ΝΙΚΟΛΑΟΥ ΙΩ. ΔΑΡΑ

ΕΠΙΚΟΥΡΟΥ ΚΑΘΗΓΗΤΗ ΣΤΡΑΤΙΩΤΙΚΗΣ ΣΧΟΛΗΣ ΕΥΕΛΠΙΔΩΝ

ΣΤΟΙΧΕΙΑ ΑΡΙΘΜΗΤΙΚΗΣ ΑΝΑΛΥΣΗΣ

ΤΟΜΟΣ 2ος

ΔΙΑΚΡΙΤΗ ΑΡΙΘΜΗΤΙΚΗ ΑΝΑΛΥΣΗ

ΕΛΛΗΝΙΚΟ ΚΕΝΤΡΟ ΕΛΕΓΧΟΥ ΟΠΛΩΝ

www.armscontrol.info

-2007-

ΠΕΡΙΕΧΟΜΕΝΑ

Πρόλογος vii

Κεφάλαιο IV

Διακριτή Αριθμητική Ανάλυση **11**

IV.1. Παρεμβολή 11

IV.1.1. Τοποθέτηση του Προβλήματος 11

IV.1.2. Πολυώνυμο Παρεμβολής Lagrange 16

IV.1.3. Σχήμα Neville-Aitken 18

IV.1.4. Διηρημμένες Διαφορές 23

IV.1.5. Ο Τύπος του Newton 28

IV.1.6. Σφάλμα Παρεμβολής 41

IV.1.7. Βέλτιστη Επιλογή των Σημείων Παρεμβολής 45

IV.1.8. Σύγκλιση Ακολουθιών Πολυωνύμων Παρεμβολής 54

IV.1.9. Πολυώνυμο Παρεμβολής Hermite 70

IV.1.10. Γενική Διατύπωση του Προβλήματος της Παρεμβολής 77

IV.1.11. Παρεμβολή με Ρητή Συνάρτηση 82

IV.2. Αριθμητικός Τετραγωνισμός 101

IV.2.1. Τοποθέτηση του Προβλήματος 101

IV.2.2. Οι Τύποι Newton-Côtes 104

IV.2.3. Οι Σύνθετες Μέθοδοι 121

IV.2.4. Μέθοδος Τραπεζίων 122

IV.2.5. Μέθοδος Simpson 124

IV.2.6. Μέθοδος Romberg 128

IV.2.6.(i). Πολυώνυμα και Αριθμοί Bernoulli 129

IV.2.6.(ii). Τύπος Euler-MacLaurin 131

IV.2.6.(iii).Αλγόριθμος Romberg 135

IV.2.7. Οι Μέθοδοι Gauss 141

IV.2.8. Ύπαρξη και Ιδιότητες των Τύπων του Gauss 143
IV.2.9. Συστήματα Ορθογωνίων Πολυωνύμων 152
IV.2.10. Κατασκευή Ορθογωνίων Πολυωνύμων 163
IV.2.11. Μέθοδος Προσδιοριστέων Συντελεστών 169
IV.2.12. Αριθμητικός Υπολογισμός Πολλαπλών Ολοκληρωμάτων 176
IV.3. Αριθμητική Παραγώγιση 177
IV.3.1. Τοποθέτηση του Προβλήματος 177
IV.3.2. Μελέτη του Σφάλματος 179
IV.3.3. Υπολογισμός των Συντελεστών 180
IV.3.4. Η Περίπτωση των Ισαπεχόντων Σημείων 182
IV.3.4.(i). Αριθμητική Ευστάθεια 184
IV.3.4.(ii). Χρήσιμοι Τύποι 186
IV.3.4.(iii).Παραγώγιση και Τελεστές Διαφορών 187
IV.3.4.(iv).Βελτίωση Ακρίβειας 190
IV.3.5. Παραγώγιση Ανώτερης Τάξης 192
IV.3.6. Παραγώγιση Συναρτήσεων Πολλών Μεταβλητών 194
IV.4. Ολοκλήρωση Διαφορικών Συστημάτων 197
IV.4.1. Το Πρόβλημα Cauchy 197
IV.4.2. Οι Μέθοδοι Ξεχωριστών Βημάτων 206
IV.4.2.(i). Θεωρητικές Έννοιες 207
IV.4.2.(ii). Συνάφεια Μεθόδου 207
IV.4.2.(iii). Θεωρητική Ευστάθεια Μεθόδου 210
IV.4.2.(iv). Σύγκλιση Μεθόδου 212
IV.4.2.(v). Τάξη Μεθόδου 213
IV.4.2.(vi). Οι Μέθοδοι Euler και Runge-Kutta 215
IV.4.2.(vii). Πρακτικές Έννοιες 223
IV.4.2.(viii).Α-Ευστάθεια των Μεθόδων Runge-Kutta 223
IV.4.2.(ix). Έλεγχος Βήματος 230

IV.4.2.(x). Βελτίωση Ακρίβειας 233
IV.4.3. Οι Μέθοδοι Συναπτών Βημάτων 234
IV.4.3.(i). Εξισώσεις Διαφορών 237
IV.4.3.(ii). Θεωρητικές Έννοιες 247
IV.4.3.(iii). Συνάφεια Μεθόδου 247
IV.4.3.(iv). Θεωρητική Ευστάθεια Μεθόδου 250
IV.4.3.(v). Σύγκλιση Μεθόδου 252
IV.4.3.(vi). Τάξη Μεθόδου 255
IV.4.3.(vii). Οι Μέθοδοι Adams 257
IV.4.3.(viii).Οι Μέθοδοι Πρόβλεψης-Διόρθωσης 266
IV.4.3.(ix). Πρακτικές Έννοιες 268
IV.4.3.(x). Α-Ευστάθεια των Μεθόδων Συναπτών Βημάτων 269
IV.4.3.(xi). Α-Ευστάθεια των Μεθόδων Πρόβλεψης-Διόρθωσης 271
IV.4.3.(xii). Εκκίνηση των Μεθόδων Συναπτών Βημάτων 272
IV.4.4. Η Μέθοδος των Αναδρομικών Διαφορίσεων 274
IV.4.4.(i). Περιγραφή της Μεθόδου 282
IV.4.4.(ii). Μελέτη της Τάξης 289
IV.4.4.(iii). Ευστάθεια των Μεθόδων Τάξεων 1 και 2 292
IV.4.4.(iv). Ευστάθεια με Αμετάβλητη Τάξη 296
IV.4.4.(v). Ευστάθεια με Μεταβλητή Τάξη και Μεταβλητό Βήμα 308
IV.5. Στοιχεία Θεωρίας Προσεγγίσεων 320
IV.5.1. Τοποθέτηση του Προβλήματος 320
IV.5.2. Περί Βέλτιστης Προσέγγισης 321
IV.5.2.(i). Χαρακτηρισμός, Ύπαρξη και Μοναδικότητα 321
IV.5.2.(ii). Κατασκευή της Βέλτιστης Προσέγγισης 328
IV.5.3. Παραδείγματα Προσεγγίσεων 334
IV.5.3.(i). Προσέγγιση με την Έννοια των Ελαχίστων Τετραγώνων 334
IV.5.3.(i).(α).Συνεχής Προσέγγιση 334

IV.5.3.(i).(β).Διακριτή Προσέγγιση 337
IV.5.3.(i).(γ).Προσέγγιση με Συναρτήσεις Splines 340
IV.5.3.(ii). Βέλτιστη Ομοιόμορφη Προσέγγιση 341
Βιβλιογραφία **347**

ΠΡΟΛΟΓΟΣ

Βασικός σκοπός των Μαθηματικών υπήρξε κατ' αρχήν η δημιουργία γενικών Θεωριών, ανεξάρτητα του κατά πόσο οι σχετικές μέθοδοι ήταν εφαρμόσιμες σε συγκεκριμένα πρακτικά προβλήματα, έτσι ώστε να παρέχουν σαφή αριθμητικά αποτελέσματα.

Η έλλειψη υπολογιστικών μέσων αποτέλεσε σοβαρό λόγο για την καθυστέρηση της τροπής των Μαθηματικών προς μία συστηματική ανάπτυξη παράλληλων και διαφορετικών μεθόδων οι οποίες οδηγούν σε συγκεκριμένα αριθμητικά αποτελέσματα. Αυτές οι μέθοδοι απαιτούν πεπερασμένο, αλλά τεράστιο, πλήθος αριθμητικών πράξεων, ώστε η ευχερής, και μέσα σε λογικά χρονικά πλαίσια, εφαρμογή τους προϋποθέτει την χρήση ταχύτατων υπολογιστικών μηχανών. Και ακριβώς, η κατά τα τελευταία χρόνια αλματώδης ανάπτυξη των ηλεκτρονικών υπολογιστών είχε πολύ ευεργετική επίδραση στην ανακάλυψη, την επεξεργασία και την εφαρμογή τέτοιων μεθόδων. Ένα κεντρικό γνωστικό αντικείμενο της Αριθμητικής Ανάλυσης έγκειται στην αναζήτηση, την δημιουργία και την μελέτη αυτών των μεθόδων, των λεγόμενων αριθμητικών μεθόδων. Τα κριτήρια επιλογής της κατάλληλης αριθμητικής μεθόδου, που πρέπει να εφαρμοστεί για την επίλυση κάποιου προβλήματος, δεν είναι πάντοτε απλά, με συνέπεια αυτή η επιλογή να συνιστά ένα από τα δυσκολότερα στάδια της όλης διερεύνησης του εκάστοτε ζητήματος.

Ανεξάρτητα από τα όσα προηγήθηκαν, πρέπει να σημειωθεί ότι η εννοιολογική δυναμική και οι συνεχώς επεκτεινόμενες δυνατότητες της Αριθμητικής Ανάλυσης δίνουν διεξόδους και θεωρητικές προσεγγίσεις σε πολλά μαθηματικά ερωτήματα που είναι άγνωστο (ή και ανέφικτο) αν και πώς μπορούν να απαντηθούν με μεθόδους της κλασσικής Μαθηματικής Ανάλυσης. Για παράδειγμα, είναι γνωστό ότι δεν μπορούμε να υπολογίσουμε κάποια ολοκληρώματα πολλών συναρτήσεων ή λύσεις ορισμένων διαφορικών εξισώσεων ή ακόμη τον μοναδικό πραγματικό αριθμό ξ που επαληθεύει την εξίσωση $\xi = e^{-\xi}$. Σε μία τέτοια περίπτωση, αντικαθιστούμε την μαθηματική επίλυση του προβλήματος με την αριθμητική

επίλυσή του. Θα δίνουμε, για παράδειγμα, μία προσεγγιστική τιμή της λύσης της διαφορικής εξίσωσης σε έναν ορισμένο αριθμό σημείων του διαστήματος ολοκλήρωσης. Έτσι, η γενική θεματολογική κατεύθυνση της Αριθμητικής Ανάλυσης συνοψίζεται στον προσδιορισμό της σαν τον κλάδο των (Εφαρμοσμένων) Μαθηματικών που πραγματεύεται τις μεθόδους αριθμητικής επίλυσης των προβλημάτων της Μαθηματικής Ανάλυσης . Κατά τον *Henrici* ([28]), Αριθμητική Ανάλυση είναι η θεωρία των κατασκευαστικών μεθόδων της Μαθηματικής Ανάλυσης . Ο όρος κατασκευαστική μέθοδος σημαίνει ένα σύνολο από κανόνες, που λέγεται αλγόριθμος και που επιτρέπει την ανεύρεση της λύσης ενός μαθηματικού προβλήματος με μία αξιόπιστη ακρίβεια και μετά από πεπερασμένο αριθμό αριθμητικών πράξεων.

Το πρώτο *Κεφάλαιο* του παρόντος Βιβλίου περιλαμβάνει μία συνοπτική αναφορά στις βοηθητικές έννοιες και εύχρηστες εφαρμογές της Θεωρίας Πινάκων. Μετά από μία σύντομη παράθεση βασικών ιδιοτήτων, εξετάζονται τρία σπουδαία προβλήματα που άπτονται της Θεωρίας αυτής: η επίλυση τετραγωνικών συστημάτων γραμμικών εξισώσεων, η αναζήτηση των ιδιοτιμών και των ιδιοδιανυσμάτων ενός τετραγωνικού πίνακα και η διάσπαση τετραγωνικών πινάκων. Κατά την αντιμετώπιση αυτών των προβλημάτων στο *Κεφάλαιο* αυτό, δόθηκε έμφαση στην κατανόηση των αφηρημένων τεχνικών επίλυσής τους, γιατί μέσα από αυτές περιγράφεται άμεσα το βαθύτερο θεωρητικό υπόβαθρο και μοντέλο που ανταποκρίνεται στην μελέτη της φύσης τους.

Στο δεύτερο *Κεφάλαιο*, αναπτύσσονται οι κύριες αριθμητικές μέθοδοι υπολογισμού των ιδιοδιανυσμάτων ενός τετραγωνικού πίνακα. Όπως θα δούμε, υπάρχουν δύο μεγάλες κατηγορίες τέτοιων μεθόδων. Στην πρώτη κατηγορία, περιλαμβάνονται όλες εκείνες οι τεχνικές που επιτρέπουν τον πρακτικό υπολογισμό ενός ιδιοδιανύσματος κάθε φορά, καθώς και όλες εκείνες οι μέθοδοι που οδηγούν στον διαδοχικό υπολογισμό όλων των ιδιοδιανυσμάτων. Οι μέθοδοι της δεύτερης κατηγορίας διακρίνονται σε δύο κύρια είδη: τις μεθόδους προσδιορισμού του χαρακτηριστικού πολυωνύμου και τις μεθόδους διάσπασης. Το

πρώτο είδος είναι το λιγότερο ενδιαφέρον (εκτός από ορισμένες ειδικές περιπτώσεις), επειδή πρέπει στην συνέχεια να υπολογισθούν οι ρίζες του χαρακτηριστικού πολυωνύμου. Το δεύτερο είδος έχει κυρίως επαναληπτικό χαρακτήρα ο οποίος στηρίζεται στην χρησιμοποίηση όμοιων μετασχηματισμών. Πέραν των μεθόδων που έχουν συμπεριληφθεί σε αυτό το δεύτερο *Κεφάλαιο*, υπάρχει και ένας μικρός αριθμός άλλων αριθμητικών μεθόδων που αναπτύσσονται αρκετά εξαντλητικά σε εργασίες και πραγματείες της Βιβλιογραφίας.

Στο τρίτο *Κεφάλαιο* παρουσιάζονται οι αριθμητικές μέθοδοι επίλυσης αλγεβρικών εξισώσεων και συστημάτων γραμμικών εξισώσεων. Μόνο λίγες αλγεβρικές εξισώσεις μπορούν να λυθούν με την χρήση κλασσικών Μαθηματικών (:οι πολυωνυμικές εξισώσεις βαθμού ≤ 4). Με την χρήση αριθμητικών μεθόδων όμως έχει κατορθωθεί ο προσεγγιστικός εντοπισμός των ριζών μίας εξίσωσης της μορφής $f(\xi)=0$, όπου $f(x)$ είναι μία συνάρτηση πραγματικής ή μιγαδικής μεταβλητής x. Στο πρώτο μέρος του *Κεφαλαίου* αυτού, θα ασχοληθούμε κυρίως με τον αριθμητικό υπολογισμό μίας απλής πραγματικής ρίζας της εξίσωσης $f(\xi)=0$. Στα υπόλοιπα μέρη του *Κεφαλαίου*, θα μελετήσουμε το πρόβλημα επίλυσης ενός γραμμικού συστήματος εξισώσεων *A.x=b* όπου *A* είναι ένας τετραγωνικός αντιστρέψιμος πίνακας με *n* γραμμές και *n* στήλες και όπου *b* είναι ένα διάνυσμα του $\mathbb{C}^n$.Θεωρητικά ένα τέτοιο σύστημα μπορεί να επιλυθεί υπολογίζοντας τον αντίστροφο πίνακα A^{-1} του A. Όμως, το γεγονός ότι ο υπολογισμός του A^{-1} προϋποθέτει τον υπολογισμό της ορίζουσας του A, υπολογισμός που, με την σειρά του, απαιτεί την εκτέλεση $(n\times n)!$ πολλαπλασιασμών, φανερώνει ότι η μαθηματική επίλυση ενός μεγάλου συστήματος γραμμικών εξισώσεων είναι πρακτικά αδύνατη. Οι αριθμητικές μέθοδοι επίλυσης συστημάτων γραμμικών εξισώσεων χωρίζονται σε τρεις μεγάλες κατηγορίες που είναι ριζικά διαφορετικές μεταξύ τους : τις άμεσες μεθόδους, που δίνουν θεωρητικά το διάνυσμα λύση x μετά από πεπερασμένο αριθμό στοιχειωδών αριθμητικών πράξεων, τις επαναληπτικές μεθόδους, που δίνουν την λύση x σαν όριο μιας ακολουθίας διανυσμάτων και τις μεθόδους προβολής .

Στο τέταρτο θεωρητικό *Κεφάλαιο* παρουσιάζεται μία εισαγωγή στην Διακριτή Αριθμητική Ανάλυση, δηλαδή στον κλάδο εκείνο της Αριθμητικής Ανάλυσης όπου η επίλυση ενός μαθηματικού προβλήματος επιτυγχάνεται με την αντικατάσταση ενός συνεχούς φαινομένου με ένα προσεγγιστικό διακριτό φαινόμενο. Οι κεντρικές γνωστικές οντότητες του *Κεφαλαίου* αναφέρονται στην μελέτη του προβλήματος παρεμβολής, της αριθμητικής ολοκλήρωσης και παραγώγισης και τέλος της αριθμητικής επίλυσης μίας συνήθους διαφορικής εξίσωσης ή ενός συστήματος διαφορικών εξισώσεων.

Στο πέμπτο *Κεφάλαιο* συγκεντρώσαμε ορισμένα χαρακτηριστικά παραδείγματα και ασκήσεις, που αναφέρονται στα θεωρητικά εδάφια που προηγήθηκαν. Η προτίμηση που δείξαμε σε στοιχειώδη παραδείγματα οφείλεται, όχι μόνο στην επιθυμία μας για άμεση κατανόηση του μηχανισμού των διάφορων μεθόδων από μέρους του αναγνώστη, αλλά και στην απρόσκοπτη ευχέρεια συγκριτικών αντιπαραθέσεων των δυνατοτήτων τους. Η συλλογή λυμένων και άλυτων ασκήσεων πιστεύουμε ότι καλύπτει όλο το φάσμα των κυριότερων αριθμητικών μεθόδων.

Παρά την προσπάθεια που έγινε για μία αναλυτική και πλήρη ανάπτυξη των πρώτων *Κεφαλαίων* ενός μαθήματος Αριθμητικής Ανάλυσης, υπάρχουν μερικά ζητήματα και κάποιες ειδικές μεθοδολογίες, που δεν ήταν δυνατόν να συμπεριλάβουμε στα πέντε αυτά γενικά *Κεφάλαια*. Προς τούτο, ο κάθε ενδιαφερόμενος αναγνώστης μπορεί να προστρέξει και να συμβουλευτεί τα συγγράμματα της προτεινόμενης Βιβλιογραφίας (στο τέλος του παρόντος βιβλίου), που όμως, όπως είναι φυσικό, αντιπροσωπεύει μόνο ένα μέρος του συνεχώς διογκούμενου καταλόγου σχετικών και συναφών πραγματειών.

Νικόλαος Ιω. Δάρας

Κεφάλαιο 4

Διακριτή Αριθμητική Ανάλυση

IV.1. ΠΑΡΕΜΒΟΛΗ

IV.1.1. ΤΟΠΟΘΕΤΗΣΗ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ

Έστω *f* μία απεικόνιση ορισμένη μέσα σε ένα υποσύνολο U του K και τέτοια ώστε να παίρνει τιμές μέσα στο K $(K = R, C)$.

Υποθέτουμε ότι γνωρίζουμε τις τιμές:

$$f(z_i)$$

της f σε $(n+1)$ σημεία $z_0, z_1, \ldots, z_n$ του U και ότι υπάρχει τουλάχιστον ένα σημείο z_k ανάμεσα στα $z_0, z_1, \ldots, z_n$, τέτοιο ώστε:

$$f(z_k) \neq 0 .$$

Επιθυμούμε να βρούμε ένα πολυώνυμο:

$$p : K \to K : z \mapsto p(z),$$

που να χαρακτηρίζεται από την ιδιότητα να παίρνει τις ίδιες τιμές με την απεικόνιση f στα σημεία $z_0, z_1, \ldots, z_n$:

$$p(z_i) = f(z_i) \quad \text{για κάθε} \quad i = 0,1,\ldots,n .$$

Ένα τέτοιο πολυώνυμο, εφόσον υπάρχει, καλείται ***πολυώνυμο παρεμβολής της*** f. Λέμε επίσης ότι ***το πολυώνυμο*** $p(z)$ ***παρεμβάλλει την*** $f(z)$ ***στα σημεία*** $z_0, z_1, \ldots, z_n$.

Προς τούτο, συμβολίζουμε με:

$$P_n(K)$$

τον διανυσματικό χώρο όλων των πολυωνύμων, με συντελεστές και τιμές στο K, των οποίων ο βαθμός είναι μικρότερος ή ίσος από n.

Αρχίζουμε με το ακόλουθο:

Θεώρημα IV.1.1.1. ***Μία αναγκαία και ικανή συνθήκη προκειμένου να υπάρχει ένα και μόνο ένα πολυώνυμο*** $p(z) \in P_n(K)$ ***που παρεμβάλει την*** $f(z)$ ***στα σημεία*** $z_0, z_1, \ldots, z_n$ ***είναι τα σημεία παρεμβολής*** z_i ***να είναι ανά δυο διάφορα μεταξύ τους.***

Απόδειξη. Για να υπάρχει ένα μοναδικό πολυώνυμο $p(z) \in P_n(K)$ που παρεμβάλλει την συνάρτηση $f(z)$ στα σημεία $z_0, z_1, \ldots, z_n$ πρέπει και αρκεί το σύστημα των $(n+1)$ γραμμικών εξισώσεων με τους $(n+1)$ άγνωστους $a_0, a_1, \ldots, a_n$:

$$\sum_{\nu=0}^{n} a_\nu z_i^\nu = f(z_i) \qquad i = 0,1,\ldots,n+1$$

να έχει μία και μοναδική λύση.

Η ορίζουσα αυτού του συστήματος:

$$det\begin{pmatrix} 1 & z_0 & \dots & z_0^n \\ 1 & z_1 & \dots & z_1^n \\ \dots & \dots & \dots & \dots \\ 1 & z_n & \dots & z_n^n \end{pmatrix}$$

ονομάζεται ***ορίζουσα Vandermonde*** και είναι ίση με:

$$\prod_{i>j}(z_i - z_j).$$

Συνεπώς, το παραπάνω σύστημα έχει μία και μοναδική λύση εάν και μόνον εάν:

$$\prod_{i>j}(z_i - z_j) \neq 0,$$

δηλαδή εάν και μόνον εάν όλα τα σημεία $z_0, z_1, \dots, z_n$ είναι ανά δύο διάφορα μεταξύ τους. ■

Παρατήρηση IV.1.1.2. Η μοναδικότητα του πολυώνυμου παρεμβολής $p(z) \in \boldsymbol{P}_n(\boldsymbol{K})$ της συνάρτησης f στα σημεία $z_0, z_1, \dots, z_n \in U$ μπορεί να αποδειχθεί και κατά άλλον τρόπο: εάν $q(z) \in \boldsymbol{P}_n(\boldsymbol{K})$ ήταν ένα δεύτερο τέτοιο πολυώνυμο, τότε η διαφορά $r(z) = p(z) - q(z)$ θα ήταν επίσης ένα στοιχείο του $\boldsymbol{P}_n(\boldsymbol{K})$ που θα είχε την ιδιότητα $r(z_i) = 0 \quad \forall i = 0,1,\dots,n,$ γιατί:

$$r(z_i) = p(z_i) - q(z_i) = f(z_i) - f(z_i) = 0,$$

δηλαδή το $r(z)$ θα ήταν ένα πολυώνυμο βαθμού το πολύ ίσου με n που θα είχε $(n+1)$ ρίζες, άρα το $r(z)$ θα ήταν το μηδενικό πολυώνυμο:

$$r(z) = 0 \Leftrightarrow p(z) = q(z).■$$

Ορμώμενοι από το *Θεώρημα ΙΙΙ.1.1.1*, στην συνέχεια, θα συμβολίζουμε με:

$$p_n(z)$$

το μοναδικό πολυώνυμο παρεμβολής της $f(z)$ στα ανά δύο διάφορα σημεία $z_0, z_1, \ldots, z_n$ του *U* αντί για $p(z)$, ακριβώς για να τονίζουμε ότι το πολυώνυμο παρεμβολής του *Θεωρήματος ΙΙΙ.1.1.1* έχει βαθμό το πολύ ίσο με n.

Θα κατασκευάσουμε τώρα το $p_n(z)$.

Είναι σαφές πως:

$$-p_n(z) + a_0 + a_1\ z + \ldots + a_n\ z^n = 0$$

$$-f(z_0) + a_0 + a_1\ z_0 + \ldots + a_n\ z_0^n = 0$$

$$-f(z_1) + a_0 + a_1\ z_1 + \ldots + a_n\ z_1^n = 0$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$-f(z_n) + a_0 + a_1\ z_n + \ldots + a_n\ z_n^n = 0,$$

από όπου έπεται η σχέση:

$$\det\begin{pmatrix} p_n(z) & 1 & z & \ldots & z^n \\ -f(z_0) & 1 & z_0 & \ldots & z_0^n \\ -f(z_1) & 1 & z_1 & \ldots & z_1^n \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ -f(z_n) & 1 & z_n & \ldots & z_n^n \end{pmatrix} = 0,$$

που μπορεί επίσης να γραφεί υπό την μορφή:

$$\det\begin{pmatrix} p_n(z) & 1 & z & \dots & z^n \\ 0 & 1 & z_0 & \dots & z_0^n \\ 0 & 1 & z_1 & \dots & z_1^n \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 1 & z_n & \dots & z_n^n \end{pmatrix} = \det\begin{pmatrix} 0 & 1 & z & \dots & z^n \\ -f(z_0) & 1 & z_0 & \dots & z_0^n \\ -f(z_1) & 1 & z_1 & \dots & z_1^n \\ \dots & \dots & \dots & \dots & \dots \\ -f(z_n) & 1 & z_n & \dots & z_n^n \end{pmatrix}.$$

Από την τελευταία ισότητα συνάγεται αμέσως ότι:

$$p_n(z) = \frac{\det\begin{pmatrix} 0 & 1 & z & \dots & z^n \\ -f(z_0) & 1 & z_0 & \dots & z_0^n \\ -f(z_1) & 1 & z_1 & \dots & z_1^n \\ \dots & \dots & \dots & \dots & \dots \\ -f(z_n) & 1 & z_n & \dots & z_n^n \end{pmatrix}}{\det\begin{pmatrix} 1 & z_0 & \dots & z_0^n \\ 1 & z_1 & \dots & z_1^n \\ \dots & \dots & \dots & \dots \\ 1 & z_n & \dots & z_n^n \end{pmatrix}}.$$

Δυστυχώς, ο αριθμητικός υπολογισμός του $p_n(z)$ με την βοήθεια αυτού του *Τύπου* είναι πρακτικά ανέφικτος, γιατί η εύρεση μίας τέτοιας ορίζουσας, με την μέθοδο της μετάθεσης γραμμών ή στηλών θα απαιτούσε την πραγματοποίηση τουλάχιστον $n \times (n!)$ πολλαπλασιασμών.

Στην πράξη, καταφεύγουμε στην χρήση του *Τύπου* που δίνει το ***πολυώνυμο παρεμβολής Lagrange***.

IV.1.2.ΠΟΛΥΩΝΥΜΟ ΠΑΡΕΜΒΟΛΗΣ LAGRANGE

Το πολυώνυμο παρεμβολής $p_n(z)$ δίνεται και από τον ***Τύπο του Lagrange***:

$$p_n(z) = \sum_{i=0}^{n} L_i^{(n)}(z)\, f(z_i),$$

με:

$$L_i^{(n)}(z) = \prod_{\substack{j=0 \\ j\neq i}}^{n} \frac{z - z_j}{z_i - z_j}$$

(όπου έχουμε διατηρήσει την παραδοχή ότι $z_i \neq z_j \quad \forall j \neq i$).

Προκειμένου να επαληθεύσουμε ότι ο τύπος του Lagrange δίνει πράγματι το πολυώνυμο, βαθμού το πολύ n, που παρεμβάλλει την συνάρτηση $f(z)$ στα σημεία $z_0, z_1, \ldots, z_n \in U$, παρατηρούμε κατ' αρχήν ότι η παράσταση

$$\sum_{i=0}^{n} L_i^{(n)}(z)\, f(z_i)$$

είναι ένα πολυώνυμο, μέσα από τον διανυσματικό χώρο $\boldsymbol{P}_n(\boldsymbol{K})$, γιατί σχηματίζεται από τον γραμμικό συνδυασμό των πολυωνύμων $L_i^{(n)}(z) \in \boldsymbol{P}_n(\boldsymbol{K})$.

Ας δείξουμε τώρα ότι:

$$p_n(z_i) = f(z_i) \qquad \forall i = 0,1,\ldots,n.$$

Είναι σαφές πως

$$L_i^{(n)}(z_i) = 1.$$

Ακόμη, είναι

$$L_i^{(n)}(z_k) = 0 \ \ \forall k \neq i,$$

γιατί, τότε, υπάρχει ένας δείκτης *j* ίσος με *k* μέσα στο γινόμενο και ο αριθμητής του $L_i^{(n)}(z_k)$ ισούται με μηδέν.

Επομένως, ισχύει:

$$L_i^{(n)}(z_k) = \delta_{i,k} = \begin{Bmatrix} 1, \text{ όταν } \ i = k \\ 0, \text{ όταν } \ i \neq k \end{Bmatrix}$$

και έπεται, αμέσως, ότι $p_n(z_i) = f(z_i) \quad \forall i = 0,1,\ldots,n$.

Επειδή, σύμφωνα με το *Θεώρημα IV.1.1.1*, το πολυώνυμο, βαθμού το πολύ n, που παρεμβάλλει την $f(z)$ στα, ανά δύο διάφορα σημεία $z_0, z_1, \ldots, z_n$ του U είναι μοναδικό, συμπεραίνουμε ότι πράγματι ο τύπος του Lagrange δίνει το πολυώνυμο που περιγράφεται στο *Θεώρημα IV.1.1.1*.

Ας παρουσιάσουμε τώρα μία άλλη έκφραση του πολυωνύμου παρεμβολής Lagrange.

Θέτουμε:

$$V_n(z) := \prod_{j=0}^{n} (z - z_j),$$

και υπολογίζουμε την τιμή $V_n'(z_i)$ της παραγώγου της συνάρτησης $V_n(z)$ στο σημείο z_i.

Έχουμε:

$$\frac{V_n(z)}{z - z_i} = \prod_{\substack{j=0 \\ j \neq i}}^{n} (z - z_j),$$

κι επειδή $V_n(z_i) = 0$ είναι:

$$\frac{V_n(z)-V_n(z_i)}{z-z_i}=\frac{V_n(z)}{z-z_i}.$$

Αφήνοντας το z να τείνει προς το z_i, διαπιστώνουμε ότι:

$$V_n'(z_i)=\prod_{\substack{j=0\\ j\neq i}}^{n}\left(z-z_j\right),$$

και επομένως ισχύει:

$$L_i^{(n)}(z)=\frac{V_n(z)}{(z-z_i)\,V_n'(z_i)}.$$

Έτσι,

το πολυώνυμο παρεμβολής Lagrange γράφεται και υπό την μορφή:

$$p_n(z)=V_n(z)\sum_{i=0}^{n}\frac{f(z_i)}{(z-z_i)\,V_n'(z_i)}.$$

IV.1.3. ΣΧΗΜΑ NEVILLE-AITKEN

Αντί να κατασκευάσουμε το πολυώνυμο παρεμβολής $p_n(z)$ της $f(z)$ χρησιμοποιώντας τον τύπο του Lagrange, μπορούμε να το κατασκευάσουμε με μαθηματική επαγωγή χρησιμοποιώντας το ***σχήμα Neville-Aitken***.

Προς τούτο, ας συμβολίσουμε με:

$$T_k^{(i)}(z)$$

το πολυώνυμο βαθμού το πολύ k που παρεμβάλλει την συνάρτηση $f(z)$ στα σημεία

$$z_i, z_{i+1}, \ldots, z_{i+k}.$$

Στο σχήμα Neville-Aitken, το πολυώνυμο $T_{k+1}^{(i)}$ δημιουργείται με την βοήθεια των πολυώνυμων $T_k^{(i)}$ και $T_k^{(i+1)}$, σύμφωνα με τον ακόλουθο τρόπο.

Θέτουμε:

$$T_0^{(i)}(z) := f(z_i) \quad \forall i = 0,1,\ldots,n\,.$$

Είναι φανερό ότι το $T_0^{(i)}(z)$ είναι το πολυώνυμο βαθμού ίσου με 0 που παρεμβάλλει την $f(z)$ στο μοναδικό σημείο z_i.

Στην συνέχεια, χρησιμοποιούμε τον ακόλουθο επαγωγικό *Τύπο*:

$$T_{k+1}^{(i)}(z) = \frac{(z_{i+k+1} - z)\, T_k^{(i)}(z) - (z_i - z)\, T_k^{(i+1)}(z)}{z_{i+k+1} - z_i}$$

$$\vdots \qquad \text{για } k = 0,1,\ldots,n-1 \text{ και } i = 0,1,\ldots,n-k-1$$

και

$$T_{k+1}^{(i)}(z) = \frac{(z_{i+k+1} - z)\, T_k^{(i)}(z) - (z_i - z)\, T_k^{(i+1)}(z)}{z_{i+k+1} - z_i}$$

$$\text{για } k = 0,1,\ldots,n-1 \text{ και } i = 0,1,\ldots,n-k-1$$

και υποθέτουμε ότι τα

$$T_k^{(i)}(z) \text{ και } T_k^{(i+1)}(z)$$

είναι πολυώνυμα παρεμβολής της $f(z)$ στα σημεία:

$$z_i, z_{i+1}, \ldots, z_{i+k} \quad \text{και} \quad z_{i+1}, z_{i+2}, \ldots, z_{i+k+1},$$

αντίστοιχα.

Θα δείξουμε ότι το πολυώνυμο:

$$T_{k+1}^{(i)}(z)$$

παρεμβάλλει την $f(z)$ στα σημεία:

$$z_i, z_{i+1}, \ldots, z_{i+k+1}.$$

Προς τούτο, κατ' αρχήν, παρατηρούμε ότι, επειδή τα

$$T_k^{(i)}(z) \text{ και } T_k^{(i+1)}(z)$$

είναι δύο πολυώνυμα βαθμών το πολύ ίσων με k, το πολυώνυμο

$$T_{k+1}^{(i)}(z) \ ,$$

που ορίζεται από τον παραπάνω επαγωγικό τύπο, έχει βαθμό το πολύ ίσο με $(k+1)$.

Στην συνέχεια, προβαίνουμε στις εξής τρεις διαπιστώσεις που συνοψίζουν τον παρεμβλητικό χαρακτήρα του πολυωνύμου

$$T_{k+1}^{(i)}(z) \quad :$$

1) Για $z = z_i$, έχουμε:

$$T_{k+1}^{(i)}(z_i) = \frac{(z_{i+k+1} - z_i)\, T_k^{(i)}(z_i) - (z_i - z_i)\, T_k^{(i+1)}(z_i)}{z_{i+k+1} - z_i} \Rightarrow$$

$$T_{k+1}^{(i)}(z_i) = T_k^{(i)}(z_i),$$

και επειδή $T_k^{(i)}(z_i) = f(z_i)$ (αφού το $T_k^{(i)}(z)$ είναι το πολυώνυμο παρεμβολής της $f(z)$ στα σημεία $z_i, z_{i+1}, \ldots, z_{i+k}$), ισχύει

$$T_{k+1}^{(i)}(z_i) = f(z_i).$$

2) Για $z = z_p \quad (\forall p = i+1, \ldots, i+k)$, έχουμε:

$$T_{k+1}^{(i)}(z_p) = \frac{(z_{i+k-1} + z_p)\, T_k^{(i)}(z_p) - (z_i - z_p)\, T_k^{(i+1)}(z_p)}{(z_{i+k+1} - z_p) - (z_i - z_p)},$$

κι επειδή $T_k^{(i)}(z_p) = T_k^{(i+1)}(z_p) = f(z_p)$, ισχύει

$$T_{k+1}^{(i)}(z_p) = f(z_p).$$

3) Για $z = z_{i+k+1}$ έχουμε:

$$T_{k+1}^{(i)}(z_{i+k+1}) = \frac{(z_{i+k+1} - z_{i+k+1})\, T_k^{(i)}(z_{i+k+1}) - (z_i - z_{i+k+1})\, T_k^{(i+1)}(z_{i+k+1})}{z_{i+k+1} - z_i},$$

κι επειδή $T_{k+1}^{(i)}(z_{i+k+1}) = f(z_{i+k+1})$, ισχύει

$$T_{k+1}^{(i)}(z_{i+k+1}) = f(z_{i+k+1}).$$

Συνοψίζοντας, η παρεμβλητική συμπεριφορά του πολυωνύμου $T_{k+1}^{(i)}(z)$ χαρακτηρίζεται από την ακόλουθη ιδιότητα:

$$T_{k+1}^{(i)}(z_\nu) = f(z_\nu),\ \forall \nu = i+1, \ldots, i+k+1.$$

Άρα, λόγω του *Θεωρήματος IV.1.1.1*, το $T_{k+1}^{(i)}(z)$ είναι πράγματι το πολυώνυμο, που έχει βαθμό το πολύ ίσο με *k*+1 και που παρεμβάλλει την συνάρτηση $f(z)$ στα, ανά δύο διάφορα μεταξύ τους, σημεία $z_i, z_{i+1}, \ldots, z_{i+k+1}$.

Από αυτό έπεται ότι η διαδοχική συνέχιση της διαδικασίας των υπολογισμών των $T_{k+1}^{(i)}(z)$ θα οδηγήσει σε ένα τελευταίο πολυώνυμο, βαθμού το πολύ n:

$$T_n^0(z),$$

το οποίο θα είναι το πολυώνυμο παρεμβολής $p_n(z)$ της $f(z)$ στα ανά δύο διάφορα μεταξύ τους, σημεία $z_0, z_1, \ldots, z_n$ του U.

Η πρακτική διάθεση της παραπάνω επαγωγικής διαδικασίας υπολογισμών του σχήματος Neville-Aitken αποδίδεται από την παρακάτω διαρθρωτική αναπαράσταση της σειράς διαδοχικών κατασκευών:

$$\begin{tikzcd}
T_0^{(0)}(z)=f(z_0) \arrow[dr] & & & \\
& T_1^{(0)}(z) \arrow[dr] & & \\
T_0^{(1)}(z)=f(z_1) \arrow[ur] \arrow[dr] & & T_2^{(0)}(z) \arrow[dr] & \\
& T_1^{(1)}(z) \arrow[ur] \arrow[dd, dotted, no head] & & T_n^{(0)}(z)=p_n(z) \\
T_0^{(2)}(z)=f(z_2) \arrow[ur] \arrow[dd, dotted, no head] & & T_2^{(n-2)}(z) \arrow[ur] \arrow[dd, dotted, no head] & \\
& & & \\
& T_1^{(n-1)}(z) \arrow[ur] & & \\
T_0^{(n)}(z)=f(z_n) \arrow[ur] & & &
\end{tikzcd}$$

Παρατήρηση IV.1.3.1. Για μία δεδομένη τιμή του $z \in K$, ο υπολογισμός της αντίστοιχης τιμής του πολυωνύμου παρεμβολής $p_n(z)$ απαιτεί:

σύμφωνα με τη μέθοδο Lagrange:

$2\,n\,(n+1)$ πολλαπλασιασμούς ή διαιρέσεις,

$\dfrac{n^2+5\,n+2}{2}$ προσθέσεις ή αφαιρέσεις,

σύμφωνα με το σχήμα Neville-Aitken:

$\dfrac{3\,n\,(n+1)}{2}$ πολλαπλασιασμούς ή διαιρέσεις,

$2\,n\,(n+1)$ προσθέσεις ή αφαιρέσεις. ■

Στην συνέχεια, θα μελετήσουμε άλλους τρόπους υπολογισμού του $p_n(z)$.

IV.1.4. ΔΙΗΡΗΜΜΕΝΕΣ ΔΙΑΦΟΡΕΣ

Στην παρούσα *Παράγραφο*, θα παρουσιάσουμε μία διαφορετική μέθοδο κατασκευής του πολυωνύμου παρεμβολής $p_n(z)$, η οποία στηρίζεται στην *έννοια των διηρημμένων διαφορών*.

Έστω f μία συνάρτηση της μεταβλητής $z \in U \subset \boldsymbol{K}$, της οποίας γνωρίζουμε τις ακριβείς αριθμητικές τιμές στα σημεία $z_0, z_1, \ldots, z_n \in U$, που είναι ανά δύο διάφορα μεταξύ τους.

Ορισμός IV.1.4.1. ***Ονομάζουμε διηρημμένες διαφορές τάξης*** $0,1,\ldots,n$ ***της συνάρτησης*** $f(z)$ ***στα σημεία*** $z_0, z_1, \ldots, z_n$, ***τις ακόλουθες εκφράσεις:***

τάξη $0: [z_i] = f(z_i)$,

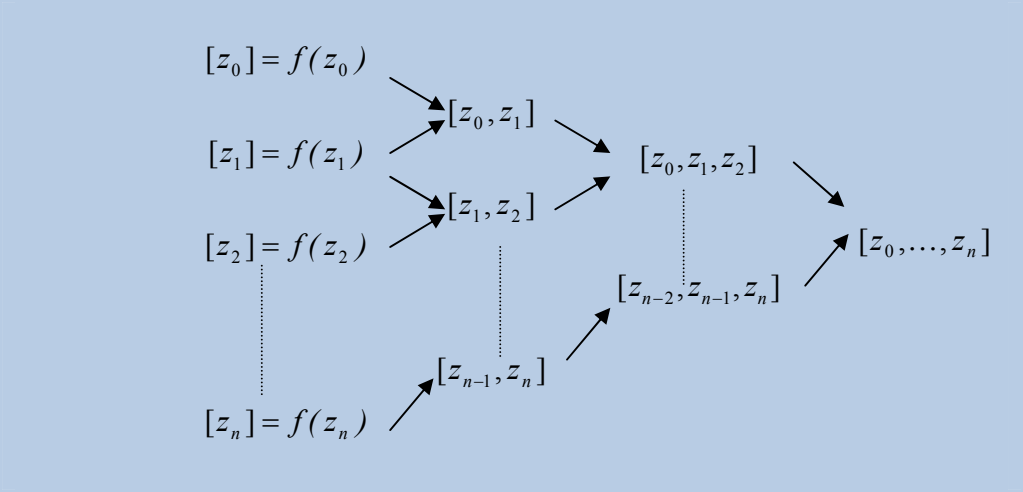
τάξη $1: [z_i, z_j] = \dfrac{[z_i] - [z_j]}{z_i - z_j} \qquad (i \neq j)$,

τάξη $2: [z_i, z_j, z_k] = \dfrac{[z_i, z_j] - [z_j - z_k]}{z_i - z_k} \qquad (i \neq j \neq k \neq i)$,

.....

τάξη $n: [z_0, \ldots z_n] = \dfrac{[z_0, \ldots, z_{n-1}] - [z_1, \ldots, z_n]}{z_0 - z_n}$.■

Ο πρακτικός υπολογισμός των διηρημμένων διαφορών μίας συνάρτησης πραγματοποιείται σύμφωνα με το εξής σχήμα:



Οι διηρημμένες διαφορές της $f(z)$ διέπονται από τις πιο κάτω ιδιότητες:

Ιδιότητα IV.1.4.2. ***Εάν*** $V(z)=(z-z_{\nu_0})\,(z-z_{\nu_1})\ldots(z-z_{\nu_k})$ ***, τότε:***

$$[z_{\nu_0},z_{\nu_1},\ldots,z_{\nu_k}]=\sum_{i=\nu_0}^{\nu_k}\frac{f(z_i)}{V'(z_i)}.$$

Απόδειξη. Η *Απόδειξη* της *Ιδιότητας* αυτής γίνεται με μαθηματική επαγωγή.

Για $k=0$, είναι $V(z)=z-z_{\nu_0}$ και $V'(z)=1$, οπότε

$$[z_{\nu}\ ]=f(z_{\nu_0}).$$

Έχοντας τώρα υποθέσει πως, για $V(z)=(z-z_{\nu_0})\ldots(z-z_{\nu_k})$ έχει αποδειχθεί ότι:

$$[z_{\nu_0},z_{\nu_1},\ldots,z_{\nu_k}]=\sum_{i=\nu_0}^{\nu_k}\frac{f(z_i)}{V'(z_i)},$$

και πως για $W(z)=(z-z_{\nu_1})\ldots(z-z_{\nu_{k+1}})$, έχει αποδειχθεί ότι:

$$[z_{\nu_1},z_{\nu_2},\ldots,z_{\nu_{k+1}}]=\sum_{i=\nu_1}^{\nu_{k+1}}\frac{f(z_i)}{W'(z_i)},$$

θα δείξουμε ότι η *Ιδιότητα* παραμένει αληθινή και για την έκφραση:

$$[z_{\nu_1},z_{\nu_2},\ldots,z_{\nu_{k+1}}].$$

Χρησιμοποιώντας τον ορισμό των διηρημμένων διαφορών, έχουμε:

$$[z_{\nu_0},\ldots,z_{\nu_{k+1}}]=\frac{1}{z_{\nu_0}-z_{\nu_{k+1}}}\left\{\frac{f(z_{\nu_0})}{V'(z_{\nu_0})}+\sum_{i=\nu_1}^{\nu_k}f(z_i)\,(\frac{1}{V'(z_i)}-\frac{1}{W'(z_i)})-\frac{f(z_{\nu_{k+1}})}{W'(z_{\nu_{k+1}})}\right\},$$ οπότε

αποδεικνύεται εύκολα πως, για $S(z)=(z-z_{\nu_0})\ldots(z-z_{\nu_{k+1}})$, ισχύει ότι:

$$[z_{\nu_0},z_{\nu_1},\ldots,z_{\nu_{k+1}}]=\sum_{i=\nu_0}^{\nu_{k+1}}\frac{f(z_i)}{S'(z_i)},$$

ισότητα που ολοκληρώνει την διαδικασία της μαθηματικής επαγωγής, και που, κατ' επέκταση, περατώνει την *Απόδειξη* της *Ιδιότητας*. ■

Μία ενδιαφέρουσα συνέπεια αυτής της *Ιδιότητας* είναι η συμμετρία που παρουσιάζουν οι διηρημμένες διαφορές μίας συνάρτησης, υπό την έννοια ότι ***η διάταξη των δομικών συμβόλων z_i μίας διηρημμένης διαφοράς μπορεί να αλλάξει αφήνοντας ανεπηρέαστη την τιμή της διηρημμένης διαφοράς***.

Ακόμη, έχουμε τις παρακάτω *Ιδιότητες*:

Ιδιότητα IV.1.4.3.

$$f(z)=f(z_0)+(z-z_0)[z_0,z_1]+(z-z_0)(z-z_1)[z_0,z_1,z_2]+\cdots$$
$$\cdots+(z-z_0)\ldots(z-z_{n-1})[z_0,z_1,\ldots,z_n]+(z-z_0)\ldots(z-z_n)[z,z_0,\ldots,z_n].$$

Απόδειξη. Η *Απόδειξη* της *Ιδιότητας* αυτής γίνεται επίσης με μαθηματική επαγωγή. Χωρίς να υπεισέλθουμε σε λεπτομέρειες αποδεικνύουμε μόνον τα πρώτα στάδια της επαγωγής.

Επειδή $[z] = f(z)$, ισχύει ότι:

$$[z, z_0] = \frac{f(z) - f(z_0)}{z - z_0},$$

απ' όπου έπεται η σχέση:

$$f(z) = f(z_0) + (z - z_0)\,[z, z_0].$$

Ακόμη, επειδή:

$$[z, z_0, z_1] = \frac{[z, z_0] - [z_0, z_1]}{z - z_1},$$

ισχύει ότι:

$$[z, z_0] = [z_0, z_1] + (z - z_1)\,[z, z_0, z_1]$$

και άρα:

$$f(z) = f(z_0) + (z - z_0)\,\{[z_0, z_1] + (z - z_1)\,[z, z_0, z_1]\} \Rightarrow$$

$$f(z) = f(z_0) + (z - z_0)\,[z_0, z_1] + (z - z_0)\,(z - z_1)\,[z, z_0, z_1]. \blacksquare$$

Ιδιότητα IV.1.4.4. ***Το πολυώνυμο παρεμβολής*** $p_n(z)$ ***της*** $f(z)$ ***στα αμοιβαία διάφορα μεταξύ τους σημεία*** $z_0, z_1, \ldots, z_n$ ***.δίνεται από τον Τύπο:***

$$p_n(z) = f(z_0) + (z - z_0)\,[z_0, z_1] + (z - z_0)\,(z - z_1)\,[z_0, z_1, z_2] + \cdots$$
$$\cdots + (z - z_0)\ldots(z - z_{n-1})\,[z_0, z_1, .., z_n].$$

Απόδειξη. Σύμφωνα με την *Ιδιότητα IV.1.4.2*, ισχύει:

$$[z, z_0, z_1, \ldots, z_n] = \frac{f(z)}{V_{n+1}(z)} + \sum_{i=0}^{n} \frac{f(z_i)}{(z_i - z)\,V_n'(z)},$$

όπου $V_{n+1}(z) = (z - z_0)\,(z - z_1)\ldots(z - z_n)$.

Εξάλλου, όπως είδαμε στην *Παράγραφο IV.1.2*, το πολυώνυμο παρεμβολής $p_n(z)$ δίνεται από τον *Τύπο*:

$$p_n(z) = V_{n+1}(z) \sum_{i=0}^{n} \frac{f(z_i)}{(z - z_i)\,V_n'(z)}.$$

Από τις δύο αυτές σχέσεις, συνεπάγεται:

$$[z, z_0, z_1, \ldots, z_n]\,V_{n+1}(z) = f(z) - p_n(z).$$

Συνδυασμός της τελευταίας ισότητας με την *Ιδιότητα IV.1.4.3* αποδεικνύει τον ισχυρισμό της Πρότασης. ■

Μία άμεση συνέπεια των παραπάνω είναι ένας πρώτος υπολογισμός του σφάλματος παρεμβολής:

Ιδιότητα IV.1.4.5. ***Εάν*** $V_{n+1}(z)=(z-z_0)(z-z_1)\ldots(z-z_n)$, ***τότε το σφάλμα παρεμβολής δίνεται από τον Τύπο:***

$$E_n(z)=p_n(z)-f(z)=[z,z_0,z_1,\ldots,z_n]\,V_{n+1}(z)\quad (z\in U).\blacksquare$$

Από αυτήν την *Ιδιότητα* καθίσταται φανερό ότι, με την βοήθεια των διηρημμένων διαφορών μπορούμε να επιτύχουμε μία πολύ γενική έκφραση για το σφάλμα παρεμβολής. Όμως, αυτή η έκφραση του σφάλματος δεν έχει εφαρμογές στα πρακτικά προβλήματα, γιατί ο υπολογισμός της διηρημμένης διαφοράς $[z,z_0,z_1,\ldots,z_n]$ προϋποθέτει υπολογιστική διαχείριση της $f(z)$. Θα δούμε λίγο αργότερα πως, όταν η συνάρτηση f είναι $(n+1)-$φορές συνεχώς διαφορίσιμη, το σφάλμα παρεμβολής μπορεί να τεθεί σε μία πιο εύχρηστη μορφή.

IV.1.5. Ο ΤΥΠΟΣ ΤΟΥ NEWTON

Ο ***Τύπος του Newton*** επιτρέπει τον υπολογισμό του πολυωνύμου παρεμβολής $p_n(z)$ μίας συνάρτησης $f:\mathbb{K}\to\mathbb{K}:z\mapsto f(z)$, στην περίπτωση που τα σημεία παρεμβολής $z_0,z_1,\ldots,z_n\in\mathbb{K}$ βρίσκονται όλα σε μία ευθεία και ισαπέχουν μεταξύ τους με την έννοια ότι:

$$z_j=z_0+j\,h\quad (j=0,1,\ldots,n),$$

όπου $h\in\mathbb{K}$ είναι μία σταθερή παράμετρος.

Μπορούμε να αποκομίσουμε αυτόν τον τύπο του Newton χρησιμοποιώντας διαφορετικούς τρόπους.

Κατ' αρχάς, εισάγουμε τον *Ορισμό*:

Ορισμός IV.1.5.1. *Έστω* F *το σύνολο των συναρτήσεων* $f : K \to K \quad (K = \mathbb{R}, \mathbb{C})$ *και έστω* $h \in K$.

(i). *Ονομάζουμε τελεστή των προς τα εμπρός διαφορών τον τελεστή* Δ *που σε κάθε* $f \in \mathsf{F}$ *αντιστοιχεί την συνάρτηση* Δf *σύμφωνα με τον Τύπο:*

$$\Delta f(z) = f(z+h) - f(z).$$

Εάν $k = 1,2,\ldots$, *τότε η* $k-$*οστή δύναμη του* Δ *είναι ο τελεστής* $\Delta^k : \mathsf{F} \to \mathsf{F}$, *που σε κάθε* $f \in \mathsf{F}$ *αντιστοιχεί την συνάρτηση* $\Delta^k f \in \mathsf{F}$ *σύμφωνα με τον επαγωγικό Τύπο:*

$$\Delta^k f(z) = \Delta(\Delta^{k-1} f(z))$$

ή

$$\Delta^k f(z) = \Delta^{k-1} f(z+h) - \Delta^{k-1} f(z).$$

Εξ άλλου, κατά σύμβαση, *θέτουμε:*

$$\Delta^0 f(z) := f(z).$$

(ii). *Ονομάζουμε τελεστή των προς τα πίσω διαφορών τον τελεστή* $\Delta : \mathsf{F} \to \mathsf{F}$ *που σε κάθε* $f \in \mathsf{F}$ *αντιστοιχεί την συνάρτηση* ∇f, *σύμφωνα με τον Τύπο:*

$$\nabla f(z) = f(z) - f(z-h).$$

Εάν $k = 1,2,\ldots$, *τότε η* $k-$*οστή δύναμη του* ∇ *είναι ο τελεστής* $\nabla^k : \mathsf{F} \to \mathsf{F}$, *που σε κάθε* $f \in \mathsf{F}$ *αντιστοιχεί την συνάρτηση* $\nabla^k f \in \mathsf{F}$, *σύμφωνα με τον επαγωγικό Τύπο:*

$$\nabla^k f(z) = \nabla(\nabla^{k-1} f(z))$$

ή

$$\nabla^k f(z) = \nabla^{k-1} f(z) - \nabla^{k-1} f(z-h).$$

Κατά σύμβαση, *θέτουμε:*

$$\nabla^0 f(z) := f(z).$$

(iii). ***Ονομάζουμε τελεστή των κεντρικών διαφορών, τον τελεστή δ:*** $\mathrm{F} \to \mathrm{F}$ ***που σε κάθε*** $f \in \mathrm{F}$ ***αντιστοιχεί την συνάρτηση*** δf ***σύμφωνα με τον Τύπο:***

$$\delta f(z) = f\left(z + \frac{h}{2}\right) - f\left(z - \frac{h}{2}\right).$$

Εάν $k = 1,2,\ldots$***, τότε η*** $k-$***οστή δύναμη του*** δ ***είναι ο τελεστής*** $\delta^k : \mathrm{F} \to \mathrm{F}$ ***που σε κάθε*** $f \in \mathrm{F}$ ***αντιστοιχεί την συνάρτηση*** $\delta^k f \in \mathrm{F}$ ***σύμφωνα με τον επαγωγικό Τύπο:***

$$\delta^k f(z) = \delta(\delta^{k-1} f(z))$$

ή

$$\delta^k f(z) = \delta^{k-1} f(z + \frac{h}{2}) - \delta^{k-1} f(z - \frac{h}{2}).$$

Κατά σύμβαση, *θέτουμε:*

$$\delta^0 f(z) := f(z). \blacksquare$$

Τα σύμβολα Δ, ∇ και δ καλούνται ***τελεστές διαφορών***. Οι τελεστές αυτοί είναι K – γραμμικοί, γιατί ικανοποιούν την σχέση:

$$T(\lambda\, f(z) + \mu\, g(z)) = \lambda\, T(f(z)) + \mu\, T(g(z)),$$

όπου $T = \Delta, \nabla$ και δ, για κάθε $\lambda, \mu \in K$ και κάθε $f, g \in \mathrm{F}$. Εάν $E : \mathrm{F} \to \mathrm{F}$ είναι ο ***τελεστής μετατόπισης***, που ορίζεται από τον Τύπο:

$$Ef(z) = f(z+h),$$

και εάν $id : \mathrm{F} \to \mathrm{F}$ είναι ο ***ταυτοτικός τελεστής***, που ορίζεται από τον Τύπο:

$$idf(z) = f(z),$$

τότε οι τελεστές διαφορών Δ και ∇ μπορούν να εκφραστούν με την βοήθεια των τελεστών E και id:

Πρόταση IV.1.5.2. ***Ισχύουν οι παρακάτω συσχετισμοί:***

(i). $\Delta = E - id$,

(ii). $E = (id - \nabla)^{-1}$ ***και*** $\nabla = id - E^{-1}$,

(iii). $\Delta = (id - \nabla)^{-1} - id$,

όπου $(id - \nabla)^{-1}$ ***είναι η αντίστροφη απεικόνιση της*** $(id - \nabla)$.

Απόδειξη. Έστω $f \in \mathrm{F}$ και $z \in K$. Τότε:

(i). Ισχύει:

$$\begin{aligned}\Delta f(z) &= f(z+h) - f(z)\\ &= Ef(z) - f(z)\\ \Rightarrow \Delta f(z) &= (E - id)\big(f(z)\big)\\ \Rightarrow \Delta &= E - id.\end{aligned}$$

(ii). Έχουμε:

$$\nabla f(z+h) = f(z+h) - f(z) = (E - id)\big(f(z)\big),$$

κι επειδή $\nabla f(z+h) = \nabla Ef(z)$ συνεπάγεται ότι:

$$\nabla Ef(z) = (E - id)\big(f(z)\big) \Rightarrow \nabla = id - E^{-1} \quad \text{και} \quad E = (id - \nabla)^{-1}.$$

(iii). Συνδυασμός των *(i)* και *(ii)* αποδεικνύει ότι

$$\Delta = (id - \nabla)^{-1} - id\,. \blacksquare$$

Πρόταση IV.1.5.3. ***Για κάθε*** $f \in \mathrm{F}$ ***και κάθε*** $\kappa = 0,1,2,\ldots$ ***, ισχύουν τα ακόλουθα:***

(i). $\Delta^k f(z) = \sum_{i=0}^{k} C_k^i\ (-1)^i\ f(z + [k-i]\,h) \qquad (z \in K),$

(ii). $\nabla^k f(z) = \sum_{i=0}^{k} C_k^i\ (-1)^{k+i}\ f(z - [k-i]\,h) \qquad (z \in K),$

όπου:

$$C_k^i = \frac{k!}{i!\,(k-i)!}.$$

Απόδειξη. (i). Σύμφωνα με την *Πρόταση IV.1.5.2.(i)*, είναι $\Delta = E - id$, και επομένως $\Delta^k = (E - id)^k$**,** οπότε, χρησιμοποιώντας τον *Διωνυμικό Τύπο*, λαμβάνουμε:

$$\Delta^k = \sum_{i=0}^{k} C_k^i \, (-1)^i \, E^{k-i} \, .$$

Οι δυνάμεις του τελεστή μετατόπισης E ορίζονται με τον ίδιο τρόπο όπως και οι δυνάμεις του τελεστή διαφορών Δ. Έτσι:

$$E^{\nu} f(z) = f(z + \nu\, h), \text{ για κάθε } \nu = 0,1,2,...,$$

και συνεπώς:

$$\Delta^k f(z) = \sum_{i=0}^{k} C_k^i \, (-1)^i \, f(z + [k-i]) \qquad (z \in \mathbb{K}).$$

(ii). Σύμφωνα με την *Πρόταση IV.1.5.2.(ii)*, είναι $\nabla = id - E^{-1} = -(E^{-1} - id)$, και επομένως $\nabla^k = (-1)^k \, (E^{-1} - id)^k$, οπότε, χρησιμοποιώντας πάλι τον *Διωνυμικό Τύπο*, λαμβάνουμε:

$$\nabla^k = (-1)^k \sum_{i=0}^{k} C_k^i \, (-1)^i \, (E^i)^{k-i} \, .$$

Είναι σαφές πως ο αντίστροφος E^{-1} του τελεστή μετατόπισης E ορίζεται από την σχέση $E^{-1} f(z) = f(z-h)$ και πως οι δυνάμεις του τελεστή E^{-1} ορίζονται από την ισότητα:

$$(E^{-1}) f(z) = f(z - \nu\, h), \quad \text{για κάθε } \nu = 0,1,2,\ldots,$$

και συνεπώς

$$\nabla^k f(z) = \sum_{i=0}^{k} C_k^i \, (-1)^{k+i} \, f(z - [k-i]\, h) \qquad (z \in \mathbb{K}).$$

Η *Απόδειξη* του *Πορίσματος* είναι πλήρης. ■

Στην συνέχεια, θεωρούμε μία τυχαία συνάρτηση $f:\mathbb{R}\to\mathbb{R}:x\mapsto f(x)$ και υποθέτουμε ότι η απόσταση καθενός από τα σημεία x_i αυτά από το επόμενό του είναι σταθερή. Κάτω από αυτές τις συνθήκες, θα εξετάσουμε τον τρόπο με τον οποίο όλα όσα προηγήθηκαν σχετικά με τους τελεστές διαφορών μπορούν να αποδώσουν διεξόδους στο πρόβλημα εύρεσης μίας πρακτικά δόκιμης έκφρασης του πολυωνύμου παρεμβολής $p_n(x)$ της $f(x)$ στα ισαπέχοντα σημεία. Υποθέτουμε ότι δίνεται ένας πίνακας τιμών της $f(x)$:

x_i	$f_i = f(x_i)$
$\vdots$	$\vdots$
x_{-4}	$f_{-4} = f(x_{-4})$
$x_{-7/2}$	$f_{-7/2} = f(x_{-7/2})$
x_{-3}	$f_{-3} = f(x_{-3})$
$x_{-5/2}$	$f_{-5/2} = f(x_{-5/2})$
x_{-2}	$f_{-2} = f(x_{-2})$
$x_{-3/2}$	$f_{-3/2} = f(x_{-3/2})$
x_{-1}	$f_{-1} = f(x_{-1})$
$x_{-1/2}$	$f_{-1/2} = f(x_{-1/2})$
x_0	$f_0 = f(x_0)$
$x_{1/2}$	$f_{1/2} = f(x_{1/2})$
x_1	$f_1 = f(x_1)$
$x_{3/2}$	$f_{3/2} = f(x_{3/2})$
x_2	$f_2 = f(x_2)$
$x_{5/2}$	$f_{5/2} = f(x_{5/2})$
x_3	$f_3 = f(x_3)$
$x_{7/2}$	$f_{7/2} = (x_{7/2})$

x_4	$f_4 = f(x_4)$
$\vdots$	$\vdots$
x_{n-2}	$f_{n-2} = f(x_{n-2})$
$x_{[n-3]/2}$	$f_{[n-3]/2} = f(x_{[n-3]/2})$
x_{n-1}	$f_{n-1} = f(x_{n-1})$
$x_{[n-1]/2}$	$f_{[n-1]/2} = f(x_{[n-1]/2})$
x_n	$f_n = f(x_n)$
$x_{[n+1]/2}$	$f_{[n+1]/2} = f(x_{[n+1]/2})$
x_{n+1}	$f_{n+1} = f(x_{n+1})$
$x_{[n+3]/2}$	$f_{[n+3]/2} = f(x_{[n+3]/2})$
x_{n+2}	$f_{n+2} = f(x_{n+2})$
$\vdots$	$\vdots$

Η ***πρώτης τάξης προς τα εμπρός πεπερασμένη διαφορά*** *της* f *στο σημείο* x_n ορίζεται από τις σχέσεις:

$$\Delta f_n = f(x_{n+1}) - f(x_n) = f_{n+1} - f_n .$$

Η ***πρώτης τάξης προς τα πίσω πεπερασμένη διαφορά*** *της* f *στο σημείο* x_n ορίζεται από τις σχέσεις:

$$\nabla f_n = f(x_n) - f(x_{n-1}) = f_n - f_{n-1} .$$

Η ***πρώτης τάξης κεντρική διαφορά*** *της* f *στο σημείο* x_n ορίζεται από τις σχέσεις:

$$\delta f_n = f(x_{n+1}) - f(x_{n-1}) = f_{n+1} - f_{n-1} .$$

Γενικότερα, έχουμε τους εξής ορισμούς (στους οποίους, κατά σύμβαση, έχουμε θέσει $\Delta^0 f_n := \nabla^0 f_n := \delta^0 f_n := f_n$):

Η $k-$***τάξης προς τα εμπρός πεπερασμένη διαφορά*** *της* f *στο σημείο* x_n ορίζεται από τις σχέσεις:

$$\Delta^k f_n = \Delta(\Delta^{k-1} f_n) = \Delta^{k-1} f_{n+1} - \Delta^{k-1} f_n \qquad (k = 1,2,\ldots).$$

Η $k-$***τάξης προς τα πίσω πεπερασμένη διαφορά*** *της* f *στο σημείο* x_n ορίζεται από τις σχέσεις:

$$\nabla^k f_n = \nabla(\nabla^{k-1} f_n) = \nabla^{k-1} f_n - \nabla^{k-1} f_{n-1} \qquad (k = 1,2,\ldots).$$

Η $k-$***τάξης κεντρική διαφορά*** *της* f *στο σημείο* x_n ορίζεται από τις σχέσεις:

$$\delta^k f_n = \delta(\delta^{k-1} f_n) = \delta^{k-1} f_{n+1} - \delta^{k-1} f_{n+1} \qquad (k = 1,2,\ldots).$$

Είναι προφανές πως όταν η απόσταση κάθε σημείου x_i της αριστερής στήλης από το επόμενο του (ή από το προηγούμενο του) είναι σταθερή, τότε τα σύμβολα

$$\Delta, \nabla \text{ και } \delta$$

παραπέμπουν στην κατά σημείο μελέτη των τελεστών διαφορών του *Ορισμού IV.1.5.1*.

Επί πλέον, αποδεικνύεται εύκολα ότι:

$$\Delta f_n = \nabla f_{n+1} = \delta f_{\frac{n+1}{2}}$$

και γενικότερα, με επαγωγή, ότι:

$$\Delta^k f_n = \nabla^k f_{n+1} = \delta^k f_{\frac{n+k}{2}} \qquad (k = 1,2,\ldots).$$

Πριν προχωρήσουμε, ας δώσουμε την κατασκευή κάποιων ενδεικτικών πινάκων πεπερασμένων διαφορών, που είναι αρκετά χρήσιμοι για την κατανόηση του διατακτικού μηχανισμού δημιουργίας των ποσοτήτων

$$\Delta^k f_n, \nabla^k f_n \quad \text{και} \quad \delta^k f_{\frac{n+m}{2}}:$$

1) Πίνακας των προς τα εμπρός πεπερασμένων διαφορών $(n=5)$:

$$
\begin{array}{lllll}
x_0 & f_0 & & & \\
 & & \Delta f_0 & & \\
x_1 & f_1 & & \Delta^2 f_0 & \\
 & & \Delta f_1 & & \Delta^3 f_0 \\
x_2 & f_2 & & \Delta^2 f_1 & \\
 & & \Delta f_2 & & \Delta^3 f_1 \\
x_3 & f_3 & & \Delta^2 f_2 & \\
 & & \Delta f_3 & & \\
x_4 & f_4 & & &
\end{array}
$$

2) Πίνακας των προς τα πίσω διαφορών $(n=5)$:

$$
\begin{array}{lllll}
x_{-4} & f_{-4} & & & \\
 & & \Delta f_{-3} & & \\
x_{-3} & f_{-3} & & \Delta^2 f_{-2} & \\
 & & \Delta f_{-2} & & \Delta^3 f_{-1} \\
x_{-2} & f_{-2} & & \Delta^2 f_{-1} & \\
 & & \Delta f_{-1} & & \Delta^3 f_0 \\
x_{-1} & f_{-1} & & \Delta^2 f_0 & \\
 & & \Delta f_0 & & \\
x_0 & f_0 & & &
\end{array}
$$

3) Πίνακας κεντρικών διαφορών

$$
\begin{array}{lllll}
x_{-2} & f_{-2} & & & \\
 & & \delta f_{-3/2} & & \\
x_{-1} & f_{-1} & & \delta^2 f_{-1} & \\
 & & \delta f_{-1/2} & & \delta^3 f_{-1/2} \\
x_0 & f_0 & & \delta^2 f_0 & \\
 & & \delta f_{1/2} & & \delta^3 f_{1/2} \\
x_1 & f_1 & & \delta^2 f_1 & \\
 & & \delta f_{3/2} & & \\
x_2 & f_2 & & &
\end{array}
$$

Θα εξετάσουμε τώρα την συσχέτιση που υπάρχει ανάμεσα στην έννοια των διηρημμένων διαφορών και την έννοια των πεπερασμένων (προς τα εμπρός) διαφορών. Έχουμε το εξής:

Θεώρημα IV.1.5.4. ***Εάν*** $x_i = x_0 + i\,h,$ ***με*** $h > 0$ ***και*** $i = 0,1,\ldots,$ ***τότε:***

$$k!\,h^k\,[x_i, x_{i+1}, \ldots, x_{i+k}] = \Delta^k f(x_i)$$

για κάθε $k = 0,1,2,\ldots$

Απόδειξη. Θα χρησιμοποιήσουμε ακόμα μία φορά μαθηματική επαγωγή. Η ιδιότητα είναι αληθής *όταν* $k = 0$ (και για κάθε i), γιατί τότε:

$$0!\,h^0\,[x_i] = [x_i] = f(x_i)\,.$$

Ας υποθέσουμε ότι η ιδιότητα αληθεύει όταν $k \le N-1$ (και για κάθε i) και ας αποδείξουμε ότι παραμένει αληθής όταν $k = N$. Είναι:

$$\Delta^N f(x_i) = \Delta^{N-1} f(x_{i+1}) - \Delta^{N-1} f(x_i)$$

$$= (N-1)!\,h^{N-1}\,([x_{i+1},\ldots,x_{i+N}] - [x_i,\ldots,x_{i+N-1}]),$$

οπότε χρησιμοποιώντας τον ορισμό των διηρημμένων διαφορών:

$$[x_i,\ldots,x_{i+N}] = \frac{[x_i,\ldots,x_{i+N-1}] - [x_{i+1},\ldots,x_{i+N}]}{x_i - x_{i+N}},$$

αλλά και το γεγονός ότι:

$$x_i - x_{i+N} = x_0 + i\,h - x_0 - (i+N)\,h = -N\,h,,$$

διαπιστώνουμε πως:

$$\Delta^N f(x_i) = (N-1)!\,h^{N-1}\,N\,h\,[x_i,\ldots,x_{i+N}],$$

που ολοκληρώνει την *Απόδειξη* του *Θεωρήματος*. ■

Συνδυάζοντας τώρα το παραπάνω *Θεώρημα* με την *Ιδιότητα IV.1.4.4* της προηγούμενης *Παραγράφου*, οδηγούμαστε άμεσα στο:

Θεώρημα IV.1.5.5. ***Εάν*** $x_i = x_0 + i\,h,$ ***με*** $h > 0$ ***και*** $i = 0,1,2,\ldots,n$ ***, τότε το πολυώνυμο παρεμβολής*** $p_n(x)$ ***της*** f ***στα σημεία*** x_i ***είναι:***

$$p_n(x) = f(x_0) + (x - x_0)\,\frac{\Delta f(x_0)}{1!\,h} + (x - x_0)\,(x - x_1)\,\frac{\Delta^2 f(x_0)}{2!\,h^2} + \ldots$$

$$\ldots + (x - x_0)\,(x - x_1)\ldots(x - x_{n-1})\,\frac{\Delta^n f(x_0)}{n!\,h^n}.\ \blacksquare$$

Παρατήρηση IV.1.5.6. Ο υπολογισμός των ποσοτήτων $\Delta^k f(x_0)$ μπορεί να πραγματοποιηθεί επαγωγικά με την βοήθεια ενός πίνακα, όπως αυτός που δόθηκε ενδεικτικά λίγο πιο πάνω, είτε απευθείας χρησιμοποιώντας το *Πόρισμα IV.1.5.3*. ■

Το *Θεώρημα IV.1.5.5* μπορεί να διασκευαστεί κατάλληλα έτσι ώστε το πολυώνυμο παρεμβολής να εκφράζεται με την βοήθεια των πεπερασμένων προς τα πίσω διαφορών $\nabla f(x_0), \nabla^2 f(x_0), \ldots, \nabla^n f(x_0)$.

Προς τούτο, αρκεί κανείς να θεωρήσει τα σημεία:

$$x_i = x_0 + i\,h, \text{ με } h > 0 \text{ και } i = 0,1,2,\ldots,n,$$

και να αξιώσει ότι τα σημεία αυτά ταυτίζονται με εκείνα της προηγούμενης περίπτωσης, δηλαδή ότι:

$$\begin{aligned} x_0 &\equiv x_0, \\ x_{-1} &\equiv x_1, \\ &\vdots \\ x_{-n} &\equiv x_n. \end{aligned}$$

Τότε, είναι εύκολο να διαπιστώσει (χρησιμοποιώντας μαθηματική επαγωγή) ότι:

$$\Delta^k f_0 = (-1)^k \nabla^k f_0 \qquad \forall\, k = 1,2,\ldots$$

και να συμπεράνει το ακόλουθο:

Θεώρημα IV.1.5.7. ***Εάν*** $x_i = x_0 + i\,h,$ ***με*** $h > 0$ ***και*** $i = 0,1,2,\ldots,n$ ***, τότε το πολυώνυμο παρεμβολής*** $p_n(x)$ ***της*** f ***στα σημεία*** x_i ***είναι:***

$$p_n(x) = f(x_0) - (x - x_0)\frac{\nabla f(x_0)}{1!\,h} + (x - x_0)(x - x_1)\frac{\nabla^2 f(x_0)}{2!\,h^2} - \cdots$$

$$\cdots + (-1)^n (x - x_0)(x - x_1)\cdots(x - x_{n-1})\frac{\nabla^n f(x_0)}{n!\,h^n}.$$

Παρατήρηση IV.1.5.8. Όπως και στο *Θεώρημα IV.1.5.5*, ο υπολογισμός των ποσοτήτων $\nabla^k f(x_0)$ μπορεί να γίνει επαγωγικά με την βοήθεια ενός πίνακα όπως αυτός που δόθηκε λίγο παραπάνω είτε απ' ευθείας χρησιμοποιώντας το *Πόρισμα IV.1.5.3*.

Παρατήρηση IV.1.5.9. Το πολυώνυμο παρεμβολής $p_n(z)$ της f μπορεί επίσης να εκφραστεί και με την βοήθεια του τελεστή δ των κεντρικών διαφορών, που όπως είναι γνωστό ορίζεται από την σχέση

$$\delta f(x) = f(x+\frac{h}{2}) - f(x-\frac{h}{2}),$$

ή ακόμα και με την βοήθεια του *τελεστή* μ, που ορίζεται από την σχέση

$$\mu f(x) = \frac{1}{2}[f(x+\frac{h}{2}) + f(x-\frac{h}{2})].$$

Οι εκφράσεις όμως αυτές στερούνται ουσιαστικού πρακτικού αντικρίσματος, γιαυτό, στην θεωρητική μελέτη του φαινομένου της παρεμβολής για ισαπέχοντα σημεία, έχουν κυριαρχήσει οι εκφράσεις που περιγράφονται στα *Θεωρήματα IV.1.5.5* και *IV.1.5.7.* ■

IV.1.6. ΣΦΑΛΜΑ ΠΑΡΕΜΒΟΛΗΣ

Υστερα από όσα αναφέρθηκαν στις προηγούμενες *Παραγράφους*, έχουμε εξαντλήσει το πρόβλημα του υπολογισμού του πολυωνύμου παρεμβολής $p_n(z)$ μίας συνάρτησης

$$f : U \to \boldsymbol{K} : z \mapsto f(z) \ \ (U \subset \boldsymbol{K} = \boldsymbol{R}, \boldsymbol{C}),$$

στην περίπτωση που τα σημεία παρεμβολής $z_0, z_1, \ldots, z_n$ είναι αμοιβαία διαφορετικά μεταξύ τους.

Μετά την εύρεση του πολυωνύμου παρεμβολής $p_n(z)$, που διέρχεται από τα ανά δύο διάφορα σημεία $z_0, z_1, \ldots, z_n$ του U, υπάρχουν τρία βασικά ερωτήματα που εγείρονται και αποσκοπούν στην κατανόηση των δυνατοτήτων και της προσεγγιστικής αποτελεσματικότητας του πολυωνύμου αυτού.

Συγκεκριμένα, το πρώτο ερώτημα, που τίθεται, είναι ο υπολογισμός (ή έστω η εκτίμηση) του σφάλματος που προέκυψε από την παρεμβολή σε κάθε σημείο του U :

$$E_n(z) = f(z) - p_n(z) \qquad (z \in U).$$

Το δεύτερο ερώτημα, που τίθεται, αφορά την βέλτιστη επιλογή των σημείων παρεμβολής, δηλαδή την επιλογή εκείνη των σημείων $z_0, z_1, \ldots, z_n$ που ελαχιστοποιεί την απόλυτη τιμή $|E_n(z)|$ του σφάλματος παρεμβολής (εάν είναι δυνατόν, σε κάθε σημείο z του συνόλου U).

Το τρίτο, και τελευταίο, ερώτημα που τίθεται, συνίσταται στην σύγκλιση των ακολουθιών πολυωνύμων παρεμβολής, με άλλα λόγια, στην αναζήτηση ενός συνεχούς διογκούμενου πλήθους σημείων $z_0^{(n)}, z_1^{(n)}, \ldots, z_n^{(n)}$ του U $(z_i^{(n)} \neq z_j^{(n)})$, τέτοιου ώστε εάν $p_n(z)$ είναι τα πολυώνυμα παρεμβολής στα $z_0^{(n)}, \ldots, z_n^{(n)}$ $(n \in \mathbb{N}_0)$ να εξασφαλίζεται η σύγκλιση

$$\lim_{n\to\infty} E_n(z) = \lim_{n\to\infty} (f(z) - p_n(z)) = 0$$

εάν είναι δυνατόν, σε κάθε σημείο $z \in U$).

Στην παρούσα και στις επόμενες δύο *Παραγράφους* θα ασχοληθούμε με την μελέτη αυτών των τριών ερωτημάτων.

Για λόγους ενίσχυσης της υπολογιστικής ευχέρειας και χωρίς σημαντικό περιορισμό της γενικότητας, θα υποθέσουμε ότι $U = [-1,1]$ και ότι η f είναι μία πραγματική συνάρτηση:

$$f : [-1,1] \to \mathbb{R} : x \mapsto f(x).$$

Ας αρχίσουμε με την μελέτη του σφάλματος παρεμβολής:

$$E_n(x) = f(x) - p_n(x) \ , (-1 \le x \le 1)$$

όπου $p_n(x)$ είναι ένα πολυώνυμο, βαθμού το πολύ n, που παρεμβάλει την f στα, ανά δύο διάφορα μεταξύ τους σημεία:

$$x_0, x_1, \ldots, x_n \in [-1,1].$$

Έστω $x \in [-1,1]$, με x: σταθερό σημείο και έστω I το κλειστό διάστημα:

$$I := [\min\{x, x_0, x_1, \ldots, x_n\}, \max\{x, x_0, x_1, \ldots, x_n\}].$$

Έχουμε το ακόλουθο:

Θεώρημα IV.1.6.1. ***Εάν η f είναι $(n+1)-$φορές συνεχώς διαφορίσιμη, μέσα σε μία ανοικτή περιοχή του διαστήματος I, τότε υπάρχει ένα σημείο $\xi \in I$ τέτοιο ώστε:***

$$E_n(x) = \frac{V_n(x)}{(n+1)!} \frac{d^{n+1} f}{dx^{n+1}}(\xi),$$

όπου $V_n(x) = (x - x_0)(x - x_1)\ldots(x - x_n)$.

Απόδειξη. Θέτουμε:

$$F(t) := f(t) - p_n(t) - c\, V_n(t) \qquad (-1 \le t \le 1),$$

όπου c είναι ο σταθερός αριθμός που δίνεται από την σχέση:

$$c = \frac{f(x) - p_n(x)}{V_n(x)}.$$

Είναι εύκολο να διαπιστώσουμε πως η συνάρτηση F έχει τις εξής ιδιότητες:

- $F(x_i) = 0 \qquad \forall\, i = 0,1,\ldots,n$,
- $F(x) = 0$ (σύμφωνα με τον ορισμό του αριθμού c), και

- η F έχει τουλάχιστον $(n+2)$ ρίζες στο *I*.

Σύμφωνα με το *Θεώρημα Rolle*, η παράγωγος F' της F έχει τουλάχιστον $(n+1)$ ρίζες στο *I*, η δεύτερη παράγωγος F'' της F έχει τουλάχιστον n ρίζες στο *I*, και συνεχίζοντας, με την επίκληση του ιδίου επιχειρήματος, η $(n+1)-$τάξης παράγωγος:

$$\frac{d^{n+1}F}{d\,t^{n+1}}$$

της F έχει τουλάχιστον μία ρίζα στο *I*.

Εάν συμβολίσουμε με ξ αυτήν την ρίζα, τότε, επειδή

$$\frac{d^{n+1}F}{d\,t^{n+1}}(t)=\frac{d^{n+1}f}{d\,t^{n+1}}(t)-\frac{d^{n+1}p_n}{d\,t^{n+1}}(t)+c\frac{d^{n+1}V_n}{d\,t^{n+1}}(t),$$

και επί πλέον επειδή:

$$\frac{d^{n+1}p_n}{dt^{n+1}}=0 \quad \text{και} \quad \frac{d^{n+1}V_n}{dt^{n+1}}(t)=(n+1)!,$$

θα είναι:

$$0=\frac{d^{n+1}F}{dt^{n+1}}(\xi)=\frac{d^{n+1}f}{dt^{n+1}}(\xi)-c\,(n+1)!,$$

απ' όπου συνεπάγεται η σχέση:

$$c=\frac{f(x)-p_n(x)}{V_n(x)}=\frac{1}{(n+1)!}\quad\frac{d^{n+1}f}{dt^{n+1}}(\xi),$$

από την οποία ο ισχυρισμός του *Θεωρήματος*. ■

Μία άμεση συνέπεια του συνδυασμού του *Θεωρήματος* αυτού με την *Ιδιότητα IV.1.4.5* εκτίθεται στο:

Πόρισμα IV.1.6.2. ***Κάτω από τις υποθέσεις του Θεωρήματος IV.1.6.1, ισχύει ότι:***

$$[x, x_0, x_1, \ldots, x_n] = \frac{1}{(n+1)!} \frac{d^{n+1} f}{d\, x^{n+1}}(\xi). \blacksquare$$

IV.1.7. ΒΕΛΤΙΣΤΗ ΕΠΙΛΟΓΗ ΤΩΝ ΣΗΜΕΙΩΝ ΠΑΡΕΜΒΟΛΗΣ

Στην *Παράγραφο* αυτή θα αναζητήσουμε μία επιλογή των σημείων παρεμβολής $x_0, x_1, \ldots, x_n \in [-1,1]$ που ελαχιστοποιούν τον αριθμό:

$$\max_{-1 \le x \le 1} |E_n(x)|,$$

όπου

$$E_n(x) = f(x) - p_n(x),$$

$f(x)$ είναι μία πραγματική συνάρτηση ορισμένη στο κλειστό διάστημα $[-1, 1]$

και

$p_n(x)$ είναι το πολυώνυμο, το πολύ βαθμού n ,που παρεμβάλλει την $f(x)$ στα σημεία $x_0, x_1, \ldots, x_n$.

Ας υποθέσουμε ότι η συνάρτηση f είναι $(n+1)-$φορές συνεχώς διαφορίσιμη, μέσα μία ανοικτή περιοχή του διαστήματος $[-1, 1]$.

Σύμφωνα με το *Θεώρημα IV.1.6.1*, το σφάλμα της παρεμβολής $E_n(x)$ περιέχει δύο παράγοντες:

- έναν παράγοντα που εξαρτάται από την f, δηλαδή τον παράγοντα:

$$\frac{d^{n+1}f}{dx^{n+1}}(\xi),$$

- έναν παράγοντα, που είναι ανεξάρτητος από την f, αλλά εξαρτάται άμεσα από τα σημεία παρεμβολής $x_0,\ldots,x_n$, δηλαδή τον παράγοντα:

$$V_n(x).$$

Επειδή ισχύει:

$$|E_n(x)| \le \sup_{x\in[-1,1]}\left|\frac{1}{(n+1)!}\frac{d^{n+1}f}{dx^{n+1}}(x)\right| \sup_{x\in[-1,1]}|V_n(x)|,$$

είναι προφανές ότι, εφόσον γνωρίζουμε την τιμή της f σε κάθε σημείο $x\in[-1,1]$, μπορούμε να επιλέξουμε τα σημεία $x_0,x_1,\ldots,x_n$ έτσι ώστε να ελαχιστοποιείται η μέγιστη τιμή

$$\sup_{x\in[-1,1]}|E(x)|$$

του σφάλματος παρεμβολής στο $[-1,1]$.

Πράγματι αρκεί να επιλύσουμε ως προς $x_0,x_1,\ldots,x_n$ το ακόλουθο πρόβλημα βελτιστοποίησης:

$$\min_{x_0,\ldots,x_n\in[-1,1]} \quad \max_{x\in[-1,1]}|V_n(x)|.$$

Η απάντηση στο πρόβλημα αυτό δίνεται με την εισαγωγή των *πολυωνύμων του Chebyshev* (ή *Tchebycheff*):

Ορισμός IV.1.7.1. ***Ονομάζουμε πολυώνυμο Chebyshev βαθμού n την απεικόνιση:***

$$T_n : [-1,1] \to [-1,1] : x \mapsto T_n(x) = \cos(n\ Arc\cos x). \ \blacksquare$$

Τα πολυώνυμα Chebyshev διέπονται και χαρακτηρίζονται από τις ιδιότητες που περιγράφονται στα επόμενα τέσσερα *Θεωρήματα*:

Θεώρημα IV.1.7.2. ***Τα πολυώνυμα επαληθεύουν την επαγωγική σχέση:***

$$T_{n+1}(x) = 2\,x\,T_n(x) - T_{n-1}(x) \qquad (n = 1,2,\ldots),$$

με:

$$T_0(x) = 1 \qquad \textbf{\textit{και}} \qquad T_1(x) = x.$$

Απόδειξη. Κατ' αρχάς, παρατηρούμε πως:

$$T_0(x) = \cos(0) = 1 \qquad \text{και} \qquad T_1(x) = \cos(Arc\cos x) = x$$

Ακόμη, έχουμε:

$$\begin{aligned}
2\,x\,T_n(x) - T_{n-1}(x) &= 2\,x\,\cos(n\ Arc\cos x) - \cos([n-1]\cos x) \\
&= 2\,x\,\cos(n\ Arc\cos x) - \cos(Arc\cos x)\cos(n\ Arc\cos x) \\
&\qquad - \sin(Arc\cos x)\sin(n\ Arc\cos x) \\
&= x\,\cos(n\ Arc\cos x) - \sin(Arc\cos x)\sin(n\ Arc\cos x) \\
&= T_1(x)\cos(n\ Arc\cos x) - \sin(Arc\cos x)\sin(n\ Arc\cos x) \\
&= \cos(Arc\cos x)\cos(n\ Arc\cos x) \\
&\qquad - \sin(Arc\cos x)\sin(n.Arc\cos x) \\
&= \cos([n+1]\ Arc\cos x) \\
&= T_{n+1}(x),
\end{aligned}$$

που ολοκληρώνει την *Απόδειξη* του *Θεωρήματος*. ■

Θεώρημα IV.1.7.3. ***Η συνάρτηση $T_n(x)$ είναι ένα πολυώνυμο βαθμού n. Ο συντελεστής του x^n στο πολυωνυμικό ανάπτυγμα του $T_n(x)$ ισούται με 2^{n-1}.***

Απόδειξη. Η *Απόδειξη* είναι μία προφανής συνέπεια του προηγούμενου *Θεωρήματος*. ■

Θεώρημα IV.1.7.4. ***Το πολυώνυμο $T_n(x)$ του Chebyshev έχει ακριβώς n απλές ρίζες, τα σημεία:***

$$X_i = \cos(\frac{2i+1}{2n}\pi) \qquad i = 0,1,\ldots,n-1.$$

Απόδειξη. Επειδή, για κάθε $i = 0,1,\ldots,n-1$, είναι:

$$T_n(X_i) = \cos(n\,Arc\cos(\frac{2i+1}{2n}\pi)) = \cos(n\,\frac{2i+1}{2n}\pi) = \cos(\frac{2i+1}{2}\pi) = 0,$$

τα n σημεία:

$$X_i = \cos(\frac{2i+1}{2n}\pi) \qquad i = 0,1,\ldots,n-1$$

είναι ρίζες του $T_n(x)$. Κι επειδή το $T_n(x)$ είναι, σύμφωνα με το *Θεώρημα IV.1.7.3*, ένα πολυώνυμο *n* βαθμού, οι *n* αυτές ρίζες του $T_n(x)$ είναι απλές. ■

Θεώρημα IV.1.7.5. ***Το πολυώνυμο $T_n(x)$ του Chebyshev παίρνει ακρότατες τιμές στα $(n+1)$ σημεία:***

$$X_j^* = \cos\frac{j\pi}{n}, \qquad i = 0,1,\ldots,n.$$

Σε αυτά τα σημεία το $T_n(x)$ παίρνει εναλλασσόμενα τις τιμές -1 και 1.

Απόδειξη. Επειδή $(Arc\cos x)' = (1-x^2)^{-\frac{1}{2}}$, έχουμε:

$$T_n'(x) = \sin(n\ Arc\cos x)\ n\ (1-x^2)^{-\frac{1}{2}},$$

κι επομένως, επαληθεύεται άμεσα ότι για κάθε $j = 0,1,\ldots,n$, ισχύει ότι:

$$T_n'(X_j^*) = n\,[1-\cos^2\frac{j\,\pi}{n}]^{-\frac{1}{2}}\ \sin(j\,\pi) = 0,$$

που αποδεικνύει ότι το πολυώνυμο $T_n(x)$ παίρνει ακρότατες τιμές στα $(n+1)$ σημεία $X_0^*, X_1^*, \ldots, X_n^*$.

Επί πλέον, οι ακρότατες τιμές αυτές είναι -1 ή 1, γιατί:

$$T_n(X_j^*) = \cos(n\ Arc\cos\cos\frac{j\,\pi}{n}) = \cos(j\,\pi) = (-1)^j$$

για κάθε $j = 0,1,\ldots,n$.

Η τιμή $T_n(X_0^*)$ είναι 1. ■

Ας επιστρέψουμε τώρα στο κεντρικό πρόβλημα της *Παραγράφου*:

Επιθυμούμε να προσδιορίσουμε τα σημεία $x_0, x_1, \ldots, x_n \in [-1,1]$ που καθιστούν ελάχιστη την ποσότητα:

$$\sup_{x\in[-1,1]}|V_n(x)|,$$

όπου $V_n(x) = (x-x_0)(x-x_1)\ldots(x-x_n)$.

Ας συμβολίσουμε Π_{n+1} το σύνολο όλων των πολυωνύμων, με πραγματικούς συντελεστές και με πραγματικές τιμές, που έχουν βαθμό ακριβώς $n+1$ και που είναι τέτοια ώστε:

. ο συντελεστής του x^{n+1} να ισούται με 1, και

. όλες οι ρίζες τους να είναι απλές και να ανήκουν στο $[-1,1]$.

Με αυτήν την ορολογία, το κεντρικό πρόβλημα της παραγράφου μπορεί να επαναδιατυπωθεί έτσι ώστε να εντοπίζεται στην αναζήτηση ενός πολυωνύμου $V_n \in \Pi_{n+1}$ που να επαληθεύει την παρακάτω ανισότητα:

$$\sup_{x\in[-1,1]}|V_n(x)| \le \sup_{x\in[-1,1]}|p(x)|,$$

για κάθε $p \in \Pi_{n+1}$.

Συναφώς έχουμε το εξής:

Θεώρημα IV.1.7.6. ***Για κάθε*** $p \in \Pi_{n+1}$***, ισχύει:***

$$\frac{1}{2^n} = \frac{1}{2^n}\sup_{x\in[-1,1]}|T_{n+1}(x)| \le \sup_{x\in[-1,1]}|p(x)|.$$

Απόδειξη. Το πολυώνυμο:

$$\frac{T_{n+1}(x)}{2^n}$$

ανήκει στο σύνολο Π_{n+1}, γιατί ο συντελεστής του x^{n+1} ισούται, σύμφωνα με το *Θεώρημα IV.1.7.3*, με 1, αλλά και γιατί όλες οι ρίζες του είναι απλές και βρίσκονται στο διάστημα $[-1,1]$, σύμφωνα με το *Θεώρημα IV.1.7.4*.

Επί πλέον, η συνάρτηση:

$$\frac{|T_{n+1}(x)|}{2^n}$$

μεγιστοποιείται $(n+1)$ φορές στο διάστημα $[-1,1]$.

Πιο συγκεκριμένα, σύμφωνα με το *Θεώρημα IV.1.7.5*, η συνάρτηση αυτή μεγιστοποιείται στα σημεία:

$$X_j^* = \cos(\frac{j\,\pi}{n+1}), \quad j = 0,1,\ldots,n+1,$$

όπου παίρνει την μέγιστή της τιμή που ισούται με $\frac{1}{2^n}$, δηλαδή:

$$\sup_{x\in[-1,1]} \frac{|T_{n+1}(x)|}{2^n} = \frac{|T_{n+1}(X_j^*)|}{2^n} = \frac{1}{2^n}.$$

Ας υποθέσουμε τώρα ότι υπάρχει $p \in \Pi_{n+1}$, τέτοιο ώστε:

$$\sup_{x\in[-1,1]} |p(x)| < \frac{1}{2^n}.$$

Τότε, το πολυώνυμο:

$$r(x) = \frac{T_{n+1}(x)}{2^n} - p(x)$$

είναι ένα πολυώνυμο βαθμού το πολύ n, γιατί οι όροι με βαθμό $(n+1)$ αλληλοαφαιρούνται.

Ακόμη, οι αριθμοί:

$$r(X_j^*) = \frac{T_{n+1}(X_j^*)}{2^n} = \frac{(-1)^j}{2^n} - p(X_j^*), \quad j = 0,1,\ldots,n+1,$$

γίνονται εναλλασσόμενα θετικοί ή αρνητικοί, γιατί $|p(X_j^*)| < \frac{1}{2^n}$. Αυτό σημαίνει ότι υπάρχουν $(n+2)$ σημεία του $[-1,1]$, γύρω από τα οποία οι τιμές της συνάρτησης $r(x)$ αλλάζουν πρόσημο.

Άρα καθώς η $r(x)$ είναι συνεχής συνάρτηση, υπάρχουν $(n+1)$ διάφορες ρίζες της $r(x)$ στο $[-1,1]$.

Επειδή όμως η $r(x)$ είναι πολυώνυμο βαθμού το πολύ n, έπεται ότι:

$$r(x) \equiv 0,$$

ή ισοδύναμα ότι:

$$p(x) \equiv \frac{T_{n+1}(x)}{2^n}.$$

Επομένως, θα ισχύει:

$$\sup_{x\in[-1,1]}|p(x)| = \frac{1}{2^n} < \frac{1}{2^n},$$

που είναι προφανώς άτοπο, και συνεπώς η *Απόδειξη* είναι πλήρης. ■

Σαν άμεση συνέπεια αυτού του *Θεωρήματος*, παραθέτουμε το επόμενο αποτέλεσμα που απαντά στο πρόβλημα της βέλτιστης επιλογής του πολυωνύμου $V_n(x)$:

Θεώρημα IV.1.7.7. ***Η επιλογή των σημείων παρεμβολής x_i που ελαχιστοποιούν τον αριθμό***

$$\sup_{x\in[-1,1]}|V_n(x)| = \sup_{x\in[-1,1]}|(x-x_0)\ldots(x-x_n)|$$

είναι:

$$x_i = X_i = \cos\left(\frac{2i+1}{2n+2}\pi\right) \quad i = 0,1,\ldots n. ■$$

Ο συνδυασμός αυτού του Θεωρήματος με το *Θεώρημα IV.1.7.6* και το *Θεώρημα IV.1.6.1* οδηγεί στην ακόλουθη εκτίμηση του σφάλματος παρεμβολής:

$$E_n(x) = f(x) - p_n(x), \text{ όταν } V_n(x) = \frac{T_{n+1}(x)}{2^n}:$$

Πόρισμα IV.1.7.8. ***Έστω*** $f(x)$ ***μία πραγματική συνάρτηση, που είναι*** $(n+1)-$ ***φορές συνεχώς διαφορίσιμη μέσα σε μία ανοιχτή περιοχή του κλειστού διαστήματος*** $[-1,1]$.

Εάν $p_n(x)$ ***είναι το μοναδικό πολυώνυμο, βαθμού το πολύ*** n, ***που παρεμβάλλει την*** $f(x)$ ***στα*** $(n+1)$ ***σημεία:***

$$X_i = \cos\left(\frac{2i+1}{2n+2}\pi\right) \qquad i = 0,1,\ldots,n,$$

τότε:

$$|E_n(x)| \le \frac{1}{2^n(n+1)} \sup_{x\in[-1,1]} \left|\frac{d^{n+1}f}{dx^{n+1}}(x)\right|$$

για κάθε $x \in [-1,1]$.■

Παρατήρηση IV.1.7.9. Εάν η πραγματική συνάρτηση $f(x)$ είναι $(n+1)-$ φορές συνεχώς διαφορίσιμη μέσα σε μία ανοικτή περιοχή ενός κλειστού διαστήματος $[\alpha,\beta]$ (και όχι του $[-1,1]$), μπορούμε να μεταφερθούμε στο διάστημα $[-1,1]$ χρησιμοποιώντας τον ακόλουθο μετασχηματισμό αλλαγής μεταβλητής:

$$y = \frac{\beta-\alpha}{2}x + \frac{\beta+\alpha}{2},$$

όπου $y \in [\alpha,\beta]$ και $x \in [-1,1]$. ■

Η εκτίμηση του σφάλματος παρεμβολής, που δόθηκε στο *Πόρισμα IV.1.7.8*, αφορά μόνον την περίπτωση που τα σημεία της παρεμβολής είναι οι ρίζες κάποιου πολυωνύμου του Chebyshev.

Στην επόμενη *Παράγραφο*, εξετάζοντας το πρόβλημα της σύγκλισης μίας ακολουθίας παρεμβολής, θα δώσουμε έναν γενικότερο τύπο για την εκτίμηση του σφάλματος παρεμβολής, στην περίπτωση που τα σημεία της παρεμβολής συνιστούν μία οποιαδήποτε συλλογή, ανά δύο διάφορων, σημείων του κλειστού διαστήματος $[-1,1]$.

IV.1.8.ΣΥΓΚΛΙΣΗ ΑΚΟΛΟΥΘΙΩΝ ΠΟΛΥΩΝΥΜΩΝ ΠΑΡΕΜΒΟΛΗΣ

Το *Πρόβλημα* που θα συζητήσουμε στην παρούσα *Παράγραφο*, είναι το εξής:

"Έστω

$$M=(x_k^{(n)})_{n\geq 0,0\leq k\leq n}=\begin{pmatrix} x_0^{(0)} & & & & \\ x_0^{(1)} & x_1^{(1)} & & & \\ x_0^{(2)} & x_1^{(2)} & x_2^{(2)} & & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_0^{(n)} & x_1^{(n)} & x_2^{(n)} & \cdots & x_n^{(n)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

ένας άπειρος τριγωνικός πίνακας σημείων παρεμβολής $x_k^{(n)}\in[-1,1]$ ***με*** $x_k^{(n)}\neq x_m^{(n)}\quad \forall k\neq m$ ***. Ο πίνακας*** M ***καλείται σχήμα παρεμβολής. Κάθε γραμμή:***

$$M_n = \{x_0^{(n)}, x_1^{(n)}, \ldots, x_n^{(n)}\}$$

του πίνακα M ορίζει ένα σύνολο $(n+1)$ σημείων παρεμβολής. Δοθείσης μίας συνεχούς πραγματικής συνάρτησης $f(x)$, ορισμένης στο κλειστό διάστημα $[-1,1]$, θεωρούμε την ακολουθία των πολυωνύμων παρεμβολής:

$$p_n(x) \qquad (n \geq 0),$$

της $f(x)$, στα σημεία της κάθε γραμμής M_n του M.

Ζητείται να διερευνηθούν οι συνθήκες, που πρέπει να πληρούν τα σημεία του σχήματος παρεμβολής M, έτσι ώστε να συμβαίνει:

$$\lim_{n\to\infty} p_n(x) = f(x),$$

για κάθε x μέσα σε ένα (υπό προσδιορισμό) αρκετά μεγάλο υποσύνολο του $[-1,1]$. Σε ένα δεύτερο στάδιο, ζητείται να εξεταστεί η ανεξαρτησία των ικανών αυτών συνθηκών από την φύση της συνεχούς συνάρτησης $f(x)$".

Προκειμένου να αντιμετωπίσουμε το πρόβλημα αυτό χρειαζόμαστε κάποιο προπαρασκευαστικό υλικό.

Ας θεωρήσουμε το σύνολο:

$$C[-1,1] = \{f : [-1,1] \to \mathbb{R} : f \text{ είναι συνεχής }\}$$

όλων των συνεχών πραγματικών συναρτήσεων που είναι ορισμένες στο $[-1,1]$.

Το σύνολο αυτό, εφοδιασμένο με τις πράξεις ***πρόσθεση κατά σημείο***:

$$+ : (f+g)(x) := f(x) + g(x) \quad (f, g \in C[-1,1] \text{ και } x \in [-1,1])$$

και ***αριθμητικό πολλαπλασιασμό κατά σημείο***:

$$\cdot : (\lambda . f)(x) := \lambda \cdot f(x) \quad (\lambda \in \mathbb{R}, f \in C[-1,1] \text{ και } x \in [-1,1])$$

είναι ένας διανυσματικός χώρος υπεράνω του $\mathbb{R}$.

Στον διανυσματικό χώρο αυτό μπορούμε να ορίσουμε την ***νόρμα της ομοιόμορφης σύγκλισης***:

$$\|f\|_\infty := \sup_{x\in[-1,1]} |f(x)| \qquad \forall f \in C[-1,1].$$

(Παρατηρήστε πως, επειδή η f είναι συνεχής στο $[-1,1]$ κι επειδή το διάστημα $[-1,1]$ είναι κλειστό και φραγμένο υποσύνολο του $\mathbb{R}$, η ποσότητα $\sup_{x\in[-1,1]} |f(x)|$ είναι πραγματικός αριθμός και όχι ∞.)

Χωρίς απόδειξη παραθέτουμε το εξής, χρήσιμο για την συνέχεια, αποτέλεσμα:

Θεώρημα IV.1.8.1. ***Ο διανυσματικός χώρος $C[-1,1]$ είναι ένας χώρος Banach ως προς την νόρμα της ομοιόμορφης σύγκλισης .***■

Για κάθε $n = 0,1,2,\ldots$, θεωρούμε τώρα την απεικόνιση:

$$\ell_n^{(M)} : C[-1,1] \to C[-1,1] : f(x) \mapsto \ell_n^{(M)}(f)(x) = p_n(x),$$

που σε κάθε $f \in C[-1,1]$ αντιστοιχεί το πολυώνυμο $p_n(x)$, βαθμού το πολύ n, που παρεμβάλλει την f στα σημεία της γραμμής M_n του σχήματος παρεμβολής M .

Η απεικόνιση αυτή είναι $\mathbb{R}$ -γραμμική υπό την έννοια ότι:

$$[\ell_n^{(M)}(f+g)](x) = [\ \ell_n^{(M)}(f)](x) + [\ \ell_n^{(M)}(g)](x) \quad \forall x \in [-1,1]$$

και

$$[\ell_n^{(M)}(\lambda \cdot f)](x) = \lambda \cdot [\ \ell_n^{(M)}(f)](x) \quad \forall x \in [-1,1],$$

για κάθε $f, g \in C[-1,1]$ και για κάθε $\lambda \in \mathbb{R}$.

Λέμε ότι αυτή η απεικόνιση είναι ένας ***γραμμικός τελεστής υπεράνω του*** $\mathbb{R}$. Επί πλέον, η απεικόνιση $\ell_n^{(M)}$ είναι συνεχής, υπό την έννοια ότι εάν $f_i \in C[-1,1]$ και $f \in C[-1,1]$ τότε:

$$\lim_{i\to\infty} \|f_i - f\|_\infty = 0 \Rightarrow \lim_{i\to\infty} \|\ \ell_n^{(M)}(f_i) - \ell_n(f)\|_\infty = 0\,.$$

Μετά από όσα προηγήθηκαν, μπορούμε να επαναδιατυπώσουμε το πρόβλημα της *Παραγράφου* ως εξής:

"Ζητείται να διερευνηθούν οι συνθήκες, που πρέπει να πληρούν τα σημεία του σχήματος παρεμβολής M ***, έτσι ώστε να συμβαίνει:***

$$\lim_{n\to\infty} [\ \ell_n^{(M)}(f)](x) = f(x), \quad \forall f \in C[-1,1],$$

για κάθε x ***μέσα σε ένα (υπό προσδιορισμό) αρκετά μεγάλο υποσύνολο του*** $[-1,1]$***."***

Μπορούμε να προχωρήσουμε ένα βήμα περαιτέρω, θεωρώντας την ταυτοτική απεικόνιση του $C[-1,1]$:

$$id : C[-1,1] \to C[-1,1] : f(x) \mapsto id(f)(x) = f(x).$$

Πρόκειται για μία $\mathbb{R}$-γραμμική απεικόνιση (που, εξάλλου, ονομάζεται και ***ταυτοτικός τελεστής***), η οποία είναι προφανώς συνεχής.

Επειδή η έκφραση:

$$\lim_{n\to\infty}[\ell_n^{(M)}(f)](x) = f(x), \quad \forall f \in C[-1,1] \text{ και } \forall x \in [-1,1]$$

είναι ισοδύναμη με την έκφραση:

$$\lim_{n\to\infty} \ell_n^{(M)}(f) = id(f), \quad \forall f \in C[-1,1]$$

(όπου $\lim_{n\to\infty} \ell_n^{(M)}(f) = id(f) \Leftrightarrow \lim_{n\to\infty}\left\|\ell_n^{(M)}(f) - id(f)\right\|_\infty = 0)$, μία ισχυρότερη τοποθέτηση του κεντρικού προβλήματος δίνεται από την ακόλουθη εκφώνηση:

"Ζητείται να διερευνηθούν οι συνθήκες, που πρέπει να πληρούν τα σημεία του σχήματος παρεμβολής M, έτσι ώστε να συμβαίνει:

$$\lim_{n\to\infty} \ell_n^{(M)}(f) = id(f), \quad \forall f \in C[-1,1]."$$

Με αυτήν την διατύπωση, μπορούμε να δώσουμε απάντηση στο πρόβλημα της *Παραγράφου*.

Πράγματι, προς αυτήν την κατεύθυνση διαθέτουμε το επόμενο *Θεώρημα*, που είναι μία συνέπεια κλασσικού *Θεωρήματος Banach-Steinhauss*:

Θεώρημα IV.1.8.2. ***Έστω E ένας χώρος Banach. Έστω επίσης $(\ell_n : n = 0,1,2,\ldots)$ μία ακολουθία συνεχών γραμμικών τελεστών από το E στο E και έστω τέλος ένας συνεχής γραμμικός τελεστής $\ell : E \to E$.***

Εάν D είναι ένας διανυσματικός υπόχωρος του E, που είναι παντού πυκνός στο E, τότε μία αναγκαία και ικανή συνθήκη για να ισχύει:

$$\lim_{n\to\infty} \ell_n(f) = \ell(f) \quad \forall f \in E$$

είναι η ικανοποίηση του συστήματος των εξής δύο ιδιοτήτων:

$$\left\{ \begin{array}{c} \lim_{n\to\infty} \ell_n(f) = \ell(f) \quad \forall f \in D, \\ \|\ell_n\| := \sup\{|\ell_n f| \ : f \in E \ \text{και} \ |f| \le 1\} \le M < \infty \quad \forall n, \end{array} \right\}$$

όπου $|\cdots|$ είναι η νόρμα του χώρου Banach E. ■

Η υπόθεση του *Θεωρήματος* ότι *το σύνολο D είναι παντού πυκνό στο E* σημαίνει ότι για κάθε $f \in E$ και για κάθε $\varepsilon > 0$ υπάρχει $g \in D$ τέτοιο ώστε

$$|f - g| < \varepsilon.$$

Στην περίπτωση μας, ο χώρος Banach E είναι το $C[-1,1]$, και το πυκνό υποσύνολο του D είναι ο διανυσματικός χώρος $\boldsymbol{P}(\boldsymbol{R})$ όλων των πραγματικών πολυωνυμικών συναρτήσεων. (Υπενθυμίζουμε ότι σύμφωνα με το κλασσικό *Θεώρημα της Ομοιόμορφης Σύγκλισης του Weierstrass*, για κάθε $f \in C[-1,1]$ και για $\varepsilon > 0$ υπάρχει πολυώνυμο $p \in \boldsymbol{P}(\boldsymbol{R})$ τέτοιο ώστε $\|f - p\|_\infty < \varepsilon$.)

Προφανώς, στην περίπτωσή μας, η πρώτη από τις ιδιότητες του συστήματος του παραπάνω *Θεωρήματος* ικανοποιείται, γιατί:

$$p(x) \in \boldsymbol{P}_k(\boldsymbol{R}) \Rightarrow \ell_n^{(M)}(p) = p \quad \forall n \geq k\,,$$

σύμφωνα με την κατασκευή του πολυωνύμου παρεμβολής.

Αντίθετα, η δεύτερη ιδιότητα δεν ισχύει, καθώς επαληθεύεται ότι:

$$\left\|\ell_n^{(M)}\right\| = \sup_{x\in[-1,1]} \sum_{i=0}^{n} \left|L_i^{(n)}(x)\right|,$$

και, επί πλέον, έχουμε το ακόλουθο:

Θεώρημα IV.1.8.3. ***Για οποιαδήποτε επιλογή σημείων:***

$$M_n = \{x_0^{(n)}, x_1^{(n)}, \ldots, x_n^{(n)}\}\,,$$

ισχύει:

$$\sup_{x\in[-1,1]} \sum_{i=0}^{n} \left|L_i^{(n)}(x)\right| \geq \frac{2}{\pi^2}\,\ell n n + \beta(n)\,,$$

όπου $\beta(n)$ ***είναι μία φραγμένη συνάρτηση του*** n **.** ■

Κατά συνέπεια, δεν υπάρχει σύγκλιση των ακολουθιών πολυωνύμων παρεμβολής για κάθε συνάρτηση του $C[-1,1]$, όταν ο αριθμός των σημείων παρεμβολής αυξάνεται απεριόριστα. Με άλλα λόγια έχουμε το εξής αποτέλεσμα:

Θεώρημα IV.1.8.4. *Για οποιοδήποτε σχήμα παρεμβολής M, υπάρχει τουλάχιστον μία συνάρτηση $f \in C[-1,1]$ για την οποία η αντίστοιχη ακολουθία πολυωνύμων παρεμβολής*

$$(p_n(x) = [\ell_n^{(M)}(f)](x) : n = 0,1,\ldots)$$

της f δεν συγκλίνει προς την f, ως προς την νόρμα της ομοιόμορφης σύγκλισης.■

Παρ' όλα αυτά, μπορούμε να αποδείξουμε τα επόμενα δύο θετικά αποτελέσματα:

Θεώρημα IV.1.8.5. ([41]) *Για κάθε συνάρτηση $f \in C[-1,1]$, υπάρχει σχήμα παρεμβολής M τέτοιο ώστε η αντίστοιχη ακολουθία πολυωνύμων παρεμβολής*

$$(p_n(x) = [\ell_n^{(M)}(f)](x) : n = 0,1,\ldots)$$

της f να συγκλίνει προς την f, ως προς την νόρμα ομοιόμορφης σύγκλισης. ■

Θεώρημα IV.1.8.6. ([5]) *Εάν η συνάρτηση $f \in C[-1,1]$ είναι αναλυτική μέσα σε μία ανοικτή περιοχή του διαστήματος $[-1,1]$, τότε για οποιοδήποτε σχήμα παρεμβολής M, η αντίστοιχη ακολουθία πολυωνύμων παρεμβολής*

$$(p_n(x) = [\ell_n^{(M)}(f)](x) : n = 0,1,\ldots)$$

της f συγκλίνει προς την f, ως προς την νόρμα της ομοιόμορφης σύγκλισης:

$$\lim_{n\to\infty} \sup_{x\in[-1,1]} |f(x) - p_n(x)| = 0. \blacksquare$$

Μία ενστικτώδης αντίδραση του αναγνώστη μπορεί να προέλθει από την αμφισβήτηση των δυνατοτήτων της νόρμας της ομοιόμορφης σύγκλισης $\|\cdot\cdot\|_\infty$, με την οποία ο διανυσματικός χώρος $C[-1,1]$ είχε καταστεί ένας χώρος Banach.

Στην συνέχεια, θα σκιαγραφήσουμε κάποια αποτελέσματα, που μπορούν να προκύψουν με την αντικατάσταση της αξίωσης μίας σύγκλισης ως προς την νόρμα $\|\cdot\cdot\|_\infty$ με την αξίωση ως προς διαφορετικά κριτήρια.

Ένα από τα αποτελέσματα, που έχουν αναφερθεί πολλές φορές μέσα σε εργασίες που αφορούν την Θεωρία Προσεγγίσεων, εμφανίστηκε το 1937 στο περίφημο μαθηματικό περιοδικό "*Annals of Mathematics*" με συγγραφείς τους διάσημους μαθηματικούς *Paul Erdös* και *Paul Turán*. Αυτοί, οι νέοι, τότε, επιστήμονες απέδειξαν το αξιοσημείωτο θετικό αποτέλεσμα ότι ***για κάθε πραγματική και συνεχή συνάρτηση f του $[-1,1]$, η ακολουθία των πολυωνύμων παρεμβολής $p_n(x)$ της f συγκλίνει κατά τετραγωνικό μέσο προς την $f(x)$, εάν το σχήμα παρεμβολής M απαρτίζεται από τις ρίζες ενός συστήματος ορθογωνίων πολυωνύμων ως προς κάποια τυχαία συνάρτηση βάρους***.

Πιο συγκεκριμένα, οι Erdös και Turán έδειξαν το επόμενο *Θεώρημα*:

Θεώρημα IV.1.8.7. ***Έστω $w:[-1,1]\to\mathbb{R}:x\mapsto w(x)$ μία συνάρτηση βάρους στο διάστημα $[-1,1]$, δηλαδή μία συνάρτηση που είναι τέτοια ώστε:***

$$w(x)\geq 0 \quad \forall x\in[-1,1] \quad \textbf{\textit{και}} \quad \int_{-1}^{1} w(x)dx>0.$$

Έστω ακόμα μία ακολουθία ορθοκανονικών πολυωνύμων του $[-1,1]$, ως προς το

βάρος $w(x)$***:***

$$(Q_0(x), Q_1(x), \ldots, Q_n(x), \ldots, Q_m(x), \ldots) \quad (Q_j(x) \in \mathbb{P}_j(\mathbb{R})),$$

δηλαδή μία ακολουθία πολυωνύμων με την ιδιότητα:

$$\int_{-1}^{1} Q_m(x)\, Q_n(x)\, w(x)dx = \delta_{m,n} = \begin{Bmatrix} 1, & \varepsilon\acute{\alpha}\nu & m = n \\ 0, & \varepsilon\acute{\alpha}\nu & m \neq n \end{Bmatrix}$$

Εάν:

$$M_n = \{x_0^{(n)}, x_1^{(n)}, \ldots, x_n^{(n)}\}$$

είναι το σύνολο των απλών ριζών του πολυωνύμου $Q_n(x)$ ***και εάν:***

$$M_n \subset [-1,1],$$

τότε, για κάθε $f \in C[-1,1]$ ***το πολυώνυμο:***

$$[\ell_n^{(M)}(f,w)](x),$$

που παρεμβάλει την f ***στα σημεία του*** M_n***, επαληθεύει την ανισότητα:***

$$\int_{-1}^{1} \{f(x) - [\ell_n^{(M)}(f,w)](x)\}^2\, w(x)dx \leq \sqrt{6}\, \mathbb{E}_n(f),$$

όπου:

$$\mathbb{E}_n(f) := \min_{p \in \mathbb{P}_n(\mathbb{R})} \|f - p\|_\infty . \blacksquare$$

Επειδή, σύμφωνα με το κλασσικό *Θεώρημα της Ομοιόμορφης Σύγκλισης του Weierstrass*, για κάθε $f \in C[-1,1]$ και για κάθε $n = 1,2,\ldots,$ υπάρχει πολυώνυμο $p_n \in \mathbb{P}_n(\mathbb{R})$ τέτοιο ώστε:

$$\|f - p_n\|_\infty < \frac{1}{n},$$

βλέπουμε αμέσως, ότι:

Πόρισμα IV.1.8.8. ***Κάτω από τις υποθέσεις του προηγούμενου Θεωρήματος, εάν, για κάθε*** $n = 0,1,2,\ldots,$ ***το σύνολο:***

$$M_n = \{x_0^{(n)}, x_1^{(n)}, \ldots, x_n^{(n)}\}$$

των απλών ριζών του πολυωνύμου $Q(x)$ ***είναι υποσύνολο του*** $[-1,1]$***, τότε, για κάθε*** $f \in C[-1,1]$***, ακολουθία των πολυωνύμων:***

$$([\ell_n^{(M)}(f,w)](x) : n = 0,1,2,\ldots),$$

που παρεμβάλλουν την f ***στα σημεία των γραμμών του σχήματος:***

$$M = \begin{pmatrix} M_0 \\ M_1 \\ \vdots \end{pmatrix},$$

συγκλίνει κατά τετραγωνικό μέσο προς την f ***:***

$$\lim_{n\to\infty} \int_{-1}^{1} [f(x) - [\ell_n^{(M)}(f,w)](x)]^2 \, w(x)dx = 0. \blacksquare$$

Για να συνειδητοποιήσουμε την αξία αυτής της κατά τετραγωνικό μέσο σύγκλισης, αναφέρουμε ένα θεμελιώδες αρνητικό αποτέλεσμα του *G.Faber*, που ενισχύει το *Θεώρημα IV.1.8.4* και που δημοσιεύτηκε το 1914, σύμφωνα με το οποίο ***για οποιοδήποτε σχήμα παρεμβολής*** M***, υπάρχει τουλάχιστον μία συνάρτηση*** $f \in C[-1,1]$ ***για την οποία η αντίστοιχη ακολουθία πολυωνύμων παρεμβολής***

$$([\ell_n^{(M)}(f)](x): \quad n = 0,1,\ldots)$$

της f ***είναι τέτοια ώστε:***

$$\lim_{n\to\infty} \sup \left\| \ell_n^{(M)}(f) \right\|_\infty = \infty.$$

Μία εύλογη ερώτηση που απασχόλησε πολλούς μαθηματικούς ήταν η δυνατότητα αντικατάστασης του εκθέτη 2 στον βασικό τύπο του *Θεωρήματος IV.1.8.7* (ή του *Πορίσματος IV.1.8.8*) με έναν μεγαλύτερο εκθέτη.

Οι απαντήσεις που δόθηκαν αφορούσαν ειδικές περιπτώσεις σχημάτων παρεμβολής.

Εάν, για παράδειγμα, το σχήμα παρεμβολής M είναι ο *πίνακας Chebyshev*:

$$M = \begin{pmatrix} M_0 \\ M_1 \\ \vdots \end{pmatrix},$$

με:

$$M_n = \left\{ \cos\left(\frac{\pi}{2n+2}\right), \cos\left(\frac{3\pi}{2n+2}\right), \ldots, \cos\left(\frac{[2j+1]\pi}{2n+2}\right), \ldots, \cos\left(\frac{[2n+1]\pi}{2n+2}\right) \right\},$$

$(n \geq 0)$, τότε, όπως αποδείχτηκε το 1936 από τους *Erdös* και *Feldheim*, ισχύει

$$\lim_{n\to\infty} \int_{-1}^{1} \left| f(x) - [\ell_n^{(M)}(f)] \right|^p (1-x)^{-\frac{1}{2}} dx = 0$$

για κάθε $f \in C[-1,1]$ και για κάθε $p > 0$.

Η αντίστοιχη τριγωνομετρική περίπτωση εξετάστηκε από τον *J.Marcinkiewicz*.

Παρ' όλες όμως τις ενθαρρυντικές αυτές μερικές απαντήσεις, σαράντα περίπου χρόνια αργότερα αποδείχτηκε από τον *P.Nevai* ότι, γενικά, ο εκθέτης 2 του βασικού τύπου του *Θεωρήματος IV.1.8.7* (ή του *Πορίσματος IV.1.8.8*) δεν μπορεί να μεταβληθεί χωρίς να αλλοτριωθούν τα συμπεράσματα αυτών των αποτελεσμάτων.

Παρόμοιου τύπου προβλήματα έχουν μελετηθεί και από άλλους σύγχρονους μαθηματικούς, όπως οι *R. Askey*, *B. Della Vecchia*, *G. Freud*, *G. Mastroianni*, *B. Muckenhoupt*, *D. S. Lubinski*, *A.K. Varma*, και *Y.Xu*.

Εάν κανείς επιθυμεί να αναζητήσει κάποια γενική εκτίμηση της διαφοράς

$$\left| f(x) - [\ell_n^{(M)}(f)](x) \right| \quad (x \in [-1,1]),$$

δηλαδή του σφάλματος παρεμβολής τότε, πρέπει να σταθεί ιδιαίτερα στην ***Ανισότητα του Lebesgue*** (που αφορά όχι μόνον τις συνεχείς συναρτήσεις f):

$$E_n(x) = \left| f(x) - [\ell_n^{(M)}](x) \right| \leq \{ \mathcal{L}_n^{(M)}(x) + 1 \} \, \mathcal{E}_n(f) \quad (x \in [-1,1]),$$

όπου $\mathcal{L}_n^{(M)}$ είναι η $n-$*οστή* ***συνάρτηση Lebesgue***, που ορίζεται σαν το άθροισμα

$$\mathcal{L}_n^{(M)}(x) := \sum_{k=0}^{n} \left| \lambda_{n,k}^{(M)}(x) \right|,$$

με

$$\mathcal{L}_n^{(M)}(x) \in [\mathcal{P}_n(\mathcal{R}) - \mathcal{P}_{n-1}(\mathcal{R})]$$

το μοναδικό πολυώνυμο, που αντιστοιχεί στο σχήμα παρεμβολής M και που επαληθεύει την σχέση:

$$\lambda_{n,k}^{(M)}(x_n^{(j)}) = \delta_{k,j} = \begin{Bmatrix} 1, & \text{εάν } j = k \\ 0, & \text{εάν } j \neq k \end{Bmatrix}$$

Η παραπάνω Ανισότητα του Lebesgue καθιστά φανερό το γεγονός ότι τόσο η συνάρτηση Lebesgue $\mathcal{L}_n^{(M)}(x)$, όσο και η ***σταθερά Lebesgue***:

$$\Lambda_n^{(M)} := \| \mathrm{L}_n^{(M)} \|_\infty ,$$

διαδραματίζουν έναν σπουδαίο ρόλο στην έρευνα του ασυμπτωτικού χαρακτήρα μίας ακολουθίας πολυωνύμων παρεμβολής.

Συναφώς, παραθέτουμε ένα θεωρητικό, αλλά και ενδεικτικό αποτέλεσμα, που απέδειξε το 1958 ο *P. Erdös: Για κάθε σχήμα παρεμβολής*

$$M \subset [-1,1]$$

και για κάθε $\varepsilon > 0$ και $A > 0$, το μέτρο του συνόλου:

$$\{x \in [-1,1] : \mathrm{L}_n^{(M)}(x) \leq A,\ n \geq n_0(A,\varepsilon)\}$$

είναι μικρότερο από το ε.

Ανάλογου τύπου αποτελέσματα βρέθηκαν και από άλλους μαθηματικούς, όπως οι *G.Halàsz*, *D. Newman*, *J. Knoppenberger*, *J. Szabados*, *A.K Varma*, *P.Vértesi* και ο *Y.G Shi*.

Όλα αυτά τα αποτελέσματα αποσκοπούν, λίγο ή πολύ, στην αναζήτηση μίας όσο το δυνατόν καλύτερης συμπεριφοράς των συναρτήσεων Lebesgue

$$\mathrm{L}_n^{(M)}(x)$$

και άλλων παρόμοιων εκφράσεων, αλλά και στην επίτευξη πολύτιμων γενικεύσεων του *Θεωρήματος του Faber*, ή και άλλων προτάσεων, όπως το *Θεώρημα των Grünwald-Marcinkiewicz*.

Είναι φυσικό να θεωρήσουμε την ακολουθία των ελάχιστων σταθερών Lebesgue:

$$(\Lambda_n^* := \min_{M\subset[-1,1]} \Lambda_n^{(M)} : \quad n = 0,1,2,\ldots t).$$

Σε μία εργασία υψηλού επιπέδου, που δημοσιεύτηκε το 1961, ο *P. Erdös* βελτίωσε ένα σχεδόν ταυτόχρονο αποτέλεσμα του *P. Turán*, δείχνοντας ότι:

$$\left|\Lambda_n^* - \frac{2}{\pi}\,\ell nn\right| \le c \quad (n \ge n_0),$$

όμως ο περίφημος *ισχυρισμός των Bernstein-Erdos*, σχετικά με την ύπαρξη ενός βέλτιστου σχήματος παρεμβολής M^* για το οποίο ισχύει:

$$\Lambda_n^{(M^*)} = \Lambda_n^*,$$

αποδείχτηκε μόλις το 1978 από τους *T. Kilgore*, *C. de Boor*, *A. Pinkus* και *L. Brutman*.

Ας στρέψουμε τέλος την προσοχή μας στην περίπτωση όπου η συνάρτηση f είναι συνεχής στο κλειστό διάστημα $[-1,1]$ αλλά ικανοποιεί την γενικότερη *συνθήκη Lipschitz*:

$$|f(x) - f(y)| \le c\,|x-y|^a \qquad (x\in[-1,1], y\in[-1,1]),$$

για κάποια θετική σταθερά *c*.

Συμβολίζουμε με $Lip(a)$ το σύνολο όλων αυτών των συναρτήσεων. Για $f \in Lip(a)$ και $0 < a < 1,$ η Ανισότητα του Lebesgue συνεπάγεται την εκτίμηση:

$$\left\|f - [\ell_n^{(M)}(f)]\right\|_\infty \le \frac{c}{n^a}\,\Lambda_n^{(M)}.$$

Το 1955, οι *Erdös* και *Turán* οδηγήθηκαν στο ακόλουθο εντυπωσιακό αποτέλεσμα. *Ας υποθέσουμε ότι:*

$$\Lambda_n^{(M)} \sim n^{\beta} \quad (\beta > 0).$$

Τότε:

(i). $a > \beta \Rightarrow \lim_{n\to\infty} \left\| f - [\ell_n^{(M)}(f)] \right\|_\infty = 0, \ \forall f \in Lip(a)$.

(ii). $a < \dfrac{\beta}{\beta+2} \Rightarrow \exists f \in Lip(a) : \lim_{n\to\infty} \sup \left\| [\ell_n^{(M)}(f)] \right\|_\infty = \infty$.

(iii). *Εάν* $\dfrac{\beta}{\beta+2} < a < \beta$, *τότε τόσο η σύγκλιση όσο και η απόκλιση της ακολουθίας*

$$([\ell_n^{(M)} f](x) : n = 0,1,2,...)$$

είναι ενδεχόμενη.

Η τελευταία διαπίστωση αυτού του αποτελέσματος σημαίνει ότι σε περίπτωση όπου είναι $\dfrac{\beta}{\beta+2} < a < \beta$ η ασυμπτωτική συμπεριφορά της ακολουθίας $([\ell_n^{(M)}(f)](x) : n = 0,1,2,\ldots)$ των πολυωνύμων παρεμβολής δεν μπορεί να προσδιοριστεί μόνον από την τάξη της σταθεράς του Lebesgue $\Lambda_n^{(M)}$. Τότε, θα πρέπει να εξετάσουμε συστηματικότερα τις δυνατότητες του σχήματος παρεμβολής M.

IV.1.9. ΠΟΛΥΩΝΥΜΟ ΠΑΡΕΜΒΟΛΗΣ HERMITE

Το πρόβλημα με το οποίο θα ασχοληθούμε σε αυτήν την *Παράγραφο* είναι το εξής:

"Έστω f μία απεικόνιση ορισμένη μέσα σε ένα υποσύνολο U του K και τέτοια ώστε να παίρνει τιμές μέσα στο K ($K = R, C$). Έστω ακόμη ότι η f είναι διαφορίσιμη μέσα σε ανοικτή περιοχή του U και έστω ότι η παράγωγός της συμβολίζεται με f'. Υποθέτουμε ότι γνωρίζουμε τις τιμές:

$$f(z_i) \text{ και } f'(z_i)$$

των f και f' αντίστοιχα σε $(n+1)$ σημεία $z_0, z_1, \ldots, z_n$ του U και ότι υπάρχουν τουλάχιστον δύο σημεία z_k και z_m ανάμεσα στα $z_0, z_1, \ldots, z_n$ τέτοια ώστε:

$$f(z_k) \neq 0 \text{ και } f'(z_m) \neq 0.$$

Επιθυμούμε να βρούμε ένα πολυώνυμο:

$$q : K \to K : z \mapsto q(z),$$

που να χαρακτηρίζεται από τις επόμενες δύο ιδιότητες:

το $q(z)$ να παίρνει τις ίδιες τιμές με $f(z)$ στα σημεία $z_0, z_1, \ldots, z_n$:

$$q(z_i) = f(z_i) \text{ για κάθε } i = 0,1,\ldots,n,$$

η παράγωγος $q'(z)$ του $q(z)$ να παίρνει τις ίδιες τιμές με την παράγωγο $f'(z)$ της $f(z)$ στα σημεία $z_0, z_1, \ldots, z_n$:

$$q'(z_i) = f'(z_i) \text{ για κάθε } i = 0,1,\ldots,n."$$

Το παραπάνω πρόβλημα ονομάζεται ***πρόβλημα παρεμβολής του Hermite***. Έχουμε το:

Θεώρημα IV.1.9.1. ***Εάν τα σημεία*** $z_0, z_1, \ldots, z_n$ ***είναι όλα ανά δύο διάφορα μεταξύ τους, τότε υπάρχει ένα και μόνον ένα πολυώνυμο*** $q(z) \in \boldsymbol{P}_{2.n+1}(\boldsymbol{K})$ ***που επιλύει το πρόβλημα παρεμβολής του Hermite.***■

Ορισμός IV.1.9.2. **Στην συνέχεια, *θα συμβολίζουμε με:***

$$q_{2n+1}(z)$$

το μοναδικό πολυώνυμο που επιλύει το πρόβλημα παρεμβολής του Hermite*, ακριβώς για να τονίζουμε ότι ο βαθμός του είναι το πολύ ίσος με** $2n+1$**. Επίσης, θα λέμε ότι *το $q_{2n+1}(z)$ ***είναι το πολυώνυμο παρεμβολής του Hermite της*** $f(z)$**.**■

Απόδειξη του Θεωρήματος IV.1.9.1. Σε ένα πρώτο στάδιο θα αποδείξουμε την μοναδικότητα του πολυωνύμου παρεμβολής του Hermite. Η απόδειξη της ύπαρξής του θα δοθεί σε δεύτερο στάδιο χρησιμοποιώντας μία κατασκευαστική μέθοδο.

Ας υποθέσουμε ότι υπάρχουν δύο πολυώνυμα παρεμβολής του Hermite της $f(z)$:

$$q_{2n+1}(z) \in \boldsymbol{P}_{2n+1}(\boldsymbol{K}) \text{ και } p_{2n+1}(z) \in \boldsymbol{P}_{2n+1}(\boldsymbol{K}).$$

Τότε, εάν συμβολίσουμε με $r_{2n+1}(z)$ την διαφορά $q_{2n+1}(z) - p_{2n+1}(z)$:

$$r_{2n+1}(z) = q_{2n+1}(z) - p_{2n+1}(z),$$

παρατηρούμε ότι το πολυώνυμο $r_{2n+1}(z) \in \boldsymbol{P}_{2n+1}(\boldsymbol{K})$ έχει $(2n+1)$ διπλές ρίζες γιατί

$$q_{2n+1}(z_i) = p_{2n+1}(z_i) = f(z_i) \text{ και } q'_{2n+1}(z_i) = p'_{2n+1}(z_i) = f'(z_i) \quad \forall i = 0,1,\ldots,n.$$

Καθώς, ο βαθμός του πολυωνύμου $r_{2n+1}(z)$ είναι το πολύ ίσος με $(2n+1)$, βλέπουμε ότι απαραίτητα πρέπει να ισχύει $r_{2n+1}(z) \equiv 0$.

Κατασκευάζουμε το πολυώνυμο παρεμβολής του Hermite με την βοήθεια του ακόλουθου τύπου:

$$q_{2n+1}(z) = \sum_{i=0}^{n} H_i^{(n)}(z)\, f(z_i) + \sum_{i=0}^{n} G_i^{(n)}(z)\, f'(z_i) \quad ,$$

όπου:

$$H_i^{(n)}(z) := \{1 - 2(z - z_i)\,[L_i^{(n)}]'(z_i)\}\,[L_i^{(n)}]^2(z) \text{ και } G_i^{(n)}(z) := (z - z_i)[L_i^{(n)}]^2(z),$$

για κάθε $i = 0,1,\ldots,n$.

Οι συναρτήσεις $L_i^{(n)}$ είναι αυτές που υπεισέρχονται στον τύπο παρεμβολής του Lagrange, όπως αυτός παρουσιάστηκε στην *Παράγραφο IV.1.2*, δηλαδή:

$$L_i^{(n)}(z) = \prod_{\substack{j=0 \\ j \neq i}}^{n} \frac{z - z_j}{z_i - z_j}.$$

Ας δείξουμε τώρα ότι το πολυώνυμο $q_{2n+1}(z)$, που κατασκευάστηκε από τον παραπάνω τύπο, είναι πράγματι το πολυώνυμο που επιλύει το πρόβλημα παρεμβολής του Hermite.

Κατ' αρχήν παρατηρούμε πως:

$$q_{2n+1}(z) \in \boldsymbol{P}_{2n+1}(\boldsymbol{K}),$$

επειδή $L_i^{(n)}(z) \in \boldsymbol{P}_n(\boldsymbol{K})$, οπότε $[L_i^{(n)}]^2(z) \in \boldsymbol{P}_{2n}(\boldsymbol{K})$ και συνεπώς

$$H_i^{(n)}(z) \in \boldsymbol{P}_{2n+1}(\boldsymbol{K}) \text{ και } G_i^{(n)}(z) \in \boldsymbol{P}_{2n+1}(\boldsymbol{K}).$$

Ύστερα, για κάθε $k = 0,1,\ldots,n$, ισχύει:

$$q_{2n+1}(z_k) = H_k^{(n)}(z_k)\, f(z_k) + G_k^{(n)}(z_k)\, f'(z_k),$$

επειδή:

$$L_i^{(n)}(z_k) = \delta_{i,k} = \begin{Bmatrix} 1, & \text{όταν } i = k \\ 0, & \text{όταν } i \neq k \end{Bmatrix}.$$

Από την ισότητα αυτή συνεπάγεται ότι:

$$q_{2n+1}(z_k) = \{1 - 2(z_k - z_k)\,[L_k^{(n)}]'(z_k)\}\, f(z_k) + (z_k - z_k)\, f'(z_k) = f(z_k).$$

Από την άλλη πλευρά, για κάθε $k = 0,1,\ldots,n$, έχουμε:

$$q_{2n+1}(z) = \sum_{i=0}^{n} [H_i^{(n)}]'(z)\;\; f(z_i) + \sum_{i=0}^{n} [G_i^{(n)}]'(z)\, f'(z_i).$$

Δεδομένου όμως ότι:

$$[H_i^{(n)}]'(z) = 2\,[L_i^{(n)}]'(z)\,\{1 - 2.(z - z_j)\,[L_i^{(n)}]'(z_i)\} - 2[L_i^{(n)}]'(z_i)\,[L_i^{(n)}]^2(z)$$

και ότι:

$$[G_i^{(n)}]'(z) = [L_i^{(n)}]^2(z) + 2\,(z - z_i)\,[L_i^{(n)}]'(z)\, L_i^{(n)}(z),$$

με συνέπεια να ισχύει:

$$[H_i^{(n)}]'(z_k) = 0 \text{ και } [G_i^{(n)}]'(z_k) = \delta_{i,k} = \begin{Bmatrix} 1, & \text{όταν } i = k \\ 0, & \text{όταν } i \neq k \end{Bmatrix}$$

από την παραπάνω ισότητα έπεται ότι:

$$q'_{2n+1}(z_k) = f'(z_k),$$

που ολοκληρώνει την *Απόδειξη* του *Θεωρήματος*. ■

Η μελέτη του σφάλματος:

$$E_{2n+1}(z) = f(z) - q_{2n+1}(z)$$

της παρεμβολής Hermite πραγματοποιείται με τρόπο ευθέως ανάλογο του τρόπου που αναπτύχθηκε στην *Παράγραφο IV.1.6*.

Έτσι, θα περιοριστούμε στην περίπτωση που

$$K = R, \quad U = [-1,1]$$

και

$$x_0 = z_0, x_1 = z_1, \ldots, x_n = z_n .$$

Έστω λοιπόν $x \in [-1,1]$, με x σταθερό και έστω I το κλειστό διάστημα:

$$I := [\min\{x, x_0, x_1, \ldots, x_n\}, \max\{x, x_0, x_1, \ldots, x_n\}] .$$

Θεώρημα IV.1.9.3. ***Εάν η συνάρτηση f είναι $(2n+1)-$φορές συνεχώς διαφορίσιμη, μέσα σε μία ανοικτή περιοχή του διαστήματος I, τότε υπάρχει ένα σημείο $\xi \in I$ τέτοιο ώστε:***

$$E_{2.n+1}(x) = \frac{[V_n(x)]^2}{(2n+2)!} \frac{d^{2n+2} f}{dx^{2n+2}}(\xi),$$

όπου $V_n(x) = (x - x_0)(x - x_1)\ldots(x - x_n)$.

Απόδειξη. Θέτουμε:

$$F(t) := f(t) - q_{2n+1}(t) - c\,[V_n(t)]^2 \quad (-1 \le t \le 1),$$

όπου c είναι ο σταθερός αριθμός που δίνεται από την σχέση:

$$c = \frac{f(x) - q_{2n+1}(x)}{[V_n(x)]^2}.$$

Είναι εύκολο να διαπιστώσουμε πως η συνάρτηση F έχει τις εξής ιδιότητες:

. $F(x_i) = 0 \quad \forall i = 0,1,\ldots,n$,

. $F(x) = 0$ (σύμφωνα με τον ορισμό του αριθμού c), και

. $F'(x_i) = 0 \quad \forall i = 0,1,\ldots,n$.

Άρα, η F έχει τουλάχιστον $(2n+3)$ ρίζες μέσα στο διάστημα I. Με διαδοχικές εφαρμογές του *Θεωρήματος Rolle*, συμπεραίνουμε ότι η $(2n+2)-$τάξης παράγωγος:

$$\frac{d^{2n+2}F}{dt^{2n+2}}$$

της F έχει τουλάχιστον μία ρίζα στο ξ. Εάν συμβολίσουμε με ξ αυτήν την ρίζα, τότε επειδή:

$$\frac{d^{2n+2}F}{dt^{2n+2}}(t) = \frac{d^{2n+2}f}{dt^{2n+2}}(t) - \frac{d^{2n+2}q_{2n+1}}{dt^{2n+2}}(t) - c\,\frac{d^{2n+2}[V_n]^2}{dt^{2n+2}}(t),$$

και επί πλέον επειδή:

$$\frac{d^{2n+2}q_{2n+1}}{dt^{2n+2}}(t) = 0 \quad \text{και} \quad \frac{d^{2n+2}[V_n]^2}{dt^{2n+2}}(t) = (2n+2)!,$$

θα είναι:

$$0 = \frac{d^{2n+2}F}{dt^{2n+2}}(\xi) = \frac{d^{2n+2}f}{dt^{2n+2}}(\xi) - c\,(2n+2)!,$$

απ' όπου συνεπάγεται η σχέση:

$$c = \frac{f(x) - q_{2n+1}(x)}{[V_n(x)]^2} = \frac{1}{(2n+2)!}\,\frac{d^{2n+2}f}{dt^{2n+2}}(\xi),$$

από την οποία έπεται αμέσως ο ισχυρισμός του *Θεωρήματος*. ■

Παρατήρηση IV.1.9.4. Μπορούμε να γενικεύσουμε την παρεμβολή Hermite και στην περίπτωση που οι ποσότητες:

$$\frac{d^k f}{dx^k}(x_i) \quad i = 0,1,\ldots,n$$

είναι γνωστές (για κάποιο $k = 0,1,2,\ldots,m$).

Εάν, για παράδειγμα, $n = 0$ και $m \neq 0$, μπορούμε να βρούμε ένα πολυώνυμο $p \in \boldsymbol{P}_m(\boldsymbol{R})$, τέτοιο ώστε:

$$p(x_0) = f(x_0),$$

$$p'(x_0) = f'(x_0),$$

.....

$$\frac{d^m p}{dx^m}(x_0) = \frac{d^m f}{dx^m}(x_0).$$

Πρόκειται για την ***παρεμβολή Taylor***, όπου το πολυώνυμο $p(x)$ δίνεται από τον *Τύπο*:

$$p(x) = f(x_0) + (x - x_0)\, f'(x_0) + \ldots + \frac{(x - x_0)^m}{m!}\, \frac{d^m f}{dx^m}(x_0).$$

Εάν η συνάρτηση f είναι $(m+1)-$ φορές συνεχώς διαφορίσιμη, μέσα σε μία ανοικτή περιοχή του κλειστού διαστήματος:

$$\boldsymbol{I} = [\min\{x, x_0, x_1, \ldots, x_n\}, \max\{x, x_0, x_1, \ldots, x_n\}],$$

τότε υπάρχει ένας αριθμός $\xi \in \boldsymbol{I}$ τέτοιος ώστε:

$$E_m(x) = f(x) - p(x) = \frac{(x - x_0)^{m+1}}{(m+1)!}\, \frac{d^{m+1} f}{dx^{m+1}}(\xi). \blacksquare$$

IV.1.10. ΓΕΝΙΚΗ ΔΙΑΤΥΠΩΣΗ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ ΤΗΣ ΠΑΡΕΜΒΟΛΗΣ

Οι διάφοροι τύποι παρεμβολής που συναντήσαμε μέχρι τώρα (όπως πολυωνυμική παρεμβολή, παρεμβολή Hermite, παρεμβολή Taylor), αλλά και άλλοι τύποι παρεμβολών, μπορούν να συγχωνευτούν μέσα σε μία μοναδική και γενική διατύπωση. Στην παρούσα *Παράγραφο* δεν θα παρουσιάσουμε παρά μόνον τις βασικές γραμμές αυτής της συγκεντρωτικής διατύπωσης. Για περισσότερες λεπτομέρειες, ο ενδιαφερόμενος αναγνώστης μπορεί να ανατρέξει στην παραπομπή $[13]$.

Έστω $V=(V,+,\cdot)$ ένας διανυσματικός χώρος, διάστασης $(n+1)$, υπεράνω του $\boldsymbol{K}$ ($\boldsymbol{K}=\boldsymbol{R}$, $\boldsymbol{C}$).

Συμβολίζουμε με

$$V^*$$

τον δυϊκό διανυσματικό χώρο του V , δηλαδή το σύνολο όλων των γραμμικών μορφών του V :

$$V^* = \{\Lambda : V \to \boldsymbol{K} : \Lambda \text{ είναι } \boldsymbol{K} \text{ } -\text{γραμμική απεικόνιση } \},$$

εφοδιασμένο με τις κατά σημείο πράξεις:

. πρόσθεση +: $V^* \times V^* \to V^* : (\Lambda, \mathrm{M}) \mapsto \Lambda + \mathrm{M}$, όπου

$$(\Lambda + \mathrm{M})(u) := \Lambda(u) + \mathrm{M}(u) \quad \forall u \in V ,$$

. αριθμητικό πολλαπλασιασμό $\cdot$: $\boldsymbol{K} \times V^* \to V^* : (\lambda, \Lambda) \mapsto \lambda \cdot \Lambda$, όπου

$$(\lambda \cdot \Lambda)(u) := \lambda \cdot \Lambda(u) \quad \forall u \in V .$$

Έστω $B^* = \{\Lambda_0, \Lambda_1, \ldots, \Lambda_n\}$ ένα σύνολο $(n+1)$ στοιχείων του V^* και έστω $\{w_0, w_1, \ldots, w_n\}$ ένα σύνολο $(n+1)$ στοιχείων του K, όχι όλων ίσων με μηδέν. Το γενικό πρόβλημα της παρεμβολής μπορεί να διατυπωθεί όπως ακολούθως:

Πρόβλημα IV.1.10.1. ***"Είναι δυνατή η εύρεση ενός στοιχείου $p \in V$ τέτοιου ώστε:***

$$\Lambda_i(p) = w_i \quad \forall i = 0,1,\ldots,n;"$$

Ας επιχειρήσουμε να δώσουμε μία απάντηση στο *Πρόβλημα* αυτό.

Έστω B_n μία βάση του V:

$$B_n = \{p_0, p_1, \ldots, p_n\}.$$

Ονομάζουμε ***ορίζουσα Gram*** του ζεύγους (B_n, B_n^*) την ορίζουσα του πίνακα του οποίου τα στοιχεία είναι οι αριθμοί $\Lambda_i(p_j) \in K$, όπου i είναι ο δείκτης που αντιπροσωπεύει την γραμμή του πίνακα και j είναι ο δείκτης που αντιπροσωπεύει την στήλη του πίνακα:

$$\det(\Lambda_i(p_j))_{0 \le i,j \le n} = \det\begin{pmatrix} \Lambda_0(p_0) & \Lambda_0(p_1) & \ldots & \Lambda_0(p_n) \\ \Lambda_1(p_0) & \Lambda_1(p_1) & \ldots & \Lambda_1(p_n) \\ \ldots & \ldots & \ldots & \ldots \\ \Lambda_n(p_0) & \Lambda_n(p_1) & \ldots & \Lambda_n(p_n) \end{pmatrix}.$$

Εάν τώρα $p \in V$ είναι ένα στοιχείο που επιλύει το *Πρόβλημα IV.1.10.1*, τότε το σύνολο B_n συνιστά μία βάση του V, θα είναι:

$$p = \sum_{j=0}^{n} a_j \cdot p_j$$

και, κατά συνέπεια, θα ισχύει:

$$\Lambda_i(p) = \sum_{j=0}^{n} a_j \ \Lambda_i(p_j) = w_i \qquad \forall i = 0,1,\ldots,n\,.$$

Με άλλα λόγια, το ζήτημα του προσδιορισμού του στοιχείου $p \in V$ που επιλύει το *Πρόβλημα IV.1.10.1* ανάγεται στο ζήτημα του προσδιορισμού των αριθμών:

$$a_0, a_1, \ldots, a_n\,.$$

Έχουμε, δηλαδή αποδείξει το εξής:

Θεώρημα IV.1.10.2. ***Μία αναγκαία και ικανή συνθήκη προκειμένου το γενικό πρόβλημα της παρεμβολής να έχει μία και μοναδική λύση, για κάθε σύνολο*** $(n+1)$ ***στοιχείων του*** K ***, όχι όλων ίσων με μηδέν, είναι η ορίζουσα Gram του ζεύγους*** (B_n, B_n^*) ***να είναι διάφορη του μηδενός.*** ■

Υπάρχει το ακόλουθο αποτέλεσμα, που αναφέρεται στο κριτήριο του μη εκφυλισμού της ορίζουσας Gram:

Θεώρημα IV.1.10.3. ***Μία αναγκαία και ικανή συνθήκη προκειμένου η ορίζουσα Gram του ζεύγους*** (B_n, B_n^*) ***να είναι διάφορη του μηδενός είναι το σύνολο*** B_n^* ***να συνιστά μία βάση του*** V^* ***.***

Ορισμός IV.1.10.4. ***Εάν το σύνολο*** B_n^* ***αποτελεί μία βάση του*** V^****, τότε λέμε ότι το σύνολο*** B_n^* ***έχει την ιδιότητα της παρεμβολής ή ότι το σύνολο*** B_n^* ***είναι ένα σύστημα Chebyshev.*** ■

Απόδειξη του Θεωρήματος IV.1.10.3. Ας θεωρήσουμε το σύστημα των εξισώσεων:

$$\left.\begin{cases} a_0\ \Lambda_0(p_0)+a_1\ \Lambda_1(p_0)+\ldots+a_n\ \Lambda_n(p_0)=0 \\ a_0\ \Lambda_0(p_1)+a_1\ \Lambda_1(p_1)+\ldots+a_n\ \Lambda_n(p_1)=0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_0\ \Lambda_0(p_n)+a_1\ \Lambda_1(p_n)+\ldots+a_n\ \Lambda_n(p_n)=0 \end{cases}\right\} \Leftrightarrow$$

$$\Leftrightarrow (a_0\ \Lambda_0+a_1\ \Lambda_1+\ldots+a_n\ \Lambda_n)(p_i)=0 \qquad \forall i=0,1,\ldots,n.$$

Επειδή το σύνολο $B_n=\{p_0,p_1,\ldots,p_n\}$ είναι μία βάση του διανυσματικού χώρου V, έπεται ότι:

$$(a_0\cdot\Lambda_0+a_1\cdot\Lambda_1+\ldots+a_n\cdot\Lambda_n)(f)=0 \qquad \forall f\in V,$$

που σημαίνει ότι:

$$a_0\cdot\Lambda_0+a_1\cdot\Lambda_1+\ldots+a_n\cdot\Lambda_n=0\ .$$

Εάν η ορίζουσα Gram του ζεύγους (B_n,B_n^*) είναι διάφορη του μηδενός, τότε το σύστημα δεν επιδέχεται παρά μόνον την μηδενική λύση και επομένως το σύνολο B_n^* συνιστά μία βάση του V^*.

Αντιστρόφως, εάν το σύνολο B_n^* είναι βάση του V^* και εάν η ορίζουσα είναι μηδενική τότε το σύστημα επιδέχεται μη μηδενική λύση $(a_0,a_1,\ldots,a_n)$ τέτοια ώστε

$$a_0\,\Lambda_0+a_1\,\Lambda_1+\ldots+a_n\,\Lambda_n=0\,.$$

Αλλά τούτο είναι αδύνατον, και συνεπώς η ορίζουσα Gram πρέπει να είναι διάφορη του μηδενός. ■

Παραδείγματα IV.1.10.5. Θα ολοκληρώσουμε την ανάπτυξη της *Παραγράφου* με την ενδεικτική επαλήθευση των όσων προηγήθηκαν.

(i). ***Πολυωνυμική παρεμβολή*:** **Το γενικό πρόβλημα της παρεμβολής εξειδικεύεται στο πρόβλημα της πολυωνυμικής παρεμβολής όταν:**

. $V = \boldsymbol{P}_n(\boldsymbol{K})$,

. $B_n^* = \{\Lambda_0(f) = f(z_0), \Lambda_1(f) = f(z_1), \ldots, \Lambda_n(f) = f(z_n)\}$

με $z_i \neq z_j$ **για** $i \neq j$**, όπου** $\Lambda_i(f) = f(z_i)$ **αναπαριστά την γραμμική μορφή η οποία αντιστοιχίζει σε κάθε πολυώνυμο** $p \in \boldsymbol{P}_n(\boldsymbol{K})$ **τον αριθμό**

$$\Lambda_i(f) = f(z_i).$$

(ii). ***Παρεμβολή Hermite:*** **Το γενικό πρόβλημα της παρεμβολής περιορίζεται στο πρόβλημα της παρεμβολής Hermite όταν:**

. $V = \boldsymbol{P}_{2n+1}(\boldsymbol{K})$,

. $$B_n^* = \{\Lambda_0(f) = f(z_0), \ldots, \Lambda_n(f) = f(z_n), \Lambda_{n+1}(f) = f'(z_0), \ldots, \Lambda_{2n+1}(f) = f'(z_n)\}$$

με $z_i \neq z_j$ **για** $i \neq j$.

(iii). ***Παρεμβολή Taylor*:** **Επιτυγχάνεται όταν:**

. $V = \boldsymbol{P}_n(\boldsymbol{K})$,

. $B_n^* = \{\Lambda_0(f) = f(z_0), \ldots, \Lambda_n(f) = \dfrac{d^n f}{dz^n}(z_0)\}$.■

IV.1.11. ΠΑΡΕΜΒΟΛΗ ΜΕ ΡΗΤΗ ΣΥΝΑΡΤΗΣΗ

Έστω f μία απεικόνιση ορισμένη μέσα σε ένα υποσύνολο U του K και τέτοια ώστε να παίρνει τιμές μέσα στο K ($K = R, C$).

Υποθέτουμε ότι γνωρίζουμε τις τιμές:

$$f(z_i)$$

της f σε $(m+n+1)$ σημεία $z_0, z_1, \ldots, z_m, z_{m+1}, \ldots, z_{m+n}$ του U.

Επιθυμούμε την εύρεση μίας *ρητής συνάρτησης*:

$$R_{m,n}(z) = \frac{P_m(z)}{Q_n(z)} = \frac{a_0 + a_1\ z + \ldots + a_m\ z^m}{\beta_0 + \beta_1\ z + \ldots + \beta_n\ z^n},$$

όπου

$$P_m(z) = a_0 + a_1\ z + \ldots + a_m\ z^m \text{ και } Q_n(z) = \beta_0 + \beta_1\ z + \ldots \beta_n\ z^n$$

είναι πολυώνυμα βαθμών m και n αντίστοιχα, που να έχει την χαρακτηριστική ιδιότητα να παίρνει τις ίδιες τιμές με την απεικόνιση f στα σημεία $z_0, \ldots, z_m, \ldots, z_{m+n}$:

$$R_{m,n}(z_i) = f(z_i) \quad \text{για κάθε } i = 0,1,\ldots,m+n.$$

Μία τέτοια ρητή συνάρτηση, εφόσον υπάρχει, καλείται ***ρητή παρεμβολή τύπου*** (m,n) της f. Λέμε επίσης ότι ***η ρητή συνάρτηση*** $R_{m,n}(z)$ ***παρεμβάλλει την*** $f(z)$ (στα σημεία $z_0, z_1 \ldots, z_m, z_{m+1}, \ldots, z_{m+n}$).

Όπως θα δούμε στην συνέχεια, μία ρητή συνάρτηση παρεμβολής $R_{m,n}(z)$, σε σύγκριση με το πολυώνυμο παρεμβολής $p_{m+n}(z)$ της $f(z)$, έχει ένα σοβαρό και απρόβλεπτο πλεονέκτημα, το οποίο έγκειται στο γεγονός ότι ο υπολογισμός της τιμής $R_{m,n}(z)$, για κάποιο δεδομένο z, δεν προϋποθέτει την εκτέλεση $(m+n)$ προσθέσεων, $(m+n-1)$ πολλαπλασιασμών, και μίας διαίρεσης (όπως φαίνεται με μία πρώτη ματιά).

Αντιθέτως, αφού μετασχηματίσουμε την $R_{m,n}(z)$ σε ένα ***συνεχές κλάσμα***:

$$R_{m,n}(z) = r_1(z) + \cfrac{c_2}{r_2(z) + \cfrac{c_3}{r_3(z) + \ddots + \cfrac{c_k}{r_k(z)}}},$$

όπου $r_j(z)$ συμβολίζει κάποιο πολυώνυμο, μπορούμε να διαπιστώσουμε πως ο υπολογισμός της τιμής $R_{m,n}(z)$ προϋποθέτει την (σημαντικά μειωμένη σε αριθμό) εκτέλεση ***μόνον*** m ή n "μακροσκελών " αριθμητικών πράξεων.

Με τον όρο " μακροσκελείς" αριθμητικές πράξεις εννοούμε πολλαπλασιασμούς είτε διαιρέσεις. Γενικότερα, έχουμε το εξής:

Θεώρημα IV.1.11.1. ***Κάθε ρητή συνάρτηση*** $R_{m,n}(z)$ ***μπορεί να εκφρασθεί υπό την μορφή συνεχούς κλάσματος:***

$$(\Sigma\mathrm{K}) \quad R_{m,n}(z) = r_1(z) + \cfrac{c_2}{r_2(z) + \cfrac{c_3}{r_3(z) + \ddots + \cfrac{c_k}{r_k(z)}}},$$

όπου $r_j(z)$ ***συμβολίζει κάποιο πολυώνυμο. Από την έκφραση αυτή, η*** $R_{m,n}(z)$ ***μπορεί να υπολογισθεί για κάθε*** z***, αφού πρώτα εκτελεστούν το πολύ*** $\max\{m,n\}$ ***μακροσκελείς αριθμητικές πράξεις.***

Απόδειξη. Έστω ότι ο αριθμητής και ο παρανομαστής της $R_{m,n}(z)$ είναι τα πολυώνυμα $r^{(0)}(z)$ και $r^{(1)}(z)$, αντίστοιχα.

Συμβολίζουμε με $\deg(\cdots)$ την έκφραση *"βαθμός του* $(\cdots)$*"* και υποθέτουμε ότι:

$$\deg(r^{(0)}) \geq \deg(r^{(1)}).$$

Με διαδοχική διαίρεση (του $r^{(j-1)}$ με τα $r^{(j)}$) βρίσκουμε πηλίκα $q^{(j)}$ και υπόλοιπα $r^{(j+1)}$, όπως ακολούθως:

$$r^{(0)} = r^{(1)}\, q^{(1)} + r^{(2)} \quad (\deg r^{(2)} < \deg r^{(1)}),$$

$$r^{(1)} = r^{(2)}\, q^{(2)} + r^{(3)} \quad (\deg r^{(3)} < \deg r^{(2)}),$$

.....

Επειδή οι βαθμοί $\deg r^{(j)}$ σχηματίζουν μία γνησίως φθίνουσα ακολουθία μη αρνητικών ακεραίων, είναι βέβαιο ότι θα καταλήξουμε σε ένα στάδιο όπου $\deg r^{(k)} = 0$:

$$r^{(k-2)} = r^{(k-1)}\, q^{(k-1)} + r^{(k)} \quad (\deg r^{(k)} = 0)$$

$$r^{(k-1)} = r^{(k)}\, q^{(k)}.$$

Από αυτό το σχήμα προκύπτει ότι:

$$R_{m,n} = \frac{r^{(0)}}{r^{(1)}} = q^{(1)} + \frac{1}{r^{(1)}/r^{(2)}}$$

$$= q^{(1)} + \cfrac{1}{q^{(2)} + \cfrac{1}{r^{(2)}/r^{(3)}}}$$

$$= q^{(1)} + \cfrac{1}{q^{(2)} + \cfrac{1}{q^{(3)} + \ddots + \cfrac{1}{q^{(k-1)} + 1/q^{(k)}}}}$$

Η τελευταία έκφραση της παραπάνω σχέσης μπορεί να γραφεί υπό την μορφή $(\Sigma\mathrm{K})$, γιατί κάθε συνάρτηση r_j (εκτός από την r_1) της $(\Sigma\mathrm{K})$ είναι ένα πολυώνυμο με συντελεστή μεγιστοβάθμιου όρου του ίσο με ένα.

Για τον υπολογισμό ενός τέτοιου πολυωνύμου απαιτούνται όχι περισσότεροι από $(\deg r_j) - 1$ πολλαπλασιασμοί, επειδή το πολυώνυμο μπορεί να γραφεί ως εξής:

$$z^{\nu} + a_{\nu-1}^{(j)} z^{\nu-1} + \ldots + a_1^{(j)} z + a_0^{(j)} = (\ldots((z + a_{\nu-1}^{(j)} z + a_{\nu-2}^{(j)})\ldots) z + a_0^{(j)} .$$

Οι μακροσκελείς αριθμητικές πράξεις, που είναι απαραίτητες για τον υπολογισμό της τιμής $R_{m,n}(z)$ από την εξίσωση $(\Sigma\mathrm{K})$, είναι τότε, οι πολλαπλασιασμοί που απαιτούνται για τον υπολογισμό των $r_j(z)$ και $k-1$ διαιρέσεις.

Ο συνολικός αριθμός αυτών των πράξεων είναι:

$$\begin{aligned}\deg r_1 + \deg r_2 + \ldots + \deg r_k &= \deg q^{(1)} + \deg q^{(2)} + \ldots + \deg q^{(k)} \\ &= [\deg r^{(0)} - \deg r^{(1)}] + [\deg r^{(1)} - \deg r^{(2)}] + \ldots \\ &\quad \ldots + [\deg r^{(k-1)} - \deg r^{(k)}] \\ &= \deg r^{(0)} \le m .\end{aligned}$$

Στον προηγούμενο συλλογισμό χρησιμοποιήσαμε τις ανισότητες $\deg r^{(j+1)} < \deg r^{(j)}$ προκειμένου να συμπεράνουμε ότι

$$\deg r^{(j)} = \deg r^{(j+1)} + \deg q^{(j+1)}.$$

Ας σημειωθεί ότι εάν ο συντελεστής του μεγιστοβάθμιου όρου του πολυωνύμου $q^{(1)}$ είναι ίσος με ένα, τότε ο συνολικός αριθμός των αριθμητικών πράξεων είναι το πολύ ίσος με $(n-1)$. Τώρα, στην περίπτωση που:

$$\deg r^{(0)} < \deg r^{(1)},$$

γράφοντας:

$$R_{m,n} = \frac{c_1}{c_1\ r^{(1)} / r^{(0)}},$$

όπου ο αριθμός c_1 επιλέγεται έτσι ώστε τα πολυώνυμα $c_1 r^{(1)}$ και $r^{(0)}$ να έχουν τον ίδιο συντελεστή των μεγιστοβάθμιων όρων τους, βλέπουμε ότι τα προηγούμενα επιχειρήματα και συλλογισμοί εγγυώνται ότι το κλάσμα

$$c_1 r^{(1)} / r^{(0)}$$

μπορεί να εκφρασθεί σαν ένα συνεχές κλάσμα και ότι για τον υπολογισμό οποιασδήποτε τιμής αυτού του κλάσματος απαιτούνται όχι περισσότερες από $(\deg r^{(1)}) - 1$ μακροσκελείς αριθμητικές πράξεις, και συνεπώς σε αυτήν την περίπτωση ο υπολογισμός των τιμών $R_{m,n}(z)$ της $R_{m,n}$ απαιτεί την εκτέλεση το πολύ $\deg r^{(1)} \le m$ μακροσκελών αριθμητικών πράξεων.

Η *Απόδειξη* του *Θεωρήματος* ολοκληρώθηκε. ■

Παρά το σημαντικό αυτό πλεονέκτημα των ρητών συναρτήσεων, οι συνθήκες ύπαρξης και μοναδικότητας ρητής παρεμβολής είναι αρκετά πιο πολύπλοκες σε σχέση με τις αντίστοιχες που αφορούν την πολυωνυμική παρεμβολή. Ο ενδιαφερόμενος αναγνώστης μπορεί να αναζητήσει περισσότερες επεξηγηματικές και σαφείς λεπτομέρειες γύρω από την μελέτη αυτών των συνθηκών στην βιβλιογραφική παραπομπή $[36]$. Εμείς θα προτιμήσουμε να περιοριστούμε σε μία αναφορά γενικού περιεχόμενου σε δυο ειδικές περιπτώσεις:

. στην περίπτωση όπου $m = n > 1$, και

. στην περίπτωση όπου η συνάρτηση f είναι μία ολόμορφη συνάρτηση ορισμένη σε έναν τόπο U του μιγαδικού επιπέδου.

Στην πρώτη περίπτωση, δοθείσας μίας συνάρτησης

$$f : U \to \boldsymbol{K} : z \mapsto f(z), \text{ με } U \subset \boldsymbol{K} \quad (\boldsymbol{K} = \boldsymbol{R}, \boldsymbol{C})$$

και των τιμών της $f(z)$ σε $(2n+1)$ σημεία

$$z_0, z_1, \ldots, z_{2n}$$

του U με $n > 1$, θα αναζητήσουμε μία ρητή συνάρτηση τύπου (n,n):

$$R_{n,n}(z) = \frac{P_n(z)}{Q_n(z)},$$

όπου $P_n(z)$ και $Q_n(z)$ είναι πολυώνυμα βαθμού ίσου με *n*, τέτοια ώστε:

$$R_{n,n}(z_i) = f(z_i), \text{ για κάθε } i = 0,1,\ldots,2n.$$

Για να επιλύσουμε αυτό το πρόβλημα, θα χρησιμοποιήσουμε:

. *τις αντίστροφες διαφορές* (αντί για τις διηρημένες διαφορές που είχαμε χρησιμοποιήσει στην πολυωνυμική παρεμβολή), και

. *τον τύπο παρεμβολής του Thiele* (αντί για τους τύπους Lagrange, Hermite κ.λ.π., που είχαμε χρησιμοποιήσει στην πολυωνυμική παρεμβολή).

Οι ***αντίστροφες διαφορές*** της f ορίζονται από τις επόμενες επαγωγικές σχέσεις:

$$p_0(z) = f(z) \quad (\forall z \in U),$$

$$p_1(z_i, z_j) = \frac{z_i - z_j}{p_0(z_i) - p_0(z_j)},$$

$$p_2(z_i, z_j, z_k) = p_0(z_j) + \frac{z_i - z_k}{p_1(z_i, z_k) - p_1(z_j, z_k)},$$

και γενικότερα:

$$p_{\nu+1}(z_{i_0}, \ldots, z_{i_{\nu+1}}) = p_{\nu-1}(z_{i_1}, \ldots, z_{i_\nu}) + \frac{z_{i_0} - z_{i_{\nu+1}}}{p_\nu(z_{i_0}, \ldots, z_{i_\nu}) - p_\nu(z_{i_1}, \ldots, z_{i_{\nu+1}})}.$$

Είναι φανερό ότι στην περίπτωσή μας οι προηγούμενες σχέσεις οδηγούν στην εξής ακολουθία τύπων:

$$f(z) = p_0(z_0) + \frac{z - z_0}{p_1(z, z_0)},$$

$$p_1(z, z_0) = p_1(z_0, z_1) + \frac{z - z_1}{p_2(z, z_0, z_1) - p_0(z_0)},$$

.....

$$p_{k-1}(z, z_0, \ldots, z_{k-2}) = p_{k-1}(z_0, \ldots, z_{k-1}) + \frac{z - z_{k-1}}{p_k(z, z_0, \ldots, z_{k-1}) - p_{k-2}(z_0, \ldots, z_{k-2})},$$

που με την σειρά της, οδηγεί στο ανάπτυγμα σε συνεχές κλάσμα της $f(z)$:

$$f(z) = p_0(z_0) + \frac{z - z_0|}{|p_1(z_0, z_1)} + \frac{z - z_1 \quad |}{|p_2(z_0, z_1, z_2) - p_0(z_0)} + \ldots,$$

όπου ο συμβολισμός:

$$t_0 + \frac{s_1|}{|t_1} + \frac{s_2|}{|t_2} + \frac{s_3|}{|t_3} + \ldots$$

σημαίνει το συνεχές κλάσμα:

$$t_0 + \cfrac{s_1}{t_1 + \cfrac{s_2}{t_2 + \cfrac{s_3}{t_3 + \cfrac{s_4}{\ddots}}}}$$

που έχει ***μερικό πηλίκο*** $k-$***τάξης*** το πηλίκο

$$\frac{s_k}{t_k},$$

και $n-$***τάξης προσέγγιση*** (ή $n-$***τάξης συγκλίνοντα όρο***) την έκφραση:

$$\Sigma_n = t_0 + \frac{s_1|}{|t_1} + \frac{s_2|}{|t_2} + \ldots + \frac{s_n|}{|t_n}.$$

Μπορούμε να υπολογίσουμε τους διαδοχικούς συγκλίνοντες όρους του συνεχούς κλάσματος με την βοήθεια των τύπων:

$$P_{-1} := -1 \ , \ Q_{-1} := 0,$$

$$P_0 := t_0 \ , \ Q_0 := 1,$$

.....

$$P_n := t_n P_{n-1} + s_n P_{n-2} \ , \ Q_n := t_n Q_{n-1} + s_n Q_{n-2},$$

γιατί τότε, όπως εύκολα αποδεικνύεται ισχύει:

$$\Sigma_n = \frac{P_n}{Q_n} \qquad n = 0,1,2,\ldots$$

Εάν εφαρμόσουμε αυτούς τους επαγωγικούς τύπους στην διαδικασία υπολογισμού των διαδοχικών συγκλινόντων όρων του συνεχούς κλάσματος που εκφράζει την συνάρτηση

$f(z)$, διαπιστώνουμε ότι ο $n-$τάξης συγκλίνων όρος του συνεχούς κλάσματος της $f(z)$ είναι της μορφής:

$$R_{n,n}(z)=\frac{P_n(z)}{Q_n(z)},$$

όπου:

$$P_{2k}(z)=a_0+a_1\ z+\ldots+a_{k-1}\ z^{k-1}+z^k,$$

$$Q_{2k}(z)=\beta_0+\beta_1\ z+\ldots+\beta_{k-2}\ z^{k-2}+p_{2k-1}(z_0,z_1,\ldots,z_{2k-1})\ z^{k-1},$$

$$P_{2k+1}(z)=\widetilde{a}_0+\widetilde{a}_1\ z+\ldots+\widetilde{a}_{k-1}\ z^{k-1}+p_{2k}(z_0,z_1,\ldots,z_{2k})\ z^k,$$

$$Q_{2k+1}(z)=\widetilde{\beta}_0+\widetilde{\beta}_1\ z+\ldots+\widetilde{\beta}_{k-1}\ z^{k-1}+z^k.$$

Συμπεραίνουμε επίσης αμέσως ότι:

$$R_{n,n}(z_i)=f(z_i)\qquad \forall i=0,1,\ldots,2.n,$$

και συνεπώς η $R_{n,n}(z)$ είναι η ζητούμενη ρητή συνάρτηση παρεμβολής της $f(z)$ στα σημεία $z_0,z_1,\ldots,z_{2.n}$ του U $(n>1)$.

Θα ασχοληθούμε τώρα με την περίπτωση όπου η συνάρτηση f είναι μία ολόμορφη συνάρτηση ορισμένη μέσα σε ένα ανοικτό συνεκτικό υποσύνολο U του μιγαδικού επιπέδου. Τότε, όπως θα δούμε, υπάρχει μεγαλύτερη ευελιξία θεωρήσεων. Συμβολίζουμε με $\mathcal{P}_n(\mathcal{C})$ και $\mathcal{R}_{m,n}(\mathcal{C})$ το σύνολο όλων των μιγαδικών πολυωνύμων με βαθμό το πολύ *n* και το σύνολο όλων των ρητών μιγαδικών συναρτήσεων με βαθμό αριθμητή και παρανομαστή το πολύ m και n αντίστοιχα.

Έστω ένας άπειρος τριγωνικός πίνακας σημείων παρεμβολής $z_k^{(m)}\in U$, που ονομάζεται ***σχήμα παρεμβολής***:

$$M = (z_k^{(m)})_{m\geq 0, 0\leq k\leq m} = \begin{pmatrix} z_0^{(0)} & & & \\ z_0^{(1)} & z_1^{(1)} & & \\ z_0^{(2)} & z_1^{(2)} & z_2^{(2)} & \\ \dots & \dots & \dots & \dots \\ z_0^{(m)} & z_1^{(m)} & \dots & z_m^{(m)} \\ \dots & \dots & \dots & \dots \end{pmatrix}.$$

Κάθε γραμμή:

$$M_m = \{z_0^{(m)}, z_1^{(m)}, \dots, z_m^{(m)}\}$$

του πίνακα *M* ορίζει ένα σύνολο παρεμβολής που περιέχει $(m+1)$ σημεία παρεμβολής. (Δεν αποκλείεται κάποια από τα σημεία αυτά να συμπίπτουν.)

Σε κάθε σύνολο παρεμβολής M_m, αντιστοιχούμε το πολυώνυμο:

$$V_{m+1}(\mathrm{z}) = \gamma \prod_{k=0}^{m} (z - z_k^{(m)}) \quad (\gamma \in \mathbb{C} - \{0\}).$$

Ορισμός IV.1.11.2.

(i). ***Μία ρητή συνάρτηση*** $R_{m,n}(z) \in \mathbb{R}_{m,n}(\mathbb{C})$ ***ονομάζεται ρητή παρεμβολή τύπου*** (m,n) ***της συνάρτησης*** $f(z)$ ***στα*** $(m+n+1)$ ***σημεία παρεμβολής του συνόλου*** M_{m+n}***, εάν:***

$$f - R_{m,n} = \mathrm{O}(V_{m+n}) \text{ σε κάθε } z_k^{(m+n)} \in M_{m+n}.$$

(ii). ***Μία ρητή συνάρτηση***

$$R_{m,n} = \frac{P_m}{Q_n} \in \mathbb{R}_{m,n}(\mathbb{C}), \text{ με } P_m \in \mathbb{P}_m(\mathbb{C}) \text{ και } Q_n \in \mathbb{P}_n(\mathbb{C})$$

ονομάζεται γραμμικοποιημένη ρητή παρεμβολή (ή πολυσημειακή προσέγγιση Padé) τύπου (m,n) ***της συνάρτησης*** $f(z)$ ***στα*** $(m+n+1)$ ***σημεία παρεμβολής του συνόλου*** M_{m+n}***, εάν:***

$$Q_n\ f - P_m = \mathrm{O}(V_{m+n})$$ ***σε κάθε*** $z_k^{(m+n)} \in M_{m+n}$***.***

(Εδώ, η έκφραση $f - P_{m,n} = \mathrm{O}(V_{m+n})$ **σε κάθε** $z_k^{(m+n)} \in M_{m+n}$ **σημαίνει ότι η συνάρτηση:**

$$\frac{f - R_{m,n}}{V_{m+n}}$$

είναι φραγμένη στην περιοχή κάθε σημείου $z_k^{(m+n)} \in M_{m+n}$**. Ομοίως, η έκφραση** $Q_n\ f - P_m = \mathrm{O}(V_{m+n})$ **σε κάθε σημείο** $z_k^{(m+n)} \in M_{m+n}$ **σημαίνει ότι η συνάρτηση:**

$$\frac{Q_n\ f - P_m}{V_{m+n}}$$

είναι φραγμένη στην περιοχή κάθε σημείου $z_k^{(m+n)} \in M_{m+n}$**. Το σύμβολο** O **καλείται *μεγάλο όμικρον του Landeau*.)** ■

Ο *Ορισμός IV.1.11.2.(i)* συνεπάγεται ότι σε κάθε ρίζα του πολυωνύμου V_{m+n} το σφάλμα παρεμβολής $f - R_{m,n}$ έχει επίσης μία ρίζα της ίδιας τουλάχιστον τάξης. Έτσι, η παρεμβολή $R_{m,n}$ καθώς και οι παράγωγοί της $R_{m,n}^{(j)}$ συμπίπτουν με την συνάρτηση f και τις παραγώγους της $f^{(j)}$ σε ένα σημείο $z \in M_{m+n}$ μέχρι μία τάξη που προσδιορίζεται από την συχνότητα εμφάνισης του z στο M_{m+n}.

Η γραμμικοποιημένη μορφή μίας ρητής παρεμβολής $R_{m,n}$, όπως παρουσιάστηκε στον *Ορισμό IV.1.11.2.(ii)*, υπάρχει πάντοτε, γιατί η σχέση

$$Q_n \cdot f - P_m = \mathrm{O}(V_{m+n})$$

ισοδυναμεί με ένα σύστημα $(m+n+1)$ γραμμικών (ομογενών) εξισώσεων, με $(m+n+2)$ άγνωστες παραμέτρους τους συντελεστές των δύο πολυωνύμων P_m και Q_n.

Έτσι, μία μη-τετριμμένη λύση υπάρχει πάντοτε, ενώ δεν είναι καθόλου δύσκολο να επαληθεύσουμε ότι, για μία τέτοια λύση, το ενδεχόμενο $Q_n \equiv 0$ αποκλείεται.

Παρ' όλα αυτά, όπως φαίνεται στο επόμενο παράδειγμα, η ύπαρξη ρητής παρεμβολής δεν είναι γενικά εγγυημένη. Όμως, επαληθεύεται εύκολα πως εάν υπάρχει μία ρητή παρεμβολή τύπου (m,n) της $f(z)$ στα $(m+n+1)$ σημεία του συνόλου M_{m+n}, τότε αυτή είναι μοναδική.

Παράδειγμα IV.1.11.3. ***Επιλέγουμε:***

$$m = n = 1,\ \mathrm{M}_2 = \{1,0,1\} \text{ και } f(z) = z^2.$$

Κάθε ρητή συνάρτηση $R_{1,1} \in \mathcal{R}_{1,1}(\mathbb{C})$ ***είναι ή ένας μετασχηματισμός Möebius ή ένας σταθερός μιγαδικός αριθμός. Εάν η*** $R_{1,1}(z)$ ***είναι ένας μετασχηματισμός Möebius, τότε είναι μονότιμη στο*** $\overline{\mathbb{C}}$ ***και γιαυτό δεν μπορεί να παρεμβάλλει την τιμή*** 1 ***στα δύο διαφορετικά σημεία*** -1 ***και*** 1***. Εάν πάλι η*** $R_{1,1}(z)$ ***είναι ένας σταθερός αριθμός, τότε, προφανώς, δεν μπορεί να παρεμβάλλει τις διαφορετικές τιμές*** 0 ***και*** 1***. Έτσι, ακόμη και σε αυτήν την απλή περίπτωση, δεν υπάρχει ρητή παρεμβολή της συνάρτησης*** $f(z) = z^2$***.*** ■

Συγκρίνοντας την ρητή με την πολυωνυμική παρεμβολή ολόμορφων συναρτήσεων, μπορούμε να παρατηρήσουμε ότι η κύρια αιτία για την μη ύπαρξη των ρητών παρεμβολών οφείλεται στην μη γραμμικότητα της παραμετρικοποίησης των εκάστοτε παρεμβολών. Ιστορικά, ο πρώτος που επισήμανε το ενδεχόμενο μη ύπαρξης ρητών παρεμβολών ήταν ο *Kronecker* το 1881.

Θα συζητήσουμε τώρα την ασυμπτωτική συμπεριφορά μίας ακολουθίας ρητών παρεμβολών της f.

Εάν όλα τα σημεία παρεμβολής συμπίπτουν, δηλαδή εάν:

$$z_0^{(m+n)} = z_1^{(m+n)} = \ldots = z_{m+n}^{(m+n)},$$

τότε η ρητή παρεμβολή τύπου (m,n) υπάρχει και ονομάζεται ***προσεγγιστής Padé*** της f, και συμβολίζεται με $[m/n]_f(z)$.

Στην ασυμπτωτική μελέτη μίας ακολουθίας Padé, συναρτήσεις της μορφής:

$$F_\mu(z) = \int \frac{d\mu(x)}{x-z},$$

όπου μ είναι ένα θετικό μέτρο με φορέα $supp(\mu)$ που περιέχεται στο $\mathbb{R}$, διαδραματίζουν έναν κυρίαρχο ρόλο. Είναι γνωστές σαν ***συναρτήσεις Markov***, ***Stietjes*** ή ***Hamburger***, ανάλογα με το εάν ο φορέας $supp(\mu)$ είναι συμπαγής, περιέχεται σε έναν από τους δυο ημιάξονες $[-\infty,0]$ ή $[0,-\infty]$, ή είναι μη φραγμένος και τέμνει τα δύο σύνολα $[-\infty,0]$ και $[0,-\infty]$, αντίστοιχα.

Διαγώνιοι προσεγγιστές Padé της F_μ :

$$[m/n]_{F_\mu}$$

συγκλίνουν, καθώς το $m \to \infty$, προς την F_μ ως προς την νόρμα της ομοιόμορφης σύγκλισης μέσα σε κάθε συμπαγές υποσύνολο του $\overline{\mathbb{C}} - \overline{supp(\mu)}$. (Ειδικότερα, για την περίπτωση των συναρτήσεων Stieltjes και Hamburger είναι, επί πλέον, αναγκαία η πρόσθετη προϋπόθεση όπως το *πρόβλημα της στιγμής, που συνδέεται με το μέτρο* μ *, είναι προσδιοριστέο*.).

Εάν Q_m είναι ο πολυωνυμικός παρανομαστής του προσεγγιστού Padé $[m/n]_{F_\mu}$ μίας συνάρτησης F_μ τότε το πολυώνυμο Q_m είναι ορθογώνιο ως προς το μέτρο μ.

Ο πολυωνυμικός παρανομαστής Q_m προσδιορίζεται από αυτήν την ορθογωνιότητα, κατά παρέκκλιση μίας πολλαπλασιαστικής σταθεράς.

Ένας παρόμοιος χαρακτηρισμός του παρανομαστή συμβαίνει και στην περίπτωση μίας ρητής παρεμβολής $R_{m,m}(z)$ μίας συνάρτησης $F_\mu(z)$, με την μόνη διαφορά ότι, τώρα, ο πολυωνυμικός παρανομαστής είναι ορθογώνιος ως προς μία επιβαρυμένη σχέση ορθογωνιότητας:

$$\int x^\nu \, Q_m(x) \, \frac{d\mu(x)}{V_{2.m}(x)} = 0 \qquad \forall \nu = 0,1,\ldots,m-1,$$

όπου $V_{2m}(x)$ είναι το πολυώνυμο

$$\prod_{k=0}^{2m-1} (x - z_k^{(2m-1)}).$$

Έτσι, το πολυώνυμο Q_m είναι ορθογώνιο ως προς το επιβαρημένο μέτρο $V_{2m}^{-1} d\mu$.

Γενικά, η θεωρία σύγκλισης ρητών παρεμβολών συναρτήσεων F_μ είναι μία σχεδόν ευθεία γενίκευση εκείνης των προσεγγιστών Padé. Έτσι, και στις δύο Θεωρίες, ο τόπος

σύγκλισης είναι το σύνολο $\overline{C} - \overline{supp(\mu)}$, ενώ ιδιαίτερα για το σφάλμα παρεμβολής έχουμε την ασυμπτωτική εκτίμηση:

$$\limsup_{m\to\infty} \left|F_\mu(z) - R_{m,m}(z)\right|^{\frac{1}{2m}} \le \exp(-\int g_{\overline{C}-\overline{supp(\mu)}}(z,x)\, da_M(x)),$$

για κάθε $z \in \overline{C} - \overline{supp(\mu)}$, όπου:

- $g_{\overline{C}-\overline{supp(\mu)}}(z,x)$ *είναι η* ***συνάρτηση Green*** του τόπου $\overline{C} - \overline{supp(\mu)}$,

και

- $da_M(\cdots)$ *είναι η* ***ασυμπτωτική κατανομή*** *του πίνακα* M, δηλαδή το μέτρο πιθανότητας a_M που ικανοποιεί την σχέση:

$$\lim_{m\to\infty}\left(\frac{1}{m+1}\sum_{k=0}^{m}\delta_{z_k^{(m)}}\right) = a_M,$$

ως προς την ασθενή τοπολογία του διανυσματικού χώρου των μέτρων Borel ($\delta_{z_k^{(m)}}$ είναι το μέτρο Dirac στο σημείο $z_k^{(m)}$).

Μία άλλη κατηγορία συναρτήσεων, για την οποία η σύγκλιση μίας ακολουθίας προσεγγιστών Padé (ως προς την νόρμα της ομοιόμορφης σύγκλισης μέσα σε κάθε συμπαγές σύνολο) έχει αποδειχτεί ότι είναι η κατηγορία των *συναρτήσεων συχνοτήτων Polya*:

$$G(z) = e^{\gamma z}\frac{\prod_j (1+a_j z)}{\prod_j (1+\beta_j z)},$$

όπου $\gamma, a_j, \beta_j \neq 0$ και $\sum_j (a_j + \beta_j) < \infty$.

Έχει αποδειχτεί ότι διαγώνιοι προσεγγιστές Padé τέτοιων συναρτήσεων συγκλίνουν ομοιόμορφα μέσα στα συμπαγή υποσύνολα του

$$\overline{\mathbb{C}} - \{-a_1^{-1}, a_2^{-1}, \ldots, \beta_1, \beta_2, \ldots\}$$

προς τις συναρτήσεις που προσεγγίζουν.

Αυτό το αποτέλεσμα έχει επεκταθεί στην περίπτωση ρητών παρεμβολών στα σημεία σχημάτων παρεμβολής που περιέχουν μόνον πραγματικούς αριθμούς $z_k^{(m)}$ και οι συναρτήσεις G απαρτίζονται από πεπερασμένο μόνον πλήθος πολλαπλασιαστικών παραγόντων που συμμετέχουν στα γινόμενα του ορισμού τους. Το γενικό όμως πρόβλημα παραμένει ακόμη ανοικτό.

Από διάφορα αντιπαραδείγματα, στα οποία χρησιμοποιούνται προσεγγιστές Padé, είναι γνωστό πως η ολομορφία της συνάρτησης f δεν είναι μία αρκετή προϋπόθεση για να εξασφαλίσει την ομοιόμορφη σύγκλιση στα συμπαγή σύνολα των ρητών παρεμβολών.

Έχει μάλιστα αποδειχτεί ότι είναι δυνατή η κατασκευή μίας ακέραιας συνάρτησης τέτοιας ώστε η ακολουθία των διαγωνίων προσεγγιστών Padé αυτής της συνάρτησης να αποκλίνει σε κάθε σημείο του διάτρητου συμπαγοποιημένου επιπέδου $\overline{\mathbb{C}} - \{0\}$.

Από αυτό το αντιπαράδειγμα έπεται ότι προκειμένου περί αποτελεσμάτων σύγκλισης που αφορούν τις κατηγορίες των συναρτήσεων F_μ και G η ειδική δομή αυτών των συναρτήσεων διαδραματίζει έναν εύρωστο και κρίσιμο ρόλο.

Έχοντας υπόψη τις όποιες αυτές δυσκολίες που αφορούν την ομοιόμορφη σύγκλιση, είναι, βεβαίως, ενδιαφέρον το να εξεταστεί το φαινόμενο της σύγκλισης σε ευρύτερες κατηγορίες συναρτήσεων, που ορίζονται κυρίως από ιδιότητες ολομορφίας, εφ' όσον θεωρηθεί ένα ασθενέστερο είδος σύγκλισης.

Ένα τέτοιο είδος σύγκλισης, το οποίο κρίνεται και αρκετά πετυχημένο, είναι η σύγκλιση ως προς την *χωρητικότητα*. Με το σύμβολο

$$cap(\cdots)$$

αναπαριστούμε την (*λογαριθμική*) *χωρητικότητα* των υποσυνόλων του $\mathbb{C}$ που έχουν χωρητικότητα.

Ας σημειωθεί ότι για κάθε υποσύνολο Borel B του $\mathbb{C}$ ισχύει $d(B) \le \pi\, cap^2(B)$ όπου $d(\cdots)$ συμβολίζει το μιγαδικό μέτρο Lebesgue.

Μία ακολουθία μιγαδικών συναρτήσεων $f_m \quad (m = 0,1,2,\ldots)$ λέγεται ότι *συγκλίνει ως προς χωρητικότητα* προς την συνάρτηση f του ανοικτού μοναδιαίου δίσκου D, εάν, για κάθε $\varepsilon > 0$ και για κάθε συμπαγές υποσύνολο E του D ισχύει:

$$\lim_{m\to\infty} cap(\{z \in E : |f_m(z) - f(z)| < \varepsilon\}) = 0 .$$

Το πρώτο αποτέλεσμα που αφορά την σύγκλιση προσεγγιστών Padé ως προς χωρητικότητα αποδείχτηκε από τους *Nuttall-Pommerenke* το 1973.

Αργότερα, το 1979, το Θεώρημα Nuttall-Pommerenke γενικεύθηκε ως εξής:

Θεώρημα IV.1.11.4. ***Έστω f μία μιγαδική (μονότιμη) συνάρτηση, που είναι ολόμορφη μέσα στον τόπο $D-K$, όπου D είναι ο ανοικτός μοναδικός επίπεδος δίσκος και K είναι ένα συμπαγές υποσύνολο του, χωρητικότητας $cap(K)$ ίσης με 0 και τέτοιο ώστε:***

$$K \cap \{z \in \mathbb{C} : z \in M_m, m \in \mathbb{N}\} = \varnothing .$$

Εάν $R_{m,m}$ είναι η γραμμικοποιημένη ρητή παρεμβολή της f στα σημεία του συνόλου M_{2m}, τότε, για κάθε $\varepsilon > 0$ και για κάθε συμπαγές υποσύνολο του E του D ισχύει:

$$\lim_{m\to\infty} cap(\{z \in E : |R_{m,m}(z) - f(z)| < \varepsilon^m\}) = 0 .\blacksquare$$

Από το *Θεώρημα* αυτό έπεται ότι η ακολουθία των ρητών παρεμβολών $R_{m,m}$ $(m = 0,1,\ldots)$ συγκλίνει ως προς την χωρητικότητα προς την συνάρτηση f. Αλλά και, επί πλέον, βλέπουμε ότι η επιτάχυνση της σύγκλισης είναι μεγαλύτερη από ότι η αντίστοιχη της γεωμετρικής σύγκλισης, με ενδεχόμενες εξαιρέσεις σε σύνολα που η χωρητικότητα τους μικραίνει απεριόριστα.

Στην πραγματικότητα, το 1979, το Θεώρημα Nuttal-Pommerenke δεν αποδείχτηκε μόνον για την διαγώνια ακολουθία $(R_{m,m} : m \in \mathbb{N})$, αλλά και για ακολουθίες

$$(R_{m,m} : m \in \mathbb{N}, n \in \mathbb{N})$$

για τις οποίες υπάρχει ένας θετικός αριθμός λ που είναι τέτοιος ώστε:

$$\lambda \le \frac{m}{n} \le \frac{1}{\lambda}, \text{ καθώς } m,n \to \infty .$$

Η υπόθεση:

$$cap(K) = 0$$

του *Θεωρήματος IV.1.11.4* είναι ουσιαστική, καθώς , όπως αποδείχτηκε το 1980 και το 1983, υπάρχουν παραδείγματα στα οποία ενώ η συνάρτηση f έχει ένα σύνολο ανωμαλιών με γνήσια θετική χωρητικότητα, η αντίστοιχη ακολουθία των διαγωνίων προσεγγιστών Padé δεν συγκλίνει (ως προς την χωρητικότητα) μέσα σε οποιοδήποτε ανοικτό υποσύνολο του D.

Όλες οι μερόμορφες συναρτήσεις f ικανοποιούν τις υποθέσεις του *Θεωρήματος IV.1.11.4*, αλλά οι συναρτήσεις που καλύπτονται από το θεώρημα σχηματίζουν μία πολύ μεγαλύτερη κλάση. Για παράδειγμα, οι συναρτήσεις *f* μπορούν να έχουν ουσιώδεις ανωμαλίες ή να είναι ακέραιες.

Τέλος, ας σημειωθεί ότι όπως αποδείχτηκε το 1976, η σύγκλιση ως προς την χωρητικότητα συνεπάγεται την σύγκλιση κατάλληλα επιλεγμένων υπακολουθιών σε κάθε σημείο του ανοικτού μοναδιαίου δίσκου εκτός από τα σημεία κάποιου υποσυνόλου του που έχει μηδενική χωρητικότητα.

IV.2. ΑΡΙΘΜΗΤΙΚΟΣ ΤΕΤΡΑΓΩΝΙΣΜΟΣ

IV.2.1. ΤΟΠΟΘΕΤΗΣΗ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ

Στο *Κεφάλαιο* αυτό, θα συζητήσουμε το πρόβλημα της εύρεσης μίας προσεγγιστικής αριθμητικής τιμής του ορισμένου ολοκληρώματος:

$$I = \int_a^\beta f(x)\, w(x)\, dx\,,$$

όπου $a < \beta$, και όπου w είναι μία δεδομένη συνάρτηση που δεν μηδενίζεται και που είναι ολοκληρώσιμη στο $[a, \beta]$.

Οι μέθοδοι, που θα μελετήσουμε, στηρίζονται στην Θεωρία Παρεμβολής, για αυτό και θα λέμε ότι αυτές είναι ***τύπου παρεμβολής***: Εάν

$$x_0, x_1, \ldots, x_n$$

είναι $(n+1)$, ανά δύο διάφορα μεταξύ τους, σημεία του κλειστού διαστήματος $[a, \beta]$ αντικαθιστούμε την συνάρτηση $f(x)$ με το πολυώνυμο $p_n(x)$ που παρεμβάλλει την f στα σημεία $x_0, x_1, \ldots, x_n$.

Γράφοντας μάλιστα:

$$f(x) = p_n(x) + E_n(x)\,,$$

όπου $E_n(x)$ είναι το αντίστοιχο σφάλμα παρεμβολής, τότε είναι φανερό ότι:

$$\int_a^\beta f(x)\, w(x) dx = \int_a^\beta p_n(x)\, w(x) dx + \int_a^\beta E_n(x)\, w(x) dx\,.$$

Χρησιμοποιώντας τον τύπο παρεμβολής του Lagrange η τελευταία ισότητα γίνεται:

$$I = \sum_{i=0}^{n} A_i^{(n)} \, f(x_i) + \int_a^{\beta} E_n(x) \, w(x) \, dx \, ,$$

όπου έχουμε θέσει:

$$A_i^{(n)} := \int_a^{\beta} L_i^{(n)}(x) \, w(x) \, dx \, .$$

Στην συνέχεια, θα λαμβάνουμε το πεπερασμένο άθροισμα:

$$\sum_{i=0}^{n} A_i^{(n)} \, f(x_i)$$

σαν μία προσέγγιση του ολοκληρώματος I και θα λέμε ότι ένα τέτοιο άθροισμα συνιστά έναν ***αριθμητικό τετραγωνισμό*** *του* I .

Το σφάλμα του αριθμητικού τετραγωνισμού θα δίδεται από τον *Τύπο*:

$$R_n(f) := \int_a^{\beta} E_n(x) \, w(x) \, dx \, .$$

Το επόμενο αποτέλεσμα είναι άμεσο και σαφές, γιατί απορρέει από τον ίδιο τον ορισμό της πολυωνυμικής παρεμβολής:

Θεώρημα IV.2.1.1. ***Μία αναγκαία και ικανή συνθήκη προκειμένου να ισχύει*** $R_n(f) = 0$ ***είναι*** $f \in \mathbb{P}_n(\mathbb{R})$. ■

Σε μία τέτοια περίπτωση, δηλαδή όταν

$$R_n(f) = 0 \, ,$$

λέμε ότι ***ο (αριθμητικός) τετραγωνισμός είναι ακριβής***.

Σύμφωνα λοιπόν με το προηγούμενο *Θεώρημα* ο τετραγωνισμός είναι ακριβής εάν και μόνο εάν

$$f \in \boldsymbol{P}_n(\boldsymbol{R}).$$

Στην συνέχεια θα εξετάσουμε τον τρόπο υπολογισμού των ποσοτήτων $A_i^{(n)}$, όταν τα σημεία x_i ισαπέχουν διαδοχικά μεταξύ τους: θα οδηγηθούμε, έτσι, στους ***Τύπους Newton-Côtes***.

Η επισταμένη μελέτη αυτών των *Τύπων* θα καταστήσει εμφανή την αριθμητική τους αστάθεια, όπως επίσης και την αδυναμία σύγκλισης τους όταν ο αριθμός των σημείων παρεμβολής αυξάνει απεριόριστα. Είναι λοιπόν φυσικώς επακόλουθη η αναζήτηση σύνθετων τύπων ολοκλήρωσης.

Τέτοιοι *Τύποι* είναι οι ***Τύποι Gauss***, στους οποίους μπορούμε να καταλήξουμε διερευνώντας την "βέλτιστη" επιλογή των σημείων παρεμβολής.

Στην διερεύνηση αυτή, θα υποχρεωθούμε να μελετήσουμε κάποιες οικογένειες ***ορθογωνίων πολυωνύμων***.

Εκτός από όλα αυτά, θα παρουσιάσουμε την ***μέθοδο των προσδιοριστέων συντελεστών*** καθώς και την ***Θεωρία των Στιγμών***, που θα μας επιτρέψουν να αποκομίσουμε όλους τους παραπάνω τύπους προσεγγιστικής ολοκλήρωσης.

Το *Κεφάλαιο* θα ολοκληρωθεί με μία αναφορά στο πρόβλημα της προσέγγισης των αριθμητικών τιμών πολλαπλών ολοκληρωμάτων.

IV.2.2. ΟΙ ΤΥΠΟΙ NEWTON-COTES

Λαμβάνουμε $w(x)=1$ και θέτουμε:

$$h := \frac{\beta - a}{n}$$

Επιλέγουμε:

$$x_i = a + i\,h, \quad \forall i = 0,1,\ldots,n\,.$$

Όπως είδαμε στην *Παράγραφο IV.1.2*, το πολυώνυμο παρεμβολής $p_n(x)$ της $f(x)$ στα σημεία $x_0, x_1, \ldots, x_n$ μπορεί να γραφεί υπό την μορφή:

$$p_n(x) = V_n(x) \sum_{i=0}^{n} \frac{f(x_i)}{(x - x_i)\,V_n'(x_i)},$$

όπου $V_n(x) = (x - x_0)\,(x - x_1)\ldots(x - x_n)$.

Σύμφωνα με όσα αναφέρθηκαν στην προηγούμενη *Παράγραφο*, ισχύει ότι:

$$A_i^{(n)} = \int_a^{\beta} \frac{V_n(x)}{(x - x_i)\,V_n'(x_i)}\,dx$$

Εάν πραγματοποιήσουμε την αλλαγή μεταβλητής:

$$y = \frac{x - a}{\beta},$$

βρίσκουμε ότι

$$y_i = i \quad (i = 0,1,2,\ldots n)$$

και διαπιστώνουμε αμέσως ότι:

$$A_i^{(n)} = \frac{\beta - a}{n}\,\frac{(-1)^{n-i}}{i!\,(n-i)!} \int_0^n \prod_{\substack{k=0 \\ k \neq i}}^{n} (y - k)\,dy\,,$$

για κάθε $i = 0,1,2,\ldots,n$ (και για $w(x)=1$).

Παρατήρηση IV.2.2.1. Είναι σαφές ότι $A_i^{(n)} = A_{n-i}^{(n)} \quad (\forall i = 0,1,\ldots,n)$. ■

Μία εφαρμογή του παραπάνω *Τύπου* για $n = 1$ δείχνει ότι:

$$A_0^{(1)} = A_1^{(1)} = \frac{\beta - a}{2},$$

και συνεπώς:

$$I = \frac{\beta - a}{2}[f(a) + f(\beta)] + R_1(f).$$

Μία δεύτερη εφαρμογή του ίδιου *Τύπου* για $n = 2$ δείχνει ότι:

$$A_0^{(2)} = A_2^{(2)} = \frac{\beta - a}{6} \quad \text{και} \quad A_1^{(2)} = \frac{4\,(\beta - a)}{6},$$

και επομένως:

$$I = \frac{\beta - a}{6}[f(a) + 4\,f(\frac{a+\beta}{2}) + f(\beta)] + R_2(f).$$

Οι συντελεστές $A_i^{(n)}$ υπολογίζονται μία τελεσίδικη φορά, και για όλες τις διαφορετικές τιμές του n.

Προς τούτο, υπάρχουν συναφείς πίνακες εύρεσης των ποσοτήτων $A_i^{(n)}$. Η γνώση όμως του περιεχομένου αυτών των πινάκων δεν είναι χρήσιμη, γιατί στα πρακτικά προβλήματα επικαλούμαστε τους παραπάνω τύπους μόνον όταν $n = 1$ ή $n = 2$. Η αιτία αυτής της προτίμησης προέρχεται από την αστάθεια και την μη σύγκλιση των τύπων Newton-Côtes. Στην συνέχεια και μέχρι το τέλος της παραγράφου θα επιχειρήσουμε να δικαιολογήσουμε αυτόν τον ισχυρισμό.

Ας αρχίσουμε με την μελέτη του σφάλματος $R_n(f)$ του αριθμητικού τετραγωνισμού Newton-Côtes.

Όπως επισημάνθηκε στην *Παράγραφο IV.1.6*, εφόσον η συνάρτηση $f(x)$ είναι $(n+1)$ φορές συνεχώς διαφορίσιμη σε κάθε σημείο μίας ανοικτής περιοχής του διαστήματος $[a,\beta]$ το σφάλμα παρεμβολής $E_n(x)$ δίνεται από την ισότητα:

$$E_n(x) = \frac{V_n(x)}{(n+1)!}\frac{d^{(n+1)}f}{dx^{(n+1)}}(\xi) = V_n(x)\,[x, x_0, x_1, \ldots, x_n],$$

για κάποιο $\xi \in [a,\beta]$ το οποίο εξαρτάται από το εκάστοτε x, και όπου $V_n(x)$ είναι πάντα η έκφραση $V_n(x) = (x-x_0)\,(x-x_1)\ldots(x-x_n)$.

Από αυτήν την ισότητα έπεται μία γενική διατύπωση του σφάλματος τετραγωνισμού:

$$R_n(f) = \int_a^\beta \frac{V_n(x)}{(n+1)!}\frac{d^{(n+1)}f}{dx^{(n+1)}}(\xi)\,dx$$

$$= \frac{1}{(n+1)!}\frac{d^{(n+1)}f}{dx^{(n+1)}}(\xi)\int_a^\beta V_n(x)\,dx\,.$$

Ο υπολογισμός όμως του τελευταίου ολοκληρώματος μπορεί να είναι αρκετά πολύπλοκος, γιατί υπάρχει αδυναμία εφαρμογής του *Θεωρήματος Μέσης Τιμής*, αφού η συνάρτηση $V_n(x)$ δεν διατηρεί το πρόσημό της σταθερό καθώς το x διατρέχει το διάστημα $[a,\beta]$. Γιαυτό, θα προτιμήσουμε να χρησιμοποιήσουμε τον τύπο του σφάλματος που προκύπτει από την εφαρμογή των διηρημμένων διαφορών:

$$R_n(f) = \int_a^\beta V_n(x)\,[x, x_0, x_1, \ldots, x_n]\,dx\,.$$

Θα χρειασθούμε το προπαρασκεαστικό υλικό των δύο επομένων Προτάσεων:

Λήμμα IV.2.2.2. ***Εάν η συνάρτηση*** f ***είναι*** $(n+2)$ ***φορές συνεχώς διαφορίσιμη μέσα σε μία ανοικτή περιοχή του διαστήματος*** $[a,\beta]$***, τότε για κάθε*** $x\in[a,\beta]$ ***υπάρχει*** $\xi\in[a,\beta]$ ***τέτοιο ώστε***

$$\frac{d}{dx}[x,x_0,x_1,\ldots,x_n]=\frac{d^{(n+2)}f}{dx^{(n+2)}}(\xi).$$

Απόδειξη. Σύμφωνα με το *Πόρισμα IV.6.2*, ισχύει:

$$[x,x_0,x_1,\ldots x_n]=\frac{1}{(n+1)!}\frac{d^{(n+1)}f}{dx^{(n+1)}}(\xi),$$

για κάποιο $\xi\in[a,\beta]$ που εξαρτάται από το x. Άρα:

$$[x,x_0,x_1,\ldots,x_n,x+h]=\frac{1}{(n+2)!}\frac{d^{(n+2)}f}{dx^{(n+2)}}(\xi).$$

Χρησιμοποιώντας τον ορισμό των διηρημμένων διαφορών, όπως επίσης, και την ιδιότητα συμμετρίας που εκδηλώνουν ως προς τα ορίσματά τους, έχουμε:

$$[x,x_0,x_1,\ldots,x_n,x+h]=-\frac{[x,x_0,\ldots,x_n]-[x+h,x_0,\ldots,x_n]}{h},$$

οπότε, αφήνοντας το h να τείνει απεριόριστα προς το μηδέν, παίρνουμε:

$$\lim_{h\to 0}[x,x_0,\ldots,x_n,x+h]=\frac{d[x,x_0,\ldots,x_n]}{dx}=\frac{d^{(n+2)}f}{dx^{(n+2)}}(\xi),$$

που ολοκληρώνει την *Απόδειξη* του *Λήμματος*. ■

Θα εισάγουμε τώρα την συνάρτηση $W_n(x)$ που ορίζεται από την σχέση:

$$W_n(x) := \int_a^x V_n(t)\ dt\,.$$

Λήμμα IV.2.2.3.. ***Εάν ο φυσικός αριθμός n είναι άρτιος, τότε:***

$$W_n(a) = W_n(\beta) = 0 \quad \textbf{\textit{και}} \quad W_n(x) > 0 \qquad \forall x \in [a,\beta]\,.$$

Απόδειξη. Επειδή ο φυσικός αριθμός n είναι άρτιος και επειδή τα σημεία x_i έχουν, ανά δύο διαδοχικά, ίσες αποστάσεις μεταξύ τους, θα ισχύει:

$$x_{n/2} = \frac{\beta + a}{2}\,,$$

και, κατά συνέπεια, η συνάρτηση $V_n(x)$ είναι αντισυμμετρική ως προς το κέντρο του διαστήματος $[a,\beta]$, με αποτέλεσμα να έπεται ότι:

$$W_n(\beta) = \int_a^\beta V_n(x)dx = 0\,.$$

Ακόμη, είναι προφανές ότι:

$$W_n(a) = 0\,.$$

Η απόδειξη του δεύτερου ισχυρισμού του *Λήμματος* είναι σαφώς πιο μακροσκελής.

Κατ' αρχήν, πραγματοποιούμε, μέσα στον τύπο της συνάρτησης $W_n(x)$, την αλλαγή μεταβλητής $t = x_0 + p\,h$ και παίρνουμε:

$$W_n(x) = \int_0^x h^{n+2}\ p\,(p-1)...(p-n)\ dp \quad \big(x \in [0,n]\big).$$

Ύστερα, θέτουμε:

$$u_n(p) := p\,(p-1)\ldots(p-n) \quad \big(p \in [0,n]\big)$$

και βλέπουμε ότι:

$$\frac{u_n(p+1)}{u_n(p)} = \frac{p+1}{p-n}.$$

Αυτός ο λόγος είναι μία συνάρτηση φθίνουσα του $p \in [0,n]$, τέτοια ώστε:

$$(\text{για } p=0)\quad \frac{u_n(1)}{u_n(0)} = \frac{1}{n},$$

και

$$(\text{για } p = \frac{n}{2}-1)\quad \frac{u_n(\frac{n}{2})}{u_n(\frac{n}{2}-1)} = -\frac{1}{1+\frac{2}{n}}.$$

Επομένως, θα είναι:

$$|u_n(p+1)| < |u_n(p)|, \text{ όταν } 0 < p+1 \le \frac{n}{2}.$$

Ακόμη, λόγω της αντισυμμετρίας της $u_n(p)$ ως προς το σημείο $\frac{n}{2}$, θα είναι:

$$|u_n(p)| < |u_n(p+1)|, \text{ όταν } \frac{n}{2} \le p < n.$$

Οι μόνες ρίζες της συνάρτησης $V_n(x)$ είναι τα σημεία

$$x_0, x_1, \ldots, x_n$$

και, από την άλλη πλευρά, ο βαθμός του πολυωνύμου $V_n(x)$ είναι περιττός. Έτσι θα έχουμε:

$$V_n(x) < 0 \text{, όταν } x < x_0,$$

$$V_n(x) > 0 \text{, όταν } x_0 < x < x_1,$$

και συνεπώς:

$$W_n(x) > 0,\ \forall x \in [x_0, x_1].$$

Όμως, όπως έγινε πρόδηλο, η αρνητική συνεισφορά του $V_n(x)$ στην διαμόρφωση της συνάρτησης $W_n(x)$ μέσα στο διάστημα $[x_1, x_2]$ είναι πιο ασθενής (κατ' απόλυτη τιμή) απ' ότι η αντίστοιχη θετική συνεισφορά μέσα στο διάστημα $[x_0, x_1]$. Από αυτό συνεπάγεται ότι θα έχουμε:

$$W_n(x) > 0,\ \ \forall x \in [x_0, x_2].$$

Συνεχίζοντας, κατ' αυτόν τον τρόπο, διαπιστώνουμε πως:

$$W_n(x) > 0,\ \forall x \in [x_0, x_{\frac{n}{2}}].$$

Η απόδειξη της ίδιας ανισότητας μέσα στο δεύτερο μισό του διαστήματος $[a, \beta]$, δηλαδή μέσα στο διάστημα $]x_{n/2}, x_n[$, είναι ανάλογη με αυτήν που προηγήθηκε.

Έτσι, τελικά:

$$W_n(x) > 0,\ \ \forall x \in [a, \beta],$$

και η *Απόδειξη* του *Λήμματος* έχει τελειώσει. ■

Είμαστε τώρα σε θέση να μελετήσουμε το σφάλμα του τύπου τετραγωνισμού Newton-Côtes, στην περίπτωση που ο φυσικός αριθμός n είναι άρτιος.

Θεώρημα IV.2.2.4. ***Εάν η συνάρτηση f είναι $(n+2)$ φορές συνεχώς διαφορίσιμη μέσα σε μία ανοικτή περιοχή του διαστήματος $[a,\beta]$, τότε το σφάλμα $R_n(f)$ που προκύπτει από τον τύπο τετραγωνισμού Newton-Côtes, για n =άρτιο, ισούται με:***

$$R_n(f) = -\frac{h^{n+3}}{(n+2)!}\frac{d^{(n+2)}f}{dx^{(n+2)}}(\xi)\int_0^n t^2\,(t-1)...(t-n)\,dt\,,$$

για κάποιο σημείο $\xi \in [a,\beta]$.

Απόδειξη. Όπως αναφέρθηκε πιο πάνω, το σφάλμα $R_n(f)$ είναι ίσο με:

$$R_n(f) = \int_a^\beta V_n(x)\,[x,x_0,...,x_n]\,dx\,,$$

που μπορεί επίσης να γραφεί υπό την μορφή:

$$R_n(f) = \int_a^\beta \frac{dW_n}{dx}(x)\,[x,x_0,...,x_n]\,dx\,.$$

Ολοκληρώνοντας κατά μέρη (για n =άρτιο), και χρησιμοποιώντας το πρώτο συμπέρασμα του *Λήμματος IV.2.2.3*, η τελευταία ισοδυναμεί με την:

$$R_n(f) = [W_n(x)\,[x,x_0,...,x_n]]_{x=a}^{x=\beta} - \int_a^\beta W_n(x)\,\frac{d[x,x_0,...,x_n]}{dx}\,dx$$

$$= -\int_a^\beta W_n(x)\,\frac{d[x, x_0, \ldots, x_n]}{dx}\,dx\,,$$

οπότε, μία εφαρμογή του *Λήμματος IV.2.2.2* δείχνει ότι:

$$R_n(f) = -\int_a^\beta W_n(x)\,\frac{d^{(n+2)} f}{dx^{(n+2)}}(\xi)\,\frac{1}{(n+2)!}\,dx$$

για κάποιο σημείο $\xi \in [a, \beta]$.

Η συνάρτηση $d^{(n+2)} f / dx^{(n+2)}$ είναι συνεχής, ενώ, σύμφωνα με το δεύτερο συμπέρασμα του *Λήμματος IV.2.2.3*, ισχύει $W_n(x) \geq 0$ για κάθε $x \in [a, \beta]$. Άρα, μπορούμε να επικαλεσθούμε το *Θεώρημα Μέσης Τιμής* και να εκφράσουμε το πιο πάνω ολοκλήρωμα υπό την μορφή:

$$R_n(f) = -\frac{1}{(n+2)!}\,\frac{d^{(n+2)} f}{dx^{(n+2)}}(\xi)\int_a^\beta W_n(x)\,dx \quad (\xi \in [a, \beta])\,.$$

Μία νέα ολοκλήρωση κατά μέρη πιστοποιεί ότι:

$$\int_a^\beta W_n(x)\,dx = [x\,W_n(x)]_{x=a}^{x=\beta} - \int_a^\beta \frac{dW_n}{dx}(x)\,dx = -\int_a^\beta x\,V_n(x)\,dx\,,$$

απ' όπου έπεται το συμπέρασμα του *Θεωρήματος*, μετά από την αλλαγή μεταβλητής $x = x_0 + t\,h$. □

Η περίπτωση που ο φυσικός αριθμός n είναι περιττός εξετάζεται χωριστά στο επόμενο:

Θεώρημα IV.2.2.5. ***Εάν η συνάρτηση f είναι $(n+1)$ φορές συνεχώς διαφορίσιμη μέσα σε μία ανοικτή περιοχή του διαστήματος $[a,\beta]$, τότε το σφάλμα $R_n(f)$ που προκύπτει από τον τύπο τετραγωνισμού Newton-Côtes, για n =περιττό, ισούται με:***

$$R_n(f)=\frac{h^{n+2}}{(n+1)!}\frac{d^{(n+1)}f}{dx^{(n+1)}}(\xi)\int_0^n t\,(t-1)\ldots(t-n)\,dt\,,$$

για κάποιο σημείο $\xi\in[a,\beta]$.

Απόδειξη. Επειδή:

$$R_n(f)=\int_a^{\beta-h}V_n(x)\,[x,x_0,\ldots,x_n]\,dx+\int_{\beta-h}^{\beta}V_n(x)\,[x,x_0,\ldots x_n]\;dx\,,$$

και καθώς η συνάρτηση $V_n(x)$ δεν αλλάζει πρόσημο μέσα στο διάστημα $[\beta-h,\beta]$, έχουμε:

$$R_n(f)=\int_a^{\beta-h}V_n(x)\,[x,x_0,\ldots,x_n]\;dx+\frac{1}{(n+1)!}\frac{d^{(n+1)}f}{dx^{(n+1)}}(\xi_1)\int_{\beta-h}^{\beta}V_n(x)\,dx\,,$$

για κάποιο $\xi_1\in[\beta-h,\beta]$.

Η διαχείριση του πρώτου ολοκληρώματος αυτής της ισότητας πραγματοποιείται θεωρώντας τις σχέσεις:

$$V_n(x)=(x-x_n).V_{n-1}(x)\ \text{και}\ W_{n-1}(x)=\int_a^x V_{n-1}(t)dt$$

και χρησιμοποιώντας τον ορισμό των διηρημμένων διαφορών:

$$[x,x_0,\ldots,x_n]=\frac{[x,x_0,\ldots,x_{n-1}]-[x_0,\ldots,x_n]}{x-x_n}\,.$$

Τότε, το πρώτο ολοκλήρωμα γίνεται:

$$\int_a^{\beta-h} V_n(x)\,[x,x_0,\ldots,x_n]\,dx = \int_a^{\beta-h} \frac{dW_{n-1}}{dx}(x)\,\{[x,x_0,\ldots,x_{n-1}]-[x_0,\ldots x_n]\}\;dx\,.$$

Εφόσον τώρα ο φυσικός αριθμός n είναι περιττός, ο αριθμός $(n-1)$ είναι άρτιος, οπότε, σύμφωνα με το *Λήμμα IV.2.2.3*, θα είναι:

$$W_{n-1}(a) = W_{n-1}(\beta-h) = 0\,,$$

και επομένως:

$$\int_a^{\beta-h} \frac{dW_{n-1}}{dx}(x)\,[x_0,\ldots,x_n]\;dx = 0\,.$$

Άρα, το πρώτο ολοκλήρωμα γράφεται υπό την μορφή:

$$\int_a^{\beta-h} V_n(x)\,[x,x_0,\ldots,x_n]\;dx = \int_a^{\beta-h} \frac{dW_{n-1}}{dx}(x)\,[x,x_0,\ldots,x_n]\;dx\,.$$

Πραγματοποιώντας μία ολοκλήρωση κατά μέρη και εφαρμόζοντας το *Θεώρημα Μέσης Τιμής* (όπως προηγουμένως), βρίσκουμε ότι:

$$\int_a^{\beta-h} V_n(x)\,[x,x_0,\ldots,x_n]\;dx = -\frac{1}{(n+1)!}\,\frac{d^{(n+1)}f}{dx^{(n+1)}}(\xi_2)\int_a^{\beta-h} W_{n-1}(x)\,dx\,,$$

για κάποιο $\xi_2 \in [a,\beta-h]$.

Συνεπώς, το σφάλμα $R_n(f)$ γράφεται υπό την μορφή:

$$R_n(f) = \frac{1}{(n+1)!}\left\{\frac{d^{(n+1)}f}{dx^{(n+1)}}(\xi_1)\int_{\beta-h}^{\beta} V_n(x)\,dx - \frac{d^{(n+1)}f}{dx^{(n+1)}}(\xi_2)\int_a^{\beta-h} W_{n-1}(x)\,dx\right\}.$$

Παρατηρούμε τώρα ότι η συνάρτηση $V_n(x)$ είναι θετική, όταν $x>\beta$, γιατί η $V_n(x)$ είναι ένα πολυώνυμο άρτιου βαθμού.

Γι' αυτό $V_n(x)\le 0$, όταν $x\in[\beta-h,\beta]$, γεγονός που συνεπάγεται ότι το πρώτο ολοκλήρωμα της παραπάνω έκφρασης είναι ένας αρνητικός αριθμός.

Από την άλλη πλευρά, το δεύτερο ολοκλήρωμα της έκφρασης αυτής είναι ένας θετικός αριθμός, επειδή, σύμφωνα με το *Λήμμα IV.2.2.3*, ισχύει $W_{n-1}(x) \geq 0$ όταν $x \in [a, \beta - h]$.

Έτσι, το σφάλμα $R_n(f)$ είναι της μορφής:

$$R_n(f) = -\left\{ A \ \frac{d^{(n+1)} f}{dx^{(n+1)}}(\xi_1) + B \ \frac{d^{(n+1)} f}{dx^{(n+1)}}(\xi_2) \right\},$$

όπου A και B είναι θετικοί αριθμοί. Καθώς η συνάρτηση $d^{(n+1)} f / dx^{(n+1)}$ είναι συνεχής, υπάρχει ένα σημείο $\xi \in [\xi_1, \xi_2]$ τέτοιο ώστε:

$$A \ \frac{d^{(n+1)} f}{dx^{(n+1)}}(\xi_1) + B \ \frac{d^{(n+1)} f}{dx^{(n+1)}}(\xi_2) = (A + B) \ \frac{d^{(n+1)} f}{dx^{(n+1)}}(\xi).$$

Επί πλέον, αφού:

$$\int_a^{\beta-h} V_n(x) \, dx = [W_{n-1}(x) \, (x - \beta)]_{x=a}^{x=\beta-h} - \int_a^{\beta-h} W_{n-1}(x) \, dx = -\int_a^{\beta-h} W_{n-1}(x) \, dx,$$

ισχύει:

$$A + B = -\frac{1}{(n-1)!} \int_a^{\beta} V_n(x) \ dx,$$

που ολοκληρώνει την *Απόδειξη* του *Θεωρήματος*. (Ο ενδιαφερόμενος αναγνώστης μπορεί επίσης να συμβουλευτεί τις βιβλιογραφικές παραπομπές [26] είτε [50]). ■

Παρατηρήσεις IV.2.2.6.(i). Είναι προτιμότερη η εφαρμογή των τύπων Newton-Côtes για n =άρτιος.

(ii). Τα μεγέθη n και h, που υπεισέρχονται στην διατύπωση των τύπων των σφαλμάτων $R_n(f)$, δεν είναι ανεξάρτητα μεταξύ τους, γιατί ισχύει $n \, h = \beta - a$. ■

Μετά από αυτήν την σχετικά εκτενή αναφορά μας στην μελέτη του σφάλματος που προκύπτει από τον τύπο τετραγωνισμού Newton-Côtes, επανερχόμαστε στην εξέταση του φαινομένου της ευστάθειας και του φαινομένου της σύγκλισης της μεθόδου που περιγράφεται από τον συγκεκριμένο τετραγωνισμό.

Όπως θα διαπιστώσουμε, η ευστάθεια και η σύγκλιση είναι δύο διαφορετικές έννοιες, που όμως δεν είναι ανεξάρτητες μεταξύ τους. Στην πραγματικότητα, η συνθήκη, που εξασφαλίζει την ευστάθεια, εγγυάται και την σύγκλιση ([41], [59]).

Είναι, εξ άλλου προφανές ότι πρόκειται για δύο θεωρητικές έννοιες, οι οποίες εξαρτώνται από την εκάστοτε χρησιμοποιούμενη μέθοδο τετραγωνισμού.

Η ευστάθεια είναι μία έννοια που εκφράζει το γεγονός ότι ο μαθηματικός τύπος, που μορφοποιεί τον τετραγωνισμό, παραμένει ελάχιστα ευαίσθητος στα σφάλματα υπολογισμού των ποσοτήτων:

$$f(x_0), f(x_1), \ldots, f(x_n).$$

Πράγματι, αντί να υπολογισθεί το άθροισμα

$$\sum_{i=0}^{n} A_i^{(n)} \cdot f(x_i),$$

υπολογίζεται ένα άθροισμα

$$\sum_{i=0}^{n} A_i^{(n)} \cdot (f(x_i) + \varepsilon_i),$$

που σημαίνει ότι υπάρχει πεπλεγμένα ένα ακόμη σφάλμα, το εξής:

$$\sum_{i=0}^{n} A_i^{(n)} \ \varepsilon_i .$$

Γι' αυτό, δίνουμε τον ακόλουθο *Ορισμό*:

Ορισμός IV.2.2.7. ***Ένας μαθηματικός τύπος τετραγωνισμού καλείται ευσταθής, εάν, για οποιοδήποτε*** $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_n \in \mathbb{R}$ ***, υπάρχουν ένας σταθερός αριθμός*** $M = M(\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_n)$ ***που είναι ανεξάρτητος του n και τέτοιος ώστε:***

$$\left|\sum_{i=0}^{n} A_i^{(n)} \varepsilon_i\right| \le M \max_{0\le i\le n} |\varepsilon_i| . \blacksquare$$

Το επόμενο αποτέλεσμα χαρακτηρίζει την ευστάθεια τετραγωνισμών:

Θεώρημα IV.2.2.8. ***Μία συνθήκη, ικανή και αναγκαία, για την ευστάθεια ενός μαθηματικού τετραγωνισμού είναι η ύπαρξη μίας σταθεράς*** M ***, ανεξάρτητης του αριθμού*** n ***και τέτοιας ώστε:***

$$\sum_{i=0}^{n} \left|A_i^{(n)}\right| \le M \quad \forall n = 0,1\ldots$$

Απόδειξη. Επειδή:

$$\left|\sum_{i=0}^{n} A_i^{(n)} \varepsilon_i\right| \le \sum_{i=0}^{n} \left|A_i^{(n)}\right| \quad \max_{0\le i\le n} |\varepsilon_i| ,$$

η συνθήκη είναι αναγκαία.

Από την άλλη πλευρά, η συνθήκη είναι και ικανή, γιατί εάν δεν υπήρχε σταθερά M, ανεξάρτητη του n και που να επαληθεύει την συνθήκη του *Θεωρήματος*, τότε θα είχαμε:

$$\lim_{n\to\infty} \sum_{i=0}^{n} \left|A_i^{(n)}\right| = \infty ,$$

οπότε, επιλέγοντας:

$$\varepsilon_i = \frac{A_i^{(n)}}{\left|A_i^{(n)}\right|} \quad (i = 0,1,\ldots,n),$$

θα βρίσκαμε ότι:

$$\left|\sum_{i=0}^{n} A_i^{(n)}\ \varepsilon_i\right| = \sum_{i=0}^{n}\left|A_i^{(n)}\right| \Rightarrow \lim_{n\to\infty}\left|\sum_{i=0}^{n} A_i^{(n)}\ \varepsilon_i\right| = \infty,$$

που είναι άτοπο, λόγω του *Ορισμού IV2.2.7*. Άρα, η συνθήκη του *Θεωρήματος* είναι ικανή και η *Απόδειξη* είναι πλήρης. ■

Η έννοια της σύγκλισης εκφράζει το γεγονός ότι του σφάλμα του τετραγωνισμού τείνει προς το μηδέν όταν το n τείνει προς το άπειρο και όταν η συνάρτηση f ανήκει σε ένα ορισμένο σύνολο. Δίνουμε τον ακόλουθο *Ορισμό*:

Ορισμός IV.2.2.9. ***Έστω V ένα ορισμένο σύνολο πραγματικών συναρτήσεων ενός κλειστού υποσυνόλου της ευθείας. Λέμε ότι ένας μαθηματικός τύπος τετραγωνισμού συγκλίνει επί του V, εάν για κάθε $f \in V$ συμβαίνει:***

$$\lim_{n\to\infty} R_n(f) = 0.$$

Ας θεωρήσουμε ένα συγκεκριμένο παράδειγμα.

Έστω ο διανυσματικός χώρος $C[a,\beta]$ όλων των πραγματικών συναρτήσεων που είναι συνεχής επί του διαστήματος $[a,\beta] \subset \mathbb{R}$.

Εφοδιάζουμε τον διανυσματικό χώρο $C[a,\beta]$ με την νόρμα της ομοιόμορφης σύγκλισης επί του $[a,\beta]$.

Με άλλα λόγια, σε κάθε $f \in C[a,\beta]$ αντιστοιχούμε τον μη αρνητικό αριθμό:

$$\|f\|_\infty := \sup_{x\in[a,\beta]} |f(x)|.$$

Η μελέτη της σύγκλισης των τύπων τετραγωνισμού Newton-Côtes επί του $C[a,\beta]$ προσδιορίζει μέσα από τις γραμμές του *Θεωρήματος IV.1.8.2*, όπως ακριβώς είχε γίνει για την μελέτη της σύγκλισης των μαθηματικών τύπων παρεμβολής.

Έτσι, ισχύει το:

Θεώρημα IV.2.2.10. ***Μία συνθήκη, ικανή και αναγκαία, για την σύγκλιση ενός μαθηματικού τύπου τετραγωνισμού επί του $C[a,\beta]$ είναι η ταυτόχρονη ικανοποίηση των δύο επόμενων αξιώσεων:***

(i). ***ΟΤύπος αυτός να συγκλίνει επί του διανυσματικού υποχώρου $P(\mathbb{R})$ του $C[a,\beta]$ (όπου $P(\mathbb{R})$ είναι ο διανυσματικός χώρος των πολυωνύμων, με πραγματικούς συντελεστές, εφοδιασμένος με την ίδια νόρμα της ομοιόμορφης σύγκλισης επί του $[a,\beta]$).***

(ii). ***Να υπάρχει μία σταθερά M, ανεξάρτητη από το n, τέτοια ώστε:***

$$\sum_{i=0}^{n} \left|A_i^{(n)}\right| \le M \quad \forall n. \blacksquare$$

Για τους μαθηματικούς Τύπους τετραγωνισμού, που στηρίζονται στην Θεωρία Παρεμβολής (και που, μελετάμε στο συγκεκριμένο *Κεφάλαιο*), η ιδιότητα *(i)* του προηγούμενου Θεωρήματος επαληθεύεται αμέσως από την ίδια την κατασκευή αυτών των τύπων τετραγωνισμού.

Έτσι, η ιδιότητα *(ii)* εμφανίζεται να είναι η ουσιαστικά αναγκαία και ικανή συνθήκη για την ευστάθεια και την σύγκλιση επί του $P(R)$, της μεθόδου τετραγωνισμού Newton-Côtes.

Ας εξετάσουμε εάν αυτή η συνθήκη επαληθεύεται κατά την εφαρμογή της μεθόδου Newton-Côtes.

Έχουμε το ακόλουθο αποτέλεσμα, το οποίο και παραθέτουμε χωρίς απόδειξη:

Θεώρημα IV.2.2.11. ***Εάν τα σημεία $x_0, x_1, \ldots, x_n$ έχουν, ανά δύο διαδοχικά ίση απόσταση μεταξύ τους, τότε οι συντελεστές $A_i^{(n)}$ της μεθόδου τετραγωνισμού Newton-Côtes είναι τέτοιοι ώστε η ποσότητα:***

$$\sum_{i=0}^{n} \left| A_i^{(n)} \right|$$

δεν είναι άνω φραγμένη για κάθε n (και για κάθε $f \in C[a,\beta]$) ■

Πόρισμα IV.2.2.12.(i). ***Η μέθοδος τετραγωνισμού Newton-Côtes είναι ασταθής.***

(ii). ***Υπάρχει τουλάχιστον μία συνάρτηση $f \in C[a,\beta]$ για την οποία η αντίστοιχη μέθοδος τετραγωνισμού Newton-Côtes δεν συγκλίνει.*** ■

Το τελευταίο αποτέλεσμα δεν πρέπει να μας εκπλήξει, επειδή στο *Θεώρημα IV.1.8.4* είχε καταστεί εμφανής η αδυναμία σύγκλισης των ακολουθιών πολυωνύμων παρεμβολής. Από αυτό το τελευταίο αποτέλεσμα, εξ άλλου, δικαιολογείται το γεγονός της περιορισμένης εφαρμογής που τυγχάνουν οι μέθοδοι τετραγωνισμού Newton-Côtes στα πρακτικά προβλήματα.

Υπάρχουν δύο τρόποι για να ανταπεξέλθουμε στην αστάθεια και την μη-σύγκλισης. Από την μία πλευρά, είναι η χρησιμοποίηση των σύνθετων μεθόδων τετραγωνισμού και από την άλλη, βρίσκονται οι δυνατότητες επιλογής των σημείων $x_0, x_1, \ldots, x_n$ με τέτοιο τρόπο ώστε να μην ισαπέχουν διαδοχικά μεταξύ τους. Στις επόμενες *Παραγράφους*, θα μελετήσουμε τους δύο αυτούς τρόπους.

IV.2.3. ΟΙ ΣΥΝΘΕΤΕΣ ΜΕΘΟΔΟΙ

Η βασική ιδέα πάνω στην οποία στηρίζονται οι ***σύνθετες μέθοδοι*** είναι απλή: *Αφού θεωρήσουμε μία διαμέριση:*

$$a = y_0 < y_1 \ldots < y_j < y_{j+1} < \ldots < y_{m-1} < y_m = \beta$$

του διαστήματος $[a, \beta]$, *γράφουμε το προς υπολογισμό ολοκλήρωμα υπό την μορφή:*

$$\int_a^\beta f(x)\,dx = \int_a^{y_1} f(x)\,dx + \int_{y_1}^{y_2} f(x)\,dx + \ldots + \int_{y_j}^{y_{j+1}} f(x)\,dx + \ldots + \int_{y_{m-1}}^{\beta} f(x)\,dx,$$

και υπολογίζουμε, σε κάθε υποδιάστημα $[y_j, y_{j+1}]$, *μία προσεγγιστική τιμή του αντίστοιχου ολοκληρώματος, χρησιμοποιώντας έναν μαθηματικό τύπο τετραγωνισμού Newton-Côtes για* n *αρκετά μικρό.* Στην πράξη, συνήθως, επιλέγουμε $n = 1$ ή $n = 2$, οπότε οδηγούμαστε στις ***μεθόδους τραπεζίων*** και ***Simpson*** αντίστοιχα.

IV.2.4. ΜΕΘΟΔΟΣ ΤΡΑΠΕΖΙΩΝ

Παίρνουμε τα σημεία της διαμέρισης σύμφωνα με τον εξής τρόπο:

$$y_j = a + j\,h\text{, με } h = \frac{\beta - a}{m} \quad (j = 0,1,\ldots,m).$$

Σε κάθε υποδιάστημα $[y_j, y_{j+1}]$ εφαρμόζουμε τον μαθηματικό τύπο τετραγωνισμού Newton-Côtes σε δύο σημεία παρεμβολής, δηλαδή με $n = 1$.

Συγκεκριμένα, σε κάθε υποδιάστημα $[y_j, y_{j+1}]$ θα έχουμε:

$$\int_{y_j}^{y_{j+1}} f(x)\,dx = \frac{h}{2}[f(y_j) + f(y_{j+1})] + R_j(f).$$

Υποθέτοντας ότι η συνάρτηση f είναι δύο φορές συνεχώς διαφορίσιμη στην ανοικτή περιοχή του (κάθε) υποδιαστήματος $[y_j, y_{j+1}]$ και επικαλούμενοι το *Θεώρημα IV.2.2.5*, βλέπουμε ότι:

$$R_j(f) = \frac{h^3}{2}\frac{d^2 f}{dx^2}(\xi)\int_0^1 t(t-1)\,dt$$

$$\Leftrightarrow$$

$$R_j(f) = -\frac{h^3}{12}\frac{d^2 f}{dx^2}(\xi_j)\text{, για κάποιο } \xi_j \in [y_j, y_{j+1}].$$

Άρα:

$$\int_a^\beta f(x)\,dx = \sum_{j=0}^{m+1}\int_{y_j}^{y_j+1} f(x)\,dx = \frac{h}{2} + \sum_{j=0}^{m-1}\left[f(y_j) + f(y_{j+1})\right] - \frac{h^3}{12}\sum_{j=0}^{m-1}\frac{d^2 f}{dx^2}(\xi_j),$$

απ' όπου προκύπτει, τελικά, ο μαθηματικός *Τύπος* της μεθόδου τραπεζίων:

$$\int_a^\beta f(x)\,dx = h\,[\frac{f(a)}{2} + f(a+h) + \ldots + f(a+(m-1)h) + \frac{f(\beta)}{2}] + R(f),$$

με:

$$h = \frac{\beta - a}{m} \text{ **και** } R(f) = -\frac{(\beta-a)^3}{12\,m^2}\frac{d^2 f}{d x^2}(\xi), \text{ **για κάποιο** } \xi \in [a,\beta].$$

Βλέπουμε αμέσως ότι:

$$lim_{m\to\infty} R(f) = 0,$$

και συνεπώς ισχύει το ακόλουθο:

Θεώρημα IV.2.4.1. ***Εάν*** $C^2[a,\beta]$ ***είναι ο διανυσματικός χώρος όλων των πραγματικών συναρτήσεων που είναι δύο φορές συνεχώς διαφορίσιμες στην ανοικτή περιοχή του διαστήματος*** $[a,\beta]$***, τότε ο μαθηματικός τύπος τετραγωνισμού της μεθόδου τραπεζίων συγκλίνει επί του*** $C^2[a,\beta]$***.*** ■

Επί πλέον, βλέπουμε ότι οι συντελεστές $A_j^{(m)}$ της μεθόδου δίνονται από τις σχέσεις:

$$A_0^{(m)} = A_m^{(m)} = \frac{h}{2} \text{ και } A_j^{(m)} = h \text{ (για } j = 1,2,\ldots,m-1).$$

Επομένως:

$$A_j^{(m)} > 0 \quad \forall\, j \text{ και } \sum_{j=0}^{m} A_j^{(m)} = m\,h = \beta - a,$$

απ' όπου, σύμφωνα με το *Θεώρημα IV.2.2.8*, έπεται το εξής αποτέλεσμα:

Θεώρημα IV.2.4.2. ***Η μέθοδος τραπεζίων είναι ευσταθής.*** ■

Ιδιαίτερα, η σχέση

$$\sum_{j=0}^{m} A_j = \beta - a$$

επιτρέπει την γενίκευση του *Θεωρήματος IV.2.4.1* επί του διανυσματικού χώρου $C[a,\beta]$.

Πράγματι, έχοντας υπ' όψη το *Θεώρημα IV.2.2.10* και τη σχέση αυτή, μπορούμε αμέσως να διαπιστώσουμε ότι:

Θεώρημα IV.2.4.3. ***Ο μαθηματικός τύπος τετραγωνισμού της μεθόδου τραπεζίων συγκλίνει επί του διανυσματικού χώρου*** $C[a,\beta]$ ***όλων των πραγματικών συναρτήσεων που είναι συνεχείς σε κάθε σημείο του κλειστού διαστήματος*** $[a,\beta]$**.**■

IV.2.5. ΜΕΘΟΔΟΣ SIMPSON

Η *μέθοδος Simpson* έγκειται στην χρησιμοποίηση του μαθηματικού τύπου τετραγωνισμού Newton-Côtes σε τρία σημεία κάθε υποδιαστήματος της διαμέρισης. Καθώς σ' αυτήν την εφαρμογή του τετραγωνισμού Newton-Côtes υπεισέρχεται το μέσο κάθε υποδιαστήματος, προτιμάμε, για λόγους απλοποίησης των συμβολισμών, να αριθμούμε με άρτιους δείκτες τα άκρα των υποδιαστημάτων και να διαφυλάττουμε τους περιττούς δείκτες για τα μέσα των υποδιαστημάτων.

Με άλλα λόγια, θεωρούμε μία διαμέριση σε ισομήκη υποδιαστήματα:

$$a = y_0 < y_2 < y_4 < \ldots < y_{2.j} < y_{2.j+2} < \ldots < y_m = \beta$$

και βρίσκουμε τα μέσα y_{2j+1} κάθε διαστήματος $[y_{2j}, y_{2j+2}]$:

$$y_1 < y_3 < \ldots < y_{2j+1} < \ldots < y_{m-1},$$

όπου m είναι κάποιος άρτιος αριθμός.

Το προς υπολογισμό ολοκλήρωμα εκφράζεται τότε ως εξής:

$$\int_a^\beta f(x)\,dx = \int_a^{y_2} f(x)\,dx + \int_{y_2}^{y_4} f(x)\,dx + \ldots + \int_{y_{m-2}}^{\beta} f(x)\,dx,$$

με:

$$\int_{y_{2.j}}^{y_{2.j+2}} f(x)\,dx = \frac{h}{3}.[f(y_{2j}) + 4.f(y_{2j+1}) + f(y_{2j+2})] + R_j(f)$$

και:

$$h = \frac{\beta - a}{m} \quad (m = \text{άρτιος}).$$

Λαμβάνοντας υπ' όψη τα αποτελέσματα του *Θεωρήματος IV.2.2.5*, και υποθέτοντας ότι η συνάρτηση *f* είναι τέσσερις φορές συνεχώς διαφορίσιμη μέσα σε μία ανοικτή περιοχή του διαστήματος $[a, \beta]$, βλέπουμε ότι για κάθε j υπάρχει $\xi_j \in [y_{2j}, y_{2j+2}]$ τέτοιο ώστε:

$$R_j(f) = \frac{h^5}{4!}\frac{d^4 f}{dx^4}\int_0^2 t^2\,(t-1)\,(t-2)\,dx$$

$$\Leftrightarrow$$

$$R_j(f) = -\frac{h^5}{90}\frac{d^4 f}{dx^4}(\xi_j).$$

Έτσι, σ' αυτήν την περίπτωση, το προς υπολογισμό ολοκλήρωμα γράφεται υπό την μορφή:

$$\int_a^\beta f(x)\,dx = \sum_{j=0}^{m/2-1}\int_{j=0}^{y_{2.j+2}} f(x)\,dx$$

$$= \frac{h}{3}\sum_{j=0}^{m/2-1}[f(y_{2j}) + 4\,f(y_{2j+1}) + f(y_{2j+2})] - \frac{h^5}{90}\sum_{j=0}^{(m/2)-1}\frac{d^4 f}{dx^4}(\xi_j),$$

με

$$h = \frac{\beta - a}{m} \quad \text{και } m = \text{άρτιος}.$$

Από αυτήν την τελευταία ισότητα προκύπτει η επεξηγηματική και αναλυτική έκφραση του μαθηματικού *Τύπου* τετραγωνισμού Simpson, στην περίπτωση όπου η συνάρτηση f είναι τέσσερις φορές συνεχώς διαφορίσιμη στην ανοικτή περιοχή του $[a,\beta]$:

$$\int_a^\beta f(x)\,dx = \frac{h}{3}\left[f(a)+4f(a+h)+2f(a+2h)+\cdots\right.$$

$$\left.\cdots+2f(a+(m-2)h)+4f(a+(m-1)h)+f(\beta)\right]+R(f),$$

με:

$$h=\frac{\beta-a}{m},\quad m=\text{άρτιος και}\quad R(f)=-\frac{(\beta-a)^5}{180\,m^4}\frac{d^4 f}{d\,x^4}(\xi),$$

για κάποιο σημείο $\xi\in[a,\beta]$.

Βλέπουμε αμέσως ότι:

$$\lim_{m\to\infty} R(f)=0,$$

και συνεπώς ισχύει το ακόλουθο:

Θεώρημα IV.2.5.1. ***Εάν $C^4[a,\beta]$ είναι διανυσματικός χώρος όλων των πραγματικών συναρτήσεων που είναι τέσσερις φορές συνεχώς διαφορίσιμες στην ανοικτή περιοχή του διαστήματος $[a,\beta]$, τότε ο μαθηματικός Τύπος τετραγωνισμού της μεθόδου Simpson συγκλίνει επί του $C^4[a,\beta]$.*** ■

Επί πλέον, βλέπουμε ότι οι συντελεστές $A_j^{(m)}$ της μεθόδου Simpson δίνονται από τις σχέσεις:

$$A_0^{(m)} = A_m^{(m)} = \frac{h}{3} \quad \text{και} \quad \left\{ \begin{array}{ll} A_{2j+1}^{(m)} = \dfrac{4h}{3} & (\text{για } j = 0,1,\ldots,\dfrac{m}{2}-1) \\ A_{2j}^{(m)} = \dfrac{2h}{3} & (\text{για } j = 1,\ldots,\dfrac{m}{2}-1) \end{array} \right\}.$$

Επομένως:

$$A_j^{(m)} > 0 \ \ \forall\, j \ \ \text{και} \ \ \sum_{j=0}^{m} A_j^{(m)} = \frac{2h}{3} + \frac{m}{2}\frac{4h}{3} + \left(\frac{m}{2}-1\right)\frac{2h}{3} = \beta - a$$

Σύμφωνα με το *Θεώρημα IV.2.2.8*, η τελευταία ιδιότητα οδηγεί στο συμπέρασμα ότι:

Θεώρημα IV.2.5.2. ***Η μέθοδος Simpson είναι ευσταθής.***■

Ιδιαίτερα, η σχέση

$$\sum_{j=0}^{m} A_j^{(m)} = \beta - a$$

επιτρέπει την επέκταση του *Θεωρήματος IV.2.5.1* επί του διανυσματικού χώρου $C[a,\beta]$. Πράγματι, έχοντας υπ' όψη το περιεχόμενο του *Θεωρήματος IV.2.2.10* και την σχέση αυτή, μπορούμε αμέσως να διαπιστώσουμε το επόμενο:

Θεώρημα IV.2.5.3. ***Ο μαθηματικός τύπος τετραγωνισμού της μεθόδου Simpson συγκλίνει επί του διανυσματικού χώρου*** $C[a,\beta]$ ***όλων των πραγματικών συναρτήσεων που είναι συνεχείς σε κάθε σημείο του κλειστού διαστήματος*** $C[a,\beta]$.■

Παρατήρηση IV.2.5.4. Τόσο στην μέθοδο τραπεζίων, όσο και στην μέθοδο Simpson, η συνθήκη:

$$\sum_{j=0}^{m} A_j^{(m)} = \beta - a$$

συνεπάγεται, απλά, ότι η μέθοδος τετραγωνισμού είναι ακριβής για $f(x) = 1$ (με την έννοια που επισημάνθηκε αμέσως μετά την εκφώνηση του *Θεωρήματος IV.2.1.1*). Πράγματι, είναι:

$$\int_a^\beta f(x)dx = \int_a^\beta dx = \beta - a$$

και:

$$\sum_{j=0}^{m} A_j^{(m)} f(x_j) = \sum_{j=0}^{m} A_j^{(m)} = \beta - a \,.\blacksquare$$

IV.2.6. ΜΕΘΟΔΟΣ ROMBERG

Η ***μέθοδος Romberg*** συνίσταται στην εύρεση κάποιων προσεγγιστικών τιμών του προς υπολογισμό ολοκληρώματος, με την βοήθεια της μεθόδου τραπεζίων, για διαφορετικά μεγέθη ευρύτητας του βήματος h:

$h = h_0, h_1, \ldots$

Στην συνέχεια, και με την βοήθεια του σχήματος Neville-Aitken, παρεμβάλλουμε τις προσεγγιστικές τιμές, που βρέθηκαν κατά το πρώτο στάδιο, κατασκευάζοντας ένα πολυώνυμο παρεμβολής.

Τέλος, υπολογίζουμε την τιμή αυτού του πολυωνύμου στο σημείο $h = 0$.

Γενικά, η τιμή του πολυωνύμου παρεμβολής στο 0 είναι μία σαφώς πιο βελτιωμένη προσέγγιση του ολοκληρώματος, συγκριτικά με όλες τις προσεγγιστικές τιμές, που είχαν προκύψει από τις διαδοχικές εφαρμογές της μεθόδου τραπεζίων και που απετέλεσαν τα σημεία παρεμβολής.

Κατ' αυτόν τον τρόπο, επιτυγχάνουμε την δημιουργία μαθηματικών *Τύπων* τετραγωνισμού, οι οποίοι είναι ακριβείς για πολυώνυμα με βαθμό όλο και περισσότερο πιο μεγάλο.

Επίσης, όπως θα δούμε αυτοί οι *Τύποι* είναι ευσταθείς και συγκλίνοντες.

Πριν προχωρήσουμε στην ανάπτυξη της μεθόδου Romberg, είναι ανάγκη να θέσουμε το θεωρητικό υπόβαθρο πάνω στο οποίο η μέθοδος θεμελιώνεται. Πρόκειται για τις έννοιες των πολυωνύμων και των αριθμών Bernoulli, καθώς και για τον τύπο Euler-Mac Laurin.

IV.2.6.(i). Πολυώνυμα και αριθμοί Bernoulli Τα πολυώνυμα και οι αριθμοί Bernoulli μπορούν να ορισθούν με διάφορους τρόπους.

Γενικά, και προς τούτο, θεωρούμε το ανάπτυγμα σε σειρά δυνάμεων της συνάρτησης:

$$\frac{t\, e^{t\,x}}{e^{t}-1}.$$

Μία τέτοια παρουσίαση δεν μας φάνηκε προσαρμοσμένη στους στόχους του παρόντος βιβλίου, γιατί επισύρει την ανάπτυξη επιχειρημάτων και εννοιών που δεν μπορούν να χρησιμεύσουν στην συνέχεια των προσπαθειών μας.

Θα πρέπει επίσης να επισημανθεί η ποικιλία των παραλλαγών που υπάρχουν στους ορισμούς των πολυωνύμων Bernoulli, και που μπορεί να συναντήσει κανείς στην σχετική βιβλιογραφία. Σε ό,τι ακολουθήσει, θα υιοθετήσουμε την θεωρητική θεμελίωσή τους, όπως αυτή προτείνεται στο παιδαγωγικό σύγγραμμα $[4]$.

Τα ***πολυώνυμα Bernoulli*** ορίζονται σύμφωνα με τις τέσσερις επόμενες ιδιότητες, που επιτρέπουν τον πλήρη προσδιορισμό τους:

- $B_0(\mathrm{x}) := 1$,
- $B_1(x) := x - \frac{1}{2}$,
- $B_n'(x) = B_{n-1}(x)$, για $n > 1$,
- $B_n(1) = B_n(0)$, για $n > 1$.

Βλέπουμε ότι, σύμφωνα με την τρίτη ιδιότητα, ο προσδιορισμός του $B_n(x)$ προέρχεται από την εκμετάλλευση του $B_{n-1}(x)$, κατά προσέγγιση μίας προσθετικής σταθεράς. Η τέταρτη επιτρέπει τον υπολογισμό αυτής της σταθεράς.

Οι ***αριθμοί Bernoulli*** B_n ορίζονται από την σχέση:

- $B_n := n!.B_n(0) \quad (n = 0,1,2,...)$.

Μπορούμε να αποδείξουμε χωρίς ιδιαίτερη δυσκολία ότι:

- $B_{2.n+1} = 0 \quad \forall \quad n = 0,1,2,...$

IV.2.6. (ii). Τύπος Euler-Mac Laurin Αρχίζουμε την *Υποπαράγραφο* αυτή με τον υπολογισμό του ολοκληρώματος:

$$R_k = \int_0^1 B_k(t)\,\frac{d^k g}{dk^k}(t)\,dt\,,$$

όπου $g(t))$ είναι μία πραγματική συνάρτηση που είναι επαρκώς διαφορίσιμη μέσα σε μία ανοικτή περιοχή του διαστήματος $[0,1]$.

Μετά από μία πραγματική ολοκλήρωση, λαμβάνουμε:

$$R_k = \left[B_k(t)\frac{d^{(k-1)}g}{d\,t^{(k-1)}}(t)\right]_{t=0}^{t=1} - \int_0^1 \frac{d\,B_k}{d\,t}(t)\frac{d^{(k-1)}g}{d\,t^{(k-1)}}(t)dt\,,$$

και κάνοντας χρήση της ιδιότητας:

$$B_k(0) = B_k(1)\ \ (k>1)\,,$$

καθώς και της ιδιότητας:

$$B_k'\ (t) = B_{k-1}(t)\ \ (k>1)\,,$$

διαπιστώνουμε πως:

$$\bullet\ R_k = \frac{B_k}{k!}\,[\frac{d^{(k-1)}g}{dt^{(k-1)}}(1) - \frac{d^{(k-1)}g}{dt^{(k-1)}}(0)] - R_{k-1}\,,$$

και, διαδοχικά ότι:

$$\bullet\ R_k = \frac{B_k}{k!}\,[\frac{d^{(k-1)}g}{dt^{(k-1)}}(1) - \frac{d^{(k-1)}g}{dt^{(k-1)}}(0)] - \frac{B_{k-1}}{(\mathrm{k}-1)!}\,[\frac{d^{(k-2)}g}{dt^{(k-2)}}(1) - \frac{d^{(k-2)}g}{dt^{(k-2)}}(0)] - R_{k-2}\,.$$

Ιδιαίτερα, για $k = 2\,p$, επειδή $B_{2\,p-1} = 0\ \ (p = 2,3,...)$, έχουμε:

$$\bullet\ R_{2p} = \frac{B_{2p}}{(2p)!}\left[\frac{d^{(2p-1)}g}{dt^{(2p-1)}}(1) - \frac{d^{(2p-1)}g}{dt^{(2p-1)}}(0)\right] + R_{2p-2}\,.$$

Η περίπτωση $k=2$ εξετάζεται μεμονωμένα, χρησιμοποιώντας τα ίδια επιχειρήματα. Συγκεκριμένα, πάλι, μετά από μία παραγοντική ολοκλήρωση, παίρνουμε:

$$R_2 = [B_2(t)\,\frac{dg}{dt}(t)]_{t=0}^{t=1} - \int_0^1 \frac{dB_k(t)}{dt}\,\frac{dg}{dt}(t)\,dt\,,$$

οπότε, εφαρμόζοντας την ιδιότητα $B_2(0) = B_2(1) = B_2$, βλέπουμε ότι:

$$R_2 = \frac{B_2}{2}\left[\frac{dg}{dt}(1) - \frac{dg}{dt}(0)\right]_{t=0}^{t=1} - R_1\,.$$

Επειδή όμως:

$$R_1 = [B_1(t)\,g(t)]_{t=0}^{t=1} - \int_0^1 \frac{dB(t)}{dt}\,g(t)\,dt = \frac{1}{2}[g(1)-g(0)] - \int_0^1 g(t)\,dt\,,$$

έχουμε, τελικά, ότι:

$$R_2 = \frac{B_2}{2}\left[\frac{dg}{dt}(1) + \frac{dg}{dt}(0)\right]_{t=0}^{t=1} - \frac{1}{2}[g(1)-g(0)] + \int_0^1 g(t)\,dt\,.$$

Έτσι, συνδυάζοντας τους δύο *Τύπους* που βρήκαμε για τις ποσότητες R_{2p} $(p = 2,3,\ldots)$ και R_2, αντίστοιχα, και ακολούθως, λύνοντας την εξίσωση που θα προκύψει ως προς

$$\int_0^1 g(t)dt\,,$$

καταλήγουμε στον γενικό *Τύπο*:

$$\int_0^1 g(t)dt = \frac{1}{2}[g(0)+g(1)] - \frac{B_2}{2}[\frac{dg}{dt}(1) - \frac{dg}{dt}(0)] - \cdots$$

$$\cdots - \frac{B_{2n}}{(2n)!}[\frac{d^{(2n-1)}g}{dt^{(2n-1)}}(0) - \frac{d^{(2n-1)}g}{dt^{(2n-1)}}(0)] + R_{2n},$$

κάτω, φυσικά, από την προϋπόθεση ότι η συνάρτηση g είναι (τουλάχιστον) $2n-$φορές συνεχώς διαφορίσιμη στην ανοικτή περιοχή του $[0,1]$.

Πραγματοποιούμε τώρα την αλλαγή μεταβλητής:

$$x := x_i + ht,$$

όπου x_i και h είναι δύο πραγματικές παράμετροι.

Επί πλέον, θέτουμε:

$$f(x) := g(t).$$

Αντικαθιστώντας στον προηγούμενο γενικό τύπο, επιτυγχάνουμε την κλασσική έκφραση του τύπου Euler- Mac Laurin:

$$\int_{x_i}^{x_i+h} f(x)dx = \frac{h}{2}[f(x_i)+f(x_{i+1})] - h^2\frac{B_2}{2}[\frac{df}{dx}(x_{i+1}) - \frac{df}{dx}(x_i)] - \cdots$$

$$\cdots - h^{2n}\frac{B_{2n}}{(2n)!}[\frac{d^{(2n-1)}f}{dx^{(2n-1)}}(x_{i+1}) - \frac{d^{(2n-1)}f}{dx^{(2n-1)}}(x_i)] + R_{2n},$$

με:

$$x_{i+1} = x_i + h.$$

Εάν γράψουμε τον τύπο αυτόν για $i = 0,1,\ldots,n-1$, με $x_0 = a$ και $x_n = \beta$, και εάν αθροίσουμε όλα τα αποτελέσματα, θα προκύψει:

$$\int_a^\beta f(x)dx = h\,[\frac{f(a)}{2} + f(a+h) + \cdots + f(a+(n-1)h) + \frac{f(\beta)}{2}] -$$

$$- h^2 \frac{B_2}{2}[f'(\beta) - f'(a)] - - h^4 \frac{B_4}{4}[f'''(\beta) - f'''(a)] - \cdots\cdots$$

Βλέπουμε λοιπόν ότι αυτός ο τύπος μας επιτρέπει, για την μέθοδο τραπεζίων, να επιτύχουμε μία έκφραση του σφάλματος τετραγωνισμού, που είναι πιο πλήρης σε σύγκριση με εκείνη που γνωρίζαμε μέχρι τώρα.

Πιο συγκεκριμένα, βλέπουμε ότι το εν λόγω σφάλμα ισούται με:

$$R(f) = a_1 h^2 + a_2 h^4 + \ldots + a_n h^{2n} + \frac{B_{2n+2}}{(2n+2)!} h^{2n+2} \frac{d^{(2n+2)} f}{dx^{(2n+2)}}(\xi),$$

για κάποιο $\xi \in [a,\beta]$, και κάτω, φυσικά από την προϋπόθεση ότι η συνάρτηση f είναι $(2n+2)$ φορές συνεχώς διαφορίσιμη μέσα σε μία ανοικτή περιοχή του κλειστού διαστήματος $[a,\beta]$. (Είναι σαφές ότι για την επίτευξη του τελευταίου όρου στο δεξιό μέλος της παραπάνω ισότητας, πρέπει να εφαρμοσθεί το *Θεώρημα Μέσης Τιμής* στην έκφραση του R_{2n}, που είχε βρεθεί νωρίτερα).

Αυτή η καινούργια πιο πλήρης μαθηματική διατύπωση του σφάλματος τετραγωνισμού για την μέθοδο τραπεζίων, αποτελεί την θεωρητική βάση, γύρω από την οποία στοιχειοθετείται και αναπτύσσεται η μέθοδος Romberg, που θα μελετήσουμε στην *Υποπαράγραφο* που ακολουθεί.

IV.2.6.(iii). Αλγόριθμος Romberg Όπως είδαμε στην προηγούμενη *Υποπαράγραφο*, το σφάλμα τετραγωνισμού της μεθόδου τραπεζίων είναι ένα πολυώνυμο ως προς h^2.

Μπορούμε, συνεπώς, να εφαρμόσουμε πολλές, διαδοχικές φορές την μέθοδο τραπεζίων, με διαφορετικά βήματα h κάθε φορά, και, ακολούθως, να αφήσουμε να διέλθει από τις διαφορετικές τιμές, που θα προκύψουν, ένα πολυώνυμο παρεμβολής, με την βοήθεια του σχήματος Neville-Aitken.

Η τιμή αυτού του πολυωνύμου στο σημείο $h = 0$ θα αποτελεί μία καλή προσέγγιση του προς υπολογισμό ολοκληρώματος.

Ας αρχίσουμε, συμβολίζοντας με:

$$T_0^{(n)}$$

το αποτέλεσμα που έχει βρεθεί κατόπιν μίας εφαρμογής της μεθόδου τραπεζίων κατά την εκτέλεση της οποίας το αντίστοιχο βήμα ήταν ίσο με h_n.

Στον *Τύπο του Σχήματος Neville-Aitken* θέτουμε:

$$x_n := h_n^2 \text{ και } x := 0.$$

Εάν:

$$T_k^{(n)} := T_k^{(n)}(0),$$

τότε:

$$T_{k+1}^{(n)} = \frac{h_{n+k+1}^2\, T_k^{(n)} - h_n^2\, T_k^{(n+1)}}{h_{n+k+1}^2 - h_n^2},$$

για κάθε $n = 0,1,2,\ldots$ και κάθε $k = 0,1,2,\ldots$

Επιλέγουμε τώρα τα βήματα h_n έτσι ώστε:

$$h_n > h_{n+1} > 0 \ \ \forall \ \ n = 0,1,2,... \text{ και } \lim_{n\to\infty} h_n = 0 .$$

Εάν η f είναι δύο φορές συνεχώς διαφορίσιμη μέσα στην ανοικτή περιοχή του $[a,\beta]$, τότε, σύμφωνα με το *Θεώρημα IV.2.4.1*, θα είναι:

$$\lim_{n\to\infty} T_0^{(n)} = \int_a^\beta f(x)dx .$$

Θα εξετάσουμε τώρα την σύγκλιση προς το ολοκλήρωμα $\int_a^\beta f(x)\,dx$:

τόσο της ακολουθίας $(T_k^{(n)} : n = 0,1,2,...) = (T_k^{(0)}, T_k^{(1)}, T_{(k)}^{(2)},...)$, για κάθε $k =$ σταθερό,

όσο και της ακολουθίας $(T_k^{(n)} : k = 0,1,2,...) = (T_0^{(n)}, T_1^{(n)}, T_2^{(n)},...)$, για κάθε $n =$ σταθερό.

Για λόγους απλοποίησης της γραφής, θέτουμε:

$$I = I(f) := \int_a^\beta f(x)dx .$$

Όσον αφορά την πρώτη περίπτωση, όπου $k =$ σταθερό, έχουμε το:

Θεώρημα IV.2.6.(iii).1. ***Μία ικανή και αναγκαία συνθήκη, προκειμένου να ισχύει:***

$$\lim_{n\to\infty} T_k^{(n)} = \mathrm{I}, \text{ για } k = 1,2,...$$

για κάθε συνάρτηση $f \in C^2[a,\beta]$***, είναι η ύπαρξη μίας σταθεράς:***

$$A > 1,$$

τέτοιας ώστε:

$$\frac{h_n}{h_{n+1}} \geq A, \ \ \forall n .$$

Απόδειξη. Παρατηρούμε ότι:

$$T_{k+1}^{(n)} = \frac{T_k^{(n)} - \dfrac{h_n^2}{h_{n+k+1}^2} T_k^{(n+1)}}{1 - \dfrac{h_n^2}{h_{n+k+1}^2}} \quad (\forall n, \forall k).$$

Έτσι, εάν

$$\lim_{n\to\infty} T_k^{(n)} = \mathrm{I},$$

τότε μία ικανή και αναγκαία συνθήκη προκειμένου να ισχύει

$$\lim_{n\to\infty} T_{k+1}^{(n)} = \mathrm{I}$$

είναι να υπάρχει μία σταθερά $A > 1$, τέτοια ώστε:

$$\frac{h_n^2}{h_{n+k+1}^2} \geq A > 1 \quad \forall n \ \text{και} \ \forall k,$$

απ' όπου έπεται η συνθήκη του *Θεωρήματος*. ■

Πριν προχωρήσουμε στην δεύτερη περίπτωση, όπου $n =$ σταθερό, θα αναφέρουμε μία βοηθητική πρόταση την οποία θα χρειαστούμε:

Λήμμα IV.2.6.(iii).2. ***Εάν*** $f \in \mathcal{P}_{2k+1}(\mathbb{R})$, ***τότε:***

$$T_n^{(k)} = \mathrm{I}, \ \textbf{\textit{για κάθε}} \ n = 0,1,2,\ldots$$

Απόδειξη. Είναι μία άμεση συνέπεια του *Τύπου Euler-Mac Laurin*, καθώς και του ίδιου του ορισμού του πολυωνύμου παρεμβολής. ■

Είμαστε τώρα σε θέση να μελετήσουμε την σύγκλιση προς το ολοκλήρωμα *I* της ακολουθίας $(T_k^{(n)} : k = 0,1,2,...)$, για $n =$ σταθερό. Έχουμε το:

Θεώρημα IV.2.6.(iii).3. ***Μία αναγκαία και ικανή συνθήκη προκειμένου να ισχύει:***

$$\lim_{k\to\infty} T_k^{(n)} = \mathrm{I}, \text{ για } n = 0,1,2,...$$

για κάθε συνάρτηση $f \in C^2[a,\beta]$***, είναι η ύπαρξη μίας σταθεράς*** $A > 1$ ***τέτοιας ώστε:***

$$\frac{h_n}{h_{n+1}} \geq A, \quad \forall n.$$

Απόδειξη. Δεν θα δώσουμε παρά μόνον την κατευθυντήρια ιδέα της *Απόδειξης*. Για τις λεπτομέρειες των υπολογισμών, ο κάθε ενδιαφερόμενος αναγνώστης παραπέμπεται στο εξαιρετικό βιβλίο $[41]$.

Γράφοντας την τιμή στο σημείο 0 του πολυωνύμου παρεμβολής Lagrange, παρατηρούμε ότι:

$$T_k^{(n)} = \sum_{i=n}^{n+k} A_i^{(k)}\, T_0^{(i)},$$

με:

$$A_i^{(k)} = \prod_{\substack{j=n \\ j\neq i}}^{n+k} \frac{1}{1-\dfrac{h_i^2}{h_j^2}}.$$

Έτσι, εάν $f \in C^2[a,\beta]$, τότε, σύμφωνα με το *Θεώρημα IV.2.4.1*, ισχύει:

$$\lim_{n\to\infty} T_0^{(n)} = \mathrm{I}.$$

Από την άλλη πλευρά, σύμφωνα με το *Θεώρημα IV.1.8.2*, βλέπουμε ότι μία αναγκαία και ικανή συνθήκη προκειμένου να είναι:

$$\lim_{n\to\infty} T_k^{(n)} = \mathrm{I}, \text{ για } n = 0,1,2,...$$

για κάθε συνάρτηση $f \in C^2[a,\beta]$ είναι η ταυτόχρονη ισχύς των επόμενων δύο ιδιοτήτων:

- $\lim_{n\to\infty} T_k^{(n)} = \mathrm{I}$, για κάθε $f \in \mathbb{P}(\mathbb{R})$, και
- $\exists M \in \mathbb{N} : \sum_{i=n}^{n+k} \left| A_i^{(k)} \right| \le M \quad \forall k$.

Η πρώτη ιδιότητα επαληθεύεται με την επίκληση του *Λήμματος IV.2.6.(iii).2*. Η δεύτερη ιδιότητα ικανοποιείται όταν και μόνον όταν υπάρχει μία σταθερά $A > 1$, τέτοια ώστε $h_n \ge h_{n+1} A, \quad \forall n$. Η επαλήθευση αυτού του ισχυρισμού παραλείπεται. Σημειώνουμε όμως ότι από τον ισχυρισμό αυτό συνεπάγεται το επόμενο:

Θεώρημα IV.2.6.(iii).4. ***Εάν τα βήματα h_n επιλεγούν έτσι ώστε:***

$$\frac{h_n}{h_{n+1}} \ge A > 1, \quad \forall n,$$

(για κάποιο σταθερό αριθμό $A > 1$), *τότε η αντίστοιχη μέθοδος τετραγωνισμού, που έγκειται στην κατασκευαστική διασύνδεση κάθε συνάρτησης $f \in C^2[a,\beta]$ με την ακολουθία:*

$$(T_0^{(n)}, T_1^{(n)}, T_2^{(n)}, \ldots)$$

είναι ευσταθής, για κάθε $n = 0,1,2,\ldots$ ■

Από πρακτικής άποψης, τα βήματα $h_0, h_1, h_2 \ldots$ δεν επιλέγονται τυχαία και με μοναδική προοπτική να ικανοποιείται η συνθήκη:

$$\frac{h_n}{h_{n+1}} \ge A > 1 \quad \forall n.$$

Αντίθετα, τα βήματα h_n δημιουργούνται με κανονικό και φυσικό τρόπο, και με κριτήριο την, όσο το δυνατόν, μεγαλύτερη οικονομία στον αριθμό των απαιτούμενων υπολογισμών.

Έτσι, αν επιλέξουμε την επόμενη ακολουθία βημάτων:

$$h_0, h_1 = \frac{h_0}{2}, h_2 = \frac{h_1}{2}, \ldots$$

τότε, κάθε φορά που εφαρμόζεται ο τύπος των τραπεζίων, το μισό πλήθος τιμών των $f(x_i)$, που χρησιμοποιούνται, είναι αυτές που υπολογίσθηκαν την προηγούμενη φορά που εφαρμόσθηκε ο ίδιος τύπος. Η διαδικασία μεταφοράς των βημάτων:

$$h_n = \frac{h_0}{2^n}, \quad n = 0,1,2,\ldots$$

στο σχήμα Neville-Aitken οδηγεί κατευθείαν στον ***αυθεντικό αλγόριθμο Romberg***:

$$T_{k+1}^{(n)} = \frac{2^{2k+2} T_k^{(n+1)} - T_k^{(n)}}{2^{2k+2} - 1}, \quad n = 0,1,2,\ldots \text{ και } k = 0,1,2\ldots$$

Παρατήρηση IV.2.6.(iii).5. Η μέθοδος που περιγράφεται από τον αλγόριθμο Romberg ανταποκρίνεται στις υποθέσεις των *Θεωρημάτων IV.2.6.(iii).1*, *IV.2.6.(iii).3* και *IV.2.6.(iii).4*. ■

Παρατήρηση IV.2.6.(iii).6. Η μέθοδος που αποδίδεται από τον αλγόριθμο Romberg είναι πολύ διαδεδομένη στις πρακτικές εφαρμογές. Ας δώσουμε ένα παράδειγμα που καταδεικνύει την αποτελεσματικότητα της.

Έστω ότι δίνεται προς υπολογισμό το ολοκλήρωμα:

$$I = \int_0^1 \frac{dx}{x + 0.01} = 4{,}615120517 .$$

Εάν

$$h_0 = \frac{1}{3},$$

τότε η μέθοδος τραπεζίων δίνει τα ακόλουθα αποτελέσματα:

$$T_0^{(0)} = 18.29 \text{ και } T_0^{(8)} = 4.616,$$

ενώ η εφαρμογή του αλγορίθμου Romberg στις *9* αυτές τιμές δίνει το αποτέλεσμα:

$$T_8^{(0)} = 4{,}615120793 . \blacksquare$$

IV.2.7. ΟΙ ΜΕΘΟΔΟΙ GAUSS

Στην αρχή αυτού του *Κεφαλαίου*, είδαμε ότι η επιλογή σημείων x_i που διαδοχικά ισαπέχουν και η αξιοποίησή τους μέσα από απόπειρες τετραγωνισμού (τύπου παρεμβολής) οδηγούσε στις μεθόδους Newton-Côtes, που, αφ' ενός, είναι ακριβείς επί του διανυσματικού χώρου $\mathcal{P}_n(\mathbb{R})$, των πραγματικών πολυωνύμων βαθμού το πολύ n, αλλά που, αφ' ετέρου, είναι ασταθείς και μη συγκλίνουσες επί του διανυσματικού χώρου $C[a,\beta]$, των πραγματικών συναρτήσεων που είναι συνεχείς στο κλειστό διάστημα $[a,\beta]$.

Στην συνέχεια, θα δοκιμάσουμε να επιλέξουμε τα σημεία x_i με βέλτιστο τρόπο, δηλαδή έτσι ώστε ο αριθμητικός τετραγωνισμός

$$\sum_{i=0}^{n} A_i^{(n)} f(x_i)$$

τύπου παρεμβολής, που θα προκύψει, να είναι ακριβής επί του διανυσματικού χώρου $\mathcal{P}_{2n+1}(\mathbb{R})$. Μία τέτοια μέθοδος τετραγωνισμού ονομάζεται ***μέθοδος Gauss***.

Στην προσπάθειά μας αυτή, θα εξετάσουμε τις συνθήκες ύπαρξης μίας τέτοιας μεθόδου, θα μελετήσουμε τις ιδιότητές της και θα επιχειρήσουμε την πραγματοποίηση της

κατασκευής της, δηλαδή θα επιχειρήσουμε να επιτύχουμε τον προσδιορισμό των σημείων x_i καθώς και των ποσοτήτων $A_i^{(n)}$.

Θα διατηρήσουμε τους συμβολισμούς των προηγούμενων *Παραγράφων*. Έτσι, θα γράφουμε:

$$I = \int_a^\beta f(x)\, w(x)\, dx\,,$$

όπου $f(x)$ είναι μία πραγματική συνάρτηση, ολοκληρώσιμη στο κλειστό διάστημα $[a,\beta]$ (ή σε μία ανοικτή περιοχή του), και όπου $w(x) \geq 0$, όταν $x \in [a,\beta]$, με την συνάρτηση $w(x)$ ολοκληρώσιμη στο $[a,\beta]$.

Θα υπολογίσουμε το ολοκλήρωμα I, με την βοήθεια του ακόλουθου τύπου τετραγωνισμού:

$$I = \sum_{i=0}^{n} A_i^{(n)}\, f(x_i) + R_n(f)\,,$$

όπου ο αριθμός $R_n(f)$ αναπαριστά το σφάλμα του τετραγωνισμού

$$R_n(f) := \int_a^\beta E_n(x)\, w(x)\, dx$$

όπου $E_n(x)$ είναι το αντίστοιχο σφάλμα παρεμβολής δηλαδή

$$E_n(x) = f(x) - p_n(x)\,.$$

Τέλος, όπως και προηγουμένως, θέτουμε:

$$V_n(x) := (x - x_0)\,(x - x_1)\ldots(x - x_n)\,.$$

IV.2.8. ΥΠΑΡΞΗ ΚΑΙ ΙΔΙΟΤΗΤΕΣ ΤΩΝ ΤΥΠΩΝ ΤΟΥ GAUSS

Ας εξετάσουμε, κατ' αρχάς, την συνθήκη που καθορίζει εάν ένας τέτοιος τύπος είναι ακριβής επί του $\mathcal{P}_{2n+1}(\mathcal{R})$:

Θεώρημα IV.2.8.1. ***Ο Τύπος Gauss είναι ακριβής επί του $\mathcal{P}_{2n+1}(\mathcal{R})$ εάν και μόνον εάν:***

$$\int_a^\beta x^p\ V_n(x)\ w(x)\ dx = 0\,, \forall p = 0,1,\ldots,n\,..$$

Απόδειξη. Θα δείξουμε πρώτα ότι η πιο πάνω συνθήκη είναι αναγκαία. Προς τούτο, ας υποθέσουμε ο τύπος Gauss είναι ακριβής επί του $\mathcal{P}_{2n+1}(\mathcal{R})$.

Τότε, για κάθε πολυωνυμική συνάρτηση $f(x) = x^p\ V_n(x)\ \ (p = 0,1,\ldots,n)$, θα πρέπει να ισχύει:

$$\int_a^\beta x^p\ V_n(x)\ w(x)\ dx = \sum_{i=0}^n A_i^{(n)}\ x_i^p\ V_n(x_i)\,.$$

Καθώς όμως, εξ ορισμού, συμβαίνει $V_n(x_i) = 0$, έπεται ότι:

$$\int_a^\beta x^p\ V_n(x)\ w(x)\ dx = 0\,,\ \forall p = 0,1,\ldots,n\,,$$

δηλαδή ότι η συνθήκη του *Θεωρήματος* είναι αναγκαία.

Θα δείξουμε τώρα ότι η ίδια συνθήκη είναι, επί πλέον, και ικανή. Έστω λοιπόν ότι είναι:

$$\int_a^\beta x^p\ V_n(x)\ w(x)\ dx = 0\,,\ \forall p = 0,1,\ldots,n\,.$$

Τότε, συνδυάζοντας γραμμικά αυτές τις ισότητες, βλέπουμε ότι θα πρέπει:

$$\int_a^\beta Q(x)\, V_n(x)\, w(x)\, dx = 0\,,$$

για κάθε $Q \in \mathbb{P}_n(\mathbb{R})$.

Ας επιλέξουμε ένα οποιοδήποτε πολυώνυμο $R(x) \in \mathbb{P}_n(\mathbb{R})$ και ας θέσουμε:

$$P(x) := V_n(x)\, Q(x) + R(x)\,.$$

Είναι φανερό πως $P(x) \in \mathbb{P}_{2n+1}(\mathbb{R})$, και ακόμη, πως, ολοκληρώνοντας, λαμβάνουμε:

$$\int_a^\beta P(x)\, w(x)\, dx = \int_a^\beta V_n(x)\, Q(x)\, w(x)\, dx + \int_a^\beta R(x)\, w(x)\, dx\,.$$

Το πρώτο ολοκλήρωμα του δεξιού μέλους της τελευταίας ισότητας είναι μηδέν, σύμφωνα με όσα προηγήθηκαν.

Όσον αφορά στο δεύτερο ολοκλήρωμα του ίδιου μέλους, παρατηρούμε ότι, επειδή ο τετραγωνισμός Gauss είναι του τύπου παρεμβολής, ισχύει:

$$\int_a^\beta R(x)\, w(x)\, dx = \sum_{i=0}^{n} A_i^{(n)}\, R(x_i)\,.$$

Καθώς όμως $R(x_i) = P(x_i)$, αφού $V_n(x_i) = 0$, έπεται ότι:

$$\int_a^\beta P(x)\, w(x)\, dx = 0\,,$$

που αποδεικνύει ότι ο τύπος τετραγωνισμού Gauss είναι ακριβής επί του $\mathbb{P}_{2n+1}(\mathbb{R})$. ■

Μπορούμε επίσης να αποδείξουμε ότι το αποτέλεσμα του *Θεωρήματος IV.2.8.1* είναι το καλύτερο δυνατόν, με την έννοια ότι δεν μπορεί να βελτιωθεί περαιτέρω:

Θεώρημα IV.2.8.2. ***Οι τύποι τετραγωνισμού Gauss δεν είναι ακριβείς επί του διανυσματικού χώρου*** $\mathcal{P}_{2n+2}(\mathbb{R})$.

Απόδειξη. Έστω ότι δίνεται προς υπολογισμό το ολοκλήρωμα:

$$I = \int_a^\beta V_n^2(x)\, w(x)\, dx .$$

Επειδή τα μεγέθη $V_n^2(x)$ και $w(x)$ είναι μη αρνητικά, για οποιαδήποτε τιμή $x \in [a, \beta]$, θα είναι υποχρεωτικά:

$$I > 0 .$$

Από την άλλη πλευρά όμως, σύμφωνα με τον ορισμό του πολυωνύμου $V_n(x)$, ισχύει:

$$\sum_{i=0}^{n} A_i^{(n)}\, V_n^2(x_i) = 0 .$$

Άρα, ένας τύπος τετραγωνισμού Gauss δεν μπορεί να είναι ακριβής επί του διανυσματικού χώρου $\mathcal{P}_{2n+2}(\mathbb{R})$, όλων των πραγματικών πολυωνύμων βαθμού το πολύ $(2n+2)$, και η *Απόδειξη* είναι πλήρης. ■

Μέχρι στιγμής, στο *Θεώρημα IV.2.8.1*, είδαμε ότι μία αναγκαία και ικανή συνθήκη για την ακρίβεια επί του $\mathcal{P}_{2n+1}(\mathcal{R})$ της μεθόδου Gauss είναι η επιλογή των σημείων παρεμβολής $x_0, x_1, \ldots x_n$ να είναι τέτοια ώστε να ισχύει:

$$\int_a^\beta x^p \, V_n(x) \, w(x) \, dx = 0 \, , \; \forall p = 0,1,\ldots n \, ,$$

όπου $V_n(x) = (x - x_0)\,(x - x_1)\ldots(x - x_n)$.

Ακόμη, στο *Θεώρημα IV.2.8.2*, είδαμε ότι για οποιαδήποτε επιλογή των σημείων $x_0, x_1, \ldots, x_n$, η αντίστοιχη μέθοδος Gauss αποκλείεται να είναι ακριβής επί του $\mathcal{P}_{2n+2}(\mathcal{R})$.

Το εύλογο ερώτημα που εγείρεται είναι εάν υπάρχει ένα πολυώνυμο $V_n(x)$ που να ικανοποιεί την συνθήκη του *Θεωρήματος IV.2.8.1*.

Επί πλέον, σε περίπτωση καταφατικής απάντησης, εάν ένα τέτοιο πολυώνυμο $V_n(x)$ υπάρχει είναι μοναδικό.

Με άλλα λόγια, *εάν υπάρχουν σημεία*:

$$x_0, x_1, x_2 \ldots, x_n \in [a, \beta],$$

ανά δύο διάφορα μεταξύ τους, τότε είναι με μοναδικό τρόπο ορισμένα και τέτοια ώστε:

$$\int_a^\beta x^p \; V_n(x) \; w(x) \; dx = 0 \, , \; \forall \; p = 0,1,\ldots, n$$

Συναφώς, έχουμε το εξής:

Θεώρημα IV.2.8.3. ***Στις μεθόδους Gauss, τα σημεία*** $x_0, x_1, \ldots, x_n$ ***υπάρχουν και είναι μοναδικά ορισμένοι, ανά δύο διάφοροι μεταξύ τους, πραγματικοί αριθμοί του διαστήματος*** $[a, \beta]$.

Απόδειξη. Η *Απόδειξη* στηρίζεται στο γεγονός ότι η συνθήκη του *Θεωρήματος IV.2.8.1* ισχύει εάν και μόνον εάν:

$$\int_a^\beta Q(x)\, V_n(x)\, w(x)\, dx = 0,$$

για κάθε πολυώνυμο $Q(x) \in \boldsymbol{P}_n(\boldsymbol{R})$.

Ας υποθέσουμε ότι το πολυώνυμο $V_n(x)$ δεν διαθέτει παρά μόνον k πραγματικές ρίζες, που βρίσκονται μέσα στο κλειστό διάστημα $[a, \beta]$.

Ας συμβολίσουμε με:

$$x_0, x_1, \ldots, x_{k-1}$$

αυτές τις ρίζες και ας υποθέσουμε ότι $k < n+1$.

Τότε, μπορούμε να γράψουμε:

$$V_n(x) = (x - x_0)\,(x - x_1)\ldots(x - x_k)\, u(x),$$

όπου $u(x)$ είναι ένα πολυώνυμο βαθμού $(n-k+1)$ που διατηρεί σταθερό το πρόσημό του μέσα στο διάστημα $[a, \beta]$.

Εάν λάβουμε:

$$Q(x) = (x - x_0)\,(x - x_1)\ldots(x - x_k),$$

τότε $Q(x) \in \boldsymbol{P}_n(\boldsymbol{R})$ και επί πλέον θα είναι:

$$\int_a^\beta Q(x)\, V_n(x)\, w(x)\, dx = 0 \text{ ή } < 0,$$

που είναι άτοπο. Άρα, πρέπει να ισχύει:

$$k = n + 1 .$$

Δηλαδή, δείξαμε ότι το πολυώνυμο $V_n(x)$ έχει $n+1$ πραγματικές ρίζες, όχι όμως υποχρεωτικά διάφορες μεταξύ τους και που βρίσκονται μέσα στο $[a, \beta]$.

Ας αιτιολογήσουμε τώρα γιατί αυτές οι ρίζες είναι, αναγκαστικά, ανά δύο διάφορες μεταξύ τους.

Εάν ήταν:

$$x_0 = x_1 = \ldots = x_{m-1} ,$$

για κάποιο $m > 1$, τότε θα μπορούσαμε να γράψουμε:

$$V_n(x) = (x - x_1)^m \; u(x) ,$$

όπου $u(x)$ είναι ένα πολυώνυμο βαθμού $n - m + 1$, οπότε παίρνοντας:

$$Q(x) = \begin{cases} u(x), & \text{εάν } m = \text{άρτιος}, \\ (x - x_0)\, u(x), & \text{εάν } m = \text{περιττός}, \end{cases}$$

θα καταλήγαμε στο ίδιο συμπέρασμα όπως και προηγουμένως, δηλαδή θα οδηγούμασταν σε άτοπο.

Άρα, πρέπει να ισχύει:

$$m = 1 .$$

Συνεπώς, μετά απ' όλα αυτά, συμπεραίνουμε ότι το πολυώνυμο $V_n(x)$ έχει $n+1$ πραγματικές ρίζες, που είναι υποχρεωτικά ανά δύο διάφορες μεταξύ τους και που βρίσκονται όλες στο διάστημα $[a, \beta]$.

Για να ολοκληρωθεί η *Απόδειξη*, μένει να αποδειχθεί ότι το πολυώνυμο $V_n(x)$, που ανταποκρίνεται στις προδιαγραφές της συνθήκης:

$$\int_a^\beta Q(x)\, V_n(x)\, w(x)\, dx = 0$$

για κάθε $Q(x) \in \mathcal{P}_{2n+1}(\mathbb{R})$, είναι μοναδικό.

Προς τούτο, ας υποθέσουμε ότι υπάρχει και δεύτερο πολυώνυμο $\widetilde{V}_n(x) \in \mathcal{P}_{n+1}(\mathbb{R})$ που ανταποκρίνεται στις ίδιες προδιαγραφές.

Τότε, θα συνέβαιναν ταυτόχρονα τα εξής δύο:

$$\int_a^\beta x^{n+1}\, V_n(x)\, w(x)\, dx = A \quad \text{και} \quad \int_a^\beta x^{n+1}\, \widetilde{V}_n(x)\, w(x)\, dx = B\,,$$

με τους αριθμούς A και B να είναι αμφότεροι διάφοροι του μηδενός, γιατί στην αντίθετη περίπτωση ο τύπος τετραγωνισμού Gauss θα ήταν ακριβής επί του $\mathcal{P}_{2n+2}(\mathbb{R})$, πράγμα που είναι αδύνατον, λόγω του *Θεωρήματος IV.2.8.2*.

Εάν τώρα θεωρούσαμε το πολυώνυμο:

$$[A\,\widetilde{V}_n(x) + B\,V_n(x)] \in \mathcal{P}_{n+1}(\mathbb{R})\,,$$

τότε θα ίσχυε:

$$\int_a^\beta x^{n+1}\,[A\widetilde{V}_n(x) + B\,V_n(x)]\, w(x)\, dx = 0\,,$$

πράγμα που δεν θα ήταν δυνατόν παρά μόνον όταν το πολυώνυμο $[A\,\widetilde{V}_n(x) + B\,V_n(x)]$ θα ήταν ταυτοτικά ίσο με μηδέν, κι αυτό για τον ίδιο όπως και προηγουμένως λόγο. Οι ρίζες $x_0, x_1, \ldots, x_n$ των δύο πολυωνύμων $V_n(x)$ και $\widetilde{V}_n(x)$ είναι κοινές και η *Απόδειξη* του *Θεωρήματος* έχει ολοκληρωθεί. ■

Επί πλέον, και σε αντίθεση με τις μεθόδους Newton-Côtes, το επόμενο αποτέλεσμα είναι ενθαρρυντικό:

Θεώρημα IV.2.8.4. ***Οι μέθοδοι Gauss είναι ευσταθείς και συγκλίνουσες επί του διανυσματικού χώρου*** $C[a,\beta]$ ***όλων των πραγματικών συναρτήσεων που είναι συνεχείς στο κλειστό διάστημα*** $[a,\beta]$.

Απόδειξη. Η *Απόδειξη* στηρίζεται στην επαλήθευση των κριτηρίων των *Θεωρημάτων IV.2.2.8* και *IV.2.2.10*.

Κατ' αρχήν, παρατηρούμε ότι οι μέθοδοι Gauss συγκλίνουν πάντοτε όταν η συνάρτηση $f(x)$ είναι πολυωνυμική, γιατί είναι μέθοδοι τύπου παρεμβολής. Απομένει λοιπόν να δειχθεί ότι υπάρχει μία σταθερά $M > 0$ τέτοια ώστε:

$$\sum_{i=0}^{n} \left| A_i^{(n)} \right| \le M \,, \quad \text{για κάθε } n = 0,1,2,\ldots$$

Προς τούτο, ας θεωρήσουμε τις πολυωνυμικές εκφράσεις:

$$L_i(x) = \prod_{\substack{j=0 \\ j\ne i}}^{n} \frac{x - x_j}{x_i - x_j} \,.$$

Είναι φανερό ότι:

$$L_i(x_k) = 0 \quad \forall k \ne i \quad \text{και } L_i(x_i) = 1$$

και ότι, επί πλέον:

$$L_i^2(x) \in \mathbb{P}_{2n}(\mathbb{R}) \,.$$

Έπεται αμέσως ότι:

$$\int_a^\beta L_i^2(x)\, w(x)\, dx = \sum_{j=0}^{n} A_j^{(n)}\, L_i^2(x_j) = A_i^{(n)} \,,$$

και επομένως:

$$A_i^{(n)} > 0 \qquad \forall i = 0,1,\ldots,n \,.$$

Ακόμη, καθώς ο τύπος τετραγωνισμού Gauss είναι ακριβής επί του $\mathbb{P}_0(\mathbb{R})$, ισχύει:

$$\int_a^\beta w(x)dx = \sum_{i=0}^n A_i^{(n)},$$

ισότητα που αποδεικνύει το *Θεώρημα*, επειδή έχουμε υποθέσει, από την αρχή του *Κεφαλαίου*, ότι η συνάρτηση $w(x)$ είναι ολοκληρώσιμη επί του $[a,\beta]$. ■

Θα κλείσουμε την παρούσα *Παράγραφο* με την μελέτη του σφάλματος που προκύπτει από την εφαρμογή της μεθόδου τετραγωνισμού Gauss.

Έστω $p_{2n+1}(x) \in \boldsymbol{P}_{2n+1}(\boldsymbol{R})$ το πολυώνυμο παρεμβολής Hermite της συνάρτησης $f(x)$ στα σημεία $x_0, x_1, \ldots, x_n \in [a,\beta]$.

Όπως είδαμε στην *Παράγραφο IV.1.9*, υπάρχει ένας αριθμός $\xi \in [a,\beta]$ τέτοιος ώστε:

$$f(x) = p_{2n+1}(x) + V_n^2(x)\frac{1}{(2n+2)!}\frac{d^{(2n+2)}f}{dx^{(2n+2)}}(\xi),$$

κάτω, φυσικά, από την προϋπόθεση ότι η $f(x)$ είναι $(2n+2)$ φορές συνεχώς διαφορίσιμη μέσα σε μία ανοικτή περιοχή του κλειστού διαστήματος $[a,\beta]$.

Έπεται ότι, κάτω από την ίδια προϋπόθεση, ισχύει:

$$\int_a^\beta f(x)w(x)\,dx = \int_a^\beta p_{2n+1}(x)\,w(x)dx + \int_a^\beta V_n^2(x)\frac{w(x)}{(2n+2)!}\frac{d^{(2n+2)}f}{dx^{(2n+2)}}(\xi)\,dx,$$

ή ακόμη, επειδή ο τύπος τετραγωνισμού Gauss είναι ακριβής επί του $\boldsymbol{P}_{2n+1}(\boldsymbol{R})$, καθώς κι επειδή $p_{2n+1}(x_i) = f(x_i)\ \ \forall i = 0,1,2,\ldots,n$:

$$\int_a^\beta f(x)w(x)dx - \sum_{i=0}^n A_i^{(n)} f(x_i) = R_n(f) = \frac{1}{(2n+2)!}\int_a^\beta V_n^2(x)\frac{d^{(2n+2)}f}{dx^{(2n+2)}}(\xi)\,w(x)\,dx.$$ Αφού όμως η συνάρτηση $V_n^2(x)\,w(x)$ διατηρεί σταθερό το θετικό πρόσημό της σε όλο το μήκος του διαστήματος $[a,\beta]$ μπορούμε να εφαρμόσουμε το *Θεώρημα Μέσης Τιμής* στο τελευταίο

ολοκλήρωμα της παραπάνω σχέσης και να επιτύχουμε την συναγωγή του ακόλουθου αποτελέσματος:

Θεώρημα IV.2.8.5. ***Εάν η συνάρτηση*** $f(x)$ ***είναι*** $(2n+2)$ ***φορές συνεχώς διαφορίσιμη μέσα σε μία ανοικτή περιοχή του*** $[a,\beta]$ ***τότε το σφάλμα*** $R_n(f)$ ***τετραγωνισμού της μεθόδου Gauss ισούται με:***

$$R_n(f)=\frac{1}{(2n+2)!}\frac{d^{(2n+2)}f}{dx^{(2n+2)}}(\xi)\int_a^{\beta}V_n^2(x)\,w(x)\,dx,$$

για κάποιο σημείο $\xi\in[a,\beta]$. ■

IV.2.9. ΣΥΣΤΗΜΑΤΑ ΟΡΘΟΓΩΝΙΩΝ ΠΟΛΥΩΝΥΜΩΝ

Μετά απ' όλα όσα προηγήθηκαν, πρέπει τώρα να εντοπίσουμε με ακρίβεια τις τιμές των μοναδικών και διακριτών σημείων $x_0,x_1,\ldots,x_n$ του $[a,\beta]$, ή ισοδύναμα την σαφή έκφραση του μοναδικού πολυωνύμου $V_n(x)$ που επαληθεύει την θεμελιώδη συνθήκη του *Θεωρήματος IV.2.8.1*.

Η επιτυχία της προσπάθειάς μας προέρχεται και διέπεται από την Θεωρία των Συστημάτων Ορθογωνίων Πολυωνύμων.

Ορισμός IV.2.9.1. ***Έστω η ακολουθία των πραγματικών πολυωνύμων:***

$$(P_0, P_1, P_2, \ldots, P_i, \ldots),$$

όπου ο δείκτης i ***συμβολίζει τον βαθμό του πολυωνύμου*** P_i***. Λέμε ότι τα πολυώνυμα αυτά σχηματίζουν ένα σύστημα ορθογωνίων πολυωνύμων επί του διαστήματος*** $[a, \beta]$***, ως προς την συνάρτηση βάρους*** $w(x)$***, εάν:***

$$\int_a^\beta P_k(x)\, P_m(x)\, w(x)\, dx = 0, \ \forall k \neq m. \blacksquare$$

Έστω τώρα $n+1 = m > k$. Μπορούμε να πραγματοποιήσουμε έναν γραμμικό συνδυασμό τέτοιων ολοκληρωμάτων, για $n =$ σταθερό και για $k = 0,1,2,\ldots,n$, και να διαπιστώσουμε πως:

$$\int_a^\beta P(x)\, P_{n+1}(x)\, w(x)\, dx = 0, \qquad \forall P(x) \in \mathbb{P}_n(\mathbb{R}).$$

Αυτό σημαίνει πως το πολυώνυμο $P_{n+1}(x)$ επαληθεύει την θεμελιώδη συνθήκη του *Θεωρήματος IV.2.8.1*, και συνεπώς μπορούμε να λάβουμε:

$$V_n(x) = P_{n+1}(x).$$

Δηλαδή, έχουμε αμέσως το εξής αποτέλεσμα:

Θεώρημα IV.2.9.2. ***Τα σημεία*** $x_0, x_1, \ldots, x_n$ ***είναι οι ρίζες του πολυωνύμου, βαθμού*** $(n+1)$***, που ανήκει στο σύστημα ορθογωνίων πολυωνύμων επί του διαστήματος*** $[a, \beta]$ ***ως προς την συνάρτηση βάρους*** $w(x)$***.*** $\blacksquare$

Παρατήρηση IV.2.9.3. Το *Θεώρημα IV.2.8.3* βεβαιώνει την ύπαρξη και την μοναδικότητα ενός τέτοιου συστήματος πολυωνύμων επί του διαστήματος $[a,\beta]$ και ως προς μία δεδομένη συνάρτηση βάρους $w(x)$. Ακόμα, σύμφωνα με το ίδιο *Θεώρημα*, όλες οι ρίζες ενός πολυωνύμου, που ανήκει σ' αυτό το σύστημα, είναι απλές και ανήκουν στο $[a,\beta]$. ■

Στην *Απόδειξη* του *Θεωρήματος IV.2.8.4* συναντήσαμε μία έκφραση των συντελεστών $A_i^{(n)}$, μέσα στην οποία υπεισέρχονται τα πολυώνυμα $L_i^2(x)$:

$$A_i^{(n)} = \int_a^\beta L_i(x)\, w(x)\, dx \qquad (L_i(x) = \prod_{\substack{j=0 \\ j\neq i}}^{n} \frac{x - x_j}{x_i - x_j}).$$

Επειδή αυτή η έκφραση δεν χρησιμοποιείται στα πρακτικά προβλήματα, θα παρουσιάσουμε μία διαφορετική και πιο δημοφιλή στις εφαρμογές:

Θεώρημα IV.2.9.4. ***Εάν η ακολουθία:***

$$(P_0(x), P_1(x), P_2(x), \ldots, P_i(x), \ldots),$$

όπου ο δείκτης i συμβολίζει τον βαθμό του πολυωνύμου $P_i(x)$, συνιστά το σύστημα ορθογωνίων πολυωνύμων επί του $[a,\beta]$, ως προς την συνάρτηση βάρους $w(x)$, τότε οι συντελεστές $A_i^{(n)}$ δίνονται από την σχέση:

$$A_i^{(n)} = \frac{1}{P'_{n+1}(x_i)} \int_a^\beta \frac{P_{n+1}(x)}{x - x_i}\, w(x)\, dx \qquad \forall i = 0,1,\ldots,n,$$

όπου τα σημεία x_i είναι οι ρίζες του $P_{n+1}(x)$.

Απόδειξη. Μέσα στο ολοκλήρωμα:

$$\int_a^\beta p(x)\, w(x)\, dx\,,$$

παίρνουμε:

$$p(x) = \frac{P_{n+1}(x)}{x - x_i}\,,$$

όπου x_i είναι μία από τις ρίζες του $P_{n+1}(x)$. Τότε, αφ' ενός:

$$p(x_i) = \lim_{x \to x_i} \frac{P_{n+1}(x)}{x - x_i} = \lim_{x \to x_i} \frac{P_{n+1}(x) - P_{n+1}(x_i)}{x - x_i} = P'_{n+1}(x_i)\,,$$

(γιατί $P_{n+1}(x_i) = 0$) και αφ' ετέρου:

$$p(x) \in \boldsymbol{P}_n(\boldsymbol{R})$$

(γιατί το σημείο x_i είναι ρίζα του $P_{n+1}(x)$).

Επειδή, τώρα, ο *Τύπος* τετραγωνισμού Gauss είναι ακριβής επί του $\boldsymbol{P}_{2n+1}(\boldsymbol{R})$, θα είναι:

$$\int_a^\beta p(x)\, w(x)\, dx = \sum_{j=0}^{n} A_j^{(n)}\, p(x_j) \Rightarrow$$

$$\int_a^\beta \frac{P_{n+1}(x)}{x - x_i}\, w(x)\, dx = \sum_{j=0}^{n} A_j^{(n)}\, \frac{P_{n+1}(x_j)}{x_i - x_j} = A_i^{(n)}\, P'_{n+1}(x_i)\,,$$

που αποδεικνύει το *Θεώρημα*, γιατί $P'_{n+1}(x_i) \neq 0$, αφού το σημείο x_i είναι απλή ρίζα του $P_{n+1}(x)$. ■

Θα δώσουμε τώρα ένα θεωρητικό αποτέλεσμα που περιγράφει την επαγωγική κατασκευή των συστημάτων ορθογωνίων πολυωνύμων:

Θεώρημα IV.2.9.5 ***Κάθε σύστημα ορθογωνίων πολυωνύμων ικανοποιεί την επόμενη σχέση μαθηματικής επαγωγής:***

$$P_{n+2}(x) = (A_{n+1}\ x + B_{n+1})\ P_{n+1}(x) - C_{n+1}\ P_n(x)\ \ \forall n = 0,1,...$$

για δεδομένα $P_0(x)$ ***και*** $P_1(x)$***.***

Απόδειξη. Συμβολίζουμε με a_n συντελεστή του x^n στην έκφραση του πολυωνύμου $P_n(x)$ και με a'_n τον συντελεστή του x^{n-1} στην ίδια έκφραση.

Θέτουμε:

$$h_n := \int_a^\beta P_n^2(x)\, w(x) dx\,.$$

Λόγω των σχέσεων ορθογωνιότητας που συνδέουν το πολυώνυμο $P_n(x)$ με κάθε άλλο πολυώνυμο, βαθμού μικρότερου του n, βλέπουμε ότι:

$$h_n = a_n \int_a^\beta x^n\ P_n(x)\ w(x)\ dx\,.$$

Ας θεωρήσουμε τώρα το πολυώνυμο:

$$P_{n+2}(x) - \frac{a_{n+2}}{a_{n+1}}\ x\ P_{n+1}(x)\,.$$

Πρόκειται για ένα πολυώνυμο βαθμού το πολύ $n+1$, που μπορούμε επίσης να γράψουμε υπό την μορφή:

$$P_{n+2}(x) - \frac{a_{n+2}}{a_{n+1}}\ x\ P_{n+1}(x) = \sum_{i=0}^{n+1} \beta_i\ P_i(x_i)\,.$$

Πολλαπλασιάζοντας και τα δύο μέλη αυτής της ισότητας με:

$$P_k(x)\ w(x)\,,\quad \text{για } k=0,1,\ldots,n-1\,,$$

και ακολούθως ολοκληρώνοντας, μεταξύ των σημείων a και β, βρίσκουμε:

$$\beta_i=0\,,\quad \text{για } i=0,1,\ldots,n-1$$

και συνεπώς:

$$P_{n+2}(x)-\frac{a_{n+2}}{a_{n+1}}\,x\,P_{n+1}(x)=\beta_{n+1}\,P_{n+1}(x)+\beta_n\,P_n(x)\,.$$

Η τελευταία αυτή σχέση είναι η επιδιωκόμενη. Απομένει η εύρεση της μορφής των σταθερών β_{n+1} και β_n.

Προς τούτο, πολλαπλασιάζουμε την ισότητα:

$$P_{n+2}(x)-\frac{a_{n+2}}{a_{n+1}}\,x\,P_{n+1}(x)=\sum_{i=0}^{n+1}\beta_i\,P_i(x_i)$$

με την συνάρτηση $(P_n(x)\ w(x))$ και ολοκληρώνουμε μεταξύ των a και β.

Προκύπτει ότι:

$$h_n-\beta_n=-\frac{a_{n+2}}{a_{n+1}}\int_a^\beta P_{n+1}(x)\,P_n(x)\,w(x)\,dx$$

$$=-\frac{a_{n+2}}{a_{n+1}}\,\frac{a_n}{a_{n+1}}\int_a^\beta a_{n+1}\,P_{n+1}(x)\,x^n\,w(x)\,dx\,,$$

απ' όπου έπεται ότι:

$$\beta_n=-\frac{h_{n+1}}{h_n}\,\frac{a_{n+2}\,a_n}{a_{n+1}^2}\,.$$

Από την άλλη πλευρά, ταυτίζοντας τους συντελεστές των όρων που περιέχουν το σύμβολο x^{n+1} μέσα στην ισότητα:

$$P_{n+2}(x) - \frac{a_{n+2}}{a_{n+1}}\, x\, P_{n+1}(x) = \sum_{i=0}^{n+1} \beta_i\; P_i(x_i)\,,$$

συμπεραίνουμε ότι:

$$\beta_{n+1} = \frac{a_{n+2}}{a_{n+1}} - \frac{a_{n+2}\; a_{n+1}}{a_{n+1}^2}\,.$$

Άρα, τελικά, εάν:

$$A_{n+1} = \frac{a_{n+2}}{a_{n+1}}\,,$$

$$B_{n+1} = \frac{a_{n+2}}{a_{n+1}} \left[\frac{a_{n+2}}{a_{n+2}} - \frac{a_{n+1}}{a_{n+1}} \right],$$

και:

$$C_{n+1} = \frac{h_{n+1}}{h_n}\, \frac{a_{n+2}\; a_n}{a_n^2}\,,$$

τότε:

$$P_{n+2}(x) = \left(A_{n+1}\, x + B_{n+1}\right) P_{n+1}(x) + C_{n+1}\; P_n(x), \quad \forall\, n\,,$$

και η *Απόδειξη* του *Θεωρήματος* είναι πλήρης. ■

Υπάρχουν διάφορα συστήματα ορθογωνίων πολυωνύμων που είναι γνωστά. Προς το παρόν, στην *Παράγραφο IV.1.7* συναντήσαμε μόνον τα πολυώνυμα Chebyshev πρώτου είδους, που αποτελούν ένα χαρακτηριστικό παράδειγμα.

Στην συνέχεια, και μέχρι το τέλος της *Παραγράφου*, θα παρουσιάσουμε επιγραμματικά τα πιο χρήσιμα (και μέχρι τώρα γνωστά) συστήματα ορθογωνίων πολυωνύμων. Για περισσότερες πληροφορίες, παραπέμπουμε στην βιβλιογραφική αναφορά [14].

Παραδείγματα IV.2.9.6.:Τα πολυώνυμα Legendre $P_n(x)$. ***Το σύστημα ορθογωνίων πολυωνύμων Legendre ορίζεται επί του διαστήματος:***

$$[-1,1]$$

και ως προς την συνάρτηση βάρους:

$$w(x)=1 \ \left(-1\le x\le 1\right).$$

Τα πολυώνυμα Legendre $P_n(x)$ ***ικανοποιούν την επαγωγική σχέση:***

$$(n+2)\,P_{n+2}(x)=(2n+2)\,x\,P_{n+1}(x)-(n+1)\,P_n(x) \qquad \left(n=0,1,2,\ldots\right),$$

με:

$$P_0(x)=1 \text{ και } P_1(x)=x\,.$$

Χρησιμοποιώντας τις ιδιότητες αυτών των πολυωνύμων, *βρίσκουμε ότι:*

$$A_i^{(n)}=\frac{2\,(1-x_i^2)}{(n+1)^2\;P_n^2(x_i)}\text{, για } i=0,1,2,\ldots,n$$

και:

$$R(f)=\frac{2^{2n+3}\,[(n+1)!]^4}{(2n+3)\,[(2n+2)!]^3}\,\frac{d^{(2n+2)}f}{dx^{(2n+2)}}(\xi),$$

για κάποιο $\xi\in[-1,1]$ ***και κάτω από την προϋπόθεση ότι η συνάρτηση*** $f(x)$ ***είναι*** $\left(2n+2\right)$ ***φορές συνεχώς διαφορίσιμη μέσα σε μία ανοικτή περιοχή του*** $[-1,1]$.

Υπάρχουν πίνακες που δίνουν τις ρίζες x_i **κάθε πολυωνύμου Legendre** $P_n(x)$. **Μπορούμε επίσης να υπολογίσουμε αυτές τις ρίζες χρησιμοποιώντας, για παράδειγμα, την μέθοδο Newton, την οποία γνωρίσαμε στην *Παράγραφο ΙΙΙ.1.2.(i)*.**■

Παραδείγματα IV.2.9.7.:Τα πολυώνυμα Laguerre $L_n(x)$. ***Το σύστημα των ορθογωνίων πολυωνύμων Laguerre*** **ορίζεται επί του διαστήματος:**

$$[0,+\infty[$$

και ως προς την συνάρτηση βάρους:

$$w(x) = e^{-x} \qquad (x \geq 0).$$

Τα πολυώνυμα Laguerre $L_n(x)$ ***επαληθεύουν την επαγωγική σχέση:***

$$(n+2)\, L_{n+2}(x) = (2n+2-x)\, L_{n+1}(x) - (n+1)\, L_n(x) \quad (n = 0,1,2,\ldots),$$

$$L_0(x) = 1 \text{ και } L_1(x) = 1 - x.$$

Αποδεικνύεται ότι:

$$A_i^{(n)} = \frac{[(n+1)!]^2}{x_i\, [L'_{n+1}(x_i)]^2}, \text{ για } i = 0,1,\ldots,n,$$

και

$$R(f) = \frac{[(n+1)!]^2}{(2n+2)!} \frac{d^{(2n+2)} f}{dx^{(2n+2)}}(\xi),$$

για κάποιο $\xi \in]0,+\infty[$ ***και κάτω από την προϋπόθεση ότι η συνάρτηση*** $f(x)$ ***είναι*** $(2n+2)$ ***φορές συνεχώς διαφορίσιμη μέσα σε μία ανοικτή περιοχή του*** $[0,+\infty[$ **.■**

Παραδείγματα IV.2.9.8.:Τα πολυώνυμα Hermite $H_n(x)$. ***Το σύστημα ορθογωνίων πολυωνύμων Hermite ορίζεται επί του διαστήματος:***

$$]-\infty,+\infty[$$

και ως προς την συνάρτηση βάρους:

$$w(x)=e^{-x^2} \qquad (x\in\mathbb{R}).$$

Τα πολυώνυμα Hermite $H_n(x)$ ***ικανοποιούν τη επαγωγική σχέση:***

$$H_{n+2}(x)=2\,x\,H_{n+1}(x)-(2\,n+2)\,H_n(x) \qquad (n=0,1,2,\ldots),$$

με:

$$H_0(x)=1 \text{ και } H_1(x)=2\,x\,.$$

Μπορούμε ν' αποδείξουμε, χωρίς σοβαρές δυσκολίες, ότι:

$$A_i^{(n)}=\frac{2^{n+2}\,(n+1)!\,\sqrt{\pi}}{[H'_{n+1}(x_i)]^2}, \qquad \text{για } i=0,1,2,\ldots,n,$$

και

$$R(f)=\frac{(n+1)!\,\sqrt{\pi}}{2^{n+1}\,(2\,n+2)!}\,\frac{d^{(2n+2)}f}{d\,x^{(2n+2)}}(\xi),$$

για κάποιο $\xi\in]-\infty,+\infty[$ ***και κάτω από την προϋπόθεση ότι η συνάρτηση*** $f(x)$ ***είναι*** $(2n+2)$ ***φορές συνεχώς διαφορίσιμη μέσα στο*** $\mathbb{R}=]-\infty,+\infty[$. ■

Παραδείγματα IV.2.9.9.: Τα πολυώνυμα Chebyshev πρώτου είδους $T_n(x)$.

Το σύστημα ορθογωνίων πολυωνύμων Chebyshev 1ου είδους ορίζεται επί του διαστήματος:

$$]-1,1[$$

και ως προς την συνάρτηση βάρους:

$$w(x) = (1-x^2)^{-\frac{1}{2}} \qquad (-1 \le x \le 1).$$

Τα πολυώνυμα Chebyshev 1ου είδους $T_n(x)$ ***επαληθεύουν την επαγωγική σχέση:***

$$T_{n+2}(x) = 2\,x\,T_{n+1}(x) - T_n(x) \quad (n = 0,1,2,\ldots),$$

με:

$$T_0(x) = 1 \text{ και } T_1(x) = x\,.$$

Ακόμη, *αποδεικνύεται ότι:*

$$x_i = \cos\left(\frac{2i+1}{2n+2}\pi\right), \text{ για } i = 0,1,\ldots,n\,,$$

$$A_i^{(n)} = \frac{\pi}{n+1}, \text{ για } i = 0,1,\ldots,n\,,$$

και

$$R(f) = \frac{2\,\pi}{2^{2n+2}\,(2\,n+2)!}\,\frac{d^{(2n+2)}f}{dx^{(2n+2)}}(\xi),$$

για κάποιο $\xi \in [-1,1]$ ***και κάτω, φυσικά, από την προϋπόθεση ότι η συνάρτηση*** $f(x)$ ***είναι*** $(2n+2)$ ***φορές συνεχώς διαφορίσιμη μέσα σε μία ανοικτή περιοχή του*** $[-1,1]$. ■

IV.2.10. ΚΑΤΑΣΚΕΥΗ ΟΡΘΟΓΩΝΙΩΝ ΠΟΛΥΩΝΥΜΩΝ.

Στην πράξη παρουσιάζεται συχνά το ενδεχόμενο να πρέπει να χρησιμοποιηθεί ένας τύπος τετραγωνισμού Gauss, για τον οποίο, όμως, δεν υπάρχει κάποιο γνωστό σύστημα ορθογωνίων πολυωνύμων επί του διαστήματος $[a,\beta]$ και ως προς κάποια δεδομένη συνάρτηση βάρους $w(x)$.

Τότε, δημιουργείται η ανάγκη κατασκευής του συστήματος ορθογωνίων πολυωνύμων που χρειαζόμαστε.

Προφανώς, για να επιτύχουμε αυτήν την κατασκευαστική επιταγή, αρκεί να υπολογίσουμε τους συντελεστές $a_i^{(n+1)}$ των πολυωνύμων $P_{n+1}(x)$ που απαρτίζουν το προς αναζήτηση ορθογώνιο σύστημα.

Έτσι, αρχίζουμε θέτοντας ένα τέτοιο πολυώνυμο υπό την μορφή:

$$P_{n+1}(x) = a_{n+1}^{(n+1)}\, x^{n+1} + a_n^{(n+1)}\, x^n + \ldots + a_1^{(n+1)}\, x + a_0^{(n+1)}.$$

Επειδή το, υπό θεώρηση, πολυώνυμο πρέπει να έχει την ιδιότητα:

$$\int_a^\beta x^k\; P_{n+1}(x)\; w(x)\; dx = 0\,,\; \forall k = 0,1,\ldots,n\,,$$

βλέπουμε ότι πρέπει να επαληθεύονται οι ακόλουθες ισότητες:

$$\int_a^\beta (a_0^{(n+1)} + \ldots + a_{n+1}^{(n+1)}\; x^{n+1})\; w(x)\; dx = 0\,,$$

$$\int_a^\beta (a_0^{(n+1)} + \ldots + a_{n+1}^{(n+1)}\; x^{n+1})\; x\; w(x)\; dx = 0\,,$$

.....

$$\int_a^\beta (a_0^{(n+1)} + \ldots + a_{n+1}^{(n+1)}\; x^{n+1})\; x^n\; w(x)\; dx = 0\,.$$

Συμβολίζοντας με:

$$c_k := \int_a^\beta x^k \; w(x)\, dx \quad (k = 0,1,\ldots,2.n+1)$$

τις *στιγμές* $k-$*τάξης της* $w(x)$, το προηγούμενο σύστημα εξισώσεων γράφεται ισοδύναμα υπό την μορφή:

$$a_0^{(n+1)}\, c_0 + a_1^{(n+1)}\, c_1 + \ldots + a_{n+1}^{(n+1)}\, c_{n+1} = 0$$

$$a_0^{(n+1)}\, c_1 + a_1^{(n+1)}\, c_2 + \ldots + a_{n+1}^{(n+1)}\, c_{n+2} = 0$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$a_0^{(n+1)}\, c_n + a_1^{(n+1)}\, c_{n+1} + \ldots + a_{n+1}^{(n+1)}\, c_{2n+1} = 0\,.$$

Εάν οι στιγμές:

$$c_0, c_1, c_2, \ldots, c_n, \ldots, c_{2n+1}$$

είναι γνωστές, τότε έχουμε προς επίλυση ένα γραμμικό σύστημα $(n+1)$ εξισώσεων με $(n+2)$ αγνώστους.

Μπορεί να αποδειχθεί πως, όταν:

$$w(x) \geq 0 \;\; \forall x \in [a, \beta],$$

τότε κάθε ορίζουσα πίνακα, τάξης $(n+1)$, που εξάγεται από το παραπάνω σύστημα είναι μη μηδενικός αριθμός.

Αυτή η απροσδιοριστία πρώτης τάξης συνεπάγεται ότι, για $a_{n+1}^{(n+1)} =$ σταθερό, το παραπάνω σύστημα έχει μία και μοναδική (μη τετριμμένη) λύση.

Επομένως, οι συνθήκες ορθογωνιότητας προσδιορίζουν το πολυώνυμο $P_{n+1}(x)$, κατά παρέκλιση ενός αυθαίρετου πολλαπλασιαστικού παράγοντα, εφόσον είναι γνωστές οι στιγμές:

$$c_0, c_1, c_2, \ldots, c_n, \ldots, c_{2n+1}$$

της μη αρνητικής (και ολοκληρώσιμης) συνάρτησης $w(x)$ $(x \in [a, \beta])$.

Στην γενική περίπτωση, δεν είναι εύκολη η χρήση της σχέσης του *Θεωρήματος IV.2.9.4* που δίνει τους συντελεστές $A_i^{(n)}$. Γι'αυτό, θα επιχειρήσουμε την ανακάλυψη μίας διαφορετικής και εύχρηστης σχέσης που θα αποδίδει αμέσως τις ποσότητες $A_i^{(n)}$.

Στο εγχείρημά μας αυτό, θα χρειαστούμε την βοήθεια δύο ανεξάρτητων Προτάσεων που παραθέτουμε αμέσως τώρα:

Θεώρημα IV.2.10.1. (Ταυτότητα Cristoffel-Darboux) Εάν:

$$h_n = \int_a^\beta P_n^2(x)\, w(x)\, dx = 1, \qquad \forall n = 0,1,2,\ldots,$$

τότε τα ορθογώνια πολυώνυμα $P_n(x)$ ικανοποιούν την ταυτότητα Cristoffel-Darboux:

$$\frac{a_n^{(n)}}{a_{n+1}^{(n+1)}}\left[P_{n+1}(x)\, P_n(t) - P_{n+1}(t)\, P_n(x)\right] = (x-t)\sum_{i=0}^{n} P_i(x)\, P_i(t),$$

για κάθε $x \in [a,\beta]$ και για κάθε $t \in [a,\beta]$.

Απόδειξη. Πολλαπλασιάζοντας την επαγωγική σχέση του *Θεωρήματος IV.2.9.5* με $P_{n+1}(t)$, βρίσκουμε ότι:

$$P_{n+1}(t)\, P_{n+2}(x) = (A_{n+1}\, x + B_{n+1})\, P_{n+1}(t) P_{n+1}(x) - C_{n+1}\, P_{n+1}(t)\, P_n(x).$$

Εναλλάσσοντας, στη συνέχεια, τους ρόλους των x και t, και αφαιρώντας κατά μέλη τις δύο ισότητες, οδηγούμαστε στην σχέση:

$$P_{n+1}(t) P_{n+2}(x) - P_{n+1}(x) P_{n+2}(t) = A_{n+1}\, (x-t)\, P_{n+1}(t)\, P_{n+1}(x)$$
$$- C_{n+1}.\left[P_{n+1}(t)\, P_n(x) - P_{n+1}(x)\, P_n(t)\right].$$

Επειδή $C_{n+1}\ A_{n+1}^{-1} = A_n^{-1}$, εφόσον $h_n = 1\ \ \forall n$, βλέπουμε ότι πολλαπλασιάζοντας την τελευταία με A_{n+1}^{-1} επιτυγχάνεται η σχέση:

$$(x-t)\,P_{n+1}(x)\,P_{n+1}(t) = A_{n+1}^{-1}\,\left(P_{n+2}(x)\,P_{n+1}(t) - P_{n+2}(t)\,P_{n+1}(x)\right)$$
$$- A_n^{-1}\,\left(P_{n+1}(x)\,P_n(t) - P_{n+1}(t)\,P_n(x)\right).$$

Γράφοντας αυτήν την ισότητα διαδοχικά για:

$$n = -1,0,1,2,\ldots,n-1,$$

και αθροίζοντας κατά μέλη όλες τις προκύπτουσες, καταλήγουμε στην απόδειξη της *ταυτότητας Cristoffel-Darboux*, κάτω από την συμβατική παραδοχή ότι $a_{-1}^{(-1)} = 0$. ■

Θεώρημα IV.2.10.2. ***Τα ορθογώνια πολυώνυμα*** **(ως προς** w**):**

$$P_{n+1}(x) \ \textbf{\textit{και}}\ P_n(x) \qquad (n = 0,1,2,\ldots)$$

δεν έχουν κοινές ρίζες.

Απόδειξη. Κατ' αρχάς παρατηρούμε, ευθύς αμέσως, ότι τα πολυώνυμα

$$P_1(x) \text{ και } P_0(x)$$

δεν είναι δυνατόν να έχουν κοινή ρίζα.

Θα αποδείξουμε το *Θεώρημα*, χρησιμοποιώντας μαθηματική επαγωγή.

Προς τούτο, ας δεχθούμε ότι τα πολυώνυμα $P_{n+1}(x)$ και $P_n(x)$ δεν έχουν κοινές ρίζες, και ας υποθέσουμε ότι τα $P_{n+2}(x)$ και $P_{n+1}(x)$ έχουν μία κοινή ρίζα. Έστω ότι αυτή η ρίζα είναι ο αριθμός z.

Χρησιμοποιώντας την επαγωγική σχέση του *Θεωρήματος IV.2.9.5*, βλέπουμε ότι:

$$P_{n+2}(z) = (A_{n+1}\ z + B_{n+1})P_{n+1}(z) - C_{n+1}\ P_n(z) = 0$$

$$\Rightarrow$$

$$C_{n+1}\ P_n(z) = 0\,,$$

και επειδή $C_{n+1} \neq 0$, θα είναι:

$$P_n(z) = 0\,.$$

Αυτό σημαίνει ότι ο αριθμός z είναι κοινή ρίζα των πολυωνύμων:

$$P_{n+1}(x) \text{ και } P_n(x)\,,$$

ισχυρισμός που είναι άτοπος, γιατί έρχεται σε αντίθεση με την υπόθεση της μαθηματικής επαγωγής. Άρα, τα ορθογώνια πολυώνυμα (ως προς $w(x)$):

$$P_{n+2}(x) \text{ και } P_{n+1}(x)$$

δεν έχουν κοινή ρίζα και η *Απόδειξη* του *Θεωρήματος* ολοκληρώθηκε. ■

Μετά από αυτήν την προπαρασκευαστική παράθεση των δύο αποτελεσμάτων που προηγήθηκαν, είμαστε έτοιμοι να παρουσιάσουμε μία πιο εύχρηστη σχέση για τους συντελεστές $A_i^{(n)}$:

Θεώρημα IV.2.10.3. ***Εάν τα πολυώνυμα $P_n(x)$ σχηματίζουν ένα ορθοκανονικό σύστημα, δηλαδή εάν:***

$$\int_a^\beta P_n(x)\ P_m(x)\ w(x)\ dx = \begin{cases} \int_a^\beta P_n^2(x)\, w(x)\, dx = h_n = 1 \\ \int_a^\beta P_n(x)\ P_m(x)\ w(x)\ dx = 0 \qquad \forall n \neq m, \end{cases}$$

τότε οι συντελεστές του τετραγωνισμού Gauss δίνονται από την σχέση:

$$A_i^{(n)} = \frac{a_{n+1}^{(n+1)}}{a_n^{(n)}} \frac{1}{P'_{n+1}(x_i)\ P_n(x_i)} \qquad (i = 0,1,2,\ldots,n),$$

όπου x_i είναι η ρίζα του $P_{n+1}(x)$.

Απόδειξη. Ας συμβολίσουμε με:

$$x_0, x_1, x_2, \ldots, x_n$$

τις ρίζες του πολυωνύμου $P_{n+1}(x)$ και ας πάρουμε, μέσα στην *Ταυτότητα Cristoffel-Darboux*, $t = x_i$.

Τότε, είναι:

$$\frac{a_n^{(n)}}{a_{n+1}^{(n+1)}} P_n(x_i) \frac{P_{n+1}(x)}{x - x_i} = \sum_{j=0}^{n} P_j(x_i) P_j(x_i),$$

και, πολλαπλασιάζοντας με $w(x)$ και στην συνέχεια ολοκληρώνοντας ως προς x στο $[a, \beta]$, μπορούμε να γράψουμε την ακόλουθη μακροσκελή σχέση:

$$\frac{a_n^{(n)}}{a_{n+1}^{(n+1)}} P_n(x_i) \int_a^\beta \frac{P_{n+1}(x)}{x - x_i} w(x)\, dx = \sum_{j=0}^{n} P_j(x_i) \int_a^\beta P_j(x)\, w(x)\, dx.$$

Επειδή ο σταθερός αριθμός 1 είναι, τετριμμένα, ένα πολυώνυμο μηδενικού βαθμού, έπεται, σύμφωνα με τις σχέσεις ορθογωνιότητας, ότι:

$$\int_a^\beta P_j(x)\, w(x)\, dx = 0 \qquad \forall \mathrm{j} = 1,2,\ldots$$

Επειδή, ακόμη, έχουμε υποθέσει πως $h_n = 1$ για κάθε n, ισχύει ότι:

$$P_0(x_i) \int_a^\beta P_0(x)\, w(x)\, dx = 1.$$

Άρα, η πιο πάνω μακροσκελής σχέση μετατρέπεται στην ισότητα:

$$\frac{a_n^{(n)}}{a_{n+1}^{(n+1)}} P_n(x_i) \int_a^\beta \frac{P_{n+1}(x)}{x - x_i} w(x)\, dx = 1.$$

Έχοντας τώρα υπ' όψη τη σχέση:

$$A_i^{(n)} = \frac{1}{P'_{n+1}(x_i)} \int_a^\beta \frac{P_{n+1}(x)}{x - x_i} w(x)\, dx$$

του *Θεωρήματος IV.2.9.4*, και χρησιμοποιώντας το γεγονός ότι σύμφωνα με το *Θεώρημα IV.2.10.2*:

$$P_n(x_i) \neq 0,$$

η *Απόδειξη* ολοκληρώνεται εύκολα. ■

IV.2.11. ΜΕΘΟΔΟΣ ΠΡΟΣΔΙΟΡΙΣΤΕΩΝ ΣΥΝΤΕΛΕΣΤΩΝ

Στην μέθοδο Newton-Côtes, η επιλογή των σημείων παρεμβολής είναι επιβεβλημένη και προκαθορισμένη, ενώ ο υπολογισμός των συντελεστών $A_i^{(n)}$ πραγματοποιείται αξιώνοντας ο αντίστοιχος μαθηματικός τύπος τετραγωνισμού να είναι ακριβής για τα πολυώνυμα που έχουν βαθμό όσο το δυνατόν μεγαλύτερο.

Στην μέθοδο Gauss, η επιλογή των σημείων παρεμβολής και των συντελεστών $A_i^{(n)}$ γίνεται με αποκλειστικό κριτήριο την ακρίβεια των αντίστοιχων τύπων τετραγωνισμού για τα πολυώνυμα που έχουν τον μεγαλύτερο δυνατόν βαθμό.

Μπορούμε να θεωρήσουμε και ένα τρίτο είδος μεθόδων τετραγωνισμού, που αποτελεί ένα ενδιάμεσο στάδιο ανάμεσα στα δύο προηγούμενα είδη, δηλαδή το είδος εκείνο που περιλαμβάνει τις μεθόδους αριθμητικής ολοκλήρωσης στις οποίες μόνον ένα ορισμένο πλήθος σημείων παρεμβολής και συντελεστών $A_i^{(n)}$ είναι προκαθορισμένο και οι υπόλοιπες άγνωστες ποσότητες υπολογίζονται έτσι ώστε οι μέθοδοι να είναι ακριβείς για τα πολυώνυμα με το μεγαλύτερο δυνατόν βαθμό. Μία τέτοια μέθοδος ονομάζεται ***μέθοδος προσδιοριστέων συντελεστών*** και αποσκοπεί στον υπολογισμό των άγνωστων σημείων παρεμβολής και των άγνωστων συντελεστών $A_i^{(n)}$, έχοντας σαν στόχο την επίτευξη του παραπάνω κριτηρίου ακριβείας. Στην *Παράγραφο* αυτή θα αναφερθούμε στην μελέτη αυτών των μεθόδων.

Έστω ότι δίνεται ο ακόλουθος *Τύπος* τετραγωνισμού:

$$I = \sum_{i=0}^{n} A_i^{(n)} \, f(x_i) + R_n(f)$$

για τον αριθμητικό υπολογισμό του ολοκληρώματος:

$$I = \int_a^\beta f(x)\, w(x) dx .$$

Ας υποθέσουμε ότι, ανάμεσα στις $(2n+2)$ ποσότητες x_i και $A_i^{(n)}$, υπάρχουν k που είναι άγνωστες, και ας αξιώσουμε ο τύπος τετραγωνισμού να είναι ακριβής επί του διανυσματικού χώρου $P_{k-1}(R)$, όλων των πραγματικών πολυωνύμων βαθμού όχι μεγαλύτερου από $k-1$.

Αυτή η αξίωσή μας ισοδυναμεί με την ύπαρξη κατάλληλων αριθμών x_i και $A_i^{(n)}$ έτσι ώστε να επαληθεύεται το επόμενο σύστημα των *k* εξισώσεων:

$$\begin{aligned}
A_0^{(n)} \quad + \cdots + A_n^{(n)} \quad &= \int_a^\beta \quad w(x)\ dx = c_0 \\
A_0^{(n)} x_0 \ + \cdots + A_n^{(n)} x_n \ &= \int_a^\beta x \quad w(x)\, dx = c_1 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
A_0^{(n)} x_0^{k-1} + \cdots + A_n^{(n)} x_n^{k-1} &= \int_a^\beta x^{k-1}\, w(x) dx = c_{k-1}
\end{aligned}$$

Ας συμβολίσουμε με $V_n(x)$ το πολυώνυμο που έχει ρίζες τα x_i:

$$V_n(x) = (x - x_0)\,(x - x_1)...(x - x_n) = a_0 + a_1\ x + ... + a_n\ x^n + a_{n+1}\ x^{n+1}.$$

Προφανώς, για κάθε $i = 0,1,...,n$, ισχύει η εξίσωση:

$$V_n(x_i) = a_0 + a_1\ x_i + ... + a_n\ x_i^n + a_{n+1}\ x_i^{n+1} = 0 .$$

Ας υποθέσουμε τώρα ότι:

$$n + 2 \le k \le 2\,n + 2 .$$

Πολλαπλασιάζοντας την πρώτη από τις k εξισώσεις του προηγούμενου συστήματος με a_0, την δεύτερη με a_1, και την $(n+2)-$οστή με a_{n+1}, και στην συνέχεια αθροίζοντας κατά μέλη, παίρνουμε:

$$A_0^{(n)}\,(a_0 + a_1\,x_0 + \ldots + a_{n+1}\,x_0^{n+1}) + \cdots + A_n^{(n)}\,(a_0 + a_1\,x_n + \ldots + a_{n+1}\,x_n^{n+1}) = a_0\,c_0 + a_1\,c_1 + \ldots + a_{n+1}\,c_{n+1},$$

απ' όπου έπεται ότι:

$$a_0\,c_0 + a_1\,c_1 + \ldots + a_{n+1}\,c_{n+1} = 0\,.$$

Ομοίως, πολλαπλασιάζοντας την δεύτερη από τις k εξισώσεις του προηγούμενου συστήματος με a_0, την τρίτη με a_1, και την $(n+3)-$οστή με a_{n+1}, και στη συνέχεια αθροίζοντας κατά μέλη, βρίσκουμε ότι:

$$a_0\,c_1 + a_1\,c_2 + \ldots + a_{n+1}\,c_{n+2} = 0\,.$$

Επαναλαμβάνοντας την ίδια διαδικασία, μπορούμε να καταλήξουμε στο ακόλουθο σύστημα $k-n-1$ εξισώσεων με $(n+2)$ αγνώστους, τους αριθμούς $a_0,\ldots,a_{n+1}$:

$$(\Sigma_1)\quad \begin{cases} a_0\,c_0 + \ldots + a_{n+1}\,c_{n+1} = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_0\,c_{k-n-1} + \ldots + a_{n+1}\,c_{k-1} = 0. \end{cases}$$

Εάν υποθέσουμε τώρα ότι οι $(2n+2-k)$ γνωστές ποσότητες χωρίζονται σε:

p γνωστά σημεία παρεμβολής, τα $x_0, x_1, \ldots, x_{p-1}$,

και

q γνωστούς συντελεστές $A_i^{(n)}$, τους $A_0^{(n)}, A_1^{(n)}, \ldots, A_{q-1}^{(n)}$,

τότε $(2n+2-k) = p+q$, και επιπλέον μπορούμε να διατυπώσουμε και τις εξής αριθμητικές ταυτότητες για τα σημεία $x_0, x_1, \ldots, x_{p-1}$:

$$(\Sigma_2)\quad \begin{cases} a_0 + a_1\ x_0 + \ldots + a_{n+1}\ x_0^{n+1} = 0 \\ a_0 + a_1\ x_1 + \ldots + a_{n+1}\ x_1^{n+1} = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_0 + a_1\ x_{p-1} + \ldots + a_{n+1}\ x_{p-1}^{n+1} = 0. \end{cases}$$

Η ένωση των συστημάτων (Σ_1) και (Σ_2) στοιχειοθετεί ένα ευρύτερο σύστημα $(k-n-1+p)$ εξισώσεων με $(n+1)$ αγνώστους, τους αριθμούς:

$$a_0, a_1, \ldots, a_n,$$

επειδή ο αριθμός a_{n+1} μπορεί να επιλεγεί αυθαίρετα και, κατά συνέπεια, χωρίς βλάβη της γενικότητας μπορεί να εκληφθεί ίσος με 1.

Οι εξισώσεις που ενδεχομένως απουσιάζουν προκειμένου το σύστημα αυτό να διογκωθεί και να σχηματισθεί ένα τετραγωνικό σύστημα $(n+1)$ εξισώσεων, με τους ίδιους αγνώστους, δημιουργούνται με την βοήθεια των αξιοσημείωτων ταυτοτήτων, που περιγράφουν τις σχέσεις που υφίστανται μεταξύ των συντελεστών και των ριζών ενός πολυωνύμου.

Θεωρώντας το διογκωμένο αυτό τετραγωνικό γραμμικό σύστημα, μπορούμε να προχωρήσουμε στην επίλυσή του και, να βρούμε τους αριθμούς:

$$a_0, a_1, \ldots, a_n,$$

δηλαδή την αναλυτική μορφή του πολυωνύμου:

$$V_n(x) = a_0 + a_1\ x + \ldots + a_n\ x^n + x^{n+1}.$$

Σε ένα δεύτερο στάδιο, υπολογίζουμε τις άγνωστες ρίζες $x_p, x_{p+1}, \ldots, x_n$ του $V_n(x)$ και τότε προσδιορίζουμε τους άγνωστους συντελεστές $A_q^{(n)}, A_{q+1}^{(n)}, \ldots, A_n^{(n)}$.

Ας δώσουμε ένα συγκεκριμένο παράδειγμα, για να πετύχουμε την εύληπτη κατανόηση της εξέλιξης της μεθόδου.

Έστω ότι πρέπει να υπολογισθούν οι ποσότητες:

$$A_0^{(3)}, A_1^{(3)}, A_2^{(3)}, A_3^{(3)}, \quad x_1 \text{ και } x_2,$$

έτσι ώστε ο *Τύπος* τετραγωνισμού:

$$\int_{-1}^{1} f(x)dx = A_0^{(3)} f(-1) + A_1^{(3)} f(x_1) + A_2^{(3)} f(x_2) + A_3^{(3)} f(1) + R_3(f)$$

να είναι ακριβής για πολυώνυμα με όσο το δυνατόν μεγαλύτερο βαθμό.

Στην προκειμένη περίπτωση, θα αξιώσουμε ο παραπάνω *Τύπος* να είναι ακριβής για πολυώνυμα μέχρι και πέμπτου βαθμού, δηλαδή για τα μονώνυμα:

$$1, x, x^2, x^3, x^4 \quad \text{και} \quad x^5$$

και τους γραμμικούς συνδυασμούς τους.

Θέτουμε:

$$V_3(x) = (x - x_1)\,(x - x_2)\,(x - 1)\,(x + 1) = a_0 + a_1\,x + a_2\,x^2 + a_3\,x^3 + x^4.$$

Χρησιμοποιώντας την μέθοδο που υποδείχθηκε παραπάνω, βρίσκουμε:

$$\begin{array}{llll} \frac{2}{5} & + \frac{2}{5} & \frac{2}{3}a_2 + & = 0 \\ \frac{2}{5}a_3 & + \frac{2}{3}a_1 & & = 0 \end{array}$$

Επίσης, επειδή

$$V_3(1) = V_3(-1) = 0,$$

έχουμε τις, επί πλέον, εξισώσεις:

$$1+\quad a_3+\quad a_2+\quad a_1+\quad a_0=0\,,$$

$$1-\quad a_3+\quad a_2-\quad a_1+\quad a_0=0\,.$$

Επιλύοντας το ενοποιημένο σύστημα, που διαμορφώνεται από την συνεύρεση των τεσσάρων τελευταίων εξισώσεων, προσδιορίζουμε τους αριθμούς:

$$a_1=a_3=0\,,$$

$$a_2=-\frac{6}{5}\,,$$

$$a_0=\frac{1}{5}\,,$$

και κατά συνέπεια εντοπίζουμε την αναγκαία (για την ακρίβεια πολυωνύμων μέχρι και πέμπτου βαθμού) μορφή του $V_3(x)$:

$$V_3(x)=\frac{1}{5}-\frac{6}{5}x^2+x^4=\frac{1}{5}(x-1)(x+1)\left(5x^2-1\right),$$

απ' όπου έπεται αμέσως ότι:

$$x_1=-\frac{\sqrt{5}}{5}\quad \text{και}\quad x_2=\frac{\sqrt{5}}{5}\,.$$

Μεταφέροντας αυτές τις τιμές στις σχέσεις που εκφράζουν την ακρίβεια του μαθηματικού τύπου τετραγωνισμού για τα μονώνυμα:

$$1,\,x,\,x^2,\,x^3,$$

βρίσκουμε ότι:

$$A_0^{(3)} = A_3^{(3)} = \frac{1}{6},$$

$$A_1^{(3)} = A_2^{(3)} = \frac{5}{6}.$$

Έτσι, στο συγκεκριμένο παράδειγμα, διαμορφώνεται ο εξής *Τύπος* τετραγωνισμού:

$$\int_{-1}^{1} f(x)\,dx = \frac{1}{6}\left[f(-1)+f(1)\right]+\frac{5}{6}\left[f\left(-\frac{\sqrt{5}}{5}\right)+f\left(\frac{\sqrt{5}}{5}\right)\right]+R_3(f).$$

Αυτός ο *Τύπος* αποτελεί μία ειδική περίπτωση ενός γενικότερου τύπου τετραγωνισμού, που είναι γνωστός σαν ***μέθοδος Lobatto*** και που συνοψίζεται στην ακόλουθη έκφραση:

$$\int_{-1}^{1} f(x)\,dx = \frac{2}{n(n+1)}\left[f(-1)+f(1)\right]+\sum_{i=1}^{n-1} A_i^{(n)} f(x_i)+R_n(f),$$

όπου τα σημεία x_i ***είναι οι ρίζες της παραγώγου*** $P_n'(x)$ ***του πολυωνύμου Legendre*** $P_n(x)$***, που έχει βαθμό ίσο με n, και όπου:***

$$A_i^{(n)} = \frac{2}{n(n+1)P_n^2(x_i)} \qquad (x_i \neq \pm 1)$$

και

$$R_n(f) = -\frac{(n+1)n^3\,2^{2n+1}\left[(n-1)!\right]^4}{(2n+1)\left[(2n)!\right]^3}\,\frac{d^{(2n)}f}{dx^{(2n+2)}}(\xi)$$

για κάποιο $\xi \in [-1,1]$ ***και κάτω φυσικά από την προϋπόθεση ότι η*** $f(x)$ ***είναι*** $2n$ ***φορές συνεχώς διαφορίσιμη μέσα σε μία ανοικτή περιοχή του*** $[-1,1]$**.**

IV.2.12. ΑΡΙΘΜΗΤΙΚΟΣ ΥΠΟΛΟΓΙΣΜΟΣ ΠΟΛΛΑΠΛΩΝ ΟΛΟΚΛΗΡΩΜΑΤΩΝ

Σ' αυτήν την *Παράγραφο* δεν θα δώσουμε παρά μόνον κάποιες γενικές κατευθύνσεις και λίγες διαδικαστικές υποδείξεις σχετικές με το πρόβλημα του αριθμητικού υπολογισμού πολλαπλών ολοκληρωμάτων.

Για λόγους απλοποίησης, και χωρίς απώλεια της γενικότητας, θα περιοριστούμε στην περίπτωση των διπλών ολοκληρωμάτων.

Έστω λοιπόν ότι δίνεται προς υπολογισμό το ολοκλήρωμα:

$$I = \int_{a}^{\beta} \int_{g_1(x)}^{g_2(x)} f(x,y)\, dy\, dx .$$

Θέτουμε:

$$F(x) := \int_{g_1(x)}^{g_2(x)} f(x,y)\, dy ,$$

όπου το μέγεθος x διαδραματίζει τον ρόλο μίας παραμέτρου.

Τότε:

$$I = \int_{a}^{\beta} F(x)\, dx .$$

Για τον υπολογισμό του I χρησιμοποιούμε έναν πρώτο *Τύπου* τετραγωνισμού:

$$I = \sum_{i=0}^{n} A_i^{(n)}\ F(x_i) + R_1(F) .$$

Για τον υπολογισμό των $F(x_i)$ χρησιμοποιούμε έναν δεύτερο *Τύπο* τετραγωνισμού:

$$F(x_i) = \int_{g_1(x_i)}^{g_2(x_i)} f(x_i, y)\, dy = \sum_{j=0}^{m} B_j^{(m)}\ f(x_i, y_j) + R_2^{(i)}(f),$$

Συνδυάζοντας τις δύο παραπάνω σχέσεις, καταλήγουμε στον ακόλουθο *Τύπο*:

$$I = \sum_{i=0}^{n} A_i^{(n)} \sum_{j=o}^{m} B_j^{(m)}\ f(x_i, y_j) + R(f),$$

όπου:

$$R(f) = R_1(\mathrm{F}) + \sum_{i=0}^{n} A_i^{(n)} \; R_2^{(\mathrm{i})}(f).$$

Μετά απ' όσα προηγήθηκαν, διαπιστώνουμε ότι μία τέτοια μέθοδος συνίσταται

- κατ' αρχάς στην πραγματοποίηση μίας διατετραγώνισης του τόπου ολοκλήρωσης,
- κατόπιν στον υπολογισμό των τιμών της $f(x,y)$ στα γωνιακά σημεία αυτής της διατετραγώνισης,
- και τέλος στην θεώρηση των γραμμικών συνδυασμών αυτών των τιμών.

IV.3. ΑΡΙΘΜΗΤΙΚΗ ΠΑΡΑΓΩΓΙΣΗ

IV.3.1. ΤΟΠΟΘΕΤΗΣΗ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ

Έστω f μία πραγματική συνάρτηση μίας πραγματικής μεταβλητής, που διατρέχει ένα ανοικτό διάστημα I του R. Υποθέτουμε ότι η f είναι διαφορίσιμη στο I.

"Δοθέντος ενός σημείου $x \in I$, θα προσπαθήσουμε τώρα να υπολογίσουμε αριθμητικά μία ικανοποιητική προσεγγιστική τιμή της τιμής $f'(x)$ της παραγώγου της f στο σημείο x."

Η μέθοδος που θα υιοθετήσουμε στηρίζεται στην Θεωρία της Πολυωνυμικής Παρεμβολής: Εάν

$$x_0, x_1, \ldots, x_n \in I$$

είναι $(n+1)$ σημεία, ανά δύο διάφορα μεταξύ τους, θα αντικαταστήσουμε την συνάρτηση *f* με το πολυώνυμο παρεμβολής της, έστω p_n, στα σημεία αυτά κι έτσι θα λάβουμε την σχέση:

$$f(x) = p_n(x) + E_n(x),$$

και, κατ' ακολουθία, την σχέση:

$$f'(x) = p'_n(x) + E'_n(x).$$

Επειδή η παραγώγιση ως προς x του τύπου παρεμβολής Lagrange οδηγεί στην εξίσωση:

$$f'(x) = \sum_{i=0}^{n} B_i^{(n)}(x)\, f(x_i)\, S_n(f,x),$$

όπου:

$$B_i^{(n)}(x) := L_i'^{(n)}(x) \text{ και } S_n(f,x) := E'_n(x),$$

θα εκλαμβάνεται ως το πεπερασμένο άθροισμα:

$$\sum_{i=0}^{n} B_i^{(n)}(x)\, f(x_i)$$

σαν προσέγγιση του αριθμού $f'(x)$.

Το σφάλμα αυτής της αριθμητικής παραγώγισης, δηλαδή το σφάλμα της παραπάνω προσέγγισης του αριθμού $f'(x)$, ισούται με:

$$S_n(f,x).$$

Το επόμενο αποτέλεσμα είναι προφανές, γιατί απορρέει άμεσα από τον ίδιο τον ορισμό του πολυωνύμου παρεμβολής:

Θεώρημα IV.3.1.1. ***Μία αναγκαία και ικανή συνθήκη προκειμένου να ισχύει:***

$$S_n(f,x)=0 \quad \forall x$$

είναι η συνάρτηση $f(x)$ ***να ανήκει στο*** $P_n(R)$***, δηλαδή να είναι ένα πραγματικό πολυώνυμο βαθμού το πολύ ίσου με*** n***.*** ■

IV.3.2. ΜΕΛΕΤΗ ΤΟΥ ΣΦΑΛΜΑΤΟΣ

Θυμίζουμε ότι το σφάλμα της πολυωνυμικής παρεμβολής δίνεται από τον *Τύπο*:

$$E_n(x)=\frac{1}{(n+1)!}V_n(x)\frac{d^{(n+1)}f}{d\,x^{(n+1)}}(\xi),$$

κάτω από την προϋπόθεση ότι η συνάρτηση f είναι $(n+1)-$φορές συνεχώς διαφορίσιμη μέσα στο ανοικτό διάστημα I και όπου:

$$V_n(x)=(x-x_0)\,(x-x_1)...(x-x_n), \quad (\xi\in[\min_{0\le i\le n}\{x,x_i\},\max_{0\le i\le n}\{x,x_i\}]).$$

Παρατηρήστε πως η επιλογή του σημείου ξ εξαρτάται από το εκάστοτε x. Για να υπολογίσουμε το σφάλμα $S_n(f,x)$ της αριθμητικής παραγώγισης, αρκεί να παραγωγίσουμε την παραπάνω έκφραση του $E_n(x)$ ως προς x:

$$S_n(f,x)=E'_n(x)=\frac{1}{(n+1)!}[V'_n(x)\,f^{(n+1)}(\xi)+V_n(x)\frac{df^{(n+1)}}{dx^{(n+1)}}(\xi)].$$

Καθώς όμως δεν μπορούμε να γνωρίσουμε το ακριβές είδος της εξάρτησης του *ξ* από το *x*, είναι αδύνατον να υπολογίσουμε την παράγωγο ως προς x της (σύνθετης) συνάρτησης $f^{(n+1)}(\xi)$.

Για αυτό, θα επιλέξουμε το σημείο *x* έτσι ώστε ο τελευταίος όρος στην προηγούμενη έκφραση του $S_n(f,x)$ να είναι μηδενικός, δηλαδή έτσι ώστε κάποιο σημείο παρεμβολής, έστω το x_j, να συμπίπτει με το δοθέν σημείο x.

Σε ό,τι ακολουθήσει κατά την εξέλιξη αυτού του *Κεφαλαίου*, θα σεβόμαστε και θα υιοθετούμε σιωπηλά αυτήν την περίπτωση.

Κάτω από μία τέτοια παραδοχή, το σφάλμα $S_n(f,x)$ της αριθμητικής παραγώγισης θα ισούται με:

$$S_n(f,x_j)=\frac{1}{(n+1)!}\frac{d^{(n+1)}f}{dx^{(n+1)}}(\xi)\prod_{\substack{k=0\\k\neq j}}^{n}(x_j-x_k),$$

για κάποιο $\xi\in[\min_{0\le i\le n}x_i,\max_{0\le i\le n}x_i]$ και κάτω από την προϋπόθεση ότι η συνάρτηση f είναι $(n+1)$ φορές συνεχώς διαφορίσιμη στο I.

IV.3.3. ΥΠΟΛΟΓΙΣΜΟΣ ΤΩΝ ΣΥΝΤΕΛΕΣΤΩΝ

Θα υπολογίσουμε τώρα τους συντελεστές:

$$B_i^{(n)}(x_j)\quad(i=0,1,\ldots,n).$$

Για να επιτύχουμε στην προσπάθειά μας, αρκεί να παραγωγίσουμε ως προς *x* τις συναρτήσεις $L_i^{(n)}(x)$ που εμφανίζονται στον τύπο παρεμβολής του Lagrange.

Επειδή, όπως αναφέρθηκε στην *Παράγραφο IV.1.2*, ισχύει:

$$L_i^{(n)}(x) = \prod_{\substack{k=0 \\ k\neq i}}^{n} \frac{x - x_k}{x_i - x_k} = \frac{V_n(x)}{(x - x_i)\, V'_n(x_i)},$$

διαπιστώνουμε πως:

$$L_i'^{(n)}(x) = \frac{1}{V'_n(x_i)} \left[\frac{V'_n(x)}{x - x_i} - \frac{V_n(x)}{(x - x_i)^2} \right].$$

Ας θεωρήσουμε την περίπτωση όπου:

$$x = x_j .$$

Εάν $j \neq i$, τότε, με την βοήθεια του προηγούμενου τύπου, βρίσκουμε:

$$B_i^{(n)}(x_j) = \frac{V'_n(x_j)}{V'_n(x_i)\,(x_j - x_i)} = \frac{1}{x_j - x_i} \frac{\prod_{\substack{k=0 \\ k\neq j}}^{n} (x_j - x_k)}{\prod_{\substack{k=0 \\ k\neq i}}^{n} (x_i - x_k)}.$$

Εάν πάλι $j = i$, τότε δεν μπορούμε να χρησιμοποιήσουμε αυτόν τον *Τύπο*. Διαφοροποιούμαστε, γράφοντας την σχέση:

$$L_i^{(n)}(x) = \prod_{\substack{k=0 \\ k\neq i}}^{n} \frac{x - x_k}{x_i - x_k} = \frac{x - x_m}{x_i - x_m} \prod_{\substack{k=0 \\ k\neq i,m}}^{n} \frac{x - x_k}{x_i - x_k},$$

και στην συνέχεια, παραγωγίζοντας, ως προς x, την ισότητα αυτή:

$$B_i^{(n)}(x) = L_i^{(n)}(x) = \sum_{\substack{m=0 \\ m\neq i}}^{n} \frac{1}{x_i - x_m} \prod_{\substack{k=0 \\ k\neq i,m}}^{n} \frac{x - x_k}{x_i - x_k}.$$

Διαπιστώνουμε τότε πως όταν $x = x_j = x_i$, έχουμε:

$$B_j^{(n)}(x_j) = \sum_{\substack{k=0 \\ k\neq j}}^{n} \frac{1}{x_j - x_k}.$$

Συνοψίζοντας, ο *Τύπος* που δίνει τους συντελεστές αριθμητικής παραγωγίσης είναι ο εξής:

$$B_i^{(n)}(x_j) = \begin{cases} \dfrac{1}{(x_j - x_i)} \dfrac{\prod_{\substack{k=0 \\ k\neq j}}^{n} (x_j - x_k)}{\prod_{\substack{k=0 \\ k\neq i}}^{n} (x_i - x_k)}, & \text{εάν } j \neq i \\ \sum_{\substack{k=0 \\ k\neq j}}^{n} \dfrac{1}{(x_j - x_k)}, & \text{εάν } j = i. \end{cases}$$

IV.3.4. Η ΠΕΡΙΠΤΩΣΗ ΤΩΝ ΙΣΑΠΕΧΟΝΤΩΝ ΣΗΜΕΙΩΝ

Οι *Τύποι* των δύο προηγούμενων *Παραγράφων* που εκφράζουν το σφάλμα $S_n(f,x)$ και τους συντελεστές $B_i^{(n)}(x_j)$ της αριθμητικής παραγώγισης μίας συνάρτησης f, που είναι $(n+1)$ φορές συνεχώς διαφορίσιμη, στο σημείο $x = x_j$, απλοποιούνται σημαντικά όταν τα σημεία παρεμβολής x_i ισαπέχουν ανά δύο διαδοχικά:

$$x_i = x_0 + i\,h\,,\ i = 0,1,2,\ldots,n\,,$$

όπου h είναι μία θετική παράμετρος που ονομάζεται ***βήμα*** της αριθμητικής παραγώγισης.

Τότε, προφανώς, θα είναι:

$$x_i - x_k = (i-k)\,h \quad \forall i = 0,1,\ldots,n \ \text{ και } \ \forall k = 0,1,\ldots,n\,,$$

με αποτέλεσμα οι προηγούμενοι *Τύποι* να διαμορφώνονται ως εξής:

$$B_i^{(n)}(x_j) = \begin{cases} \dfrac{(-1)^{i+j}}{(i-j)\,h}\,\dfrac{i\,!\,(n-i)\,!}{j\,!\,(n-j)\,!}, & \text{εάν } \ j \neq i \\[2ex] \dfrac{1}{h}\sum_{\substack{k=0\\k\neq j}}^{n}\dfrac{1}{j-k}, & \text{εάν } \ j = i \end{cases}$$

και επίσης:

$$S_n(f,x_j) = \frac{f^{(n+1)}(\xi)}{(n+1)\,!}\,\Big(\prod_{\substack{k=0\\k\neq j}}^{n} h\,j-k\Big) = \frac{(-1)^{n-j}\,h^n\,j\,!(n-j)\,!\,f^{(n+1)}(\xi)}{(n+1)!}\,.$$

Γενικά, δεν έχουμε καμμία πληροφορία σχετική με τον αριθμό $f^{(n+1)}(\xi)$, και συνεπώς δεν μπορούμε να προβλέψουμε την συμπεριφορά της τιμής $f^{(n+1)}(\xi)$.

Αντιθέτως, είναι δυνατή η επιλογή του δείκτη j του σημείου $x_j = x$, πάνω στο οποίο υπολογίζεται η παράγωγος, έτσι ώστε να ελαχιστοποιηθεί η ποσότητα:

$$j!\,(n-j)!.$$

Πράγματι, ισχύει:

$$j!\,(n-j)! = \frac{C_n^j}{n!},$$

όπου C_n^j είναι αριθμός των συνδυασμών n αντικειμένων ανά κλάσεις που περιέχουν j απ' αυτά.

Επειδή, όπως είναι γνωστό,

$$C_n^j = C_n^{(n-j)}$$

και το C_n^j γίνεται ελάχιστο όταν το j ισούται με τον ακέραιο που βρίσκεται πιο κοντά στο $\frac{n}{2}$, βλέπουμε ότι έχουμε ιδιαίτερο όφελος εάν επιλέξουμε:

$$n = 2p \text{ (δηλαδή, } n = \text{άρτιος), και } \quad j = \frac{n}{2} = p.$$

Τότε, το σφάλμα $S_n(f, x_p)$ της αριθμητικής παραγώγισης της f στο σημείο x_p θα ισούται με:

$$S_n(f, x_p) = \frac{(-1)^p h^{2p} (p!)^2}{(2p+1)!} f^{(2p+1)}(\xi),$$

για κάποιο $\xi \in [\min_{0 \le i \le 2p} x_i, \max_{0 \le i \le 2p} x_i]$, και φυσικά πάντα κάτω από την προϋπόθεση ότι η συνάρτηση f είναι $(2p+1)-$φορές συνεχώς διαφορίσιμη στο σημείο ξ.

IV.3.4.(i). Αριθμητική Ευστάθεια Η έννοια της αριθμητικής ευστάθειας για την παραγώγιση είναι ανάλογη με αυτήν που μελετήσαμε στην *Παράγραφο IV.2.2* για τον τετραγωνισμό.

Συγκεκριμένα, στην πραγματικότητα, αντί να υπολογίσουμε το άθροισμα:

$$\sum_{i=0}^{n} B_i^{(n)}(x_j)\, f(x_i),$$

για να σχηματίσουμε μία προσέγγιση της τιμής $f'(x_j)$, υπολογίζουμε μία ποσότητα της μορφής:

$$\sum_{i=0}^{n} B_i^{(n)}(x_j)\,[f(x_i)+\varepsilon_i],$$

δηλαδή, συμπεριλαμβάνουμε στο αποτέλεσμα μας και ένα σφάλμα του οποίου το μέγεθος είναι:

$$\sum_{i=0}^{n} B_i^{(n)}(x_j)\,\varepsilon_i\,.$$

Δίνουμε λοιπόν τον ακόλουθο *Ορισμό*:

Ορισμός IV.3.4.(i).1. ***Ένας μαθηματικός Τύπος αριθμητικής παραγώγισης, που θεμελιώνεται σε*** $(n+1)$ ***σημεία παρεμβολής, λέγεται ευσταθής εάν για οποιαδήποτε*** $\varepsilon_0,\varepsilon_1,\ldots,\varepsilon_n \in \mathbb{R}$ ***υπάρχει μία σταθερά*** $M = M(\varepsilon_0,\varepsilon_1,\ldots,\varepsilon_n)$***, ανεξάρτητη από το n, τέτοια ώστε:***

$$\left|\sum_{i=0}^{n} B_i^{(n)}(x_j)\,\varepsilon_i\right| \le M\ \max_{0\le i\le n}|\varepsilon_i|\,. \blacksquare$$

Είναι εύκολο να επαληθεύσουμε πως, ***στην περίπτωση των ισαπεχόντων σημείων παρεμβολής με βήμα h, ένας Μαθηματικός Τύπος αριθμητικής παραγώγισης είναι ευσταθής εάν, για οποιαδήποτε $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_n \in \mathbb{R}$, υπάρχει μία σταθερά $M = M(\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_n)$ που είναι ανεξάρτητη από το βήμα h και τέτοια ώστε να ικανοποιείται η ανισότητα του προηγούμενου Ορισμού***.

Όμως, για την περίπτωση αυτή, έχουμε το εξής αρνητικό αποτέλεσμα:

Θεώρημα IV.3.4.(i).2. ***Κάθε μαθηματικός τύπος αριθμητικής παραγώγισης, που θεμελιώνεται σε $(n+1)$ σημεία παρεμβολής τα οποία ισαπέχουν, ανά δύο διαδοχικά, με βήμα h, είναι ασταθής.***

Απόδειξη. Εάν πάρουμε:

$$\varepsilon_i = \begin{cases} 1, & \text{εάν } i = k \\ 0, & \text{εάν } i \neq k \end{cases}$$

τότε, θα έχουμε:

$$\left| \sum_{i=0}^{n} B_i^{(n)}(x_j)\, e_i \right| = \left| B_k^{(n)}(x_j) \right| .$$

Σύμφωνα όμως με όσα εκτέθηκαν παραπάνω ισχύει ότι:

$$B_k^{(n)}(x_j) = \begin{cases} \dfrac{(-1)^{k+j}}{(k-j)\,h} \cdot \dfrac{k!\,(n-k)!}{j!\,(n-j)!}, & \text{εάν } j \neq k \\ \dfrac{1}{h} \displaystyle\sum_{\substack{m=0 \\ m \neq j}}^{n} \dfrac{1}{j-m}, & \text{εάν } j = k. \end{cases}$$

Άρα, η ποσότητα

$$B_k^{(n)}(x_j)$$

δεν είναι δυνατόν να φράσσεται από μία σταθερά που είναι ανεξάρτητη από το βήμα h και με το συμπέρασμα αυτό ολοκληρώνεται η *Απόδειξη* του *Θεωρήματος*. ■

IV.3.4.(ii). Χρήσιμοι Τύποι Σαν συνέπεια του αρνητικού αποτελέσματος που εμφανίσθηκε στο *Θεώρημα IV.3.4.(i).1*, επισημαίνουμε ότι, στην περίπτωση των ισαπεχόντων σημείων παρεμβολής, οι Μαθηματικοί *Τύποι* αριθμητικής παραγώγισης που εφαρμόζονται στα διάφορα προβλήματα πρέπει να επιλέγονται με ιδιαίτερη σύνεση και προσοχή.

Στην ***πράξη***, οι πλέον διαδεδομένοι είναι οι μαθηματικοί τύποι αριθμητικής παραγώγισης, που θεμελιώνονται σε τρία ισαπέχοντα σημεία παρεμβολής. Τότε,

$$n = 2$$

και οι Τύποι που χρησιμοποιούνται είναι οι εξής:

$$f'(x_1) = \frac{1}{2h}[f(x_2) - f(x_0)] - \frac{h^2}{6} f'''(\xi),$$

$$f'(x_0) = \frac{1}{2h}[-3f(x_0) + 4f(x_1) - f(x_2)] + \frac{h^2}{3} f'''(\xi),$$

$$f'(x_2) = \frac{1}{2h}[f(x_0) - 4f(x_1) + 3f(x_2)] - \frac{h^2}{3} f'''(\xi),$$

όπου

$$x_i = x_0 + ih \quad (i = 0,1,2),$$

$h > 0$, ***και***

για κάποιο $\xi \in [x_0, x_2]$,

και κάτω από την προϋπόθεση ότι η f ***είναι τρεις φορές συνεχώς διαφορίσιμη στο σημείο*** ξ.

IV.3.4.(iii). Παραγώγιση και Τελεστές Διαφορών Θα δείξουμε τώρα πώς μπορούμε να επιτύχουμε την διαμόρφωση των Μαθηματικών Τύπων της αριθμητικής παραγώγισης, στην περίπτωση των ισαπεχόντων σημείων παρεμβολής, με τη βοήθεια της Θεωρίας των Τελεστών Διαφορών, που συναντήσαμε στην *Παράγραφο IV.1.5*.

Ας υποθέσουμε ότι η συνάρτηση f αναπτύσσεται σε σειρά Taylor, ως εξής:

$$f(x+h) = f(x) + \frac{h}{1!} f'(x) + \frac{h^2}{2!} f''(x) + \frac{h^3}{3!} f'''(x) + \cdots$$

Έστω D ο τελεστής παραγώγισης $\frac{d}{dx}$, που ορίζεται σύμφωνα με τον *Τύπο*:

$$Df(x) = f'(x).$$

Οι διαδοχικές δυνάμεις του D ορίζονται σύμφωνα με τον επαγωγικό *Τύπο*:

$$D^{\nu} f(x) = D\,(D^{\nu-1} f)(x) \qquad (\nu = 2,3,\ldots).$$

Χρησιμοποιώντας τους ορισμούς αυτούς, καθώς και τον συντελεστή μετατόπισης E:

$$E\,f(x) = f(x+h),$$

όπως αυτός είχε εισαχθεί στην *Παράγραφο IV.1.5*, το ανάπτυγμα σε σειρά Taylor της f, γύρω από το x, μπορεί να γράφει υπό την μορφή:

$$Ef(x) = (id + \frac{h}{1!} D + \frac{h^2}{2!} D^2 + \frac{h^3}{3!} D^3 + \ldots)\, f(x).$$

Άρα, μπορούμε να συμπεράνουμε την ακόλουθη τυπική σχέση:

$$E = id + \frac{h}{1!} D + \frac{h^2}{2!} D^2 + \frac{h^3}{3!} D^3 + \ldots$$

(όπου id είναι ο ταυτοτικός τελεστής), ή ακόμα την ισοδύναμη και πιο συμπυκνωμένη τυπική ταυτότητα:

$$E = \exp(h\ D) \Leftrightarrow h.D = LogE.$$

Θυμίζοντας ότι ο τελεστής διαφορών Δ, που ορίσαμε επίσης στην *Παράγραφο IV.1.5*, ικανοποιεί την σχέση $\Delta = E - id$, βλέπουμε ότι, επειδή:

$$E = \Delta + id,$$

το δεξιό μέλος της παραπάνω τυπικής ταυτότητας επαναδιατυπώνεται ως εξής:

$$h\ D = Log(id + \Delta).$$

Χρησιμοποιώντας το ανάπτυγμα σε σειρά της συνάρτησης $Log(1+z)$, η τελευταία ισότητα γράφεται υπό την μορφή:

$$h\ D = \Delta - \frac{\Delta^2}{2} + \frac{\Delta^3}{3} - \ldots + (-1)^{n+1}\frac{\Delta^n}{n} + \ldots,$$

απ' όπου έπεται ο ακόλουθος *Τύπος*:

$$f'(x_0) = \frac{1}{h}[\Delta f(x_0) - \frac{1}{2}\Delta^2 f(x_0) + \frac{1}{3}\Delta^3 f(x_0) - \ldots].$$

Από τον *Τύπο* αυτόν δημιουργούμε διάφορες προσεγγίσεις της τιμής $f'(x_0)$, απομονώνοντας μόνον τους πρώτους όρους της σειράς. Έτσι, διατηρώντας μόνον τον πρώτο όρο, σχηματίζεται η προσέγγιση:

$$f'(x_0) \approx \frac{1}{h}[f(x_1) - f(x_0)],$$

ενώ διατηρώντας μόνον τους δύο πρώτους όρους, έχουμε την προσέγγιση:

$$f'(x_0) \approx \frac{1}{2h}[-3\,f(x_0) + 4\,f(x_1) - f(x_2)].$$

Παρατηρήστε πως εάν $f \in \mathcal{P}_n(\mathbb{R})$, τότε, σύμφωνα με το *Θεώρημα IV.1.5.4* και το *Πόρισμα IV.1.6.2*, ισχύει ότι:

$$\Delta^{n+1} f(x_0) = 0.$$

Επομένως, ο προσεγγιστικός *Τύπος* της αριθμητικής παραγώγισης, που προκύπτει από το άθροισμα των πρώτων n όρων της παραπάνω σειρά, και μόνον αυτών, είναι ακριβής επί του διανυσματικού χώρου $\mathcal{P}_n(\mathbb{R})$, όλων των πραγματικών πολυωνύμων βαθμού το πολύ ίσου με n.

Επίσης, παρατηρήστε πως οι τύποι, που προκύπτουν κατ' αυτόν τον τρόπο, επιτρέπουν την προσέγγιση της τιμής της παραγώγου της f μόνον στο σημείο x_0. Εάν κανείς ήθελε να έχει παρόμοια αποτελέσματα για το σημείο x_n, θα έπρεπε να είχε χρησιμοποιήσει έναν ανάλογο *Τύπο*, όμως αυτός ο τύπος δεν θα περιείχε τον τελεστή Δ αλλά τον τελεστή ∇ των προς τα πίσω πεπερασμένων διαφορών, που γνωρίσαμε πάλι στην *Παράγραφο IV.1.5*.

Ομοίως, εάν κανείς ήθελε να προσεγγίσει την τιμή της f' στο σημείο x_p, στην περίπτωση όπου $n = 2\,p$, θα μπορούσε πάλι να χρησιμοποιήσει έναν ανάλογο τύπο, που όμως θα εκφραζόταν με την βοήθεια του τελεστή δ των κεντρικών διαφορών.

IV.3.4.(iv). Βελτίωση Ακριβείας Η ιδέα είναι η ίδια με αυτήν της μεθόδου Romberg. Υπενθυμίζουμε πως η μέθοδος Romberg αποσκοπεί στην βελτίωση των αποτελεσμάτων της μεθόδου τραπεζίων, κατά τον υπολογισμό ορισμένων ολοκληρωμάτων.

Έτσι, θα υπολογίσουμε διαδοχικές προσεγγίσεις του αριθμού $f'(x_j)$, που αντιστοιχούν σε διαδοχικές διαφορετικές επιλογές βημάτων

$$h_0, h_1, h_2, \ldots$$

Στη συνέχεια, θα θεωρήσουμε ένα πολυώνυμο παρεμβολής που διέρχεται από τις αυτών των διαδοχικών προσεγγίσεων και τέλος, θα υπολογίσουμε την τιμή αυτού του πολυωνύμου στο σημείο μηδέν.

Η μεταβλητή αυτού του πολυωνύμου θα επιλέγεται σύμφωνα με τον *Τύπο* που εκφράζει το σφάλμα της αριθμητικής παραγώγισης.

Έστω, για παράδειγμα, ότι πρέπει να προσεγγιστεί αριθμητικά η τιμή:

$$f'(x_1),$$

με έναν *Τύπο* τριών σημείων. Τότε, γράφουμε:

$$f(x_0) = f(x_1) - h\, f'(x_0) + \frac{h^2}{2} f''(x_1) - \ldots + (-1)^p\, h^p\, o(h),$$

και:

$$f(x_2) = f(x_1) - h\, f'(x_0) + \frac{h^2}{2} f''(x_1) - \ldots + h^p\, o(h)$$

όπου το σύμβολο $o(h)$ αναπαριστά μία πραγματική συνάρτηση του h τέτοια ώστε:

$$\lim_{h\to 0} \frac{o(h)}{h} = 0.$$

Προσθέτοντας την δεύτερη από τις παραπάνω σχέσεις με την αντίθετη της πρώτης, βρίσκουμε:

$$f'(x_1) = \frac{1}{2h}[f(x_2) - f(x_0)] - \frac{h^3}{3!} f^{(3)}(x_1) - \frac{h^4}{5!} f^{(5)}(x_1) - \ldots - h^{2k}\, o(h).$$

Τότε, για $i = 0,1,2,...,$ θέτουμε:

$$T_0^{(i)} := \frac{1}{2\,h_i}\,[f(x_1 + h_i) - f(x_1 - h_i)],$$

και συνδυάζουμε αυτές τις τιμές με την βοήθεια του τύπου παρεμβολής Neville-Aitken ως προς την μεταβλητή $x = h^2$, επειδή το σφάλμα της αριθμητικής παραγώγισης είναι ένα άρτιο πολυώνυμο ως προς h.

Θα καταλήξουμε εντοπίζοντας *Τύπους* προσεγγιστικής παραγώγισης τάξης όλο και περισσότερο πιο αυξημένης.

Ας υπολογίσουμε, για παράδειγμα την τιμή

$$T_1^{(0)}.$$

Επειδή:

$$T_1^{(0)} = \frac{h_0^2\ T_0^{(1)} - h_1^2\ T_0^{(0)}}{h^2 - h_1^2}$$

$$= f'(x_1) - \frac{h_0^2\ h_1^2}{5!}\,f^{(5)}(x_1) + \frac{h_0^2\ h_1^2}{h_0^2 - h_1^2}\,[h_1^2\ o(h_1) + h_2^2\ o(h_2)],$$

θέτοντας:

$$h_1 := h \quad \text{και} \quad h_0 := a\,h,$$

λαμβάνουμε:

$$T_0^{(0)} = f'(x_1) - \frac{a^2\,h^4}{5!}\,f^{(5)}(x_1) + h^4\ o(h).$$

Πρέπει, τέλος, να σημειωθεί ότι μπορούμε να αποδείξουμε αποτελέσματα που είναι ανάλογα με εκείνα που αναφέρθηκαν κατά την παρουσίαση της μεθόδου Romberg, στην περίπτωση της βελτίωσης της ακρίβειας του αριθμητικού τετραγωνισμού, όπως παρουσιάσθηκαν στην *Παράγραφο IV.2.6*.

Δεν πρόκειται να προχωρήσουμε σε κάποια παρουσίαση αυτών των αποτελεσμάτων. Συνιστούμε όμως σε κάθε ενδιαφερόμενο αναγνώστη να ανατρέξει, για μία ολοκληρωμένη μελέτη, στην βιβλιογραφική παραπομπή $[40]$.

IV.3.5 ΠΑΡΑΓΩΓΙΣΗ ΑΝΩΤΕΡΗΣ ΤΑΞΗΣ.

Στην *Παράγραφο* αυτή θα αντιμετωπίσουμε το πρόβλημα της ανεύρεσης μίας προσεγγιστικής αριθμητικής τιμής της ποσότητας $f^{(k)}(x_j)$, που αναπαριστά την τιμή της $k-$οστής παραγώγου της συνάρτησης f στο σημείο $x_j \quad (k>1)$.

Προς τούτο, μπορούμε να μετέλθουμε δύο μεθόδων:

- *ή να παραγωγίσουμε k φορές το πολυώνυμο που παρεμβάλλει την f στα σημεία:*

$$x_0, x_1, \ldots, x_j, \ldots, x_n,$$

- *ή να συνδυάσουμε τους προαναφερθέντες Τύπους, χρησιμοποιώντας το γεγονός ότι:*

$$f^{(k)}(x_j) = \frac{d}{dx} f^{(k-1)}(x_j).$$

Η πρώτη μέθοδος είναι δύσχρηστη, γιατί ο υπολογισμός των μεγεθών:

$$[L_i^{(n)}(x)]^{(k)}$$

είναι δύσκολος και συνήθως πολύπλοκος και μακροσκελής.

Αντίθετα, η δεύτερη μέθοδος προσφέρεται στις διάφορες εφαρμογές. Ας δώσουμε μία σαφή και περιεκτική εικόνα του δεύτερου τρόπου αντιμετώπισης του προβλήματος, στην περίπτωση που τα σημεία παρεμβολής ισαπέχουν ανά δύο διαδοχικά. Ας πάρουμε, για ενδεικτικό παράδειγμα, την περίπτωση που:

$$n = 2 \quad \text{και} \quad k = 2.$$

Τότε, αντικαθιστώντας την f με f', μέσα στις σχέσεις της *Υποπαραγράφου IV.3.4(ii)*, βλέπουμε ότι ισχύει:

$$f''(x_1) = \frac{1}{2h}[f'(x_2) - f'(x_0)] - \frac{h^2}{6} f^{(4)} + \mathrm{O}(h^4),$$

όπου το σύμβολο $\mathrm{O}(h^4)$ αναπαριστά μία πραγματική συνάρτηση του h που είναι τέτοια ώστε:

$$\frac{|\mathrm{O}(h^4)|}{|h^4|} \le C \text{, για κάθε } |h| \text{ αρκετά μικρό.}$$

Επειδή όμως:

$$f'(x_2) = \frac{1}{2h}[f(x_0) - 4\,f(x_1) + 3\,f(x_2)] + \frac{h^2}{3}\,f'''(x_2) + \mathrm{O}(h^4)$$

και:

$$f'(x_0) = \frac{1}{2h}[-3.f(x_0) + 4\,f(x_1) - f(x_2)] + \frac{h^2}{3}\,f'''(x_0) + \mathrm{O}(h^4),$$

η προηγούμενη σχέση συνεπάγεται ότι:

$$f''(x_1) = \frac{1}{h^2}[f(x_0) - 2\,f(x_1) + f(x_2)] - \frac{h^2}{12}\,f^{(4)}(\xi),$$

για κάποιο $\xi \in [x_0, x_2]$ και κάτω από την προϋπόθεση ότι η συνάρτηση f είναι τέσσερις φορές συνεχώς διαφορίσιμη μέσα σε μία ανοικτή περιοχή του κλειστού διαστήματος $[x_0, x_2]$.

Στην περίπτωση των ισαπεχόντων σημείων παρεμβολής, μπορούμε επίσης να επιτύχουμε την δημιουργία κατάλληλων προσεγγιστικών τύπων αριθμητικής παραγώγισης, υψώνοντας στην δύναμη k τα μέλη της εξίσωσης:

$$h\,D = Log(id + \Delta).$$

Τότε, σχηματίζεται η ισότητα:

$$D^k = \frac{1}{h^k}\left(\Delta^k - \frac{k}{2}\Delta^{k+1} + \frac{k(3k+5)}{24}\Delta^{k+2} - \cdots\right).$$

Απομονώνοντας μόνον τους πρώτους όρους της σειράς που βρίσκεται στο δεξιό μέλος της τελευταίας ισότητας, μπορούμε να δημιουργήσουμε διάφορες προσεγγίσεις του διαφορικού τελεστή D^k, και κατά συνέπεια, της τιμής

$$f^{(k)}(x_0) = D^k f(x_0)$$

της συνάρτησης f στο σημείο x_0.

Οι *Τύποι* αυτού του είδους χρησιμοποιούνται κατά την αριθμητική ολοκλήρωση διαφορικών εξισώσεων δεύτερης τάξης, όταν είναι γνωστές οι τιμές της λύσεις σε δύο διαφορετικά σημεία.

IV.3.6.ΠΑΡΑΓΩΓΙΣΗ ΣΥΝΑΡΤΗΣΕΩΝ ΠΟΛΛΩΝ ΜΕΤΑΒΛΗΤΩΝ

Στην τελευταία *Παράγραφο* αυτού του *Κεφαλαίου*, θα υποθέσουμε ότι η f είναι μία πραγματική συνάρτηση πολλών (πραγματικών) μεταβλητών.

Η κάθε τάξης παραγώγιση κάθε τάξης της f, ως προς κάποια από τις πολλές αυτές μεταβλητές πραγματοποιείται χρησιμοποιώντας τις μεθόδους των προηγούμενων *Παραγράφων*.

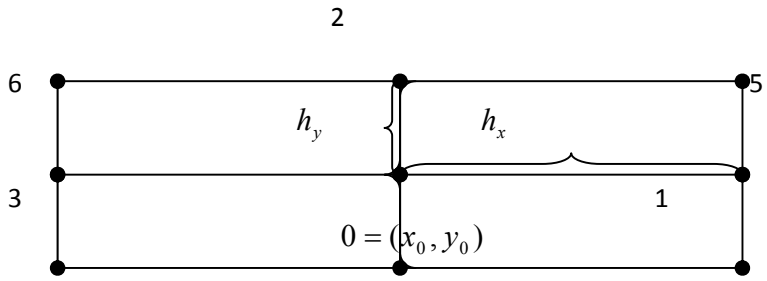
Ο υπολογισμός των διασταυρούμενων παραγώγων γίνεται συνθέτοντας τους *Τύπους* αριθμητικής παραγώγισης των προηγούμενων *Παραγράφων* για κάθε μία από τις μεταβλητές.

Έστω, για παράδειγμα, ότι η f είναι πραγματική συνάρτηση των δύο μεταβλητών $x \in \mathbb{R}$ και $y \in \mathbb{R}$. Έστω ακόμη ότι ζητείται να υπολογισθεί η τιμή της δεύτερης μερικής παραγώγου:

$$\frac{\partial^2 f}{\partial x \partial y}$$

της f στο σημείο (x_0, y_0).

Προς τούτο, σχηματίζουμε ένα ορθογώνιο δικτυωτό πλέγμα σημείων, στο επίπεδο που σχηματίζουν τα ζεύγη $(x, y) \in \mathbb{R}^2$, και γύρω από το σημείο (x_0, y_0), όπως στο παρακάτω σχήμα:



Στο σχήμα αυτό, το ζεύγος (x_0, y_0) ταυτίζεται με το σημείο του επιπέδου που έχουμε αναπαραστήσει με το σύμβολο 0, το μήκος του ευθυγράμμου τμήματος μεταξύ των σημείων 0 και 1 συμβολίζεται με h_x, και το μήκος του ευθυγράμμου τμήματος μεταξύ των σημείων 0 και 2 συμβολίζεται με h_y.

Σε ό,τι ακολουθήσει, θα συμβολίζουμε με f_i την τιμή της f στο σημείο i του πλέγματος.

Επειδή η δεύτερη μερική παράγωγος της f ως προς το x και y ισούται με την πρώτη μερική παράγωγο ως προς y της f, μπορούμε να γράψουμε:

$$\frac{\partial^2 f_0}{\partial x\,\partial y} = \frac{1}{2\,h_x}\left[\frac{\partial f_1}{\partial y} - \frac{\partial f_3}{\partial y}\right] + R_1 \quad .$$

Καθώς όμως, ισχύει:

$$\frac{\partial f_1}{\partial y} = \frac{1}{2\,h_y}[f_5 - f_8] + R_2 \qquad \text{και} \qquad \frac{\partial f_3}{\partial y} = \frac{1}{2\,h_y}[f_6 - f_7] + R_3\,,$$

συμπεραίνουμε πως:

$$\frac{\partial^2 f_0}{\partial x \partial y} = \frac{1}{4\,h_x\,h_y}[f_5 - f_8 - f_6 + f_7] + R\,,$$

με:

$$R = R_1 + \frac{1}{2\,h_x}[R_2 - R_3].$$

Οι υπόλοιπες διασταυρούμενες μερικές παράγωγοι:

$$\frac{\partial^2 f_i}{\partial x \partial y}$$

της f στα σημεία του πλέγματος βρίσκονται με τον ίδιο τρόπο.

Για παράδειγμα, χρησιμοποιώντας ανάλογα επιχειρήματα και σκέψεις, μπορούμε να διαπιστώσουμε ότι:

$$\frac{\partial^2 f_5}{\partial x\,\partial y} = \frac{1}{4\,h_x\,h_y}[f_7 - 4\,f_4 + 3\,f_8 + 16\,f_0 - 12\,f_1 + 3\,f_6 - 12\,f_2 + 3\,f_5] + \widetilde{R}\,.$$

IV.4. ΟΛΟΚΛΗΡΩΣΗ ΔΙΑΦΟΡΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

IV.4.1. ΤΟ ΠΡΟΒΛΗΜΑ CAUCHY

Έστω $[a,\beta]$ ένα κλειστό διάστημα του $\mathbb{R}$.

Έστω, επίσης, F μία απεικόνιση του καρτεσιανού γινομένου $\mathbb{R}\times\mathbb{R}^k$ μέσα στον Ευκλείδειο χώρο $\mathbb{R}^k$ και y μία διαφορίσιμη απεικόνιση του $\mathbb{R}$ μέσα στον $\mathbb{R}^k$.

Ορισμός IV.4.1.1. ***Ονομάζουμε διαφορικό σύστημα πρώτης τάξης την σχέση:***

$$\frac{dy(x)}{dx}=F(x,y(x)).$$

Λέμε ότι η απεικόνιση y είναι λύση αυτού του διαφορικού συστήματος επί του $[a,\beta]$, εάν η $y(x)$ επαληθεύει την παραπάνω σχέση, για κάθε $x\in[a,\beta]$. Δοθέντος ενός σημείου $y_0\in\mathbb{R}^n$, ονομάζουμε πρόβλημα Cauchy το σύστημα του προηγούμενου διαφορικού συστήματος και της αρχικής συνθήκης $y(a)=y_0$:

$$\begin{cases}\dfrac{dy(x)}{dx}=F(x,y(x))\\ y(a)=y_0.\end{cases}\ \blacksquare$$

Το θεμελιώδες αποτέλεσμα είναι το επόμενο:

Θεώρημα IV.4.1.2. ***Εάν η απεικόνιση F ικανοποιεί μία συνθήκη Lipschitz μέσα στο $[a,\beta]\times\mathbb{R}^k$, δηλαδή εάν υπάρχει μία σταθερά $L>0$ τέτοια ώστε:***

$$\|F(x,y)-F(x,z)\| \le L\,\|y-z\|, \quad \forall x\in[a,\beta] \quad \textbf{και} \quad \forall y,z\in\mathbb{R}^k$$

(όπου το σύμβολο $\|\ \|$ αναπαριστά, όπως συνήθως, την Ευκλείδεια νόρμα του $\mathbb{R}^k$), *τότε το πρόβλημα Cauchy επιδέχεται μία και μοναδική λύση*. (Η σταθερά L καλείται *σταθερά Lipschitz* της F).

Απόδειξη. Για λόγους απλοποίησης των πράξεων, θα περιοριστούμε στην περίπτωση όπου $k=1$.

Έστω $C[a,\beta]$ ο χώρος Banach όλων των πραγματικών συναρτήσεων *g* που είναι ορισμένες και συνεχείς επί του $[a,\beta]$ με νόρμα:

$$\|g\| := \sup_{t\in[a,\beta]}|g(t)|.$$

Η λύση y του προβλήματος Cauchy ικανοποιεί την σχέση:

$$y(x) = y_0 + \int_a^x F(t,y(t))\,dt \qquad \forall x\in[a,\beta].$$

Θα αποδείξουμε πρώτα την μοναδικότητα της λύσης.

Προς τούτο, ας υποθέσουμε ότι υπάρχει και δεύτερη λύση $z(x)\in C[a,\beta]$ του προβλήματος Cauchy. Τότε, πρέπει να είναι:

$$y(x)-z(x) = \int_a^x [F(t,y(t))-F(t,z(t))]\,dt,$$

και, κατ' ακολουθία, επειδή η F ικανοποιεί τη συνθήκη Lipschitz μέσα στο $[a,\beta]\times\mathbb{R}^1$ με σταθερά L, πρέπει να ισχύει η σχέση:

$$\left|y(x)-z(x)\right| \le L\int_a^x \left|y(t)-z(t)\right| dt\,, \qquad (22)$$

από την οποία, περνώντας σε μεγιστικές τιμές, έπεται η ανισότητα:

$$\left|y(x)-z(x)\right| \le L\,(x-a)\,\|y-z\| \qquad \forall x\in[a,\beta]. \qquad (23)$$

Μεταφέροντας την (23) στην (22), βρίσκουμε ότι:

$$\left|y(x)-z(x)\right| \le L^2\,\|y-z\|\int_a^x (t-a)dt = L^2\,\frac{(x-a)^2}{2}\,\|y-z\|,$$

αλλά και γενικότερα, χρησιμοποιώντας μαθηματική επαγωγή, μπορούμε να δείξουμε:

$$|\,y(x)-z(x)\,| \le L^n\,\frac{(x-a)^n}{n!}\,\|y-z\|, \forall n\ge 0 \text{ και } x\in[a,\beta]. \qquad (24)$$

(Πράγματι, εάν δεχτούμε την (24), τότε αφού την μεταφέρουμε μέσα στην ανισότητα (22) βλέπουμε ότι:

$$|\,y(x)-z(x)\,| \le \frac{L^{n+1}}{n!}\,\|y-z\|\,\int_a^x (t-a)^n\,dt = L^{n+1}\,\frac{(x-a)^{n+1}}{(n+1)!}\,\|y-z\|.)$$

Λαμβάνοντας την μέγιστη τιμή του αριστερού μέλους της (24), καθώς το x διατρέχει το διάστημα $[a,\beta]$, διαπιστώνουμε ότι:

$$\|y-z\| \le L^n\,\frac{(\beta-a)^n}{n!}\,\|y-z\|$$

ή ισοδύναμα ότι:

$$\|y-z\| \; (1-L^n \frac{(\beta-a)^n}{n!}) \le 0 .$$

Θεωρώντας το όριο, καθώς το *n* τείνει προς το άπειρο, συμπεραίνουμε πως:

$$\|y-z\| \le 0 ,$$

που συμβαίνει μόνο όταν:

$$\|y-z\| = 0 ,$$

δηλαδή όταν:

$$y = z .$$

Άρα, η λύση του προβλήματος Cauchy είναι μοναδική.

Θα αποδείξουμε τώρα την ύπαρξη αυτής της λύσης.

Στην προσπάθειά μας αυτή, το σκεπτικό, στο οποίο θα στηριχτούμε, βασίζεται στην *Αρχή των Επαναλήψεων Picard*: Ας θεωρήσουμε την ακολουθία συναρτήσεων $(\tilde{y}_m \in C[a,\beta]: \; m = 0,1,2,\ldots)$, που ορίζονται από τις σχέσεις:

$$\tilde{y}_0(x) = y_0 \;\; \text{και} \;\; \tilde{y}_{m+1}(x) = y_0 + \int_a^x F(t,\tilde{y}_m(t))dt, \;\; m = 0,1,2,\ldots$$

Τότε, επειδή:

$$\tilde{y}_{m+1}(x) - \tilde{y}_m(x) = \int_a^x [F(t,\tilde{y}_m(t)) - F(t,\tilde{y}_{m-1}(t))]\, dt ,$$

και εκμεταλλευόμενοι το γεγονός ότι η απεικόνιση F ικανοποιεί μία συνθήκη Lipschitz με σταθερά L, βλέπουμε ότι:

$$\left|\tilde{y}_{m+1}(x)-\tilde{y}_m(x)\right| \le L\int_a^x \left|\tilde{y}_m(t)-\tilde{y}_{m-1}(t)\right| dt \qquad \forall x\in[a,\beta].$$

Εφαρμόζοντας τώρα την μέθοδο της μαθηματικής επαγωγής, όπως ακριβώς πράξαμε και κατά την απόδειξη της μοναδικότητας της λύσης του προβλήματος Cauchy, βρίσκουμε ότι:

$$\left|\tilde{y}_{m+1}(x)-\tilde{y}_m(x)\right| \le L^m \frac{(x-a)^m}{m!}\left\|\tilde{y}_1-\tilde{y}_0\right\| \quad \forall m\ge 0 \text{ και } x\in[a,\beta],$$

από όπου, λαμβάνοντας την μέγιστη τιμή του αριστερού μέλους, καθώς το *x* διατρέχει το κλειστό διάστημα $[a,\beta]$, συνεπάγεται ότι:

$$\left|\tilde{y}_{m+1}(x)-\tilde{y}_m(x)\right| \le L^m \frac{(\beta-a)^m}{m!}\left\|\tilde{y}_1-\tilde{y}_0\right\| \quad \forall m\ge 0.$$

Εάν, τώρα, p είναι ένας οποιοσδήποτε ακέραιος, μεγαλύτερος ή ίσος του 1, τότε:

$$\left\|\tilde{y}_{m+p}-\tilde{y}_m\right\| \le \sum_{j=0}^{p-1}\left\|\tilde{y}_{m+j+1}-\tilde{y}_{m+j}\right\| \le L^m \frac{(\beta-a)^m}{m!}\sum_{j=0}^{p-1} L^j \frac{(\beta-a)^j}{j!}\left\|\tilde{y}_1-\tilde{y}_0\right\|,$$

γιατί, για κάθε $m\ge 1$ και κάθε $j\ge 0$ έχουμε $(m+j)!\ge m!\,j!$. Επί πλέον:

$$\sum_{j=0}^{p-1} L^j \frac{(\beta-a)^j}{j!} \le e^{L(\beta-a)}.$$

Άρα, τελικά ισχύει:

$$\left\|\tilde{y}_{m+p}-\tilde{y}_m\right\| \le L^m \frac{(\beta-a)^m}{m!} e^{L(\beta a)} \left\|\tilde{y}_1-\tilde{y}_0\right\|,$$

και συνεπώς, για κάθε $\varepsilon > 0$ είναι δυνατή η εύρεση ενός ακεραίου αριθμού $M = M(\varepsilon)$ τέτοιου ώστε, εάν $m \ge M$ και $p \ge 1$, τότε:

$$\left\|\tilde{y}_{m+p}-\tilde{y}_m\right\| \le \left\|\tilde{y}_{m+p}-\tilde{y}_m\right\| \le \varepsilon .$$

Επομένως, η ακολουθία:

$$(\tilde{y}_m \in C[a,\beta] : m = 0,1,2,\ldots).$$

είναι μία ακολουθία Cauchy.

Επειδή ο διανυσματικός χώρος $C[a,\beta]$ είναι πλήρης, ως προς την νόρμα || || της ομοιόμορφης σύγκλισης, η ακολουθία αυτή συγκλίνει ομοιόμορφα επί του $[a,\beta]$, καθώς $m \to \infty$, προς μία συνάρτηση:

$$y \in C[a,\beta].$$

Οριακά μάλιστα ισχύει:

$$y(x) = y_0 + \int_a^x F(t, y(t))\, dt ,$$

που πιστοποιεί ότι *y* είναι λύση του προβλήματος Cauchy. Η *Απόδειξη* είναι πλήρης. ■

Στην πράξη είναι συχνά δύσκολο να αποφανθούμε εάν η συνθήκη Lipschitz ικανοποιείται. Για τον λόγο αυτό, παρουσιάζουμε το ακόλουθο κριτήριο:

Θεώρημα IV.4.1.3. ***Εάν η απεικόνιση F είναι διαφορίσιμη, με την έννοια του Fréchet, ως προς την εξαρτημένη μεταβλητή y, εάν το διαφορικό της F είναι συνεχές και εάν υπάρχει θετικός αριθμός $K > 0$ τέτοιος ώστε***

$$\left\| \frac{\partial f}{\partial y}(x, y) \right\| \le K \quad \forall x \in [a, \beta] \text{ και } \forall y \in \mathbb{R}^k,$$

τότε η F ικανοποιεί μία συνθήκη Lipschitz και το αντίστοιχο πρόβλημα Cauchy έχει μοναδική λύση. **(Το σύμβολο $\|\cdots\|$ αναπαριστά την νόρμα πίνακα.)**

Απόδειξη. Έπεται αμέσως από το γεγονός ότι, για κάθε $x \in [a, \beta]$ και για κάθε $y, z \in \mathbb{R}^k$, ισχύει η ισότητα:

$$F(x, y) - F(x, z) = \frac{\partial F}{\partial y}(x, \tilde{\xi})\,(y - x),$$

για κάποιο $\tilde{\xi} \in \mathbb{R}^k$. ■

Παρατήρηση IV.4.1.4. Τα δύο *Θεωρήματα*, που μόλις είδαμε, βεβαιώνουν την ύπαρξη και την μοναδικότητα μίας ολικής λύσης επί του $[a, \beta]$.

Εάν όμως η συνθήκη Lipschitz δεν ικανοποιείται παρά μόνον τοπικά, τότε το αντίστοιχο αποτέλεσμα δεν θα αφορά παρά μόνον την ύπαρξη και την μοναδικότητα μίας τοπικής λύσης του προβλήματος Cauchy.

Αξίζει πάντως να σημειωθεί πως δεν είναι δυνατόν να απαλλαγούμε τελείως από την υπόθεση της συνθήκης Lipschitz δίχως να διακινδυνεύσουμε την μοναδικότητα της λύσης.

Για του λόγου το αληθές, ας δώσουμε ένα παράδειγμα ([52]). Στο διαφορικό σύστημα:

$$y' = 3\, y^{2/3}$$

η συνθήκη Lipschitz δεν επαληθεύεται στην περιοχή του $y = 0$, γιατί:

$$\frac{|\, y^{2/3} - 0\,|}{|\, y - 0\,|} = \frac{1}{|\, y\,|^{1/3}},$$

που δεν μπορεί να παραμένει φραγμένο όταν η ποσότητα y προσεγγίζει απεριόριστα το 0.

Από την άλλη πλευρά, όπως είναι γνωστό, η γενική λύση της παραπάνω διαφορικής εξίσωσης είναι:

$$y(x) = (x - c)^3 \text{ ή } y(x) = 0.$$

Εάν όμως θελήσουμε να εντοπίσουμε την μερική λύση της εξίσωσης αυτής, που επαληθεύει την αρχική συνθήκη:

$$y(0) = 0,$$

τότε βλέπουμε ότι υπάρχει μία απειρία τέτοιων μερικών λύσεων και ειδικότερα, όλες οι συναρτήσεις:

$$y(x) = 0, \qquad y(x) = x^3, \qquad y(x) = \begin{cases} 0, & \text{εάν } x \le 0 \\ x^3, & \text{εάν } x > 0 \end{cases} \quad \text{κ.λ.π.}$$

είναι λύσεις του προβλήματος Cauchy:

$$\{\ y' = 3\, y^{2/3} \text{ και } y(0) = 0\ \}. \blacksquare$$

Στο *Κεφάλαιο* αυτό θα ενδιαφερθούμε για την αριθμητική επίλυση του προβλήματος Cauchy.

Θα υποθέτουμε πάντοτε ότι οι συνθήκες ύπαρξης και μοναδικότητας της λύσης ικανοποιούνται. Με άλλα λόγια, σε ό,τι ακολουθήσει θα δεχόμαστε, εκ των προτέρων, ότι η απεικόνιση *F* επαληθεύει μία συνθήκη Lipschitz.

Από την σκοπιά της Αριθμητικής Ανάλυσης, το ζήτημα της αναζήτησης της λύσης $y(x)$ του προβλήματος Cauchy υποκαθίσταται από εκείνο της αναζήτησης προσεγγίσεων y_n της τιμής $y(x_n)$ της λύσης σε ορισμένα σημεία x_n του διαστήματος $[a,\beta]$.

Στην συνέχεια, θα συμβολίζουμε με

$$y(x_n)$$

την ακριβή τιμή της λύσης στο σημείο x_n και με

$$y_n$$

την προσεγγιστική τιμή, στην οποία θα οδηγούμαστε μετά από εφαρμογή μίας από τις μεθόδους αριθμητικής ολοκλήρωσης που θα μελετήσουμε.

Οι μέθοδοι προσεγγιστικής ολοκλήρωσης διαφορικών συστημάτων, που θα συζητήσουμε, διαχωρίζονται σε δύο μεγάλες κατηγορίες:

- *τις **μεθόδους ξεχωριστών βημάτων***, κατά τις οποίες οι προσεγγιστικές τιμές y_{n+1} υπολογίζονται με την βοήθεια μόνον των τιμών y_n,
- *τις **μεθόδους συναπτών μεθόδων***, κατά τις οποίες οι προσεγγιστικές τιμές y_{n+1} υπολογίζονται με την βοήθεια των τιμών $y_n, y_{n-1}, \ldots, y_{n-k}$.

Όλες αυτές οι μέθοδοι διαχωρίζονται επίσης σε μεθόδους σαφείς και σε μεθόδους πεπλεγμένες, ανάλογα με το εάν οι των τιμές y_{n+1} δίνονται με τρόπο σαφή ή πεπλεγμένο, σαν λύσεις συστημάτων, γραμμικών ή όχι εξισώσεων, που απορρέουν από την γραμμικότητα ή όχι της F ως προς την εξαρτημένη μεταβλητή y.

IV.4.2. ΟΙ ΜΕΘΟΔΟΙ ΞΕΧΩΡΙΣΤΩΝ ΒΗΜΑΤΩΝ

Οι ***μέθοδοι ξεχωριστών βημάτων*** είναι της γενικής μορφής:

$$y_{n+1} = y_n + h_n \, \Phi(x_n, y_n, h_n), \tag{25}$$

όπου:

$n = 0,1,2,\ldots,$

y_0 : ***είναι η δεδομένη αρχική τιμή στο πρόβλημα Cauchy,***

$x_0 = a$,

$x_{n+1} = x_n + h_n$,

$(h_n : n = 0,1,2,\ldots)$ ***είναι η ακολουθία των βημάτων ολοκλήρωσης,***

και Φ ***είναι μία συνεχής απεικόνιση του*** $[a,\beta] \times \mathbb{R}^k \times [a, \beta - a]$ ***μέσα στο*** $\mathbb{R}^k$.

Όπως θα δούμε πιο κάτω, οι διάφορες μέθοδοι ξεχωριστών βημάτων διαφοροποιούνται μαζί με την αλλαγή επιλογής της απεικόνισης Φ.

Είναι προφανές ότι η Φ πρέπει να ανταποκρίνεται σε κάποιες θεωρητικές προδιαγραφές, προκειμένου οι τιμές y_n να είναι ουσιαστικές προσεγγίσεις των $y(x_n)$.

Αυτές οι προδιαγραφές απαρτίζουν τις θεωρητικές έννοιες που θα εξετάσουμε αμέσως τώρα.

Για απλοποίηση της παρουσίασης, θα περιοριστούμε στην περίπτωση $k = 1$, επειδή η γενίκευση στην πολυδιάστατη τοποθέτηση όπου $k > 1$, είναι άμεση και δεν εμφανίζει καμιά θεωρητική, τεχνική ή πρακτική δυσκολία.

Επίσης, θα δεχόμαστε ανελλιπώς ότι το βήμα ολοκλήρωσης είναι σταθερό δηλαδή ότι:

$$h_n = h = \frac{\beta - a}{N} > 0, \quad \forall n = 0,1,2,\ldots,N,$$

για κάποιο φυσικό αριθμό N.

IV.4.2.(i). Θεωρητικές Έννοιες Οι τρεις σημαντικές έννοιες είναι η συνάφεια, η θεωρητική ευστάθεια και η σύγκλιση.

Η ***συνάφεια*** εκφράζεται από το γεγονός ότι η απεικόνιση *Φ* πρέπει να συνδέεται λειτουργικά με την προς ολοκλήρωση διαφορική εξίσωση.

Η ***ευστάθεια*** εξετάζεται από το αποτέλεσμα που μπορούν να επιφέρουν κάποιες μικρές διαταραχές των προσεγγιστικών τιμών. Με άλλα λόγια, πρόκειται για την μελέτη της διάδοσης των σφαλμάτων.

Η ***σύγκλιση*** εκφράζεται από το ποιοτικό γεγονός ότι οι προσεγγιστικές λύσεις πρέπει να πλησιάζουν απεριόριστα την ακριβή λύση σε κάθε σημείο του $[a,\beta]$, όταν το βήμα *h* τείνει προς το μηδέν. Η έννοια της τάξης προσδιορίζει το ποσοτικό περιεχόμενο αυτής της σύγκλισης.

IV.4.2.(ii). Συνάφεια Μεθόδου Η μέθοδος ξεχωριστών βημάτων (25) λέγεται ***συναφής προς την διαφορική εξίσωση***, εάν:

$$\lim_{h\to 0} \max_{0\le n\le N} \left| \frac{1}{h}[y(x_{n+1}) - y(x_n)] - \Phi(x_n, y(x_n), h) \right| = 0,$$

για κάθε λύση $y \in C[a,\beta]$ της διαφορικής εξίσωσης.

Ισχύει το ακόλουθο θεμελιώδες αποτέλεσμα:

Θεώρημα IV.4.2.(ii).1. ***Μία συνθήκη αναγκαία και ικανή προκειμένου η μέθοδος ξεχωριστών βημάτων*** (25) ***να είναι συναφής προς την διαφορική εξίσωση είναι:***

$$\Phi(\mathrm{x},\mathrm{y},0) = F(x,y) \quad \forall x \in [a,\beta] \text{ και } \forall y \in \mathbb{R}.$$

Απόδειξη. Θα δείξουμε πρώτα ότι η συνθήκη είναι αναγκαία.

Εάν η μέθοδος είναι συναφής, τότε έχουμε:

$$\lim_{h\to 0} \max_{0\le n\le N} \left| \frac{1}{h}[y(x_{n+1}) - y(x_n)] - \Phi(x_n, y(x_n), h) \right| = 0,$$

ή ακόμα:

$$\lim_{h\to 0} \max_{0\le n\le N} \left| \frac{1}{h}\int_{x_n}^{x_{n+1}} f(t, y(t))\, dt - \Phi(x_n, y(x_n), h) \right| = 0 \quad .$$

Έστω x ένα τυχαίο σημείο του $[a,\beta]$ και έστω $(x_n : n = 0,1,2,\ldots)$ μία ακολουθία σημείων του $[a,\beta]$, με τον δείκτη n να εξαρτάται από το h, και τέτοια ώστε:

$$\lim_{h\to 0} x_n = x.$$

Σύμφωνα με την προηγούμενη ισότητα, αλλά και λόγω της συνέχειας των F και Φ, θα είναι:

$$\lim_{h\to 0} \left| \frac{1}{h}\int_{x_n}^{x_{n+1}} F(t, y(t))dt - \Phi(x_n, y(x_n), h) \right| = F(x, y(x)) - \Phi(x, y(x), 0) = 0.$$

Από τη άλλη πλευρά, δοθέντων δύο σημείων $x \in [a,\beta]$ και $z \in \mathbb{R}$, υπάρχει πάντοτε μία λύση $y \in C[a,\beta]$ της διαφορικής εξίσωσης που επαληθεύει την συνθήκη $y(x) = z$.

Άρα, μέσα στην προηγούμενη ισότητα, το $y(x)$ διατρέχει όλο το $\mathbb{R}$ όταν το *y* διατρέχει το σύνολο των λύσεων της διαφορικής εξίσωσης. Αυτό αποδεικνύει ότι η σχέση:

$$\Phi(x, y, 0) = F(x, y) \quad \forall x \in [a, \beta] \text{ και } \forall y \in \mathbb{R}$$

είναι πράγματι μία αναγκαία συνθήκη της συνάφειας.

Θα δείξουμε τώρα ότι αυτή η συνθήκη είναι, επί πλέον, και ικανή.

Προς τούτο, παρατηρούμε ότι η συνθήκη:

$$\Phi(t, y(t), 0) = F(t, y(t)) \quad \forall t \in [a, \beta]$$

εγγυάται την ισχύ της ισότητας:

$$\frac{1}{h}[y(x_{n+1}) - y(x_n)] - \Phi(x_n, y(x_n), h) = \frac{1}{h}\int_{x_n}^{x_{n+1}} [\Phi(t, y(t), 0) - \Phi(x_n, y(x_n), h)]dt.$$

Επειδή όμως η ομοιόμορφη συνέχεια της Φ επί του συμπαγούς συνόλου:

$$\{(x, y(x)) : x \in [a, \beta]\} \times [0, \beta - a]$$

εξασφαλίζει ότι:

$$\lim_{h\to 0} \max_{0\le n\le N} \left|\Phi(t, y(t), 0) - \Phi(x_n, y(x_n), h)\right| = 0 \quad \forall t \in [x_n, x_{n+1}],$$

έπεται ότι:

$$\lim_{h\to 0} \max_{0\le n\le N} \frac{1}{h}\int_{x_n}^{x_{n+1}} \left[\Phi(t, y(t), 0) - \Phi(x_n, y(x_n), h)\right]dt = 0,$$

που αποδεικνύει ότι η συνθήκη του *Θεωρήματος* είναι ικανή. ■

IV.4.2.(iii).Θεωρητική Ευστάθεια Μεθόδου Έστωσαν y_n και z_n $(n = 0,1,\ldots,N)$ οι αντίστοιχες λύσεις των προβλημάτων:

$$\begin{cases} y_{n+1} = y_n + h\,\Phi(x_n, y_n, h) \\ y_0 \in \mathbb{R}, \text{ τυχαίο και δεδομένο} \end{cases}$$

και

$$\begin{cases} z_{n+1} = z_n + h\,[\Phi(x_n, z_n, h) + \varepsilon_n] \\ z_0 \in \mathbb{R}, \text{ τυχαίο και δεδομένο} \end{cases}$$

με:

$$x_0 = a \text{ και } x_{n+1} = x_n + h.$$

Λέμε ότι η μέθοδος (25) είναι ***ευσταθής***, εάν υπάρχουν δύο θετικές σταθερές M_1 και M_2 που είναι ανεξάρτητες του *h* και τέτοιες ώστε:

$$\max_{0\le n\le N} |\, y_n - z_n \,| \le M_1 |\, y_0 - z_0 \,| + M_2 \max_{0\le n\le N} |\, \varepsilon_n \,|.$$

Αυτή η έννοια της ευστάθειας σημαίνει πως μία μικρή διαταραχή επί της τιμής του y_0 και επί της επιλογής της απεικόνισης Φ δεν επιφέρει παρά μόνον μικρή διαταραχή επί όλων των τιμών y_n, και ακόμη πως μία τέτοια διαπίστωση δεν επηρεάζεται από το μέγεθος του *h*.

Θα αποκαλούμε ***θεωρητική ευστάθεια*** την έννοια ευστάθειας που μελετάται σε αυτήν την *Παράγραφο*, σε αντίθεση με την έννοια της *αριθμητικής ευστάθειας*, που θα συναντήσουμε αργότερα.

Ισχύει το επόμενο βασικό αποτέλεσμα:

Θεώρημα IV.4.2.(iii).1. ***Εάν η απεικόνιση Φ ικανοποιεί μία συνθήκη Lipschitz, για κάθε h αρκετά μικρό, δηλαδή εάν:***

$$\left|\Phi(x,y,h)-\Phi(x,z,h)\right| \le M\,|\,y-z\,|,\quad \forall x\in[a,\beta],\forall y,z\in\mathbb{R},\forall h<\varepsilon,$$

όπου M είναι μία σταθερά ανεξάρτητη του h, τότε η μέθοδος ξεχωριστών βημάτων (25) είναι ευσταθής.

Απόδειξη. Επειδή:

$$|\,y_{n+1}-z_{n+1}\,|\le|\,y_n-z_n\,|+h\,|\,\Phi(x_n,y_n,h)-\Phi(x_n,z_n,h)\,|+h\,|\,\varepsilon_n\,|,$$

μπορούμε να δείξουμε, χρησιμοποιώντας μαθηματική επαγωγή, ότι:

$$|\,y_{n+1}-z_{n+1}\,|\le(1+hM)^{n+1}\,|\,y_0-z_0\,|+\frac{(1+hM)^{n+1}-1}{M}\ \max_{0\le n\le N}|\,\varepsilon_n\,|.$$

(Πράγματι, είναι:

$$|\,y_{n+2}-z_{n+2}\,|\le(1+hM)\,|\,y_{n+1}-z_{n+1}\,|+h|\,\varepsilon_n\,|,$$

απ' όπου έπεται (εφόσον το h είναι αρκετά μικρό, έτσι ώστε: $h\,M\le 1$) ότι:

$$|\,y_{n+2}-z_{n+2}\,|\le(1+hM)^{n+2}\,|\,y_0-z_0\,|$$
$$+\frac{(1+hM)^{n+2}}{M}\max_{0\le n\le N}|\,\varepsilon_n\,|+\frac{hM\,|\,\varepsilon_{n+1}\,|-hM\max_{0\le n\le N}|\,\varepsilon_n\,|-1}{M}\Rightarrow$$
$$|\,y_{n+2}-z_{n+2}\,|\le(1+hM)^{n+2}\,|\,y_0-z_0\,|+\frac{(1+hM)^{n+2}-1}{M}\max_{0\le n\le N}|\,\varepsilon_n\,|.)$$

Καθώς όμως:

$$(1+hM)^{n+1}\le e^{(n+1)hM}\le e^{(\beta-a)M},$$

συμπεραίνουμε πως:

$$|\,y_{n+1}-z_{n+1}\,|\le e^{(\beta-a)M}\,|\,y_0-z_0\,|+\frac{e^{(\beta-a)M}-1}{M}\max_{0\le n\le N}|\,\varepsilon_n\,|.$$

Η ανισότητα αυτή αποδεικνύει την ευστάθεια της μεθόδου, εφόσον πάρουμε:

$$M_1=e^{(\beta-a)M}\qquad\text{και}\qquad M_2=\frac{e^{(\beta-a)M}-1}{M}.\ \blacksquare$$

IV.4.2.(iv). Σύγκλιση Μεθόδου Λέμε ότι η μέθοδος ξεχωριστών βημάτων (25) είναι ***συγκλίνουσα*** (ή απλά ότι ***συγκλίνει***), εάν:

$$\lim_{h\to 0} \max_{0\le n\le N} | y_n - y(x_n) | = 0 .$$

Σχετικά με τη σύγκλιση της μεθόδου (25), έχουμε το εξής αποτέλεσμα:

Θεώρημα IV.4.2.(iv).1. ***Η μέθοδος ξεχωριστών βημάτων (25) είναι συγκλίνουσα, εάν είναι ευσταθής και συναφής προς την διαφορική εξίσωση.***

Απόδειξη. Η λύση y της διαφορικής εξίσωσης ικανοποιεί την σχέση:

$$y(x_{n+1}) = y(x_n) + h\,[\Phi(x_n, y(x_n), h) + \varepsilon_n], \text{ με } y(a) = y_0 .$$

Επειδή η μέθοδος είναι συναφής προς την διαφορική εξίσωση, ισχύει:

$$\lim_{h\to 0} \max_{0\le n\le N} | \varepsilon_n | = 0 ,$$

ενώ, επειδή η μέθοδος είναι ευσταθής, υπάρχει μία σταθερά M_2, που είναι ανεξάρτητη από την επιλογή του βήματος h και τέτοια ώστε:

$$\max_{0\le n\le N} | y_n - y(x_n) | \le M_2 \max_{0\le n\le N} | \varepsilon_n | .$$

Η σύγκλιση της μεθόδου (25) έπεται αμέσως, εάν αφήσουμε το h να πλησιάσει απεριόριστα προς το μηδέν. ■

IV.4.2.(ν). Τάξη Μεθόδου Η τάξη μίας μεθόδου ξεχωριστών βημάτων προσδιορίζει τον τρόπο και τον ρυθμό με τον οποίο η ποσότητα

$$\max_{0\le n\le N} | y_n - y(x_n) |$$

τείνει προς το μηδέν καθώς το h μικραίνει ταυτόχρονα απεριόριστα.

Λέμε ότι ***η μέθοδος ξεχωριστών βημάτων*** (25) ***είναι τάξης p***, εάν:

$$\max_{0\le n\le N}\left|\frac{1}{h}[y(x_{n+1})-y(x_n)]-\Phi(x_n,y(x_n),h)\right| = \mathrm{O}(h^P),$$

όπου το σύμβολο $\mathrm{O}(h^P)$ αναπαριστά μία συνάρτηση του h που είναι τέτοια ώστε:

$$\left|\frac{\mathrm{O}(h^P)}{h^P}\right| \le C \text{, καθώς } h\to 0.$$

Παρατηρήστε ότι η έννοια της τάξης της μεθόδου παρέχει μία ποσοτική πληροφορία σχετικά με την προσέγγιση της F από την Φ, καθώς το h ελαττώνεται συνεχώς.

Ισχύει το επόμενο θεμελιώδες αποτέλεσμα:

Θεώρημα IV.4.2.(ν).1. ***Εάν η μέθοδος*** (25) ***είναι ευσταθής και τάξης*** p ***, τότε:***

$$\max_{0\le n\le N} | y_n - y(x_n) | = \mathrm{O}(h^P).$$

Απόδειξη. Επειδή:

$$y(x_{n+1})-y(x_n)=\int_{x_n}^{x_{n+1}} F(t,y(t))\,dt,$$

κι επειδή, από την άλλη πλευρά:

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

αφαιρώντας κατά μέλη, και θέτοντας $\sigma_n := y_n - y(x_n)$, διαπιστώνουμε ότι:

$$\sigma_{n+1} = \sigma_n + h\,[\Phi(x_n, y_n, h) - \frac{1}{h}\int_{x_n}^{x_{n+1}} F(t, y(t))dt]$$

ή ακόμα ότι:

$$\sigma_{n+1} = \sigma_n + h\,[\Phi(x_n, y_n, h) - \Phi(x_n, y(x_n), h)]$$
$$+ h\,[\Phi(x_n, y(x_n), h) - \frac{1}{h}\int_{x_n}^{x_{n+1}} F(t, y(t))dt].$$

Χρησιμοποιώντας το γεγονός ότι η μέθοδος (25) είναι ευσταθής, για να φράξουμε το περιεχόμενο της πρώτης αγκύλης, αλλά και τον ορισμό της τάξης της μεθόδου, προκειμένου να φράξουμε το περιεχόμενο της δεύτερης αγκύλης, βλέπουμε ότι:

$$|\sigma_{n+1}| \le |\sigma_n| + h\,M\,|\sigma_n| + K\,h^{p+1} = (1 + h\,M)|\sigma_n| + K\,h^{p+1}.$$

Από την σχέση αυτή, έπεται, με μαθηματική επαγωγή, ότι:

$$|\sigma_n| \le \frac{K}{M}[\,(1 + hM)^{n+1} - 1]\,h^p.$$

(Πράγματι, εκμεταλλευόμενοι την επαγωγική υπόθεση, βρίσκουμε διαδοχικά:

$$|\sigma_{n+1}| \le (1 + h\,M)\,\frac{K}{M}\,[(1 + h\,M)^{n+1} - 1]\,h^p + K\,h^{p+1}$$
$$\Rightarrow$$
$$|\sigma_{n+1}| \le \frac{K}{M}(1 + h\,M)^{n+2}\,h^p - \frac{K}{M}h^p - K\,h^{p+1} + K\,h^{p+1}$$
$$\Rightarrow$$
$$|\sigma_{n+1}| \le \frac{K}{M}[(1 + h\,M)^{n+2} - 1]h^p.)$$

Καθώς όμως, για κάθε *h* αρκετά μικρό, ισχύει:

$$(1 + h\,M)^{n+1} \le e^{(n+1)hM} \le e^{(\beta - a)M},$$

η τελευταία ανισότητα συνεπάγεται ότι:

$$|\sigma_n| \le \frac{K}{M}\,[e^{(\beta - a)M} - 1]\,h^p = \mathrm{O}(h^p),$$

που ολοκληρώνει την *Απόδειξη* του *Θεωρήματος*. ■

Το σφάλμα $\sigma_n = y_n - y_n(x_n)$, που μελετήσαμε σ' αυτήν την *Παράγραφο*, είναι το ολικό σφάλμα της μεθόδου στο σημείο x_n.

Έτσι, στο προηγούμενο *Θεώρημα*, δείξαμε ότι, στο σημείο x_n, το ολικό σφάλμα είναι μία συνάρτηση του h, που τείνει πιο γρήγορα προς το 0 (καθώς το $h \to 0$), απ' ότι η συνάρτηση h^p, δηλαδή το ολικό σφάλμα σ_n είναι της μορφής $\mathrm{O}(h^p)$. Αυτό το ολικό σφάλμα είναι το άθροισμα των διαδοχικών σφαλμάτων, που προηγήθηκαν σε κάθε βήμα της μεθόδου ολοκλήρωσης. Αυτό το *κατά βήμα σφάλμα* είναι της μορφής $\mathrm{O}(h^{p+1})$.

Πράγματι, όπως είδαμε για τα ολικά σφάλματα ισχύει:

$$|\sigma_{n+1}| \le |\sigma_n| + h\, M\, |\sigma_n| + K\, h^{p+1}.$$

Εάν υποθέσουμε πως δεν υπάρχει σφάλμα μέχρι το σημείο x_n, δηλαδή εάν υποθέσουμε πως $\sigma_n = 0$, τότε βλέπουμε αμέσως ότι το κατά βήμα σφάλμα της μεθόδου στο σημείο x_{n+1} δεν μπορεί να υπερβεί την ποσότητα $K\, h^{p+1}$ και άρα θα είναι της μορφής $\mathrm{O}(h^{p+1})$.

IV.4.2.(vi). Οι Μέθοδοι Euler και Runge-Kutta

Η ***μέθοδος Euler*** είναι ιδιαίτερα απλή, γιατί συνίσταται στην λήψη:

$$\Phi(x, y, h) = F(x, y).$$

Έτσι, σύμφωνα μ' αυτήν την επιλογή, η μέθοδος αυτή περιγράφεται από την ακόλουθη σχέση:

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

(με την τιμή $y_0 = y(x_0) = y(a)$ δεδομένη).

Βλέπουμε αμέσως ότι, σύμφωνα με το *Θεώρημα IV.4.2.(ii).1*, ***η μέθοδος Euler είναι συναφής προς την διαφορική εξίσωση.***

Επίσης, ***είναι ευσταθής***, επειδή έχουμε υποθέσει από την αρχή του Κεφαλαίου ότι η απεικόνιση *F* ικανοποιεί μία συνθήκη Lipschitz. Άρα σύμφωνα με το *Θεώρημα IV.4.2.(iv).1*, ***η μέθοδος Euler είναι, επί πλέον, και συγκλίνουσα.***

Ας εξετάσουμε την τάξη της. Αφού:

$$\frac{y(x+h)-y(x)}{h} - \Phi(x, y(x)h) = \frac{y(x+y)-y(x)}{h} - F(x, y(x)),$$

κι επειδή:

$$y(x+h) = y(x) + h\, F(x, y(x)) + h\, \mathrm{O(h)},$$

είναι:

$$\frac{1}{h}[(y(x+h)-y(x)] - \Phi(x, y(x), h) = \mathrm{O(h)},$$

και επομένως ***η τάξη της μεθόδου Euler είναι ίση με*** 1.

Οι ***μέθοδοι Runge-Kutta*** μπορούν να θεωρούνται σαν γενικεύσεις της μεθόδου Euler. Η κεντρική ιδέα πάνω στην οποία στηρίζονται συνίσταται στην θεώρηση ενδιάμεσων σημείων μεταξύ των x_n και x_{n+1}. Έτσι, θέτουμε:

$$u_1 = u_1(x, y, h) = F(x, y),$$

$$u_2 = u_2(x, y, h) = F(x + h\,\theta_2, y + h\, a_{2,1}\, u_1),$$

$$u_3 = u_3(x, y, h) = F(x + h\,\theta_3, y + h\, a_{3,1}\, u_1 + h\, a_{3,2}\, u_2),$$

……

$$u_r = u_r(x,y,h) = F(x + h\,\theta_r, y + h\sum_{j=1}^{r-1} a_{r,j}\,u_j),$$

$$\Phi(x,y,h) = \sum_{i=1}^{r} c_i\,u_i\,.$$

Ο αριθμός r ονομάζεται ***βαθμός*** της μεθόδου Runge-Kutta. Είναι προφανές ότι η μέθοδος Euler αντιστοιχεί στην ειδική περίπτωση που:

$$r = 1 \text{ και } c_1 = 1\,.$$

Ας εξετάσουμε τώρα τις θεωρητικές έννοιες που προσάπτονται μίας μεθόδου Runge-Kutta. Κατ' αρχάς, επειδή:

$$\Phi(x,y,0) = F(x,y) = \sum_{i=1}^{r} c_i\,,$$

μία ικανή και αναγκαία συνθήκη προκειμένου μία μέθοδος Runge-Kutta να είναι συναφής προς την διαφορική εξίσωση είναι η ισχύς της ισότητας:

$$\sum_{i=1}^{r} c_i = 1\,.$$

Ακόμη, ***οι μέθοδοι Runge-Kutta είναι ευσταθείς.*** Πράγματι, έχουμε:

$$|u_1(x,y,h) - u_1(x,z,h)| = |F(x,y) - F(x,z)| \le L\,|y-z| \;\; \forall x \in [a,\beta] \text{ και } y,z \in \mathbb{R}$$

και ομοίως:

$$|u_2(x,y,h) - u_2(x,z,h)|$$

$$= |F(x + h\,\theta_2, y + h\,a_{2,1}\,F(x,y)) - F(x + h\,\theta_2, z + h\,a_{2,1}\,F(x,z))|$$

$$\le L\,|y + h\,a_{2,1}\,F(x,y) - z - h\,a_{2,1}\,F(x,z)|$$

$$\le L\,|y-z| + h\,L\,|a_{2,1}|\,|F(x,y) - F(x,z)|$$

$$\le [L\,(1 + h\,|a_{2,1}|\,L)]\;|y-z|,\; \forall x \in [a,\beta] \text{ και } \forall y,z \in \mathbb{R}\,,$$

και ούτω καθ' εξής για τα υπόλοιπα u_i.

Συνδυάζοντας όλες τις ανισότητες που θα προκύψουν, καταλήγουμε στο συμπέρασμα ότι η απεικόνιση:

$$\Phi(x, y, h) = \sum_{i=1}^{r} c_i \; u_i$$

πληροί τις προϋποθέσεις του *Θεωρήματος IV.4.2.(iii).1*, και συνεπώς κάθε μέθοδος Runge-Kutta είναι ευσταθής.

Μετά απ' όλα όσα προηγήθηκαν, και έχοντας υπ' όψη το *Θεώρημα IV.4.2.(iv).1*, μπορούμε να αποφανθούμε ότι ***η μέθοδος Runge-Kutta είναι συγκλίνουσα, όταν:***

$$\sum_{i=1}^{r} c_i = 1 .$$

Η μελέτη της τάξης των μεθόδων Runge-Kutta είναι, γενικά, πιο δύσκολη, συγκριτικά με την μελέτη των άλλων θεωρητικών εννοιών.

Στην πραγματικότητα, οι συντελεστές:

$$a_{i,j}, \theta_i \text{ και } c_i$$

επιλέγονται κατά τέτοιον τρόπο ώστε η τάξη της μεθόδου να είναι όσο το δυνατόν πιο μεγάλη.

Αυτό επιτυγχάνεται χρησιμοποιώντας αναπτύγματα σε σειρές Taylor και, φυσικά, τον ορισμό της τάξης που δόθηκε στην *Υποπαράγραφο IV.4.2.(v)*, και αφού καταλήξουμε σε έναν ορισμένο αριθμό σχέσεων που αυτοί οι συντελεστές πρέπει να ικανοποιούν. Αυτές οι σχέσεις δεν είναι γραμμικές και επί πλέον υστερούν, γενικά, σε πλήθος έναντι του συνολικού αριθμού των συντελεστών. Το γεγονός τούτο επιτρέπει την αυθαίρετη επιλογή κάποιων συντελεστών, με εύλογο παρεπόμενο την ύπαρξη απειρίας μεθόδων Runge-Kutta που είναι ίδιας τάξης.

Για να γίνουν πιο κατανοητά τα όσα προηγήθηκαν, περί της τάξης των μεθόδων Runge-Kutta, ας εξετάσουμε ενδελεχώς τις μεθόδους που έχουν βαθμό ίσο με 2. Θεωρώντας το (περιορισμένο) ανάπτυγμα σε σειρά Taylor της $y(x+h)$, λαμβάνουμε την σχέση:

$$\frac{1}{h}.[y(x+h)-(x)] = y'(x) + \frac{h}{2}\, y''(x) + \mathrm{O}(h^2)\,.$$

Καθώς όμως:

$$y''(x) = \frac{d}{dx} F(x, y(x)) = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\,\frac{dy}{dx} = F_x' + F\; F_y'\,,$$

η προηγούμενη γίνεται:

$$\frac{1}{h}\,[y(x+h)-y(x)] = F + \frac{h}{2}\,(F_x' + F\; F_y') + \mathrm{O}(h^2)\,.$$

Από την άλλη πλευρά επειδή:

$$\Phi(x, y(x,h) = c_1\; F(x,y) + c_2\; F(x + h\,\theta_2, y + h\,a_{2,1}\; F(x,y))\,,$$

θα πρέπει να γράψουμε το ανάπτυγμα σε σειρά Taylor μίας συνάρτησης δύο μεταβλητών.

Θυμίζουμε την έκφραση της σειράς Taylor μίας συνάρτησης $g(x,y)$ δύο πραγματικών μεταβλητών x και y:

$$g(x+s, y+t) = \sum_{k\geq 0} \frac{1}{k!}\,(s\,\frac{\partial}{\partial x} + t\,\frac{\partial}{\partial y})^k\, g(x,y)$$

$$= g(x,y) + s\,\frac{\partial}{\partial x} g(x,y) + t\,\frac{\partial}{\partial y} g(x,y)$$

$$+ \frac{s^2}{2}\,\frac{\partial^2}{\partial x^2} g(x,y) + s\,t\,\frac{\partial^2}{\partial x \partial y} g(x,y) + \frac{t^2}{2}\,\frac{\partial^2}{\partial y^2} g(x,y) + \ldots$$

Στην περίπτωσή μας, θα έχουμε:

$$F(x+h\,\theta_2, y+h\,a_{2,1}\,F) = F + h\,\theta_2\,F'_x + h\,a_{2,1}\,F'_y + \mathrm{O}(h^2),$$

και επομένως:

$$\Phi(x, y(x), h) = (c_1 + c_2)\,F + c_2\,h\,\theta_2\,F'_x + c_2\,h\,a_{2,1}\,F\,F'_y + \mathrm{O}(h^2).$$

Εξισώνοντας την τελευταία με το (περιορισμένο) ανάπτυγμα Taylor της συνάρτησης:

$$\frac{1}{h}[y(x+h) - y(x)],$$

βρίσκουμε ότι πρέπει να ισχύει:

$$c_1 + c_2 = 1,$$

$$c_2\,\theta_2 = \frac{1}{2},$$

$$c_2\,a_{2,1} = \frac{1}{2}.$$

Δεν είναι δυνατόν να σχηματίσουμε περισσότερες τέτοιες σχέσεις και, συνεπώς, να υπερβούμε την τάξη δύο. Αυτό είναι, εξ άλλου, ένα χαρακτηριστικό γνώρισμα των μεθόδων Runge-Kutta: ***Η τάξη δεν μπορεί να είναι πιο μεγάλη από τον βαθμό***. Εάν επανέλθουμε στο παραπάνω σύστημα, βλέπουμε ότι πρόκειται για μία παράθεση τριών εξισώσεων με τέσσερις αγνώστους.

Έτσι, λαμβάνοντας αυθαίρετα:

$$\theta_2 = \frac{1}{2},$$

συμπεραίνουμε ότι πρέπει:

$$c_1 = \frac{1}{2},$$

$$c_2 = a_{2,1} = 1,$$

και, κατ' ακολουθία, αυτή η μέθοδος Runge-Kutta θα είναι της μορφής:

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

Η συγκεκριμένη μέθοδος είναι γνωστή με το όνομα ***βελτιωμένη μέθοδος Euler***.

Για τις περιπτώσεις των μεθόδων Runge-Kutta που έχουν βαθμούς ίσους με 3 και 4, είναι πάλι δυνατή η επίλυση δύο συστημάτων εξισώσεων, (που όμως δεν είναι γραμμικές), εξισώνοντας τους κατάλληλους όρους στα αναπτύγματα Taylor μέχρι εκείνους που είναι της μορφής

$O(h^3)$ και $O(h^4)$, αντίστοιχα.

Όμως, αποδεικνύεται επίσης ότι είναι αδύνατη η κατασκευή μίας μεθόδου Runge-Kutta βαθμού ίσου με 5 και τάξης 5 ($[8]$).

Προκειμένου να επιτύχουμε την τάξη 5, θα πρέπει να ανέλθουμε σε μέθοδο βαθμού ίσου με 6. Αυτός είναι ο λόγος, που εξηγεί γιατί οι μέθοδοι Runge-Kutta τάξης και βαθμού 4 είναι οι πιο διαδεδομένες στις διάφορες εφαρμογές.

Ο κλασσικός μαθηματικός *Τύπος* μίας μεθόδου Runge-Kutta τάξης 4 σχηματίζεται από τις παρακάτω σχέσεις:

$$u_1 = F(x, y),$$

$$u_2 = F(x + \frac{h}{2}, y + h\,\frac{u_1}{2}),$$

$$u_3 = F(x + \frac{h}{2}, y + h\,\frac{u_2}{2}),$$

$$u_4 = F(x, y, y + h\,u_3),$$

και:

$$\Phi(x, y, h) = \frac{1}{6}\,(u_1 + 2\,u_2 + 2\,u_3 + u_4),$$

και, κατ' ακολουθία, η αντίστοιχη πλήρης έκφραση αυτού του μαθηματικού *Τύπου* είναι της μορφής:

$$y_{n+1} = y_n + h\,\Phi(x_n, y_n, h) = y_n + \frac{h}{6}\,(u_1 + 2\,u_2 + 2\,u_3 + u_4)\,.$$

Σαν μία τελική παρατήρηση σχετικά με την πρώτη γνωριμία μας με τις μεθόδους Runge-Kutta, αξίζει να επισημάνουμε την ύπαρξη κριτηρίων βελτιωμένης επιλογής των συντελεστών, εκείνων στους οποίους είχαν δοθεί αυθαίρετες τιμές.

Τα κίνητρα για την θεώρηση τέτοιων κριτηρίων οριοθετούνται από την αναγκαιότητα απλοποίησης ή ελαχιστοποίησης διαφόρων παραγόντων που υπεισέρχονται στις μαθηματικές διατυπώσεις του σφάλματος.

IV.4.2.(vii). Πρακτικές Έννοιες Όταν ολοκληρώνουμε (δηλαδή, όταν επιλύουμε) ένα διαφορικό σύστημα με την βοήθεια των μεθόδων Runge-Kutta, εμφανίζονται δύο στοιχειώδη ερωτήματα, τα οποία είναι, από την φύση τους, ανεξάρτητα μεταξύ τους:

- *Το πρώτο πρόβλημα, που εγείρεται, αφορά το μέγεθος του βήματος* h. **Η επιλογή του μεγέθους εξαρτάται από την αριθμητική ευστάθεια (που ονομάζεται επίσης και** $A-$ ***ευστάθεια*****) της μεθόδου ολοκλήρωσης που χρησιμοποιούμε.**
- **Το δεύτερο ζήτημα, που παρουσιάζεται, συνίσταται στην δυνατότητα ελέγχου της διάδοσης των σφαλμάτων καθ' όλη την διάρκεια της εξέλιξης της αριθμητικής ολοκλήρωσης, έτσι ώστε να υπάρχει η ευχέρεια ελάττωσης του μεγέθους του βήματος, εάν το σφάλμα είναι σημαντικό ή η πολυτέλεια αύξησης του μεγέθους του βήματος, όταν η ακρίβεια του αποτελέσματος υπερβαίνει το σφάλμα.** *Το δεύτερο πρόβλημα, λοιπόν, αποσκοπεί στον έλεγχο (του μεγέθους) του βήματος.*

IV.4.2.(viii). A- Ευστάθεια των Μεθόδων Runge-Kutta. Η έννοια της ***αριθμητικής ευστάθειας*** ενός σχήματος ολοκλήρωσης διαφορικού συστήματος θεωρήθηκε για πρώτη φορά το 1956 από τον *Dahlquist*, ο οποίος και της απέδωσε την ονομασία $A-$ ***ευστάθεια.***

Σύμφωνα με τον Dahlquist, μία μέθοδος αριθμητικής ολοκλήρωσης (δηλαδή μία μέθοδος επίλυσης διαφορικών συστημάτων), καλείται $A-$ ***ευσταθής***, εάν:

$$\lim_{n\to\infty} y_n = 0 \ \forall h,$$

για κάθε πρόβλημα Cauchy που είναι της μορφής:

$$y' = -A\, y$$

$$y(0) = y_0,$$

όπου A είναι μία μιγαδική σταθερά της οποίας το πραγματικό μέρος είναι θετικό. Υπενθυμίζουμε ότι η λύση ενός τέτοιου διαφορικού προβλήματος είναι:

$$y(x) = y_0 \; e^{-A\,x}$$

και κατά συνέπεια η ασυμπτωτική συμπεριφορά αυτής της λύσης δίνεται από την σχέση:

$$\lim_{x\to\infty} y(x) = 0 \,.$$

Γενικά, η συνθήκη $\lim_{n\to\infty} y_n = 0$, του ορισμού της $A-$ευστάθειας, δεν επαληθεύεται για κάθε τιμή του βήματος h, αλλά μόνον όταν η ποσότητα $A\,h$ ανήκει σε κάποιον ορισμένο τόπο του μιγαδικού επιπέδου. Αυτός ο τόπος ονομάζεται ***τόπος αριθμητικής ευστάθειας της μεθόδου***.

Είναι λοιπόν δικαιολογημένη η επιβολή περιοριστικών όρων στην επιλογή του βήματος h, το οποίο δεν πρέπει να θεωρείται αυθαίρετα ορισμένο.

Ο Dahlquist έδειξε, εξ' άλλου, ότι μία μέθοδος $A-$ευσταθής είναι αναγκαστικά πεπλεγμένη.

Ας εξετάσουμε τώρα λεπτομερειακά την ιδιότητα της $A-$ευστάθειας. Προς τούτο, ας υποθέσουμε δεδομένη την παρακάτω μέθοδο ξεχωριστών βημάτων:

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

Έστω ότι η τάξη αυτής της μεθόδου είναι ίση με p. Εάν εφαρμόσουμε το σχήμα της μεθόδου, στην διαφορική εξίσωση $y' = -A\,y$, θα διαπιστώσουμε πως:

$$y_{n+1} = P(A\,h)\,y_n\,,$$

ή ακόμα πως:

$$y_{n+1} = [P\,(A\,h)]^{n+1}\;y_0\,,$$

όπου P είναι ένα πολυώνυμο βαθμού p ως προς την μεταβλητή $(A\,h)$. Ας θεωρήσουμε την συνάρτηση:

$$\tilde{y}(x) = y_n \; e^{-A(x-x_n)}.$$

Είναι σαφές ότι η διαφορά $\tilde{y}(x_{n+1}) - y_{n+1}$ αναπαριστά το σφάλμα για κάθε βήμα.

Όπως αναφέρθηκε στην *Υποπαράγραφο IV.4.2.(ν)*, για την μέθοδο τάξης p, αυτό το σφάλμα είναι της μορφής $\mathrm{O}(h^{p+1})$.

Εάν τώρα αναπτύξουμε την συνάρτηση $\tilde{y}(x)$ σε σειρά Taylor μέσα στην περιοχή του x_n, θα πάρουμε:

$$\tilde{y}(x_{n+1}) = y_n \sum_{j=0}^{\infty} (-1)^j \frac{A^j \; h^j}{j!}.$$

Έτσι, συνοψίζοντας, θα έχουμε:

$$\tilde{y}(x_{n+1}) - y_{n+1} = y_n \; (\sum_{j=0}^{\infty} (-1)^j \frac{A^j \; h^j}{j!} - P(A\,h)) = \mathrm{O}(h^{p+1}),$$

Επομένως, θα είναι:

$$P(A\,h) = \sum_{j=0}^{p} (-1)^j \frac{A^j \; h^j}{j!},$$

οπότε χρησιμοποιώντας, αφ' ενός τον ορισμό της $A-$ευστάθειας και αφ' ετέρου το γεγονός ότι $y_n = [P(A\,h)]^n \; y_0$, συμπεραίνουμε ότι:

Θεώρημα IV.4.2.(viii).1. ***Μία αναγκαία και ικανή συνθήκη, προκειμένου μία μέθοδος ξεχωριστών βημάτων να είναι $A-$ευσταθής, δίνεται από την παρακάτω ανισοτική σχέση:***

$$|P(A\,h)| < 1. \; \blacksquare$$

Για να κατανοήσουμε την τεχνική λειτουργικότητα του περιεχομένου του *Θεωρήματος IV.4.2.(iii).1*, θα αναφερθούμε σε συγκεκριμένες μεθόδους:

Παραδείγματα IV.4.2.(viii).2. ***Αν θεωρήσουμε την μέθοδο Euler, τότε:***

$$|P(A\,h)| = 1 - A\,h\,,$$

και συνεπώς, σύμφωνα με το προηγούμενο Θεώρημα, αυτή η μέθοδος είναι $A-$***ευσταθής όταν η ποσότητα*** $A\,h$ ***ανήκει μέσα στον ανοικτό επίπεδο δίσκο που έχει κέντρο το σημείο*** $1 \in \mathbb{C}$ ***και ακτίνα ίση με*** 1***, δηλαδή όταν:***

$$A\,h \in \Delta(1,1) = \{z \in \mathbb{C} : |z-1| < 1\}\,.\blacksquare$$

Ας επανέλθουμε τώρα στον σχολιασμό της βασικής συνθήκης του *Θεωρήματος IV.4.2.(viii).1.*

Εάν, στην ανισότητα $|P(A\,h)| < 1$, πάρουμε $A = 1$, τότε η σχέση $|P(h)| < 1$, θα ισχύει για κάθε h, που ανήκει μέσα σ' ένα ανοικτό διάστημα της μορφής $[0, R]$. Ο θετικός αριθμός R ονομάζεται ***ακτίνα ευστάθειας*** της υπό θεώρηση μεθόδου.

Σύμφωνα με το προηγούμενο *Παράδειγμα*, η ακτίνα ευστάθειας R της μεθόδου Euler είναι ίση με 2.

Ομοίως, μπορούμε να αποδείξουμε ότι η ακτίνα ευστάθειας της βελτιωμένης μεθόδου Euler είναι ίση με 2, ενώ για την κλασσική μέθοδο Runge-Kutta τάξης 4, η ακτίνα *R* ισούται με 8.

Παρατήρηση IV.4.2.(viii).3. Όπως είδαμε στην *Υποπαράγραφο IV.4.2.(vi)*, για δεδομένους βαθμούς $1, 2, 3$ και 4, είναι δυνατή η κατασκευή μαθηματικών *Τύπων* Runge-Kutta, που έχουν τάξεις ίσες με τους αντίστοιχους βαθμούς.

Εάν όμως θελήσουμε να σχηματίσουμε μαθηματικούς *Τύπους* ολοκλήρωσης, που έχουν τάξη μικρότερη από τον βαθμό, τότε μπορούμε να αυξήσουμε το μέγεθος της ακτίνας ευστάθειας.

Προς τούτο, αρκεί να πραγματοποιηθεί μία κατάλληλη επιλογή των αυθαίρετων συντελεστών, που συμμετέχουν στις δομικές εκφράσεις των σχημάτων Runge-Kutta. Έτσι, για την μέθοδο Runge-Kutta βαθμού 2 και τάξης 1:

$$y_{n+1} = \left(1 - h + a\, h^2\right) y_n ,$$

εάν πάρουμε $a = \frac{1}{8}$, θα έχουμε $R = 8$ (κάτω από την προϋπόθεση ότι $h \neq 4$), σε αντιδιαστολή με την βελτιωμένη μέθοδο Euler, όπου έχουμε $R = 2$. ■

Παρατήρηση IV.4.2.(viii).4. Στην περίπτωση της μεθόδου Euler, ο τόπος αριθμητικής ευστάθειας είναι ένας ανοικτός δίσκος. Δεν συμβαίνει όμως πάντα το ίδιο και στις άλλες μεθόδους.

Στην *Υποπαράγραφο IV.4.3.(x)*, θα περιγράψουμε έναν αποτελεσματικό τρόπο αναζήτησης του τόπου αριθμητικής ευστάθειας. ■

Παρατήρηση IV.4.2.(viii).5. Γενικά, οι διαφορικές εξισώσεις που πρέπει να ολοκληρώσουμε δεν είναι της μορφής: $y' = -A\,y$, αλλά περιγράφονται από την ευρεία θεωρητική διατύπωση $y' = F(x, y)$.

Εάν η απεικόνιση F είναι διαφορίσιμη ως προς την εξαρτημένη μεταβλητή *y*, τότε θα εφαρμόζουμε την προηγηθείσα θεωρία επί της εξίσωσης:

$$y' = F'_y\ y\,.$$

Στην πραγματικότητα, από την Θεωρία αυτήν δεν μπορούμε να προσδοκούμε αξιόλογα αποτελέσματα παρά μόνον όταν η F_y' δεν μεταβάλλεται πολύ, καθώς το *x* διατρέχει το διάστημα $[a, \beta]$, και δεν εξαρτάται αποκλειστικά από το y. ■

Παρατήρηση IV.4.2.(viii).6. Προκειμένου περί ενός συστήματος διαφορικών εξισώσεων, η ποσότητα A δεν είναι πλέον μία μιγαδική σταθερά, αλλά ένας τετραγωνικός πίνακας $k \times k$. Τότε, η αντίστοιχη συνθήκη ευστάθειας είναι η ακόλουθη:

$$h\,\lambda_j \in \boldsymbol{D}\,,\quad j = 1,2,\ldots,k\,,$$

όπου $\boldsymbol{D}$ είναι ο τόπος αριθμητικής ευστάθειας της μεθόδου και όπου λ_j είναι οι ιδιοτιμές του πίνακα *A*.

Επίσης, σ' αυτήν την περίπτωση, στην οποία πλέον αναζητούμε την διανυσματική συνάρτηση y που επαληθεύει το σύστημα $y' = F(x, y)$, ενεργούμε όπως και στην ειδική περίπτωση όπου $k = 1$, δηλαδή θεωρούμε την διανυσματική εξίσωση:

$$y' = F'_y\ y\,,$$

όπου το σύμβολο F'_y αναπαριστά τον Ιακωβιανό πίνακα της F ως προς την μεταβλητή y, και επιδιώκουμε τα ίδια αποτελέσματα, κάτω όμως από τις ίδιες περιοριστικές προϋποθέσεις που αναφέρθηκαν στην *Παρατήρηση IV.4.2.(viii).5*. ■

Παρατήρηση IV.4.2.(viii).7. Η έννοια της ευστάθειας, που εξετάσαμε στην παρούσα *Υποπαράγραφο*, δεν είναι χρήσιμη μόνον στην διαδικασία ολοκλήρωσης διαφορικών εξισώσεων των οποίων η λύση τείνει προς το 0, καθώς το x τείνει προς το $+\infty$.

Πράγματι, για κάθε διαφορική εξίσωση, συνεπάγεται ότι:

$$\lim_{n\to\infty} | y_n - y(x_n) | = 0$$

και, συνεπώς, η ύπαρξη ευστάθειας εμποδίζει την μεγάλη απομάκρυνση της προσεγγιστικής λύσης από την αληθινή λύση της διαφορικής εξίσωσης. Πρόκειται λοιπόν για μία πρακτική έννοια, που διαδραματίζει σημαίνοντα ρόλο στην προσπάθειά μας για σωστή αριθμητική ολοκλήρωση διαφορικών εξισώσεων.

Το γεγονός ότι η έννοια αυτή δεν είχε εμφανιστεί σε προηγούμενες *Παραγράφους* οφείλεται στην έμφαση που είχαμε δώσει στην προτεραιότητα των θεωρητικών εννοιών: έτσι, η συνάφεια και η σύγκλιση είχαν οριστεί για μηδενικά βήματα, ενώ η θεωρητική ευστάθεια είχε θεμελιωθεί για μικροσκοπικά βήματα.

Όμως, στην πραγματικότητα, δεν ολοκληρώνουμε αριθμητικά μία διαφορική εξίσωση χρησιμοποιώντας μηδενικά ή πολύ μικροσκοπικά βήματα. Έτσι, οι θεωρητικές συνθήκες δεν μπορούν να επαληθευτούν επακριβώς και η θεώρηση της πρακτικής και επιπρόσθετης έννοιας της *A*-ευστάθειας είναι αναγκαία. ■

IV.4.2.(ix). Έλεγχος Βήματος Υπάρχουν ακόμη δύο πρακτικά προβλήματα:

- ***Το πρώτο αφορά στην επιλογή του αρχικού βήματος***, λαμβανομένου υπ' όψη του μέγιστου σφάλματος που μπορούμε να ανεχθούμε, αλλά και του σεβασμού προς την διατήρηση της ιδιότητας της $A-$ευστάθειας που πρέπει να δείξουμε.
- ***Το δεύτερο αφορά στον μεταγενέστερο έλεγχο του βήματος***, λαμβανομένων υπ' όψη των όποιων αποτελεσμάτων έχουν, εν τω μεταξύ, προκύψει από την αριθμητική ολοκλήρωση του διαφορικού συστήματος.

Ας εξετάσουμε πρώτα την επιλογή του αρχικού βήματος.

Έστω λοιπόν μία μέθοδος ολοκλήρωσης, που είναι τάξης p.

Όπως αναφέρθηκε στην *Υποπαράγραφο IV.4.2.(v)*, το σφάλμα σε κάθε βήμα είναι μία συνάρτηση της μορφής $O(h^{p+1})$.

Άρα το ολικό σφάλμα σ_n μετά από n βήματα θα είναι μία συνάρτηση της μορφής $K\ n\ h^{p+1}$.

Ας συμβολίσουμε με $y_{n\,h}$ την προσεγγιστική τιμή της λύσης στο σημείο $x_0+n\,h$. Φυσικά, αντί να ολοκληρώσουμε αριθμητικά από το σημείο $x_0=a$ έως το σημείο $x_0+n\,h=a+n\,h$ (πραγματοποιώντας n βήματα αριθμητικής ολοκλήρωσης), μπορούμε να ολοκληρώσουμε αριθμητικά έως το ίδιο σημείο, σε μία μόνον φορά και με βήμα ολοκλήρωσης ίσο με $n\,h$. Σε μία τέτοια περίπτωση, θα προκύψει ένα σφάλμα περίπου ίσο με $K\ (n\,h)^{p+1}$.

Εάν αναπαραστήσουμε με $y_{(nh)}$ την προσεγγιστική τιμή της λύσης στο σημείο $x_0+n\,h$, όπως αυτή θα σχηματιστεί, κατά αυτόν τον τρόπο, τότε θα είναι:

$$\frac{y_{(nh)} - y_{nh}}{\sigma_n} = \frac{K\,(n\,h)^{p+1} - K\,n\,h^{p+1}}{K\,n\,h^{p+1}},$$

από όπου έπεται ότι:

$$\sigma_n = \frac{y_{(nh)} - y_{nh}}{n^p - 1}.$$

Το σφάλμα σ_1, που διαπράττεται μετά από ένα βήμα ολοκλήρωσης, είναι λοιπόν ίσο με:

$$\sigma_1 = \frac{\sigma_n}{n},$$

ενώ *το μοναδιαίο σφάλμα* για ένα βήμα ορίζεται από την σχέση:

$$\tau_1 = \frac{\sigma_n}{h}.$$

Στην πράξη, επιθυμούμε το μοναδιαίο σφάλμα να παραμένει σταθερό σε κάθε βήμα, και ίσο με μία προκαθορισμένη τιμή τ.

Για αυτό, υπολογίζοντας το τ_1, όπως πιο πάνω, για ένα βήμα h_0 και για n γενικά ίσο με 2, προκειμένου να έχουμε ένα μοναδιαίο σφάλμα, σε κάθε βήμα, ίσο με τ, θα πρέπει να λάβουμε:

$$h = h_0 \left(\frac{\tau}{\tau_1} \right)^{\frac{1}{p}}.$$

Θέλοντας να αποφύγουμε τα σφάλματα στρογγύλευσης, δεν διατηρούμε παρά μόνον δύο ή τρία δεκαδικά ψηφία στην αριθμητική έκφραση του h, που πετυχαίνουμε κατά αυτόν τον τρόπο.

Εάν η συμπεριφορά της θεωρητικής λύσης ποικίλλει θεαματικά μέσα στο διάστημα $[a,\beta]$, ίσως να πρέπει να επιδιώξουμε την μεταβολή του h ταυτόχρονα με την εξέλιξη της διαδικασίας της αριθμητικής ολοκλήρωσης.

Συνήθως, οδηγούμαστε στην σμίκρυνση του h, όταν η ποσότητα $y'(x)$ αυξάνεται πολύ, κατ' απόλυτη τιμή, και σε αύξηση του h, σε αντίθετη περίπτωση.

Για παράδειγμα στην διαφορική εξίσωση $y' = -10^9\ y$, πρέπει να επιλέξουμε ένα μικρό βήμα, όταν το x είναι γειτονικό με το μηδέν και στην συνέχεια, ένα πολύ πιο μεγάλο βήμα. Αυτός ο μεταγενέστερος έλεγχος του βήματος της αριθμητικής ολοκλήρωσης πραγματοποιείται κατά την διάρκεια της ολοκλήρωσης, σύμφωνα με τον τρόπο που εκτέθηκε πιο πάνω.

Όμως, η χρήση μίας τέτοιας μεθόδου απαιτεί την, εκ των προτέρων, γνώση της συμπεριφοράς της θεωρητικής λύσης.

Ένας άλλος τρόπος έγκειται στην διαδοχική εφαρμογή δύο μεθόδων με διαφορετικές τάξεις, και στην, εν συνεχεία, θεώρηση των αριθμητικών διαφορών των προσεγγιστικών τιμών, που θα έχουν βρεθεί. Αυτές οι διαφορές μπορούν να θεωρούνται σαν ικανοποιητικές προσεγγίσεις των διαδοχικών τιμών του σφάλματος, τουλάχιστον στην περίπτωση που οι μαθηματικοί τύποι της αριθμητικής ολοκλήρωσης έχουν πολύ ασθενείς τάξεις.

IV.4.2.(x). Βελτίωση ακρίβειας Η ιδέα είναι η ίδια με εκείνη που διέπει την διάρθρωση της μεθόδου Romberg για την βελτίωση της ακρίβειας του μαθηματικού τύπου τετραγωνισμού τω τραπεζίων: Ολοκληρώνουμε αριθμητικά την διαφορική εξίσωση, χρησιμοποιώντας ένα βήμα h και ένα βήμα $(k\,h)$, όπου k είναι ένας ακέραιος αριθμός μεγαλύτερος του ένα, και συμβολίζουμε με y_n και $\bar{y}_n$, αντίστοιχα, τις προσεγγιστικές τιμές της θεωρητικής λύσης στο κοινό σημείο x_n:

$$y_n = y(x_n) + h^p\ \sigma(x_n) + \mathrm{O}(h^{p+1}),$$

$$\bar{y}_n = y(x_n) + k^p\ h^{\ p}\ \sigma(x_n) + \mathrm{O}(h^{p+1}).$$

Μπορούμε τότε να χρησιμοποιήσουμε το σχήμα Neville-Aitken για να προεκβάλλουμε, στο σημείο $x_0 = 0$, το πολυώνυμο παρεμβολής, πρώτου βαθμού ως προς την μεταβλητή h^p, που διέρχεται από αυτές τις δύο τιμές. Καταλήγουμε έτσι σε μία νέα προσέγγιση y_n^* του $y(x_n)$, που είναι τέτοια ώστε:

$$y_n^* = \frac{\bar{y}_n - k^p\ y_n}{1 - k^p} \text{ και } y_n^* = y(x_n) + \mathrm{O}(h^{p+1}).$$

IV.4.3. ΟΙ ΜΕΘΟΔΟΙ ΣΥΝΑΠΤΩΝ ΒΗΜΑΤΩΝ

Στα πρακτικά προβλήματα, οι μέθοδοι Runge-Kutta είναι πολύ χρήσιμες και έχουν συχνές και διαδεδομένες εφαρμογές. Αυτό οφείλεται στα έξοχα αποτελέσματα, που δίνουν, καθώς και στην σχετικά μεγάλη ακτίνα ευστάθειας, που παρουσιάζουν, όπως επίσης και στην αμεσότητα της διαδικασίας εκκίνησής τους, που δεν εξαρτάται παρά μόνον από την γνώση της αρχικής τιμής y_0.

Επί πλέον, ο έλεγχος του σφάλματος μπορεί να πραγματοποιηθεί εύκολα κατά την διάρκεια της ολοκλήρωσης του διαφορικού συστήματος, ενώ η αλλαγή της επιλογής του μεγέθους του βήματος συνιστά, ταυτόχρονα, μία απλή δυνατότητα. Το σοβαρότερο μειονέκτημα των μεθόδων Runge-Kutta έγκειται στην δομική τους εσωτερική ανάγκη να προϋποθέτουν κάποιες βοηθητικές εκτιμήσεις της απεικόνισης F. Ο σκοπός των μεθόδων συναπτών βημάτων συνίσταται στην αποφυγή αυτού του μειονεκτήματος.

Η γενική μορφή των ***μεθόδων συναπτών βημάτων*** είναι η ακόλουθη:

$$a_k\ y_{n+k} + \ldots + a_0\ h_n = h\,(\beta_k\ F_{n+k} + \ldots + \beta_0 \cdots F_n)\,, \qquad (26)$$

όπου:

$n \in \boldsymbol{N}$,

$k = 0,1,2,\ldots,$

y_0 : ***είναι η δεδομένη αρχική τιμή στο πρόβλημα Cauchy,***

$x_0 = a$,

$x_{n+1} = x_n + h$,

$a_k \neq 0$,

και:

$F_i = F(x_i, y_i)$.

Βλέπουμε ότι στις μεθόδους συναπτών βημάτων πραγματοποιείται ο υπολογισμός της προσεγγιστικής τιμής y_{n+k}, μετά από συνδυασμό αποτελεσμάτων των προγενεστέρων βημάτων.

Έτσι, εφόσον είναι γνωστές οι τιμές $y_n, y_{n+1}, \ldots, y_{n+k-1}$, ο τύπος (26) επιτρέπει τον προσδιορισμό του y_{n+k}:

- απ' ευθείας, εάν $\beta_k = 0$, οπότε λέμε ότι ***πρόκειται περί μίας σαφούς μεθόδου***,
- εμμέσως, σαν λύση της εξίσωσης

$$a_k\ y_{n+k} = h\,\beta_k\ F(x_{n+k}, y_{n+k}) + \ldots + h\,\beta_0\ F(x_n, y_n) - a_{k-1}\ y_{n+k+1} - \ldots - a_0\ y_n,$$

όταν $\beta_k \neq 0$, οπότε λέμε ότι ***πρόκειται περί μίας μεθόδου κατά την οποία το y_{n+k} προκύπτει με πεπλεγμένο τρόπο*** σαν λύση μίας εξίσωσης.

Ισχύει η εξής Πρόταση:

Θεώρημα IV.4.3.1. ***Έστω ότι $\beta_k \neq 0$. Τότε, η πεπλεγμένη εξίσωση (26) έχει μοναδική λύση, εάν:***

$$h < \frac{1}{L}\left|\frac{a_k}{\beta_k}\right|,$$

όπου L είναι η σταθερά Lipschitz της απεικόνισης F.

Απόδειξη. Προφανώς, σύμφωνα με την σχέση (26), είναι:

$$y_{n+k} = h\,\frac{\beta_k}{a_k}\,F(x_{n+k}, y_{n+k}) + \ldots$$

Μπορούμε να προσδιορίσουμε την τιμή y_{n+k}, χρησιμοποιώντας την μέθοδο των διαδοχικών προσεγγίσεων, που συναντήσαμε στην *Παράγραφο ΙΙΙ.1*.

Πράγματι, για κάθε $p = 0,1,2,\ldots$, θέτουμε:

$$y_{n+k}^{(p+1)} := h\,\frac{\beta_k}{a_k}\,F(x_{n+k}, y_{n+k}^{(p)}) + \ldots$$

Τότε, είναι σαφές ότι:

$$y_{n+k}^{(p+2)} - y_{n+k}^{(p+1)} = h\,\frac{\beta_k}{a_k}\,(F(x_{n+k}, y_{n+k}^{(p+1)}) - F(x_{n+k}, y_{n+k}^{(p)}))\,,$$

οπότε, εκμεταλλευόμενοι την ιδιότητα Lipschitz της F, διαπιστώνουμε πως:

$$\left|y_{n+k}^{(p+2)} - y_{n+k}^{(p+1)}\right| \le h\left|\frac{\beta_k}{a_k}\right| L\,\left|y_{n+k}^{(p+1)} - y_{n+k}^{(p)}\right|.$$

Μία εφαρμογή του *Θεωρήματος ΙΙΙ.1.2* (, με σημείο εκκίνησης της επαναληπτικής διαδικασίας των διαδοχικών προσεγγίσεων το σημείο $(x_{n+k}, y_{n+k}^{(0)})$), μπορεί τώρα να μας πείσει ότι η επαναληπτική ακολουθία:

$$(y_{n+k}^{(p)} : p = 0,1,2,\ldots)$$

συγκλίνει προς την μοναδική λύση y_{n+k} της εξίσωσης (26), εάν ισχύει

$$h\left|\frac{\beta_k}{a_k}\right| L < 1.$$

Το συμπέρασμα αυτό ολοκληρώνει την *Απόδειξη* του *Θεωρήματος*. ■

Είναι φανερό ότι, όπως οι μέθοδοι ξεχωριστών βημάτων, έτσι και οι μέθοδοι συναπτών βημάτων οφείλουν να σέβονται τις προϋποθέσεις που εξασφαλίζουν τις ιδιότητες της συνάφειας, της ευστάθειας και της σύγκλισης. Αυτές οι προϋποθέσεις μεταφράζονται σε συνθήκες, που περιγράφονται από έναν ορισμένο αριθμό σχέσεων, που αναφέρονται στον αποτελεσματικό προσδιορισμό των συντελεστών a_j και β_j. Η μελέτη αυτών των εννοιών στηρίζεται στην πρότερη γνώση της Θεωρίας των Εξισώσεων Διαφορών, που θα εξετάσουμε αμέσως τώρα.

IV.4.3.(i). Εξισώσεις Διαφορών Στην *Υποπαράγραφο* αυτή, δεν θα πραγματευτούμε παρά μόνον τις εξισώσεις διαφορών, που είναι γραμμικές, ομογενείς και με σταθερούς συντελεστές. Ο ενδιαφερόμενος αναγνώστης, που επιζητεί μία πλήρη ανάπτυξη της γενικής Θεωρίας των Εξισώσεων Διαφορών, μπορεί να προστρέξει στην βιβλιογραφική παραπομπή [25].

Έστω:

$$y : \mathbb{Z} \to \mathbb{R} : n \mapsto y(n)$$

μία πραγματική συνάρτηση μίας ακέραιας μεταβλητής.

Ορισμός IV.4.3.(i).1. ***Κάθε σχέση της μορφής:***

$$G(y(n),\ldots,y(n+k),n)=0$$

λέγεται εξίσωση διαφορών. Ο αριθμός k ***ονομάζεται τάξη της εξίσωσης διαφορών. Εάν η απεικόνιση*** G ***είναι γραμμική, ως προς τις μεταβλητές*** $y(n),\ldots,y(n+k)$***, τότε η εξίσωση διαφορών καλείται γραμμική.***

Επί πλέον, εάν η εξίσωση είναι της μορφής:

$$\sum_{j=0}^{k} c_j \; y(n+j) = g(n),$$

όπου οι αριθμοί c_j ***είναι πραγματικοί και ανεξάρτητοι του*** n***, τότε λέμε ότι η εξίσωση έχει σταθερούς συντελεστές ή ότι είναι με σταθερούς συντελεστές.***

Όταν

$$g(n) = 0, \forall n,$$

λέμε ότι η εξίσωση είναι ομογενής. ■

Η Θεωρία των Εξισώσεων Διαφορών είναι ενδιαφέρουσα και απαραίτητη για την Εξαγωγή αριθμητικών συμπερασμάτων, που αναφέρονται σε διάφορους κλάδους των εφαρμοσμένων Μαθηματικών.

Υπάρχει, μάλιστα, μία εντυπωσιακή ομοιότητα μεταξύ της Θεωρίας των Γραμμικών Ομογενών Εξισώσεων Διαφορών με σταθερούς συντελεστές και που έχουν τάξη ίση με k, και της Θεωρίας των Γραμμικών Ομογενών Διαφορικών Εξισώσεων με σταθερούς συντελεστές και που έχουν τάξη ίση με k.

Συγκεκριμένα, και για να γίνει κατανοητή η επινόηση αυτού του ισχυρισμού, παραθέτουμε τα ακόλουθα ενδεικτικά αποτελέσματα:

Θεώρημα IV.4.3.(i).2. ***Εάν*** $y_1, y_2, \ldots, y_p$ ***είναι λύσεις της γραμμικής, ομογενούς, τάξης k, και με σταθερούς συντελεστές εξίσωσης διαφορών:***

$$\sum_{j=0}^{k} c_j \; y(n+j) = 0, \qquad (27)$$

τότε κάθε γραμμικός συνδυασμός των $y_1, y_2, \ldots, y_p$ ***είναι επίσης λύση της*** (27).

Απόδειξη. Επειδή οι συναρτήσεις $y_\nu : \mathbb{Z} \to \mathbb{R}$ αποτελούν λύσεις της (27), ισχύουν οι εξισώσεις

$$\sum_{j=0}^{k} c_j \; y_\nu \; (n+j) = 0, \text{ για κάθε } \nu = 1, \ldots, p \, .$$

Πολλαπλασιάζοντας την πρώτη εξίσωση με λ_1, την δεύτερη με $\lambda_2, \ldots,$ και την τελευταία με λ_p όπου $\lambda_1, \lambda_2, \ldots, \lambda_p$ είναι οποιεσδήποτε πραγματικές σταθερές, και, στην συνέχεια, αθροίζοντας κατά μέλη, λαμβάνουμε διαδοχικά:

$$\sum_{\nu=0}^{p} \lambda_\nu \; \sum_{j=0}^{k} c_j \; y_\nu (n+j) = 0$$

$$\Leftrightarrow$$

$$\sum_{j=0}^{k} c_j \; \sum_{\nu=0}^{p} \lambda_\nu \; y_\nu (n+j) = 0 \, .$$

Η τελευταία σχέση σημαίνει ότι οποιοσδήποτε γραμμικός συνδυασμός:

$$\sum_{\nu=0}^{p} \lambda_\nu \; y_\nu (n+j)$$

των λύσεων $y_1, y_2, \ldots, y_p$ της (27) είναι επίσης λύση της ίδιας εξίσωσης. ■

Θεώρημα IV.4.3.(i).3. ***Εάν*** $y_1, y_2, \ldots, y_k$ ***είναι λύσεις της*** (27) ***και εάν η ορίζουσα:***

$$D = \det \begin{pmatrix} y_1(0) & y_1(1) & \ldots & y_1(k-1) \\ y_2(0) & y_2(1) & \ldots & y_2(k-1) \\ \ldots & \ldots & \ldots & \ldots \\ y_k(0) & y_k(1) & \ldots & y_k(k-1) \end{pmatrix}$$

είναι διαφορετική από το μηδέν, τότε η γενική λύση της (27) ***είναι:***

$$y = \sum_{\nu=1}^{k} \lambda_\nu \; y_\nu \qquad (\lambda_1, \ldots, \lambda_k \in \mathbb{R}) \, .$$

Απόδειξη. Οι k λύσεις $y_1, y_2, \ldots, y_k$ της εξίσωσης (27) προσδιορίζονται πλήρως από τις αρχικές τιμές:

$$y_1(0), \ldots, y_1(k-1),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots,$$
$$y_k(0), \ldots, y_k(k-1).$$

Πράγματι, αρκεί να χρησιμοποιήσουμε την (27) προκειμένου να υπολογίσουμε την ποσότητα $y_j(n)$,για κάθε ακέραια θετική τιμή του *n* και για κάθε $j = 1, \ldots, n$.

Επί πλέον, μπορούμε πάντα να επιλέξουμε τις αρχικές τιμές, με τέτοιον τρόπο ώστε η ορίζουσα D να είναι διάφορη από το μηδέν. Σύμφωνα με το προηγούμενο *Θεώρημα* κάθε συνάρτηση

$$y = \sum_{\nu=1}^{k} \lambda_\nu \; y_\nu$$

είναι επίσης λύση της εξίσωσης (27). Θα δείξουμε ότι όλες οι λύσεις της (27) είναι αυτής της μορφής.

Προς τούτο, ας υποθέσουμε ότι $\phi(n)$ είναι μία τυχαία λύση της (27). Είναι σαφές πως η $\phi(n)$ είναι πλήρως προσδιορισμένη από τις τιμές $\phi(0), \ldots, \phi(\mathrm{k}-1)$. Εάν αναζητήσουμε μέσα στο σύνολο των συναρτήσεων

$$y = \sum_{\nu=1}^{k} \lambda_\nu \; y_\nu$$

μία συνάρτηση, που να έχει τις ίδιες αρχικές τιμές, τότε η αναζήτηση αυτή ανάγεται στον προσδιορισμό πραγματικών σταθερών $\lambda_1, \ldots, \lambda_k$ που είναι τέτοιες ώστε:

$$y(0) = \phi(0) = \sum_{\nu=1}^{k} \lambda_\nu \; y_\nu(0),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,$$
$$y(k-1) = \phi(\mathrm{k}-1) = \sum_{\nu=1}^{k} \lambda_\nu \; y_\nu(k-1).$$

Επειδή η ορίζουσα D είναι διαφορετική από το μηδέν, το παραπάνω σύστημα επιδέχεται μία και μόνο λύση, για οποιεσδήποτε τιμές των $\phi(0),\ldots,\phi(k-1)$. Καθώς όμως οι αρχικές τιμές προσδιορίζουν πλήρως και μονοσήμαντα την λύση της εξίσωσης, έπεται ότι θα είναι $y(n)=\phi(n)$, για κάθε n. Η *Απόδειξη* του *Θεωρήματος* είναι πλήρης. ■

Θα στρέψουμε τώρα την προσοχή μας προς την ανεύρεση της ακριβούς μορφής της γενικής λύσης της εξίσωσης διαφορών (27). Εάν δεχθούμε ότι η έκφραση της γενικής αυτής λύσης είναι:

$$y(n)=t^n \text{, για κάποιο } t\neq 0,$$

τότε, μεταφέροντάς την μέσα στην (27), θα λάβουμε:

$$t^n\,(c_0+c_1\,t+\ldots+c_k\,t^k)=0,$$

ή ακόμα, επειδή $t\neq 0$:

$$c_0+c_1\,t+\ldots+c_k\,t^k=0. \qquad (28)$$

Αυτή η εξίσωση ονομάζεται ***χαρακτηριστική εξίσωση*** της εξίσωσης πεπερασμένων διαφορών (27). Η σημασία της έγκειται στην διάγνωση ότι η συνάρτηση t^n είναι λύση της (27) εάν ο αριθμός t είναι ρίζα της εξίσωσης (28).

Προφανώς, η επίλυση της χαρακτηριστικής εξίσωσης διέπεται από μία ποικιλία ενδεχομένων: Εάν όλες οι ρίζες της (28) είναι απλές, έστω $t_1,\ldots t_k$, τότε:

$$D=\det\begin{pmatrix} 1 & t_1 & \ldots & t_1^{k-1} \\ 1 & t_2 & \ldots & t_2^{k-1} \\ \ldots & \ldots & \ldots & \ldots \\ 1 & t_k & \ldots & t_k^{k-1} \end{pmatrix},$$

κι επειδή η ορίζουσα αυτή είναι μία ορίζουσα Vandermonde, θα έχουμε, σύμφωνα με τα όσα αναφέρθηκαν στην *Παράγραφο IV.1.1*, την ακόλουθη σχέση:

$$D = \Pi_{\nu > j}(t_\nu - t_j)\,.$$

Καθώς όμως οι ρίζες $t_1, \ldots t_k$ είναι όλες ανά δύο διάφορες, έπεται ότι:

$$D \neq 0\,,$$

οπότε, σύμφωνα με το *Θεώρημα IV.4.3.(i).3*, η γενική λύση της (27) θα είναι:

$$y(n) = \sum_{\nu=1}^{k} \lambda_\nu \; t_\nu^n \qquad (\lambda_1, \ldots \lambda_k \in \mathbb{R})\,.$$

Εάν η εξίσωση (28) έχει μιγαδικές ρίζες, τότε, επειδή οι συντελεστές c_j έχουν υποτεθεί ότι είναι πραγματικοί αριθμοί, οι ρίζες αυτές μπορούν να εκλαμβάνονται ανά ζεύγη: κάθε ζεύγος θα απαρτίζεται από δύο συζυγείς ρίζες της (28).

Επί πλέον, οι συντελεστές λ_ν και $\overline{\lambda}_\nu$, που αντιστοιχούν σε ένα τέτοιο ζεύγος, θα είναι και αυτοί συζυγείς μεταξύ τους, προκειμένου το τελικό άθροισμα να είναι πραγματικός αριθμός, για κάθε n. Σε αυτήν την περίπτωση, η γενική λύση θα σχηματίζεται από έναν γραμμικό συνδυασμό των συναρτήσεων:

t^n εάν t είναι μία πραγματική λύση της (28),

και:

$\begin{Bmatrix} \rho^n \cos(\omega\, n) \\ \rho^n \sin(\omega\, n) \end{Bmatrix}$, εάν $\rho\,(\cos\omega + i\,\sin\omega)$ είναι μία μιγαδική λύση της (28).

Τέλος, η περίπτωση, όπου η εξίσωση (28) επιδέχεται πολλαπλές ρίζες, είναι περισσότερο δύσκολη, γιατί τότε δεν υπάρχουν k διαφορετικές λύσεις της μορφής t^n και τέτοιες ώστε η αντίστοιχη ορίζουσα D να διαφοροποιείται από το μηδέν. Θα πρέπει τότε να χρησιμοποιήσουμε την πολλαπλότητα των ριζών προκειμένου να κατασκευάσουμε πλεονάζουσες λύσεις.

Προς τούτο, ας υποθέσουμε ότι t είναι μία ρίζα πολλαπλότητας r της (28). Τότε:

$$P(t) = c_0\ t^n + \ldots + c_k\ t^{n+k} = 0,$$

και συνεπώς, η συνάρτηση:

$$t^n$$

είναι λύση της (27).

Ακόμη:

$$P_1(t) = t\ P'(t) = c_0\ n\ t^n + \ldots + c_k\ (n+k)\ t^{n+k} = 0,$$

επειδή $P'(t) = 0$, και επομένως, η συνάρτηση:

$$n\ t^n$$

είναι λύση της (27) ομοίως:

$$P_2(t) = t\ P_1'(t) = c_0\ n^2\ t^n + \ldots + c_k\ (n+k)^2\ t^{n+k} = 0,$$

επειδή $P''(t) = 0$, και άρα, η συνάρτηση:

$$n^2\ t^n$$

είναι λύση της (27).

Συνεχίζοντας κατά αυτόν τον τρόπο, καταλήγουμε βρίσκοντας:

$$P_{r-1}(t) = t\ P_{r-2}'(t) = c_0\ n^{r-1}\ t^n + \cdots + c_k\ (n+k)^{r-1} t^{n+k} = 0,$$

επειδή $P^{(r-1)}(t) = 0$, και έτσι, η συνάρτηση:

$$n^{r-1}\ t^n$$

είναι λύση της (27).

Τελικά, θα έχουν βρεθεί k λύσεις, γιατί το άθροισμα των πολλαπλοτήτων των ριζών της (28) ισούται με k.

Έστωσαν λοιπόν:

$$t_1, t_2, \ldots, t_p$$

όλες οι ανά δύο διάφορες ρίζες της (28) και έστωσαν:

$$r_1, r_2, \ldots r_p$$

οι αντίστοιχες πολλαπλότητες τους. Εκείνο που απομένει να δειχθεί είναι ότι η ορίζουσα:

$$D = \det \begin{pmatrix} 1 & t_1 & \ldots & t_1^{k-1} \\ 0 & t_1 & \ldots & (k-1)\, t_1^{k-1} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & t_1 & \ldots & (k-1)^{r_1-1}\, t_1^{k-1} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ 1 & t_p & \ldots & t_p^{k-1} \\ 0 & t_p & \ldots & (k-1)\, t_p^{k-1} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & t_p & \ldots & (k-1)^{r_p-1}\, t_p^{k-1} \end{pmatrix}$$

είναι διάφορη του μηδενός, έτσι ώστε να είναι δυνατή η εφαρμογή του *Θεωρήματος IV.4.3.(i).3*, προκειμένου να σχηματιστεί η γενική λύση της (27).

Η απόδειξη αυτού του ισχυρισμού μπορεί να γίνει απ' ευθείας ή ακόμη πιστοποιώντας ότι οι k λύσεις, που προέκυψαν παραπάνω, είναι γραμμικά ανεξάρτητες μεταξύ τους. (Η γραμμική ανεξαρτησία προέρχεται απλά από το γεγονός ότι οι δυνάμεις του *n* είναι γραμμικά ανεξάρτητες.) Οι λεπτομέρειες της επί μέρους αυτής απόδειξης αφήνονται στην ευχέρεια του αναγνώστη.

Συνοψίζοντας, παραθέτουμε το επόμενο *Θεώρημα*:

Θεώρημα IV.4.3.(i).4. ***Ας θεωρήσουμε την εξίσωση διαφορών:***

$$\sum_{j=0}^{k} c_j \; y(n+j) = 0,$$

όπου οι συντελεστές c_j ***είναι πραγματικοί αριθμοί, ανεξάρτητοι από το*** n***, και τέτοιοι ώστε:***

$$c_0 \; c_k \neq 0.$$

Τότε, η γενική λύση $y(n)$ ***αυτής της εξίσωσης δίνεται από τον τύπο:***

$$y(n) = \sum_{\nu=1}^{p} A_\nu(n) t_\nu^n + \sum_{\nu=p+1}^{q} [B_\nu(n)\cos(\beta_\nu n) + C_\nu(n)\sin(\beta_\nu n)]\,\lambda_\nu^n,$$

όπου οι ποσότητες:

$$t_\nu \qquad (\nu = 1,\ldots,p)$$

και:

$$t_\nu \; (\cos\beta_\nu \pm i\sin\beta_\nu) \qquad (\nu = p+1,\ldots,q \quad \text{και} \quad i = \sqrt{-1})$$

είναι οι ρίζες της εξίσωσης της (27)***:***

$$\sum_{j=0}^{k} c_j \; t^j = 0.$$

και όπου οι συναρτήσεις:

$$A_\nu(n),\; B_\nu(n) \;\; \textbf{και} \;\; C_\nu(n)$$

αναπαριστούν πολυώνυμα, ως προς n***, που έχουν βαθμούς ίσους με την πολλαπλότητα των αντιστοίχων ριζών της*** (28) ***ελαττωμένης κατά ένα. Εάν συμβολίσουμε με:***

$$\deg(A_\nu),\;\; \deg(B_\nu) \;\; \textbf{και} \;\; \deg(C_\nu)$$

τους βαθμούς αυτών των πολυωνύμων και εάν θέσουμε:

$$d_\nu = \begin{cases} \deg(A_\nu)+1, \text{για } \nu = 1,\ldots,p \\ \max\{\deg(B_\nu),\deg(C_\nu)\}+1, \text{για } \nu = p+1,\ldots,q, \end{cases}$$

τότε ισχύει:

$$\sum_{\nu=1}^{p} d_\nu + 2\sum_{\nu=p+1}^{q} d_\nu = k. \;\blacksquare$$

Ας εξετάσουμε τέλος την περίπτωση των εξισώσεων διαφορών που έχουν μη τετριμμένη υπόσταση στο δεξιό τους μέλος:

$$\sum_{j=0}^{k} c_j \; y(n+j) = g(n) \qquad (29)$$

Η γενική λύση της μη ομογενούς αυτής εξίσωσης είναι το άθροισμα μίας μερικής λύσης της (29) και της γενικής λύσης της αντίστοιχης ομογενούς (27).

Πράγματι, έστω $y^*(n)$ μία μερική λύση της μη ομογενούς εξίσωσης. Εάν αντικαταστήσουμε την συνάρτηση $y(n)$ με το άθροισμα $y^*(n)+\phi(n)$ όπου ϕ είναι κάποια συνάρτηση του $\mathbb{Z}$ μέσα στο $\mathbb{R}$, τότε παίρνουμε:

$$\sum_{j=0}^{k} c_j \; [y^*(n+j)+\phi(n+j)] = g(n).$$

Επειδή η συνάρτηση $y^*(n)$ είναι λύση της (29), έπεται ότι:

$$\sum_{j=0}^{k} c_j \; y^*(n+j) = g(n),$$

και, κατά συνέπεια, ότι:

$$\sum_{j=0}^{k} c_j \; \phi(n+j) = 0.$$

Άρα, η συνάρτηση ϕ πρέπει να είναι, πράγματι, γενική λύση της αντίστοιχης ομογενούς εξίσωσης διαφορών.

IV.4.3.(ii). Θεωρητικές Έννοιες Όπως στην περίπτωση των μεθόδων των ξεχωριστών βημάτων, έτσι και στην περίπτωση των μεθόδων των συναπτών βημάτων μπορούμε να ορίσουμε τις θεμελιώδεις έννοιες της ***συνάφειας***, της ***ευστάθειας***, της ***σύγκλισης*** και της ***τάξης*** και να συνάγουμε τα βασικά αποτελέσματα που τις αφορούν.

Υπενθυμίζουμε ότι η γενική μορφή των μεθόδων συναπτών βημάτων είναι η ακόλουθη:

$$\sum_{i=0}^{k} a_i \; y_{n+i} = h \sum_{i=0}^{k} \beta_i \; F_{n+i} \; ,$$

με:

$a_k \neq 0$,

και όπου:

$x_{n+1} = x_n + h$,

x_0 : δεδομένο,

$y_0 = y(x_0)$: δεδομένο,

και:

$F_j = F(x_j, y_j)$.

IV.4.3.(iii). Συνάφεια Μεθόδου Λέμε ότι μία μέθοδος συναπτών βημάτων είναι ***συναφής προς την διαφορική εξίσωση***, εάν

$$\lim_{h \to 0} \max_{0 \le n \le N} \left| \frac{1}{h} \sum_{i=0}^{k} a_i \; y(x_{n+i}) - \sum_{i=0}^{k} \beta_i \; F(x_{n+i}, y(x_{n+i})) \right| = 0 \, ,$$

για κάθε λύση $y \in C[a, \beta]$ της διαφορικής εξίσωσης.

Στις *Υποπαραγράφους IV.4.3.(v)* και *IV.4.3.(vi)*, θα δούμε πως η έννοια της συνάφειας συνδέεται με τις έννοιες της σύγκλισης και της τάξης.

Προς το παρόν, θα επισυνάψουμε στην υπό θεώρηση μέθοδο συναπτών μεθόδων τα πολυώνυμα $a(s)$ και $\beta(s)$ που ορίζονται από τις σχέσεις:

$$a(s)=\sum_{i=0}^{k} a_i\ s^i \text{ και } \beta(s)=\sum_{i=0}^{k} \beta_i\ s^i\ .$$

Ισχύει το ακόλουθο θεμελιώδες αποτέλεσμα:

Θεώρημα IV.4.3.(iii).1. ***Μία συνθήκη ικανή και αναγκαία προκειμένου η μέθοδος συναπτών βημάτων να είναι συναφής προς την διαφορική εξίσωση είναι η επαλήθευση των εξισώσεων:***

$$a(1)=0 \text{ και } a'(1)=\beta(1).$$

Απόδειξη. Θα δείξουμε, κατ' αρχάς, ότι η συνθήκη είναι αναγκαία.

Προς τούτο, θα υποθέσουμε ότι η μέθοδος είναι συναφής, και θα δείξουμε ότι, τότε, αναγκαστικά, η συνθήκη του *Θεωρήματος* θα επαληθεύεται για οποιαδήποτε διαφορική εξίσωση. Ας θεωρήσουμε, λοιπόν, το πρόβλημα αρχικών τιμών:

$$y'(x)=0, \text{ με } y(0)=1,$$

του οποίου η μοναδική λύση είναι η συνάρτηση $y(x)=1$.

Εάν χρησιμοποιήσουμε τον ορισμό της συνάφειας, θα αποκομίσουμε την σχέση:

$$\lim_{h\to 0}\frac{1}{h}\lambda\sum_{i=0}^{k} a_i=0,$$

από την οποία έπεται, κατ' ανάγκη, η διαπίστωση ότι:

$$\sum_{i=0}^{k} a_i=a(1)=0.$$

Εάν, επί πλέον, θεωρήσουμε το πρόβλημα αρχικών τιμών:

$$y'(x)=1, \text{ με } y(0)=0,$$

του οποίου η μοναδική λύση είναι η συνάρτηση $y(x)=x$, τότε, χρησιμοποιώντας, και πάλι, τον ορισμό της συνάφειας λαμβάνουμε:

$$\sum_{i=1}^{k} i\, a_i - \sum_{i=0}^{k} \beta_i = a'(1) - \beta(1) = 0 .$$

Στην συνέχεια, θα δείξουμε ότι η συνθήκη του *Θεωρήματος* είναι ικανή, δηλαδή ότι η επαλήθευση των δύο ισοτήτων του *Θεωρήματος* αρκεί για να εγγυηθεί την συνάφεια της μεθόδου προς κάθε διαφορική εξίσωση (και όχι μόνο προς κάποια ειδική διαφορική εξίσωση).

Επειδή η y είναι διαφορίσιμη, παρατηρούμε ότι, σύμφωνα με το περιορισμένο ανάπτυγμα Taylor της y, ισχύει:

$$y(x_{n+i}) = y(x_n) + i\, h\, y'(x_n) + \mathrm{O}(h^2)$$

και:

$$y'(x_{n+i}) = y'(x_n) + \mathrm{O}(h) ,$$

οπότε, μεταφέροντας τις δύο σχέσεις αυτές μέσα στον τυποποιημένο ορισμό της συνάφειας, βλέπουμε πως:

$$\lim_{h\to 0} \max_{0\le n\le N} \left| \frac{y(x_n)}{h} \sum_{i=0}^{k} a_i + y'(x_n) \sum_{i=1}^{k} i\, a_i + \mathrm{O}(h) - y'(x_n) \sum_{i=0}^{k} \beta_i + \mathrm{O}(h) \right| = 0, \text{ ή}$$

ακόμη, πως:

$$\lim_{h\to 0} \max_{0\le n\le N} \left| \frac{y(x_n)}{h} \sum_{i=0}^{k} a_i + y'(x_n) \sum_{i=1}^{k} (i\, a_i - \beta_i) + \mathrm{O}(h) \right| = 0 .$$

Κατ' ισοδύναμη έκφραση:

$$\lim_{h\to 0} \max_{0\le n\le N} \left| \frac{y(x_n)}{h} .a(1) + y'(x_n).(a'(1) - \beta(1)) + \mathrm{O}(h) \right| = 0 ,$$

που φανερώνει πως η επαλήθευση των δύο εξισώσεων:

$$a(1) = 0 \text{ και } a'(1) = \beta(1)$$

είναι αρκετή για να εξασφαλίσει ότι η μέθοδος συναπτών βημάτων είναι συναφής προς την διαφορική εξίσωση. Η *Απόδειξη* του *Θεωρήματος* είναι ολοκληρωμένη. ■

IV.4.3.(iv). Θεωρητική Ευστάθεια Μεθόδου Έστωσαν y_n και z_n ($n = 0,1,\ldots,N$) οι αντίστοιχες λύσεις της εξίσωσης:

$$\sum_{i=0}^{k} a_i y_{n+1} = h \sum_{i=0}^{k} \beta_i \, F(x_{n+i}, y_{n+i}),$$

με τους πραγματικούς αριθμούς $y_0,\ldots,y_{k-1}$ δεδομένους και τυχαίους, και της εξίσωσης:

$$\sum_{i=0}^{k} a_i \, z_{n+i} = h\,[\sum_{i=0}^{k} \beta_i \, F(x_{n+i}, y_{n+i}) + \varepsilon_n]$$

με τους πραγματικούς αριθμούς $z_0,\ldots,z_{k-1}$ δεδομένους και τυχαίους, όπου:

$$x_0 = a \text{ και } x_{n+1} = x_n + h.$$

Λέμε ότι η μέθοδος συναπτών βημάτων είναι ***ευσταθής***, εάν υπάρχουν δύο σταθερές M_1 και M_2, ανεξάρτητες από το βήμα h, που είναι τέτοιες ώστε:

$$\max_{0\le n\le N} |y_n - z_n| \le M_1 \max_{0\le i\le k-1} |y_i - z_i| + M_2 \max_{0\le n\le N} |\varepsilon_n|.$$

Η έννοια αυτή της ευστάθειας θα ονομάζεται ***θεωρητική ευστάθεια***, σε αντίθεση με την έννοια της αριθμητικής ευστάθειας ή $A-$ ευστάθειας, που θα συζητήσουμε αργότερα.

Το θεμελιώδες αποτέλεσμα, που μπορούμε να συνάγουμε με την βοήθεια της Θεωρίας των Εξισώσεων Διαφορών της *Υποπαραγράφου IV.4.3.(i)*, είναι το επόμενο:

Θεώρημα IV.4.3.(vi).1. ***Μία συνθήκη ικανή και αναγκαία προκειμένου η μέθοδος συναπτών βημάτων να είναι ευσταθής είναι η ταυτόχρονη πλήρωση των ακόλουθων ιδιοτήτων:***

- ***όλες οι ρίζες του πολυωνύμου*** $a(s)$ ***να έχουν απόλυτες τιμές μικρότερες ή ίσες του αριθμού*** 1
- ***όλες οι ρίζες του πολυωνύμου*** $a(s)$***, που έχουν απόλυτες τιμές ίσες με ένα, είναι απλές ρίζες.***

Απόδειξη. Θα δείξουμε πρώτα την αναγκαιότητα της συνθήκης.

Προς τούτο, ας θεωρήσουμε το πρόβλημα Cauchy:

$$y'(x) = 0\,,\ \text{με}\ y(0) = 0\,,$$

του οποίου η μοναδική λύση είναι η συνάρτηση $y(x) = 0$. Τότε, θα είναι:

$$\sum_{i=0}^{k} a_i\ y_{n+i} = 0 \ \text{ και } \sum_{i=0}^{k} a_i\ z_{n+i} = h\ \varepsilon_n\ ,$$

από όπου έπεται ότι:

$$\sum_{i=0}^{k} a_i\ (z_{n+i} - y_{n+i}) = h\ \varepsilon_n\,,$$

ή ακόμα, ισοδύναμα και εφόσον θέσουμε $u_n = z_n - y_n$, ότι:

$$\sum_{i=0}^{k} a_i\ u_{n+i} = h\ \varepsilon_n\ .$$

Όπως τονίστηκε στο τέλος της *Υποπαραγράφου IV.4.3.(i)*, η γενική λύση αυτής της εξίσωσης διαφορών ισούται με το άθροισμα της γενικής λύσης της αντίστοιχης ομογενούς εξίσωσης και μίας οποιασδήποτε μερικής λύσης της.

Επειδή η μέθοδος είναι συναφής, το μέγεθος $|u_n|$ οφείλει να είναι φραγμένο. Κατά συνέπεια, όλες οι ρίζες του πολυωνύμου $a(s)$ οφείλουν να έχουν απόλυτες τιμές που είναι το πολύ ίσες με ένα και όσες από αυτές έχουν απόλυτη τιμή ακριβώς ίση με ένα πρέπει να είναι απλές ρίζες,.

Η απόδειξη της ικανότητας της συνθήκης του *Θεωρήματος* είναι πολύ πιο μακροσκελής, γιατί, τότε, η συνάφεια θα πρέπει να επαληθευτεί για κάθε διαφορική εξίσωση, και όχι μόνο για μία ειδική διαφορική εξίσωση.

Για λόγους εξοικονόμησης χώρου, θα αποφύγουμε να δώσουμε αυτήν την απόδειξη. Όμως, ο ενδιαφερόμενος αναγνώστης μπορεί να αποταθεί στις βιβλιογραφικές παραπομπές [28] και [54]. Εκεί θα βρει μία αυστηρή απόδειξη, η οποία πραγματοποιείται με την βοήθεια διαδοχικών ανισοτικών εκτιμήσεων, που πηγάζουν από την κεντρική ισότητα:

$$\sum_{i=0}^{k} a_i\ (z_{n+i} - y_{n+i}) = h \sum_{i=0}^{k} \beta_i\ [F(x_{n+i}, z_{n+i}) - F(x_{n+i}, y_{n+i})] + h\ \varepsilon_n\ . \blacksquare$$

Παρατήρηση IV.4.3.(iv).2. Εάν η μέθοδος είναι συναφής, τότε ο αριθμός 1 είναι πάντοτε ρίζα του πολυωνύμου $a(s)$, γιατί, τότε, $a(1)=0$. Σύμφωνα με το προηγούμενο *Θεώρημα*, προκειμένου η μέθοδος να είναι, επί πλέον, ευσταθής, πρέπει η ρίζα αυτή να είναι απλή. ■

IV.4.3.(v). Σύγκλιση Μεθόδου

Λέμε ότι η μέθοδος συναπτών βημάτων είναι ***συγκλίνουσα*** (ή απλά ότι ***συγκλίνει***) εάν:

$$\lim_{h\to 0} \max_{0\le n\le N} | y_n - y(x_n) | = 0,$$

όταν $\lim_{h\to 0} y_i = 0$ για κάθε $i = 0,1,\ldots,k-1$.

Σύμφωνα με τον ορισμό αυτόν, η προσεγγιστική λύση οφείλει να συγκλίνει προς την λύση της διαφορικής εξίσωσης σε κάθε σημείο του διαστήματος ολοκλήρωσης, όταν οι k αρχικές τιμές $y_0, y_1, \ldots, y_{k-1}$ τείνουν προς την αρχική συνθήκη της υπό ολοκλήρωση διαφορικής εξίσωσης.

Όπως για την περίπτωση των μεθόδων ξεχωριστών βημάτων, ισχύει το:

Θεώρημα IV.4.3.(v).1. ***Εάν η μέθοδος συναπτών βημάτων είναι ευσταθής και συναφής, τότε είναι επί πλέον και συγκλίνουσα.***

Απόδειξη. Είναι φανερό ότι η λύση $y(x)$ της διαφορικής εξίσωσης επαληθεύει τις ακόλουθες δύο σχέσεις:

$$\sum_{i=0}^{k} a_i \; y(x_{n+i}) = h \sum_{i=0}^{k} \beta_i \; F(x_{n+i}, y(x_{n+i})) + h \; \varepsilon_n$$

και:

$$y(a) = y_0 .$$

Εάν η μέθοδος συναπτών βημάτων είναι συναφής, τότε σύμφωνα με όσα αναφέρθηκαν στην *Υποπαράγραφο IV.4.3.(iii)*, πρέπει να είναι:

$$\lim_{h\to 0} \max_{0\le n\le N} |\varepsilon_n| = 0 .$$

Επειδή η μέθοδος έχει υποτεθεί ότι είναι ευσταθής, υπάρχουν, σύμφωνα με τον ορισμό της ευστάθειας που δόθηκε στην *Υποπαράγραφο IV.4.3.(iv)*, δύο σταθερές M_1 και M_2, που είναι ανεξάρτητες του βήματος h και τέτοιες ώστε:

$$\max_{0\le n\le N} |y_n - z_n| \le M_1 \max_{0\le i\le k-1} |y_i - z_i| + M_2 \max_{0\le n\le N} |\varepsilon_n| .$$

Η ιδιότητα της σύγκλισης έπεται τώρα αμέσως, εφόσον αφήσουμε το h να τείνει απεριόριστα προς το 0 και χρησιμοποιώντας το γεγονός ότι:

$$\lim_{h\to 0} y_i = y_0 = y(a), \text{ για κάθε } i = 0,1,\ldots,k-1 . \blacksquare$$

Ισχύει επίσης το αντίστροφο αποτέλεσμα ([44]):

Θεώρημα IV.4.3.(v).2. ***Εάν η μέθοδος συναπτών βημάτων είναι συγκλίνουσα, τότε είναι ευσταθής και συναφής.***

Απόδειξη. Θα δείξουμε πρώτα ότι η σύγκλιση συνεπάγεται την ευστάθεια.

Προς τούτο, ας θεωρήσουμε το πρόβλημα Cauchy:

$$y'(x) = 0 ,\ y(0) = 0 ,$$

του οποίου η μοναδική λύση είναι η συνάρτηση $y(x)=0$. Τότε, η μέθοδος συναπτών βημάτων συγχωνεύεται στην ακόλουθη σχέση:

$$\sum_{i=0}^{k} a_i \; y_{n+i} = 0,$$

και, από το γεγονός ότι $\lim_{h\to 0} y_n = 0$ (επειδή η μέθοδος συγκλίνει), έπεται πως όλες οι ρίζες του πολυωνύμου $a(s)$ έχουν απόλυτες τιμές το πολύ ίσες με ένα και, ακόμη, πως όσες από αυτές τις ρίζες έχουν απόλυτες τιμές ίσες με ένα είναι, υποχρεωτικά, απλές ρίζες.

Απομένει τώρα να δειχθεί ότι η σύγκλιση συνεπάγεται την συνάφεια.

Προς τούτο, ας θεωρήσουμε, κατ' αρχάς, το πρόβλημα Cauchy:

$$y'(x)=0 \;,\; \mathrm{y}(0)=1,$$

του οποίου η μοναδική λύση είναι η συνάρτηση $y(x)=1$. Τότε, η μέθοδος συνοψίζεται στην ισότητα:

$$\sum_{i=0}^{k} a_i \; y_{n+i} = 0,$$

κι επειδή $\lim_{h\to 0} y_n = 1$ συμπεραίνουμε ότι:

$$\sum_{i=0}^{k} a_i = a(1) = 0.$$

Εάν πάρουμε τέλος το πρόβλημα:

$$y'(x)=1 \;,\; \mathrm{y}(0)=0,$$

του οποίου η μοναδική λύση είναι η συνάρτηση $y(x)=x$, τότε η μέθοδος γίνεται:

$$\sum_{i=0}^{k} a_i \; y_{n+i} = h\;\beta(1),$$

και καθώς $\lim_{n\to 0}\left|y_n - n\,h\right| = 0$ (επειδή η μέθοδος είναι συγκλίνουσα), βλέπουμε ότι:

$$a'(1) = \beta(1),$$

που ολοκληρώνει την *Απόδειξη* του *Θεωρήματος*. ■

IV.4.3.(vi). Τάξη Μεθόδου Λέμε ότι ***η μέθοδος συναπτών βημάτων είναι τάξης p***, εάν:

$$\max_{0\le n\le N}\left|\frac{1}{h}\sum_{i=0}^{k} a_i\ y(x_{n+i})-\sum_{i=0}^{k}\beta_i\ y'(x_{n+i})\right| = \mathrm{O}(h^p)\,.$$

Σχετικά παρουσιάζουμε το ακόλουθο αποτέλεσμα, το οποίο θα μας φανεί χρήσιμο στην συνέχεια:

Θεώρημα IV.4.3.(vi).1. ***Εάν η μέθοδος συναπτών βημάτων είναι ευσταθής, τάξης p, και εάν:***

$$\max_{0\le i\le k-1}\left|y_i - y(x_i)\right| = \mathrm{O}(h^p),$$

τότε:

$$\max_{0\le n\le N}\left|y_n - y(x_n)\right| = \mathrm{O}(h^p).$$

Απόδειξη. Κατ' αρχάς, επισημαίνουμε ότι ισχύει:

$$\frac{1}{h}\sum_{i=0}^{k} a_i\ y(x_{n+i})-\sum_{i=0}^{k}\beta_i\ y'(x_{n+i}) = \varepsilon_n\,.$$

Εάν η μέθοδος είναι τάξης *p*, τότε, σύμφωνα με τον πιο πάνω Ορισμό, έχουμε:

$$\max_{0\le n\le N}\left|\varepsilon_n\right| = \mathrm{O}(h^p)\,.$$

Από την άλλη πλευρά, εάν η μέθοδος είναι ευσταθής, τότε υπάρχουν δύο σταθερές M_1 και M_2, που είναι ανεξάρτητες από το βήμα h και τέτοιες ώστε:

$$\max_{0\le n\le N}\left|y_n - y(x_n)\right| \le M_1\ \max_{0\le i\le k-1}\left|y_i - y(x_i)\right| + M_2\ \max_{0\le n\le N}\left|\varepsilon_n\right|.$$

Το συμπέρασμα του *Θεωρήματος* έπεται τώρα αμέσως, εφόσον δεχθούμε ότι:

$$\max_{0\le i\le k-1}\left|y_i - y(x_i)\right| = \mathrm{O}(h^p)\,.\ \blacksquare$$

Παρατήρηση IV.4.3.(vi).2. Το *Θεώρημα* αυτό υποδεικνύει ότι οι ποσότητες y_i $(i = 1,2,\ldots,k-1)$ πρέπει να υπολογισθούν χρησιμοποιώντας μία μέθοδο τουλάχιστον τάξης *p* προκειμένου να εξασφαλισθεί ότι:

$$\max_{0\le n\le N}\left|y_n - y(x_n)\right| = \mathrm{O}(h^p)\,. \blacksquare$$

Θα εξετάσουμε τώρα με μεγαλύτερη επιμέλεια την βασική σχέση του παραπάνω *Ορισμού* της τάξης μίας μεθόδου συναπτών βημάτων.

Αν υποθέσουμε ότι η λύση *y(x)* είναι μία συνάρτηση αρκετές φορές διαφορίσιμη επί του $\mathbb{R}$, και εάν χρησιμοποιήσουμε το ανάπτυγμά της σε σειρά Taylor, τότε διαμορφώνεται η εξής σχέση:

$$\frac{1}{h}\sum_{i=0}^{k} a_i\; y(x_{n+i}) - \sum_{i=0}^{k}\beta_i\; y'(x_{n+i})$$

$$= \frac{A_0}{h}\, y(x_n) + A_1\; y'(x_n) + A_2\; h\; y''(x_n) + \cdots + A_p\; h^{p-1}\; y^{(p)}(x_n) + \mathrm{O}(h^p)\,,$$

όπου:

$$A_0 := \sum_{i=0}^{k} a_i\,,$$

$$A_1 := \sum_{i=0}^{k}(i\; a_i - \beta_i)\,,$$

και:

$$A_q := \frac{1}{q!}\sum_{i=0}^{k} i^q\; a_i - \frac{1}{(q-1)!}\sum_{i=0}^{k} i^{q-1}\;\beta_i \qquad (q = 2,\ldots,p)\,.$$

Επομένως, η μέθοδος συναπτών βημάτων είναι τουλάχιστον τάξης p, εάν:

$$A_0 = A_1 = \ldots = A_p = 0\,.$$

Παρατηρήστε ότι η συνθήκη $A_0 = 0$ δεν είναι άλλη από την $a(1) = 0$, ενώ η συνθήκη $A_1 = 0$ δεν είναι άλλη από την $a'(1) = \beta(1)$. Για τον λόγο αυτό, βεβαιωνόμαστε για την αλήθεια του παρακάτω *Θεωρήματος*:

Θεώρημα IV.4.3.(vi).3.. ***Μία συνθήκη, ικανή και αναγκαία, προκειμένου μία μέθοδος συναπτών βημάτων να είναι συναφής επικεντρώνεται στην ιδιότητα της μεθόδου να έχει τάξη τουλάχιστον ίση με ένα.*** □

IV.4.3.(vii). Οι Μέθοδοι Adams

Στον απόηχο των αποτελεσμάτων που παρατέθηκαν στις προηγούμενες *Υποπαραγράφους*, βλέπουμε ότι το πρόβλημα που εγείρεται τώρα συνοψίζεται στην ακόλουθη διατύπωση:

Δοθέντος ενός σταθερού k, να βρεθούν πολυώνυμα $a(s)$ και $\beta(s)$, τέτοια ώστε:

- ***το πολυώνυμο $a(s)$ να είναι ευσταθές, και***
- ***η αντίστοιχη μέθοδος συναπτών βημάτων να είναι τάξης $p \geq 1$, όσο το δυνατόν πιο υψηλής.***

Γνωρίζουμε ότι εάν η συνάρτηση $y(x)$ είναι πολυωνυμική βαθμού p, τότε το ανάπτυγμά της σε σειρά Taylor καταλήγει σ' ένα όρο που περιέχει την ποσότητα h^p. Κατά συνέπεια, τα πολυώνυμα $a(s)$ και $\beta(s)$ πρέπει να επιλεγούν έτσι ώστε το $a(s)$ να είναι ευσταθές και η αντίστοιχη μέθοδος συναπτών βημάτων να είναι ακριβής για όλα τα πολυώνυμα που έχουν τον μεγαλύτερο δυνατό βαθμό.

Οι ***μέθοδοι Adams*** κατασκευάστηκαν σύννομα προς αυτήν την αρχή. Η βασική τους ιδέα περιγράφεται στην συνέχεια. Προφανώς, ισχύει:

$$y(x_q) - y(x_j) = \int_{x_j}^{x_q} F(t, y(t))dt \qquad (x_q \geq x_j),$$

όπου τα σημεία x_q και x_j ανήκουν στο διάστημα ολοκλήρωσης $[a, \beta]$.

Βεβαίως, δεν μπορούμε να γνωρίζουμε τον τρόπο υπολογισμού του ολοκληρώματος, που βρίσκεται στο δεξιό μέλος της προηγούμενης ισότητας, γιατί η συνάρτηση y είναι άγνωστη. **Οι *μέθοδοι Adams συνίστανται στην αντικατάσταση της απεικόνισης* F *με το πολυώνυμο παρεμβολής της, που λαμβάνει τις τιμές:***

$$F_n, \ldots, F_{n+k}$$

***στα διαδοχικά ισαπέχοντα σημεία*:**

$$x_n, \ldots, x_{n+k}.$$

Με άλλα λόγια, το παραπάνω ολοκλήρωμα υπολογίζεται προσεγγιστικά χρησιμοποιώντας μαθηματικούς *Τύπους* αριθμητικού τετραγωνισμού, ανάλογους με εκείνους που εξετάσθηκαν στο *Κεφάλαιο IV.2*.

Έτσι, ***μία μέθοδος συναπτών βημάτων, που βασίζεται στην ιδέα Adams,*** είναι της ακόλουθης γενικής μορφής:

$$y_q - y_j = h \sum_{i=0}^{k} \beta_i \, F_{n+i} \qquad (q \geq j).$$

Παρατηρήστε, αμέσως, ότι $a(1) = 0$ και το πολυώνυμο $a(s)$ είναι ευσταθές. Προκειμένου, επί πλέον, μία τέτοια μέθοδος να είναι συναφής και συγκλίνουσα, πρέπει οι συντελεστές β_i να ανταποκρίνονται στη σχέση $a'(1) = \beta(1)$, δηλαδή πρέπει να επαληθεύουν την ισότητα:

$$q - j = \sum_{i=0}^{k} \beta_i .$$

Οι σχετικές μέθοδοι, που επιτυγχάνονται κατ' αυτόν τον τρόπο, διαφοροποιούνται από τις εκάστοτε θέσεις των αριθμών x_j και x_q ως προς τα σημεία παρεμβολής $x_n, \ldots, x_{n+k}$. Αυτές οι διάφορες μέθοδοι φέρουν τις ακόλουθες ονομασίες:

- ***Μέθοδοι Adams – Bashforth, όταν:***

$$x_j = x_{n+k} \text{ και } x_q = x_{n+k+1}.$$

- ***Μέθοδοι Adams – Moulton, όταν:***

$$x_j = x_{n+k-1} \text{ και } x_q = x_{n+k}.$$

- ***Μέθοδοι Nyström, όταν:***

$$x_j = x_{n+k-1} \text{ και } x_q = x_{n+k+1}.$$

- ***Μέθοδοι Milne – Simpson, όταν:***

$$x_j = x_{n+k-2} \text{ και } x_q = x_{n+k}.$$

Ο υπολογισμός των συντελεστών β_i μπορεί να πραγματοποιηθεί σύμφωνα με δύο εκδοχές:

ή αξιώνουμε την ακρίβεια της μεθόδου συναπτών βημάτων, για όλα τα πολυώνυμα με βαθμούς όσο το δυνατόν πιο μεγάλους, οπότε επιλύουμε το αντίστοιχο σύστημα γραμμικών εξισώσεων που προκύπτει. (Για παράδειγμα, προκειμένου περί του *Τύπου*:

$$y_{n+2} - y_{n+1} = h\,\beta_0\,f_n + h\,\beta_1\,f_{n+1},$$

εάν πάρουμε $F(x,y)=0$, με αντίστοιχη λύση την συνάρτηση $y(x)=1$, βλέπουμε ότι αυτή η διαφορική εξίσωση ολοκληρώνεται με ακρίβεια από την μέθοδο συναπτών βημάτων που περιγράφεται από τον παραπάνω τύπο, ενώ, εάν πάρουμε $F(x,y)=1$, με αντίστοιχη λύση την συνάρτηση $y(x)=x$, βλέπουμε ότι ο παραπάνω τύπος γίνεται:

$$x_{n+2} - x_{n+1} = h\,(\beta_0 + \beta_1) = h,$$

ενώ, τέλος, εάν πάρουμε $F(x,y)=x$, με αντίστοιχη λύση την συνάρτηση $y(x)=x^2/2$, βλέπουμε ότι ο ίδιος *Τύπος* γίνεται:

$$\frac{x_{n+2}^2}{2} - \frac{x_{n+1}^2}{2} = h\,(\beta_0\,x_n + \beta_1\,x_{n+1}).$$

Στο σύστημα των γραμμικών εξισώσεων, ως προς το β_0 και β_1:

$$x_{n+2} - x_{n+1} = h\,(\beta_0 + \beta_1) = h$$

$$\frac{x_{n+2}^2}{2} - \frac{x_{n+1}^2}{2} = h\,(\beta_0\; x_n + \beta_1\; x_{n+1}),$$

εάν λάβουμε $x_n = 0$, βρίσκουμε πως πρέπει να είναι:

$$\beta_1 = \frac{3}{2} \quad \text{και (συνεπώς)} \quad \beta_0 = -\frac{1}{2}\ .)$$

ή αντικαθιστώντας την συνάρτηση $F(t, y(t))$ με το πολυώνυμο Newton που την παρεμβάλλει στα σημεία $x_n, \ldots, x_{n+k}$, ολοκληρώνουμε μεταξύ των αριθμών x_j και x_q. Ο ουσιαστικός προσδιορισμός των συντελεστών β_i πραγματοποιείται στην συνέχεια, με επαγωγικό τρόπο και χρησιμοποιώντας τις συναρτήσεις γεννήτορες.

Σε ό,τι ακολουθήσει, θα αναπτύξουμε την διαδικαστική εξέλιξη των πιο πάνω ιδεών, επικεντρωνόμενοι στην ενδεικτική περίπτωση των μεθόδων Adams-Bashforth. Επειδή, τότε, τα σημεία παρεμβολής είναι διαδοχικά ισαπέχοντα, θα επικαλεσθούμε το πολυώνυμο Newton, του οποίου η αναλυτική έκφραση διατυπώθηκε στο *Θεώρημα IV.1.5.5* και στο *Θεώρημα IV.1.5.7*.

Έτσι, εάν στην έκφραση του *Θεωρήματος IV.1.5.7* αντικαταστήσουμε το σύμβολο n με το σύμβολο k, και το σύμβολο 0 με το σύμβολο n, και εάν πραγματοποιήσουμε την αλλαγή μεταβλητής $u = (x - x_{n+k})/h$, τότε, μετά από μία αναδιάταξη των όρων, θα βρούμε την ακόλουθη αποδοτική παράσταση του πολυωνύμου παρεμβολής Newton:

$$p_k(u) = \sum_{i=0}^{k} (-1)^i\, C_{-u}^i\; \nabla^i F_{n+k}\,,$$

όπου έγινε χρήση του συμβατικού συμβολισμού:

$$C_w^i = \frac{w\,(w-1)\cdots(w-i+1)}{i!} \qquad (\forall w \in \mathbb{R})\,.$$

Παρατηρήστε ότι ισχύει:

$$p_k(-i) = F_i, \text{ για κάθε } i = 0,1,\ldots,k.$$

Επομένως, οι *Τύποι* Adams-Bashforth διαμορφώνονται όπως ακολούθως:

$$\begin{aligned} y_{n+k+1} - y_{n+k} &= \int_{x_{n+k}}^{x_{n+k+1}} p_k(u)du \\ &= h\int_0^1 \sum_{i=0}^{k} (-1)^i C_{-u}^i \ \nabla^i F_{n+k} du \\ &= h\sum_{i=0}^{k} \left\{ \int_0^1 (-1)^i C_{-u}^i du \right\} \nabla^i F_{n+k} \\ &= h\sum_{i=0}^{k} \gamma_i \ \nabla^i F_{n+k}, \end{aligned}$$

όπου θέσαμε:

$$\gamma_i = (-1)^i \int_0^1 C_{-u}^i du \ (i = 0,1,\ldots).$$

Οι συντελεστές γ_i υπολογίζονται χρησιμοποιώντας μαθηματική επαγωγή και με την βοήθεια της μεθόδου των συναρτήσεων γεννητόρων. Ας σημειωθεί πως η προσφυγή στην διαχείριση των συναρτήσεων γεννητόρων είναι μία γενική και συνηθισμένη διεξοδική αντιμετώπιση, που έχει και ένα ιδιαίτερο πρακτικό αντίκρυσμα στην διαδικασία επίλυσης ορισμένων εξισώσεων πεπερασμένων διαφορών.

Ονομάζουμε ***συνάρτηση γεννήτορα*** της ακολουθίας αριθμών $(\gamma_0, \gamma_1, \gamma_2 \ldots)$ την συνάρτηση Γ που ορίζεται από την σειρά:

$$\Gamma(t) = \sum_{i=0}^{\infty} \gamma_i \ t^i \qquad (t \in \mathbb{C}).$$

Θα υποθέτουμε ότι:

$$|t| < \min\{R,1\},$$

όπου R είναι η ακτίνα σύγκλισης αυτής της δυναμοσειράς.

Επειδή:

$$\Gamma(t) = \sum_{i=0}^{\infty} (-t)^i \int_0^1 C_{-u}^i \, du \,,$$

η ομοιόμορφη σύγκλιση της σειράς επιτρέπει την αμοιβαία μετάθεση των συμβόλων της ολοκλήρωσης και της άθροισης, έτσι ώστε να ισχύει:

$$\begin{aligned}\Gamma(t) &= \int_0^1 \sum_{i=0}^{\infty} (-t)^i \; C_{-u}^i \; du \\ &= \int_0^1 (1-t)^{-u} \; du \\ &= \int_0^1 e^{-u\, Log(1-t)} \, du \\ &= -\frac{t}{(1-t)\, Log(1-t)} \,.\end{aligned}$$

Με άλλα λόγια, δείξαμε ότι:

$$-\frac{1}{t} Log(1-t)\,\Gamma(t) = \frac{1}{1-t} \qquad (|t| < \min\{R,1\}) \,.$$

Καθώς όμως είναι:

$$-\frac{1}{t} Log(1-t) = \sum_{i=0}^{\infty} \frac{t^i}{1+i} \,,$$

$$\Gamma(t) = \sum_{i=0}^{\infty} \gamma_i \; t^i \,,$$

και:

$$\frac{1}{1-t} = \sum_{i=0}^{\infty} t^i \,,$$

η πιο πάνω ισότητα γράφεται:

$$\left(\sum_{i=0}^{\infty} \frac{t^i}{1+t}\right) \left(\sum_{i=0}^{\infty} \gamma_i \; t^i\right) = \sum_{i=0}^{\infty} t^i \quad (|t| < \min\{R,1\}) \,.$$

Εάν τώρα εξισώσουμε, μετά από πράξεις, τους ομοβάθμιους (ως προς t) όρους, θα διαπιστώσουμε πως οι συντελεστές γ_i πρέπει να υπακούουν στις ακόλουθες σχέσεις:

$$\gamma_0 = 1$$

$$\gamma_1 + \frac{1}{2}\gamma_0 = 1$$

.....

$$\gamma_i + \frac{1}{2}\gamma_{i-1} + \frac{1}{3}\gamma_{i-2} + \ldots + \frac{1}{i+1}\gamma_0 = 1$$

.....

από τις οποίες έπονται οι τιμές των συντελεστών γ_i. Για παράδειγμα, έχουμε:

i	0	1	2	3	4	5
γ_i	1	$\frac{1}{2}$	$\frac{5}{12}$	$\frac{3}{8}$	$\frac{251}{720}$	$\frac{95}{288}$

- ***Καταλήγουμε λοιπόν στις εξής εκφράσεις για τις μεθόδους συναπτών βημάτων Adams-Bashforth:***

  - $y_{n+k+1} = y_{n+k} + h\,F_{n+k}$, ***που είναι τάξης*** 1 ***(:μέθοδος Euler),***
  - $y_{n+k+1} = y_{n+k} + \frac{h}{2}\,(3\,F_{n+k} - F_{n+k-1})$, ***που είναι τάξης*** 2,
  - $y_{n+k+1} = y_{n+k} + \frac{h}{12}\,(23\,F_{n+k} - 16\,F_{n+k-1} + 5\,F_{n+k-2})$, ***που είναι τάξης*** 3,
  - $y_{n+k+1} = y_{n+k} + \frac{h}{24}\,(55\,F_{n+k} - 59\,F_{n+k-1} + 37\,F_{n+k-2} - 9\,F_{n+k-3})$, ***με τάξη*** 4.

Ανάλογα, μπορούν να βρεθούν και οι εκφράσεις των μαθηματικών τύπων των υπολοίπων κλασσικών μεθόδων συναπτών βημάτων.

- Έτσι, ***για τις μεθόδους Adams-Moulton, έχουμε την ακόλουθη γενική έκφραση:***

$$y_{n+k} = y_{n+k-1} + h \sum_{i=0}^{k} \gamma_i^* \nabla^i F_{n+k} ,$$

όπου:

$$\gamma_i^* = (-1)^i \int_0^1 C_{-u}^i \, du .$$

Η συνάρτηση γεννήτορας $\Gamma^*(t)$ ***των μεθόδων Adams-Moulton δίνεται από τον Τύπο:***

$$.\Gamma^*(t) = -\frac{t}{Log(1-t)}$$

- ***Για τις μεθόδους Nyström, έχουμε την ακόλουθη γενική μαθηματική διατύπωση:***

$$y_{n+k+1} = y_{n+k-1} + h \sum_{i=0}^{k} X_i \nabla^i F_{n+k} ,$$

όπου:

$$X_i = (-1)^i \int_{-1}^{1} \binom{-u}{i} du .$$

Η συνάρτηση γεννήτορας $X(t)$ ***των μεθόδων Nyström δίνεται από τον Τύπο:***

$$X(t) = -\frac{2-t}{1-t} \frac{t}{Log(1-t)} .$$

- ***Τέλος, για τις μεθόδους Milne-Simpson, ισχύει η εξής περιγραφή:***

$$y_{n+k} = y_{n+k-2} + h \sum_{i=0}^{k} X_i^* \, \nabla^i F_{n+k} \,,$$

όπου:

$$X_i^* = (-1)^i \int_{-2}^{0} C_{-u}^i \, du \,,$$

ενώ η συνάρτηση γεννήτορας $X^*(t)$ ***αυτών των μεθόδων δίνεται από τον Τύπο:***

$$X^*(t) = -(2-t)\frac{t}{Log(1-t)}$$

Ενδεικτικά, αναφέρουμε κάποια μεμονωμένα παραδείγματα.

Η μέθοδος Adams-Moulton τάξης 4 παρουσιάζεται από την έκφραση:

$$y_{n+k} = y_{n+k-1} + \frac{h}{24}(9\,F_{n+k} + 19\,F_{n+k-1} - 5\,F_{n+k-2} + F_{n+k-3}),$$

ενώ η μέθοδος Nyström τάξης 2 έχει την μορφή:

$$y_{n+k+1} = y_{n+k-1} + 2\,h\,F_{n+k}\,.$$

Η μέθοδος Milne-Simpson τάξης 4 είναι:

$$y_{n+k} = y_{n+k-2} + \frac{h}{3}(F_{n+k} + 4\,F_{n+k-1} + F_{n+k-2}),$$

που συμπίπτει με την έκφραση της μεθόδου ολοκλήρωσης Simpson (:Newton – Côtes, με $n = 2$) για τον αριθμητικό υπολογισμό ενός ορισμένου ολοκληρώματος.

IV.4.3.(viii). Οι Μέθοδοι Πρόβλεψης-Διόρθωσης Εάν η μέθοδος συναπτών βημάτων είναι πεπλεγμένη (με $\beta_k \neq 0$), τότε η τιμή y_{n+k} είναι ακριβής λύση της πεπλεγμένης εξίσωσης. Ο υπολογισμός της ποσότητας y_{n+k} γίνεται, όπως ήδη είπαμε, με επαναλήψεις.

Τίθενται, ως εκ τούτου,λοιπόν δύο προβλήματα:

το πρώτο αναφέρεται στην επιλογή της «αρχικής τιμής» των επαναλήψεων,

ενώ

το δεύτερο αφορά στον αριθμό των προς πραγματοποίηση επαναλήψεων.

Στην πράξη, αυτά τα δύο προβλήματα αντιμετωπίζονται με την βοήθεια μίας ***μεθόδου πρόβλεψης-διόρθωσης***.

Σύμφωνα με μία τέτοια μέθοδο, η «αρχική τιμή» των επαναλήψεων προέρχεται από μία σαφή μέθοδο συναπτών βημάτων, που λέγεται ***πρόβλεψη*** (με άλλα λόγια, προβλέπουμε την τιμή της ποσότητας y_{n+k}). Στην συνέχεια, πραγματοποιούμε μόνον μία επανάληψη, χρησιμοποιώντας την πεπλεγμένη μέθοδο συναπτών βημάτων, που ονομάζεται ***διόρθωση*** (με άλλα λόγια, διορθώνουμε την τιμή της ποσότητας y_{n+k}, που προέκυψε από την πρόβλεψη).

Έχουμε δηλαδή το ακόλουθο σχήμα:

$$\left\{\begin{array}{ll} a_k^* \; y_{n+k}^* + \sum_{i=0}^{k-1} a_i^* \; y_{n+i} = h \sum_{i=0}^{k-1} \beta_i^* \; F_{n+i} & (: \text{πρόβλεψη}) \\ \sum_{i=0}^{k} a_i \; y_{n+i} = h\,[\beta_k \; F(x_{n+k}^*, y_{n+k}) + \sum_{i=0}^{k-1} \beta_i \; F_{n+i}] & (: \text{διόρθωση}) \end{array}\right\}$$

Θα υποθέσουμε ότι η πρόβλεψη είναι τάξης q και ότι η διόρθωση είναι τάξης p, και θα εξετάσουμε τις προϋποθέσεις, κάτω από τις οποίες ισχύει:

$$\max_{0 \le n \le N} \left| y_n - y(x_n) \right| = \mathrm{O}(h^p),$$

και οι οποίες μπορούν να εξασφαλίσουν την επάρκεια μίας μόνον επανάληψης για την διόρθωση. Σχετικά, είμαστε σε θέση να παραθέσουμε το επόμενο αποτέλεσμα:

Θεώρημα IV.4.3.(viii).1.. ***Εάν η πρόβλεψη είναι τάξης q και η διόρθωση τάξης p, τότε μία ικανή συνθήκη προκειμένου να συμβαίνει:***

$$\max_{0\le n\le N}\left|y_n - y(x_n)\right| = \mathrm{O}(h^p),$$

μετά από την πραγματοποίηση μίας μόνον επανάληψης κατά το στάδιο της διόρθωσης, είναι η ισχύς της παρακάτω σχέσης:

$$q = p-1$$

(και φυσικά, εφ' όσον οι ποσότητες $y_0,\ldots,y_{k-1}$ έχουν υπολογισθεί χρησιμοποιώντας μία μέθοδο που είναι τουλάχιστον τάξης p).

Απόδειξη. Σύμφωνα με όσα προηγήθηκαν, εάν η πρόβλεψη είναι τάξης q τότε:

$$\left|y^*_{n+k} - y(x_{n+k})\right| = \mathrm{O}(h^q),$$

ενώ εάν η διόρθωση είναι τάξης p, τότε:

$$\sum_{i=0}^{k} a_i\ (y_{n+i} - y(x_{n+i}))$$
$$= h\,[\beta_k\ F(x_{n+k}, y^*_{n+k}) - \beta_k\ y'(x_{n+k})] + h\sum_{i=0}^{k-1}\beta_i\ [F_{n+i} - y'(x_{n+i})] + \mathrm{O}(h^{p+1}).$$

Κατά συνέπεια, και επικαλούμενοι το γεγονός ότι η F ικανοποιεί μία συνθήκη Lipschitz, μπορούμε να βεβαιωθούμε πως:

$$|a_k|\left|y_{n+k} - y(x_{n+k})\right| \le \mathrm{O}(h^p) + h\,\mathrm{O}(h^q) + h.\mathrm{O}(h^p) + \mathrm{O}(h^{p+1}),$$

και, μετά απ' αυτό, να καταλήξουμε στο συμπέρασμα πως μία ικανή συνθήκη προκειμένου να συμβαίνει:

$$\max_{0\le n\le N}\left|y_n - y(x_n)\right| = \mathrm{O}(h^p)$$

είναι η ισχύς της ισότητας:

$$q = p-1.$$

Η συναγωγή του συμπεράσματος αυτού ολοκληρώνει την *Απόδειξη* του *Θεωρήματος*.■

Παρατήρηση IV.4.3.(viii).2. Είναι δυνατή η συμπληρωματική αποκομιδή περισσότερων πληροφοριών, σχετικά με το αποτέλεσμα αυτού του *Θεωρήματος*. Οι πληροφορίες αυτές αφορούν την ασυμπτωτική συμπεριφορά του σφάλματος, και αναφέρονται επιγραμματικά αμέσως τώρα. Για τις λεπτομέρειες και τις επεξηγήσεις, που παραλείπονται, ο ενδιαφερόμενος αναγνώστης μπορεί να συμβουλευτεί την βιβλιογραφική παραπομπή $[54]$.

Συγκεκριμένα, λοιπόν, ισχύουν οι επόμενοι κανόνες:

• εφ' όσον $q > p-1$, ο ασυμπτωτικός χαρακτήρας του σφάλματος παραμένει αναλλοίωτος, άσχετα με το εάν η διόρθωση έχει μία ή άπειρες επαναλήψεις,

• εφ' όσον $q = p-1$, η τάξη του σφάλματος δεν αλλάζει όταν διορθώσουμε μία μόνο φορά (πρόκειται για την Πρόταση που μόλις αποδείξαμε),

• εφ' όσον $q < p-1$, το σφάλμα ελαττώνεται όταν διορθώσουμε μία μόνον φορά. ■

IV.4.3.(ix). Πρακτικές Έννοιες Τα πρακτικά προβλήματα, που παρουσιάζονται κατά την ολοκλήρωση μίας διαφορικής εξίσωσης με μία μέθοδο συναπτών βημάτων είναι, μεταξύ άλλων, κυρίως αυτά που εξετάσαμε στην *Παράγραφο IV.4.2.(vi)*, στην περίπτωση των μεθόδων ξεχωριστών βημάτων, ήτοι:

- ***η επιλογή του βήματος ολοκλήρωσης (: $A-$ευστάθεια),***
- ***ο έλεγχος της διάδοσης των σφαλμάτων,***
- ***η βελτίωση της ακρίβειας.***

Στα τρία αυτά προβλήματα πρέπει να προστεθεί εκείνο της εκκίνησης των μεθόδων συναπτών βημάτων, δεδομένου ότι ο υπολογισμός της ποσότητας y_{n+k} προϋποθέτει την

γνώση των τιμών $y_n, \ldots, y_{n+k+1}$, ενώ ο αντίστοιχος υπολογισμός, με την μέθοδο των ξεχωριστών βημάτων, προαπαιτούσε την γνώση μόνο της τιμής y_n.

Το πρόβλημα του ελέγχου της διάδοσης των σφαλμάτων επιλύεται με τρόπο ανάλογο εκείνου που χρησιμοποιήθηκε στην περίπτωση των μεθόδων ξεχωριστών βημάτων (στην *Υποπαράγραφο IV.4.2.(ix)*). Η βελτίωση της ακρίβειας πραγματοποιείται επίσης με ανάλογο τρόπο εκείνου που εξετάστηκε στην *Υποπαράγραφο IV.4.2.(x)*. Δεν θα μελετήσουμε λοιπόν παρά μόνον τα δύο άλλα προβλήματα, που απομένουν.

IV.4.3.(x). Α-Ευστάθεια των Μεθόδων Συναπτών Βημάτων

Ο ορισμός της $A-$ευστάθειας μίας μεθόδου συναπτών βημάτων στηρίζεται στην ιδέα της ολοκλήρωσης της διαφορικής εξίσωσης:

$$y' = -A\, y \quad \text{με } \mathrm{Re}(A) > 0 .$$

Γι' αυτήν την διαφορική εξίσωση, έχουμε:

$$\sum_{i=0}^{k} a_i \; y_{n+i} = -A\, h \sum_{i=0}^{k} \beta_i \; y_{n+i}$$

απ' όπου έπεται ότι:

$$\sum_{i=0}^{k} (a_i + A\, h\, \beta_i)\, y_{n+i} = 0 .$$

Για να ισχύει

$$\lim_{n\to\infty} y_n = 0 ,$$

δηλαδή για να είναι η *μέθοδος* $A-$ *ευσταθής, πρέπει και αρκεί οι ρίζες της εξίσωσης:*

$$\sum_{i=0}^{k} (a_i + A\, h\, \beta_i)\, \lambda^{\,i} = 0$$

να έχουν απόλυτες τιμές όχι μεγαλύτερες του αριθμού ένα, και ακόμη, οι ρίζες, που έχουν απόλυτες τιμές ίσες με 1, *να είναι απλές.*

Παρατηρεί κανείς ότι αυτή η διπλή συνθήκη δεν επαληθεύεται για κάθε τιμή του βήματος h, αλλά μόνον όταν το μέγεθος $(A\,h)$ ανήκει σε έναν ορισμένο τόπο του μιγαδικού επιπέδου, που ονομάζεται ***τόπος ευστάθειας*** της μεθόδου συναπτών βημάτων.

Ας δούμε, λοιπόν, τον τρόπο με τον οποίο προσδιορίζεται αποτελεσματικά αυτός ο τόπος ευστάθειας. Ο τρόπος, που θα περιγράψουμε εφαρμόζεται επίσης και στην περίπτωση αναζήτησης του τόπου ευστάθειας μεθόδων ξεχωριστών βημάτων.

Η προηγούμενη χαρακτηριστική εξίσωση γράφεται ισοδύναμα υπό την μορφή:

$$\sum_{j=0}^{k} a_j\,\lambda^{\,j} = -A\,h\sum_{j=0}^{k}\beta_j\,\lambda^{\,j}\,,$$

απ' όπου έπεται ότι:

$$A\,h = -\frac{\sum_{j=0}^{k} a_j\,\lambda^{\,j}}{\sum_{j=0}^{k}\beta_j\,\lambda^{\,j}}$$

Θα πρέπει να ευρεθούν οι τιμές της ποσότητας $(A\,h)$ για τις οποίες η χαρακτηριστική εξίσωση έχει όλες τις απόλυτες τιμές των ριζών της ίσες με ένα.

Προς, τούτο θέτοντας

$$\lambda = e^{i\theta},$$

η παραπάνω εξίσωση επαναδιατυπώνεται ως εξής:

$$g(\theta) = A\,h = -\frac{\sum_{j=0}^{k} a_j\,e^{i\,j\,\theta}}{\sum_{j=0}^{k}\beta_j\,e^{i\,j\,\theta}}\,,$$

και, τότε, δίνοντας στην μεταβλητή θ διάφορες τιμές, μέσα από το διάστημα $[0,\pi]$, μπορούμε να σχεδιάσουμε (εμπειρικά) το σχήμα της επίπεδης καμπύλης $g(\theta)$, καθώς επίσης και το σχήμα της συζυγούς προς αυτήν καμπύλης.

Επειδή απομένει το ζήτημα του προσδιορισμού της σχετικής θέσης, ως προς το μόλις κατασκευασμένο σύνορό του, του τόπου ευστάθειας της μεθόδου (δηλαδή η επιβεβαίωση του κατά πόσον ο τόπος ευστάθειας είναι το εσωτερικό ή το εξωτερικό μέρος της κλειστής καμπύλης του συνόρου), πρέπει (και αρκεί) να προσδιορίσουμε τις ρίζες του χαρακτηριστικού πολυωνύμου, για μία δεδομένη και σταθερή τιμή της ποσότητας $(A\,h)$ που όμως δεν βρίσκεται επί του συνόρου του τόπου ευστάθειας.

IV.4.3.(xi). Α- Ευστάθεια των Μεθόδων Πρόβλεψης–Διόρθωσης

Ας θεωρήσουμε πάλι τους συμβολισμούς και την ορολογία της *Υποπαραγράφου IV.4.3.(viii)*, και ας υποθέσουμε ότι το στάδιο της διόρθωσης δεν περιλαμβάνει παρά μόνον μία επανάληψη.

Τότε, η πρόβλεψη διατυπώνεται ως εξής:

$$y^*_{n+k} = -\frac{1}{a^*_k} \sum_{i=0}^{k-1} (a^*_i + A\,h\,\beta^*_i)\, y_{n+i}\,,$$

και μεταφέροντας αυτήν την τιμή στην μαθηματική έκφραση της πρόβλεψης λαμβάνουμε:

$$\sum_{i=o}^{k} a_i\; y_{n+i} = -A\,h\,[-\beta_k \sum_{i=0}^{k-1} (a^*_i + A\,h\,\beta^*_i)\, y_{n+i} + \sum_{i=0}^{k-1} \beta_i\; y_{n+i}]\,,$$

από όπου έπεται η χαρακτηριστική εξίσωση:

$$\sum_{i=0}^{k-1} [a_i - A\,h\,\beta_k\,(a^*_i + A\,h\,\beta^*_i) + A\,h\,\beta_i]\,\lambda^{\,i} + a_k\;\lambda^{\,k} = 0\,.$$

Πρέπει να βρεθούν οι τιμές της ποσότητας $(A\,h)$, για τις οποίες αυτή η χαρακτηριστική εξίσωση έχει ρίζες των οποίων οι απόλυτες τιμές είναι ίσες με 1.

Εάν διατάξουμε την πιο πάνω εξίσωση, ως προς τις φθίνουσες δυνάμεις της ποσότητας $(A\,h)$, θα προκύψει η τριωνυμική έκφραση:

$$P\,(A\,h)^2 + S\,(A\,h) + T = 0$$

όπου:

$$P := -\beta_k \sum_{i=0}^{k-1} \beta_i^* \ \lambda^{\,i}\,,\ S := \sum_{i=0}^{k-1} (\beta_i - \beta_k\ a_i^*)\ \lambda^{\,i}\,,\ \text{και}\ T := \sum_{i=0}^{k} a_i\ \lambda^{\,i}\,.$$

Όπως στην προηγούμενη *Υποπαράγραφο*, θέτοντας

$$\lambda := e^{i\,\theta}\,,$$

μπορούμε να προσδιορίσουμε τις δύο καμπύλες, που περιγράφουν το σύνορο του τόπου ευστάθειας της μεθόδου πρόβλεψης – διόρθωσης:

$$g_1(\theta) = \frac{-S - \sqrt{S^2 - 4\,P\,T}}{2\,P} \quad \text{και} \quad g_2(\theta) = \frac{-S + \sqrt{S^2 - 4\,P\,T}}{2\,P}\,.$$

Χαράζοντας τις καμπύλες

$$g_1(\theta)\ \text{και}\ g_2(\theta)\ \text{για}\ 0 \le \theta \le \pi\,,$$

όπως επίσης και τις συζυγείς τους, κατασκευάζουμε το σύνορο του τόπου ευστάθειας.

IV.4.3.(xii). Εκκίνηση των Μεθόδων Συναπτών Βημάτων Σε μία μέθοδο συναπτών βημάτων ή σε μία μέθοδο πρόβλεψης-διόρθωσης, ο υπολογισμός της τιμής y_{n+k} προϋποθέτει δεδομένες τις τιμές

$$y_n, y_{n+1}, \ldots, y_{n+k+1}\,.$$

Συνεπώς, όταν $n = 0$, πρέπει να γνωρίζουμε, εκ των προτέρων, τις ποσότητες

$$y_0, \ldots, y_{k-1}\,.$$

Όμως, η μόνη γνωστή τιμή είναι αυτή που προέρχεται από την αρχική συνθήκη

$$y(a) = y_0\,.$$

Θα υπολογίσουμε τις υπολειπόμενες απαραίτητες τιμές

$$y_1, \ldots, y_{k-1}\,,$$

με την βοήθεια μίας μεθόδου ξεχωριστών βημάτων (όπως, για παράδειγμα, μίας μεθόδου Runge-Kutta).

Σύμφωνα με το *Θεώρημα IV.4.3.(vi).1*, η τάξη αυτής της μεθόδου ξεχωριστών βημάτων οφείλει να είναι τουλάχιστον ίση με την τάξη της μεθόδου συναπτών βημάτων του προβλήματος. Στην πραγματικότητα, είναι πλεονασμός να επιλεγεί μία μέθοδος ξεχωριστών βημάτων με υψηλότερη τάξη. Εν πάσει περιπτώσει, συνήθως θεωρούμε ότι οι δύο μέθοδοι έχουν την ίδια τάξη.

Βλέπουμε επίσης ότι, κατά την διάρκεια μίας ολοκλήρωσης μεθόδου συναπτών βημάτων, είναι πολύ δύσκολη η οποιαδήποτε μεταβολή του μήκους του βήματος. Αυτό οφείλεται στο γεγονός ότι αυτές οι μέθοδοι κατασκευάζονται επί σημείων $x_n \in [a, \beta]$, τα οποία ισαπέχουν διαδοχικά μεταξύ τους. Εάν, παρ' όλα αυτά, επιθυμούμε την αλλαγή του μεγέθους του βήματος, κατά την εξέλιξη της διαδικασίας της ολοκλήρωσης, θα υποχρεωθούμε να ξαναξεκινήσουμε τους υπολογισμούς, χρησιμοποιώντας την μέθοδο ξεχωριστών βημάτων, που περιέχει k βήματα και, φυσικά, την νέα τιμή του μήκους του βήματος που επιλέξαμε.

Στις μεθόδους πρόβλεψης-διόρθωσης, ο έλεγχος της διάδοσης των σφαλμάτων πραγματοποιείται σχετικά εύκολα: πράγματι, εάν πρόβλεψη και διόρθωση είναι δύο μέθοδοι που έχουν την ίδια τάξη, τότε η ποσότητα:

$$| y^*_{n+k} - y_{n+k} |$$

είναι μία καλή προσέγγιση του σφάλματος:

$$| y(x_{n+k}) - y_{n+k} | .$$

Τέλος, αξίζει να επισημάνουμε ότι το μεγάλο πλεονέκτημα των μεθόδων συναπτών βημάτων, έναντι των μεθόδων Runge-Kutta, είναι ότι δεν προαπαιτούν (εκτός από το στάδιο της εκκίνησης) τον υπολογισμό των τιμών της F σε ενδιάμεσα σημεία. Έτσι, ο όγκος και ο χρόνος των υπολογισμών, που είναι αναγκαίοι για να εκτελεστούν προκειμένου να προχωρήσει ικανοποιητικά η διαδικασία της ολοκλήρωσης, μειώνονται ουσιαστικά και αποτελεσματικά.

IV.4.4. Η ΜΕΘΟΔΟΣ ΤΩΝ ΑΝΑΔΡΟΜΙΚΩΝ ΔΙΑΦΟΡΙΣΕΩΝ

Ένα μαθηματικό πρόβλημα λέγεται ότι είναι ***καλά τοποθετημένο*** (ή ***καλά ορισμένο***), εάν επιδέχεται μοναδικής λύσης και εάν αυτή η λύση εξαρτάται συνεχώς από τα δεδομένα του προβλήματος.

Γενικά, το ζήτημα της μη-μοναδικότητας της λύσης δεν είναι το πιο ενοχλητικό, γιατί μπορεί να υπερκεραστεί, εφ' όσον προσδιοριστεί ακριβώς το είδος της λύσης (ή του συνόλου των λύσεων) που επιδιώκουμε να βρούμε. Όμως, το ζήτημα της συνέχειας της λύσης, είναι σίγουρα, το πλέον καθοριστικό, τουλάχιστον όσον αφορά την αριθμητική προσέγγιση του προβλήματος. Πράγματι, το γεγονός ότι η εισαγωγή δεδομένων σε έναν ηλεκτρονικό υπολογιστή ή, ακόμη, η αντικατάσταση του θεωρητικού προβλήματος με ένα προσεγγιστικό πρόβλημα συνεπάγονται (γενικά) διαταραχές επί των αρχικών δεδομένων δείχνει ότι εάν δεν έχουμε σιγουρευτεί πως ελαφρές διαταραχές επί των αρχικών δεδομένων, δημιουργούν όχι αξιόλογες διαταραχές επί της λύσης, τότε είναι πολύ δύσκολος ο υπολογισμός μίας προσεγγιστικής λύσης του προβλήματος. Παρ' όλα αυτά, δεν είναι πάντοτε αδύνατη η αποτελεσματική αριθμητική αντιμετώπιση ενός *άσχημα τοποθετημένου* μαθηματικού προβλήματος, κυρίως όταν διαθέτουμε συμπληρωματικές πληροφορίες, όπως είναι η μεγάλη κανονικότητα της υπό αναζήτηση λύσης ή, ακόμη, όπως είναι η πιθανοθεωρητική φύση των διαταραχών και των συνεπειών τους.

Παραδείγματα IV.4.4.1. **Είδαμε στην *Παράγραφο IV.4.1*, ότι, σύμφωνα με τις υποθέσεις του *Θεωρήματος IV.4.1.2*, το πρόβλημα Cauchy είναι καλά τοποθετημένο.■**

Δυστυχώς, στο πιο πάνω *Παράδειγμα*, η έννοια του εάν το πρόβλημα Cauchy είναι καλά τοποθετημένο εμφανίζεται ανεπαρκής για να εξασφαλίσει την ύπαρξη προσεγγιστικής λύσης του προβλήματος. Για να ελπίζουμε βάσιμα ότι μπορούμε να προσεγγίσουμε ικανοποιητικά την θεωρητική λύση, πρέπει να έχουμε βεβαιωθεί ότι το πρόβλημα αυτό είναι, επί πλέον, ***αριθμητικά καλά τοποθετημένο***, δηλαδή ότι, αφ' ενός, είναι καλά τοποθετημένο και ότι, αφ' ετέρου, η συνέχεια της λύσης, σαν συνάρτηση των δεδομένων, είναι *επαρκώς καλή*.

Ας θεωρήσουμε, για παράδειγμα, το ακόλουθο πρόβλημα Cauchy:

$$y'(x) = 3\ y(x) - 3\ x, \quad x \in [0,5],$$

με :

$$y(0) = \frac{1}{3},$$

του οποίου η θεωρητική λύση είναι η συνάρτηση:

$$y(x) = \frac{1}{3} + x.$$

Είναι σαφές ότι το πρόβλημα αυτό είναι *μαθηματικά καλά τοποθετημένο.*Αλλά, εάν η συνθήκη $y(0) = \frac{1}{3}$ αντικατασταθεί με την αρχική συνθήκη:

$$y(0) = \frac{1}{3} + \varepsilon,$$

(που μπορεί να «δημιουργηθεί» την στιγμή που εισάγουμε το δεδομένο $\frac{1}{3}$ στον υπολογιστή), τότε η παραπάνω θεωρητική λύση υποκαθίσταται από την:

$$y(x) + \varepsilon\ e^{3x}.$$

Ενδεικτικά, η αληθινή τιμή της λύσης στο σημείο 5:

$$y(5) = \frac{16}{3}.$$

υποκαθίσταται από την τιμή:

$$\frac{16}{3}+\varepsilon\, e^{15} \approx \frac{16}{3}+3\,\varepsilon\, 10^{6}.$$

Έτσι, ένα σχετικό σφάλμα της τάξης

$$\frac{1}{10000}$$

επί της αρχικής τιμής συνεπάγεται ένα σχετικό σφάλμα της τάξης

$$\frac{5000}{100}$$

επί της αληθινής τιμής της λύσης $y(5)$.

Αυτό σημαίνει πως εάν, για τους υπολογισμούς μας, χρησιμοποιήσουμε έναν ηλεκτρονικό υπολογιστή σχετικής ακρίβειας της τάξης του

$$10^{-6},$$

τότε, όποια αριθμητική μέθοδο κι αν εφαρμόσουμε, είναι αδύνατον να προσεγγίσουμε την τιμή $y(5)$. Αντίθετα, χρησιμοποιώντας έναν ηλεκτρονικό υπολογιστή σχετικής ακρίβειας της τάξης του

$$10^{-15},$$

μπορούμε να προσδοκούμε μία αξιόλογη προσέγγιση της τιμής $y(5)$. Με την πρώτη υπολογιστική μηχανή, το παραπάνω πρόβλημα θα θεωρείται ότι είναι *αριθμητικά άσχημα τοποθετημένο,* ενώ με την δεύτερη υπολογιστική μηχανή, το ίδιο πρόβλημα θα θεωρείται ότι είναι *αριθμητικά καλά τοποθετημένο.*

Χωρίς απόδειξη, αναφέρουμε ότι, σύμφωνα με τις υποθέσεις του *Θεωρήματος IV.4.1.2*, το πρόβλημα Cauchy:

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

με:

$y(a)$: δεδομένη αρχική τιμή

είναι αριθμητικά καλά τοποθετημένο, εάν η ποσότητα:

$$\sup_{x\in[a,\beta]} e^{L(x)}$$

δεν είναι πολύ μεγάλη. (Το σύμβολο $L(x)$ αναπαριστά την συνάρτηση

$$L(x) = \int_a^x \ell(s)\, ds\,,$$

όπου $\ell(s)$ είναι μία ολοκληρώσιμη συνάρτηση του [*a*,*β*] τέτοια ώστε:

$$\langle F(s,y) - F(s,z) / y - z\rangle \le \ell(s)\, \|y-z\|^2$$

και όπου το σύμβολο $\langle X / Y\rangle$ αναπαριστά το συνήθες εσωτερικό γινόμενο των οποιονδήποτε δύο στοιχείων $X = (X_1,\ldots,X_k)^T$ και $Y = (Y_1,\ldots,Y_k)^T$ του $\mathbb{R}^k$ δηλαδή:

$$\langle X / Y\rangle = \sum_{i=1}^k X_i\, Y_i\ .)$$

Για τις τεχνικές ιδέες της απόδειξης αυτού του αποτελέσματος, ο ενδιαφερόμενος αναγνώστης παραπέμπεται στην βιβλιογραφική αναφορά [12] ή ακόμη στις εργασίες του *G.Dahlquist*: *Gonvergence and Stability in the Numeric Integration of Ordinary Different Equations* (Math.Scand. 4,33-53,1956), *A Special Stability Problem for Linear Multistep Methods* (BIT 3,27,1963), *Error Analysis for a Class of Methods for Stiff Non-Linear Initial Value Problems* (Proceedings of the Dundee Conference on Numerical Analysis, Lecture Notes in Mathematics,Springer-Verlag,1976).

Μολοταύτα, τίποτα δεν εγγυάται την εύκολη επίλυση ενός αριθμητικά καλά τοποθετημένου προβλήματος.

Ας θεωρήσουμε, για παράδειγμα, το πρόβλημα:

$$y'(x) = -150\, y(x) + 49, \quad x \in [0,1]\,,$$

με:

$$y(0)=\frac{1}{3},$$

του οποίου η θεωρητική λύση είναι η συνάρτηση:

$$y(x)=\frac{1}{3}-x.$$

Το πρόβλημα αυτό είναι αριθμητικά καλά τοποθετημένο, γιατί

$$\sup_{x\in[a,\beta]} e^{L(x)}=1.$$

Εάν προσεγγίσουμε το πρόβλημα με την μέθοδο Euler, βλέπουμε ότι για σταθερό βήμα

$$h=\frac{1}{50},$$

ισχύει:

$$y(x_i)-y_i=(1-\frac{150}{50})^n\ (y(x_0)-y_0)=(-2)^n\ (\frac{1}{3}-y_0),$$

και, κατά συνέπεια, στο τελευταίο σημείο $x=1$ θα δημιουργηθεί ένα σφάλμα:

$$|\,y(1)-y_{50}\,|=2^{50}\,|\,\frac{1}{3}-y_0\,|.$$

Γενικότερα, μπορούμε να διαπιστώσουμε ότι, για

$$h\le\frac{1}{75},$$

η μέθοδος Euler δίνει μία ικανοποιητική προσέγγιση.

Εν τούτοις, μπορούμε, παράλληλα, να εκτιμήσουμε πως, τότε, το κόστος της μεθόδου είναι υψηλό, παρ' ότι το πρόβλημα είναι αριθμητικά καλά τοποθετημένο και η θεωρητική λύση είναι πολύ κανονική (δηλαδή έχει συνεχείς παραγώγους πολύ υψηλής τάξης, μέσα σε μία περιοχή του $[0,1]$).

Μετά από αυτό το παράδειγμα, οδηγούμαστε, δικαιολογημένα, στον επόμενο *Ορισμό*:

Ορισμός IV.4.4.2. ***Λέμε ότι ένα πρόβλημα είναι υπό κανονικές συνθήκες καλά κατασκευασμένο, εάν η εφαρμογή σε αυτό των συνηθισμένων αριθμητικών μεθόδων παρέχει, κάτω από ένα λογικό κόστος, μία ικανοποιητική προσέγγιση της θεωρητικής λύσης. Στην αντίθεση περίπτωση λέμε ότι τι πρόβλημα είναι υπό κανονικές συνθήκες άσχημα κατασκευασμένο.*** ■

Μόλις πριν από λίγο είδαμε ότι το πρόβλημα:

$$y'(x) = -150\, y(x) + 49 \qquad x \in [0,1],$$

με:

$$y(0) = \frac{1}{3}$$

δεν είναι υπό συνθήκες καλά κατασκευασμένο (χωρίς όμως να είναι υπό συνθήκες πολύ άσχημα κατασκευασμένο, αφού μπορούμε να το προσεγγίσουμε σωστά όταν $h \le \frac{1}{15}$).

Σε συνάφεια με τα παραπάνω, δίνουμε, χωρίς απόδειξη, το εξής αποτέλεσμα ($[12]$):

Θεώρημα IV.4.4.3. ***Ένα κριτήριο προκειμένου ένα πρόβλημα Cauchy να θεωρείται υπό κανονικές συνθήκες καλά κατασκευασμένο είναι η ιδιότητα της ποσότητας***

$$e^{L(\beta - a)}$$

να μην είναι πολύ μεγάλη. **(L είναι η σταθερά Lipschitz της απεικόνισης F στο πρόβλημα Cauchy.)** ■

Ας σημειωθεί ότι υπάρχουν προβλήματα τα οποία, ενώ είναι αριθμητικά καλά τοποθετημένα, παρουσιάζονται υπό κανονικές συνθήκες πολύ άσχημα κατασκευασμένα. Σε τέτοια προβλήματα, η εφαρμογή των συνηθισμένων αριθμητικών μεθόδων (όπως η μέθοδος Euler), ακόμα και κάτω από ένα σημαντικό κόστος, εμφανίζεται αποδυναμωμένη και ανίκανη στο να οδηγήσει σε προσεγγιστικές λύσεις. Τότε, λέμε ότι έχουμε να επιλύσουμε ***δύσμορφα προβλήματα***.

Σ' αυτά τα προβλήματα, η απεικόνιση F είναι Lipschitz, οι ποσότητες $e^{L(x)}$ δεν είναι πολύ μεγάλες (αρκετές φορές μάλιστα ισχύει $L(x) \le 0$), όμως, από την άλλη πλευρά, η σταθερά

$$e^{L(\beta-a)}$$

είναι πολύ μεγάλη (υπάρχουν μάλιστα και περιπτώσεις όπου $L(\beta-a)=10^4$). Τα προβλήματα αυτού του τύπου έχουν συχνά μεγάλη σπουδαιότητα στην πράξη. Για παράδειγμα, τα συναντούμε στην κινητική χημεία (:πολύ γρήγορες αντιδράσεις, πυρηνικές αντιδράσεις,…), στην μελέτη των ηλεκτρικών κυκλωμάτων (:τρανζίστορς υψηλών ταχυτήτων), κ.α. Ένα επιστημονικό πεδίο, ιδιαίτερα σημαντικό, μέσα στο οποίο υπεισέρχονται είναι εκείνο της προσέγγισης προβλημάτων εξισώσεων με μερικές παραγώγους (όπως οι εξισώσεις θερμότητας, κυμάτων, Schröndinger, κ.λ.π.).

Η μελέτη των δύσμορφων προβλημάτων αναπτύχθηκε κυρίως τα τελευταία τριάντα χρόνια, και σήμερα διαθέτουμε μεθόδους που αποδίδουν ικανοποιητικές προσεγγίσεις σε πάρα πολλές περιπτώσεις. Τέτοιες μέθοδοι είναι οι πεπλεγμένες μέθοδοι Runge-Kutta και η μέθοδος των αναδρομικών διαφορίσεων, την οποία θα εξετάσουμε αμέσως τώρα.

Οι μέθοδοι Adams είναι πολύ αποτελεσματικές όταν πρόκειται να προσεγγίσουμε ένα διαφορικό πρόβλημα, που είναι υπό κανονικές συνθήκες καλά κατασκευασμένο, αλλά μπορούν να έχουν πολύ υψηλό κόστος κατά την επίλυση άλλων τύπων προβλημάτων. Ιδιαίτερα, για τα δύσμορφα προβλήματα οι μέθοδοι Adams είναι τελείως αναποτελεσματικές.

Γι'αυτόν τον λόγο, υπάρχει ενδιαφέρον εισαγωγής της μεθόδου των αναδρομικών διαφορίσεων, η οποία συγκεντρώνει βελτιωμένες ιδιότητες ευστάθειας. Πρόκειται για μία μέθοδο που μπορεί να χρησιμοποιηθεί και σε άλλα προβλήματα, που είναι υπό κανονικές συνθήκες καλά κατασκευασμένα, αλλά, τότε, παρουσιάζεται λιγότερο αποτελεσματική σε σχέση με τις μεθόδους Adams.

Έχουμε επιλέξει να μελετήσουμε εδώ την μέθοδο των αναδρομικών διαφορίσεων, κάτω από υποθέσεις που αντιστοιχούν στο πλαίσιο των δύσμορφων προβλημάτων. Θα περιγράψουμε, έτσι, το σχήμα της μεθόδου, θα δείξουμε την ύπαρξη και την μοναδικότητα της λύσης της και θα εξετάσουμε την τάξη της, κάτω από μία υπόθεση που προσαρμόζεται σε μία ευρεία κλάση δύσμορφων προβλημάτων.

Η μελέτη της ευστάθειας της μεθόδου τάξης ένα, στην γενική περίπτωση, καθώς και της ευστάθειας της μεθόδου τάξης δύο, στη περίπτωση του σταθερού βήματος, πραγματοποιείται κάτω από την ίδια υπόθεση. Οι εργασίες του G.Dahlquist έδειξαν πως αυτή η υπόθεση αποδεικνύεται ανεπαρκής για να εγγυηθεί την ευστάθεια των μεθόδων αναδρομικών μεθόδων που έχουν τάξη γνήσια μεγαλύτερη του δύο. Μπορούμε να θεωρήσουμε μία πιο ισχυρή υπόθεση, που εξασφαλίζει την ευστάθεια και ταυτόχρονα επιτρέπει την επεξεργασία παρά πολλών δύσμορφων προβλημάτων, όμως, τότε η παρουσίαση της αντίστοιχης μελέτης θα αποβεί πολύπλοκη και λεπτεπίλεπτη, σε τέτοιο σημείο που να ξεπερνά τους εισαγωγικούς στόχους του παρόντος συγγράμματος. Γιαυτό προτιμήσαμε να εξετάσουμε την ευστάθεια των μεθόδων, που έχουν τάξη μεγαλύτερη ή ίση από το τρία, στην περίπτωση των προβλημάτων που είναι υπό συνθήκες καλά κατασκευασμένα.

Δεν θα συμπεριλάβουμε μία μελέτη ενός σχήματος πρόβλεψης-διόρθωσης για την μέθοδο των αναδρομικών διαφορίσεων, καθώς ένα τέτοιο σχήμα (που, είναι, αναγκαστικά, σαφές) δεν αφομοιώνεται σε προβλήματα που είναι υπό κανονικές συνθήκες άσχημα κατασκευασμένα.

Ακόμη, πρόθεσή μας είναι να διεξάγουμε μία μελέτη της ευστάθειας που είναι ανεξάρτητη από την μελέτη της τάξης, αφήνοντας στον ενδιαφερόμενο αναγνώστη την πρωτοβουλία και την επιμέλεια να ταξινομήσει τις υποθέσεις που διασφαλίζουν την σύγκλιση της μεθόδου.

Τέλος, δεν θα συμπεριλάβουμε την ασυμπτωτική μελέτη του σφάλματος της μεθόδου, η οποία, εξ άλλου, οδηγεί σε ένα αποτέλεσμα που είναι ανάλογο με εκείνο των μεθόδων Adams.

IV.4.4.(i). Περιγραφή της Μεθόδου

Ας θεωρήσουμε πάλι το πρόβλημα Cauchy:

$$y'(x) = F(x, y(x)) \qquad x \in [a, \beta],$$

με:

$$y(x_0) = \xi ,$$

και ας υποθέσουμε ότι έχει δοθεί μία διαμέριση:

$$a = x_0 < x_1 < \ldots < x_N = \beta$$

του διαστήματος $[a, \beta]$.

Συμβολίζουμε με:

$$h_n = x_{n+1} - x_n \quad (n = 0,1,\ldots,N-1)$$

το $n-$οστό βήμα και δεχόμαστε ότι, μετά από κάποια διαδικασία, γνωρίζουμε τις προσεγγίσεις:

$$y_n, y_{n-1}, \ldots$$

των τιμών της θεωρητικής λύσης:

$$y(x_n), y(x_{n-1}), \ldots,$$

αντίστοιχα.

Για να βρούμε μία προσέγγιση:

$$y_{n+1}$$

της τιμής $y(x_{n+1})$, θεωρούμε το πολυώνυμο παρεμβολής $p_r^{(n)}(x)$, βαθμού το πολύ r, που είναι τέτοιο ώστε:

$$p_r^{(n)}(x_{n+1-i}) = y_{n+1-i} \qquad \forall i = 0,1,\ldots,r\,.$$

Η μέθοδος των αναδρομικών διαφορίσεων σχηματίζεται γράφοντας:

$$\frac{dp_r^{(n)}}{dx}(x_{n+1}) = F(x_{n+1}, y_{n+1})\,.$$

Πραγματοποιώντας την αλλαγή μεταβλητής:

$$s = \frac{x_{n+1} - x}{h_n}\,,$$

και θέτοντας:

$$a_i(n) := \frac{h_n}{x_{n+1} - x_{n+1-i}} \quad \text{και} \quad P_{n,r}(s) := p_r^{(n)}(x)\,,$$

η παραπάνω εξίσωση γράφεται υπό την μορφή:

$$P'_{n,r}(0) = -h_n\ F(x_{n+1}, y_{n+1})\,,$$

όπου το πολυώνυμο $P_{n,r}$ είναι το μοναδικό πολυώνυμο (ως προς την μεταβλητή s), βαθμού το πολύ r, που είναι τέτοιο ώστε:

$$P_{n,r}(\frac{1}{a_i(n)}) = y_{n+1-i}, \qquad \forall i = 0,1,...,r .$$

Θεωρώντας, μάλιστα, την πολυωνυμική συνάρτηση:

$$\omega_{n,r}(s) = (1 - a_1(n)\, s)...(1 - a_r(n)\, s),$$

επαληθεύεται εύκολα ότι:

$$P_{n,r}(s) = \omega_{n,r}(s)\, y_{n+1} - \sum_{i=1}^{r} \frac{a_i^2(n)\, s\, \omega_{n,r}(s)}{\omega'_{n,r}(1/a_i(n))\,(1 - a_i(n).s)}\, y_{n+1-i},$$

με συνέπεια η εξίσωση

$$P'_{n,r}(0) = -h_n\; F(x_{n+1}, y_{n+1})$$

να διατυπώνεται σύμφωνα με την ακόλουθη έκφραση:

$$\omega'_{n,r}(0)\, y_{n+1} - \sum_{i=1}^{r} \frac{a_i^2(n)\, \omega_{n,r}(0)}{\omega'_{n,r}(1/a_i(n))}\, y_{n+1-i} = -h_n\; F(x_{n+1}, y_{n+1}).$$

Θέτοντας, ακόμα μία φορά:

$$\mathrm{a}_{i,r}(n) := \frac{a_i^2(n)}{\omega'_{n,r}(\frac{1}{a_i(n})\, \omega'_{n,r}(0)} \quad \text{και} \quad \mathrm{b}_r(n) := -\frac{1}{\omega'_{n,r}(0)},$$

η έκφραση αυτή γράφεται υπό την γενική μορφή:

$$y_{n+1} - h_n\; \mathrm{b}_r(n)\, F(x_{n+1}, y_{n+1}) = \sum_{i=1}^{r} \mathrm{a}_{i,r}(n)\, y_{n+1-i} .$$

Ας σημειωθεί πως, όταν $r = 1$, τότε:

$$\mathrm{a}_{1,1}(n) = \mathrm{b}_1(n) = 1,$$

ενώ, όταν $r > 1$, τότε:

$$\mathrm{a}_{i,r}(n) = \frac{a_i(n)\,\mathrm{b}_r(n)}{\prod_{\substack{j=1\\ j\neq i}}^{r}\left(1 - \frac{a_j(n)}{a_i(n)}\right)}, \text{ με } \mathrm{b}_r(n) = \frac{1}{\sum_{i=1}^{r} a_i(n)}.$$

Ισχύει η ακόλουθη βοηθητική Πρόταση, της οποίας η απόδειξη παραλείπεται και αφήνεται σαν εύκολη άσκηση για τον αναγνώστη:

Λήμμα IV.4.4.(i).1. ***Υποθέτουμε ότι η ακολουθία των βημάτων*** h_n ***επαληθεύει την ανισοτική σχέση:***

$$h_{n+1} \le \delta\, h_n \qquad \forall n = 0,1,\ldots,N-1,$$

για κάποιο $\delta > 0$***. Τότε, υπάρχει μία σταθερά*** $\mathrm{a}_r(\delta) > 0$ ***τέτοια ώστε***

$$\sum_{i=1}^{r} \left|\mathrm{a}_{i,r}(n)\right| \le \mathrm{a}_r(\delta).$$

Επί πλέον, ισχύει:

$$\frac{1}{r} \le \mathrm{b}_r(n) \le 1. \blacksquare$$

Εάν η απεικόνιση F ικανοποιεί μία συνθήκη Lipschitz, με σταθερά L, τότε μπορούμε ν' αποδείξουμε την ύπαρξη και την μοναδικότητα της λύσης της:

$$y_{n+1} - h_n \, \mathrm{b}_r(n) \, F(x_{n+1}, y_{n+1}) = \sum_{i=1}^{r} \mathrm{a}_{i,r}(n) \, y_{n+1-i} \, ,$$

εφ' όσον είναι:

$$h_n \, \mathrm{b}_r(n) \, L < 1 \, .$$

Στην συνέχεια, θα δείξουμε ένα ασθενέστερο αποτέλεσμα, χρησιμοποιώντας το εξής:

Λήμμα IV.4.4.(i).2. ***Έστω ϕ μία συνεχής συνάρτηση από το $\mathbb{R}$ μέσα στο $\mathbb{R}$, για την οποία υπάρχει μία σταθερά $c < 1$ τέτοια ώστε:***

$$\phi(x) - \phi(y) \le c \, (x - y) \qquad (\forall x \in \mathbb{R}, \forall y \in \mathbb{R}) \, .$$

Τότε, για κάθε $b \in \mathbb{R}$ και κάθε $\theta \in [0,1]$ υπάρχει μοναδικό $y \in \mathbb{R}$ που επιλύει την εξίσωση:

$$y = \theta \, \phi(y) + b \, .$$

Απόδειξη. Θα δείξουμε πρώτα την μοναδικότητα της λύσης.

Προς τούτο, ας υποθέσουμε ότι υπάρχουν δύο λύσεις y και z της παραπάνω εξίσωσης. Τότε, θα έχουμε:

$$z - y = \theta \, [\phi(z) - \phi(y)] \Rightarrow |z - y|^2 = \theta \, [\phi(z) - \phi(y)] \, [z - y] \le c \, \theta \, |z - y|^2 \, ,$$

πράγμα που συνεπάγεται ότι $|z - y| = 0$ και, επομένως, ότι $z = y$.

Όσον αφορά στην ύπαρξη, θα δείξουμε, κατ' αρχάς, ότι οι λύσεις της εξίσωσης $y = \theta \, \phi(y) + \mathrm{b}$ συνιστούν ένα φραγμένο σύνολο του $\mathbb{R}$.

Προς τούτο, παρατηρούμε ότι εάν y είναι μία λύση αυτής της εξίσωσης τότε:

$$\begin{aligned} |y-\mathrm{b}|^2 &= \theta\,\phi(y)\,(y-\mathrm{b}) \\ &= \theta\,(\phi(y)-\phi(\mathrm{b}))\,(y-\mathrm{b}) + \theta\,\phi(\mathrm{b})\,(y-\mathrm{b}) \\ &\le c\,|y-\mathrm{b}|^2 + |\phi(\mathrm{b})|\,|y-\mathrm{b}| \end{aligned}$$

και άρα:

$$|y-\mathrm{b}| \le \frac{\phi(\mathrm{b})}{1-c}.$$

Έπεται, λόγω συμπάγειας, ότι το σύνολο:

$$E = \{\theta \in [0,1] : \textit{η εξίσωση}\ \ y = \theta\,\phi(y) + \mathrm{b}\ \ \textit{έχει λύση}\}$$

είναι κλειστό μέσα στο $[0,1]$.

Από την άλλη πλευρά, στην περίπτωση που η συνάρτηση ϕ είναι μία (μη σταθερή) συνεχώς διαφορίσιμη απεικόνιση παντού στο $\mathbb{R}$, η ανισότητα της υπόθεσης του *Θεωρήματος* συνεπάγεται ότι

$$|\phi'(y)| \le c \quad \forall y \in \mathbb{R},$$

απ' όπου έπεται το σύνολο *E* είναι, επίσης, ανοικτό μέσα στο $[0,1]$. Επειδή $0 \in E \Rightarrow E \ne \varnothing$ και επειδή το διάστημα $[0,1]$ είναι συνεκτικό σύνολο, συμπεραίνουμε πως $\mathrm{E} = [0,1]$.

Η περίπτωση που η συνάρτηση ϕ δεν είναι διαφορίσιμη αντιμετωπίζεται με ένα πέρασμα στην οριακή κατάσταση της διαφορίσιμης περίπτωσης. ■

Εάν το *Λήμμα* αυτό εφαρμοσθεί για:

$$\theta = 1 \text{ και } \phi(y) = h_n \ \mathrm{b}_r(n)\ F(x_{n+1}, y),$$

θα συνάγουμε το επόμενο βασικό *Θεώρημα*:

Θεώρημα IV.4.4.(i).3. ***Εάν υπάρχει μία σταθερά*** $\ell \in \mathbb{R}$ ***τέτοια ώστε:***

$$F(x,y) - F(x,z) \le \ell\,(y-z) \quad (\forall y \in \mathbb{R}, \forall z \in \mathbb{R}, \forall x \in [a,\beta]),$$

τότε για κάθε h_n***, με*** $h_n\ \mathrm{b}_r(n)\ \ell < 1$***, το σχήμα:***

$$y_{n+1} - h_n\ \mathrm{b}_r(n)\ F(x_{n+1}, y_{n+1}) = \sum_{i=1}^{r} \mathrm{a}_{i,r}(n)\ y_{n+1-i}$$

έχει μία και μοναδική λύση. ■

Στην περίπτωση που $r = 1$, το σχήμα του *Θεωρήματος IV.4.4.(i).3* γράφεται υπό την μορφή:

$$y_{n+1} - y_n = h_n\ F(x_{n+1}, y_{n+1}),$$

και ξαναβρίσκουμε την μέθοδο Euler (που είναι, επίσης, η μέθοδος Adams-Moulton πρώτης τάξης). Ακόμα, όταν:

$$h_n = h_{n-1} = \ldots = h_{n+1-r} \ ,$$

οι συντελεστές $\mathrm{a}_{i,r}(n)$ και $\mathrm{b}_r(n)$ δεν εξαρτώνται από το *n*, και δίνονται στον επόμενο πίνακα:

	b_r	$a_{1,1}$	$a_{2,1}$	$a_{3,1}$	$a_{4,1}$	$a_{5,1}$	$a_{6,1}$
$r=1$	1	1	-	-	-	-	-
$r=2$	$\frac{2}{3}$	$\frac{4}{3}$	$-\frac{1}{3}$	-	-	-	-
$r=3$	$\frac{6}{11}$	$\frac{18}{11}$	$-\frac{9}{11}$	$\frac{2}{11}$	-	-	-
$r=4$	$\frac{12}{25}$	$\frac{48}{25}$	$-\frac{36}{25}$	$\frac{16}{25}$	$-\frac{3}{25}$	-	-
$r=5$	$\frac{60}{137}$	$\frac{300}{137}$	$-\frac{300}{137}$	$\frac{200}{137}$	$-\frac{75}{137}$	$\frac{12}{137}$	-
$r=6$	$\frac{60}{147}$	$\frac{360}{147}$	$-\frac{450}{147}$	$\frac{400}{147}$	$-\frac{225}{147}$	$\frac{72}{147}$	$-\frac{10}{147}$

IV.4.4.(ii). Μελέτη της Τάξης Ας θεωρήσουμε το σφάλμα της μεθόδου στο σημείο x_n:

$$\sigma_n = y(x_{n+1}) - h_n\, b_r(n)\, y'(x_{n+1}) - \sum_{i=1}^{r} a_{i,r}(n)\, y(x_{n+1-r}).$$

Σύμφωνα με την κατασκευή της μεθόδου, το σφάλμα σ_n ισούται με μηδέν, εάν η συνάρτηση *y* είναι πολυώνυμο βαθμού το πολύ r. Άρα, ***η μέθοδος των αναδρομικών διαφορίσεων, που διαρθρώνεται σε r βήματα, είναι τάξης r.***

Χρησιμοποιώντας τον τύπο Taylor, για τάξη μέχρι $r+1$, στο σημείο x_{n+1}, βλέπουμε ότι:

$$\sigma_n = \sum_{i=1}^{r} a_{i,r}(n) \int_{x_{n+1-i}}^{x_{n+1}} \frac{(x_{n+1-i} - s)^r}{r!} y^{(r+1)}(s)\, ds ,$$

απ' όπου έπεται ότι, κάτω από τις υποθέσεις του *Λήμματος IV.4.4.(i).1*, υπάρχει μία θετική σταθερά $C_r = C_r(\delta)$ που είναι τέτοια ώστε:

$$|\sigma_n| \le C_r\, h_n^r \int_{x_{n+1-r}}^{x_{n+1}} \left| y^{(r+1)}(\mathrm{s}) \right| ds .$$

Κατά συνέπεια:

$$\sum_{i=r-1}^{n-1} |\sigma_i| \le r\, C_r\, h_n^r \int_0^{x_n} \left| y^{(r+1)}(\mathrm{s}) \right| ds .$$

Έχουμε, ακόμη, το εξής αποτέλεσμα:

Λήμμα IV.4.4.(ii).1. ***Υπάρχει σημείο*** $\xi_n \in [x_{n+1-r}, x_{n+1}]$ ***τέτοιο ώστε:***

$$\sigma_n = \varepsilon_r(n)\, y^{(r-1)}(\xi_n),$$

όπου:

$$\varepsilon_r(n) = -h_n\, b_r(n) \frac{(x_{n+1} - x_n)\ldots(x_{n+1} - x_{n+1-r})}{(r+1)!} .$$

Απόδειξη. Έστω $p_r(x)$ το πολυώνυμο, βαθμού το πολύ r, που είναι τέτοιο ώστε:

$$p_r(x_{n+1-i}) = y(x_{n+1-i}) \quad \forall i = 0,1,\ldots,r .$$

Τότε, σύμφωνα με την κατασκευή της μεθόδου, θα είναι:

$$\sigma_n = h_n\, b_r(n)\, [p_r'\, (x_{n+1}) - y'(x_{n+1})] .$$

Ας θεωρήσουμε τώρα το πολυώνυμο $q(x)$, βαθμού το πολύ $r=1$, που ορίζεται από τον *Τύπο*:

$$q(x)=p_r(x)-\frac{p'_r\,(x_{n+1})-y'(x_{n+1})}{(x_{n+1}-x_n)\ldots(x_{n+1}-x_{n+1-r})}\,(x-x_{n+1})\,(x-x_n)\ldots(x-x_{n+1-r})\,.$$

Προφανώς, επειδή ισχύει:

$$q(x_{n+1-i})=p_r(x_{n+1-r})=y(x_{n+1-r})\qquad \forall i=0,1,\ldots,r\,,$$

και

$$q'(x_{n+1})=p'_r(x_{n+1})-(p'_r\,(x_{n+1})-y'(x_{n+1}))=y'(x_{n+1})\,,$$

η συνάρτηση $f(x)=y(x)-q(x)$ επιδέχεται το πολύ $(r-1)$ διακριτών ριζών, από τις οποίες η μία είναι διπλή, επί του διαστήματος $[x_{n+1-r},x_{n+1}]$.

Κατά συνέπεια, η παράγωγος $f^{(r+1)}(x)$ επιδέχεται το πολύ μίας ρίζας ξ_n, και κατ' ακολουθία, είναι:

$$0=y^{(r+1)}(\xi_n)-q^{(r+1)}(\xi_n)=y^{(r+1)}(\xi_n)+\frac{p'(x_{n+1})-y'(x_{n+1})}{(x_{n+1}-x_n)\ldots(x_{n+1}-x_{n+1-r})}\,(r+1)!,$$

απ' όπου έπεται ο ισχυρισμός του *Λήμματος*. ■

Παρατήρηση IV.4.4.(ii).2. Όταν τα βήματα h_n είναι όλα σταθερά και όταν $r \neq 1$ (δηλαδή όταν η μέθοδος των αναδρομικών διαφορίσεων ταυτίζεται με την μέθοδο Adams-Moulton), η μέθοδος Adams-Moulton είναι τουλάχιστον 4 φορές πιο ακριβής από την μέθοδο των αναδρομικών διαφορίσεων. Γι' αυτόν τον λόγο, χρησιμοποιούμε, κατά προτίμηση, την μέθοδο Adams-Moulton, εφ' όσον, βέβαια, έχει εξασφαλισθεί η επάρκεια της ευστάθειάς της.■

IV.4.4.(iii). Ευστάθεια των Μεθόδων Τάξεων 1 και 2

Ας θεωρήσουμε, κατ' αρχάς, την μέθοδο αναδρομικών διαφορίσεων τάξης 1:

$$y_{n+1} = y_n + h_n \; F(x_{n+1}, y_{n+1}) \quad (n \geq 0),$$

και ας επισυνάψουμε σ' αυτήν το διαταρασσόμενο σχήμα:

$$z_{n+1} = z_n + h_n \; F(x_{n+1}, z_{n+1}) + \varepsilon_n .$$

Θα υποθέσουμε ότι η συνάρτηση F ικανοποιεί την υπόθεση του *Θεωρήματος IV.4.4.(i).3*. Τότε, τα δύο προηγούμενα σχήματα επιδέχονται μοναδικής λύσης το καθένα, εφ' όσον ισχύει η ανισοτική συνθήκη $h_n \; \ell < 1$. Επί πλέον, τότε, θα ισχύει:

$$|y_{n+1} - z_{n+1}| \leq (|y_n - z_n| + h_n \; \ell \, |y_{n+1} - z_{n+1}| + |\varepsilon_n|) ,$$

που συνεπάγεται ότι:

$$|y_{n+1} - z_{n+1}| \leq \frac{1}{1 - h_n \; \ell} \; (|y_n - z_n| + |\varepsilon_n|) .$$

Με την επίκληση των ανισοτικών σχέσεων:

$$\frac{1}{1 - \ell \, h_n} \leq 1 + \frac{h_n \; \ell}{1 - h_n \; \ell} \leq \exp(\frac{h_n \; \ell}{1 - h_n \; \ell}) ,$$

μπορούμε να δείξουμε, χρησιμοποιώντας μαθηματική επαγωγή, ότι:

$$|y_n - z_n| \le \exp[\frac{(x_n - x_0)\,\ell}{1 - h_n\,\ell}]\,|y_0 - z_0| + \sum_{i=0}^{n-1} \exp[\frac{(x_n - x_i)\,\ell}{1 - h_n.\ell}]\,|\varepsilon_i|,$$

απ' όπου έπεται η ευστάθεια του σχήματος των αναδρομικών διαφορίσεων 1ης τάξης. Ας σημειωθεί ότι η πιο πάνω τελευταία ανισότητα παραμένει σε ισχύ ακόμη και όταν $\ell < 0$, γιατί τότε το αρχικό σφάλμα φθίνει κατά των εκτέλεση των υπολογισμών.

Εάν, μάλιστα, δεχθούμε ότι η συνάρτηση $y(x)$ πρέπει να είναι δύο φορές συνεχώς διαφορίσιμη μέσα σε μία ανοικτή περιοχή του διαστήματος $[a, \beta]$ και αν λάβουμε $z_n = y(x_n)$, τότε θα διαπιστώσουμε ότι:

$$|\varepsilon_i| \le h_\iota \int_{x_i}^{x_{i+1}} |y''(x)|dx\,,$$

απ' όπου έπεται η ακόλουθη εκτίμηση του σφάλματος

$$\sigma_n = y(x_n) - y_n (= \varepsilon_n)$$

των αναδρομικών διαφορίσεων:

$$|\sigma_n| \le \exp\left[\frac{(x_n - x_0)\ell}{1 - h_n\,\ell}\right]|y(x_0) - y_0| + \begin{cases} h_n \int_{x_0}^{x_n} \exp\left[\frac{(x_n - s)\ell}{1 - h_n\,\ell}\right]|y''(s)|ds, \; av\; \ell \le 0 \\ \frac{h_n}{1 - h_n\,\ell} \int_{x_0}^{x_n} \exp\left[\frac{(x_n - s)}{1 - h_n\,\ell}\right]|y''(s)|ds, \; av\; \ell \ge 0. \end{cases}$$

Ας θεωρήσουμε, τώρα, την μέθοδο αναδρομικών διαφορίσεων τάξης 2.

Εδώ ας υποθέσουμε ότι τα βήματα h_n είναι σταθερά, και όλα ίσα με h. Σε αυτήν την περίπτωση, το σχήμα της μεθόδου διατυπώνεται ως εξής:

$$y_{n+1} = \frac{4}{3} y_n - \frac{1}{3} y_{n-1} + \frac{2}{3} h\, F(x_{n+1}, y_{n+1}) \qquad (n \ge 1).$$

Ας επισυνάψουμε στο σχήμα αυτό το διαταραγμένο σχήμα:

$$z_{n+1} = \frac{4}{3} z_n - \frac{1}{3} z_{n-1} + \frac{2}{3} h\, F(x_{n+1}, y_{n+1}) + \varepsilon_n.$$

Εάν δεχθούμε ότι η συνάρτηση F ικανοποιεί την υπόθεση του *Θεωρήματος IV.4.4.(i).3*, και, ακόμα, ότι ισχύει:

$$\frac{2}{3} h \ell < 1,$$

τα δύο παραπάνω σχήματα επιδέχονται ακριβώς μίας λύσης το κάθε ένα. Θέτοντας:

$$t_n := z_n - y_n,$$

βλέπουμε ότι:

$$(t_{n+1} - \frac{4}{3} t_n + \frac{1}{3} t_{n-1})\, t_{n+1} = \frac{2}{3} h (F(x_{n+1}, z_{n+1}) - F(x_{n+1}, y_{n+1})) t_{n+1} + \varepsilon_n\, t_{n+1}$$

$$\leq \frac{2}{3}\, h\, \ell\, t_{n+1}^2 + |\varepsilon_n|\, |t_{n+1}|.$$

Χρησιμοποιώντας τον συμβολισμό:

$$\theta_n^2 := t_n^2 + (2\, t_n - t_{n-1})^2,$$

μπορούμε να διαπιστώσουμε ότι, καθώς

$$\theta_{n+1}^2 - \theta_n^2 = 6\, (t_{n+1} - \frac{4}{3} t_n + \frac{1}{3} t_{n-1})\, t_{n+1} - (t_{n+1} - 2\, t_n + t_{n-1})^2,$$

ισχύει, εφ' όσον $\ell \geq 0$, ότι:

$$\theta_{n+1}^2 - \theta_n^2 \leq 4\, h\, \ell\, t_{n+1}^2 + 6 |\varepsilon_n|\, |t_{n+1}| \leq 4\, h\, \ell\, \theta_{n+1}^2 + 6 |\varepsilon_n|\, \theta_{n+1},$$

από την οποία συνεπάγεται, εφ' όσον $0 \leq 4\, h\, \ell \leq 1$, ότι:

$$(\sqrt{1 - 4\, h\, \ell\, \theta_{n+1}} - \frac{3 |\varepsilon_n|}{\sqrt{1 - 4\, h\, \ell}})^2 \leq \theta_n^2 + \frac{9}{1 - 4 h\, \ell} |\varepsilon_n|^2 \leq (\theta_n + \frac{3 |\varepsilon_n|}{\sqrt{1 - 4 h\, \ell}})^2$$

απ' όπου έπεται, εφ' όσον $0 \leq 4\, h \ell \leq 1$, ότι:

$$\theta_{n+1} \leq \frac{1}{\sqrt{1 - 4\, h\, \ell}}\, \theta_n + \frac{6 |\, \varepsilon_n\, |}{1 - 4\, h\, \ell}, \quad (n \geq 1)$$

και, περαιτέρω, κάνοντας χρήση της ανισότητας:

$$\frac{1}{\sqrt{1-4h\ell}} \le \exp(\frac{4h\ell}{1-4.h.\ell})$$

ότι:

$$\theta_n \le \exp[\frac{2\ell}{1-4h\ell}(x_n - x_1)]\theta_1 + \frac{6}{1-4h\ell}\sum_{i=1}^{n-1}\exp[\frac{2\ell}{1-4h\ell}(x_n - x_{i+1})]|\varepsilon_i|$$

$(\ell \ge 0)$.

Στην περίπτωση που η συνθήκη του *Θεωρήματος IV.4.4.(i).3* επαληθεύεται για $\ell < 0$, τότε η ίδια συνθήκη θα επαληθεύεται, κατά πλείστον λόγο, για $\ell = 0$, και, κατά συνέπεια, τότε θα είναι:

$$\theta_n \le \theta_1 + 6\sum_{i=1}^{n-1}|\varepsilon_i| \quad (\ell < 0).$$

Τελικά, επειδή:

$$|t_n| \le \theta_n \quad \text{και} \quad |\theta_1| \le 3|t_1| + |t_0|,$$

καταλήγουμε στην ακόλουθη εκτίμηση:

$$|y_n - z_n| \le \exp[\frac{2\ell_+}{1-4h\ell_+}(x_n - x_0)][3|y_1 - z_1| + |y_0 - z_0|]$$

$$+\frac{6}{1-4h\ell_+}\sum_{i=1}^{n-1}\exp[\frac{2\ell_+}{1-4h\ell_+}(x_n - x_{i+1})]|\varepsilon_i|,$$

από την οποία πιστοποιείται η ευστάθεια του σχήματος των αναδρομικών διαφορίσεων τάξης 2 (και με σταθερό βήμα).

Εάν, επί πλέον, δεχθούμε ότι η συνάρτηση $y(x)$ πρέπει να είναι τρεις φορές συνεχώς διαφορίσιμη μέσα σε μία ανοικτή περιοχή του διαστήματος $[a,\beta]$, και εάν πάρουμε $z_n = y(x_n)$, τότε θα διαπιστώσουμε ότι:

$$\varepsilon_n = \frac{4}{3}\int_{x_n}^{x_{n+1}}\frac{(s-x_n)^2}{2}y'''(s)\,ds - \frac{1}{3}\int_{x_{n-1}}^{x_{n+1}}\frac{(s-x_{n-1})^2}{2}y'''(s)\,ds,$$

από όπου συνεπάγεται ότι:

$$|\varepsilon_n| \le \frac{2h^2}{9}\int_{x_n}^{x_{n+1}} |y'''(s)|\,ds + \frac{h^2}{6}\int_{x_{n-1}}^{x_n} |y'''(s)|\,ds\,.$$

Από αυτήν την ανισότητα, και έχοντας υπ' όψη την παραπάνω εκτίμηση, συμπεραίνουμε ότι μπορούμε να φράξουμε το απόλυτο σφάλμα

$$|\,\sigma_n\,| = |\,y(x_n) - y_n\,| = |\,\varepsilon_n\,|$$

της μεθόδου των αναδρομικών διαφορίσεων 2ης τάξης (με σταθερό βήμα h) κατά τον εξής τρόπο (και κάτω από την προϋπόθεση ότι η συνάρτηση F ικανοποιεί την συνθήκη του *Θεωρήματος IV.4.4.(i).3*):

$$|\,\sigma_n\,| \le \begin{cases} \exp\left[\dfrac{2\ell(x_n - x_0)}{1-4h\ell}\right][3|\,y(x_1) - y_1\,| + |\,y(x_0) - y_0\,|] \\ \qquad + \dfrac{7h^2}{3}\,\dfrac{1}{1-4h\ell}\displaystyle\int_{x_0}^{x_n} \exp\left[\dfrac{2\ell(x_n - s)}{1-4h\ell}\right]|y'''(s)|\,ds, \;\; av\ \ell \ge 0 \\ 3|\,y(x_1) - y_1\,| + |\,y(x_0) - y_0\,| + \dfrac{7h^2}{3}\displaystyle\int_{x_0}^{x_n} |\,y'''(s)\,|\,ds, \;\; av\ \ell \le 0\,. \end{cases}$$

IV.4.4.(iv). Ευστάθεια με Αμετάβλητη Τάξη

Θα εξετάσουμε τώρα την γενική περίπτωση κάτω από την υπόθεση Lipschitz:

$$|\,F(x,y) - F(x,z)\,| \le L\,|\,y - z\,| \quad (\forall y \in \mathbb{R}, \forall z \in \mathbb{R}, \forall x \in [a,\beta])\,.$$

Προς τούτο, στο σχήμα της μεθόδου των αναδρομικών διαφορίσεων:

$$y_{n+1} - h_n\ \mathrm{b}_r(n)\ F(x_{n+1}, y_{n+1}) = \sum_{i=1}^{r} \mathrm{a}_{i,r}(n)\ y_{n+1-i}$$

επισυνάπτουμε το διαταρασσόμενο σχήμα:

$$z_{n+1} - h_n \, \mathrm{b}_r(n) \, F(x_{n+1}, z_{n+1}) = \sum_{i=1}^{r} \mathrm{a}_{i,r}(n) \, z_{n+1-i} + \varepsilon_n \qquad (n \geq r-1) .$$

Λέμε ότι το σχήμα της μεθόδου των αναδρομικών διαφορίσεων είναι ***ευσταθές*** (ή απλά ότι ***η μέθοδος αυτή είναι ευσταθής***) εάν υπάρχει μία σταθερά $C_1 > 0$ τέτοια ώστε:

$$\max_{r-1 \leq n \leq N} | y_n - z_n | \leq C_1 \, \{ \sum_{i=0}^{r-1} | y_i - z_i | + \sum_{j=r-1}^{N-1} | \varepsilon_j | \} .$$

Θέτοντας:

$$G_{n+1} := \int_0^1 \frac{\partial F}{\partial y} (x_{n+1}, y_{n+1} + \theta \, (z_{n+1} - y_{n+1})) \, d\theta ,$$

βλέπουμε αμέσως ότι είναι:

$$| G_{n+1} | \leq L \text{ και } F(x_{n+1}, z_{n+1}) - F(x_{n+1}, y_{n+1}) = G_{n+1} \, (z_{n+1} - y_{n+1}) ,$$

από όπου έπεται η ισότητα:

$$(1 - h_n \, \mathrm{b}_r(n) \, G_{n+1}) \, (z_{n+1} - y_{n+1}) = \sum_{i=1}^{r} \mathrm{a}_{i,r}(n) \, (z_{n+1-i} - y_{n+1-i}) + \varepsilon_n \quad (n \geq r-1) .$$

Εάν, μάλιστα, εισάγουμε τους συμβολισμούς:

$$u_n = z_n - y_n ,$$

$$U_n = \begin{pmatrix} u_n \\ u_{n-1} \\ \vdots \\ u_{n-r+1} \end{pmatrix}, \qquad R_n = \begin{pmatrix} \mathrm{a}_1 & \mathrm{a}_2 & \cdots & & \mathrm{a}_r \\ 1 & 0 & \cdots & & 0 \\ & 1 & \ddots & & \vdots \\ & & \ddots & & \\ & \mathrm{O} & & & \\ & & & 1 & 0 \end{pmatrix} \quad (\text{με } \mathrm{a}_i = \mathrm{a}_{i,r}(n)),$$

$$D = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & \mathrm{O} & \\ 0 & & & \end{pmatrix} \quad \text{και} \quad E_n = \begin{pmatrix} \varepsilon_n \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

η παραπάνω ισότητα γίνεται:

$$(\mathrm{I} - h_n \, \mathrm{b}_r(n) G_{n+1} D) \cdot U_{n+1} = R_n \cdot U_n + E_n ,$$

η οποία γράφεται πιο απλά ως εξής:

$$U_{n+1} = S_n \cdot U_n + \widetilde{E}_n \qquad (n \geq r-1),$$

όπου χρησιμοποιήσαμε τους, επί πλέον, συμβολισμούς:

$$S_n = (\mathrm{I} - h_n \, \mathrm{b}_r(n) G_{n+1} \, D)^{-1} \cdot R_n ,$$

$$\widetilde{E}_n = (\mathrm{I} - h_n \, \mathrm{b}_r(n) \, G_{n+1} \, D)^{-1} \cdot E_n = \frac{1}{\mathrm{I} - h_n \, \mathrm{b}_r(n) G_{n+1}} \cdot E_n .$$

Μπορούμε, κατ' αυτόν τον τρόπο και χρησιμοποιώντας μαθηματική επαγωγή, να οδηγηθούμε στον ακόλουθο *Τύπο*:

$$U_n = S_{n-1} \cdot S_{n-2} \cdot \ldots \cdot S_{r-1} \cdot U_{r-1} + \sum_{i=r-1}^{n-1} S_{n-1} \cdot S_{n-2} \cdot \ldots \cdot S_{i+1} \cdot \widetilde{E}_i$$

(κάτω από την, κατά σύμβαση, παραδοχή ότι:

$$S_{n-1} \cdot S_{n-2} \cdot \ldots \cdot S_{i+1} = \mathrm{I}, \text{ εάν } \; i = n-1).$$

Ας θεωρήσουμε, κατ' αρχάς, την ειδική περίπτωση που η απεικόνιση $F(x,y)$ δεν εξαρτάται από την μεταβλητή y και ισχύει $\varepsilon_i = 0 \;\; \forall i$. Τότε, επειδή $G_n = 0$, θα είναι $S_n = R_n$ και άρα:

$$U_n = R_{n-1} \cdot R_{n-2} \cdot \ldots \cdot R_{r-1} \cdot U_{r-1} .$$

Εάν το σχήμα της μεθόδου είναι ευσταθές, τότε θα είναι:

$$\|U_n\| \leq r \, C_1 \, \|U_{r-1}\|.$$

Εδώ ο συμβολισμός $\|w\|$, για το διάνυσμα στήλη:

$$w = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_q \end{pmatrix} \in \mathbb{R}^q ,$$

αναπαριστά τον αριθμό:

$$\|w\| = \sum_{i=1}^{q} |\, w_i \,|.$$

Ακόμη, δοθέντος ενός $p \times q$ πίνακα $M = (m_{i,j})$, με $m_{i,j} \in \mathbb{R}$, θα θέτουμε:

$$\|M\| := \sup_{w \neq 0} \left(\frac{\|M \cdot w\|}{\|w\|} \right).$$

Επειδή, όπως μπορεί, σχετικά εύκολα, να αποδειχτεί, ισχύει:

$$\|M\| = \sup_{1 \le j \le q} \sum_{i=1}^{p} \left| m_{i,j} \right|,$$

έχουμε την εύλογη δυνατότητα να εισάγουμε τον παρακάτω συμβολισμό, στην περίπτωση που ο πίνακας M περιορίζεται σε ένα διάνυσμα γραμμή $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_q)$:

$$\|\gamma\| = \max_{1 \le j \le q} \left| \gamma_j \right|.$$

Τέλος, για αργότερη χρήση, αξίζει να σημειωθεί πως όταν δύο πίνακες M και N, με πραγματικά στοιχεία, είναι τέτοιοι ώστε το γινόμενό τους να έχει έννοια, τότε έχουμε:

$$\|M \cdot N\| \le \|M\| \, \|N\|.)$$

Ας επανέλθουμε στην ροή της κεντρικής ανάλυσης της παρούσης *Παραγράφου*. Καθώς το διάνυσμα U_{r-1}, που αποτελείται από τις αρχικές διαταραχές $z_i - y_i \quad (i = 0,1,\ldots,r-1)$, είναι αυθαίρετο, συμπεραίνουμε, έχοντας υπ' όψη τα προηγηθέντα, ότι:

$$\|R_{n-1} \cdot R_{n-2} \cdot \ldots \cdot R_{r-1}\| \le r \, C_1.$$

Οδηγούμαστε, έτσι, στον ακόλουθο *Ορισμό*:

Ορισμός IV.4.4. (iv).1. ***Λέμε ότι το σχήμα της μεθόδου των αναδρομικών διαφορίσεων ικανοποιεί την συνθήκη ευστάθειας (S), εάν υπάρχει ένας θετικός αριθμός K τέτοιος ώστε***

$$\|R_n \cdot R_{n-1} \cdot \ldots \cdot R_m\| \le K, \quad \text{εάν } r-1 \le m \le n \le N-1. \blacksquare$$

Η επαλήθευση της συνθήκης (S) συνεπάγεται ορισμένες ενδιαφέρουσες και κρίσιμες ιδιότητες, που επηρεάζουν την ευστάθεια της μεθόδου. Για την επιβεβαίωση του ισχυρισμού αυτού, θα χρειαστούμε την ακόλουθη βοηθητική *Πρόταση*:

Λήμμα IV.4.4. (iv).2. ***Εάν η μέθοδος ικανοποιεί την συνθήκη (S), τότε για κάθε m και n που είναι τέτοια ώστε $r-1 \le m \le n \le N-1$, ισχύει η ανισότητα:***

$$\|S_n \cdot S_{n-1} \cdot \ldots \cdot S_m\| \le K \exp(K \sum_{i=m}^{n} \|S_i - R_i\|).$$

Απόδειξη. Κατ' αρχάς, παρατηρούμε ότι $K \ge 1$ και ότι ισχύει η ταυτότητα:

$$S_n \cdot S_{n-1} \cdot \ldots \cdot S_m = R_n \cdot R_{n-1} \cdot \ldots \cdot R_m + \sum_{i=m}^{n} R_n \cdot \ldots R_{i+1} \cdot (S_i - R_i) \cdot S_{i-1} \cdot \ldots \cdot S_m .$$

Συμπεραίνουμε, λοιπόν, πως:

$$\|S_n \cdot S_{n-1} \cdot \ldots \cdot S_m\| \le K \, \beta_n ,$$

όπου χρησιμοποιήσαμε τον συμβολισμό:

$$\beta_n = 1 + \sum_{i=m}^{n} \|S_i - R_i\| \quad \|S_{i-1} \cdot S_{i-2} \cdot \ldots \cdot S_m\| .$$

Επειδή όμως είναι:

$$\beta_m \le 1 + \|S_m - R_m\| \le 1 + K\|S_m - R_m\| \le \exp\left(K \, \|S_m - R_m\|\right)$$

και:

$$\beta_{n+1} - \beta_n = \|S_{n+1} - R_{n+1}\| \, \|S_n \cdot S_{n-1} \cdot \ldots \cdot S_m\| \leq K \, \beta_n \, \|S_{n+1} - R_{n+1}\|,$$

θα είναι και:

$$\beta_{n+1} \leq \beta_n \, \exp(K \, \|S_{n+1} - R_{n+1}\|),$$

ή ακόμη:

$$\beta_n \leq \exp(K \sum_{i=m}^{n} \|S_i - R_i\|),$$

από όπου έπεται το αποτέλεσμα του *Λήμματος* (εφ' όσον φυσικά ληφθεί υπ' όψη η ανισότητα $\|S_n \cdot S_{n-1} \cdot \ldots \cdot S_m\| \leq K \, \beta_n$, που δείξαμε πιο πάνω). ■

Είμαστε τώρα σε θέση να δείξουμε το πρώτο βασικό αποτέλεσμα της *Παραγράφου*:

Θεώρημα IV.4.4. (iv).3. ***Εάν το σχήμα της μεθόδων των αναδρομικών διαφορίσεων ικανοποιεί την συνθήκη (S), τότε το σχήμα αυτό είναι ευσταθές.***

Απόδειξη. Αρχίζουμε την *Απόδειξη* με την επισήμανση ότι, επειδή:

$$(\mathrm{I} - h_n \, \mathrm{b}_r(n) \, G_{n+1} \cdot D)^{-1} = \mathrm{I} + \frac{h_n \, \mathrm{b}_r(n) \, G_{n+1}}{1 - h_n \, \mathrm{b}_r(n) \, G_{n+1}} \cdot D,$$

έχουμε:

$$S_n - R_n = \frac{h_n \, \mathrm{b}_r(n) \, G_{n+1}}{1 - h_n \, \mathrm{b}_r(n) \, G_{n+1}} \cdot D \cdot R_n,$$

και, επομένως, ισχύει:

$$\| S_n - R_n \| = \left(\frac{h_n \, \mathrm{b}_r(n) \, G_{n+1}}{1 - h_n \, \mathrm{b}_r(n) \, G_{n+1}}\right) \max_i |\, \mathrm{a}_{i,r}(n) \,|.$$

Στην συνέχεια, θέτοντας:

$$h\mathrm{b} := \max_n h_n \ \mathrm{b}_r(n),$$

και παρατηρώντας πως η συνθήκη (S) συνεπάγεται ότι:

$$\left|\mathrm{a}_{i,r}(n)\right| \le K,$$

καταλαβαίνουμε ότι:

$$\left\|S_n - R_n\right\| \le K \frac{h_n \ L}{1 - h\mathrm{b}\ L},$$

από όπου, χρησιμοποιώντας το προηγούμενο *Λήμμα*, συνάγεται η ακόλουθη ανισότητα:

$$\left\|S_n \cdot S_{n-1} \cdot \ldots \cdot S_m\right\| \le K \ \exp\left(\frac{K^2 \ L\,(x_{n+1} - x_n)}{1 - h\mathrm{b}\ L}\right).$$

Μεταφέροντας την ανισοτική αυτή σχέση στον *Τύπο*:

$$U_n = S_{n-1} \cdot S_{n-2} \cdot \ldots S_{r-1} \cdot U_{r-1} + \sum_{i=r-1}^{n-1} S_{n-1} \cdot S_{n-2} \cdot \ldots S_{i+1} \cdot \widetilde{E}_i$$

(που τοποθετήθηκε λίγο πριν από την εκφώνηση του *Ορισμού IV.4.4.(iv).1)* συμπεραίνουμε ότι:

$$\left\|U_n\right\| \le K\left\{\exp[K' L\,(x_n - x_0)]\,\left\|U_{r-1}\right\|\right\} + \frac{K}{1 - h\mathrm{b}\,L}\left\{\sum_{i=r-1}^{n-1} \exp[K' L\,(x_n - x_{i+1})]\,|\,\varepsilon_i\,|\right\}$$

όπου:

$$K' = \frac{K^2}{1 - h\mathrm{b}\ L}.$$

Ο ισχυρισμός του *Θεωρήματος* έπεται τώρα αμέσως, αφού ληφθεί υπ' όψη η προφανής σχέση:

$$|\,z_n - y_n\,| \le \left\|U_n\right\|. \blacksquare$$

Σύμφωνα με το αποτέλεσμα αυτό, η συνθήκη (S) εμφανίζεται σαν μία αναγκαία και ικανή συνθήκη για την ευστάθεια του σχήματος της μεθόδου των αναδρομικών διαφορίσεων.

Θα εξετάσουμε τώρα, λοιπόν, τις προϋποθέσεις εκείνες, κάτω από τις οποίες η συνθήκη αυτή ικανοποιείται. Πριν ξεκινήσουμε, ας σημειώσουμε ότι η συνθήκη αυτή δεν εξαρτάται παρά μόνο από το σχήμα και από την επιλογή της διαμέρισης $x_0, x_1, \ldots, x_n, \ldots$

Θα εξετάσουμε τώρα την περίπτωση που τα σημεία της διαμέρισης ισαπέχουν διαδοχικά μεταξύ τους. Συναφώς, παραθέτουμε το εξής:

Θεώρημα IV.4.4. (iv).4. ***Υποθέτουμε ότι το βήμα h_n είναι σταθερό.***

(α). ***Εάν $r > 6$, τότε το σχήμα της μεθόδου των αναδρομικών διαφορίσεων δεν ικανοποιεί την συνθήκη (S).***

(β). ***Εάν $1 \le r \le 6$, τότε το σχήμα της μεθόδου των αναδρομικών διαφορίσεων ικανοποιεί την συνθήκη (S).***

Απόδειξη. Επειδή, κάτω από την υπόθεση του *Θεωρήματος*, ο πίνακας R_n είναι ανεξάρτητος από το n, μπορούμε να τον συμβολίζουμε απλά με R και να διατυπώνουμε την συνθήκη (S) υπό την μορφή:

$$\left\| R^m \right\| \le K \ , \ \forall m > 0 \, .$$

Το χαρακτηριστικό πολυώνυμο του πίνακα R γράφεται:

$$P_R(\lambda) = \lambda^{\,r} - \sum_{i=1}^{r} \mathrm{a}_{i,r} \ \lambda^{\,r-i} \, .$$

Ο αριθμητικός υπολογισμός των ριζών του πολυωνύμου $P_R(\lambda)$ πιστοποιεί ότι, για $7 \le r \le 15$, το συγκεκριμένο πολυώνυμο επιδέχεται τουλάχιστον μίας ρίζας, της οποίας η απόλυτη τιμή υπερέχει γνησίως του αριθμού ένα. Κατά συνέπεια, τότε, ο πίνακας R^m δεν

μπορεί να είναι φραγμένος για κάθε $m > 0$, και, κατ' ακολουθία, τότε η συνθήκη (S) δεν ικανοποιείται. Έτσι, το μέρος *(α)* του *Θεωρήματος* αποδείχτηκε.

Ας τοποθετηθούμε τώρα στην περίπτωση $1 \le r \le 6$. Παρατηρούμε, κατ' αρχάς ότι, επειδή $P_R(1) = 1 - \sum_{i=1}^{r} a_{i,r} = 0$, ο αριθμός 1 είναι ιδιοτιμή του πίνακα R και ότι το αντίστοιχο ιδιοδιάνυσμα του R είναι το διάνυσμα:

$$\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^r .$$

Ο υπολογισμός των ριζών του πολυωνύμου $P_R(\lambda)$ καταδεικνύει ότι οι ρίζες αυτές είναι απλές και, ακόμη, ότι όσες από αυτές τις ρίζες είναι διάφορες του αριθμού ένα έχουν απόλυτη τιμή < 1. Για τον λόγο τούτο, πρέπει να υπάρχει ένας πίνακας αντιστρέψιμος H, τέτοιος ώστε το γινόμενο $H^{-1} \cdot R \cdot H$ να είναι της μορφής:

$$H^{-1} \cdot R \cdot H = \begin{pmatrix} 1 & | & \\ 0 & | & \\ \vdots & | & R' \\ 0 & | & \end{pmatrix}$$

όπου ο πίνακας R' να έχει r γραμμές και $(r-1)$ στήλες και να είναι τέτοιος ώστε $\|R'\| < 1$. (Πράγματι, προς τούτο, αρκεί να λάβουμε τον πίνακα H που καθιστά διαγώνιο τον πίνακα R.) Αυτό σημαίνει ότι πρέπει να ισχύει: $\|H^{-1} \cdot R \cdot H\| \le 1$, από όπου έπεται ότι: $\|H^{-1} \cdot R^n \cdot H\| \le 1$ και, κατ' ακολουθία, ότι, για κάθε $m \ge 0$: $\|R^m\| \le \|H^{-1}\| \, \|H\|$, που ολοκληρώνει την απόδειξη του μέρους *(β)* του *Θεωρήματος*.■

Απομένει να εξεταστεί η περίπτωση που το βήμα της διαμέρισης δεν παραμένει σταθερό. Τότε έχουμε το εξής:

Θεώρημα IV.4.4. (iv).5. ***Υποθέτουμε ότι το βήμα h_n δεν είναι σταθερό.***

Εάν $1 \le r \le 6$, τότε υπάρχουν δύο πραγματικοί αριθμοί β_r και γ_r, τέτοιοι ώστε αφενός:

$$0 \le \beta_r < 1 < \gamma_r$$

και αφετέρου η συνθήκη (S) να ικανοποιείται για κάθε διαμέριση

$$x_0 < x_1 < \ldots < x_n < \ldots$$

που έχει την ακόλουθη ιδιότητα:

$$\beta_r \le \frac{h_{n+1}}{h_n} \le \gamma_r \quad \forall n \ge 0.$$

Απόδειξη. Ας δεχτούμε ότι $r \ge 2$. (Η περίπτωση $r = 1$, είναι τετριμμένη, καθώς, τότε, ο πίνακας R_n εκφυλίζεται σε πίνακα με μοναδικό στοιχείο τον αριθμό 1: $R_n = (1)$ και ως εκ τούτου μπορούμε να πάρουμε $\beta_1 = 0$ και $\gamma_1 \to +\infty$.)

Ας ξαναθεωρήσουμε τον πίνακα H, της *Απόδειξης* του προηγούμενου *Θεωρήματος*. Επειδή όπως είδαμε:

$$H^{-1} \cdot R \cdot H = \begin{pmatrix} 1 & | & \\ 0 & | & \\ \vdots & | & R' \\ 0 & | & \end{pmatrix}$$

(όπου ο R' είναι ένας $r \times (r-1)$ πίνακας), η πρώτη στήλη του H συνιστά το ιδιοδιάνυσμα του R που αντιστοιχεί στην ιδιοτιμή 1.

Πολλαπλασιάζοντας, εφ' όσον υπάρχει ανάγκη, τον πίνακα H με έναν μη-μηδενικό πραγματικό αριθμό, μπορούμε να υποθέσουμε ότι αυτή η πρώτη στήλη του H είναι της μορφής:

$$e = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^r .$$

Ακόμη, παρατηρούμε πως είναι $R_n \cdot e = e$, γιατί $\sum_{i=1}^{r} \mathrm{a}_{i,r}(n) = 1$. Έτσι, το γινόμενο $H^{-1} \cdot R_n \cdot H$ θα έχει την μορφή:

$$H^{-1} \cdot R_n \cdot H = \begin{pmatrix} 1 & | & \\ 0 & | & \\ \vdots & | & R'_n \\ 0 & | & \end{pmatrix}$$

(όπου ο R'_n είναι ένας $r \times (r-1)$ πίνακας).

Από την άλλη πλευρά, οι συντελεστές $\mathrm{a}_{i,r}(n)$ εξαρτώνται, με συνεχή τρόπο, από το διάνυσμα:

$$s_n = \begin{pmatrix} h_n / h_{n-1} \\ h_{n-1} / h_{n-2} \\ \vdots \\ h_{n-r+2} / h_{n-r+1} \end{pmatrix} .$$

Έπεται ότι ο πίνακας R'_n τείνει προς τον πίνακα R', όταν το διάνυσμα s_n τείνει προς το διάνυσμα e. Καθώς, μάλιστα, $\|R'\| < 1$, υπάρχουν δύο σταθερές β_r και γ_r, τέτοιες ώστε εάν:

$$\beta_r \le \frac{h_{n+1}}{h_n} \le \gamma_r\,, \quad \forall n \ge 0\,,$$

τότε:

$$\|R'_n\| \le 1\,.$$

Κατά συνέπεια:

$$\left\|H^{-1} \cdot R_n \cdot H\right\| = \max\{1, \|R'_n\|\} = 1, \quad \forall n \ge 0\,,$$

από όπου συνάγεται ότι:

$$\|R_n \cdot R_{n-1} \cdot \ldots \cdot R_m\| \le \left\|H^{-1}\right\| \|H\| \quad (r-1 \le m \le n)\,,$$

που ολοκληρώνει την *Απόδειξη* του *Θεωρήματος*. ■

Παραδείγματα IV.4.4.(iv).6. ***Έστω*** $r = 2$ ***. Τότε:***

$$\mathrm{a}_{2,2}(n) = 1 - \mathrm{a}_{1,2}(n) = -\frac{(h_n/h_{n-1})^2}{(1 + 2\,[h_n/h_{n-1}])}\,,$$

και εάν λάβουμε:

$$H = \begin{pmatrix} 1 & 0 \\ 1 & \varepsilon \end{pmatrix},$$

θα διαπιστώσουμε ότι:

$$H^{-1} \cdot R_n \cdot H = \begin{pmatrix} 1 & \varepsilon\,\mathrm{a}_{2,2}(n) \\ 0 & \mathrm{a}_{2,2}(n) \end{pmatrix},$$

από όπου συνεπάγεται ότι:

$$\left\|H^{-1} \cdot R_n \cdot H\right\| \le 1,$$

εφ' όσον:

$$\left|\mathrm{a}_{2,2}(n)\right|(1+|\varepsilon|) \le 1.$$

Η ανισότητα $\left|\mathrm{a}_{2,2}(n)\right| < 1$ ***ισοδυναμεί με την σχέση:***

$$\frac{h_n}{h_{n-1}} < 1+\sqrt{2}.$$

Για τον λόγο αυτό, εάν επιλέξουμε:

$$\gamma_2 = 1+\sqrt{2}, \quad \varepsilon = (1+2\,\gamma_2 - \gamma_2^2)/\gamma_2^2 \qquad \textbf{\textit{και}} \qquad \beta_2 = 0,$$

βλέπουμε ότι η συνθήκη ευστάθειας (S) ***ικανοποιείται για***

$$0 < \frac{h_{n+1}}{h_n} \le \gamma_2. \blacksquare$$

IV.4.4.(ν). Ευστάθεια με Μεταβλητή Τάξη και Μεταβλητό Βήμα

Θα υποθέσουμε τώρα ότι η τάξη r_n μεταβάλλεται μαζί με το n και ότι

$$1 \le r_n \le 6 \text{ και } r_n \le n+1.$$

Τότε, είναι δυνατόν να επιταχύνουμε, μην γνωρίζοντας παρά μόνον το y_0.

Θέτοντας:

$$\mathrm{a}_i(n) := \mathrm{a}_{i,r_n}(n) \quad (1 \le i \le r_n) \text{ και } \mathrm{b}(n) := \mathrm{b}_{r_n}(n),$$

η μέθοδος των αναδρομικών διαφορίσεων διατυπώνεται ως εξής:

$$y_{n+1} - h_n\,\mathrm{b}(n)\,F(x_{n+1}, y_{n+1}) = \sum_{i=1}^{r_n} \mathrm{a}_i(n)\,y_{n+1+i} \quad (n \ge 0).$$

Επισυνάπτουμε σε αυτήν την μέθοδο το διαταραγμένο σχήμα:

$$z_{n+1} - h_n\,\mathrm{b}(n)\,F(x_{n+1}, z_{n+1}) = \sum_{i=1}^{r_n} \mathrm{a}_i(n)\,z_{n+1+i} + \varepsilon_n \quad (n \ge 0).$$

Εξακολουθούμε να λέμε ότι η μέθοδος είναι ***ευσταθής***, εάν υπάρχει σταθερά $C_1 > 0$ τέτοια ώστε:

$$\max_{0\le n\le N} |\, y_n - z_n \,| \le C_1 (|\, y_0 - z_0 \,| + \sum_{j=0}^{N-1} |\, \varepsilon_j \,|).$$

Θέτουμε:

$$u_n := z_n - y_n \;\;,\;\; u_i := 0 \;\;\; \forall i < 0,$$

$$\mathrm{a}_i(n) := 0 \;\;\;,\;\;\; \text{εάν } i > r_n$$

και:

$$r := \max_{0\le n\le N} r_n .$$

Τότε, χρησιμοποιώντας τους συμβολισμούς της προηγούμενης *Υποπαραγράφου*, μπορούμε να γράψουμε την ακόλουθη σχέση:

$$(1 - h_n \,\mathrm{b}(n) G_{n+1}) u_{n+1} = \sum_{i=1}^{r} \mathrm{a}_i(n) u_{n+1-i} + \varepsilon_n \quad (n \ge 0),$$

από την οποία συνεπάγεται ότι:

$$U_{n+1} = S_n \cdot U_n + \widetilde{E}_n \quad (n \ge 0).$$

Αν αναπαραστήσουμε με U_0 το διάνυσμα:

$$U_0 = \begin{pmatrix} z_0 - y_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^r .$$

τότε το ίδιο σκεπτικό με αυτό της προηγούμενης *Υποπαραγράφου* μπορεί να οδηγήσει στην διαπίστωση ότι η *μέθοδος είναι ευσταθής εάν η επόμενη συνθήκη ευστάθειας* (S) *ικανοποιείται:*

"υπάρχει ένας θετικός αριθμός K *τέτοιος ώστε:*

$$\left\| R_n \cdot R_{n-1} \cdot \ldots \cdot R_m \right\| \le K \quad (0 \le m \le n)".$$

Στηριζόμενοι σε αυτήν την διαπίστωση, παρουσιάζουμε τώρα ένα θετικό αποτέλεσμα:

Θεώρημα IV.4.4. (ν).1. ***Έστω*** $r_n \le 3$. ***Τότε υπάρχουν δύο πραγματικοί αριθμοί*** β^* ***και*** γ^****, με***

$$0 \le \beta^* < 1 < \gamma^*,$$

τέτοιοι ώστε η μέθοδος με μεταβλητή τάξη και μεταβλητό βήμα να είναι ευσταθής, εφ' όσον η διαμέριση της μεθόδου έχει την παρακάτω ιδιότητα:

$$\beta^* \le \frac{h_{n+1}}{h_n} \le \gamma^* \quad (\forall n \ge 0).$$

Απόδειξη. Επιλέγοντας:

$$H = \begin{pmatrix} 1 & 0 & 0 \\ 1 & (1/12) & 0 \\ 1 & (1/12) & (1/3) \end{pmatrix},$$

βλέπουμε ότι:

$$H^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -12 & 12 & 0 \\ 0 & -3 & 3 \end{pmatrix},$$

και, κατά συνέπεια, το γινόμενο $H^{-1} \cdot R_n \cdot H$ θα είναι της μορφής:

$$H^{-1} \cdot R_n \cdot H = \begin{pmatrix} 1 & | & \\ 0 & | & \\ \vdots & | & R'_n \\ 0 & | & \end{pmatrix},$$

όπου ο πίνακας R'_n θα επαληθεύει τις παρακάτω σχέσεις:

- $\|R'_n\| = \frac{1}{4}$, εάν $r_n = 1$,

- $\|R_n'\| = \frac{13}{12}\left|a_{2,2}(n)\right| + \frac{1}{4}$, εάν $r_n = 2$,
- $\|R_n'\| = \max\left\{[\frac{1}{4} + |a_{2,3}(n) + a_{3,3}(n)|\frac{13}{12}],\ [(4+\frac{1}{3})|a_{3,3}(n)|]\right\}$ εάν $r_n = 3$.

Όπως είδαμε στο *Παράδειγμα IV.4.4.(iv).6*:

$$|a_{2,2}| = \frac{(h_n / h_{n-1})^2}{1 + 2\, h_n / h_{n-1}}.$$

Επομένως, θα είναι $\|R_n'\| \le 1$, εάν $r_n = 2$ και εάν:

$$\frac{h_n}{h_{n-1}} \le \frac{9}{13}(1 + \frac{\sqrt{22}}{3}) \approx 1.77.$$

Ακόμη, εάν $r_n = 3$ και εάν $h_n = h_{n-1} = h_{n-2}$, τότε

$$\|R_n'\| = 31/33,$$

ενώ εάν $h_n \ne h_{n-1} \ne h_{n-2}$, τότε, από την συνέχεια των $a_{2,3}(n)$ και $a_{3,3}(n)$ ως προς τις ποσότητες h_n / h_{n-1} και h_{n-1} / h_{n-2}, έπεται ότι υπάρχουν πραγματικοί αριθμοί β^* και γ^* τέτοιοι ώστε

$$0 \le \beta^* < 1 < \gamma \le 1.77$$

και:

$$\beta^* \le \frac{h_n}{h_{n-1}}, \frac{h_{n-1}}{h_{n-2}} \le \gamma^* \Rightarrow \|R_n'\| \le 1,$$

Σε κάθε περίπτωση, λοιπόν, θα ισχύει:

$$\left\|H^{-1} \cdot R_n \cdot H\right\| \le 1,$$

και άρα η συνθήκη (S) θα επαληθεύεται. ■

Παρατήρηση IV.4.4.(ν).2. Η συνθήκη:

$$\beta^* \le \frac{h_{n+1}}{h_n} \le \gamma^* \quad (\forall n \ge 0)$$

του προηγούμενου *Θεωρήματος* μπορεί να υποκατασταθεί, αναλυόμενη ως εξής:

- $h_n / h_{n-1} \le 1.77$, εάν $r_n = 2$,
- $\beta^* \le h_n / h_{n-1} \le \gamma^*$ και $\beta^* \le h_{n-1} / h_{n-2} \le \gamma^*$, εάν $r_n = 3$.

(Δεν υπάρχουν αντίστοιχες σχέσεις όταν $r_n = 1$). ■

Παρατήρηση IV.4.4.(ν).3. Στην *Απόδειξη* του προηγούμενου *Θεωρήματος*, επιλέξαμε έναν πίνακα H, του οποίου η φύση επέτρεπε την διεξαγωγή αρκετά εύκολων υπολογισμών. Μία διαφορετική επιλογή του πίνακα H μπορεί να οδηγήσει σε έναν βελτιωμένο προσδιορισμό των σταθερών β^* και γ^*. ■

Στην συνέχεια, θα τοποθετηθούμε στην περίπτωση που:

$$r_n = 6 .$$

Ας θεωρήσουμε, λοιπόν, τον 6×6 πίνακα:

$$H = \begin{pmatrix} 1 & & & \\ 1 & 1 & \text{O} & \\ \vdots & \vdots & \ddots & \\ 1 & 1 & \dots & 1 \end{pmatrix},$$

που έχει αντίστροφό του τον πίνακα:

$$H^{-1}=\begin{pmatrix} 1 & & & & \\ -1 & 1 & \mathrm{O} & & \\ & -1 & \ddots & & \\ & & \ddots & 1 & \\ & \mathrm{O} & & -1 & 1 \end{pmatrix}.$$

Χωρίς κάποια αξιόλογη δυσκολία, μπορούμε, τότε, να επαληθεύσουμε τους πιο κάτω *Τύπους*:

$$\widetilde{R}_n = H^{-1}\cdot R_n \cdot H = \left(\begin{array}{c|c} 1 & d_n \\ \hline 0 & \\ \vdots & R'_n \\ 0 & \end{array}\right), \text{ με } d_n = (-\mathrm{a}'_1, -\mathrm{a}'_2, -\mathrm{a}'_3, -\mathrm{a}'_4, -\mathrm{a}'_5)$$

και:

$$R'_n = \begin{pmatrix} \mathrm{a}'_1 & \mathrm{a}'_2 & \cdots & \mathrm{a}'_5 \\ 1 & \ddots & \mathrm{O} & \\ & \ddots & \ddots & \\ \mathrm{O} & & 1 & 0 \end{pmatrix}, \text{ με } \mathrm{a}'_i = \mathrm{a}'_i(n) = (\mathrm{a}'_{i+1}(n) + \ldots + \mathrm{a}_6(n))$$

Εξ άλλου, έχουμε και την ακόλουθη βοηθητική Πρόταση:

Λήμμα IV.4.4.(ν).4. ***Εάν υπάρχει ένας θετικός αριθμός*** K ***και ένας πραγματικός αριθμός*** $c \in]-\infty, 1[$ ***, τέτοιοι ώστε:***

$$\left\| R'_n \cdot R'_{n-1} \cdot \ldots \cdot R'_m \right\| \le K\, c^{n-m},$$

τότε η συνθήκη (S) ***ικανοποιείται.***

Απόδειξη. Χωρίς βλάβη της γενικότητας, θα πραγματοποιήσουμε την *Απόδειξη* μόνον στην περίπτωση που $m = 0$.

Προς τούτο, επισημαίνουμε, κατ' αρχάς, ότι επειδή:

$$\|R'_n\| = \max_{1\le i\le 5}(1+|a'_i|) \le K\,,$$

θα είναι:

$$\|d_n\| = \max_{1\le i\le 5}(|a'_i|) \le K-1\,.$$

Μετά, θέτουμε:

$$A_n := \widetilde{R}_n \cdot \widetilde{R}_{n-1} \cdot \ldots \cdot \widetilde{R}_0 = \begin{pmatrix} 1 & | & & b_n & \\ - & - & - & - & - \\ 0 & | & & & \\ \vdots & | & & A'_n & \\ 0 & | & & & \end{pmatrix}$$

Επειδή, προφανώς:

$$A'_n = R'_n \cdot R'_{n-1} \cdot \ldots \cdot R'_0\,,$$

μπορούμε να γράψουμε:

$$b_n = d_n\ A'_{n-1} + d_{n-1}\ A'_{n-2} + \ldots + d_1\ A'_0 + d_0\,,$$

και να διαπιστώσουμε ότι:

$$\|A'_n\| \le K\ c^n\,.$$

Από αυτά, μπορούμε, κατ' ακολουθία, να συμπεράνουμε πως:

$$\begin{aligned} \|b_n\| &\le \|d_n\|\,\|A'_{n-1}\| + \|d_{n-1}\|\,\|A'_{n-2}\| + \ldots + \|d_1\|\,\|A'_0\| + \|d_0\| \\ &\le (K-1)\ K\ (c^{n-1} + c^{n-2} + \ldots + 1) \\ &\le K\,\frac{K-1}{c-1}\,. \end{aligned}$$

Καθώς είναι σαφές ότι:

$$\|A_n\| \le \max\{1, \|b_n\| + \|A'_n\|\},$$

και επομένως ότι:

$$\|A_n\| \le \max\left\{1,\ K\frac{K-1}{1-c} + K\,c^n\right\} \le K^2\frac{1}{1-c},$$

μπορούμε τώρα να συνάγουμε ότι:

$$\|R_n \cdot R_{n-1} \cdot \ldots \cdot R_0\| = \|H \cdot A_n \cdot H^{-1}\| \le \|H\|\,\|H^{-1}\|\,\frac{K^2}{1-c},$$

απ' όπου έπεται ο ισχυρισμός του *Λήμματος*. ■

Είμαστε τώρα σε θέση να αποδείξουμε το δεύτερο βασικό αποτέλεσμα της παρούσης *Υποπαραγράφου*:

Θεώρημα IV.4.4.(v).5. ***Υπάρχουν πραγματικοί αριθμοί*** β^* ***και*** γ^****:***

$$0 \le \beta^* < 1 < \gamma^*,$$

και ακέραιοι αριθμοί:

$$\pi_{i,j}\ (1 \le i\ \ \kappa\alpha\iota\ \ j \le 6),$$

τέτοιο ώστε οι δύο επόμενες ιδιότητες:

- $\beta^* \le \dfrac{h_{n+1}}{h_n} \le \gamma^*\ \ \forall n \ge 0,$

και:

- ***δεν υπάρχει αλλαγή της τάξης*** r_n ***σε τάξη*** r_{n+1} ***παρά μόνον όταν η μέθοδος τάξης*** r_n ***έχει χρησιμοποιηθεί*** $\pi_{r_n, r_{n+1}}$ ***φορές διαδοχικά***

να εξασφαλίζουν την ισχύ της συνθήκης ευστάθειας (S)***.***

Απόδειξη. Συμβολίζουμε με $\Re_r$ τον πίνακα R'_n, που αντιστοιχεί στα δεδομένα

$$r_n = r \text{ και } h_n = h_{n-1} = \ldots = h_{n+1-r}.$$

Όταν $1 \le r \le 6$, η φασματική ακτίνα $\boldsymbol{\rho}(\Re_r)$ του πίνακα $\Re_r$ επαληθεύει την ανισότητα:

$$\boldsymbol{\rho}(\Re_r) < 1,$$

επειδή οι ιδιοτιμές του R'_n είναι εκείνες οι ιδιοτιμές του πίνακα R_n που διαφέρουν από τον αριθμό ένα.

Επιλέγουμε πραγματικούς αριθμούς c_r ώστε:

$$\boldsymbol{\rho}(\Re_r) < c_r < 1,$$

και αντιστρέψιμους πίνακες H_r, τέτοιους ώστε:

$$\left\| H_r^{-1} \cdot \Re_r \cdot H_r \right\| < c_r.$$

Χρησιμοποιώντας την συνέχεια του πίνακα R'_n, ως προς τις διατεταγμένες ποσότητες $\left(h_n / h_{n-1}, h_{n-1} / h_{n-2}, \ldots\right)$ $\left(\text{για } r_n = \text{σταθερό}\right)$, συνάγουμε ότι υπάρχουν πραγματικοί αριθμοί β^* και γ^*, με $0 \le \beta^* < 1 \le \gamma^*$, τέτοιοι ώστε η πρώτη ιδιότητα να συνεπάγεται ότι:

$$\left\| H_{r_n}^{-1} \cdot R'_n \cdot H_{r_n} \right\| \le c_{r_n}, \ \forall n \ge 0.$$

Ας εξετάσουμε τώρα το μέγεθος $R'_n \cdot R'_{n-1} \cdot \ldots \cdot R'_m$. Προς τούτο, θα θέσουμε $n_0 = m$ και θα υποθέσουμε ότι χρησιμοποιούμε (μετά το σημείο x_m)

$(n_1 - n_0)$ φορές την μέθοδο τάξης s_1,

μετά $(n_2 - n_1)$ φορές την μέθοδο τάξης s_2,

.....

και τέλος $(n_\ell - n_{\ell-1})$ φορές την μέθοδο τάξης s_ℓ,

μέχρι να φτάσουμε στο σημείο x_{n+1}.

Έχουμε λοιπόν:

$$n_\ell = n+1 \text{ και } r_i = s_j, \quad \text{για} \quad n_{j-1} \le i \le n_j .$$

Θέτουμε

$$s_{\ell+1} = s_\ell = r_n$$

και θεωρούμε τους πίνακες:

$$A_j = H_{s_{j+1}}^{-1} \cdot R'_{n_j-1} \cdot R'_{n_j-2} \cdot \ldots \cdot R'_{n_{j-1}} \cdot H_{s_j} .$$

Σύμφωνα με τη σχέση $\left\| H_{r_n}^{-1} \cdot R'_n \cdot H_{r_n} \right\| \le c_{r_n} \quad (\forall n \ge 0)$, που δείξαμε παραπάνω, πρέπει να ισχύει:

$$\left\| A_j \right\| \le \left\| H_{s_{j+1}}^{-1} \cdot H_{s_j} \right\| c_{s_j}^{(n_j - n_{j-1})} .$$

Έστω τώρα ένας πραγματικός αριθμός c, τέτοιος ώστε:

$$c_r < c < 1, \qquad \forall r = 1,\ldots,6 .$$

Είναι τότε σαφές ότι υπάρχουν ακέραιοι $\pi_{i,j}$, τέτοιος ώστε:

$$\left\| H_j^{-1} \cdot H_i \right\| \le \left(\frac{c}{c_i}\right)^{\pi_{i,j}} ,$$

και, έτσι, σύμφωνα με την δεύτερη ιδιότητα της εκφώνησης, πρέπει να είναι:

$$\pi_{s_j,s_{j+1}} \le n_j - n_{j-1}, \qquad (\forall j = 2,\ldots,\ell-1),$$

απ' όπου έπεται ότι:

$$\|A_j\| \le (\frac{c}{c_{s_j}})^{(n_j-n_{j-1})}\ c_{s_j}^{(n_j-n_{j-1})} = c^{(n_j-n_{j-1})}.$$

Για $j=1$, η σχέση αυτή γίνεται:

$$\|A_1\| \le \|H_{s_2}^{-1}\cdot H_{s_1}\|\ c_{s_1}^{(n_1-n_0)} \le \|H_{s_2}^{-1}\cdot H_{s_1}\|\ \frac{c_{s_2}}{c}\ c^{(n_1-n_0)}.$$

Για $j=\ell$, η ίδια σχέση γίνεται:

$$\|A_\ell\| \le c_{s_1}^{(n_\ell-n_{\ell-1})} \le c^{(n_\ell-n_{\ell-1})},$$

γιατί $s_{\ell+1} = s_\ell$. Άρα:

$$\|R'_n\cdot R'_{n-1}\cdot R'_m\| = \|H_{s_\ell}\cdot A_\ell\cdot A_{\ell-1}\cdot\ldots\cdot A_1\cdot H_{s_1}^{-1}\|,$$

απ' όπου συνεπάγεται η ανισοτική σχέση:

$$\|R'_n\cdot R'_{n-1}\cdot R'_m\| \le \|H_{s_\ell}\|\ \|H_{s_1}^{-1}\|\ \|H_{s_2}^{-1}\cdot H_{s_1}\|\ c_{s_1}\ c^{(n-m)}.$$

Θέτοντας:

$$K := \max_{1\le r,s,t\le 6}(\|H_r\|\ \|H_s^{-1}\|\ \|H_t^{-1}.H_s\|\ c_s),$$

προκύπτει αμέσως ότι:

$$\|R'_n\cdot R'_{n-1}\cdot R'_m\| \le K\ c^{(n-m)}.$$

Ο ισχυρισμός του *Θεωρήματος* είναι τώρα μία απλή και ευθεία εφαρμογή του *Λήμματος IV.4.4.(v).4.* ■

Παρατήρηση IV.4.4.(ν).6. Πολλαπλασιάζοντας, εφ' όσον είναι αναγκαίο, κάθε πίνακα H_r με μία κατάλληλα επιλεγμένη σταθερά, μπορούμε να οδηγηθούμε στην ανισότητα:

$$\left\| H_r^{-1} \cdot H_{r+1} \right\| \le \frac{c_{r+1}}{c_r}, \quad \forall r = 1,2,\ldots,5,$$

και, κατ' ακολουθία, στην γενικότερη εκτίμηση:

$$\left\| H_r^{-1} \cdot H_s \right\| \le \frac{c_s}{c_r} \le \frac{c}{c_r}, \ \forall r \le s,$$

με συνέπεια να μπορούμε να πάρουμε:

$$\pi_{s,r} = 1, \quad \forall r \le s,$$

και βεβαιωνόμαστε ότι η δεύτερη ιδιότητα του προηγούμενου *Θεωρήματος* δεν υπεισέρχεται, όταν πρόκειται να αυξηθεί η τάξη. ■

Παρατήρηση IV.4.4.(ν).7. Μπορούμε, επίσης, να πάρουμε:

$$c_1 = c_2 = c_3 \qquad \text{και} \qquad H_1 = H_2 = H_3,$$

με αποτέλεσμα να έχουμε, τότε:

$$\pi_{i,j} = 1, \text{ εάν } 1 \le i, j \le 3.$$

Εάν, τέλος, επιθυμούμε να αυξήσουμε την τάξη κατά μία μόνον μονάδα την φορά, τότε οι μόνοι ακέραιοι

$$\pi_{i,j} \ne 1,$$

που μπορούν να υπεισέλθουν είναι οι:

$$\pi_{3,4}, \quad \pi_{4,5}, \text{ και } \pi_{5,6}. \ ■$$

IV.5. ΣΤΟΙΧΕΙΑ ΘΕΩΡΙΑΣ ΠΡΟΣΕΓΓΙΣΕΩΝ

IV.5.1. ΤΟΠΟΘΕΤΗΣΗ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ

Η *Θεωρία Προσεγγίσεων* αποτελεί ένα θεμελιώδες Κεφάλαιο της Αριθμητικής Ανάλυσης. Ο, τι εξετάσαμε, μέχρι τώρα, θα μπορούσε να τυποποιηθεί και να αντιμετωπιστεί με την βοήθεια αυτής της Θεωρίας. Για παράδειγμα, η προσέγγιση μίας συνάρτησης, μέσα σ' ένα διάστημα, με την χρησιμοποίηση ενός πολυωνύμου παρεμβολής, η προσέγγιση του συναρτησοειδούς της ολοκλήρωσης με ένα πεπερασμένο άθροισμα, η προσέγγιση της λύσης μίας διαφορικής εξίσωσης κ.λ.π. είναι ειδικά θέματα της γενικότατης αυτής Θεωρίας.

Σε ό, τι ακολουθήσει, θα επιχειρήσουμε να παρουσιάσουμε τις βασικές έννοιες της *Θεωρίας Προσεγγίσεων*. Εξυπακούεται ότι η εκτενής ανάπτυξη της Θεωρίας, όπως επίσης και της συγγενούς προς αυτήν *Θεωρίας της Βελτιστοποίησης*, διαφεύγει από το περιορισμένο πλαίσιο ανάπτυξης το οποίο είμαστε υποχρεωμένοι, για ευνόητους λόγους, να σεβαστούμε. Όποιος αναγνώστης ενδιαφέρεται να επιδιώξει μία πιο εμπεριστατωμένη άποψη επί των παραπάνω θεμάτων μπορεί να συμβουλευτεί την βιβλιογραφική παραπομπή [41], από την οποία έχει εξαχθεί το γνωστικό υλικό αυτού του *Κεφαλαίου*. Μία απλή παρουσίαση των ίδιων ζητημάτων βρίσκεται επίσης στο βιβλίο [26].

IV.5.2. ΠΕΡΙ ΒΕΛΤΙΣΤΗΣ ΠΡΟΣΕΓΓΙΣΗΣ

Το πρόβλημα της αναζήτησης της βέλτιστης προσέγγισης μπορεί να διατυπωθεί με τον ακόλουθο γενικό τρόπο:

Ορισμός IV.5.2.1. ***Έστω $V=(V,+,\cdot)$ ένας διανυσματικός χώρος με νόρμα $\|\ldots\|$, υπεράνω του $K=(\mathbb{R},\mathbb{C})$ και έστω $E\subset V$. Λέμε ότι το σημείο $\bar{g}\in E$ είναι η βέλτιστη προσέγγιση του σημείου $f\in V$ μέσα στο V, εάν:***

$$\|\bar{g}-f\|=\min_{g\in E}\|g-f\|.$$

Το σημείο $\bar{g}$ είναι η ορθογώνια προβολή f επί του E. ■

Ανάλογα με τις υποθέσεις, που κάνουμε για τα σύνολα V *και* E, καταλήγουμε σε διάφορα θεωρητικά αποτελέσματα, πού είναι κλασσικές προτάσεις της Μαθηματικής Ανάλυσης ([55]).

IV.5.2.(i). Χαρακτηρισμός, Υπαρξη και Μοναδικότητα

Θεώρημα IV.5.2.(i).1. ***Εάν το σύνολο E είναι ένα συμπαγές κομμάτι του V, τότε για οποιοδήποτε $f\in V$, υπάρχει τουλάχιστον μία βέλτιστη προσέγγιση $\bar{g}$ του f επί του V.***

Απόδειξη. Το μέγεθος $\|g-f\|$ είναι οπωσδήποτε μη αρνητικό, για οποιαδήποτε $g \in E$ και $f \in V$.

Ας σταθεροποιήσουμε το σημείο f του V και ας συμβολίσουμε με $d(\geq 0)$ το μέγιστο κάτω φράγμα της ποσότητας $\|g-f\|$, καθώς το g διατρέχει τα σημεία του E.

Από τον ορισμό της έννοιας του μέγιστου κάτω φράγματος, βλέπουμε ότι είναι δυνατή η εξαγωγή μιάς ακολουθίας σημείων $(g_n : n = 0,1,2,...)$ του E τέτοιων ώστε:

$$\lim_{n\to\infty}\|g_n - f\| = d.$$

Επειδή το σύνολο E είναι συμπαγές, υπάρχει μία υπακολουθία της ακολουθίας $(g_n : n = 0,1,2,...)$ η οποία συγκλίνει προς ένα σημείο $\bar{g}$ του E. Είναι τότε προφανές ότι:

$$\|\bar{g} - f\| \geq d.$$

Από την άλλη πλευρά, καθώς για κάθε $n \geq 0$ ισχύει:

$$\|\bar{g} - f\| \leq \|g_n - f\| + \|\bar{g} - g_n\|,$$

λαμβάνουμε, αφού αφήσουμε το n να αυξηθεί απεριόριστα:

$$\|\bar{g} - f\| \leq d.$$

Κατά συνέπεια, ισχύει ότι:

$$\|\bar{g} - f\| = d = \min_{g\in E}\|g - f\|,$$

και επομένως το σημείο $\bar{g}$ συνιστά μία βέλτιστη προσέγγιση του f επί του E. ■

Σαν σημαντική συνέπεια του πιο πάνω *Θεωρήματος*, αποδεικνύουμε το αποτέλεσμα που ακολουθεί:

Θεώρημα IV.5.2.(i).2. ***Εάν το σύνολο E είναι ένας διανυσματικός υπόχωρος πεπερασμένης διάστασης του V, τότε, για οποιοδήποτε σημείο $f \in V$, υπάρχει τουλάχιστον μία βέλτιστη προσέγγιση $\bar{g}$ του f επί του V.***

Απόδειξη. Σύμφωνα με το προηγούμενο *Θεώρημα*, αρκεί να δείξουμε ότι το $\bar{g}$ ανήκει σε ένα συμπαγές κομμάτι του E.

Προς τούτο, ας θεωρήσουμε ένα σταθερό σημείο f του V, και ας θέσουμε:

$$\|f\| = \|f - 0\| = a,$$

όπου 0 είναι το μηδενικό στοιχείο του διανυσματικού χώρου V.

Τότε, για κάθε σημείο g του E, που είναι τέτοιο ώστε $\|g\| = \|g - 0\| > 2a$, ισχύει:

$$\|f - g\| \geq \|f - 0\| - \|g - 0\| > a.$$

Επομένως, το σημείο $\bar{g}$ δεν μπορεί να βρίσκεται παρά μόνον στο εσωτερικό της κλειστής μπάλας, που έχει κέντρο το $0 \in E \subset V$ και ακτίνα ίση με $2a$. Αυτή η μπάλα αποτελεί ένα κλειστό και φραγμένο μέρος του διανυσματικού χώρου, πεπερασμένης διάστασης, E, και άρα, είναι ένα συμπαγές υποσύνολο του E. ■

Θα προχωρήσουμε τώρα δίνοντας ένα κριτήριο χαρακτηρισμού της βέλτιστης προσέγγισης σε μία ειδική περίπτωση.

Προηγουμένως, χρειαζόμαστε τον εξής *Ορισμό*:

Ορισμός IV.5.2.(i).3. ***Έστω*** $V=(V,+,\cdot)$ ***ένας διανυσματικός χώρος υπεράνω του*** $\mathbb{R}$***. Λέμε ότι το σύνολο*** V ***είναι ένας προ-Hilbert χώρος, εάν είναι εφοδιασμένο με ένα εσωτερικό γινόμενο, δηλαδή με μία απεικόνιση:***

$$\langle\cdot/\cdot\rangle : V\times V\to\mathbb{R} : (f,g)\mapsto\langle f/g\rangle,$$

που έχει τις ακόλουθες ιδιότητες:

$$\langle f/\lambda\cdot g+\mu\cdot h\rangle=\lambda\langle f/g\rangle+\mu\langle f/h\rangle \quad \forall\lambda,\mu\in\mathbb{R},$$

$$\langle f/g\rangle=\langle g/h\rangle \text{ (:συμμετρία),}$$

$$\langle f/f\rangle\geq 0 \text{ (:θετικότητα),}$$

$$\langle f/f\rangle=0\Leftrightarrow f=0.$$

Εφοδιασμένος με την νόρμα $\|f\|:=\langle f/f\rangle^{\frac{1}{2}}$***, ο προ-Hilbert χώρος είναι ένας χώρος με νόρμα. Εάν, επί πλέον, είναι πλήρης, ως προς αυτήν την νόρμα, τότε λέγεται χώρος Hilbert.*** ■

Έχουμε το ακόλουθο αποτέλεσμα:

Θεώρημα IV.5.2.(i).4. ***Εάν ο διανυσματικός χώρος*** V ***είναι ένας χώρος προ-Hilbert και εάν το σύνολο*** E ***είναι ένας διανυσματικός υπόχωρος, πεπερασμένης διάστασης, του*** V***, τότε μία αναγκαία και ικανή συνθήκη προκειμένου ένα σημείο*** $\bar{g}\in E$ ***να συνιστά την βέλτιστη προσέγγιση προς το στοιχείο*** $f\in V$ ***είναι η ιδιότητα:***

$$\langle\bar{g}-f/\bar{g}-g\rangle\leq 0 \qquad \forall g\in E.$$

Απόδειξη. Κατ' αρχάς, θα αποδείξουμε την αναγκαιότητα της συνθήκης. Έστω ότι σημείο $\bar{g} \in E$ αποτελεί την βέλτιστη προσέγγιση του f επί του V.

Ας υποθέσουμε ότι υπάρχει ένα στοιχείο $g_1 \in E$ τέτοιο ώστε:

$$\langle \bar{g} - f / \bar{g} - g_1 \rangle > 0,$$

και ας θεωρήσουμε το σημείο $g_2 \in E$ που ορίζεται από τον *Τύπο*:

$$g_2 = \bar{g} + \lambda\,(g_1 - g),$$

με:

$$\lambda = \frac{\langle \bar{g} - f / \bar{g} - g_1 \rangle}{\|g_1 - \bar{g}\|^2}.$$

Τότε θα είναι:

$$\begin{aligned} \|f - g_2\|^2 &= \langle f - g_2 / f - g_2 \rangle \\ &= \langle f - \bar{g} - \lambda \cdot (g_1 - g) / f - \bar{g} - \lambda \cdot (g_1 - g) \rangle \\ &= \|f - \bar{g}\|^2 - 2\,\lambda\,\langle f - \bar{g} / g_1 - \bar{g} \rangle + \lambda^2\,\|g_1 - \bar{g}\|^2, \end{aligned}$$

και, αφού αντικαταστήσουμε το σύμβολο λ με την τιμή του, θα προκύψει ότι:

$$\|f - g_2\|^2 = \|f - \bar{g}\|^2 - \frac{\langle f - \bar{g} / g_1 - \bar{g} \rangle^2}{\|g_1 - \bar{g}\|^2} < \|f - \bar{g}\|^2,$$

που είναι αντίθετη διαπίστωση προς την αρχική αξίωση ότι το σημείο $\bar{g} \in E$ συνιστά την βέλτιστη προσέγγιση του f επί του V.

Κατά συνέπεια, εάν το $\bar{g}$ είναι βέλτιστη προσέγγιση του f επί του V, πρέπει να ισχύει:

$$\langle \bar{g} - f / \bar{g} - g \rangle \le 0 \qquad \forall g \in E.$$

Θα δείξουμε τώρα την επάρκεια (δηλαδή την ικανότητα) της συνθήκης του *Θεωρήματος*.

Επειδή για κάθε $\overline{g} \in E, g \in E$ και κάθε $f \in V$, είναι:

$$\begin{aligned}\|f-g\|^2 &= \|(f-\overline{g})-(g-\overline{g})\|^2 \\ &= \langle f-\overline{g}-(g-\overline{g})/f-\overline{g}-(g-\overline{g})\rangle \\ &= \|f-g\|^2 - 2\langle \overline{g}-f/\overline{g}-g\rangle + \|g-\overline{g}\|^2,\end{aligned}$$

βλέπουμε ότι εάν $\langle \overline{g}-f/\overline{g}-g\rangle \le 0 \ \ \forall g \in E$ τότε:

$$\|f-g\|^2 \ge \|f-\overline{g}\|^2 \qquad \forall g \in E,$$

δηλαδή το σημείο $\overline{g}$ είναι η βέλτιστη προσέγγιση του f επί του E. ■

Παρατήρηση IV.5.2.(i).5. Η συνθήκη:

$$\langle \overline{g}-f/\overline{g}-g\rangle = 0 \qquad \forall g \in E.$$

του προηγούμενου *Θεωρήματος* ισοδυναμεί με την συνθήκη:

$$\langle \overline{g}-f/g\rangle = 0 \qquad \forall g \in E.$$

Πράγματι, ας θέσουμε $h := \overline{g}-g$.

Αφού η συνθήκη του *Θεωρήματος IV.5.2.(i).4* αφορά κάθε σημείο $g \in E$, θα αφορά ειδικότερα και κάθε σημείο h που ανήκει στο E. Με άλλα λόγια, θα μπορεί να διατυπωθεί υπό την μορφή:

$$\langle \overline{g}-f/h\rangle \le 0 \qquad \forall h \in E.$$

Καθώς, όμως, το σύνολο E είναι ένας διανυσματικός υπόχωρος του V, θα ισχύει ότι $-h \in E$, και, κατά συνέπεια, από την παραπάνω ανισότητα θα έπεται ότι:

$$\langle \overline{g}-f/-h\rangle \le 0 \qquad \forall h \in E.$$

Επομένως, συμπεραίνουμε πως θα είναι:

$$\langle \overline{g}-f/g\rangle = 0 \qquad \forall g \in E. \ ■$$

Μετά από την απόδειξη της ύπαρξης της βέλτιστης προσέγγισης, αλλά και την ανακάλυψη της χαρακτηριστικής της ιδιότητας, θα στρέψουμε το ενδιαφέρον μας στην μοναδικότητά της. Σχετικά, έχουμε το ακόλουθο:

Θεώρημα IV.5.2.(i).6. ***Εάν ο διανυσματικός χώρος V είναι ένας χώρος προ-Hilbert και εάν το σύνολο E είναι ένας διανυσματικός υπόχωρος, πεπερασμένης διάστασης του V, τότε η βέλτιστη προσέγγιση $\bar{g} \in E$ του σημείου $f \in V$ είναι μοναδική.***

Απόδειξη. Ας υποθέσουμε ότι υπάρχουν $\bar{g}_1 \in E$ και $\bar{g}_2 \in E$, που είναι ταυτόχρονα δύο βέλτιστες προσεγγίσεις ενός σημείου $f \in V$ επί του E, και ας θέσουμε:

$$g := \bar{g}_1 - \bar{g}_2 .$$

Τότε, θα ήταν:

$$\begin{aligned} \left\| \bar{g}_1 - \bar{g}_2 \right\|^2 &= \langle \bar{g}_1 - f - (\bar{g}_2 - f) / g \rangle \\ &= \langle \bar{g}_1 - f / g \rangle - \langle \bar{g}_2 - f / g \rangle , \end{aligned}$$

κι επειδή, σύμφωνα με την πιο πάνω *Παρατήρηση* και το *Θεώρημα IV.5.2.(i).4*:

$$\langle \bar{g}_1 - f / g \rangle = \langle \bar{g}_2 - f / g \rangle = 0 ,$$

θα είχαμε:

$$\left\| \bar{g}_1 - \bar{g}_2 \right\|^2 = 0 ,$$

και άρα $g_1 - g_2$, που ολοκληρώνει την *Απόδειξη* του *Θεωρήματος*. ■

IV.5.2.(ii). Κατασκευή της Βέλτιστης Προσέγγισης Θα προχωρήσουμε τώρα στην εξέταση του τρόπου αποτελεσματικής κατασκευής της βέλτιστης προσέγγισης $\overline{g} \in E$ ενός σημείου $f \in V$, στην περίπτωση που το σύνολο E είναι ένας διανυσματικός υπόχωρος, πεπερασμένης διάστασης, ενός χώρου προ-Hilbert V.

Προς τούτο, ας υποθέσουμε ότι τα στοιχεία:

$$g_1, g_2, \ldots, g_n$$

του προ-Hilbert χώρου V είναι γραμμικά ανεξάρτητα, και ας θεωρήσουμε τον διανυσματικό υπόχωρο E του V, που παράγεται από όλους τους πεπερασμένους γραμμικούς συνδυασμούς των $g_1, g_2, \ldots, g_n$.

Θεώρημα IV.5.2.(ii).1. ***Μία αναγκαία και ικανή συνθήκη προκειμένου τα στοιχεία $g_1, g_2, \ldots, g_n$ του V να αποτελούν ένα γραμμικά ανεξάρτητο σύστημα είναι η ορίζουσα Gram:***

$$\Delta(g_1, \ldots, g_n) = \det \begin{pmatrix} \langle g_1 / g_1 \rangle & \ldots & \langle g_1 / g_n \rangle \\ \ldots & \ldots & \ldots \\ \langle g_n / g_1 \rangle & \ldots & \langle g_n / g_n \rangle \end{pmatrix}$$

να είναι διάφορη του μηδενός.

Απόδειξη. Ας υποθέσουμε, κατ' αρχάς ότι $\Delta(g_1, g_2, \ldots, g_n) \neq 0$. Εάν οι πραγματικοί αριθμοί $\lambda_1, \lambda_2, \ldots, \lambda_n$ είναι τέτοιοι ώστε να ισχύει η ισότητα:

$$\sum_{i=1}^{n} \lambda_i \cdot g_i = 0,$$

τότε πολλαπλασιάζοντας εσωτερικά από τα αριστερά με το στοιχείο g_j, παίρνουμε:

$$\sum_{i=1}^{n} \lambda_i \langle g_i / g_j \rangle = 0 \text{, με } j = 1,2,\ldots,n .$$

Οι n εξισώσεις, που προκύπτουν κατ' αυτόν τον τρόπο, συνιστούν ένα σύστημα που απαρτίζεται από n ομογενείς γραμμικές εξισώσεις με n αγνώστους: τους αριθμούς

$$\lambda_1, \lambda_2, \ldots, \lambda_n .$$

Επειδή η ορίζουσά του είναι διάφορη του μηδενός, το σύστημα αυτό δεν επιδέχεται παρά μόνον την μηδενική λύση:

$$\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0 .$$

Το συμπέρασμα αυτό πιστοποιεί την γραμμική ανεξαρτησία των

$$g_1, g_2, \ldots, g_n .$$

Απομένει να αποδειχτεί η αναγκαιότητα της συνθήκης του *Θεωρήματος*. Προς τούτο, παρατηρούμε ότι εάν τα στοιχεία $g_1, g_2, \ldots, g_n$ είναι γραμμικά ανεξάρτητα, τότε το παραπάνω σύστημα:

$$\sum_{i=1}^{n} \lambda_i \langle g_i / g_j \rangle = 0 \text{, με } j = 1,2,\ldots,n$$

δεν επιδέχεται παρά μόνον την μηδενική λύση:

$$\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0 .$$

Πράγματι, εάν υπήρχε μία μη μηδενική λύση, τότε το σημείο $\sum_{i=1}^{n} \lambda_i \cdot g_i$ θα ήταν ορθογώνιο προς κάθε σημείο g_j ($j = 1,2,\ldots,n$) και, από αυτό, θα συνεπάγονταν ότι:

$$\langle \sum_{i=1}^{n} \lambda_i \cdot g_i / \sum_{j=1}^{n} \lambda_j \cdot g_j \rangle = \left\| \sum_{i=1}^{n} \lambda_i \cdot g_i \right\|^2 = 0 ,$$

που θα σήμαινε ότι θα υπήρχαν $\lambda_1, \ldots, \lambda_n$, όχι όλα μηδέν, τέτοια ώστε:

$$\sum_{i=1}^{n} \lambda_i \cdot g_i = 0 ,$$

δηλαδή θα σήμαινε ότι τα στοιχεία $g_1, g_2, \ldots, g_n$ δεν είναι γραμμικά ανεξάρτητα. Άρα, η μοναδική λύση του συστήματος:

$$\sum_{i=1}^{n} \lambda_i \langle g_i / g_j \rangle = 0 \quad \text{με} \quad j = 1,2,\ldots,n$$

είναι η μηδενική, και, κατά συνέπεια, πρέπει να ισχύει:

$$\Delta(g_1, g_2, \ldots, g_n) \neq 0.$$

Η *Απόδειξη* του *Θεωρήματος* είναι πλήρης. ∎

Θεώρημα IV.5.2.(ii).2. ***Η βέλτιστη προσέγγιση $\bar{g}$ του σημείου f επί του E $(f \notin E)$ γράφεται υπό την μορφή:***

$$\bar{g} = \sum_{i=1}^{n} \lambda_i \cdot g_i \,,$$

όπου οι πραγματικοί αριθμοί λ_i είναι λύσεις του συστήματος εξισώσεων:

$$\sum_{i=1}^{n} \lambda_i \langle g_i / g_j \rangle = \langle f / g_j \rangle \,, \text{ με } j = 1,2,\ldots,n.$$

Επί πλέον, ισχύει:

$$\|f - \bar{g}\|^2 = \frac{\Delta(g_1, \ldots, g_n, f)}{\Delta(g_1, \ldots, g_n)}.$$

Απόδειξη. Σύμφωνα με το *Θεώρημα IV.5.2.(i).4* και την *Παρατήρηση IV.5.2.(i).5*, η βέλτιστη προσέγγιση $\bar{g}$ είναι η προβολή της f επί του E. Συνεπώς, η διαφορά $\bar{g} - f$ είναι ορθογώνια επί του E και, ειδικότερα επί των στοιχείων $g_1, \ldots g_n$ του E:

$$\langle \sum_{i=1}^{n} \lambda_i g_i - f / g_j \rangle = 0 \quad \forall j = 1,2,\ldots,n$$

$$\Leftrightarrow$$

$$\sum_{i=1}^{n} \lambda_i \langle g_i / g_j \rangle = \langle f / g_j \rangle \quad \forall j = 1,2,\ldots,n.$$

Από την άλλη πλευρά, ισχύει:

$$\|f - \overline{g}\|^2 = \langle f - \overline{g} / f - \overline{g} \rangle = \langle f - \overline{g} / f \rangle,$$

επειδή τα στοιχεία $f - \overline{g}$ και $\overline{g}$ είναι ορθογώνια μεταξύ τους. Θέτοντας

$$d := \|f - g\|,$$

βλέπουμε ότι:

$$d^2 = \|f\|^2 - \sum_{i=1}^{n} \lambda_i \langle g_i / f \rangle$$

ή, ακόμη, ότι:

$$\sum_{i=1}^{n} \lambda_i \langle g_i / f \rangle = \|f\|^2 - d^2.$$

Επισυνάπτοντας αυτήν την εξίσωση στο παραπάνω σύστημα, επιτυγχάνουμε τον σχηματισμό ενός ευάρμοστου συστήματος, με $(n+1)$ γραμμικές εξισώσεις και n αγνώστους, του οποίου η ορίζουσα είναι μηδενική:

$$\det \begin{pmatrix} \langle g_1 / g_1 \rangle & \ldots & \langle g_n / g_1 \rangle & \langle f / g_1 \rangle \\ \ldots & \ldots & \ldots & \ldots \\ \langle g_1 / g_n \rangle & \ldots & \langle g_n / g_n \rangle & \langle f / g_n \rangle \\ \langle g_1 / f \rangle & \ldots & \langle g_n / f \rangle & \langle f / f \rangle - d^2 \end{pmatrix} = 0,$$

από όπου έπεται το αποτέλεσμα του *Θεωρήματος*, εφ' όσον χρησιμοποιηθεί η γραμμικότητα της ορίζουσας ως προς την τελευταία στήλη της. ■

Παρατήρηση IV.5.2.(ii).3. Είναι σαφές ότι το σύστημα εξισώσεων του *Θεωρήματος IV.5.2.(ii).2*, που επιτρέπει τον υπολογισμό των αριθμών λ_i, ανάγεται σε ένα διαγώνιο σύστημα γραμμικών εξισώσεων, όταν:

$$\langle g_i / g_j \rangle = 0 \quad \forall i \neq j \,.$$

Σε μία τέτοια περίπτωση λέμε ότι τα σημεία $g_1, g_2, \ldots, g_n$ συνιστούν ένα ***ορθογώνιο σύστημα***.■

Μετά από την παραπάνω επισήμανση, καταλαβαίνουμε την σπουδαιότητα των ορθογωνίων συστημάτων στην Θεωρία Προσεγγίσεων.

Έχουμε ήδη διαμορφώσει μία μερική συνείδηση της διαπίστωσης αυτής, στην *Παράγραφο IV.2.9* και στην *Παράγραφο IV.2.10*, όταν εξετάζαμε τις αριθμητικές μεθόδους τετραγωνισμού Gauss.

Από την άλλη πλευρά, είναι εύκολο να επιβεβαιωθεί ότι εάν τα $g_1, g_2, \ldots, g_n$ συνιστούν ένα ορθογώνιο σύστημα τότε είναι γραμμικά ανεξάρτητα μεταξύ τους, και ισχύει:

$$\lambda_i = \frac{\langle f / g_i \rangle}{\langle g_i / g_i \rangle} \quad \forall i = 1,2,\ldots,n \,.$$

Εάν, επί πλέον, το ορθογώνιο σύστημα των g_i είναι και ***ορθοκανονικό***, δηλαδή εάν:

$$\langle g_i / g_i \rangle = 1 \quad \forall i = 1,2,\ldots,n \,,$$

τότε ισχύει:

$$\lambda_i = \langle f / g_i \rangle \quad \forall i = 1,2,\ldots,n \quad \text{και} \quad \|f - \bar{g}\|^2 = \|f\|^2 - \sum_{i=1}^{n} \lambda_i^2 \ ,$$

και οι αριθμοί $\lambda_1, \lambda_2, \ldots, \lambda_n$ ονομάζονται ***συντελεστές Fourier του στοιχείου*** f ***ως προς το ορθοκανονικό σύστημα*** $g_1, g_2, \ldots, g_n$, ενώ η βέλτιστη προσέγγιση $\bar{g}$ καλείται ***άθροισμα Fourier***.

Υπενθυμίζουμε, τέλος, ότι, δοθείσης μίας τυχαίας βάσης $g_1, g_2, \ldots, g_n$ του *E*, είναι πάντοτε δυνατή η μετατροπή της σε ορθοκανονική βάση $h_1, h_2, \ldots, h_n$ του *E*, με την βοήθεια της ***μεθόδου ορθογωνοποίησης Schmidt***, σύμφωνα με την οποία:

$$h_1 := \frac{g_1}{\|g_1\|},$$

$$u_2 := g_2 - \langle g_2 / h_1 \rangle \cdot h_1,$$

$$h_2 := \frac{u_2}{\|u_2\|},$$

και γενικότερα:

$$u_i := g_i - \sum_{j=1}^{i} < g_i / h_j > \cdot h_j \ ,$$

$$h_i := \frac{u_i}{\|u_i\|} \qquad \forall i = 1,2,\ldots,n.$$

IV.5.3. ΠΑΡΑΔΕΙΓΜΑΤΑ ΠΡΟΣΕΓΓΙΣΕΩΝ

Στην *Παράγραφο* αυτή θα παρουσιάσουμε δύο βασικές κατευθύνσεις παραδειγμάτων προσεγγίσεων.

Στην πρώτη κατεύθυνση, ο διανυσματικός χώρος V θα είναι ένας χώρος προ-Hilbert.

Θυμίζουμε ότι, σε αυτήν την περίπτωση, έχουμε ήδη παρουσιάσει το χαρακτηριστικό γνώρισμα της βέλτιστης προσέγγισης, όπως επίσης και την προϋπόθεση που εξασφαλίζει την μοναδικότητά της.

Ακόμη, σ'αυτήν την περίπτωση, έχουμε ήδη προτείνει έναν τρόπο κατασκευής μίας τέτοιας προσέγγισης.

Κατά την δεύτερη κατεύθυνση, το σύνολο V θα είναι απλά ένας διανυσματικός χώρος με νόρμα. Και στις δύο κατευθύνσεις, το πρόβλημα της πολυωνυμικής παρεμβολής θα εμφανιστεί σαν ειδική περίπτωση.

IV.5.3.(i). Προσέγγιση με την έννοια των Ελαχίστων Τετραγώνων

Θα διαχωρίσουμε την συνεχή περίπτωση, που θα εξετάσουμε πρώτα, και την διακριτή περίπτωση, που θα ακολουθήσει.

IV.5.3.(i).(α). Συνεχής Προσέγγιση

Έστω $C[a,\beta]$ ο διανυσματικός χώρος όλων των πραγματικών συναρτήσεων, που είναι συνεχείς επί του διαστήματος $[a,\beta]$ και έστω $w(x)$ μία θετική συνάρτηση βάρους επί

του $[a,\beta]$***, τέτοια ώστε το ολοκλήρωμα:***

$$\int_a^\beta f(x)\, w(x)\, dx$$

να υπάρχει, για κάθε $f \in C[a,\beta]$***. Ορίζουμε εσωτερικό γινόμενο στον χώρο*** $C[a,\beta]$***, θέτοντας:***

$$\langle f / g \rangle = \int_a^\beta f(x)\, g(x)\, w(x)\, dx,$$

για κάθε $f \in C[a,\beta]$ ***και για κάθε*** $g \in C[a,\beta]$***. Ας υποθέσουμε ότι τα*** $(n+1)$ ***στοιχεία του*** $C[a,\beta]$***:***

$$g_0, g_1, \ldots, g_n$$

είναι γραμμικά ανεξάρτητα μεταξύ τους, και ας θεωρήσουμε τον διανυσματικό υπόχωρο E ***που αυτά παράγουν. Η βέλτιστη προσέγγιση*** $\bar{g} \in E$ ***μίας συνάρτησης*** $f \in C[a,\beta] - E$ ***ορίζεται, τότε, από την σχέση:***

$$\int_a^\beta [f(x) - \bar{g}(x)]^2\, w(x)\, dx = \min_{g \in E} \int_a^\beta [f(x) - g(x)]^2\, w(x)\, dx,$$

και, σύμφωνα με τα αποτελέσματα της προηγούμενης Παραγράφου:

η συνάρτηση $\bar{g}(x)$ ***υπάρχει, είναι μοναδική, και ισούται με:***

$$\bar{g}(x) = \lambda_0\, g_0(x) + \ldots + \lambda_n\, g_n(x) \quad (x \in [a,\beta]),$$

όπου οι συντελεστές $\lambda_i \in \mathbb{R}$ ***είναι οι λύσεις του συστήματος:***

$$\left\| \sum_{i=0}^n \lambda_i \int_a^\beta g_i(x)\, g_j(x)\, w(x)\, dx = \int_a^\beta f(x)\, g_j(x)\, w(x)\, dx \quad (j = 0,1,\ldots,n).$$

Παράδειγμα IV.5.3.(i).(α).1. Πολυωνυμική Προσέγγιση *Εάν οι συναρτήσεις* $g_0(x),\ldots,g_n(x)$ *είναι πολυωνυμικές, τότε η βέλτιστη προσέγγιση* $\bar{g}$ *είναι επίσης πολυώνυμο. Συνήθως, λαμβάνουμε:*

$$g_0(x)=1\ ,\quad g_1(x)=x,\ldots,g_n(x)=x^n,$$

όμως, όπως είναι φανερό, η καλύτερη επιλογή των πολυωνυμικών συναρτήσεων είναι εκείνη που περιέχει ορθογώνιες πολυωνυμικές συναρτήσεις:

$$\int_a^\beta g_i(x)\ g_j(x)\ w(x)\ dx=0 \quad \forall i\neq j\,.$$

Το προφανές αυτό επιχείρημα καταμαρτυρά την σπουδαιότητα που αποδίδεται στα συστήματα ορθογωνίων πολυωνύμων, όσον αφορά τον ρόλο που διαδραματίζουν στην αποτελεσματικότητα και, γενικότερα, στην Θεωρία των Πολυωνυμικών Προσεγγίσεων.■

Παράδειγμα IV.5.3.(i).(α).2: Τριγωνομετρική Προσέγγιση *Εάν* $[a,\beta]=[-\pi,\pi]$*, και οι συναρτήσεις* $g_i(x)$ *είναι οι τριγωνομετρικές:*

$$g_0(x)=1\ \ ,\ \ g_1(x)=\sin x\ \ ,\ \ g_2(x)=\cos x\,,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$g_{2n-1}(x)=\sin(n\,x)\ \ ,\ \ g_{2n}(x)=\cos(n\,x),$$

τότε, η βέλτιστη προσέγγιση γράφεται υπό την μορφή:

$$\bar{g}(x)=a_0+\sum_{k=1}^{n}[a_k\ \cos(k\ x)+\beta_k\ \sin(k\,x)]\,.$$

Αξίζει να σημειωθεί ότι σε αυτήν την περίπτωση, οι συναρτήσεις

$$g_0(x),g_1(x),\ldots,g_{2n}(x)$$

συνιστούν ένα ορθοκανονικό σύστημα. ■

IV.5.3.(i).(β). Διακριτή Προσέγγιση

Θα εξετάσουμε τώρα την περίπτωση της διακριτής προσέγγισης, με την έννοια των ελαχίστων τετραγώνων. Τούτη την φορά, ο διανυσματικός χώρος V θα είναι το σύνολο $F[a,\beta]$ ***όλων των πραγματικών συναρτήσεων που ορίζονται στο κλειστό διάστημα*** $[a,\beta]$***, εφοδιασμένο με τις συνηθισμένες σημειακές πράξεις πρόσθεσης και αριθμητικού πολλαπλασιασμού.***

Ας υποθέσουμε ότι η συνάρτηση βάρους $w(x) \in F\,[a,\beta]$ ***είναι θετική επί του*** $[a,\beta]$ ***και ας θεωρήσουμε τα σημεία:***

$$x_0, x_1, \ldots, x_N$$

του $[a,\beta]$***. Στον διανυσματικό χώρο*** $F[a,\beta]$ ***μπορούμε να εισάγουμε εσωτερικό γινόμενο, θέτοντας:***

$$\langle f / g \rangle := \sum_{i=0}^{N} f(x_i)\, g(x_i)\, w(x_i),$$

για κάθε $f \in F[a,\beta]$ ***και κάθε*** $g \in F[a,\beta]$***.***

Εάν τώρα E ***είναι ένας οποιοσδήποτε διανυσματικός υπόχωρος του*** $F[a,\beta]$***, τότε η βέλτιστη προσέγγιση*** $\bar{g}$ ***μίας συνάρτησης*** $f \in F[a,\beta] - E$ ***επί του*** E ***ορίζεται από την σχέση:***

$$\sum_{i=0}^{N} \{f(x_i) - \bar{g}(x_i)\}^2\, w(x_i) = \min_{g \in E} \sum_{i=0}^{N} \{f(x_i) - g(x_i)\}^2\, w(x_i).$$

Σύμφωνα με τα αποτελέσματα της Παραγράφου IV.5.2, μία τέτοια βέλτιστη προσέγγιση

υπάρχει και είναι μοναδική, όταν το σύνολο E ***είναι ο διανυσματικός υπόχωρος που παράγεται από*** $(n+1)$ ***στοιχεία του*** $F[a,\beta]$***:***

$$g_0, g_1, \cdots, g_n,$$

που είναι γραμμικά ανεξάρτητα μεταξύ τους. Εξ' άλλου, τότε ισχύει ότι:

$$\bar{g}(x) = a_0\ g_1(x) + \ldots + a_n\ g_n(x) \qquad \forall x \in [a,\beta],$$

όπου οι αριθμοί a_i ***είναι λύση του συστήματος:***

$$\left\| \sum_{i=0}^{n} a_i \sum_{k=0}^{N} g_i(x_k)\, g_j(x_k)\, w(x_k) = \sum_{k=0}^{N} f(x_k)\, g_j(x_k)\, w(x_k) \quad (j = 0,1,\ldots,n).\right.$$

Ακόμη, χρησιμοποιώντας τους ακόλουθους τρείς συμβολισμούς:

$$A = \begin{pmatrix} \sqrt{w(x_0)}\ g_0(x_0) & \ldots & \sqrt{w(x_0)}\ g_n(x_0) \\ \ldots\ldots & \ldots\ldots & \ldots\ldots \\ \sqrt{w(x_N)}\ g_0(x_N) & \ldots & \sqrt{w(x_N)}\ g_n(x_N) \end{pmatrix}$$

$$\tilde{\beta} = \begin{pmatrix} \sqrt{w(x_0)}\ f(x_0) \\ \vdots \\ \sqrt{w(x_N)}\ f(x_N) \end{pmatrix} \quad \textbf{και} \quad \tilde{a} = \begin{pmatrix} a_0 \\ \vdots \\ a_n \end{pmatrix},$$

το παραπάνω σύστημα γράφεται υπό την μορφή:

$$A^T \cdot A \cdot \tilde{a} = A^T \cdot \tilde{\beta}.$$

Παράδειγμα IV.5.3.(i).(β).1. Πολυωνυμική Προσέγγιση *Εάν οι συναρτήσεις* $g_0(x),\ldots,g_n(x)$ *είναι το μονώνυμα:*

$$g_0(x)=1, g_1(x)=x,\ldots,g_n(x)=x^n,$$

τότε οδηγούμαστε στην διακριτή πολυωνυμική προσέγγιση. Ιδιαίτερα, όταν $N=n$*, ξαναβρίσκουμε την πολυωνυμική παρεμβολή.*

Σημειώστε ότι σε αυτήν την περίπτωση, είναι:

$$\|f-\bar{g}\|=0$$

(όπου, όπως πάντα, η νόρμα $\|u\|$ *του* $u\in F[a,\beta]$ *ορίζεται από τον Τύπο:*

$$\|u\|^2=\langle u/u\rangle$$ *).■*

Παράδειγμα IV.5.3.(i).(α).2: Τριγωνομετρική Προσέγγιση *Εάν* $N=2n,\ [a,\beta]=[-\pi,\pi]$ *και ξαναθεωρήσουμε τις συναρτήσεις* $g_i(x)$ *του Παραδείγματος IV.5.3.(i).(α).2, τότε οδηγούμαστε στο μερικό άθροισμα του αναπτύγματος σε σειρά Fourier της* f *.■*

IV.5.3.(i).(γ). Προσέγγιση με Συναρτήσεις Splines

Ας θεωρήσουμε τώρα, σαν διανυσματικό χώρο V, τον χώρο:

$$H^q[a,\beta] \quad (q \geq 1),$$

όλων των πραγματικών συναρτήσεων f επί του διαστήματος $[a,\beta]$, που έχουν την $(q-1)-$τάξης παράγωγό τους απολύτως συνεχή στην ανοικτή περιοχή του $[a,\beta]$ και την τετραγωνική δύναμη της q-τάξης παράγωγό τους (στην περιοχή του $[a,\beta]$) ολοκληρώσιμη στο $[a,\beta]$:

$$\int_a^\beta [f^{(q)}(x)]^2\, dx < \infty.$$

Στον διανυσματικό χώρο $H^q[a,\beta]$, εισάγουμε το ακόλουθο εσωτερικό γινόμενο:

$$\langle f/g \rangle = \int_a^\beta f^{(q)}(x)\, g^{(q)}(x)\, dx + \rho \sum_{i=1}^N f(x_i)\, g(x_i),$$

για κάθε $f \in H^q[a,\beta]$, $g \in H^q[a,\beta]$ και όπου ρ είναι μία μη-αρνητική παράμετρος και τα σημεία $x_1, x_2, \ldots, x_N$ ανήκουν στο $[a,\beta]$. Εάν E είναι ένας διανυσματικός υπόχωρος του $H^q[a,\beta]$ και εάν $f \in H^q[a,\beta] - E$ τότε η βέλτιστη προσέγγιση $\overline{g}$ της f επί του E ορίζεται από τον Τύπο:

$$\int_a^\beta [f^{(q)}(x) - \overline{g}^{(q)}(x)]^2\, dx + \rho \sum_{i=1}^N [f(x_i) - \overline{g}(x_i)]^2$$

$$\leq \int_a^\beta [f^{(q)}(x) - g^{(q)}(x)]^2 dx + \rho \sum_{i=1}^N [f(x_i) - g(x_i)]^2, \quad \forall g \in E.$$

Τα θεωρητικά αποτελέσματα του χαρακτηρισμού, της μοναδικότητας και της κατασκευής μίας τέτοιας προσέγγισης, που παρουσιάστηκαν στην Παράγραφο IV.5.2, μπορούν να εφαρμοστούν και στην προκείμενη περίπτωση . Λέμε ότι $\bar{g}$ είναι μία συνάρτηση spline διευθέτησης της f. Δεν θα αναπτύξουμε όμως περαιτέρω εδώ το ζήτημα των συναρτήσεων splines.

IV.5.3.(ii). Βέλτιστη Ομοιόμορφη Προσέγγιση Έστω πάλι ο διανυσματικός χώρος $C[a,\beta]$, όλων των πραγματικών συναρτήσεων, που είναι συνεχείς επί του κλειστού διαστήματος $[a,\beta]$, τούτη την φορά, όμως, και σε αντίθεση με την περίπτωση της *Υποπαραγράφου IV.5.3.(i).(α)*, εφοδιασμένος με την νόρμα της ομοιόμορφης σύγκλισης:

$$\|f\| = \sup_{x\in[a,\beta]} |f(x)| \quad (f \in C[a,\beta]).$$

Εάν E είναι ένας οποιοσδήποτε διανυσματικός υπόχωρος του $C[a,\beta]$, τότε, σύμφωνα με τον *Ορισμό IV.5.2.1*, μία συνάρτηση $\bar{g} \in E$ είναι βέλτιστη προσέγγιση της $f \in C[a,\beta] - E$ επί του E εάν επαληθεύει την εξίσωση:

$$\|f - \bar{g}\| = \inf_{g\in E} \sup_{x\in[a,\beta]} |f(x) - g(x)|.$$

Σε αυτήν την περίπτωση, λέμε, επίσης, ότι η συνάρτηση $\bar{g}$ είναι ***βέλτιστη ομοιόμορφη προσέγγιση της f επί του E***.

Έχουμε ήδη συναντήσει ένα παράδειγμα μίας τέτοιας βέλτιστης ομοιόμορφης προσέγγισης στην *Παράγραφο IV.1.7*, όταν μελετούσαμε τα πολυώνυμα Chebyshev.

Το σύνολο $C[a,\beta]$, εφοδιασμένο με την παραπάνω νόρμα, δεν είναι χώρος προ-Hilbert ούτε, και κατ' επέκταση, χώρος Hilbert.

Θα πρέπει λοιπόν να δώσουμε έναν χαρακτηρισμό της βέλτιστης ομοιόμορφης σύγκλισης (που κάποιες φορές, ονομάζεται, επίσης, ***προσέγγιση με την έννοια του Chebyshev***) και, ακόμη, να αποδείξουμε αποτελέσματα μοναδικότητας και να μορφοποιήσουμε ένα, τουλάχιστον θεωρητικό, τρόπο κατασκευής της.

Προς τούτο, ας θεωρήσουμε ότι ο διανυσματικός υπόχωρος E είναι πεπερασμένης διάστασης, και ας υποθέσουμε, ότι παράγεται από n στοιχεία:

$$g_1, g_2, \ldots, g_n \in C[a,\beta]$$

που είναι γραμμικά ανεξάρτητα μεταξύ τους.

Ισχύει το ακόλουθο *Θεώρημα*, του οποίου δεν θα δώσουμε την απόδειξη, αλλά που θα χρησιμοποιήσουμε σαν βοηθητική πρόταση για να διατυπώσουμε, στην συνέχεια, έναν βασικό χαρακτηρισμό της βέλτιστης ομοιόμορφης προσέγγισης:

Θεώρημα IV.5.3.(ii).1. ***Μία ικανή και αναγκαία συνθήκη προκειμένου οι βέλτιστες ομοιόμορφες προσεγγίσεις όλων των συναρτήσεων*** $f \in C[a,\beta] - E$ ***να ορίζονται μονοσήμαντα επί του*** E ***είναι η ιδιότητα κάθε συνάρτησης*** $g \in E$ ***να έχει το πολύ*** $(n-1)$ ***ρίζες στο διάστημα*** $[a,\beta]$. ■

Η συνθήκη του *Θεωρήματος* αυτού ονομάζεται ***συνθήκη Haar***, και, όπως θα φανεί αμέσως, η επαλήθευσή της παρέχει διεξοδικές δυνατότητες στο πρόβλημα της αναζήτησης των βασικών ιδιοτήτων της βέλτιστης ομοιόμορφης προσέγγισης.

Θεώρημα IV.5.3.(ii).2. ***Εάν η συνθήκη Haar επαληθεύεται επί του*** E ***και εάν, για κάθε*** $g \in E$***, ισχύει:***

$$f(x_i) - g(x_i) = \theta\,(-1)^i\,\lambda_i \qquad (i = 0,1,\ldots,n),$$

όπου:

$$f \in C[a,\beta] - E,$$

$$\theta = \pm 1,$$

$$\lambda_i > 0 \qquad (\forall i = 0,1,2,\ldots,n),$$

και:

$$a \le x_0 < x_1 < \ldots < x_n \le \beta,$$

τότε η βέλτιστη ομοιόμορφη προσέγγιση $\bar{g}$ ***της*** f ***επί του*** E ***επαληθεύει την ανισοτική σχέση:***

$$\|f - \bar{g}\| \ge \min_{0 \le i \le n} \lambda_i .$$

Απόδειξη. Ας υποθέσουμε, για να οδηγηθούμε σε άτοπο, ότι είναι:

$$\|f - \bar{g}\| < \min_{0 \le i \le n} \lambda_i .$$

Τότε, θα έχουμε:

$$|\,f(x_i) - \bar{g}(x_i)\,| < \lambda_i = |\,f(x_i) - g(x_i)\,|,$$

για κάθε $i = 0,1,\ldots,n$ και κάθε $g \in E$, και, συνεπώς, η παράσταση:

$$g(x_i) - \bar{g}(x_i) = (f(x_i) - \bar{g}(x_i)) - (f(x_i) - g(x_i))$$

θα έχει το ίδιο πρόσημο με αυτό της παράστασης

$$-(f(x_i) - g(x_i)) .$$

Αυτό σημαίνει ότι η συνάρτηση $g - \bar{g}$ θα μεταβάλλει το πρόσημο της n φορές μέσα στο διάστημα $[a,\beta]$. Το συμπέρασμα αυτό αντιτίθεται στην συνθήκη Haar του E. ■

Θεώρημα IV.5.3.(ii).3. ***Εάν η συνάρτηση*** $g \in E$ ***ικανοποιεί τις επόμενες δύο συνθήκες:***

$$|f(x_i) - g(x_i)| = \|f - \bar{g}\|$$

και:

$$f(x_i) - g(x_i) = -(f(x_{i+1}) - g(x_{i+1}))$$

σε τουλάχιστον $(n+1)$ ***σημεία:***

$$x_0, x_1, \ldots, x_n$$

του $[a, \beta]$***, και για κάποια συνάρτηση*** $f \in C[a, \beta] - E$***, τότε η*** g ***είναι η βέλτιστη ομοιόμορφη προσέγγιση της*** f ***επί του*** E***.***

Απόδειξη. Έστω $\bar{g}$ η βέλτιστη ομοιόμορφη προσέγγιση της f επί του E. Επειδή:

$$f(x_i) - g(x_i) = \theta \ (-1)^i \ \|f - \bar{g}\| \quad (\text{με } \theta = \pm 1),$$

το προηγούμενο *Θεώρημα* εγγυάται ότι:

$$\|f - \bar{g}\| \geq \|f - g\|.$$

Από την άλλη πλευρά όμως, από τον ορισμό της βέλτιστης προσέγγισης, ισχύει:

$$\|f - \bar{g}\| \leq \|f - g\|.$$

Άρα τελικά:

$$g = \bar{g},$$

και η *Απόδειξη* είναι πλήρης. ■

Μπορούμε να αποδείξουμε και το αντίστροφο αποτέλεσμα του προηγούμενου *Θεωρήματος*. Η ανάπτυξη και τα επιχειρήματα που χρησιμοποιούνται, επί τούτου, είναι αρκετά τεχνικά, για αυτό και θα αποφύγουμε να δώσουμε την απόδειξη.

Θα προτιμήσουμε να διατυπώσουμε την συγχώνευση της ευθείας και της αντίστροφης Πρότασης, που είναι γνωστή σαν το ***Θεώρημα Εναλλαγής Chebyshev***:

Θεώρημα IV.5.3.(ii).4. ***Μία ικανή και αναγκαία συνθήκη προκειμένου η συνάρτηση*** $g \in E$ ***να ταυτίζεται με την βέλτιστη ομοιόμορφη προσέγγιση*** $\bar{g}$ ***μίας*** $f \in C[a,\beta] - E$ ***επί του*** E ***είναι η ικανοποίηση των επόμενων δυο συνθηκών:***

$$|f(x_i) - g(x_i)| = \|f - \bar{g}\|$$

και:

$$f(x_i) - g(x_i) = -(f(x_{i+1}) - g(x_{i+1}))$$

σε τουλάχιστον $(n+1)$ ***σημεία*** $x_0, x_1, \ldots, x_n$ ***του διαστήματος*** $[a,\beta]$**.** ■

Ας πάρουμε τώρα:

$$g_1(x) = 1,\ g_2(x) = x, \ldots, g_N(x) = x^{N-1},$$

δηλαδή ας θεωρήσουμε την περίπτωση της βέλτιστης πολυωνυμικής ομοιόμορφης προσέγγισης.

Μπορούμε να διαπιστώσουμε αμέσως ότι η συνθήκη Haar ικανοποιείται, και άρα, σύμφωνα με το *Θεώρημα IV.5.3.(ii).1*, η βέλτιστη πολυωνυμική ομοιόμορφη προσέγγιση είναι μοναδικά ορισμένη. Η κατασκευή της βέλτιστης αυτής προσέγγισης πραγματοποιείται με επαναληπτική μέθοδο, και με την βοήθεια του ***αλγορίθμου Rémes*** ([41]).

Ας σημειωθεί πως όταν η διάσταση του διανυσματικού χώρου E αυξάνεται, δηλαδή, όταν αυξάνεται ο βαθμός του πολυωνύμου, που αναπαριστά την βέλτιστη πολυωνυμική ομοιόμορφη προσέγγιση, τότε η ομοιόμορφη σύγκλιση της ακολουθίας αυτών των πολυωνύμων εξασφαλίζεται από το *Θεώρημα Weierstrass*. Όταν, ειδικότερα,

$$n = N + 1,$$

ξαναβρίσκουμε την περίπτωση της πολυωνυμικής παρεμβολής, που εξετάσαμε ήδη στην *Παράγραφο IV.1*.

Βιβλιογραφία

[1] AHO, A.V., HOPCROFT, J.E. and ULLMAN, J.D.: *The Design and Analysis of Computer Algorithms,* Addison-Wesley, Reading, *Mass.,* 1974.

[2] ΑΠΟΣΤΟΛΑΤΟΣ, Ν.Θ.: *Αριθμητική Ανάλυση,* Τεύχος 1 και 2, 1983.

[3] ATKINSDON, K.E.: *An Introduction to Numerical Analysis (second edition),* John Wiley & Sons, New York, 1989.

[4] BASS, J.: *Cours de Mathématiques,* 1964.

[5] BEREZIN, I.S. and ZHIDKOV, N.P.: *Computing Methods,* Pergamon Press 1965.

[6] BREZINSKI, C.: *Accélération de la Convergence en Analyse Numérique,* Publication 41, Laboratoire de Calcul, Université de Lille I, 1973.

[7] BURDEN, R.L. and FAIRES, J.D.: *Numerical Analysis,* PWS-KENT Publ., Boston, 1993.

[8] CESHINO, F. and KUNTZMANN, J.: *Problémes Différentiels de Conditions Initiales,* Dunod, 1963.

[9] CHATELIN-LABORDE, F.: *Méthodes Numériques de Calcul des Valeurs Propres et Vecteurs Propres d' un Opérateur Linéaire,* Thése, Grenoble, 1971.

[10] COLEMAN, T.F. and C.VAN LOAN: *Handbook for Matrix Computations,* SIAM Publications, Philadelphia, 1988.

[11] CONTE, S.D. and C. de BOOR: *Elementary Numerical Analysis,* 3rd ed., McGraw-Hill, 1980.

[12] DAHLQUIST, G. and BJÖRCK, A.: *Numerical Methods*, Prentice-Hall, Inc., 1974.

[13] DAVIS, P.J.: *Interpolation and Approximation,* Blaisdell Publ. Comp., 1961.

[14] DENNERY, P and KRZYWICKI, A.: *Mathematics for Physicists,* Harper and Row, 1967.

[15] ΔΙΟΝΥΣΙΟΥ, ΔΗΜ. Δ.: *Μαθηματική Ανάλυση Ι-ΙΙ και Εφαρμοσμένα Μαθηματικά,* Εκδόσεις Συμεών, Αθήνα 1997.

[16] FADDEEVA, N.V.: *Computational Methods of Linear Algebra,* Dover, 1959.

[17] FORSYTHE, G.E., MALCOLM, M.A. and MOLER, C.A.: *Computer Methods for Mathematics Computations,* Prentice Hall, Englewood Cliffs, N.J., 1977.

[18] FORSYTHE, G.E. and MOLER, C.A.: *Computer Solution of Linear Algebraic Systems,* Prentice Hall, Englewood Cliffs, N.J., 1967.

[19] FRÖBERG, G.E.: *Introduction to Numerical Analysis,* 2nd ed., Addison-Wersley, 1969.

[20] GANTMACHER, FR.: *Theorie des Matrices.* Tome 1: Théorie Générale, Dunod.

[21] GASTINEL, N.: *Analyse Numérique Linérique, Hermann,* 1966.

[22] GEORGE, J.A. and LIU, J.W.H.: *Computer Solution of Large Sparse Positive Definite Systems,* Prentice Hall, Englewood Cliffs, N.J., 1981.

[23] GERALD, C.F.: *Applied Numerical Analysis,* Addison – Wesley, 1970.

[24] GOLUB, G.H. and C.F.VAN LOAN: *Matrix Computations,* 2nd ed., Johns Hopkins University Press, Baltimore, 1989.

[25] GUELFOND, A.O.: *Calcul des Différences Finies,* Dunod, 1963.

[26] HACQUES, G.: *Mathématiques pour l' Informatique. 3-Algorithmique Numérique,* Armand Colin, 1971.

[27] HAGEMAN, L.A. and YOUNG, D.M.: *Applied Iterative Methods,* Academic Press, New York, 1981.

[28] HENRICI, P.: *Elements of Numerical Analysis,* Wiley, 1964.

[29] HILDEBRAND, F.B.: *Introduction to Numerical Analysis,* 2nd ed., McGraw-Hill, 1974.

[30] HOUSEHOLDER, A.S.: *The Theory of Matrices in Numerical Analysis,* Blaisdell Publ. Comp., 1964.

[31] HOUSEHOLDER, A.S.: *The Numerical Treatment of a Single Nonlinear Equation*

[32] ISSACSON, E. and KELLER, H.B.: *Analysis of Numerical Methods,* John Wiley & Sons, New York, 1966.

[33] JOHNSON, L.W. and RIESS, R.D.: *Numerical Analysis,* Addison-Wesley, 1977.

[34] KAHANER, D., MOLER, C. and NASH, S.: *Numerical Methods and Software,* Prentice Hall, Englewood Cliffs, N.J., 1989.

[35] KINCAID, D. and CHENEY, W.: *Numerical Analysis: Mathematics of Scientific Computing,* Brooks/Cole Publishing Company, Pacific Grove, California, 1991.

[36] KJELLBERG, G.: *On the convergence of successive overrelaxation applied to a class of linear systems of equations with complex eigen-values,* Ericson Technics 2 (1958), σελ. 245-258.

[37] KREDELL, B.: *On complex successive overrelaxation*, BIT 2 (1962), σελ. 143-152.

[38] KUNTZMANN, J.: *Méthodes Numériques, Interpolation, Dérivées,* Dunod, 1959.

[39] LASCOUX, A.: *Interpolation de Lagrange,* Αδημοσίευτη εργασία, 1986.

[40] LAURENT, P.J.: *Etude des Procédés d' Interpolation en Analyse Numérique,* Thése, Grenoble, 1964.

[41] LAURENT, P.J.: *Approximation et Optimisation,* Hermann, 1972.

[42] LEBAUD, C.: *Contribution à l' Etude de l' Algorithme QR,* Thése, Rennes, 1971.

[43] LEGRAND, G.: *Algèbre Linéaire et Multilinéaire et Géométrie Différentielle,* Masson, 1971.

[44] LIONS, J.L.: *Cours d' Analyse Numérique, Ecole Polytechnique,* 1965-1966.

[45] LORENTZ, G.G., M.VON GOLITSCHEK and MAKOVOZ, Y.: *Constructive Approximation: Advanced Problems,* Springer-Verlag, Berlin and New York, 1996.

[46] ΜΙΣΥΡΛΗΣ, Ν.: *Εισαγωγή στην Αριθμητική Ανάλυση,* Αθήνα, 1995.

[47] MITCHELL, A.R.: *Computational Methods for Partial Differential Equations,* John Wiley & Sons, London, 1969.

[48] ORTEGA, J.M.: *Numerical Analysis: A Second Course,* Academic Press, 1972.

[49] ORTEGA, J.M. and PHEINBOLDT, W.C.: *Iterative Solution of Nonlinear Equations in Several Variables,* Academic Press, 1970.

[50] PARLETT, B.: *The Symmetric Eigenvalue Problem,* Prentice Hall, Englewood Cliffs, N.J., 1980.

[51] PISOT, C. et ZAMANSKY, M.: *Mathématiques Générales,* Dunod, 1966.

[52] POUZET, P.: *Analyse Numérique,* Publication du Laboratoire de Calcul, Université de Lille I.

[53] PALSTON, A. and RABINOWITZ, P.: *A First Course in Numerical Analysis,* 2nd ed., McGraw-Hill, New York, 1978.

[54] RAVIART, P.A.: *Cours d' Analyse Numérique,* Université de Paris VI.

[55] SCHWARTZ, L.: *Analyse, Topologie Générale et Analyse Fonctionnelle,* Hermann, 1970.

[56] STEWART, G.W.: *Introduction to Matrix Computations,* Academic Press, 1973.

[57] STOER, J. and BULIRSCH, R.: *Introduction to Numerical Analysis,* Springer-Verlag, 1980.

[58] STRANG, G.: *Linear Algebra and its Applications,* 2nd ed., Academic Press, 1980.

[59] SZABADOS, J. and VERTESI, P.: *Interpolation of Functions,* World Sci. Publ., London and Singapore, 1990.

[60] SZABADOS, J. and VERTESI, P.: *A survey of mean convergence of interpolatory processses,* J. Comput. Appl. Math. 43 (1992), σελ. 3-18.

[61] THEODOR, R.: *Analyse Numérique,* Editions scientifiques Riber.

[62] VARGA, R.S.: *Matrix Iterative Analysis,* Prentice Hall, Englewood Cliffs, N.J., 1962.

[63] WILKINSON, J.H.: *Rounding Errors in Algebraic Processes,* H.M. Stationery Office, London, 1963.

[64] WILKINSON, J.H.: *The Algebraic Eigevalue Problem,* Clarendon Press, Oxford, 1965.

[65] WOLF, J.: *Méthodes de Calcul des Values Propres d' une Matrice Quelconque par Utilisation de Transformation de Transformations Unitaires,* Thèse, Grenoble, 1971.

[66] ΧΑΤΖΗΔΗΜΟΣ, Α.: *Εισαγωγή στην Αριθμητική Ανάλυση,* Ιωάννινα, 1977.

[67] ΧΑΤΖΗΔΗΜΟΣ, Α.: *Αριθμητική Ανάλυση Ι και ΙΙ,* Ιωάννινα 1978 και 1979.

[68] YOUNG, D.M.: *Iterative Solution of Large Linear Systems,* Academic Press, 1971.

[69] YOUNG, D.M. and GRECORY, R.T.: *A Survey of Numerical Mathematics,* Volumes I and II, Addison Welsey, 1972 and 1973.

[70] ZAMANSKY, M.: *Introduction à l' Algèbre et à l' Analyse Moderne,* Dunod, 1967.