

Αριθμητική Επίλυση Διαφορικών Εξισώσεων

Ζαχαρούλα Καλογηράτου - Θεόδωρος Μονοβασίλης

5 Απριλίου 2022

Περιεχόμενα

1	Εισαγωγή	7
1.1	Διαφορικές Εξισώσεις - Γενικά	7
1.1.1	Διαφορικές εξισώσεις 1ης τάξης	8
1.1.2	Συστήματα ΔΕ 1ης τάξης	9
1.2	Αριθμητικές μέθοδοι	11
1.2.1	Η μέθοδος του Euler	12
1.2.2	Πολυβηματικές Μέθοδοι	13
1.3	Προβλήματα Δοκιμής	18
1.3.1	Πρόβλημα 1	18
1.3.2	Πρόβλημα 2	23
1.3.3	Πρόβλημα 3	25
1.3.4	Πρόβλημα 4	29
1.3.5	Πρόβλημα 5	32
1.4	Βιβλιογραφία	37
2	Μέθοδοι Runge-Kutta	41
2.1	Εισαγωγή	41
2.2	Γενική μορφή των μεθόδων Runge-Kutta	43
2.3	Κατασκευή μεθόδων	46
2.3.1	Άμεσες μέθοδοι δεύτερης τάξης	47
2.3.2	Άμεσες μέθοδοι τρίτης τάξης	53
2.3.3	Άμεσες μέθοδοι τέταρτης τάξης	64
2.3.4	Συνθήκες με χρήση πινάκων	75
2.4	Βιβλιογραφία	76
3	Μέθοδοι Runge-Kutta	77
3.1	Πεπλεγμένες (implicit) μέθοδοι	82
3.2	Μέθοδοι και αποτελέσματα	84
3.3	Βιβλιογραφία	85

Κατάλογος σχημάτων

1.1	Πρόβλημα 1. Αριθμητική και αναλυτική λύση.	21
1.2	Πρόβλημα 1. Απόλυτο σφάλμα.	22
1.3	Πρόβλημα 2. Η αριθμητική και η αναλυτική λύση στο $[0, 4\pi]$ για βήματα $\pi/4$, $\pi/8$, $\pi/16$, $\pi/32$	24
1.4	Πρόβλημα 2. Σφάλμα της λύσης στο $[0, 4\pi]$ για βήματα $\pi/16$, $\pi/32$	25
1.5	Πρόβλημα 3. Αναλυτική λύση.	27
1.6	Πρόβλημα 3. Αριθμητική και αναλυτική λύση.	28
1.7	Πρόβλημα 4. Αναλυτική λύση.	31
1.8	Πρόβλημα 4. Σφάλμα της λύσης για $N = 512$, $N = 1280$, $N = 5120$	32
1.9	Πρόβλημα 4. Αριθμητική και αναλυτική λύση για $N = 12800$ σημεία.	33
1.10	Πρόβλημα 5 ($e = 0$) Αριθμητική και αναλυτική λύση για την $y_1(x)$, $[0, 20]$, $N = 10^4$	37
1.11	Πρόβλημα 5 ($e = 0$) Αριθμητική και αναλυτική λύση για την $y_3(x)$, $[0, 20]$, $N = 10^4$	38
1.12	Πρόβλημα 5 ($e = 0$) Διάγραμμα $(y_1(x), y_3(x))$, $[0, 20]$ για $N = 1000$, $N = 5000$, $N = 10^4$, $N = 10^5$	39
1.13	Πρόβλημα 5 Διάγραμμα $(y_1(x), y_3(x))$, $[0, 20]$ για $e = 0$, $e = 0.5$, $e = 0.8$, $e = 0.95$	40

Κεφάλαιο 1

Εισαγωγή στις Αριθμητικές Μεθόδους Επίλυσης Διαφορικών Εξισώσεων

1.1 Διαφορικές Εξισώσεις - Γενικά

Μια διαφορική εξίσωση είναι μια σχέση μεταξύ μιας συνάρτησης και των παραγώγων της. Έστω I ένα διάστημα του $\mathbb{R}$ και U ένα υποσύνολο του $\mathbb{R}^{n+1}$, $g : I \times U \rightarrow \mathbb{R}$ μια συνάρτηση $n+2$ μεταβλητών και $y : I \rightarrow \mathbb{R}$ μια άγνωστη συνάρτηση, τότε η εξίσωση

$$g(x, y(x), y'(x), y''(x), \dots, y^{(n-1)}(x), y^{(n)}(x)) = 0$$

λέγεται συνήθως διαφορική εξίσωση (ΣΔΕ) ή διαφορική εξίσωση (ΔΕ). Τάξη της διαφορικής εξίσωσης ονομάζεται η μεγαλύτερη τάξη παραγώγου που εμφανίζεται στην εξίσωση.

Κάθε συνάρτηση $y : I \rightarrow \mathbb{R}$ που έχει παραγώγους έως και n τάξης για την οποία ισχύει

$$(x, y(x), y'(x), y''(x), \dots, y^{(n-1)}(x), y^{(n)}(x)) \in I \times U$$

και

$$g(x, y(x), y'(x), y''(x), \dots, y^{(n-1)}(x), y^{(n)}(x)) = 0$$

λέγεται λύση της (1).

Η σχέση (1) λέγεται πεπλεγμένη μορφή της ΔΕ, αν μπορεί να λυθεί ως προς τη παράγωγο μεγαλύτερης τάξης $y^{(n)}$ τότε γράφεται

$$y^{(n)} = f(x, y(x), y'(x), y''(x), \dots, y^{(n-1)}(x)).$$

και λέμε ότι έχουμε τη ΔΕ σε λυμένη μορφή.

Πρόβλημα Αρχικών Τιμών ΠΑΤ (initial value problem IVP) είναι μια ΔΕ μαζί με τις αρχικές συνθήκες $[x_0, b)$ ή $[x_0, b]$ ή $[x_0, +\infty)$

$$y(x_0) = a_0, \quad y^{(k)}(x_0) = a_k, \quad k = 1, 2, \dots, n-1$$

$a_0, a_1, a_2, \dots, a_{n-1}$ πραγματικοί αριθμοί

1.1.1 Διαφορικές εξισώσεις 1ης τάξης

Η γενική μορφή μιας διαφορικής εξίσωσης πρώτης τάξης είναι

$$y' = f(x, y) \quad (1.1)$$

όπου f μια συνάρτηση δύο μεταβλητών.

Όταν η συνάρτηση στο δεξι μέλος είναι ανεξάρτητη του x το πρόβλημα λέγεται αυτόνομο

$$y' = f(y)$$

Το θεμελιώδες ερώτημα κατά την μοντελοποίηση είναι κατά πόσο ένα πρόβλημα αρχικών τιμών μπορεί να χρησιμοποιηθεί αξιόπιστα για να προβλέψει τη συμπεριφορά της τροχιάς σε μελλοντικό χρόνο. Στην περίπτωση που αυτό επιτυγχάνεται λέμε ότι έχουμε ένα καλά τοποθετημένο πρόβλημα. Τα ερωτήματα που μας απασχολούν είναι τα εξής:

- ύπαρξη της λύσης
- μοναδικότητα της λύσης εφόσον υπάρχει
- ευσταθία της λύσης σε μικρές αλλαγές της αρχικής πληροφορίας

Θεώρημα 1.1.1 Έστω f και $\partial f / \partial y$ συνεχείς στο

$$R : x_0 \leq x \leq x_0 + a, \quad |y - y_0| \leq b$$

και

$$M = \max_{(t,y) \in R} |f(x, y)|$$

Τότε το ΠΑΤ

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

έχει μοναδική λύση $y(t)$ στο διάστημα $x_0 \leq x \leq x_0 + \alpha$ όπου $\alpha = \min(a, b/M)$.

Ορισμός 1.1.1 Μια συνάρτηση $f(x, y)$ λέμε ότι πληροί μια συνθήκη Lipschitz σε μια περιοχή $U \subset \mathbb{R}^2$, ως προς τη μεταβλητή y , αν υπάρχει σταθερά L , έτσι ώστε να ισχύει η ανισότητα:

$$|f(x, y_1) - f(x, y_2)| \leq L |y_1 - y_2|,$$

για οποιαδήποτε $(x, y_1), (x, y_2) \in U$.

Θεώρημα 1.1.2 Έστω $f : [x_0 - a, x_0 + a] \times [y_0 - b, y_0 + b] \rightarrow \mathbb{R}$. Αν για το ΠΑΤ

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

ισχύει ότι η f

- είναι συνεχής, και κατά συνέπεια φραγμένη με $|f(x, y)| \leq M$
- ικανοποιεί μια συνθήκη Lipschitz με σταθερά L

τότε το ΠΑΤ έχει ακριβώς μια λύση $y = y(x)$ στο διάστημα $[x_0 - \alpha, x_0 + \alpha]$ όπου $\alpha = \min(a, b/M)$.

1.1.2 Συστήματα ΔΕ 1ης τάξης

Επίσης μπορούμε να έχουμε ένα σύστημα δύο διαφορικών εξισώσεων με δύο άγνωστες συναρτήσεις $y_1(x)$ και $y_2(x)$,

$$\begin{aligned}y_1' &= f_1(x, y_1, y_2), \\y_2' &= f_2(x, y_1, y_2).\end{aligned}$$

όπως το γραμμικό σύστημα παρακάτω

$$\begin{aligned}y_1' &= a_1 y_1 + a_2 y_2, \\y_2' &= b_1 y_1 + b_2 y_2.\end{aligned}$$

Μπορούμε να γενικεύσουμε για n άγνωστες συναρτήσεις $y_1(x), y_2(x), \dots, y_n(x)$

$$\begin{aligned}y_1' &= a_{11}y_1 + a_{12}y_2 + \dots + a_{1n}y_n, \\y_2' &= a_{21}y_1 + a_{22}y_2 + \dots + a_{2n}y_n, \\&\vdots \\y_n' &= a_{n1}y_1 + a_{n2}y_2 + \dots + a_{nn}y_n.\end{aligned}$$

Το παραπάνω είναι ένα σύστημα n γραμμικών διαφορικών εξισώσεων, το οποίο μπορούμε να γράψουμε σε διανυσματική μορφή

$$Y' = AY$$

όπου $Y(x)$ η διανυσματική συνάρτηση που ορίζεται σαν

$$Y(x) = \begin{pmatrix} y_1(x) \\ y_2(x) \\ \vdots \\ y_n(x) \end{pmatrix}$$

και ο A πίνακας πραγματικών αριθμών

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

Παρακάτω δίνουμε τη γενική μορφή ενός συστήματος διαφορικών εξισώσεων πρώτης τάξης

$$\begin{aligned}y_1' &= f_1(x, y_1, y_2, \dots, y_n), \\y_2' &= f_2(x, y_1, y_2, \dots, y_n), \\&\vdots \\y_n' &= f_n(x, y_1, y_2, \dots, y_n).\end{aligned} \tag{1.2}$$

σε διανυσματική μορφή

$$Y' = F(x, Y)$$

όπου

$$F(x, Y) = \begin{pmatrix} f_1(x, Y) \\ f_2(x, Y) \\ \vdots \\ f_n(x, Y) \end{pmatrix}$$

Αντίστοιχα το αυτόνομο σύστημα είναι

$$\begin{aligned} y'_1 &= f_1(y_1, y_2, \dots, y_n), \\ y'_2 &= f_2(y_1, y_2, \dots, y_n), \\ &\vdots \\ y'_n &= f_n(y_1, y_2, \dots, y_n). \end{aligned}$$

Μια διαφορική εξίσωση n τάξης μπορεί να γραφεί σαν ένα σύστημα n διαφορικών εξισώσεων πρώτης τάξης με τον εξής τρόπο.

Θέτουμε τις συναρτήσεις

$$\begin{aligned} y_1(x) &= y(x) \\ y_2(x) &= y'(x) \\ y_3(x) &= y''(x) \\ &\vdots \\ y_{n-1}(x) &= y^{(n-2)}(x) \\ y_n(x) &= y^{(n-1)}(x) \end{aligned}$$

τότε

$$\begin{aligned} y'_1(x) &= y_2(x) \\ y'_2(x) &= y_3(x) \\ &\vdots \\ y'_{n-1}(x) &= y_n(x) \\ y'_n(x) &= f(x, y_1(x), y_2(x), \dots, y_{n-1}(x)) \end{aligned}$$

Δηλαδή είναι ένα σύστημα της μορφής (1.2) με

$$\begin{aligned} f_1(x, y_1, y_2, \dots, y_n) &= y_2(x) \\ f_2(x, y_1, y_2, \dots, y_n) &= y_3(x) \\ &\vdots \\ f_{n-1}(x, y_1, y_2, \dots, y_n) &= y_n(x) \\ f_n(x, y_1, y_2, \dots, y_n) &= f(x, y_1(x), y_2(x), \dots, y_{n-1}(x)). \end{aligned}$$

Μια ειδική περίπτωση διαφορικής εξίσωσης δεύτερης τάξης είναι αυτή στην οποία δεν εμφανίζεται η πρώτη παράγωγος

$$y''(x) = f(x, y) \quad (1.3)$$

αυτή γράφεται σαν σύστημα θέτοντας $y_1(x) = y(x)$ και $y_2(x) = y'(x)$

$$\begin{aligned} y_1'(x) &= y_2(x) \\ y_2'(x) &= f(x, y_1(x)). \end{aligned}$$

1.2 Αριθμητικές μέθοδοι

Η αριθμητική επίλυση των διαφορικών εξισώσεων είναι τόσο παλιά όσο και η αναλυτική θεωρία τους, χαρακτηριστικά αναφέρουμε τον Isaac Newton (1642 - 1727), B. Taylor (1685 - 1731), Euler (1707 - 1783). Οι J.C. Adams και F. Bashforth το 1883 παρουσίασαν μεθόδους πολλαπλού βήματος που είναι ιδιαίτερα δημοφιλείς μέχρι και σήμερα. Ο C. Runge το 1895 δημοσίευσε την πρώτη αριθμητική μέθοδο απλού βήματος μαζί με τον Kutta θεωρούνται οι θεμελιωτές των φημισμένων μεθόδων Runge-Kutta.

Αφορμή για την ανάπτυξη αριθμητικών μεθόδων στάθηκαν προβλήματα της ουράνιας μηχανικής που απασχόλησαν τους επιστήμονες από το δέκατο έβδομο αιώνα. Ένα από τα προβλήματα είναι η εμφάνιση του κομήτη Halley ο οποίος παρατηρήθηκε ήδη το 240 πΧ, μελετήθηκε από τον Βρετανό αστρονόμο Edmond Halley σύγχρονο και φίλο του Newton. Είναι ο μόνος μικρής περιόδου κομήτης που είναι καθαρά ορατός από τη γη, επιστρέφει κάθε 75 με 76 χρόνια, τελευταία φορά εμφανίστηκε το 1986 και θα ξαναεμφανιστεί στα μέσα του 2061. Οι γάλλοι μαθηματικοί Clairaut, Lalande και Lepaute το 1748 ασχολήθηκαν με τον χρονικό προσδιορισμό της επιστροφής του, η πρόβλεψη τους ήταν ότι θα μπει στο ηλιακό μας σύστημα στις 13 Απριλίου 1749, το λάθος ήταν 31 ημέρες. Η επόμενη επιστροφή του το 1835 προβλέφθηκε από τους επιστήμονες με σφάλμα 5 ημερών. Η εμφάνιση του το 1910 υπολογίστηκε με σφάλμα 2.7 ημερών από τους βρετανούς αστρονόμους P.H. Cowell και A.C.D. Crommelin στο άρθρο με τίτλο 'Investigation of the Motion of Halley's Comet from 1759-1910', Greenwich Observations 1909. Η αριθμητική μέθοδος που χρησιμοποιήθηκε είναι σήμερα γνωστή ως μέθοδος Numeron και οφείλει το όνομα της στον Ρώσο αστρονόμο B. V. Numeron (1891-1941) που την κατασκεύασε ανεξάρτητα από τους Βρετανούς συναδέλφους του. Η μέθοδος Numeron είναι μέθοδος δύο βημάτων για προβλήματα ΔΕ δεύτερης τάξης.

Μια αριθμητική μέθοδος υπολογίζει εκτιμήσεις της λύσης σε ένα πεπερασμένο πλήθος σημείων του διαστήματος ολοκλήρωσης $[a, b]$. Θεωρούμε μια διαμέριση n σημείων του διαστήματος ολοκλήρωσης

$$a = x_0 < x_1 < x_2 < \dots < x_{N-1} < x_N = b.$$

Η απόσταση δύο διαδοχικών σημείων ονομάζεται βήμα της μεθόδου και συμβολίζεται με h

$$h_n = x_{n+1} - x_n.$$

Αν τα υποδιαστήματα θεωρηθούν ίδιου μήκους τότε το βήμα είναι σταθερό και ίσο με

$$h = \frac{b - a}{N}$$

Θεωρήσουμε τη ΔΕ πρώτης τάξης (1.1)

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

με διάστημα ολοκλήρωσης $[a, b]$ και αρχική συνθήκη $y(a) = y(x_0) = y_0$.

Το παραπάνω ονομάζεται αρχικών τιμών (ΠΑΤ) (initial value problem (IVP)) σε συνήθειες διαφορικές εξισώσεις πρώτης τάξης.

Από την αρχική συνθήκη γνωρίζουμε την τιμή της άγνωστης συνάρτησης στο σημείο x_0 , από αυτό υπολογίζουμε προσέγγιση της αναλυτικής λύσης στο σημείο x_1 ($y(x_1)$) την οποία συμβολίζουμε με y_1 . Στη συνέχεια από την προσέγγιση x_1 υπολογίζουμε την εκτίμηση y_2 της αναλυτικής λύσης $y(x_2)$, μια τέτοια μέθοδος ονομάζεται μέθοδος απλού βήματος. Θα μπορούσαμε να χρησιμοποιήσουμε τα y_0 και y_1 για να υπολογίσουμε την εκτίμηση y_2 της αναλυτικής λύσης $y(x_2)$, μια τέτοια μέθοδος ονομάζεται μέθοδος δύο βημάτων. Γενικά μια αριθμητική μέθοδος που στηρίζεται μόνο στην εκτίμηση y_n για τον υπολογισμό της y_{n+1} ονομάζεται μέθοδος απλού βήματος ή μονοβηματική μέθοδος (one step method ή single step method), ενώ μια μέθοδος που απαιτεί περισσότερες από μία προηγούμενες εκτιμήσεις ονομάζεται μέθοδος πολλαπλού βήματος ή πολυβηματική μέθοδος (multistep method). Μια μέθοδος k βημάτων απαιτεί k προηγούμενες εκτιμήσεις $y_{n-k+1}, \dots, y_{n-2}, y_{n-1}, y_n$ για τον υπολογισμό της y_{n+1} .

Η γενική μορφή μιας μεθόδου απλού βήματος είναι:

$$y_{n+1} = y_n + h\Phi(x_n, y_n, h). \quad (1.4)$$

Ενώ η γενική μορφή μιας πολυβηματικής μεθόδου είναι:

$$\sum_{i=0}^{k-1} a_i y_{n+1-i} = h\Phi(x_n, y_{n+1}, y_n, y_{n-1}, \dots, y_{n-k+1}, h). \quad (1.5)$$

1.2.1 Η μέθοδος του Euler

Θα παρουσιάσουμε πρώτα την μέθοδο του Euler που είναι η παλαιότερη αριθμητική μέθοδος.

Μπορούμε να εκτιμήσουμε την παράγωγο της $y(x)$ σαν το πηλίκο διαφορών

$$y'(x_n) \approx \frac{y(x_{n+1}) - y(x_n)}{x_{n+1} - x_n} = \frac{y(x_{n+1}) - y(x_n)}{h} = \frac{y_{n+1} - y_n}{h}$$

Αντικαθιστούμε στην ΔΕ και έχουμε

$$\frac{y_{n+1} - y_n}{h} \approx f(x_n, y_n)$$

από όπου έχουμε τον τύπο

$$y_{n+1} = y_n + h f(x_n, y_n)$$

Θα μπορούσαμε να κατασκευάσουμε τη μέθοδο χρησιμοποιώντας το ανάπτυγμα Taylor γύρω από το x_n

$$y(x) = y(x_n) + (x - x_n)y'(x_n) + \frac{(x - x_n)^2}{2!}y''(x_n) + \dots$$

$$y(x_{n+1}) = y(x_n) + (x_{n+1} - x_n)y'(x_n) + \frac{(x_{n+1} - x_n)^2}{2!}y''(x_n) + \dots$$

αντικαθιστούμε το $x_{n+1} - x_n$ με το h

$$y(x_{n+1}) = y(x_n) + h y'(x_n) + \frac{h^2}{2} y''(x_n) + \dots$$

αποκόπτουμε τη σειρά από τον όρο h^2 και πέρα

$$y(x_{n+1}) \approx y(x_n) + h y'(x_n)$$

αντικαθιστούμε την παράγωγο $y'(x_n)$ η οποία είναι ίση με $f(x_n, y(x_n))$

$$y(x_{n+1}) \approx y(x_n) + h f(x_n, y(x_n))$$

τέλος γράφουμε τη μέθοδο

$$y_{n+1} = y_n + h f(x_n, y_n)$$

1.2.2 Πολυβηματικές Μέθοδοι

Οι Adams και Bashforth το σε άρθρο τους το 1883 εισήγαγαν πολυβηματικές μεθόδους. Η γενική μορφή μιας γραμμικής πολυβηματικής μεθόδου είναι

$$y_{n+1} = \sum_{i=1}^k a_i y_{n+1-i} + h \sum_{i=0}^k b_i f(x_{n+1-i}, y_{n+1-i}), \quad (1.6)$$

Για να είναι η μέθοδος συνεπής (consistent) πρέπει να είναι ακριβής για $y(x) = 1$ ($f(x, y(x)) = 0$) δηλαδή να ισχύει

$$\sum_{i=1}^k a_i = 1$$

Θεωρούμε την μέθοδο για $k = 1$

$$y_{n+1} = a_1 y_n + h (b_0 f(x_{n+1}, y_{n+1}) + b_1 f(x_n, y_n)),$$

παρατηρούμε ότι στο δεξί μέλος εμφανίζεται ο όρος $f(x_{n+1}, y_{n+1})$ ο οποίος εμπλέκει την άγνωστη εκτίμηση y_{n+1} . Μια τέτοια μέθοδος λέγεται έμμεση, εδώ θα θεωρήσουμε την άμεση μέθοδο που είναι πιο εύκολη στην εφαρμογή της. Για το λόγο αυτό θέτουμε $b_0 = 0$. Για να είναι η μέθοδος ακριβής για $y(x) = 1$ ($f(x, y(x)) = 0$) πρέπει να ισχύει $a_1 = 1$.

Για να είναι η μέθοδος ακριβής για $y(x) = x$ ($f(x, y) = y'(x) = 1$) πρέπει

$$x + h = x + h b_1$$

άρα $b_1 = 1$. Συνεπώς η μέθοδος που προκύπτει είναι

$$y_{n+1} = y_n + h f(x_n, y_n)$$

δηλαδή η μέθοδος Euler.

Θεωρούμε την μέθοδο για $k = 2$

$$y_{n+1} = a_1 y_n + a_2 y_{n-1} + h \left(b_0 f(x_{n+1}, y_{n+1}) + b_1 f(x_n, y_n) + b_2 f(x_{n-1}, y_{n-1}) \right),$$

Θέτοντας $a_1 = 1$, $a_2 = 0$ και $b_0 = 0$ η μέθοδος γράφεται

$$y_{n+1} = y_n + h \left(b_1 f(x_n, y_n) + b_2 f(x_{n-1}, y_{n-1}) \right),$$

μπορούμε να υπολογίσουμε τους συντελεστές b_1 και b_2 ώστε η μέθοδος να είναι ακριβής για πολυώνυμα όσο το δυνατό μεγαλύτερου βαθμού.

Αν $y(x) = x$ τότε $f(x, y) = y'(x) = 1$

$$x + h = x + h(b_1 + b_2)$$

οι συντελεστές πρέπει να ικανοποιούν τη σχέση $b_1 + b_2 = 1$.

Αν $y(x) = x^2$ τότε $f(x, y) = y'(x) = 2x$

$$(x + h)^2 = x^2 + h(b_1 2x + b_2 2(x - h))$$

$$x^2 + 2xh + h^2 = x^2 + 2h(b_1 + b_2)x - 2h^2 b_2$$

παίρνουμε επιπλέον της $b_1 + b_2 = 1$ τη σχέση $2b_2 = -1$ από όπου προκύπτει $b_2 = -1/2$ και $b_1 = 3/2$.

Η μέθοδος γράφεται

$$y_{n+1} = y_n + h \left(\frac{3}{2} f(x_n, y_n) - \frac{1}{2} f(x_{n-1}, y_{n-1}) \right). \quad (1.7)$$

Η παραπάνω μέθοδος είναι ακριβής για πολυώνυμα δεύτερου βαθμού.

Παρατηρούμε ότι χρησιμοποιεί τις εκτιμήσεις y_{n-1} και y_n για να βρει την εκτίμηση y_{n+1} , μια τέτοια μέθοδος λέγεται δύο βημάτων. Μέθοδοι Adams - Bashforth είναι άμεσες μέθοδοι $b_0 = 0$ της μορφής (1.6) για τις οποίες ισχύει $a_1 = 1$ και $a_i = 0$ για $i = 2, \dots, k$. Δηλαδή έχουν τη γενική μορφή

$$y_{n+1} = y_n + h \sum_{i=1}^k b_i f(x_{n+1-i}, y_{n+1-i}), \quad (1.8)$$

ή

$$y_{n+1} = y_n + h \left(b_1 f(x_n, y_n) + b_2 f(x_{n-1}, y_{n-1}) + \dots + b_k f(x_{n-k}, y_{n-k}) \right).$$

Οι παραπάνω μέθοδοι είναι της μορφής αυτής. Θα βρούμε τους συντελεστές της μεθόδου τριών βημάτων $k = 3$

$$y_{n+1} = y_n + h \left(b_1 f(x_n, y_n) + b_2 f(x_{n-1}, y_{n-1}) + b_3 f(x_{n-2}, y_{n-2}) \right),$$

Θεωρούμε $y(x) = x$ τότε $f(x, y) = y'(x) = 1$

$$x + h = x + h(b_1 + b_2 + b_3)$$

οι συντελεστές πρέπει να ικανοποιούν τη σχέση $b_1 + b_2 + b_3 = 1$

Θεωρούμε $y(x) = x^2$ τότε $f(x, y) = y'(x) = 2x$

$$\begin{aligned}(x+h)^2 &= x^2 + h(b_1 2x + b_2 2(x-h) + b_3 2(x-2h)) \\ x^2 + 2xh + h^2 &= x^2 + 2h(b_1 + b_2 + b_3) - 2h^2(b_2 + 2b_3)\end{aligned}$$

οπότε έχουμε τη σχέση $b_2 + 2b_3 = -1/2$.

Θεωρούμε $y(x) = x^3$ τότε $f(x, y) = y'(x) = 3x^2$

$$\begin{aligned}(x+h)^3 &= x^3 + h(b_1 3x^2 + b_2 3(x-h)^2 + b_3 3(x-2h)^2) \\ x^3 + 3x^2h + 3xh^2 + h^3 &= x^3 + 3x^2h(b_1 + b_2 + b_3) - 6xh^2(b_2 + 2b_3) + 3h^3(b_2 + 4b_3)\end{aligned}$$

οπότε έχουμε τη σχέση $b_2 + 4b_3 = 1/3$.

Συνοψίζοντας η μέθοδος 3 βημάτων για να είναι 3ης τάξης πρέπει να ικανοποιούνται οι σχέσεις

$$\begin{aligned}b_1 + b_2 + b_3 &= 1 \\ b_2 + 2b_3 &= -\frac{1}{2} \\ b_2 + 4b_3 &= \frac{1}{3}\end{aligned}\tag{1.9}$$

από τις οποίες παίρνουμε τους συντελεστές

$$b_1 = \frac{23}{12}, \quad b_2 = -\frac{16}{12}, \quad b_3 = \frac{5}{12}$$

$$y_{n+1} = y_n + \frac{h}{12} (23f(x_n, y_n) - 16f(x_{n-1}, y_{n-1}) + 5f(x_{n-2}, y_{n-2}))\tag{1.10}$$

Μια ειδική περίπτωση ΔΕ δεύτερης τάξης είναι αυτή στην οποία δεν εμφανίζεται η πρώτη παράγωγος, της οποίας η μορφή είναι

$$y''(x) = f(x, y(x)), \quad x \in [x_0, X], \quad y(x_0) = y_0, \quad y'(x_0) = y'_0.\tag{1.11}$$

Παρότι η παραπάνω εξίσωση μπορεί να γραφεί σαν ένα σύστημα δύο ΔΕ πρώτης τάξης για την επίλυση της αναπτύχθηκαν ειδικές μέθοδοι.

Η πιο δημοφιλής είναι η μέθοδος Numeron στην οποία αναφερθήκαμε παραπάνω. Η μέθοδος είναι της μορφής

$$y_{n+1} + a_1 y_n + a_0 y_{n-1} = h^2(b_2 f_{n+1} + b_1 f_n + b_0 f_{n-1})\tag{1.12}$$

Η μέθοδος Numeron είναι έμμεση μέθοδος διότι εμπλέκει στο δεξί μέλος το f_{n+1} ο υπολογισμός του οποίου απαιτεί το y_{n+1} που είναι άγνωστο και ζητούμενο. Αυτό το πρόβλημα αντιμετωπίζεται με τη χρήση μιας επαναληπτικής μεθόδου (όπως η Newton) για τη λύση μιας εξίσωσης ή ενός συστήματος.

Ζητώντας η μέθοδος να είναι ακριβής για πολώνυμα όσο το δυνατό μεγαλύτερου βαθμού θεωρούμε το σύνολο

$$1, x, x^2, x^3, \dots\tag{1.13}$$

Έστω $y(x) = 1$ τότε $y'(x) = 0$ και $y''(x) = 0$. αντικαθιστούμε στην (1.16) και έχουμε:

$$1 + a_1 + a_0 = 0$$

από όπου παίρνουμε τη σχέση για τα a

$$a_1 + a_0 = -1$$

Συνεχίζουμε παίρνοντας $y(x) = x$ τότε $y'(x) = 1$ και $y''(x) = 0$.

$$\begin{aligned}(x+h) + a_1x + a_0(x-h) &= 0 \Rightarrow \\ (1+a_1+a_0)x + (1-a_0)h &= 0\end{aligned}$$

από όπου παίρνουμε

$$1 + a_1 + a_0 = 0 \text{ και } 1 - a_0 = 0 \quad (1.14)$$

Παρατηρούμε ότι την πρώτη από τις δύο σχέσεις την πήραμε ήδη από το $y(x) = 1$ άρα

$$a_0 = 1 \text{ και } a_1 = -2 \quad (1.15)$$

Συνεπώς το αριστερό μέλος έχει προσδιοριστεί και μπορούμε να γράψουμε για τη μέθοδο

$$y_{n+1} - 2y_n + y_{n-1} = h^2(b_2f_{n+1} + b_1f_n + b_0f_{n-1})$$

Ζητώντας η μέθοδος να είναι ακριβής για

$$y(x) = x^2 \text{ τότε } y'(x) = 2x \text{ και } y''(x) = 2$$

και βρίσκουμε

$$\begin{aligned}(x+h)^2 - 2x^2 + (x-h)^2 &= h^2(b_2 \cdot 2 + b_1 \cdot 2 + b_0 \cdot 2) \Rightarrow \\ (x+h)^2 - 2x^2 + (x-h)^2 &= 2h^2(b_2 + b_1 + b_0) \Rightarrow \\ 2h^2 &= 2h^2(b_2 + b_1 + b_0) \Rightarrow \\ b_2 + b_1 + b_0 &= 1\end{aligned}$$

Ζητώντας η μέθοδος να είναι ακριβής για

$$y(x) = x^3 \text{ τότε } y'(x) = 3x^2 \text{ και } y''(x) = 6x$$

το πρώτο μέλος γίνεται:

$$\begin{aligned}A &= (x+h)^3 - 2x^3 + (x-h)^3 = \\ &= x^3 + 3x^2h + 3xh^2 + h^3 - 2x^3 - 3x^2h + 3xh^2 - h^3 = \\ &= 6xh^2\end{aligned}$$

και το δεύτερο μέλος

$$\begin{aligned}B &= h^2(b_0 \cdot 6(x+h) + b_1 \cdot 6x + b_0 \cdot 6(x-h)) \\ &= 6h^2((b_2 + b_1 + b_0)x + (b_2 - b_0)h) \\ &= 6xh^2(b_2 + b_1 + b_0) + 6h^3(b_2 - b_0)\end{aligned}$$

και παίρνουμε τη σχέση $b_2 = b_0$ συνεπώς $b_1 = 1 - 2b_0$.

Ζητώντας η μέθοδος να είναι ακριβής για

$$y(x) = x^4 \text{ τότε } y'(x) = 4x^3 \text{ και } y''(x) = 12x^2,$$

το πρώτο μέλος γίνεται:

$$\begin{aligned} A &= (x+h)^4 - 2x^4 + (x-h)^4 \\ &= x^4 + 4x^3h + 6x^2h^2 + 4xh^3 + h^4 - 2x^4 + x^4 - 4x^3h + 6x^2h^2 - 4xh^3 + h^4 = \\ &= 12x^2h^2 + 2h^4 \end{aligned}$$

και το δεύτερο μέλος

$$\begin{aligned} B &= h^2(b_0 12(x+h)^2 + b_1 12x^2 + b_0 12(x-h)^2) = \\ &= 12h^2(b_0(x^2 + 2xh + h^2) + b_1 x^2 + b_0(x^2 - 2xh + h^2)) \\ &= 12h^2((2b_0 + b_1)x^2 + 2b_0h^2) \\ &= 12(2b_0 + b_1)x^2h^2 + 24b_0h^4 \end{aligned}$$

από όπου παίρνουμε τη νέα σχέση $24b_0 = 2$ συνεπώς

$$b_0 = b_2 = \frac{1}{12} \text{ και } b_1 = \frac{10}{12}$$

Τελικά έχουμε τη μέθοδο του Numerov

$$y_{n+1} - 2y_n + y_{n-1} = h^2 \left(\frac{1}{12} f_{n+1} + \frac{10}{12} f_n + \frac{1}{12} f_{n-1} \right)$$

ή

$$y_{n+1} - 2y_n + y_{n-1} = \frac{h^2}{12} (f_{n+1} + 10f_n + f_{n-1})$$

Η παραπάνω μέθοδος είναι έμμεση και είναι ακριβής για πολυώνυμα τέταρτου βαθμού. Προκειμένου να πάρουμε μια άμεση μέθοδο θέτουμε $b_2 = b_0 = 0$ τότε η μέθοδος είναι της μορφής

$$y_{n+1} + a_1 y_n + a_0 y_{n-1} = h^2 b f_n \quad (1.16)$$

Σύμφωνα με τα προηγούμενα ισχύει

$$a_0 = 1 \text{ και } a_1 = -2 \quad (1.17)$$

Συνεπώς το αριστερό μέλος έχει προσδιοριστεί και μπορούμε να γράψουμε για τη μέθοδο

$$y_{n+1} - 2y_n + y_{n-1} = h^2 b f_n$$

Ζητώντας η μέθοδος να είναι ακριβής για

$$y(x) = x^2 \text{ τότε } y'(x) = 2x \text{ και } y''(x) = 2$$

και βρίσκουμε

$$\begin{aligned}(x+h)^2 - 2x^2 + (x-h)^2 &= h^2 b^2 \Rightarrow \\ (x+h)^2 - 2x^2 + (x-h)^2 &= 2bh^2 \Rightarrow b = 1\end{aligned}$$

Τελικά έχουμε τη μέθοδο

$$y_{n+1} - 2y_n + y_{n-1} = h^2 f_n$$

η οποία είναι ακριβής για πολυώνυμα δεύτερου βαθμού.

1.3 Προβλήματα Δοκιμής

Θα παρουσιάσουμε κάποια προβλήματα στα οποία θα εφαρμόσουμε τις αριθμητικές μεθόδους.

1.3.1 Πρόβλημα 1

Να εφαρμοσθεί η μέθοδος Euler στο πρόβλημα αρχικών τιμών

$$y' = x y + 2 x, \quad x \in [0, 1], \quad y(0) = 1.$$

Θα εφαρμόσουμε τη μέθοδο του Euler με $n = 10$ σημεία τότε το βήμα της μεθόδου είναι $h = 0.1$. Έχουμε $y_0 = 1$, υπολογίζουμε τις τρεις επόμενες προσεγγίσεις

$$\begin{aligned}y_1 &= y_0 + h f(x_0, y_0) = 1 + 0.1(0 \cdot 1 + 2 \cdot 0) = 1 \\ y_2 &= y_1 + h f(x_1, y_1) = 1 + 0.1(0.1 \cdot 1 + 2 \cdot 0.1) = 1.03 \\ y_3 &= y_2 + h f(x_2, y_2) = 1.03 + 0.1(0.2 \cdot 1.03 + 2 \cdot 0.2) = 1.0906\end{aligned}$$

Η αναλυτική λύση είναι:

$$y(x) = -2 + 3 e^{\frac{x^2}{2}}.$$

στα σημεία που υπολογίστηκαν οι προσεγγίσεις έχουμε τις τιμές της αναλυτικής λύσης

$$\begin{aligned}y(x_1) &= y(0.1) = 1.0150 \\ y(x_2) &= y(0.2) = 1.0606 \\ y(x_3) &= y(0.3) = 1.1381\end{aligned}$$

Γράφουμε μια συνάρτηση στο MATLAB η οποία υλοποιεί την μέθοδο Euler

```
function [x,y] = myEuler(fun,y0,a,b,N)
h = (b-a)/N ;
x = zeros(N+1,1) ;
y = zeros(N+1,1) ;
x(1) = a ;
```

```

y(1) = y0 ;
for n=1:N
    x(n+1) = x(n) + h ;
    y(n+1) = y(n) + h*feval(fun,x(n),y(n)) ;
end
end

```

Τα ορίσματα εισόδου είναι:

- fun η συνάρτηση $f(x, y)$,
- y0 η αρχική συνθήκη,
- a, b τα άκρα του διαστήματος ολοκλήρωσης $[a, b]$,
- N το πλήθος των υποδιαστημάτων.

Τα ορίσματα εξόδου είναι:

- x πίνακας διάστασης $(N \times 1)$ και περιέχει τα σημεία υπολογισμού της αριθμητικής λύσης (σημεία της διαμέρισης),
- y πίνακας διάστασης $(N \times 1)$ και περιέχει την αριθμητική λύση στα σημεία του x.

Ορίζουμε τη συνάρτηση δεξιού μέλους $f(x, y) = xy + 2x$ ως εξής

```

function f = ode1(x,y)
f = x*y+2*x ;
end

```

Λύνουμε το πρόβλημα στο $[0, 1]$ με 11 σημεία ($N = 10$ υποδιαστήματα)

```
>> [x,y]=myEuler('ode1',1,0,1,10);
```

Μπορούμε να γράψουμε τη συνάρτηση δεξιού μέλους και σαν inline function ως εξής:

```
[x,y]=myEuler(@(x,y) x*y+2*x,1,0,1,10)
```

```
>> [x,y]
```

```
ans =
```

0	1.0000
0.1000	1.0000
0.2000	1.0300
0.3000	1.0906
0.4000	1.1833
0.5000	1.3107
0.6000	1.4762
0.7000	1.6848
0.8000	1.9427
0.9000	2.2581
1.0000	2.6413

Βλέπουμε ότι οι τιμές συμφωνούν με αυτές που υπολογίσαμε παραπάνω. Δημιουργούμε διάγραμμα y_i το οποίο περιέχει την αναλυτική λύση στα σημεία x

```
>> yt=3*exp(x.^2/2)-2;
>> [x,y,yt]
```

x_i	y_i	$y(x_i)$
0	1.0000	1.0000
0.1000	1.0000	1.0150
0.2000	1.0300	1.0606
0.3000	1.0906	1.1381
0.4000	1.1833	1.2499
0.5000	1.3107	1.3994
0.6000	1.4762	1.5917
0.7000	1.6848	1.8329
0.8000	1.9427	2.1314
0.9000	2.2581	2.4979
1.0000	2.6413	2.9462

Μπορούμε να δημιουργήσουμε μια επιπλέον στήλη με την απόλυτη τιμή του σφάλματος

```
>> [x,y,yt,abs(y-yt)]
```

x_i	y_i	$y(x_i)$	$ y_i - y(x_i) $
0	1.0000	1.0000	0
0.1000	1.0000	1.0150	0.0150
0.2000	1.0300	1.0606	0.0306
0.3000	1.0906	1.1381	0.0475
0.4000	1.1833	1.2499	0.0665
0.5000	1.3107	1.3994	0.0888
0.6000	1.4762	1.5917	0.1155
0.7000	1.6848	1.8329	0.1481
0.8000	1.9427	2.1314	0.1887
0.9000	2.2581	2.4979	0.2398
1.0000	2.6413	2.9462	0.3048

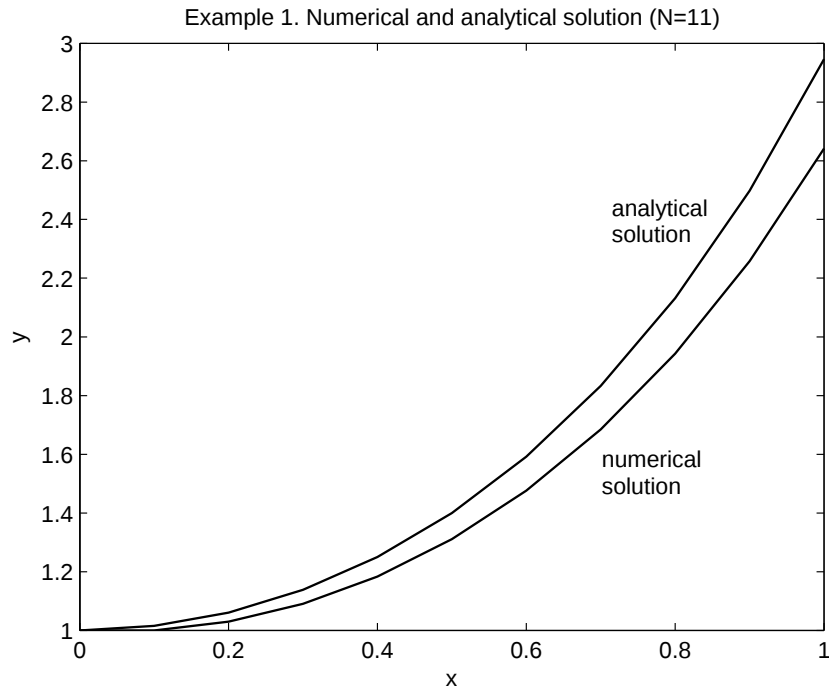
Το μέγιστο κατά απόλυτη τιμή σφάλμα είναι 0.3048 στο τέλος του διαστήματος ολοκλήρωσης.

Για να πάρουμε τις γραφικές παραστάσεις της αριθμητικής και της αναλυτικής λύσης (??) για τα παραπάνω σημεία γράφουμε

```
>> plot(x,y,x,yt)
```

Για τη γραφικές παράσταση του απόλυτου σφάλματος $|y_i - y(x_i)|$ (??)γράφουμε

```
>> plot(x,abs(y-yt))
```

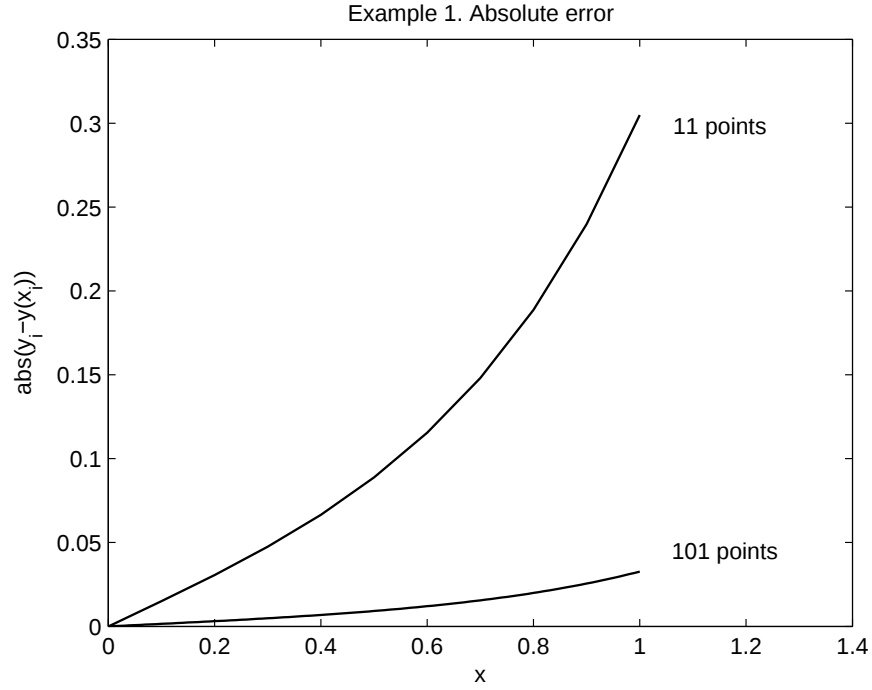


Σχήμα 1.1: Πρόβλημα 1. Αριθμητική και αναλυτική λύση.

Αν πάρουμε $N = 100$ το μέγιστο κατά απόλυτη τιμή σφάλμα είναι 0.0327 στο τέλος του διαστήματος ολοκλήρωσης, για $N = 1000$ είναι 0.0033. Στον πίνακα που ακολουθεί βλέπουμε τις προγεγρείσεις της λύσης με 10, 100 και 1000 διαστήματα στα σημεία x_i και το πως αυτές πλησιάζουν την αναλυτική λύση.

x_i	$y_i(N = 10)$	$y_i(N = 100)$	$y_i(N = 1000)$	$y(x_i)$
0	1.0000	1.0000	1.0000	1.0000
0.1000	1.0000	1.0135	1.0149	1.0150
0.2000	1.0300	1.0575	1.0603	1.0606
0.3000	1.0906	1.1332	1.1376	1.1381
0.4000	1.1833	1.2430	1.2492	1.2499
0.5000	1.3107	1.3903	1.3985	1.3994
0.6000	1.4762	1.5796	1.5904	1.5917
0.7000	1.6848	1.8173	1.8313	1.8329
0.8000	1.9427	2.1115	2.1294	2.1314
0.9000	2.2581	2.4724	2.4953	2.4979
1.0000	2.6413	2.9135	2.9429	2.9462

αντίστοιχα το απόλυτο σφάλμα είναι:



Σχήμα 1.2: Πρόβλημα 1. Απόλυτο σφάλμα.

x_i	error ($N = 10$)	error ($N = 100$)
0	0	0
0.1000	0.0150	0.0015
0.2000	0.0306	0.0031
0.3000	0.0475	0.0048
0.4000	0.0665	0.0068
0.5000	0.0888	0.0092
0.6000	0.1155	0.0120
0.7000	0.1481	0.0155
0.8000	0.1887	0.0199
0.9000	0.2398	0.0255
1.0000	0.3048	0.0327

Σημειώνεται ότι πρέπει να ανανεώσουμε το διάνυσμα που περιέχει την αναλυτική λύση για να ταιριάζουν τα μεγέθη των διανυσμάτων. Βλέπουμε ότι όσο λεπταίνει η διαμέριση (μειώνεται το βήμα) το σφάλμα γίνεται μικρότερο δηλαδή η εκτίμηση μας περισσότερο ακριβής.

Θα εφαρμόσουμε την μέθοδο Adams Bashforth (1.10) στην επίλυση του προβλήματος θεωρώντας 10 υποδιαστήματα.

Για να ξεκινήσουμε χρειαζόμαστε δύο εκτιμήσεις την y_0 και την y_1 , για να βρούμε την y_2 . Επειδή η τιμή y_1 είναι άγνωστη θα χρησιμοποιήσουμε τη μέθοδο Euler για να την εκτιμήσουμε οπότε παίρνουμε $y_1 = 1$. Από όπου βρίσκουμε

$$f(x_1, y_1) = x_1(y_1 + 2) = 0.1 \times (1 + 2) = 0.3$$

. Στη συνέχεια υπολογίζουμε το y_2

$$\begin{aligned} y_2 &= y_1 + \frac{h}{2} (3f(x_1, y_1) - f(x_0, y_0)) = \\ &= 1 + \frac{0.1}{2} (3 \times 0.3 - 0) = 1.0450 \end{aligned}$$

έχουμε

$$f(x_2, y_2) = x_2(y_2 + 2) = 0.2 \times (1.045 + 2) = 0.6090$$

υπολογίζουμε το y_3

$$\begin{aligned} y_3 &= y_2 + \frac{h}{2} (3f(x_2, y_2) - f(x_1, y_1)) = \\ &= 1.0450 + \frac{0.1}{2} (3 \times 0.6090 - 0.3) = 1.1213 \end{aligned}$$

θ	1.0000
0.1000	1.0000
0.2000	1.0450
0.3000	1.1213
0.4000	1.2314
0.5000	1.3784
0.6000	1.5672
0.7000	1.8038
0.8000	2.0961
0.9000	2.4545
1.0000	2.8921

Το μέγιστο απόλυτο σφάλμα για $N = 10$ είναι 0.0541, ενώ για $N = 100$ είναι $6.0149e - 04$.

1.3.2 Πρόβλημα 2

Ας θεωρήσουμε ένα ακόμα πρόβλημα αρχικών τιμών το

$$y' + y = x \sin(x)$$

στο διάστημα $[0, 4\pi]$ με αρχική συνθήκη $y(0) = 1$, και πάλι είναι μια γραμμική ΔΕ πρώτης τάξης με σταθερούς συντελεστές της οποίας η αναλυτική λύση είναι

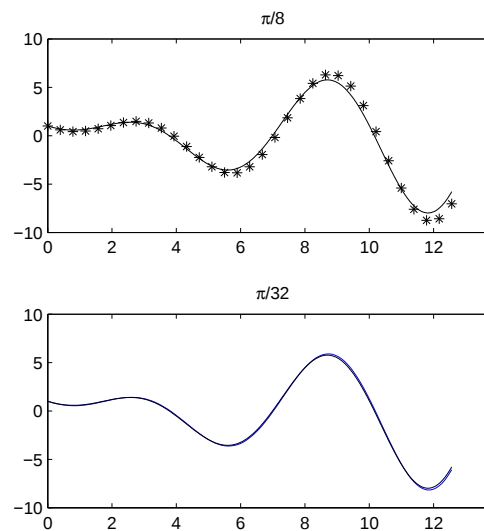
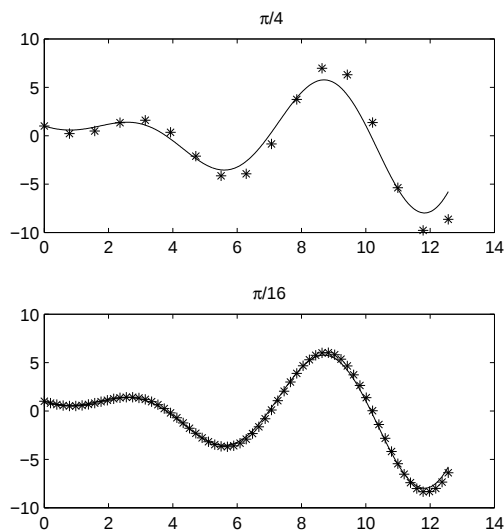
$$y(x) = \frac{1}{2} (e^{-x} + \cos(x) - x \cos(x) + x \sin(x))$$

Καλούμε την myEuler με $N = 16$, $N = 32$, $N = 64$, $N = 128$ διαστήματα, δηλαδή βήμα $h = \pi/4$, $h = \pi/8$, $h = \pi/16$ και $h = \pi/32$. Θα χρησιμοποιήσουμε την εντολή subplot και θα κατασκευάσουμε έναν πίνακα 2×2 με τις γραφικές παραστάσεις της αριθμητικής και της αναλυτικής λύσης

```

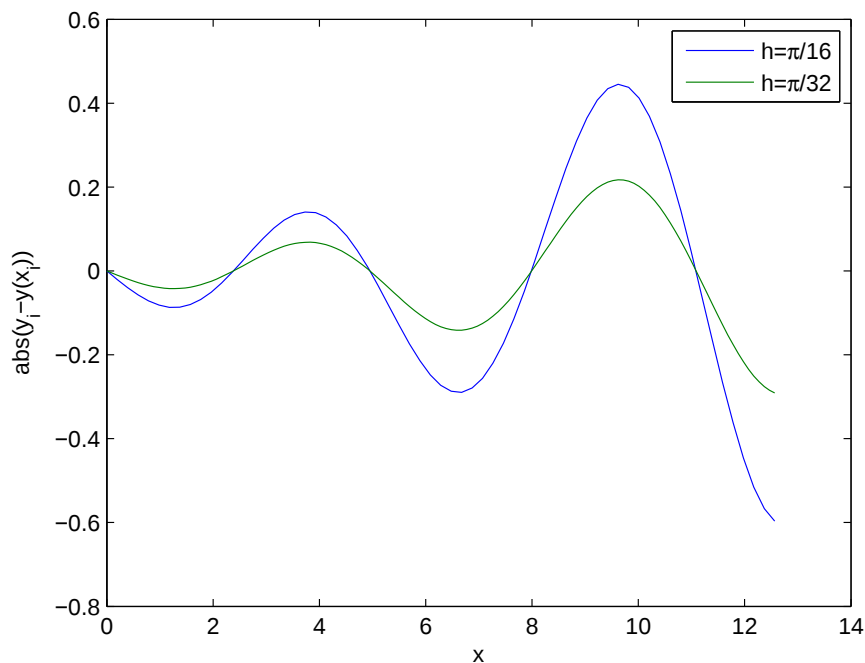
>> xt = 0:pi/128:4*pi;
>> yt=(exp(-xt)+cos(xt)-xt.*(cos(xt)-sin(xt)))/2;
>> [x,y]=myEuler(@(x,y) -y+x*sin(x),1,0,4*pi,16);
>> subplot(2,2,1)
>> plot(x,y,'*',xt,yt)
>> [x,y]=myEuler(@(x,y) -y+x*sin(x),1,0,4*pi,32);
>> subplot(2,2,2)
>> plot(x,y,'*',xt,yt)
>> [x,y]=myEuler(@(x,y) -y+x*sin(x),1,0,4*pi,64);
>> subplot(2,2,3)
>> plot(x,y,'*',xt,yt)
>> [x,y]=myEuler(@(x,y) -y+x*sin(x),1,0,4*pi,128);
>> subplot(2,2,4)
>> plot(x,y,xt,yt)
>>

```



Σχήμα 1.3: Πρόβλημα 2. Η αριθμητική και η αναλυτική λύση στο $[0, 4\pi]$ για βήματα $\pi/4$, $\pi/8$, $\pi/16$, $\pi/32$.

Επίσης εμφανίζουμε το σφάλμα της αριθμητικής λύσης $\text{error}=\text{abs}(y-y_t)$ για βήματα $\pi/16$, $\pi/32$.



Σχήμα 1.4: Πρόβλημα 2. Σφάλμα της λύσης στο $[0, 4\pi]$ για βήματα $\pi/16, \pi/32$.

1.3.3 Πρόβλημα 3

Θα δούμε τώρα ένα σύστημα με 3 διαφορικές εξισώσεις (εξισώσεις Euler) που περιγράφουν την κίνηση του στερεού σώματος (rigid body) στο οποίο δεν ασκούνται δυνάμεις.

$$\begin{aligned} y_1' &= (\alpha - \beta) y_2 y_3 \\ y_2' &= (1 - \alpha) y_1 y_3 \\ y_3' &= (\beta - 1) y_1 y_2 \end{aligned}$$

με αρχικές συνθήκες

$$y_1(0) = 0, \quad y_2(0) = 1, \quad y_3(0) = 1$$

και παραμέτρους

$$\alpha = 1 + \frac{1}{\sqrt{1.51}}, \quad \beta = 1 - \frac{0.51}{\sqrt{1.51}}$$

Η αναλυτική λύση είναι

$$y_1(x) = \sqrt{1.51} \operatorname{sn}(x, 0.51), \quad y_2(x) = \operatorname{cn}(x, 0.51), \quad y_3(x) = \operatorname{dn}(x, 0.51)$$

όπου sn , cn , dn είναι οι ελλειπτικές συναρτήσεις Jacobi. Το συναρτήσεις είναι περιοδικές με περίοδο $T = 7.45056320933095$.

Θα γράψουμε τη μέθοδο Euler για ένα σύστημα τριών διαφορικών εξισώσεων

```
function [x,y1,y2,y3]=Euler3(fun,y10,y20,y30,x0,xn,N)
% Euler method for a system of 3 odes
%
% step length
h = (xn-x0)/N ;
%
% creates the vectors to be used
x = zeros(N+1,1);
y1 = zeros(N+1,1);
y2 = zeros(N+1,1);
y3 = zeros(N+1,1);
% starting point of the integration interval
x(1) = x0 ;
%
% initial condition
y1(1) = y10 ;
y2(1) = y20;
y3(1) = y30;
%
for i=1:N
    x(i+1) = x(i) + h ;
    [f1,f2,f3]=feval(fun,y1(i),y2(i),y3(i));
    y1(i+1) = y1(i) + h*f1 ;
    y2(i+1) = y2(i) + h*f2 ;
    y3(i+1) = y3(i) + h*f3 ;
end
end
```

σε ένα mfile που εδώ ονομάζουμε rigid.m θα γράψουμε τις συναρτήσεις δεξιού μέλους

```
function [f1,f2,f3] = rigid(y1,y2,y3)
% rigid body
a = 1 + 1/sqrt(1.51) ;
b = 1 - 0.51/sqrt(1.51) ;
f1 = (a-b)*y2*y3 ;
f2 = (1-a)*y1*y3 ;
f3 = (b-1)*y1*y2 ;
end
```

Εφαρμόζουμε το function με αρχικές συνθήκες $(0, 1, 1)$, διάστημα ολοκλήρωσης το $[0, 2\pi]$ και $N + 1 = 10001$ σημεία διαμέρισης

```
>>[x,y1,y2,y3]=Eulers('rigid',0,1,1,0,2*pi,10000);
```

υπολογίζουμε την αναλυτική λύση στα σημεία της διαμέρισης x_i με τη συνάρτηση `ellipj` του MATLAB

```
>> [sn,cn,dn]=ellipj(x,0.51) ;
```

τότε το μέγιστο απόλυτο σφάλμα για κάθε συντεταγμένη της λύσης είναι

```
>> max(abs(sqrt(1.51)*sn-y1))
```

```
ans = 0.209809750974931
```

```
>> max(abs(cn-y2))
```

```
ans = 0.223034154080894
```

```
>> max(abs(dn-y3))
```

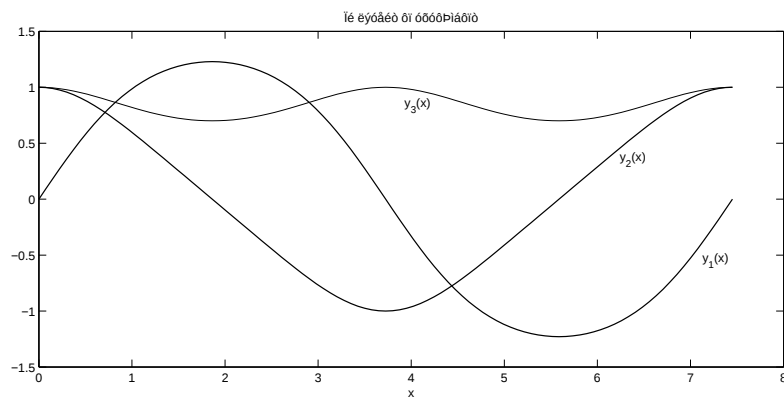
```
ans = 0.080089029586601
```

Για να εκτιμήσουμε το συνολικό σφάλμα υπολογίζουμε την νόρμα

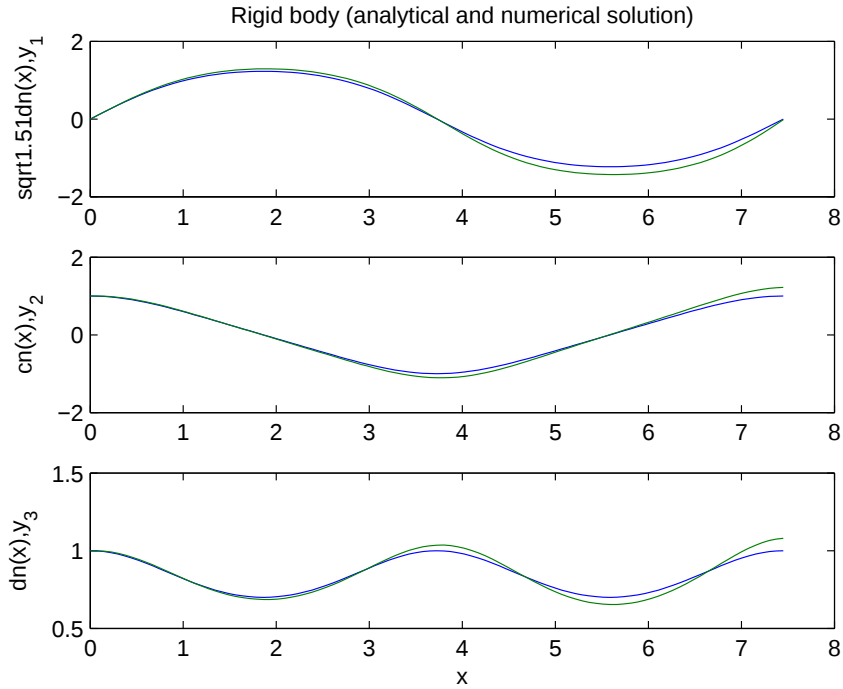
$$\sqrt{(y_1i - y_1(x_i))^2 + (y_2i - y_2(x_i))^2 + (y_3i - y_3(x_i))^2}$$

```
>> max(sqrt((sqrt(1.51)*sn-y1).^2+(cn-y2).^2+(dn-y3).^2))
```

```
ans = 0.237688030916049
```



Σχήμα 1.5: Πρόβλημα 3. Αναλυτική λύση.



Σχήμα 1.6: Πρόβλημα 3. Αριθμητική και αναλυτική λύση.

Θα κατασκευάσουμε ένα function το οποίο να είναι ανεξάρτητο του αριθμού των διαφορικών εξισώσεων που θέλουμε να λύσουμε. Ο αριθμός των εξισώσεων n δεν θα είναι όρισμα εισόδου αλλά θα προκύπτει από το μήκος του διάνυσματος των αρχικών συνθηκών ($n = \text{lenght}(y0)$), οι προσεγγίσεις των συναρτήσεων $y_i(x)$ θα αποθηκεύονται σαν στήλες στον πίνακα y έτσι ώστε το στοιχείο $y(i, j)$ να είναι η προσέγγιση της συνάρτησης $y_j(x)$ στο σημείο x_i . Ανάλογα θα πρέπει να τροποποιηθεί και η συνάρτηση ώστε να δέχεται ως όρισμα το διάνυσμα των συναρτήσεων στο κάθε σημείο x_i

$$(y_1(x_i), y_2(x_i), \dots, y_n(x_i))$$

και να επιστρέφει το διάνυσμα των συναρτήσεων δεξιού μέλους

$$(f_1, f_2, \dots, f_n)$$

.

Ονομάζουμε το function Eulers και είναι το ακόλουθο

```
function [x,y]=Eulers(fun,y0,x0,xn,N)
% Euler method for system of odes
%
n = length(y0) ;
```

```

%
% step length
h = (xn-x0)/N ;
%
% creates the vectors to be used
x = zeros(N+1,1);
y = zeros(N+1,n);
% starting point of the integration interval
x(1) = x0 ;
%
% initial condition
y(1,:) = y0 ;
%
for i=1:N
    x(i+1) = x(i) + h ;
    f=feval(fun,y(i,:));
    y(i+1,:) = y(i,:) + h*f ;
end
end

function f = rigid_a(y)
% rigid body
n = length(y) ;
f = zeros(1,n);
a = 1 + 1/sqrt(1.51) ;
b = 1 - 0.51/sqrt(1.51) ;
f(1) = (a-b)*y(2)*y(3) ;
f(2) = (1-a)*y(1)*y(3) ;
f(3) = (b-1)*y(1)*y(2) ;
end

```

1.3.4 Πρόβλημα 4

Το ταλαντωτικό πρόβλημα $[?], [?]$

$$y'' = -\omega^2 y + (\omega^2 - 1)\sin x, \quad y(0) = 1, \quad y'(0) = \omega + 1, \quad x \geq 0.$$

Η ακριβής λύση είναι

$$y(x) = \cos(\omega x) + \sin(\omega x) + \sin(x),$$

Εδώ έχουμε μια διαφορική εξίσωση δεύτερης τάξης την οποία θα μετασχηματίσουμε σε ένα σύστημα δύο διαφορικών εξισώσεων πρώτης τάξης και στη συνέχεια θα το λύσουμε με τη μέθοδο Euler. Θέτουμε

$$\begin{aligned} y_1(x) &= y(x) \\ y_2(x) &= y'(x) \end{aligned}$$

και παίρνουμε το σύστημα

$$\begin{aligned}y_1' &= y_2 \\y_2' &= -\omega^2 y_1 + (\omega^2 - 1)\sin x\end{aligned}$$

με αρχικές συνθήκες

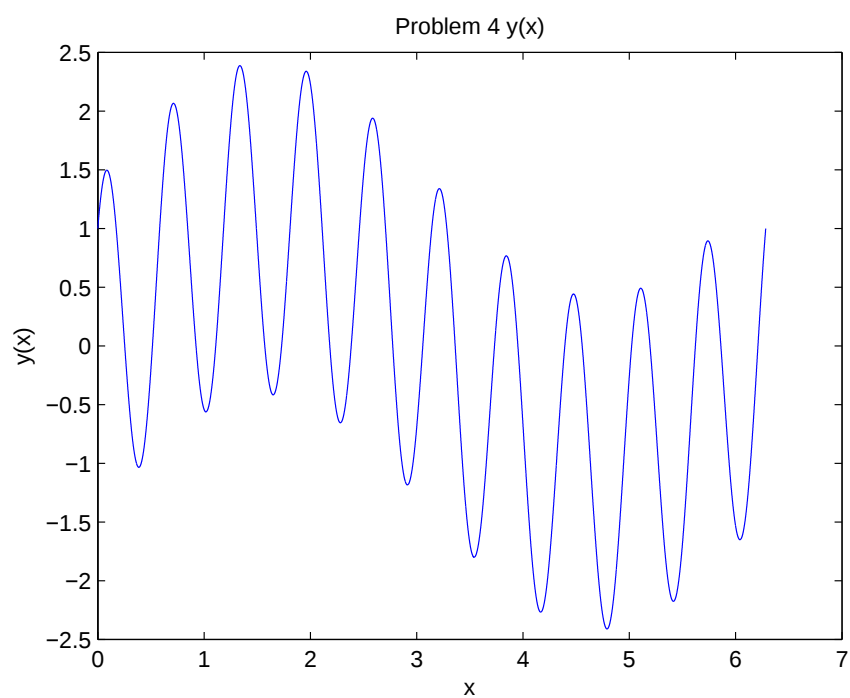
$$\begin{aligned}y_1(0) &= 1 \\y_2(0) &= \omega + 1\end{aligned}$$

Θα λύσουμε το σύστημα για $\omega = 10$ στο διάστημα $[0, 1]$

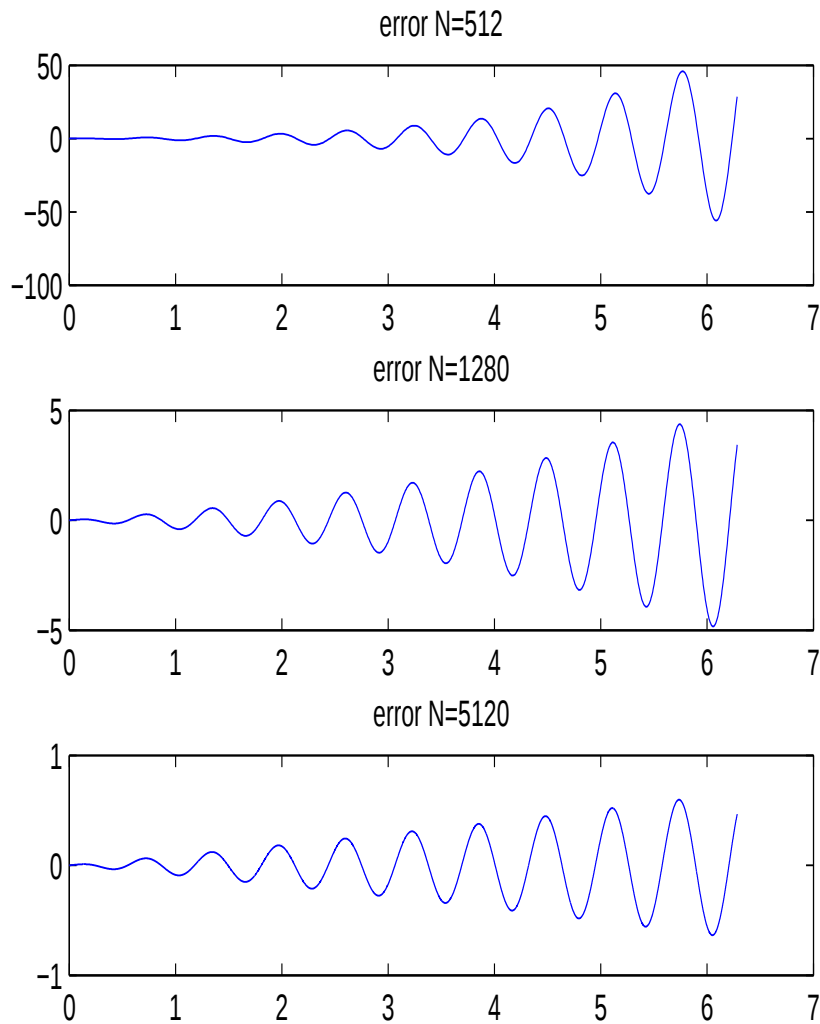
$$\begin{aligned}y_1' &= y_2, \quad y_1(0) = 1 \\y_2' &= -100y_1 + 99\sin x, \quad y_2(0) = 11\end{aligned}$$

Η λύση αποτελείται από μια αργά και μία γρήγορα ταλαντωτική συνάρτηση εξαιτίας του μη ομογενούς όρου.

```
[x,y]=Eulersx(@(x,y) [y(2) -100*y(1)+99*sin(x)], [1 11], 0, 2*pi, 12800);
```



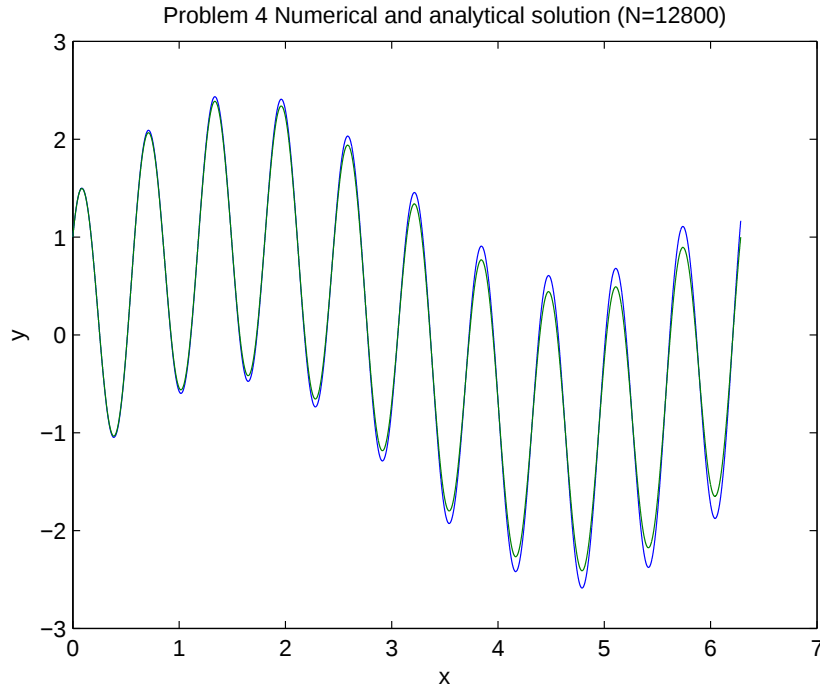
Σχήμα 1.7: Πρόβλημα 4. Αναλυτική λύση.



Σχήμα 1.8: Πρόβλημα 4. Σφάλμα της λύσης για $N = 512$, $N = 1280$, $N = 5120$.

1.3.5 Πρόβλημα 5

Το πρόβλημα δύο σωμάτων που αρχικά μελετήθηκε από τον Kepler αποτελεί ένα τυπικό πρόβλημα δοκιμής. Περιγράφει την κίνηση ενός πλανήτη γύρω από τον ήλιο όταν αγνοήσουμε την ύπαρξη των άλλων πλανητών. Τότε η κίνηση περιγράφεται από το νόμο της βαρύτητας του Νεύτωνα για δύο σώματα



Σχήμα 1.9: Πρόβλημα 4. Αριθμητική και αναλυτική λύση για $N = 12800$ σημεία.

$$mr'' = -\frac{GmM}{||r||^3}r$$

όπου M η μάζα του ήλιου, m η μάζα του πλανήτη, G η βαρυτική σταθερά και r η απόσταση των δύο σωμάτων. Αν τοποθετήσουμε τον ήλιο στην αρχή των αξόνων τότε το διάνυσμα $r(t) = (x(t), y(t))$ είναι η θέση του πλανήτη τη χρονική στιγμή t . Από τον πρώτο νόμο του Kepler γνωρίζουμε ότι η τροχιά είναι ελλειπτική με τον ήλιο στη μία εστία.

Θεωρούμε το κανονικοποιημένο πρόβλημα όπου οι μάζες και η βαρυτική σταθερά είναι ίσες με τη μονάδα. Θέτουμε $q_1(t) = x(t)$ και $q_2(t) = y(t)$ οπότε προκύπτει το ακόλουθο σύστημα

$$\begin{aligned} q_1'' &= -\frac{q_1}{\sqrt{(q_1^2 + q_2^2)^3}}, \\ q_2'' &= -\frac{q_2}{\sqrt{(q_1^2 + q_2^2)^3}}. \end{aligned}$$

Μπορούμε να το γράψουμε και σαν ένα σύστημα διαφορικών εξισώσεων πρώτης τάξης

$$\begin{aligned} q_1' &= p_1 \\ p_1' &= -\frac{q_1}{\sqrt{(q_1^2 + q_2^2)^3}} \\ q_2' &= p_2 \\ p_2' &= -\frac{q_2}{\sqrt{(q_1^2 + q_2^2)^3}}. \end{aligned}$$

Θεωρούμε τις αρχικές συνθήκες

$$\begin{aligned} q_1(0) &= 1 - e \\ p_1(0) &= 0 \\ q_2(0) &= 0 \\ p_2(0) &= \sqrt{\frac{1+e}{1-e}} \end{aligned}$$

όπου $e, 0 \leq e < 1$ είναι η εκκεντρότητα της τροχιάς.

Η ελλειπτική τροχιά που αναπαριστά τη λύση του προβλήματος αρχικών τιμών στο επίπεδο (q_1, q_2) είναι η έλλειψη

$$(q_1 + e)^2 + \frac{q_2^2}{1 - e^2} = 1$$

Η λύση του προβλήματος είναι περιοδική με περίοδο 2π και δίνεται από τους τύπους

$$\begin{aligned} q_1(t) &= \cos(E) - e \\ q_2(t) &= \sqrt{1 - e^2} \sin(E) \end{aligned}$$

όπου E είναι η ανωμαλία της εκκεντρότητας και συνδέεται με το χρόνο μέσω της πεπλεγμένης εξίσωσης

$$t = E - e \sin(E).$$

Τότε μετά από παραγωγή προκύπτει

$$\begin{aligned} p_1(t) &= \frac{\sin(E)}{e \cos(E) - 1} \\ p_2(t) &= \frac{\sqrt{1 - e^2} \cos(E)}{1 - e \cos(E)} \end{aligned}$$

Για $e = 0$ η τροχιά είναι κυκλική και η αναλυτική λύση είναι

$$\begin{aligned} q_1(t) &= \cos(t), \\ q_2(t) &= \sin(t). \end{aligned}$$

Η χαμιλτονιανή του προβλήματος που εκφράζει τη συνολική ενέργεια του συστήματος είναι το άθροισμα της κινητικής ενέργειας $T(p_1, p_2)$ και της δυναμικής ενέργειας $V(q_1, q_2)$

$$H(p_1, p_2, q_1, q_2) = T(p_1, p_2) + V(q_1, q_2),$$

όπου

$$T(p_1, p_2) = \frac{1}{2}(p_1^2 + p_2^2),$$

$$V(q_1, q_2) = -\frac{1}{\sqrt{q_1^2 + q_2^2}}.$$

Θα εφαρμόσουμε τη μέθοδο Euler στην αριθμητική ολοκλήρωση του προβλήματος αυτού. Θεωρούμε το διάνυσμα λύσης y το οποίο έχει 4 συντεταγμένες τα q_1, p_1, q_2, p_2 ως εξής

$$y(t) = \begin{pmatrix} q_1(t) \\ p_1(t) \\ q_2(t) \\ p_2(t) \end{pmatrix}$$

τότε

$$y'(t) = \begin{pmatrix} y_2(t) \\ -\frac{y_1(t)}{\sqrt{((y_1(t))^2 + y_3(t)^2)^3}} \\ y_4(t) \\ -\frac{y_3(t)}{\sqrt{((y_1(t))^2 + y_3(t)^2)^3}} \end{pmatrix}$$

Τα ορίσματα εξόδου είναι το x (ένας πίνακας διάστασης $N \times 1$) που περιέχει τα σημεία (της διαμέρισης) στα οποία υπολογίζεται η αριθμητική λύση και το διάνυσμα y (ένας πίνακας διάστασης $N \times 4$) που περιέχει σε κάθε στήλη κάθε συντεταγμένη της αριθμητικής λύσης (στην πρώτη στήλη την $y_1(t)$, στην δεύτερη την $y_2(t)$, κοκ.) σε κάθε στοιχείο του x .

t	$y_1(t)$	$y_2(t)$	$y_3(t)$	$y_4(t)$
x	$y(:, 1)$	$y(:, 2)$	$y(:, 3)$	$y(:, 4)$
$x(1)$	$y(1, 1)$	$y(1, 2)$	$y(1, 3)$	$y(1, 4)$
$x(2)$	$y(2, 1)$	$y(2, 2)$	$y(2, 3)$	$y(2, 4)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x(N)$	$y(N, 1)$	$y(N, 2)$	$y(N, 3)$	$y(N, 4)$

Υπολογίζουμε το μέγιστο κατά απόλυτη τιμή σφάλμα της πρώτης και στη συνέχεια της τρίτης συντεταγμένης. Τέλος υπολογίζουμε το σφάλμα της ευκλείδειας νόρμας

$$\sqrt{((y(:, 1) - \cos(x))^2 + (y(:, 3) - \sin(x))^2)}$$

```
>> [x,y]=Euler-kepler('Fun',0,2*pi,1024);
>> max(abs(y(:,1)-cos(x)))
ans =
    0.1896
>> max(abs(y(:,3)-sin(x)))
ans =
    0.3502
>> max(sqrt(y(:,1)-cos(x)).^2+(y(:,3)-sin(x)).^2)
ans =
    0.1391
>>
```

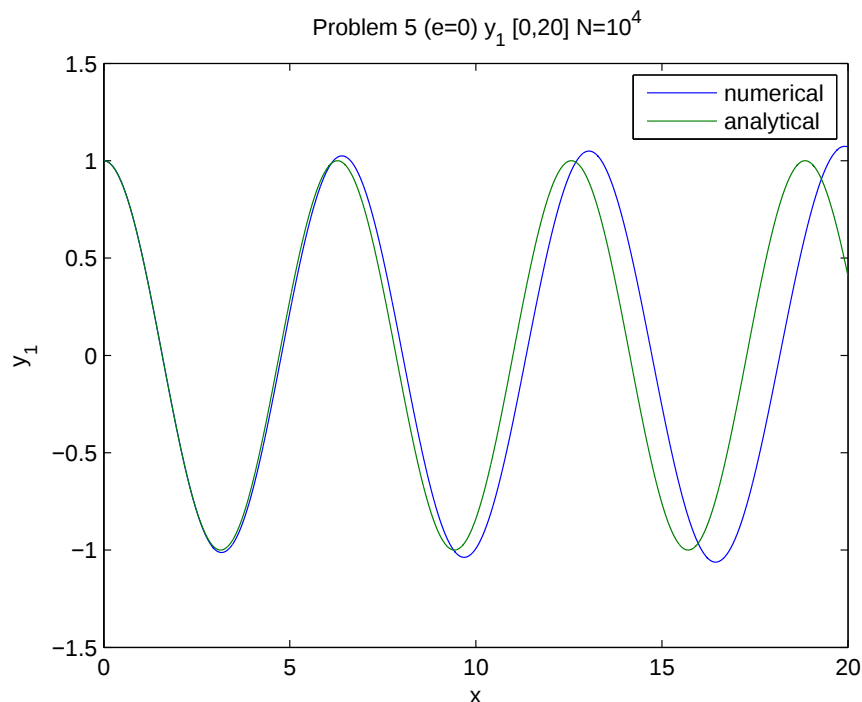
Αν πάρουμε περισσότερα σημεία $N = 10^4$ παρατηρούμε ότι τα σφάλματα μειώνονται

```
>> [x,y]=Euler-kepler('Fun',0,2*pi,10000);
>> max(abs(y(:,1)-cos(x)))
ans =
    0.0194
>> max(abs(y(:,3)-sin(x)))
ans =
    0.0371
>> max(sqrt(y(:,1)-cos(x)).^2+(y(:,3)-sin(x)).^2)
ans =
    0.0086
```

Παρουσιάζουμε την πρώτη και τρίτη συντεταγμένη της αριθμητικής λύσης $y(1)$ και $y(3)$ και την αντίστοιχη αναλυτική λύση $\cos(t)$ και $\sin(t)$ για το διάστημα ολοκλήρωσης $[0, 2\pi]$ με $N = 1024$ υποδιαστήματα. Χρειάζεται μικρό βήμα γιατί η μέθοδος Euler έχει πολύ μικρή ακρίβεια είναι μόλις πρώτης αλγεβρικής τάξης όπως θα δούμε στο τρίτο κεφάλαιο.

Το διάγραμμα (q_1, q_2) στην αναλυτική λύση είναι κύκλος $(\cos(t), \sin(t))$ και αυτό περιμένουμε και από την αριθμητική λύση. Το διάγραμμα (q_1, q_2) για την αριθμητική λύση με $N = 10^4$ το παίρνουμε γράφοντας

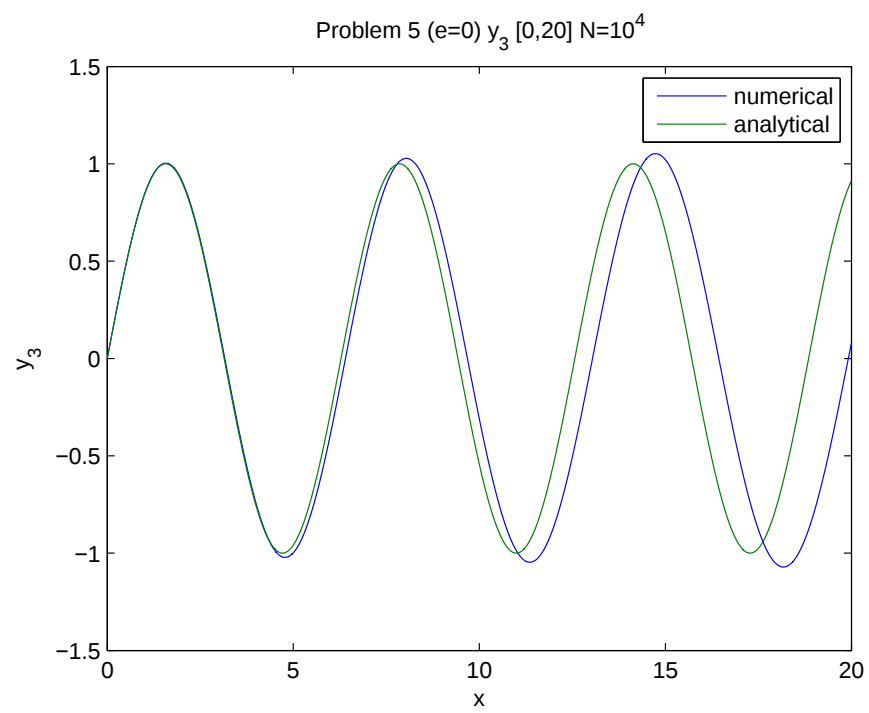
```
>> plot(y(:,1),y(:,3))
```



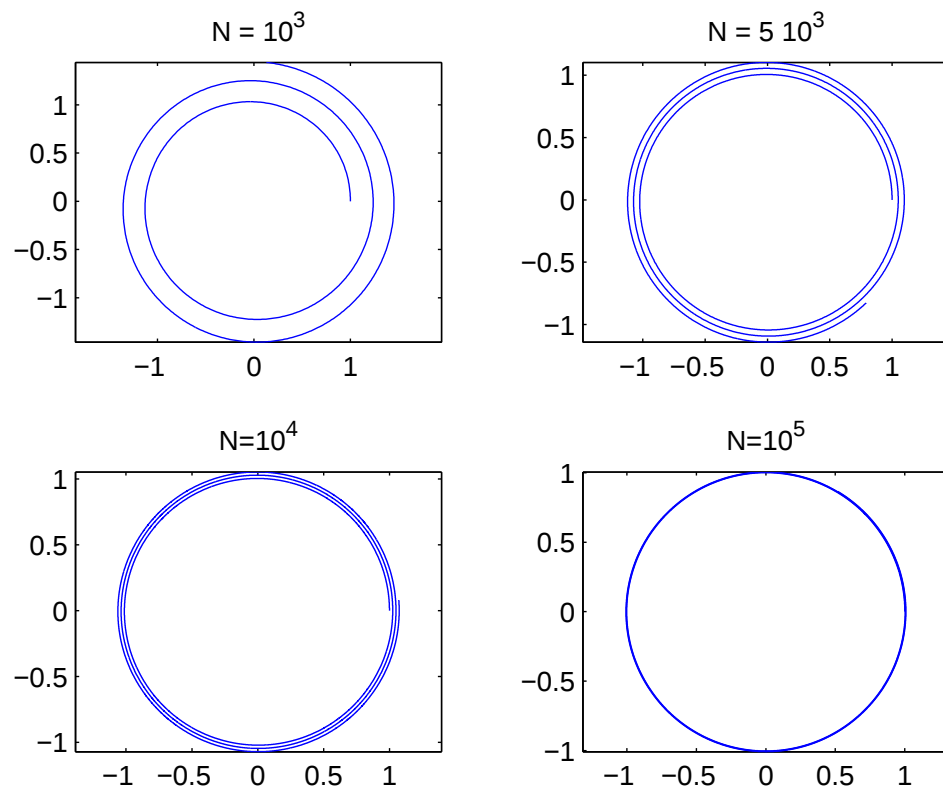
Σχήμα 1.10: Πρόβλημα 5 ($e = 0$) Αριθμητική και αναλυτική λύση για την $y_1(x)$, $[0, 20]$, $N = 10^4$.

1.4 Βιβλιογραφία

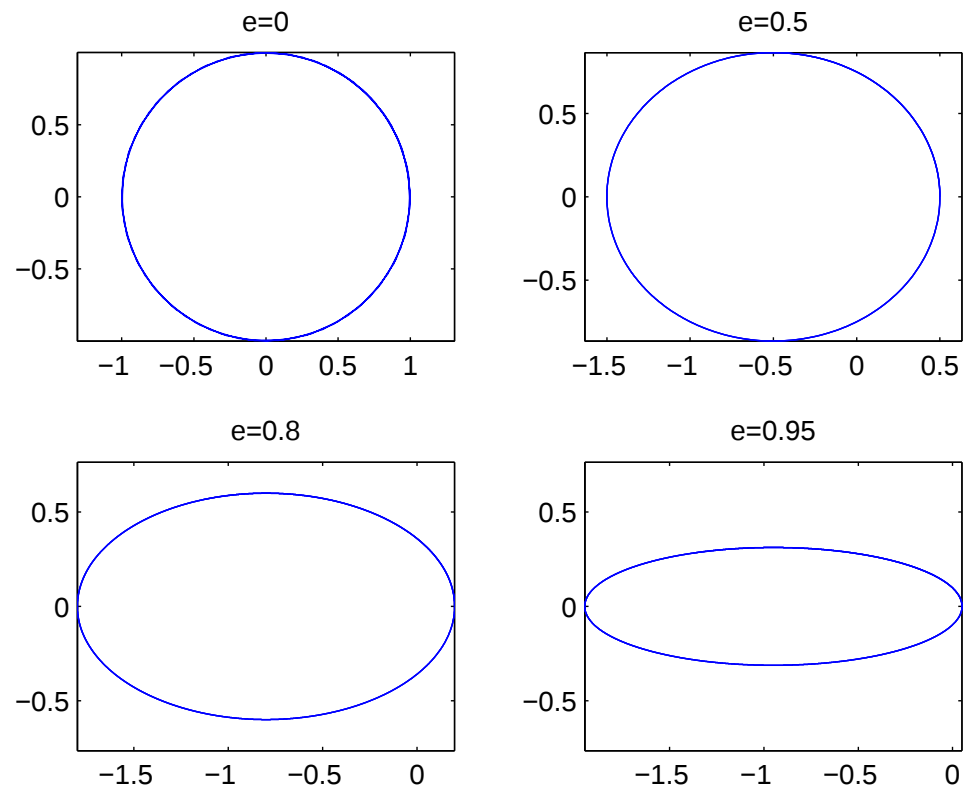
- Αβδελάς Γ., Σίμος Θ., Αριθμητική Ανάλυση, Εκδόσεις Συμεών, 2004.
 Βραχάτης Μ., Αριθμητική Ανάλυση, Εκδόσεις Κλειδάριθμος, 2012.
 Βραχάτης Μ., Αριθμητική Ανάλυση Συνήθεις Διαφορικές Εξισώσεις, Εκδόσεις Κλειδάριθμος, 2012.
 Γουσιδου - Κουτίτα Μ., Αριθμητική Ανάλυση, Εκδόσεις Αφοι Κυριακίδη, 2017.
 Μισυρλής Ν., Αριθμητική Ανάλυση, Εκδόσεις ΕΚΠΑ, 2017.
 Παπαγεωργίου Γ.Σ. και Τσίτουρας Χ., Αριθμητική Ανάλυση με εφαρμογές σε MATLAB και Mathematica, Εκδόσεις Τσότρας, 2015.
 Φαμέλης Ι., Υπολογιστικά Μαθηματικά, Εκδόσεις Κριτική, 2021.
 Moler C., Αριθμητικές Μέθοδοι με το MATLAB, Εκδόσεις Κλειδάριθμος, 2010.
 Hairer E., Norset S.P., Wanner G., Solving Ordinary Differential Equations, Springer 2000
 Butcher J., Numerical Methods for Ordinary Differential Equations, Wiley, 2016



Σχήμα 1.11: Πρόβλημα 5 ($e = 0$) Αριθμητική και αναλυτική λύση για την $y_3(x)$, $[0, 20]$, $N = 10^4$.



Σχήμα 1.12: Πρόβλημα 5 ($e = 0$) Διάγραμμα $(y_1(x), y_3(x))$, $[0, 20]$ για $N = 1000$, $N = 5000$, $N = 10^4$, $N = 10^5$.



Σχήμα 1.13: Πρόβλημα 5 Διάγραμμα $(y_1(x), y_3(x))$, $[0, 20]$ για $e = 0, e = 0.5, e = 0.8, e = 0.95$.

Κεφάλαιο 2

Μέθοδοι Runge-Kutta

2.1 Εισαγωγή

Θεωρούμε το πρόβλημα αρχικών τιμών

$$y' = f(x, y), \quad y(x_0) = y_0. \quad (2.1)$$

όπου y είναι ένα διάνυσμα διάστασης D από πραγματικές συναρτήσεις και f ένα διάνυσμα διάστασης D από πραγματικές συναρτήσεις των x και y . Όπως είδαμε στο κεφάλαιο 1 για να καταλήξουμε σε μια αριθμητική μέθοδο παίρνουμε μια προσέγγιση για το ορισμένο ολοκλήρωμα

$$\int_{x_n}^{x_{n+1}} f(x, y(x)) dx$$

από τον απλό τύπο του ορθογωνίου

$$\int_{x_n}^{x_{n+1}} f(x, y(x)) dx \approx h f(x_n, y(x_n)).$$

Από τη διαφορική εξίσωση έχουμε $y'(x) = f(x, y(x))$ και μπορούμε να γράψουμε

$$\int_{x_n}^{x_{n+1}} f(x, y(x)) dx = \int_{x_n}^{x_{n+1}} y'(x) dx = y(x_{n+1}) - y(x_n)$$

από τις δυο παραπάνω σχέσεις προκύπτει και πάλι η μέθοδος Euler

$$y(x_{n+1}) - y(x_n) \approx h f(x_n, y(x_n)).$$

Αν χρησιμοποιήσουμε τον απλό τύπο του τραπεζίου για την προσέγγιση του ολοκληρώματος

$$\int_{x_n}^{x_{n+1}} f(x, y(x)) dx \approx \frac{1}{2} h (f(x_n, y(x_n)) + f(x_{n+1}, y(x_{n+1}))).$$

$$y(x_{n+1}) - y(x_n) \approx \frac{1}{2} h (f(x_n, y(x_n)) + f(x_{n+1}, y(x_{n+1}))).$$

Αντικαθιστούμε $y(x_{n+1})$ και $y(x_n)$ με τις προσεγγίσεις y_{n+1} και y_n αντίστοιχα και παίρνουμε την μέθοδο

$$y_{n+1} = y_n + \frac{1}{2} h (f(x_n, y_n) + f(x_{n+1}, y_{n+1})).$$

παρατηρούμε ότι στο δεξί μέλος εμφανίζεται το y_{n+1} μια τέτοια μέθοδος καλείται έμμεση. Μπορούμε να την μετατρέψουμε σε άμεση αν αντικαταστήσουμε το y_{n+1} με ένα βήμα Euler

$$y_{n+1} = y_n + \frac{1}{2} h (f(x_n, y_n) + f(x_{n+1}, y_n + hf(x_n, y_n))).$$

αυτή είναι γνωστή σαν μέθοδος του Heun (1900).

Παρατηρούμε ότι με την μέθοδο Heun ο υπολογισμός του y_{n+1} απαιτεί τον υπολογισμό της συνάρτησης δεξιού μέλους $f(x, y)$ σε δύο σημεία (x_n, y_n) και $(x_{n+1}, y_n + hf(x_n, y_n))$. Συμβολίζουμε με k_1 και k_2 τις τιμές της συνάρτησης στα σημεία αυτά:

$$\begin{aligned} k_1 &= f(x_n, y_n) \\ k_2 &= f(x_{n+1}, y_n + hf(x_n, y_n)) = f(x_n + h, y_n + hk_1). \end{aligned}$$

τότε για τον υπολογισμό της εκτίμησης y_{n+1} μπορούμε να γράψουμε

$$y_{n+1} = y_n + \frac{h}{2} (k_1 + k_2).$$

Η μέθοδος Euler απαιτεί τον υπολογισμό της συνάρτησης $f(x, y)$ σε ένα μόνο σημείο και σύμφωνα με τα παραπάνω μπορεί να γραφεί

$$\begin{aligned} k_1 &= f(x_n, y_n) \\ y_{n+1} &= y_n + hk_1 \end{aligned}$$

τα k_1 και k_2 ονομάζονται στάδια της μεθόδου.

Για την κατασκευή της μεθόδου Euler προσεγγίσαμε το ολοκλήρωμα με το εμβαδό του ορθογωνίου με ύψος

$$f(x_n, y(x_n)).$$

Θα μπορούσαμε να κατασκευάσουμε μια άλλη μέθοδο χρησιμοποιώντας το εμβαδό του ορθογωνίου με ύψος

$$f\left(x_n + \frac{h}{2}, y\left(x_n + \frac{h}{2}\right)\right).$$

και να προσεγγίσουμε το άγνωστο

$$y\left(x_n + \frac{h}{2}\right) \approx y(x_n) + \frac{h}{2} y'(x_n) = y(x_n) + \frac{h}{2} f(x_n, y_n)$$

η μέθοδος αυτή γράφεται

$$y_{n+1} = y_n + hf\left(x_n + \frac{h}{2}, y_n + \frac{h}{2} f(x_n, y_n)\right)$$

Αυτή είναι γνωστή και σαν βελτιωμένη μέθοδος Euler ή σαν μέθοδος του Runge (1895) δεύτερης τάξης.

$$\begin{aligned}k_1 &= f(x_n, y_n) \\k_2 &= f\left(x_n + \frac{h}{2}, y_n + \frac{h}{2} k_1\right) \\y_{n+1} &= y_n + h k_2\end{aligned}$$

Οι μέθοδοι Heun και Runge που κατασκευάσαμε έχουν την γενική μορφή

$$\begin{aligned}k_1 &= f(x_n, y_n) \\k_2 &= f(x_n + ch, y_n + ha k_1) \\y_{n+1} &= y_n + h(b_1 k_1 + b_2 k_2)\end{aligned}$$

Όπου για την μέθοδο Heun οι συντελεστές είναι

$$c = 1, \quad a = 1, \quad b_1 = b_2 = \frac{1}{2}$$

και για την μέθοδο Runge

$$c = \frac{1}{2}, \quad a = \frac{1}{2}, \quad b_1 = 0, \quad b_2 = 1.$$

2.2 Γενική μορφή των μεθόδων Runge-Kutta

Ο Kutta το 1901 έγραψε τις μεθόδους στη μορφή που σήμερα ονομάζουμε Runge-Kutta.

$$\begin{aligned}k_1 &= f(x_n, y_n) \\k_2 &= f(x_n + c_2 h, y_n + h a_{21} k_1) \\k_3 &= f(x_n + c_3 h, y_n + h(a_{31} k_1 + a_{32} k_2)) \\&\vdots \\k_s &= f(x_n + c_s h, y_n + h(a_{s1} k_1 + a_{s2} k_2 + \cdots + a_{s,s-1} k_{s-1})) \\y_{n+1} &= y_n + h(b_1 k_1 + b_2 k_2 + \cdots + b_s k_s)\end{aligned} \tag{2.2}$$

Αυτή είναι μια μέθοδος s σταδίων. Ο Kutta έκανε την υπόθεση

$$\begin{aligned}c_2 &= a_{21} \\c_3 &= a_{31} + a_{32} \\&\vdots \\c_s &= a_{s1} + a_{s2} + \cdots + a_{s,s-1}\end{aligned}$$

Η θεωρητική τεκμηρίωση των μεθόδων Runge-Kutta έγινε από τον Νεοζηλανδό μαθηματικό John Butcher με μια σειρά άρθρων που ξεκίνησε να δημοσιεύει από το 1963. Ο Butcher χρησιμοποίησε έναν πίνακα για την αναπαράσταση των μεθόδων γνωστό ως Butcher array ή Butcher

tableau

0					
c_2	a_{21}				
c_3	a_{31}	a_{32}			
$\vdots$	$\vdots$	$\vdots$	$\ddots$		
c_s	a_{s1}	a_{s2}	$\cdots$	$a_{s,s-1}$	
	b_1	b_2	$\cdots$	b_{s-1}	b_s

Οι μέθοδοι που κατασκευάσαμε στην προηγούμενη παράγραφο ανήκουν στην οικογένεια των μεθόδων που πρότεινε ο Kutta και τα Butcher arrays τους είναι: Μέθοδος Euler

$$\begin{array}{c|c} 0 & 0 \\ \hline & 1 \end{array}$$

Μέθοδοι 2 σταδίων Heun και Runge

$$\begin{array}{c|cc} 0 & & \\ \hline 1 & 1 & \\ \hline & 1/2 & 1/2 \end{array} \quad \begin{array}{c|cc} 0 & & \\ \hline 1/2 & 1/2 & \\ \hline & 0 & 1 \end{array}$$

Μια μέθοδος Runge-Kutta s σταδίων ορίζεται ως εξής

$$y_{n+1} = y_n + h \sum_{i=1}^s b_i f(c_i h, Y_i), \quad (2.3)$$

$$k_i = y_n + h \sum_{j=1}^s a_{ij} f(x + c_j h, Y_j), \quad i = 1, \dots, s.$$

Τα διανύσματα k_i καλούνται τα εσωτερικά στάδια της μεθόδου.

$$\begin{aligned} k_1 &= f(x_n + c_1 h, y_n + h(a_{11} k_1 + a_{12} k_2 + \cdots + a_{1s} k_s)) \\ k_2 &= f(x_n + c_2 h, y_n + h(a_{21} k_1 + a_{22} k_2 + \cdots + a_{2s} k_s)) \\ k_3 &= f(x_n + c_3 h, y_n + h(a_{31} k_1 + a_{32} k_2 + \cdots + a_{3s} k_s)) \\ &\vdots \\ k_s &= f(x_n + c_s h, y_n + h(a_{s1} k_1 + a_{s2} k_2 + \cdots + a_{ss} k_s)) \end{aligned}$$

Το αντίστοιχο Butcher array της μεθόδου είναι

c_1	a_{11}	a_{12}	$\cdots$	a_{1s}
c_2	a_{21}	a_{22}	$\cdots$	a_{2s}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
c_s	a_{s1}	a_{s2}	$\cdots$	a_{ss}
	b_1	b_2	$\cdots$	b_s

Μπορούμε να γράψουμε το Butcher array με χρήση πινάκων

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

όπου A πίνακας πραγματικών αριθμών διάστασης $s \times s$, b και c είναι διανύσματα διάστασης s .

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1s} \\ a_{21} & a_{22} & \cdots & a_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ a_{s1} & a_{s2} & \cdots & a_{ss} \end{pmatrix}$$

$$c = (c_1, c_2, \dots, c_s)^T$$

$$b = (b_1, b_2, \dots, b_s)^T$$

Πολλές φορές είναι χρήσιμο να θεωρούμε τον διαγώνιο πίνακα C με διαγώνια στοιχεία τα στοιχεία του διανύσματος c

$$C = \begin{pmatrix} c_1 & & \cdots & \\ & c_2 & \cdots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \cdots & c_s \end{pmatrix}$$

Τα b_i ονομάζονται βάρη (weights) της μεθόδου ενώ τα c_i κόμβοι (abscissae) της μεθόδου. Η μέθοδος που διατύπωσε ο Kutta (2.2) είναι μια άμεση μέθοδος Runge-Kutta (explicit Runge-Kutta method) ή ERK. Ο πίνακας A είναι αυστηρά κάτω τριγωνικός ($a_{ij} = 0$ για $i \leq j$). Το πρώτο στάδιο εμπλέκει μόνο τον υπολογισμό της συνάρτησης f στο προηγούμενο βήμα (x_n, y_n)

$$k_1 = f(x_n, y_n)$$

Το δεύτερο στάδιο k_2 εμπλέκει το πρώτο στάδιο k_1 που έχει ήδη υπολογισθεί

$$k_2 = f(x_n + c_2 h, y_n + h a_{21} k_1)$$

Το τρίτο στάδιο εμπλέκει τα k_1 και k_2 που είναι τώρα γνωστά

$$k_3 = f(x_n + c_3 h, y_n + h(a_{31} k_1 + a_{32} k_2))$$

Με τον τρόπο αυτόν όταν έχουν υπολογισθεί όλα τα στάδια $k_1, k_2, \dots, k_s$ υπολογίζεται η εκτίμηση της λύσης y_{n+1} στο επόμενο βήμα (x_{n+1})

$$y_{n+1} = y_n + h(b_1 k_1 + b_2 k_2 + \cdots + b_s k_s)$$

Βλέπουμε ότι σε κάθε βήμα χρειαζόμαστε s υπολογισμούς της συνάρτησης f .

Αν ο πίνακας A είναι κάτω τριγωνικός ($a_{ij} = 0$ για $i < j$), δηλαδή επιτρέπονται μη μηδενικά στοιχεία και στη κύρια διαγώνιο, η μέθοδος ονομάζεται διαγώνια πεπλεγμένη Runge-Kutta μέθοδος (diagonally implicit Runge-Kutta method ή DIRK). Τα στάδια γράφονται

$$\begin{aligned} k_1 &= y_n + h a_{11} f(x + c_1 h, k_1), \\ k_2 &= y_n + h a_{21} f(x + c_1 h, k_1) + h a_{22} f(x + c_2 h, k_2), \\ &\vdots \\ k_s &= y_n + h \sum_{j=1}^s a_{sj} f(x + c_j h, k_j). \end{aligned}$$

Για τον υπολογισμό του πρώτου σταδίου έχουμε

$$k_1 = y_n + ha_{11}f(x + c_1h, k_1)$$

Από αυτή την μη γραμμική εξίσωση θα βρεθεί το k_1 . Στην περίπτωση που η f είναι πραγματική συνάρτηση (παίρνει τιμές στους πραγματικούς αριθμούς) αυτή είναι μια μη γραμμική εξίσωση και μπορεί να λυθεί με μια επαναληπτική μέθοδο, συνήθως τη μέθοδο Newton. Για τον υπολογισμό του δεύτερου σταδίου έχουμε

$$k_2 = y_n + ha_{21}f(x + c_1h, k_1) + ha_{22}f(x + c_2h, k_2),$$

από όπου θα βρεθεί το k_2 . Τέλος από την εξίσωση

$$k_s = y_n + h \sum_{j=1}^s a_{sj}f(x + c_jh, k_j)$$

θα βρεθεί το k_s .

Δηλαδή συνολικά λύνουμε s μη γραμμικές εξισώσεις. Αν η f είναι διανυσματική συνάρτηση τότε σε κάθε στάδιο απαιτείται η επίλυση ενός συστήματος μη γραμμικών αλγεβρικών εξισώσεων. Αν η f έχει D συνιστώσες θα λύσουμε s συστήματα αλγεβρικών εξισώσεων διάστασης $D \times D$ το κάθε ένα.

Τα πράγματα είναι ακόμα πιο περίπλοκα όταν ο πίνακας A δεν είναι τριγωνικός (έχει μη μηδενικά στοιχεία και πάνω από την κύρια διαγώνιο). Βλέπουμε πως ο υπολογισμός του k_1 εμπλέκει όλα τα k_i τα οποία δεν μας είναι γνωστά. Τότε πρέπει να αντιμετωπίσουμε τον υπολογισμό των σταδίων $k_1, k_2, \dots, k_s$ σαν ένα σύστημα μη γραμμικών εξισώσεων

$$\begin{aligned} k_1 &= f(x_n + c_1h, y_n + h(a_{11}k_1 + a_{12}k_2 + \dots + a_{1s}k_s)) \\ k_2 &= f(x_n + c_2h, y_n + h(a_{21}k_1 + a_{22}k_2 + \dots + a_{2s}k_s)) \\ k_3 &= f(x_n + c_3h, y_n + h(a_{31}k_1 + a_{32}k_2 + \dots + a_{3s}k_s)) \\ &\vdots \\ k_s &= f(x_n + c_sh, y_n + h(a_{s1}k_1 + a_{s2}k_2 + \dots + a_{ss}k_s)) \end{aligned}$$

Τότε η μέθοδος ονομάζεται πεπλεγμένη (implicit Runge Kutta method ή IRK). Στην περίπτωση που η f είναι πραγματική συνάρτηση λύνουμε ένα σύστημα από s μη γραμμικές αλγεβρικές εξισώσεις. Αν η f είναι διανυσματική συνάρτηση (με D συνιστώσες) σε κάθε βήμα χρειάζεται να λύσουμε ένα σύστημα μη γραμμικών αλγεβρικών εξισώσεων διάστασης $sD \times sD$.

2.3 Κατασκευή μεθόδων

Η κατασκευή των μεθόδων RK στηρίζεται στην αντιστοίχιση του αναπτύγματος Taylor της ακριβούς λύσης $y(x_n + h)$ και του αναπτύγματος Taylor της αριθμητικής λύσης y_{n+1} . Οι Hairer, Norset και Wanner στο κλασικό βιβλίο τους [?] δίνουν τον ακόλουθο ορισμό της τάξης μιας αριθμητικής μεθόδου (Definition 1.2.)

Ορισμός 2.3.1 Μία μέθοδος Runge-Kutta (2.3) έχει αλγεβρική τάξη p αν για ικανοποιητικά ομαλά προβλήματα (sufficiently smooth problems) (2.1)

$$\|y(x_n + h) - y_n\| \leq Kh^{p+1}$$

Δηλαδή, αν τα αναπτύγματα Taylor της ακριβούς λύσης $y(x_n + h)$ και της αριθμητικής λύσης y_{n+1} συμπίπτουν για όλους τους όρους συμπεριλαμβανομένου και του όρου h^p .

Αυτή η σύμπτωση οδηγεί σε συνθήκες που εμπλέκουν τους συντελεστές a_{ij} , c_i , b_i της μεθόδου που ονομάζονται συνθήκες τάξης (order conditions).

Ο προσδιορισμός των συνθηκών τάξης είναι μια τετριμμένη διαδικασία χρησιμοποιώντας κανόνες διαφορικού λογισμού. Η διαδικασία αυτή γίνεται περίπλοκη όσο αυξάνεται η τάξη. Παρακάτω θα δείξουμε πως προκύπτουν οι συνθήκες τάξης για άμεσες μεθόδους έως και τρίτης τάξης.

2.3.1 Άμεσες μέθοδοι δεύτερης τάξης

Αναπτύσσουμε κατά Taylor την αναλυτική λύση:

$$\begin{aligned} y(x_{n+1}) &= y(x_n + h) = y(x_n) + hy'(x_n) + \frac{h^2}{2}y''(x_n) + \\ &\quad + \frac{h^3}{3!}y^{(3)}(x_n) + O(h^4) \end{aligned} \quad (2.4)$$

από τη διαφορική εξίσωση έχουμε ότι

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

Για την δεύτερη παράγωγο έχουμε

$$y''(x) = \frac{d}{dx}f(x, y) = \frac{\partial f}{\partial x}f(x, y) + \frac{\partial f}{\partial y}f(x, y)\frac{dy}{dx} = f_x + f_y f$$

Για την τρίτη παράγωγο έχουμε

$$\begin{aligned} y^{(3)}(x) &= \frac{d}{dx} \left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} f \right) \\ &= \frac{d}{dx} \frac{\partial f}{\partial x} + \frac{d}{dx} \left(\frac{\partial f}{\partial y} f \right) = \\ &= \frac{\partial}{\partial x} \frac{\partial f}{\partial x} + \frac{\partial}{\partial y} \frac{\partial f}{\partial x} \frac{dy}{dx} + \left(\frac{d}{dx} \frac{\partial f}{\partial y} \right) f + \frac{\partial f}{\partial y} \frac{d}{dx} f = \\ &= f_{xx} + f_{xy} f + \left(\frac{\partial}{\partial x} f_y + \frac{\partial}{\partial y} f_y \frac{dy}{dx} \right) f + f_y (f_x + f_y f) = \end{aligned}$$

ή

$$\begin{aligned} y^{(3)}(x) &= f_{xx} + f_{xy} f + (f_{yx} + f_{yy} f) f + f_y (f_x + f_y f) = \\ &= f_{xx} + f_{xy} f + f_{yx} f + f_{yy} f^2 + f_y f_x + f_y^2 f = \\ &= f_{xx} + 2f_{xy} f + f_{yy} f^2 + f_y f_x + f_y^2 f \end{aligned}$$

Αντικαθιστούμε στην (2.4) και έχουμε

$$\begin{aligned} y(x_n + h) &= y(x_n) + hf + \frac{h^2}{2} (f_x + f_y f) + \\ &+ \frac{h^3}{6} (f_{xx} + 2f_{xy}f + f_{yy}f^2 + f_x f_y + f_y^2 f) + O(h^4) \end{aligned} \quad (2.5)$$

Θεωρούμε μια μέθοδο Runge-Kutta με δύο στάδια

$$\begin{aligned} k_1 &= f(x_n, y_n) \\ k_2 &= f(x_n + c_2 h, y_n + h a_{21} k_1) \\ y_{n+1} &= y_n + h(b_1 k_1 + b_2 k_2) \end{aligned} \quad (2.6)$$

Θα χρησιμοποιήσουμε το ανάπτυγμα Taylor για συναρτήσεις δύο μεταβλητών για το k_2

$$\begin{aligned} k_2 &= f(x_n + c_2 h, y_n + h a_{21} k_1) = \\ &= f + (h c_2 f_x + h a_{21} k_1 f_y) + \\ &\quad \left(\frac{1}{2} (c_2 h)^2 f_{xx} + (c_2 h)(h a_{21} k_1) f_{xy} + \frac{1}{2} (h a_{21} k_1)^2 f_{yy} \right) + O(h^3) \end{aligned}$$

αντικαθιστούμε στην μέθοδο (2.6) και παίρνουμε

$$\begin{aligned} y_{n+1} &= y_n + h(b_1 k_1 + b_2 k_2) \\ &= y_n + h(b_1 + b_2)f + h^2 b_2 (h c_2 f_x + h a_{21} k_1 f_y) + \\ &\quad \frac{1}{2} h^3 b_2 (c_2^2 f_{xx} + 2c_2 a_{21} f f_{xy} + a_{21}^2 f^2 f_{yy}) + O(h^4) \end{aligned} \quad (2.7)$$

Εξισώνοντας τις παραπάνω (2.5) και (2.7) εκφράσεις για την αναλυτική λύση $y(x_n + h)$ και την αριθμητική λύση y_{n+1} έχουμε

- για τους συντελεστές του h

$$b_1 + b_2 = 1 \quad (2.8)$$

- για τους συντελεστές του h^2

$$b_2 c_2 = \frac{1}{2}, \quad b_2 a_{21} = \frac{1}{2} \quad (2.9)$$

τους συντελεστές του h^3 δεν μπορούμε να τους εξισώσουμε αφού στην αριθμητική y_{n+1} εμπλέκονται οι όροι

$$f_{xx}, f_{xy}f, f_{yy}f^2$$

στην αναλυτική λύση $y(x_n + h)$ εμπλέκονται οι όροι

$$f_{xx}, f_{xy}f, f_{yy}f^2, f_x f_y, f_y^2 f.$$

Παρατηρούμε ότι οι μέθοδοι Heun και Runge είναι δεύτερης τάξης γιατί ικανοποιούν τις εξισώσεις (2.8) και (2.9), οι οποίες καλούνται και εξισώσεις τάξης.

Για την κατασκευή λοιπόν μιας μεθόδου δεύτερης τάξης με δύο στάδια ($s=2$) απαιτείται ο προσδιορισμός των συντελεστών από την επίλυση του συστήματος:

$$\begin{aligned} b_1 + b_2 &= 1 \\ b_2 c_2 &= \frac{1}{2} \\ c_2 &= a_{21} \end{aligned}$$

Παρατηρούμε ότι έχουμε να λύσουμε ένα σύστημα 3 εξισώσεων με 4 αγνώστους. Συνεπώς πάντα θα έχουμε έναν ελεύθερο άγνωστο, ως προς τον οποίο μπορούμε να λύσουμε όλους τους υπόλοιπους συντελεστές. Προφανώς ένα τέτοιο σύστημα θα έχει άπειρες λύσεις.

Λύνοντας το σύστημα ως προς τον συντελεστή c_2 (για $c_2 \neq 0$) προκύπτουν οι ακόλουθες λύσεις:

$$\begin{aligned} b_1 &= 1 - \frac{1}{2c_2} \\ b_2 &= \frac{1}{2c_2} \\ a_{21} &= c_2 \end{aligned}$$

Το αντίστοιχο ταμπλό Butcher τώρα γράφεται:

$$\begin{array}{c|c} 0 & \\ c_2 & c_2 \end{array} \quad \begin{array}{cc} & \\ \hline & 1 - \frac{1}{2c_2} \quad \frac{1}{2c_2} \end{array}$$

Οι πλέον γνωστές μέθοδοι είναι οι ακόλουθες: Για $c_2 = 1$ (γνωστή ως κανόνας του τραπεζίου ή βελτιωμένη μέθοδος Euler (RK2a)), για $c_2 = 1/2$ (γνωστή ως κανόνας του ενδιάμεσου σημείου (RK2b)) και για $c_2 = 2/3$ (RK2c).

$$\begin{array}{c|c} 0 & \\ 1 & 1 \end{array}, \quad \begin{array}{c|c} 0 & \\ \frac{1}{2} & \frac{1}{2} \end{array}, \quad \begin{array}{c|c} 0 & \\ \frac{2}{3} & \frac{2}{3} \end{array}$$

$$\begin{array}{cc} \frac{1}{2} & \frac{1}{2} \\ \hline \end{array}, \quad \begin{array}{cc} 0 & 1 \\ \hline \end{array}, \quad \begin{array}{cc} \frac{1}{4} & \frac{3}{4} \\ \hline \end{array}$$

Θα μπορούσαμε βέβαια να κατασκευάσουμε δεύτερης τάξης μεθόδους με 3 στάδια ($s=3$). Τότε θα έπρεπε να λύσουμε το ακόλουθο σύστημα εξισώσεων:

$$\begin{aligned} b_1 + b_2 + b_3 &= 1 \\ b_2 c_2 + b_3 c_3 &= \frac{1}{2} \\ c_2 &= a_{21} \\ c_3 &= a_{31} + a_{32} \end{aligned}$$

Λύνοντας το παραπάνω σύστημα ως προς c_2, c_3, b_3, a_{32} έχουμε:

$$\begin{aligned} b_1 &= -\frac{1-2c_2+2b_3c_2-2b_3c_3}{2c_2} \\ b_2 &= -\frac{2b_3c_3-1}{2c_2} \\ a_{21} &= c_2 \\ a_{31} &= c_3 - a_{32} \end{aligned}$$

με $c_2 \neq 0$. Μια δημοφιλής μέθοδος που ικανοποιεί τις παραπάνω συνθήκες είναι η:

$$\begin{array}{c|ccc} 0 & & & \\ \hline \frac{1}{3} & \frac{1}{3} & & \\ \frac{2}{3} & -\frac{1}{3} & 1 & \\ \hline & 0 & \frac{1}{2} & \frac{1}{2} \end{array}$$

Παρατηρούμε ότι οι μέθοδοι Heun και Runge είναι δεύτερης τάξης γιατί ικανοποιούν τις εξισώσεις (2.8) και (2.9), οι οποίες καλούνται και εξισώσεις τάξης.

Για να δημιουργήσουμε μεθόδους μεγαλύτερης τάξης αν θέλουμε να είναι άμεσες θα πρέπει να εισάγουμε περισσότερα ενδιάμεσα σημεία εκτίμησης της $f(x, y)$ (στάδια της μεθόδου). Μπορούμε όμως όπως θα δούμε παρακάτω να κατασκευάσουμε πεπλεγμένες μεθόδους με δύο στάδια και αλγεβρική τάξη έως και τέταρτη.

Ένα κριτήριο επιλογής των ελεύθερων παραμέτρων μιας μεθόδου Runge-Kutta, με δεδομένη τάξη ακρίβειας p , είναι η ελαχιστοποίηση της νόρμας του σφάλματος της επόμενης τάξης, η οποία εκφράζεται από την σχέση:

$$A^{(p+1)} = \|\tau^{(p+1)}\|_2 = \sqrt{\sum \left(\tau_j^{(p+1)}\right)^2}$$

όπου με $\tau^{(p+1)} = 0$ συμβολίζουμε τις εξισώσεις τάξης $p+1$.

Αν εξισώσουμε τους αντίστοιχους όρους τρίτης τάξης θα προέκυπταν οι συνθήκες:

$$\begin{aligned} \frac{1}{2}b_2c_2^2 &= \frac{1}{6} \\ b_2c_2a_{21} &= \frac{2}{6} \\ \frac{1}{2}b_2a_{21}^2 &= \frac{1}{6} \end{aligned}$$

Έχουμε όμως και τους συντελεστές των όρων $f_x f_y$ και $f_y^2 f$ οι οποίοι είναι ίσοι με $1/6$. Προκειμέ-

νου λοιπόν να ελαχιστοποιήσουμε την νόρμα των σφαλμάτων τρίτης τάξης ορίζουμε τους όρους:

$$\begin{aligned}\tau_1 &= \frac{1}{2}b_2c_2^2 - \frac{1}{6} \\ \tau_2 &= b_2c_2a_{21} - \frac{1}{3} \\ \tau_3 &= \frac{1}{2}b_2a_{21}^2 - \frac{1}{6} \\ \tau_4 &= -\frac{1}{6} \\ \tau_5 &= -\frac{1}{6}\end{aligned}$$

Συνεπώς ορίζουμε την συνάρτηση:

$$f(c_2) = \sqrt{\tau_1^2 + \tau_2^2 + \tau_3^2 + \tau_4^2 + \tau_5^2}$$

ή αν αντικαταστήσουμε τα b_1, b_2, a_{21} ως συνάρτηση του c_2 προκύπτει η συνάρτηση:

$$A(c_2) = \frac{1}{6} \sqrt{\frac{27c_2^2}{2} - 18c_2 + 8}$$

Ελαχιστοποιώντας την παραπάνω συνάρτηση βρίσκουμε ότι η μέθοδος δεύτερης αλγεβρικής τάξης, δύο σταδίων με το μικρότερο σφάλμα επόμενης τάξης προκύπτει για $c_2 = \frac{2}{3}$ (RK2c).

Για να δούμε την χρήση των μεθόδων Runge Kutta θα λύσουμε το πρόβλημα 1 με την μέθοδο RK2a που είδαμε παραπάνω.

Θέλουμε λοιπόν να λύσουμε την εξίσωση

$$y' = xy + 2x, \quad x \in [0, 1], \quad y(0) = 1.$$

του οποίου ακριβής λύση είναι η $y(x) = -2 + 3e^{\frac{x^2}{2}}$.

Η μέθοδος RK2a είναι η ακόλουθη:

$$\begin{aligned}k_1 &= f(x_n, y_n) \\ k_2 &= f(x_n + h, y_n + hk_1) \\ y_{n+1} &= y_n + h \left(\frac{k_1}{2} + \frac{k_2}{2} \right)\end{aligned}$$

Αν επιλέξουμε διαμέριση $N = 10$ τότε το μέγεθος του βήματος είναι $h = 0.1$ και το x παίρνει τις τιμές $0, 0.1, 0.2, 0.3, \dots, 0.9, 1$

Για $n = 0$ έχουμε:

$$\begin{aligned}k_1 &= f(x_0, y_0) = f(0, 1) = 0 \\ k_2 &= f(x_0 + h, y_0 + hk_1) = f(0 + 0.1, 1 + 0.1 \times 0) \\ &= f(0.1, 1) = 0.1 \times 1 + 2 \times 0.1 = 0.3\end{aligned}$$

$$y_1 = y_0 + h \left(\frac{k_1}{2} + \frac{k_2}{2} \right) = 1 + 0.1 \left(\frac{0}{2} + \frac{0.3}{2} \right) = 1.015$$

Συνεχίζοντας για $n = 1$ έχουμε:

$$k_1 = f(x_1, y_1) = f(0.1, 1.015) = 0.1 \times 1.015 + 2 \times 0.1 = 0.3015$$

$$\begin{aligned} k_2 &= f(x_1 + h, y_1 + hk_1) = f(0.1 + 0.1, 1.015 + 0.1 \times 0.3015) = \\ &= f(0.2, 1.04515) = 0.2 \times 1.04515 + 2 \times 0.2 = 0.60903 \end{aligned}$$

$$y_2 = y_1 + h \left(\frac{k_1}{2} + \frac{k_2}{2} \right) = 1.015 + 0.1 \left(\frac{0.3015}{2} + \frac{0.60903}{2} \right) = 1.0605265$$

για $n = 2$ έχουμε:

$$k_1 = f(x_2, y_2) = f(0.2, 1.0605265) = 0.2 \times 1.0605265 + 2 \times 0.2 = 0.6121053$$

$$\begin{aligned} k_2 &= f(x_2 + h, y_2 + hk_1) = f(0.2 + 0.1, 1.0605265 + 0.1 \times 0.6121053) = \\ &= f(0.3, 1.12173703) = 0.3 \times 1.12173703 + 2 \times 0.3 = 0.936521109 \end{aligned}$$

$$y_3 = y_2 + h \left(\frac{k_1}{2} + \frac{k_2}{2} \right) = 1.0605265 + 0.1 \left(\frac{0.6121053}{2} + \frac{0.936521109}{2} \right) = 1.1379578$$

και συνεχίζοντας με τον ίδιο τρόπο βρίσκουμε όλες τις τιμές του y_i .

Υλοποιούμε τη μέθοδο στο function RK2a

```
function [x,y] = RK2a(fun,y0,x0,xn,N)
```

```
n = length(y0) ;
```

```
% step length
```

```
h = (xn-x0)/N ;
```

```
% creates the vectors to be used
```

```
x = zeros(N+1,1);
```

```
y = zeros(N+1,n);
```

```
% starting point of the integration interval
```

```
x(1) = x0 ;
```

```
% initial condition
```

```
y(1,:) = y0 ;
```

```
for i=1:N
```

```
    x(i+1) = x(i) + h ;
```

```
    Y1 = y(i,:) ;
```

```
    F1 = feval(fun,Y1) ;
```

```
    Y2 = y(i,:) + h*F1 ;
```

```
    F2 = feval(fun,x+h,Y2) ;
```

```
    y(i+1,:) = y(i,:) + h*(F1+F2)/2 ;
```

```
end
```

```
end
```

Στον επόμενο πίνακα εμφανίζονται τα αποτελέσματα της μεθόδου.

x_i	y_i	$y(x_i)$	$ y_i - y(x_i) $
0	1.0000000	1.0000000	0.0000000
0.1000	1.0150000	1.0150376	0.0000376
0.2000	1.0605265	1.0606040	0.0000775
0.3000	1.1379578	1.1380836	0.0001258
0.4000	1.2496691	1.2498612	0.0001921
0.5000	1.3991539	1.3994454	0.0002915
0.6000	1.5912061	1.5916521	0.0004460
0.7000	1.8321760	1.8328639	0.0006879
0.8000	2.1303193	2.1313833	0.0010640
0.9000	2.4962656	2.4979075	0.0016419
1.0000	2.9436440	2.9461638	0.0025198

Το μέγιστο κατά απόλυτη τιμή σφάλμα είναι 2.52×10^{-3} στο τέλος του διαστήματος ολοκλήρωσης.

Αν είχαμε επιλέξει να πάρουμε $N = 100$ σημεία διαμέρισης το μέγιστο κατά απόλυτη τιμή σφάλμα θα ήταν ίσο με 2.11×10^{-5} στο τέλος του διαστήματος ολοκλήρωσης, ενώ αν επιλέγαμε $N = 1000$ το μέγιστο σφάλμα ισούται με 2.07×10^{-7} .

Στον παρακάτω πίνακα παρουσιάζουμε το μέγιστο απόλυτο σφάλμα της αριθμητικής λύσης για το πρόβλημα 1 για τις 3 μεθόδους δεύτερης τάξης που είδαμε παραπάνω με διάστημα ολοκλήρωσης το $[0, 1]$ για διάφορες τιμές του βήματος h .

h	$RK2a$	$RK2b$	$RK2c$
0.2	$1.17 \cdot 10^{-2}$	$2.88 \cdot 10^{-2}$	$2.31 \cdot 10^{-2}$
0.1	$2.52 \cdot 10^{-3}$	$7.71 \cdot 10^{-3}$	$5.98 \cdot 10^{-3}$
0.05	$5.75 \cdot 10^{-4}$	$1.99 \cdot 10^{-3}$	$1.52 \cdot 10^{-3}$
0.02	$8.63 \cdot 10^{-5}$	$3.25 \cdot 10^{-4}$	$2.46 \cdot 10^{-4}$
0.01	$2.11 \cdot 10^{-5}$	$8.19 \cdot 10^{-5}$	$6.16 \cdot 10^{-5}$

2.3.2 Άμεσες μέθοδοι τρίτης τάξης

Θεωρούμε μέθοδο Runge-Kutta με τρία στάδια:

$$\begin{aligned}
 k_1 &= f(x_n, y_n) \\
 k_2 &= f(x_n + c_2 h, y_n + h a_{21} k_1) \\
 k_3 &= f(x_n + c_3 h, y_n + h(a_{31} k_1 + a_{32} k_2)) \\
 y_{n+1} &= y_n + h(b_1 k_1 + b_2 k_2 + b_3 k_3)
 \end{aligned} \tag{2.10}$$

Αναπτύσσουμε κατά Taylor την αναλυτική λύση:

$$\begin{aligned}
 y(x_n + h) &= y(x_n) + hf + \frac{h^2}{2} (f_x + f_y f) + \\
 &\quad + \frac{h^3}{6} (f_{xx} + 2f_{xy}f + f_{yy}f^2 + f_x f_y + f_y^2 f) \\
 &\quad + \frac{h^4}{24} y^{(4)}(x) + O(h^5)
 \end{aligned} \tag{2.11}$$

Για την τέταρτη παράγωγο έχουμε

$$\begin{aligned}
 y^{(4)}(x) &= (y'''(x))' \\
 &= \frac{d}{dx} (f_{xx} + 2f_{xy}f + f_{yy}f^2 + f_x f_y + f_y^2 f) \\
 &= \frac{d}{dx} f_{xx} + 2 \frac{d}{dx} (f_{xy}f) + \frac{d}{dx} (f_{yy}f^2) + \frac{d}{dx} (f_x f_y) + \frac{d}{dx} (f_y^2 f)
 \end{aligned}$$

Οι παραπάνω παράγωγοι υπολογίζονται ως εξής:

$$\frac{d}{dx} f_{xx} = \frac{\partial f}{\partial x} f_{xx} + \frac{\partial f}{\partial y} f_{xx} \frac{dy}{dx} = f_{xxx} + f_{xxy}f$$

$$\begin{aligned}
 \frac{d}{dx} (f_{xy}f) &= \left(\frac{d}{dx} f_{xy} \right) f + f_{xy} \frac{df}{dx} = \\
 &= \left(\frac{\partial}{\partial x} f_{xy} + \frac{\partial}{\partial y} f_{xy} \frac{dy}{dx} \right) f + f_{xy} (f_x + f_y f) = \\
 &= (f_{xyx} + f_{xyy}f) f + f_{xy} f_x + f_{xy} f_y f = \\
 &= f_{xxy}f + f_{xyy}f^2 + f_{xy} f_x + f_{xy} f_y f
 \end{aligned}$$

$$\begin{aligned}
 \frac{d}{dx} (f_{yy}f^2) &= \left(\frac{d}{dx} f_{yy} \right) f^2 + f_{yy} \frac{d}{dx} (f^2) = \\
 &= \left(\frac{\partial}{\partial x} f_{yy} + \frac{\partial}{\partial y} f_{yy} \frac{dy}{dx} \right) f^2 + f_{yy} 2f \frac{df}{dx} = \\
 &= (f_{yyx} + f_{yyy}f) f^2 + f_{yy} 2f (f_x + f_y f) = \\
 &= f_{xyy}f^2 + f_{yyy}f^3 + 2f_{yy} f_x f + 2f_{yy} f_y f^2
 \end{aligned}$$

$$\begin{aligned}
 \frac{d}{dx} (f_x f_y) &= \left(\frac{d}{dx} f_x \right) f_y + f_x \left(\frac{d}{dx} f_y \right) \\
 &= \left(\frac{\partial}{\partial x} f_x + \frac{\partial}{\partial y} f_x \frac{dy}{dx} \right) f_y + f_x \left(\frac{\partial}{\partial x} f_y + \frac{\partial}{\partial y} f_y \frac{dy}{dx} \right) = \\
 &= (f_{xx} + f_{xy}f) f_y + f_x (f_{yx} + f_{yy}f) = \\
 &= f_{xx} f_y + f_{xy} f_y f + f_{xy} f_x + f_{yy} f_x f
 \end{aligned}$$

$$\begin{aligned}
\frac{d}{dx} (f_y^2 f) &= \left(\frac{d}{dx} f_y^2 \right) f + f_y^2 \frac{df}{dx} = \\
&= 2f_y \left(\frac{d}{dx} f_y \right) f + f_y^2 (f_x + f_y f) = \\
&= 2f_y (f_{yx} + f_{yy} f') f + f_y^2 (f_x + f_y f) = \\
&= 2f_{xy} f_y f + 2f_{yy} f_y f^2 + f_y^2 f_x + f_y^3 f
\end{aligned}$$

Άρα τελικά από τον υπολογισμό των παραγώγων η $y^{(4)}(x)$ γίνεται:

$$\begin{aligned}
y^{(4)}(x) &= f_{xxx} + f_{xxy}f \\
&\quad + 2f_{xxy}f + 2f_{xyy}f^2 + 2f_{xy}f_x + 2f_{xy}f_y f \\
&\quad + f_{xyy}f^2 + f_{yyy}f^3 + 2f_{yy}f_x f + 2f_{yy}f_y f^2 \\
&\quad + f_{xx}f_y + f_{xy}f_y f + f_{xy}f_x + f_{yy}f_x f \\
&\quad + 2f_{xy}f_y f + 2f_{yy}f_y f^2 + f_y^2 f_x + f_y^3 f
\end{aligned}$$

ή

$$\begin{aligned}
y^{(4)}(x) &= f_{xxx} + 3f_{xxy}f + 3f_{xyy}f^2 + f_{yyy}f^3 \\
&\quad + 3f_{xy}f_x + 5f_{xy}f_y f + 3f_{yy}f_x f + 4f_{yy}f_y f^2 + f_{xx}f_y + f_y^2 f_x + f_y^3 f
\end{aligned}$$

Αντικαθιστώντας στην (2.11) την $y^{(4)}(x)$ θα πάρουμε:

$$\begin{aligned}
y(x_n + h) &= y(x_n) + hf + \frac{h^2}{2} (f_x + f_y f) + \\
&\quad + \frac{h^3}{6} (f_{xx} + 2f_{xy}f + f_{yy}f^2 + f_x f_y + f_y^2 f) + \\
&\quad + \frac{h^4}{24} (f_{xxx} + f_y^3 f + 3f_x f_{xy} + f_{xx}f_y + f_x f_y^2 + f_y^3 f_{yy} + \\
&\quad + 3f_{xxy}f + 4f_y^2 f_{yy} + 5f_{xy}f_y f + 3f_y^2 f_{xyy} + 3f_x f_{yyy}) + \\
&\quad + O(h^5)
\end{aligned} \tag{2.12}$$

Από το θεώρημα Taylor τα k_2 και k_3 θα γίνουν:

$$\begin{aligned}
k_2 &= f + ((c_2 h)f_x + (h a_{21} k_1)f_y) + \\
&\quad + \frac{1}{2} ((c_2^2 h^2)f_{xx} + (2c_2 a_{21} k_1 h^2)f_{xy} + (h^2 a_{21} k_1^2)f_{yy}) + \\
&\quad + \frac{1}{6} ((h^3 a_{21}^3 k_1^3)f_{yyy} + (3a_{21}^2 c_2 k_1^2)f_{xyy} + (3a_{21} c_2^2 k_1)f_{xxy} + c_2^3 f_{xxx}) + \\
&\quad + O(h^4)
\end{aligned}$$

και

$$\begin{aligned}
k_3 &= f + ((c_3 h)f_x + h(a_{31} k_1 + a_{32} k_2)f_y) + \\
&\quad + \frac{1}{2} ((c_3^2 h^2)f_{xx} + (2c_3(a_{31} k_1 + a_{32} k_2)h^2)f_{xy} + (h^2(a_{31}^2 k_1^2 + a_{32}^2 k_2^2))f_{yy}) + \\
&\quad + O(h^3)
\end{aligned}$$

Στις δύο παραπάνω εξισώσεις αντικαθιστούμε τα k_1, k_2 και εφαρμόζουμε στη μέθοδο RK τριών σταδίων (2.10)

$$\begin{aligned}
 y_{n+1} &= y_n + h(b_1 + b_2 + b_3)f + \\
 &+ h^2((b_2c_2 + b_3c_3)f_x + (a_{21}b_2 + (a_{31} + a_{32})b_3)ff_x) + \\
 &+ \frac{h^3}{2}((b_2c_2^2 + b_3c_3^2)f_{xx} + (a_{21}^2b_2 + (a_{31} + a_{32})^2b_3)f^2f_{yy} \\
 &\quad + 2a_{32}b_3c_2f_xf_y + 2(a_{21}a_{32}b_3)f_y^2 \\
 &\quad + 2(a_{21}b_2c_2 + (a_{31} + a_{32})b_3c_3)f_{xy})f + O(h^4)
 \end{aligned} \tag{2.13}$$

Εξισώνουμε την αναλυτική (2.12) με την αριθμητική λύση (2.13)

- για τους συντελεστές του hf

$$b_1 + b_2 + b_3 = 1 \tag{2.14}$$

- για τους συντελεστές του h^2f_x

$$b_2c_2 + b_3c_3 = \frac{1}{2} \tag{2.15}$$

- για τους συντελεστές του h^2ff_y

$$a_{21}b_2 + (a_{31} + a_{32})b_3 = \frac{1}{2} \tag{2.16}$$

- για τους συντελεστές του h^3f_{xx}

$$\frac{1}{2}((b_2c_2^2 + b_3c_3^2) = \frac{1}{6} \tag{2.17}$$

- για τους συντελεστές του $h^3f_{yy}f^2$

$$\frac{1}{2}(a_{21}^2b_2 + (a_{31} + a_{32})^2b_3) = \frac{1}{6} \tag{2.18}$$

- για τους συντελεστές του $h^3f_xf_y$

$$a_{32}b_3c_2 = \frac{1}{6} \tag{2.19}$$

- για τους συντελεστές του $h^3f_y^2$

$$a_{21}a_{32}b_3 = \frac{1}{6} \tag{2.20}$$

- για τους συντελεστές του $h^3f_{xy}f$

$$a_{21}b_2c_2 + (a_{31} + a_{32})b_3c_3 = \frac{2}{6} \tag{2.21}$$

Από τις συνθήκες γραμμής

$$\begin{aligned}c_2 &= a_{21}, \\c_3 &= a_{31} + a_{32},\end{aligned}$$

η συνθήκη δεύτερης τάξης (2.16) είναι ισοδύναμη με την (2.15). Επίσης στην τρίτη τάξη οι συνθήκες (2.18) και (2.21) είναι ισοδύναμες με την (2.17) και η (2.20) είναι ισοδύναμη με την (2.19).

Έτσι έχουμε μία συνθήκη πρώτης τάξης, μία συνθήκη δεύτερης τάξης και δύο συνθήκες τρίτης τάξης ως εξής:

$$\begin{aligned}b_1 + b_2 + b_3 &= 1 \\b_2c_2 + b_3c_3 &= \frac{1}{2} \\b_2c_2^2 + b_3c_3^2 &= \frac{1}{3} \\a_{32}b_3c_2 &= \frac{1}{6}\end{aligned}\tag{2.22}$$

από τη συνθήκη γραμμής έχουμε $c_2 = a_{21}$ και $c_3 = a_{31} + a_{32}$.

Έχουμε τέσσερις εξισώσεις από τις οποίες θέλουμε να υπολογίσουμε τα c_i τα b_i και ένα από τα a_{31} ή a_{32} δηλαδή έξι συντελεστές. Αν θέσουμε $c_3 = 1$ και $b_1 = b_3$ παίρνουμε την πρώτη μέθοδο και αν θέσουμε $a_{31} = 0$ και $c_2 = 1/2$ παίρνουμε τη δεύτερη μέθοδο

$$\begin{array}{c|ccc} 0 & & & \\ \hline \frac{1}{2} & \frac{1}{2} & & \\ \hline 1 & -1 & 2 & \\ \hline & \frac{1}{6} & \frac{4}{6} & \frac{1}{6} \end{array}, \begin{array}{c|ccc} 0 & & & \\ \hline \frac{1}{2} & \frac{1}{2} & & \\ \hline \frac{3}{4} & 0 & \frac{3}{4} & \\ \hline & \frac{2}{9} & \frac{1}{3} & \frac{4}{9} \end{array}$$

Έτσι λοιπόν για να κατασκευάσουμε μια μέθοδο τρίτης τάξης εξισώνουμε τους συντελεστές και προκύπτει το ακόλουθο σύστημα εξισώσεων:

$$\begin{aligned}b_1 + b_2 + b_3 &= 1 \\b_2c_2 + b_3c_3 &= \frac{1}{2} \\b_2c_2^2 + b_3c_3^2 &= \frac{1}{3} \\b_3a_{32}c_2 &= \frac{1}{6} \\c_2 &= a_{21} \\c_3 &= a_{31} + a_{32}\end{aligned}$$

Όπως μπορούμε να παρατηρήσουμε έχουμε 6 εξισώσεις με 8 αγνώστους, δηλαδή έχουμε δύο βαθμούς ελευθερίας οπότε το σύστημα έχει άπειρες λύσεις. Μπορούμε να λύσουμε τις παραπάνω συνθήκες ως προς δύο συντελεστές και με βάση τις λύσεις να κατασκευάσουμε μεθόδους Runge-Kutta τρίτης αλγεβρικής τάξης. Όπως προκύπτει από τις σχέσεις που προκύπτουν από τα αναπτύγματα Taylor, δεν μπορούμε να κατασκευάσουμε μεθόδους Runge-Kutta τέταρτης αλγεβρικής τάξης με τρία στάδια (Δεν μπορούμε να αντιστοιχίσουμε όλους τους όρους). Για να μπορέσουμε να το καταφέρουμε χρειαζόμαστε τουλάχιστον τέσσερα στάδια για να υπάρχει η δυνατότητα να εξισώσουμε τους συντελεστές του h^4 .

Λύνοντας το παραπάνω σύστημα ως προς τους συντελεστές c_2, c_3 και θεωρώντας ότι: $c_2 c_3 \neq 0$, $c_2 \neq c_3$ και $c_2 \neq 2/3$ βρίσκουμε:

$$\begin{aligned} b_1 &= \frac{2 - 3c_2 - 3c_3 + 6c_2c_3}{6c_2c_3} \\ b_2 &= \frac{3c_3 - 2}{6c_2(c_3 - c_2)} \\ b_3 &= \frac{3c_2 - 2}{6(c_2 - c_3)c_3} \\ a_{21} &= c_2 \\ a_{31} &= \frac{c_3(3c_2^2 + c_3 - 3c_2)}{c_2(3c_2 - 2)} \\ a_{32} &= \frac{c_3(c_2 - c_3)}{c_2(3c_2 - 2)} \end{aligned}$$

Γνωστές μέθοδοι προκύπτουν για $c_2 = 1/2$ και $c_3 = 1$ (RK3b), για $c_2 = 1/3$ και $c_3 = 2/3$ (RK3c). Επίσης γνωστή μέθοδος προκύπτει αν θέσουμε $c_2 = c_3 = 2/3$ (RK3a) και στην συνέχεια λύσουμε το σύστημα των συνθηκών.

$$\begin{array}{c|ccc} 0 & & & \\ 2 & 2 & & \\ 3 & 1 & 1 & \\ 3 & 1 & 3 & \\ \hline & 1 & 0 & 3 \\ & 4 & & 4 \end{array}$$

$$\begin{array}{c|ccc} 0 & & & \\ 1 & 1 & & \\ 2 & 2 & & \\ 1 & -1 & 2 & \\ \hline & 1 & 2 & 1 \\ & 6 & 3 & 6 \end{array}$$

$$\begin{array}{c|ccc} 0 & & & \\ 1 & 1 & & \\ 2 & 3 & 2 & \\ 3 & 0 & 3 & \\ \hline & 1 & 0 & 3 \\ & 4 & & 4 \end{array}$$

Για να δούμε την αποτελεσματικότητα των μεθόδων Runge Kutta τρίτης αλγεβρικής τάξης θα λύσουμε το πρόβλημα 2 με την μέθοδο RK2b που είδαμε παραπάνω.

Θέλουμε λοιπόν να λύσουμε την εξίσωση

$$y' + y = x \sin(x), \quad x \in [0, 5], \quad y(0) = 1.$$

του οποίου ακριβής λύση είναι η $y(x) = \frac{1}{2}(e^{-x} + \cos(x) - x \cos(x) + x \sin(x))$.

Η μέθοδος RK3b είναι η ακόλουθη:

$$k_1 = f(x_n, y_n)$$

$$k_2 = f\left(x_n + \frac{1}{2}h, y_n + h\frac{1}{2}k_1\right)$$

$$k_3 = f\left(x_n + h, y_n + h(-k_1 + 2k_2)\right)$$

$$y_{n+1} = y_n + h\left(\frac{1}{6}k_1 + \frac{4}{6}k_2 + \frac{1}{6}k_3\right)$$

με $f(x, y) = x \sin(x) - y$. Αν επιλέξουμε διαμέριση $N = 50$ τότε το μέγεθος του βήματος είναι $h = 0.1$ και το x παίρνει τις τιμές $0, 0.1, 0.2, 0.3, \dots, 4.9, 5$

Για $n = 0$ έχουμε:

$$k_1 = f(x_0, y_0) = f(0, 1) = -1$$

$$k_2 = f\left(x_0 + \frac{1}{2}h, y_0 + h\frac{1}{2}k_1\right) = f\left(0 + \frac{1}{2}0.1, 1 + 0.1\frac{1}{2} \times (-1)\right) = f(0.05, 0.95) = -0.947501$$

$$\begin{aligned} k_3 &= f\left(x_0 + h, y_0 + h(-k_1 + 2k_2)\right) = f(0 + 0.1, 1 + 0.1(1 - 2 \times 0.947501)) \\ &= f(0.1, 0.9104998) = -0.900516 \end{aligned}$$

$$y_1 = y_0 + h\left(\frac{1}{6}k_1 + \frac{4}{6}k_2 + \frac{1}{6}k_3\right) = 1 + 0.1\left(-\frac{1}{6} - \frac{4}{6}0.947501 - \frac{1}{6}0.900516\right) = 0.905158$$

Για $n = 1$ έχουμε:

$$k_1 = f(x_1, y_1) = f(0.1, 0.905158) = -0.895175$$

$$\begin{aligned} k_2 &= f\left(x_1 + \frac{1}{2}h, y_1 + h\frac{1}{2}k_1\right) = f\left(0.1 + \frac{1}{2}0.1, 0.905158 + 0.1\frac{1}{2} \times (-0.895175)\right) \\ &= f(0.15, 0.86039925) = -0.837984 \end{aligned}$$

$$\begin{aligned} k_3 &= f\left(x_1 + h, y_1 + h(-k_1 + 2k_2)\right) = f(0.1 + 0.1, 0.905158 + 0.1(0.895175 - 2 \times 0.837984)) \\ &= f(0.2, 0.8270787) = -0.787345 \end{aligned}$$

$$\begin{aligned} y_2 &= y_1 + h\left(\frac{1}{6}k_1 + \frac{4}{6}k_2 + \frac{1}{6}k_3\right) = \\ &= 0.905158 + 0.1\left(-\frac{1}{6}0.895175 - \frac{4}{6}0.837984 - \frac{1}{6}0.787345\right) = 0.82125 \end{aligned}$$

Για $n = 2$ έχουμε:

$$k_1 = f(x_2, y_2) = f(0.2, 0.82125) = -0.781516$$

$$\begin{aligned} k_2 &= f\left(x_2 + \frac{1}{2}h, y_2 + h\frac{1}{2}k_1\right) = f\left(0.2 + \frac{1}{2}0.1, 0.82125 + 0.1\frac{1}{2} \times (-0.781516)\right) \\ &= f(0.25, 0.782174) = -0.720323 \end{aligned}$$

$$\begin{aligned} k_3 &= f\left(x_2 + h, y_2 + h(-k_1 + 2k_2)\right) = f(0.2 + 0.1, 0.82125 + 0.1(0.781516 - 2 \times 0.720323)) \\ &= f(0.3, 0.755337) = -0.666681 \end{aligned}$$

$$\begin{aligned} y_3 &= y_2 + h\left(\frac{1}{6}k_1 + \frac{4}{6}k_2 + \frac{1}{6}k_3\right) = \\ &= 0.82125 + 0.1\left(-\frac{1}{6}0.781516 - \frac{4}{6}0.720323 - \frac{1}{6}0.666681\right) = 0.749092 \end{aligned}$$

Υλοποιούμε τη μέθοδο στη function RK3b

```
function [x,y] = RK3b(fun,y0,x0,xn,N)

n = length(y0) ;

% step length
h = (xn-x0)/N ;

% creates the vectors to be used
x = zeros(N+1,1);
y = zeros(N+1,n);
% starting point of the integration interval
x(1) = x0 ;

% initial condition
y(1,:) = y0 ;

for i=1:N
    x(i+1) = x(i) + h ;
    Y1 = y(i,:) ;
    F1 = feval(fun,Y1) ;
    Y2 = y(i,:) + h*F1/2 ;
    F2 = feval(fun,x+h/2,Y2) ;
    Y3 = y(i,:) + h*(-F1+2*F2) ;
    F3 = feval(fun,x+h,Y3) ;
    y(i+1,:) = y(i,:) + h*(F1+4F2+F3)/6 ;
end
end
```

Στον επόμενο πίνακα εμφανίζονται ενδεικτικά τιμές στο διάστημα ολοκλήρωσης

x_i	y_i
0	1.0000000
0.1	0.9051580
0.2	0.8212504
0.3	0.7490922
0.4	0.6893453
0.5	0.6424968
0.6	0.6088415
0.7	0.5884677
0.8	0.5812472
0.9	0.5868296
1.0	0.6046404
$\vdots$	$\vdots$
2.0	1.1850170
$\vdots$	$\vdots$
3.0	1.2266003
$\vdots$	$\vdots$
4.0	-0.5239232
$\vdots$	$\vdots$
5.0	-2.9612675

Προκειμένου να ελέγξουμε την ακρίβεια της μεθόδου προσθέτουμε από δίπλα την ακριβή λύση για τις διάφορες τιμές του x και έχουμε:

x_i	y_i	$y(x_i)$
0	1.0000000	1.0000000
1.0000	0.6046404	0.6046752
2.0000	1.1850170	1.1850385
3.0000	1.2266003	1.2265660
4.0000	-0.5239232	-0.5239817
5.0000	-2.9612675	-2.9612661

Αν τώρα προσθέσουμε και μια στήλη με το απόλυτο σφάλμα έχουμε:

x_i	y_i	$y(x_i)$	$ y_i - y(x_i) $
0	1.0000000	1.0000000	0
1.0000	0.6046404	0.6046752	$3.48 \cdot 10^{-5}$
2.0000	1.1850170	1.1850385	$2.15 \cdot 10^{-5}$
3.0000	1.2266003	1.2265660	$3.42 \cdot 10^{-5}$
4.0000	-0.5239232	-0.5239817	$5.85 \cdot 10^{-5}$
5.0000	-2.9612675	-2.9612661	$1.41 \cdot 10^{-6}$

Το μέγιστο κατά απόλυτη τιμή σφάλμα είναι 5.99×10^{-5} στο 39ο διάστημα διαμέρισης ($y(39)$).

Αν είχαμε επιλέξει να πάρουμε $N = 100$ σημεία διαμέρισης το μέγιστο κατά απόλυτη τιμή σφάλμα θα ήταν ίσο με 7.28×10^{-6} , ενώ αν επιλέγαμε $N = 1000$ το μέγιστο σφάλμα ισούται με 7.10×10^{-9} .

Προκειμένου να κατασκευάσουμε μια μέθοδο 3ης αλγεβρικής τάξης ελαχιστοποιώντας την νόρμα του σφάλματος της επόμενης τάξης (4ης) έχουμε: Αν αναπτύξουμε σε σειρά Taylor και αντιστοιχίσουμε τους αντίστοιχους συντελεστές θα έχουμε: Συντελεστής του $h^4 f_{xxx}$: $\frac{1}{6}(b_2 c_2^3 + b_3 c_3^3) = \frac{1}{24}$
 Συντελεστής του $h^4 f_{xy}$: $a_{32} b_3 c_2 c_3 = \frac{1}{8}$
 Συντελεστής του $h^4 f_{yxx}$: $\frac{1}{2}(a_{32} b_3 c_2^2) = \frac{1}{24}$

Συνεπώς έχουμε:

$$\begin{aligned}\tau_1 &= \frac{1}{6}(b_2 c_2^3 + b_3 c_3^3) - \frac{1}{24} \\ \tau_2 &= a_{32} b_3 c_2 c_3 - \frac{1}{8} \\ \tau_3 &= \frac{1}{2}(a_{32} b_3 c_2^2) - \frac{1}{24} \\ \tau_4 &= -\frac{1}{24}\end{aligned}$$

Όπως είδαμε παραπάνω λύσαμε τις συνθήκες τάξης ως προς c_2 και c_3 . Αντικαθιστώντας τις σχέσεις αυτές στα τ_i βρίσκουμε:

$$\begin{aligned}\tau_1 &= \frac{4(c_2 + c_3) - 6c_2 c_3 - 3}{72} \\ \tau_2 &= \frac{4c_3 - 3}{24} \\ \tau_3 &= \frac{2c_2 - 1}{24}\end{aligned}$$

Αν θεωρήσουμε συνάρτηση τη νόρμα του σφάλματος θα έχουμε:

$$A(c_2, c_3) = \sqrt{\tau_1^2 + \tau_2^2 + \tau_3^2 + \tau_4^2}$$

Η συνάρτηση που μας δίνει το σφάλμα της επόμενης τάξης λοιπόν θα είναι η:

$$A(c_2, c_3) = \frac{1}{36} \sqrt{(13 - 12c_3 + 9c_3^2)c_2^2 + (-15 + 17c_3 - 12c_3^2)c_2 + 27 - c_3 + 40c_3^2}$$

Θα μπορούσαμε, για αποφυγή πολλών πράξεων, ισοδύναμα να ελαχιστοποιήσουμε την ακόλουθη συνάρτηση:

$$A_1(c_2, c_3) = 9c_2^2 c_3^2 - 12c_2 c_3^2 - 12c_2^2 c_3 + 40c_3^2 + 17c_2 c_3 + 13c_2^2 - 15c_2 - 60c_3 + 27$$

Η πρώτη μερική παράγωγος ως προς c_2 είναι:

$$\frac{\partial A_1(c_2, c_3)}{\partial c_2} = 18c_2 c_3^2 - 12c_3^2 - 24c_2 c_3 + 26c_2 + 17c_3 - 15$$

και η πρώτη μερική παράγωγος ως προς c_3 είναι:

$$\frac{\partial A_1(c_2, c_3)}{\partial c_3} = 18c_2^2 c_3 - 12c_2^2 - 24c_2 c_3 + 80c_3 + 17c_2 - 60$$

Λύνοντας το σύστημα των δύο συνθηκών πρώτης τάξης βρίσκουμε την πραγματική λύση $c_2 = 0.496505$ και $c_3 = 0.7517475$ (οι υπόλοιπες είναι μιγαδικές).

Προκειμένου να χαρακτηρίσουμε το ακρότατο υπολογίζουμε τις δεύτερες παραγώγους και έχουμε:

$$\frac{\partial^2 A_1(c_2, c_3)}{\partial c_2^2} = 18c_3^2 - 24c_3 + 26, \quad \frac{\partial^2 A_1(c_2, c_3)}{\partial c_2 \partial c_3} = 36c_2 c_3 - 24c_2 - 24c_3 + 17, \quad \frac{\partial^2 A_1(c_2, c_3)}{\partial c_3^2} = 18c_2^2 - 24c_2 + 80$$

Για τις τιμές $c_2 = 0.496505$ και $c_3 = 0.7517475$ ο πίνακας Hess είναι θετικά ορισμένος ($H_1 = 18.1303$, $H_2 = 1314.6$).

Η τιμή του σφάλματος της μεθόδου είναι $A(0.496505, 0.7517475) = 0.0418091$.

Αντίστοιχα το σφάλμα για τις άλλες δύο μεθόδους που αναφέραμε παραπάνω είναι για την $RK3a$ είναι $A(0.5, 1) = 0.0589256$ ενώ για την $RK3b$ είναι $A(0.5, 0.75) = 0.0418111$.

Παρατηρούμε ότι το σφάλμα της μεθόδου $RK3b$ είναι πάρα πολύ κοντά σε αυτό της βέλτιστης μεθόδου. Άλλωστε οι τιμές των c_2, c_3 είναι σχεδόν ίσες.

Στον παρακάτω πίνακα παρουσιάζουμε το μέγιστο απόλυτο σφάλμα της αριθμητικής λύσης για το πρόβλημα 1 για τις 3 μεθόδους τρίτης τάξης που είδαμε παραπάνω με διάστημα ολοκλήρωσης το $[0, 1]$ για διάφορες τιμές του βήματος h .

h	$RK3a$	$RK3b$	$RK3c$
0.2	$1.17 \cdot 10^{-3}$	$1.07 \cdot 10^{-3}$	$8.21 \cdot 10^{-4}$
0.1	$1.51 \cdot 10^{-4}$	$1.48 \cdot 10^{-4}$	$1.07 \cdot 10^{-4}$
0.05	$1.92 \cdot 10^{-5}$	$1.95 \cdot 10^{-5}$	$1.36 \cdot 10^{-5}$
0.02	$1.24 \cdot 10^{-6}$	$1.29 \cdot 10^{-6}$	$8.75 \cdot 10^{-7}$
0.01	$1.55 \cdot 10^{-7}$	$1.63 \cdot 10^{-7}$	$1.10 \cdot 10^{-7}$

2.3.3 Άμεσες μέθοδοι τέταρτης τάξης

Για την κατασκευή άμεσων μεθόδων τέταρτης τάξης απαιτούνται τέσσερα στάδια. Η μέθοδος γράφεται:

$$\begin{aligned} k_1 &= f(x_n, y_n) \\ k_2 &= f(x_n + c_2 h, y_n + h a_{21} k_1) \\ k_3 &= f(x_n + c_3 h, y_n + h(a_{31} k_1 + a_{32} k_2)) \\ k_4 &= f(x_n + c_4 h, y_n + h(a_{41} k_1 + a_{42} k_2 + a_{43} k_3)) \\ y_{n+1} &= y_n + h(b_1 k_1 + b_2 k_2 + b_3 k_3 + b_4 k_4) \end{aligned} \quad (2.23)$$

Εργαζόμενοι ανάλογα με τις περιπτώσεις των δύο και τριών σταδίων παίρνουμε τις συνθήκες έως και τρίτης τάξης

$$\begin{aligned}
 b_1 + b_2 + b_3 + b_4 &= 1 \\
 b_2 c_2 + b_3 c_3 + b_4 c_4 &= \frac{1}{2} \\
 b_2 c_2^2 + b_3 c_3^2 + b_4 c_4^2 &= \frac{1}{3} \\
 b_3 a_{32} c_2 + b_4 (a_{42} c_2 + a_{43} c_3) &= \frac{1}{6}
 \end{aligned} \tag{2.24}$$

επιπλέον οι συνθήκες τέταρτης τάξης

$$\begin{aligned}
 b_2 c_2^3 + b_3 c_3^3 + b_4 c_4^3 &= \frac{1}{4} \\
 b_3 c_3 a_{32} c_2 + b_4 c_4 (a_{42} c_2 + a_{43} c_3) &= \frac{1}{8} \\
 b_3 c_2^2 a_{32} + b_4 (a_{42} c_2^2 + a_{43} c_3^2) &= \frac{1}{12} \\
 b_4 c_2 a_{32} a_{43} &= \frac{1}{24}
 \end{aligned} \tag{2.25}$$

οι συνθήκες γραμμής είναι

$$\begin{aligned}
 c_2 &= a_{21} \\
 c_3 &= a_{31} + a_{32} \\
 c_4 &= a_{41} + a_{42} + a_{43}
 \end{aligned} \tag{2.26}$$

Η κατασκευή μιας μεθόδου τέταρτης τάξης απαιτεί την επίλυση του παραπάνω συστήματος εξισώσεων. Δεχόμενοι τις συνθήκες γραμμής (2.26) έχουμε να υπολογίσουμε ένα συντελεστή για το δεύτερο στάδιο, δύο συντελεστές για το τρίτο στάδιο, $s - 1$ συντελεστές για το s στάδιο, τέλος έχουμε και τα b_i . Το πλήθος των συντελεστών που υπολογίζονται είναι

$$1 + 2 + 3 + \dots + (s - 1) + s = \frac{s(s + 1)}{2}$$

Έτσι για τα $s = 4$ στάδια έχουμε να υπολογίσουμε 10 συντελεστές από οκτώ εξισώσεις (2.24) και (2.25), συνεπώς έχουμε δύο ελεύθερες παραμέτρους.

Ο Butcher (1963) [?] έδωσε την ακόλουθη απλοποιητική παραδοχή για την ευκολότερη επίλυση των συνθηκών.

$$\sum_{i=1}^s b_i a_{ij} = b_j (1 - c_j)$$

για $j = 2, 3, 4$.

Για $j = 4$ προκύπτει $0 = b_4(1 - c_4)$ άρα $c_4 = 1$.

Για $j = 3$ προκύπτει

$$b_4 a_{43} = b_3(1 - c_3)$$

Για $j = 2$ προκύπτει

$$b_3 a_{32} + b_4 a_{42} = b_2(1 - c_2).$$

Μετά την εφαρμογή των παραπάνω η τελευταία από τις συνθήκες τρίτης τάξης (2.24) και οι δύο τελευταίες από τις συνθήκες (2.25) προκύπτουν από τις υπόλοιπες. Πράγματι έχουμε:

$$\begin{aligned} b_3 a_{32} c_2 + b_4 a_{42} c_2 + b_4 a_{43} c_3 &= (b_3 a_{32} + b_4 a_{42}) c_2 + b_4 a_{43} c_3 = \\ &= b_2(1 - c_2) c_2 + b_3(1 - c_3) c_3 = \\ &= b_2 c_2 + b_3 c_3 - b_2 c_2^2 - b_3 c_3^2 = \\ &= \frac{1}{2} - b_4 c_4 - \frac{1}{3} + b_4 c_4^2 = \\ &= \frac{1}{6} + b_4 c_4 (c_4 - 1) = \frac{1}{6} \end{aligned}$$

$$\begin{aligned} b_3 c_2^2 a_{32} + b_4 a_{42} c_2^2 + b_4 a_{43} c_3^2 &= (b_3 a_{32} + b_4 a_{42}) c_2^2 + b_4 a_{43} c_3^2 = \\ &= b_2(1 - c_2) c_2^2 + b_3(1 - c_3) c_3^2 = \\ &= b_2 c_2^2 - b_2 c_2^3 + b_3 c_3^2 - b_3 c_3^3 = \\ &= \frac{1}{3} - b_4 c_4^2 - \frac{1}{4} + b_4 c_4^3 = \\ &= \frac{1}{12} - b_4 c_4^2 (c_4 - 1) = \frac{1}{12} \end{aligned}$$

$$\begin{aligned} b_4 c_2 a_{32} a_{43} &= b_3(1 - c_3) a_{32} c_2 = \\ &= b_3 a_{32} c_2 - b_3 a_{32} c_2 c_3 = \\ &= b_3 a_{32} c_2 - \frac{1}{8} + b_4 c_4 (a_{42} c_2 + a_{43} c_3) = \\ &= b_3 a_{32} c_2 - \frac{1}{8} + b_4 a_{42} c_2 + b_4 a_{43} c_3 = \frac{1}{6} - \frac{1}{8} = \frac{1}{24} \end{aligned}$$

Το παρακάτω σύστημα εξισώσεων είναι γραμμικό ως προς τα b_i

$$b_1 + b_2 + b_3 + b_4 = 1$$

$$b_2 c_2 + b_3 c_3 + b_4 c_4 = \frac{1}{2}$$

$$b_2 c_2^2 + b_3 c_3^2 + b_4 c_4^2 = \frac{1}{3}$$

$$b_2 c_2^3 + b_3 c_3^3 + b_4 c_4^3 = \frac{1}{4}$$

(2.27)

έτσι μπορούμε να εκφράσουμε τα βάρη b_i ως προς τους κόμβους c_i ως εξής

$$\begin{aligned}
 b_1 &= \frac{1 - 2c_2 - 2c_3 + 6c_2c_3}{12c_2c_3} \\
 b_2 &= \frac{2c_3 - 1}{12c_2(c_2 - 1)(c_2 - c_3)} \\
 b_3 &= \frac{1 - 2c_2}{12c_3(c_2 - c_3)(c_3 - 1)} \\
 b_4 &= \frac{3 - 4c_3 + c_2(6c_3 - 4)}{12(c_2 - 1)(c_3 - 1)}.
 \end{aligned} \tag{2.28}$$

Έχουμε δύο βαθμούς ελευθερίας, λύνουμε το παραπάνω σύστημα ως προς τους συντελεστές c_2 και c_3 , θεωρώντας ότι $c_4 = 1$, $c_2c_3 \neq 0$, $c_2 \neq \frac{1}{2}$, $c_3 \neq 1$ έχουμε:

$$\begin{aligned}
 a_{21} &= c_2 \\
 a_{31} &= c_3 - \frac{c_3(c_2 - c_3)}{c_2(4c_2 - 2)} \\
 a_{32} &= \frac{c_3(c_2 - c_3)}{c_2(4c_2 - 2)} \\
 a_{41} &= \frac{2 - 5c_3 + 4c_3^2 + 4c_2^2(3c_3^2 - 3c_3 + 1) - 3c_2(4c_3^2 - 5c_3 + 2)}{2c_2c_3(3 - 4c_3 + c_2(6c_3 - 4))} \\
 a_{42} &= \frac{2 + c_2^2 - 5c_3 + 4c_3^2 + c_2(4c_3^2 + 5c_3 - 3)}{2c_2(c_2 - c_3)(3 - 4c_3 + c_2(6c_3 - 4))} \\
 a_{43} &= \frac{(c_3 - 1)(2c_2^2 - 3c_2 + 1)}{c_3(c_2 - c_3)(3 - 4c_3 + c_2(6c_3 - 4))}
 \end{aligned}$$

Για τις διάφορες τιμές των c_2 και c_3 προκύπτουν μέθοδοι RK τέταρτης τάξης, από τις οποίες

μερικές από τις πιο δημοφιλείς είναι οι ακόλουθες:

0				0			
$\frac{1}{4}$	$\frac{1}{4}$			$\frac{1}{3}$	$\frac{1}{3}$		
$\frac{1}{2}$	0	$\frac{1}{2}$		$\frac{2}{3}$	$-\frac{1}{3}$	1	
1	1	-2	2	1	1	-1	1
	$\frac{1}{6}$	0	$\frac{2}{3}$	$\frac{1}{6}$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$
					$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Η δεύτερη μέθοδος (κανόνας 3/8) κατασκευάστηκε από τον Kutta το 1901.

Επίσης στο ίδιο άρθρο ο Kutta κατασκεύασε την γνωστή σαν κλασική μέθοδο Runge-Kutta τέταρτης τάξης (RK44). Θέτοντας $a_{31} = a_{41} = a_{42} = 0$ έχουμε $c_3 = a_{32}$ και $c_4 = a_{43}$ εκτός από το γραμμικό σύστημα (2.27) μένει μόνο η σχέση

$$b_3 c_3^2 c_2 + b_4 c_3 = \frac{1}{8}$$

και προκύπτει η μέθοδος

0			
$\frac{1}{2}$	$\frac{1}{2}$		
$\frac{1}{2}$	0	$\frac{1}{2}$	
1	0	0	1
	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{2}{6}$
		$\frac{1}{6}$	$\frac{1}{6}$

Για να δούμε την αποτελεσματικότητα των μεθόδων αυτών θα λύσουμε το πρόβλημα 1 με την μέθοδο RK4b που είδαμε παραπάνω.

Θέλουμε λοιπόν να λύσουμε την εξίσωση

$$y' = xy + 2x, \quad x \in [0,1], \quad y(0) = 1.$$

του οποίου ακριβής λύση είναι η $y(x) = -2 + 3e^{\frac{x^2}{2}}$.

η RKb είναι η ακόλουθη:

$$\begin{aligned}
 k_1 &= f(x_n, y_n) \\
 k_2 &= f\left(x_n + \frac{1}{3}h, y_n + h\frac{1}{3}k_1\right) \\
 k_3 &= f\left(x_n + \frac{2}{3}h, y_n + h\left(-\frac{1}{3}k_1 + k_2\right)\right) \\
 k_4 &= f\left(x_n + h, y_n + h(k_1 - k_2 + k_3)\right) \\
 y_{n+1} &= y_n + h\left(\frac{1}{8}k_1 + \frac{3}{8}k_2 + \frac{3}{8}k_3 + \frac{1}{8}k_4\right)
 \end{aligned}$$

Αν επιλέξουμε διαμέριση $N = 10$ τότε το μέγεθος του βήματος είναι $h = 0.1$ και το x παίρνει τις τιμές $0, 0.1, 0.2, 0.3, \dots, 0.9, 1$

Για $n = 0$ έχουμε:

$$\begin{aligned}
 k_1 &= f(x_0, y_0) = f(0, 1) = 0 \\
 k_2 &= f\left(x_0 + \frac{1}{3}h, y_0 + h\frac{1}{3}k_1\right) = f\left(0 + 0.1 \times \frac{1}{3}, 1 + 0.1 \times \frac{1}{3} \times 0\right) \\
 &= f(0.033, 1) = 0.033 \times 1 + 2 \times 0.033 = 0.1 \\
 k_3 &= f\left(x_0 + \frac{2}{3}h, y_0 + h\left(-\frac{1}{3}k_1 + k_2\right)\right) = f\left(0 + \frac{2}{3}0.1, 1 + 0.1\left(-\frac{1}{3}0 + 0.1\right)\right) \\
 &= f(0.0667, 1.01) = 0.0667 \times 1.01 + 2 \times 0.067 = 0.200667 \\
 k_4 &= f\left(x_0 + h, y_0 + h(k_1 - k_2 + k_3)\right) = f(0 + 0.1, 1 + 0.1(0 - 0.1 + 0.200667)) \\
 &= f(0.1, 1.01007) = 0.1 \times 1.01007 + 2 \times 0.1 = 0.301007 \\
 y_1 &= y_0 + h\left(\frac{1}{8}k_1 + \frac{3}{8}k_2 + \frac{3}{8}k_3 + \frac{1}{8}k_4\right) \\
 &= 1 + 0.1\left(\frac{1}{8}0 + \frac{3}{8}0.1 + \frac{3}{8}0.200667 + \frac{1}{8}0.301007\right) = 1.01504
 \end{aligned}$$

Συνεχίζοντας για $n = 1$ έχουμε:

$$k_1 = f(x_1, y_1) = f(0.1, 1.01504) = 0.301504$$

$$\begin{aligned} k_2 &= f\left(x_1 + \frac{1}{3}h, y_1 + h\frac{1}{3}k_1\right) = f\left(0.1 + 0.1 \times \frac{1}{3}, 1.01504 + 0.1 \times \frac{1}{3} \times 0.301504\right) \\ &= f(0.133, 1.02509) = 0.133 \times 1.02509 + 2 \times 0.133 = 0.402337 \end{aligned}$$

$$\begin{aligned} k_3 &= f\left(x_1 + \frac{2}{3}h, y_1 + h\left(-\frac{1}{3}k_1 + k_2\right)\right) = f\left(0.1 + \frac{2}{3} \times 0.1, 1.01504 + 0.1\left(-\frac{1}{3} \times 0.301504 + 0.402337\right)\right) \\ &= f(0.1667, 1.04522) = 0.167 \times 1.10693 + 2 \times 0.167 = 0.507638 \end{aligned}$$

$$\begin{aligned} k_4 &= f\left(x_1 + h, y_1 + h(k_1 - k_2 + k_3)\right) = f(0.1 + 0.1, 1.01504 + 0.1(0.301504 - 0.402337 + 0.507638)) \\ &= f(0.2, 1.05572) = 0.2 \times 1.05572 + 2 \times 0.2 = 0.611144 \end{aligned}$$

$$\begin{aligned} y_2 &= y_1 + h\left(\frac{1}{8}k_1 + \frac{3}{8}k_2 + \frac{3}{8}k_3 + \frac{1}{8}k_4\right) \\ &= 1.01504 + 0.1\left(\frac{1}{8} \times 0.301504 + \frac{3}{8} \times 0.402337 + \frac{3}{8} \times 0.507638 + \frac{1}{8} \times 0.611144\right) = 1.06057 \end{aligned}$$

για $n = 2$ έχουμε:

$$k_1 = f(x_2, y_2) = f(0.2, 1.06057) = 0.612114$$

$$\begin{aligned} k_2 &= f\left(x_2 + \frac{1}{3}h, y_2 + h\frac{1}{3}k_1\right) = f\left(0.2 + 0.1 \times \frac{1}{3}, 1.06057 + 0.1 \times \frac{1}{3} \times 0.612114\right) \\ &= f(0.2333, 1.08097) = 0.2333 \times 1.08097 + 2 \times 0.2333 = 0.71879 \end{aligned}$$

$$\begin{aligned} k_3 &= f\left(x_2 + \frac{2}{3}h, y_2 + h\left(-\frac{1}{3}k_1 + k_2\right)\right) = f\left(0.2 + \frac{2}{3} \times 0.1, 1.06057 + 0.1\left(-\frac{1}{3} \times 0.612114 + 0.71879\right)\right) \\ &= f(0.2667, 1.11205) = 0.2667 \times 1.11205 + 2 \times 0.2667 = 0.829984 \end{aligned}$$

$$\begin{aligned} k_4 &= f\left(x_2 + h, y_2 + h(k_1 - k_2 + k_3)\right) = f(0.2 + 0.1, 1.06057 + 0.1(0.612114 - 0.71879 + 0.829984)) \\ &= f(0.3, 1.1329) = 0.3 \times 1.1329 + 2 \times 0.3 = 0.93987 \end{aligned}$$

$$\begin{aligned} y_3 &= y_2 + h\left(\frac{1}{8}k_1 + \frac{3}{8}k_2 + \frac{3}{8}k_3 + \frac{1}{8}k_4\right) \\ &= 1.06057 + 0.1\left(\frac{1}{8} \times 0.612114 + \frac{3}{8} \times 0.71879 + \frac{3}{8} \times 0.829984 + \frac{1}{8} \times 0.93987\right) = 1.13805 \end{aligned}$$

και συνεχίζοντας με τον ίδιο τρόπο βρίσκουμε όλες τις τιμές του y_i .

Στην παρακάτω συνάρτηση υλοποιούμε τη μέθοδο RK4b

```
function [x,y] = RK4b(fun,y0,x0,xn,N)

n = length(y0) ;

% step length
h = (xn-x0)/N ;

% creates the vectors to be used
x = zeros(N+1,1);
y = zeros(N+1,n);
% starting point of the integration interval
x(1) = x0 ;

% initial condition
y(1,:) = y0 ;

for i=1:N
    x(i+1) = x(i) + h ;
    Y1 = y(i,:) ;
    F1 = feval(fun,Y1) ;
    Y2 = y(i,:) + h*F1/3 ;
    F2 = feval(fun,x+h/3,Y2) ;
    Y3 = y(i,:) + h*(-F1/3+F2) ;
    F3 = feval(fun,x+2*h/3,Y3) ;
    Y4 = y(i,:) + h*(F1-F2+F3) ;
    F4 = feval(fun,x+h,Y4) ;
    y(i+1,:) = y(i,:) + h*(F1+3*F2+3*F3+F4)/8 ;
end
end
```

Τα αποτελέσματα φαίνονται στον ακόλουθο πίνακα.

x_i	y_i	$y(x_i)$	$\ y_i - y(x_i)\ $
0	1.0000000	1.0000000	0
0.1000	1.0150000	1.0150376	$2.08 \cdot 10^{-8}$
0.2000	1.0605265	1.0606040	$8.31 \cdot 10^{-8}$
0.3000	1.1379578	1.1380836	$1.89 \cdot 10^{-7}$
0.4000	1.2496691	1.2498612	$3.42 \cdot 10^{-7}$
0.5000	1.3991539	1.3994454	$5.49 \cdot 10^{-7}$
0.6000	1.5912061	1.5916521	$8.13 \cdot 10^{-7}$
0.7000	1.8321760	1.8328639	$1.14 \cdot 10^{-6}$
0.8000	2.1303193	2.1313833	$1.53 \cdot 10^{-6}$
0.9000	2.4962656	2.4979075	$1.96 \cdot 10^{-6}$
1.0000	2.9436440	2.9461638	$2.41 \cdot 10^{-6}$

Το μέγιστο κατά απόλυτη τιμή σφάλμα είναι 2.41×10^{-6} στο τέλος του διαστήματος ολοκλήρωσης.

Με την επόμενη function υλοποιούμε μεθόδους με τέσσερα στάδια όταν οι συντελεστές c_{ij} , a_{ij} και b_j δίνονται σαν ορίσματα εισόδου.

```
function [x,y] = RK_s4(fun,a,b,c,y0,x0,xn,N)

% Runge Kutta method for systems
n = length(y0) ;

% step length
h = (xn-x0)/N ;

% creates the vectors to be used
x = zeros(N+1,1);
y = zeros(N+1,n);
% starting point of the integration interval
x(1) = x0 ;

% initial condition
y(1,:) = y0 ;

for i=1:N
    x(i+1) = x(i) + h ;
    Y1 = y(i,:) ;
    F1 = feval(fun,Y1) ;
    Y2 = y(i,:) + a(2,1)*h*F1 ;
    F2 = feval(fun,x+c(2)*h,Y2) ;
    Y3 = y(i,:) + (a(3,1)*F1+a(3,2)*F2)*h ;
    F3 = feval(fun,x+c(3)*h,Y3) ;
    Y4 = y(i,:) + (a(4,1)*F1+a(4,2)*F2+a(4,3)*F3)*h ;
    F4 = feval(fun,x+c(4)*h,Y4) ;
    y(i+1,:) = y(i,:) + h*(b(1)*F1+b(2)*F2+b(3)*F3+b(4)*F4) ;
end
end
```

Αν είχαμε επιλέξει να πάρουμε $N = 100$ σημεία διαμέρισης το μέγιστο κατά απόλυτη τιμή σφάλμα θα ήταν ίσο με 2.71×10^{-10} στο τέλος του διαστήματος ολοκλήρωσης, ενώ αν επιλέγαμε $N = 1000$ το μέγιστο σφάλμα ισούται με 2.89×10^{-14} .

Στον παρακάτω πίνακα παρουσιάζουμε το μέγιστο απόλυτο σφάλμα της αριθμητικής λύσης για το πρόβλημα 1 για τις 3 μεθόδους τέταρτης τάξης που είδαμε παραπάνω με διάστημα ολοκλήρωσης το $[0,1]$ για διάφορες τιμές του βήματος h .

h	$RK4a$	$RK4b$	$RK4c$
0.2	$4.89 \cdot 10^{-6}$	$3.39 \cdot 10^{-5}$	$1.38 \cdot 10^{-5}$
0.1	$5.25 \cdot 10^{-7}$	$2.41 \cdot 10^{-6}$	$7.91 \cdot 10^{-7}$
0.05	$4.10 \cdot 10^{-8}$	$1.61 \cdot 10^{-7}$	$4.65 \cdot 10^{-8}$
0.02	$1.19 \cdot 10^{-9}$	$4.28 \cdot 10^{-9}$	$1.14 \cdot 10^{-9}$
0.01	$7.78 \cdot 10^{-11}$	$2.71 \cdot 10^{-10}$	$6.99 \cdot 10^{-11}$

Τέλος, δίνουμε μια function που υλοποιεί οποιαδήποτε μέθοδο Runge Kutta ανεξάρτητα από τον αριθμό των σταδίων ο οποίος λαμβάνεται από το μήκος του διανύσματος b .

```
function [x,y] = RK_gen(fun,a,b,y0,x0,xn,N)

% Runge Kutta method
s = length(b) ;    % number of stages
c = sum(a') ;      % row condition
n = length(y0) ;  % number of equations
h = (xn-x0)/N ;   % step length

x = zeros(N+1,1);
y = zeros(N+1,n);

x(1) = x0 ;
y(1,:) = y0 ;
for i=1:N
    x(i+1) = x(i) + h ;
    F = zeros(s,n) ;
    for j = 1:s
        F(j,:) = feval(fun, y(i,:) + h*a(j,:)*F) ;
    end
    y(i+1,:) = y(i,:) + h*b*F ;
end
end
```

2.3.4 Συνθήκες με χρήση πινάκων

Ένας τρόπος για να γράψουμε τις συνθήκες είναι με τη χρήση πινάκων. Για μια μέθοδο με δύο στάδια έχουμε

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad b = (b_1, b_2)^T, \quad C = \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix}$$

στην περίπτωση άμεσης μεθόδου οι πίνακες γίνονται

$$A = \begin{pmatrix} 0 & 0 \\ a_{21} & 0 \end{pmatrix} \quad b = (b_1, b_2) \quad C = \begin{pmatrix} 0 & 0 \\ 0 & c_2 \end{pmatrix}$$

Τότε η συνθήκη (2.8) μπορεί να γραφεί

$$b.e = 1$$

και η συνθήκη (2.9)

$$b.C.e = \frac{1}{2}$$

Για μια μέθοδο με τρία στάδια οι πίνακες είναι

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \quad b = (b_1, b_2, b_3)^T, \quad C = \begin{pmatrix} c_1 & 0 & 0 \\ 0 & c_2 & 0 \\ 0 & 0 & c_3 \end{pmatrix}$$

στην περίπτωση άμεσης μεθόδου οι πίνακες γίνονται

$$A = \begin{pmatrix} 0 & 0 & 0 \\ a_{21} & 0 & 0 \\ a_{31} & a_{32} & 0 \end{pmatrix}, \quad b = (b_1, b_2, b_3)^T, \quad C = \begin{pmatrix} 0 & 0 & 0 \\ 0 & c_2 & 0 \\ 0 & 0 & c_3 \end{pmatrix}$$

τότε οι συνθήκες (2.22) γράφονται

$$b.e = 1$$

$$b.C.e = \frac{1}{2}$$

$$b.C^2.e = \frac{1}{3} \tag{2.29}$$

$$b.A.C.e = \frac{1}{6}$$

Για μια μέθοδο με τέσσερα στάδια οι πίνακες είναι

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}, \quad b = (b_1, b_2, b_3, b_4)^T, \quad C = \begin{pmatrix} c_1 & 0 & 0 & 0 \\ 0 & c_2 & 0 & 0 \\ 0 & 0 & c_3 & 0 \\ 0 & 0 & 0 & c_4 \end{pmatrix}$$

στην περίπτωση άμεσης μεθόδου οι πίνακες γίνονται

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ a_{21} & 0 & 0 & 0 \\ a_{31} & a_{32} & 0 & 0 \\ a_{41} & a_{42} & a_{43} & 0 \end{pmatrix}, \quad b = (b_1, b_2, b_3, b_4)^T, \quad C = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & c_2 & 0 & 0 \\ 0 & 0 & c_3 & 0 \\ 0 & 0 & 0 & c_4 \end{pmatrix}$$

τότε οι συνθήκες έως και τρίτης τάξης (2.24) γράφονται όπως και στα τρία στάδια σαν (2.29) και οι συνθήκες τέταρτης τάξης (2.25) γράφονται

$$\begin{aligned} b.C^3.e &= \frac{1}{4} \\ b.C.A.C.e &= \frac{1}{8} \\ b.A.C^2.e &= \frac{1}{12} \\ b.A.A.C.e &= \frac{1}{24} \end{aligned} \tag{2.30}$$

2.4 Βιβλιογραφία

- Αβδελάς Γ., Σίμος Θ., Αριθμητική Ανάλυση, Εκδόσεις Συμεών, 2004.
 Βραχάτης Μ., Αριθμητική Ανάλυση, Εκδόσεις Κλειδάριθμος, 2012.
 Βραχάτης Μ., Αριθμητική Ανάλυση Συνήθεις Διαφορικές Εξισώσεις, Εκδόσεις Κλειδάριθμος, 2012.
 Γουσίδου - Κουτίτα Μ., Αριθμητική Ανάλυση, Εκδόσεις Αφοι Κυριακίδη, 2017.
 Μισυρλής Ν., Αριθμητική Ανάλυση, Εκδόσεις ΕΚΠΑ, 2017.
 Παπαγεωργίου Γ.Σ. και Τσίτουρας Χ., Αριθμητική Ανάλυση με εφαρμογές σε MATLAB και Mathematica, Εκδόσεις Τσότρας, 2015.
 Φαμέλης Ι., Υπολογιστικά Μαθηματικά, Εκδόσεις Κριτική, 2021.
 Moler C., Αριθμητικές Μέθοδοι με το MATLAB, Εκδόσεις Κλειδάριθμος, 2010.
 Hairer E., Norset S.P., Wanner G., Solving Ordinary Differential Equations, Springer 2000
 Butcher J., Numerical Methods for Ordinary Differential Equations, Wiley, 2016

Κεφάλαιο 3

Μέθοδοι Runge-Kutta

Τα δέντρα είναι συνδυαστικά γραφήματα τα οποία είναι συνδεδεμένα και δεν έχουν κύκλους. Ένα δέντρο με ρίζα (rooted tree) είναι ένα δέντρο στο οποίο μια συγκεκριμένη κορυφή (vertex) θεωρείται η ρίζα (root).

Ορισμός 3.0.1 Ένα γράφημα t λέγεται δέντρο με ρίζα (rooted tree) αν υπάρχει μια κορυφή του την οποία αποκαλούμε ρίζα και έχει τις εξής ιδιότητες:

1. Οι κορυφές χωρίζονται σε επίπεδα $0, 1, \dots, H$ με H το ύψος του δέντρου.
2. Στο μηδενικό επίπεδο έχουμε μόνο τη ρίζα του δέντρου.
3. Στο κάθε επίπεδο πλην το μηδενικό βρίσκονται τα παιδιά των κορυφών του προηγούμενου επιπέδου ορίζοντας τα κλαδιά που ξεκινούν από τον γονέα και καταλήγουν στο παιδί.
4. Κάθε κορυφή εκτός από την ρίζα έχει γονέα, η ρίζα όμως δεν έχει κανέναν.

Η τάξη ενός δέντρου $r(t)$ είναι ο αριθμός των κορυφών του.

Η συνάρτηση $\gamma(t)$ είναι η πυκνότητα του δέντρου και ορίζεται αναδρομικά. Η πυκνότητα ενός δέντρου με ρίζα ορίζεται σαν το γινόμενο των πυκνοτήτων των δέντρων με ρίζα που προκύπτουν αν από το δέντρο αποκοπεί η ρίζα επί την τάξη του δέντρου. Η πυκνότητα του δέντρου πρώτης τάξης ορίζεται ίση με τη μονάδα.

Υποδέντρα ονομάζονται τα δέντρα τα οποία προκύπτουν αν αφαιρέσουμε την ρίζα μαζί με τα κλαδιά που ξεκινούν από αυτή. Σε κάθε δέντρο αντιστοιχούν πολλά δέντρα με ρίζα και μπορούν να θεωρηθούν σαν μια κλάση ισοδυναμίας. Στο εξής, εκτός και αν αναφέρεται διαφορετικά, όταν λέμε δέντρο θα εννοούμε ένα δέντρο με ρίζα.

Σε κάθε δέντρο t αντιστοιχεί ένα πολυώνυμο ως προς τους συντελεστές της μεθόδου $\Phi(t)$

$$\Phi(t) = \sum_i b_i \Phi_i(t),$$

όπου

$$\Phi_i(t) = \sum_{j,k,l,\dots} a_{ij} a_{jk} a_{kl} \dots$$

Ισχύει το ακόλουθο αποτέλεσμα (βλέπε [?])

Θεώρημα 3.0.1 Μια μέθοδος Runge-Kutta (2.3) έχει τάξη p αν

$$\Phi(t) = \frac{1}{\gamma(t)}$$

για όλα τα δέντρα τάξης $\leq p$.

Η ρίζα αντιστοιχεί στο διάνυσμα b και η τελευταία κορυφή κάθε υποδέντρου, δηλαδή η κορυφή που δεν έχει παιδί αντιστοιχεί στον πίνακα C . Όλες οι άλλες κορυφές αντιστοιχούν στον πίνακα A .

Συγκεκριμένα:

- Υπάρχει ένα μόνο δέντρο 1ης τάξης (τ) που αποτελείται από μόνο μία κορυφή την ρίζα. Στο δέντρο αυτό αντιστοιχεί η συνθήκη $b.e = 1$.

- Υπάρχει επίσης ένα δέντρο 2ης τάξης (t_{21}) που έχει δύο κορυφές την ρίζα και το παιδί



έχει πυκνότητα $\gamma(t_{21}) = r(t_{21})\gamma(\tau) = 2$ και η συνθήκη που αντιστοιχεί είναι η $b.C.e = 1/2$.

- Υπάρχουν δύο δέντρα 3ης τάξης, τα δέντρα τρίτης τάξης έχουν 3 κορυφές και εδώ διακρίνονται οι περιπτώσεις:

- Η ρίζα έχει δύο παιδιά (t_{31})



με πυκνότητα $\gamma(t_{31}) = r(t_{31})\gamma(\tau)\gamma(\tau) = 3$ και η συνθήκη που αντιστοιχεί είναι η $b.C^2.e = 1/3$.

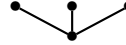
- Η ρίζα έχει ένα παιδί και αυτό το παιδί έχει ένα δικό του παιδί (t_{32})



έχει πυκνότητα $\gamma(t_{32}) = r(t_{32})\gamma(t_{21}) = 6$ και η συνθήκη που αντιστοιχεί είναι η $b.A.C.e = 1/6$.

- Υπάρχουν τέσσερα δέντρα 4ης τάξης, εδώ διακρίνονται οι περιπτώσεις:

- Η ρίζα έχει τρία παιδιά (t_{41})



έχει πυκνότητα $\gamma(t_{41}) = r(t_{41})\gamma(\tau)\gamma(\tau)\gamma(\tau) = 4$ και η συνθήκη που αντιστοιχεί είναι η $b.C^3.e = 1/4$.

- Η ρίζα έχει δύο παιδιά, το ένα από αυτά τα παιδιά έχει ένα δικό του παιδί (t_{42})



έχει πυκνότητα $\gamma(t_{42}) = r(t_{42})\gamma(\tau)\gamma(t_{21}) = 4 \cdot 1 \cdot 2 = 8$ και η συνθήκη που αντιστοιχεί είναι η $b.C.A.C.e = \frac{1}{8}$

- Το δέντρο έχει ένα παιδί και αυτό έχει δύο παιδιά (t_{43})



έχει πυκνότητα $\gamma(t_{43}) = r(t_{43})\gamma(t_{31}) = 4 \cdot 3 = 12$ και η συνθήκη που αντιστοιχεί είναι η $b.A.C^2.e = 1/12$.

- Η ρίζα έχει ένα παιδί, αυτό έχει παιδί και αυτό επίσης έχει ένα παιδί.



έχει πυκνότητα $\gamma(t_{44}) = r(t_{44})\gamma(t_{32}) = 4 \cdot 6 = 24$ και η συνθήκη που αντιστοιχεί είναι η $b.A.A.C.e = \frac{1}{24}$.

Ο αριθμός των δέντρων έως και δέκατης τάξης φαίνεται στον ακόλουθο πίνακα

τάξη	1	2	3	4	5	6	7	8	9	10
πλήθος	1	1	2	4	9	20	48	115	286	719

Στον πίνακα που ακολουθεί αναγράφονται τα δέντρα και οι αντίστοιχες συνθήκες έως και τέταρτης τάξης.

$r(t)$	t	$\Phi(t)$	$\Phi(t)$	$\gamma(t)$
1		$\sum_{i=1}^s b_i$	$b^T e$	1
2		$\sum_{i=1}^s b_i c_i$	$b^T C e$	2
3		$\sum_{i=1}^s b_i c_i^2$	$b^T C^2 e$	3
3		$\sum_{i,j=1}^s b_i a_{ij} c_j$	$b^T A C e$	6
4		$\sum_{i=1}^s b_i c_i^3$	$b^T C^3 e$	4
4		$\sum_{i,j=1}^s b_i c_i a_{ij} c_j$	$b^T C A C e$	8
4		$\sum_{i,j=1}^s b_i a_{ij} c_j^2$	$b^T A C^2 e$	12
4		$\sum_{i,j,k=1}^s b_i a_{ij} a_{jk} c_k$	$b^T A^2 C e$	24

Πίνακας 1: Δέντρα με ρίζα, στοιχειώδη βάρη και η συνάρτηση πυκνότητας έως και τέταρτης τάξης.

Ο Kutta (1901) ασχολήθηκε και με μεθόδους πέμπτης τάξης. Υπάρχουν εννέα συνθήκες πέμπτης τάξης όπως φαίνονται στον πίνακα 2.

$r(t)$	t	$\Phi(t)$	$\Phi(t)$	$\gamma(t)$
5		$\sum_{i,j,k,l=1}^s b_i a_{ij} a_{jk} a_{kl} c_l$	$b^T A^3 C e$	120
5		$\sum_{i,j,k=1}^s b_i c_i a_{ij} a_{jk} c_k$	$b^T C A^2 C e$	30
5		$\sum_{i,j=1}^s b_i c_i^2 a_{ij} c_j$	$b^T C^2 A C e$	10
5		$\sum_{i,j,k=1}^s b_i c_i a_{ij} c_j^2$	$b^T C A C^2 e$	15
5		$\sum_{i,j,k=1}^s b_i a_{ij} c_j^3$	$b^T A C^3 e$	20
5		$\sum_{i,j,k=1}^s b_i a_{ij} c_j a_{jk} c_k$	$b^T A C A C e$	40
5		$\sum_{i=1}^s b_i \left(\sum_{j=1}^4 a_{ij} c_j \right)^2$	$b^T ((A C e) * (A C e))$	20
5		$\sum_{i,j,k=1}^s b_i a_{ij} a_{jk} c_k^2$	$b^T A^2 C^2 e$	60
5		$\sum_{i,j,k=1}^s b_i c_i^4$	$b^T C^4 e$	5

Πίνακας 2: Δέντρα με ρίζα, στοιχειώδη βάρη και η συνάρτηση πυκνότητας πέμπτης τάξης.

Συνεπώς για μια άμεση μέθοδο πέμπτης τάξης πρέπει να ικανοποιούνται 17 εξισώσεις. Δεν μπορούμε να κατασκευάσουμε άμεση μέθοδο πέμπτης τάξης με $s = 5$ στάδια γιατί έχουμε να υπολογίσουμε 15 συντελεστές, οπότε πρέπει να χρησιμοποιήσουμε $s = 6$ στάδια. Οι μέθοδοι που κατασκεύασε ο Kutta και διορθώθηκαν από τον Nystrom είναι οι ακόλουθες

0						
$\frac{1}{5}$	$\frac{1}{5}$					
$\frac{2}{5}$	0	$\frac{2}{5}$				
1	$\frac{9}{4}$	-5	$\frac{15}{4}$			
$\frac{3}{5}$	$-\frac{63}{100}$	$\frac{9}{5}$	$-\frac{13}{20}$	$\frac{2}{25}$		
$\frac{4}{5}$	$-\frac{6}{25}$	$\frac{4}{5}$	$\frac{2}{15}$	$\frac{8}{75}$	0	
	$\frac{17}{144}$	0	$\frac{25}{36}$	$\frac{1}{72}$	$-\frac{25}{72}$	$\frac{25}{48}$

και

0						
$\frac{1}{3}$	$\frac{1}{3}$					
$\frac{2}{5}$	$\frac{4}{25}$	$\frac{6}{25}$				
1	$\frac{1}{4}$	-3	$\frac{15}{4}$			
$\frac{2}{3}$	$\frac{2}{27}$	$\frac{10}{9}$	$-\frac{50}{81}$	$\frac{8}{81}$		
$\frac{4}{5}$	$\frac{2}{25}$	$\frac{12}{25}$	$\frac{2}{15}$	$\frac{8}{75}$	0	
	$\frac{23}{192}$	0	$\frac{125}{192}$	0	$-\frac{27}{64}$	$\frac{125}{192}$

Για την κατασκευή μεθόδου έκτης τάξης απαιτούνται επτά στάδια, γενικά ο απαιτούμενος αριθμός σταδίων για την κατασκευή μεθόδων έως και όγδοης τάξης είναι:

τάξη	1	2	3	4	5	6	7	8
στάδια	1	2	3	4	6	7	9	11

Βέβαια μπορούμε να μειώσουμε τον αριθμό των σταδίων αν θεωρήσουμε πεπλεγμένες μεθόδους.

3.1 Πεπλεγμένες (implicit) μέθοδοι

Θα ασχοληθούμε τώρα με την κατασκευή πεπλεγμένων (έμμεσων) μεθόδων και θα θεωρήσουμε μέθοδο με δύο στάδια

$$\begin{aligned} k_1 &= f(x_n + c_1 h, y_n + h(a_{11}k_1 + a_{12}k_2)) \\ k_2 &= f(x_n + c_2 h, y_n + h(a_{21}k_1 + a_{22}k_2)) \\ y_{n+1} &= y_n + h(b_1k_1 + b_2k_2) \end{aligned}$$

ο αντίστοιχος πίνακας Butcher είναι

$$\begin{array}{c|cc} c_1 & a_{11} & a_{12} \\ c_2 & a_{21} & a_{22} \\ \hline & b_1 & b_2 \end{array}$$

Έχουμε οκτώ παραμέτρους συνεπώς μπορούμε να ικανοποιήσουμε τις οκτώ συνθήκες τάξης. Λύνουμε το σύστημα των οκτώ εξισώσεων και παίρνουμε δύο λύσεις, η πρώτη είναι

$$c_1 = \frac{3 - \sqrt{3}}{6}, \quad c_2 = 1 - c_1, \quad b_1 = b_2 = \frac{1}{2},$$

και

$$\begin{aligned} a_{11} &= \frac{3 - \sqrt{3} - 24a_{12} - 12\sqrt{3}a_{12}}{12}, \\ a_{21} &= \frac{1 + \sqrt{3} + 8a_{12} + 4\sqrt{3}a_{12}}{4}, \\ a_{22} &= \frac{3 - \sqrt{3} - 6a_{12}}{6}, \end{aligned}$$

και η δεύτερη είναι

$$c_1 = \frac{3 + \sqrt{3}}{6}, \quad c_2 = 1 - c_1, \quad b_1 = b_2 = \frac{1}{2},$$

και

$$\begin{aligned} a_{11} &= \frac{3 + \sqrt{3} - 24a_{12} + 12\sqrt{3}a_{12}}{12}, \\ a_{21} &= \frac{1 - \sqrt{3} + 8a_{12} - 4\sqrt{3}a_{12}}{4}, \\ a_{22} &= \frac{3 + \sqrt{3} - 6a_{12}}{6}. \end{aligned}$$

Οι μέθοδοι τύπου Gauss προκύπτουν αν θέσουμε $a_{11} = a_{22}$ για την πρώτη μέθοδο οι συντελεστές a_{ij} γίνονται

$$a_{11} = a_{22} = \frac{1}{4}, \quad a_{12} = \frac{3 - 2\sqrt{3}}{12}, \quad a_{21} = \frac{3 + 2\sqrt{3}}{12}$$

και για τη δεύτερη μέθοδο

$$a_{11} = a_{22} = \frac{1}{4}, \quad a_{12} = \frac{3+2\sqrt{3}}{12}, \quad a_{21} = \frac{3-2\sqrt{3}}{12}.$$

ο αντίστοιχος πίνακας Butcher είναι

$\frac{3-\sqrt{3}}{6}$	$\frac{1}{4}$	$\frac{3-2\sqrt{3}}{12}$	$\frac{3+\sqrt{3}}{6}$	$\frac{1}{4}$	$\frac{3+2\sqrt{3}}{12}$
$\frac{3+\sqrt{3}}{6}$	$\frac{3+2\sqrt{3}}{12}$	$\frac{1}{4}$	$\frac{3-\sqrt{3}}{6}$	$\frac{3-2\sqrt{3}}{12}$	$\frac{1}{4}$
	$\frac{1}{2}$	$\frac{1}{2}$		$\frac{1}{2}$	$\frac{1}{2}$

Αν θέσουμε $a_{12} = 0$ έχουμε διαγώνια πεπλεγμένη μέθοδο γενικά η μέθοδος

m	m	
$1-m$	$1-2m$	m
	$\frac{1}{2}$	$\frac{1}{2}$

έχει αλγεβρική τάξη δύο εκτός από την περίπτωση

$$m = \frac{3 \pm \sqrt{3}}{6}$$

που έχει τρίτη αλγεβρική τάξη.

Μπορούμε με όμοιο τρόπο να κατασκευάσουμε πεπλεγμένη μέθοδο έκτης τάξης με τρία στάδια και μέθοδο όγδοης τάξης με τέσσερα στάδια.

$\frac{1}{2} - \frac{\sqrt{15}}{10}$	$\frac{5}{36}$	$\frac{2}{9} - \frac{\sqrt{15}}{15}$	$\frac{5}{36} - \frac{\sqrt{15}}{30}$
$\frac{1}{2}$	$\frac{5}{36} + \frac{\sqrt{15}}{24}$	$\frac{2}{9}$	$\frac{5}{36} - \frac{\sqrt{15}}{24}$
$\frac{1}{2} + \frac{\sqrt{15}}{10}$	$\frac{5}{36} + \frac{\sqrt{15}}{30}$	$\frac{2}{9} + \frac{\sqrt{15}}{15}$	$\frac{5}{36}$
	$\frac{5}{10}$	$\frac{4}{9}$	$\frac{5}{18}$

Γενικά έχει αποδειχθεί ότι υπάρχει μια μέθοδος τάξης $2s$ με s στάδια.

3.2 Μέθοδοι και αποτελέσματα

Η μέθοδος πέμπτης τάξης με έξι στάδια των Dormand και Prince (*RKs6DP*)

0						
$\frac{1}{5}$	$\frac{1}{5}$					
$\frac{3}{10}$	$\frac{3}{40}$	$\frac{9}{40}$				
$\frac{3}{5}$	$\frac{3}{10}$	$-\frac{9}{10}$	$\frac{6}{5}$			
$\frac{2}{3}$	$\frac{226}{729}$	$-\frac{25}{27}$	$\frac{880}{729}$	$\frac{55}{729}$		
1	$-\frac{181}{27}$	$\frac{5}{2}$	$-\frac{266}{297}$	$-\frac{91}{27}$	$\frac{189}{55}$	
	$\frac{19}{216}$	0	$\frac{1000}{2079}$	$-\frac{125}{216}$	$\frac{81}{88}$	$\frac{5}{56}$

Η μέθοδος πέμπτης τάξης με έξι στάδια (RK56)

0						
$\frac{1}{4}$	$\frac{1}{4}$					
$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$				
$\frac{1}{2}$	0	0	$\frac{1}{2}$			
$\frac{3}{4}$	$\frac{3}{16}$	$-\frac{3}{8}$	$\frac{3}{8}$	$\frac{9}{16}$		
1	$-\frac{3}{7}$	$\frac{8}{7}$	$\frac{6}{7}$	$-\frac{12}{7}$	$\frac{8}{7}$	
	$\frac{7}{90}$	0	$\frac{16}{45}$	$\frac{2}{15}$	$\frac{16}{45}$	$\frac{7}{90}$

Η μέθοδος έκτης τάξης με επτά στάδια (RK67)

0							
$\frac{1}{3}$	$\frac{1}{3}$						
$\frac{2}{3}$	0	$\frac{2}{3}$					
$\frac{1}{3}$	$\frac{1}{12}$	$\frac{1}{3}$	$-\frac{1}{12}$				
$\frac{5}{6}$	$\frac{25}{48}$	$-\frac{55}{24}$	$\frac{35}{48}$	$\frac{15}{8}$			
$\frac{1}{6}$	$\frac{3}{20}$	$-\frac{11}{24}$	$-\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{10}$		
1	$-\frac{261}{260}$	$\frac{33}{13}$	$\frac{43}{156}$	$-\frac{118}{39}$	$\frac{32}{195}$	$\frac{80}{39}$	
	$\frac{13}{200}$	0	$\frac{11}{40}$	$\frac{11}{40}$	$\frac{4}{25}$	$\frac{4}{25}$	$\frac{13}{200}$

3.3 Βιβλιογραφία

Βραχάτης Μ., Αριθμητική Ανάλυση Συνήθεις Διαφορικές Εξισώσεις, Εκδόσεις Κλειδάριθμος, 2012.

Παπαγεωργίου Γ.Σ. και Τσίτουρας Χ., Αριθμητική Ανάλυση με εφαρμογές σε MATLAB και Mathematica, Εκδόσεις Τσότρας, 2015.

Hairer E., Norset S.P., Wanner G., Solving Ordinary Differential Equations, Springer 2000

Butcher J., Numerical Methods for Ordinary Differential Equations, Wiley, 2016