\partial y_i} \frac{\partial f}{\partial y_j} f_j \right) + O(h^4)$$

ενώ από την 2-3-7

$$h\Phi_{RK}(x_n, y_n, h) = hb_1 f + hb_2 (f + a_{21} h \sum_{i=1}^m \frac{\partial f}{\partial y_i} f_i^1 + \frac{1}{2}h^2 a_{21} \sum_{j=1}^m \sum_{j=1}^m f_i^1 f_j^1) +$$

$$+ b_3 h (f + h \sum_{i=1}^m \frac{\partial f}{\partial y_i} (a_{31}f_i^1 + a_{32}f_i^2)) + \quad 2-3-9$$

$$+ b_3 h \left(\frac{1}{2}h^2 a_{21} \sum_{i=1}^m \sum_{j=1}^m \frac{\partial^2 f}{\partial y_i \partial y_j} (a_{31}f_i^1 + a_{32}f_i^2) (a_{31}f_j^1 + a_{32}f_j^2) \right) +$$

Οι σχέσεις 2-3-8 και 2-3-9 σε συνεπτυγμένη μορφή μπορεί να γραφτούν

$$y(x_{n+1}) = y_n + \sum_{k \geq 1} \frac{h^k}{k!} \left(\sum_{i=1}^{i_k} a_{ki} P_{ki} \right)$$

$$y_{n+1} = y_n + \sum_{k \geq 1} h^k \sum_{i=1}^{i_k} \beta_{ki} P_{ki}$$

αντίστοιχα, όπου P_{ki} είναι υπό κατάλληλη μορφή τα στοιχειώδη διαφορικά που εμφανίζονται στις 2-3-8 και 2-3-9, a_{ki} είναι θετικοί αριθμοί και τα β_{ki} είναι αλγεβρικές εξισώσεις των παραμέτρων, ενώ i_k παριστάνει το πλήθος των εξισώσεων κ-τάξης όπως προκύπτει από την παραπάνω διαδικασία. Το τοπικό σφάλμα αποκοπής Τ.Σ.Α ορίζεται να είναι

$$T.S.A = y_{n+1} - y(x_n)$$

ή

$$T.S.A = h \sum_{i=1}^{i_1} T_{1i} P_{1i} + h^2 \sum_{i=1}^{i_2} T_{2i} P_{2i} + h^3 \sum_{i=2}^{i_3} T_{3i} P_{3i} + \dots \quad 2-3-10$$

$$\text{όπου} \quad T_{ki} = \beta_{ki} - a_{ki}/k! .$$

Η μέθοδος λέμε ότι έχει τάξη p αν ισχύουν οι σχέσεις

$$T_{ki} = 0 \quad \text{για} \quad k=1,2,\dots,p \quad \text{και} \quad i=1,2,\dots,i_k$$

Τελικά για μία μέθοδο p -τάξης έχουμε

$$T.S.A = h^{p+1} \cdot \psi(x_n, y_n) + O(h^{p+2}) \quad 2-3-11$$

όπου το $\psi(x_n, y_n)$ λέγεται πρωτεύων όρος του T.S.A.

2.4 Εξισώσεις συνθηκών Runge-Kutta μέχρι 6^{ης} τάξης

Στη συνέχεια δίνουμε τις εξισώσεις συνθηκών για τις παραμέτρους της μεθόδου Runge-Kutta μέχρι και 7^{ης} τάξης. Χρησιμοποιούμε τη διανυσματική μορφή 2-2-4b και θεωρούμε το $e = e_1 + e_2 + \dots + e_m$ με e_i η i -στη συντεταγμένη της ορθοκανονικής βάσης του χώρου \mathbb{R}^m . Επίσης θεωρούμε $B = b^T$ και τον πίνακα C που έχει διαγώνιο τα στοιχεία του διανύσματος c και τα υπόλοιπα στοιχεία μηδέν, δηλ $C = \text{diag}(c)$. Οι εξισώσεις εμφανίζονται σαν γινόμενα πινάκων και διανυσμάτων που περιέχουν τις άγνωστες παραμέτρους.

Είναι η πιο συμπαγής μορφή των εξισώσεων που απο κάποια τάξη και πάνω είναι δύσκολο να παρακολουθήσει με νηφαλιότητα κανείς. Παράσταση των εξισώσεων με θεωρία δέντρων έχει επιχειρηθεί επίσης από Butcher [1987], Hairer & Wanner [1976], αλλά δεν την συνιστούμε για κάποιον που απλώς θέλει τη λίστα με τις εξισώσεις.

Οι δύο πράξεις που εμφανίζονται στις εξισώσεις ορίζονται ως εξής

(i) πράξη (.) : $\mathbb{R}^{m \times L} \times \mathbb{R}^{L \times m} \rightarrow \mathbb{R}^{m \times m}$

που ορίζεται από τη σχέση $A \cdot B = C$ αν $A \in \mathbb{R}^{m \times L}, B \in \mathbb{R}^{L \times m}, C \in \mathbb{R}^{m \times m}$

$$\text{όπου } c_{ij} = \sum_{k=1}^L a_{ik} \cdot b_{kj}$$

αν οι πίνακες είναι οι $A = \{a_{ij}\}, B = \{b_{ij}\}, C = \{c_{ij}\}$

(ii) πράξη (*) : $\mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m$

που ορίζεται από τη σχέση $a * b = c$ αν $a, b, c \in \mathbb{R}^m$

όπου $c_i = a_i \cdot b_i \quad i=1,2,\dots,m$ αν τα διανύσματα είναι τα $a = \{a_i\}, b = \{b_i\}, c = \{c_i\}$.

Δηλαδή η πρώτη πράξη είναι το γνωστό γινόμενο μεταξύ πινάκων που στον περιορισμό του σε διανύσματα ($L=1$) είναι το συνηθισμένο εσωτερικό γινόμενο. Η δε δεύτερη πράξη που ορίζεται μεταξύ διανυσμάτων δίνει σαν αποτέλεσμα ένα διάνυσμα με ίδια διάσταση που προκύπτει από πολλαπλασιασμό των αντίστοιχων συντεταγμένων. Οι εξισώσεις σε μορφή πινάκων είναι.

α/α τάξη α/α/τάξης εξίσωση

$$1 \quad 1 \quad 1 \quad B \cdot e = 1$$

$$2 \quad 2 \quad 1 \quad B \cdot C \cdot e = \frac{1}{2}$$

$$3 \quad 3 \quad 1 \quad \frac{1}{2} \cdot B \cdot C \cdot C \cdot e = \frac{1}{6}$$

$$4 \quad 3 \quad 2 \quad B \cdot A \cdot C \cdot e = \frac{1}{6}$$

5	4	1	$\frac{1}{6} \cdot \text{B.C.C.C.e} = \frac{1}{24}$
6	4	2	$\frac{1}{2} \cdot \text{B.A.C.C.e} = \frac{1}{24}$
7	4	3	$\text{B.C.A.C.e} = \frac{1}{8}$
8	4	4	$\text{B.A.A.C.e} = \frac{1}{24}$
9	5	1	$\frac{1}{24} \cdot \text{B.C.C.C.C.e} = \frac{1}{120}$
10	5	2	$\frac{1}{2} \cdot \text{B.C.C.A.C.e} = \frac{1}{20}$
11	5	3	$\frac{1}{2} \cdot \text{B} \cdot (\text{A.C.e}) \cdot (\text{A.C.e}) = \frac{1}{40}$
12	5	4	$\frac{1}{2} \cdot \text{B.C.A.C.C.e} = \frac{1}{30}$
13	5	5	$\frac{1}{6} \cdot \text{B.A.C.C.C.e} = \frac{1}{120}$
14	5	6	$\text{B.C.A.A.C.e} = \frac{1}{30}$
15	5	7	$\text{B.A.C.A.C.e} = \frac{1}{40}$
16	5	8	$\frac{1}{2} \cdot \text{B.A.A.C.C.e} = \frac{1}{120}$
17	5	9	$\text{B.A.A.A.C.e} = \frac{1}{120}$
18	6	1	$\frac{1}{120} \cdot \text{B.C.C.C.C.C.e} = \frac{1}{720}$
19	6	2	$\frac{1}{6} \cdot \text{B.C.C.C.A.C.e} = \frac{1}{72}$
20	6	3	$\frac{1}{2} \cdot \text{B.C.} \cdot ((\text{A.C.e}) \cdot (\text{A.C.e})) = \frac{1}{48}$
21	6	4	$\frac{1}{4} \cdot \text{B.C.C.A.C.C.e} = \frac{1}{72}$

22	6	5	$\frac{1}{2} \cdot B.((A.C.C.e) \cdot (A.C.e)) = \frac{1}{72}$
23	6	6	$\frac{1}{6} \cdot B.C.A.C.C.C.e = \frac{1}{144}$
24	6	7	$\frac{1}{24} \cdot B.A.C.C.C.C.e = \frac{1}{720}$
25	6	8	$\frac{1}{2} \cdot B.C.C.A.A.C.e = \frac{1}{72}$
26	6	9	$B.((A.A.C.e) \cdot (A.C.e)) = \frac{1}{72}$
27	6	10	$B.C.A.C.A.C.e = \frac{1}{48}$
28	6	11	$\frac{1}{2} \cdot B.A.C.C.A.C.e = \frac{1}{120}$
29	6	12	$\frac{1}{2} \cdot B.A.((A.C.e) \cdot (A.C.e)) = \frac{1}{240}$
30	6	13	$\frac{1}{2} \cdot B.C.A.A.C.C.e = \frac{1}{144}$
31	6	14	$\frac{1}{2} \cdot B.A.C.A.C.C.e = \frac{1}{180}$
32	6	15	$\frac{1}{6} \cdot B.A.A.C.C.C.e = \frac{1}{720}$
33	6	16	$B.C.A.A.A.C.e = \frac{1}{144}$
34	6	17	$B.A.C.A.A.C.e = \frac{1}{180}$
35	6	18	$B.A.A.C.A.C.e = \frac{1}{240}$
36	6	19	$\frac{1}{2} \cdot B.A.A.A.C.C.e = \frac{1}{720}$
37	6	20	$B.A.A.A.A.C.e = \frac{1}{720}$

Σημειώνουμε ότι δεν κάνουμε τις αριθμητικές απλοποιήσεις διότι οι εξισώσεις χρειάζονται σε αυτή τη μορφή αν θέλουμε να εκτιμήσουμε σωστά το τοπικό σφάλμα αποκοπής. Πάντως στον πίνακα που ακολουθεί δίνουμε τη σχέση τάξης - αριθμού εξισώσεων Shampine & Watts [1976]

τάξη k	1	2	3	4	5	6	7	8	9	10
$O(h^k)$, i_k εξισώσεις	1	1	2	4	9	20	48	115	286	719
σύνολο εξισώσεων	1	2	4	8	17	37	85	200	486	1205

Πίνακας 2.4a

Θα πρέπει να σημειώσουμε ότι ο αριθμός των βαθμίδων και η τάξη είναι διαφορετικοί αριθμοί που είναι αδύνατο να συμπίπτουν αν υπερβαίνουν το τέσσερα. Αλλωστε από τον ορισμό βλέπουμε ότι η μέθοδος έχει $s(s+1)/2$ παραμέτρους όταν έχει s -βαθμίδες. Στον επόμενο πίνακα δίνουμε τη σχέση βαθμίδων-τάξης.

αριθμός βαθμίδων	1	2	3	4	5	6	7	8	9	10
αριθμός παραμέτρων	1	3	6	10	15	21	28	36	45	55
ανώτερη τάξη	1	2	3	4	4	5	6	6	7	7

Πίνακας 2-4b

Δεν πρέπει να ξεκινάει ο μικρός αριθμός παραμέτρων σε σχέση με τις εξισώσεις διότι με απλοποιητικές παραδοχές μπορούν να λιγοστέψουν και το σύστημα να επιλύεται. 0

Butcher [1964], απέδειξε ότι η ανώτερη ταξη που μπορεί να επιτευχθεί σε μία s -βαθμίδων μέθοδο είναι $s-2$. Παντως μέχρι σήμερα έχουν κατασκευασθεί μεθοδοι ογδός τάξης με 11-βαθμίδες Curtis [1969] και δεκάτης τάξης απο τον Hairer [1979] με 17-βαθμίδες που είναι και η ανώτερη τάξης μέθοδος που έχει κατασκευαστεί μέχρι στιγμής.

2.5 Embedded Runge-Kutta

Η γενική μορφή μιας embedded Runge-Kutta μεθόδου s -βαθμίδων μπορεί να εκφραστεί με τον παρακάτω πίνακα συντελεστών

$$\begin{array}{c|cccc}
 c_1 & & & & \\
 c_2 & a_{21} & & & \\
 c_3 & a_{31} & a_{32} & & \\
 \vdots & \vdots & \vdots & \ddots & \vdots \\
 c_s & a_{s1} & a_{s2} & \dots & a_{s,s-1} \\
 \hline
 & b_1 & b_2 & \dots & b_{s-1} & b_s \\
 \hline
 & \bar{b}_1 & \bar{b}_2 & \dots & \bar{b}_{s-1} & \bar{b}_s
 \end{array} \tag{2-5-1}$$

ή σε διανυσματική μορφή ,

$$\begin{array}{c|c}
 c & A \\
 \hline
 & b^T \\
 & \bar{b}^T
 \end{array} \tag{2-5-2}$$

με $A \in \mathbb{R}^{s \times s}$, $\bar{b}, b, c \in \mathbb{R}^s$ και $c = A \cdot e$. Τα διανύσματα b και \bar{b}

ορίζουν δύο προσεγγίσεις της $y(x_{n+1})$ τάξης p και $p+1$ αντίστοιχα με το ίδιο σύνολο υπολογισμών της συνάρτησης f . Embedded Runge-Kutta μεθόδους έχουν προτείνει διάφοροι ερευνητές. Μεταξύ αυτών οι πιο γνωστές είναι αυτές που προτάθηκαν από τους Fehlberg [1969], Verner [1978] και τους Dormand & Prince [1980].

Ετσι έχουμε μια εκτίμηση του Τ.Σ.Α της μεθόδου χαμηλής τάξης, ή τουλάχιστον του κυρίαρχου όρου του Τ.Σ.Α. αυτής.

Ας θεωρήσουμε τώρα το τοπικό πρόβλημα αρχικών τιμών μετα από n βήματα.

$$u' = f(x, u)$$

2-5-3

$$u(x_n) = y_n$$

Αν y_{n+1} , \bar{y}_{n+1} οι δύο προσεγγίσεις τάξης p και $p+1$ αντίστοιχα του τοπικού προβλήματος (2-5-3) θα έχουμε

$$y_{n+1} - \bar{y}_{n+1} = y_{n+1} - u(x_{n+1}) - (\bar{y}_{n+1} - u(x_{n+1})) =$$

$$= h^{p+1} \sum_{i=1}^{p+1} T_{p+1,i} P_{p+1,i} + h^{p+2} \sum_{i=1}^{p+1} (T_{p+1,i} - \bar{T}_{p+1,i}) P_{p+1,i} + O(h^{p+3}) =$$

$$= h^{p+1} \cdot \Psi(x_n, y_n) + O(h^{p+2}) = \text{Τ.Σ.Α} + O(h^{p+2})$$

Ενας σύγχρονος αλγόριθμος εκμεταλλεύεται την εκτίμηση του σφάλματος αποκοπής της μεθόδου για αυτόματη αλλαγή του βήματος ολοκλήρωσης. Τούτο προφανώς έχει σαν αποτέλεσμα την βελτιστοποίηση του αλγόριθμου όσον αφορά το συνολικό υπολογιστικό κόστος.

Ο αλγόριθμος απαιτεί σε κάθε βήμα να ισχύει η σχέση

$$\| y_{n+1} - \bar{y}_{n+1} \| = \| \text{EST} \| < \text{TOL} \quad 2-5-4$$

όπου TOL εκφράζει την ακρίβεια των υπολογισμών που απαιτούμε. Αν η σχέση αυτή ισχύει, το βήμα θεωρείται επιτυχές, και το επόμενο βήμα υπολογίζεται από τον τύπο.

$$h_{n+1} = \text{fact } h_n \left(\frac{\text{TOL}}{\text{EST}} \right)^{1/(p+1)} \quad 2-5-5$$

όπου h_n το προηγούμενο βήμα, fact ένας συντελεστής ασφαλείας με $\text{fact} \leq 1$ και p η τάξη της μεθόδου. Αν η σχέση αυτή δεν ισχύει το βήμα θεωρείται ανεπιτυχές και απορρίπτεται. Τότε επαναλαμβάνεται ο υπολογισμός της λύσης με τις ίδιες αρχικές συνθήκες αλλά ελαττωμένο το μήκος του βήματος σύμφωνα με τον τύπο 2-5-5 αλλά θεωρώντας ότι το h_{n+1} είναι το νέο μήκος του n -βήματος.

Ο τύπος 2-5-5 προκύπτει πρακτικά θεωρώντας ότι το εκτιμώμενο σφάλμα EST σε μια μέθοδο p τάξης είναι h_n^{p+1} . Αν απαιτήσουμε το επόμενο βήμα να έχει σφάλμα TOL,

ΕΚΤΙΜΩΝΤΑΣ όΤΙ αυτο θα είναι περίπου h_{n+1}^{p+1} , τότε ισχύουν οι σχέσεις

$$EST \approx h_n^{p+1} \quad TOL \approx h_{n+1}^{p+1} \quad 2-5-6$$

Διαιρώντας τις σχέσεις αυτές κατά μέλη και λαμβάνοντας υπ' όψη τον συντελεστή ασφαλείας fact προκύπτει η 2-5-5.

Η ποιότητα τώρα μιάς μεθόδου *Runge-Kutta* τάξης p εξαρτάται απο το πόσο μικρά είναι τα $T_{p+1,i}$, $i=1,2,\dots,i_{p+1}$. Ετσι αφού αποφασίσει κανείς την κατασκευή μιάς τέτοιας μεθόδου και μηδενίσει τα $T_{j,i}$, $j=1,2,\dots,p$, $i=1,2,\dots,i_j$ θα πρέπει να ελαχιστοποιήσει τα αντίστοιχα $T_{p+1,j}$, $j=1,2,\dots,(p+1)_k$ για να παράγει μέθοδο καλύτερη. Διάφορες νόρμες έχουν προταθεί για να μετρήσουμε πόσο καλή ή κακή είναι μια μέθοδος. Σε όλες το i_{p+1} -διάστατο διάνυσμα T_{p+1} που αποτελείται απο τα $T_{p+1,j}$, $j=1,2,\dots,(p+1)_k$ αντιπροσωπεύεται απο έναν θετικό αριθμό

(i) Ευκλειδειος νόρμα:

$$\| T_{p+1} \|_2 = \sqrt{\sum_{i=1}^{i_{p+1}} (T_{p+1,i})^2}$$

(ii) Την νόρμα μέγιστης απόλυτης τιμής

$$\| T_{p+1} \|_{\infty} = \max_{i=1,\dots,i_{p+1}} | T_{p+1,i} |$$

(iii) Τη νόρμα που χρησιμοποιούν οι Enright et. al. [1986]

$$\| T_{p+1} \|_E = \max_{i=1, \dots, i_{p+1}} \left\{ | T_{p+1, i} | \cdot \frac{(p+1)!}{a_{p+1, i}} \right\}$$

με $a_{p+1, i}$ όπως ορίσθηκε στις 2-3-10

Βέβαια στην περίπτωση κατασκευής embedded μεθόδων δεν είναι δυνατό να ακολουθηθούν τα παραπάνω γιά τη χαμηλής τάξης μέθοδο διότι τότε η ελαχιστοποίηση θα μας οδηγήσει σε μηδενισμό και τελικά θα έχουμε δύο τύπους ίδιας τάξης και λάθος εκτίμηση. Σ' αυτή την περίπτωση υπάρχουν τα κριτήρια των Dormand-Prince [1980].

Τέλος αναφέρουμε ότι οι σύγχρονοι αλγόριθμοι χρησιμοποιούν την τεχνική της τοπικής παρεκβολής στην διαδικασία υπολογισμού της λύσης που σημαίνει ότι προχωράμε με την προσέγγιση υψηλής τάξης \bar{y}_{n+1} παρ' όλο που η εκτίμηση του Τ.Σ.Α αφορά τη χαμηλής τάξης προσέγγιση.

2.6 Μέθοδοι Runge-Kutta-Nystrom - Γενική θεωρία.

Ο Nystrom [1925] ανέπτυξε τις ομώνυμες μεθόδους γιά προβλήματα β' τάξης

$$y'' = f(x, y, y')$$

2-6-1

$$y(x_0) = y_0, \quad y'(x_0) = y'_0$$

Αν παραλείψουμε πάλι όπως πριν το x , έχουμε:

$$\begin{bmatrix} y \\ y' \end{bmatrix}' = \begin{bmatrix} y' \\ f(x, y, y') \end{bmatrix} \quad 2-6-2$$

Αν εφαρμόσουμε τώρα την μέθοδο *Runge-Kutta* για το πρόβλημα 2-6-2 στο σημείο (x_n, y_n, y'_n) θα έχουμε

$$f^i = y'_n + h \sum_j a_{ij} (f^j)' \quad 2-6-3$$

$$(f^i)' = f(x_n + c_i h, y_n + h \sum_j a_{ij} f^j, y'_n + h \sum_j a_{ij} (f^j)')$$

$$y_{n+1} = y_n + h \sum_l b_l f^l, \quad y'_{n+1} = y'_n + h \sum_l b_l (f^l)'$$

όπου οι εκθέτες δείχνουν την βαθμίδα και όταν εμφανισθούν οι δείκτες θα υπονοούν την συντεταγμένη, αφού το y και τα f, f' είναι m -διάστατα διανύσματα.

Εισάγοντας την πρώτη σχέση μέσα στις άλλες τότε παίρνουμε:

$$(f^i)' = f(x_n + c_i h, y_n + c_i h y'_n + h^2 \sum_j \bar{a}_{ij} (f^j)', y'_n + h \sum_j a_{ij} (f^j)')$$

$$y_{n+1} = y_n + h y'_n + h^2 \sum_l \bar{b}_l (f^l)' \quad , \quad y'_{n+1} = y'_n + h \sum_l b_l (f^l)' \quad 2-6-4$$

όπου

$$\bar{a}_{ij} = \sum_k a_{ik} a_{kj}, \quad \bar{b}_i = \sum_j b_j a_{ji} \quad 2-6-5$$

Οι σχέσεις 2-6-4 και 2-6-5 ορίζουν την γενική μορφή μιας μεθόδου Runge-Kutta-Nystrom s-βαθμίδων για το γενικό πρόβλημα 2-6-1.

Ο Nystrom [1925] κατασκεύασε τις μεθόδους του χωρίς αναγκαστικά να ισχύουν οι σχέσεις 2-6-4 & 2-6-5. Ακολουθώντας την φιλοσοφία των μεθόδων *Runge-Kutta* εκφράζουμε την συνάρτηση $\Phi(x,y,y')$ για το πρόβλημα β' τάξης σε ανάπτυγμα Taylor. Επίσης τα αντίστοιχα f εκφράζονται σε ανάπτυγμα Taylor και προσπαθούμε να συγκρίνουμε τις δύο σειρές που προκύπτουν. Ο Nystrom ισχυριζόταν ότι είχε υπολογιστικό κόστος ελαττωμένο κατά 25%. Αυτό όμως δεν είναι αλήθεια αν ακολουθηθεί η τεχνική που βλέπουμε παραπάνω.

Για το ειδικό πρόβλημα αρχικών τιμών 2^{ης} τάξης

$$y''=f(x,y), \quad y(x_0)=y_0, \quad y'(x_0)=y'_0 \quad 2-6-6$$

η έλλειψη της y' από την f παίζει σημαντικό ρόλο. Γι' αυτό πρέπει να αντιμετωπισθεί κατασκευάζοντας μεθόδους μορφής *Runge-Kutta* εξ αρχής, διότι η έλλειψη του y' από το f , παίζει σημαντικό ρόλο. Με τον τρόπο αυτό επιτυγχάνεται σημαντική ελάττωση του υπολογιστικού κόστους. Η 2-6-4 τώρα παίρνει τη μορφή

$$g^i = f(x_n + c_i h, y_n + c_i h y'_n + h^2 \sum_{j>0}^{i-1} \bar{a}_{ij} g^j) \quad 2-6-7a$$

$$y_{n+1} = y_n + h y'_n + h^2 \sum_{i>0}^s \bar{b}_i g^i \quad 2-6-7b$$

$$y'_{n+1} = y'_n + h \sum_{i>0}^s b_i g^i \quad 2-6-7c$$

Οι σχέσεις 2-6-7 ορίζουν την γενική μορφή της μεθόδου *Runge - Kutta - Nystrom* για προβλήματα του τύπου 2-6-6. Σε διανυσματική μορφή η μέθοδος μπορεί να εκφραστεί

$$c \begin{vmatrix} A \\ \bar{b}^T \\ b^T \end{vmatrix}, (\bar{a}_{ij}) = A \in \mathbb{R}^{s \times s}, \bar{b}, b, c \in \mathbb{R}^s$$

Περιορίζουμε τους εαυτούς μας σε συστήματα αυτόνομα, προσθέτοντας εν ανάγκη την $x''=0$. Η j -συντεταγμένη του συστήματος των διαφορικών εξισώσεων $y''=f(y)$ είναι :

$$(y_j)'' = (y_j)^{[2]} = f_j(y_1, \dots, y_m) = f_j(y) \quad 2-6-8a$$

με τους δείκτες να υποδεικνύουν την αντίστοιχη συντεταγμένη στο m -διάστατο πρόβλημα.

Οι παράγωγοι ανώτερης τάξης οι οποίες εμφανίζονται στο ανάπτυγμα Taylor δίνονται από τις σχέσεις:

$$(y_j)^{[3]} = \sum_k \frac{\partial f_j}{\partial y_k}(y) \cdot y'_k \quad 2-6-8b$$

$$(y_j)^{[4]} = \sum_{k,L} \frac{\partial^2 f_j}{\partial y_k \partial y_L}(y) \cdot y'_k \cdot y'_L + \sum_k \frac{\partial f_j}{\partial y_k}(y) \cdot f_k(y) \quad 2-6-8c$$

⋮

Αναπτύσσοντας τώρα και τα $g^i, i=1,2,\dots,s$ της μεθόδου RKN s -βαθμίδων γύρω από το y και λαμβάνοντας υπ' όψη ότι

$$\sum_j a_{ij} = \frac{1}{2} c_i^2 \quad 2-6-9a$$

παίρνουμε

$$g_j^1 = f_j(y) \quad 2-6-9b$$

$$g_j^2 = f_j(y + c_2 h y' + h^2 \bar{a}_{21} g^1) = f_j + c_2 h \sum_i \frac{\partial f_j}{\partial y_i} + \quad 2-6-9c$$

$$+ \frac{1}{2} h^2 c_2^2 \left(\sum_i \frac{\partial^2 f_j}{\partial y_i^2} + 2 \sum_i \sum_k \frac{\partial^2 f_j}{\partial y_i \partial y_k} \right) + O(h^3)$$

⋮

Βάσει των προηγουμένων τα αναπτύγματα Taylor τάξης p για τη j -συντεταγμένη των y & y' γράφονται :

$$y_j(x+h) = y_j(x) + \sum_{r=1}^p \frac{h^r}{r!} y_j^{[r]}(x) \quad 2-6-10a$$

$$y'_j(x+h) = \sum_{r=1}^p \frac{h^r}{r!} y_j^{[r]}(x) \quad 2-6-10b$$

Επίσης οι τιμές της j -συντεταγμένης των y & y' όπως δίνονται από τη μέθοδο Runge-Kutta-Nystrom γράφονται:

$$y_j(x+h) = y_j(x) + hy'_j(x) + h^2 \sum_{r=1}^s \bar{b}_r g_j^r \quad 2-6-11a$$

$$y'_j(x+h) = y'_j(x) + h \sum_{r=1}^s b_r g_j^r \quad 2-6-11b$$

Οι τελευταίες σχέσεις μπορούν να εκφραστούν σε δυνάμεις του h με τη βοήθεια των 2-6-9. Οι εξισώσεις των συντελεστών της μεθόδου Runge-Kutta-Nystrom τάξης p για το ειδικό πρόβλημα 2-6-6 προκύπτουν από τη σύγκριση των αναπτυγμάτων 2-6-10a και 2-6-11a καθώς και των 2-6-10b και 2-6-11b μέχρι και την p -στη δύναμη του h .

Ακολουθώντας τη μεθοδολογία που αναπτύξαμε για τις απλές Runge-Kutta και συγκεκριμένα ορίζοντας δύο σφάλματα αποκοπής, ένα για την y και ένα για την y' , βρίσκουμε ότι οι εξισώσεις συνθηκών έχουν την παρακάτω μορφή.

$$T_{ki} = B^T \cdot \beta_{ki} - a_{ki} / k! = 0 \quad 2-6-12$$

$$T'_{k-1,i} = (b')^T \cdot \beta_{ki} - a_{ki} / (k-1)! = 0 \quad 2-6-13$$

Η μέθοδος είναι τάξης p αν οι 2-6-12 ισχύουν για $k=2,3,\dots,p$ και $i=1,2,\dots,i_k$, με $i_k = | \{ T_{k1}, T_{k2}, \dots, T_{ki_k} \} |$

Επίσης οι 2-6-13 πρέπει να ισχύουν για $k=2,3,\dots,p+1$ και $i=1,2,\dots,i'_k$. Ομως μπορεί να παρατηρήσει κανείς ότι τα β_{ki} και τα α_{ki} έχουν συμμετοχή και στις δύο σχέσεις άρα $i_k = i'_{k-1}$

Στη συνέχεια δίνουμε τις εξισώσεις συνθηκών για τους συντελεστές της προσέγγισης από τους οποίους υπολογίζονται οι αντίστοιχες εξισώσεις για τους συντελεστές της προσέγγισης της παραγώγου.

α/α	τάξη	α/α τάξης	εξίσωση
1	2	1	$B \cdot e = 1/2$
2	3	1	$B \cdot C \cdot e = 1/6$
3	4	1	$\frac{1}{2} \cdot B \cdot C \cdot C \cdot e = 1/24$
4	5	1	$(1/6) \cdot B \cdot C \cdot C \cdot C \cdot e = 1/120$
5	5	2	$B \cdot A \cdot C \cdot e = 1/120$
6	6	1	$(1/24) \cdot B \cdot C \cdot C \cdot C \cdot C \cdot e = 1/720$
7	6	2	$(1/2) \cdot B \cdot A \cdot C \cdot C \cdot e = 1/720$
8	6	3	$B \cdot C \cdot A \cdot C \cdot e = 1/180$
9	7	1	$(1/120) \cdot B \cdot C \cdot C \cdot C \cdot C \cdot C \cdot e = 1/5040$
10	7	2	$(1/6) \cdot B \cdot A \cdot C \cdot C \cdot C \cdot e = 1/5040$

11	7	3	$(1/2) \cdot \bar{B}.C.A.C.C.e = 1/1008$
12	7	4	$(1/2) \cdot \bar{B}.C.C.A.C.e = 1/504$
13	7	5	$\bar{B}.A.A.C.e = 1/5040$
14	8	1	$(1/720) \cdot \bar{B}.C.C.C.C.C.C.e = 1/40320$
15	8	2	$(1/24) \cdot \bar{B}.A.C.C.C.C.e = 1/40320$
16	8	3	$(1/6) \cdot \bar{B}.C.A.C.C.C.e = 6/40320$
17	8	4	$(1/4) \cdot \bar{B}.C.C.A.C.C.e = 15/40320$
18	8	5	$(1/6) \cdot \bar{B}.C.C.C.A.C.e = 20/40320$
19	8	6	$(1/2) \cdot \bar{B}.(A.C.e) \cdot (A.C.e) = 10/40320$
20	8	7	$(1/2) \cdot \bar{B}.A.A.C.C.e = 1/40320$
21	8	8	$\bar{B}.A.C.A.C.e = 4/40320$
22	8	9	$\bar{B}.C.A.A.C.e = 6/40320$
23	9	1	$(1/5040) \cdot \bar{B}.C.C.C.C.C.C.C.e = 1/362880$
24	9	2	$(1/120) \cdot \bar{B}.A.C.C.C.C.C.e = 1/362880$
25	9	3	$(1/24) \cdot \bar{B}.C.A.C.C.C.C.e = 7/362880$
26	9	4	$(1/12) \cdot \bar{B}.C.C.A.C.C.C.e = 21/362880$
27	9	5	$(1/12) \cdot \bar{B}.C.C.C.A.C.C.e = 35/362880$

28	9	6	$(1/24) \cdot \bar{B}.C.C.C.C.A.C.e = 35/362880$
29	9	7	$(1/2) \cdot \bar{B}.(A.C.e) \cdot (A.C.C.e) = 35/362880$
30	9	8	$(1/2) \cdot \bar{B}.C.(A.C.e) \cdot (A.C.e) = 70/362880$
31	9	9	$(1/6) \cdot \bar{B}.A.A.C.C.C.e = 1/362880$
32	9	10	$(1/2) \cdot \bar{B}.A.C.A.C.C.e = 5/362880$
33	9	11	$(1/2) \cdot \bar{B}.A.C.C.A.C.e = 10/362880$
34	9	12	$(1/2) \cdot \bar{B}.C.A.A.C.C.e = 7/362880$
35	9	13	$\bar{B}.C.A.C.A.C.e = 28/362880$
36	9	14	$(1/2) \cdot \bar{B}.C.C.A.A.C.e = 21/362880$
37	9	15	$\bar{B}.A.A.A.C.e = 1/362880$

Σημειώνουμε ότι οι εξισώσεις που προκύπτουν είναι περισσότερες από αυτές που αναφέρουμε εδώ. Αποτελούν όμως απλή επανάληψη τους και έτσι τις αποφεύγουμε. Αν όμως σε τυχούσα μελλοντική ανάλυση σφάλματος χρησιμοποιήσει κανείς την Ευκλείδειο νόρμα θα πρέπει να τις λάβει υπ' όψιν. Εμείς στο υπόλοιπο κείμενο χρησιμοποιούμε την πρόταση των Enright et.al.[1986], οπότε δεν χρειάζονται οι εξισώσεις που επαναλαμβάνονται.

Η σύγκλιση των μεθόδων εξασφαλίζεται από τη μορφή 2-6-2 που

είναι ισοδύναμη με *Runge-Kutta* μορφή.

2.7 Απόλυτη ευστάθεια.

Με βάση τα αναφερόμενα στις 1-1-8 ως την 1-1-11 και εφαρμόζοντας μια *s*-βαθμίδων *Runge-Kutta* μέθοδο στο πρόβλημα 1-1-8 έχουμε :

$$f^1 = \lambda y_n$$

$$f^2 = \lambda(y_n + a_{21} f^1 h) \quad 2-7-1$$

$$f^3 = \lambda(y_n + h(a_{31} f^1 + a_{32} f^2))$$

⋮

Τελικά επειδή

$$y_{n+1} - y_n = h \cdot (b_1 f^1 + b_2 f^2 + \dots + b_s f^s) \Rightarrow$$

2-7-2

$$y_{n+1} = (1 + z \sum_1 b_1 + z^2 \sum_1 b_i \sum_j a_{ij} + \dots + z^s \sum_{j_1} b_{j_1} \sum_{j_1 j_2} a_{j_1 j_2} \sum_{j_2 j_3} a_{j_2 j_3} \dots \sum_{j_{s-1} j_s} a_{j_{s-1} j_s}) \cdot y_n$$

με $z = \lambda h$. Δηλαδή το $\mu(z)$ είναι ένα πολυώνυμο *s*-βαθμού, οι δε συντελεστές του z^k αν είναι τα $T_{k1_k} + \frac{1}{k!} = \tilde{T}_{k1_k}$, τότε η εξίσωση 2-7-2 μπορεί να γραφεί σε άλλη μορφή σαν

$$y_{n+1} = (1 + z \cdot \tilde{T}_{11} + z^2 \cdot \tilde{T}_{21} + z^3 \cdot \tilde{T}_{32} + z^4 \cdot \tilde{T}_{44} + z^5 \cdot \tilde{T}_{59} + \dots + z^s \cdot \tilde{T}_{s1_s}) \quad 2-7-3$$

Το μιγαδικό χωρίο για το οποίο ισχύει η σχέση $|\mu(z)| < 1$

είναι το χωρίο απόλυτης ευστάθειας. Πολλές φορές όμως αναφερόμαστε μόνο στο μέρος του πραγματικού άξονα για τον οποίο ισχύει η σχέση 1-2-10.

Καμιά άμεση μέθοδος Runge-Kutta δεν είναι A-ευσταθής. Ετσι αυτές δεν μπορούν να χρησιμοποιηθούν για άκαμπτα (stiff) προβλήματα. Κατάλληλες γι' αυτό είναι οι λεγόμενες έμμεσες Runge-Kutta μέθοδοι. Αυτές διαφέρουν από τις άμεσες μεθόδους Runge-Kutta διότι οι συντελεστές τους a_{ij} με $i \geq j$ είναι μη μηδενικοί. Ετσι όμως οι τιμές της f δεν μπορούν να υπολογιστούν άμεσα αλλά μόνο μετά απο την επίλυση ενός μη γραμμικού συστήματος εξισώσεων.

Εμείς εδώ δεν θα ασχοληθούμε με έμμεσες Runge-kutta μεθόδους. Ως εκ' τούτου πολλή μικρή σημασία έχει η απόλυτη ευστάθεια στη εξέλιξη της διατριβής.

Για το πρόβλημα αρχικών τιμών 2^{ης} τάξης ειδικής μορφής δεν υπάρχει σαφώς καθορισμένη έννοια απόλυτης ευστάθειας. Από τις εργασίες των Chawla [1981], Dormand et.al [1987] βλέπουμε ότι ορίζεται με δύο διαφορετικούς τρόπους η απόλυτη ευστάθεια για τη περίπτωση αυτή.

Κ Ε Φ Α Λ Α Ι Ο 3^{ΟΝ}

SCALED RUNGE-KUTTA ΜΕΘΟΔΟΙ

3.1 Γενικά περί scaling.

Η μεγάλη ακρίβεια και η ευκολία εφαρμογής των αλγόριθμων *Runge-Kutta* τους τοποθετούν ανάμεσα στους πιο δημοφιλείς για την επίλυση των συνήθων διαφορικών εξισώσεων. Κάθε αλγόριθμος όμως δεν αποτελεί και πανάκεια. Έτσι οι μέθοδοι *Runge-Kutta* δεν παράγουν προσεγγίσεις της λύσης σε ενδιάμεσα σημεία. Αν λοιπόν κανείς ήθελε κάτι τέτοιο θα έπρεπε να μικρύνει κατάλληλα το βήμα για να περνά από τα σημεία που χρειάζεται. Κάτι τέτοιο είναι υπέρμετρα δαπανηρό σε υπολογιστικό χρόνο και πρέπει να αποφεύγεται (Shampine et.al.[1975]). Έτσι για να αντιμετωπιστεί το πρόβλημα πρέπει να κατασκευασθούν *Runge-Kutta* μέθοδοι με παρεμβολικές ιδιότητες.

Πρώτη με το θέμα αυτό ασχολήθηκε η Horn [1983] που εισήγαγε την τεχνική του scaling για τη αντιμετώπιση του προβλήματος αυτού. Έπειτα οι Dormand & Prince [1986],[1987] χρησιμοποίησαν την ίδια τεχνική προσαρμοσμένη στις μεθόδους που αυτοί έχουν προτείνει, τακτική που ακολουθεί πιστά και ο Verner [1990α],[1990β],[1990c] για τις δικές του όμως μεθόδους. Επίσης οι Parageorgiou et.al. [1989] επέκτειναν τη

μέθοδο για περισσότερες μεθόδους και τέλος οι Calvo et.al [1990a],[1990β] εργαζονται σε αυτή την κατεύθυνση.

Βέβαια πολλοί ερευνητές ασχολούνται με το θέμα αλλά χρησιμοποιώντας και τεχνικές παρεμβολής όπως θα δούμε στο επόμενο κεφάλαιο.

Στο 3^ο κεφάλαιο θα αναπτύξουμε τη θεωρία και θα κατασκευάσουμε Scaled Runge-Kutta αλγόριθμους οι οποίοι θα προσεγγίζουν τη λύση μιας συνήθους διαφορικής εξίσωσης σε ενδιάμεσα σημεία. Συγκεκριμένα θα κατασκευαστούν scaled μέθοδοι 4^{ης} τάξης για προσεγγιστική λύση σε οποιοδήποτε ενδιάμεσα σημεία με μικρότερο υπολογιστικό κόστος και ελαχιστοποιημένα σφάλματα αποκοπής. Επίσης μέθοδοι scaled 5^{ης} τάξης θα προταθούν για a-priori γνωστά ενδιάμεσα σημεία με επίσης ελαττωμένο υπολογιστικό κόστος.

3.2 Scaling - Γενική θεωρία.

Στην παράγραφο αυτή θα αναπτύξουμε την γενική θεωρία της τεχνικής scaling για Runge-Kutta μεθόδους. Υποθέτουμε ότι η λύση του προβλήματος (0-1) έχει υπολογισθεί στο τέλος του βήματος $[x_n, x_{n+1}]$ εφαρμόζοντας την μέθοδο Runge-Kutta και είναι αποδεκτή. Εστω τώρα ότι ζητάμε την λύση στο ενδιάμεσο σημείο,

$$x^* = x_n + \sigma \cdot h_n \quad 3-2-1$$

όπου σ είναι ένας συντελεστής με $\sigma \in (0,1)$. Εχοντας στην



διάθεση μας τις τιμές f^i , $i=1,2,\dots,s$ από την εφαρμογή της Runge-Kutta μεθόδου στο διάστημα $[x_n, x_{n+1}]$ και ενδεχόμενα μερικές πρόσθετες τιμές f^i , $i=s+1, s+2, \dots, s^*$, μπορούμε να προσδιορίσουμε την scaled λύση του προβλήματος. Έτσι με δοσμένη μια μέθοδο Runge-Kutta είναι δυνατόν να κατασκευαστεί μια οικογένεια scaled αλγορίθμων με την επιθυμητή ιδιότητα.

Ανακαλούμε από το προηγούμενο κεφάλαιο τη μορφή των μεθόδων Runge-Kutta. Εστω ότι η λύση στο $x_{n+1} = x_n + h$ έχει γίνει αποδεκτή και θέλουμε λύση στο ενδιάμεσο σημείο 3-2-1, με

$$h^* = \sigma h \quad 3-2-2$$

και $\sigma \in (0,1)$

Η γενική μορφή μιας scaled Runge-Kutta μεθόδου είναι η ακόλουθη.

$$y(x_n + \sigma h) \cong y_{n+1}^* = y_n + h^* \sum_{j=0}^{s^*} b_j^* f_j^* \quad 3-2-3a$$

ή

$$y_{n+1}^* = y_n + h^* \cdot \Phi^*(x_n, y_n, \sigma h_n)$$

με

$$f^1 = f(x_n, y_n) \quad 3-2-3b$$

$$f^i = f(x_n + c_i^* h^*, y_n + h^* \sum_{j=1}^{i-1} a_{ij}^* f_j^*) \quad 3-2-4c$$

με $i=1,2,3,\dots,s^*$ και $s^* \geq s$. Αν τα a_{ij}^* και τα c_i^* επιλεγθούν έτσι ώστε:

$$f^{i^*} = f^i, i=1,2,\dots,s \quad 3-2-4$$

τότε το μεγαλύτερο κόστος στην εφαρμογή της μεθόδου θα αποφευχθεί αφού τα f^1, f^2, \dots, f^s έχουν ήδη υπολογισθεί κατά τη διάρκεια του υπολογισμού της προσέγγισης y_{n+1} . Έτσι μόνο τα f^{s+1}, \dots, f^{s^*} μένουν να υπολογισθούν.

Η στρατηγική που προτάθηκε μας αναγκάζει να συσχετίσουμε τα a^*, c^* με τα αντίστοιχά τους a και c . Τα εναπομένοντα c_i^* και a_{ij}^* $i=s+1, \dots, s^*$ χρησιμοποιούνται για να κατασκευάσουμε scaled μέθοδο ακρίβειας συγκρινόμενης με την ακρίβεια της δοσμένης μεθόδου. Συγκρίνοντας τις σχέσεις 3-2-3b και 3-2-3c με τις αντίστοιχες σχέσεις της αρχικής μεθόδου βρίσκουμε :

$$c_i^* = \frac{c_i h}{h^*} = \frac{c_i}{\sigma} \quad 3-2-5$$

και

$$a_{ij}^* = \frac{a_{ij} h}{h^*} = \frac{a_{ij}}{\sigma} \quad 3-2-6$$

γιά $i=2,3,\dots,s$ και $j=1,2,\dots,i-1$ οπότε $f^{i^*} = f^i$ για $i=1,2,\dots,s$. Έτσι όταν $i \leq s$ οι παράμετροι a^*, c^* έχουν ήδη καθορισθεί, σχετιζόμενοι μάλιστα με τις αντίστοιχές τους του κυρίως αλγορίθμου με τον συντελεστή $1/\sigma$.

Οι παράμετροι των επιπρόσθετων υπολογισμών της f , είναι τέτοιες που να ικανοποιούν τις εξισώσεις συνθηκών μεθόδων *Runge-Kutta*, επιθυμητής τάξης. Οι εξισώσεις όμως αυτές πρέπει να λαμβάνουν υπ' όψιν το σ . Γι' αυτό πρέπει να κατασκευασθούν εξ' αρχής όπως στο προηγούμενο κεφάλαιο με βήμα $h^* = \sigma h$. Η μορφή τους όμως δεν διαφέρει ιδιαίτερα από τις ήδη υπάρχουσες εξισώσεις. Αν την 3-2-3α την μετασχηματίσουμε σε

$$y_{n+1}^* = y_n + h \sum_{j=0}^s b_j^* f_j^* \quad 3-2-7$$

τότε η k -στη εξίσωση τάξης i

$$T_{ik} = \beta_{ik} - a_{ik} / i! = 0 \quad 3-2-8$$

γίνεται

$$T_{ik}^* = \beta_{ik} - \sigma^i \cdot a_{ik} / i! = 0 \quad 3-2-9$$

Οι μέθοδοι που θα κατασκευασθούν μπορεί να έχουν δύο στόχους. Ο πρώτος στόχος είναι να κατασκευαστεί μια γενική μέθοδος που θα δίνει με s^* -s επι πλέον υπολογισμούς της f , λύση σε οποιοδήποτε σημείο του διαστήματος $[x_n, x_{n+1}]$, χωρίς να χρειάζεται επαναπροσδιορισμός των f για άλλο σημείο. Προφανώς για να πετύχουμε αυτό πρέπει τα a_{ij}, c_i όταν $i \geq s+1$ να είναι ανεξάρτητα του σ . Όπως φαίνεται από την 3-2-3c

απαλοΐφοντας το h^* έχουμε

$$f^i = f^{i*} = f(x_0 + c_i h, y_0 + h \sum a_{ij} \cdot f^j) . \quad 3-2-10$$

Δηλαδή, τα επί πλέον f , υπολογΐζονται άπαξ διά παντός και μετά, ανάλογα το σ , οι παράμετροι b_i^* καθορίζουν τις διαφορετικές τιμές της προσέγγισης. Η δυσκολία εδώ είναι εμφανής διότι στην ουσία τα b_i είναι οι ελεύθερες παράμετροι που βοηθούν στην λύση των εξισώσεων μια και αυτά μόνο είναι συναρτήσεις του σ , ενώ οι άλλες είναι απλοί αριθμοί. Στην περίπτωση όμως που χρειάζεται επαναληπτικός υπολογισμός ενδιάμεσων σημείων όπως π.χ. στον προσδιορισμό της περιόδου μιας τροχιάς το κόστος είναι σχετικά αμελητέο.

Αντίθετα, αν ο στόχος είναι κάποιο συγκεκριμένο σημείο *a-priori* καθορισμένο τότε στην λύση των εξισώσεων όλες οι παράμετροι μπορεί να επιλεχθούν σα συναρτήσεις (πολυώνυμα) του σ . Τότε δεν μας ενδιαφέρει αν τα επί πλέον f , είναι άχρηστα για τον υπολογισμό άλλου σημείου. Αυτή είναι επίσης ενδιαφέρουσα περίπτωση, διότι επιτρέπει με μικρότερο κόστος τον υπολογισμό σημείων που μπορεί να χρησιμεύσουν σαν σημεία στήριξης για την κατασκευή πολυωνύμων παρεμβολής.

Το scaling δεν έχει γενικά καλά αποτελέσματα σε υψηλής τάξης μεθόδους. Η πρώτη περίπτωση αποδίδει σε μεθόδους μέχρι και 4^{ης} τάξης. Η δεύτερη περίπτωση είναι αναγκαία παντού ανεξάρτητα αποτελέσματος, αλλά πάντως το κόστος της δεν

είναι απαγορευτικό.

Εδώ θα παρουσιασθούν μέθοδοι scaling 4^{ης} τάξης για συνεχή λύση (1^η περίπτωση) με ελάχιστο συντελεστή σφάλματος 5^{ης} τάξης χωρίς κόστος. Οι μέθοδοι αυτές κατασκευάστηκαν έτσι ώστε να ανήκουν στην κλάση C¹. Εγινε εφαρμογή στις Runge-Kutta μεθόδους των Fehlberg [1969] και Dormand & Prince [1980]. Σε αντίθεση, οι μέθοδοι της Horn δεν ήταν C¹, απαιτούσαν ένα επι πλέον υπολογισμό της f και τα τοπικά σφάλματα αποκοπής δεν είχαν βελτιστοποιηθεί.

Επίσης κατασκευάζουμε μεθόδους 5^{ης} τάξης για συγκεκριμένο σημείο με ένα μονο επί πλέον υπολογισμό της f, αντί για δύο υπολογισμούς που χρειαζόταν η μέθοδος της Horn [1984].

3.3 Scaling 4^{ης} τάξης.

Από τον πίνακα εξισώσεων που βρίσκονται στο προηγούμενο κεφάλαιο παρατήρηση 3-2-8 για να κατασκευάσουμε scaling μέθοδο 4^{ης} τάξης πρέπει να λύσουμε τις εξισώσεις

$$\text{I-1} \quad \sum b_i^* = \sigma$$

$$\text{II-1} \quad \sum b_i^* \cdot c_i = \frac{\sigma^2}{2}$$

$$\text{III-1} \quad \sum b_i^* \cdot c_i^2 = \frac{\sigma^3}{3}$$

$$\text{III-2} \quad \sum b_i^* \sum a_{ij} \cdot c_j = \frac{\sigma^3}{6}$$

$$\text{IV-1} \quad \sum b_i^* \cdot c_i^3 = \frac{\sigma^4}{4}$$

$$\text{IV-2} \quad \sum b_i^* c_i \sum a_{ij} c_j = \frac{\sigma^4}{8}$$

$$\text{IV-3} \quad \sum b_i^* \sum a_{ij} c_j^2 = \frac{\sigma^4}{12}$$

$$\text{IV-4} \quad \sum b_i^* \sum a_{ij} \sum a_{jk} c_k = \frac{\sigma^4}{24}$$

Σε όλα τα αθροίσματα οι δείκτες πηγαίνουν απο το 1 στο s^* . Εμείς θα ασχοληθούμε με embedded μεθόδους τάξεων 4 και 5 που χρησιμοποιούν 6 βαθμίδες. Σ' αυτές τις περιπτώσεις μας χρειάζεται ένα επί πλέον f για να επιτύχουμε την επίλυση των εξισώσεων I-1 έως IV-4. Η μέθοδος που θα προκύψει θα δίδει προσέγγιση για οποιοδήποτε σ .

Στην παρούσα περίπτωση έχουμε 8 εξισώσεις και $s^*=7$. Δηλαδή οι άγνωστοι είναι τα a_{7i} $i=2,3,\dots,6$ το c_7 και τα επτά b_i , $i=1,2,\dots,7$. Το a_{71} προκύπτει απο την 2-3-3 και δεν το θεωρούμε άγνωστο. Ομως στην ουσία μόνο τα b_i $i=1,2,\dots,7$ είναι αυτά που χρησιμεύουν στην επίλυση των εξισώσεων διότι μόνο αυτά μπορεί να παρθούν σαν πολυώνυμα του σ . Τα a_{7i} και το c_7 πρέπει να είναι σταθεροί αριθμοί έτσι ώστε να μπορεί να ξαναχρησιμοποιηθεί το f^7 που ορίζουν σε άλλη επιλογή του σ . Εν τέλει έχουμε μόνο επτά άγνωστους. Πάντα

όμως χρησιμοποιούμε απλοποιητικές παραδοχές που λιγοστεύουν τις εξισώσεις. Οι παραδοχές που κάνουμε είναι:

$$\sum b_i^* a_{i2} = 0 \quad 3-3-1$$

$$\sum_{j=1}^{i-1} a_{ij} c_j^k = \frac{c_i^{k+1}}{k+1}, \quad i=3,4,\dots,s^* \quad k=1,2 \quad 3-3-2$$

$$b_2^* = 0 \quad 3-3-3$$

Επίσης σαν f^0 παίρνουμε το f^1 του επόμενου βήματος το οποίο έτσι κι' αλλιώς δεν αποτελεί επί πλέον κόστος. Δηλαδή ζητάμε $a_{7i} = b_i$ $i=1,2,\dots,7$ και $c_7 = 1$. Οι 3-3-2 ισχύουν έτσι κι' αλλιώς πάντα διότι όλοι τις χρησιμοποιούν για να κατασκευάσουν τις μεθόδους τους. Σε συνδυασμό με την 3-3-1 απαλοΐφονται μερικές εξισώσεις και στο τέλος μένουν προς επίλυση οι εξισώσεις τυπου Vandermode I-1, II-1, III-1, IV-1 και η 3-3-1. Τώρα πια οι άγνωστοι είναι έξι, δηλ. οι $b_1^*, b_3^*, b_4^*, b_5^*, b_6^*, b_7^*$ ενώ οι εξισώσεις μόνο πέντε. Τις επιλύουμε λοιπόν με παράμετρος τα σ, b_7^* .

Εχοντας λοιπόν τις εκφράσεις που μας εξασφαλίζουν την επιθυμητή τάξη, ο επόμενος στόχος είναι η λύση να είναι C^1 δηλ. με συνεχείς παραγώγους. Απο 3-2-3a προκύπτει μια λύση στο δοθέν διάστημα που είναι στήν ουσία ένα πολυώνυμο του σ . Παραγωγίζοντας το πολυώνυμο αυτό θα έχουμε μια προσέγγιση της $y'(x_n + \sigma h)$ η οποία δίνεται από τη σχέση

$$y'_{n+1}^* = \sum \frac{db_i^*}{d\sigma} f^i = \sum (b_i')^* \cdot f^i \quad 3-3-4$$

Παρατηρούμε ότι τα b_i^* θα μπορούσαν να προκύψουν απ' ευθείας από τις (I-1 ως IV-4) αν παραγωγίζαμε τις εξισώσεις ως προς σ και θεωρούσαμε τα b^* πολυώνυμα του σ τις δε υπόλοιπες παραμέτρους απλούς αριθμούς. Είναι προφανές τώρα ότι αν τα a, c είναι ανεξάρτητα του σ , τότε η 3-3-4 είναι μια σχέση που μας ικανοποιεί και στην περίπτωση αυτή πράγματι τα f είναι απλοί αριθμοί. Γενικότερα όμως αν τα a, c εξαρτώνται από το σ , τότε δεν μπορούμε να χρησιμοποιήσουμε την 3-3-4 αλλά να λύσουμε τις (I-1 έως IV-4) εξ' αρχής για όλους τους αγνώστους (a, b, c). Βέβαια τότε $db_i^*(\sigma)/d\sigma \neq (b_i')^*$ Περιοριζόμενοι στην περίπτωση του scaling 4ης τάξης πρέπει λοιπόν θεωρώντας μόνο τα b σαν πολυώνυμα του σ να ισχύουν οι εξισώσεις:

$$b_i^*(0) = 0 \quad i=1,2,\dots,s^* \quad 3-3-5a$$

$$b_i^{*'}(0) = \begin{cases} 1 & \text{αν } i=1 \\ 0 & \text{αν } i \neq 1 \end{cases} \quad 3-3-5b$$

$$b_i^*(1) = b_i \quad i=1,2,\dots,s^* \quad 3-3-5c$$

$$b_i^{*'}(1) = \begin{cases} 0 & \text{αν } i \neq 7 \\ 1 & \text{αν } i=7 \end{cases} \quad 3-3-5d$$

όπου βέβαια $s^*=7$ και b_i οι γνωστές παράμετροι του αρχικού αλγόριθμου.

Για να επιτευχθούν τα ανωτέρω πρέπει να θεωρήσουμε ότι το b_7^* είναι ένα πολυώνυμο του σ π.χ. $b_7^* = \gamma_4 \sigma^4 + \gamma_3 \sigma^3 + \gamma_2 \sigma^2 + \gamma_1 \sigma + \gamma_0$ και να επιλύσουμε τις υπόλοιπες εξισώσεις λαμβάνοντας υπ' όψη τις 3-3-5. Όπως θα παρατηρήσει κανείς τελικά θα τού μείνει μια ελεύθερη παράμετρος (η γ_4) με την οποία θα μπορούσε να ελαχιστοποιήσει το σφάλμα 5^{ης} τάξης.

Απο τις καταστατικές εξισώσεις μπορούμε να υπολογίσουμε τα T_{5k}^* $k=1,2,\dots,9$ λαμβάνοντας υπ' όψη την 3-2-8. Αφού επιλέξουμε μια νόρμα εργαζόμεθα με σκοπό την ελαχιστοποίησή της για κάθε μέθοδο. Δηλαδή προσπαθούμε να επιλέξουμε το γ_4 έτσι ώστε τα T_{5k}^* να γίνουν όσο πίο μικρά κατ' απόλυτο τιμή γίνεται.

Στο scaled πολυώνυμο ευστάθειας συμμετέχει η παράμετρος σ . Αυτό συμβαίνει διότι τα b_i είναι πολυώνυμα του σ . Ετσι για μια scaled μέθοδο 4^{ης} τάξης το πολυώνυμο ευστάθειας είναι :

$$\begin{aligned} \mu(z) = & 1 + \sigma z + \frac{\sigma^2 z^2}{2} + \frac{\sigma^3 z^3}{6} + \frac{\sigma^4 z^4}{24} + z^5 \cdot (T_{59}^* + \frac{\sigma^5}{120}) + \\ & z^6 \cdot (T_{6,20}^* + \frac{\sigma^6}{720}) + z^7 \cdot (T_{7,48}^* + \frac{\sigma^7}{5040}) \end{aligned} \quad 3-3-6\alpha$$

και σε συμπαγή μορφή

$$\mu(z) = 1 + \sum_i (T_{i,i_k}^* + \frac{\sigma^i}{i!}) \quad 3-3-6\beta$$

Για $i=1,2,3,4$ τα $T_{i i_k}^*$ είναι πλήρως καθορισμένα από την μέθοδο και έχουμε $T_{i i_k}^* = \sigma^i / i!$. Στις υψηλότερες τάξεις οι όροι του $\mu(z)$ καθορίζονται από τις παραμέτρους του scaling. Σημειώνουμε ότι η επιλογή του γ_4 μπορεί να γίνει έτσι ώστε να αυξηθεί το διάστημα της απόλυτης ευστάθειας.

3.4 Κατασκευή scaled μεθόδων 4^{ης} τάξης.

Στην παράγραφο αυτή θα δώσουμε τις scaled επεκτάσεις δύο μεθόδων 4^{ης} τάξης.

(1) Μέθοδος Fehlberg #1

Η μέθοδος αυτή αναπτύχθηκε από τον Fehlberg [1969] ο οποίος προτείνει δύο μεθόδους 4^{ης} τάξης. Ο πίνακας των συντελεστών της μεθόδου είναι ο 3-4α

Πρόκειται δηλ. για embedded μέθοδο με δύο προσεγγίσεις 4^{ης} και 5^{ης} τάξης. Επιλύοντας τις εξισώσεις σύμφωνα με τα παραπάνω έχουμε:

$$b_1^* = (-\sigma \cdot (396 \cdot \sigma^3 - 1148 \cdot \sigma^2 + 1155 \cdot \sigma - 450) / 450)$$

$$b_3^* = 6 \cdot \sigma^2 \cdot (9 \cdot \sigma^2 - 22 \cdot \sigma + 15) / 25$$

$$b_4^* = 16 \cdot \sigma^2 \cdot (9 \cdot \sigma^2 - 22 \cdot \sigma + 15) / 225 \quad 3-4-1$$

$$b_5^* = (-\sigma^3 \cdot (3 \cdot \sigma - 4) / 30)$$

$$b_8^* = (-6 \cdot \sigma^2 \cdot (18 \cdot \sigma^2 - 34 \cdot \sigma + 15) / 25)$$

$$b_7^* = \sigma^2 \cdot (5 \cdot \sigma^2 - 8 \cdot \sigma + 3) / 2$$

0							
$\frac{2}{9}$	$\frac{2}{9}$						
$\frac{1}{3}$	$\frac{1}{12}$	$\frac{1}{4}$					
$\frac{3}{4}$	$\frac{69}{128}$	$-\frac{243}{128}$	$\frac{135}{64}$				
1	$-\frac{17}{12}$	$\frac{27}{4}$	$-\frac{27}{5}$	$\frac{16}{15}$			
$\frac{5}{6}$	$\frac{65}{432}$	$-\frac{5}{16}$	$\frac{13}{16}$	$\frac{4}{27}$	$\frac{5}{144}$		
	$\frac{47}{450}$	0	$\frac{12}{25}$	$\frac{32}{225}$	$\frac{1}{30}$	$\frac{6}{25}$	5 ^η τάξη
	$\frac{1}{9}$	0	$\frac{9}{20}$	$\frac{16}{45}$	$\frac{1}{12}$	0	4 ^η τάξη

Πίνακας 3-4α μέθοδος Fehlberg #1

Για την ανάλυση σφάλματος πρέπει να εξετάσουμε τους εννέα συντελεστές των σφαλμάτων 5^{ης} τάξης. Δεδομένης της γραμμικής εξάρτησης που έχουν όμως τα T_{5i}^* από τις απλοποιητικές παραδοχές αυτά που πρέπει να εξετάσουμε είναι τα:

$$T_{51}^* = (-\sigma^2 \cdot (48 \cdot \sigma^3 - 131 \cdot \sigma^2 + 118 \cdot \sigma - 35) / 48)$$

$$T_{55}^* = (-\sigma^3 \cdot (\sigma^2 - 2 \cdot \sigma + 1))$$

3-4-2

$$T_{56}^* = (-\sigma^2 \cdot (16 \cdot \sigma^3 - 47 \cdot \sigma^2 + 46 \cdot \sigma - 15) / 16)$$

Χρησιμοποιώντας τη νορμα των Enright et.al [1986] ,δεν είναι ανάγκη να αναφέρουμε τους γραμμικά εξαρτημένους συντελεστές. Τελικά ισχύει η σχέση.

$$\sup_{\sigma \in (0,1)} \| T_5^* \|_{\epsilon} = \frac{108}{3125} \quad 3-4-3$$

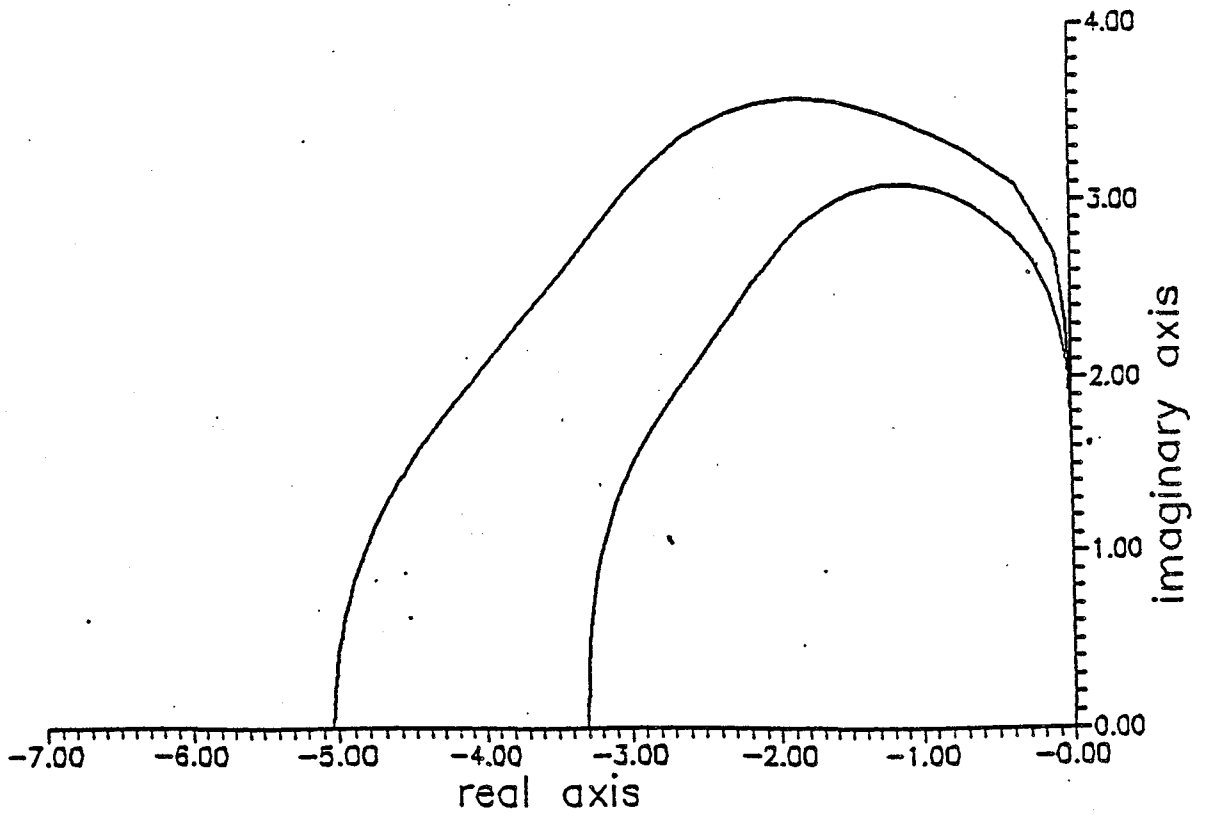
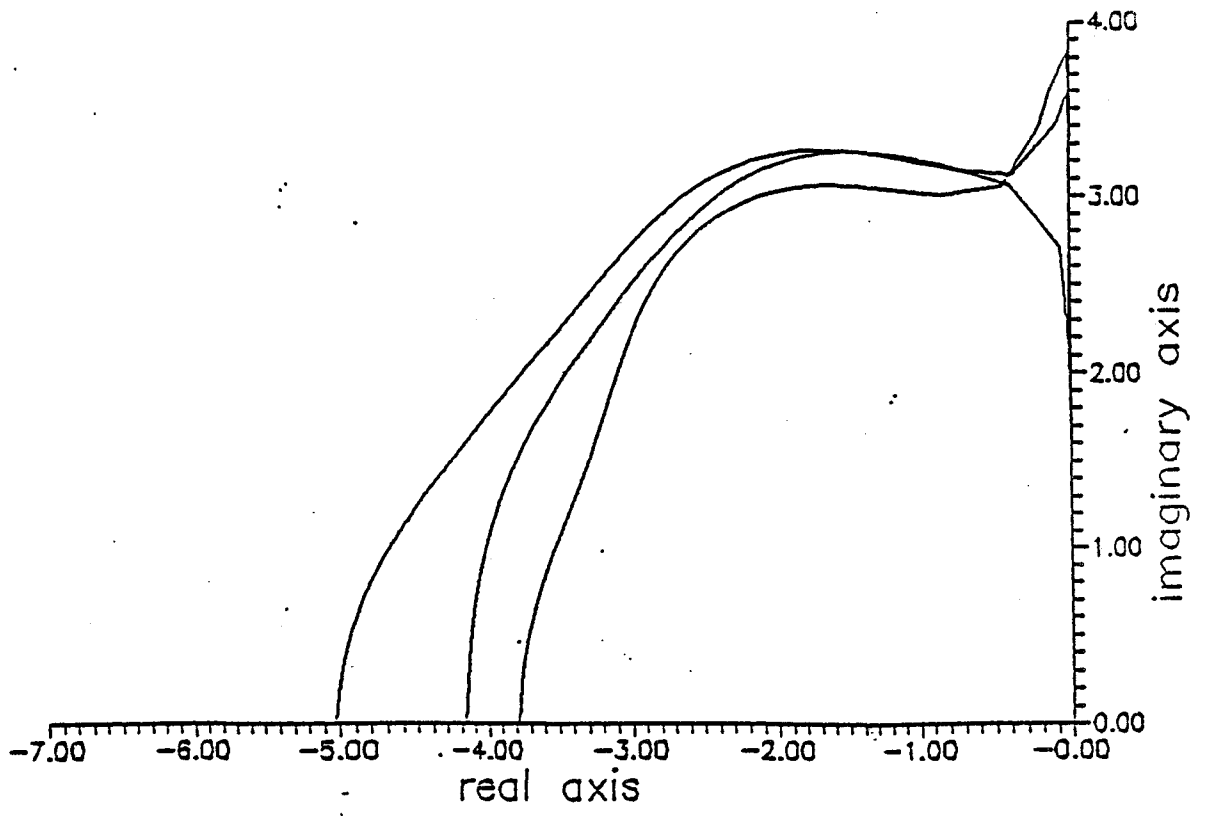
ενώ για τον βασικό αλγόριθμο έχουμε ότι η νόρμα του πρωτεύοντα όρου του τοπικού σφάλματος αποκοπής (Δηλ. ο συντελεστής 5^{ης} τάξης) δίνεται από τη σχέση:

$$\| T_5 \| = \frac{1}{4} \quad 3-4-4$$

Η τιμή που αναφέρεται στην σχέση 3-4-3 είναι η μικρότερη που μπορεί να πετύχει κανείς. Ηδη όμως είναι σημαντικά μικρότερη απο την τιμή που δίνεται στην σχέση 3-4-4. Η λύση για $\sigma=1$ είναι η λύση 5^{ης} τάξης που συνοδεύει την αρχική μέθοδο.

Το πολυώνυμο ευστάθειας της scaled μεθόδου είναι:

$$\begin{aligned} \mu(z) = & \sigma^2 \cdot z^7 \cdot (5 \cdot \sigma^2 - 8 \cdot \sigma + 3) / 1920 + \sigma^2 \cdot z^6 \cdot (2 \cdot \sigma^2 + 2 \cdot \sigma - 3) / 960 + \\ & + \sigma^3 \cdot z^5 \cdot (2 \cdot \sigma - 1) / 120 + \sigma^4 \cdot z^4 / 24 + \sigma^3 \cdot z^3 / 6 + \sigma^2 \cdot z^2 / 2 + \sigma \cdot z + 1 \end{aligned} \quad 3-4-5$$



Παρατηρούμε το σχήμα-1 που δείχνει δύο χωρία ευστάθειας της scaled μεθόδου (για $\sigma=0.4$ και $\sigma=0.8$), καθώς και το χωρίο ευστάθειας της αρχικής μεθόδου. Είναι φανερό πως ανάλογα την επιλογή του σ υπάρχει μια μικρή υπεροχή της μεθόδου του Fehlberg στα κεντρικά σημεία, αλλά υστερεί έναντι της scaled μεθόδου στα σημεία κοντά στους κόμβους.

(ii) Η μέθοδος RKDP5(4)7M των Dormand & Prince [1980]

Οι Dormand-Prince [1980] πρότειναν μία embedded μέθοδο με δύο προσεγγίσεις 4^{ης} & 5^{ης} τάξης. Το ιδιαίτερο χαρακτηριστικό αυτής της μεθόδου είναι ότι ενώ στον πίνακα των συντελεστών της βλέπουμε να απαιτεί επτά υπολογισμούς της f σε κάθε βήμα, στην ουσία δεν συμβαίνει κάτι τέτοιο. Πράγματι παρατηρεί κανείς ότι ο τελευταίος υπολογισμός της f μπορεί μετά από ένα επιτυχημένο βήμα να χρησιμοποιηθεί σαν να είναι ο πρώτος υπολογισμός της f στο επόμενο βήμα. Έτσι τελικά πρόκειται για μία μέθοδο 6-βαθμίδων. Οι συντελεστές της δίνονται στον πίνακα 3-4β

Επιλύουμε τις εξισώσεις ακολουθώντας την ίδια διαδικασία με προηγουμένως και παίρνουμε

$$b_1^* = (-\sigma \cdot (78025 \cdot \sigma^3 - 212884 \cdot \sigma^2 + 198028 \cdot \sigma - 69504)) / 69504$$

$$b_2^* = 0$$

$$b_3^* = 100 \cdot \sigma^2 \cdot (5359 \cdot \sigma^2 - 12528 \cdot \sigma + 8074) / 201453$$

$$b_4^* = (-25 \cdot \sigma^2 \cdot (7719 \cdot \sigma^2 - 13628 \cdot \sigma + 5004)) / 34752$$

$$b_5^* = 2187 \cdot \sigma^2 \cdot (1875 \cdot \sigma^2 - 3388 \cdot \sigma + 1332) / 1227904$$

$$b_6^* = (-11 \cdot \sigma^2 \cdot (2235 \cdot \sigma^2 - 4108 \cdot \sigma + 1692)) / 15204$$

$$b_7^* = \sigma^2 \cdot (415 \cdot \sigma^2 - 649 \cdot \sigma + 234) / 181$$

0							
$\frac{1}{5}$	$\frac{1}{5}$						
$\frac{3}{10}$	$\frac{3}{40}$	$\frac{9}{40}$					
$\frac{4}{5}$	$\frac{44}{45}$	$-\frac{56}{15}$	$\frac{32}{9}$				
$\frac{8}{9}$	$\frac{19372}{6561}$	$-\frac{25360}{2187}$	$\frac{64448}{6561}$	$-\frac{212}{729}$			
1	$\frac{9017}{3168}$	$-\frac{355}{33}$	$\frac{46732}{5247}$	$\frac{49}{176}$	$-\frac{5103}{18656}$		
1	$\frac{35}{384}$	0	$\frac{500}{1113}$	$\frac{125}{192}$	$-\frac{2187}{6784}$	$\frac{11}{84}$	
5 ^η τάξη	$\frac{35}{384}$	0	$\frac{500}{1113}$	$\frac{125}{192}$	$-\frac{2187}{6784}$	$\frac{11}{84}$	
4 ^η τάξη	$\frac{5179}{57600}$	0	$\frac{7571}{16695}$	$\frac{393}{640}$	$-\frac{92097}{339200}$	$\frac{187}{2100}$	$\frac{1}{40}$

Πίνακας 3-4β μέθοδος Dormand-Prince

Τα T_{5i}^* που απαιτούνται για την ανάλυση σφάλματος είναι:

$$T_{51}^* = (-\sigma^2 \cdot (362 \cdot \sigma^3 - 919 \cdot \sigma^2 + 752 \cdot \sigma - 195)) / 362$$

$$T_{55}^* = (-\sigma^2 \cdot (181 \cdot \sigma^3 - 434 \cdot \sigma^2 + 325 \cdot \sigma - 72) / 181) \quad 3-4-7$$

$$T_{56}^* = (-\sigma^2 \cdot (181 \cdot \sigma^3 - 488 \cdot \sigma^2 + 433 \cdot \sigma - 126) / 181)$$

Για τη μέθοδο αυτή έχουμε ότι

$$\max_{\sigma \in (0,1)} \| T_5^* \|_{\infty} \cong 0.01793 . \quad 3-4-8$$

Σημειώνουμε ότι ο παραπάνω αριθμός βρέθηκε σαν ο μέγιστος απο τους αριθμούς που ορίζει το σύνολο :

$$\{ 120 \cdot |T_{51}^*(\sigma)| , 120 \cdot |T_{55}^*(\sigma)| , 30 \cdot |T_{56}^*(\sigma)| \quad \forall \sigma \in (0,1) \}$$

Δεν συμμετέχουν τα υπόλοιπα T_{51}^* λόγω της γραμμικής εξάρτησης αλλά και λόγω του ότι είναι μικρότερα ή ίσα σε μέγεθος από τα αναφερόμενα.

Αντίθετα για την scaled μέθοδο που προτάθηκε από τους Dormand & Prince [1987] ισχύει η σχέση,

$$\max_{\sigma \in (0,1)} \| T_5^* \| \cong 0.043014 \quad 3-4-9$$

δηλαδή είναι κατά πολύ χειρότερη η τιμή της νόρμας. Να σημειωθεί ότι προτιμάμε τη χρήση αυτής της νόρμας και όχι της Ευκλείδειας διότι μετα την απλοποιητική παραδοχή 2-3-3 πολλοί συντελεστές στοιχειωδών διαφορικών παραλείπονται και ως

εκ' τούτου μέρος μόνον αυτών αναφέρονται. Επίσης σημαντικό είναι το ότι για να επιτευχθεί η ελαχιστοποίηση δεν χρειάζεται ψάξιμο με τον υπολογιστή αλλά η εξίσωση $T_{55}^* = T_{59}^*$ που καθορίζει το γ_4 .

Το πολυώνυμο ευστάθειας της scaled μεθόδου είναι:

$$\mu(z) = \sigma^2 \cdot z^7 \cdot (415 \cdot \sigma^2 - 649 \cdot \sigma + 234) / 108600 - \sigma^2 \cdot z^6 (160 \cdot \sigma^2 - 863 \cdot \sigma + 522) / 108600 + \sigma^2 \cdot z^5 (488 \cdot \sigma^2 - 433 \cdot \sigma + 126) / 21720 + \sigma^4 \cdot z^4 / 24 + \sigma^3 \cdot z^3 / 6 + \sigma^2 \cdot z^2 / 2 + \sigma \cdot z + 1$$

Στο σχήμα-2 βλέπουμε το χωρίο ευστάθειας της scaled μεθόδου για $\sigma=0.65$ και το χωρίο ευστάθειας της αρχικής μεθόδου. Υπάρχει μια σαφής υπεροχή της scaled μεθόδου για κάθε επιλογή του σ , που δικαιολογείται και από την κατασκευή της RKDP54, (Dormand-Prince [1980]), όπου η προσέγγιση 4^{ης} τάξης είχε πολύ καλύτερα χαρακτηριστικά ευστάθειας από την προσέγγιση 5^{ης} τάξης.

Με την ίδια διαδικασία μπορεί κανείς να παράγει scaled μεθόδους για τις δημοφιλείς μεθόδους 4^{ης} & 5^{ης} τάξης.

3.5 Μέθοδοι 5^{ης} τάξης

Στην παράγραφο αυτή θα ασχοληθούμε με τη παραγωγή προσεγγιστικής λύσης 5^{ης} τάξης σε προκαθορισμένο ενδιάμεσο σημείο έστω το $x_n + t h$ ενός επιτυχούς βήματος $[x_n, x_{n+1}]$. Δηλαδή όταν δεν μάς ενδιαφέρει η συνεχής λύση και η κατασκευή piece-wise παρεμβολικών πολυωνύμων. Η φιλοσοφία που δουλεύουμε

σε αυτή την περίπτωση είναι διαφορετική. Δεν μας απασχολεί αν οι παράμετροι a_{ij} και c_i εξαρτώνται από το t . Αυτό σημαίνει όμως ότι τα επί πλέον f που θα χρειαστούμε είναι πλήρως εξαρτημένα από το t και έτσι δεν μπορούν να χρησιμοποιηθούν για την παραγωγή άλλης προσέγγισης σε διαφορετικό σημείο. Βέβαια τώρα οι νέες παράμετροι που εμπλέκονται στις εξισώσεις που έχουμε να λύσουμε είναι πιά πολλές και έτσι είναι δυνατόν να έχουμε τις επιθυμητές προσεγγίσεις με μικρότερο υπολογιστικό κόστος.

Η Horn [1983], ασχολήθηκε πρώτη με το θέμα αυτό δίνοντας αλγόριθμο για την εύρεση προσεγγίσεων 5^{ης} τάξης για ενδιάμεσα a-priori ορισμένα σημεία. Ο αλγόριθμος αυτός συνόδευε τη μέθοδο RKF45 και απαιτούσε 2 επί πλέον υπολογισμούς της f . Εμείς εδώ χρησιμοποιούμε τον πρώτο υπολογισμό της f στο επόμενο βήμα ο οποίος έτσι και αλλιώς θα υπολογιστεί. Οι παράμετροι που χαρακτηρίζουν αυτή τη βαθμίδα (δηλαδή τα $a_{7i} = b_i$ $i=1,2,\dots,6$ και το $c_7=1$) μαζί με τις παράμετρους μιας ακόμα βαθμίδας που είναι αναγκαία, είναι αρκετές για να ικανοποιήσουν τις scaled εξισώσεις 5^{ης} τάξης. Έτσι χρειάζεται ένας υπολογισμός της f λιγότερος, σύμφωνα με την δική μας τεχνική. Αναλυτικοί τύποι που καθορίζουν τις νέες παραμέτρους δίνονται για πρώτη φορά και δεν είμαστε αναγκασμένοι να τους παράγουμε δυναμικά για κάθε σημείο όπως προτείνει η Horn. [1983]

Από τους άλλους ερευνητές οι Calvo et.al [1990] και ο

Shampine [1986] έχουν ενδιαφερθεί για point-wise παρεμβολή με κύριο στόχο όμως την παραγωγή σημείων στήριξης για Hermite παρεμβολικά πολυώνυμα για την μέθοδο των Dormand & Prince [1980]. Σημειώνεται ότι στην παρούσα περίπτωση δεν έχει καμιά έννοια η συνέχεια καθώς δεν παράγονται συνεχείς λύσεις σε όλο το διάστημα αλλά απλές προσεγγίσεις.

Οι 9 εξισώσεις 5^{ης} τάξης που πρέπει να ικανοποιούνται μαζί με τις 8 εξισώσεις (I-1 έως IV-4) που παρατίθενται στην παράγραφο 3-3 είναι :

$$V-1 \quad \sum b_i^* \cdot c_i^4 = \frac{\sigma^5}{5}$$

$$V-2 \quad \sum b_i^* \cdot c_i^2 \cdot \sum a_{ij} \cdot c_j = \frac{\sigma^5}{10}$$

$$V-3 \quad \sum b_i^* \cdot (\sum a_{ij} \cdot c_j)^2 = \frac{\sigma^5}{20}$$

$$V-4 \quad \sum b_i^* \cdot c_i \cdot \sum a_{ij} \cdot c_j^2 = \frac{\sigma^5}{15}$$

$$V-5 \quad \sum b_i^* \cdot \sum a_{ij} \cdot c_j^3 = \frac{\sigma^5}{20}$$

$$V-6 \quad \sum b_i^* \cdot c_i \cdot \sum a_{ij} \cdot \sum a_{jk} \cdot c_k = \frac{\sigma^5}{30}$$

$$V-7 \quad \sum b_i^* \cdot \sum a_{ij} \cdot c_j \cdot \sum a_{jk} \cdot c_k = \frac{\sigma^5}{40}$$

$$V-8 \quad \sum b_i^* \cdot \sum a_{ij} \cdot \sum a_{jk} \cdot c_k^2 = \frac{\sigma^5}{60}$$

$$V-9 \quad \sum b_i^* \cdot \sum a_{ij} \cdot \sum a_{jk} \cdot \sum a_{kl} \cdot c_l = \frac{\sigma^5}{120}$$

Αν η μέθοδος είναι μια 6-βάθμιδη embedded μεθοδος 4(5) τάξεων τότε η 7^η βαθμίδα εξαρτάται από τους συντελεστές $c_7=1$ και $a_{7i}=b_i$ $i=1,2,\dots,7$ που στην ουσία είναι η πρώτη βαθμίδα του επόμενου βήματος. Η 8^η βαθμίδα είναι και η επιπρόσθετη με τους συντελεστές a_{8i} $i=1,2,\dots,7$ και το c_8 ελεύθερες παράμετρος.

Με τις απλοποιητικές παραδοχές 3-3-1 έως 3-3-3 και τις παραδοχές

$$\sum a_{8i} c^k = \frac{c_8^{k+1}}{k+1}, k=1,2 \quad 3-5-1$$

$$\sum b_i^* c_i a_{i2} = 0 \quad 3-5-2$$

$$\sum b_i^* \sum a_{ij} a_{j2} = 0 \quad 3-5-3$$

έχει να λύσει τελικά κανείς το σύστημα των εξισώσεων 3-3-1, 3-5-1, 3-5-2, 3-5-3, και I-1, II-1, III-1, IV-1, V-1, V-5 με άγνωστους τα έξι a_{8i} (το a_{81} καθορίζεται από την 2-3-3) τα έξι b_i^* ($b_2^*=0$) και το c_8 . Δηλαδή έχουμε 11 εξισώσεις και 14 άγνωστους. Το σ στις εξισώσεις αντικαθίσταται από το δοθέν t . Έτσι μπορούμε θεωρώντας παράμετρος τα c_8, a_{82}, a_{83} να

κατασκευάσουμε την λύση y_{n+1}^* που προσεγγίζει την $y(x_n+th)$ από τη σχέση:

$$y_{n+1}^* = y_n + h \sum_{i=1}^8 b_i^*(t) f^i \quad 3-5-4$$

με $f^7 = f^1$ (επόμενου βήματος) και f^8 το επιπρόσθετο f . Είναι προφανές ότι μπορούμε να κατασκευάσουμε λύση σε k σημεία με τουλάχιστον k επιπλέον f . Η περίπτωση που μας ενδιαφέρει ιδιαίτερα είναι να βρούμε δύο σημεία που μας κοστίζουν δύο f , διότι χρειάζονται για να στηριχθεί πολυώνυμο παρεμβολής. Η προσέγγιση στο σημείο x_n+th βρίσκεται χρησιμοποιώντας τους παρακάτω συντελεστές και τη σχέση 3-5-4

$$b_8^* = 4 \cdot t^2 \cdot (4 \cdot t^3 - 11 \cdot t^2 + 10 \cdot t - 3) / ((a_{82} + 8 \cdot c_8 \cdot (10 \cdot c_8^2 - 12 \cdot c_8 + 3)) \cdot (c_8 - 1))$$

$$b_7^* = t^2 \cdot (t-1) \cdot (c_8 \cdot (5 \cdot t - 3) - 2 \cdot t \cdot (2 \cdot t - 1)) / (2 \cdot (c_8 - 1))$$

$$b_6^* = 2 \cdot t^2 \cdot (a_{82} \cdot (416 \cdot t^3 - 1195 \cdot t^2 + 1140 \cdot t - 360) - 8 \cdot c_8 \cdot (10 \cdot c_8^2 \cdot (51 \cdot t^2 - 100 \cdot t + 48) - 6 \cdot c_8 \cdot (68 \cdot t^3 - 85 \cdot t^2 - 30 \cdot t + 45) + 3 \cdot t \cdot (64 \cdot t^2 - 125 \cdot t + 60))) / (55 \cdot (a_{82} + 8 \cdot c_8 \cdot (10 \cdot c_8^2 - 12 \cdot c_8 + 3)))$$

$$b_5^* = t^2 \cdot (a_{82} \cdot (c_8 \cdot (416 \cdot t^3 - 1060 \cdot t^2 + 890 \cdot t - 255) - 316 \cdot t^3 + 5 \cdot (157 \cdot t^2 - 128 \cdot t + 36)) + 24 \cdot c_8 \cdot (10 \cdot c_8^3 \cdot (28 \cdot t^2 - 50 \cdot t + 19) - 2 \cdot c_8^2 \cdot (112 \cdot t^3 - 270 \cdot t + 125) + 15 \cdot c_8 \cdot (20 \cdot t^3 - 27 \cdot t^2 + 4) - t \cdot (76 \cdot t^2 - 125 \cdot t + 40))) / (50 \cdot (a_{82} + 8 \cdot c_8 \cdot (10 \cdot c_8^2 - 12 \cdot c_8 + 3)) \cdot (c_8 - 1))$$

$$b_4^* = (-28561 \cdot t^2 \cdot (a_{82} \cdot (96 \cdot t^3 - 225 \cdot t^2 + 170 \cdot t - 45) + 16 \cdot c_8 \cdot (5 \cdot c_8^2 \cdot (39 \cdot t^2 - 70 \cdot t + 27) - 3 \cdot c_8 \cdot (52 \cdot t^3 - 65 \cdot t^2 - 10 \cdot t + 15) + 6 \cdot t \cdot (9 \cdot t^2 - 15 \cdot t + 5)))) / (225720 \cdot (a_{82} + 8 \cdot c_8 \cdot (10 \cdot c_8^2 - 12 \cdot c_8 + 3)))$$

$$b_3^* = (-512 \cdot t^2 \cdot (a_{82} \cdot (312 \cdot t^3 - 945 \cdot t^2 + 980 \cdot t - 360) - 8 \cdot c_8 \cdot (10 \cdot c_8^2 \cdot (87 \cdot t^2 - 200t + 126) - 6 \cdot c_8 \cdot (116t^3 - 145t^2 - 110t + 165) + 3 \cdot t \cdot (168 \cdot t^2 - 375 \cdot t + 220)))) / (12825 \cdot (a_{82} + 8 \cdot c_8 \cdot (10 \cdot c_8^2 - 12 \cdot c_8 + 3)))$$

$$b_1^* = t \cdot (a_{82} \cdot (1248 \cdot t^4 - 4365 \cdot t^3 + 5810 \cdot t^2 - 3645 \cdot t + 1080) - 16 \cdot (5 \cdot c_8^3 \cdot (933 \cdot t^3 - 2690 \cdot t^2 + 2709 \cdot t - 1080) - 3 \cdot c_8^2 \cdot (1244 \cdot t^4 - 1555 \cdot t^3 - 2270 \cdot t^2 + 4485 \cdot t - 2160) + 6 \cdot c_8 \cdot (723 \cdot t^4 - 1755 \cdot t^3 + 1135 \cdot t^2 + 135 \cdot t - 270) - 270 \cdot t \cdot (4 \cdot t^3 - 11 \cdot t^2 + 10 \cdot t - 3))) / (1080 \cdot (a_{82} + 8 \cdot c_8 \cdot (10 \cdot c_8^2 - 12 \cdot c_8 + 3)))$$

3-5-5

$$a_{84} = (-169 \cdot (4 \cdot a_{82} \cdot (2 \cdot c_8 \cdot (52 \cdot t - 33) - 284 \cdot t + 201) - 195 \cdot a_{83} \cdot (4 \cdot t - 3) + 128 \cdot c_8 \cdot (c_8^3(52 \cdot t - 9) - 2 \cdot c_8^2(64t - 15) + c_8 \cdot (88t - 21) - 12 \cdot t))) / (33792 \cdot (4 \cdot t - 3))$$

$$a_{85} = (12 \cdot a_{82} \cdot (6 \cdot c_8 \cdot (16 \cdot t - 11) - 244 \cdot t + 177) - 513 \cdot a_{83} \cdot (4 \cdot t - 3) + 128 \cdot c_8 \cdot (3 \cdot c_8^3 \cdot (32t - 9) - 2 \cdot c_8^2 \cdot (118 \cdot t - 39) + 54 \cdot c_8 \cdot (3 \cdot t - 1) - 18 \cdot t))) / (768 \cdot (4 \cdot t - 3))$$

$$a_{86} = (12 \cdot a_{82} \cdot (2 \cdot c_8 \cdot (52 \cdot t - 33) - 812 \cdot t + 597) - 2565 \cdot a_{83} \cdot (4 \cdot t - 3) + 128 \cdot c_8 \cdot (3c_8^3(52t - 9) - 4c_8^2(118t - 39) + 18 \cdot c_8 \cdot (22 \cdot t - 9) - 36 \cdot t))) / (2112 \cdot (4 \cdot t - 3))$$

$$a_{87} = t \cdot (c_8 - 1) \cdot (a_{82} + 8 \cdot c_8 \cdot (10 \cdot c_8^2 - 12 \cdot c_8 + 3)) / (8 \cdot (4 \cdot t - 3))$$

Απο εδώ και πέρα πρέπει να ελαχιστοποιήσει κανείς το συντελεστή του σφάλματος 6^{ης} τάξης δεδομένου ότι έχει 3 ελεύθερες παραμέτρους. Επειδή δεν έχει πρακτική αξία κάτι τέτοιο το αντιπαρερχόμαστε. Βέβαια δεν έγινε όλη αυτή η αναφορά για το τίποτα. Στο επόμενο κεφάλαιο θα δούμε το πόσο χρήσιμες είναι αυτές οι εκφράσεις για τη κατασκευή παρεμβολικών πολυωνύμων. Ετσι καλό θα ήταν η βελτιστοποίηση να γίνει απ' ευθείας εκεί που θά κατασκευάσουμε τα *riese-wise* πολυώνυμα συνεχούς λύσης. Αλλωστε το να επιχειρήσει κανείς κάτι τέτοιο εδώ δεν έχει έννοια διότι θα έπρεπε να το κάνει για μεγάλο φάσμα τιμών του t και εκτός από ασύμφορο θα ήταν και αναποτελεσματικό.

3.7 Αριθμητικά αποτελέσματα.

Οι Hull et al. [1973] υπέδειξαν ένα σύνολο προβλημάτων στα οποία μπορεί κανείς να κάνει έλεγχο των διαφόρων αριθμητικών μεθόδων. Τα βασικά μεγέθη που μετρώνται είναι το πλήθος των υπολογισμών της f και το μέγιστο ολικό σφάλμα που παρουσιάζεται στους κόμβους κάθε βήματος. Αυτό γίνεται για διάφορες ανοχές (TOL) .

Στη συγκεκριμένη περίπτωση μας ενδιαφέρει τι γίνεται με το σφάλμα του scaled αλγόριθμου σε ενδιάμεσα σημεία. Ετσι λοιπόν χωρίζουμε το κάθε βήμα σε 10 μέρη $z_{ni} = x_n + ih_n/10$, $i=1,2,\dots,10$ και εφαρμόζουμε τις νέες μεθόδους 4^{ης} τάξης γι' αυτά τα σημεία. Οι λόγοι του μέγιστου scaled σφάλματος προς

το σφάλμα της κυρίως μεθόδου είναι ενδεικτικοί για την αξία των νέων μεθόδων και παρουσιάζονται στον πίνακα 3-6. Το υπολογιστικό κόστος που σώζεται είναι πολύ μεγάλο. Για την παραπάνω εφαρμογή θα χρειαζόταν τουλάχιστον 9-πλάσιος υπολογιστικός χρόνος αν προσαρμόζαμε το βήμα στα σημεία που θέλουμε λύση.

Τα προβλήματα αρχικών τιμών στα οποία θα γίνει ο έλεγχος είναι:

Πρόβλημα 1^ο

$$y' = -y, y(0) = 1, x \in [0, 20]$$

ακριβής λύση $y(x) = e^{-x}$

Πρόβλημα 2^ο

$$y' = -y^3/2, y(0)=1, x \in [0, 20]$$

ακριβής λύση $y(x) = \frac{1}{\sqrt{1+x}}$

Πρόβλημα 3^ο

$$y' = \frac{y}{4} \left(1 - \frac{y}{20} \right), y(0) = 1, x \in [0, 20]$$

ακριβής λύση

$$y(x) = \frac{20}{(1+19 \cdot e^{-x/4})}$$

Πρόβλημα 4^ο

$$\begin{aligned}y_1' &= y_3 & y_1(0) &= 0.5 \\y_2' &= y_4 & y_2(0) &= 0 \\y_3' &= -y_1/r^3 & y_3(0) &= 0 \\y_4' &= -y_2/r^3 & y_4(0) &= \sqrt{0.75} \\x &\in [0, 20]\end{aligned}$$

ακριβής λύση

$$y_1(x) = \cos(u) - 0.5$$

$$y_2(x) = \sqrt{0.75} \cdot \sin(u)$$

$$y_3(x) = \frac{-\sin(u)}{(1 - 0.5 \cdot \cos(u))}$$

$$y_4(x) = \frac{\sqrt{0.75} \cdot \cos(u)}{(1 - 0.5 \cdot \cos(u))}$$

με

$$u - 0.5 \sin(u) - x = 0 \quad \text{και} \quad r = \sqrt{y_1^2 + y_2^2}.$$

Οι πίνακες με τους λόγους των σφαλμάτων είναι οι παρακάτω

Πίνακας 3-6α

(i) μέθοδος Fehlberg							
TOL	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}
Π1	1.54	1.41	1.38	1.30	1.28	1.25	1.22
Π2	1.01	1.01	1.00	1.00	1.00	1.00	1.00
Π3	1.06	1.31	2.00	2.01	2.11	2.14	2.34
Π4	1.01	1.00	1.01	1.00	1.00	1.00	1.00

Πίνακας 3-6β

(ii) μέθοδος Dormand-Prince							
TOL	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}
Π1	1.10	1.01	1.00	1.00	1.00	1.00	1.00
Π2	1.55	1.48	1.32	1.06	1.01	1.10	1.09
Π3	1.12	1.33	1.35	1.44	1.41	1.36	1.29
Π4	1.00	1.00	1.01	1.01	1.00	1.00	1.00

Κ Ε Φ Α Λ Α Ι Ο 4^{ΟΝ}

ΠΑΡΕΜΒΟΛΙΚΑ ΠΟΛΥΩΝΜΑ ΓΙΑ ΜΕΘΟΔΟΥΣ RUNGE-KUTTA

4.1 Εισαγωγή

Στο προηγούμενο κεφάλαιο εξετάσαμε τις Scaled Runge-Kutta μεθόδους και προτείναμε συγκεκριμένες μεθόδους αυτής της κατηγορίας. Οι μέθοδοι αυτές δίνουν την δυνατότητα προσέγγισης της λύσης του προβλήματος αρχικών τιμών

$$y' = f(x,y) , x \geq a , y \in \mathbb{R}^m$$

4-1-1

$$y(a) = y_0$$

σε ενδιάμεσα σημεία ενός επιτυχούς βήματος. Είναι φανερό ότι οι μέθοδοι αυτής της κατηγορίας μπορεί να παράγουν αριθμητικές λύσεις που είναι ολικά (globally) συνεχείς, αλλά με μεγάλο υπολογιστικό κόστος. Για μεθόδους Runge-Kutta 5^{ης} τάξης και πάνω είναι τελείως ασύμφορη η κατασκευή scaled μεθόδων.

Στο κεφάλαιο αυτό θα ασχοληθούμε με το πρόβλημα της κατασκευής μεθόδων οι οποίες δίνουν ολικά συνεχείς προσεγγίσεις με μικρό υπολογιστικό κόστος, αλλά η βάση κατασκευής τους είναι διαφορετική από την scaled τεχνική. Τέτοιες λύσεις για παράδειγμα απαιτούνται, όταν θέλουμε να

κατασκευάσουμε το γράφημα της αριθμητικής λύσης του προβλήματος 4-1-1. Συγκεκριμένα θα ασχοληθούμε με την κατασκευή παρεμβολικών πολυωνύμων (interpolants) τα οποία βασίζονται σε ικανό αριθμό προσεγγίσεων της $y(x)$ και της παραγώγου $f(x,y)$ όπως καθορίζονται από το πρόβλημα 4-1-1. Για μεθόδους Runge-Kutta υψηλής τάξης απαιτούνται αρκετές πληροφορίες .

Τα ερωτήματα που τίθενται τώρα είναι τα εξής. Μπορούμε να κατασκευάσουμε interpolants για μεθόδους Runge-Kutta χρησιμοποιώντας μόνο τις διαθέσιμες τοπικά τιμές της $y(x)$ και της $y'(x)$; Αν αυτό δεν είναι δυνατόν πόσες επί πλέον τιμές της f (function evaluations) απαιτούνται σε κάθε βήμα για να το επιτύχουμε; Μπορούμε να χρησιμοποιούμε τιμές από το προηγούμενο ή το επόμενο βήμα;

Διάφοροι ερευνητές έχουν ασχοληθεί πρόσφατα με το πρόβλημα αυτό, όπως οι Enright et.al [1986], Shampine [1985],[1986], Calvo et.al [1989],[1990]. Όλοι όμως έχουν κατασκευάσει παρεμβολικά πολυώνυμα που βασίζονται σε πληροφορίες από ένα βήμα. Τούτο έχει σαν συνέπεια για υψηλής τάξης μεθόδους να απαιτούνται πρόσθετοι υπολογισμοί της f , οι οποίοι επιβαρύνουν το κόστος της μεθόδου.

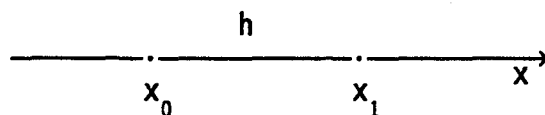
Εμείς εδώ προτείνουμε νέα πολυώνυμα παρεμβολής βασιζόμενα σε τιμές από ένα βήμα τα οποία έχουν μικρότερα σφάλματα αποκοπής από τα μέχρι σήμερα προταθέντα. Αλλά η κύρια συμβολή της διατριβής στη προσπάθεια αντιμετώπισης του θέματος είναι τα διβηματικά interpolants που προτείνονται, τα

οποία εκτός από το μηδενικό σχεδόν υπολογιστικό κόστος που απαιτούν έχουν και τα καλύτερα σφάλματα αποκοπής.

4.2 Γενική θεωρία.

Ας θεωρήσουμε το πρόβλημα αρχικών τιμών 4-1-1 και το αντίστοιχο τοπικό πρόβλημα 2-5-3 μετά από n βήματα. Υστερα από ένα επιτυχημένο βήμα εφαρμογής μιας μεθόδου Runge-Kutta $[x_n, x_{n+1}]$ οι διαθέσιμες τιμές είναι οι $y_n, y'_n, y_{n+1}, y'_{n+1}$. Οι δύο πρώτες θεωρούνται απολύτως ακριβείς αφού άλλωστε χρησιμοποιούνται σαν αρχικές τιμές. Οι δύο δεύτερες έχουν ακρίβεια τάξης $p+1$ αν η μέθοδος που χρησιμοποιούμε είναι τάξης p . Η συμμετοχή των τιμών αυτών σε ένα παρεμβολικό πολυώνυμο εγγυάται ότι η λύση μας θα είναι C^1 . Συνήθως δουλεύουμε με μια *embedded* μέθοδο που δίνει δύο προσεγγίσεις τάξεων p και $p+1$.

Είναι προφανές ότι για μικρής τάξης μεθόδους τα πράγματα είναι εύκολα. Για υψηλότερης τάξης μεθόδους πρέπει να υπολογίσουμε ενδιάμεσα σημεία τόσα στο πλήθος όσο και το p . Για να κατασκευάσουμε το πολυώνυμο παρεμβολής στο διάστημα $[x_n, x_{n+1}]$ θεωρούμε το προσανατολισμένο διάστημα



με x_0 : αρχή του άξονα

h : μοναδιαίο ίσο με το μήκος τού βήματος,

όπου για λόγους ευκολίας έγινε αναγωγή του διαστήματος που

μας ενδιαφέρει στο διάστημα $[x_0, x_1]$ με βήμα h .

Υποθέτουμε ότι στο διάστημα αυτό τώρα έχουμε διαθέσιμες τις τιμές:

$$y_r^{[n]} \quad r \in \mathbb{R}, \quad n \in \{0, 1, 2, \dots\}$$

στα αντίστοιχα σημεία x_r .

Ορίζουμε για τις διάφορες τιμές των r και n τα $q+1$ διανύσματα γραμμή,

$$V(y_r^{[n]}) = \left[\frac{d^n 1}{dx^n} \Big|_{x=r}, \frac{d^n x}{dx^n} \Big|_{x=r}, \frac{d^n x^2}{dx^n} \Big|_{x=r}, \dots, \frac{d^n x^q}{dx^n} \Big|_{x=r} \right]$$

4-2-1

όπου $q+1$ το πλήθος των διαθέσιμων σημείων τα οποία ορίζονται από τα διατεταγμένα ζεύγη $(r, n) \in \mathbb{R} \times \mathbb{N}_0$. Το σύνολο των ζευγών που ορίζουν τα διαθέσιμα σημεία είναι

$$S = \{ (r, n) / y_r^{[n]} \text{ διαθέσιμο σημείο} \}.$$

Κατόπιν ορίζουμε τον $(q+1) \times (q+1)$ πίνακα D που αποτελείται από τα διανύσματα 4-2-1. Επίσης ορίζουμε τους πίνακες $D_{r,n}$ που προκύπτουν από τον D αν στη θέση του διανύσματος $V(y_r^{[n]})$ τοποθετήσουμε το διάνυσμα $[1, \sigma, \sigma^2, \sigma^3, \dots, \sigma^q]$. Το σ ορίζει το σημείο $x = x_0 + \sigma h$, $\sigma \in (0, 1)$, στο οποίο θέλουμε ενδιάμεση λύση. Το πολυώνυμο που μας προσεγγίζει τη λύση $u(x_0 + \sigma h)$ είναι το $p(x)$ και δίνεται από την σχέση

$$p(x) = \sum_{(r,n) \in S} h^n \frac{D_{rn}}{D} y_r^{[n]} \quad 4-2-2$$

Αν $P(x)$ το πολυώνυμο που παρεμβάλει την ακριβή τοπική λύση $u(x)$ στα διαθέσιμα σημεία

$$P(x) = \sum_{(r,n) \in S} h^n \frac{D_{rn}}{D} u(x_r)^{[n]} \quad 4-2-3$$

τότε το σφάλμα του $p(x)$ που άλλωστε μας ενδιαφέρει είναι Kansy [1973]

$$\|u(x) - p(x)\| \leq \|u(x) - P(x)\| + \|P(x) - p(x)\| \quad 4-2-4$$

Ο πρώτος όρος ορίζει το σφάλμα της μεθόδου παρεμβολής αν συμμετείχαν απολύτως ακριβείς τιμές. Το σφάλμα αυτό είναι πάντα τάξης $O(h^{q+1})$ ενώ ο δεύτερος όρος έχει τάξη όση και η ακρίβεια των σημείων που χρησιμοποιούμε. Δηλαδή ισχύει.

$$\|u(x) - p(x)\| = O(h^{q+1}) + O(h^{p+1}) \quad 4-2-5$$

Ερμηνεύοντας τις 4-2-4 και 4-2-5 βλέπουμε πως αν έχουμε μια μέθοδο τάξης p , τότε οι τυχόν επιπρόσθετες τιμές πρέπει να έχουν την ίδια τάξη. Επίσης το πλήθος των σημείων που συμμετέχουν πρέπει να είναι $p+1$ διότι διαφορετικά δεν θα επιτύχουμε την επιθυμητή ακρίβεια. Περισσότερα από $p+1$ σημεία είναι ασύμφορο να χρησιμοποιηθούν όπως φαίνεται από την

4-2-4. Γενικά δηλαδή πρέπει $q+1=p+1$ και δεδομένου ότι 4 τιμές των άκρων είναι πάντα διαθέσιμες πρέπει κανείς να υπολογίζει $p-3$ επί πλέον τιμές. Συνήθως οι επιπλέον απαιτούμενες τιμές υπολογίζονται από την αντίστοιχη point-wise Scaled Runge-Kutta μέθοδο ίδιας τάξης.

Κάθε παρεμβολική μέθοδος καταλήγει σε ένα τύπο της μορφής 4-2-2. Αυτός όμως μπορεί να μετασχηματισθεί εύκολα σε τύπο Runge-Kutta μορφής όπως στον scaled αλγόριθμο. Δεν έχουμε παρά να εκφράσουμε όλες τις τιμές συναρτήσεων των x_n, y_n, h . Στον τύπο 4-2-2 τα y_r δίνονται από τον αντίστοιχο scaled τύπο αν είναι επιπρόσθετα, τα δε y'_r μπορούν απλά να θεωρηθούν σαν επιπλέον τιμές της f . Στην ουσία δηλαδή η παρεμβολή είναι ένας πιά συστηματικός τρόπος scaling μια και μας δίνει μια μεθοδολογία να ικανοποιούμε τις εξισώσεις τύπου Runge-Kutta.

Το ολικό σφάλμα της scaled μεθόδου δίνεται από την σχέση 4-2-6α το δε ολικό σφάλμα της μεθόδου παρεμβολής δίνεται από τη σχέση 4-2-6b,

$$\varepsilon_{n+\sigma}^* = y_{n+\sigma}^* - y(x_n + \sigma h_n) \quad 4-2-6a$$

$$\varepsilon_{n+\sigma} = p(x_n + \sigma h_n) - y(x_n + \sigma h_n) \quad 4-2-6b$$

Ας περιοριστούμε στην πρώτη σχέση που ισχύει και για τις δύο περιπτώσεις αφού μια μέθοδος παρεμβολής μπορεί να μετασχηματιστεί σε Runge-Kutta μορφή. Ορίζοντας κατά τα

γνωστά το τοπικό σφάλμα διακριτοποίησης $\delta_{n+\sigma}^*$ έχουμε

$$\varepsilon_{n+\sigma}^* = \varepsilon_n + \sigma h (\Phi^*(x_n, y_n, \sigma h) - \Phi^*(x_n, y(x_n), \sigma h)) - \delta_{n+\sigma}^* \quad 4-2-7$$

Ας υποθέσουμε ότι η συνάρτηση $\Phi^*(x, y, \sigma h)$ ικανοποιεί μια συνθήκη Lipschitz τότε από το λήμμα 2-1 έχουμε,

$$\|\varepsilon_{n+\sigma}^*\| \leq \|\varepsilon_n\| + \sigma h_n L \|\varepsilon_n\| + \|\delta_{n+\sigma}^*\| \quad 4-2-8$$

Επειδή $\|\varepsilon_n\| = O(h^p)$ αφού η μέθοδος RK είναι τάξης p και $\|\delta_{n+\sigma}^*\| = O(h^{q+1})$ εφ' όσον διαθέτουμε $q+1$ σημεία, συνάγεται ότι

$$\|\varepsilon_{n+\sigma}^*\| = O(h^{\min(q+1, p)})$$

Δηλαδή τελικά, για μια μέθοδο p -τάξης που το ολικό σφάλμα είναι $O(h^p)$, με p -σημεία μπορούμε να κατασκευάσουμε interpolant που διατηρεί την τάξη σφάλματος. Υπάρχουν όμως περιπτώσεις που κάποιος θέλει τοπικά στη μέθοδο p -τάξης να διατηρεί σφάλμα τάξης $O(h^{p+1})$ και τότε χρειάζεται $p+1$ τιμές αντίστοιχης τάξης.

4.3 Παρεμβολικά πολυώνυμα δύο βημάτων για μεθόδους 5ης τάξης

Πολλοί συγγραφείς έχουν ασχοληθεί με τη μελέτη και κατασκευή παρεμβολικών πολυωνύμων των πίο δημοφιλών μεθόδων *Runge-Kutta* όπως οι Enright et.al [1986], Shampine [1985],[1986],

Calvo et.al. [1989],[1990]. Έχουν περιορισθεί στη λεγόμενη τοπική παρεμβολή που χρησιμοποιεί μόνο σημεία μέσα στο παρόν διάστημα $[x_n, x_{n+1}]$. Τούτο είχε σαν αποτέλεσμα το αυξημένο υπολογιστικό κόστος.

Στην παράγραφο αυτή προτείνουμε νέα παρεμβολικά πολυώνυμα για τις πιο γνωστές μεθόδους Runge-Kutta διαφόρων τάξεων στηριζόμενοι σε τιμές που ανήκουν σε δύο διαδοχικά βήματα επιτυχούς ολοκλήρωσης της μεθόδου.

$$\begin{array}{ccc}
 y'_{n-1} & y'_n & y'_{n+1} \\
 y_{n-1} & y_n & y_{n+1} \\
 \hline
 x_{n-1} & x_n & x_{n+1} \\
 h_{n-1} = \varepsilon \cdot h_n & h_n &
 \end{array}$$

Οι προτεινόμενες μέθοδοι δεν απαιτούν επί πλέον υπολογισμούς και για $\varepsilon < 2$ δίνουν αποδεκτά σφάλματα αποκοπής.

- (i) Runge-Kutta-Fehlberg 4(5) #2 και
Runge-Kutta-Dormand-Prince 5(4)7S

Η πρώτη μέθοδος προτάθηκε από τον Fehlberg [1969], και είναι η πιο δημοφιλής μέθοδος 4(5) τάξης. Οι παράμετροί της δίδονται από τον πίνακα 4.3

Η δεύτερη είναι μία μέθοδος που παρουσιάστηκε στον πίνακα-3 του άρθρου Dormand-Prince [1980]. Για τις μεθόδους αυτές θα παρουσιάσουμε τα νέα διβηματικά πολυώνυμα παρεμβολής

τα οποία δίνουν συνεχή λύση στο διάστημα $[x_n, x_{n+1}]$ χρησιμοποιώντας τιμές και από το προηγούμενο διάστημα $[x_{n-1}, x_n]$. Οι μέθοδοι που προτείνουμε ισχύουν χωρίς μετατροπές και για τις δύο μεθόδους αλλά και για όλες τις μεθόδους 4^{ης} ή 5^{ης} τάξης.

0							
$\frac{1}{4}$	$\frac{1}{4}$						
$\frac{3}{8}$	$\frac{3}{32}$	$\frac{9}{32}$					
$\frac{12}{13}$	$\frac{1932}{2197}$	$-\frac{7200}{2197}$	$\frac{7296}{2197}$				
1	$\frac{439}{216}$	- 8	$\frac{3680}{513}$	$-\frac{845}{4104}$			
$\frac{1}{2}$	$-\frac{8}{27}$	2	$-\frac{3544}{2565}$	$\frac{1859}{4104}$	$-\frac{11}{40}$		
	$\frac{16}{135}$	0	$\frac{6656}{12825}$	$\frac{28561}{56430}$	$-\frac{9}{50}$	$\frac{2}{55}$	5 ^η τάξη
	$\frac{25}{216}$	0	$\frac{1408}{2565}$	$\frac{2197}{4104}$	$-\frac{1}{5}$	0	4 ^η τάξη

Πίνακας 4.3 μέθοδος Fehlberg #2

Για να κατασκευάσουμε το νέο πολυώνυμο παρεμβολής χρησιμοποιούμε τις τιμές $y_{n-1}, y'_{n-1}, y_n, y'_n, y_{n+1}, y'_{n+1}$ οι οποίες είναι διαθέσιμες χωρίς κόστος. Εφαρμόζουμε τα αναφερθέντα στην παράγραφο 4-2 και αν

$$h_{n-1} = \epsilon h_n$$

4-3-1

τότε το πολυώνυμο παρεμβολής δίνεται από τη σχέση

$$U_6^1(x_n + \sigma h_n) = d_1 y_{n-1} + d_2 h y'_{n-1} + d_3 y_n + d_4 h y'_n + d_5 y_{n+1} + d_6 h y'_{n+1} \quad 4-3-2$$

με

$$d_1(\sigma) = \left\{ 1 + \frac{2 \cdot (\sigma + \varepsilon) \cdot (2\varepsilon + 1)}{\varepsilon \cdot (\varepsilon + 1)} \right\} \cdot \frac{\sigma^2 \cdot (\sigma - 1)^2}{(\varepsilon + 1)^2 \cdot \varepsilon^2}$$

$$d_2(\sigma) = \left\{ 1 - \frac{2 \cdot \sigma \cdot (1 - \varepsilon)}{\varepsilon} \right\} \cdot \frac{(\sigma + \varepsilon)^2 \cdot (1 - \sigma)^2}{\varepsilon^2}$$

4-3-3

$$d_3(\sigma) = \left\{ 1 + \frac{2 \cdot (1 - \sigma) \cdot (2 + \varepsilon)}{(1 + \varepsilon)} \right\} \cdot \frac{(\sigma + \varepsilon)^2 \cdot \sigma^2}{(1 + \varepsilon)^2}$$

$$d_4(\sigma) = \frac{\sigma^2 \cdot (\sigma - 1)^2 \cdot (\sigma + \varepsilon)}{(\varepsilon + 1)^2 \cdot \varepsilon^2}$$

$$d_5(\sigma) = \frac{\sigma \cdot (\sigma + \varepsilon)^2 \cdot (1 - \sigma)^2}{\varepsilon^2}$$

$$d_6(\sigma) = \frac{(\sigma - 1) \cdot \sigma^2 \cdot (\sigma + \varepsilon)^2}{(1 + \varepsilon)^2}$$

Το παρεμβολικό πολυώνυμο αυτό με βάση όσα ειπώθηκαν πριν είναι $O(h^6)$ για τις μεθόδους που μας απασχολούν. Στο πρώτο βήμα εφαρμόζουμε ένα γνωστό παρεμβολικό πολυώνυμο διότι δεν έχουμε προηγούμενες πληροφορίες. Στα επόμενα βήματα η εφαρμογή του νέου πολυωνύμου βασίζεται στην τιμή του ε :

Το ίδιο πολυώνυμο κατάλληλα προσαρμοσμένο μπορεί να χρησιμοποιηθεί για παρεμβολή χρησιμοποιώντας τιμές από το επόμενο βήμα. Σ' αυτή τη περίπτωση μας ενδιαφέρει να έχουμε ενδιάμεσες τιμές στο $[x_{n-1}, x_n]$ χρησιμοποιώντας δεδομένα από

το διάστημα $[x_n, x_{n+1}]$. Το καινούργιο πολυώνυμο μπορεί να προκύψει λοιπόν από το παλιό ως εξής:

$$U_6^N(x_{n-1} + \sigma h_{n-1}) = U_6^N(x_n + \varepsilon \cdot (\sigma - 1) \cdot h_n) = U_6^P(x_n + \sigma^* h_n)$$

με

$$\sigma^* = \varepsilon \cdot (\sigma - 1) \quad 4-3-4$$

4.4 Ανάλυση σφάλματος για τις μεθόδους 5ης τάξης.

Στην συνέχεια κάνουμε ανάλυση του τοπικού σφάλματος αποκοπής του πολυωνύμου παρεμβολής στην περίπτωση που χρησιμοποιούνται οι τιμές του προηγούμενου βήματος και ο τύπος 4-3-2.

Για το λόγο αυτό λαμβάνουμε υπ' όψη και τα δύο επιτυχημένα βήματα $[x_{n-1}, x_n], [x_n, x_{n+1}]$ όπου είναι διαθέσιμες 12 τιμές της f , καθώς και το $f(x_{n+1}, y_{n+1}) = f_{\text{next}}^1$ που είναι η πρώτη τιμή της συνάρτησης f του επόμενου βήματος. Οι 13 αυτές τιμές της f εμφανίζονται και στις δύο μεθόδους αλλά έχουν παραχθεί από διαφορετικές παραμέτρους. Όλοι οι συντελεστές των μεθόδων και για τα δύο βήματα εκφράζονται συναρτήσει του αρχικού σημείου (x_n, y_n) και του βήματος h_n που μας ενδιαφέρει. Έτσι οι τιμές της f υπολογίζονται από τις σχέσεις,

$$f^i = f(x_n + (c_i - 1)\varepsilon h_n, y_n + h_n \varepsilon \sum_{j=1}^6 (a_{ij} - b_j) \cdot f^j) \quad i=1, 2, \dots, 6$$

$$f^7 = f(x_n, y_n)$$

$$f^i = f(x_n + c_{i-6} h_n, y_n + h_n \sum_{j=7}^{i-6} a_{i-6, j-6} f^j) \quad i=8, \dots, 12 \quad 4-4-1$$

$$f^{13} = f(x_n + h_n, y_n + h_n \sum_{i=1}^6 b_i f^{i+6})$$

Η έκφραση του νέου interpolant σε Runge-Kutta μορφή που είναι απαραίτητη για την ανάλυση σφάλματος είναι

$$U_6^1(x_n + \sigma h_n) = y_n + h_n \sum_{i=1}^{13} b_i^* f^i \quad 4-4-2$$

διότι $d_1 + d_3 + d_5 = 1$. Οι συντελεστές τώρα της μεθόδου 4-4-1 & 4-4-2 σε συμπαγή μορφή είναι

$$c_i^* = \varepsilon (c_i - 1), \quad i=1(1)6$$

$$c_i^* = c_{i-6}, \quad i=7(1)12$$

$$c_{13}^* = 1$$

$$a_{ij}^* = \varepsilon \cdot (-b_j + a_{ij}) \quad i=1(1)6, j=1(1)6$$

4-4-3

$$a_{7j}^* = 0 \quad \forall j$$

$$a_{ij}^* = a_{i-6, j-6} \quad i=8(1)12, j=7(1)i-1$$

$$a_{13j}^* = b_{j-6} \quad j=7(1)12$$

τα δε b^* γίνονται:

$$b_1^* = (-\varepsilon \cdot d_1 \cdot b_1 + d_4)$$

$$b_i^* = -\varepsilon \cdot d_1 \cdot b_i \quad i=2(1)6$$

4-4-4

$$b_7^* = d_3 \cdot b_1 + d_5$$

$$b_i^* = d_3 \cdot b_{i-6} \quad i=8(1)12$$

$$b_{13}^* = d_6$$

Η ανάλυση σφάλματος καθορίζει τις τιμές του ε για τις οποίες το φράγμα του τοπικού σφάλματος του πολυωνύμου δεν είναι χειρότερο από το αντίστοιχο φράγμα του τοπικού σφάλματος του y_{n+1} . Για τις δύο μεθόδους που αναφέρονται εδώ, οι αντίστοιχες νόρμες του πρωτεύοντα όρου του Τ.Σ.Α είναι:

$$\| T_6 \|_E = \frac{17}{26} \quad \text{για τη Fehlberg}$$

$$\| T_6 \|_E = \frac{4}{9} \quad \text{για τη Dormand-Prince}$$

Το διβηματικό interpolant 4-3-2 σε Runge-Kutta μορφή και

περιορισμένο στο τρέχον βήμα $[x_n, x_{n+1}]$ καθορίζεται από τις παραμέτρους 4-4-3 και 4-4-4. Μπορεί να προχωρήσει κανείς κατόπιν, στην εκτίμηση των φραγμάτων του $\Psi(x_n, y_n, h_n, \sigma)$ για διάφορες τιμές του ε .

Παρατηρούμε ότι γενικά $\max \|T_6^*\| = K$ για $\sigma \in (0, 1)$, και διάφορες ενδεικτικές τιμές του K δίνονται για ένα φάσμα τιμών του ε . Φαίνεται από τους πίνακες 4-4α και 4-4γ ότι για $\varepsilon \leq 2$ η νέα μέθοδος μπορεί να εφαρμοσθεί άφοβα καθότι το φράγμα του σφάλματος δεν υπερβαίνει το φράγμα της κυρίως μεθόδου. Ομως και για μεγαλύτερες τιμές του ε δεν υπάρχει πρόβλημα, Fine [1985], και μπορεί να είναι αποδεκτές. Επειδή παρουσιάζουμε μόνο τα αποτελέσματα για μερικές τιμές του ε , αναφέρουμε ότι για οποιαδήποτε ενδιάμεση τιμή του ε το μέγεθος του K παραμένει στα ίδια επίπεδα.

Η ανάλυση σφάλματος στη περίπτωση που χρησιμοποιούμε τιμές από το επόμενο βήμα γίνεται όπως και πριν, θεωρώντας πάλι τις ίδιες 13 τιμές της f και εκφράζοντας τα πάντα συναρτήσει των $x_{n-1}, y_{n-1}, h_{n-1}$. Πάλι διάφορες τιμές του K δίνονται για ένα φάσμα τιμών του ε . Είναι φανερό ότι εδώ το πρόβλημα δημιουργείται καθώς το ε μικραίνει, διότι τότε το άκρο του επόμενου βήματος απομακρύνεται. Πάντως για $\varepsilon > \frac{2}{3}$ το μέγεθος του φράγματος είναι αποδεκτό.

Η ορθή χρήση των διβηματικών interpolants απαιτεί την χρήση του U_6^D ή U_6^N ανάλογα με τους πίνακες των φραγμάτων. Γενικά η κακή περίπτωση είναι ένα μικρό βήμα ανάμεσα σε δύο μεγάλα.

Ακολουθούν πίνακες που δίνουν για ένα φάσμα τιμών του ε το μέγεθος του φράγματος

Πίνακας 4-4α

Ανάλυση σφάλματος για το U_6^P με την Fehlberg

ε	K	παρατηρήθηκε για σ
0.5	17/26	1
1.0	17/26	1
1.5	17/26	1
2.0	17/26	1
2.3	17/26	1
2.4	0.68967	0.498
2.5	0.75847	0.498

Πίνακας 4-4β

Ανάλυση σφάλματος για το U_6^N με την Fehlberg

ε	K	παρατηρήθηκε για σ
2.0	0.65385	1
1.33	0.65798	0.817
1.0	0.68662	0.702
0.83	0.73573	0.643
0.66	0.85446	0.588

Πίνακας 4-4γ

Ανάλυση σφάλματος για το U_6^D για την Dormand-Prince

ϵ	K	παρατηρήθηκε για σ
0.5	4/9	1
1.0	0.44504	0.92
1.5	0.45784	0.831
2.0	0.46612	0.81
2.1	0.468	0.51
2.5	0.67443	0.512

Πίνακας 4-4δ

Ανάλυση σφάλματος για το U_6^N με την Dormand-Prince

ϵ	K	παρατηρήθηκε για σ
2.0	0.44504	0.866
1.33	0.45736	0.721
1.0	0.49483	0.638
0.83	0.54484	0.595
0.66	0.65222	0.557

4.5 Μέθοδοι ανώτερης τάξης.**(iii) Runge-Kutta-Verner 5(6)**

Ο Verner [1978a] έχει προτείνει μιά 8-βαθμίδων embedded RK μέθοδο τάξης πέντε και έξι, που έχει χρησιμοποιηθεί στη subroutine DVERK που προτάθηκε από τους Hull et.al [1979]. Οι παράμετροι της, δοσμένες στη γνωστή μορφή πινάκων έχουν ως

εξής:

0								
$\frac{1}{6}$	$\frac{1}{6}$							
$\frac{4}{15}$	$\frac{4}{75}$	$\frac{16}{75}$						
$\frac{2}{3}$	$\frac{5}{6}$	$-\frac{8}{3}$	$\frac{5}{2}$					
$\frac{5}{6}$	$-\frac{165}{64}$	$\frac{55}{6}$	$-\frac{425}{64}$	$\frac{85}{96}$				
1	$\frac{12}{5}$	- 8	$\frac{4015}{612}$	$-\frac{11}{36}$	$\frac{88}{255}$			
$\frac{1}{15}$	$-\frac{8263}{15000}$	$\frac{124}{75}$	$-\frac{643}{680}$	$-\frac{81}{250}$	$\frac{2484}{10625}$	0		
1	$\frac{3501}{1720}$	$-\frac{300}{43}$	$\frac{297275}{52632}$	$-\frac{319}{2322}$	$\frac{24068}{84065}$	0	$\frac{3850}{26703}$	
5 ^η	$\frac{13}{160}$	0	$\frac{2375}{5984}$	$\frac{5}{16}$	$\frac{12}{85}$	$\frac{3}{44}$	0	0
6 ^η	$\frac{3}{40}$	0	$\frac{875}{2244}$	$\frac{23}{72}$	$\frac{264}{1955}$	0	$\frac{125}{11592}$	$\frac{43}{616}$

Πίνακας 4-5 μέθοδος Verner 6(5)

Για την 5^ης τάξης μέθοδο η νόρμα του πρωτεύοντα όρου του σφάλματος ικανοποιεί την σχέση,

$$\|T_6\|_E = \frac{1}{3}$$

ενώ για την 6^ης τάξης προσέγγιση έχουμε

$$\|T_7\|_E \cong 0.972931$$

Για τη μέθοδο αυτή προτείνουμε δύο παρεμβολικά πολυώνυμα. Το πρώτο είναι τάξης $O(h^6)$ και βασίζεται στις τιμές $y'_{n-1}, y_n, y'_n, \bar{y}_{n+1/2}, y_{n+1}, y'_{n+1}$ όπου $\bar{y}_{n+1/2}$ 5ης τάξης προσέγγιση του μέσου που δίδεται από τον τύπο

$$\bar{y}_{n+1/2} = y_n + h_n \cdot \left\{ \frac{121}{1536} \cdot f^1 + \frac{317125}{861696} \cdot f^3 + \frac{397}{6912} \cdot f^4 - \frac{1}{68} \cdot f^5 + \right. \\ \left. + \frac{17}{704} \cdot f^6 + \frac{125}{12096} \cdot f^7 - \frac{817}{14784} \cdot f^8 + \frac{1}{32} \cdot f^1_{\text{next}} \right\}$$

όπου f^1_{next} η πρώτη τιμή του επόμενου βήματος.

Αυτό το πολυώνυμο προσεγγίζει τη λύση σε ενδιάμεσα σημεία του διαστήματος $[x_n, x_{n+1}]$ και δεν χρειάζεται κανένα επί πλέον υπολογισμό. Αν $h_{n-1} = \epsilon h_n$ τότε σε Hermite μορφή το πολυώνυμο είναι.

$$U_6^V(x_n + \sigma h_n) = d_1 y_n + d_2 \bar{y}_{n+1/2} + d_3 y_{n+1} + h_n \cdot (d_4 y'_{n-1} + d_5 y'_n + d_6 y'_{n+1}) \quad 4-5-1$$

όπου V σημαίνει Verner και

$$d_1(\sigma) = \left\{ \epsilon (-1 + 26\sigma^3 - 47\sigma^4 + 22\sigma^5) + 3\epsilon^2 (-2 + 13\sigma^2 - 29\sigma^4 + 18\sigma^5) \right. \\ \left. + 2\epsilon^2 (-5 + 47\sigma^2 - 58\sigma^3 + 16\sigma^5) + 5\epsilon^4 (-1 - 11\sigma^2 - 18\sigma^3 + 8\sigma^4) \right\} / (16 \cdot \Delta)$$

$$d_2(\sigma) = \sigma^2 \left\{ -2\epsilon (1-\sigma)^2 + 3\epsilon^2 (-1 + 3\sigma^2 - 2\sigma^3) \right. \\ \left. - 4\epsilon^3 (2 - 3\sigma + \sigma^2) - 5\epsilon^4 (1-\sigma)^2 \right\} / \Delta$$

$$d_3(\sigma) = \sigma^2(\epsilon(6\sigma-17\sigma^2+10\sigma^3)+3\epsilon^2(3-19\sigma^2+14\sigma^3)+2\epsilon^3(17-38\sigma+16\sigma^3) \\ +5\epsilon^4(5-14\sigma+8\sigma^2))/(16\cdot\Delta)$$

4-5-2

$$d_4(\sigma) = \sigma^2(1-4\sigma+5\sigma^2-2\sigma^3)/(32\cdot\Delta)$$

$$d_5(\sigma) = \sigma\{2\epsilon(-1+11\sigma^2-18\sigma^3+8\sigma^4)+3\epsilon^2(-4+11\sigma-17\sigma^3+10\sigma^4) \\ +4\epsilon^3(-5+18\sigma-17\sigma^2+4\sigma^4)+10\epsilon^4(-1+4\sigma-5\sigma^2+2\sigma^3) \\ +\sigma(-1+4\sigma-5\sigma^2+2\sigma^3)\}/(32\cdot\Delta)$$

$$d_6(\sigma) = \sigma^2\{2\sigma\epsilon(-1+3\sigma-2\sigma^2)+\epsilon^2(-3+21\sigma^2-18\sigma^3)-4\epsilon^3(-3+7\sigma-4\sigma^3) \\ +10\epsilon^4(-1+3\sigma-2\sigma^2)\}/(32\cdot\Delta)$$

και

$$\Delta = -(\epsilon + 6\epsilon^2 + 10\epsilon^3 + 5\epsilon^4)/16$$

Φέρνοντας τώρα το παρεμβολικό πολυώνυμο σε RK μορφή και εκφράζοντας τα πάντα συναρτήσει των x_n, y_n, h_n προχωράμε στην ανάλυση σφάλματος. Αυτή χρειάζεται τις 16 τιμές της f των δύο βημάτων καθώς και την $y'_{n+1} = f(x_{n+1}, y_{n+1})$. Το αποτέλεσμα είναι πολύ ικανοποιητικό αφού και για $\epsilon=20$ το σφάλμα ικανοποιεί την σχέση,

$$\max_{\sigma \in (0,1)} \|T_6^*\|_{\epsilon} \cong \frac{1}{3}$$

Τα πολυώνυμα που προτάθηκαν στην προηγούμενη παράγραφο 4-4

για μεθόδους 5^{ης} τάξης, επίσης μπορούν να χρησιμοποιηθούν. Δεν είναι όμως τόσο αποτελεσματικά όσο το 4-5-1.

Το δεύτερο παρεμβολικό πολυώνυμο που προτείνουμε για τη μέθοδο του Verner είναι τάξης $O(h^7)$ και βασίζεται στις τιμές y'_{n-1} , $y'_{n-1/2}$, y_n , y'_n , $y'_{n+1/2}$, y_{n+1} , y'_{n+1} όπου $y'_{n+1/2} = f(x_{n+1/2}, \bar{y}_{n+1/2})$ είναι το μοναδικό επί πλέον f που πρέπει να υπολογισθεί σε κάθε βήμα. Σε μορφή Hermite το νέο πολυώνυμο είναι:

4-5-3

$$U_7^V(x_n + \sigma h_n) = d_1 y_n + d_2 y_{n+1} + h (d_3 y'_{n-1} + d_4 y'_{n-1/2} + d_5 y'_n + d_6 y'_{n+1/2} + d_7 y'_{n+1})$$

με d_i $i=1,2,\dots,7$ που δίδονται από τους παρακάτω τύπους.

$$\begin{aligned} d_1(\sigma) = & 3\epsilon^3 \{ (4(1+15\sigma^4 - 36\sigma^5 + 20\sigma^6) + 24\epsilon^2(1+5\sigma^3 - 21\sigma^5 + 15\sigma^6) \\ & + 5\epsilon^3(11+12\sigma^2 + 84\sigma^3 - 72\sigma^4 - 147\sigma^5 + 112\sigma^6) \\ & + 30\epsilon^4(2+9\sigma^2 + 10\sigma^3 - 45\sigma^4 + 12\sigma^5 + 12\sigma^6) \\ & + \epsilon^5(31+420\sigma^2 - 300\sigma^3 - 735\sigma^4 + 504\sigma^5 + 80\sigma^6) + 6\epsilon^6(1+45\sigma^2 - 70\sigma^3 + 24\sigma^4) \\ & + 60\epsilon^7\sigma^2(1-\sigma)^2 \} / (128 \cdot \Delta) \end{aligned}$$

$$\begin{aligned} d_2(\sigma) = & 3\epsilon^3 \sigma^2 \{ 4\sigma^2(-15+36\sigma - 20\sigma^2) + 24\epsilon\sigma(-5+21\sigma^2 - 15\sigma^3) \\ & + 5\epsilon^2(12-84\sigma + 147\sigma^2 + 72\sigma^3 - 112\sigma^4) + 30\epsilon^3(-9-10\sigma + 45\sigma^2 - 12\sigma^3 - 12\sigma^4) \\ & + \epsilon^4(-420+300\sigma + 735\sigma^2 - 504\sigma^3 - 80\sigma^4) + 6\epsilon^5(-45+70\sigma - 24\sigma^2) \\ & - 60\epsilon^6(1-\sigma) \} / (128 \cdot \Delta) \end{aligned}$$

$$\begin{aligned} d_3(\sigma) = & \epsilon\sigma^2 \{ 16\sigma(1-3\sigma+3\sigma^2 - \sigma^3) + 12\epsilon(1-5\sigma^2+6\sigma^3 - 2\sigma^4) + 8\epsilon^2(3-5\sigma+3\sigma^2 - \sigma^4) \\ & + 15\epsilon^3(1-\sigma)^2 + 3\epsilon^4(1-\sigma)^2 \} / (256 \cdot \Delta) \end{aligned}$$

$$d_4(\sigma) = \varepsilon\sigma^2(2\sigma(-1+3\sigma-3\sigma^2+\sigma^3)+3\varepsilon(-1+5\sigma^2-6\sigma^3+2\sigma^4)+4\varepsilon^2(-3+5\sigma-3\sigma^2+\sigma^3) - 15\varepsilon^3(1-\sigma)^2-6\varepsilon^4(1-\sigma)^2)/(16\cdot\Delta)$$

4-5-4

$$d_5(\sigma) = \varepsilon\sigma(16\sigma^2(1-3\sigma+3\sigma^2-\sigma^3)+36\varepsilon\sigma(1-5\sigma^2+6\sigma^3-2\sigma^4) + 8\varepsilon^2(3+21\sigma-35\sigma^2-9\sigma^3+24\sigma^4-4\sigma^5)+9\varepsilon^3(16+25\sigma-66\sigma^2+25\sigma^3-32\sigma^4+32\sigma^5) + \varepsilon^4(330+21\sigma-256\sigma^2-327\sigma^3-312\sigma^4+544\sigma^5)+45\varepsilon^5(8-\sigma+2\sigma^2-25\sigma^3+8\sigma^4 + 8\sigma^5)+\varepsilon^6(186+183\sigma-260\sigma^2-693\sigma^3+504\sigma^4+80\sigma^5) + 36\varepsilon^7(1+6\sigma-11\sigma^2+4\sigma^4)+60\varepsilon^8(1-\sigma)^2)/(256\cdot\Delta)$$

$$d_6(\sigma) = \varepsilon^3\sigma^2(24\sigma^2(1-\sigma)^2+48\varepsilon(1-3\sigma^2+2\sigma^3)+2\varepsilon^2(12+70\sigma-105\sigma^2-48\sigma^3+71\sigma^4) + 90\varepsilon^3(1+\sigma-4\sigma^2+\sigma^3+\sigma^4)+\varepsilon^4(123-80\sigma-189\sigma^2+126\sigma^3+20\sigma^4) + 36\varepsilon^5(2-3\sigma+\sigma^3)+15\varepsilon^6(1-\sigma)^2) / (16\cdot\Delta)$$

$$d_7(\sigma) = \varepsilon^3\sigma^2(24\sigma^2(2-5\sigma+3\sigma^2)+48\varepsilon(2-9\sigma^2+7\sigma^3)+2\varepsilon^2(24+175\sigma-315\sigma^2-156\sigma^3 + 272\sigma^4)+45\varepsilon^3(5+6\sigma-27\sigma^2+8\sigma^3+8\sigma^4)+\varepsilon(369-260\sigma-693\sigma^2+504\sigma^3+80\sigma^4) + 36\varepsilon^5(7-11\sigma+4\sigma^3)+60\varepsilon^6(1-\sigma)^2)/(256\cdot\Delta)$$

και

$$\Delta = \frac{3\cdot(4\varepsilon^3+24\varepsilon^4+55\varepsilon^5+60\varepsilon^6+31\varepsilon^7+6\varepsilon^8)}{128}$$

Το υπολογιστικό κόστος είναι κάπως μεγάλο αλλά σε περίπτωση που χρησιμοποιούνται σταθερές τιμές για ε ή σ μπορεί να ελαττωθεί. Επίσης σε περίπτωση που η συνάρτηση f δεν είναι στοιχειώδης η εύκολη στον υπολογισμό τότε δεν έχει νόημα να μιλάμε για κόστος υπολογισμού πολυωνύμων.

Η ανάλυση σφάλματος εδώ γίνεται κατα τόν ίδιο τρόπο με

τα προηγούμενα και τα αποτελέσματα δίνονται στον πίνακα 4-5α

Πίνακας 4-5α

Ανάλυση σφάλματος για την RK-Verner 6ης τάξης

ϵ	K	παρατηρήθηκε για σ
0.5	0.973	1
0.75	0.973	1
1.0	0.973	1
1.72	0.982	0.54
2.0	1.134	0.535

4.6 Αποτελέσματα - Συμπεράσματα

Για τα 4 προβλήματα του προηγούμενου κεφαλαίου δίνονται οι λόγοι του σφάλματος του πολυωνύμου παρεμβολής προς το σφάλμα της αντίστοιχης μεθόδου Runge-Kutta.

Πίνακας 4-6α (i) Μέθοδος RKF45, interpolant U_6^D

πρόβλημα	ανοχή	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}
A		1.040	1.034	1.029	1.021	1.016	1.011	1.008
B		0.909	0.894	1.287	1.389	1.258	1.008	0.926
C		1.318	1.281	1.463	1.347	1.539	2.212	1.496
D		1.004	1.000	1.000	1.000	1.000	1.000	1.000

Πίνακας 4-6β (iii) Μέθοδος RKF45, interpolant U_6^N

πρόβλημα	ανοχή	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}
A		1.209	1.169	1.131	1.097	1.070	1.050	1.035
B		1.039	0.935	1.071	1.296	1.234	1.015	0.945
C		1.110	1.091	1.209	1.189	1.420	1.871	1.461
D		1.001	1.000	1.000	1.000	1.000	1.000	1.000

Πίνακας 4-6γ(iii) Μέθοδος RKDP45, interpolant U_6^D

πρόβλημα	ανοχή	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}
A		1.587	1.068	1.057	1.045	1.034	1.025	1.018
B		12.657	10.909	19.715	15.950	11.921	8.259	5.695
C		1.649	1.648	1.619	1.621	1.473	1.574	1.627
D		1.026	1.009	1.001	1.004	1.000	1.000	1.000

Πίνακας 4-6δ (iv) Μέθοδος RKDP45, interpolant U_6^D

πρόβλημα	ανοχή	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}
A		0.876	1.239	1.195	1.150	1.111	1.080	1.057
B		4.679	3.823	5.964	9.962	8.838	6.879	5.107
C		1.725	1.412	1.673	1.591	1.447	1.538	1.657
D		1.024	1.006	0.999	1.003	1.000	1.000	1.000

Πίνακας 4-6ε (v) Μέθοδος RKV45 ,interpolant U_6^V

πρόβλημα	ανοχή	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}
A		0.956	7.204	13.478	14.220	14.795	16.310	15.854
B		0.973	0.960	0.984	1.035	1.159	1.236	1.301
C		1.677	1.340	3.777	1.162	1.480	1.633	1.232
D		1.056	1.054	1.001	1.000	1.000	1.091	1.000

Πίνακας 4-6ζ (vi) Μέθοδος RKV45 ,interpolant U_7^V

πρόβλημα	ανοχή	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}
A		7.343	7.898	7.163	7.075	7.038	7.535	7.075
B		0.926	0.946	0.962	0.968	1.041	1.076	1.083
C		1.658	2.421	5.214	2.391	3.221	2.672	2.773
D		1.048	1.050	1.001	1.000	1.000	1.001	1.001

Το σφάλμα σε σχέση με το σφάλμα των μονοβηματικών interpolant είναι στο ίδιο μέγεθος κάτι που άλλωστε αναμένει κανείς. Οι κύριες μέθοδοι που έχουν προταθεί μέχρι τώρα δεν είχαν θεωρητικά φράγματα μικρότερα από αυτά που φάνηκαν εδώ. Το δε κέρδος όπως φαίνεται από τον επόμενο πίνακα σε υπολογισμούς είναι σημαντικό

Πίνακας 4-6η

Μέθοδος	τάξη	επί πλέον f Tsitouras et.al	επί πλέον f Enright et.al	% μείωση σε f
RKF45	$O(h^6)$	0	2	25%
RKV56	$O(h^6)$	0	1	11%
RKV56	$O(h^7)$	1	3	18%

Επίσης και τα τοπικά σφάλματα αποκοπής που δίνονται από εμάς για $\epsilon=1$, που είναι η συνηθισμένη περίπτωση σε αντιπαράθεση με τα αντίστοιχα Τ.Σ.Α που δίνονται από τους Enright et.al [1986], φαίνονται στον επόμενο πίνακα.

Πίνακας 4-6θ

Μέθοδος	τάξη	φράγμα Tsitouras et.al	φράγμα Enright et.al
RKF45	$O(h^6)$	0.65	0.66
RKV56	$O(h^6)$	0.10	0.10
RKV56	$O(h^7)$	0.97	1.05

Ακόμα όμως, και για $\epsilon=1.5$ ή $\epsilon=2$, τα διβηματικά interpolants που χρησιμοποιούν τις προηγούμενες τιμές δεν φαίνεται να χάνουν την υπεροχή τους. Υπάρχει δηλαδή μια σαφής υπεροχή των νέων interpolants που βγαίνει από τη μείωση του υπολογιστικού κόστους από 11% έως 25%. Οι μέθοδοι βέβαια

που προτάθηκαν από τον Enright et.al [1986] δεν είναι οι βέλτιστες για μονοβηματική παρεμβολή.

Το κόστος σε υπολογισμούς της f δεν είναι δυνατόν να ελατωθεί σε σχέση με ότι φαίνεται στον πίνακα 4-6η. Μπορούν όμως να βελτιωθούν τα φράγματα των αντίστοιχων σφαλαμάτων. Αυτό μπορεί να γίνει με την κατασκευή ενός interpolant που να βασίζεται εκτός των άκρων σε δύο ενδιάμεσα σημεία έστω τα σ_1, σ_2 . Η κατασκευή ενός τέτοιου παρεμβολικού πολυωνύμου αναλύεται στην επόμενη παράγραφο.

4.7 Μονοβηματικό interpolant.

Αν κάποιος θα ήθελε να περιοριστεί σε ένα βήμα τότε μπορεί να χρησιμοποιήσει μεθόδους που έχουν προτείνει άλλοι ερευνητές. Εμείς εδώ με βάση τα αναφερθέντα στο προηγούμενο κεφάλαιο θα κατασκευάσουμε πολυώνυμο παρεμβολής για τη μέθοδο RKF5(4) που θα χρησιμοποιεί τις τιμές $y_n, y'_n, y_{n+\sigma_1}, y_{n+\sigma_2}, y_{n+1}, y'_{n+1}$ από το βήμα $[x_n, x_{n+1}]$ με δύο ενδιάμεσες τιμές $y_{n+\sigma_1}$ και $y_{n+\sigma_2}$. Σύμφωνα με την παράγραφο 3-5 η τιμή $y(x_{n+\sigma_1}h_n)$ υπολογίζεται με παραμέτρους τα a_{82}, a_{83}, c_8 το δε $y(x_{n+\sigma_2}h_n)$ υπολογίζεται με παραμέτρους τα $a_{92}, a_{93}, a_{98}, c_9, b_8^*(\sigma_2)$. Έτσι όμως το interpolant που στηρίζεται στα $y_n, y'_n, y_{n+\sigma_1}, y_{n+\sigma_2}, y_{n+1}, y'_{n+1}$ έχει 10 ελεύθερες παραμέτρους που δίνουν καλύτερες δυνατότητες ελαχιστοποίησης από τη μέθοδο των Enright et.al. [1986] που έχει μόνο το c_8 και το c_9 . Πράγματι για άπειρες επιλογές της ελεύθερης 10-άδας βρήκαμε αντίστοιχα πολυώνυμα που επιτυγχάνουν φράγμα

το ίδιο με το φράγμα του σφάλματος 6^{ης} τάξης της μεθόδου RKF45 στο άκρο που είναι και το ελάχιστο δυνατό. Οι Enright et.al[1986] δεν κατάφεραν κάτι τέτοιο.

Αναφέρουμε μια περίπτωση όπου τα σημεία σ_1 και σ_2 ισαπέχουν από τα άκρα. Θεωρούμε τις τιμές

$$\sigma_1 = \frac{1}{3}, \sigma_2 = \frac{2}{3}, a_{92} = a_{93} = a_{98} = a_{82} = a_{83} = 0$$

4-7-1

$$c_8 = \frac{1}{2-0}, c_9 = 0.425, b_8^*(\sigma_2) = 0.$$

Με την επιλογή αυτή υπολογίζουμε από τις 3-5-5 τα υπόλοιπα a_{8i}, a_{9i} $i=4,5,6,7$ και τα $b_i(\sigma_1), b_i(\sigma_2)$ που προσδιορίζουν τις προσεγγίσεις $y_{n+\sigma_1}$ και $y_{n+\sigma_2}$. Τέλος κατασκευάζουμε το πολυώνυμο τάξης $O(h^6)$ που παρεμβάλει τις έξι δοσμένες τιμές. Για τον υπολογισμό της τιμής $y_{n+\sigma_1}$ έχουμε τις τιμές.

$$b_1^*(1/3) = -292657/2121390, b_2^*(1/3) = 0, b_3^*(1/3) = 14825984/100766025,$$

$$b_4^*(1/3) = -2941783/221685255, b_5^*(1/3) = 433/130950,$$

$$b_6^*(1/3) = -12622/432135, b_7^*(1/3) = 28/4617, b_8^*(1/3) = 160000/447849$$

$$a_{84} = -266513/3840000,$$

$$a_{85} = 5761/192000,$$

$$a_{86} = 5891/240000$$

$$a_{87} = 1843/80000$$

Για τον υπολογισμό της τιμής $y_{n+\sigma_2}$ οι παράμετροι είναι:

$$b_1^*(2/3)=397493/3494826, \quad b_2^*(2/3)=0, \quad b_3^*(2/3)=6565888/9764955,$$

$$b_4^*(2/3)=7511543/42965802, \quad b_5^*(2/3)=-422/6345, \quad b_6^*(2/3)=-3104/41877$$

$$b_7^*(2/3)=-218/5589, \quad b_9^*(2/3)=-512000/4465611$$

$$a_{94}=-8788507/675840000, \quad a_{95}=1760129/15360000$$

$$a_{96}=11091769/42240000, \quad a_{97}=-18377/128000$$

Αφού λοιπόν υπολογίσουμε τις ενδιάμεσες προσεγγίσεις το πολυώνυμο παρεμβολής παίρνει την παρακάτω μορφή:

$$u(x_n+sh_n)=d_1 y_n + d_2 h y_n' + d_3 y_{n+\sigma_1} + d_4 y_{n+\sigma_2} + d_5 y_{n+1} + d_6 h y_{n+1}'$$

4-7-2

όπου

$$d_1(\sigma)=(117 \cdot \sigma^5 - 333 \cdot \sigma^4 + 323 \cdot \sigma^3 - 111 \cdot \sigma^2 + 4)/4$$

$$d_2(\sigma)=\sigma \cdot (9 \cdot \sigma^4 - 27 \cdot \sigma^3 + 29 \cdot \sigma^2 - 13 \cdot \sigma + 2)/2$$

$$d_3(\sigma)=(-81 \cdot \sigma^2 \cdot (3 \cdot \sigma^3 - 8 \cdot \sigma^2 + 7 \cdot \sigma - 2))/4$$

4-7-3

$$d_4(\sigma)=81 \cdot \sigma^2 \cdot (\sigma - 1) \cdot (3 \cdot \sigma^2 - 4 \cdot \sigma + 1)/4$$

$$d_5(\sigma)=(-\sigma^2 \cdot (117 \cdot \sigma^3 - 252 \cdot \sigma^2 + 161 \cdot \sigma - 30))/4$$

$$d_6(\sigma)=\sigma^2 \cdot (9 \cdot \sigma^3 - 18 \cdot \sigma^2 + 11 \cdot \sigma - 2)/2$$

Εκφράζοντας το πολυώνυμο 4-7-2 σε RK μορφή εφαρμόζουμε τη γνωστή διαδικασία για την ανάλυση σφάλματος που μας εξασφαλίζει για την μέθοδο που κατασκευάστηκε ότι :

$$\max_{\sigma \in (0,1)} \| T_6^* \|_E = \frac{17}{26} \quad 4-7-4$$

δηλαδή το ίδιο φράγμα με την αρχική μέθοδο και σαφώς μικρότερο από το φράγμα που πέτυχαν οι Enright et.al [1986]. Το φράγμα 4-7-4 είναι και το καλύτερο που μπορεί να πετύχουμε αν θέλουμε C^1 παρεμβολή

Άλλοι ερευνητές τελευταία ,Calvo[1990], Verner [1990a],[1990b], [1990c] προσπάθησαν να βελτιώσουν τις μεθόδους παρεμβολής για τις πιο δημοφιλείς μεθόδους που εμφανίζονται στη παρούσα διατριβή. Εμείς νομίζουμε ότι το διβηματικό interpolation που πρωτοπαρουσιάζουμε όμως είναι υπερβολικά φθηνό σε σχέση με όλες αυτές και το συνιστούμε.

153

106

SCALED RUNGE-KUTTA-NYSTROM ΜΕΘΟΔΟΙ

5.1 Εισαγωγή

Η αριθμητική επίλυση συνήθων διαφορικών εξισώσεων 2^{ης} τάξης της μορφής

$$y'' = f(x, y) \quad , \quad x \in [x_0, x_{\text{end}}]$$

5-1-1

$$y(x_0) = y_0 \quad , \quad y'(x_0) = y'_0$$

με $y \in \mathbb{R}^m$, $f : [x_0, x_{\text{end}}] \times \mathbb{R}^m \longrightarrow \mathbb{R}^m$, αποτελεί μια από τις πιο ενδιαφέρουσες περιοχές εφαρμογής για τις μεθόδους προβλημάτων αρχικών τιμών. Η Ουράνιος Μηχανική και η Δυναμική, για παράδειγμα, καταλήγουν σε προβλήματα της μορφής 5-1-1 και η ανάπτυξη ικανών και αξιόπιστων αριθμητικών μεθόδων αποτελεί ουσιαστικό παράγοντα για την ανάπτυξη αυτών των περιοχών.

Αριθμητικές μέθοδοι για την προσέγγιση του προβλήματος 5-1-1 εφαρμόζονται για περισσότερο από πενήντα χρόνια, και συνεχώς αναπτύσσονται νέες μέθοδοι ή βελτιώνονται οι ήδη υπάρχουσες. Μια ευρεία κλάση άμεσων μεθόδων είναι τύπου Runge-Kutta και έχουν προταθεί αρχικά από τον Nystrom [1925]. Εκτοτε αναφέρονται σαν μέθοδοι Runge-Kutta-Nystrom (RKN). Στην συνέχεια διάφοροι ερευνητές έχουν προτείνει μεθόδους αυτής

της κατηγορίας . Οι πιο γνωστές οφείλονται στους Fehlberg [1972],[1973], Bettis [1971], Filippi & Graff [1986],[1987], Fehlberg et.al. [1980], Dormand et.al. [1987a],[1987b]

Όπως και για τις μεθόδους Runge-Kutta (RK) έτσι και για τις μεθόδους (RKN) το βασικό μειονέκτημα είναι ότι δεν έχουν δυνατότητα να μας δώσουν μια συνεχή προσέγγιση της λύσης $y(x)$ και της πρώτης παραγώγου της $y'(x)$, χωρίς μια ουσιαστική αύξηση του υπολογιστικού κόστους.

Στο κεφάλαιο αυτό θα αναπτύξουμε Scaled Runge-Kutta-Nystrom (SRKN) μεθόδους οι οποίες προσεγγίζουν την λύση του προβλήματος 5-1-1 καθώς και την παράγωγό της σε οποιοδήποτε σημείο $x^* = x + \sigma \cdot h$, όπου $\sigma \in (0,1)$, αφού πρώτα αυτές έχουν υπολογιστεί στο σημείο $x_{n+1} = x_n + h$. Οι νέοι αυτοί αλγόριθμοι έχουν σχεδιαστεί να χρησιμοποιηθούν με υπάρχοντες (RKN) μεθόδους, παίρνοντας τούς υπολογισμούς της συνάρτησης (function evaluations) που απαιτούν αυτές οι μέθοδοι σαν τον πυρήνα του νέου συστήματος.

5.2 Scaled Runge-Kutta-Nystrom μέθοδοι. Γενική θεωρία.

Θεωρούμε το πρόβλημα 2^{ης} τάξης της μορφής 5-1-1. Το σύστημα μπορεί να λυθεί εφαρμόζοντας συνήθως μια embedded μέθοδο Runge-Kutta-Nystrom τάξης q και p ($q > p$) της μορφής.

$$\bar{y}_{n+1} = \bar{y}_n + h \bar{y}'_n + h^2 \sum_{i=1}^s b_i f^i \quad 5-2-1a$$

$$\bar{y}'_{n+1} = \bar{y}'_n + h_n \sum_{i=1}^s b_i' f^i \quad 5-2-1b$$

και

$$\bar{y}_{n+1} = \bar{y}_n + h_n \bar{y}'_n + h_n^2 \sum_{i=1}^s b_i f^i \quad 5-2-2a$$

$$\bar{y}'_{n+1} = \bar{y}'_n + h_n \sum_{i=1}^s b_i' f^i \quad 5-2-2b$$

αντίστοιχα , όπου

$$f^i = f(x_n + c_i h_n, y_n + c_i h_n y'_n + h_n^2 \sum_{j=1}^{i-1} a_{ij} f^j), \quad i=1,2,\dots,s \quad 5-2-3$$

με $x_{n+1} = x_n + h_n$, $h_n = \theta(x_n)h$, $0 \leq \theta(x) \leq 1$. Οι \bar{y}_n , \bar{y}'_n , y_n , y'_n δηλώνουν προσεγγίσεις της θεωρητικής λύσης $y(x)$ και της $y'(x)$. Τέλος θεωρούμε ότι προχωρούμε με τις προσεγγίσεις τάξης q . Οπως ήδη αναφέραμε στη εισαγωγή μια (RKN) μέθοδος δεν είναι κατάλληλη για πυκνή έξοδο προσεγγίσεων ή για προσεγγίσεις προβλημάτων που παρουσιάζουν ασυνέχειες. Οι Dormand & Prince [1987] ο Fine [1985],[1987] και οι Parageorgiou & Tsitouras [1987],[1988] πρώτοι ασχολήθηκαν με το πρόβλημα της κατασκευής μεθόδων (SRKN). Ομως οι άλλοι ερευνητές περιορίστηκαν στην επέκταση των μεθόδων που οι ίδιοι είχαν πριν παράγει. Ο Fine στην ουσία ασχολήθηκε μόνο με το γενικό πρόβλημα δεύτερης τάξης .

Εμείς εδώ αναπτύσσουμε (SRKN) μεθόδους μέχρι και 7^{ης} τάξης

που προσεγγίζουν την λύση και την παράγωγό της με το ελάχιστο δυνατόν κόστος σε οποιοδήποτε ενδιαμέσο σημείο ενός επιτυχούς βήματος $[x_n, x_{n+1}]$.

Οι λύσεις σε ενδιάμεσα σημεία μπορούν να υπολογιστούν απο ένα νέο ζεύγος μεθόδων (SRKN) οι οποίες χρησιμοποιούνται για την προσέγγιση απο το σημείο x_n με ένα βήμα μήκους sh_n . Η γενική μορφή της Scaled μεθόδου είναι

$$F1: \quad y_{n+\sigma}^* = \bar{y}_n + \sigma h_n \bar{y}'_n + h_n^2 \sum_{i=1}^{s_1} b_i^* f^{i*} \quad 5-2-4a$$

$$F2: \quad y_{n+\sigma}^{**} = \bar{y}'_n + h_n \sum_{i=1}^{s_2} b_i^{**} f^{i**} \quad 5-2-4b$$

με

$$f^{1*} = f(x_n, y_n) \quad 5-2-5a$$

$$f^{i*} = f(x_n + c_i^* h_n^*, y_n + c_i^* h_n^* \bar{y}'_n + h_n^{*2} \sum a_{ij}^* f^{*j}) \quad 5-2-5b$$

και με αντίστοιχες εκφράσεις για τα f^{i**} όπου $s_1^* \geq s$ $s_2^* \geq s$ και $h^* = \sigma h$, $\sigma \in (0,1)$.

Αν θεωρήσουμε ότι

$$c_i^* = \frac{c_i}{\sigma}, \quad a_{ij}^* = \frac{a_{ij}}{\sigma^2} \quad 5-2-6$$

τότε $f^i = f^{i*} = f^{i**}$ για $i=1,2,\dots,s$ στις 5-2-4a και 5-2-4b. Έτσι μεγάλο μέρος των υπολογισμών αποφεύγεται αφού τα f^i

$i=1,2,\dots,s$ έχουν ήδη υπολογισθεί κατά την εφαρμογή του βασικού αλγόριθμου (RKN) Μόνο τα f^{i*} $i=s+1,\dots,s_1^*$ και f^{i**} $i=s+1,\dots,s_2^*$ μένουν να υπολογισθούν. Είναι προφανές όμως ότι αν $s_m = \min\{s_1^*, s_2^*\}$ τότε πρέπει κανείς να επιδιώξει τα s_m πρώτα f να είναι ίδια και να χρησιμοποιούνται για scaling στο y αλλά και στο y' .

Γενικά, για χαμηλής τάξης μέθοδο δεν υπάρχει ανάγκη επί πλέον υπολογισμών της f για να κατασκευασθεί SRKN αλγόριθμος. Για μεθόδους υψηλότερης τάξης η κατάσταση είναι πίο πολύπλοκη και λεπτομέρειες θα φανούν στον πίνακα 5-2β. Εδώ σημειώνουμε ότι οι εξισώσεις είναι δύσκολο να λυθούν χωρίς να γίνουν απλοποιητικές παραδοχές που τις λιγοστεύουν. Η πίο αναγκαία παραδοχή σύμφωνα με τον Fehlberg [1972] είναι:

$$(A_i): \quad P_{ji} = \sum_{k=2}^{j-1} a_{jk} \cdot c_k^i = \frac{c_i^{j+1}}{(j+1)(j+2)} \quad 5-2-7$$

με $j=3,4,\dots,s^* = \max\{s_1^*, s_2^*\}$, $i=1,2,\dots,j-1$

Οι καταστατικές εξισώσεις που πρέπει να ικανοποιούνται για ένα SRKN αλγόριθμο φαίνονται στον πίνακα 5-2α. Οι εξισώσεις του πίνακα ισχύουν μέχρι και για 9^η τάξη. Οι εξισώσεις στο αριστερό μέρος είναι για τη προσέγγιση του $y_{n+\sigma}^*$ ενώ στο δεξιό για το $y'_{n+\sigma}$. Οι εξισώσεις E1, E2 και E1', E2' είναι οι εξισώσεις τύπου Vandermode για τις δύο περιπτώσεις. Οι εξισώσεις E3, E3' εμφανίζονται σε υψηλής τάξης αλγόριθμους ενώ οι εξισώσεις E4, E5 και E4', E5' χρησιμοποιούνται σε συνδιασμό

με τις απλοποιητικές παραδοχές Α1 για να μικρύνουν τον αριθμό των εξισώσεων. Όλες οι εξισώσεις μέχρι και 10^{ης} τάξης βρίσκονται στο παράρτημα καθώς και στον Fehlberg [1972]
Πίνακας 5-2α

$$(E1) \quad \sum_{i=1}^{s^*} c_i^j \cdot b_i^* = \frac{\sigma^{j+2}}{(j+1) \cdot (j+2)} \quad \sum_{i=1}^{s^*} c_i^j \cdot b_i^{*'} = \frac{\sigma^{j+1}}{j+1} \quad (E1')$$

$$(E2) \quad \sum_{i=2}^{s^*} p_{ij} \cdot b_i^* = \frac{\sigma^5}{z_j} \quad \sum_{i=2}^{s^*} p_{ij} \cdot b_i^{*'} = \frac{\sigma^4}{g_j} \quad (E2')$$

με

$$z_1=120, z_2=360, z_{i+2} = 2 \cdot z_{i+1} - z_i + \frac{(i+4)(i+5)}{2}$$

και

$$g_i = z_i / (i+4), \quad i=0,1,2,\dots,s^*$$

$$(E3) \quad \sum_{i=2}^{s^*} p_{ij} \cdot c_i \cdot b_i^* = \frac{\sigma^9}{2160} \quad \sum_{i=2}^{s^*} p_{ij} \cdot c_i \cdot b_i^{*'} = \frac{\sigma^8}{240} \quad (E3')$$

$$(E4) \quad \sum_{i=2}^{s^*} a_{ik} c_i^j \cdot b_j^* = 0 \quad \sum_{i=2}^{s^*} a_{ik} c_i^j \cdot b_j^{*'} = 0 \quad (E4')$$

με $(k,j) \in \{ (2,0), (2,1), (2,2), (3,0), (4,0) \}$

$$(E5) \quad \sum_{i=5}^{s^*} b_i^* \cdot \sum_{k=4}^{i-1} a_{ik} a_{k2} = 0 \quad \sum_{i=5}^{s^*} b_i^{*'} \cdot \sum_{k=4}^{i-1} a_{ik} a_{k2} = 0 \quad (E5')$$

Στον πίνακα 5-2b δίνουμε για κάθε μέθοδο που αργότερα θα μας απασχολήσει τις εξισώσεις τύπου A,E,E' που πρέπει να ικανοποιεί. Επίσης δίνεται ο αριθμός των f που απαιτούνται για την κατασκευή του scaled αλγόριθμου. Ακόμα έγινε προσπάθεια να καταχωρηθούν για κάθε τάξη οι εξισώσεις που χρειάζονται, χωρίς όμως αυτό να είναι πάντα δυνατό. Π.χ. για το 5^{ης} τάξης scaling της παραγώγου της μεθόδου του Bettis [1973] η E2' απαλείφεται και έτσι δεν χρειαζόμαστε επί πλέον f .

Σημειώνουμε ακόμα ότι για τις μεθόδους του Fehlberg [1972] το scaling της παραγώγου για την υψηλότερη τάξη είναι άχρηστο διότι, και πολλά επί πλέον f χρειάζονται και οι ίδιες οι μέθοδοι δεν έχουν τέτοια προσέγγιση για την παράγωγο. Στα επόμενα παρουσιάζουμε τις scaled μεθόδους για τη προσέγγιση της λύσης και της παραγώγου της για τις πιο γνωστές RKN μεθόδους διαφόρων τάξεων

Ακολουθεί ο πίνακας 5-2β

Πίνακας 5-2β

Μέθοδος	Τάξη	A_i	Εξισώσεις για την $y(x)$	f.e	Εξισώσεις για την $y'(x)$	f.e
RKNF45	4	-	$E1, j=0,1,2$	3	$E1', j=0,1,2,3$ $E2, j=1$	5
RKNF45	5	-	$E1, j=0,1,2,3$ $E2, j=1$	5	$E1', j=0,1,2,3$ $E2', j=1.V1, V2, V3(*)$	8
RKNF56	5	1	$E1, j=0,1,2,3$	5	$E1', j=0,1,2,3,4$ $E2', j=2$	7
RKNF56	6	1	$E1, j=0,1,2,3,4$ $E2, j=2$	7	-----	--
RKNF67	6	1,2	$E1, j=0,1,2,3,4$ $E4, (k,j)=(1,0)$	7	$E1', j=0,1,2,3,4,5$ $E2', j=3.E4', k=1, j=0$	9
RKNF67	7	1,2	$E1, j=0,1,2,3,4,5$ $E2, j=3.E4, k=1, j=0$	9	-----	--
RKNF78	7	1,2 3	$E1, j=0,1,2,3,4,5$ $E4, (k,j) \in$ $\{(1,0), (1,1), (2,0)\}$	11	$E1', j=0(1)6 E4, j=4$ $E4', (k,j) \in$ $\{(1,0), (1,1), (2,0)\}$	
RKNF78	8	1,2 3	$E1, j=0(1)6 E2, j=4$ $E4, (k,j) \in$ $\{(1,0), (1,1), (2,0)\}$	13	-----	
RKNB45	4	1,2	$E1, j=0,1,2$	4	$E1', j=0,1,2,3$ $E2', j=1$	6
RKNB45	5	1,2	$E1, j=0,1,2,3 E2, j=1$	6	$E1', j=0,1,2,3,4$ $E2', j=2$	7

5.3 Scaled Runge-Kutta-Nystrom μεθόδων χαμηλής τάξης.

(i) Fehlberg 4(5)

Ο Fehlberg [1972] πρότεινε την RKNF4(5) μέθοδο που αποτελείται από ένα ζευγάρι μεθόδων με τάξεις 4 και 5 οι οποίες βασίζονται στο ίδιο σύνολο τιμών της f . Στη μέθοδο αυτή η παράγωγος υπολογίζεται μια φορά και η προσέγγισή της είναι 4^{ης} τάξης. Η εκτίμηση δηλαδή του σφάλματος γίνεται μόνο για τη y .

i	c_i	a_{ij}				
1	0					
2	$\frac{1}{3}$	$\frac{1}{18}$				
3	$\frac{2}{3}$	0	$\frac{2}{9}$			
4	1	$\frac{1}{3}$	0	$\frac{1}{6}$		
5	1	$\frac{13}{120}$	$\frac{3}{10}$	$\frac{3}{40}$	$\frac{1}{6}$	
	b	$\frac{13}{120}$	$\frac{3}{10}$	$\frac{3}{40}$	$\frac{1}{60}$	4 ^η τάξη
	\bar{b}	$\frac{13}{120}$	$\frac{3}{10}$	$\frac{3}{40}$	0	$\frac{1}{60}$ 5 ^η τάξη
	b'	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	παράγωγος

Πίνακας 5-3a μέθοδος Fehlberg 45

Για την παραπάνω μέθοδο κατασκευάσαμε 5^{ης} τάξης scaled μέθοδο $y_{n+\sigma}^*$ για την προσέγγιση της λύσης και 4^{ης} τάξης scaled μέθοδο $y_{n+\sigma}^{*}$ για την προσέγγιση της παραγώγου. Αυτά τα πετύχαμε χρησιμοποιώντας μόνο τους 5 υπολογισμούς f της κύριας μεθόδου. Από τους πίνακες βρίσκουμε τις αντίστοιχες εξισώσεις και αφού τις λύσουμε προκύπτουν τα b_i^* $i=1,2,3,4,5$ σαν συναρτήσεις του σ , και έχουμε:

$$b_1^* = \frac{1}{2} + \left(-\frac{11}{12}\sigma + \frac{3}{4}\sigma^2 - \frac{9}{40}\sigma^3 \right) \sigma^2$$

$$b_2^* = \left(-\frac{3}{2}\sigma - \frac{15}{8}\sigma^2 - \frac{27}{40}\sigma^3 \right) \sigma^2$$

$$b_3^* = \left(-\frac{3}{4}\sigma + \frac{3}{2}\sigma^2 - \frac{27}{40}\sigma^3 \right) \sigma^2$$

5-3-1

$$b_4^* = \left(-\frac{1}{2}\sigma + \frac{7}{8}\sigma^2 - \frac{3}{8}\sigma^3 \right) \sigma^2$$

$$b_5^* = \left(\frac{2}{3}\sigma - \frac{5}{4}\sigma^2 + \frac{3}{5}\sigma^3 \right) \sigma^2$$

Για να καθοριστεί η ποιότητα του scaling υπολογίζουμε κατα τα γνωστά το συντελεστή του h^6 στο σφάλμα αποκοπής της μεθόδου. Επειδή όμως σε αυτόν συμμετέχουν 3 στοιχειώδη διαφορικά αναφερόμαστε και στα τρία. Οι τύποι σαν συναρτήσεις του σ είναι:

$$T_{61}^* = \frac{1}{24} \cdot \sum_{i=2}^5 b_i^* c_i^4 - \frac{\sigma^6}{720}$$

$$T_{62}^* = \frac{1}{2} \sum_{i=2}^5 b_i^* P_{i2} - \frac{\sigma^6}{720} \quad 5-3-2$$

$$T_{63}^* = \sum_{i=2}^5 b_i^* c_i P_{i1} - \frac{\sigma^6}{720}$$

Για την παράγωγο ενεργώντας με τον ίδιο τρόπο λύνουμε τις εξισώσεις του πίνακα 5-2α και έχουμε

$$b_1^{*'} = 1 + \sigma \left(-\frac{11}{4} \sigma + 3 \sigma^2 - \frac{9}{8} \sigma^3 \right)$$

$$b_2^{*'} = \sigma \left(\frac{9}{8} \sigma - \frac{15}{8} \sigma^2 + \frac{27}{8} \sigma^3 \right)$$

$$b_3^{*'} = \sigma \left(-\frac{9}{4} \sigma + 6 \sigma^2 - \frac{27}{8} \sigma^3 \right) \quad 5-3-3$$

$$b_4^{*'} = \sigma \left(-\frac{3}{2} \sigma + \frac{7}{2} \sigma^2 - \frac{15}{8} \sigma^3 \right)$$

$$b_5^{*'} = \sigma \left(2 \sigma - 5 \sigma^2 + 3 \sigma^3 \right)$$

Στην περίπτωση αυτή μας ενδιαφέρουν οι συντελεστές του σφάλματος αποκοπής 5ης τάξης για τη παράγωγο που είναι:

$$T_{51}^{*'} = \frac{1}{24} \sum_{i=1}^5 b_i^{*'} c_i^4 - \frac{\sigma^5}{120}$$

$$T_{52}^{*'} = \sum_{i=2}^5 b_i^{*'} c_i P_{i1} - \frac{\sigma^5}{30} \quad 5-3-4$$

$$T_{53}^{*'} = \sum_{i=2}^5 b_i^* P_{i2} - \frac{\sigma^5}{120}$$

Για τη μέθοδο αυτή ισχύουν

$$\|T_6\|_E = \frac{1}{6} \quad \text{και} \quad \|T_5'\|_E = \frac{1}{9}$$

ενώ για τις scaled μεθόδους έχουμε

$$\max_{\sigma \in (0,1)} \|T_6^*\|_E \cong 0.18846 \quad \text{και} \quad \max_{\sigma \in (0,1)} \|T^{*'}\|_E = \frac{1}{9}.$$

Δηλαδή υπάρχει μόνο μια μικρή υπέρβαση στο μέγεθος του πρωτεύοντα όρου του Τ.Σ.Α. για την scaled προσέγγιση της λύσης. Η υπέρβαση αυτή είναι σε φυσιολογικά πλαίσια αφού άλλοι ερευνητές έχουν προτείνει μεθόδους παρεμβολής σε Runge-Kutta μεθόδους με πολύ μεγαλύτερες υπερβάσεις.

Αφού πάντως έχουμε επιτύχει την κατασκευή της μεθόδου χωρίς επί πλέον κόστος έχουμε επιτύχει τον πρώτο αντικειμενικό στόχο του ερευνητή που ασχολείται με την Αριθμητική Ανάλυση. Το να μικρύνουμε το τοπικό σφάλμα αποκοπής θα απαιτούσε την χρήση μιας ακόμα βαθμίδας ,δηλαδή αύξηση του υπολογιστικού κόστους κατά 20%. Η μείωση όμως του σφάλματος αποκοπής ελαττώνοντας απλά κατά, π.χ. 10% το μέγεθος του βήματος ολοκλήρωσης θά είναι πολύ μεγαλύτερη.

(ii) Runge-Kutta-Nystrom-Bettis (4)5

Ο Bettis [1973], πρότεινε μια μέθοδο 5^{ης} τάξης που παράγει προσεγγίσεις 4^{ης} και 5^{ης} τάξης για τη λύση και την παράγωγό της. Έτσι έχουμε στη διάθεσή μας εκτιμήσεις του σφάλματος και για τις δύο προσεγγίσεις. Οι συντελεστές της μεθόδου βρίσκονται στον πίνακα 5-3β

i	c _i	a _{ij}				
1	0					
2	$\frac{1}{8}$	$\frac{1}{128}$				
3	$\frac{1}{4}$	$\frac{1}{96}$	$\frac{1}{48}$			
4	$\frac{1}{2}$	$\frac{1}{24}$	0	$\frac{1}{12}$		
5	$\frac{3}{4}$	$\frac{9}{128}$	0	$\frac{9}{64}$	$\frac{9}{128}$	
6	1	$\frac{7}{90}$	0	$\frac{4}{15}$	$\frac{1}{15}$	$\frac{4}{45}$
τάξη 4	b	$\frac{1}{6}$	0	0	$\frac{1}{3}$	
τάξη 4	b'	0	0	$\frac{2}{3}$	$-\frac{1}{3}$	$\frac{2}{3}$
τάξη 5	b	$\frac{7}{90}$	0	$\frac{4}{15}$	$\frac{1}{15}$	$\frac{4}{45}$
τάξη 5	b'	$\frac{7}{90}$	0	$\frac{16}{45}$	$\frac{2}{15}$	$\frac{16}{45} - \frac{7}{90}$

Πίνακας 5-3β

Για την παραπάνω μέθοδο κατασκευάσαμε δύο scaled προσεγγίσεις $y_{n+\sigma}^*$ και $y_{n+\sigma}^{* \prime}$ 5ης τάξης χωρίς επί πλέον κόστος σε υπολογισμούς της f παρ' όλο που ο πίνακας 5-2β μας λέει το αντίθετο. Για τη μέθοδο αυτή οι εξισώσεις E2 και E2' δεν χρειάζονται διότι εκτός της A1 ισχύει και η A2. Έτσι για να προσεγγίσουμε το $y(x_n+h^*)$ χρειαζόμαστε μόνο 5 υπολογισμούς της f . Είναι καλύτερα όμως να χρησιμοποιήσουμε τη τιμή f_6 . Τότε μας δίνεται η ευκαιρία να παράγουμε 6ης τάξης scaled αλγόριθμο λύνοντας την E1 $j=4$ μαζί με τις άλλες εξισώσεις. Οι scaled συντελεστές της μεθόδου δίνονται από τις σχέσεις:

$$b_1^* = \sigma^2 \left(\frac{1}{2} - \frac{25}{18} \sigma + \frac{35}{18} \sigma^2 - \frac{4}{3} \sigma^3 + \frac{16}{45} \sigma^4 \right)$$

$$b_2^* = 0$$

$$b_3^* = \sigma^2 \left(\frac{8}{3} \sigma - \frac{52}{9} \sigma^2 + \frac{24}{5} \sigma^3 - \frac{64}{45} \sigma^4 \right)$$

5-3-5

$$b_4^* = \sigma^2 \left(-2 \sigma + \frac{19}{3} \sigma^2 - \frac{32}{5} \sigma^3 + \frac{32}{15} \sigma^4 \right)$$

$$b_5^* = \sigma^2 \left(\frac{8}{9} \sigma - \frac{28}{9} \sigma^2 + \frac{56}{15} \sigma^3 - \frac{64}{45} \sigma^4 \right)$$

$$b_6^* = \sigma^2 \left(\frac{1}{6} \sigma + \frac{11}{18} \sigma^2 - \frac{4}{5} \sigma^3 + \frac{16}{45} \sigma^4 \right)$$

Για τη μέθοδο αυτή όλοι οι 6ης τάξης συντελεστές του τοπικού σφάλματος αποκοπής είναι μηδέν και δεν έχει νόημα να παρουσιάσουμε τους συντελεστές του Τ.Σ.Α 7ης τάξης.

Η 5^{ης} τάξης scaled προσέγγιση της παραγώγου προκύπτει βάζοντας $b_2^* = 0$ και απαλοίφοντας την εξίσωση Ε2'. Οι συντελεστές αυτής της προσέγγισης δίνονται από τις σχέσεις:

$$b_1^{*'} = \sigma \left(\frac{32}{15} \sigma^4 - \frac{20}{3} \sigma^3 + \frac{70}{9} \sigma^2 - \frac{25}{6} \sigma + 1 \right)$$

$$b_2^{*'} = 0$$

$$b_3^{*'} = \sigma \left(\frac{128}{15} \sigma^4 - 24 \sigma^3 + \frac{208}{9} \sigma^2 + 8 \sigma \right)$$

5-3-6

$$b_4^{*'} = \sigma \left(\frac{64}{5} \sigma^4 - 32 \sigma^3 + \frac{76}{3} \sigma^2 - 6 \sigma \right)$$

$$b_5^{*'} = \sigma \left(\frac{128}{15} \sigma^4 + \frac{56}{3} \sigma^3 - \frac{112}{9} \sigma^2 + \frac{8}{3} \sigma \right)$$

$$b_6^{*'} = \sigma \left(\frac{32}{15} \sigma^4 - 4 \sigma^3 + \frac{22}{9} \sigma^2 - \frac{1}{2} \sigma \right)$$

Οι συντελεστές του τοπικού σφάλματος αποκοπής 6^{ης} τάξης για την παράγωγο δίνονται από τους παρακάτω τύπους.

$$T_{61}^{*'} = \frac{1}{120} \sum_{i=1}^6 b_i^{*'} c_i^5 = \frac{\sigma^6}{720}$$

$$T_{62}^{*'} = \frac{1}{6} \sum_{i=1}^6 b_i^{*'} P_{i3} = \frac{\sigma^6}{720}$$

5-3-7

$$T_{63}^{*'} = \frac{1}{2} \sum_{i=1}^6 b_i^{*'} c_i P_{i2} = \frac{\sigma^6}{144}$$

$$T_{64}^{*'} = \frac{1}{2} \sum_{i=1}^6 b_i^{*'} c_i^2 P_{i1} = \frac{\sigma^6}{72}$$

$$T_{65}^{*'} = \sum_{i=1}^6 b_i^{*'} \sum_{j=1}^{i-1} a_{ij} P_{j1} = \frac{\sigma^6}{720}$$

Στην περίπτωση αυτή έχουμε

$$\max_{\sigma \in (0,1)} \| T_6^{*'} \|_E = \| T_6' \| \cong 0.04375 .$$

Δηλαδή η scaled μέθοδος δίνει πολύ ικανοποιητικό αποτέλεσμα σε αυτήν την περίπτωση.

5.4 SRKN μέθοδοι υψηλής τάξης.

(i) Runge-Kutta-Fehlberg 5(6)

Ο Fehlberg [1972] πρότεινε επίσης ένα ζευγάρι μεθόδων τάξης πέμπτης και έκτης με 5^{ης} τάξης προσέγγιση για την παράγωγο. Οι συντελεστές της μεθόδου δίνονται στον παρακάτω πίνακα 5-4α

Για τη μέθοδο αυτή έχουμε κατασκευάσει μια 6^{ης} τάξης scaled προσέγγιση $y_{n+\sigma}^*$ της λύσης και μια 5^{ης} τάξης προσέγγιση της παραγώγου στα ενδιάμεσα σημεία. Οι συντελεστές b_i^* είναι:

$$b_1^* = \sigma^2 \left(\frac{1}{2} \sigma^4 - \frac{37}{20} \sigma^3 + \frac{125}{48} \sigma^2 - \frac{41}{24} \sigma + \frac{1}{2} \right)$$

$$b_2^* = 0$$

i	c _i	a _{ij}					
1	0						
2	$\frac{1}{12}$	$\frac{1}{288}$					
3	$\frac{1}{6}$	$\frac{1}{216}$	$\frac{1}{108}$				
4	$\frac{1}{2}$	0	0	$\frac{1}{8}$			
5	$\frac{4}{5}$	$\frac{16}{125}$	0	$\frac{4}{125}$	$\frac{4}{45}$		
6	1	$\frac{-247}{1152}$	0	$\frac{12}{19}$	$\frac{7}{432}$	$\frac{4375}{65664}$	
7	1	$\frac{11}{240}$	0	$\frac{108}{475}$	$\frac{8}{45}$	$\frac{125}{2736}$	$\frac{1}{300}$
b _i		$\frac{11}{240}$	0	$\frac{108}{475}$	$\frac{8}{45}$	$\frac{125}{2736}$	$\frac{1}{300}$
B _i		$\frac{11}{240}$	0	$\frac{108}{475}$	$\frac{8}{45}$	$\frac{125}{2736}$	0 $\frac{1}{300}$
b' _i		$\frac{1}{24}$	0	$\frac{27}{95}$	$\frac{1}{3}$	$\frac{125}{456}$	$\frac{1}{15}$

Πίνακας 5-4α

$$b_3^* = \sigma^2 \left(-\frac{324}{285} \sigma^4 + \frac{1863}{475} \sigma^3 - \frac{5049}{1045} \sigma^2 + \frac{1944}{855} \sigma \right)$$

$$b_4^* = \sigma^2 \left(\frac{4}{3} \sigma^4 - \frac{59}{15} \sigma^3 + \frac{11}{3} \sigma^2 - \frac{8}{9} \sigma \right)$$

5-4-1

$$b_5^* = \sigma^2 \left(-\frac{125}{114} \sigma^4 + \frac{625}{228} \sigma^3 - \frac{625}{304} \sigma^2 + \frac{625}{1368} \sigma \right)$$

$$b_6^* = \sigma^2 \left(-\frac{84}{165} \sigma + \frac{73}{55} \sigma - \frac{35}{33} \sigma^2 + \frac{8}{33} \sigma \right)$$

$$b_7^* = \sigma^2 \left(\frac{10}{11} \sigma^4 - \frac{607}{275} \sigma^3 + \frac{1107}{660} \sigma^2 - \frac{62}{165} \sigma \right)$$

Οι συντελεστές του Τ.Σ.Α 7^{ης} τάξης ικανοποιούν τις σχέσεις:

$$\|T_7\|_E = 0.05 \quad \max_{\sigma \in (0,1)} \|T_7^*\|_E \cong 0.0517$$

Δηλαδή το σφάλμα της scaled μεθόδου είναι στο ίδιο επίπεδο με το σφάλμα της αρχικής μεθόδου.

Για τη scaled λύση της παραγώγου 5^{ης} τάξης έχουμε τούς παρακάτω συντελεστές.

$$b_i^{*'} = db_i^*/d\sigma \quad i=1,2,\dots,6$$

για το δε σφάλμα ισχύουν οι σχέσεις

$$\|T_6^{\prime}\|_E = 0.041666 \quad \|T_6^{*'}\|_E = 0.021$$

υπάρχει δηλαδή μια διαφορά σε βάρος της scaled προσέγγισης η οποία όμως δεν είναι σημαντική.

(ii) Runge-Kutta-Nystrom 6(7)

Πάλι ο Fehlberg [1972] πρότεινε μέθοδο 6^{ης} τάξης με 7^{ης} τάξης συνοδό για την προσέγγιση του σφάλματος και 6^{ης} τάξης προσέγγιση της παραγώγου. Η μέθοδος έχει 8 βαθμίδες αλλά όπως σε όλες τις μεθόδους του Fehlberg η 8^η βαθμίδα είναι και η πρώτη του επομένου βήματος. Έτσι στην ουσία είναι

σαν να έχουμε μέθοδο με εφτά βαθμίδες. Οι συντελεστές της μεθόδου είναι στον πίνακα 5-4γ.

Για τη μέθοδο αυτή έχουμε παράγει μια 6^{ης} τάξης scaled λύση χωρίς επί πλέον κόστος και μια 7^{ης} τάξης λύση με ένα ακόμα υπολογισμό της f . Για τη πρώτη περίπτωση βλέποντας τις εξισώσεις που πρέπει να ικανοποιηθούν βλέπουμε ότι πρέπει ένα b^* να παραμείνει ελεύθερο. Προτείνουμε $b_6^* = 0$ καθώς αν λύσουμε ταυτόχρονα την $E1, j=5$ θα έχουμε πολυώνυμα υψηλότερης τάξης ενώ η $E4$ με $(k,j)=(1,1)$ ανεβάζει τους συντελεστές του Τ.Σ.Α. Επίσης με $b_1^* = 0$ έχουμε τους συντελεστές $b_i^* \quad i=1,2,\dots,8$ που δίνονται στον πίνακα 5-4δ.

Χρησιμοποιώντας τώρα ένα επί πλέον f μπορούμε να κατασκευάσουμε scaled προσέγγιση για την λύση 7^{ης} τάξης. Με την ίδια τιμή της f μπορούμε να κατασκευάσουμε scaled προσέγγιση για την παράγωγο 6^{ης} τάξης. Το επί πλέον f καθορίζεται από τις εξής παραμέτρους,

$$c_8 = 1 \quad a_{92} = a_{93} = a_{94} = a_{95} = a_{97} = 0 \quad 5-4-2$$

$$a_{91} = \frac{11}{48} \quad a_{96} = \frac{25}{48} \quad a_{98} = -\frac{1}{4}$$

Οι εκφράσεις των $b_i^* \quad i=1,2,\dots,9$ δίνονται στον πίνακα 5-4ε. Από εκεί υπολογίζουμε τις εκφράσεις $b_i^{*'} = db_i(\sigma)/d\sigma \quad i=1,2,\dots,9$. Για τον πρωτευοντα όρο των Τ.Σ.Α και για τις διάφορες τάξεις ισχύουν οι τιμές που παρουσιάζονται στον πίνακα 5-4β.

Πίνακας 5-4β

$$\|T_7\|_E = 0.02505$$

$$\max_{\sigma \in (0,1)} \|T_7^*\|_E = 0.02597$$

$$\|T_8\|_E = 0.037$$

$$\max_{\sigma \in (0,1)} \|T_8^*\|_E = 0.0894$$

$$\|T_7'\|_E = 0.0658$$

$$\max_{\sigma \in (0,1)} \|T_7'^*\|_E = 0.0658$$

Πίνακας 5-4δ συντελεστές 6ης τάξης.

$$b_1^* = ((325/504)\sigma^4 - (45/21)\sigma^3 + (2175/1008)\sigma^2 - (103/65)\sigma + (1/2))\sigma^2$$

$$b_2^* = 0$$

$$b_3^* = ((-2375/1008)\sigma^4 + (7125/1008)\sigma^3 - (14725/2016)\sigma^2 + (1425/504)\sigma)\sigma^2$$

$$b_4^* = ((125/42)\sigma^4 - (1325/168)\sigma^3 + (1625/252)\sigma^2 - (125/84)\sigma)\sigma^2$$

$$b_5^* = ((-125/504)\sigma^4 + (75/28)\sigma^3 - (1425/1008)\sigma^2 + (25/252)\sigma)\sigma^2$$

$$b_6^* = 0$$

$$b_7^* = ((25/112)\sigma^4 - (155/336)\sigma^3 + (695/21016)\sigma^2 - (17/168)\sigma)\sigma^2$$

Πίνακας 5-4ε συντελεστές 7ης τάξης.

$$b_1^* = \sigma^2 ((-3125/5040)\sigma^5 + (125/48)\sigma^4 - (425/96)\sigma^3 + (125/32)\sigma^2 - (137/72)\sigma + 1/2)$$

$$b_2^* = 0$$

$$b_3^* = \sigma^2 ((3125/1008)\sigma^5 - (875/72)\sigma^4 + (1775/96)\sigma^3 - (1925/144)\sigma^2 + (25/6)\sigma)$$

$$b_4^* = \sigma^2 ((-3125/504)\sigma^5 + (1625/72)\sigma^4 - (1475/48)\sigma^3 + (2675/144)\sigma^2 - (25/6)\sigma)$$

$$b_5^* = \sigma^2 ((3125/504)\sigma^5 - (125/6)\sigma^4 + (1225/48)\sigma^3 - (325/24)\sigma^2 + (25/9)\sigma)$$

$$b_6^* = \sigma^2 ((-3125/1008)\sigma^5 + (1375/144)\sigma^4 - (1025/96)\sigma^3 + (1525/288)\sigma^2 - (25/24)\sigma)$$

$$b_7^* = \sigma^2 ((-2057225/1525104)\sigma^5 + (905065/217872)\sigma^4 - (4643239/1016736)\sigma^3 +$$

$$(6505733/3050208)\sigma^2 - (94373/254184)\sigma$$

$$b_8^* = \sigma^2 \left((1477751/762552)\sigma^5 - (6319387/1089360)\sigma^4 + (2004353/317730)\sigma^3 + (-45473359/15251040)\sigma^2 + (688669/1270920)\sigma \right)$$

$$b_9^* = \sigma^2 \left((19/612)\sigma^5 - (91/1020)\sigma^4 + (83/1020)\sigma^3 - (59/3060)\sigma^2 - (1/255)\sigma \right)$$

Πίνακας 5-4γ Μέθοδος RKNF 6(7)

i	ζ_i	α_{ij}							
0	0								
1	$\frac{1}{10}$	$\frac{1}{200}$							
2	$\frac{1}{5}$	$\frac{1}{150}$	$\frac{1}{75}$						
3	$\frac{2}{5}$	$\frac{2}{75}$	0	$\frac{4}{75}$					
4	$\frac{3}{5}$	$\frac{9}{200}$	0	$\frac{9}{100}$	$\frac{9}{200}$				
5	$\frac{4}{5}$	$\frac{199}{3600}$	$-\frac{19}{150}$	$\frac{47}{120}$	$-\frac{119}{1200}$	$\frac{89}{900}$			
6	1	$-\frac{179}{1824}$	$\frac{17}{38}$	0	$-\frac{37}{152}$	$\frac{219}{456}$	$-\frac{157}{1824}$		
7	1	$\frac{61}{1008}$	0	$\frac{475}{2016}$	$\frac{25}{504}$	$\frac{125}{1008}$	$\frac{25}{1008}$	$\frac{11}{2016}$	
	b_i	$\frac{61}{1008}$	0	$\frac{475}{2016}$	$\frac{25}{504}$	$\frac{125}{1008}$	$\frac{25}{1008}$	$\frac{11}{2016}$	
	\bar{b}_i	$\frac{61}{1008}$	0	$\frac{475}{2016}$	$\frac{25}{504}$	$\frac{125}{1008}$	$\frac{25}{1008}$	0 $\frac{11}{2016}$	
	b'_i	$\frac{19}{288}$	0	$\frac{25}{96}$	$\frac{25}{144}$	$\frac{25}{144}$	$\frac{25}{96}$	$\frac{19}{288}$	

5.5 Αριθμητικά αποτελέσματα.

Σε αυτή την παράγραφο παρουσιάζουμε δύο προβλήματα αρχικών τιμών και τα αποτελέσματα των μεθόδων που προτείναμε όταν εφαρμόζονται σε αυτά. Τα προβλήματα είναι,

Πρόβλημα 1^ο

$$y'' = -4x^2y - \frac{2 \cdot z}{(y^2 + z^2)^{1/2}} \quad z'' = -4x^2 - \frac{2 \cdot y}{(y^2 + z^2)^{1/2}}$$

$$x \in [(\pi/2)^{1/2}, 10]$$

με αρχικές τιμές

$$x_0 = (\pi/2)^{1/2} \quad y(x_0) = 0 \quad z(x_0) = 1$$

$$y'(x_0) = - (2\pi)^{1/2} \quad z'(x_0) = 0$$

και ακριβή λύση

$$y(x) = \cos(x^2), \quad z(x) = \sin(x^2)$$

Πρόβλημα 2^ο

$$y'' = -y / r^3 \quad z'' = -z / r^3 \quad x \in [0, 10]$$

όπου

$$r = (y^2 + z^2)^{1/2} \text{ και αρχικές συνθήκες}$$

$$x_0 = 0 \quad y(x_0) = 1 - e^{-1} \quad z(x_0) = 0$$

$$y'(x_0)=0 \quad z'(x_0)=\left\{\frac{1+e^{-1}}{1-e^{-1}}\right\}^{1/2}$$

και ακριβή λύση

$$y=\cos u - e^{-1}, \quad y' = -\sin u / (1-e^{-1} \cos u)$$

$$z=(1-e^{-2})^{1/2} \sin u, \quad z' = (1-e^{-2})^{1/2} \cos u / (1-e^{-1} \cos u)$$

όπου $x=u-e^{-1} \sin u$.

Για κάθε μέθοδο και τον αντίστοιχό της scaled αλγόριθμο έχουμε υπολογίσει το μέγιστο ολικό σφάλμα ϵ^* σε 9 ενδιάμεσα σημεία

$$\xi_{ni} = x_n + \frac{i}{10} \cdot h_n, \quad i=1,2,\dots,9. \quad n=0,1,\dots$$

κάθε βήματος $[x_n, x_{n+1}]$ που χρησιμοποιεί η μέθοδος. Επίσης το μέγιστο σφάλμα ϵ που εμφανίζεται στους κόμβους της αρχικής μεθόδου υπολογίζεται. Κατόπιν υπολογίζουμε τον λόγο $g=\epsilon^*/\epsilon$ για κάθε συντεταγμένη και την παράγωγό της. Οι λόγοι αυτοί θα συμβολίζονται $g_y, \dot{g}_y, g_z, \dot{g}_z$ και αποτελούν μια καλή ένδειξη για το πόσο επιτυχημένα προσεγγίζει ο scaled αλγόριθμος την λύση. Ειδικά για την μέθοδο RKF 6(7) εμφανίζονται και οι λόγοι g_y^* και g_z^* που αφορούν την χαμηλής τάξης προσέγγιση της λύσης.

Πίνακας 5-5

Πρόβλημα 1 RKNF45

TOL	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}
g_y	1.047	1.001	1.000	1.001	1.000
g_z	1.012	1.001	1.006	1.000	1.000
\dot{g}_y	1.000	0.999	1.002	1.001	1.002
\dot{g}_z	1.009	1.000	1.000	1.000	1.001

Πρόβλημα 2 RKNF45

TOL	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}
g_y	0.911	0.943	0.987	0.977	0.993
g_z	1.007	1.010	1.002	1.002	1.000
\dot{g}_y	1.004	0.985	0.998	0.999	0.999
\dot{g}_z	1.000	0.998	0.999	1.000	0.999

Πρόβλημα 1 RKNB45

TOL	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}
g_y	1.034	1.010	1.006	1.000	1.000
g_z	1.023	1.000	1.001	1.002	1.000
\dot{g}_y	1.022	0.999	1.001	1.001	1.000
\dot{g}_z	1.026	1.010	1.003	1.000	1.000

Πρόβλημα 2 RKNB45

TOL	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}
g_y	0.973	0.972	0.976	0.990	0.999
g_z	1.015	1.001	0.999	0.999	1.000
\dot{g}_y	1.035	1.007	1.007	1.004	1.003
\dot{g}_z	1.003	1.008	1.013	1.004	1.002

Πρόβλημα 1 RKNF56

TOL	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}
g_y	1.003	1.000	1.002	1.004	0.999
g_z	1.026	1.003	1.002	0.999	1.004
\dot{g}_y	1.028	1.005	1.003	1.005	1.005
\dot{g}_z	1.018	1.002	1.004	1.006	1.000

Πρόβλημα 2 RKNF56

TOL	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}
g_y	1.007	1.008	0.941	0.965	0.977
g_z	1.075	1.000	1.004	1.009	1.003
\dot{g}_y	1.075	1.006	1.027	1.031	1.014
\dot{g}_z	1.075	1.044	1.019	1.000	1.004

Πρόβλημα 1 RKNF67

TOL	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}
g_y	1.084	0.998	1.000	0.999	1.009
g_z	1.107	1.090	1.013	0.999	1.013
\dot{g}_y	1.151	1.080	1.006	0.999	1.013
\dot{g}_z	1.072	0.995	0.999	0.998	1.006
g_y^*	1.086	0.998	1.000	0.999	1.009
g_z^*	1.113	1.092	1.015	0.999	1.013

Πρόβλημα 2 RKNF67

TOL	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}
g_y	1.002	1.049	0.969	0.931	0.986
g_z	1.013	1.001	1.081	1.025	1.006
\dot{g}_y	1.045	0.992	1.097	1.100	1.024
\dot{g}_z	1.066	1.032	1.016	1.005	0.994
g_y^*	1.002	1.057	0.971	0.926	0.984
g_z^*	1.022	1.001	1.039	1.012	1.001

Κ Ε Φ Α Λ Α Ι Ο 6^{ΟΝ}

ΠΑΡΕΜΒΟΛΙΚΑ ΠΟΛΥΩΝΥΜΑ ΓΙΑ RKN ΜΕΘΟΔΟΥΣ ΠΟΥ ΑΝΗΚΟΥΝ ΣΤΗΝ ΚΛΑΣΗ C^2

6.1 C^2 scaled Runge-Kutta-Nystrom αλγόριθμοι.

Στα μέχρι τώρα αναφερθέντα μας απασχόλησε μόνο η τάξη του υπο κατασκευή (SRKN) αλγόριθμου. Πολλοί ερευνητές όμως θεωρούν σημαντικό το να έχει η νέα μέθοδος καλές ιδιότητες όσον αφορά τη συνέχεια, Champagne [1986], Dormand & Prince [1987]. Στο προηγούμενο κεφάλαιο καμιά από τις μεθόδους που προτάθηκαν δεν ήταν συνεχής με την έννοια ότι οι προσεγγίσεις της λύσης ή της παραγώγου στους κόμβους ενός διαστήματος ολοκλήρωσης $[x_n, x_{n+1}]$ είναι ασυνεχείς. Στο κεφάλαιο αυτό θα ασχοληθούμε αρχικά με την κατασκευή scaled Runge-Kutta-Nystrom μεθόδων που ανήκουν στην κλάση C^2 και στη συνέχεια με την κατασκευή πολυωνύμων παρεμβολής για RKN μεθόδους με την ίδια ιδιότητα.

Ειδικά το scaling στις Runge-Kutta-Nystrom μεθόδους μπορεί να παράγει παρεμβολικά πολυώνυμα που είναι C^2 όσον αφορά την προσέγγιση της λύσης και C^1 όσον αφορά την προσέγγιση της παραγώγου.

Θεωρούμε ότι έχουμε παράγει μια scaled Runge-Kutta-Nystrom μέθοδο της μορφής

$$y_{n+1}^* = y_n + h\alpha y_n' + h^2 \sum_{i=1}^{s^*} b_i^* f^i \quad 6-1-1$$

Τα b_i^* είναι πολυώνυμα του σ . Για να είναι η μέθοδος C^0 πρέπει για $\sigma=1$ να προκύπτουν οι συντελεστές b_i $i=1,2,\dots,s$ της αρχικής μεθόδου. Επίσης για $\sigma=0$ πρέπει τα b_i^* $i=1,\dots,s^*$ να μηδενίζονται. Επομένως για C^0 προσέγγιση πρέπει να ισχύουν οι σχέσεις:

$$b_i^*(0) = 0, i=1,2,\dots,s^* \quad 6-1-2\alpha$$

$$b_i^*(1) = b_i, i=1,2,\dots,s^* \quad 6-1-2\beta$$

Γίνεται εύκολα κατανοητό ότι οι scaled Runge-Kutta-Nystrom μέθοδοι για τα ζεύγη του Fehlberg που παρουσιάστηκαν στο προηγούμενο κεφάλαιο δεν είναι C^0 .

Ο scaled αλγόριθμος για την παράγωγο δίνεται από τη σχέση

$$y_{n+\sigma}^* = y_n' + h \sum_{i=1}^{s^*} b_i^{*\prime} f^i \quad 6-1-3$$

όπου $b_i^{*\prime}(\sigma)$ είναι πολυώνυμα του σ και

$$b_i^{*\prime}(\sigma) = \frac{db_i^*(\sigma)}{d\sigma}$$

Για να είναι η 6-1-1 C^1 πρέπει να ισχύουν οι σχέσεις

$$b_i^{*'}(0) = 0 \quad i=1,2,\dots,s^* \quad 6-1-4a$$

$$b_i^{*'}(1) = b_i' \quad i=1,2,\dots,s^* \quad 6-1-4b$$

όπου b_i' είναι οι συντελεστές με τους οποίους γίνεται η προσέγγιση της παραγώγου.

Αν παραγωγίσουμε την 5-1-3 ως προς σ προκύπτει η προσέγγιση της 2^{ης} παραγώγου η οποία χρησιμοποιείται για να παράγουμε λύση που είναι C^2 . Πράγματι έχουμε

$$y_{n+1}^{*''} = \sum_{s=1}^{s^*} b_i^{*''} f^i \quad 6-1-5$$

όπου

$$b_i^{*''} = d^2 b_i^*(\sigma) / d\sigma^2$$

οπότε οι απαραίτητες συνθήκες για μια C^2 προσέγγιση είναι

$$\frac{d^2 b_i^*(\sigma)}{d\sigma^2} \Big|_{\sigma=0} = \begin{cases} 0 & \text{αν } i \neq 1 \\ 1 & \text{αν } i=1 \end{cases} \quad 6-1-6a$$

$$\frac{d^2 b_i^*(\sigma)}{d\sigma^2} \Big|_{\sigma=1} = \begin{cases} 0 & \text{αν } i \neq s \\ 1 & \text{αν } i=s \end{cases} \quad 6-1-6b$$

Στις μεθόδους που έχουν αναπτυχθεί στο προηγούμενο κεφάλαιο το $f^1=y''(x_n)$ ενώ $f^s=y''(x_{n+1})$. Αν η μέθοδος δεν διαθέτει

την τελευταία τιμή τότε πρέπει να χρησιμοποιήσουμε το f^1 του επομένου βήματος. Στην περίπτωση του scaling για την προσέγγιση της παραγώγου μπορούμε να έχουμε λύση που είναι το πολύ C^1 . Θεωρούμε την scaled προσέγγιση της παραγώγου.

$$y_{n+1}^{*'} = y_n' + h \sum_{s=1}^s b_i^{*'} f^i \quad 6-1-6$$

Σημειώνουμε ότι οι παράμετροι $b_i^*(\sigma)$ που χρησιμεύουν για την C^2 προσέγγιση του $y(x)$ και οι παράμετροι $b_i^{*'}(\sigma)$ που χρησιμοποιούνται για την προσέγγιση της $y'(x)$ δεν έχουν γενικά καμιά σχέση στη γενική περίπτωση. Δηλαδή $db_i^*(\sigma)/d\sigma \neq b_i^*(\sigma)$. Οι συντελεστές που μας εξασφαλίζουν C^1 προσέγγιση είναι οι παρακάτω:

$$b_i^{*'}(0) = 0 \quad \forall i, \quad b_i^{*'}(1) = b_i' \quad 6-1-7$$

που εξασφαλίζουν το C^0 ενώ με

$$db_i^{*'}(\sigma) / d\sigma \Big|_{\sigma=0} = \begin{cases} 1 & \text{αν } i=1 \\ 0 & \text{αν } i \neq 1 \end{cases} \quad 6-1-8a$$

$$db_i^{*'}(\sigma) / d\sigma \Big|_{\sigma=1} = \begin{cases} 1 & \text{αν } i=s \\ 0 & \text{αν } i \neq s \end{cases} \quad 6-1-9b$$

εξασφαλίζεται το C^1 .

6.2 Κατασκευή scaled Runge-Kutta-Nystrom-Fehlberg 4(5) μεθόδου.

Εδώ μπορεί να κατασκευαστεί μια μέθοδος 4^{ης} τάξης για την λυση στο $x_n + sh_n$ η οποία χαρακτηρίζεται από τις παρακάτω παράμετρους.

$$b_1^* = -9 \cdot \sigma^5 / 40 + 3 \cdot \sigma^4 / 4 - 11 \cdot \sigma^3 / 12 + \sigma^2 / 2$$

$$b_2^* = 27 \cdot \sigma^5 / 40 - 15 \cdot \sigma^4 / 8 + 3 \cdot \sigma^3 / 2$$

$$b_3^* = -27 \cdot \sigma^5 / 40 + 3 \cdot \sigma^4 / 2 - 3 \cdot \sigma^3 / 4$$

6-2-1

$$b_4^* = -11 \cdot \sigma^5 / 40 + 5 \cdot \sigma^4 / 8 - \sigma^3 / 3$$

$$b_5^* = \sigma^5 / 2 - \sigma^4 + \sigma^3 / 2$$

Παρατηρούμε ότι ισχύουν όλες οι σχέσεις που αφορούν τη συνέχεια καθώς και οι τρεις εξισώσεις που εξασφαλίζουν την τάξη. Για την παράγωγο μας εξυπηρετούν οι σχέσεις του προηγούμενου κεφαλαίου για να έχουμε προσέγγιση C^1 . Εδώ είναι μια χαρακτηριστική περίπτωση που $\partial b_i^*(\sigma) / \partial \sigma \neq b_i^*$, το οποίο πρέπει να περιμένει κανείς όταν οι τάξεις των δύο προσεγγίσεων συμπίπτουν.

6.3 Παρεμβολή για Runge-Kutta-Nystrom μεθόδους.

Γενικά το scaling και σε αυτές τις μεθόδους είναι χρήσιμο μάλλον στό να μας βρεί τιμές στις οποίες θα στηρίξουμε

κάποιο πολυώνυμο παρεμβολής. Επίσης με τελείως φυσικό τρόπο μας εξασφαλίζει τη συνεχή διαφορισιμότητα των παραγωμένων μεθόδων. Αυτό γίνεται διότι αν π.χ. θέλουμε να παρεμβάλουμε πολυώνυμο σε ένα βήμα είναι διαθέσιμες οι τιμές $y_n, y_n', y_n'', y_{n+1}, y_{n+1}', y_{n+1}''$. Είναι φανερό ότι μια μέθοδος που θα στηρίζεται στα σημεία αυτά είναι C^2 αναγκαστικά η δέ τάξη της είναι το $\min\{5, \text{τάξη μεθόδου}\}$. Παραγωγίζοντας απλά την μέθοδο αυτή ως προς σ έχει μια μέθοδο παρεμβολής για την παράγωγο τάξης $\min\{4, \text{τάξη μεθόδου}\}$ που είναι C^1 . Τελικά ανακαλύπτει κανείς ότι στην περίπτωση της μεθόδου του Fehlberg τα πολυώνυμα παρεμβολής δεν είναι άλλα από αυτά που χαρακτηρίζουν οι 5-1-10 και 4-2-5.

Από ότι φαίνεται μέχρι τώρα για μια μέθοδο 7^{ης} τάξης απαιτείται ένα ακόμα f και αυτό είναι δυνατό να το πετύχουμε με scaling σχετικά εύκολα. Οι μέθοδοι παρεμβολής αποδίδουν για μεθόδους 8^{ης} τάξης και πάνω. Συγκεκριμένα μπορούμε με 3 επί πλέον f να κατασκευάσουμε interpolant ογδόνης τάξης για τη προσέγγιση της λύσης και εβδόμης για τη προσέγγιση της παραγώγου. Αυτό γίνεται αφού πρώτα με scaling κατασκευάσουμε τρεις ενδιάμεσες λύσεις 6^{ης} τάξης

$\bar{y}_{n+s_1}, \bar{y}_{n+s_2}, \bar{y}_{n+s_3}$ τις χρησιμοποιούμε για την κατασκευή των 6^{ης} τάξης προσεγγίσεων των δευτέρων παραγώγων

$f_{s_q} = f(x_{n+s_q}, \bar{y}_{n+s_q}), q=1,2,3$. Αυτές μαζί με τις έξι τιμές που έχουμε στα άκρα συμπληρώνουν τον απαραίτητο αριθμό των εννιά

τιμών που χρειάζονται για να φτειαζουμε interpolant $O(h^9)$. Υπενθυμίζω ότι τα επί πλέον f πολλαπλασιάζονται στον τελικό τύπο με h^2 και έτσι εξηγείται η χαμηλή τους ακρίβεια.

Μιά από τις πιο γνωστές μεθόδους ψηλής τάξης είναι η μέθοδος των Dormand-Prince που έχει ένα ζευγάρι μεθόδων τάξεων οκτώ και έξι, για τη προσέγγιση της y και όμοια για την προσέγγιση της y'' . Οι συντελεστές που την χαρακτηρίζουν δίνονται στον πίνακα 6-3.

Με τη μέθοδο του scaling βρίσκω πως

$$\bar{y}_{n+\sigma_k} = y_n + h\sigma_k y'_n + h^2 \sigma_k^2 \sum B_i(\sigma_k) f^i \quad 5-3-1$$

όπου τα $B_i(\sigma_k)$ δίνονται από τους τύπους

$$B_1 = (63 + 4 \cdot \sigma_k \cdot (-88 + 5 \cdot \sigma_k \cdot (43 + 4 \cdot \sigma_k \cdot (-12 + 5 \cdot \sigma_k)))) / 126$$

$$B_3 = 25 \cdot \sigma_k \cdot (21 + \sigma_k \cdot (-71 + 10 \cdot \sigma_k \cdot (9 - 4 \cdot \sigma_k))) / 144$$

$$B_4 = 25 \cdot \sigma_k \cdot (-7 + \sigma_k \cdot (47 + 2 \cdot \sigma_k \cdot (-39 + 20 \cdot \sigma_k))) / 144 \quad 5-3-2$$

$$B_5 = \sigma_k \cdot (21 + 5 \cdot \sigma_k \cdot (-31 + 2 \cdot \sigma_k \cdot (33 - 20 \cdot \sigma_k))) / 48$$

$$B_6 = 25 \cdot \sigma_k \cdot (-3 + \sigma_k \cdot (23 + 2 \cdot \sigma_k \cdot (-27 + 20 \cdot \sigma_k))) / 1008$$

$$B_2 = B_7 = B_8 = B_9 = 0$$

TABLE 1
RKN8(6)9FM

c_i	a_{ij}	b_i	b_i	b_i	b_i	b_i
0		223	223	7987313	7987313	7987313
1		7938	7938	109941300	109941300	109941300
20	$\frac{1}{800}$	0	0	0	0	0
1	$\frac{1}{600}$	1175	5875	1610737	1610737	1610737
10	$\frac{1}{300}$	8064	36288	44674560	44674560	40207104
3	$\frac{9}{100}$	925	4625	10023263	10023263	10023263
10	$\frac{9}{100}$	6048	21168	33505920	33505920	23454144
1	$\frac{-66701}{197352}$	41	41	-497221	-497221	-497221
2	$\frac{28325}{32892}$	448	224	12409600	12409600	6204800
7	$\frac{-54897451}{30425100}$	925	4625	10023263	10023263	10023263
10	$\frac{30425100}{10141700}$	14112	21168	78180480	78180480	23454144
9	$\frac{-1131891597}{901789000}$	1175	5875	1610737	1610737	1610737
10	$\frac{901789000}{12882700}$	72576	36288	402071040	402071040	402071040
1	$\frac{13836959}{3667458}$	0	223	0	0	-4251941
	$\frac{3667458}{1833729}$	1645	7938	57088	57088	54970650
1	$\frac{223}{7938}$	0	0	0	0	3
	$\frac{223}{7938}$	0	0	0	0	20

Η καλύτερη επιλογή των σ_k είναι

$$s_1 = 0.05, \quad s_2 = 0.15, \quad s_3 = 0.27 \quad 6-3-3$$

Εχουμε λοιπόν

$$f(x_{n+0.05}, \bar{y}_{n+0.05}) =$$

$$= f(x_n + \frac{h}{20}, y_n + \frac{h}{20}) + h^2 \cdot \left(\frac{18973}{50400} f^1 + \frac{1767}{11520} f^3 - \frac{121}{2880} f^4 + \frac{281}{19200} f^5 - \frac{11}{4480} f^6 \right) / 400$$

Ομοια βγαίνουν και τα επόμενα δύο f τα οποία έχουν διαλεχτεί έτσι που να δίνουν το μικρότερο δυνατό σφάλμα στο interpolant που ακολουθεί. Η μορφή του interpolant είναι

$$y_{n+1}^* = d_1 y_n + d_2 y_n' h + d_3 y_n'' h^2 + h^2 \cdot (d_4 f_{n+0.05} + d_5 f_{n+0.15} + d_6 f_{n+0.27}) +$$

$$+ d_7 y_{n+1} + d_8 y_{n+1}' h + d_9 y_{n+1}'' h^2 = u(x_n + \sigma h_n) + O(h^9)$$

6-3-4

με $u(x)$ η τοπική λύση στο $[x_n, x_{n+1}]$ του $u' = f(x, u)$, $u(x_n) = y_n$ και $u''(x_n) = y_n''$ που λύνουμε μετά απο n βήματα. Τα d_i $i=1, 2, \dots, 9$ είναι

$$d_1 = -121800000 \cdot \sigma^8 / 1810909 + 3976008000 \cdot \sigma^7 / 19919999 -$$

$$4107560240 \cdot \sigma^6 / 19919999 + 1745112726 \cdot \sigma^5 / 19919999 - 28437885 \cdot \sigma^4 /$$

$$1810909 + 19136250 \cdot \sigma^3 / 19919999 + 1$$

$$d_2 = -1012500000 \cdot \sigma^8 / 19919999 + 33408620000 \cdot \sigma^7 / 219119989 - 34879713960 \cdot \sigma^6 / 219119989 + 14905061913 \cdot \sigma^5 / 219119989 - 243611088 \cdot \sigma^4 / 19919999 + 164134026 \cdot \sigma^3 / 219119989 + \sigma$$

$$d_3 = 29309050000 \cdot \sigma^8 / 1613519919 - 3729694874000 \cdot \sigma^7 / 53246157327 + 1870137392260 \cdot \sigma^6 / 17748719109 - 2767784576669 \cdot \sigma^5 / 35497438218 + 287550277151 \cdot \sigma^4 / 9681119514 - 202925793847 \cdot \sigma^3 / 35497438218 + \sigma^2 / 2$$

$$d_4 = -195170000000 \cdot \sigma^8 / 4163279791 + 7923287200000 \cdot \sigma^7 / 45796077701 - 11237119692000 \cdot \sigma^6 / 45796077701 + 7640335876000 \cdot \sigma^5 / 45796077701 - 229828316000 \cdot \sigma^4 / 4163279791 + 348478092000 \cdot \sigma^3 / 45796077701$$

6-4-5

$$d_5 = 117265000000 \cdot \sigma^8 / 3047759847 - 14222571200000 \cdot \sigma^7 / 100576074951 + 6563747138000 \cdot \sigma^6 / 33525358317 - 4145249014000 \cdot \sigma^5 / 33525358317 + 303575422000 \cdot \sigma^4 / 9143279541 - 8962882000 \cdot \sigma^3 / 3725039813$$

$$d_6 = -36453125000000 \cdot \sigma^8 / 1295656494957 + 4032662500000000 \cdot \sigma^7 / 42756664333581 + 1638093781250000 \cdot \sigma^6 / 14252221444527 + 858800093750000 \cdot \sigma^5 / 14252221444527 - 47803718750000 \cdot \sigma^4 / 3886969484871 + 11337531250000 \cdot \sigma^3 / 14252221444527$$

$$d_7 = 121800000 \cdot \sigma^8 / 1810909 - 3976008000 \cdot \sigma^7 / 19919999 + 4107560240 \cdot \sigma^6 / 19919999 - 1745112726 \cdot \sigma^5 / 19919999 + 28437885 \cdot \sigma^4 / 1810909 - 19136250 \cdot \sigma^3 / 19919999$$

$$d_8 = -327300000 \cdot \sigma^8 / 19919999 + 10327468000 \cdot \sigma^7 / 219119989 - \\ 10303448680 \cdot \sigma^6 / 219119989 + 4291178073 \cdot \sigma^5 / 219119989 - \\ 69205647 \cdot \sigma^4 / 19919999 + 46364724 \cdot \sigma^3 / 219119989$$

$$d_9 = 551660650000 \cdot \sigma^8 / 469693656421 - 16552583574000 \cdot \sigma^7 / 5166630220631 \\ + 15900459258100 \cdot \sigma^6 / 5166630220631 - 12970465439969 \cdot \sigma^5 / \\ 10333260441262 + 103462382879 \cdot \sigma^4 / 469693656421 - \\ 137992651569 \cdot \sigma^3 / 10333260441262$$

Η ανάλυση σφάλματος για την κύρια μέθοδο δίνει

$$\| T_9 \|_E \cong 0.004866$$

σύμφωνα με όσα φαίνονται στις προηγούμενες παραγράφους. Για το interpolant που μόλις κατασκευάσαμε βλέπουμε μετά από τη συνηθισμένη διαδικασία πως

$$\max_{\sigma \in (0,1)} \| T_9^* \|_E \cong 0.004866$$

6-3-6

που σημαίνει πως για κανένα σ δεν έχουμε υπέρβαση του φράγματος που ισχύει για τη μέθοδο, κάτι που αναμενόταν μια και πρόσθεσαμε 3 f για να κατασκευάσουμε το παραπάνω interpolant και διάλεξαμε τα σ_k , $k=1,2,3$ έτσι ώστε να πετύχουμε την 6-3-6.

Όμως για το παραπάνω πολυώνυμο έγινε ελαχιστοποίηση της $\max \| T_9^{*'} \|_E$ $\sigma \in (0,1)$ και η βέλτιστη τιμή είναι

$$\max_{\sigma \in (0,1)} \| T_9^{*'} \|_E \cong 0.0013 .$$

που επιτυγχάνεται με την επιλογή των σημείων 5-3-3.

Βέβαια σαν επέκταση της παραπάνω εργασίας θα μπορούσε να εξετάσει κανείς interpolants με τιμές από δύο βήματα οπότε με τις 9 τιμές που θα έχει διαθέσιμες είναι δυνατό να κατασκευαστεί παρεμβολικό πολυώνυμο χωρίς κόστος. Μια πρώτη προσέγγιση αυτής της ιδέας έχουμε κάνει από άλλη θέση. (Κόσσυβας [1988]) και επιφυλασόμεθα για το μέλλον.

Σ Υ Μ Π Ε Ρ Α Σ Μ Α Τ Α - Π Ρ Ο Ο Π Τ Ι Κ Ε Σ

Στη διατριβή αυτή ασχοληθήκαμε με τη μελέτη και κατασκευή συνεχών επεκτάσεων της διακριτής λύσης του προβλήματος αρχικών τιμών, όπως αυτή προκύπτει από την εφαρμογή των Runge-Kutta μεθόδων. Εξετάστηκαν τέσσερις διαφορετικές περιπτώσεις. Συγκεκριμένα για το πρόβλημα αρχικών τιμών 1^{ης} τάξης (0-1), οι συνεχείς επεκτάσεις των Runge-Kutta (RK) μεθόδων δημιουργήθηκαν εφαρμόζοντας αφ'ενός την τεχνική του scaling και αφ'ετέρου χρησιμοποιώντας παρεμβολικά πολυώνυμα (interpolation). Ομοίως, για το ειδικό πρόβλημα 2^{ης} τάξης (0-2), οι δύο προαναφερθείσες τεχνικές εφαρμόστηκαν σε συνδυασμό με τις Runge-Kutta-Nystrom (RKN) μεθόδους.

Αναλυτικά τα συμπεράσματα για τα αποτελέσματα που αναφέρονται στην διατριβή μας είναι:

[1] (RK)-scaling : Οι scaled αλγόριθμοι που κατασκευάζονται για τις μεθόδους 4^{ης} τάξης του Fehlberg [1969] και των Dormand-Prince [1980], με βάση την ανάλυση σφάλματος που παρουσιάζουμε είναι οι βέλτιστοι δυνατοί. Επίσης για την 5^{ης} τάξης προσέγγιση της λύσης παρουσιάζονται για πρώτη φορά αναλυτικοί τύποι των παραμέτρων , και δείχνουμε πως η κατασκευή scaled αλγόριθμου απαιτεί μόνο έναν επί πλέον υπολογισμό της τιμής της f αντί για δύο που πρότεινε η Horn [1983].

[2] (RK)-παρεμβολή (interpolation): Εδώ οι μέθοδοι που προτείνουμε λαμβάνουν υπ' όψιν τους προσεγγιστικές τιμές και εκτός του βήματος ολοκλήρωσης. Το υπολογιστικό κόστος είναι

κατά πολύ μικρότερο από τις μεθόδους των άλλων ερευνητών (Enright et.al. [1986], Shampine [1986], Dormand-Prince [1987], Horn [1983], Calvo et.al. [1990], Zennaro [1986]). Η ανάλυση σφάλματος έδειξε ότι τα φράγματα του τοπικού σφάλματος αποκοπής (Τ.Σ.Α) είναι αν όχι μικρότερα τουλάχιστον ίσα με αυτά των μεθόδων που προτείνουν άλλοι ερευνητές.

Επίσης περιοριζόμενοι αυστηρά σε ένα βήμα ολοκλήρωσης προτείνουμε επιλογή ενδιάμεσων σημείων για στήριξη interpolant με ελαχιστοποιημένο το Τ.Σ.Α. χωρίς αύξηση του υπολογιστικού κόστους.

3 (RKN)-scaling : Η τεχνική του scaling εισήχθη και στις μεθόδους (RKN) για το ειδικό πρόβλημα 2^{ης} τάξης. Για μια σειρά από μεθόδους κατασκευάζονται scaled αλγόριθμοι για την παραγωγή συνεχούς προσέγγισης της λύσης και της παραγώγου της και γίνεται ανάλυση του Τ.Σ.Α. Επίσης προτείνονται scaled αλγόριθμοι που ανήκουν στην κλάση C^2 .

4 (RKN)-παρεμβολή (interpolation) : Εδώ ασχολούμεθα με μεθόδους παρεμβολής υψηλής τάξης. Για την μέθοδο (RKN) των Dormand et.al [1987] 8^{ης} τάξης κατασκευάζεται interpolant ίδιας τάξης για την προσέγγιση της λύσης και interpolant 7^{ης} τάξης για την προσέγγιση της παραγώγου της. Στην περίπτωση αυτή χρειάζονται τρεις επί πλέον υπολογισμοί της f . Επειτα από επίπονη λόγω πλήθους υπολογισμών και ελεύθερων παραμέτρων, ελαχιστοποίηση του Τ.Σ.Α. καταλήξαμε στη μέθοδο που προτείναμε.

Προοπτικές για παραπέρα έρευνα υπάρχουν πολλές πάνω σε αυτό το αντικείμενο. Αν περιοριστούμε στις τεχνικές του

scaling και του interpolation τότε οι άμεσες δυνατότητες που υπάρχουν είναι:

(i) Υψηλής τάξης interpolation χρησιμοποιώντας τιμές από δύο βήματα για μεθόδους (RK) και (RKN) με ανάλυση αφάλατος για βέλτιστη επιλογή των σημείων στήριξης.

(ii) Υπολογισμός των υψηλής τάξης παραγώγων στο τέλος κάθε βήματος με βάση το παραχθέν πολυώνυμο παρεμβολής. Χρήση των τιμών αυτών στην κατασκευή interpolant στο επόμενο βήμα. Έλεγχος της σύγκλισης της προσεγγιστικής λύσης για τις παραγώγους.

(iii) Κατασκευή συνεχών επεκτάσεων για έμμεσες (implicit) Runge-Kutta μεθόδους με έλεγχο της Απόλυτης ευστάθειας του παραγομένου παρεμβολικού πολυωνύμου. Οι τυχόν επί πλέον τιμές της f που θα απαιτούσε μια τέτοια μέθοδος υπολογίζονται και αυτές έμμεσα οπότε το υπολογιστικό κόστος είναι πολύ μεγάλο Bellen-Zennaro [1987].

Όμως και σε άλλες περιοχές της Αριθμητικής Επίλυσης των Συνήθων Διαφορικών εξισώσεων μπορούν να εφαρμοστούν τα αναφερθέντα σε αυτή την εργασία. Συγκεκριμένα σε,

(i) Περιοδικά προβλήματα. Εδώ θα μπορούσε κανείς να παράγει συνεχείς επεκτάσεις μεθόδων Runge-Kutta οι οποίες να έχουν όσο το δυνατόν μικρότερη υστέρηση φάσης (phase-lag). Οι Houwen-Sommeijer [1987] πρότειναν μεθόδους Runge-Kutta (-Nystrom) με ελαχιστοποιημένη υστέρηση φάσης. Χρησιμοποιώντας την τεχνική του scaling μπορούμε να κατασκευάσουμε συνεχείς επεκτάσεις τους, αφού λύσουμε από κοινού τις καταστατικές εξισώσεις και τις εξισώσεις για την ικανοποίηση της τάξης

του phase-lag.

(iii) Διαφορικές εξισώσεις με υστέρηση (Delay O.D.E's). Οι εξισώσεις αυτές έχουν την μορφή

$$y'(x) = f(x, y(x), y(x-\tau)) \quad 7-1$$

ή

$$y'(x) = f(x, y(x), y(x-\tau_1), y(x-\tau_2)) \quad 7-2$$

ή ακόμα γενικότερες μορφές. Αν εφαρμόσουμε μια Runge-Kutta μέθοδο στην εξίσωση 7-1 τότε χρειάζονται προσεγγίσεις της $y(x-\tau)$ σε όλο το διάστημα ολοκλήρωσης. Αυτές μπορεί να παραχθούν με τις τεχνικές που αναπτύχθηκαν παραπάνω. Bellen-Zennaro [1985].

(iii) Διαφορικές εξισώσεις με ασυνέχειες. Στην περίπτωση αυτή δεν συνίσταται το να αγνοήσουμε την ασυνέχεια και να ολοκληρώσουμε όλο το διάστημα αδιαφορώντας για τις συνέπειες. Γενικά το αποτέλεσμα είναι απρόβλεπτο. Αν η ασυνέχεια είναι δυνατόν να εντοπιστεί καλό είναι να ληφθεί υπ' όψιν και αφού σταματήσει η ολοκλήρωση να ξαναρχίσει με νέους υπολογισμούς της f . Έτσι π.χ. στη διαφορική εξίσωση που ορίζεται από τον τύπο

$$y' = \begin{cases} -1 & y \geq 0 \\ 1 & y < 0 \end{cases}, y(0) = 0, \quad 7-3$$

όταν εντοπίζεται αλλαγή προσήμου της y πρέπει να ξαναρχίζουμε την ολοκλήρωση εξ' αρχής. Για να γίνει αυτό πρέπει η Runge-Kutta μέθοδος να είναι εφοδιασμένη με συνεχή επέκταση. Enright et.al. [1988]

(iv) Defect Correction τεχνική. Με την περίπτωση αυτή έχουν

ασχοληθεί οι Enright [1989a],[1989b] και Dormand et.al. [1984],[1985]. Χρησιμοποιεί τα επί πλέον υπολογιζόμενα f κατά την εφαρμογή του scaling ή της παρεμβολής σε μια μέθοδο Runge-Kutta για να αυξήσει την τάξη της. Τα προτεινόμενα στην διατριβή μας διβηματικά interpolant που δεν έχουν κόστος υπόσχονται πολλά στην περίπτωση αυτή. Φαίνεται ότι μπορούμε να αυξήσουμε την τάξη μιας Runge-Kutta μεθόδου ανέξοδα Enright [1990]. Όμως το defect correction μπορεί να χρησιμοποιηθεί και για την κατασκευή καλύτερων εκτιμητριών του τοπικού σφάλματος. Μια πλήρης ενημέρωση επί του θέματος μπορεί να πάρει κανείς στο άρθρο του Skeel [1983].

Πρόσφατα έχει εφαρμοστεί τροποποιημένη η παραπάνω τεχνική σε παράλληλους αλγόριθμους. Κατασκευάζεται το γειτονικό πρόβλημα

$$z' = f(x,y)+d(x) , z(x) = y$$

7-4

$$d(x) = P'(x) - f(x,P(x))$$

με $P(x)$ πολυώνυμο που παρεμβάλλει την αριθμητική λύση $\{y_n\}$. Η αριθμητική λύση $\{z_n\}$ του 7-4 μπορεί να παραχθεί από την ίδια αριθμητική μέθοδο που παράγαγε την λύση του αρχικού προβλήματος. Το ολικό σφάλμα $y_n - y(x_n)$ μπορεί να προσεγγιστεί από την τιμή $z_n - P(x)$ που είναι το ολικό σφάλμα του γειτονικού προβλήματος. Έτσι αυξάνουμε την ακρίβεια του αποτελέσματος προσθέτοντας συντεταγμένες στο σύστημα των διαφορικών εξισώσεων και όχι άλλες βαθμίδες, που αναγκαστικά υπολογίζονται σειριακά. Το αποτέλεσμα είναι πολύ ικανοποιητικό σύμφωνα με την εργασία Enright & Higham [1990].

Τέλος πολλές εφαρμογές, οι τεχνικές αυτές, μπορούν να βρουν στη Μηχανική όταν απαιτείται το γράφημα της λύσης μιας διαφορικής εξίσωσης. Τότε χρειαζόμαστε λύσεις σε πολλά ενδιάμεσα σημεία και αν μικρύνουμε το βήμα για να τις υπολογίσουμε αυξάνει υπέρμετρα το υπολογιστικό κόστος, Shampine et.al.[1976] Επίσης όταν θέλουμε έμμεση έξοδο, π.χ. την τιμή της ανεξάρτητης μεταβλητής για την οποία μηδενίζεται η λύση, τότε με τις συνεχείς επεκτάσεις των μεθόδων Runge-Kutta μπορούμε να τις εντοπίσουμε με μικρό κόστος.

A N A Φ O P E Σ

- Bellen,A., Zennaro,M. [1985], Numerische Mathematik 47,301-316
- Bellen,A., Zennaro,M. [1988], SIAM J.Num.Anal. 25,411-432
- Bettis,D.G. [1973], Celestial Mechanics 8,229-233
- Butcher,J.C. [1963], J. Austral. Math. Soc., 3,185-201
- Butcher,J.C. [1964], J. Austral. Math. Soc., 4,179-194
- Butcher,J.C. [1965], Mathematics of Computation 19,408-417
- Butcher,J.C. [1966], Mathematics of Computation 20,1-10
- Butcher,J.C. [1972], Mathematics of Computation 26,79-106
- Butcher,J.C. [1987], John Wiley & Sons
- Calvo,M.,Montijano,J.I,Randez,L.[1990a],J.Comp.Appl.Math. 29,91-100
- Calvo,M., Montijano,J.I, Randez,L. [1990b], to appear
- Chawla,M.M., Sharma.S.R. [1981], BIT 21,455-464
- Curtis,A.R. [1970], Numerische Mathematik 16,268-277
- Dormand,J.R., Prince,P.J. [1978],Celestial Mechanics 18,223-232
- Dormand,J.R., Prince,P.J. [1980], J.Comp.Appl.Math. 6,19-26
- Dormand,J.R., Duckers,R.R., Prince,P.J. [1984], SIAM J. Numerical Analysis 42,169-184
- Dormand,J.R., Prince,P.J. [1985], IMA J.Num. Anal. 5,481-497

Dormand, J.R., Prince, P.J. [1986], *Comp. & Maths. Appls.* **12**, 1007-1017

Dormand, J.R., Prince, P.J. [1987], *Comp. & Maths. Appls.* **13**, 937-949

Dormand, J.R., El-Mikkaway, M.E.A., Prince, P.J. [1987a], *IMA Journal of Numerical Analysis* **7**, 235-250

Dormand, J.R., El-Mikkaway, M.E.A., Prince, P.J. [1987b], *IMA Journal of Numerical Analysis* **7**, 423-430

Δουγαλής, Β. [1987], *Αριθμητική Ανάλυση (Σημειώσεις)*, Ηράκλειο

Enright, W.H., Jackson, K.R., Norsett, S.P., Thomsen, P.G. [1986], *A.C.M. Trans. Math. Software* **12**, 193-218

Enright, W.H., Jackson, K.R., Norsett, S.P., Thomsen, P.G. [1988], *Applied Mathematics and Computation* **27**, 313-335

Enright, W.H. [1989], *Applied Mathematics and Computation* **31**, 288-301

Enright, W.H. [1989], *SIAM J. Numerical Analysis* **26**, 588-599

Enright, W.H., Higham, D.J. [1989], *Technical Rep 237/90*, Univ. of Toronto, Canada

Enright, W.H. [1990], paper in preparation.

Euler, L. [1768], *Opera Omnia vol. XII*

Fatunla, S.O. [1988], Academic Press, Inc.

Fehlberg, E. [1968], NASA TR-287

Fehlberg, E. [1970], *Computing* **4**, 93-106; Corrigendum: *Computing* **5**, 184

Fehlberg, E. [1969], NASA TR-315

Fehlberg, E. [1970], *Computing* 4, 61-71

Fehlberg, E. [1972], NASA TR-381

Fehlberg, E. [1972a], *Computing* 10, 305-315

Fehlberg, E. [1973], NASA TR-410

Fehlberg, E. [1974], NASA TR-432

Fehlberg, E., Filippi, S., Graff, J. [1987], *ZAMM*

Filippi, S. Graff, J. [1986], *J. Comput. and Applied Math.*, 14, 361-370

Filippi, S. Graff, J. [1987], *Computing*

Fine, J. [1985], Technical Report, Univ. of Toronto, Toronto Canada

Fine, J. [1987], *Computing* 39, 27-42

Gear, C.W. [1971], Prentice-Hall, Englewood Cliffs, New Jersey, USA

Gill, S. [1951], *Proc. Cambridge Philos. Soc.*, 47, 95-108

Hairer, E., Wanner, G. [1976], *Numerische Mathematik* 25, 383-400

Hairer, E. [1978], *J. Inst. Mathematics and its Applications* 21, 47-59

Hairer, E. [1982], *I.M.A J. Numer. Anal.* 2, 83-94

Hairer, E., Norsett, S.P., Wanner, G. [1987], Springer Verlag, Heidelberg

Henrici, P. [1962], John Wiley & Sons, Inc., New York, USA

Heun, K. [1900], *Zeitschr. für Math. u. Phys.* 45, 23-38

- Horn, M.K. [1984], SIAM J. Numerical Analysis **20**, 558-568
- Hull, T.E., Enright, W.H., Fellen, B.M., Sedgwick, A.E. [1972], SIAM J. Num. Analysis **9**, 603-637
- Hull, T.E., Enright, W.H., Jackson, K.R. [1980], Report, Dept. of Computer Science, Univ. of Toronto, Canada
- Huta, A. [1957], Acta Fac. Rerum Natur. Univ. Comenian. Math **2**, 21-24
- Κόσσυβας, Ι. [1988], Διπλωματική Εργασία, Ε.Μ.Π., Αθήνα
- Kutta, W. [1901], Zeitschr. für Math. u. Phys. **46**, 435-453
- Merson, R.H. [1957], Proc. Symp. Data Processing, Weapons Research Establishments, Salisbury, S. Australia
- Nyström, E.J. [1925], Acta Soc. Sci. Fenn. vol **50**, no 13, 1-54
- Lambert, J.D. [1973], John Willey, New York, USA
- Owren, B., Zennaro, M. [1989], Technical Report 2/89, Norges Tekniske Hogskole, Norway
- Παπαγεωργίου, Γ. [1985], Αθήνα
- Papageorgiou, G., Tsitouras, Ch. [1987], TR #1/87, National Technical Univ. Athens, Greece
- Papageorgiou, G., Simos, Th., Tsitouras, Ch. [1988], Celestial Mechanics **44**, 167-177
- Papageorgiou, G., Tsitouras, Ch. [1989], Int. J. Comp. Maths **28**, 139-150
- Runge, C. [1895], Math. Ann., **46**, 167-178

- Shampine,L.F.,Watts,H.A.,Davenport,S.M. [1976],SIAM Rev. 18,376-411
- Shampine,L.F., Watts,H.A. [1976],Report SAND-76-0585, N.Mexico, USA
- Shampine,L.F. [1985], SIAM J.Num.Anal. 22,1014-1027
- Shampine,L.F. [1986], Mathematics of Computation 46,135-150
- Shintani,H. [1966],J.Sci. Hiroshima Univ. Ser. A-I Math. 30,91-107
- Skeel,R.D. [1986], Numerische Mathematik 48,1-20
- Tsitouras,Ch. Papageorgiou,G. [1990], Computing 43,255-266
- Verner,J.H. [1978a], SIAM J.Numer. Analysis 15,772-790
- Verner,J.H. [1978b], SIAM J.Numer. Analysis 16,857-875
- Verner,J.H. [1990a], Conference on numerical solution of O.D.E., Helsinki, Finland
- Verner,J.H. [1990b], Mathematical Preprint #1990-9, Queen's Univ., Canada
- Verner,J.H. [1990c], Mathematical Preprint #1990-11, Queen's Univ., Canada
- Zennaro,M. [1986], Mathematics of Computation 46,119-133
- Zonneveld,J.A. [1964],Mathematical Centre Tracts 8,Amsterdam,Holland

