

Πολυτεχνείο Κρήτης Γενικό Τμήμα



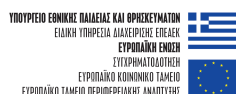
ΔΙΑΔΟΧΙΚΕΣ ΠΡΟΣΕΓΓΙΣΕΙΣ ΚΑΙ ΠΑΡΑΛΛΗΛΟΙ ΥΠΟΛΟΓΙΣΜΟΙ ΔΙΚΤΥΟΥ/ΠΛΕΓΜΑΤΟΣ ΓΙΑ ΤΗΝ ΚΑΙΝΟΤΟΜΟ ΕΠΙΛΥΣΗ ΕΛΛΕΙΠΤΙΚΩΝ ΜΕΡΙΚΩΝ ΔΙΑΦΟΡΙΚΩΝ ΕΞΙΣΩΣΕΩΝ.

Αναστάσης Γ. Σηφαλάκης

Διδακτορική Διατριβή

Επιβλέπων: Καθηγητής Ιωάννης Σαριδάκης

*Χανιά
Δεκέμβριος 2007*



*ΕΠΕΑΕΚ/ΗΡΑΚΛΕΙΤΟΣ: ΥΠΟΤΡΟΦΙΕΣ ΕΡΕΥΝΑΣ ΜΕ
ΠΡΟΤΕΡΑΙΟΤΗΤΑ ΣΤΗ ΒΑΣΙΚΗ ΕΡΕΥΝΑ
Με συγχρηματοδότηση της Ε.Ε.*

Ευχαριστίες

Κατά την διάρκεια της εκπόνησης αυτής της διδακτορικής διατριβής, χρειάστηκα συχνά την βοήθεια και υποστήριξη αρκετών ανθρώπων, τους οποίους θα ήθελα να ευχαριστήσω ιδιαίτερα.

Δικαιωματικά και ουσιαστικά θα ευχαριστήσω πρώτο τον σύμβουλο καθηγητή και δάσκαλο μου, Ιωάννη Σαριδάκη ο οποίος μου παρείχε την άρτια επιστημονική καθοδήγηση για την ολοκλήρωση της παρούσας διατριβής. Για τις αμέτρητες ώρες που μου αφιέρωσε.

Την καθηγήτρια μου κυρία Έλενα Παπαδοπούλου για την επιστημονική και ηθική υποστήριξη της. Για το ότι σε κρίσιμες στιγμές μου έδωσε το κουράγιο να συνεχίσω την παρούσα διατριβή.

Τον καθηγητή Αθανάσιο Φωκά (και εμπνευστή της μεθόδου πάνω στην οποία είναι βασισμένη η εργασία) για τις πολύτιμες συμβουλές και κατευθύνσεις που μου έδωσε.

Τον καθηγητή Θεόδωρο Παπαθεοδώρου για την υποστήριξη και εμπύχωση του κατά την διάρκεια της πρώτης μου ομιλίας σε διεθνές συνέδριο.

Τον καθηγητή μου στο Πανεπιστήμιο Κρήτης καθηγητή Απόστολο Χατζηδήμο ο οποίος με μια και μόνο φράση του με οδήγησε στο να αποφασίσω να ακολουθήσω τον δρόμο των μεταπτυχιακών σπουδών.

Τον Πρύτανη του Πολυτεχνείου Κρήτης καθηγητή Ιωακείμ Γρυσπολάκη που παρόλο το φόρτο εργασίας του μπόρεσε να συμμετάσχει στην επταμελή επιτροπή.

Τον λέκτορα Εμμανουήλ Μαθιουδάκη, για την ειλικρινή βοήθειά του πάνω στα υπολογιστι-

κά συστήματα του Εργαστηρίου Εφαρμοσμένων Μαθηματικών και Υπολογιστών. Χωρίς την βοήθεια του δεν θα μπορούσα να ασχοληθώ με την παράλληλη επεξεργασία.

Τον λέκτορα Ανάργυρο Δελή, για τις επιστημονικές συζητήσεις που είχαμε και τις πάντα εύστοχες παρατηρήσεις του.

Το προσωπικό και τον διευθυντή (καθηγητή Ιωάννη Σαριδάκη) του εργαστηρίου Εφαρμοσμένων Μαθηματικών και Υπολογιστών, για την δυνατότητα χρησιμοποίησης όλων των πόρων του εργαστηρίου και την φιλοξενία που μου παρείχαν κατά την διάρκεια των μεταπτυχιακών σπουδών μου.

Δεν μπορώ βεβαίως σε καμία περίπτωση να ξεχάσω να ευχαριστήσω τους συναδέλφους μου μεταπτυχιακούς φοιτητές του τομέα Μαθηματικών και ιδιαίτερα την κυρία Μαριάννα Παπαδομανωλάκη, για την ηθική συμπαράσταση τους και το κουράγιο που μου έδιναν όλα αυτά τα χρόνια.

Θα ήθελα επίσης να ευχαριστήσω την κυρία Τόνια Σγουρομάλλη και την κυρία Δωροθέα Φραγκομιχελάκη από την γραμματεία του Γενικού Τμήματος του Πολυτεχνείου Κρήτης για την όλη υποστήριξη, διοικητική και ανθρώπινη, κατά την διάρκεια των μεταπτυχιακών σπουδών μου και την κυρία Δέσποινα Παντελάκη από τον Ειδικό Λογαριασμό Κονδυλίων Έρευνας για την βοήθεια της με το πρόγραμμα Ηράκλειτος.

Τέλος, αφιερώνοντας την διατριβή αυτή στην οικογένειά μου, θέλω να ευχαριστήσω και να εκφράσω την ευγνωμοσύνη μου πρώτον στους γονείς μου Γεώργιο και Ιωάννα για την αμέριστη, κάθε είδους και άνευ όρων, υποστήριξη και συμπαράσταση που μου παρείχαν πάντα, χωρίς τα οποία δεν θα μπορούσα έχω ολοκληρώσει τις σπουδές μου, και δεύτερον τις αδελφές μου Ρωζάνη και Βασιλική (και την οικογένεια της) για την συμπαράστασή τους και την υπομονή που επέδειξαν στην αντιμετώπιση των δυσκολιών που συνάντησα.

Περιεχόμενα

Ευχαριστίες	i
Περίληψη	vii
1 Βασικές Έννοιες	1
1.1 Εισαγωγή	1
1.2 Κλασσικές Επαναληπτικές Μέθοδοι	1
1.2.1 SOR	3
1.3 Μέθοδοι Προβολής (Projection methods).	6
1.3.1 Krylov Subspace Μέθοδοι.	8
1.3.2 Κατασκευή ορθοκανονικής βάσης του χώρου $\mathbf{K}_m(\mathbf{A}, \mathbf{r}_0)$.	9
1.3.3 Generalized Minimum Residual Method (GMRES)	10
1.3.4 Biorthogonalization methods	13
1.3.5 Biconjugate Gradient Algorithm (BCG)	14
1.4 Προρύθμιση (Preconditioning).	17
1.5 Ο Ταχύς Μετασχηματισμός Fourier (FFT)	19
1.5.1 Ο Διακριτός Μετασχηματισμός Fourier (DFT)	19
1.5.2 FFT	20
1.6 Circulant Πίνακες	21
2 Η Μέθοδος Φωκά	25
2.1 Η Μέθοδος Φωκά	25
2.2 Η Αριθμητική Μέθοδος	29
2.2.1 Παραμετροποίηση	30
2.2.2 Η Επιλογή του k	31
2.2.3 Διακριτοποίηση	36
2.3 Σχόλια	40
3 Η μέθοδος SFFS	41
3.1 Αρίθμηση των Εξισώσεων/Αγνώστων	41
3.2 Ημιτονοειδής Βάση	42

3.3	Η μέθοδος FFX	43
3.4	Η μέθοδος SFFS	45
3.5	Διαμέριση του Collocation Πίνακα	51
3.6	Ασυμπτωτική Συμπεριφορά του Collocation Πίνακα	52
4	Σύγκριση Μεθόδων SFFS και FFX	55
4.1	Το Αριθμητικό Πρόβλημα - Μοντέλο	55
4.2	Αριθμητική Υλοποίηση	56
4.3	Αριθμητικά Αποτελέσματα	58
4.3.1	Σφάλμα - Κανονικά Πολύγωνα	59
4.3.2	Σφάλμα - Μη Κανονικά Πολύγωνα	64
4.3.3	Δείκτης Κατάστασης	69
4.4	Συμπεράσματα	71
5	Αριθμητική Μελέτη και Συμπεριφορά Επαναληπτικών Μεθόδων	73
5.1	Κανονικά Πολύγωνα	74
5.2	Μη Κανονικά Πολύγωνα	82
5.3	Συμπεράσματα	85
6	Δομή και Ιδιότητες του Collocation Πίνακα	87
6.1	Κανονικά Πολύγωνα με Ίδιες Συνοριακές Συνθήκες σε Όλες τις πλευρές	87
6.2	Γενικότερες Περιπτώσεις	98
6.2.1	Μεταφορά, Στροφή και Μεγέθυνση του Πολυγώνου	98
6.2.2	Ομοιόμορφη Μεταβολή των Συνοριακών Συνθηκών	100
6.2.3	Διαφορετικές Συνοριακές Συνθήκες σε Κάθε Πλευρά	101
6.3	Τετράγωνο	104
6.4	Συμπεράσματα	111
7	Επίλυση της Dirichhlet to Neumann Απεικόνισης με Χρήση FFT	113
7.1	Block – Circulants και FFT	113
7.1.1	FFTmul	115
7.1.2	FFTinv	116
7.1.3	FFTinv(Iterative)	119
7.1.4	FFT για Διαφορετικές Συνοριακές Συνθήκες σε Κάθε Πλευρά	120
7.2	Συμπεράσματα	124
8	Παράλληλοι Υπολογισμοί / Υπολογισμοί Πλέγματος	125
8.1	Ο Παράλληλος Αλγόριθμος	125
8.2	Αριθμητικά Αποτελέσματα	129
8.3	Συμπεράσματα	151

9 Υψηλής Τάξης Σύγκλιση	153
9.1 Chebyshev Βάση	153
9.2 Αριθμητικά Αποτελέσματα	154
9.3 Συμπεράσματα	159
10 Συμπεράσματα	161
Παράρτημα	169

Περίληψη

Η επίλυση προβλημάτων συνοριακών τιμών για μερικές διαφορικές εξισώσεις (PDEs) κατέχει κεντρική θέση στα εφαρμοσμένα μαθηματικά. Αναλυτικές μέθοδοι (π.χ. χωριζόμενων μεταβλητών και τεχνικές μετασχηματισμών) παράγουν ακριβείς λύσεις και μας δίνουν και μια ισχυρή αίσθηση για την διαφορική εξίσωση. Όμως το φάσμα των προβλημάτων που λύνουν είναι περιορισμένο. Αριθμητικές μέθοδοι (π.χ. πεπερασμένων στοιχείων, πεπερασμένων διαφορών και φασματικές) είναι ικανές να επιλύσουν ένα ευρύτερο φάσμα προβλημάτων αλλά μόνο προσεγγιστικά. Μερικές μέθοδοι (π.χ. boundary integral methods) συνδυάζουν συγκεκριμένες αναλυτικές πληροφορίες για την λύση μαζί με αριθμητικές προσεγγίσεις. Η παρούσα εργασία ασχολείται με μια μέθοδο που ανήκει στην τελευταία αυτή κατηγορία.

Πρόσφατα μια νέα αναλυτική μέθοδος για την μελέτη προβλημάτων συνοριακών τιμών για γραμμικές και για ολοκληρώσιμες μη γραμμικές PDEs σε δυο διαστάσεις εισήχθη από τον καθηγητή Φωκά (βλ. [FOK97, FOK01]). Αυτή η μέθοδος έχει χρησιμοποιηθεί για την επίλυση γραμμικών Ελλειπτικών Μερικών Διαφορικών Εξισώσεων σε κυρτά πολυγωνικά χωρία [FOK01] παράγοντας αναλυτικές λύσεις σε περιπτώσεις μη αντιμετωπίσιμες από τις κλασικές τεχνικές μετασχηματισμών. Κλειδί σε αυτή την μέθοδο είναι η **ολική συνθήκη** [PGA99] η οποία συνδέει γνωστές και άγνωστες τιμές της λύσης ή των παραγώγων της λύσης στο σύνορο του χωρίου. Σε περιπτώσεις όπου η ολική συνθήκη μπορεί να λυθεί αναλυτικά η μέθοδος παράγει την λύση σε κλειστή μορφή γενικεύοντας έτσι τις κλασικές μεθόδους μετασχηματισμών (π.χ. Fourier και Bessel). Στην γενική όμως περίπτωση συνοριακών προβλημάτων η ολική συνθήκη πρέπει να λυθεί αριθμητικά.

Η παρούσα διατριβή ασχολείται με την ανάπτυξη και την μελέτη αποδοτικών αριθμητικών μεθόδων τόσο για την αριθμητική προσέγγιση της ολικής συνθήκης όσο και για την επίλυση του διακριτοποιημένου αναλόγου της. Πιο συγκεκριμένα:

- Στο 1ο κεφάλαιο παραθέτουμε το βασικό μαθηματικό υπόβαθρο που χρησιμοποιούμε. Κάνουμε μια μικρή ανασκόπηση στις επαναληπτικές μεθόδους επίλυσης γραμμικών συστημάτων, αναφέρουμε την έννοια της προρύθμισης και δίνουμε τους ορισμούς και τις βασικές ιδιότητες των circulant πινάκων και των Διακριτών Μετασχηματισμών Fourier.
- Στο 2ο κεφάλαιο περιγράφουμε την μέθοδο του καθηγητή Φωκά για την επίλυση της εξίσωσης Laplace σε κυρτά πολυγωνικά χωρία δίνοντας έμφαση στην αριθμητική υλοποίηση της. Εξετάζουμε τους βαθμούς ελευθερίας της μεθόδου, δηλαδή ποια σημεία μπορούμε να

αλλάζουμε στην αριθμητική υλοποίηση έτσι ώστε να επιτύχουμε καλύτερη αριθμητική συμπεριφορά. Ακόμα αιτιολογούμε την επιλογή της γενικής μορφής των σημείων ταξινόμησης (collocation points), δηλαδή των σημείων του μιγαδικού επιπέδου πάνω στα οποία αποτιμούμε μια γενικευμένη συνθήκη για να παράξουμε ένα σύστημα γραμμικών εξισώσεων το οποίο θα μας οδηγήσει στην επίλυση της εξίσωσης του Laplace, έχοντας ως γνώμονα τους περιορισμούς που παρουσιάζονται στα σύγχρονα υπολογιστικά περιβάλλοντα.

- Στο 3ο κεφάλαιο εξετάζουμε την υπάρχουσα επιλογή των collocation σημείων και παράλληλα προτείνουμε μια διαφορετική επιλογή σημείων για την επίλυση της ολικής συνθήκης σε κυρτά πολυγωνικά χωρία, η οποία οδηγεί σε πίνακα συντελεστών των αγνώστων του οποίου τα διαγώνια μπλοκ είναι σημειακοί διαγώνιοι πίνακες.
- Στο 4ο κεφάλαιο συγκρίνουμε, αριθμητικά, τις δυο επιλογές των collocation σημείων και συμπεραίνουμε ότι η νέα επιλογή οδηγεί σε αύξηση του ρυθμού σύγκλισης της μεθόδου (τετραγωνική σύγκλιση) και σε πίνακα συντελεστών του γραμμικού συστήματος με πολύ μικρότερο δείκτη κατάστασης.
- Στο 5ο κεφάλαιο εξετάζουμε την συμπεριφορά των επαναληπτικών μεθόδων για την επίλυση του γραμμικού συστήματος με την νέα επιλογή των collocation σημείων, για κανονικά και μη κανονικά κυρτά πολύγωνα.
- Στο 6ο κεφάλαιο εξετάζουμε, θεωρητικά, την δομή και τις ιδιότητες του collocation πίνακα.
- Στο 7ο κεφάλαιο χρησιμοποιούμε τις ιδιότητες του collocation πίνακα, όπως παρουσιάστηκαν στο 6ο κεφάλαιο (π.χ. block-circulant), για την αποδοτική (μέσω Ταχέων Μετασχηματισμών Fourier) επίλυση του collocation συστήματος.
- Στο 8ο κεφάλαιο επιλύουμε το collocation σύστημα σε παράλληλα υπολογιστικά συστήματα. Παρουσιάζουμε και αναλύουμε ένα αλγόριθμο για την παράλληλη επίλυση γενικών block-circulant γραμμικών συστημάτων.
- Στο 9ο κεφάλαιο προτείνουμε μια νέα επιλογή συναρτήσεων βάσης (Chebyshev) η οποία οδηγεί σε, σχεδόν, εκθετικούς ρυθμούς σύγκλισης της αριθμητικής μεθόδου.
- Στο 10ο κεφάλαιο ανακεφαλαιώνουμε και αναφέρουμε τα ανοιχτά ζητήματα που προέκυψαν και θα μας απασχολήσουν στο μέλλον.

Τέλος στο παράρτημα παραθέτουμε τις εργασίες μας:

1. A. G. Sifalakis, A. S. Fokas, S. Fulton, and Y. G. Saridakis, **The Generalized Dirichlet-Neumann Map for Linear Elliptic PDEs and its Numerical Implementation**, *Journal of Computational and Applied Mathematics*, (in press).

2. A. G. Sifalakis, E. P. Papadopoulou and Y. G. Saridakis, **Numerical Study of Iterative Methods for the Solution of the Dirichlet-Neumann Map for Linear Elliptic PDEs on Regular Polygon Domains**, *International Journal of Applied Mathematics and Computer Sciences*, 4(3), 2007.
3. A. G. Sifalakis, S. Fulton, E. P. Papadopoulou and Y. G. Saridakis, **Direct and Iterative Solution of the Generalized Dirichlet-Neumann Map for Elliptic PDEs on Square Domains**, *J. Comp. and Appl. Math.*, (submitted).
4. Y. G. Saridakis, A. G. Sifalakis and E. P. Papadopoulou, **Efficient Numerical Solution of the Generalized Dirichlet-Neumann Map for Linear Elliptic PDEs in Regular Polygon Domains**, *IMA Journal of Numerical Analysis*, (submitted).

που έχουν δημοσιευθεί ή έχουν υποβληθεί προς δημοσίευση σε διεθνή επιστημονικά περιοδικά.

Κεφάλαιο 1

Βασικές Έννοιες

1.1 Εισαγωγή

Σε αυτό το κεφάλαιο θα εκθέσουμε το βασικό μαθηματικό υπόβαθρο της παρούσης εργασίας. Πιο συγκεκριμένα, θα εξετάσουμε τις βασικές (stationary) επαναληπτικές μεθόδους, δίνοντας έμφαση στους p -κυκλικούς (p -cyclic) πίνακες [VAR00, HYO81, YOU81, HAD00] καθώς και τις Krylov επαναληπτικές μεθόδους [SAA00, VOR03, KEL95, GRE97, DEM96, DDS98, TEM, TBA97]. Κατόπιν θα μιλήσουμε για την τεχνική της προρύθμισης και θα αναφέρουμε τους προρρυθμιστές που προκύπτουν από τις stationary επαναληπτικές μεθόδους. Στην συνέχεια θα αναφερθούμε στους Ταχείς Μετασχηματισμούς Fourier [LOA92, HIG96, GCL96] και στην ειδική κατηγορία των circulant πινάκων [DAV79, TIA00].

1.2 Κλασσικές Επαναληπτικές Μέθοδοι

Για την αριθμητική επίλυση μη ιδιόμορφων γραμμικών συστημάτων της μορφής

$$Ax = b \quad , \quad A \in \mathbb{R}^{n,n} \quad , \quad b \in \mathbb{R}^n \setminus \{0\} \quad (1.1)$$

θεωρούμε μια διάσπαση του πίνακα A

$$A = M - N \quad (1.2)$$

με περιορισμούς

- (α) Ο πίνακας M να είναι αντιστρέψιμος, και
- (β) Ένα γραμμικό σύστημα με πίνακα συντελεστών αγνώστων M να λύνεται με πολύ λιγότερες πράξεις από ένα άλλο με πίνακα συντελεστών αγνώστων A .

Ο πίνακας M είναι γνωστός ως πίνακας διάσπασης (splitting matrix). Χρησιμοποιώντας την (1.2) στην (1.1) και αναδιατάσσοντας έχουμε

$$x = Tx + c \quad , \quad T := M^{-1}N = I - M^{-1}A \quad , \quad c = M^{-1}b \quad (1.3)$$

Η τελευταία εξίσωση είναι ισοδύναμη με την (1.1) και υποδεικνύει την κατασκευή του αλγορίθμου

$$x^{m+1} = Tx^m + c, \quad k = 0, 1, 2, \dots, \quad (1.4)$$

με $x^{(0)} \in \mathbb{R}^n$ τυχαίο διάνυσμα ή, οποτεδήποτε δυνατόν, αρχική προσέγγιση της μοναδικής λύσης $x = A^{-1}b$ του συστήματος (1.1).

Ο πίνακας T ονομάζεται επαναληπτικός πίνακας του αλγορίθμου (ή της μεθόδου). Ο αλγόριθμος (1.4) παράγει μια ακολουθία διανυσμάτων

$$\{x^{(m)}\}_{m=0}^{\infty}.$$

η οποία κάτω από συγκεκριμένες προϋποθέσεις συγκλίνει στην μοναδική λύση $x = A^{-1}b$ του γραμμικού συστήματος (1.3).

Θεώρημα 1.1 *Αναγκαία και ικανή συνθήκη για την σύγκλιση της ακολουθίας των παραγόμενων από τον αλγόριθμο (1.4) διανυσμάτων στην λύση $x = A^{-1}b$ του συστήματος (1.3) είναι η*

$$\rho(T) < 1, \quad (1.5)$$

όπου $\rho(\cdot)$ η φασματική ακτίνα (το μέτρο της μέγιστης κατά μέτρο ιδιοτιμής).

Οι κλασικές επαναληπτικές μέθοδοι βασίζονται στην ακόλουθη, μονοσήμαντη, διάσπαση του πίνακα συντελεστών αγνώστων A

$$A = D_A - L_A - U_A, \quad (1.6)$$

όπου $D_A = \text{diag}(A)$, δηλαδή, διαγώνιος πίνακας με διαγώνια στοιχεία τα αντίστοιχα του A , L_A αυστηρά κάτω τριγωνικός και U_A αυστηρά άνω τριγωνικός. Ανακαλώντας τώρα την διάσπαση του πίνακα A από την (1.2) έχουμε

Μέθοδος	M	N	T
Jacobi	D_A	$L_A + U_A$	$D_A^{-1}(L_A + U_A)$
Gauss-Seidel	$D_A - L_A$	U_A	$(D_A - L_A)^{-1}U_A$
SOR	$\frac{1}{\omega}(D_A - \omega L_A)$	$(\frac{1}{\omega} - 1)D_A + U_A$	$(D_A - \omega L_A)^{-1}[(1 - \omega)D_A + \omega U_A]$

Ο επαναληπτικός πίνακας του Jacobi συμβολίζεται, συνήθως, με J και ο επαναληπτικός πίνακας της SOR με $\mathcal{L}_\omega$. Στην SOR το $\omega \in \mathbb{C} \setminus \{0\}$ καλείται παράμετρος της υπερχαλάρωσης (overrelaxation). Παρατηρήστε ότι για $\omega = 1$ η SOR ταυτίζεται με την μέθοδο GS.

Οι παραπάνω τρεις μέθοδοι καλούνται point (σημειακές) σε αντιδιαστολή με τις block. Για τις τελευταίες θεωρούμε μια διαμέριση του A στην ακόλουθη block μορφή :

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1p} \\ A_{21} & A_{22} & \cdots & A_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ A_{p1} & A_{p2} & \cdots & A_{pp} \end{bmatrix}, \quad (1.7)$$

όπου $A_{ii} \in \mathbb{R}^{n_i, n_i}$, $i = 1(1)p$ και $\sum_{i=1}^p n_i = n$. Αν ορίσουμε $D_A = \text{diag}(A_{11}, A_{22}, \dots, A_{pp})$ θεωρώντας ότι $\det(A_{ii}) \neq 0$, $i = 1, 2, \dots, p$ και

$$A = D_A - L_A - U_A, \quad (1.8)$$

με τους L_A και U_A να είναι αυστηρά κάτω και άνω τριγωνικοί πίνακες τότε οι block Jacobi, block Gauss-Seidel και block SOR μέθοδοι ορίζονται από τα ίδια επαναληπτικά σχήματα όπως και τα point ζευγάρια τους.

Παρατήρηση 1.1 *Ο επαναληπτικοί πίνακες των κλασσικών επαναληπτικών μεθόδων γράφονται και ως*

$$T = I - M^{-1}A. \quad (1.9)$$

Δηλαδή

$$\begin{aligned} \text{Jacobi: } T_J &= I - D_A^{-1}A \\ \text{Gauss-Seidel: } T_{GS} &= I - (D_A - L_A)^{-1}A \\ \text{SOR: } T_{SOR} &= I - \omega (D_A - \omega L_A)^{-1}A \end{aligned} \quad (1.10)$$

1.2.1 SOR

Η SOR αποτελεί μια μονοπαραμετρική γενίκευση της Gauss-Seidel, αφού για $\omega = 1$ η SOR είναι η Gauss-Seidel. Η SOR δεν συγκλίνει για κάθε $\omega \in \mathbb{C} \setminus \{0\}$. Το παρακάτω θεώρημα μας βοηθάει να διατυπώσουμε μια αναγκαία συνθήκη για την σύγκλιση της SOR που είναι ανεξάρτητη του πίνακα συντελεστών των αγνώστων.

Θεώρημα 1.2 *Έστω $\mathcal{L}_\omega$ ο επαναληπτικός πίνακας της SOR μεθόδου. Για οποιοδήποτε ω (πραγματικό ή μιγαδικό) ισχύει η σχέση :*

$$\rho(\mathcal{L}_\omega) \geq |\omega - 1|. \quad (1.11)$$

Άμεση συνέπεια του Θεωρήματος 1.2 είναι το παρακάτω πόρισμα.

Πόρισμα 1.1 (Kahan) Αναγκαία συνθήκη για την σύγκλιση της *SOR* είναι η

$$|\omega - 1| < 1, \omega \in \mathbb{C} \implies \omega \in (0, 2), \omega \in \mathbb{R}. \quad (1.12)$$

Ορισμός 1.3 (Ταχύτητα σύγκλισης) Έστω $A, B \in \mathbb{C}^{n,n}$. Αν για κάποιο θετικό ακέραιο k και για μια φυσική νόρμα $\|\cdot\|_\beta$ είναι $\|A^k\|_\beta < 1$, τότε η ποσότητα

$$\mathcal{R}(A^k) = -\ln\left(\|A^k\|_\beta^{\frac{1}{k}}\right) = -\frac{1}{k} \ln \|A^k\|_\beta \quad (1.13)$$

καλείται μέση ταχύτητα σύγκλισης για k επαναλήψεις. Αν $\|B^k\|_\beta < 1$ και $\mathcal{R}(A^k) < \mathcal{R}(B^k)$, τότε ο B είναι επαναληπτικά ταχύτερος του A για k επαναλήψεις.

Όταν $k \rightarrow \infty$ μπορεί να αποδειχθεί ότι

$$\mathcal{R}_\infty(A) = \lim_{k \rightarrow \infty} \mathcal{R}(A^k) = -\ln(p(A)). \quad (1.14)$$

Η $\mathcal{R}_\infty(A)$ καλείται μέση ασυμπτωτική ταχύτητα σύγκλισης. Δηλαδή το πόσο γρήγορα συγκλίνει η μέθοδος (1.4) είναι συνάρτηση μόνο της $p(T)$. Μάλιστα ισχύει η παρακάτω πρόταση :

Όσο μικρότερη είναι η φασματική ακτίνα του επαναληπτικού πίνακα T τόσο ταχύτερα (ασυμπτωτικά) η ακολουθία $\{x^{(m)}\}_{m=0}^\infty$ συγκλίνει στην λύση του $Ax = b$.

Εν γένει η εύρεση του ω_{opt} (του ω για το οποίο η φασματική ακτίνα του $\mathcal{L}_\omega$ γίνεται ελάχιστη) είναι πρόβλημα δύσκολο. Πάντως, για μια κατηγορία πινάκων (τους (block) 2-cyclic consistently ordered [VAR00] ή έχοντες την ιδιότητα A (Young)) έχουν παραχθεί πολλά αποτελέσματα για την αντιμετώπισή τους (π.χ. [YOU54, YEI70, HPS88, SPA89, HSA92, SIF02]). Παρακάτω θα αναφέρουμε τα βασικά στοιχεία της θεωρίας αυτών των πινάκων.

Ορισμός 1.4 Έστω $A \in \mathbb{R}^{n,n}$. Ο A ονομάζεται *weakly cyclic of index k* ($k > 1$) αν υπάρχει μεταθετικός πίνακας $P \in \mathbb{R}^{n,n}$ τέτοιος ώστε

$$PAP^t = \begin{bmatrix} 0 & 0 & \cdots & 0 & A_{1k} \\ A_{21} & 0 & \cdots & 0 & 0 \\ 0 & A_{32} & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & \cdots & A_{k(k-1)} & 0 \end{bmatrix}, \quad (1.15)$$

όπου οι διαγώνιοι μηδενικοί πίνακες είναι τετραγωνικοί.

Θεώρημα 1.5 (Romanovsky 1936) Αν ο $A \in \mathbb{R}^{n,n}$ είναι *weakly cyclic of index k* ($k > 1$) ισχύει ότι

$$\phi(\lambda) = \det(\lambda I - A) = \lambda^m \prod_{i=1}^r (\lambda^k - \sigma_i^k), \quad (1.16)$$

όπου $m + rk = n$ και σ_i μη μηδενική ιδιοτιμή του A .

Πόρισμα 1.2 Οι μη μηδενικές ιδιοτιμές ενός *weakly cyclic of index 2* πίνακα εμφανίζονται σε αντίθετα ζεύγη.

Ορισμός 1.6 Εάν ο επαναληπτικός πίνακας του Jacobi που αντιστοιχεί στον πίνακα A , με την διαμέριση που φαίνεται στην (1.7), είναι *weakly cyclic of index 2* τότε ο πίνακας A ονομάζεται *p-cyclic* (αναφορικά πάντα με την διαμέριση (1.7)).

Ορισμός 1.7 Έστω ότι ο A είναι (αναφορικά με την διαμέριση (1.7)) *p-cyclic* και A_D, A_L, A_U όπως ορίζονται στην (1.8). Ο A ονομάζεται συνεπώς διατεταγμένος (*consistently ordered*) αν το

$$\sigma(\alpha L_A + \alpha^{-(p-1)} U_A)$$

είναι ανεξάρτητο του α για κάθε $\alpha \neq 0$. Το $\sigma(\cdot)$ συμβολίζει το φάσμα (το σύνολο των ιδιοτιμών).

Πόρισμα 1.3 Κάθε block τριδιαγώνιος πίνακας, με τους διαγώνιους block υποπίνακες του να είναι μη ιδιόμορφοι (ομαλοί), είναι *2-cyclic consistently ordered* δηλαδή *consistently ordered* και *2-cyclic* ταυτόχρονα.

Το βασικότερο θεώρημα που αφορά στους *2-cyclic consistently ordered* πίνακες είναι το παρακάτω το οποίο συνδέει τις ιδιοτιμές του επαναληπτικού πίνακα του Jacobi (J) με τις ιδιοτιμές του επαναληπτικού πίνακα της SOR ($\mathcal{L}_\omega$) [YOU81].

Θεώρημα 1.8 Έστω ότι ο A είναι *2-cyclic consistently ordered* πίνακας. Αν $\mu \in \sigma(J)$ και λ ικανοποιεί την σχέση

$$(\lambda + \omega - 1)^2 = \omega^2 \mu^2 \lambda \quad (1.17)$$

τότε $\lambda \in \sigma(\mathcal{L}_\omega)$. Αντίστροφα, αν $\lambda \in \sigma(\mathcal{L}_\omega) \setminus \{0\}$ και μ ικανοποιεί την (1.17) τότε $\mu \in \sigma(J)$.

Συμπέρασμα 1.9 Εάν ο A είναι *2-cyclic consistently ordered* τότε $\rho(T_{GS}) = (\rho(J))^2$. Δηλαδή η μέθοδος GS είναι δυο φορές πιο γρήγορη από την μέθοδο του Jacobi - αν φυσικά η μέθοδος του Jacobi συγκλίνει.

Θεώρημα 1.10 Έστω ότι ο A είναι *2-cyclic consistently ordered* πίνακας, ο επαναληπτικός πίνακας του Jacobi έχει πραγματικές ιδιοτιμές και ότι $\bar{\mu} = \rho(J) < 1$. Τότε η βέλτιστη τιμή της παραμέτρου ω για την SOR μέθοδο δίνεται από την σχέση

$$\omega_{opt} = \frac{2}{1 + \sqrt{1 - \bar{\mu}^2}},$$

και η φασματική ακτίνα της βέλτιστης SOR ισούται με

$$\rho(\mathcal{L}_{\omega_{opt}}) = \omega_{opt} - 1 = \frac{\mu^2}{(1 + \sqrt{1 - \bar{\mu}^2})^2}.$$

Επίσης ισχύει

$$\rho(\mathcal{L}_{\omega}) = \begin{cases} \omega - 1, & \omega_{opt} \leq \omega \leq 2 \\ 1 - \omega + \frac{1}{2}\omega^2\bar{\mu}^2 + \omega\bar{\mu}\sqrt{1 - \omega + \frac{1}{4}\omega^2\bar{\mu}^2}, & 0 < \omega \leq \omega_{opt} \end{cases}.$$

1.3 Μέθοδοι Προβολής (Projection methods).

Έστω ότι έχουμε να λύσουμε το ακόλουθο γραμμικό σύστημα

$$Ax = b, \quad (1.18)$$

όπου $A \in \mathbb{R}^{n,n}$ και $b \in \mathbb{R}^n$. Ας υποθέσουμε επίσης ότι ο A είναι αντιστρέψιμος.

Με τις direct (άμεσες) μεθόδους υπολογίζουμε την λύση $x \in \mathbb{R}^n$ απευθείας από την σχέση $x = A^{-1}b$ (χωρίς βέβαια να υπολογίζουμε τον A^{-1}). Η ιδέα των μεθόδων προβολής είναι να αναζητήσουμε μια προσεγγιστική λύση ($\tilde{x} \approx A^{-1}b$) όχι σε ολόκληρο τον $\mathbb{R}^n$ αλλά σε κάποιον υπόχωρο του $\mathcal{K}$ διάστασης $m \leq n$ απαιτώντας το $b - A\tilde{x}$ (υπόλοιπο ή residual) να είναι κάθετο σε ένα υπόχωρο $\mathcal{L}$ του $\mathbb{R}^n$ διάστασης m . Δηλαδή έχουμε να λύσουμε το πρόβλημα :

$$\text{Βρές } \tilde{x} \in \mathcal{K} \text{ τέτοιο ώστε } b - A\tilde{x} \perp \mathcal{L}. \quad (1.19)$$

Αν $\mathcal{K} \equiv \mathcal{L}$ η μέθοδος ονομάζεται orthogonal αλλιώς ονομάζεται oblique. Αν ξέρουμε ότι το $x_0 \in \mathbb{R}^n$ βρίσκεται "κοντά" στην λύση $A^{-1}b$ είναι λογικό να ψάχνουμε προσεγγίσεις της λύσης στον χώρο $x_0 + \mathcal{K}$ οπότε το πρόβλημα (1.19) μετασχηματίζεται στο εξής:

$$\text{Βρές } \tilde{x} \in x_0 + \mathcal{K} \text{ τέτοιο ώστε } b - A\tilde{x} \perp \mathcal{L}. \quad (1.20)$$

Το x_0 θα καλείται εφεξής αρχική προσέγγιση. Έστω τώρα ότι τα διανύσματα $\{v_i\}_{i=1}^m$ και $\{w_i\}_{i=1}^m$ αποτελούν βάσεις των χώρων $\mathcal{K}$ και $\mathcal{L}$ αντίστοιχα. Για να συνεχίσουμε ορίζουμε τους $n \times m$ πίνακες

$$V = \begin{bmatrix} | & | & & | \\ | & | & & | \\ v_1 & v_2 & \cdots & v_m \\ | & | & & | \\ | & | & & | \end{bmatrix} \quad \text{και} \quad W = \begin{bmatrix} | & | & & | \\ | & | & & | \\ w_1 & w_2 & \cdots & w_m \\ | & | & & | \\ | & | & & | \end{bmatrix}.$$

Η προσεγγιστική λύση μπορεί τώρα να γραφεί ως

$$\tilde{x} = x_0 + Vy, \quad (1.21)$$

όπου y διάνυσμα στήλης του $\mathbb{R}^m$. Ορίζοντας $r_0 = b - Ax_0$ και λαμβάνοντας υπόψιν την συνθήκη καθετότητας, διαδοχικά έχουμε :

$$\begin{aligned} b - A\tilde{x} &\perp \mathcal{L} \Leftrightarrow \\ b - A(x_0 + Vy) &\perp \mathcal{L} \Leftrightarrow \\ (b - Ax_0) - AVy &\perp \mathcal{L} \Leftrightarrow \\ r_0 - AVy &\perp \mathcal{L} \Leftrightarrow \\ \left. \begin{array}{l} r_0 - AVy \perp w_1 \\ r_0 - AVy \perp w_2 \\ \vdots \\ r_0 - AVy \perp w_m \end{array} \right\} &\Leftrightarrow \\ \left. \begin{array}{l} \langle r_0 - AVy, w_1 \rangle = 0 \\ \langle r_0 - AVy, w_2 \rangle = 0 \\ \vdots \\ \langle r_0 - AVy, w_m \rangle = 0 \end{array} \right\} &\Leftrightarrow \\ \left. \begin{array}{l} \langle AVy, w_1 \rangle = \langle r_0, w_1 \rangle \\ \langle AVy, w_2 \rangle = \langle r_0, w_2 \rangle \\ \vdots \\ \langle AVy, w_m \rangle = \langle r_0, w_m \rangle \end{array} \right\} &\Leftrightarrow \\ W^t AVy = W^t r_0 &\Leftrightarrow \left(\text{αν υπάρχει ο } (W^t AV)^{-1} \right) \\ y = (W^t AV)^{-1} W^t r_0, & \end{aligned}$$

και επομένως

$$\tilde{x} = x_0 + V (W^t AV)^{-1} W^t r_0. \quad (1.22)$$

Για τις μεθόδους, με τις οποίες θα ασχοληθούμε δεν είναι απαραίτητος ο σχηματισμός του $(W^t AV)^{-1}$. Ακόμα, πρέπει να σημειωθεί ότι, η ύπαρξη του αντιστρόφου του A δεν συνεπάγεται αναγκαστικά και την ύπαρξη του αντιστρόφου του $W^t AV$.

Παρατήρηση 1.2 Δυο παρατηρήσεις σχετικά με την γενικότητα των *projection* μεθόδων.

1. Η *projection* μέθοδος με $\mathcal{K} \equiv \mathcal{L} \equiv \mathbb{R}^n$ είναι άμεση. Αυτό φαίνεται από την σχέση (1.22) αν θέσουμε $W = V = I$ οπότε και παίρνουμε $\tilde{x} = A^{-1}b$ που είναι η ακριβής λύση της (1.18).
2. Η *projection* μέθοδος με $\mathcal{K} \equiv \mathcal{L} = \langle e_i \rangle^1$ και $x_0 = x^{(k)}$, όπου $x^{(k)}$ η προσεγγιστική λύση της μεθόδου *Gauss-Seidel* στο k βήμα, παράγει τη i -οστή συνιστώσα της προσεγγιστικής λύσης της *Gauss-Seidel* στο $k+1$ βήμα. Επομένως αν εφαρμόσουμε την *projection* μέθοδο διαδοχικά για $i = 1, \dots, n$ παράγουμε όλη την λύση $x^{(k+1)}$.

1.3.1 Krylov Subspace Μέθοδοι.

Μια μεγάλη κατηγορία (και ίσως η πλέον αναπτυσσόμενη στις μέρες μας) των μεθόδων προβολής είναι οι μέθοδοι υποχώρων Krylov (Krylov subspace).

Ορισμός 1.11 Έστω $A \in \mathbb{R}^{n \times n}$ και $u \in \mathbb{R}^n$. Ως *Krylov* (υπο)χώρο του A σε σχέση με το διάνυσμα u ορίζουμε τον υπόχωρο του $\mathbb{R}^n$ που παράγεται από τα διανύσματα $\{u, Au, A^2u, \dots, A^{m-1}u\}$ δηλαδή

$$K_m := K_m(A, u) = \langle u, Au, A^2u, \dots, A^{m-1}u \rangle. \quad (1.23)$$

Ο K_m έχει διάσταση το πολύ m . Krylov subspace μέθοδοι ονομάζονται οι *projection* μέθοδοι στις οποίες $\mathcal{K} \equiv K_m(A, r_0)$. Προφανώς $x_m = x_0 + \vec{w}$, όπου $\vec{w}$ είναι ένας γραμμικός συνδυασμός των διανυσμάτων παραγωγής, και επομένως

$$A^{-1}b \approx \tilde{x} = x_m = x_0 + q_{m-1}(A) r_0, \quad (1.24)$$

όπου q_{m-1} είναι κάποιο πολυώνυμο του πίνακα A βαθμού το πολύ $m-1$. Αν για αρχική προσέγγιση πάρουμε το $x_0 = 0$ η σχέση (1.24) απλοποιείται στην :

$$A^{-1}b \approx \tilde{x} = x_m = q_{m-1}(A) b. \quad (1.25)$$

Γιατί όμως ο $K_m(A, b)$ είναι ένας χώρος ο οποίος περιέχει την ακριβή λύση $A^{-1}b$ ή έστω μια καλή προσέγγιση της;

Ας το δούμε για την περίπτωση $x_0 = 0$ και για τυχαίο δεύτερο μέλος b . Αρκεί να υπάρχει πολυώνυμο q_{m-1} τέτοιο ώστε : $q_{m-1}(A) b = A^{-1}b$ για κάθε διάνυσμα b . Δηλαδή αρκεί ο A^{-1} να γράφεται ως γραμμικός συνδυασμός δυνάμεων του A . Ξέρουμε ότι σε κάθε πίνακα $A \in \mathbb{R}^{n \times n}$

¹ Με $\langle e_i \rangle$ συμβολίζουμε τον υπόχωρο που παράγεται από το e_i .

αντιστοιχεί ένα μοναδικό πολυώνυμο $r(A)$ βαθμού $t \leq n$ τέτοιο ώστε $r(A) = 0$, το ονομαζόμενο ελάχιστο ² (minimal) πολυώνυμο του A :

$$r(A) = a_0 I + a_1 A + a_2 A^2 + \cdots + a_t A^t = 0. \quad (1.26)$$

Αν ο A είναι αντιστρέψιμος $a_0 \neq 0$ και η παραπάνω εξίσωση γράφεται

$$A^{-1} = -\frac{1}{a_0} \sum_{j=0}^{t-1} a_{j+1} A^j. \quad (1.27)$$

Άρα ισχύει το παρακάτω θεώρημα :

Θεώρημα 1.12 Αν το ελάχιστο πολυώνυμο ενός αντιστρέψιμου πίνακα είναι βαθμού t , τότε η λύση του $Ax = b$ περιέχεται στον χώρο $K_t(A, b)$.

Επομένως, όχι μόνο είμαστε σίγουροι ότι η λύση που ψάχνουμε υπάρχει στον Krylov χώρο αλλά για κατάλληλους πίνακες μπορούμε να περιμένουμε ότι ο χώρος αυτός έχει μικρή διάσταση με αποτέλεσμα σχετικά γρήγορη σύγκλιση των μεθόδων.

1.3.2 Κατασκευή ορθοκανονικής βάσης του χώρου $K_m(A, r_0)$.

Αφού θα προσεγγίσουμε την λύση από τον υπόχωρο $K_m(A, r_0)$ θα πρέπει να ξέρουμε μια βάση του. Διαδεδομένη λύση σε αυτή την κατεύθυνση αποτελεί η ορθοκανονικοποίηση των διανυσμάτων παραγωγής $\{r_0, A r_0, \dots, A^{m-1} r_0\}$ του $K_m(A, r_0)$. Πολύ γνωστές διαδικασίες ορθοκανονικοποίησης αποτελούν οι μέθοδοι (modified) Gram-Schmidt και Householder. Ο αλγόριθμος που χρησιμοποιείται συνήθως για την παραγωγή βάσης υπόχωρων Krylov είναι ο modified Gram-Schmidt (MGS) και η διαδικασία είναι γνωστή με το όνομα Arnoldi MGS.

Αλγόριθμος 1.1 (Arnoldi - MGS)

² Το ελάχιστο πολυώνυμο είναι το πολυώνυμο ελάχιστου βαθμού από όλα τα πολυώνυμα p τα οποία ικανοποιούν την σχέση $p(A) = 0$.

```

1  Διάλεξε διάνυσμα  $u_1$  τέτοιο ώστε  $\|u_1\|_2 = 1$  ( $u_1 = r_0 / \|r_0\|_2$ )
2  For  $j = 1$  to  $m$ 
3       $w_j = A u_j$ 
4      For  $i = 1$  to  $j$ 
5           $h_{ij} = (w_j, u_i)$ 
6           $w_j = w_j - h_{ij} u_i$ 
7      End
8       $h_{j+1,j} = \|w_j\|_2$ 
9      If  $h_{j+1,j} = 0$  STOP
10      $u_{j+1} = w_j / h_{j+1,j}$ 
11 End

```

Ο αλγόριθμος τερματίζει όταν δημιουργήσει μια ορθοκανονική βάση του $K_m(A, r_0)$ άσχετα με το αν τερματίσει στο j βήμα οπότε η διάσταση του $K_m(A, r_0)$ είναι j και όχι m .

Θεώρημα 1.13 Αν συμβολίσουμε με V_m τον $n \times m$ πίνακα με στήλες τα διανύσματα $u_1, u_2, \dots, u_m$ με $\bar{H}_m$ τον $(m+1) \times m$ άνω Hessenberg πίνακα με μη μηδενικά στοιχεία του τα h_{ij} και με H τον $m \times m$ πίνακα που προκύπτει από τον $\bar{H}_m$ αν αφαιρέσουμε την τελευταία του γραμμή προκύπτουν οι παρακάτω σχέσεις

$$A V_m = V_m H_m + w_m e_m^t, \quad (1.28)$$

$$A V_m = V_{m+1} \bar{H}_m, \quad (1.29)$$

$$V_m^t A V_m = H_m. \quad (1.30)$$

Παρατήρηση 1.3 Ο αλγόριθμος του Arnoldi, όπως φαίνεται από την σχέση (1.28), μπορεί να χρησιμοποιηθεί και για τον υπολογισμό μερικών ιδιοτιμών του πίνακα A (για αυτόν τον λόγο εξάλλου δημιουργήθηκε). Οι ιδιοτιμές του Hessenberg πίνακα που παράγει (ritz values του A) είναι προσεγγίσεις των ιδιοτιμών του A . Αυτό ισχύει ακόμα και για μικρές τιμές του m οπότε οι ritz values είναι προσεγγίσεις των πιο περιφερειακών ιδιοτιμών του A .

1.3.3 Generalized Minimum Residual Method (GMRES)

Είναι η projection μέθοδος με

$$\begin{cases} \mathcal{K} \equiv K_m(A, r_0) \\ \mathcal{L} \equiv A \mathcal{K} \end{cases}, \quad (1.31)$$

όπου στον αλγόριθμο του Arnoldi επιλέγουμε $u_1 = r_0 / \|r_0\|_2$.

Θεώρημα 1.14 Έστω $\tilde{x} \in x_0 + K_m$ η λύση που δίνει η GMRES. Ισχύει ότι :

$$\|b - A\tilde{x}\|_2 = \min_{x \in x_0 + K_m} \|b - Ax\|_2. \quad (1.32)$$

Δηλαδή η GMRES ελαχιστοποιεί την ευκλείδια νόρμα του υπόλοιπου $b - A\tilde{x}$ πάνω στον χώρο $x_0 + K_m$.

Σύμφωνα με την σχέση (1.21) κάθε διάνυσμα $x \in x_0 + K_m$ γράφεται ως $x = x_0 + V_m y$ με $V = V_m$ ορθογώνιο πίνακα, αφού τα διανύσματα u_i είναι από κατασκευής ορθοκανονικά. Ακόμα ορίζουμε την συνάρτηση

$$J(y) = \|b - Ax\|_2 = \|b - A(x_0 + V_m y)\|_2. \quad (1.33)$$

Όμως :

$$\begin{aligned} b - Ax &= \\ &= b - A(x_0 + V_m y) \\ &= r_0 - AV_m y \\ &= \beta u_1 - V_{m+1} \bar{H}_m y \quad (\beta = \|r_0\|_2, \text{ (1.29)}) \\ &= \beta V_{m+1} e_1 - V_{m+1} \bar{H}_m y \\ &= V_{m+1} (\beta e_1 - \bar{H}_m y). \end{aligned}$$

Άρα

$$\begin{aligned} \|b - Ax\|_2 &= \\ &= \|V_{m+1} (\beta e_1 - \bar{H}_m y)\|_2 \\ &= \|V_{m+1}\|_2 \|\beta e_1 - \bar{H}_m y\|_2 \\ &= \|\beta e_1 - \bar{H}_m y\|_2. \end{aligned} \quad (1.34)$$

Επομένως η $J(y)$ ελαχιστοποιείται για

$$y_m = \min_y \|\beta e_1 - \bar{H}_m y\|_2.$$

Η εύρεση του y_m απαιτεί την λύση ενός $(m+1) \times m$ προβλήματος ελάχιστων τετραγώνων που είναι εύκολο να λυθεί (όταν το m είναι σχετικά μικρό). Τελικά η προσεγγιστική λύση που παράγει η GMRES είναι η

$$\tilde{x} = x_0 + V_m y_m. \quad (1.35)$$

Σύμφωνα με όσα έχουμε πει ως τώρα ο αλγόριθμος της GMRES θα είναι :

Αλγόριθμος 1.2 (GMRES)

```

1  For  $m = 1, 2, \dots$ 
2      βήμα  $m$  του Arnoldi
3      Βρες το  $y$  που ελαχιστοποιεί την  $\|\beta e_1 - \bar{H}_m y\|_2$ 
4       $x_m = x_0 + V_m y$ 
6      check convergence ( $\|x_m - x_{m-1}\| < \epsilon$ )
7  End

```

Το πρόβλημα ελάχιστων τετραγώνων μπορεί να λυθεί με παραγοντοποίηση QR με κόστος $O(m^2)$. Πάντως αν χρησιμοποιήσουμε περιστροφές Givens (Givens rotations) μπορούμε να παίρνουμε συνεχώς την QR παραγοντοποίηση του $\bar{H}_m$ από την QR παραγοντοποίηση του $\bar{H}_{m-1}$. Με αυτή την προσέγγιση το κόστος πέφτει σε $O(m)$ και επιπλέον έχουμε μια απλή σχέση για το υπόλοιπο $b - Ax$ που μας δίνει ένα κριτήριο για το πότε θα σταματήσουμε την επαναληπτική διαδικασία. Για περισσότερες λεπτομέρειες [TBA97, SAA00].

Καθώς προχωράμε στις επαναλήψεις οι απαιτήσεις σε μνήμη ($\bar{H}_m$ και V_m) αλλά και σε υπολογιστικό κόστος (ορθοκανονικοποίηση) αυξάνονται απαγορευτικά. Μια προφανής λύση είναι να επανεκκινούμε τον αλγόριθμο κάθε k βήματα [DDS98]. Αυτό μπορεί να επηρεάζει αρνητικά την σύγκλιση (χρειαζόμαστε περισσότερες επαναλήψεις) αλλά κάνει την μέθοδο πρακτικά εφαρμόσιμη. Η παραλλαγή αυτή της μεθόδου λέγεται GMRES(restart= k) σε αντιδιαστολή με την full GMRES (χωρίς επανεκκίνηση). Να σημειωθεί ότι δεν μπορούμε να ξέρουμε εκ' των προτέρων την βέλτιστη τιμή του k η οποία εξαρτάται από το πρόβλημα που έχουμε να λύσουμε (τον πίνακα των συντελεστών A) καθώς και από την υπολογιστική ισχύ (μνήμη) που έχουμε στην διάθεση μας.

Σύγκλιση της GMRES

Έστω x_m η προσεγγιστική λύση που δίνει η GMRES στην επανάληψη m . Ορίζουμε $r_m = b - Ax_m$ το υπόλοιπο (residual).

Θεώρημα 1.15 Η GMRES συγκλίνει μονότονα υπό την έννοια ότι :

$$\|r_{m+1}\|_2 \leq \|r_m\|_2. \quad (1.36)$$

Θεώρημα 1.16 Έστω $\mathcal{P}_m = \{\text{πολυώνυμα πινάκων } p, \text{ βαθμού το πολύ } m : p(0) = I\}$.

Το πρόβλημα εύρεσης της προσεγγιστικής λύσης x_m της GMRES με αρχική προσέγγιση x_0 είναι ισοδύναμο με το πρόβλημα :

$$\text{Βρες πολυώνυμο } p_m \in \mathcal{P}_m \text{ τέτοιο ώστε η } \|p_m(A) r_0\|_2 \text{ να γίνεται ελάχιστη.} \quad (1.37)$$

Εν γένει το δεύτερο μέλος b (υπόλοιπο $r_0 = b - Ax_0$) δεν έχει κάποια ιδιαίτερη δομή και επομένως η σύγκλιση της GMRES καθορίζεται από την $\|p_m(A)\|_2$. Ποιές όμως ιδιότητες του πίνακα A καθορίζουν την $\|p_m(A)\|_2$; Η κατάσταση δεν είναι ξεκάθαρη ειδικά για τους μη normal πίνακες¹. Πάντως πολύ γενικά μπορούμε να πούμε ότι η GMRES συγκλίνει (ασυμπτωτικά) ταχύτερα όσο μικρότερη είναι η ακτίνα της συστάδας (cluster) των ιδιοτιμών του πίνακα A και όσο πιο μακριά βρίσκεται το μηδέν από το κέντρο αυτής της συστάδας [CIK95].

1.3.4 Biorthogonalization methods

Στην GMRES (full) όσο αυξάνει η διάσταση του Krylov υποχώρου τόσο αυξάνονται και οι απαιτήσεις για μνήμη και υπολογιστικό κόστος. Αυτό οφείλεται στην επιλογή της βάσης του Krylov υποχώρου (ορθοκανονική). Αν χαλαρώσουμε την ορθοκανονικότητα της βάσης μπορούμε να πάρουμε μεθόδους που το κόστος τους σε μνήμη και πράξεις, ανά επανάληψη, είναι ανεξάρτητο του αριθμού των επαναλήψεων. Μια τέτοια (μη ορθοκανονική) βάση μας δίνει ο παρακάτω αλγόριθμος.

Αλγόριθμος 1.3 (Lanczos biorthogonalization)

```

1  Διάλεξε δυο διανύσματα  $u_1, w_1$  τέτοια ώστε  $\langle u_1, w_1 \rangle = 0$ 
2  Θεσε  $\beta_1 = \delta_1 = 0$ ,  $w_0 = u_0 = \vec{0}$ 
3  For  $j = 1, 2, \dots, m$ 
4       $\alpha_j = \langle Au_j, w_j \rangle$ 
5       $\hat{u}_{j+1} = Au_j - \alpha_j u_j - \beta_j u_{j-1}$ 
6       $\hat{w}_{j+1} = A^t w_j - \alpha_j w_j - \delta_j w_{j-1}$ 
7       $\delta_{j+1} = |\langle \hat{u}_{j+1}, \hat{w}_{j+1} \rangle|^{1/2}$  If  $\delta_{j+1} = 0$  STOP
8       $\beta_{j+1} = \langle \hat{u}_{j+1}, \hat{w}_{j+1} \rangle / \delta_{j+1}$ 
9       $w_{j+1} = \hat{w}_{j+1} / \beta_{j+1}$ 
10      $u_{j+1} = \hat{u}_{j+1} / \delta_{j+1}$ 
11 End
```

Ο άνω Hessenberg πίνακας H_m του Arnoldi έχει δώσει την θέση στον τριδιαγώνιο πίνακα T_m

¹ Ένας, πραγματικός, πίνακας ονομάζεται normal αν αντιμετατίθεται με τον ανάστροφο του, δηλαδή αν ικανοποιεί την σχέση $AA^T = A^T A$.

όπου

$$T_m = \begin{bmatrix} \alpha_1 & \beta_2 & & & \\ \delta_2 & \alpha_2 & \beta_3 & & \\ & \ddots & \ddots & \ddots & \\ & & \delta_{m-1} & \alpha_{m-1} & \beta_m \\ & & & \delta_m & \alpha_m \end{bmatrix}. \quad (1.38)$$

Ο αλγόριθμος (1.3) κατασκευάζει ένα ζεύγος biorthogonal βάσεων

$$\{u_i\}_{i=1}^m, \{w_j\}_{j=1}^m, \quad (1.39)$$

των χώρων $K_m(A, u_1)$ και $K_m(A^t, w_1)$ αντίστοιχα. Δηλαδή $\langle u_i, w_j \rangle = \delta_{ij}$ ή σε μορφή πινάκων

$$W_m^t V_m = V_m^t W_m = I. \quad (1.40)$$

Ακόμα ισχύουν οι σχέσεις [SAA00]:

$$AV_m = V_m T_m + \delta_{m+1} u_{m+1} e_m^t \quad (1.41)$$

$$AW_m = W_m T_m^t + \beta_{m+1} w_{m+1} e_m^t \quad (1.42)$$

$$W_m^t AV_m = T_m. \quad (1.43)$$

Παρατήρηση 1.4 (Κατάρευση - Breakdown)

Αν παρατηρήσουμε τον αλγόριθμο (1.3) βλέπουμε ότι στην γραμμή 8 υπάρχει μια διαίρεση με το δ_{j+1} . Αν $\delta_{j+1} = 0$ είμαστε αναγκασμένοι να σταματήσουμε την διαδικασία (*breakdown*). Αυτό μπορεί να συμβεί αν κάποιο από τα $\hat{u}_{j+1}, \hat{w}_{j+1}$ (ή και τα δυο) είναι μηδέν ή ακόμα και όταν κανένα από τα $\hat{u}_{j+1}, \hat{w}_{j+1}$ δεν είναι μηδέν. Στην πρώτη περίπτωση αν $\hat{u}_{j+1} = 0$ έχουμε φτάσει σε μια βάση του $K_m(A, u_1)$ ενώ αν $\hat{w}_{j+1} = 0$ έχουμε φτάσει σε μια βάση του $K_m(A^t, w_1)$. Τότε λέμε ότι έχουμε κανονικό τερματισμό (*regular termination*) ή τυχερή κατάρευση (*lucky breakdown*). Στην δεύτερη περίπτωση είναι αδύνατον να κατασκευάσουμε τις ζητούμενες βάσεις (*serious breakdown*).

Επειδή δουλεύουμε σε αριθμητική πεπερασμένης ακρίβειας πρέπει να θεωρήσουμε και την πιθανότητα να έχουμε $\delta_{j+1} \approx 0$ (*near breakdown*).

1.3.5 Biconjugate Gradient Algorithm (BCG)

Είναι η projection μέθοδος με

$$\begin{cases} \mathcal{K} \equiv \langle u_1, Au_1, \dots, A^{m-1}u_1 \rangle \\ \mathcal{L} \equiv \langle w_1, Aw_1, \dots, A^{m-1}w_1 \rangle \end{cases}. \quad (1.44)$$

με συνηθέστερες επιλογές $u_1 = r_0 / \|r_0\|$ και $w_1 = u_1$ αν και οποιοδήποτε w_1 με $\langle u_1, w_1 \rangle \neq 0$ θα ήταν κατάλληλο. Έστω $T_m = L_m U_m$ η LU παραγοντοποίηση του T_m . Από την εξίσωση (1.22) έχουμε για την προσεγγιστική λύση x_m :

$$x_m = x_0 + V_m (W_m^t A V_m)^{-1} W_m^t r_0$$

$$x_m = x_0 + V_m T_m^{-1} W_m^t r_0 \quad (1.43)$$

$$x_m = x_0 + V_m T_m^{-1} \beta e_1 \quad (u_i, w_j \text{ biorthogonal})$$

$$x_m = x_0 + V_m U_m^{-1} L_m^{-1} \beta e_1$$

$$x_m = x_0 + P_m L_m^{-1} \beta e_1 \quad (P_m = V_m U_m^{-1}),$$

Το ότι η παραπάνω σχέση μπορεί να υλοποιηθεί ως επαναληπτικός αλγόριθμος είναι συνέπεια του ότι οι πίνακες U_m και L_m είναι τριγωνικοί.

Αλγόριθμος 1.4 (BCG)

```

1   $r_0 = b - Ax_0$ . Διαλεξε  $r_0^*$  τέτοιο ώστε  $\langle r_0, r_0^* \rangle \neq 0$  (συνήθως  $r_0^* = r_0$ )
2   $p_0 = r_0$ ,  $p_0^* = r_0^*$ 
3  For  $j = 1, 2, \dots$  do
4       $a_j = \langle r_j, r_j^* \rangle / \langle Ap_j, p_j^* \rangle$ 
5       $x_{j+1} = x_j + a_j p_j$ 
6       $r_{j+1} = r_j - a_j A p_j$ 
7       $r_{j+1}^* = r_j^* - a_j A^t p_j^*$ 
8       $\beta_j = \langle r_{j+1}, r_{j+1}^* \rangle / \langle r_j, r_j^* \rangle$ 
9       $p_{j+1} = r_{j+1} + \beta_j p_j$ 
10      $p_{j+1}^* = r_{j+1}^* + \beta_j p_j^*$ 
11 End
```

Η μέθοδος BCG έχει, εκτός της περίπτωσης του breakdown, δυο ακόμα μειονεκτήματα: Την ανάγκη πολλαπλασιασμού με τον A^t (σε μερικά προβλήματα δεν έχουμε τέτοια δυνατότητα) και συχνά μια μη ομαλή (erratic) συμπεριφορά του υπόλοιπου (που παρόλο στο τέλος φαίνεται να μην δημιουργεί πρόβλημα στην σύγκλιση μπορεί σε κάποιες περιπτώσεις να είναι αιτία σοβαρών σφαλμάτων στογγύλευσης [VOR03]). Από την δεκαετία του 80 αρχίζουν να εμφανίζονται διάφορες παραλλαγές της BCG οι οποίες προσπαθούν να θεραπεύσουν αυτά τα μειονεκτήματα. Θα

ασχοληθούμε με τέσσερις από αυτές τις μεθόδους Conjugate Gradients Squared (CGS), BCG Stabilized (BICGSTAB) , BCG Stabilized(2) (BICGSTAB(2)) και Transpose Free QMR (TFQMR)

CGS Ο Sonneveld [SON89] απέφυγε τους πολλαπλασιασμούς με τον A^t στην μέθοδο BCG "τετραγωνίζοντας" το πολυώνυμο $\phi_m(A)$ όπου

$$b - Ax_m^{BCG} = \phi_m(A) (b - Ax_0), \quad (1.45)$$

και παράγοντας προσεγγίσεις της λύσης που ικανοποιούν την σχέση

$$b - Ax_{2m}^{CGS} = (\phi_m(A))^2 (b - Ax_0). \quad (1.46)$$

Σε πολλές περιπτώσεις φαίνεται η CGS να συγκλίνει ταχύτερα από την BCG. Πάντως, σε αρκετές περιπτώσεις, παρατηρείται η επιδείνωση της ήδη erratic συμπεριφοράς της BCG.

BICGSTAB Ο Vorst [VOR92] κατασκεύασε μια παραλλαγή της CGS όπου έχει αντικαταστήσει το $(\phi_m(A))^2$ με το $\psi_m(A) \phi_m(A)$, όπου το $\psi_m(A)$ έχει σκοπό να ελαττώσει την erratic συμπεριφορά της CGS με αποτέλεσμα ομαλότερη σύγκλιση.

BICGSTAB(2) Οι Fokkema και Sleijpen [SFO93] δημοσίευσαν μια γενίκευση της BICGSTAB την BICGSTAB(l) στην οποία έχουν αντικαταστήσει το πολυώνυμο $\psi_m(A)$ με ένα ακριβέστερο (μεγαλύτερου βαθμού). Βέβαια όσο μεγαλύτερο είναι το l τόσο αυξάνεται και το υπολογιστικό κόστος της μεθόδου. Από αυτή την ομάδα θα χρησιμοποιήσουμε την BICGSTAB(2).

TFQMR Εισήχθη από τον Freund στο [FRE93] και είναι άλλη μια παραλλαγή της CGS που προσπαθεί να βελτιώσει την erratic συμπεριφορά της CGS ελαχιστοποιώντας κάποια "χοντινή" της $\|b - Ax_m\|_2$ ποσότητα.

Παρατήρηση 1.5 (Look ahead Lanczos)

Υπάρχουν τεχνικές που όταν ο αλγόριθμος του Lanczos βρίσκεται μπροστά από πιθανό *breakdown* υλοποιούν δυο (στην αρχική τους μορφή), ή όσα χρειαστούν, βήματα μαζί με αποτέλεσμα να παρακάμπτουν το *breakdown* χαλαρώνοντας την *biorthogonal* ιδιότητα των βάσεων σε *block biorthogonal*. Τέτοιες τεχνικές ονομάζονται *look ahead*. Οι αλγόριθμοι που χρησιμοποιούμε εδώ δεν χρησιμοποιούν τέτοιες τεχνικές.

Παρατήρηση 1.6 (Quasi-Minimal Residual)

Η συχνά erratic συμπεριφορά των *Biorthogonalization* μεθόδων οφείλεται στο ότι αυτές οι μέθοδοι δεν ελαχιστοποιούν το $\|b - Ax_m\|_2$ (όπως π.χ. η GMRES) σε κάποιον υπόχωρο. Αυτή η ελαχιστοποίηση είναι υπολογιστικά ασύμφορη και αντί αυτής μέθοδοι όπως η QMR ελαχιστοποιούν κάποια παραπλήσια (*problem depended*) ποσότητα της $\|b - Ax_m\|_2$, και για αυτό απαντώνται ως *Quasi-Minimal Residual* μέθοδοι.

Παρατήρηση 1.7 Αν και έχουν εφευρεθεί πολλές μέθοδοι που προσπαθούν να εξαλείψουν τα τρία μειονεκτήματα της BCG καμιά, μέχρις στιγμής, δεν φαίνεται να τα καταφέρνει εξίσου ικανοποιητικά και στα τρία ταυτόχρονα.

Παρατήρηση 1.8 Για γενικούς πίνακες, κριτήρια για την επιλογή της ταχύτερης Krylov επαναληπτικής μεθόδου δεν υπάρχουν. Το μόνο που γνωρίζουμε, βλέπε [DDS98], από πειραματικά αποτελέσματα και μόνο, είναι ένα κριτήριο που αφορά στην μικρότερη έλλειψη που περιέχει τις ιδιοτιμές του πίνακα: Αν ο κύριος άξονας αυτής της έλλειψης βρίσκεται πάνω στον πραγματικό άξονα τότε χρησιμοποιούμε την Bicgstab ενώ αν βρίσκεται πάνω στον φανταστικό άξονα χρησιμοποιούμε την Gmres. Σε περιπτώσεις όπου το φανταστικό μέρος των ιδιοτιμών του A είναι σχετικά μεγάλο η Bicgstab ($l > 1$) τείνει να συμπεριφέρεται καλύτερα από την Bicgstab [SFO93]. Ειδικά, τώρα, στην περίπτωση μη normal πινάκων οι ιδιοτιμές του A ίσως να μην παίζουν ρόλο στην σύγκλιση των μεθόδων [SSZ07]. Έχει δειχθεί από τους Greenbaum, Pták και Strakoš στα [GPS96, GST94] ότι το φάσμα του A από μόνο του μπορεί να οδηγήσει σε παραπλανητικές πληροφορίες στην μη normal περίπτωση. Ο συνδυασμός του δεξιού μέλους και των invariant υποχώρων που σχετίζονται με τον A φαίνεται να παρέχει πιο χρήσιμες πληροφορίες από ότι το φάσμα του A [SSZ05].

1.4 Προρύθμιση (Preconditioning).

Έστω ότι θέλουμε να λύσουμε το γραμμικό σύστημα

$$Ax = b \quad (1.47)$$

με μια επαναληπτική μέθοδο. Η ταχύτητα σύγκλισης της μεθόδου έχει άμεση σχέση με κάποιες (συνήθως φασματικές) ιδιότητες του πίνακα A . Η ιδέα της προρύθμισης (preconditioning) είναι αντί του αρχικού (1.47) να λύσουμε το προρυθμισμένο γραμμικό σύστημα

$$M^{-1}Ax = M^{-1}b, \quad (1.48)$$

έτσι ώστε ο πίνακας $M^{-1}A$ να έχει καλύτερες ιδιότητες από τον αρχικό πίνακα A . Ο πίνακας M ονομάζεται προρυθμιστής (preconditioner). Η ποιότητα του προρυθμιστή εξαρτάται από το πόσο καλά μπορεί να προσεγγίσει τον A . Εν γένει, όσο ποιοτικότερος είναι ο προρυθμιστής τόσο και πιο υπολογιστικά ακριβός είναι π.χ. αν $M = A$ το σύστημα έχει μετασχηματιστεί στο $Ix = b$ αλλά ο υπολογισμός του M^{-1} είναι τουλάχιστον τόσο δαπανηρός όσο η επίλυση του αρχικού προβλήματος, ενώ αν $M = I$ ο υπολογισμός του M^{-1} δεν κοστίζει τίποτα αλλά και καθόλου δεν βελτιώνει την σύγκλιση του αρχικού συστήματος. Τις περισσότερες φορές, για να είναι αποτελεσματικές οι Krylov μέθοδοι, η προρύθμιση είναι απαραίτητη.

Ένας καλός προρυθμιστής για μια μέθοδο δεν είναι απαραίτητα καλός και για μια διαφορετική μέθοδο.

Αν γράψουμε τον προρρυθμιστή πίνακα ως γινόμενο δυο μη ιδιόμορφων πινάκων $M = M_1 M_2$ μια γενική υλοποίηση της προρρυθμισης είναι η ακόλουθη :

$$M_1^{-1} A M_2^{-1} z = M_1^{-1} b, \text{ όπου } z = M_2 x. \quad (1.49)$$

Αν πάρουμε $M_2 = I$ έχουμε

$$M_1^{-1} A x = M_1^{-1} b \text{ (Left preconditioning)} \quad (1.50)$$

ενώ αν $M_1 = I$ έχουμε

$$A M_2^{-1} (M_2 x) = b \text{ (Right preconditioning)}. \quad (1.51)$$

Στα αριθμητικά παραδείγματα χρησιμοποιούμε left preconditioning. Εδώ πρέπει να σημειωθεί ότι δεν χρειάζεται να πολλαπλασιάσουμε εξ' αρχής τον πίνακα A με τον M^{-1} αλλά κάνουμε τις εξής αλλαγές στις μη προρρυθμισμένες μεθόδους :

Αντικαθιστούμε το	με το
b	b_{new} όπου : $M b_{new} = b$
$y = Ax$	y_{new} όπου : $M y_{new} = Ax$

Τρεις δημοφιλείς προρρυθμιστές προκύπτουν από την διάσπαση $A = D - L - U$ ενός πίνακα A όπου ο D είναι διαγώνιος, ο L είναι κάτω τριγωνικός και ο U είναι άνω τριγωνικός.

Jacobi Ο προρρυθμιστής πίνακας είναι ο

$$M_{JAC} = D. \quad (1.52)$$

GS Ο προρρυθμιστής πίνακας είναι ο

$$M_{GS} = (D - L). \quad (1.53)$$

SSOR Ο προρρυθμιστής πίνακας είναι ο

$$M_{SSOR} = \frac{1}{\omega(2-\omega)} (D - \omega L) D^{-1} (D - \omega U). \quad (1.54)$$

Αν θέσουμε $\omega = 1$ έχουμε τον SGS προρρυθμιστή

$$M_{SGS} = (D - L) D^{-1} (D - U), \quad (1.55)$$

ο οποίος χρησιμοποιείται περισσότερο δεδομένου ότι (α) ο βέλτιστο ω δεν είναι γνωστό και (β) η επιλογή του ω στην περιοχή της μονάδας δεν φαίνεται να επηρεάζει και τόσο την ποιότητα του προρρυθμιστή, τουλάχιστον για κάποιες κατηγορίες 2-cyclic πινάκων. [MPS98, MPS06, EVA93].

1.5 Ο Ταχύς Μετασχηματισμός Fourier (FFT)

1.5.1 Ο Διακριτός Μετασχηματισμός Fourier (DFT)

Ο Διακριτός Μετασχηματισμός Fourier (DFT) στον $\mathbb{C}^n$ είναι ένα γινόμενο πίνακα-διανύσματος. Αν $y = [y_0, y_1, \dots, y_n]^T$ ο DFT του διανύσματος $x = [x_0, x_1, \dots, x_n]^T$ τότε ισχύει

$$y_k = \sum_{j=0}^{n-1} \omega_n^{kj} x_j, \quad k = 0, 1, \dots, n-1, \quad (1.56)$$

όπου

$$\omega_n = e^{-i\frac{2\pi}{n}} = \cos \frac{2\pi}{n} - i \sin \frac{2\pi}{n}. \quad (1.57)$$

Ο αντίστροφος Διακριτός Μετασχηματισμός Fourier (IDFT) του διανύσματος y είναι το διάνυσμα x . Παρατηρήστε ότι η ω_n είναι μια n -οστή ρίζα της μονάδος ($\omega_n^n=1$) την οποία θα συμβολίζουμε και με ω , χωρίς τον δείκτη, όταν δεν υπάρχει περίπτωση σύγχυσης.

Στην γλώσσα των πινάκων ο DFT περιγράφεται από την σχέση

$$y = F_n x, \quad (1.58)$$

όπου

$$F_n = [f_{pq}]_{p,q=0}^{n-1}, \quad f_{pq} = \omega_n^{pq} = e^{-i\frac{2\pi}{n}pq}, \quad (1.59)$$

είναι ο $n \times n$ DFT πίνακας. Π.χ.

$$F_1 = [1], \quad F_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad F_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{bmatrix},$$

και γενικά

$$F_n = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & \omega & \omega^2 & \omega^3 & \omega^4 & \dots & \omega^{(n-1)} \\ 1 & \omega^2 & \omega^4 & \omega^6 & \omega^8 & \dots & \omega^{2(n-1)} \\ 1 & \omega^3 & \omega^6 & \omega^9 & \omega^{12} & \dots & \omega^{3(n-1)} \\ 1 & \omega^4 & \omega^8 & \omega^{12} & \omega^{16} & \dots & \omega^{4(n-1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{(n-1)} & \omega^{2(n-1)} & \omega^{3(n-1)} & \omega^{4(n-1)} & \dots & \omega^{(n-1)(n-1)} \end{bmatrix}. \quad (1.60)$$

Σχόλιο 1.17 Σε πολλά βιβλία η n -οστή ρίζα της μονάδος που χρησιμοποιείται είναι η $\omega_n = e^{i\frac{2\pi}{n}}$. Ακόμα ο Διακριτός Μετασχηματισμός Fourier είναι κανονικοποιημένος (πολλαπλασιασμένος με το $1/n$). Στην παρούσα εργασία ο DFT πίνακας θα αναφέρεται και ως Fourier πίνακας.

Θεώρημα 1.18 Για τον πίνακα F_n όπως ορίστηκε στην (1.59) ισχύουν

- Ο πίνακας *Fourier* είναι συμμετρικός

$$F_n^t = F_n \quad (1.61)$$

- Ο πίνακας $\frac{1}{\sqrt{n}}F_n$ είναι ορθομοναδιαίος (*unitary*), δηλαδή

$$F_n^* F_n = nI \quad (1.62)$$

όπου $F_n^* := \overline{F_n^t}$.

- Ο αντίστροφος του πίνακα *Fourier* είναι ο συζυγοανάστροφος του (πολλαπλασιασμένος με το $1/n$).

$$F_n^{-1} = \frac{1}{n} F_n^* = \frac{1}{n} \overline{F_n^t} \quad (1.63)$$

1.5.2 FFT

Το κόστος του υπολογισμού του DFT με τον προφανή τρόπο (απευθείας πολλαπλασιασμός του πίνακα F_n με το διάνυσμα x) είναι της τάξεως $\mathcal{O}(n^2)$. Ο Ταχύς Μετασχηματισμός Fourier (FFT) είναι ένας τρόπος να υπολογίσουμε τον DFT με κόστος $\mathcal{O}(n \log_2 n)$. Δεν θα δώσουμε περαιτέρω λεπτομέρειες για τον τρόπο υλοποίησης του FFT εκτός από το παρακάτω "διάσημο" θεώρημα των Cooley-Tukey για την περίπτωση όπου η διάσταση του Fourier πίνακα είναι δύναμη του 2.

Θεώρημα 1.19 (Cooley-Tukey radix 2 factorization) Αν $n = 2^r$ τότε ο πίνακας *Fourier* F_n μπορεί να παραγοντοποιηθεί ως εξής:

$$F_n = A_t \dots A_1 P_n, \quad (1.64)$$

όπου ο P_n είναι ένας μεταθετικός πίνακας και

$$A_k = I_{2^{t-k}} \otimes B_{2^k}, \quad B_{2^k} = \begin{bmatrix} I_r & \Omega_r \\ I_r & -\Omega_r \end{bmatrix}, \quad r = 2^{k-1}$$

$$\Omega_r = \text{diag}(1, \omega_k, \dots, \omega_k^{r-1}), \quad \omega_k = e^{-2\pi i/2^k}.$$

Το παραπάνω θεώρημα μας λέει ότι μπορούμε να γράψουμε το $y = F_n x$ στην μορφή

$$y = A_t \dots A_1 P_n x$$

το οποίο είναι μια ακολουθία από πολλαπλασιασμούς πίνακα-διανύσματος. Η πολυπλοκότητα $\mathcal{O}(n \log_2 n)$ προέρχεται από το γεγονός ότι οι πίνακες A_k είναι αραιοί (δυο μη-μηδενικά στοιχεία ανά γραμμή).

Σχόλιο 1.20 Ο αντίστροφος DFT μπορεί να παραχθεί από ένα ευθύ FFT αντικαθιστώντας όλες τις ρίζες της μονάδας με τις συζυγοανάστροφες τους και πολλαπλασιάζοντας με ένα παράγοντα $1/n$ στο τέλος.

Σχόλιο 1.21 Πολλές φορές θα αναφέρουμε σαν $FFT(n)$ ή FFT μεγέθους n τον Ταχύ Μετασχηματισμό Fourier για τον υπολογισμό του Διακριτού Μετασχηματισμού Fourier ενός διανύσματος n θέσεων.

Σχόλιο 1.22 Πόσο καλύτερο όμως είναι το $\mathcal{O}(n \log_2 n)$ από το $\mathcal{O}(n^2)$; Ας κάνουμε, την απλοϊκή, εικασία ότι ο κανονικός πολλαπλασιασμός χρειάζεται n^2 πράξεις και ο FFT χρειάζεται $n \log_2 n$ και ας κατασκευάσουμε ένα πίνακα με την αναλογία n^2 προς $n \log_2 n$ για διάφορα μήκη διανυσμάτων.

n	$n^2/(n \log_2 n)$
200	≈ 27
2 000	≈ 183
20 000	$\approx 1\,400$
200 000	$\approx 11\,358$
2 000 000	$\approx 95\,550$

Από τον παραπάνω πίνακα βλέπουμε ότι αν ο FFT μεγέθους 2 000 000 ¹ χρειάζεται ένα δευτερόλεπτο ο συμβατικός πολλαπλασιασμός θα απαιτούσε πάνω από μία ολόκληρη μέρα!

1.6 Circulant Πίνακες

Ορισμός 1.23 Ένας $n \times n$ πίνακας T ονομάζεται (σημειακός) Toeplitz αν έχει σταθερά στοιχεία κατά μήκος κάθε διαγωνίου, αν δηλαδή έχει την παρακάτω μορφή

$$T := \begin{bmatrix} t_1 & t_2 & t_3 & \cdots & t_n \\ t_{-2} & t_1 & t_2 & \ddots & \vdots \\ t_{-3} & t_{-2} & t_1 & \ddots & t_3 \\ \vdots & \ddots & \ddots & \ddots & t_2 \\ t_{-n} & \cdots & t_{-3} & t_{-2} & t_1 \end{bmatrix}. \quad (1.65)$$

Στην περίπτωση που τα $t_1, t_2, t_3, \dots, t_n, t_{-2}, t_{-3}, \dots, t_{-n}$ δεν είναι αριθμοί αλλά τετραγωνικοί πίνακες ο πίνακας T ονομάζεται block Toeplitz.

Ορισμός 1.24 Ένας $n \times n$ πίνακας C ονομάζεται σημειακός Circulant, ή απλά Circulant, τάξεως n αν έχει σταθερά στοιχεία κατά μήκος κάθε διαγωνίου του και τα στοιχεία της κάθε, πέραν της

¹ Ένα τέτοιου μεγέθους διάνυσμα είναι σύνθηδες στους σημερινούς υπολογισμούς αφού αν είναι μιγαδικό και διπλής ακρίβειας σε ένα pc μπορεί να αποθηκευθεί σε ένα κλάσμα της μνήμης RAM ($\approx 31MB$).

αρχικής, γραμμής του είναι τα στοιχεία της προηγούμενης γραμμής κυκλικά μετατοπισμένα κατά μια θέση προς τα δεξιά, αν δηλαδή έχει την παρακάτω μορφή

$$C := \begin{bmatrix} c_1 & c_2 & c_3 & \cdots & c_n \\ c_n & c_1 & c_2 & \ddots & \vdots \\ c_{n-1} & c_n & c_1 & \ddots & c_3 \\ \vdots & \ddots & \ddots & \ddots & c_2 \\ c_2 & \cdots & c_{n-1} & c_n & c_1 \end{bmatrix}. \quad (1.66)$$

Ένας *Circulant* πίνακας της μορφής (1.66) συμβολίζεται με

$$C = \text{circ}(c_1, c_2, \dots, c_n)$$

Ορισμός 1.25 Έστω $C_1, C_2, \dots, C_N$ τετραγωνικοί πίνακες τάξεως n . Ένας πίνακας C , της μορφής

$$C := \begin{bmatrix} C_1 & C_2 & C_3 & \cdots & C_n \\ C_n & C_1 & C_2 & \ddots & \vdots \\ C_{n-1} & C_n & C_1 & \ddots & C_3 \\ \vdots & \ddots & \ddots & \ddots & C_2 \\ C_2 & \cdots & C_{n-1} & C_n & C_1 \end{bmatrix}, \quad (1.67)$$

ονομάζεται *block-circulant* πίνακας και συμβολίζεται με

$$C = \text{bcirc}(C_1, C_2, \dots, C_n).$$

Οι *circulant* πίνακες μπορούν να οριστούν μονοσήμαντα από την πρώτη γραμμή (ή στήλη) τους. Το ίδιο ισχύει και για τους (*block-circulant*) οι οποίοι ορίζονται μονοσήμαντα μόνο από την πρώτη *block* γραμμή (ή στήλη) τους. Για τους *circulant* πίνακες ισχύουν οι παρακάτω βασικές ιδιότητες

Θεώρημα 1.26 Εάν A και B είναι *circulant* πίνακες τάξεως n και a_1, a_2 είναι βαθμωτές ποσότητες τότε

- Οι πίνακες $A^t, A^*, a_1 A + a_2 B$ και AB είναι *circulant*.
- Οι πίνακες A, B αντιμετατίθενται δηλαδή $AB = BA$.
- Εάν ο A είναι αντιστρέψιμος τότε ο αντίστροφος του είναι *circulant*.
- Οι ιδιοτιμές $\{\lambda_k\}_{k=1}^n$ του $A = \text{circ}(c_1, c_n, c_{n-1}, \dots, c_2)$ δίνονται από την σχέση

$$\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = F_n \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

- Ο πίνακας A διαγωνοποιείται από τον *Fourier* πίνακα F_n δηλαδή

$$A = F_n^{-1} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} F_n$$

Κεφάλαιο 2

Η Μέθοδος Φωκά

Πρόσφατα μια νέα αναλυτική μέθοδος για την μελέτη προβλημάτων συνοριακών τιμών για ολοκληρώσιμες PDEs σε δυο διαστάσεις εισήχθη από τον καθηγητή Φωκά (βλ. [FOK97, FOK01]). Στο παρόν κεφάλαιο θα περιγράψουμε αυτή την μέθοδο έχοντας ως πρόβλημα-μοντέλο την εξίσωση του Laplace σε ένα κυρτό πολυγωνικό χωρίο. Θα εμβαθύνουμε κυρίως σε θέματα που αφορούν στην αριθμητική υλοποίηση της μεθόδου. Σε πιθανά προβλήματα που μπορεί να αντιμετωπίσουμε στην υλοποίηση (overflows-underflows) καθώς και σε πιθανές παραλλαγές της βασικής αριθμητικής μεθόδου, όπως διαφορετικές συναρτήσεις βάσης - διαφορετική επιλογή των collocation σημείων, που προτάθηκε στο [FFX03].

2.1 Η Μέθοδος Φωκά

Μια διαφορική εξίσωση σε δυο διαστάσεις (x, y) ονομάζεται ολοκληρώσιμη (*integrable*) αν και μόνο αν μια σχετιζόμενη διαφορική μορφή-1 $W(x, y, k)$, $k \in \mathbb{C}$ είναι κλειστή, για παράδειγμα $dW = 0$ (μια εισαγωγή στις διαφορικές μορφές μπορεί να βρεθεί στο [MTR92]). Παραδείγματα ολοκληρώσιμων εξισώσεων αποτελούν οι γραμμικές μερικές διαφορικές εξισώσεις με σταθερούς συντελεστές και οι μη-γραμμικές Schrödinger και Korteweg-de Vries εξισώσεις. Μερικές φορές, για ελλειπτικές εξισώσεις είναι, βολικό να αντικαθιστούμε τις καρτεσιανές μεταβλητές (x, y) με τις μιγαδικές $(z, \bar{z}) = (x + iy, x - iy)$. Το επόμενο λήμμα μας παρέχει την ισοδύναμη εξίσωση του Laplace στο μιγαδικό επίπεδο.

Λήμμα 2.1 Η εξίσωση του Laplace

$$q_{xx}(x, y) + q_{yy}(x, y) = 0 \quad (2.1)$$

στο καρτεσιανό επίπεδο μπορεί να γραφεί στο μιγαδικό επίπεδο ως

$$q_{z\bar{z}}(z, \bar{z}) = 0. \quad (2.2)$$

Απόδειξη. Οι μιγαδικές μεταβλητές $z, \bar{z}$ ορίζονται ως

$$\begin{cases} z = x + iy \\ \bar{z} = x - iy \end{cases}.$$

Διαδοχικά έχουμε

$$\begin{aligned} \begin{cases} \frac{\partial}{\partial x} = \frac{\partial}{\partial z} \frac{\partial z}{\partial x} + \frac{\partial}{\partial \bar{z}} \frac{\partial \bar{z}}{\partial x} \\ \frac{\partial}{\partial y} = \frac{\partial}{\partial z} \frac{\partial z}{\partial y} + \frac{\partial}{\partial \bar{z}} \frac{\partial \bar{z}}{\partial y} \end{cases} &\Leftrightarrow \begin{cases} \frac{\partial}{\partial x} = \frac{\partial}{\partial z} + \frac{\partial}{\partial \bar{z}} \\ \frac{\partial}{\partial y} = i \frac{\partial}{\partial z} - i \frac{\partial}{\partial \bar{z}} \end{cases} \Leftrightarrow \begin{cases} \frac{\partial^2}{\partial x^2} = \left(\frac{\partial}{\partial z} + \frac{\partial}{\partial \bar{z}} \right) \left(\frac{\partial}{\partial z} + \frac{\partial}{\partial \bar{z}} \right) \\ \frac{\partial^2}{\partial y^2} = \left(i \frac{\partial}{\partial z} - i \frac{\partial}{\partial \bar{z}} \right) \left(i \frac{\partial}{\partial z} - i \frac{\partial}{\partial \bar{z}} \right) \end{cases} \Leftrightarrow \\ \begin{cases} \frac{\partial^2}{\partial x^2} = \frac{\partial^2}{\partial z^2} + 2 \frac{\partial^2}{\partial z \partial \bar{z}} + \frac{\partial^2}{\partial \bar{z}^2} \\ \frac{\partial^2}{\partial y^2} = -\frac{\partial^2}{\partial z^2} + 2 \frac{\partial^2}{\partial z \partial \bar{z}} - \frac{\partial^2}{\partial \bar{z}^2} \end{cases} &\Leftrightarrow \\ \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} &= 4 \frac{\partial^2}{\partial z \partial \bar{z}}. \end{aligned} \quad (2.3)$$

Άρα $q_{xx} + q_{yy} = 0 \Leftrightarrow q_{z\bar{z}} = 0$ □

Συνεπώς από το λήμμα (2.1) ξέρουμε ότι η μιγαδική μορφή της Laplace είναι

$$q_{z\bar{z}}(z, \bar{z}) = 0. \quad (2.4)$$

Ας υποθέσουμε ότι η συνάρτηση δυο μιγαδικών μεταβλητών $q(z, \bar{z})$, με σύνολο τιμών το $\mathbb{R}$, ικανοποιεί την (2.4) σε ένα απλό συνεκτικό χωρίο $\mathcal{D}$ με σύνορο το $\partial\mathcal{D}$. Η (2.4) είναι ισοδύναμη με την

$$(e^{-ikz} q_z)_{\bar{z}} = 0, \text{ για κάθε τυχαίο } k \in \mathbb{C}, \quad (2.5)$$

και επομένως

$$\oint_{\mathcal{D}} (e^{-ikz} q_z)_z dz = 0. \quad (2.6)$$

Εφαρμόζοντας την μιγαδική μορφή του θεωρήματος του Green παίρνουμε την σχέση

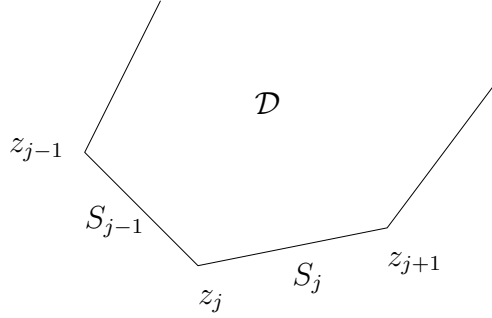
$$\int_{\partial\mathcal{D}} e^{-ikz} q_z dz = 0 \quad (2.7)$$

την οποία και ονομάζουμε ολική συνθήκη-σχέση (*Global Relation*) για την εξίσωση Laplace.

Έστω ότι το χωρίο $\mathcal{D}$ είναι ένα (κλειστό) κυρτό πολύγωνο με n κορυφές

$$z_1, z_2, \dots, z_n, z_{n+1} \equiv z_1,$$

αριθμημένες αντίθετα με την φορά των δεικτών του ρολογιού. Η z_{n+1} είναι εικονική κορυφή που ταυτίζεται με την z_1 . Έστω ακόμα ότι S_j η πλευρά του πολυγώνου από την κορυφή z_j έως την κορυφή z_{j+1} μη συμπεριλαμβανομένων των κορυφών αυτών. Τμήμα αυτού του πολυγωνικού χωρίου φαίνεται στο σχήμα (2.1).



Σχήμα 2.1: Τμήμα του (κλειστού) κυρτού πολυγώνου με κορυφές z_j , πλευρές S_j , και εσωτερικό D .

Επομένως η Global Relation για το πολυγωνικό αυτό χωρίο γράφεται

$$\sum_{j=1}^n \int_{S_j} e^{-ikz} q_z^{(j)} dz = 0. \quad (2.8)$$

Ορίζουμε ως

$$\alpha_j = \arg(z_{j+1} - z_j) \quad (2.9)$$

την γωνία που σχηματίζει η πλευρά S_j με τον άξονα των πραγματικών αριθμών, στο μιγαδικό επίπεδο, κατά την ορθή φορά. Έστω ότι η $g^{(j)}$ συμβολίζει την παράγωγο της q στην κατεύθυνση που σχηματίζει γωνία β_j , $0 \leq \beta_j \leq \pi$ με την πλευρά S_j . Δηλαδή (βλέπε σχήμα 2.2)

$$g^{(j)} = \cos(\beta_j) q_s^{(j)} + \sin(\beta_j) q_n^{(j)}, \quad (2.10)$$

όπου q_s^j και q_n^j συμβολίζουν την εφαπτομενική και την κάθετη (προς τα έξω) συνιστώσα της q_z στην πλευρά S_j αντίστοιχα. Έστω ακόμα ότι η $f^{(j)}$ συμβολίζει την παράγωγο της q στην κατεύθυνση που είναι κάθετη στην προηγούμενη. Δηλαδή

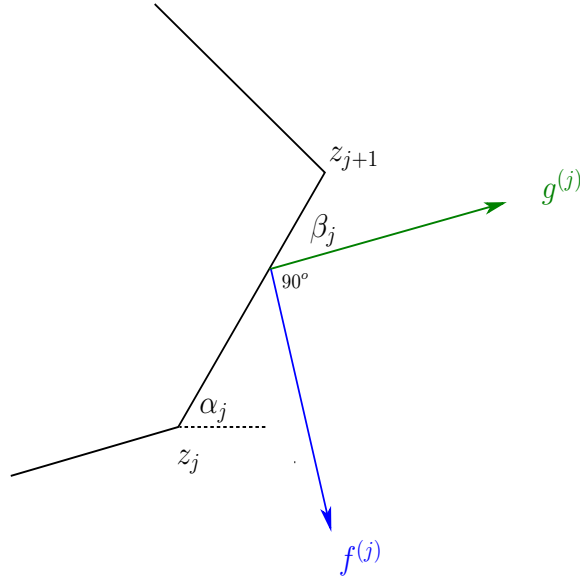
$$f^{(j)} = -\sin(\beta_j) q_s^{(j)} + \cos(\beta_j) q_n^{(j)}. \quad (2.11)$$

Το σχήμα (2.2) δείχνει την γωνία β_j και τις κατευθύνσεις πάνω στις οποίες έχουμε τις $g^{(j)}$, $f^{(j)}$ για την πλευρά S_j .

Από τις (2.10) και (2.11) παίρνουμε τις

$$\begin{cases} q_n^{(j)} &= \sin(\beta_j) g^{(j)} + \cos(\beta_j) f^{(j)} \\ q_s^{(j)} &= \cos(\beta_j) g^{(j)} - \sin(\beta_j) f^{(j)} \end{cases}. \quad (2.12)$$

Το επόμενο λήμμα μας δίνει την σχέση της μερικής παραγώγου της $q(z, \bar{z})$ ως προς z με τις ποσότητες $q_n^{(j)}$ και $q_s^{(j)}$.



Σχήμα 2.2: Οι κατευθύνσεις των $g^{(j)}, f^{(j)}$ για την πλευρά S_j .

Λήμμα 2.2 Η μερική παράγωγος ως προς z , της συνάρτησης $q(z, \bar{z})$ μπορεί να γραφεί ως

$$q_z^{(j)} = \frac{1}{2} e^{-i\alpha_j} (q_s^{(j)} + i q_n^{(j)}) \quad , \quad z \in S_j, \quad (2.13)$$

όπου $q_s^{(j)}$ η κατά κατεύθυνση παράγωγος της q στην κατεύθυνση του διανύσματος $\vec{\mu}_s = (\cos \alpha_j, \sin \alpha_j)$ και $q_n^{(j)}$ η κατά κατεύθυνση παράγωγος της q στην κατεύθυνση του διανύσματος $\vec{\mu}_n = (\sin \alpha_j, -\cos \alpha_j)$.

Απόδειξη. Από τον ορισμό της κατά κατεύθυνση παραγώγου έχουμε

$$\begin{cases} q_s^{(j)} = \vec{\mu}_s \cdot \nabla q^{(j)} = (\cos \alpha_j, \sin \alpha_j) \cdot (q_x^{(j)}, q_y^{(j)}) = q_x^{(j)} \cos \alpha_j + q_y^{(j)} \sin \alpha_j \\ q_n^{(j)} = \vec{\mu}_n \cdot \nabla q^{(j)} = (\sin \alpha_j, -\cos \alpha_j) \cdot (q_x^{(j)}, q_y^{(j)}) = q_x^{(j)} \sin \alpha_j - q_y^{(j)} \cos \alpha_j \end{cases} \Leftrightarrow$$

$$\begin{bmatrix} q_s^{(j)} \\ q_n^{(j)} \end{bmatrix} = \begin{bmatrix} \cos \alpha_j & \sin \alpha_j \\ \sin \alpha_j & -\cos \alpha_j \end{bmatrix} \begin{bmatrix} q_x^{(j)} \\ q_y^{(j)} \end{bmatrix} \Leftrightarrow \begin{bmatrix} q_x^{(j)} \\ q_y^{(j)} \end{bmatrix} = \begin{bmatrix} \cos \alpha_j & \sin \alpha_j \\ \sin \alpha_j & -\cos \alpha_j \end{bmatrix} \begin{bmatrix} q_s^{(j)} \\ q_n^{(j)} \end{bmatrix}.$$

Αλλά ξέρουμε ότι

$$q_z^{(j)} = \frac{1}{2} (q_x^{(j)} - i q_y^{(j)}), \quad (2.14)$$

και επομένως

$$q_z^{(j)} = \frac{1}{2} (q_s^{(j)} \cos \alpha_j + q_n^{(j)} \sin \alpha_j - i q_s^{(j)} \sin \alpha_j + i q_n^{(j)} \cos \alpha_j)$$

$$q_z^{(j)} = \frac{1}{2} [(\cos \alpha_j - i \sin \alpha_j) q_s^{(j)} + (\sin \alpha_j + i \cos \alpha_j) q_n^{(j)}]$$

$$q_z^{(j)} = \frac{1}{2} (\cos \alpha_j - i \sin \alpha_j) (q_s^{(j)} + i q_n^{(j)})$$

$$q_z^{(j)} = \frac{1}{2} e^{-i\alpha_j} (q_s^{(j)} + i q_n^{(j)})$$

□

Χρησιμοποιώντας τις (2.12) μαζί με το λήμμα (2.2) παίρνουμε την παρακάτω σχέση για την $q_z^{(j)}$

$$q_z^{(j)} = \frac{1}{2} e^{-i(\alpha_j - \beta_j)} (g^{(j)} + i f^{(j)}), \quad (2.15)$$

την οποία και αντικαθιστούμε στην Global Relation (2.8) παίρνοντας την παρακάτω ισοδύναμη μορφή της

$$\sum_{j=1}^n \int_{\tilde{S}_j} e^{-ikz} e^{-i(\alpha_j - \beta_j)} (g^{(j)}(z) + i f^{(j)}(z)) dz = 0, \quad (2.16)$$

την οποία και ονομάζουμε "Γενικευμένη Dirichlet-Neumann Απεικόνιση" (*Generalized Dirichlet-Neumann Map*).

Επομένως μπορούμε να δούμε την γενικευμένη Dirichlet-Neumann απεικόνιση ως μια σχέση μεταξύ δυο κάθετων συνιστωσών της παραγώγου της q στο σύνορο του πολυγωνικού χωρίου. Αυτή η σχέση μας δίνει την δυνατότητα γνωρίζοντας την παράγωγο της λύσης σε μια τυχαία κατεύθυνση (g) να μπορούμε να υπολογίσουμε την παράγωγο της λύσης σε μια κατεύθυνση κάθετη προς την προηγούμενη (f) χωρίς να λύσουμε στο εσωτερικό του χωρίου.

Από εδώ και στο εξής θα θεωρούμε την συνάρτηση $g(z)$ ως γνωστή συνοριακή συνθήκη και την συνάρτηση $f(z)$ ως την άγνωστη συνιστώσα της λύσης την οποία και αναζητούμε.

Παρατήρηση 2.1

- (α) Η γνώση των $g(z), f(z)$ στο σύνορο του πολυγώνου είναι ικανή να μας δώσει την λύση της Laplace στο σύνορο του πολυγώνου μέσω της γενικευμένης Dirichlet-Neumann απεικόνισης, και κατ' επέκταση και στο εσωτερικό του πολυγώνου (βλέπε [KAN06] για μια αριθμητική υλοποίηση). Δηλαδή, μιλώντας γενικά, έχουμε καταφέρει να ανάγουμε ένα διδιάστατο πρόβλημα (Laplace στο χωρίο) σε ένα μονοδιάστατο (Laplace στο σύνορο του πολυγώνου).
- (β) Βασικό στοιχείο για την αριθμητική επίλυση της Dirichlet-Neumann απεικόνισης αποτελεί το γεγονός ότι η Dirichlet-Neumann απεικόνιση ισχύει για κάθε μιγαδική τιμή της παραμέτρου k .
- (γ) Αν θεωρήσουμε $\beta_j = 0$ έχουμε Dirichlet συνοριακές συνθήκες ενώ αν $\beta_j = \pi/2$ έχουμε Neumann συνοριακές συνθήκες. Διαφορετικές τιμές της γωνίας β_j οδηγούν σε μεικτές (Poincaré) συνοριακές συνθήκες.

2.2 Η Αριθμητική Μέθοδος

Παρατηρώντας την (2.16) βλέπουμε ότι έχουμε n άγνωστες, μιγαδικές, συναρτήσεις $\{f^{(j)}(z)\}_{j=1}^n$. Υπενθυμίζουμε ότι η Dirichlet-Neumann απεικόνιση ισχύει για κάθε $k \in \mathbb{C}$ επομένως μπορούμε,

επιλέγοντας n διαφορετικές τιμές για το k , να κατασκευάσουμε ένα σύστημα n ολοκληρωτικών εξισώσεων του οποίου η λύση θα είναι οι άγνωστες συναρτήσεις $\{f^{(j)}(z)\}_{j=1}^n$.

Για να προχωρήσουμε στην κατασκευή της αριθμητικής μεθόδου θα χρησιμοποιήσουμε μια παραμετροποίηση για τις πλευρές του πολυγώνου έτσι ώστε να υπολογίσουμε ευκολότερα τα ολοκληρώματα που εμφανίζονται στην Dirichlet-Neumann απεικόνιση.

2.2.1 Παραμετροποίηση

Ο σκοπός μας είναι να απεικονίσουμε την κάθε πλευρά S_j του πολυγώνου στο διάστημα $[-\pi, \pi]$ έτσι ώστε

(α) Οι μιγαδικές συναρτήσεις $f^{(j)}(z)$ και $g^{(j)}(z)$ να αντιστοιχηθούν σε πραγματικές συναρτήσεις και

(β) η ολοκλήρωση να γίνεται πάνω στο $[-\pi, \pi]$ και όχι πάνω στο $[z_j, z_{j+1}]$.

Θέτουμε

$$m_j = \frac{z_{j+1} + z_j}{2} \quad (2.17)$$

και

$$h_j = \frac{z_{j+1} - z_j}{2\pi}. \quad (2.18)$$

Αμέσως έπεται ότι για $z \in S_j$ ισχύει

$$z = m_j + sh_j, \quad s \in (-\pi, \pi). \quad (2.19)$$

Ακόμα ισχύουν

$$h_j = |h_j| e^{i\alpha_j}, \quad (2.20\alpha')$$

$$s = -\pi \rightarrow z = z_j, \quad (2.20\beta')$$

$$s = \pi \rightarrow z = z_{j+1}, \quad (2.20\gamma')$$

$$\frac{dz}{ds} = h_j \rightarrow dz = h_j ds. \quad (2.20\delta')$$

Ξεκινώντας από την (2.16) και εφαρμόζοντας την παραπάνω παραμετροποίηση διαδοχικά, έχουμε

$$\begin{aligned}
\sum_{j=1}^n \int_{S_j} e^{-ikz} e^{-i(\alpha_j - \beta_j)} (g^{(j)}(z) + i f^{(j)}(z)) dz &= 0 & \Leftrightarrow \\
\sum_{j=1}^n \int_{z_j}^{z_{j+1}} e^{-ikz} e^{-i(\alpha_j - \beta_j)} (g^{(j)}(z) + i f^{(j)}(z)) dz &= 0 & \Leftrightarrow \\
\sum_{j=1}^n \int_{-\pi}^{\pi} h_j e^{-ik(m_j + sh_j)} e^{-i(\alpha_j - \beta_j)} (g^{(j)}(s) + i f^{(j)}(s)) ds &= 0 & \Leftrightarrow \\
\sum_{j=1}^n h_j e^{-ikm_j} e^{-i(\alpha_j - \beta_j)} \int_{-\pi}^{\pi} e^{-iksh_j} (g^{(j)}(s) + i f^{(j)}(s)) ds &= 0 & \Leftrightarrow \\
\sum_{j=1}^n |h_j| e^{i\alpha_j} e^{-ikm_j} e^{-i(\alpha_j - \beta_j)} \int_{-\pi}^{\pi} e^{-iksh_j} (g^{(j)}(s) + i f^{(j)}(s)) ds &= 0 & \Leftrightarrow \\
\sum_{j=1}^n |h_j| e^{i\beta_j} e^{-ikm_j} \int_{-\pi}^{\pi} e^{-iksh_j} (g^{(j)}(s) + i f^{(j)}(s)) ds &= 0 & \Leftrightarrow \\
\sum_{j=1}^n |h_j| e^{i(\beta_j - km_j)} \int_{-\pi}^{\pi} e^{-iksh_j} (f^{(j)}(s) - i g^{(j)}(s)) ds &= 0 & (2.21)
\end{aligned}$$

Ας σταθούμε λίγο εδώ. Η τελευταία μορφή της γενικευμένης Dirichlet-Neumann απεικόνισης (2.21) μας δείχνει ότι πλέον το πρόβλημα μας έχει γίνει η εύρεση n το πλήθος πραγματικών συναρτήσεων $\{f^{(j)}(s)\}_{j=1}^n$. Η (2.21) μπορεί να ειπωθεί σαν μια γεννήτρια εξισώσεων: επιλέγουμε μια τιμή για το k και παίρνουμε μια εξίσωση, επιλέγουμε διαφορετική τιμή για το k και παίρνουμε μια διαφορετική εξίσωση κ.ο.κ. .

2.2.2 Η Επιλογή του k

Η (2.21) ισχύει για κάθε μιγαδική τιμή του k . Εμείς χρειαζόμαστε ένα πεπερασμένο πλήθος εξισώσεων της μορφής (2.21) άρα και ένα πεπερασμένο πλήθος από k . Το ερώτημα που τίθεται είναι: **Ποιά k πρέπει να επιλέξουμε;** Το ότι έχουμε ένα πολύγωνο με n πλευρές μας ωθεί να επιλέξουμε n οικογένειες από k : την οικογένεια k_1 που συσχετίζουμε με την $f^{(1)}$ (ή αλλιώς με την πλευρά S_1), την οικογένεια k_2 που συσχετίζουμε με την $f^{(2)}$ (ή αλλιώς με την πλευρά S_2) και γενικά την οικογένεια k_p που συσχετίζουμε με την $f^{(p)}$ (ή αλλιώς με την πλευρά S_p). Η κάθε οικογένεια μπορεί να αποτελείται από όσα k μας χρειάζονται.

Άρα η (2.21) γίνεται

$$\left\{ \sum_{j=1}^n |h_j| e^{i(\beta_j - k_p m_j)} \int_{-\pi}^{\pi} e^{-ik_p h_j s} (f^{(j)}(s) - i g^{(j)}(s)) ds = 0 \right\}_{p=1, \dots, n} \quad (2.22)$$

Διαιρώντας με τον συντελεστή του p -οστού όρου της (2.22) παίρνουμε

$$\left\{ \sum_{j=1}^n \frac{|h_j|}{|h_p|} \frac{e^{i(\beta_j - k_p m_j)}}{e^{i(\beta_p - k_p m_p)}} \int_{-\pi}^{\pi} e^{-ik_p h_j s} (f^{(j)}(s) - ig^{(j)}(s)) ds = 0 \right\}_{p=1, \dots, n} \Leftrightarrow$$

$$\left\{ \sum_{j=1}^n \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{ik_p(m_p - m_j)} \int_{-\pi}^{\pi} e^{-ik_p h_j s} (f^{(j)}(s) - ig^{(j)}(s)) ds = 0 \right\}_{p=1, \dots, n}. \quad (2.23)$$

Στην αριθμητική υλοποίηση της μεθόδου είναι πιθανόν να αντιμετωπίσουμε ένα σημαντικό πρόβλημα: **Η κατασκευή του τελικού αλγεβρικού συστήματος μπορεί να είναι δύσκολη ή ακόμα και αδύνατη εξαιτίας των εκθετικών παραγόντων.** Η δυσκολία στην κατασκευή έγκειται στο γεγονός ότι καθώς το k απομακρύνεται από την αρχή των αξόνων (αυξάνει το μέτρο του) οι εκθετικοί όροι μπορεί να ξεφεύγουν από το, πεπερασμένο, εύρος των αριθμών της μηχανής. Ας μην ξεχνάμε ότι το k μπορεί να είναι οποιοσδήποτε μιγαδικός αριθμός. Οι δυο "υπολογιστικά" επικίνδυνοι όροι της (2.23) είναι οι:

$$\int_{-\pi}^{\pi} e^{-ik_p h_j s} (f^{(p)}(s) - ig^{(p)}(s)) ds \quad (2.24)$$

και

$$e^{ik_p(m_p - m_j)}. \quad (2.25)$$

Έστω ότι

$$k_p = |k_p| e^{i\phi_p} \quad (2.26)$$

όπου $|k_p|$ και ϕ_p συμβολίζουν το μέτρο και το πρωτεύον όρισμα του k_p αντίστοιχα.

Ας εξετάσουμε τον κάθε όρο ξεχωριστά:

Ο όρος $\int_{-\pi}^{\pi} e^{-ik_p h_p s} (f^{(p)}(s) - ig^{(p)}(s)) ds$.

Είναι της μορφής $\int_{-\pi}^{\pi} e^{-ik_p h_j s} w(s) ds$ με $w(s)$ μια πραγματική συνάρτηση. Διαδοχικά έχουμε

$$\begin{aligned} \int_{-\pi}^{\pi} e^{-ik_p h_j s} w(s) ds &= \\ \int_{-\pi}^{\pi} e^{-i|k_p||h_j|e^{i\phi_p}e^{ia_j}s} w(s) ds &= \\ \int_{-\pi}^{\pi} e^{-i|k_p||h_j|e^{i(\phi_p+a_j)}s} w(s) ds &= \\ \int_{-\pi}^{\pi} e^{-i|k_p||h_j|\cos(\phi_p+a_j)s} e^{|k_p||h_j|\sin(\phi_p+a_j)s} w(s) ds \end{aligned}$$

Ο όρος $e^{-i|k_p||h_j|\cos(\phi_p+a_j)s}$ είναι φραγμένος και επομένως, για να είναι η προς ολοκλήρωση

ποσότητα φραγμένη, αρκεί το $\sin(\phi_p + \alpha_j)$ να είναι αρνητικό. Στην γενική περίπτωση δεν μπορούμε να υποθέσουμε ότι

$$\text{sign}(\sin(\phi_p + \alpha_j)) = \text{sign}(\sin(\phi_p + \alpha_t))$$

για κάθε ζευγάρι (j, t) και έτσι δεν είμαστε σε θέση να επιλέξουμε ένα τέτοιο ϕ_p που να μας εξασφαλίζει αρνητικό $\sin(\phi_p + \alpha_j)$ για κάθε j και επομένως όχι σχετικά μεγάλες τιμές για την ολοκληρώσιμη ποσότητα όταν $|k| \rightarrow \infty$. Δηλαδή στην γενική περίπτωση για κάθε k θα υπάρχει τουλάχιστον ένα j τέτοιο ώστε $e^{-ik_p h_j} \rightarrow \infty$ όταν $|k_p| \rightarrow \infty$.

Ο όρος $e^{ik_p(m_p - m_j)}$. Αναλύοντας αυτόν τον όρο διαδοχικά έχουμε:

$$\begin{aligned} e^{-ik_p(m_j - m_p)} &= \\ e^{-i|k_p||m_j - m_p|e^{i\phi_p}e^{i\arg(m_j - m_p)}} &= \\ e^{-i|k_p||m_j - m_p|e^{i(\phi_p + \arg(m_j - m_p))}} &= \\ e^{-i|k_p||m_j - m_p|\cos(\phi_p + \arg(m_j - m_p))}e^{i|k_p||m_j - m_p|\sin(\phi_p + \arg(m_j - m_p))} \end{aligned}$$

Άρα πρέπει να επιλέξουμε ένα ϕ_p τέτοιο ώστε $\sin(\phi_p + \arg(m_j - m_p)) < 0$. Το επόμενο λήμμα μας είναι απαραίτητο.

Λήμμα 2.3 Η ανισότητα $\sin(\phi_p + \arg(m_j - m_p)) < 0$ ισχύει αν και μόνο αν $\phi_p = \pi - \alpha_p$.

Απόδειξη. Δουλεύουμε σε ένα κυρτό πολύγωνο επομένως για κάθε ζεύγος p, j ισχύουν διαδοχικά οι σχέσεις

$$\begin{aligned} \alpha_p &< \arg(m_j - m_p) &< \alpha_p + \pi & \quad [\text{σχήμα (2.3)}] &\Leftrightarrow \\ 0 &< -\alpha_p + \arg(m_j - m_p) &< \pi & &\Leftrightarrow \\ \pi &< (\pi - \alpha_p) + \arg(m_j - m_p) &< 2\pi & &\Leftrightarrow \\ -1 &< \sin[(\pi - \alpha_p) + \arg(m_j - m_p)] &< 0 \end{aligned}$$

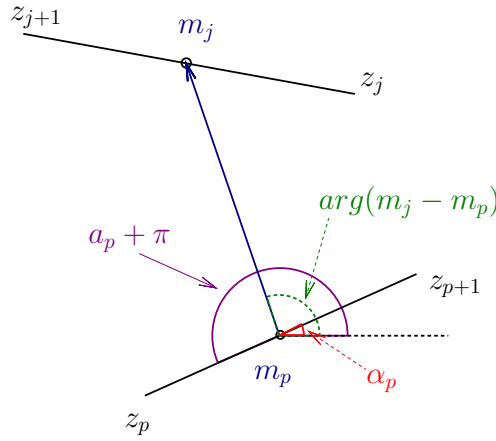
□

Επομένως η κατάλληλη επιλογή για το ϕ_p είναι

$$\phi_p = \pi - \alpha_p. \quad (2.27)$$

Η ξεχωριστή ανάλυση των δυο υπολογιστικά όρων της (2.23) μας δείχνει ότι ναι μεν για το ολοκλήρωμα (2.24) δεν υπάρχει κατάλληλη επιλογή για το k_p αλλά για τον όρο (2.24), ο οποίος πολλαπλασιάζει το ολοκλήρωμα, υπάρχει επιλογή που οδηγεί σε εκθετικούς όρους που δεν τείνουν στο ∞ όσο μεγάλο και να είναι το μέτρο του k_p . Αν δηλαδή επιλέξουμε

$$k_p = |k_p|e^{i(\pi - \alpha_p)}, \quad (2.28)$$



Σχήμα 2.3: Σχήμα στο οποίο αποδυνκνείται γραφικά η σχέση $\alpha_p < \arg(m_j - m_p) < \alpha_p + \pi$.

οι όροι της (2.23) θα είναι εν γένει της μορφής

$$(small\ value) * \int_{-\pi}^{\pi} (large\ value) =$$

$$\int_{-\pi}^{\pi} (small\ value) * (large\ value)$$

το οποίο, όπως αριθμητικά έχουμε παρατηρήσει, οδηγεί σε

$$\int_{-\pi}^{\pi} (normal\ value\ \acute{\eta}\ small\ value).$$

Επομένως κάνουμε την εξής επιλογή για τα k : Επιλέγουμε n οικογένειες από k . Η οικογένεια k_j αντιστοιχεί στην πλευρά S_j . Το κάθε k που ανήκει στην οικογένεια k_j βρίσκεται στην ημιευθεία που ξεκινάει από την αρχή των αξόνων και σχηματίζει γωνία $\pi - \alpha_j$ με τον θετικό ημιάξονα των πραγματικών αριθμών.

Με την παραπάνω επιλογή για τα k η Dirichlet-Neumann απεικόνιση (2.23) γράφεται

$$\left\{ \sum_{j=1}^n \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{i|k_p|e^{i(\pi - \alpha_p)}(m_p - m_j)} \int_{-\pi}^{\pi} e^{-i|k_p|e^{i(\pi - \alpha_p)}h_j s} (f^{(j)}(s) - ig^{(j)}(s)) ds = 0 \right\}_{p=1, \dots, n} \Leftrightarrow \quad (2.29)$$

$$\left\{ \sum_{j=1}^n \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-i|k_p|e^{-i\alpha_p}(m_p - m_j)} \int_{-\pi}^{\pi} e^{i|k_p|e^{-i\alpha_p}h_j s} (f^{(j)}(s) - ig^{(j)}(s)) ds = 0 \right\}_{p=1, \dots, n}. \quad (2.30)$$

Το μέτρο του κάθε k θα το υπολογίσουμε με την εξής λογική: Ας κάνουμε τον p -οστό όρο της

p -οστής εξίσωσης της (2.30) όσο πιο απλό μπορούμε. Μοιάζει κάπως αυθαίρετο αλλά δεν μας περιορίζει καθόλου όπως θα φανεί παρακάτω. Ο παραπάνω όρος είναι ο

$$\int_{-\pi}^{\pi} e^{i|k_p|e^{-i\alpha_p}h_p s} (f^{(j)}(s) - ig^{(j)}(s)) ds.$$

Εμείς μπορούμε να επηρεάσουμε μόνο τον εκθέτη

$$i|k_p|e^{-i\alpha_p}h_p s = i|k_p|e^{-i\alpha_p}|h_p|e^{i\alpha_p}s = i|k_p||h_p|s.$$

Η προφανής επιλογή για το μέτρο του k_p είναι λοιπόν $|k_p| = \frac{l}{|h_p|}$, όπου $l \in \mathbb{R}^+$. Δηλαδή η τελική επιλογή για τα k_p θα είναι

$$k_p = \frac{-l}{h_p}. \quad (2.31)$$

Όπως προείπαμε η επιλογή για το μέτρο του κάθε k δεν είναι περιοριστική καθώς το μέτρο μπορεί να είναι οποιοσδήποτε θετικός αριθμός.

Πλέον η Dirichlet-Neumann απεικόνιση έχει πάρει την μορφή

$$\left\{ \sum_{j=1}^n \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-il \frac{m_p - m_j}{h_p}} \int_{-\pi}^{\pi} e^{il \frac{h_j}{h_p} s} (f^{(j)}(s) - ig^{(j)}(s)) ds = 0 \right\}_{p=1, \dots, n}. \quad (2.32)$$

Παρατήρηση 2.2 Γιατί η αριθμητική μέθοδος πιθανόν να έχει πρόβλημα όταν πρέπει να υπολογίσουμε ποσότητες της μορφής e^x όταν ο x είναι σχετικά μεγάλος αριθμός; Με την υπάρχουσα τεχνολογία ο κάθε υπολογιστής μπορεί να αναπαραστήσει και να εκτελέσει αριθμητικές πράξεις με ένα πεπερασμένο υποσύνολο Ω του $\mathbb{C}$ ¹. Συναρτήσεις τώρα όπως η e^x για να υπολογιστούν σωστά πρέπει $e^x \in \Omega$ γεγονός το οποίο συνεπάγεται ότι το x πρέπει να ανήκει σε ένα ακόμα μικρότερο υποσύνολο του $\mathbb{C}$. Π.χ. το 100.000 είναι ένας σχετικά μεγάλος αριθμός τον οποίο όλοι οι σύγχρονοι H/Y μπορούν να "αντιμετωπίσουν". Αν προσπαθήσουμε όμως να υπολογίσουμε τον $e^{100.000}$, μάλλον δεν θα τα καταφέρουμε ή αν ακόμα και αν τα καταφέρουμε τότε θα υπάρχει πρόβλημα με τον $e^{e^{100.000}}$. Δηλαδή στην σύγχρονη αριθμητική αντιμετώπιση των διαφόρων προβλημάτων είμαστε αναγκασμένοι να δουλεύουμε με ένα πεπερασμένο υποσύνολο των αριθμών που γνωρίζουμε. Αυτός είναι και ο λόγος που μας αναγκάζει να επιλέξουμε τις τιμές του k που θα χρησιμοποιήσουμε. Θέματα σχετικά με την αριθμητική των υπολογιστών ή αλλιώς των αριθμών κινητής υποδιαστολής (floating point arithmetic) πραγματεύονται στο [SUN94].

Παρατήρηση 2.3 Ένα απλό κίνητρο [FFX03, SFFS] για την επιλογή του k_p όπως στην (2.31) είναι το εξής: Υποθέστε ότι υπάρχει μόνο μία άγνωστη συνάρτηση η $f^{(p)}(s)$. Τότε προφανώς θα

¹ Υπάρχουν διάφορα συμβολικά πακέτα υπολογισμών (π.χ. Maple, Mathematica, ...) τα οποία είναι ικανά να εκτελέσουν υπολογισμούς σε όλο το $\mathbb{C}$ όμως δεν χρησιμοποιούνται για την αριθμητική επίλυση προβλημάτων λόγω της μειωμένης απόδοσής τους.

επιλέγαμε το k με τέτοιο τρόπο ώστε το kh_p να είναι πραγματικός, δηλαδή $kh_p = l$ ή $kh_p = -l$. Η τελευταία επιλογή είναι συμβατή με την αναλυτική αναπαράσταση της λύσης [FOK01].

2.2.3 Διακριτοποίηση

Οι άγνωστες συναρτήσεις $f^{(j)}$ ορίζονται στο διάστημα $[-\pi, \pi]$. Οι τιμές τους στα άκρα αυτού του διαστήματος μπορούν να υπολογιστούν, αναλυτικά, από τις συνθήκες συνέχειας που ικανοποιεί η q_z στις κορυφές του πολυγώνου. Πράγματι, θέτοντας $\delta_j = \alpha_j - \beta_j$ και υποθέτοντας ότι $\delta_j \neq \delta_{j-1}$ [FFX03] μπορούμε να αποδείξουμε το παρακάτω λήμμα.

Λήμμα 2.4 Αν $\delta_j \neq \delta_{j-1}$ τότε ισχύει ότι

$$\begin{cases} f^{(j)}(\pi) = \frac{\cos(\delta_{j+1} - \delta_j) g^{(j)}(\pi) - g^{(j+1)}(-\pi)}{\sin(\delta_{j+1} - \delta_j)} \\ f^{(j)}(-\pi) = \frac{g^{(j-1)}(\pi) - \cos(\delta_j - \delta_{j-1}) g^{(j)}(-\pi)}{\sin(\delta_j - \delta_{j-1})} \end{cases} \quad (2.33)$$

Απόδειξη. Σε κάθε κορυφή z_j οι $q_z^{(j-1)}(z_j)$ και $q_z^{(j)}(z_j)$ αναπαριστούν την ίδια ποσότητα επομένως είναι ίσες.

$$q_z^{(j-1)}(z_j) = q_z^{(j)}(z_j) \quad (2.34)$$

Διαδοχικά έχουμε

$$\begin{aligned} \frac{1}{2} e^{-i(\alpha_{j-1} - \beta_{j-1})} (g^{(j-1)}(z_j) + i f^{(j-1)}(z_j)) &= \frac{1}{2} e^{-i(\alpha_j - \beta_j)} (g^{(j)}(z_j) + i f^{(j)}(z_j)) && \Leftrightarrow \\ e^{-i\delta_{j-1}} (g^{(j-1)}(z_j) + i f^{(j-1)}(z_j)) &= e^{-i\delta_j} (g^{(j)}(z_j) + i f^{(j)}(z_j)) && \Leftrightarrow \\ e^{-i\delta_{j-1}} (g^{(j-1)}(\pi) + i f^{(j-1)}(\pi)) &= e^{-i\delta_j} (g^{(j)}(-\pi) + i f^{(j)}(-\pi)) && \Leftrightarrow \\ (\cos \delta_{j-1} - i \sin \delta_{j-1}) (g^{(j-1)}(\pi) + i f^{(j-1)}(\pi)) &= (\cos \delta_j - i \sin \delta_j) (g^{(j)}(-\pi) + i f^{(j)}(-\pi)) && \Leftrightarrow \end{aligned}$$

$$\begin{aligned} \cos \delta_{j-1} g^{(j-1)}(\pi) + i \cos \delta_{j-1} f^{(j-1)}(\pi) - i \sin \delta_{j-1} g^{(j-1)}(\pi) + \sin \delta_{j-1} f^{(j-1)}(\pi) &= \\ \cos \delta_j g^{(j)}(-\pi) + i \cos \delta_j f^{(j)}(-\pi) - i \sin \delta_j g^{(j)}(-\pi) + \sin \delta_j f^{(j)}(-\pi) && \Leftrightarrow \end{aligned}$$

$$\begin{aligned} (\cos \delta_{j-1} g^{(j-1)}(\pi) + \sin \delta_{j-1} f^{(j-1)}(\pi)) + i (\cos \delta_{j-1} f^{(j-1)}(\pi) - \sin \delta_{j-1} g^{(j-1)}(\pi)) &= \\ (\cos \delta_j g^{(j)}(-\pi) + \sin \delta_j f^{(j)}(-\pi)) + i (\cos \delta_j f^{(j)}(-\pi) - \sin \delta_j g^{(j)}(-\pi)) && \Leftrightarrow \end{aligned}$$

$$\begin{cases} \cos \delta_{j-1} g^{(j-1)}(\pi) + \sin \delta_{j-1} f^{(j-1)}(\pi) = \cos \delta_j g^{(j)}(-\pi) + \sin \delta_j f^{(j)}(-\pi) \\ \cos \delta_{j-1} f^{(j-1)}(\pi) - \sin \delta_{j-1} g^{(j-1)}(\pi) = \cos \delta_j f^{(j)}(-\pi) - \sin \delta_j g^{(j)}(-\pi) \end{cases} \quad \Leftrightarrow$$

$$\begin{cases} \sin \delta_{j-1} f^{(j-1)}(\pi) - \sin \delta_j f^{(j)}(-\pi) = -\cos \delta_{j-1} g^{(j-1)}(\pi) + \cos \delta_j g^{(j)}(-\pi) \\ \cos \delta_{j-1} f^{(j-1)}(\pi) - \cos \delta_j f^{(j)}(-\pi) = \sin \delta_{j-1} g^{(j-1)}(\pi) - \sin \delta_j g^{(j)}(-\pi) \end{cases}.$$

Ισοδύναμα, σε μορφή πινάκων

$$\begin{aligned} \begin{bmatrix} \sin \delta_{j-1} & -\sin \delta_j \\ \cos \delta_{j-1} & -\cos \delta_j \end{bmatrix} \begin{bmatrix} f^{(j-1)}(\pi) \\ f^{(j)}(-\pi) \end{bmatrix} &= \begin{bmatrix} -\cos \delta_{j-1} & \cos \delta_j \\ \sin \delta_{j-1} & -\sin \delta_j \end{bmatrix} \begin{bmatrix} g^{(j-1)}(\pi) \\ g^{(j)}(-\pi) \end{bmatrix} \Leftrightarrow \\ \begin{bmatrix} f^{(j-1)}(\pi) \\ f^{(j)}(-\pi) \end{bmatrix} &= \begin{bmatrix} \sin \delta_{j-1} & -\sin \delta_j \\ \cos \delta_{j-1} & -\cos \delta_j \end{bmatrix}^{-1} \begin{bmatrix} -\cos \delta_{j-1} & \cos \delta_j \\ \sin \delta_{j-1} & -\sin \delta_j \end{bmatrix} \begin{bmatrix} g^{(j-1)}(\pi) \\ g^{(j)}(-\pi) \end{bmatrix} \Leftrightarrow \\ \begin{bmatrix} f^{(j-1)}(\pi) \\ f^{(j)}(-\pi) \end{bmatrix} &= \frac{1}{\sin(\delta_j - \delta_{j-1})} \begin{bmatrix} \cos(\delta_j - \delta_{j-1}) & -1 \\ 1 & -\cos(\delta_j - \delta_{j-1}) \end{bmatrix} \begin{bmatrix} g^{(j-1)}(\pi) \\ g^{(j)}(-\pi) \end{bmatrix} \Leftrightarrow \\ \begin{cases} f^{(j-1)}(\pi) = \frac{\cos(\delta_j - \delta_{j-1}) g^{(j-1)}(\pi) - g^{(j)}(-\pi)}{\sin(\delta_j - \delta_{j-1})} \\ f^{(j)}(-\pi) = \frac{g^{(j-1)}(\pi) - \cos(\delta_j - \delta_{j-1}) g^{(j)}(-\pi)}{\sin(\delta_j - \delta_{j-1})} \end{cases} \end{aligned} \quad (2.35)$$

Επειδή η παραπάνω σχέση ισχύει για κάθε j έχουμε

$$\begin{cases} f^{(j)}(\pi) = \frac{\cos(\delta_{j+1} - \delta_j) g^{(j)}(\pi) - g^{(j+1)}(-\pi)}{\sin(\delta_{j+1} - \delta_j)} \\ f^{(j)}(-\pi) = \frac{g^{(j-1)}(\pi) - \cos(\delta_j - \delta_{j-1}) g^{(j)}(-\pi)}{\sin(\delta_j - \delta_{j-1})} \end{cases} \quad (2.36)$$

□

Παρατήρηση 2.4 Παρατηρήστε ότι $g^{(j-1)}(\pi) \neq g^{(j)}(-\pi)$ το οποίο παραβιάζεται μόνο αν $\alpha_{j-1} - \beta_{j-1} = \alpha_j - \beta_j \Leftrightarrow \delta_{j-1} = \delta_j$. Ας θυμηθούμε ότι η $g^{(j)}$ είναι η παράγωγος της q_z σε κάποια συγκεκριμένη κατεύθυνση ως προς την πλευρά S_j και όχι ως προς τον άξονα των πραγματικών. Φυσικά ισχύει και το αντίστοιχο $f^{(j-1)}(\pi) \neq f^{(j)}(-\pi)$.

Για την ανάπτυξη αριθμητικής μεθόδου προσέγγισης των άγνωστων συναρτήσεων $f^{(j)}$ θεωρούμε ότι $f^{(j)} \approx f_{N_j}^{(j)}$ όπου

$$f_{N_j}^{(j)}(s) = f_*^{(j)}(s) + \sum_{r=1}^{N_j} U_r^j \varphi_r^{(j)}(s), \quad (2.37)$$

με $f_*^{(j)}(s)$ γραμμική συνάρτηση (πολυώνυμο 1ου βαθμού) που ικανοποιεί

$$\begin{cases} f_*^{(j)}(-\pi) = f^{(j)}(-\pi) \\ f_*^{(j)}(\pi) = f^{(j)}(\pi) \end{cases},$$

και $\{\varphi_r^{(j)}(s)\}_{r=1}^{N_j}$ μια βάση συναρτήσεων ενός υποχώρου

$$\mathbb{S} \text{ του } C_0[-\pi, \pi] \text{ διάστασης } N_j, \quad (2.38)$$

όπου $C_0[-\pi, \pi]$ είναι το σύνολο των, πραγματικών, συνεχών συναρτήσεων στο διάστημα $[-\pi, \pi]$.

Τα U_r^j είναι πραγματικοί αριθμοί και είναι οι άγνωστοι μας πλέον. Αναλυτικός τύπος για την $f_*^{(j)}(s)$ υπολογίζεται από τις σχέσεις

$$\begin{cases} f_*^{(j)}(s) = As + B \\ f_*^{(j)}(-\pi) = f^{(j)}(-\pi) \\ f_*^{(j)}(\pi) = f^{(j)}(\pi) \end{cases},$$

οπότε διαδοχικά έχουμε

$$\begin{aligned} \begin{cases} f_*^{(j)}(-\pi) = f^{(j)}(-\pi) \\ f_*^{(j)}(\pi) = f^{(j)}(\pi) \end{cases} &\Leftrightarrow \begin{cases} -A\pi + B = f^{(j)}(-\pi) \\ A\pi + B = f^{(j)}(\pi) \end{cases} \Leftrightarrow \\ \begin{cases} -A\pi + B = f^{(j)}(-\pi) \\ 2B = f^{(j)}(-\pi) + f^{(j)}(\pi) \end{cases} &\Leftrightarrow \begin{cases} A = -\frac{f^{(j)}(-\pi)}{\pi} + \frac{f^{(j)}(-\pi) + f^{(j)}(\pi)}{2\pi} \\ B = \frac{f^{(j)}(-\pi) + f^{(j)}(\pi)}{2} \end{cases} \Leftrightarrow \\ \begin{cases} A = \frac{f^{(j)}(\pi) - f^{(j)}(-\pi)}{2\pi} \\ B = \frac{f^{(j)}(-\pi) + f^{(j)}(\pi)}{2} \end{cases}. \end{aligned}$$

Άρα

$$f_*^{(j)}(s) = \frac{f^{(j)}(\pi) - f^{(j)}(-\pi)}{2\pi}s + \frac{f^{(j)}(-\pi) + f^{(j)}(\pi)}{2},$$

ή

$$f_*^{(j)}(s) = \frac{1}{2\pi} [(s + \pi) f^{(j)}(\pi) - (s - \pi) f^{(j)}(-\pi)]. \quad (2.39)$$

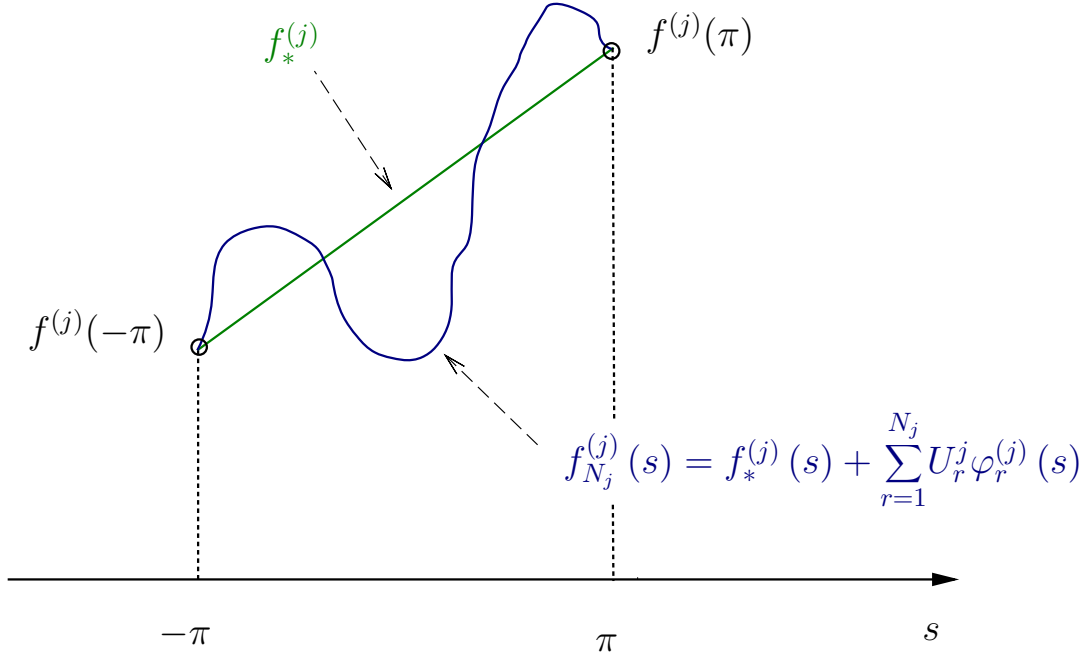
Ας σημειωθεί ότι για οι συναρτήσεις βάσης πρέπει να μηδενίζονται στα άκρα του διαστήματος $[-\pi, \pi]$, δηλαδή

$$\phi_r^{(j)}(-\pi) = \phi_r^{(j)}(\pi) = 0, \quad (2.40)$$

αφού μόνο έτσι θα ισχύει η σχέση

$$f_{N_j}^{(j)}(s) = f_*^{(j)}(s).$$

Η παραπάνω προσέγγιση της $f^{(j)}(s)$ (2.37) οπτικοποιείται στο σχήμα (2.2.3).



Σχήμα 2.4: Προσέγγιση της $f^{(j)}(s)$ από μια γραμμική συνάρτηση $f_*^{(j)}(s)$ και ένα γραμμικό συνδυασμό συναρτήσεων βάσης.

Η ενσωμάτωση της (2.37) στην (2.32) οδηγεί στην παρακάτω, προσεγγιστική πλέον, μορφή για την Dirichlet-Neumann απεικόνιση .

$$\left\{ \sum_{j=1}^n \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-il \frac{m_p - m_j}{h_p}} \int_{-\pi}^{\pi} e^{il \frac{h_j}{h_p} s} \left[\left(f_*^{(j)}(s) + \sum_{r=1}^{N_j} U_r^j \varphi_r^{(j)}(s) \right) - i g^{(j)}(s) \right] ds = 0 \right\}_{p=1, \dots, n} \quad (2.41)$$

Χωρίζοντας γνωστούς από αγνώστους έχουμε

$$\left\{ \begin{aligned} & \sum_{j=1}^n \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-il \frac{m_p - m_j}{h_p}} \int_{-\pi}^{\pi} e^{il \frac{h_j}{h_p} s} \sum_{r=1}^{N_j} U_r^j \varphi_r^{(j)}(s) ds = \\ & \sum_{j=1}^n \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-il \frac{m_p - m_j}{h_p}} \int_{-\pi}^{\pi} e^{il \frac{h_j}{h_p} s} \left[-f_*^{(j)}(s) + i g^{(j)}(s) \right] ds \end{aligned} \right\}_{p=1, \dots, n}$$

και ανακατατάσσοντας τους όρους καταλήγουμε στην

$$\left\{ \sum_{j=1}^n \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-il \frac{m_p - m_j}{h_p}} \sum_{r=1}^{N_j} U_r^j \int_{-\pi}^{\pi} e^{il \frac{h_j}{h_p} s} \varphi_r^{(j)}(s) ds = G_p(l) \right\}_{p=1, \dots, n}, \quad (2.42)$$

όπου

$$G_p(l) = i \sum_{j=1}^n \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-il \frac{m_p - m_j}{h_p}} \int_{-\pi}^{\pi} e^{il \frac{h_j}{h_p} s} (g^{(j)}(s) + i f_*^{(j)}(s)) ds. \quad (2.43)$$

Το πρόβλημα μας πλέον έχει γίνει η εύρεση των άγνωστων συντελεστών U_r^j . Σε κάθε συνάρτηση $f^{(j)}$ αντιστοιχούν N_j άγνωστοι. Το συνολικό πλήθος των αγνώστων είναι

$$N_1 + N_2 + \dots + N_n = \sum_{r=1}^n N_j. \quad (2.44)$$

2.3 Σχόλια

- (α) Η μέθοδος μπορεί να θεωρηθεί ότι ανήκει στην κατηγορία των collocation μεθόδων υπό την έννοια ότι έχουμε μια συνθήκη (την Dirichlet-Neumann απεικόνιση) που περιέχει τις άγνωστες ποσότητες (τους U_r^j) η οποία απαιτούμε να ισχύει σε συγκεκριμένα collocation points (τα k). Πάντως τα collocation points δεν είναι σημεία στα οποία ικανοποιείται η διαφορική εξίσωση αλλά εκτείνονται σε όλο το $\mathbb{C}$ και έχουν σχέση με μία αυθαίρετη παράμετρο k , ξένη προς την διαφορική εξίσωση. Έτσι θα αναφέρουμε το τελικό γραμμικό σύστημα που θα προκύψει από την διακριτοποίηση της Dirichlet-Neumann απεικόνισης ως collocation σύστημα και τον πίνακα συντελεστών των αγνώστων ως collocation πίνακα.
- (β) Η (2.42) παράγει μιγαδικές εξισώσεις ενώ οι άγνωστοι μας είναι πραγματικοί αριθμοί. Επομένως έχουμε την δυνατότητα να επιλέξουμε $N/2$ το πλήθος διαφορετικά k παίρνοντας $N/2$ μιγαδικές εξισώσεις από τις οποίες μπορούμε να εξαγάγουμε N πραγματικές εξισώσεις.
- (γ) Μπορούμε να πούμε ότι η μέθοδος έχει δυο βαθμούς ελευθερίας. Έχουμε ευχέρεια και στην επιλογή των συναρτήσεων βάσης $\varphi_r^{(j)}$ και στην επιλογή των k (του μέτρου των k στην πραγματικότητα εφόσον επιλέγουμε k της μορφής (2.31)). Μάλιστα λόγω της δυνατότητας που έχουμε από κάθε μιγαδική εξίσωση να παράξουμε δυο πραγματικές εξισώσεις μπορούμε με τα ίδια collocation points να παράγουμε διαφορετικά σύνολα πραγματικών εξισώσεων.

Κεφάλαιο 3

Η μέθοδος SFFS

Όπως προαναφέραμε το αριθμητικό ανάλογο της μεθόδου του καθηγητή Φωκά έχει δυο βαθμούς ελευθερίας. Διαφορετικές επιλογές συναρτήσεων βάσης καθώς και διαφορετικές επιλογές των collocation points οδηγούν σε διαφορετικές παραλλαγές της μεθόδου. Στην βιβλιογραφία υπάρχει μια μόνο πρόταση για την επιλογή των collocation points. Είναι αυτή των Fulton, Φωκά, Ξενοφώντος στο [FFX03] και θα αναφερόμαστε στην παραγόμενη μέθοδο με την συντομογραφία FFX. Στην παρούσα εργασία προτείνουμε μια διαφορετική επιλογή για τα collocation points [SFFS]. Θα αναφερόμαστε σε αυτή την νέα επιλογή με την συντομογραφία SFFS. Θα αποδείξουμε ότι η νέα επιλογή για τα collocation points οδηγεί σε ένα γραμμικό σύστημα με πίνακα συντελεστών των αγνώστων που έχει point-diagonal διαγώνια μπλοκ.

3.1 Αρίθμηση των Εξισώσεων/Αγνώστων

Η εξίσωση (2.42) είναι γεννήτρια εξισώσεων. Έχουμε θεωρήσει n οικογένειες από k τις $\{k_j\}_{j=1,\dots,n}$. Η οικογένεια k_j παράγει εξισώσεις που αντιστοιχούν στην j πλευρά. Το πλήθος των στοιχείων κάθε οικογένειας μπορεί να κυμαίνεται μεταξύ $N_j/2$ και N_j . Το κάθε στοιχείο της κάθε οικογένειας χαρακτηρίζεται από έναν θετικό πραγματικό αριθμό τον l_t και είναι το $k_j^{(t)} = \frac{-l_t}{h_j}$. Θεωρούμε ότι $l_{t_i} < l_{t_j}$ αν $t_i < t_j$ και χρησιμοποιούμε τον παρακάτω αλγόριθμο για την παραγωγή των μιγαδικών εξισώσεων.

```
for  $p = 1, 2, \dots, n$ 
    for  $t = t_1, t_2, \dots$ 
        χρησιμοποίησε το  $k_p^{l_t}$  στην (2.42)
    end
end
```

Με αυτόν τον τρόπο παράγουμε p οικογένειες από μιγαδικές εξισώσεις. Η αρίθμηση των αγνώ-

στων σε κάθε μιγαδική εξίσωση είναι η ακόλουθη:

$$U_1^1, U_1^2, \dots, U_1^{N_1}, \quad U_2^1, U_2^2, \dots, U_2^{N_2}, \quad \dots, \quad U_j^1, U_j^2, \dots, U_j^{N_j}, \quad \dots, \quad U_n^1, U_n^2, \dots, U_n^{N_n}. \quad (3.1)$$

Με τις παραπάνω αριθμήσεις, των εξισώσεων και των αγνώστων, παράγεται ένα γραμμικό σύστημα του οποίου ο πίνακας, των μιγαδικών, συντελεστών έχει μια φυσική μπλοκ μορφή. Το (p, j) μπλοκ του, περιέχει τις εξισώσεις που αντιστοιχούν στην p πλευρά και εμπλέκει τους αγνώστους που αντιστοιχούν στην j πλευρά.

Οι μιγαδικές εξισώσεις που παράγονται από την (2.42) έχουν την γενική μορφή

$$\sum_{j=1}^N (a_j + ib_j) x_j = (c_j + id_j), \quad (3.2)$$

όπου οι $a, b, c, d \in \mathbb{R}$ είναι οι γνωστοί συντελεστές και $x \in \mathbb{R}$ είναι οι άγνωστοι. Από την (3.2) μπορούμε να εξαγάγουμε δυο πραγματικές εξισώσεις, τις

$$\sum_{j=1}^N a_j x_j = c_j \quad (3.3\alpha')$$

και

$$\sum_{j=1}^N b_j x_j = d_j. \quad (3.3\beta')$$

Στην παρούσα εργασία, χάριν απλότητας, θεωρούμε το ίδιο σύνολο συναρτήσεων βάσης και το ίδιο πλήθος τους, N_j , σε κάθε πλευρά¹. Θα συμβολίζουμε με N αυτό το πλήθος των συναρτήσεων βάσης σε κάθε πλευρά. Σε αυτό το κεφάλαιο ως συναρτήσεις βάσης θα χρησιμοποιήσουμε ημιτονοειδείς, την φυσική επιλογή, συναρτήσεις και θα ονομάσουμε την συγκεκριμένη βάση sine βάση.

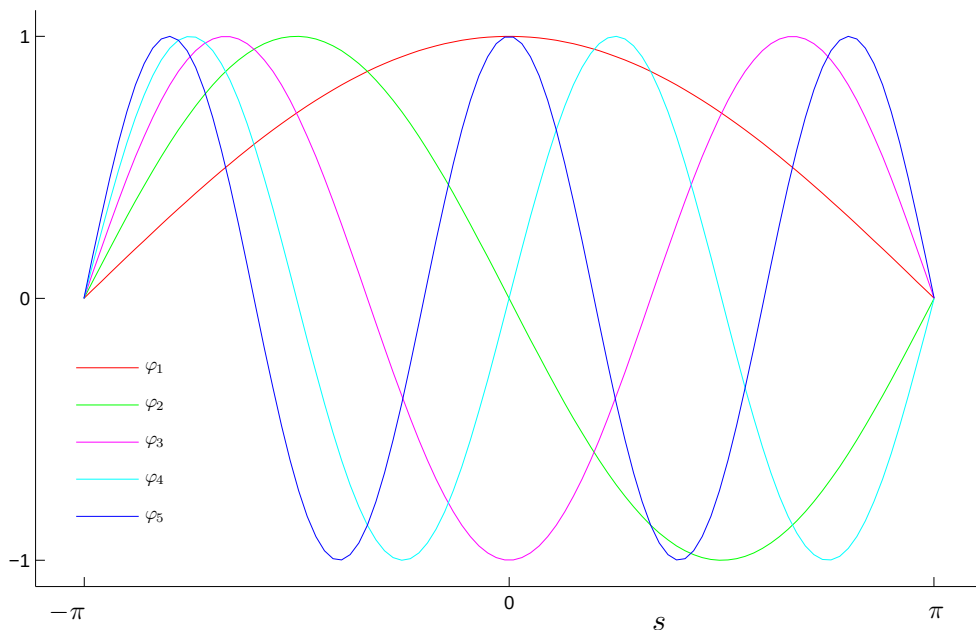
3.2 Ημιτονοειδής Βάση

Ως συναρτήσεις βάσης χρησιμοποιούνται οι ημιτονοειδείς συναρτήσεις

$$\varphi_r(s) = \sin\left(r \frac{s + \pi}{2}\right). \quad (3.4)$$

Παρατηρούμε ότι οι συναρτήσεις μηδενίζονται στα σημεία $-\pi, \pi$, όπως άλλωστε απαιτεί ο περιορισμός (2.40). Το σχήμα (3.1) δείχνει τις 5 αρχικές ($r = 1, 2, 3, 4, 5$) συναρτήσεις βάσης.

¹ Εν γένει δεν έχουμε κάποιο ιδιαίτερο λόγο να επιλέξουμε διαφορετικές συναρτήσεις βάσης σε κάθε πλευρά. Μια adaptive παραλλαγή της μεθόδου θα μπορούσε κάλλιστα να κάνει χρήση της δυνατότητας διαφορετικών συναρτήσεων βάσης ή/και διαφορετικής διάστασης του υποχώρου $\mathbb{S}$ (2.38).



Σχήμα 3.1: Οι ημιτονοειδείς συναρτήσεις βάσης $\varphi_r(s) = \sin\left(r \frac{s + \pi}{2}\right)$. Παρατηρήστε ότι όσο αυξάνει το r τόσο μειώνεται η περίοδος των συναρτήσεων.

3.3 Η μέθοδος FFX

Η επιλογή των collocation points και των εξισώσεων.

Η μέθοδος FFX προκύπτει από την συγκεκριμένη επιλογή της παραμέτρου l στον τύπο (2.31). Για κάθε πλευρά S_j επιλέγονται τα παρακάτω l

$$l = 1, 2, 3, \dots, N/2. \quad (3.5)$$

Στην συνέχεια από κάθε μιγαδική εξίσωση της (2.42) θα πάρουμε σαν μια πραγματική εξίσωση το πραγματικό μέρος και σαν μια άλλη πραγματική εξίσωση το φανταστικό μέρος.

Με τις παραπάνω επιλογές η σχέση (2.42) γίνεται

$$\left\{ \left\{ \sum_{j=1}^n \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-il \frac{m_p - m_j}{h_p}} \sum_{r=1}^{N_j} U_r^j \int_{-\pi}^{\pi} e^{il \frac{h_j}{h_p} s} \sin\left(r \frac{s + \pi}{2}\right) ds = G_p(l) \right\}_{l=1,2,\dots,\frac{N_j}{2}} \right\}_{p=1,\dots,n} \quad (3.6)$$

όπου $G_p(l)$ ορίζονται στην (2.43).

Η σχέση (3.6) αποτελείται από $(N * n)/2$ μιγαδικές εξισώσεις από τις οποίες μπορούμε να κατασκευάσουμε ένα γραμμικό σύστημα με $n * N$ πραγματικές εξισώσεις και $n * N$ αγνώστους U_r^j .

Για κάθε p και j , με $p, j = 1, 2, \dots, N$

με C_p^j συμβολίζουμε τον μιγαδικό πίνακα διάστασης $\frac{N}{2} \times N$ με στοιχεία

$$c_p^j(l, r) := \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-il \frac{m_p - m_j}{h_p}} \int_{-\pi}^{\pi} e^{il \frac{h_j}{h_p} s} \sin\left(r \frac{s + \pi}{2}\right) ds, \quad l = 1, 2, \dots, N/2, \quad r = 1, 2, \dots, N, \quad (3.7)$$

με X_p το πραγματικό διάνυσμα διάστασης N με στοιχεία

$$x_p(r) := U_r^p, \quad r = 1, 2, \dots, N \quad (3.8)$$

και με T_p το μιγαδικό διάνυσμα διάστασης $\frac{N}{2}$ με στοιχεία

$$t_p(l) = i \sum_{j=1}^n \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-il \frac{m_p - m_j}{h_p}} \int_{-\pi}^{\pi} e^{il \frac{h_j}{h_p} s} (g^{(j)}(s) + i f_*^{(j)}(s)) ds, \quad l = 1, 2, \dots, N/2. \quad (3.9)$$

Εάν οι πίνακες C, X, T είναι τέτοιοι ώστε

$$C = \begin{bmatrix} C_1^1 & & C_1^n \\ & \ddots & \\ C_n^1 & & C_n^n \end{bmatrix} \in \mathbb{R}^{nN/2 \times nN}, \quad X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} \in \mathbb{R}^{nN \times 1}, \quad T = \begin{bmatrix} T_1 \\ \vdots \\ T_n \end{bmatrix} \in \mathbb{R}^{nN/2 \times 1}, \quad (3.10)$$

τότε η (3.6) είναι ισοδύναμη με το σύστημα

$$CX = T. \quad (3.11)$$

Το σύστημα (3.11) μπορεί να μετασχηματιστεί σε ένα αμιγώς πραγματικό γραμμικό σύστημα

$$AX = B \quad (3.12)$$

όπου οι πραγματικοί πίνακες A, B διάστασης $nN \times nN$ και $nN \times 1$, αντίστοιχα ορίζονται ως εξής

$$A(row, :) := \begin{cases} \text{Re} \left[C \left(\frac{row+1}{2}, : \right) \right] & , \text{ αν } row : \text{ περιττός} \\ \text{Im} \left[C \left(\frac{row}{2}, : \right) \right] & , \text{ αν } row : \text{ άρτιος} \end{cases} \quad (3.13)$$

και¹

$$B(row) := \begin{cases} \text{Re} \left[T \left(\frac{row+1}{2} \right) \right] & , \text{ αν } row : \text{ περιττός} \\ \text{Im} \left[T \left(\frac{row}{2} \right) \right] & , \text{ αν } row : \text{ άρτιος} \end{cases} \quad (3.14)$$

όπου με $\text{Re}(v)$ και $\text{Im}(v)$ συμβολίζουμε το πραγματικό μέρος και το μιγαδικό μέρος του διανύσματος v αντίστοιχα. Επομένως, εάν A_p^j συμβολίζει τον πραγματικό πίνακα $N \times N$ με στοιχεία

$$a_p^j(l, r) = \begin{cases} \text{Re} [c_p^j(l, r)] \\ \text{Im} [c_p^j(l, r)] \end{cases}$$

¹ Ο συμβολισμός $A(i, :)$ δηλώνει την i -οστή γραμμή του πίνακα A και είναι δανεισμένος από το Matlab.

και B_p συμβολίζει τον πραγματικό πίνακα $N \times 1$ με στοιχεία

$$b_p(l) = \begin{cases} \operatorname{Re}[t_p(l)] \\ \operatorname{Im}[t_p(l)] \end{cases},$$

τότε

$$A = \begin{bmatrix} A_1^1 & & A_1^n \\ & \ddots & \\ A_n^1 & & A_n^n \end{bmatrix},$$

και

$$B = \begin{bmatrix} B_1 \\ \vdots \\ B_n \end{bmatrix}.$$

Παρατηρείστε ότι όλα τα διαγώνια μπλοκ $A_1^1, \dots, A_n^n$ είναι ίσα μεταξύ τους αφού $c_p^r(l, r) := \int_{-\pi}^{\pi} e^{ils} \sin\left(r \frac{s + \pi}{2}\right)$, ανεξάρτητο του p για κάθε $l = 1, 2, \dots, N/2$, $r = 1, 2, \dots, N$.

Ορισμός 3.1 (FFX) Η αριθμητική μέθοδος που προκύπτει από την διακριτοποίηση της *Dirichlet-Neumann* απεικόνισης με την παρακάτω επιλογή *collocation points*

$$k_p := \left\{ \frac{-l}{h_p} \right\}_{l=1, \dots, N/2} = \left\{ \frac{-1}{h_p}, \frac{-2}{h_p}, \frac{-3}{h_p}, \dots, \frac{-N/2}{h_p} \right\} \quad (3.15)$$

και επιλογή εξισώσεων: το πραγματικό και το φανταστικό μέρος της κάθε μιας από τις εξισώσεις της σχέσης (2.42), για $l = 1, \dots, N/2$ ονομάζεται FFX.

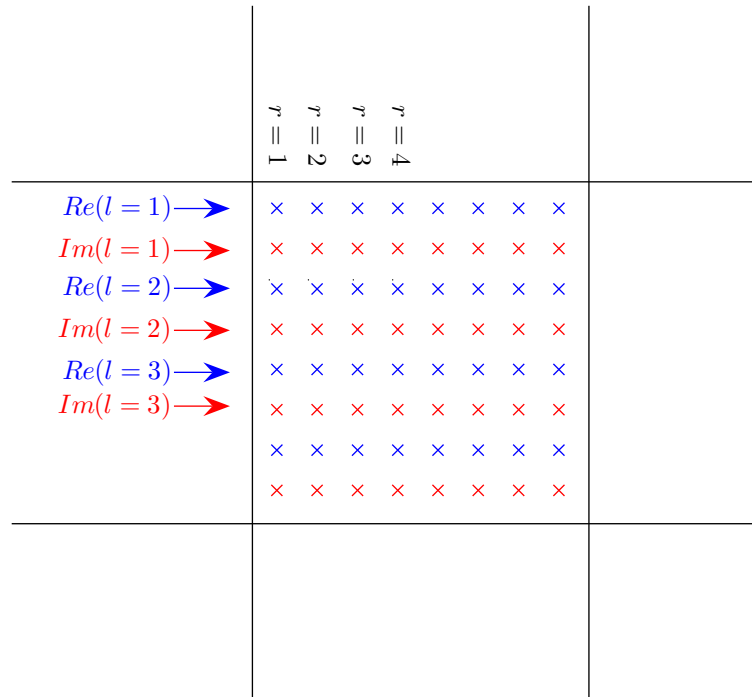
Το σχήμα (3.2) μας βοηθάει να καταλάβουμε τις παραπάνω επιλογές. Αναφέρεται σε κάθε τυχαίο μπλόκ διάστασης $(N \times N)$ με $N = 8$ του A .

3.4 Η μέθοδος SFFS

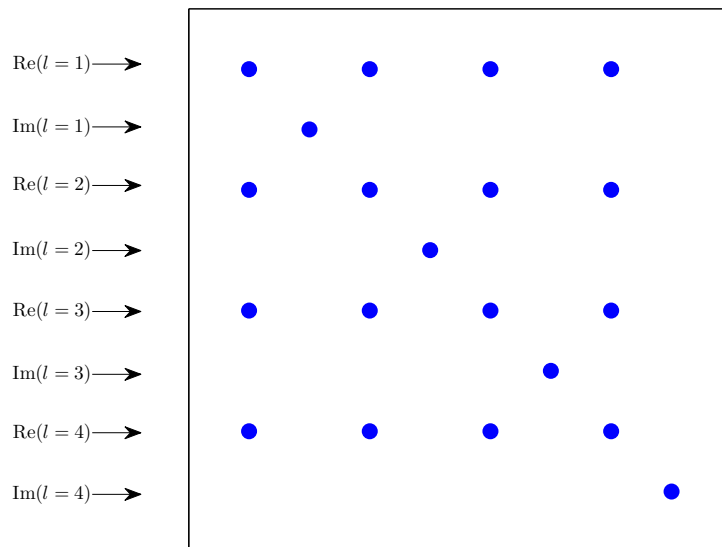
Η προσεκτική παρατήρηση και η μελέτη της δομής του πίνακα του γραμμικού συστήματος της FFX μεθόδου μας έδωσε το έναυσμα για την αναζήτηση μιας καλύτερης επιλογής των *collocation points*. Πιο συγκεκριμένα η δομή των διαγωνίων μπλοκ. Στο σχήμα (3.3) απεικονίζεται το πρώτο διαγώνιο μπλοκ του πίνακα A της (3.12) για την περίπτωση του ισόπλευρου τριγώνου με $N = 8$.

Από το γράφημα φαίνεται ότι οι άρτιες γραμμές έχουν μη-μηδενικό μόνο το διαγώνιο στοιχείο τους. Οι άρτιες γραμμές προέρχονται από το φανταστικό μέρος της κάθε εξίσωσης της (3.6) για $l = 1, 2, \dots$

Ερώτηση 3.2 Μήπως μπορούμε, κρατώντας τις άρτιες γραμμές του πίνακα A , να βρούμε κάποιο καινούργιο σύνολο από l (άλλα *collocation points* στην πραγματικότητα) τα οποία θα μας οδηγήσουν σε ένα πίνακα με διαγώνια μπλοκ που θα είναι διαγώνιοι πίνακες;



Σχήμα 3.2: Οπτική επίδειξη του A_p^j μπλοκ της μεθόδου FFX, στην περίπτωση όπου το πλήθος των συναρτήσεων βάσης είναι 8 ($N = 8$).



Σχήμα 3.3: Η δομή του πρώτου διαγώνιου μπλοκ της FFX για ισόπλευρο τρίγωνο με 8 συναρτήσεις βάσης ανά πλευρά.

Η απάντηση είναι ναι. Και μάλιστα για συναρτήσεις βάσης τις (3.4) η επιλογή των collocation points, που οδηγούν σε πίνακα με τα διαγώνια του μπλοκ να είναι διαγώνια, είναι μοναδική.

Λήμμα 3.1 Έστω r φυσικός διάφορος του μηδενός και k θετικός πραγματικός αριθμός. Συμβολίζοντας I_R και I_I το πραγματικό και το φανταστικό μέρος του ολοκληρώματος

$$I = I(k, r) = \int_{-\pi}^{\pi} e^{iks} \sin\left(r \frac{s + \pi}{2}\right) ds$$

αντίστοιχα, έχουμε τις παρακάτω περιπτώσεις.

$k = \frac{r}{2}$	(α)	$r = 1, 5, 9, \dots$	$I_R = \pi$	$I_I = 0$
	(β)	$r = 2, 6, 10, \dots$	$I_R = 0$	$I_I = -\pi$
	(γ)	$r = 3, 7, 11, \dots$	$I_R = -\pi$	$I_I = 0$
	(δ)	$r = 4, 8, 12, \dots$	$I_R = 0$	$I_I = \pi$
r περιττός, $k \neq \frac{r}{2}$	(α)	$k = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots$	$I_R = 0$	$I_I = 0$
	(β)	k περιττός	$I_R = \frac{4r}{4k^2 - r^2}$	$I_I = 0$
	(γ)	k άρτιος	$I_R = \frac{-4r}{4k^2 - r^2}$	$I_I = 0$
	(δ)	διαφορετικά	$I_R = \frac{-4r}{4k^2 - r^2} \cos(k\pi)$	$I_I = 0$
r άρτιος, $k \neq \frac{r}{2}$	(α)	k ακέραιος	$I_R = 0$	$I_I = 0$
	(β)	$k = \frac{1}{2}, \frac{5}{2}, \frac{9}{2}, \dots$	$I_R = 0$	$I_I = \frac{4r}{4k^2 - r^2}$
	(γ)	$k = \frac{3}{2}, \frac{7}{2}, \frac{11}{2}, \dots$	$I_R = 0$	$I_I = \frac{-4r}{4k^2 - r^2}$
	(δ)	διαφορετικά	$I_R = 0$	$I_I = \frac{4r}{4k^2 - r^2} \sin(k\pi)$

Απόδειξη. Παίρνουμε περιπτώσεις για το k

ΠΕΡΙΠΤΩΣΗ I ($k \neq \frac{r}{2}$)

$$\begin{aligned}
 I &= \int_{-\pi}^{\pi} e^{iks} \sin\left(r \frac{s + \pi}{2}\right) ds \\
 &= \frac{1}{ik} \int_{-\pi}^{\pi} (e^{iks})' \sin\left(r \frac{s + \pi}{2}\right) ds \\
 &= \frac{1}{ik} \left[e^{iks} \sin\left(r \frac{s + \pi}{2}\right) \right]_{-\pi}^{\pi} - \frac{r}{i2k} \int_{-\pi}^{\pi} e^{iks} \cos\left(r \frac{s + \pi}{2}\right) ds \\
 &= \frac{r}{2k^2} \left[e^{iks} \cos\left(r \frac{s + \pi}{2}\right) \right]_{-\pi}^{\pi} + \frac{r^2}{4k^2} \int_{-\pi}^{\pi} e^{iks} \sin\left(r \frac{s + \pi}{2}\right) ds.
 \end{aligned}$$

Άρα
$$\frac{4k^2 - r^2}{4k^2} I = \frac{r}{2k^2} \left[e^{iks} \cos \left(r \frac{s + \pi}{2} \right) \right]_{-\pi}^{\pi}.$$

Δηλαδή

$$\begin{aligned} I &= \frac{2r}{4k^2 - r^2} \left[e^{iks} \cos \left(r \frac{s + \pi}{2} \right) \right]_{-\pi}^{\pi} \\ &= \frac{2r}{4k^2 - r^2} [e^{ik\pi} (-1)^r - e^{-ik\pi}] \\ &= \frac{2r}{4k^2 - r^2} [((-1)^r - 1) (\cos k\pi) + i((-1)^r + 1) (\sin k\pi)]. \end{aligned}$$

Διακρίνοντας το r σε περιττό ή άρτιο παίρνουμε αντίστοιχα

r περιττό

$$I = \frac{-4r}{4k^2 - r^2} (\cos k\pi) \quad \Leftrightarrow$$

$$I = \begin{cases} 0 & , \quad k = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots \\ \frac{4r}{4k^2 - r^2} & , \quad k \text{ περιττό} \\ \frac{-4r}{4k^2 - r^2} & , \quad k \text{ άρτιο} \\ \frac{-4r}{4k^2 - r^2} \cos(k\pi) & , \quad \text{διαφορετικά} \end{cases}$$

r άρτιο

$$I = i \frac{4r}{4k^2 - r^2} (\sin k\pi) \quad \Leftrightarrow$$

$$I = \begin{cases} 0 & , \quad k \text{ φυσικός} \\ i \frac{4r}{4k^2 - r^2} & , \quad k = \frac{1}{2}, \frac{5}{2}, \dots \\ \frac{-4r}{4k^2 - r^2} & , \quad k = \frac{3}{2}, \frac{7}{2}, \dots \\ i \frac{4r}{4k^2 - r^2} \sin(k\pi) & , \quad \text{διαφορετικά} \end{cases}.$$

ΠΕΡΙΠΤΩΣΗ II ($k = \frac{r}{2}$)

$$\begin{aligned} I &= \int_{-\pi}^{\pi} e^{i \frac{r}{2} s} \sin \left(r \frac{s + \pi}{2} \right) ds \\ &= \int_{-\pi}^{\pi} \cos \left(\frac{r}{2} s \right) \sin \left(r \frac{s + \pi}{2} \right) ds + i \int_{-\pi}^{\pi} \sin \left(\frac{r}{2} s \right) \sin \left(r \frac{s + \pi}{2} \right) ds \\ &= \pi \sin \left(\frac{\pi r}{2} \right) + i \pi \cos \left(\frac{\pi r}{2} \right) \end{aligned}$$

Δηλαδή

$$I = \begin{cases} \pi & , \quad r = 1, 5, 9, \dots \\ -i\pi & , \quad r = 2, 6, 10, \dots \\ -\pi & , \quad r = 3, 7, 11, \dots \\ i\pi & , \quad r = 4, 8, 12, \dots \end{cases}.$$

□

Θεώρημα 3.3 Για συναρτήσεις βάσης τις $\varphi_r^{(j)}(s) = \sin\left(r\frac{s+\pi}{2}\right)$ μόνο η επιλογή $l = \frac{r}{2}$ των *collocation points* έχει ως αποτέλεσμα ο πίνακας του τελικού παραγόμενου γραμμικού συστήματος να έχει διαγώνια μπλοκ που είναι διαγώνιοι πίνακες.

Απόδειξη. Ανακαλώντας τους συμβολισμούς της παραγράφου 3.3 και χρησιμοποιώντας το λήμμα (3.1) παρατηρούμε ότι: Σε κάθε γραμμή του C_p^j το t -οστό στοιχείο (δηλαδή ο συντελεστής του αγνώστου U_p^t) θα είναι το πραγματικό ή το φανταστικό μέρος του ολοκληρώματος

$$J(t) := \int_{-\pi}^{\pi} e^{ils} \sin\left(t\frac{s+\pi}{2}\right) ds. \quad (3.16)$$

Ψάχνουμε το κατάλληλο l , που αντιστοιχεί στην r -οστή εξίσωση του μπλοκ C_p^j , τέτοιο ώστε το r -οστό στοιχείο της r -οστής γραμμής να είναι το μόνο μη-μηδενικό στοιχείο της γραμμής. Δηλαδή αν ορίσουμε

$$J_r := J(t=r) = \int_{-\pi}^{\pi} e^{ils} \sin\left(r\frac{s+\pi}{2}\right) ds \quad (3.17)$$

$$J_t := J(t \neq r) = \int_{-\pi}^{\pi} e^{ils} \sin\left(t\frac{s+\pi}{2}\right) ds \quad (3.18)$$

ψάχνουμε τέτοιο l ώστε τουλάχιστον μία από τις παρακάτω δυο προτάσεις να ισχύει.

Πρόταση I :	$\text{Re}(J_r) \neq 0 \quad \wedge \quad \text{Re}(J_t) = 0$ (συνθήκη I) (συνθήκη II)	1
Πρόταση II :	$\text{Im}(J_r) \neq 0 \quad \wedge \quad \text{Im}(J_t) = 0$ (συνθήκη III) (συνθήκη IV)	

(3.19)

Έστω $n \in \mathbb{N}$ (συμπεριλαμβανομένου του μηδενός). Τότε έχουμε δυο περιπτώσεις για το r .

Περίπτωση 1 (r : περιττός)

$$\text{Αν } l = \frac{r}{2} \implies \begin{cases} J_r = \pm\pi \\ J_t = \begin{cases} 0 \\ \pm i \frac{4t}{4l^2 - t^2} \end{cases} \end{cases}, \begin{matrix} t \text{ περιττός} \\ t \text{ άρτιος} \end{matrix}$$

Τότε η Πρόταση I είναι αληθής.

¹ Το σύμβολο $\wedge$ αναπαριστά το λογικό "και".

$$\mathbf{A}\mathbf{v} \quad l \neq \frac{r}{2} \implies \begin{cases} J_r = \frac{-4r}{4l^2 - r^2} \cos l\pi \\ J_t = \begin{cases} \frac{-4t}{4l^2 - t^2} \cos l\pi & , \quad t \text{ περιττός} \\ i \frac{4t}{4l^2 - t^2} \sin l\pi & , \quad t \text{ άρτιος} \end{cases} \end{cases}$$

Τότε η συνθήκη III παραβιάζεται και η πρόταση II είναι ψευδής. Επίσης παρατηρούμε ότι αν η συνθήκη I είναι αληθής τότε $\cos l\pi \neq 0$ που συνεπάγεται ότι η συνθήκη II παραβιάζεται και επομένως και η πρόταση I είναι ψευδής.

Περίπτωση 2 (r : άρτιος)

$$\mathbf{A}\mathbf{v} \quad l = \frac{r}{2} \implies \begin{cases} J_r = \pm i\pi \\ J_t = \begin{cases} \pm \frac{4t}{4l^2 - t^2} & , \quad t \text{ περιττός} \\ 0 & , \quad t \text{ άρτιος} \end{cases} \end{cases}$$

Η πρόταση II είναι αληθής.

$$\mathbf{A}\mathbf{v} \quad l \neq \frac{r}{2} \implies \begin{cases} J_r = i \frac{4r}{4l^2 - r^2} \sin l\pi \\ J_t = \begin{cases} \frac{-4t}{4l^2 - t^2} \cos l\pi & , \quad t \text{ περιττός} \\ i \frac{4t}{4l^2 - t^2} \sin l\pi & , \quad t \text{ άρτιος} \end{cases} \end{cases}$$

Τότε η συνθήκη I παραβιάζεται έτσι και η πρόταση I είναι ψευδής. Επίσης παρατηρούμε ότι αν η συνθήκη III είναι αληθής τότε $\sin l\pi \neq 0$ που συνεπάγεται ότι η συνθήκη IV παραβιάζεται και επομένως και η πρόταση II είναι ψευδής.

Το συμπέρασμα είναι ότι για κάθε r παίρνουμε $l = \frac{r}{2}$ και διαλέγουμε το πραγματικό μέρος της αντίστοιχης εξίσωσης εάν το r είναι περιττός και το φανταστικό μέρος εάν το r είναι άρτιος. Η επιλογή αυτή μάλιστα, λόγω της κατασκευαστικής μορφής της απόδειξης, είναι μοναδική. $\square$

Ορισμός 3.4 (SFFS) Η αριθμητική μέθοδος που προκύπτει από την διακριτοποίηση της *Dirichlet-Neumann* απεικόνιση με την παρακάτω επιλογή *collocation points*

$$k_p := \left\{ \frac{-l}{h_p} \right\}_{l=\frac{1}{2}, 1, \frac{3}{2}, \dots, N_j/2} = \left\{ \frac{-0.5}{h_p}, \frac{-1}{h_p}, \frac{-1.5}{h_p}, \dots, \frac{-N_j/2}{h_p} \right\} \quad (3.20)$$

και επιλογή εξισώσεων: το πραγματικό μέρος για $l = 0.5, 1.5, \dots$ και το φανταστικό μέρος για $l = 1, 2, \dots$ της κάθε μιας από τις εξισώσεις της σχέσης (2.42), ονομάζεται *SFFS*.

	$r = 1$	$r = 2$	$r = 3$	$r = 4$				
$Re(l = 0.5) \rightarrow$	x	x	x	x	x	x	x	x
$Im(l = 1) \rightarrow$	x	x	x	x	x	x	x	x
$Re(l = 1.5) \rightarrow$	x	x	x	x	x	x	x	x
$Im(l = 2) \rightarrow$	x	x	x	x	x	x	x	x
$Re(l = 2.5) \rightarrow$	x	x	x	x	x	x	x	x
$Im(l = 3) \rightarrow$	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x

Σχήμα 3.4: Οπτική επίδειξη του A_p^j μπλοκ της μεθόδου SFFS, στην περίπτωση όπου το πλήθος των συναρτήσεων βάσης είναι 8 ($N = 8$).

Το σχήμα (3.4) μας βοηθάει να καταλάβουμε τις παραπάνω επιλογές. Αναφέρεται σε κάθε τυχαίο μπλοκ διάστασης ($N \times N$) με $N = 8$ του A .

3.5 Διαμέριση του Collocation Πίνακα

Ως collocation πίνακα θεωρούμε τον πίνακα του τελικού γραμμικού συστήματος που προκύπτει μετά την διακριτοποίηση της Dirichlet-Neumann απεικόνιση. Ας σημειωθεί ότι είναι δυνατή η διαίρεση της κάθε γραμμής του πίνακα $A(row, :)$ με το διαγώνιο στοιχείο της $A(row, row)$. Φυσικά με το ίδιο στοιχείο πρέπει να διαιρεθεί και το αντίστοιχο στοιχείο του δεξιού μέλους. Αυτή η διαίρεση γίνεται με σκοπό τα διαγώνια μπλοκ του πίνακα της sine-SFFS¹ να είναι μοναδιαίοι πίνακες. Την ίδια πρακτική μπορούμε να ακολουθήσουμε και για την sine-FFX. Για ευκολότερη

¹ Λέγοντας sine-SFFS θα εννοούμε ότι χρησιμοποιούμε τις ημιτονοειδείς συναρτήσεις (3.4) ως συναρτήσεις βάσης και την μέθοδο SFFS για την επιλογή των collocation σημείων.

αναφορά στον collocation πίνακα A θεωρούμε την παρακάτω μπλοκ μορφή του

$$A = \begin{bmatrix} A_1^1 & & A_1^n \\ & \ddots & \\ A_n^1 & & A_n^n \end{bmatrix} \quad (3.21)$$

Το κάθε μπλοκ είναι διάστασης $N \times N$. Το μπλοκ A_p^j συνδέει τις πλευρές p και j του πολυγώνου υπό την εξής έννοια: Στο μπλοκ A_p^j τα collocation σημεία $(-l/h_p)$ "ανήκουν" στην πλευρά p ενώ οι άγνωστοι (U_r^j) αναφέρονται στην πλευρά j .

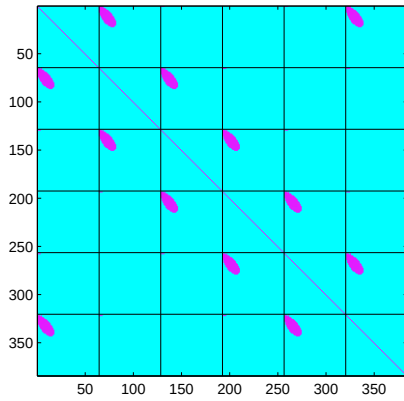
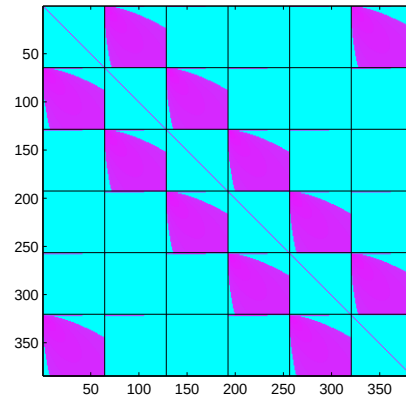
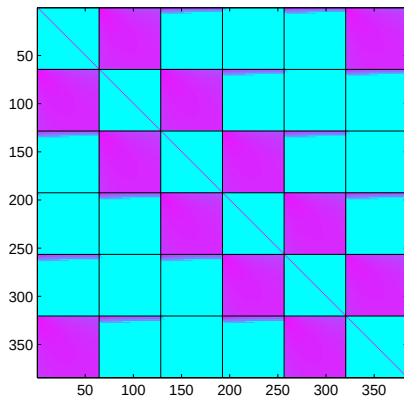
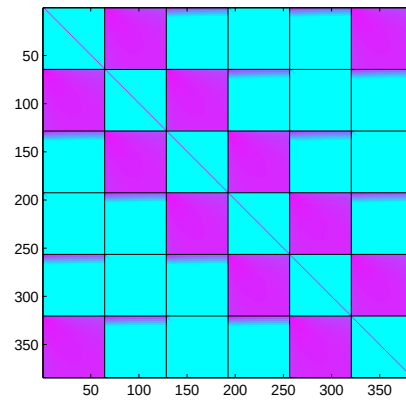
3.6 Ασυμπτωτική Συμπεριφορά του Collocation Πίνακα

Στην ανάλυση που δώσαμε στην παράγραφο (2.2.2) είχαμε καταλήξει στο συμπέρασμα ότι για $k_p = \frac{-l}{h_p}$ οι ποσότητες που εμφανίζονται στο αριστερό μέλος της Dirichlet-Neumann απεικόνισης (τα στοιχεία του collocation πίνακα) τείνουν στο μηδέν όσο αυξάνει το μέτρο του k , όσο μεγαλύτερα l δηλαδή χρησιμοποιούμε. Αυτή η συμπεριφορά περιμένουμε να φαίνεται, σε γενικές γραμμές, και στον collocation πίνακα για την sine-SFSS. Για να τονίσουμε τις διαφορές στο "μέγεθος" του κάθε στοιχείου θα χρησιμοποιήσουμε μια μεταβλητή threshold και θα θεωρούμε όλα τα στοιχεία του πίνακα που είναι, κατά απόλυτη τιμή, μικρότερα του threshold μηδενικά. Στο σχήμα (3.6) φαίνεται η γραφική αναπαράσταση του collocation πίνακα για ένα κανονικό εξαγώνο με 64 συναρτήσεις βάσης ανά πλευρά, για διάφορες τιμές του threshold.

Βλέποντας το γράφημα (3.6) μπορούμε να παρατηρήσουμε:

- Τα διαγώνια μπλοκ του πίνακα είναι διαγώνιοι πίνακες. Αυτό οφείλεται στην επιλογή των collocation points στην sine-SFSS.
- Τα μεγαλύτερα, κατά απόλυτη τιμή, στοιχεία του κάθε μπλοκ, πέρα των διαγωνίων, βρίσκονται στις πρώτες γραμμές του κάθε μπλοκ. Οι γραμμές αυτές αντιστοιχούν στα μικρότερα l .
- Μπλοκ που συνδέουν παράλληλες πλευρές του εξαγώνου, $A_1^4, A_2^5, \dots$, έχουν πολύ μικρότερα στοιχεία από τα υπόλοιπα. Ας κοιτάξουμε το γράφημα (;;) που δείχνει την συμπεριφορά της $\frac{e^{-x}}{x}$ σε σχέση με τις άλλες συναρτήσεις και ας θυμηθούμε ότι στην απόδειξη του θεωρήματος (;;) δείξαμε ότι για $\sin(\alpha_j - \alpha) = 0$ (δηλαδή για παράλληλες πλευρές p, j) το M τείνει στο μηδέν τόσο γρήγορα όσο η $\frac{e^{-x}}{x}$ καθώς το x τείνει στο ∞ .

Αν διπλασιάσουμε το πλήθος των συναρτήσεων βάσης (από N σε $2N$) πώς θα αλλάξει ο

(α') $threshold = 1E - 02$ (β') $threshold = 1E - 03$ (γ') $threshold = 1E - 10$ (δ') $threshold = 1E - 15$

Σχήμα 3.5: Γραφική αναπαράσταση του collocation πίνακα για διάφορες τιμές του $threshold$. Το ανοικτό γαλάζιο δηλώνει μηδενικό στοιχείο. Τα γραφήματα έγιναν με την εντολή `imagesc(log10(abs(A)))` του Matlab.

πίνακας της sine-SFFS (από A_N σε A_{2N}); Θα το δείξουμε σχηματικά για ένα τρίγωνο.

$$A_N \quad A_N = \begin{bmatrix} I & A_1^2 & A_1^3 \\ A_2^1 & I & A_2^3 \\ A_3^1 & A_3^2 & I \end{bmatrix}, \quad \text{τότε} \quad A_{2N} = \left[\begin{array}{cc|cc|cc} I & 0 & A_1^2 & \times & A_1^3 & \times \\ 0 & I & \times & \times & \times & \times \\ \hline A_2^1 & \times & I & 0 & A_2^3 & \times \\ \times & \times & 0 & I & \times & \times \\ \hline A_3^1 & \times & A_3^2 & \times & I & 0 \\ \times & \times & \times & \times & 0 & I \end{array} \right] \quad (3.22)$$

οπού $\times$ τα νέα μπλοκ που εμφανίζονται.

Παρατήρηση 3.1 Ίσως κάποιος αναρωτηθεί: Δεδομένου ότι (α) έχουμε ένα τρόπο να επιλέγουμε τα *collocation points* έτσι ώστε τα διαγώνια μπλοκ να είναι διαγώνιοι πίνακες ($l = \frac{r}{2}$

βλ. *sine-SFFS*) και (β) τα στοιχεία των υπολοίπων μπλοκ τείνουν στο μηδέν όσο αυξάνει το l γιατί να μην επιλέξουμε ως συναρτήσεις βάσης τις $\sin\left((M+r)\frac{s+\pi}{2}\right)$ όπου M ένας αρκετά μεγάλος ακέραιος έτσι ώστε ο *collocation* πίνακας να είναι πρακτικά διαγώνιος και το σύστημα να λύνεται άμεσα; Αυτό δυστυχώς δεν είναι υπολογιστικά εφικτό διότι οι συναρτήσεις βάσης $\sin\left((M+1)\frac{s+\pi}{2}\right), \dots, \sin\left((M+N_j)\frac{s+\pi}{2}\right)$ είναι πολύ χειρότερη επιλογή, ως αριθμητική βάση, από τις $\sin\left(1\frac{s+\pi}{2}\right), \dots, \sin\left(N_j\frac{s+\pi}{2}\right)$ και ο αριθμητικός υπολογισμός των ποσοτήτων $G_p(l=M+r)$ θα ήταν εξαιρετικά ασταθής για πολύ μεγάλα M .

Κεφάλαιο 4

Σύγκριση Μεθόδων SFFS και FFX

Σε αυτό το κεφάλαιο θα συγκρίνουμε, αριθμητικά, τις δυο μεθόδους sine-SFFS και sine-FFX. Τα δυο κριτήρια για την σύγκριση θα είναι ο ρυθμός σύγκλισης των δυο μεθόδων καθώς και ο δείκτης κατάστασης των δυο collocation πινάκων. Θα μελετήσουμε δυο κατηγορίες πολυγώνων: κανονικά και μη κανονικά. Τα κανονικά είναι τα ίδια που είχαν χρησιμοποιήσει και οι Fulton, Φωκάς, Ξενοφώντος στο [FFX03] ενώ για μη κανονικά έχουμε επιλέξει μια νέα κλάση πολυγώνων: πολύγωνα με κορυφές πάνω σε μια συγκεκριμένη έλλειψη. Μέσω των αριθμητικών παραδειγμάτων θα φανεί η ανωτερότητα της νέας μεθόδου όσο αφορά στην τάξη σύγκλισης και τον δείκτη κατάστασης του collocation πίνακα [SFFS].

4.1 Το Αριθμητικό Πρόβλημα - Μοντέλο

Για την αριθμητική μελέτη των μεθόδων SFFS και FFX θα επιλύσουμε την εξίσωση του Laplace σε μια ποικιλία πολυγωνικών χωρίων κανονικών και μη, με Dirichlet, Neumann και Mixed($\beta = \frac{\pi}{3}$) συνοριακές συνθήκες. Ειδικότερα τα πολύγωνα που θα χρησιμοποιήσουμε είναι τα

Κανονικά: Βρίσκονται πάνω στον κύκλο με ακτίνα $\rho = \sqrt{2}$. Οι κορυφές του κανονικού n -γώνου $z_1, z_2, \dots, z_n$ δίνονται από τον τύπο

$$z_k = \sqrt{2}e^{i[2(k-1)\frac{\pi}{n} - \frac{1}{5}]} \quad (4.1)$$

Μη-κανονικά: Είναι πολύγωνα με κορυφές πάνω στην έλλειψη

$$\frac{x^2}{5} + \frac{y^2}{2} = 1 \quad (4.2)$$

στραμμένα κατά 0.2 (κατά την θετική φορά). Οι τετμημένες τους, πριν τη στροφή, και αριθμώντας κατά την θετική φορά, είναι

Τρίγωνο	x	$=$	$-4, -1, 3$	(4.3)
Τετράπλευρο	x	$=$	$-4, 4, 4, 1$	
Πεντάγωνο	x	$=$	$-5, -2, 4, 3, 0$	
Εξάγωνο	x	$=$	$-4, -1, 2, 4.5, 1, -4.5$	
Οκτάγωνο	x	$=$	$-5, -4, -1, 2, 3, 1, -2, -3$	

Τα πολύγωνα φαίνονται στο σχήμα (4.1).

Σε όλες τις περιπτώσεις η ακριβής λύση της εξίσωσης του Laplace δίνεται από

$$q(x, y) = \sinh(3x) \sin(3y) \quad (4.4)$$

Γνωρίζοντας την ακριβή λύση οι συνοριακές συνθήκες και οι συναρτήσεις

$$\{f^{(j)}(s)\}_{j=1}^n, \quad \{g^{(j)}(s)\}_{j=1}^n$$

μπορούν να παραχθούν αναλυτικά.

Το σχετικό σφάλμα E_∞ που χρησιμοποιούμε για την αξιολόγηση της μεθόδου δίνεται από

$$E_\infty = \frac{\|f - f_N\|_\infty}{\|f\|_\infty} \quad (4.5)$$

όπου

$$\|f\|_\infty = \max_{1 \leq j \leq n} \left\{ \max_{-\pi \leq s \leq \pi} |f^{(j)}(s)| \right\}$$

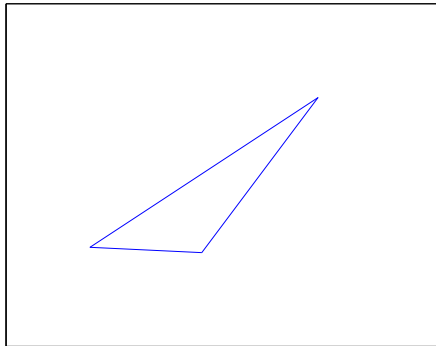
και

$$\|f - f_N\|_\infty = \max_{1 \leq j \leq n} \left\{ \max_{-\pi \leq s \leq \pi} |f^{(j)}(s) - f_N^{(j)}(s)| \right\}$$

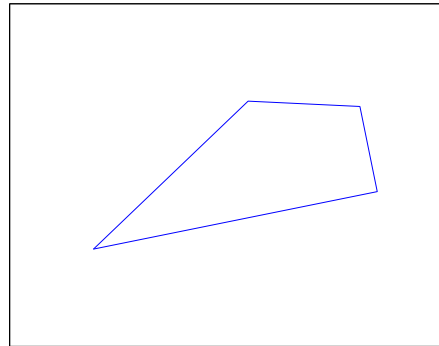
Η νόρμα απείρου υπολογίζεται πάνω σε μια πυκνή διαμέριση του $[-\pi, \pi]$.

4.2 Αριθμητική Υλοποίηση

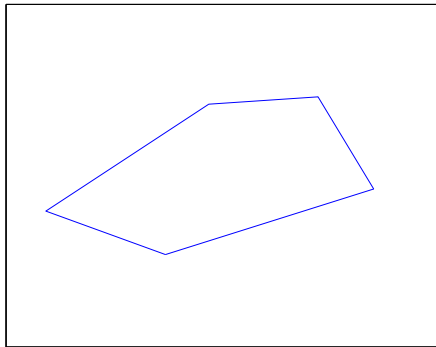
Όλοι οι αριθμητικοί υπολογισμοί έγιναν με αριθμητική διπλής ακρίβειας στην γλώσσα προγραμματισμού Fortran [MRE00]. Ο δείκτης κατάστασης υπολογίστηκε με την Matlab. Ο υπολογισμός των ολοκληρωμάτων στο πρώτο μέλος της (2.42) γίνεται αναλυτικά κάνοντας χρήση του λήμματος (3.1). Για τον υπολογισμό των ολοκληρωμάτων στο δεύτερο μέλος χρησιμοποιήσαμε την ρουτίνα dqawo του πακέτου προγραμμάτων QUADPACK [QUADPACK]. Η συγκεκριμένη ρουτίνα στηρίζεται στην τεχνική Clenshaw-Curtis για να υπολογίζει ορισμένα ολοκληρώματα



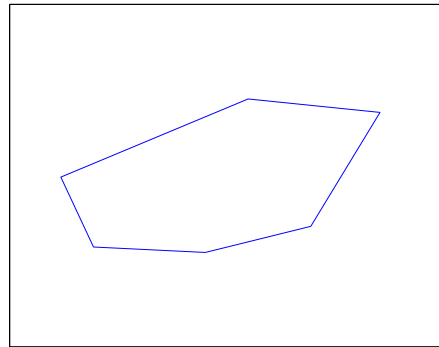
(α') Τρίγωνο



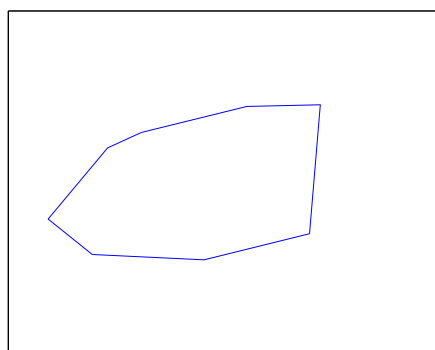
(β') Τετράπλευρο



(γ') Πεντάγωνο



(δ') Εξάγωνο



(ε') Οκτάγωνο

Σχήμα 4.1: Τα μη-κανονικά πολύγωνα που χρησιμοποιούμε για την σύγκριση των μεθόδων *SFFS* και *FFX*.

της μορφής $\cos(\omega x)F(x)$ και $\sin \omega x)F(x)$. Η υπολογιστική δυσκολία στον υπολογισμό τέτοιου είδους ολοκληρωμάτων, λόγω του εκθετικού που υπάρχει μέσα στο ολοκλήρωμα (2.43), έγκειται στο γεγονός ότι η προς ολοκλήρωση ποσότητα παρουσιάζει πολύ μεγάλες διακυμάνσεις.

Ο (αριθμητικός) ρυθμός σύγκλισης, που παρουσιάζουμε στους πίνακες (Ο.ο.Ο.) ορίζεται σύμφωνα με την παρακάτω λογική. Έστω E_x το σφάλμα όταν χρησιμοποιούμε x συναρτήσεις βάσης ανά πλευρά και E_y το σφάλμα όταν χρησιμοποιούμε y ($> x$) συναρτήσεις βάσης ανά πλευρά. Έστω ότι η μέθοδος έχει ρυθμό σύγκλισης ίσο με t . Τότε υπάρχει σταθερά c τέτοια ώστε $E_y = ch^t$ όπου το h , χωρίς βλάβη της γενικότητας, είναι $\frac{1}{y}$. Άρα

$$\frac{E_x}{E_y} = \frac{c \left(\frac{1}{x}\right)^t}{c \left(\frac{1}{y}\right)^t} = \left(\frac{y}{x}\right)^t \Rightarrow$$

$$t = \log_{y/x} \frac{E_x}{E_y}.$$

4.3 Αριθμητικά Αποτελέσματα

Οι πίνακες (4.1)-(4.5) δείχνουν το σχετικό σφάλμα (E_∞) και τον ρυθμό σύγκλισης (Ο.ο.Ο.) ως προς το πλήθος των συναρτήσεων βάσης (N) ανά πλευρά των μεθόδων SFFS και FFX για ένα ισόπλευρο τρίγωνο, ένα τετράγωνο, ένα κανονικό πεντάγωνο, ένα κανονικό εξάγωνο και για ένα κανονικό οκτάγωνο αντίστοιχα. Σε κάθε περίπτωση οι συνοριακές συνθήκες (BC) είναι ίδιες σε όλες τις πλευρές του πολυγώνου ενώ έχουμε χρησιμοποιήσει Dirichlet, Neumann και Mixed($\beta = \frac{\pi}{3}$) συνοριακές συνθήκες σε όλα τα πολύγωνα.

Τα σχήματα (4.2)-(4.7) απεικονίζουν το σχετικό σφάλμα των μεθόδων SFFS και FFX για συνοριακές συνθήκες Dirichlet, Neumann και Mixed($\beta = \frac{\pi}{3}$) αντίστοιχα. Οι διακεκομμένες γραμμές στα σχήματα είναι οι ευθείες $y = \frac{1}{N}x$, $y = \left(\frac{1}{N}\right)^2 x$ και $y = \left(\frac{1}{N}\right)^3 x$.

Ανάλογοι πίνακες ((4.6)-(4.10)) και ανάλογα σχήματα ((4.8)-(4.13)) υπάρχουν και για τα μη κανονικά πολύγωνα.

Με επισκόπηση των παραπάνω πινάκων και σχημάτων μπορούμε να συμπεράνουμε τα εξής:

- Και οι δυο μέθοδοι συγκλίνουν.
- Το σφάλμα της SFFS μεθόδου είναι πάντα μικρότερο από το σφάλμα της FFX μεθόδου. Μάλιστα υπάρχουν περιπτώσεις όπου το σφάλμα της SFFS είναι μικρότερο σχεδόν κατά μία τάξη μεγέθους στα μη κανονικά πολύγωνα (π.χ. Τετράπλευρο $N = 256$) ή και παραπάνω από μια τάξη μεγέθους στα κανονικά πολύγωνα (π.χ. Πεντάγωνο $N = 256$).
- Ο ρυθμός σύγκλισης της FFX δεν είναι σταθερός αλλά κυμαίνεται από 1 (τετράγωνο-Mixed BC) έως και 2 (π.χ. εξάγωνο-Mixed BC).

- Ο ρυθμός σύγκλισης της FFX, στις περισσότερες περιπτώσεις, είναι συνάρτηση του πολυγώνου και δεν φαίνεται να είναι ισχυρά εξαρτώμενος από τις συνοριακές συνθήκες που χρησιμοποιούνται (π.χ. στο ισόπλευρο τρίγωνο είναι 1.5 για όλες τις συνοριακές συνθήκες).
- Η SFFS μέθοδος έχει σταθερά τετραγωνική σύγκλιση για όλα τα πολύγωνα (κανονικά και μη) και για όλες τις συνοριακές συνθήκες.

Στους πίνακες (4.11) και (4.12) παρουσιάζουμε τον δείκτη κατάστασης του πίνακα των συντελεστών του γραμμικού συστήματος για κανονικά και μη κανονικά πολύγωνα αντίστοιχα. Οι δείκτες κατάστασης φαίνονται και στο σχήμα (4.14) κατηγοριοποιημένοι ανά μέθοδο και ανά κατηγορία πολυγώνων¹.

Σχετικά με τον δείκτη κατάστασης μπορούμε να παρατηρήσουμε τα εξής:

- Ο δείκτης κατάστασης είναι αύξουσα συνάρτηση του πλήθους των συναρτήσεων βάσης (N) ανά πλευρά και για τις δυο μεθόδους και για τις δυο κατηγορίες πολυγώνων. Εξαίρεση αποτελεί το τετράγωνο για την SFFS μέθοδο όπου ο δείκτης κατάστασης είναι σταθερός και ανεξάρτητος του N .
- Στην SFFS ο ρυθμός αύξησης του δείκτη κατάστασης ως προς N είναι πολύ μικρότερος από ότι στην FFX εκτός από την περίπτωση του μη κανονικού οκταγώνου και για $N = 4$.
- Στην SFFS μέθοδο και στο μη κανονικό οκτάγωνο παρατηρούμε μια ραγδαία αύξηση του δείκτη κατάστασης σε σχέση με τα υπόλοιπα μη κανονικά πολύγωνα. Και σε αυτή την περίπτωση όμως για μεγάλα N ο δείκτης κατάστασης της SFFS είναι αρκετά μικρότερος από αυτόν της FFX.

4.3.1 Σφάλμα - Κανονικά Πολύγωνα

¹ Στην περίπτωση που χρησιμοποιούμε τις ίδιες συνοριακές συνθήκες σε όλες τις πλευρές ο collocation πίνακας είναι ανεξάρτητος του τύπου των συνοριακών συνθηκών. Αυτό θα φανεί καθαρά σε επόμενο κεφάλαιο όπου και εξετάζουμε λεπτομερέστερα τις ιδιότητες του collocation πίνακα.

reg $n=3$	Dirichlet BC				Neumann BC				Mixed BC ($\pi/3$)			
	SFFS		FFX		SFFS		FFX		SFFS		FFX	
N	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C
4	6.94E-02		2.48E-01		1.72E-02		6.88E-02		5.18E-02		1.93E-01	
8	2.04E-02	1.77	9.83E-02	1.33	4.03E-03	2.09	2.55E-02	1.43	1.33E-02	1.97	7.45E-02	1.37
16	5.44E-03	1.90	3.61E-02	1.44	1.06E-03	1.92	9.21E-03	1.47	3.39E-03	1.97	2.73E-02	1.45
32	1.40E-03	1.95	1.30E-02	1.48	2.72E-04	1.97	3.29E-03	1.48	8.67E-04	1.97	9.80E-03	1.48
64	3.56E-04	1.98	4.61E-03	1.49	6.86E-05	1.99	1.17E-03	1.49	2.19E-04	1.98	3.49E-03	1.49
128	8.96E-05	1.99	1.64E-03	1.50	1.72E-05	1.99	4.15E-04	1.50	5.51E-05	1.99	1.24E-03	1.49
256	2.25E-05	2.00	5.79E-04	1.50	4.31E-06	2.00	1.47E-04	1.50	1.38E-05	2.00	4.38E-04	1.50

Πίνακας 4.1: Sine-collocation: Ισόπλευρο Τρίγωνο.

reg $n=4$	Dirichlet BC				Neumann BC				Mixed BC ($\pi/3$)			
	SFFS		FFX		SFFS		FFX		SFFS		FFX	
N	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C
4	3.94E-02		2.59E-01		3.20E-02		4.51E-02		4.23E-02		1.54E-01	
8	1.04E-02	1.93	1.27E-01	1.03	9.78E-03	1.71	1.53E-02	1.56	1.24E-02	1.76	7.55E-02	1.03
16	2.70E-03	1.94	6.34E-02	1.00	2.66E-03	1.88	7.04E-03	1.12	3.35E-03	1.89	3.78E-02	1.00
32	6.91E-04	1.97	3.19E-02	0.99	6.88E-04	1.95	3.53E-03	0.99	8.64E-04	1.95	1.91E-02	0.99
64	1.75E-04	1.98	1.60E-02	0.99	1.75E-04	1.98	1.77E-03	0.99	2.19E-04	1.98	9.58E-03	0.99
128	4.40E-05	1.99	8.05E-03	0.99	4.40E-05	1.99	8.90E-04	1.00	5.52E-05	1.99	4.81E-03	1.00
256	1.10E-05	1.99	4.03E-03	1.00	1.10E-05	1.99	4.46E-04	1.00	1.39E-05	1.99	2.41E-03	1.00

Πίνακας 4.2: Sine-collocation: Τετράγωνο.

reg $n=5$	Dirichlet BC				Neumann BC				Mixed BC ($\pi/3$)			
	SFFS		FFX		SFFS		FFX		SFFS		FFX	
N	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C
4	2.26E-02		3.83E-02		3.09E-02		4.15E-02		2.85E-02		4.04E-02	
8	5.96E-03	1.92	1.21E-02	1.67	8.89E-03	1.80	1.18E-02	1.81	7.99E-03	1.83	1.19E-02	1.77
16	1.61E-03	1.89	4.31E-03	1.48	2.38E-03	1.90	3.26E-03	1.86	2.12E-03	1.91	3.24E-03	1.87
32	4.15E-04	1.95	1.63E-03	1.40	6.13E-04	1.96	1.20E-03	1.44	5.46E-04	1.96	8.49E-04	1.93
64	1.05E-04	1.98	6.28E-04	1.38	1.56E-04	1.98	4.48E-04	1.42	1.39E-04	1.98	2.18E-04	1.96
128	2.66E-05	1.99	2.43E-04	1.37	3.92E-05	1.99	1.68E-04	1.42	3.49E-05	1.99	7.22E-05	1.59
256	6.67E-06	1.99	9.42E-05	1.37	9.84E-06	1.99	6.31E-05	1.41	8.76E-06	1.99	2.57E-05	1.49

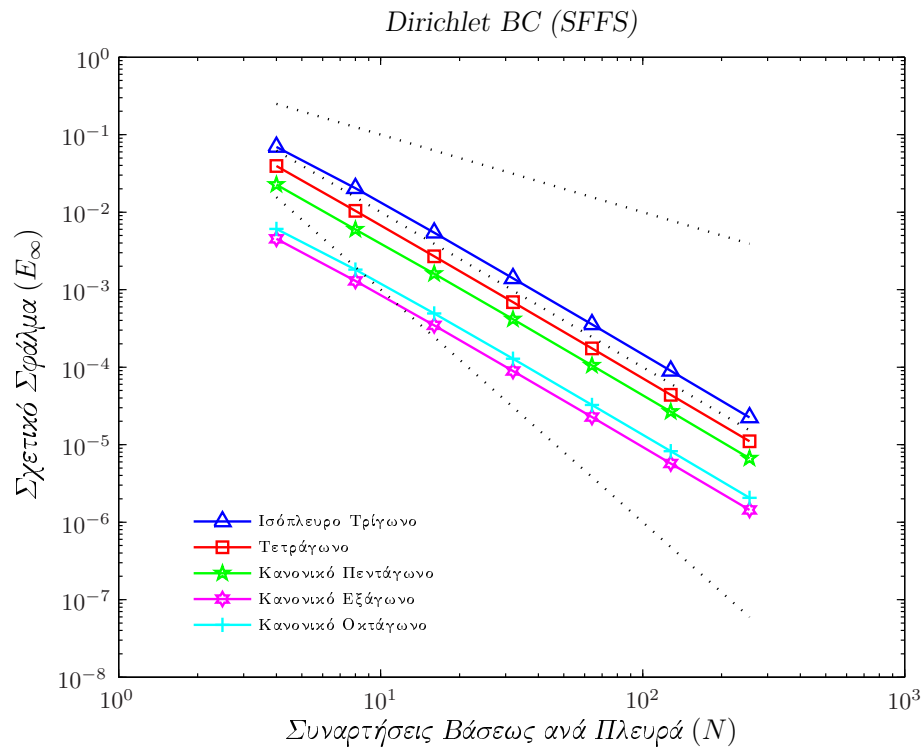
Πίνακας 4.3: Sine-collocation: Κανονικό Πεντάγωνο.

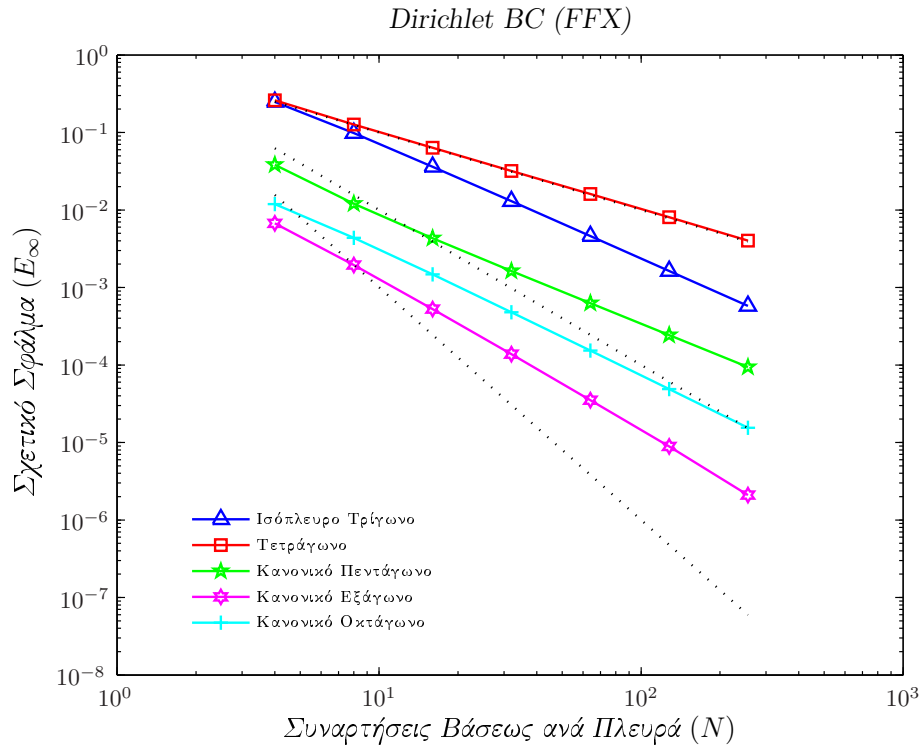
reg $n=6$	Dirichlet BC				Neumann BC				Mixed BC ($\pi/3$)			
	SFFS		FFX		SFFS		FFX		SFFS		FFX	
N	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C
4	4.52E-03		6.72E-03		2.38E-02		3.04E-02		1.73E-02		2.25E-02	
8	1.30E-03	1.80	1.94E-03	1.79	6.64E-03	1.84	8.95E-03	1.77	4.76E-03	1.86	6.49E-03	1.79
16	3.47E-04	1.90	5.27E-04	1.88	1.77E-03	1.91	2.43E-03	1.88	1.26E-03	1.92	1.75E-03	1.89
32	8.94E-05	1.96	1.38E-04	1.94	4.56E-04	1.96	6.33E-04	1.94	3.25E-04	1.96	4.56E-04	1.94
64	2.27E-05	1.98	3.50E-05	1.97	1.16E-04	1.98	1.62E-04	1.97	8.24E-05	1.98	1.16E-04	1.97
128	5.71E-06	1.99	8.87E-06	1.98	2.92E-05	1.99	4.02E-05	2.00	2.08E-05	1.99	2.90E-05	2.00
256	1.43E-06	1.99	2.10E-06	2.08	7.32E-06	1.99	9.83E-06	2.03	5.21E-06	1.99	7.07E-06	2.04

Πίνακας 4.4: Sine-collocation: Κανονικό Εξάγωνο.

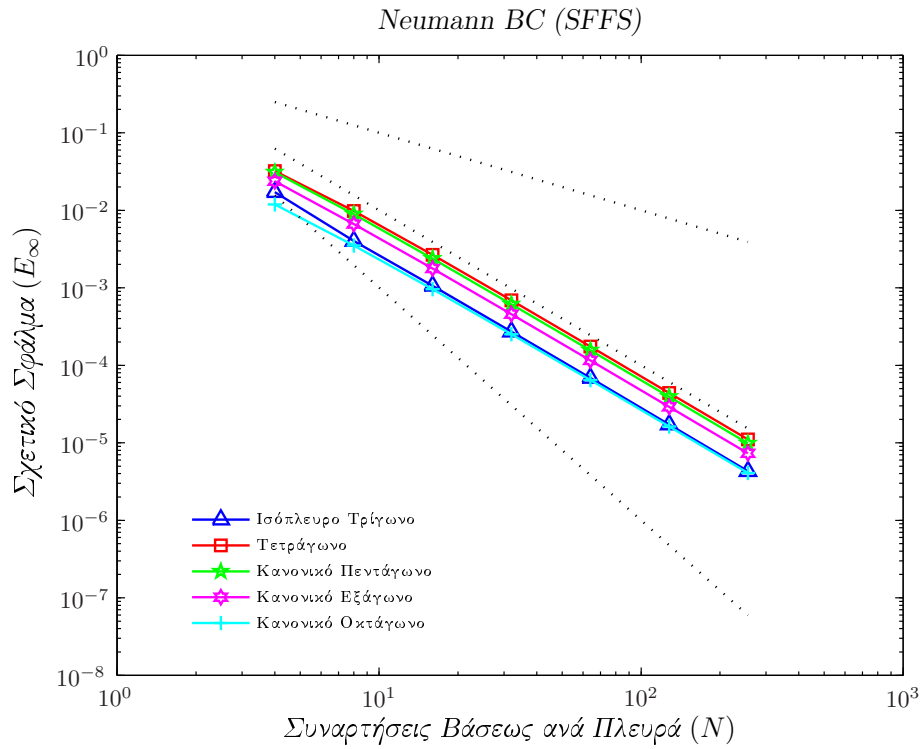
$n=8$	<i>SFFS</i>		<i>FFX</i>		<i>SFFS</i>		<i>FFX</i>		<i>SFFS</i>		<i>FFX</i>	
N	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C
4	6.09E-03		1.19E-02		1.19E-02		1.59E-02		1.10E-02		1.56E-02	
8	1.81E-03	1.75	4.37E-03	1.44	3.51E-03	1.77	4.53E-03	1.81	3.25E-03	1.76	4.64E-03	1.75
16	4.93E-04	1.88	1.47E-03	1.57	9.59E-04	1.87	1.21E-03	1.90	8.88E-04	1.87	1.28E-03	1.86
32	1.28E-04	1.94	4.79E-04	1.62	2.51E-04	1.93	3.14E-04	1.95	2.33E-04	1.93	3.66E-04	1.80
64	3.26E-05	1.98	1.54E-04	1.64	6.43E-05	1.97	9.30E-05	1.76	5.95E-05	1.97	1.14E-04	1.68
128	8.22E-06	1.99	4.88E-05	1.65	1.63E-05	1.98	2.88E-05	1.69	1.51E-05	1.98	3.59E-05	1.67
256	2.06E-06	1.99	1.54E-05	1.66	4.09E-06	1.99	9.01E-06	1.68	3.78E-06	1.99	1.13E-05	1.67

Πίνακας 4.5: Sine-collocation: Κανονικό Οκτάγωνο.

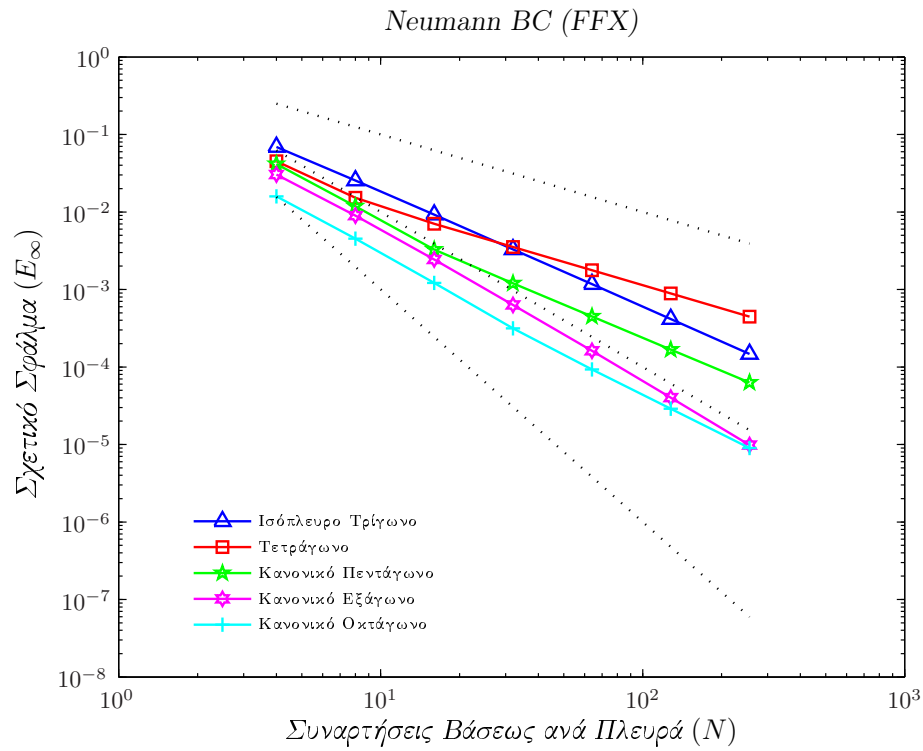
Σχήμα 4.2: Σφάλμα της μεθόδου *SFFS* για κανονικά πολύγωνα με ημιτονοειδή συναρτήσεις βάσης και *Dirichlet* συνοριακές συνθήκες.



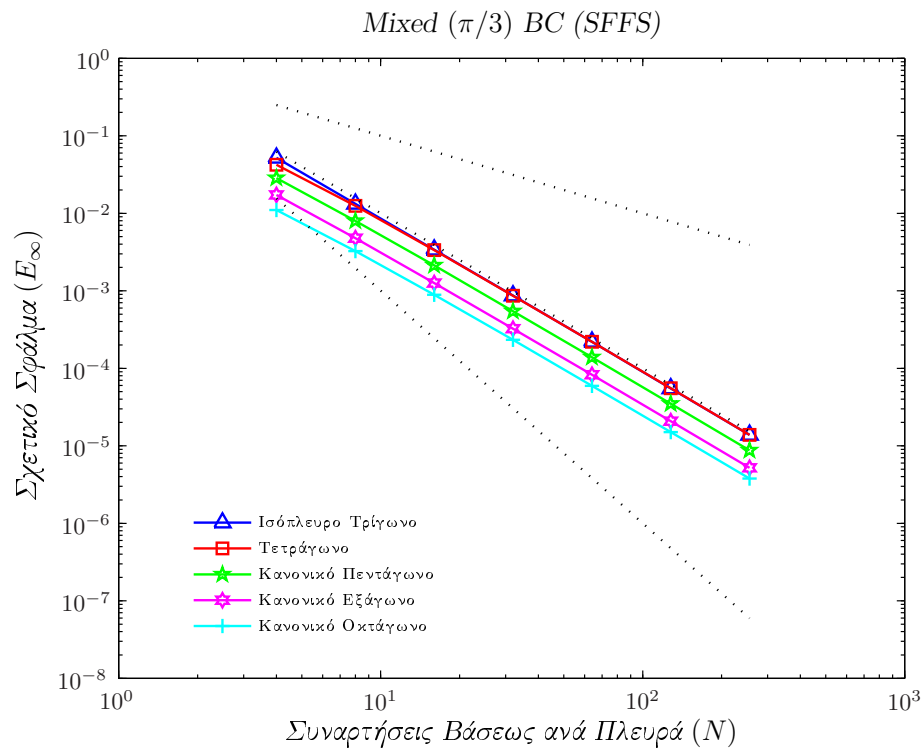
Σχήμα 4.3: Σφάλμα της μεθόδου FFX για κανονικά πολύγωνα με ημιτονοειδή συναρτήσεις βάσης και *Dirichlet* συνοριακές συνθήκες.



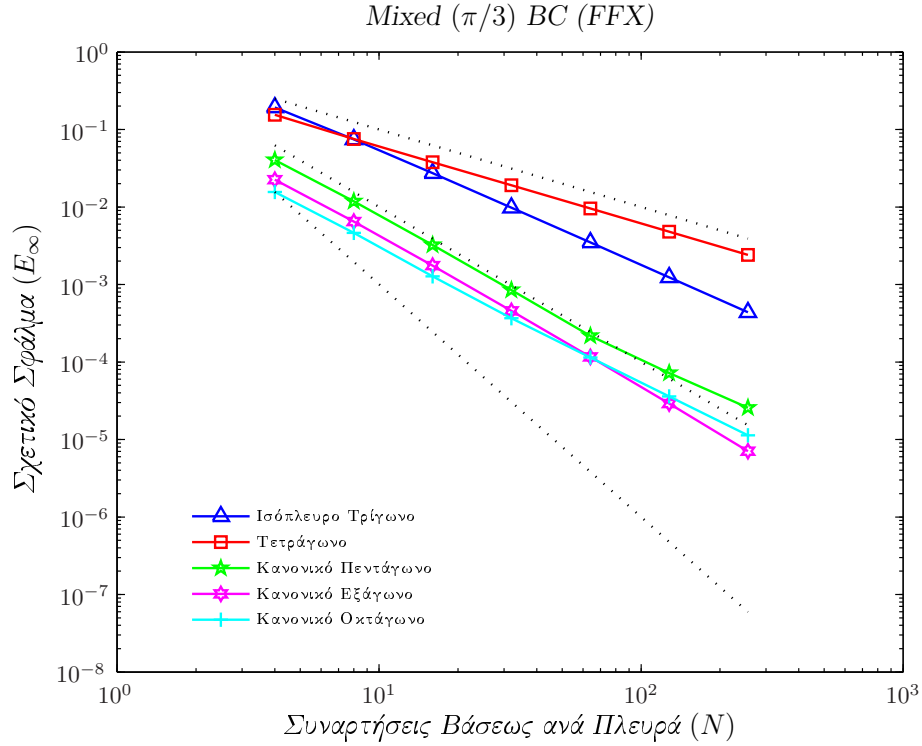
Σχήμα 4.4: Σφάλμα της μεθόδου SFFS για κανονικά πολύγωνα με ημιτονοειδή συναρτήσεις βάσης και *Neumann* συνοριακές συνθήκες.



Σχήμα 4.5: Σφάλμα της μεθόδου FFX για κανονικά πολύγωνα με ημιτονοειδή συναρτήσεις βάσης και Neumann συνοριακές συνθήκες.



Σχήμα 4.6: Σφάλμα της μεθόδου SFFS για κανονικά πολύγωνα με ημιτονοειδή συναρτήσεις βάσης και Mixed ($\pi/3$) συνοριακές συνθήκες.



Σχήμα 4.7: Σφάλμα της μεθόδου FFX για κανονικά πολύγωνα με ημιτονοειδή συναρτήσεις βάσης και Mixed ($\pi/3$) συνοριακές συνθήκες.

4.3.2 Σφάλμα - Μη Κανονικά Πολύγωνα

irr $n=3$	Dirichlet BC				Neumann BC				Mixed BC ($\pi/3$)			
	SFFS		FFX		SFFS		FFX		SFFS		FFX	
N	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C
4	2.47E-01		3.02E-01		5.97E-01		6.51E-01		4.86E-01		5.52E-01	
8	7.22E-02	1.77	9.77E-02	1.63	3.23E-01	0.89	3.81E-01	0.77	2.21E-01	1.13	2.70E-01	1.03
16	1.38E-02	2.39	1.93E-02	2.34	1.08E-01	1.58	1.49E-01	1.35	6.70E-02	1.72	8.48E-02	1.67
32	2.84E-03	2.28	3.99E-03	2.28	2.95E-02	1.88	5.31E-02	1.49	1.76E-02	1.93	2.84E-02	1.58
64	6.61E-04	2.10	9.31E-04	2.10	7.56E-03	1.97	1.83E-02	1.54	4.46E-03	1.98	9.70E-03	1.55
128	1.63E-04	2.02	2.29E-04	2.02	1.91E-03	1.98	6.12E-03	1.58	1.12E-03	1.99	3.25E-03	1.58
256	4.07E-05	2.00	5.54E-05	2.05	4.79E-04	1.99	2.03E-03	1.59	2.82E-04	1.99	1.08E-03	1.59

Πίνακας 4.6: Sine-collocation: Τρίγωνο.

irr $n=4$	<i>Dirichlet BC</i>				<i>Neumann BC</i>				<i>Mixed BC ($\pi/3$)</i>									
	<i>SFFS</i>		<i>FFX</i>		<i>SFFS</i>		<i>FFX</i>		<i>SFFS</i>		<i>FFX</i>							
	N	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C					
4	2.70E-01			3.33E-01			2.47E-01			3.90E-01			8.00E-02			2.19E-01		
8	1.35E-01	1.00		1.81E-01	0.88		1.02E-01	1.28		1.85E-01	1.08		2.58E-02	1.63		7.57E-02	1.53	
16	4.79E-02	1.50		7.14E-02	1.34		3.10E-02	1.72		6.47E-02	1.51		8.62E-03	1.58		2.39E-02	1.66	
32	1.36E-02	1.81		2.32E-02	1.62		8.22E-03	1.91		2.23E-02	1.54		2.41E-03	1.84		7.75E-03	1.62	
64	3.56E-03	1.94		8.05E-03	1.52		2.10E-03	1.97		7.99E-03	1.48		6.24E-04	1.95		2.74E-03	1.50	
128	9.05E-04	1.98		2.90E-03	1.47		5.28E-04	1.99		2.83E-03	1.50		1.58E-04	1.98		9.50E-04	1.53	
256	2.28E-04	1.99		1.00E-03	1.53		1.33E-04	1.99		9.70E-04	1.54		3.97E-05	1.99		3.23E-04	1.56	

Πίνακας 4.7: Sine-collocation: Τετράπλευρο.

irr $n=5$	<i>Dirichlet BC</i>				<i>Neumann BC</i>				<i>Mixed BC ($\pi/3$)</i>				
	<i>SFFS</i>		<i>FFX</i>		<i>SFFS</i>		<i>FFX</i>		<i>SFFS</i>		<i>FFX</i>		
	N	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C
4	1.03E-01			1.65E-01		4.19E-01		7.01E-01		2.74E-01		4.33E-01	
8	1.73E-02	2.57		3.16E-02	2.39	1.65E-01	1.35	3.15E-01	1.15	9.41E-02	1.54	1.74E-01	1.31
16	4.15E-03	2.06		1.26E-02	1.33	4.77E-02	1.79	1.04E-01	1.59	2.56E-02	1.88	5.44E-02	1.68
32	1.03E-03	2.01		4.16E-03	1.60	1.25E-02	1.93	3.16E-02	1.73	6.57E-03	1.96	1.62E-02	1.75
64	2.57E-04	2.00		1.30E-03	1.68	3.19E-03	1.97	9.40E-03	1.75	1.66E-03	1.98	4.82E-03	1.75
128	6.46E-05	1.99		4.03E-04	1.69	8.03E-04	1.99	3.09E-03	1.60	4.18E-04	1.99	1.74E-03	1.47
256	1.62E-05	2.00		1.25E-04	1.69	2.02E-04	1.99	1.12E-03	1.46	1.05E-04	2.00	6.27E-04	1.47

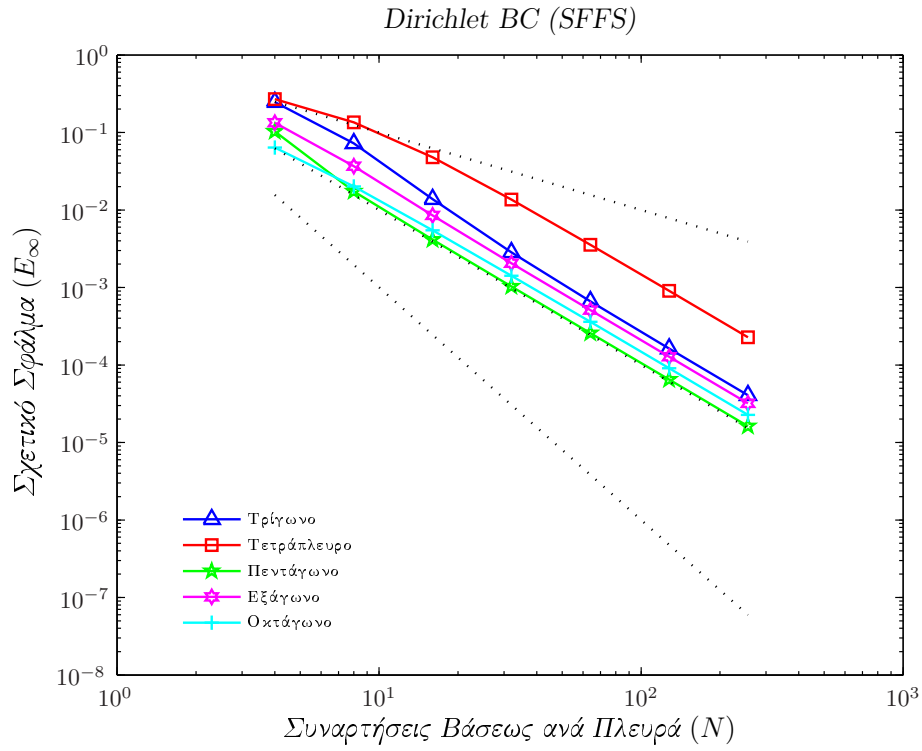
Πίνακας 4.8: Sine-collocation: Πεντάγωνο.

irr $n=6$	$Dirichlet\ BC$				$Neumann\ BC$				$Mixed\ BC\ (\pi/3)$									
	$SFFS$		FFX		$SFFS$		FFX		$SFFS$		FFX							
	N	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C					
4	1.33E-01			3.30E-01			2.16E-01			4.00E-01			2.51E-01			4.96E-01		
8	3.67E-02	1.86		1.17E-01	1.50		9.04E-02	1.26		2.14E-01	0.91		9.56E-02	1.39		2.39E-01	1.06	
16	8.52E-03	2.10		3.47E-02	1.75		2.77E-02	1.71		8.22E-02	1.38		2.80E-02	1.77		8.74E-02	1.45	
32	2.06E-03	2.05		1.07E-02	1.70		7.41E-03	1.90		2.72E-02	1.59		7.37E-03	1.92		2.86E-02	1.61	
64	5.10E-04	2.01		3.36E-03	1.67		1.90E-03	1.97		8.61E-03	1.66		1.88E-03	1.97		9.04E-03	1.66	
128	1.28E-04	2.00		1.09E-03	1.63		4.79E-04	1.99		2.67E-03	1.69		4.74E-04	1.99		2.81E-03	1.68	
256	3.20E-05	2.00		3.64E-04	1.58		1.20E-04	1.99		8.22E-04	1.70		1.19E-04	1.99		8.69E-04	1.69	

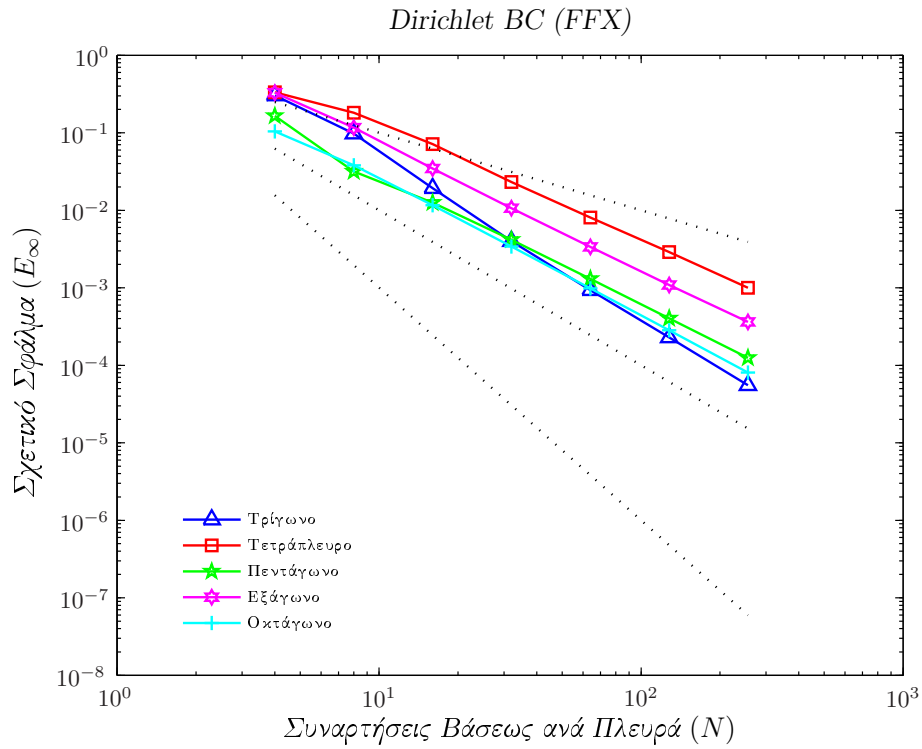
Πίνακας 4.9: Sine-collocation: Εξάγωνο.

irr $n=8$	Dirichlet BC				Neumann BC				Mixed BC ($\pi/3$)				
	SFFS		FFX		SFFS		FFX		SFFS		FFX		
	N	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C	E_∞	O.o.C
4	6.38E-02			1.04E-01		7.23E-02		1.82E-01		2.76E-02		8.27E-02	
8	2.01E-02	1.66		3.79E-02	1.46	1.81E-02	2.00	5.34E-02	1.77	7.86E-03	1.81	2.12E-02	1.97
16	5.48E-03	1.88		1.16E-02	1.71	4.61E-03	1.97	1.57E-02	1.77	2.10E-03	1.91	6.11E-03	1.79
32	1.42E-03	1.95		3.41E-03	1.77	1.17E-03	1.98	4.62E-03	1.76	5.40E-04	1.96	1.82E-03	1.75
64	3.60E-04	1.98		9.85E-04	1.79	2.96E-04	1.98	1.35E-03	1.78	1.37E-04	1.98	5.42E-04	1.75
128	9.07E-05	1.99		2.83E-04	1.80	7.45E-05	1.99	3.91E-04	1.79	3.45E-05	1.99	1.61E-04	1.75
256	2.28E-05	1.99		8.08E-05	1.81	1.87E-05	2.00	1.13E-04	1.79	8.66E-06	1.99	4.78E-05	1.75

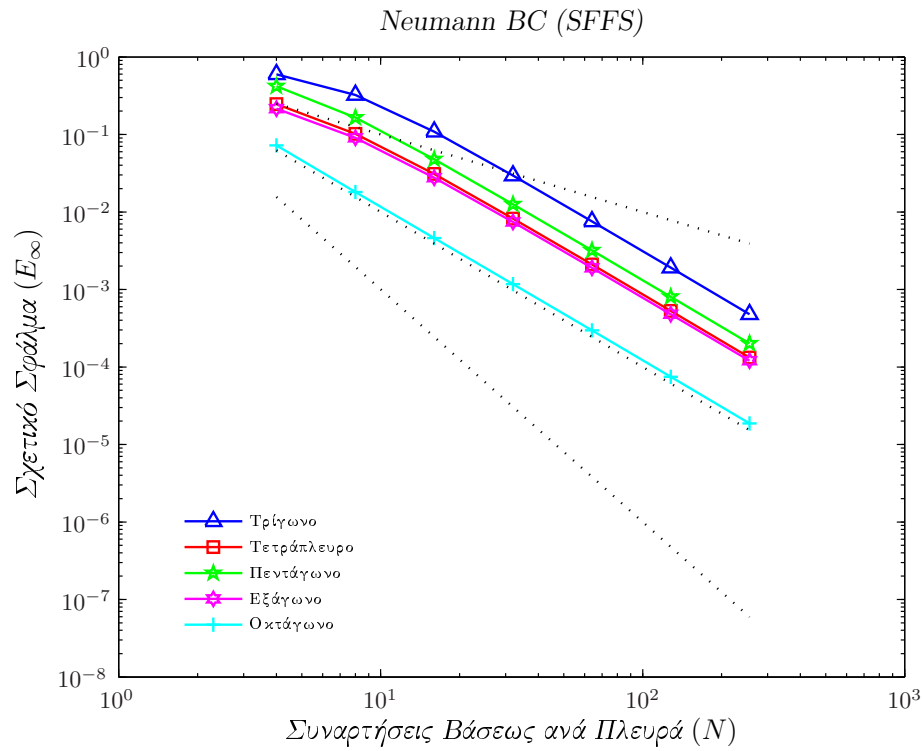
Πίνακας 4.10: Sine-collocation: Οκτάγωνο.



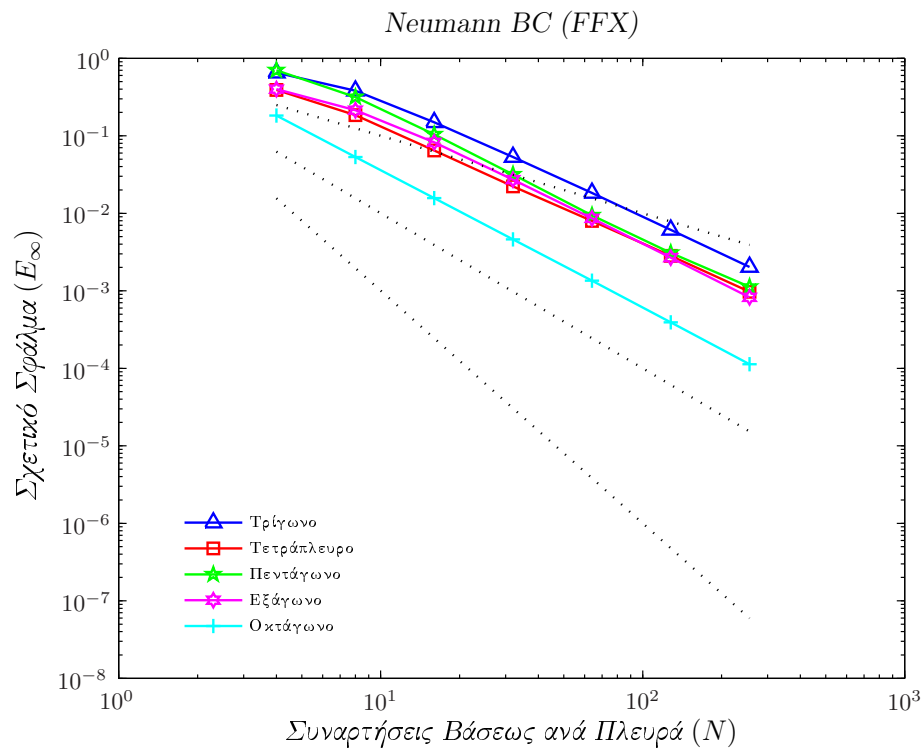
Σχήμα 4.8: Σφάλμα της μεθόδου SFFS για μη κανονικά πολύγωνα με ημιτονοειδή συναρτήσεις βάσης και *Dirichlet* συνοριακές συνθήκες.



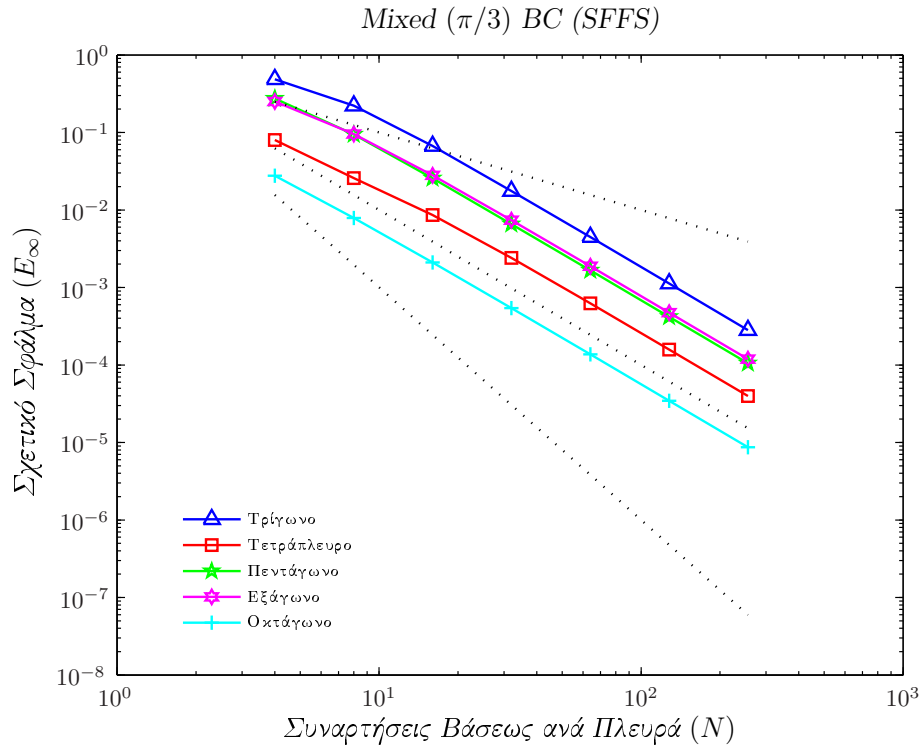
Σχήμα 4.9: Σφάλμα της μεθόδου FFX για μη κανονικά πολύγωνα με ημιτονοειδή συναρτήσεις βάσης και *Dirichlet* συνοριακές συνθήκες.



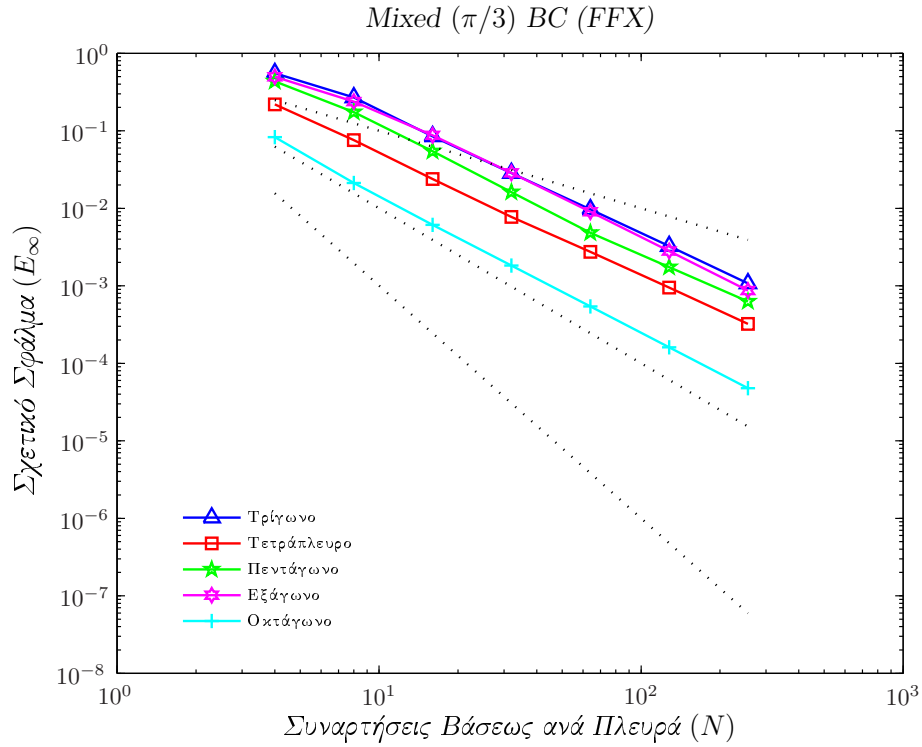
Σχήμα 4.10: Σφάλμα της μεθόδου *SFFS* για μη κανονικά πολύγωνα με ημιτονοειδή συναρτήσεις βάσης και *Neumann* συνοριακές συνθήκες.



Σχήμα 4.11: Σφάλμα της μεθόδου *FFX* για μη κανονικά πολύγωνα με ημιτονοειδή συναρτήσεις βάσης και *Neumann* συνοριακές συνθήκες.



Σχήμα 4.12: Σφάλμα της μεθόδου SFFS για μη κανονικά πολύγωνα με ημιτονοειδή συναρτήσεις βάσης και Mixed ($\pi/3$) συνοριακές συνθήκες.



Σχήμα 4.13: Σφάλμα της μεθόδου FFX για μη κανονικά πολύγωνα με ημιτονοειδή συναρτήσεις βάσης και Mixed ($\pi/3$) συνοριακές συνθήκες.

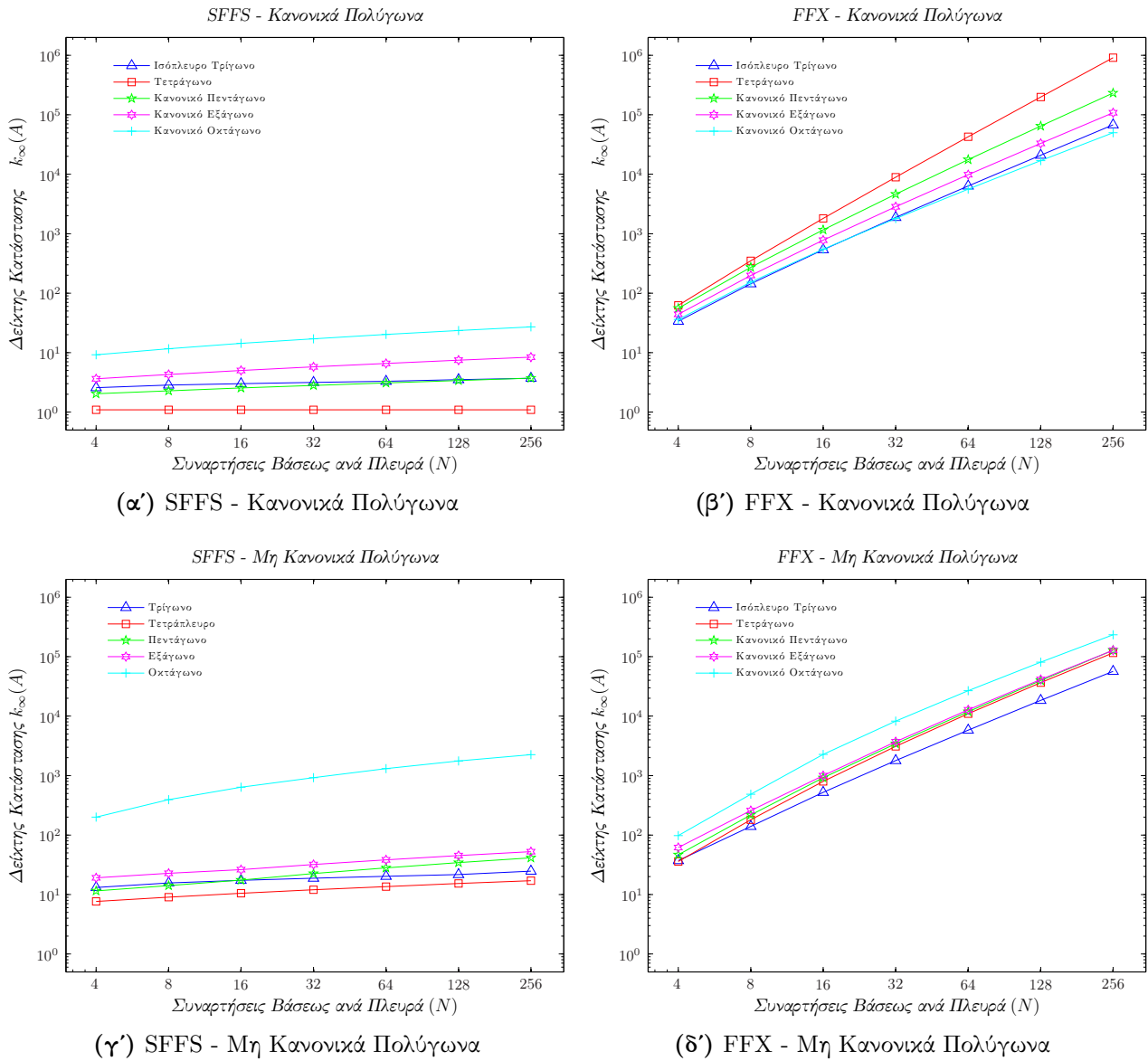
4.3.3 Δείκτης Κατάστασης

Κανονικά Πολύγωνα με Dirichlet, Neumann and Mixed($\pi/3$) BC															
reg	N = 4		N = 8		N = 16		N = 32		N = 64		N = 128		N = 256		
n	SFFS	FFX	SFFS	FFX	SFFS	FFX	SFFS	FFX	SFFS	FFX	SFFS	FFX	SFFS	FFX	
3	2.57	33.80	2.85	143.77	3.00	536.60	3.15	1 879.99	3.30	6 350.58	3.51	20 934.22	3.71	67 755.87	
4	1.09	62.19	1.09	349.88	1.09	1 806.44	1.09	8 909.85	1.09	42 604.39	1.09	198 999.00	1.09	912 320.22	
5	2.04	55.79	2.28	272.01	2.54	1 163.71	2.81	4 626.31	3.10	17 556.28	3.40	64584.31	3.72	232 458.57	
6	3.64	46.99	4.29	198.22	5.00	778.92	5.77	2 837.29	6.59	9 841.61	7.47	33 003.16	8.40	107 984.30	
8	9.19	36.08	11.63	153.24	14.26	548.78	17.12	1 801.57	20.20	5 618.93	23.52	16 933.95	27.07	49 773.99	

Πίνακας 4.11: Ο δείκτης κατάστασης, του πίνακα των συντελεστών του γραμμικού συστήματος, για τα κανονικά πολύγωνα, υπολογισμένος με την εντολή `cond(A, inf)` της *Matlab*.

Μη Κανονικά Πολύγωνα με Dirichlet, Neumann and Mixed($\pi/3$) BC															
irr	N = 4		N = 8		N = 16		N = 32		N = 64		N = 128		N = 256		
n	SFFS	FFX	SFFS	FFX	SFFS	FFX	SFFS	FFX	SFFS	FFX	SFFS	FFX	SFFS	FFX	
3	13.06	37.11	15.65	139.86	17.30	522.74	18.81	1 789.44	20.23	5 818.26	21.73	18 364.32	24.70	56 517.95	
4	7.63	35.91	9.03	180.76	10.46	797.20	11.98	3 104.16	13.60	10 995.07	15.32	36 571.76	17.15	116 490.00	
5	11.46	46.46	14.08	217.68	17.50	919.57	22.50	3 434.11	28.14	11 938.38	34.42	39 698.21	41.35	128 180.63	
6	19.10	61.83	22.80	257.92	26.24	1 000.48	32.04	3 715.31	38.33	12 752.55	45.13	41 306.55	52.47	12 846.52	
8	199.54	98.03	393.38	486.42	637.14	2 268.98	921.64	8 242.33	1 313.77	26 741.55	1 757.73	80 850.83	2 252.60	232 522.37	

Πίνακας 4.12: Ο δείκτης κατάστασης, του πίνακα των συντελεστών του γραμμικού συστήματος, για τα μη κανονικά πολύγωνα, υπολογισμένος με την εντολή `cond(A, inf)` της *Matlab*.



Σχήμα 4.14: Ο δείκτης κατάστασης του πίνακα των συντελεστών του γραμμικού συστήματος ως προς την νόρμα απείρου. Τα γραφήματα (α') και (β') αφορούν στα κανονικά πολύγωνα ενώ τα γραφήματα (γ') και (δ') στα μη κανονικά πολύγωνα.

4.4 Συμπεράσματα

Στα αριθμητικά παραδείγματα του κεφαλαίου φάνηκε ξεκάθαρα ότι η μέθοδος SFFS είναι σαφέστατα καλύτερη μέθοδος από την FFX. Αυτό στηρίζεται στο ότι:

- Παρουσιάζει ομαλότερη σύγκλιση από την FFX με σταθερό τετραγωνικό ρυθμό σύγκλισης σε όλες τις περιπτώσεις.
- Το σφάλμα της είναι πάντα καλύτερο από αυτό της FFX.
- Ο δείκτης κατάστασης του collocation πίνακα, αυξάνει εξαιρετικά πιο αργά από ότι ο αντίστοιχος στην FFX μέθοδο.

Η παραπάνω ανάλυση μας οδηγεί στο συμπέρασμα ότι η μέθοδος SFFS που προτείνουμε, είναι η μέθοδος που πρέπει να χρησιμοποιείται κατά την αριθμητική επίλυση της γενικευμένης Dirichlet-Neumann απεικόνισης.

Κεφάλαιο 5

Αριθμητική Μελέτη και Συμπεριφορά Επαναληπτικών Μεθόδων

Ο collocation πίνακας που αντιστοιχεί στην μέθοδο SFFS, με ημιτονοειδείς συναρτήσεις βάσης, έχει εκ κατασκευής διαγώνια μπλοκ που είναι point διαγώνιοι πίνακες. Αυτό είναι ένα εξαιρετικό συγκριτικό πλεονέκτημα της μεθόδου όσο αφορά στην επαναληπτική επίλυση του collocation συστήματος με Krylov μεθόδους [SPS]. Ο πίνακας είναι ήδη μπλοκ Jacobi προρρυθμισμένος με αποτέλεσμα σημαντική μείωση του χρόνου επίλυσης όχι μόνο στην περίπτωση που χρησιμοποιούμε μπλοκ Jacobi προρύθμιση αλλά και όταν ο προρρυθμιστής μας είναι ο Gauss-Seidel ή ο SGS. Αυτό συμβαίνει διότι (βλ. παράγραφο 1.4) η επίλυση των συστημάτων $Mx = y$ που εμπλέκονται στην προρύθμιση γίνεται πλέον πολύ γρήγορα.

Σε αυτό το κεφάλαιο θα μελετήσουμε την συμπεριφορά των βασικών επαναληπτικών στατικών (stationary) και Krylov μεθόδων για την επίλυση του collocation συστήματος ($Ax = b$). Η μέθοδος επιλογής των collocation σημείων είναι η SFFS και ως συναρτήσεις βάσης θα χρησιμοποιήσουμε τις ημιτονοειδείς. Σε όλα τα πολύγωνα θεωρούμε τις ίδιες συνοριακές συνθήκες (Dirichlet).

Θα επιλύσουμε την εξίσωση του Laplace σε 6 κανονικά πολύγωνα (συγκεκριμένα για πλήθος πλευρών 8, 12, 19, 24, 31 και 36) και σε 5 μη κανονικά πολύγωνα και για πλήθος συναρτήσεων βάσης $N = 8, 16, 32, 64$ για τα κανονικά και $N = 32, 64$ για τα μη κανονικά. Οι κορυφές των κανονικών πολυγώνων που θα χρησιμοποιήσουμε δίνονται από τον τύπο (4.1) ενώ τα μη κανονικά πολύγωνα είναι αυτά που φαίνονται στο σχήμα (4.1). Η ακριβής λύση της διαφορικής εξίσωσης, καθώς και το σχετικό σφάλμα E_∞ ορίζονται όπως στην παράγραφο (4.1). Πριν την εφαρμογή των μεθόδων επίλυσης του collocation συστήματος έχουμε διαιρέσει την κάθε γραμμή του collocation πίνακα με το διαγώνιο στοιχείο της (και αντίστοιχα το δεύτερο μέλος). Έτσι τα διαγώνια μπλοκ του collocation πίνακα είναι ίσα με τον μοναδιαίο πίνακα διάστασης N και ο πίνακας είναι ήδη μπλοκ-Jacobi προρρυθμισμένος. Η προρύθμιση γίνεται κατά μπλοκ όπως αυτά ορίζονται από την διαμέριση (3.21). Για την επαναληπτική μέθοδο Gmres η παράμετρος επαννεκίνησης έχει οριστεί σε 10. Ο μέγιστος αριθμός επαναλήψεων που επιτρέπουμε είναι 200 και η ανοχή του κριτηρίου

τερματισμού (tolerance) σε όλες τις μεθόδους είναι το $1.0E - 06$ ¹. Η αρχική προσέγγιση ισούται με το δεύτερο μέλος του γραμμικού συστήματος (δηλαδή $x^{(0)} = b$). Όλα τα πειράματα εκτελέστηκαν σε ένα πολυχρηστικό SUN V240 με την χρήση ενός επεξεργαστή. Ως γλώσσα προγραμματισμού χρησιμοποιήσαμε την Fortran και όλοι οι υπολογισμοί έγιναν με αριθμητική διπλής ακρίβειας. Παραθέτουμε, εκτός από τις επαναληπτικές μεθόδους, και αποτελέσματα για την άμεση μέθοδο επίλυσης Dgesv² του πακέτου Lapack [LAPACK].

5.1 Κανονικά Πολύγωνα

Κανονικό 8-γωνο												
	N = 8			N = 16			N = 32			N = 64		
	Error_oo	Iter.	Time	Error_oo	Iter.	Time	Error_oo	Iter.	Time	Error_oo	Iter.	Time
Dgesv	1.81E-03		2.62E-03	4.93E-04		2.26E-02	1.28E-04		1.87E-01	3.26E-05		1.83E+00
Jacobi	1.81E-03	40	5.83E-03	4.92E-04	41	1.75E-02	1.28E-04	42	5.68E-02	3.20E-05	42	3.93E-01
GS	1.81E-03	24	3.97E-03	4.92E-04	24	1.04E-02	1.28E-04	24	3.49E-02	3.24E-05	25	2.20E-01
Bicgstab	1.81E-03	6	9.67E-04	4.93E-04	6	1.54E-03	1.29E-04	6	3.43E-03	3.67E-05	6	2.09E-02
Gmres	1.81E-03	10	1.40E-03	4.92E-04	10	2.13E-03	1.28E-04	11	4.81E-03	3.27E-05	12	2.55E-02
Bicgstab2	1.81E-03	4	9.75E-04	4.92E-04	4	1.53E-03	1.28E-04	4	3.53E-03	3.29E-05	4	2.08E-02
CGS	1.81E-03	6	9.63E-04	4.93E-04	7	1.72E-03	1.27E-04	7	3.90E-03	2.96E-05	7	2.56E-02
Tfqmr	1.81E-03	8	1.33E-03	4.93E-04	7	2.20E-03	1.28E-04	8	4.64E-03	3.25E-05	9	2.98E-02
Bicgstab (GS)	1.81E-03	7	2.15E-03	4.93E-04	7	5.09E-03	1.28E-04	7	1.47E-02	3.26E-05	7	1.00E-01
Gmres (GS)	1.81E-03	10	2.13E-03	4.93E-04	10	4.56E-03	1.28E-04	10	1.28E-02	3.26E-05	10	8.28E-02
Bicgstab2 (GS)	1.81E-03	6	2.63E-03	4.93E-04	6	6.05E-03	1.28E-04	6	1.79E-02	3.26E-05	6	1.27E-01
CGS (GS)	1.81E-03	7	2.53E-03	4.93E-04	7	6.30E-03	1.28E-04	7	1.94E-02	3.25E-05	7	1.37E-01
Tfqmr (GS)	1.81E-03	7	3.06E-03	4.93E-04	7	5.18E-03	1.28E-04	7	1.50E-02	3.25E-05	7	1.02E-01
Bicgstab (SGS)	1.81E-03	3	1.91E-03	4.93E-04	3	4.35E-03	1.28E-04	4	1.65E-02	3.26E-05	5	1.26E-01
Gmres (SGS)	1.81E-03	6	2.18E-03	4.93E-04	6	4.84E-03	1.28E-04	7	1.88E-02	3.26E-05	7	1.17E-01
Bicgstab2 (SGS)	1.81E-03	2	2.38E-03	4.93E-04	2	5.91E-03	1.28E-04	2	1.81E-02	3.26E-05	4	1.83E-01
CGS (SGS)	1.81E-03	4	2.66E-03	4.93E-04	4	6.80E-03	1.28E-04	5	2.57E-02	3.26E-05	5	1.74E-01
Tfqmr (SGS)	1.81E-03	4	2.20E-03	4.93E-04	5	5.82E-03	1.28E-04	5	1.84E-02	3.26E-05	5	1.28E-01

Πίνακας T1: Επαναληπτικές μέθοδοι για την επίλυση του *sine-collocation* συστήματος σε ένα κανονικό 8-γωνο. Στον πίνακα φαίνονται το σχετικό σφάλμα (Error_oo), ο αριθμός των επαναλήψεων (Iter.) καθώς και ο χρόνος επίλυσης του γραμμικού συστήματος σε δευτερόλεπτα (Time).

Η συμπεριφορά των επαναληπτικών μεθόδων αλλάζει όσο αλλάζει το πλήθος των πλευρών του πολυγώνου. Έτσι θα θεωρήσουμε τρεις βασικές κατηγορίες πολυγώνων: τα μικρά (μικρό πλήθος πλευρών), τα μέτρια (μέτριο πλήθος πλευρών) και τα μεγάλα (μεγάλο πλήθος πλευρών). Θα αποφύγουμε να ορίσουμε επακριβώς τα όρια μεταξύ αυτών των κατηγοριών επιτρέποντας μια κάποια ρευστότητα.

- Ο αριθμός των επαναλήψεων αυξάνει όσο αυξάνει και το πλήθος των πλευρών του πολυγώνου για όλες τις κατηγορίες των πολυγώνων.

¹ Συζητήσεις περί των διαφόρων κριτηρίων τερματισμού των επαναληπτικών μεθόδων γίνονται στα [SSZ07, TEM].

² Η συγκεκριμένη υπορουτίνα επιλύει το γραμμικό σύστημα χρησιμοποιώντας την διάσπαση $A = LU$ με μερική οδήγηση.

Κανονικό 12-γωνο												
	N = 8			N = 16			N = 32			N = 64		
	Error_oo	Iter.	Time	Error_oo	Iter.	Time	Error_oo	Iter.	Time	Error_oo	Iter.	Time
Dgesv	1.43E-03		9.74E-03	3.94E-04		7.92E-02	1.04E-04		6.30E-01	2.64E-05		9.65E+00
Jacobi	1.49E-03	200	5.63E-02	9.86E-04	200	1.79E-01	9.63E-04	200	8.12E-01	9.61E-04	200	8.08E+00
GS	1.43E-03	87	2.46E-02	3.95E-04	94	8.43E-02	1.04E-04	98	3.91E-01	2.69E-05	100	3.88E+00
Bicgstab	1.43E-03	10	1.50E-03	3.95E-04	10	3.15E-03	1.04E-04	12	1.63E-02	2.72E-05	11	4.97E-01
Gmres	1.43E-03	19	2.55E-03	3.94E-04	21	5.44E-03	1.04E-04	21	2.10E-02	2.68E-05	21	4.95E-01
Bicgstab2	1.43E-03	8	1.54E-03	3.96E-04	8	3.19E-03	1.04E-04	10	1.78E-02	2.64E-05	10	5.41E-01
CGS	1.43E-03	11	1.69E-03	3.94E-04	13	4.05E-03	1.03E-04	12	1.76E-02	2.65E-05	12	5.47E-01
Tfqmr	1.43E-03	12	2.08E-03	3.94E-04	13	4.80E-03	1.04E-04	13	1.84E-02	2.65E-05	13	5.84E-01
Bicgstab (GS)	1.43E-03	13	5.64E-03	3.94E-04	13	1.59E-02	1.03E-04	13	7.81E-02	2.63E-05	13	1.17E+00
Gmres (GS)	1.43E-03	29	7.77E-03	3.95E-04	29	2.19E-02	1.04E-04	30	1.16E-01	2.68E-05	30	1.42E+00
Bicgstab2 (GS)	1.43E-03	12	6.68E-03	3.94E-04	12	1.98E-02	1.04E-04	12	9.53E-02	2.65E-05	12	1.41E+00
CGS (GS)	1.43E-03	14	7.52E-03	3.94E-04	14	2.39E-02	1.04E-04	14	1.08E-01	2.63E-05	14	1.53E+00
Tfqmr (GS)	1.43E-03	14	6.51E-03	3.94E-04	16	2.04E-02	1.04E-04	16	8.91E-02	2.64E-05	16	1.39E+00
Bicgstab (SGS)	1.43E-03	7	5.84E-03	3.94E-04	7	1.62E-02	1.04E-04	7	7.64E-02	2.65E-05	7	9.69E-01
Gmres (SGS)	1.43E-03	12	5.93E-03	3.94E-04	12	1.77E-02	1.04E-04	12	8.04E-02	2.65E-05	13	9.75E-01
Bicgstab2 (SGS)	1.43E-03	6	7.09E-03	3.94E-04	6	2.12E-02	1.04E-04	6	1.03E-01	2.64E-05	6	1.25E+00
CGS (SGS)	1.43E-03	7	7.40E-03	3.94E-04	8	2.58E-02	1.04E-04	8	1.23E-01	2.67E-05	9	1.60E+00
Tfqmr (SGS)	1.43E-03	8	6.34E-03	3.94E-04	8	1.80E-02	1.04E-04	8	9.18E-02	2.65E-05	10	1.36E+00

Πίνακας T2: Επαναληπτικές μέθοδοι για την επίλυση του *sine-collocation* συστήματος σε ένα κανονικό 12-γωνο. Στον πίνακα φαίνονται το σχετικό σφάλμα (*Error_oo*), ο αριθμός των επαναλήψεων (*Iter.*) καθώς και ο χρόνος επίλυσης του γραμμικού συστήματος σε δευτερόλεπτα (*Time*).

Κανονικό 19-γωνο												
	N = 8			N = 16			N = 32			N = 64		
	Error_oo	Iter.	Time	Error_oo	Iter.	Time	Error_oo	Iter.	Time	Error_oo	Iter.	Time
Dgesv	8.41E-04		3.94E-02	2.21E-04		3.14E-01	5.66E-05		3.32E+00	1.42E-05		8.18E+01
Jacobi	x	x	x	x	x	x	x	x	x	x	x	x
GS	8.51E-04	200	1.44E-01	4.89E-04	200	4.37E-01	6.76E-04	200	2.70E+00	8.06E-04	200	4.69E+01
Bicgstab	8.41E-04	11	2.44E-03	2.20E-04	11	6.40E-03	5.88E-05	11	5.52E-02	2.02E-05	11	4.38E+00
Gmres	8.39E-04	20	3.93E-03	2.20E-04	23	1.04E-02	5.60E-05	24	7.09E-02	1.39E-05	24	4.92E+00
Bicgstab2	8.40E-04	8	2.35E-03	2.19E-04	10	7.09E-03	5.43E-05	12	7.08E-02	1.38E-05	12	5.56E+00
CGS	8.40E-04	14	3.03E-03	2.20E-04	14	8.49E-03	5.58E-05	15	7.30E-02	1.47E-05	14	5.52E+00
Tfqmr	8.40E-04	14	3.69E-03	2.20E-04	15	9.86E-03	5.59E-05	16	7.42E-02	1.41E-05	15	5.94E+00
Bicgstab (GS)	8.41E-04	35	3.06E-02	2.18E-04	35	9.29E-02	5.66E-05	36	6.62E-01	1.52E-05	36	2.36E+01
Gmres (GS)	8.38E-04	105	5.72E-02	2.17E-04	130	2.14E-01	5.80E-05	157	1.73E+00	1.96E-05	171	5.81E+01
Bicgstab2 (GS)	8.41E-04	34	3.75E-02	2.21E-04	34	1.16E-01	5.66E-05	34	7.82E-01	1.52E-05	34	2.60E+01
CGS (GS)	8.41E-04	36	4.42E-02	2.21E-04	40	1.55E-01	5.69E-05	40	1.06E+00	1.44E-05	41	3.12E+01
Tfqmr (GS)	8.41E-04	40	3.63E-02	2.21E-04	41	1.16E-01	5.66E-05	41	8.18E-01	1.44E-05	42	2.66E+01
Bicgstab (SGS)	8.41E-04	15	2.44E-02	2.21E-04	16	8.01E-02	5.64E-05	18	6.07E-01	1.44E-05	18	1.67E+01
Gmres (SGS)	8.41E-04	34	3.33E-02	2.21E-04	40	1.17E-01	5.57E-05	52	1.01E+00	1.56E-05	54	2.56E+01
Bicgstab2 (SGS)	8.41E-04	14	3.07E-02	2.21E-04	18	1.21E-01	5.65E-05	18	8.30E-01	1.37E-05	16	1.89E+01
CGS (SGS)	8.41E-04	19	4.42E-02	2.21E-04	20	1.50E-01	5.65E-05	22	1.11E+00	1.42E-05	23	2.61E+01
Tfqmr (SGS)	8.41E-04	19	3.15E-02	2.21E-04	22	1.09E-01	5.66E-05	23	7.93E-01	1.43E-05	24	2.09E+01

Πίνακας T3: Επαναληπτικές μέθοδοι για την επίλυση του *sine-collocation* συστήματος σε ένα κανονικό 19-γωνο. Στον πίνακα φαίνονται το σχετικό σφάλμα (*Error_oo*), ο αριθμός των επαναλήψεων (*Iter.*) καθώς και ο χρόνος επίλυσης του γραμμικού συστήματος σε δευτερόλεπτα (*Time*).

Κανονικό 24-γωνο												
	N = 8			N = 16			N = 32			N = 64		
	Error_oo	Iter.	Time	Error_oo	Iter.	Time	Error_oo	Iter.	Time	Error_oo	Iter.	Time
Dgesv	5.36E-04		7.93E-02	1.40E-04		6.41E-01	3.58E-05		9.89E+00	9.04E-06		1.59E+02
Jacobi	x	x	x	x	x	x	x	x	x	x	x	x
GS	2.05E-03	200	2.03E-01	3.06E-03	200	9.11E-01	3.77E-03	200	9.01E+00	4.15E-03	200	6.45E+01
Bicgstab	5.37E-04	15	4.17E-03	1.41E-04	14	1.77E-02	3.71E-05	15	8.76E-01	9.96E-06	16	8.52E+00
Gmres	5.37E-04	28	6.71E-03	1.41E-04	29	2.50E-02	3.66E-05	31	9.00E-01	9.74E-06	32	8.77E+00
Bicgstab2	5.36E-04	14	4.47E-03	1.41E-04	12	1.75E-02	3.74E-05	12	8.21E-01	9.81E-06	14	8.64E+00
CGS	5.36E-04	15	4.65E-03	1.40E-04	16	2.06E-02	3.62E-05	15	8.35E-01	1.59E-05	15	7.71E+00
Tfqmr	5.36E-04	16	5.60E-03	1.40E-04	17	2.37E-02	3.59E-05	18	1.16E+00	9.30E-06	18	9.24E+00
Bicgstab (GS)	5.36E-04	60	7.71E-02	1.39E-04	63	3.64E-01	4.71E-05	65	7.96E+00	9.79E-06	68	6.38E+01
Gmres (GS)	5.51E-04	200	1.64E-01	2.81E-04	200	6.79E-01	7.12E-04	200	1.07E+01	1.12E-03	200	9.10E+01
Bicgstab2 (GS)	5.35E-04	52	8.41E-02	1.41E-04	64	4.47E-01	3.58E-05	66	9.91E+00	8.80E-06	66	7.14E+01
CGS (GS)	5.36E-04	70	1.34E-01	1.40E-04	72	5.84E-01	3.58E-05	74	9.50E+00	9.03E-06	77	7.93E+01
Tfqmr (GS)	5.36E-04	70	9.85E-02	1.40E-04	72	4.38E-01	3.58E-05	76	7.98E+00	9.04E-06	79	6.72E+01
Bicgstab (SGS)	5.39E-04	21	5.02E-02	1.40E-04	25	2.60E-01	3.47E-05	24	4.12E+00	1.15E-05	25	3.21E+01
Gmres (SGS)	5.37E-04	50	7.34E-02	1.43E-04	78	4.64E-01	3.90E-05	102	8.21E+00	2.30E-05	114	7.29E+01
Bicgstab2 (SGS)	5.36E-04	20	6.30E-02	1.40E-04	22	3.00E-01	3.60E-05	24	5.25E+00	9.07E-06	30	4.80E+01
CGS (SGS)	5.36E-04	32	1.15E-01	1.40E-04	33	5.13E-01	3.58E-05	33	7.15E+00	9.19E-06	34	5.34E+01
Tfqmr (SGS)	5.36E-04	32	8.09E-02	1.40E-04	33	3.63E-01	3.58E-05	33	5.48E+00	9.05E-06	35	4.21E+01

Πίνακας T4: Επαναληπτικές μέθοδοι για την επίλυση του *sine-collocation* συστήματος σε ένα κανονικό 24-γωνο. Στον πίνακα φαίνονται το σχετικό σφάλμα (*Error_oo*), ο αριθμός των επαναλήψεων (*Iter.*) καθώς και ο χρόνος επίλυσης του γραμμικού συστήματος σε δευτερόλεπτα (*Time*).

Κανονικό 31-γωνο												
	N = 8			N = 16			N = 32			N = 64		
	Error_oo	Iter.	Time	Error_oo	Iter.	Time	Error_oo	Iter.	Time	Error_oo	Iter.	Time
Dgesv	3.53E-04		1.72E-01	9.03E-05		1.50E+00	2.28E-05		3.39E+01	5.73E-06		4.47E+02
Jacobi	x	x	x	x	x	x	x	x	x	x	x	x
GS	1.26E-03	200	3.57E-01	1.41E-03	200	1.87E+00	1.52E-03	200	3.07E+01	1.58E-03	200	1.33E+02
Bicgstab	3.53E-04	23	8.29E-03	9.03E-05	27	7.57E-02	2.30E-05	24	5.42E+00	5.61E-06	31	3.25E+01
Gmres	3.54E-04	91	2.79E-02	9.09E-05	89	1.60E-01	2.34E-05	85	1.13E+01	6.27E-06	87	4.98E+01
Bicgstab2	3.53E-04	22	8.83E-03	9.04E-05	24	7.20E-02	2.30E-05	24	5.74E+00	5.81E-06	24	2.73E+01
CGS	3.53E-04	29	1.06E-02	9.04E-05	30	8.90E-02	2.28E-05	40	9.01E+00	5.86E-06	32	3.33E+01
Tfqmr	3.53E-04	31	1.42E-02	9.03E-05	35	9.77E-02	2.28E-05	41	9.11E+00	5.73E-06	43	4.59E+01
Bicgstab (GS)	3.53E-04	98	2.07E-01	9.12E-05	106	1.27E+00	5.27E-05	109	3.91E+01	1.13E-04	114	2.01E+02
Gmres (GS)	5.46E-04	200	2.62E-01	7.38E-04	200	1.40E+00	1.02E-03	200	4.31E+01	1.20E-03	200	1.90E+02
Bicgstab2 (GS)	3.53E-04	100	2.65E-01	9.09E-05	112	1.53E+00	2.56E-05	112	4.50E+01	2.22E-05	128	2.60E+02
CGS (GS)	3.53E-04	120	3.72E-01	9.07E-05	133	2.18E+00	2.47E-05	140	6.19E+01	5.70E-06	195	4.05E+02
Tfqmr (GS)	3.53E-04	121	2.61E-01	9.03E-05	164	2.02E+00	2.27E-05	181	6.52E+01	5.73E-06	200	3.47E+02
Bicgstab (SGS)	3.54E-04	45	1.73E-01	9.01E-05	52	1.11E+00	2.69E-05	56	2.81E+01	5.64E-06	67	1.65E+02
Gmres (SGS)	3.57E-04	128	2.90E-01	9.65E-05	177	2.21E+00	3.00E-05	178	5.44E+01	6.58E-04	200	2.66E+02
Bicgstab2 (SGS)	3.53E-04	48	2.38E-01	9.04E-05	52	1.37E+00	2.28E-05	64	3.77E+01	5.72E-06	66	1.92E+02
CGS (SGS)	3.53E-04	59	3.38E-01	9.03E-05	65	2.09E+00	2.28E-05	69	4.66E+01	5.73E-06	71	2.21E+02
Tfqmr (SGS)	3.53E-04	54	2.12E-01	9.03E-05	65	1.43E+00	2.28E-05	69	3.61E+01	5.73E-06	72	1.75E+02

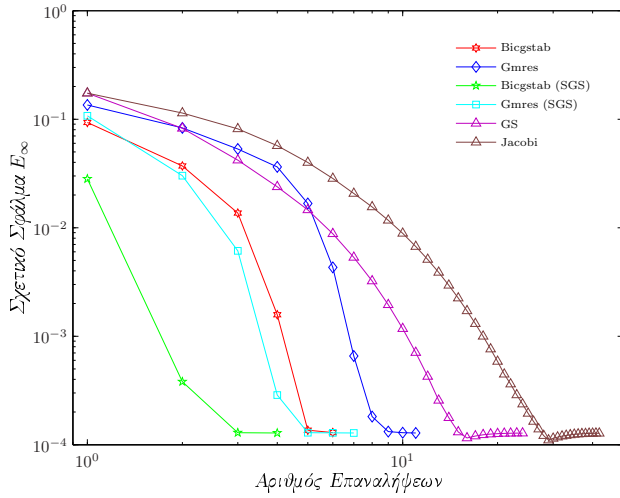
Πίνακας T5: Επαναληπτικές μέθοδοι για την επίλυση του *sine-collocation* συστήματος σε ένα κανονικό 31-γωνο. Στον πίνακα φαίνονται το σχετικό σφάλμα (*Error_oo*), ο αριθμός των επαναλήψεων (*Iter.*) καθώς και ο χρόνος επίλυσης του γραμμικού συστήματος σε δευτερόλεπτα (*Time*).

Κανονικό 36-γωνο												
	N = 8			N = 16			N = 32			N = 64		
	Error_oo	Iter.	Time	Error_oo	Iter.	Time	Error_oo	Iter.	Time	Error_oo	Iter.	Time
Dgesv	2.85E-04		2.66E-01	7.12E-05		2.54E+00	1.77E-05		6.63E+01	4.41E-06		6.30E+02
Jacobi	x	x	x	x	x	x	x	x	x	x	x	x
GS	2.70E-03	200	4.82E-01	3.37E-03	200	2.64E+00	3.77E-03	200	4.17E+01	3.98E-03	200	1.54E+02
Bicgstab	2.85E-04	34	1.51E-02	7.13E-05	39	1.49E-01	1.77E-05	38	1.25E+01	4.50E-06	37	4.37E+01
Gmres	2.87E-04	200	7.31E-02	2.08E-04	200	4.71E-01	1.42E-04	200	3.58E+01	1.89E-04	200	1.26E+02
Bicgstab2	2.85E-04	30	1.43E-02	7.12E-05	36	1.59E-01	1.77E-05	38	1.35E+01	4.62E-06	34	4.33E+01
CGS	x	x	x	x	x	x	x	x	x	x	x	x
Tfqmr	5.49E-03	200	1.02E-01	5.64E-03	79	3.38E-01	5.19E-03	200	6.35E+01	5.93E-03	200	1.73E+02
Bicgstab (GS)	2.85E-04	132	3.68E-01	7.51E-05	136	2.35E+00	2.38E-05	148	8.53E+01	3.72E-05	148	3.18E+02
Gmres (GS)	3.50E-04	200	3.46E-01	4.62E-04	200	2.03E+00	6.35E-04	200	5.87E+01	7.61E-04	200	2.15E+02
Bicgstab2 (GS)	2.86E-04	144	4.79E-01	7.22E-05	148	3.04E+00	2.17E-05	154	1.01E+02	1.41E-05	160	3.90E+02
CGS (GS)	2.86E-04	170	7.06E-01	7.15E-05	181	4.38E+00	1.86E-05	200	1.31E+02	1.89E-05	200	4.95E+02
Tfqmr (GS)	2.85E-04	200	5.65E-01	7.19E-05	200	3.43E+00	7.07E-05	200	1.09E+02	1.02E-04	200	4.11E+02
Bicgstab (SGS)	2.85E-04	68	3.55E-01	7.66E-05	75	2.32E+00	1.87E-05	92	7.21E+01	4.02E-05	94	2.79E+02
Gmres (SGS)	2.93E-04	200	6.16E-01	8.15E-05	200	3.53E+00	1.10E-03	200	8.22E+01	3.76E-03	200	3.04E+02
Bicgstab2 (SGS)	2.86E-04	68	4.29E-01	7.11E-05	76	2.96E+00	2.03E-05	90	8.27E+01	4.86E-06	96	3.35E+02
CGS (SGS)	2.85E-04	76	5.88E-01	7.12E-05	90	4.11E+00	1.77E-05	95	9.15E+01	4.42E-06	105	3.82E+02
Tfqmr (SGS)	2.85E-04	77	3.89E-01	7.12E-05	91	2.79E+00	1.77E-05	97	7.25E+01	4.41E-06	101	2.81E+02

Πίνακας T6: Επαναληπτικές μέθοδοι για την επίλυση του *sine-collocation* συστήματος σε ένα κανονικό 36-γωνο. Στον πίνακα φαίνονται το σχετικό σφάλμα (*Error_oo*), ο αριθμός των επαναλήψεων (*Iter.*) καθώς και ο χρόνος επίλυσης του γραμμικού συστήματος σε δευτερόλεπτα (*Time*).

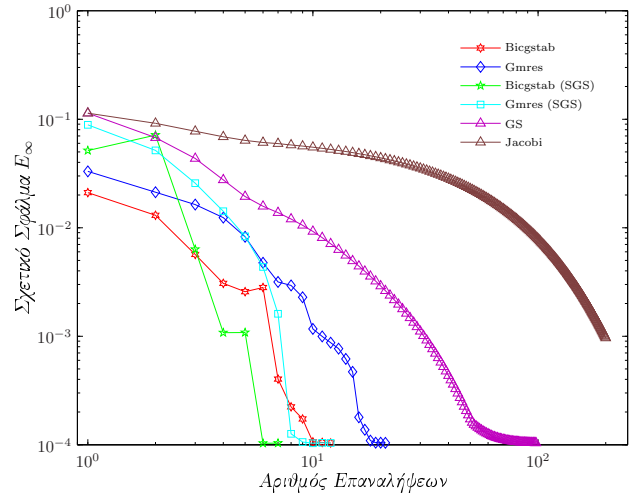
- Η μέθοδος Jacobi συγκλίνει για μικρά πολύγωνα ενώ αποκλίνει για μέτρια και μεγάλα πολύγωνα (παρατηρήστε και τα σχήματα (5.2β', 5.3)).
- Η μέθοδος GS συγκλίνει σε όλες τις περιπτώσεις αλλά για μέτρια και μεγάλα πολύγωνα συγκλίνει πολύ αργά. Είναι αξιοπερίεργο ότι η φασματική ακτίνα του επαναληπτικού πίνακα της GS αν και αυξάνει όσο αυξάνει το πλήθος των πλευρών του πολυγώνου ποτέ δεν γίνεται 1 αλλά πλησιάζει ασυμπτωτικά την μονάδα, όπως φαίνεται και από το σχήμα (5.3).
- Οι μέθοδοι Bicgstab και Bicgstab2 συγκλίνουν σε όλες τις περιπτώσεις σχετικά γρήγορα. Μάλιστα οι δύο αυτές μέθοδοι έχουν σχεδόν την ίδια συμπεριφορά. Στα γραφήματα (5.1) βλέπουμε ότι η μη προρρυθμισμένη Bicgstab συγκλίνει ομαλά για όλα τα πολύγωνα.
- Η μέθοδοι Gmres, CGS και Tfqmr συγκλίνουν για μικρά και μέτρια πολύγωνα αλλά σε μεγάλα συγκλίνουν πολύ αργά. Μάλιστα η CGS αποκλίνει για το 36-γωνο.
- Η προρύθμιση με SGS βελτιώνει το ρυθμό σύγκλισης περισσότερο από ότι η προρύθμιση με GS.
- Για μικρά πολύγωνα η προρύθμιση με SGS βελτιώνει το ρυθμό σύγκλισης των μεθόδων, για μέτρια πολύγωνα οι μη προρρυθμισμένες μέθοδοι συγκλίνουν ταχύτερα ενώ για μεγάλα πολύγωνα η προρύθμιση με SGS καταφέρνει να βελτιώσει μόνο την μέθοδο Tfqmr.
- Η μέθοδος GS είναι ταχύτερη από την άμεση για μεγάλα N αλλά για μέτρια και μεγάλα πολύγωνα δεν συγκλίνει στην επιθυμητή λύση.

Μείωση του Σχετικού Σφάλματος σε ένα Κανονικό 8-γωνο



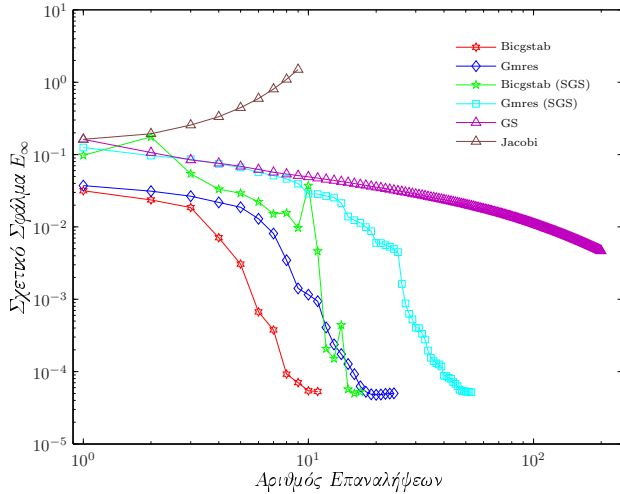
(α') Κανονικό 8-γωνο.

Μείωση του Σχετικού Σφάλματος σε ένα Κανονικό 12-γωνο



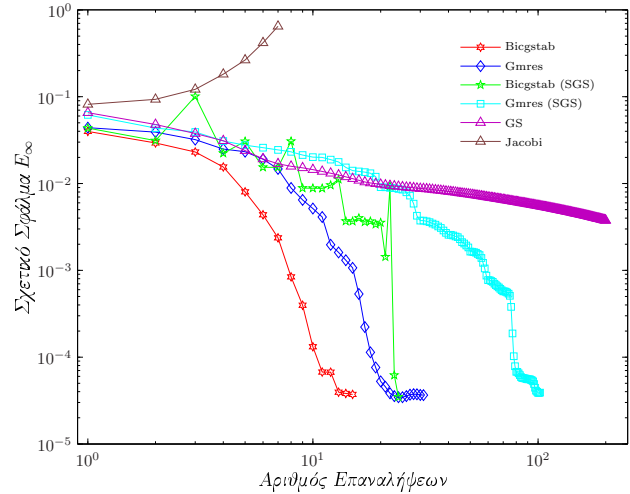
(β') Κανονικό 12-γωνο.

Μείωση του Σχετικού Σφάλματος σε ένα Κανονικό 19-γωνο



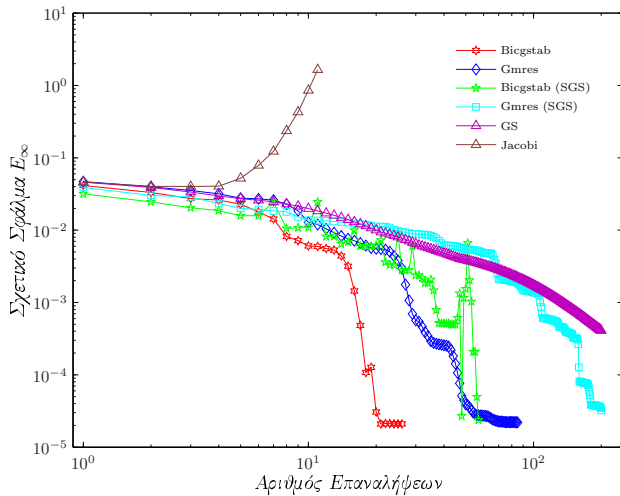
(γ') Κανονικό 19-γωνο.

Μείωση του Σχετικού Σφάλματος σε ένα Κανονικό 24-γωνο



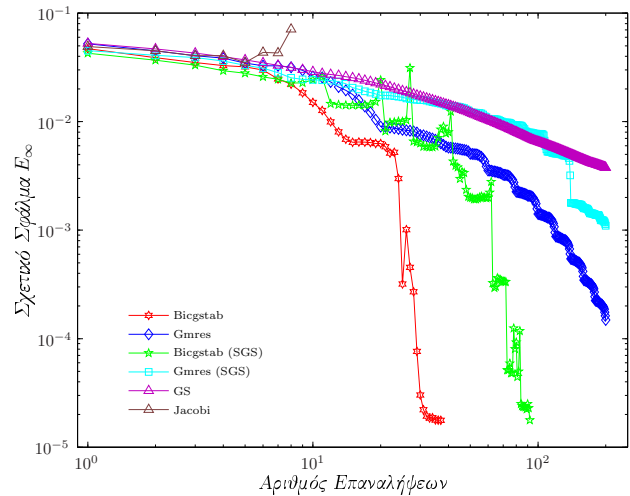
(δ') Κανονικό 24-γωνο.

Μείωση του Σχετικού Σφάλματος σε ένα Κανονικό 31-γωνο



(ε') Κανονικό 31-γωνο.

Μείωση του Σχετικού Σφάλματος σε ένα Κανονικό 36-γωνο



(ς') Κανονικό 36-γωνο.

Σχήμα 5.1: Η μείωση του σχετικού σφάλματος E_{∞} , κατά την επίλυση του collocation γραμμικού συστήματος σε κανονικά πολύγωνα, συναρτήσει των επαναλήψεων για τις επαναληπτικές μεθόδους Jacobi, GS, Bicgstab, Gmres, Bicgstab(SGS) και Gmres(SGS). Το πλήθος των συναρτήσεων βάσης είναι 32 ανά πλευρά.

Κανονικά Πολύγωνα - Ο Αριθμός των Επαναλήψεων																								
N -->	8-γωνο				12-γωνο				19-γωνο				24-γωνο				31-γωνο				36-γωνο			
	8	16	32	64	8	16	32	64	8	16	32	64	8	16	32	64	8	16	32	64	8	16	32	64
Jacobi	40	41	42		200	200	200	200	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
GS	24	24	24	25	87	94	98	100	200	200	200	200	200	200	200	200	200	200	200	200	200	200	200	200
Bicgstab	6	6	6	6	10	10	12	11	11	11	11	11	15	14	15	16	23	27	24	31	34	39	38	37
Gmres	10	10	11	12	19	21	21	21	20	23	24	24	28	29	31	32	91	89	85	87	200	200	200	200
Bicgstab2	4	4	4	4	8	8	10	10	8	10	12	12	14	12	12	14	22	24	24	24	30	36	38	34
CGS	6	7	7	7	11	13	12	12	14	14	15	14	15	16	15	15	29	30	40	32	x	x	x	x
Tfqmr	8	7	8	9	12	13	13	13	14	15	16	15	16	17	18	18	31	35	41	43	200	79	200	200
Bicgstab (GS)	7	7	7	7	13	13	13	13	35	35	36	36	60	63	65	68	98	106	109	114	132	136	148	148
Gmres (GS)	10	10	10	10	29	29	30	30	105	130	157	171	200	200	200	200	200	200	200	200	200	200	200	200
Bicgstab2 (GS)	6	6	6	6	12	12	12	12	34	34	34	34	52	64	66	66	100	112	112	128	144	148	154	160
CGS (GS)	7	7	7	7	14	14	14	14	36	40	40	41	70	72	74	77	120	133	140	195	170	181	200	200
Tfqmr (GS)	7	7	7	7	14	16	16	16	40	41	41	42	70	72	76	79	121	164	181	200	200	200	200	200
Bicgstab (SGS)	3	3	4	5	7	7	7	7	15	16	18	18	21	25	24	25	45	52	56	67	68	75	92	94
Gmres (SGS)	6	6	7	7	12	12	12	13	34	40	52	54	50	78	102	114	128	177	178	200	200	200	200	200
Bicgstab2 (SGS)	2	2	2	4	6	6	6	6	14	18	18	16	20	22	24	30	48	52	64	66	68	76	90	96
CGS (SGS)	4	4	5	5	7	8	8	9	19	20	22	23	32	33	33	34	59	65	69	71	76	90	95	105
Tfqmr (SGS)	4	5	5	5	8	8	8	10	19	22	23	24	32	33	33	35	54	65	69	72	77	91	97	101

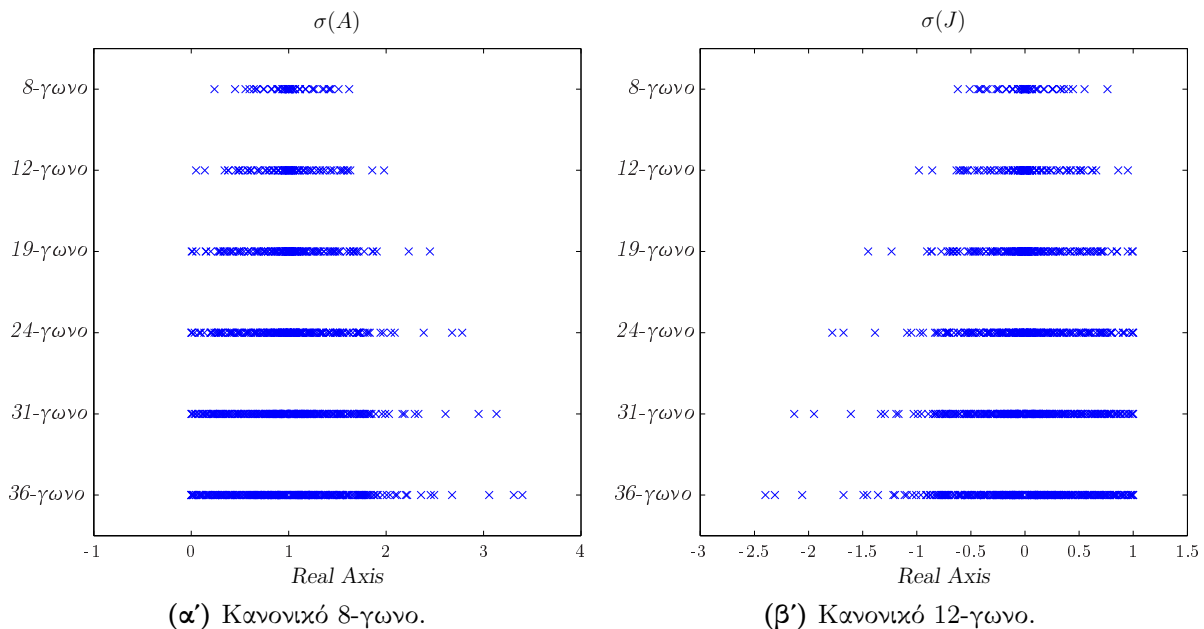
Πίνακας T7: Ο αριθμός των επαναλήψεων για τις επαναληπτικές μεθόδους σε όλα τα κανονικά πολύγωνα.

Κανονικά Πολύγωνα - Ο Χρόνος Επίλυσης των Επαναληπτικών Μεθόδων προς το Χρόνο Επίλυσης της Άμεσης Μεθόδου (%)																								
N -->	8-γωνο				12-γωνο				19-γωνο				24-γωνο				31-γωνο				36-γωνο			
	8	16	32	64	8	16	32	64	8	16	32	64	8	16	32	64	8	16	32	64	8	16	32	64
Jacobi	223	77	30	21	578	226	129	84	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
GS	152	46	19	12	253	106	62	40	365	139	81	57	256	142	91	41	208	125	91	30	181	104	63	24
Bicgstab	37	7	2	1	15	4	3	5	6	2	2	5	5	3	9	5	5	5	16	7	6	6	19	7
Gmres	53	9	3	1	26	7	3	5	10	3	2	6	8	4	9	6	16	11	33	11	27	19	54	20
Bicgstab2	37	7	2	1	16	4	3	6	6	2	2	7	6	3	8	5	5	5	17	6	5	6	20	7
CGS	37	8	2	1	17	5	3	6	8	3	2	7	6	3	8	5	6	6	27	7	x	x	x	x
Tfqmr	51	10	2	2	21	6	3	6	9	3	2	7	7	4	12	6	8	7	27	10	38	13	96	27
Bicgstab (GS)	82	23	8	5	58	20	12	12	78	30	20	29	97	57	80	40	120	85	115	45	138	93	129	50
Gmres (GS)	81	20	7	5	80	28	18	15	145	68	52	71	207	106	108	57	152	93	127	43	130	80	89	34
Bicgstab2 (GS)	100	27	10	7	69	25	15	15	95	37	24	32	106	70	100	45	154	102	133	58	180	120	152	62
CGS (GS)	97	28	10	7	77	30	17	16	112	49	32	38	169	91	96	50	216	145	183	91	265	172	198	79
Tfqmr (GS)	117	23	8	6	67	26	14	14	92	37	25	33	124	68	81	42	152	135	192	78	212	135	164	65
Bicgstab (SGS)	73	19	9	7	60	20	12	10	62	26	18	20	63	41	42	20	101	74	83	37	133	91	109	44
Gmres (SGS)	83	21	10	6	61	22	13	10	85	37	30	31	93	72	83	46	169	147	160	60	232	139	124	48
Bicgstab2 (SGS)	91	26	10	10	73	27	16	13	78	39	25	23	79	47	53	30	138	91	111	43	161	117	125	53
CGS (SGS)	102	30	14	10	76	33	20	17	112	48	33	32	145	80	72	34	197	139	137	49	221	162	138	61
Tfqmr (SGS)	84	26	10	7	65	23	15	14	80	35	24	26	102	57	55	26	123	95	106	39	146	110	109	45

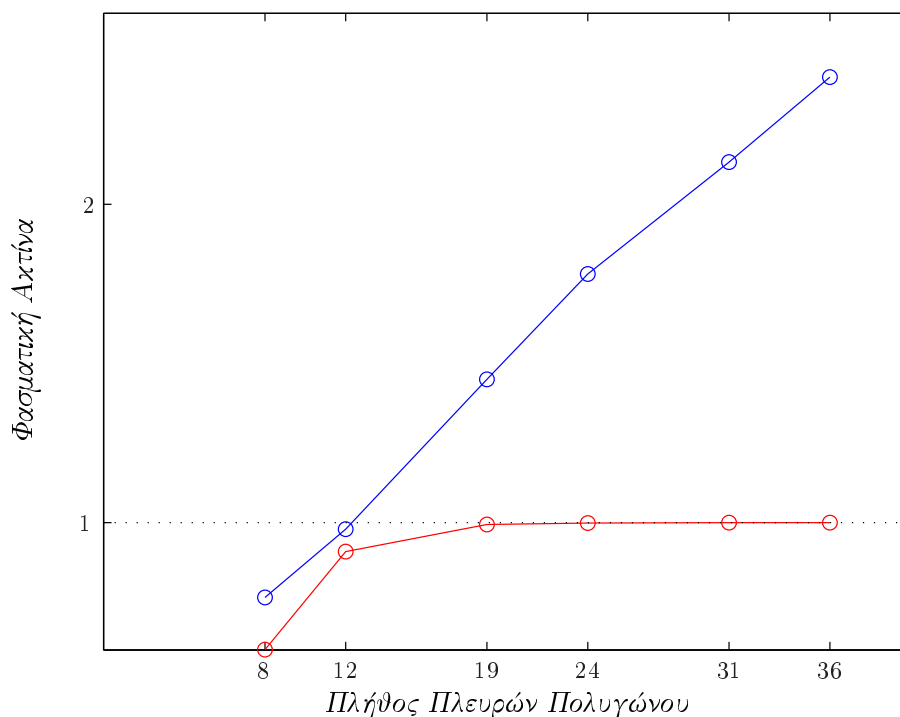
Πίνακας T8: Ο χρόνος επίλυσης των επαναληπτικών μεθόδων προς τον χρόνο επίλυσης της άμεσης μεθόδου ως ποσοστό επί τοις εκατό.

Κανονικά Πολύγωνα - Οι πιο Γρήγορες και οι πιο Αργές Μέθοδοι																											
	8-γωνο				12-γωνο				19-γωνο				24-γωνο				31-γωνο				36-γωνο						
N -->	8	16	32	64	8	16	32	64	8	16	32	64	8	16	32	64	8	16	32	64	8	16	32	64			
Dgesv	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100		
Jacobi	223	77	30	21	578	226	129	84	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		
GS	152	46	19	12	253	106	62	40	365	139	81	57	256	142	91	41	208	125	91	30	181	104	63	24			
Bicgstab	37	7	2	1	15	4	3	5	6	2	2	5	5	3	9	5	5	5	16	7	6	6	6	19	7		
Gmres	53	9	3	1	26	7	3	5	10	3	2	6	8	4	9	6	16	11	33	11	27	19	54	20			
Bicgstab2	37	7	2	1	16	4	3	6	6	2	2	7	6	3	8	5	5	5	17	6	5	6	20	7			
CGS	37	8	2	1	17	5	3	6	8	3	2	7	6	3	8	5	6	6	27	7	x	x	x	x			
Tfqmr	76	12	3	2	33	8	4	6	13	4	2	7	9	5	12	6	10	8	27	10	49	14	95	28			
Bicgstab (GS)	82	23	8	5	58	20	12	12	78	30	20	29	97	57	80	40	120	85	115	45	138	93	129	50			
Gmres (GS)	81	20	7	5	80	28	18	15	145	68	52	71	207	106	108	57	152	93	127	43	130	80	89	34			
Bicgstab2 (GS)	100	27	10	7	69	25	15	15	95	37	24	32	106	70	100	45	154	102	133	58	180	120	152	62			
CGS (GS)	97	28	10	7	77	30	17	16	112	49	32	38	169	91	96	50	216	145	183	91	265	172	198	79			
Tfqmr (GS)	117	26	9	6	81	28	16	15	133	40	25	33	129	72	81	43	161	137	191	77	231	139	164	65			
Bicgstab (SGS)	73	19	9	7	60	20	12	10	62	26	18	20	63	41	42	20	101	74	83	37	133	91	109	44			
Gmres (SGS)	83	21	10	6	61	22	13	10	85	37	30	31	93	72	83	46	169	147	160	60	232	139	124	48			
Bicgstab2 (SGS)	91	26	10	10	73	27	16	13	78	39	25	23	79	47	53	30	138	91	111	43	161	117	125	53			
CGS (SGS)	102	30	14	10	76	33	20	17	112	48	33	32	145	80	72	34	197	139	137	49	221	162	138	61			
Tfqmr (SGS)	100	29	10	8	73	24	15	14	86	38	25	26	101	60	56	27	129	97	106	39	159	114	110	45			

Πίνακας T9: Οι πιο γρήγορες και οι πιο αργές μέθοδοι για τα κανονικά πολύγωνα. Τα νούμερα που αναγράφονται είναι οι χρόνοι επίλυσης των επαναληπτικών μεθόδων προς τον χρόνο επίλυσης της άμεσης μεθόδου ως ποσοστό επί τοις εκατό. Έχουμε επιτρέψει μια ανοχή +4% σε ότι αφορά το ποιά είναι η ταχύτερη μέθοδος.



Σχήμα 5.2: Το φάσμα του collocation πίνακα (αριστερά) και του επαναληπτικού πίνακα του Jacobi (δεξιά) για $N = 16$ και για όλα τα κανονικά πολύγωνα που χρησιμοποιήσαμε. Παρατηρήστε ότι όλες οι ιδιοτιμές είναι πραγματικές.



Σχήμα 5.3: Οι φασματικές ακτίνες των επαναληπτικών πινάκων των μεθόδων Jacobi (μπλέ) και GS (κόκκινο) ως συνάρτηση του πλήθους των πλευρών του πολυγώνου. Στην περίπτωση της GS η φασματική ακτίνα πλησιάζει ασυμπτωτικά την ευθεία $y = 1$.

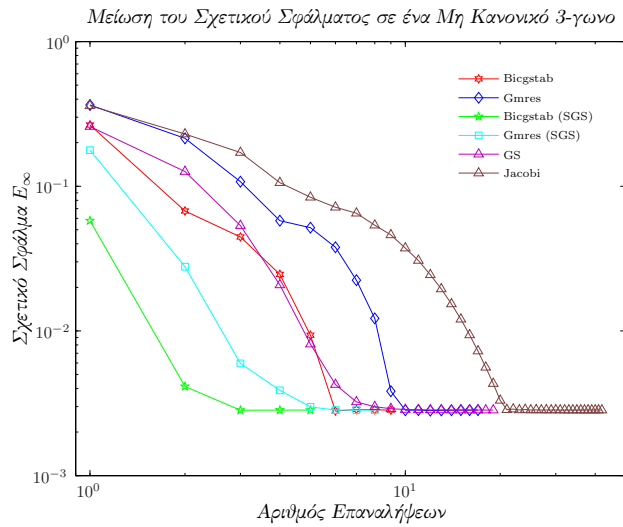
- Για μικρά πολύγωνα αν και η προρύθμιση με SGS βελτιώνει το ρυθμό σύγκλισης δεν οδηγεί σε ταχύτερες μεθόδους.
- Για όλα τα πολύγωνα και για $N = 64$ η άμεση μέθοδος είναι η πιο αργή. Στις υπόλοιπες περιπτώσεις, για μικρά πολύγωνα οι πιο αργές μέθοδοι είναι η άμεση και η Jacobi ενώ για μέτρια πολύγωνα και για μικρό πλήθος συναρτήσεων βάσης ($N = 8, 16$) η πιο αργή μέθοδος είναι η GS. Για το 24-γωνο και για $N = 32$ η πιο αργή μέθοδος είναι η Gmres με GS προρύθμιση. Για το 31-γωνο και πλήθος συναρτήσεων βάσης $N = 32$ η πιο αργή μέθοδος είναι η Tfqr με GS προρύθμιση. Για το 36-γωνο και για $N = 16, 32, 64$ η πιο αργή μέθοδος είναι η CGS με GS προρύθμιση.
- Για όλα τα πολύγωνα οι πιο γρήγορες μέθοδοι είναι οι επαναληπτικές χωρίς προρύθμιση. Για μεγάλα πολύγωνα πιο γρήγορες είναι οι Bicgstab και Bicgstab2.
- Για το 36-γωνο οι μόνες επαναληπτικές μέθοδοι που συγκλίνουν στην ακριβή λύση υπό τον περιορισμό των 200 επαναλήψεων είναι οι Bicgstab και Bicgstab2.

5.2 Μη Κανονικά Πολύγωνα

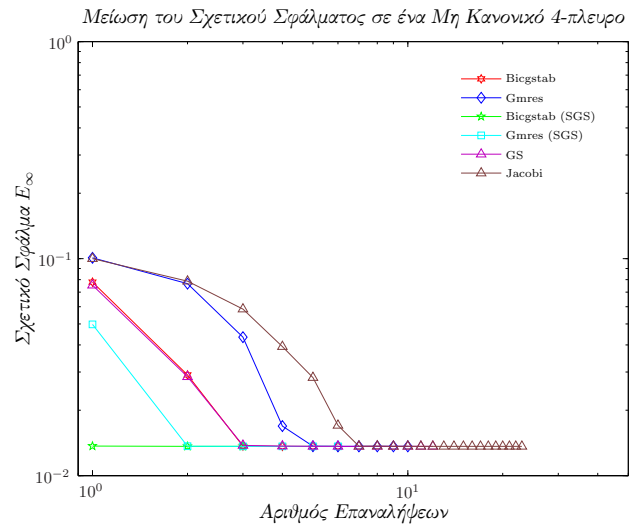
Μελετήσαμε, ως προς την σύγκλιση, τις επαναληπτικές μεθόδους Jacobi, GS, Biscstab, Gmres, Bicgstab(SGS) και Gmres(SGS) για μή κανονικά πολύγωνα με πλήθος πλευρών 3,4,5,6 και 8. Υπενθυμίζουμε ότι ο collocation πίνακας της μεθόδου sine-SFFS έχει διαγώνια μπλοκ που είναι point διαγώνιοι πίνακες και στην περίπτωση των μη κανονικών πολυγώνων.

Ο πίνακας (T10) δείχνει των αριθμό των επαναλήψεων και τον χρόνο επίλυσης του collocation συστήματος για πλήθος συναρτήσεων βάσης $N = 32$ και $N = 64$. Στο σχήμα (5.2) παρουσιάζουμε την μείωση του σχετικού σφάλματος ως προς των αριθμό των επαναλήψεων για $N = 32$. Μελετώντας τα αποτελέσματα μπορούμε να συμπεράνουμε τα εξής:

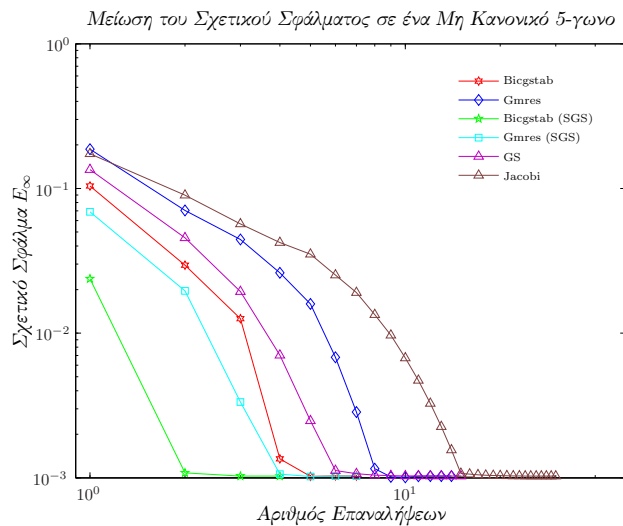
- Όλες οι μέθοδοι συγκλίνουν ομαλά σε όλα τα πολύγωνα.
- Ο αριθμός των επαναλήψεων φαίνεται να είναι ανεξάρτητος από το πλήθος των συναρτήσεων βάσης σε όλες τις μεθόδους και για όλα τα πολύγωνα.
- Η μέθοδος Jacobi συγκλίνει σε όλα τα πολύγωνα αν και για το 8-γωνο δεν καταφέρνει επιτύχει το επιθυμητό σφάλμα υπό τον περιορισμό των 200 επαναλήψεων.
- Η μέθοδος GS φαίνεται να ακολουθεί την συμπεριφορά της Jacobi απαιτώντας όμως τις μισές επαναλήψεις από την Jacobi. Η GS συγκλίνει, στο επιθυμητό σφάλμα, και στο 8-γωνο.
- Όλες οι Krylov μέθοδοι συγκλίνουν πιο γρήγορα από τις stationary.
- Η Gmres απαιτεί σχεδόν τις διπλάσιες επαναλήψεις από την Bicgstab για να επιτύχει το ίδιο σφάλμα. Το ίδιο ισχύει και για τα προρρυθμισμένα ζευγάρια τους.
- Η προρύθμιση βελτιώνει σημαντικά την σύγκλιση. Οι προρρυθμισμένες μέθοδοι συγκλίνουν στις μισές επαναλήψεις από ότι οι μη προρρυθμισμένες.
- Ο χρόνος των Krylov επαναληπτικών είναι μικρότερος από τον χρόνο της άμεσης μεθόδου ακόμα και στις μικρές περιπτώσεις.
- Αν και η προρύθμιση βελτιώνει τον ρυθμό σύγκλισης αυξάνει πάντα το υπολογιστικό κόστος.
- Η ταχύτερη μέθοδος, σε όλες τις περιπτώσεις, είναι η μη προρρυθμισμένη Bicgstab.



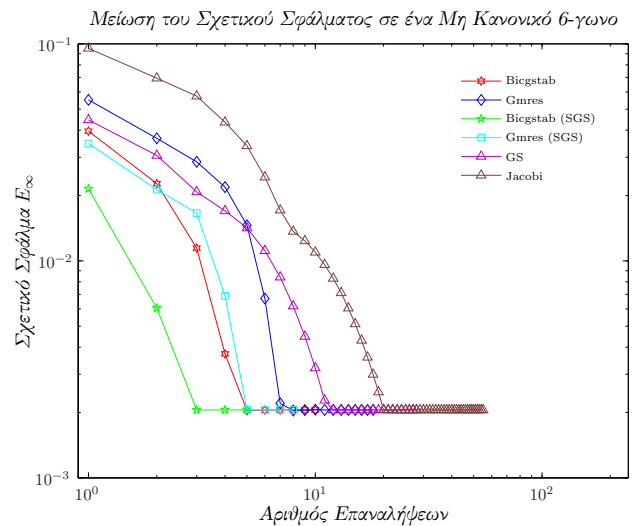
(α') Μη κανονικό 3-γωνο.



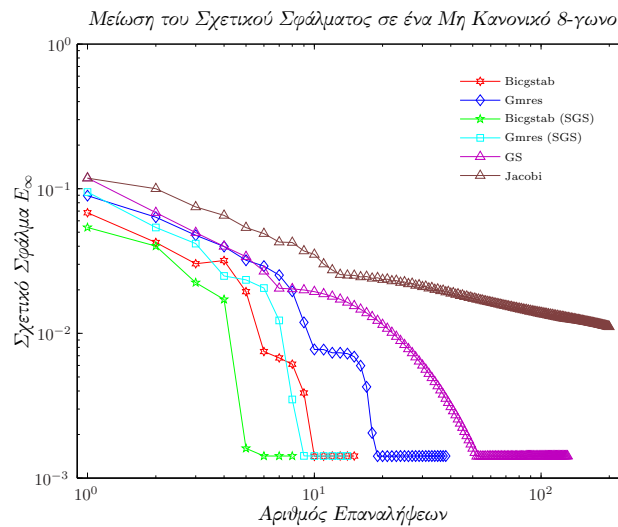
(β') Μη κανονικό 4-πλευρο.



(γ') Μη κανονικό 5-γωνο.



(δ') Μη κανονικό 6-γωνο.



(ε') Μη κανονικό 8-γωνο.

Σχήμα 5.4: Η μείωση του σχετικού σφάλματος E_∞ , κατά την επίλυση του *collocation* γραμμικού συστήματος σε μη κανονικά πολύγωνα, συναρτήσει των επαναλήψεων για τις επαναληπτικές μεθόδους *Jacobi*, *GS*, *Bicgstab*, *Gmres*, *Bicgstab(SGS)* και *Gmres(SGS)*. Το πλήθος των συναρτήσεων βάσης είναι 32 ανά πλευρά.

Μη Κανονικά Πολύγωνα										
	Ο Αριθμός των Επαναλήψεων					Ο Χρόνος Επίλυσης του Συστήματος				
	3-γωνο	4-πλευρο	5-γωνο	6-γωνο	8-γωνο	3-γωνο	4-πλευρο	5-γωνο	6-γωνο	8-γωνο
N=32										
Dgesv						9.16E-03	2.22E-02	4.36E-02	7.83E-02	1.86E-01
Jacobi	42	23	30	55	200	7.70E-03	7.87E-03	1.60E-02	4.03E-02	2.61E-01
GS	19	12	15	28	131	4.11E-03	4.78E-03	8.67E-03	2.12E-02	1.69E-01
Bicgstab	9	6	7	10	15	1.53E-03	1.54E-03	2.04E-03	3.05E-03	6.62E-03
Gmres	17	10	14	18	38	2.31E-03	2.29E-03	3.09E-03	4.65E-03	1.36E-02
Bicgstab (SGS)	5	3	4	5	8	3.19E-03	3.80E-03	6.49E-03	1.09E-02	2.81E-02
Gmres (SGS)	8	6	7	8	14	3.23E-03	3.90E-03	6.48E-03	9.85E-03	2.76E-02
N=64										
Dgesv						7.51E-02	1.81E-01	3.79E-01	6.11E-01	1.49E+00
Jacobi	42	23	30	56	200	2.48E-02	2.67E-02	5.52E-02	1.92E-01	2.29E+00
GS	19	12	16	29	139	1.22E-02	1.57E-02	3.16E-02	1.03E-01	1.59E+00
Bicgstab	9	6	7	10	16	3.15E-03	3.52E-03	5.84E-03	1.34E-02	1.29E-01
Gmres	17	10	14	18	40	4.57E-03	4.56E-03	8.26E-03	1.65E-02	1.94E-01
Bicgstab (SGS)	5	3	4	5	9	9.40E-03	1.19E-02	2.52E-02	5.99E-02	3.28E-01
Gmres (SGS)	8	6	7	8	17	8.67E-03	1.24E-02	2.29E-02	4.91E-02	3.25E-01

Πίνακας T10: Ο αριθμός των επαναλήψεων (αριστερά) και ο χρόνος επίλυσης του *collocation* συστήματος (δεξιά) για μη κανονικά πολύγωνα και για τις επαναληπτικές μεθόδους *Jacobi*, *GS*, *Bicgstab*, *Gmres*, *Bicgstab(SGS)* και *Gmres(SGS)*. Έχουμε χρησιμοποιήσει 32 (πάνω) και 64 (κάτω) συναρτήσεις βάσης ανά πλευρά.

5.3 Συμπεράσματα

Το τελικό συμπέρασμα που βγαίνει από την προηγηθείσα ανάλυση είναι ότι η μη προρυθμισμένη Bicgstab είναι η ασφαλέστερη και πιο αποδοτική μέθοδος για την επίλυση της γενικευμένης Dirichlet to Neumann απεικόνισης σε κανονικά πολύγωνα με τις ίδιες Dirichlet, Neumann ή Mixed συνοριακές συνθήκες σε όλες τις πλευρές.

Στην περίπτωση των μη κανονικών πολυγώνων, αν και παρατηρούμε παρόμοια συμπεριφορά με τα κανονικά πολύγωνα, είναι πρόωρο να βγάλουμε ένα γενικό συμπέρασμα παρόμοιο με αυτό των κανονικών πολυγώνων. Οι ποικίλες μορφές ενός n -γώνου, καθώς και οι διαφορετικοί τρόποι προσέγγισής του, είναι δυο από τους πολλούς παράγοντες που συνηγορούν σε αυτή την κατεύθυνση. Σίγουρα το θέμα αυτό θα μας απασχολήσει στο μέλλον. Πάντως, για την κλάση των πολυγώνων που εξετάσαμε μπορούμε να πούμε ότι η μη προρυθμισμένη Bicgstab είναι η πιο αποδοτική μέθοδος για την επίλυση της γενικευμένης Dirichlet-Neumann απεικόνισης. Σε επέκταση αυτού του συμπεράσματος, μπορούμε να ισχυριστούμε ότι, σε περιπτώσεις μη κανονικών πολυγώνων που έχουν προέλθει από μικρές διαταραχές στην γεωμετρία των κανονικών πολυγώνων, η συμπεριφορά σύγκλισης των επαναληπτικών μεθόδων θα είναι παρόμοια με αυτή που παρατηρήσαμε στα κανονικά πολύγωνα.

Κεφάλαιο 6

Δομή και Ιδιότητες του Collocation Πίνακα

Γνωρίζουμε ότι ο collocation πίνακας της sine-SFFS έχει μια ιδιαίτερη δομή: τα διαγώνια του μπλοκ είναι point διαγώνιοι πίνακες [SFFS]. Αυτή η δομή έχει προέλθει εκ' κατασκευής της SFFS μεθόδου. Όμως η γενική δομή του collocation πίνακα εξαρτάται άμεσα από (α) την γεωμετρία του πολυγώνου (β) τις συνοριακές συνθήκες που χρησιμοποιούμε σε κάθε πλευρά και (γ) τις συναρτήσεις βάσης. Αλλάζοντας αυτές τις παραμέτρους αλλάζει, εν γένει, και η δομή του πίνακα [SSP, SFFS].

Στο παρόν κεφάλαιο θα μελετήσουμε διεξοδικά την δομή και τις ιδιότητες του collocation πίνακα.

6.1 Κανονικά Πολύγωνα με Ίδιες Συνοριακές Συνθήκες σε Όλες τις Πλευρές

Σε αυτή την παράγραφο θα μελετήσουμε την δομή και τις ιδιότητες του collocation πίνακα στην περίπτωση στην οποία το πολυγωνικό χωρίο $\mathcal{D}$, πάνω στο οποίο λύνουμε την διαφορική εξίσωση, είναι ένα κανονικό πολύγωνο και οι συνοριακές συνθήκες είναι ίδιες ($\beta_j = \beta$) σε όλες τις πλευρές του πολυγώνου.

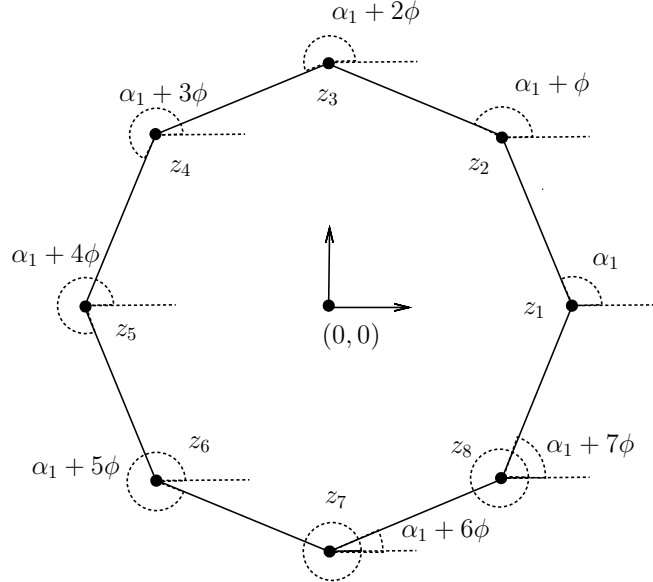
Χωρίς βλάβη της γενικότητας θα θεωρήσουμε ότι το κέντρο του πολυγώνου βρίσκεται στην αρχή των αξόνων και ότι η κορυφή z_1 βρίσκεται πάνω στον xx' άξονα¹. Επειδή το πολύγωνο είναι κανονικό θα ισχύει ότι $\alpha_{j+1} = \alpha_j + \phi$ όπου $\alpha_j = \arg(z_{j+1} - z_j)$ και ϕ μια γωνία που εξαρτάται, αποκλειστικά, από το πλήθος των πλευρών του πολυγώνου και ισούται με $2\pi/n$. Η γεωμετρία των κανονικών πολυγώνων για την περίπτωση ενός κανονικού οκταγώνου φαίνεται στο σχήμα (6.1). Θα παρακολουθήσουμε τους συμβολισμούς του 1ου κεφαλαίου ορίζοντας επιπλέον τις εξής

¹ Σε επόμενη παράγραφο θα αποσαφηνίσουμε το γιατί αυτοί οι δυο περιορισμοί δεν περιορίζουν την γενικότητα.

ποσότητες:

$$M := |m_j| \quad \text{και} \quad H := |h_j|,$$

οι οποίες είναι ανεξάρτητες του j .



Σχήμα 6.1: Γεωμετρία ενός κανονικού 8-γώνου.

Το μπλοκ A_p^j (3.21) του collocation πίνακα για την SFFS μέθοδο και για οποιοδήποτε σύνολο συναρτήσεων βάσης $\varphi_r(s)$ θα είναι το

$$A_p^j = \begin{bmatrix} \operatorname{Re} [c_p^j(\frac{1}{2}, 1)] & \operatorname{Re} [c_p^j(\frac{1}{2}, 2)] & \operatorname{Re} [c_p^j(\frac{1}{2}, 3)] & \operatorname{Re} [c_p^j(\frac{1}{2}, 4)] & \operatorname{Re} [c_p^j(\frac{1}{2}, 5)] & \cdots \\ \operatorname{Im} [c_p^j(1, 1)] & \operatorname{Im} [c_p^j(1, 2)] & \operatorname{Im} [c_p^j(1, 3)] & \operatorname{Im} [c_p^j(1, 4)] & \operatorname{Im} [c_p^j(1, 5)] & \cdots \\ \operatorname{Re} [c_p^j(\frac{3}{2}, 1)] & \operatorname{Re} [c_p^j(\frac{3}{2}, 2)] & \operatorname{Re} [c_p^j(\frac{3}{2}, 3)] & \operatorname{Re} [c_p^j(\frac{3}{2}, 4)] & \operatorname{Re} [c_p^j(\frac{3}{2}, 5)] & \cdots \\ \operatorname{Im} [c_p^j(2, 1)] & \operatorname{Im} [c_p^j(2, 2)] & \operatorname{Im} [c_p^j(2, 3)] & \operatorname{Im} [c_p^j(2, 4)] & \operatorname{Im} [c_p^j(2, 5)] & \cdots \\ \operatorname{Re} [c_p^j(\frac{5}{2}, 1)] & \operatorname{Re} [c_p^j(\frac{5}{2}, 2)] & \operatorname{Re} [c_p^j(\frac{5}{2}, 3)] & \operatorname{Re} [c_p^j(\frac{5}{2}, 4)] & \operatorname{Re} [c_p^j(\frac{5}{2}, 5)] & \cdots \\ \operatorname{Im} [c_p^j(3, 1)] & \operatorname{Im} [c_p^j(3, 2)] & \operatorname{Im} [c_p^j(3, 3)] & \operatorname{Im} [c_p^j(3, 4)] & \operatorname{Im} [c_p^j(3, 5)] & \cdots \\ \vdots & & & & & \ddots \end{bmatrix} \quad (6.1)$$

όπου

$$c_p^j(l, r) := \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-il \frac{m_p - m_j}{h_p}} \int_{-\pi}^{\pi} e^{il \frac{h_j}{h_p} s} \varphi_r(s) ds. \quad (6.2)$$

Στην (6.2) οι παράγοντες $|h_j| / |h_p|$ και $e^{i(\beta_j - \beta_p)}$ μπορούν να απαλειφθούν αφού $|h_j| = |h_p|$ (ίσο μήκος πλευρών) και $\beta_j = \beta_p$ (ίδιες συνοριακές συνθήκες σε όλες τις πλευρές) αντίστοιχα. Επομένως η σχέση (6.2) ισοδύναμα γράφεται

$$c_p^j(l, r) := e^{-il \frac{m_p - m_j}{h_p}} \int_{-\pi}^{\pi} e^{il \frac{h_j}{h_p} s} \varphi_r(s) ds. \quad (6.3)$$

Λήμμα 6.1 Έστω κανονικό πολύγωνο με n κορυφές και κέντρο του την αρχή των αξόνων. Τότε ισχύουν τα παρακάτω

(α)

$$\arg(m_j) = \alpha_j - \frac{\pi}{2} \quad (6.4)$$

(β)

$$\frac{M}{H} = \cot\left(\frac{\pi}{n}\right) \pi \quad (6.5)$$

(γ)

$$\frac{m_p - m_j}{h_p} = -i\pi \cot\left(\frac{\pi}{n}\right) (1 - e^{i(\alpha_j - \alpha_p)}). \quad (6.6)$$

Απόδειξη.

(α) Το διάνυσμα m_j είναι κάθετο, προς τα έξω, στην πλευρά S_j (σχήμα 6.2). Άρα

$$\arg(m_j) = \arg(h_j) - \frac{\pi}{2} = \alpha_j - \frac{\pi}{2}$$

(β) Θέτοντας $\omega = \frac{2\pi}{2n}$ όπως στο σχήμα (6.3) έχουμε

$$\tan(\omega) = \frac{\pi H}{M} \quad \Leftrightarrow$$

$$\frac{\tan(\omega)}{\pi} = \frac{H}{M} \quad \Leftrightarrow$$

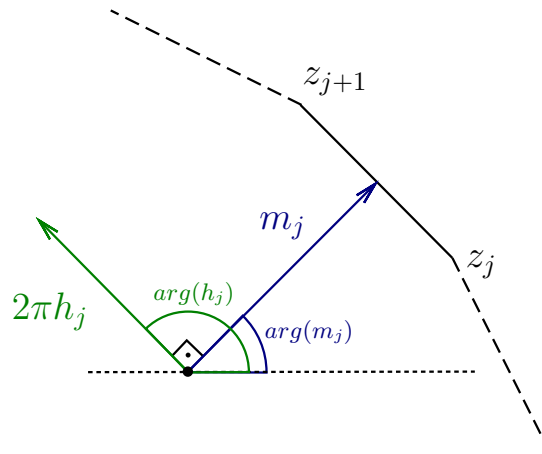
$$\frac{M}{H} = \cot(\omega) \pi$$

Δηλαδή

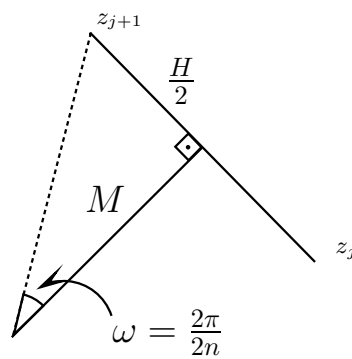
$$\frac{M}{H} = \cot\left(\frac{\pi}{n}\right) \pi.$$

(γ) Χρησιμοποιώντας τα (α) και (β) καθώς και τη σχέση $h_p = |h_p|e^{i\alpha_p}$ έχουμε

$$\begin{aligned} \frac{m_p - m_j}{h_p} &= \frac{M e^{i(\alpha_p - \frac{\pi}{2})} - M e^{i(\alpha_j - \frac{\pi}{2})}}{H e^{i\alpha_p}} \\ &= \frac{M e^{i(-\pi/2)} e^{i\alpha_p} - e^{i(-\pi/2)} e^{i\alpha_j}}{H e^{i\alpha_p}} \\ &= -i \frac{M}{H} \frac{e^{i\alpha_p} - e^{i\alpha_j}}{e^{i\alpha_p}} \\ &= -i \frac{M}{H} (1 - e^{i(\alpha_j - \alpha_p)}) \\ &= -i\pi \cot\left(\frac{\pi}{n}\right) (1 - e^{i(\alpha_j - \alpha_p)}) \end{aligned}$$



Σχήμα 6.2: Κανονικά πολύγωνα: $\arg(m_j) = \alpha_j - \pi/2$.



Σχήμα 6.3: Κανονικά πολύγωνα: $\frac{M}{H} = \cot\left(\frac{\pi}{n}\right) \pi$.

□

Με την χρήση του Λήμματος 6.1 η σχέση (6.3) διαδοχικά γράφεται:

$$\begin{aligned}
 c_p^j(l, r) &= e^{-il \frac{m_p - m_j}{h_p}} \int_{-\pi}^{\pi} e^{il \frac{h_j}{h_p} s} \varphi_r(s) ds \\
 &= e^{-il \left[-i\pi \cot\left(\frac{\pi}{n}\right) \quad 1 - e^{i(\alpha_j - \alpha_p)} \right]} \int_{-\pi}^{\pi} e^{ile^{i(\alpha_j - \alpha_p)} s} \varphi_r(s) ds \\
 &= e^{-l\pi \cot\left(\frac{\pi}{n}\right)} e^{l\pi \cot\left(\frac{\pi}{n}\right)} e^{i(\alpha_j - \alpha_p)} \int_{-\pi}^{\pi} e^{ile^{i(\alpha_j - \alpha_p)} s} \varphi_r(s) ds \\
 &= e^{-l\pi \cot\left(\frac{\pi}{n}\right)} \int_{-\pi}^{\pi} e^{l\pi \cot\left(\frac{\pi}{n}\right)} e^{i(\alpha_j - \alpha_p)} e^{ile^{i(\alpha_j - \alpha_p)} s} \varphi_r(s) ds.
 \end{aligned}$$

Δηλαδή

$$c_p^j(l, r) = e^{-l\pi \cot\left(\frac{\pi}{n}\right)} \int_{-\pi}^{\pi} e^{[\pi \cot\left(\frac{\pi}{n}\right) + is]} l e^{i(\alpha_j - \alpha_p)} \varphi_r(s) ds. \quad (6.7)$$

Όμως

$$\alpha_j - \alpha_p = [\alpha_1 + (j - 1)\phi] - [\alpha_1 + (p - 1)\phi] = (j - p)\phi.$$

Άρα

$$c_p^j = e^{-l\pi \cot\left(\frac{\pi}{n}\right)} \int_{-\pi}^{\pi} e^{[\pi \cot\left(\frac{\pi}{n}\right) + is]} l e^{i(j-p)\phi} \varphi_r(s) ds. \quad (6.8)$$

Θεώρημα 6.1 Στην περίπτωση των κανονικών πολυγώνων με ίδιες συνοριακές συνθήκες σε όλες τις πλευρές ο *collocation* πίνακας A της *SFFS* είναι *block-Toeplitz* για οποιοδήποτε σύνολο συναρτήσεων βάσης.

Απόδειξη. Αρκεί να δείξουμε ότι για κάθε τετράδα φυσικών αριθμών p_1, p_2, j_1, j_2 μεταξύ του 1 και του n με $p_1 - j_1 = p_2 - j_2$ να ισχύει, για τα μπλοκ του πίνακα A , ότι:

$$A_{p_1}^{j_1} = A_{p_2}^{j_2}. \quad (6.9)$$

Τα στοιχεία του μπλοκ A_p^j δίνονται από το πραγματικό ή το φανταστικό μέρος των ποσοτήτων $c_p^j(l, r)$, όπως αυτές ορίζονται στην (6.8). Επομένως αρκεί να δείξουμε ότι για κάθε l, r

$$c_{p_1}^{j_1}(l, r) = c_{p_2}^{j_2}(l, r),$$

ή ισοδύναμα

$$e^{-l\pi \cot(\frac{\pi}{n})} \int_{-\pi}^{\pi} e^{[\pi \cot(\frac{\pi}{n}) + is] l e^{i(j_1 - p_1)\phi}} \varphi_r(s) ds = e^{-l\pi \cot(\frac{\pi}{n})} \int_{-\pi}^{\pi} e^{[\pi \cot(\frac{\pi}{n}) + is] l e^{i(j_2 - p_2)\phi}} \varphi_r(s) ds$$

που προφανώς ισχύει αφού $j_1 - p_1 = j_2 - p_2$. □

Θεώρημα 6.2 Στην περίπτωση των κανονικών πολυγώνων με ίδιες συνοριακές συνθήκες σε όλες τις πλευρές ο collocation πίνακας A της SFPS είναι block-circulant για οποιοδήποτε σύνολο συναρτήσεων βάσης.

Απόδειξη. Σύμφωνα με το θεώρημα (6.1) ο πίνακας A είναι block-Toeplitz. Επομένως ικανή συνθήκη για να είναι και block-circulant είναι η

$$A_j^1 = A_1^{n-j+2},$$

για κάθε $j = 2, 3, \dots, n$. Αρκεί να δείξω λοιπόν ότι για κάθε l, r

$$c_j^1(l, r) = c_1^{n-j+2}(l, r),$$

ή ισοδύναμα

$$e^{-l\pi \cot(\frac{\pi}{n})} \int_{-\pi}^{\pi} e^{[\pi \cot(\frac{\pi}{n}) + is] l e^{i(1-j)\phi}} \varphi_r(s) ds = e^{-l\pi \cot(\frac{\pi}{n})} \int_{-\pi}^{\pi} e^{[\pi \cot(\frac{\pi}{n}) + is] l e^{i(n-j+2-1)\phi}} \varphi_r(s) ds.$$

Αρκεί λοιπόν να δείξω ότι

$$e^{i(1-j)\phi} = e^{i(n-j+1)\phi} \quad \text{ή} \quad e^{in\phi} = 1.$$

Έστω ω η κάθε εσωτερική γωνία του πολυγώνου. Παρατηρώντας τώρα τις γωνίες στην κορυφή z_2 του πολυγώνου στο σχήμα (6.1) και έχοντας υπόψιν ότι το άθροισμα των εσωτερικών γωνιών ενός n -γώνου ισούται με $(n-2)\pi$ έχουμε

$$\begin{aligned} \omega + \phi &= \pi \Rightarrow \\ n\omega + n\phi &= n\pi \Rightarrow \\ (n-2)\pi + n\phi &= n\pi \Rightarrow \\ n\phi &= 2\pi. \end{aligned}$$

Επομένως $e^{in\phi} = 1$. □

Τα επόμενα λήμματα θα μας φανεί χρήσιμα

Λήμμα 6.2 Μελέτη της $\sin r \frac{s+\pi}{2}$

Η συνάρτηση $\sin r \frac{s+\pi}{2}$ είναι άρτια αν το r είναι περιττό και περιττή αν το r είναι άρτιο.

Απόδειξη. Θα πάρουμε περιπτώσεις για το r .

- Έστω ότι $r = 1, 3, 5, \dots$ ($r = 2k + 1$, $k = 0, 1, 2, \dots$)

$$\sin \left(r \frac{s+\pi}{2} \right) = \sin \left[(2k+1) \frac{s+\pi}{2} \right] = \sin \left(ks + k\pi + \frac{1}{2}s + \frac{1}{2}\pi \right) = \cos \left(ks + k\pi + \frac{1}{2}s \right) =$$

$$\begin{cases} \cos \left[\left(k + \frac{1}{2} \right) s \right], & \text{αν } k = 0, 2, 4, \dots \\ -\cos \left[\left(k + \frac{1}{2} \right) s \right], & \text{αν } k = 1, 3, 5, \dots \end{cases}$$
 Άρα η $\sin r \frac{s+\pi}{2}$ είναι άρτια.
- Έστω ότι $r = 2, 4, 6, \dots$ ($r = 2k$, $k = 0, 1, 2, \dots$)

$$\sin \left(r \frac{s+\pi}{2} \right) = \sin \left(2k \frac{s+\pi}{2} \right) = \sin (ks + k\pi) = \begin{cases} \sin (ks), & \text{αν } k = 0, 2, 4, \dots \\ -\sin (ks), & \text{αν } k = 1, 3, 5, \dots \end{cases}$$
 Άρα η $\sin r \frac{s+\pi}{2}$ είναι περιττή.

□

Λήμμα 6.3 Έστω ότι

$$M_p^j(l, r) := e^{-il \frac{m_p - m_j}{h_p}},$$

και

$$I_p^j(l, r) := \int_{-\pi}^{\pi} e^{il \frac{h_j}{h_p} s} \sin r \frac{s+\pi}{2} ds.$$

Τότε ισχύουν τα παρακάτω

(α)

$$\operatorname{Re} [M_j^p(l, r)] = \operatorname{Re} [M_p^j(l, r)] \quad (6.10)$$

(β)

$$\operatorname{Im} [M_j^p(l, r)] = \operatorname{Im} [M_p^j(l, r)] \quad (6.11)$$

(γ)

$$\operatorname{Re} [I_j^p(l, r)] = \begin{cases} \operatorname{Re} [I_p^j(l, r)], & \text{αν } r \text{ περιττός} \\ -\operatorname{Re} [I_p^j(l, r)], & \text{αν } r \text{ άρτιος} \end{cases} \quad (6.12)$$

(δ)

$$\operatorname{Im} [I_j^p(l, r)] = \begin{cases} -\operatorname{Im} [I_p^j(l, r)], & \text{αν } r \text{ περιττός} \\ \operatorname{Im} [I_p^j(l, r)], & \text{αν } r \text{ άρτιος} \end{cases}. \quad (6.13)$$

Απόδειξη. Παρατηρούμε ότι

$$\begin{aligned}
 M_p^j(l, r) &= e^{-l\pi \cot \frac{\pi}{n}} e^{i(\alpha_j - \alpha_p)} = \\
 &= e^{-l\pi \cot \frac{\pi}{n} [\cos(\alpha_j - \alpha_p) + i \sin(\alpha_j - \alpha_p)]} = \\
 &= e^{-l\pi \cot \frac{\pi}{n} \cos(\alpha_j - \alpha_p)} e^{-il\pi \cot \frac{\pi}{n} \sin(\alpha_j - \alpha_p)} = \\
 &= e^{-l\pi \cot \frac{\pi}{n} \cos(\alpha_j - \alpha_p)} \left[\cos \left(-l\pi \cot \frac{\pi}{n} \sin(\alpha_j - \alpha_p) \right) + i \sin \left(-l\pi \cot \frac{\pi}{n} \sin(\alpha_j - \alpha_p) \right) \right].
 \end{aligned}$$

Δηλαδή

$$\begin{cases} \operatorname{Re} [M_p^j(l, r)] = \cos \left(-l\pi \cot \frac{\pi}{n} \sin(\alpha_j - \alpha_p) \right) e^{-l\pi \cot \frac{\pi}{n} \cos(\alpha_j - \alpha_p)} \\ \operatorname{Im} [M_p^j(l, r)] = \sin \left(-l\pi \cot \frac{\pi}{n} \sin(\alpha_j - \alpha_p) \right) e^{-l\pi \cot \frac{\pi}{n} \cos(\alpha_j - \alpha_p)} \end{cases} \quad (6.14)$$

Επίσης

$$\begin{aligned}
 I_p^j &= \int_{-\pi}^{\pi} e^{ile^{i(\alpha_j - \alpha_p)}s} \sin r^{\frac{s+\pi}{2}} ds = \\
 &= \int_{-\pi}^{\pi} e^{il[\cos(\alpha_j - \alpha_p) + i \sin(\alpha_j - \alpha_p)]s} \sin r^{\frac{s+\pi}{2}} ds = \\
 &= \int_{-\pi}^{\pi} e^{il \cos(\alpha_j - \alpha_p)s} e^{-l \sin(\alpha_j - \alpha_p)s} \sin r^{\frac{s+\pi}{2}} ds = \\
 &= \int_{-\pi}^{\pi} e^{-l \sin(\alpha_j - \alpha_p)s} [\cos(l \cos(\alpha_j - \alpha_p)s) + i \sin(l \cos(\alpha_j - \alpha_p)s)] \sin r^{\frac{s+\pi}{2}} ds,
 \end{aligned}$$

δηλαδή

$$\begin{cases} \operatorname{Re} [I_p^j(l, r)] = \int_{-\pi}^{\pi} e^{-l \sin(\alpha_j - \alpha_p)s} \cos(l \cos(\alpha_j - \alpha_p)s) \sin r^{\frac{s+\pi}{2}} ds \\ \operatorname{Im} [I_p^j(l, r)] = \int_{-\pi}^{\pi} e^{-l \sin(\alpha_j - \alpha_p)s} \sin(l \cos(\alpha_j - \alpha_p)s) \sin r^{\frac{s+\pi}{2}} ds \end{cases} \quad (6.15)$$

(α), (β) Συνεπάρχονται άμεσα από την (6.15) αφού $(\alpha_p - \alpha_j) = -(\alpha_j - \alpha_p)$, $\cos(-A) = \cos(A)$ και $\sin(-A) = -\sin(A)$.

(γ) Για r περιττό έχουμε

$$\begin{aligned}
& \operatorname{Re} [I_p^j(l, r)] - \operatorname{Re} [I_j^p(l, r)] = \\
& \int_{-\pi}^{\pi} e^{-l \sin(\alpha_j - \alpha_p) s} \cos(l \cos(\alpha_j - \alpha_p) s) \sin r \frac{s+\pi}{2} ds - \\
& \int_{-\pi}^{\pi} e^{-l \sin(\alpha_p - \alpha_j) s} \cos(l \cos(\alpha_p - \alpha_j) s) \sin r \frac{s+\pi}{2} ds = \\
& \int_{-\pi}^{\pi} [e^{-l \sin(\alpha_j - \alpha_p) s} \cos(l \cos(\alpha_j - \alpha_p) s) - e^{-l \sin(\alpha_p - \alpha_j) s} \cos(l \cos(\alpha_p - \alpha_j) s)] \sin r \frac{s+\pi}{2} ds = \\
& \int_{-\pi}^{\pi} [e^{l \sin(\alpha_p - \alpha_j) s} - e^{-l \sin(\alpha_p - \alpha_j) s}] \cos(l \cos(\alpha_p - \alpha_j) s) \sin r \frac{s+\pi}{2} ds = \\
& 2 \int_{-\pi}^{\pi} \sinh[l \sin(\alpha_p - \alpha_j) s] \cos(l \cos(\alpha_p - \alpha_j) s) \sin r \frac{s+\pi}{2} ds = 0
\end{aligned}$$

αφού το $\sinh$ είναι περιττή συνάρτηση και τα $\cos$ και $\sin$ είναι άρτιες συναρτήσεις (βλ. Λήμμα 6.2).

Άρα

$$\operatorname{Re} [I_p^j(l, r)] = \operatorname{Re} [I_j^p(l, r)], \text{ αν } r \text{ περιττός.}$$

Για r άρτιο έχουμε

$$\begin{aligned}
& \operatorname{Re} [I_p^j(l, r)] + \operatorname{Re} [I_j^p(l, r)] = \\
& 2 \int_{-\pi}^{\pi} \cosh[l \sin(\alpha_p - \alpha_j) s] \cos(l \cos(\alpha_p - \alpha_j) s) \sin r \frac{s+\pi}{2} ds = 0,
\end{aligned}$$

αφού τα $\cosh$ και $\cos$ είναι άρτιες συναρτήσεις ενώ το $\sin$ είναι περιττή συνάρτηση (βλ. Λήμμα 6.2).

Άρα

$$\operatorname{Re} [I_p^j(l, r)] = -\operatorname{Re} [I_j^p(l, r)], \text{ αν } r \text{ άρτιος.}$$

(δ) Όμοια με το (γ)

□

Θεώρημα 6.3 Στην περίπτωση των κανονικών πολυγώνων με ίδιες συνοριακές συνθήκες σε όλες τις πλευρές τα συμμετρικά μπλοκ A_p^j και A_j^p του *collocation* πίνακα της *sine-SFFS* έχουν την ιδιότητα

$$A_j^p = W A_p^j W, \quad (6.16)$$

όπου W διαγώνιος πίνακας στον $\mathbb{R}^{N \times N}$

$$W = \text{diag}(1, -1, \dots, 1, -1).^1 \quad (6.17)$$

Απόδειξη. Καταρχήν ας δούμε την δράση του μετασχηματισμού WBW σε ένα τυχαίο πίνακα B . Αν

$$B = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \cdots \\ a_{21} & a_{22} & a_{23} & a_{24} & \cdots \\ a_{31} & a_{32} & a_{33} & a_{34} & \cdots \\ a_{41} & a_{42} & a_{43} & a_{44} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

τότε

$$WBW = \begin{bmatrix} a_{11} & -a_{12} & a_{13} & -a_{14} & \cdots \\ -a_{21} & a_{22} & -a_{23} & a_{24} & \cdots \\ a_{31} & -a_{32} & a_{33} & -a_{34} & \cdots \\ -a_{41} & a_{42} & -a_{43} & a_{44} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Επομένως, για την απόδειξη του θεωρήματος, αρκεί να δείξουμε ότι για κάθε p, j, l, r ισχύει

$$\begin{cases} \text{Re}[c_j^p(l, r)] = \begin{cases} \text{Re}[c_p^j(l, r)], & \text{αν } r \text{ περιττός} \\ -\text{Re}[c_p^j(l, r)], & \text{αν } r \text{ άρτιος} \end{cases} \\ \text{Im}[c_j^p(l, r)] = \begin{cases} -\text{Im}[c_p^j(l, r)], & \text{αν } r \text{ περιττός} \\ \text{Im}[c_p^j(l, r)], & \text{αν } r \text{ άρτιος} \end{cases} \end{cases}.$$

Ισχύει ότι

$$c_p^j(l, r) = M_p^j(l, r)I_p^j(l, r),$$

όπου $M_p^j(l, r)$ και $I_p^j(l, r)$ όπως ορίστηκαν στο λήμμα (6.3) Έστω ότι, για σταθερά l και r ,

$$R_1 := \text{Re}[M_p^j(l, r)]$$

$$I_1 := \text{Im}[M_p^j(l, r)]$$

$$R_2 := \text{Re}[I_p^j(l, r)]$$

$$I_2 := \text{Im}[I_p^j(l, r)].$$

Από το λήμμα (6.3) ξέρουμε ότι

$$M_j^p = R_1 - iI_1,$$

¹ Ο συμβολισμός $D = \text{diag}(d_1, d_2, \dots, d_n)$ σημαίνει ότι ο πίνακας D είναι διαγώνιος και τα διαγώνια στοιχεία του είναι, με την σειρά που αναγράφονται, τα $d_1, d_2, d_3, \dots, d_n$.

καθώς και ότι

$$I_j^p = \begin{cases} R_2 - iI_2, & \text{αν } r = 1, 3, 5, \dots \\ -R_2 + iI_2, & \text{αν } r = 2, 4, 6, \dots \end{cases}$$

Επομένως διακρίνοντας δυο περιπτώσεις για το r έχουμε

Για r περιττό

$$c_p^j(l, r) = (R_1 + iI_1)(R_2 + iI_2) = (R_1R_2 - I_1I_2) + i(R_1I_2 + I_1R_2) \text{ και}$$

$$c_j^p(l, r) = (R_1 - iI_1)(R_2 - iI_2) = (R_1R_2 - I_1I_2) - i(R_1I_2 + I_1R_2).$$

Δηλαδή

$$\begin{cases} \operatorname{Re}(c_j^p(l, r)) = \operatorname{Re}(c_p^j(l, r)) \\ \operatorname{Im}(c_j^p(l, r)) = -\operatorname{Im}(c_p^j(l, r)) \end{cases} \quad (6.18)$$

Για r άρτιο

$$c_p^j(l, r) = (R_1 + iI_1)(R_2 + iI_2) = (R_1R_2 - I_1I_2) + i(R_1I_2 + I_1R_2) \text{ και}$$

$$c_j^p(l, r) = (R_1 - iI_1)(-R_2 + iI_2) = -(R_1R_2 - I_1I_2) + i(R_1I_2 + I_1R_2).$$

Δηλαδή

$$\begin{cases} \operatorname{Re}(c_j^p(l, r)) = -\operatorname{Re}(c_p^j(l, r)) \\ \operatorname{Im}(c_j^p(l, r)) = \operatorname{Im}(c_p^j(l, r)) \end{cases} \quad (6.19)$$

□

Παρατήρηση 6.1 Για να γίνει πιο κατανοητή η μορφή του πίνακα, της *sine-collocation* στην περίπτωση των κανονικών πολυγώνων με ίδιες συνοριακές συνθήκες σε όλες τις πλευρές, παραθέτουμε τον *collocation* πίνακα για δυο περιπτώσεις $n = 6$ και $n = 7$.

Για $n = 6$

$$A = \begin{bmatrix} D & B & C & E & \tilde{C} & \tilde{B} \\ \tilde{B} & D & B & C & E & \tilde{C} \\ \tilde{C} & \tilde{B} & D & B & C & E \\ \tilde{E} & \tilde{C} & \tilde{B} & D & B & C \\ C & \tilde{E} & \tilde{C} & \tilde{B} & D & B \\ B & C & \tilde{E} & \tilde{C} & \tilde{B} & D \end{bmatrix} \quad (6.20)$$

Για $n = 7$

$$A = \begin{bmatrix} D & B & C & E & \tilde{E} & \tilde{C} & \tilde{B} \\ \tilde{B} & D & B & C & E & \tilde{E} & \tilde{C} \\ \tilde{C} & \tilde{B} & D & B & C & E & \tilde{E} \\ \tilde{E} & \tilde{C} & \tilde{B} & D & B & C & E \\ E & \tilde{E} & \tilde{C} & \tilde{B} & D & B & C \\ C & E & \tilde{E} & \tilde{C} & \tilde{B} & D & B \\ B & C & E & \tilde{E} & \tilde{C} & \tilde{B} & D \end{bmatrix} \quad (6.21)$$

όπου D διαγώνιος πίνακας και $\tilde{X} = WXW$ με τον W όπως ορίστηκε στο θεώρημα (6.3).

Παρατήρηση 6.2 Αναφερόμενοι στο θεώρημα (6.3) μπορούμε να παρατηρήσουμε ότι τα συμμετρικά μπλοκ του collocation πίνακα έχουν

- Το ίδιο φάσμα και
- Την ίδια ορίζουσα.

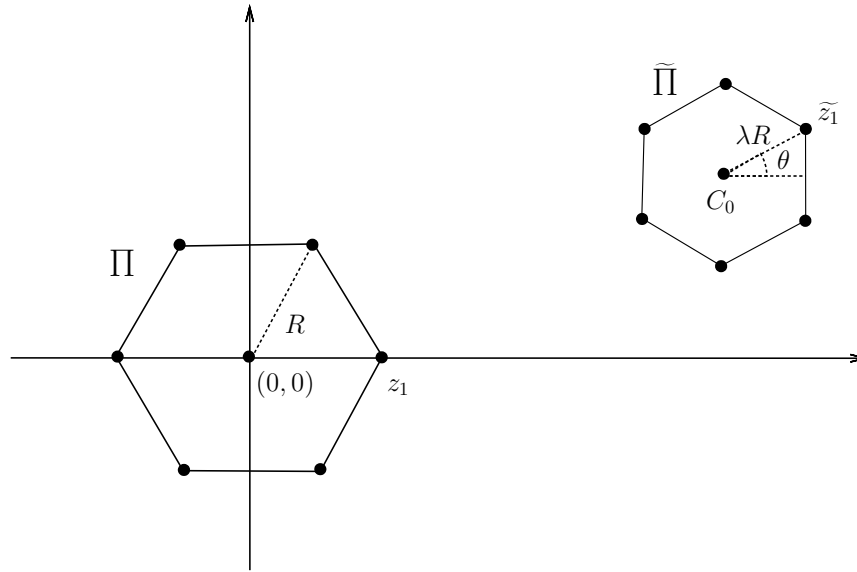
Αυτό προκύπτει από το γεγονός ότι $W^{-1} = W$ οπότε ο μετασχηματισμός που συνδέει τα συμμετρικά μπλοκ είναι μετασχηματισμός ομοιότητας.

6.2 Γενικότερες Περιπτώσεις

Εκτός από την περίπτωση των κανονικών πολυγώνων με τις ίδιες συνοριακές συνθήκες σε όλες τις πλευρές μπορούμε να εξαγάγουμε συμπεράσματα όχι για την μορφή, αλλά για το πώς μεταβάλλεται ο collocation πίνακας όταν μεταφέρουμε, στρέψουμε ή μεγεθύνουμε το πολύγωνο. Ακόμα μπορούμε να συμπεράνουμε την μη μεταβολή του collocation πίνακα καθώς αλλάζουμε ομοιόμορφα τις συνοριακές συνθήκες σε όλες τις πλευρές. Στην περίπτωση δε που έχουμε διαφορετικές συνοριακές συνθήκες σε κάθε πλευρά υπάρχει τρόπος να γράψουμε τον collocation ως συνδυασμό δυο "απλούστερων" πινάκων.

6.2.1 Μεταφορά, Στροφή και Μεγέθυνση του Πολυγώνου

Το ερώτημα που τίθεται προς διερεύνηση σε αυτή την παράγραφο είναι το πώς θα αλλάξει, ο collocation πίνακας αν μεταφέρουμε, στρέψουμε και μεγεθύνουμε το αρχικό μας πολύγωνο. Θα αναφερόμαστε στο αρχικό πολύγωνο με Π και στο μετασχηματισμένο με $\tilde{\Pi}$. Την ίδια πρακτική θα ακολουθήσουμε και για τα άλλα στοιχεία του μετασχηματισμένου πολυγώνου (θα τοποθετούμε μια "τίλντα" πάνω στο κάθε στοιχείο) για να τα ξεχωρίζουμε από τα στοιχεία του αρχικού πολυγώνου. Έστω λοιπόν ότι το μετασχηματισμένο πολύγωνο έχει μεγεθυνθεί κατά λ , έχει στραφεί θετικά κατά γωνία θ (rad) και (το κέντρο του) έχει μεταφερθεί κατά C_0 . Τα δύο πολύγωνα φαίνονται, για την περίπτωση των κανονικών πολυγώνων με ίδιες συνοριακές συνθήκες σε όλες τις πλευρές και για $n = 6$, στο σχήμα (6.4).



Σχήμα 6.4: Μεταφορά, στροφή και μεγέθυνση ενός κανονικού εξαγώνου με κέντρο την αρχή των αξόνων και ακτίνα R .

Θεώρημα 6.4 Ο *collocation* πίνακας για την *SFFS* μέθοδο παραμένει αμετάβλητος αν μεταφέρουμε, στρέφουμε ή μεγεθύνουμε ένα πολύγωνο (κανονικό ή μη) για οποιοδήποτε σύνολο συναρτήσεων βάσης.

Απόδειξη. Έστω ότι ξεκινάμε από ένα αρχικό πολύγωνο Π με κέντρο την αρχή των αξόνων το οποίο μετασχηματίζουμε σε ένα νέο πολύγωνο $\tilde{\Pi}$ μεταφέροντας το κέντρο του στο σημείο C_0 μεγεθύνοντας το κατά λ και στρέφοντας το θετικά κατά γωνία θ . Δηλαδή η κάθε κορυφή του αρχικού πολυγώνου z_j θα μετασχηματιστεί στην $\tilde{z}_j$ όπου

$$\tilde{z}_j = C_0 + \lambda e^{i\theta} z_j \quad (6.22)$$

Το μπλοκ A_p^j του *collocation* πίνακα για το αρχικό πολύγωνο θα είναι το (6.1) όπου τα στοιχεία $c_p^j(l, r)$ ορίζονται στην (6.2). Τα αντίστοιχα στοιχεία $\tilde{c}_p^j(l, r)$ για το μετασχηματισμένο πολύγωνο ορίζονται όπως

$$\tilde{c}_p^j(l, r) := \frac{|\tilde{h}_j|}{|\tilde{h}_p|} e^{i(\beta_j - \beta_p)} e^{-il \frac{\tilde{m}_p - \tilde{m}_j}{\tilde{h}_p}} \int_{-\pi}^{\pi} e^{il \frac{\tilde{h}_j}{\tilde{h}_p} s} \varphi_r(s) ds \quad {}^1 \quad (6.23)$$

Για να δείξουμε την ισότητα των δυο *collocation* πινάκων που αντιστοιχούν στα πολύγωνα Π και $\tilde{\Pi}$ αντίστοιχα αρκεί να δείξουμε ότι για κάθε p, j, l, r ισχύει

$$c_p^j(l, r) = \tilde{c}_p^j(l, r)$$

¹ Οι συνοριακές συνθήκες (β_j) δεν επηρεάστηκαν από αυτό τον μετασχηματισμό.

ή ισοδύναμα

$$\frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-il \frac{m_p - m_j}{h_p}} \int_{-\pi}^{\pi} e^{il \frac{h_j}{h_p} s} \varphi_r(s) ds = \frac{|\tilde{h}_j|}{|\tilde{h}_p|} e^{i(\beta_j - \beta_p)} e^{-il \frac{\tilde{m}_p - \tilde{m}_j}{\tilde{h}_p}} \int_{-\pi}^{\pi} e^{il \frac{\tilde{h}_j}{\tilde{h}_p} s} \varphi_r(s) ds \Leftrightarrow \quad (6.24)$$

$$\frac{|h_j|}{|h_p|} e^{-il \frac{m_p - m_j}{h_p}} \int_{-\pi}^{\pi} e^{il \frac{h_j}{h_p} s} \varphi_r(s) ds = \frac{|\tilde{h}_j|}{|\tilde{h}_p|} e^{-il \frac{\tilde{m}_p - \tilde{m}_j}{\tilde{h}_p}} \int_{-\pi}^{\pi} e^{il \frac{\tilde{h}_j}{\tilde{h}_p} s} \varphi_r(s) ds \quad (6.25)$$

Όμως

$$\begin{cases} \tilde{h}_j := \frac{\tilde{z}_{j+1} - \tilde{z}_j}{2\pi} = \frac{C_0 + \lambda e^{i\theta} z_{j+1} - (C_0 + \lambda e^{i\theta} z_j)}{2\pi} = \lambda e^{i\theta} \frac{z_{j+1} - z_j}{2\pi} = \lambda e^{i\theta} h_j \\ \tilde{m}_j := \frac{\tilde{z}_{j+1} + \tilde{z}_j}{2} = \frac{C_0 + \lambda e^{i\theta} z_{j+1} + (C_0 + \lambda e^{i\theta} z_j)}{2} = C_0 + \lambda e^{i\theta} \frac{z_{j+1} + z_j}{2} = C_0 + \lambda e^{i\theta} m_j \end{cases}$$

και

$$\begin{cases} \frac{\tilde{h}_j}{\tilde{h}_p} = \frac{\lambda e^{i\theta} h_j}{\lambda e^{i\theta} h_p} = \frac{h_j}{h_p} \\ \frac{\tilde{m}_p - \tilde{m}_j}{\tilde{h}_p} = \frac{C_0 + \lambda e^{i\theta} m_p - (C_0 + \lambda e^{i\theta} m_j)}{\lambda e^{i\theta} h_p} = \frac{\lambda e^{i\theta} (m_p - m_j)}{\lambda e^{i\theta} h_p} = \frac{m_p - m_j}{h_p} \end{cases}$$

Επομένως η (6.25) φανερά ισχύει. □

Παρατήρηση 6.3

Ο *collocation* πίνακας που αντιστοιχεί στο πρώτο μέλος της *Dirichlet-Neumann* απεικόνισης παραμένει αμετάβλητος σε μεταφορές, στροφές και μεγεθύνσεις του πολυγώνου γεγονός το οποίο δεν ισχύει για το δεύτερο μέλος της *Dirichlet-Neumann* απεικόνισης το οποίο μεταβάλλεται. Δηλαδή αν η διακριτοποίηση της *Dirichlet-Neumann* απεικόνισης για ένα συγκεκριμένο πολύγωνο μας οδηγήσει σε ένα γραμμικό σύστημα $Ax = b$ η διακριτοποίηση της *Dirichlet-Neumann* απεικόνισης για ένα μετασχηματισμένο πολύγωνο θα μας οδηγήσει σε ένα γραμμικό σύστημα $Ax = \tilde{b}$ με ίδιο τον πίνακα των συντελεστών των αγνώστων αλλά διαφορετικό δεύτερο μέλος.

6.2.2 Ομοιόμορφη Μεταβολή των Συνοριακών Συνθηκών

Θεώρημα 6.5 Έστω A ο *collocation* πίνακας που προκύπτει για ένα συγκεκριμένο σύνολο συνοριακών συνθηκών $\{\beta_j\}_{j=1}^n$. Αν μεταβάλουμε ομοιόμορφα το σύνολο των συνοριακών συνθηκών σε όλες τις πλευρές, αν δηλαδή

$$\{\tilde{\beta}_j\}_{j=1}^n = \{\beta_j + \gamma\}_{j=1}^n, \quad (6.26)$$

ο *collocation* πίνακας παραμένει αμετάβλητος.

Απόδειξη. Ο collocation πίνακας παράγεται από το πρώτο μέλος της Dirichlet-Neumann απεικόνιση (2.42). Παρατηρώντας προσεκτικά την Dirichlet-Neumann απεικόνιση βλέπουμε ότι οι συνοριακές συνθήκες εμφανίζονται, στο πρώτο μέλος, στην μορφή $e^{i(\beta_j - \beta_p)}$. Είναι προφανές ότι αν $\tilde{\beta}_j = \beta_j + \gamma$ ο παράγοντας $e^{i(\tilde{\beta}_j - \tilde{\beta}_p)}$ ισούται με $e^{i(\beta_j - \beta_p)}$. $\square$

Άμεση συνέπεια του προηγούμενου θεωρήματος είναι το επόμενο συμπέρασμα.

Συμπέρασμα 6.6 *Αν χρησιμοποιήσουμε τις ίδιες συνοριακές συνθήκες σε κάθε πλευρά του πολυγώνου ($\beta_j = \beta$), ο collocation πίνακας για την SFPS μέθοδο είναι ανεξάρτητος της τιμής του β για οποιοδήποτε σύνολο συναρτήσεων βάσης.*

Παρατήρηση 6.4 Το δεύτερο μέλος του γραμμικού συστήματος, που προκύπτει από την διακριτοποίηση της Dirichlet-Neumann απεικόνισης, δεν είναι ανεξάρτητο από την επιλογή των συνοριακών συνθηκών άσχετα με το αν έχουμε τις ίδιες συνοριακές συνθήκες σε όλες τις πλευρές. Αυτό οφείλεται στις συναρτήσεις $g^{(j)}$ οι οποίες εμφανίζονται στο δεύτερο μέλος και εξαρτώνται άμεσα από τις συνοριακές συνθήκες (β_j).

6.2.3 Διαφορετικές Συνοριακές Συνθήκες σε Κάθε Πλευρά

Σε αυτή την παράγραφο θα θεωρήσουμε (α) τον collocation πίνακα για ένα πολύγωνο με τις ίδιες συνοριακές συνθήκες σε όλες τις πλευρές και (β) τον collocation πίνακα για το ίδιο πολύγωνο αλλά με διαφορετικές συνοριακές συνθήκες σε κάθε πλευρά και θα δείξουμε την σχέση μεταξύ τους.

Θεωρούμε την ποσότητα

$$c_p^j(l, r) := \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-il \frac{m_p - m_j}{h_p}} \int_{-\pi}^{\pi} e^{il \frac{h_j}{h_p} s} \varphi_r(s) ds \quad (6.27)$$

καθώς και το μιγαδικό διάνυσμα

$$E_p^j(l) := \begin{bmatrix} c_p^j(l, 1) & c_p^j(l, 2) & \cdots & c_p^j(l, N_j) \end{bmatrix} \quad (6.28)$$

Όπως γνωρίζουμε ακολουθούμε το εξής σχήμα για την κατασκευή, των γραμμών, του μπλοκ A_p^j του collocation πίνακα A για την SFPS:

$$\begin{aligned} 1^{\text{η}} \text{ γραμμή του } A_p^j &\longmapsto \text{Re} \left[E_p^j \left(l = \frac{1}{2} \right) \right] \\ 2^{\text{η}} \text{ γραμμή του } A_p^j &\longmapsto \text{Im} \left[E_p^j (l = 1) \right] \\ 3^{\text{η}} \text{ γραμμή του } A_p^j &\longmapsto \text{Re} \left[E_p^j \left(l = \frac{3}{2} \right) \right] \\ 4^{\text{η}} \text{ γραμμή του } A_p^j &\longmapsto \text{Im} \left[E_p^j (l = 2) \right] \\ &\vdots \\ N_{j-1}^{\text{η}} \text{ γραμμή του } A_p^j &\longmapsto \text{Re} \left[E_p^j \left(l = \frac{N_j-1}{2} \right) \right] \\ N_j^{\text{η}} \text{ γραμμή του } A_p^j &\longmapsto \text{Im} \left[(E_p^j (l = N_j/2)) \right] \end{aligned}$$

Ταυτόχρονα θα θεωρήσουμε και ένα δεύτερο πίνακα B , τον οποίο θα ονομάσουμε "συγγενή *collocation* πίνακα", του οποίου τα μπλοκ B_p^j κατασκευάζονται εντελώς ανάλογα με αυτά του *collocation* πίνακα A όπως παρακάτω:

$$\begin{aligned}
 1^{\text{η}} \text{ γραμμή του } B_p^j &\longmapsto \text{Im} [E_p^j (l = \frac{1}{2})] \\
 2^{\text{η}} \text{ γραμμή του } B_p^j &\longmapsto \text{Re} [E_p^j (l = 1)] \\
 3^{\text{η}} \text{ γραμμή του } B_p^j &\longmapsto \text{Im} [E_p^j (l = \frac{3}{2})] \\
 4^{\text{η}} \text{ γραμμή του } B_p^j &\longmapsto \text{Re} [E_p^j (l = 2)] \\
 &\vdots \\
 N_{j-1}^{\text{η}} \text{ γραμμή του } B_p^j &\longmapsto \text{Im} [E_p^j (l = \frac{N_j-1}{2})] \\
 N_j^{\text{η}} \text{ γραμμή του } B_p^j &\longmapsto \text{Re} [(E_p^j (l = N_j/2))]
 \end{aligned}$$

Παρατήρηση 6.5

Και ο *collocation* πίνακας και ο συγγενής *collocation* πίνακας είναι πραγματικοί και προέρχονται από τον ίδιο μιγαδικό πίνακα. Ο συγγενής *collocation* πίνακας B θα μπορούσε να θεωρηθεί ως ο *collocation* πίνακας, όχι της *SFFS* αλλά μιας άλλης μεθόδου που θα μπορούσε κάλλιστα να ονομαστεί "συζυγής-SFFS" λόγω της στενής της σχέσης με την *SFFS*.

Έστω λοιπόν ότι έχουμε κατασκευάσει τον *collocation* πίνακα (A) καθώς και τον συγγενή *collocation* πίνακα (B) για ένα πολύγωνο θεωρώντας τις ίδιες συνοριακές συνθήκες σε κάθε πλευρά. Έστω τώρα ότι θέλουμε να κατασκευάσουμε τον *collocation* πίνακα K για την περίπτωση που έχουμε διαφορετικές συνοριακές συνθήκες β_j σε κάθε πλευρά j .

Χάριν απλότητας και αναφερόμενοι στα μπλοκ (p, j) των *collocation* πινάκων ορίζουμε, για την περίπτωση των ιδίων συνοριακών συνθηκών σε κάθε πλευρά, την ποσότητα

$$\gamma_p^j(l, r) := c_p^j(l, r) = \frac{|h_j|}{|h_p|} e^{-il \frac{m_p - m_j}{h_p}} \int_{-\pi}^{\pi} e^{il \frac{h_j}{h_p} s} \varphi_r(s) ds$$

Παρατηρούμε ότι οι ποσότητες $c_p^j(l, r)$ που χρειαζόμαστε για την κατασκευή των μπλοκ του K σχετίζονται άμεσα με τις ποσότητες $\gamma_p^j(l, r)$ που έχουμε χρησιμοποιήσει, αντίστοιχα, για την κατασκευή των μπλοκ του A . Ειδικότερα μιλώντας για τα μπλοκ (p, j) και μη χρησιμοποιώντας τους αυτονόητους δείκτες θα έχουμε

$$\begin{aligned}
 c(l, r) &= e^{i(\beta_j - \beta_p)} \gamma(l, r) = \\
 &= (\cos(\beta_j - \beta_p) + i \sin(\beta_j - \beta_p)) (\text{Re}(\gamma(l, r)) + i \text{Im}(\gamma(l, r))) = \\
 &= [\cos(\beta_j - \beta_p) \gamma_R(l, r) - \sin(\beta_j - \beta_p) \gamma_I(l, r)] + i [\cos(\beta_j - \beta_p) \gamma_I(l, r) + \sin(\beta_j - \beta_p) \gamma_R(l, r)]
 \end{aligned}$$

Επομένως για την κατασκευή των γραμμών του K_p^j βλέπουμε ότι

$$\begin{aligned} \text{περιττές γραμμές} &\longmapsto \cos(\beta_j - \beta_p) \gamma_R(l, r) - \sin(\beta_j - \beta_p) \gamma_I(l, r) \\ \text{άρτιες γραμμές} &\longmapsto \cos(\beta_j - \beta_p) \gamma_I(l, r) + \sin(\beta_j - \beta_p) \gamma_R(l, r) \end{aligned}$$

Όμως οι ποσότητες γ_R για τις περιττές γραμμές και γ_I για τις άρτιες γραμμές είναι αυτές που χρησιμοποιήσαμε στην κατασκευή του A ενώ οι γ_I για τις περιττές γραμμές και γ_R για τις άρτιες γραμμές χρησιμοποιήθηκαν για την κατασκευή του B . Ορίζοντας διαγώνιο πίνακα $M_N \in \mathbb{R}^{N \times N}$ τέτοιο ώστε

$$M_N = \text{diag}(-1, 1, -1, 1, \dots, -1, 1)$$

μπορούμε να πούμε ότι

$$K_p^j = \cos(\beta_j - \beta_p) A_p^j + \sin(\beta_j - \beta_p) M_N B_p^j \quad (6.29)$$

Δηλαδή ισχύει ότι

$$K = K1 + K2$$

όπου

$$K1 = \begin{bmatrix} \cos(\beta_1 - \beta_1) & A_1^1 & \cos(\beta_2 - \beta_1) & A_1^2 & \cos(\beta_3 - \beta_1) & A_1^3 & \cdots \\ \cos(\beta_1 - \beta_2) & A_2^1 & \cos(\beta_2 - \beta_2) & A_2^2 & \cos(\beta_3 - \beta_2) & A_2^3 & \cdots \\ \cos(\beta_1 - \beta_3) & A_3^1 & \cos(\beta_2 - \beta_3) & A_3^2 & \cos(\beta_3 - \beta_3) & A_3^3 & \cdots \\ \vdots & & \vdots & & \vdots & & \ddots \end{bmatrix}$$

και

$$K2 = \begin{bmatrix} \sin(\beta_1 - \beta_1) M_N B_1^1 & \sin(\beta_2 - \beta_1) M_N B_1^2 & \sin(\beta_3 - \beta_1) M_N B_1^3 & \cdots \\ \sin(\beta_1 - \beta_2) M_N B_2^1 & \sin(\beta_2 - \beta_2) M_N B_2^2 & \sin(\beta_3 - \beta_2) M_N B_2^3 & \cdots \\ \sin(\beta_1 - \beta_3) M_N B_3^1 & \sin(\beta_2 - \beta_3) M_N B_3^2 & \sin(\beta_3 - \beta_3) M_N B_3^3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Με την χρήση της ταυτότητας $\cos(a - b) = \cos a \cos b + \sin a \sin b$ μπορούμε να αναλύσουμε περαιτέρω τον $K1$. Για αυτήν την ανάλυση ορίζουμε επιπλέον τους κάτωθι διαγώνιους πίνακες C_n , S_n και I_N

$$C_n = \begin{bmatrix} \cos \beta_1 & & \\ & \ddots & \\ & & \cos \beta_n \end{bmatrix}, \quad S_n = \begin{bmatrix} \sin \beta_1 & & \\ & \ddots & \\ & & \sin \beta_n \end{bmatrix}, \quad I_N := \text{Μοναδιαίος στον } \mathbb{R}^{N \times N} \quad (6.30)$$

Έτσι ο $K1$ γράφεται

$$\begin{aligned}
 K1 &= \begin{bmatrix} (\cos \beta_1 \cos \beta_1 + \sin \beta_1 \sin \beta_1) A_1^1 & (\cos \beta_2 \cos \beta_1 + \sin \beta_2 \sin \beta_1) A_1^2 & (\cos \beta_3 \cos \beta_1 + \sin \beta_3 \sin \beta_1) A_1^3 & \cdots \\ (\cos \beta_1 \cos \beta_2 + \sin \beta_1 \sin \beta_2) A_2^1 & (\cos \beta_2 \cos \beta_2 + \sin \beta_2 \sin \beta_2) A_2^2 & (\cos \beta_3 \cos \beta_2 + \sin \beta_3 \sin \beta_2) A_2^3 & \cdots \\ (\cos \beta_1 \cos \beta_3 + \sin \beta_1 \sin \beta_3) A_3^1 & (\cos \beta_2 \cos \beta_3 + \sin \beta_2 \sin \beta_3) A_3^2 & (\cos \beta_3 \cos \beta_3 + \sin \beta_3 \sin \beta_3) A_3^3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \\
 &= \begin{bmatrix} \cos \beta_1 \cos \beta_1 A_1^1 & \cos \beta_2 \cos \beta_1 A_1^2 & \cos \beta_3 \cos \beta_1 A_1^3 & \cdots \\ \cos \beta_1 \cos \beta_2 A_2^1 & \cos \beta_2 \cos \beta_2 A_2^2 & \cos \beta_3 \cos \beta_2 A_2^3 & \cdots \\ \cos \beta_1 \cos \beta_3 A_3^1 & \cos \beta_2 \cos \beta_3 A_3^2 & \cos \beta_3 \cos \beta_3 A_3^3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} + \begin{bmatrix} \sin \beta_1 \sin \beta_1 A_1^1 & \sin \beta_2 \sin \beta_1 A_1^2 & \sin \beta_3 \sin \beta_1 A_1^3 & \cdots \\ \sin \beta_1 \sin \beta_2 A_2^1 & \sin \beta_2 \sin \beta_2 A_2^2 & \sin \beta_3 \sin \beta_2 A_2^3 & \cdots \\ \sin \beta_1 \sin \beta_3 A_3^1 & \sin \beta_2 \sin \beta_3 A_3^2 & \sin \beta_3 \sin \beta_3 A_3^3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \\
 &= \begin{bmatrix} \cos \beta_1 A_1^1 & \cos \beta_1 A_1^2 & \cos \beta_1 A_1^3 & \cdots \\ \cos \beta_2 A_2^1 & \cos \beta_2 A_2^2 & \cos \beta_2 A_2^3 & \cdots \\ \cos \beta_3 A_3^1 & \cos \beta_3 A_3^2 & \cos \beta_3 A_3^3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} (C_n \otimes I_N) + \begin{bmatrix} \sin \beta_1 A_1^1 & \sin \beta_1 A_1^2 & \sin \beta_1 A_1^3 & \cdots \\ \sin \beta_2 A_2^1 & \sin \beta_2 A_2^2 & \sin \beta_2 A_2^3 & \cdots \\ \sin \beta_3 A_3^1 & \sin \beta_3 A_3^2 & \sin \beta_3 A_3^3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} (S_n \otimes I_N) \\
 &= (C_n \otimes I_N) \begin{bmatrix} A_1^1 & A_1^2 & A_1^3 & \cdots \\ A_2^1 & A_2^2 & A_2^3 & \cdots \\ A_3^1 & A_3^2 & A_3^3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} (C_n \otimes I_N) + (S_n \otimes I_N) \begin{bmatrix} A_1^1 & A_1^2 & A_1^3 & \cdots \\ A_2^1 & A_2^2 & A_2^3 & \cdots \\ A_3^1 & A_3^2 & A_3^3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} (S_n \otimes I_N).
 \end{aligned}$$

Δηλαδή

$$K1 = (C_n \otimes I_N) A (C_n \otimes I_N) + (S_n \otimes I_N) A (S_n \otimes I_N). \quad (6.31)$$

Με εντελώς ανάλογο τρόπο μπορούμε να δείξουμε ότι

$$K2 = (C_n \otimes I_N) MB (S_n \otimes I_N) - (S_n \otimes I_N) MB (C_n \otimes I_N), \quad (6.32)$$

όπου

$$M = M_N \otimes I_n = \text{diag}(-1, 1, -1, 1, \dots, -1, 1). \quad (6.33)$$

Επομένως έχουμε αποδείξει το παρακάτω θεώρημα

Θεώρημα 6.7 Για την *SFFS* μέθοδο και για οποιοδήποτε σύνολο συναρτήσεων βάσης, ο *collocation* πίνακας K , στην περίπτωση που έχουμε διαφορετικές συνοριακές συνθήκες σε κάθε πλευρά ενός πολυγώνου, γράφεται ως συνδυασμός του *collocation* πίνακα A για την περίπτωση που έχουμε ίδιες συνοριακές συνθήκες σε όλες τις πλευρές ενός πολυγώνου και του συγγενή *collocation* πίνακα B ως εξής

$$K = (C_n \otimes I_N) A (C_n \otimes I_N) + (S_n \otimes I_N) A (S_n \otimes I_N) + \quad (6.34)$$

$$+ (C_n \otimes I_N) MB (S_n \otimes I_N) - (S_n \otimes I_N) MB (C_n \otimes I_N), \quad (6.35)$$

όπου C_n , S_n και I_N όπως έχουν οριστεί στην (6.30).

6.3 Τετράγωνο

Στην περίπτωση που το πολυγωνικό χωρίο, στην περίπτωση των κανονικών πολυγώνων με ίδιες συνοριακές συνθήκες σε όλες τις πλευρές, έχει τέσσερις πλευρές (τετράγωνο) τότε

εμφανίζονται επιπλέον ιδιότητες όσο αφορά στον collocation πίνακα, οι οποίες μας επιτρέπουν την αναλυτική εύρεση του φάσματος αυτού.

Λήμμα 6.4 Σε ένα τετράγωνο με ίδιες συνοριακές συνθήκες σε όλες τις πλευρές τα μπλοκ του collocation πίνακα της *SFFS* που αντιστοιχούν σε κάθετες πλευρές είναι μηδενικά για οποιοδήποτε σύνολο συναρτήσεων βάσης.

Απόδειξη. Έστω ότι οι πλευρές p, j είναι κάθετες. Τα στοιχεία του μπλοκ A_p^j δίνονται από το πραγματικό ή το φανταστικό μέρος των ποσοτήτων $c_p^j(l, r)$ (6.7) οι οποίες γράφονται ως

$$c_p^j(l, r) = e^{-l\pi} \int_{-\pi}^{\pi} e^{(\pi+is)le^{i(\alpha_j-\alpha_p)}} \varphi_r(s) ds \quad (6.36)$$

αφού για $n = 4 \Rightarrow \cot(\pi/n) = 1$. Χωρίς βλάβη της γενικότητας υποθέτουμε ότι $j > p$. Τότε ισχύει

$$\alpha_j - \alpha_p = \left(\alpha_p + \frac{\pi}{2} \right) - \alpha_p = \frac{\pi}{2} \quad (6.37)$$

και η σχέση (6.36) διαδοχικά γράφεται

$$\begin{aligned} c_p^j(l, r) &= e^{-l\pi} \int_{-\pi}^{\pi} e^{(\pi+is)le^{i\pi/2}} \varphi_r(s) ds \\ &= e^{-l\pi} \int_{-\pi}^{\pi} e^{il\pi} e^{-ls} \varphi_r(s) ds \\ &= e^{il\pi} \int_{-\pi}^{\pi} e^{-l\pi} e^{-ls} \varphi_r(s) ds. \end{aligned}$$

Όμως το $\int_{-\pi}^{\pi} e^{-l\pi} e^{-ls} \varphi_r(s) ds$ είναι ένας πραγματικός αριθμός έστω $R(l, r)$. Άρα

$$c_p^j(l, r) = e^{il\pi} R(l, r), \quad (6.38)$$

και επομένως

$$\begin{cases} \operatorname{Re}[c_p^j(l, r)] = \cos(l\pi) R(l, r) \\ \operatorname{Im}[c_p^j(l, r)] = \sin(l\pi) R(l, r) \end{cases}. \quad (6.39)$$

Οι περιττές γραμμές—εξισώσεις, του μπλοκ A_p^j , δίνονται από το πραγματικό μέρος του $c_p^j(l, r)$ για $l = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots$. Όμως για αυτά τα l ισχύει

$$\operatorname{Re}[c_p^j(l, r)] = \cos(l\pi) R(l, r) = 0$$

Δηλαδή όλες οι περιττές γραμμές στο μπλοκ A_p^j είναι μηδενικές.

Οι άρτιες γραμμές—εξισώσεις, του μπλοκ A_p^j , δίνονται από το φανταστικό μέρος του $c_p^j(l, r)$ για $l = 1, 2, \dots$. Όμως για αυτά τα l ισχύει

$$\text{Im}[c_p^j(l, r)] = \sin(l\pi) R(l, r) = 0$$

Δηλαδή και όλες οι άρτιες γραμμές στο μπλοκ A_p^j είναι μηδενικές. $\square$

Παρατήρηση 6.6 Το να μηδενίζεται τελείως ένα μπλοκ του collocation πίνακα ισχύει, μόνο για τετράγωνα αφού το $\cot(\pi/n)$ δεν είναι ακέραιος τουλάχιστον για τιμές του $n \leq 100000$.

Συμπέρασμα 6.8 Σε ένα τετράγωνο, με τις ίδιες συνοριακές συνθήκες σε όλες τις πλευρές, τα συμμετρικά μπλοκ του collocation πίνακα, της SFFS είναι ίσα για οποιοδήποτε σύνολο συναρτήσεων βάσης.

Απόδειξη. Είναι άμεση συνέπεια του λήμματος (6.4) δεδομένου ότι ο πίνακας είναι block-circulant. $\square$

Θεώρημα 6.9 Σε ένα τετράγωνο με τις ίδιες συνοριακές συνθήκες σε όλες τις πλευρές ο collocation πίνακας της sine-SFFS έχει την παρακάτω μπλοκ μορφή για οποιοδήποτε σύνολο συναρτήσεων βάσης:

$$A = \begin{bmatrix} A_1 & 0 & DA_1 & 0 \\ 0 & A_1 & 0 & DA_1 \\ DA_1 & 0 & A_1 & 0 \\ 0 & DA_1 & 0 & A_1 \end{bmatrix}, \quad (6.40)$$

όπου D διαγώνιος πίνακας

$$D := \text{diag}(e^{-1\pi}, -e^{-2\pi}, e^{-3\pi}, -e^{-4\pi} \dots e^{-(N_j-1)\pi}, -e^{-N_j\pi}). \quad (6.41)$$

Απόδειξη. Αρκεί να βρούμε την σχέση του κάθε στοιχείου του μπλοκ $A_1^3 (= A_3)$ με το αντίστοιχο στοιχείο του μπλοκ $A_1^1 (= A_1)$.

Τα στοιχεία των μπλοκ A_1 και A_3 δίνονται από το πραγματικό ή το φανταστικό μέρος των ποσοτήτων

$$c_1^1(l, r) := e^{-l\pi} \int_{-\pi}^{\pi} e^{(\pi+is)le^{i(\alpha_1-\alpha_1)}} \varphi_r(s) ds \quad (6.42)$$

και

$$c_1^3(l, r) := e^{-l\pi} \int_{-\pi}^{\pi} e^{(\pi+is)le^{i(\alpha_3-\alpha_1)}} \varphi_r(s) ds \quad (6.43)$$

αντίστοιχα.

Δεδομένου ότι $\alpha_1 - \alpha_1 = 0$ και $\alpha_3 - \alpha_1 = \pi$ οι παραπάνω σχέσεις γράφονται

$$c_1^1(l, r) = \int_{-\pi}^{\pi} e^{ils} \varphi_r(s) ds \quad (6.44)$$

και

$$c_1^3(l, r) = e^{-l\pi} \int_{-\pi}^{\pi} e^{(\pi+is)le^{i\pi}} \varphi_r(s) ds = e^{-l\pi} \int_{-\pi}^{\pi} e^{-l(\pi+is)} \varphi_r(s) ds = e^{-2l\pi} \int_{-\pi}^{\pi} e^{-ils} \varphi_r(s) ds \quad (6.45)$$

αντίστοιχα.

Από τις δύο σχέσεις (6.44) και (6.45) εύκολα συμπεραίνουμε ότι

$$\begin{cases} \operatorname{Re}[c_1^3(p, l)] = e^{-2l\pi} \operatorname{Re}[c_1^1(p, l)] \\ \operatorname{Im}[c_1^3(p, l)] = -e^{-2l\pi} \operatorname{Im}[c_1^1(p, l)] \end{cases} \quad (6.46)$$

Οι περιττές γραμμές–εξισώσεις, των μπλοκ A_1 και A_3 , οι οποίες αντιστοιχούν στο πραγματικό μέρος των ποσοτήτων $c_1^1(l, r)$ και $c_1^3(l, r)$ αντίστοιχα, δίνονται για $l = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots$ ενώ οι άρτιες γραμμές–εξισώσεις, οι οποίες αντιστοιχούν στο φανταστικό μέρος των ποσοτήτων $c_1^1(l, r)$ και $c_1^3(l, r)$ αντίστοιχα, δίνονται για $l = 1, 2, 3, \dots$. Επομένως για τον πίνακα A_3 ισχύει

$$A_3 = \begin{bmatrix} e^{-1\pi} A_1(1, :) \\ -e^{-2\pi} A_1(2, :) \\ e^{-3\pi} A_1(3, :) \\ -e^{-4\pi} A_1(4, :) \\ \vdots \\ e^{-(N_j-1)\pi} A_1(N_j-1, :) \\ -e^{-N_j\pi} A_1(N_j, :) \end{bmatrix} = D A_1 \quad (6.47)$$

□

Παρατήρηση 6.7 Σε ένα τετράγωνο ο *collocation* πίνακας της *SFFS* επιδέχεται την εξής παραγοντοποίηση

$$A = GT := \begin{bmatrix} I & 0 & D & 0 \\ 0 & I & 0 & D \\ D & 0 & I & 0 \\ 0 & D & 0 & I \end{bmatrix} \begin{bmatrix} A_1 & 0 & 0 & 0 \\ 0 & A_1 & 0 & 0 \\ 0 & 0 & A_1 & 0 \\ 0 & 0 & 0 & A_1 \end{bmatrix} \quad (6.48)$$

για οποιοδήποτε σύνολο συναρτήσεων βάσης. Ο πίνακας D είναι ο διαγώνιος πίνακας που ορίζεται στην (6.41), επομένως ο πίνακας G είναι ανεξάρτητος των συναρτήσεων βάσης ενώ ο πίνακας T

εξαρτάται αποκλειστικά από τις συναρτήσεις βάσης. Ειδικότερα τα διαγώνια μπλοκ του T είναι

$$A_1 = \begin{bmatrix} \operatorname{Re} \left(\int_{-\pi}^{\pi} e^{i\frac{1}{2}s} \varphi_1(s) ds \right) & \operatorname{Re} \left(\int_{-\pi}^{\pi} e^{i\frac{1}{2}s} \varphi_2(s) ds \right) & \cdots & \operatorname{Re} \left(\int_{-\pi}^{\pi} e^{i\frac{1}{2}s} \varphi_{N_j}(s) ds \right) \\ \operatorname{Im} \left(\int_{-\pi}^{\pi} e^{i\frac{1}{2}s} \varphi_1(s) ds \right) & \operatorname{Im} \left(\int_{-\pi}^{\pi} e^{i\frac{1}{2}s} \varphi_2(s) ds \right) & \cdots & \operatorname{Im} \left(\int_{-\pi}^{\pi} e^{i\frac{1}{2}s} \varphi_{N_j}(s) ds \right) \\ \operatorname{Re} \left(\int_{-\pi}^{\pi} e^{i\frac{3}{2}s} \varphi_1(s) ds \right) & \operatorname{Re} \left(\int_{-\pi}^{\pi} e^{i\frac{3}{2}s} \varphi_2(s) ds \right) & \cdots & \operatorname{Re} \left(\int_{-\pi}^{\pi} e^{i\frac{3}{2}s} \varphi_{N_j}(s) ds \right) \\ \operatorname{Im} \left(\int_{-\pi}^{\pi} e^{i\frac{3}{2}s} \varphi_1(s) ds \right) & \operatorname{Im} \left(\int_{-\pi}^{\pi} e^{i\frac{3}{2}s} \varphi_2(s) ds \right) & \cdots & \operatorname{Im} \left(\int_{-\pi}^{\pi} e^{i\frac{3}{2}s} \varphi_{N_j}(s) ds \right) \\ \vdots & \vdots & \vdots & \vdots \\ \operatorname{Re} \left(\int_{-\pi}^{\pi} e^{i\frac{N_j-1}{2}s} \varphi_1(s) ds \right) & \operatorname{Re} \left(\int_{-\pi}^{\pi} e^{i\frac{N_j-1}{2}s} \varphi_2(s) ds \right) & \cdots & \operatorname{Re} \left(\int_{-\pi}^{\pi} e^{i\frac{N_j-1}{2}s} \varphi_{N_j}(s) ds \right) \\ \operatorname{Im} \left(\int_{-\pi}^{\pi} e^{i\frac{N_j-1}{2}s} \varphi_1(s) ds \right) & \operatorname{Im} \left(\int_{-\pi}^{\pi} e^{i\frac{N_j-1}{2}s} \varphi_2(s) ds \right) & \cdots & \operatorname{Im} \left(\int_{-\pi}^{\pi} e^{i\frac{N_j-1}{2}s} \varphi_{N_j}(s) ds \right) \end{bmatrix} \quad (6.49)$$

Θεώρημα 6.10 Σε ένα τετράγωνο ο επαναληπτικός πίνακας της μπλοκ-Jacobi επαναληπτικής μεθόδου που αντιστοιχεί στον collocation πίνακα της SFFS έχει το ίδιο φάσμα για οποιοδήποτε σύνολο συναρτήσεων βάσης.

Απόδειξη. Ορίζουμε πίνακα

$$B_J := \begin{bmatrix} 0 & 0 & D & 0 \\ 0 & 0 & 0 & D \\ D & 0 & 0 & 0 \\ 0 & D & 0 & 0 \end{bmatrix}, \quad (6.50)$$

όπου D διαγώνιος πίνακας, όπως ορίζεται στην (6.41). Σύμφωνα με την παρατήρηση (1.1) ο επαναληπτικός πίνακας του Jacobi μπορεί να γραφεί και ως $J = I - D_A^{-1}A$. Με απευθείας υπολογισμό του J , λαμβάνοντας υπόψιν τη δομή του A σύμφωνα με το θεώρημα (6.9), έχουμε

$$\begin{aligned} J &= I - D_A^{-1}A = \\ &= \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix} - \begin{bmatrix} A_1^{-1} & 0 & 0 & 0 \\ 0 & A_1^{-1} & 0 & 0 \\ 0 & 0 & A_1^{-1} & 0 \\ 0 & 0 & 0 & A_1^{-1} \end{bmatrix} \begin{bmatrix} A_1 & 0 & DA_1 & 0 \\ 0 & A_1 & 0 & DA_1 \\ DA_1 & 0 & A_1 & 0 \\ 0 & DA_1 & 0 & A_1 \end{bmatrix} = \\ &= \begin{bmatrix} 0 & 0 & A_1^{-1}DA_1 & 0 \\ 0 & 0 & 0 & A_1^{-1}DA_1 \\ A_1^{-1}DA_1 & 0 & 0 & 0 \\ 0 & A_1^{-1}DA_1 & 0 & 0 \end{bmatrix} = \end{aligned}$$

$$= \begin{bmatrix} A_1^{-1} & 0 & 0 & 0 \\ 0 & A_1^{-1} & 0 & 0 \\ 0 & 0 & A_1^{-1} & 0 \\ 0 & 0 & 0 & A_1^{-1} \end{bmatrix} \begin{bmatrix} 0 & 0 & D & 0 \\ 0 & 0 & 0 & D \\ D & 0 & 0 & 0 \\ 0 & D & 0 & 0 \end{bmatrix} \begin{bmatrix} A_1 & 0 & 0 & 0 \\ 0 & A_1 & 0 & 0 \\ 0 & 0 & A_1 & 0 \\ 0 & 0 & 0 & A_1 \end{bmatrix}.$$

Δηλαδή

$$J = D_A^{-1} B_J D_A. \quad (6.51)$$

Επομένως ο J έχει το ίδιο φάσμα με τον B_J (όμοιοι πίνακες) που είναι ανεξάρτητος της βάσης. $\square$

Θεώρημα 6.11 Σε ένα τετράγωνο ο επαναληπτικός πίνακας του μπλοκ-Jacobi που αντιστοιχεί στον *collocation* πίνακα της *SFFS*, για οποιοδήποτε σύνολο συναρτήσεων βάσης, έχει τις εξής ιδιοτιμές

$$\{\pm e^{-r\pi}\}_{r=1}^{N_j} \quad (6.52)$$

αλγεβρικής πολλαπλότητας 2 έκαστη. Η φασματική ακτίνα του είναι ανεξάρτητη του πλήθους των συναρτήσεων βάσης και είναι ίση με $e^{-\pi} \approx 0.0432$.

Απόδειξη. Θα υπολογίσουμε τις ιδιοτιμές του πίνακα B_j όπως ορίζεται στην (6.50) με τους υποπίνακες D όπως ορίζονται στην (6.41). Ο πίνακας

$$P = \begin{bmatrix} I & -I & 0 & 0 \\ 0 & 0 & I & -I \\ I & I & 0 & 0 \\ 0 & 0 & I & I \end{bmatrix}$$

είναι αντιστρέψιμος με τον αντίστροφο του να είναι ο

$$P^{-1} = \frac{1}{2} \begin{bmatrix} I & 0 & I & 0 \\ -I & 0 & I & 0 \\ 0 & I & 0 & I \\ 0 & -I & 0 & I \end{bmatrix}.$$

Παρατηρούμε ότι

$$P^{-1} B_j P = \begin{bmatrix} D & 0 & 0 & 0 \\ 0 & -D & 0 & 0 \\ 0 & 0 & D & 0 \\ 0 & 0 & 0 & -D \end{bmatrix}. \quad (6.53)$$

Επομένως

$$\sigma(B_j) = \{\sigma(D), \sigma(-D), \sigma(D), \sigma(-D)\} = \quad (6.54)$$

$$= \{\pm e^{-r\pi}\}_{r=1}^{N_j} \cup \{\pm e^{-r\pi}\}_{r=1}^{N_j}. \quad (6.55)$$

Η μέγιστη, κατά απόλυτη τιμή, ιδιοτιμή του B_j είναι η $e^{-\pi}$ η οποία και μας δίνει και την φασματική ακτίνα. $\square$

Ο collocation πίνακας της SFFS είναι 2-cyclic consistently ordered (βλέπε πόρισμα (1.3)) επομένως άμεση συνέπεια του πορίσματος (1.9) και του θεωρήματος (1.10) σε συνδυασμό με το θεώρημα (6.11) είναι τα παρακάτω δυο πορίσματα

Πόρισμα 6.1 Σε ένα τετράγωνο ο επαναληπτικός πίνακας της μπλοκ-GS που αντιστοιχεί στον collocation πίνακα της SFFS έχει φασματική ακτίνα

$$\rho(T_{GS}) = e^{-2\pi} \approx 0.0019$$

ανεξάρτητη του πλήθους των συναρτήσεων βάσης, για οποιοδήποτε σύνολο συναρτήσεων βάσης.

Πόρισμα 6.2 Σε ένα τετράγωνο η βέλτιστη μπλοκ-SOR επαναληπτική μέθοδος που αντιστοιχεί στον collocation πίνακα της SFFS δίνεται για

$$\omega_{opt} = \frac{2}{1 + \sqrt{1 - e^{-2\pi}}} \approx 1.0005$$

και έχει φασματική ακτίνα

$$\rho(\mathcal{L}_{\omega_{opt}}) = \omega_{opt} - 1 = \frac{1 - \sqrt{1 - e^{-2\pi}}}{1 + \sqrt{1 - e^{-2\pi}}} \approx 0.0005$$

ανεξάρτητα από το πλήθος των συναρτήσεων βάσης, και για οποιοδήποτε σύνολο συναρτήσεων βάσης.

Παρατήρηση 6.8 Στην μέθοδο *sine-SFFS* οι σημειακές και οι μπλοκ κλασικές επαναληπτικές μέθοδοι ταυτίζονται. Αυτό οφείλεται στο ότι τα διαγώνια μπλοκ του collocation πίνακα είναι διαγώνιοι πίνακες. Επομένως το θεώρημα (6.11) και τα πορίσματα (6.1) και (6.2) ισχύουν ως έχουν και για τις *point* επαναληπτικές μεθόδους.

Παρατήρηση 6.9 Για την μέθοδο *sine-SFFS*, λόγω του ότι ο collocation πίνακας αποτελείται από τρεις διαγωνίους, ο αντίστροφος του (6.40) είναι ο

$$A^{-1} = \begin{bmatrix} C & 0 & F & 0 \\ 0 & C & 0 & F \\ F & 0 & C & 0 \\ 0 & F & 0 & C \end{bmatrix},$$

όπου

$$\begin{cases} C(i) = \frac{1}{A_1(i) [1 - D^2(i)]} \\ F(i) = -D(i) C(i) \end{cases},$$

Αυτό οφείλεται στο ότι τα διαγώνια μπλοκ του collocation πίνακα είναι διαγώνιοι πίνακες. Επομένως το collocation σύστημα μπορεί να λυθεί άμεσα με κόστος της τάξεως του $\mathcal{O}(N)$.

6.4 Συμπεράσματα

Σε αυτό το κεφάλαιο μελετήσαμε την δομή και τις ιδιότητες του collocation πίνακα που προέρχεται από την διακριτοποίηση της Dirichlet to Neumann συνθήκης με την μέθοδο SFES. Καταλήξαμε σε ορισμένα συμπεράσματα που μπορούν να βοηθήσουν στην κατεύθυνση της αποδοτικότερης επίλυσης του collocation συστήματος.

Το βασικό συμπέρασμα είναι ότι στην περίπτωση των κανονικών πολυγώνων με ίδιες συνοριακές συνθήκες σε όλες τις πλευρές ο collocation πίνακας είναι block-circulant. Επίσης παρατηρήσαμε μια επιπλέον, κάπως ιδιότυπη, μορφή συμμετρίας ανάμεσα στα μπλοκ που "ισαπέχουν" από το κεντρικό μπλοκ της πρώτης (μπλοκ) γραμμής του πίνακα.

Ο collocation πίνακας δεν επηρεάζεται κατά την μεταφορά, στροφή ή μεγέθυνση του πολυγώνου ή ακόμα και κατά την ομοιόμορφη μεταβολή των συνοριακών συνθηκών σε όλες τις πλευρές.

Ο collocation πίνακας είναι ανεξάρτητος των συνοριακών συνθηκών αν αυτές είναι ίδιες σε όλες τις πλευρές.

Όταν το πολυγωνικό χωρίο στο οποίο λύνουμε την εξίσωση του Laplace είναι το τετράγωνο ο collocation πίνακας είναι, με την φυσική του διαμέριση, 2-cyclic. Αυτό σε συνδυασμό με το γεγονός ότι έχει ιδιαίτερα απλή μορφή μας επέτρεψε την αναλυτική εύρεση των ιδιοτιμών των επαναληπτικών πινάκων των μεθόδων Jacobi, GS και SOR και τον προσδιορισμό της βέλτιστης overrelaxation παραμέτρου της SOR.

Στην περίπτωση που δουλεύουμε με διαφορετικές συνοριακές συνθήκες σε κάθε πλευρά ο collocation πίνακας μπορεί να γραφεί ως συνδυασμός δυο απλούστερων block-circulant collocation πινάκων.

Κεφάλαιο 7

Επίλυση της Dirichlet to Neumann Απεικόνισης με Χρήση FFT

Στο κεφάλαιο (6) δείξαμε ότι ο collocation πίνακας που προκύπτει από την διακριτοποίηση της γενικευμένης Dirichlet to Neumann Map, στην περίπτωση των κανονικών πολυγώνων με τις ίδιες συνοριακές συνθήκες σε όλες τις πλευρές, είναι block-circulant. Αυτή η ιδιαίτερη δομή μπορεί να μας οδηγήσει σε ταχύτερες μεθόδους για την επίλυση του collocation συστήματος. Σε αυτό το κεφάλαιο εξετάζουμε τρόπους βελτίωσης των άμεσων και των επαναληπτικών μεθόδων χρησιμοποιώντας ταχύς μετασχηματισμούς Fourier που είναι στενά συνδεδεμένοι με την block-circulant ιδιότητα του collocation πίνακα [SSP, SFPS, VES97, MGE83].

7.1 Block – Circulants και FFT

Έστω A ένας block-circulant πίνακας

$$A = \begin{bmatrix} A_1 & A_2 & \cdots & A_n \\ A_n & A_1 & \cdots & A_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ A_2 & A_3 & \cdots & A_1 \end{bmatrix}, \quad (7.1)$$

όπου $A_i \in \mathbb{R}^{N \times N}$. Είναι γνωστό [TIA00] ότι για κάθε block-circulant πίνακα ισχύει η παρακάτω σχέση

$$(F_n \otimes I_N)^* A (F_n \otimes I_N) = nJ \quad (7.2)$$

όπου

$$J = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_n \end{bmatrix}, \quad (7.3)$$

και

$$J_t = A_1 + \omega^{(t-1)} A_2 + \omega^{(t-1)^2} A_3 + \omega^{(t-1)^3} A_4 + \cdots + \omega^{(t-1)(n-1)} A_n, \quad t = 1, \dots, n. \quad (7.4)$$

με

$$\omega_n = e^{-i\frac{2\pi}{n}} = \cos \frac{2\pi}{n} - i \sin \frac{2\pi}{n} \quad (7.5)$$

Έυκολα μπορεί να παρατηρήσει κανείς ότι

$$\begin{bmatrix} J_1 \\ J_2 \\ \vdots \\ J_n \end{bmatrix} = (F_n \otimes I_N) \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix}. \quad (7.6)$$

Από την σχέση (7.2) προκύπτει η σχέση

$$A = \frac{1}{n} (F_n \otimes I_N) J (F_n \otimes I_N)^*. \quad (7.7)$$

Εδώ πρέπει να παρατηρήσουμε ότι και η σχέση (7.6) για τον υπολογισμό των J_t και η σχέση (7.7) περιέχουν πολλαπλασιασμούς με τον πίνακα $(F_n \otimes I_N)$ ¹. Το επόμενο λήμμα θα μας δώσει έναν υπολογιστικά αποδοτικό τρόπο υπολογισμού τέτοιων πολλαπλασιασμών.

Λήμμα 7.1 Για κάθε φυσικό n και για κάθε άρτιο φυσικό N υπάρχει μεταθετικός πίνακας P τέτοιος ώστε

$$F_n \otimes I_N = P^t \left(\bigoplus_{r=1}^N F_n \right) P \quad (7.8)$$

και

$$(F_n \otimes I_N)^* = P^t \left(\bigoplus_{r=1}^N F_n^* \right) P. \quad (7.9)$$

Απόδειξη. Ορίζω τον μεταθετικό πίνακα ως εξής

$$P = \begin{bmatrix} P_1 \\ \vdots \\ P_r \\ \vdots \\ P_N \end{bmatrix}, \quad P_r = I_n \otimes e_r^T \in \mathbb{R}^{n \times nN}, \quad r = 1, \dots, n, \quad j = 1, \dots, n,$$

¹ Και με τον συζυγοανάστροφο του.

όπου το e_r συμβολίζει το r -οστό μοναδιαίο διάνυσμα του $\mathbb{R}^{N \times 1}$. Παρατηρούμε ότι η (7.8) ισχύει αφού

$$\begin{aligned}
 P(F_n \otimes I_N)P^T &= \begin{bmatrix} I_n \otimes e_1^T \\ \vdots \\ I_n \otimes e_N^T \end{bmatrix} (F_n \otimes I_N) (I_n \otimes e_1 \cdots I_n \otimes e_N) \\
 &= \begin{bmatrix} I_n \otimes e_1^T \\ \vdots \\ I_n \otimes e_N^T \end{bmatrix} (F_n \otimes e_1 \cdots F_n \otimes e_N) \\
 &= \begin{bmatrix} F_n & 0 & \cdots & 0 \\ 0 & F_n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & F_n \end{bmatrix} \\
 &= \bigoplus_{k=1}^N F_n.
 \end{aligned}$$

Αν στην προηγούμενη σχέση αντί για τον F_n χρησιμοποιήσουμε τον F_n^* αποδεικνύεται και η (7.9). $\square$

Το κέρδος που αποκομίζουμε από το προηγούμενο λήμμα είναι ότι έχουμε μειώσει το κόστος των γινομένων της μορφής $(F_n \otimes I_N) \vec{x}$ από $\mathcal{O}(n^2 N^2)$ σε $\mathcal{O}(Nn \log n)$ ^{1 2}.

7.1.1 FFTmul

Η παραγοντοποίηση (7.7) του πίνακα A μπορεί να γραφεί ως

$$A = nP^t \left(\bigoplus_{r=1}^N F_n \right) PJP^t \left(\bigoplus_{r=1}^N F_n^* \right) P. \quad (7.10)$$

Η τελευταία σχέση μπορεί να χρησιμοποιηθεί στις Krylov επαναληπτικές μεθόδους για την μείωση του κόστους τους. Ας θυμηθούμε ότι σε αυτή την κατηγορία των μεθόδων το κυρίως κόστος προέρχεται από τον πολλαπλασιασμό του πίνακα των συντελεστών του γραμμικού συστήματος με διανύσματα. Ο πίνακας (T11) μας δείχνει την διαφορά στον χρόνο επίλυσης με επαναληπτικές μεθόδους, χωρίς προρύθμιση, με τους δυο τρόπους υπολογισμού των γινομένων $A\vec{x}$: με απευθείας υπολογισμό με χρήση του πίνακα A (Mvmul) και με χρήση της παραγοντοποίησης (7.10) (FFTmul). Από αυτόν τον πίνακα βλέπουμε καθαρά ότι σε μεγάλες περιπτώσεις η μείωση του χρόνου, χρησιμοποιώντας ταχείς μετασχηματισμούς Fourier, ξεπερνάει το 95%.

¹ Θεωρούμε ότι το κόστος των μεταθέσεων είναι μηδαμινό ως προς το κόστος των πράξεων. ² Σε αυτό το κεφάλαιο θεωρούμε ότι $\log := \log_2$.

FFTmul vs Mvmul														
Κανονικό 8-γωνο							Κανονικό 24-γωνο							
N = 32		N = 64		N = 128			N = 32		N = 64		N = 128			
	diff	%	diff	%	diff		%		diff	%	diff	%		
Bicgstab	-0.008	318	-0.015	169	0.338		38	Bicgstab	0.822	7	8.196	4	28.500	4
Gmres	-0.008	263	-0.015	160	0.369		38	Gmres	0.817	9	8.432	4	31.490	5
Κανονικό 12-γωνο							Κανονικό 31-γωνο							
N = 32		N = 64		N = 128			N = 32		N = 64		N = 128			
	diff	%	diff	%	diff		%		diff	%	diff	%		
Bicgstab	-0.007	143	0.413	16	5.598		9	Bicgstab	5.102	6	31.100	4	92.540	5
Gmres	-0.007	134	0.408	18	5.209		9	Gmres	10.753	5	47.950	4	157.680	4
Κανονικό 19-γωνο						Κανονικό 36-γωνο								
N = 32		N = 64		N = 128		N = 32		N = 64		N = 128				
	diff	%	diff	%	diff	%		diff	%	diff	%			
Bicgstab	-0.014	124	4.074	8	18.040	7	Bicgstab	12.297	2	42.960	3	145.720	3	
Gmres	-0.013	119	4.598	7	17.370	8	Gmres	35.225	2	123.930	2	385.900	3	

Πίνακας T11: *FFTmul vs Mvmul*: Διαφορά στον χρόνο επίλυσης του γραμμικού συστήματος με επαναληπτικές μεθόδους (α) χρησιμοποιώντας απευθείας πολλαπλασιασμό πίνακα επί διάνυσμα (*Mvmul*) και (β) χρησιμοποιώντας ταχείς μετασχηματισμούς Fourier (*FFTmul*) για τον πολλαπλασιασμό πίνακα επί διάνυσμα. Σε κάθε περίπτωση η πρώτη στήλη (*diff*) δείχνει την διαφορά στους χρόνους επίλυσης σε δευτερόλεπτα δηλαδή $diff = 0$ χρόνος επίλυσης χρησιμοποιώντας *Mvmul* μείον το χρόνο επίλυσης χρησιμοποιώντας *FFTmul*, ενώ η δεύτερη στήλη δείχνει τον λόγο του χρόνου επίλυσης με *FFTmul* προς τον χρόνο επίλυσης με *Mvmul* εκφρασμένο ως ποσοστό επί τοις εκατό.

7.1.2 FFTinv

Χρησιμοποιώντας την σχέση (7.7) μπορούμε να παράξουμε μια άμεση μέθοδο για την επίλυση γενικών block circulant γραμμικών συστημάτων. Έστω λοιπόν ότι ζητάμε την λύση του block circulant συστήματος $A\vec{x} = \vec{b}$.

$$\begin{aligned}
A\vec{x} &= \vec{b} \Leftrightarrow \\
\frac{1}{n} (F_n \otimes I_N) J (F_n \otimes I_N)^* \vec{x} &= \vec{b} \Leftrightarrow \\
(F_n \otimes I_N) J (F_n \otimes I_N)^* \vec{x} &= n \vec{b} \Leftrightarrow \\
J (F_n \otimes I_N)^* \vec{x} &= n (F_n \otimes I_N)^{-1} \vec{b} \Leftrightarrow \\
J (F_n \otimes I_N)^* \vec{x} &= n (F_n^{-1} \otimes I_N) \vec{b} \Leftrightarrow \\
J (F_n \otimes I_N)^* \vec{x} &= n \left(\frac{1}{n} F_n^* \otimes I_N \right) \vec{b} \Leftrightarrow \\
J (F_n \otimes I_N)^* \vec{x} &= n \left(\frac{1}{n} F_n \otimes I_N \right)^* \vec{b} \Leftrightarrow \\
J (F_n \otimes I_N)^* \vec{x} &= (F_n \otimes I_N)^* \vec{b}.
\end{aligned}$$

Επομένως ο αλγόριθμος που προτείνουμε (FFTin_v) για την λύση block circulant γραμμικών συστημάτων της μορφής

$$\begin{bmatrix} A_1 & A_2 & \cdots & A_n \\ A_n & A_1 & \cdots & A_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ A_2 & A_3 & \cdots & A_1 \end{bmatrix} \vec{x} = \vec{b} \quad (7.11)$$

μπορεί να περιγραφεί ως εξής:

Αλγόριθμος 7.1 (FFTin_v)

$$\begin{aligned} 1 & \text{ Υπολόγισε τα } J_t \text{ από την σχέση } \begin{bmatrix} J_1 \\ \vdots \\ J_n \end{bmatrix} = P^t \left(\bigoplus_{k=1}^N F_n \right) P \begin{bmatrix} A_1 \\ \vdots \\ A_n \end{bmatrix} \\ 2 & \text{ Θέσε } \vec{c} = \frac{1}{n} P^t \left(\bigoplus_{k=1}^N F_n^* \right) P \vec{b} \\ 3 & \text{ Λύσε (άμεσα) } \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_n \end{bmatrix} \vec{y} = \vec{c} \\ 4 & \vec{x} = P^t \left(\bigoplus_{k=1}^N F_n \right) P \vec{y} \end{aligned}$$

Το βασικό κόστος, σε πράξεις, της μεθόδου αυτής (FFTin_v) είναι $\mathcal{O}(N^2 n \log n)$ για την κατασκευή του πίνακα J (βήμα 1) και $\mathcal{O}(nN^3)$ για την επίλυση των n μιγαδικών συστημάτων (βήμα 3)¹. Δηλαδή το συνολικό κόστος είναι $\mathcal{O}(N^2 n \log n + nN^3)$. Το κόστος στην περίπτωση της άμεσης επίλυσης του γραμμικού συστήματος (dgesv) είναι $\mathcal{O}(n^3 N^3)$. Για να διαπιστώσουμε, χονδρικά, το μέγεθος της διαφοράς σε πράξεις μεταξύ των δυο αυτών μεθόδων παραθέτουμε ένα απλό πίνακα που δείχνει τον λόγο $r = \frac{N^2 n \log n + nN^3}{n^3 N^3}$ επί τοις εκατό για διάφορες τιμές των n, N .

¹ Θεωρούμε ότι η, αριθμητική, επίλυση των γραμμικών συστημάτων γίνεται με την άμεση μέθοδο zgesv του πακέτου Lapack

	$N = 16$	$N = 128$
$n = 8$	1.85	1.60
$n = 16$	0.48	0.40
$n = 24$	0.22	0.18
$n = 32$	0.13	0.10
$n = 40$	0.08	0.07
$n = 48$	0.06	0.05

Βέβαια πρέπει να προσθέσουμε ότι οι πράξεις που απαιτούνται στην FFTinv είναι μεταξύ μιγαδικών αριθμών και αυτό θα αυξήσει τον λόγο r (κατά μια σταθερά μεγαλύτερη από 1 και μικρότερη από 4 ανάλογα με την αρχιτεκτονική του επεξεργαστή) αλλά και πάλι περιμένουμε σημαντική βελτίωση στον χρόνο. Ένα άλλο πλεονέκτημα της FFTinv είναι ότι χρειαζόμαστε ένα κλάσμα μόνο του χώρου που απαιτείται για την αποθήκευση του πίνακα των συντελεστών A σε σχέση με την dgesv αφού κατ' ουσίαν χρησιμοποιούμε μόνο την πρώτη μπλοκ γραμμή του πίνακα $A_1, \dots, A_n$.¹ Ο πίνακας (T12) δείχνει τους χρόνους επίλυσης με τις δυο μεθόδους για διάφορες περιπτώσεις των n και N . Από αυτόν τον πίνακα εύκολα βλέπουμε ότι ακόμα και στις μικρές περιπτώσεις (π.χ. $n = 8$, $N = 16$) το κέρδος από την χρήση της FFTinv είναι σημαντικό. Στις μεγάλες περιπτώσεις βέβαια φαίνεται καθαρά η ανωτερότητα της FFTinv. Για να γίνει πιο κατανοητή η διαφορά στους χρόνους απλά αναφέρουμε ότι αν για κάποιον υπολογιστή η επίλυση του collocation συστήματος για το 48-γωνο με 128 συναρτήσεις βάσης ανά πλευρά διαρκεί 15 δευτερόλεπτα με την μέθοδο FFTinv με την μέθοδο dgesv θα διαρκούσε μια ολόκληρη μέρα!

FFTinv vs Dgesv						
	N = 16			N = 128		
	Dgesv	FFTinv	%	Dgesv	FFTinv	%
Κανονικό 8-γωνο	0.023	0.002	8.63	26	0.345	1.35
Κανονικό 16-γωνο	0.184	0.004	2.01	470	0.742	0.16
Κανονικό 24-γωνο	0.622	0.006	0.99	1540	1.15	0.07
Κανονικό 32-γωνο	1.490	0.008	0.54	4390	1.57	0.04
Κανονικό 40-γωνο	3.090	0.010	0.33	7560	1.99	0.03
Κανονικό 48-γωνο	9.180	0.013	0.14	13900	2.61	0.02

Πίνακας T12: FFTinv vs Dgesv: Οι χρόνοι επίλυσης του collocation συστήματος για διάφορα κανονικά πολύγωνα και για $N = 16, 128$ (α) με την Dgesv του πακέτου Lapack και (β) με την FFTinv. Η τρίτη στήλη δείχνει τον λόγο του χρόνου επίλυσης της FFTinv προς τον χρόνο επίλυσης της Dgesv εκφρασμένο ως ποσοστό επί τοις εκατό.

¹ Οι μεταβλητές J_t και A_t μπορούν να αποθηκευτούν στον ίδιο, μιγαδικό, χώρο αποθήκευσης.

7.1.3 FFTinv(Iterative)

Η μέθοδος FFTinv είναι άμεση διότι στο βήμα 3 όπου επιλύονται τα n μιγαδικά γραμμικά συστήματα χρησιμοποιούμε μια άμεση μέθοδο. Είναι στην διακριτική μας ευχέρεια να μην χρησιμοποιήσουμε άμεση μέθοδο για την επίλυση αυτών των μιγαδικών συστημάτων αλλά κάποια επαναληπτική. Σε αυτή την περίπτωση η FFTinv μέθοδος μπορεί να ειπωθεί ως επαναληπτική. Ειδικότερα έστω $Ax = b$ το κάθε ένα από τα n μιγαδικά συστήματα που έχουμε να επιλύσουμε στο τρίτο βήμα του αλγορίθμου (7.1). Δηλαδή έστω ότι θέλουμε να επιλύσουμε το μιγαδικό σύστημα

$$Ax = b, \quad A \in \mathbb{C}^{N \times N}, \quad x \in \mathbb{C}^{N \times 1}, \quad b \in \mathbb{C}^{N \times 1}. \quad (7.12)$$

Για να μετατρέψουμε το μιγαδικό σύστημα (7.12) σε ένα ισοδύναμο πραγματικό θεωρούμε τις παρακάτω διασπάσεις των A, x και b στο πραγματικό και το φανταστικό τους μέρος

$$\begin{cases} A_R = \operatorname{Re}(A) \\ A_I = \operatorname{Im}(A) \end{cases}, \quad \begin{cases} x_R = \operatorname{Re}(x) \\ x_I = \operatorname{Im}(x) \end{cases}, \quad \begin{cases} b_R = \operatorname{Re}(b) \\ b_I = \operatorname{Im}(b) \end{cases}.$$

Επομένως το μιγαδικό σύστημα (7.12) είναι ισοδύναμο με το

$$(A_R + iA_I)(x_R + ix_I) = b_R + ib_I, \quad (7.13)$$

ή

$$\begin{cases} A_R x_R - A_I x_I = b_R \\ A_I x_R + A_R x_I = b_I \end{cases},$$

ή

$$\begin{bmatrix} A_R & -A_I \\ A_I & A_R \end{bmatrix} \begin{bmatrix} x_R \\ x_I \end{bmatrix} = \begin{bmatrix} b_R \\ b_I \end{bmatrix}. \quad (7.14)$$

Οπότε η επίλυση του μιγαδικού γραμμικού συστήματος (7.12) είναι ισοδύναμη με την επίλυση του πραγματικού (7.14). Εδώ πρέπει να παρατηρήσουμε ότι το πραγματικό σύστημα (7.14) είναι διάστασης $2N \times 2N$ ενώ το μιγαδικό (7.12) ήταν διάστασης $N \times N$.¹

Ο πίνακας (T13) δείχνει τους χρόνους επίλυσης του collocation συστήματος με την μέθοδο FFTinv χρησιμοποιώντας, για το 3ο βήμα του αλγορίθμου (α) άμεση επίλυση των μιγαδικών συστημάτων και (β) επαναληπτική επίλυση μέσω της σχέσης (7.13), για διάφορες περιπτώσεις των n και N . Παρατηρούμε ότι για την περίπτωση $N = 16$ η επαναληπτική μέθοδος χρειάζεται τον διπλάσιο χρόνο από την άμεση ενώ για $N = 128$ χρειάζεται τον μισό χρόνο. Δηλαδή για μεγάλα N η χρήση επαναληπτικών μεθόδων μπορεί να βελτιώσει την ταχύτητα της FFTinv. Εδώ πρέπει να σημειώσουμε ότι υπάρχουν περιθώρια βελτίωσης της επαναληπτικής διαδικασίας είτε με την χρήση complex Krylov μεθόδων είτε με την χρήση κατάλληλων προορυθμιστών.

¹ Η μετατροπή στο πραγματικό σύστημα (7.14) δεν είναι ο μοναδικός τρόπος επίλυσης ενός μιγαδικού συστήματος της μορφής (7.12). Υπάρχουν και άλλα ισοδύναμα πραγματικά συστήματα [DHE01] ακόμα και υλοποιήσεις των Krylov επαναληπτικών μεθόδων που επιλύουν μιγαδικά συστήματα.

FFTinv (Iterative) vs FFTinv (Direct)						
	N = 16			N = 128		
	Bicgstab	Zgesv	%	Bicgstab	Zgesv	%
Κανονικό 8-γωνο	0.005	0.002	252.85	0.13	0.36	37.05
Κανονικό 16-γωνο	0.008	0.004	225.14	0.30	0.74	40.54
Κανονικό 24-γωνο	0.013	0.006	208.20	0.51	1.17	43.68
Κανονικό 32-γωνο	0.016	0.008	202.99	0.72	1.57	45.54
Κανονικό 40-γωνο	0.021	0.010	204.81	0.94	1.99	47.39
Κανονικό 48-γωνο	0.026	0.013	202.34	1.35	2.59	52.12

Πίνακας T13: *FFTinv(Iterative) vs FFTinv(Direct)*: Οι χρόνοι επίλυσης του collocation συστήματος για διάφορα κανονικά πολύγωνα και για $N = 16, 128$ με την μέθοδο *FFTinv* επιλύοντας τα μιγαδικά συστήματα $J_t y = c$ (α) με την *zgesv* του πακέτου *Lapack* και (β) με την επαναληπτική μέθοδο *Bicgstab* χωρίς προρύθμιση. Η τρίτη στήλη, σε κάθε περίπτωση, δείχνει τον λόγο του χρόνου επίλυσης χρησιμοποιώντας την επαναληπτική μέθοδο προς τον χρόνο επίλυσης χρησιμοποιώντας την άμεση μέθοδο εκφρασμένο ως ποσοστό επί τοις εκατό.

7.1.4 FFT για Διαφορετικές Συνοριακές Συνθήκες σε Κάθε Πλευρά

Σε αυτή την παράγραφο θα εξετάσουμε την δυνατότητα χρήσης FFTs στην επίλυση collocation συστημάτων που προκύπτουν από την διακριτοποίηση της Dirichlet-Neumann απεικόνιση σε κανονικά πολύγωνα με διαφορετικές συνοριακές συνθήκες σε κάθε πλευρά. Στην τρέχουσα παράγραφο θα παρακολουθήσουμε τους συμβολισμούς της παραγράφου (6.2.3).

Στην παράγραφο (6.2.3) είδαμε ένα τρόπο γραφής του collocation πίνακα (K) στην περίπτωση των διαφορετικών συνοριακών συνθηκών σε κάθε πλευρά συναρτήσει του collocation πίνακα (A) στην περίπτωση των ίδιων συνοριακών συνθηκών και του συγγενή του collocation πίνακα (B). Υπενθυμίζουμε αυτήν την σχέση (6.34)

$$K = (C_n \otimes I_N) A (C_n \otimes I_N) + (S_n \otimes I_N) A (S_n \otimes I_N) + \quad (7.15)$$

$$+ (C_n \otimes I_N) M B (S_n \otimes I_N) - (S_n \otimes I_N) M B (C_n \otimes I_N) \quad (7.16)$$

Η παραπάνω διάσπαση του collocation πίνακα (7.15) μπορεί να χρησιμοποιηθεί για να εισάγουμε FFTs στην περίπτωση των Krylov επαναληπτικών μεθόδων αφού και οι δύο πίνακες A και B είναι block-circulant. Γυρνώντας όμως λίγο πίσω στην παράγραφο (2.2) όπου κατασκευάσαμε την αριθμητική μέθοδο και πιο συγκεκριμένα στην σχέση (2.42) παρατηρούμε ότι από κάθε εξίσωση μπορεί να απαλειφθεί ο όρος $e^{-i\beta_p}$. Η σχέση (6.29)

$$K_p^j = \cos(\beta_j - \beta_p) A_p^j + \sin(\beta_j - \beta_p) M_N B_p^j,$$

απαλείφοντας τον όρο β_p , γράφεται ως

$$K_p^j = \cos(\beta_j) A_p^j + \sin(\beta_j) M_N B_p^j$$

Δηλαδή ισχύει ότι

$$\begin{aligned}
 K &= \begin{bmatrix} \cos \beta_1 A_1^1 & \cos \beta_2 A_1^2 & \cdots & \cos \beta_n A_1^n \\ \cos \beta_1 A_2^1 & \cos \beta_2 A_2^2 & \cdots & \cos \beta_n A_2^n \\ \vdots & \vdots & \ddots & \vdots \\ \cos \beta_1 A_n^1 & \cos \beta_2 A_n^2 & \cdots & \cos \beta_n A_n^n \end{bmatrix} + \begin{bmatrix} \sin \beta_1 M_N B_1^1 & \sin \beta_2 M_N B_1^2 & \cdots & \sin \beta_n M_N B_1^n \\ \sin \beta_1 M_N B_2^1 & \sin \beta_2 M_N B_2^2 & \cdots & \sin \beta_n M_N B_2^n \\ \vdots & \vdots & \ddots & \vdots \\ \sin \beta_1 M_N B_n^1 & \sin \beta_2 M_N B_n^2 & \cdots & \sin \beta_n M_N B_n^n \end{bmatrix} \Leftrightarrow \\
 K &= \begin{bmatrix} A_1^1 & A_1^2 & \cdots & A_1^n \\ A_2^1 & A_2^2 & \cdots & A_2^n \\ \vdots & \vdots & \ddots & \vdots \\ A_n^1 & A_n^2 & \cdots & A_n^n \end{bmatrix} \begin{bmatrix} \cos \beta_1 & & & \\ & \cos \beta_2 & & \\ & & \ddots & \\ & & & \cos \beta_n \end{bmatrix} + \bullet \\
 &+ \begin{bmatrix} M_N B_1^1 & M_N B_1^2 & \cdots & M_N B_1^n \\ M_N B_2^1 & M_N B_2^2 & \cdots & M_N B_2^n \\ \vdots & \vdots & \ddots & \vdots \\ M_N B_n^1 & M_N B_n^2 & \cdots & M_N B_n^n \end{bmatrix} \begin{bmatrix} \sin \beta_1 & & & \\ & \sin \beta_2 & & \\ & & \ddots & \\ & & & \sin \beta_n \end{bmatrix} \Leftrightarrow \\
 K &= \begin{bmatrix} A_1^1 & A_1^2 & \cdots & A_1^n \\ A_2^1 & A_2^2 & \cdots & A_2^n \\ \vdots & \vdots & \ddots & \vdots \\ A_n^1 & A_n^2 & \cdots & A_n^n \end{bmatrix} (C_n \otimes I_N) + (M_N \otimes I_N) \begin{bmatrix} B_1^1 & B_1^2 & \cdots & B_1^n \\ B_2^1 & B_2^2 & \cdots & B_2^n \\ \vdots & \vdots & \ddots & \vdots \\ B_n^1 & B_n^2 & \cdots & B_n^n \end{bmatrix} (S_n \otimes I_N) \Leftrightarrow \\
 K &= \begin{bmatrix} A_1^1 & A_1^2 & \cdots & A_1^n \\ A_2^1 & A_2^2 & \cdots & A_2^n \\ \vdots & \vdots & \ddots & \vdots \\ A_n^1 & A_n^2 & \cdots & A_n^n \end{bmatrix} (C_n \otimes I_N) + (M_N \otimes I_N) \begin{bmatrix} B_1^1 & B_1^2 & \cdots & B_1^n \\ B_2^1 & B_2^2 & \cdots & B_2^n \\ \vdots & \vdots & \ddots & \vdots \\ B_n^1 & B_n^2 & \cdots & B_n^n \end{bmatrix} (S_n \otimes I_N)
 \end{aligned}$$

Δηλαδή ο "απαλλαγμένος από τον όρο $e^{-i\beta_p}$ " collocation πίνακας K μπορεί να γραφεί ως

$$K = A(C_n \otimes I_N) + MB(S_n \otimes I_N) \quad (7.17)$$

Η σχέση (7.17) είναι προτιμότερη από την (7.15) ως προς τη χρήση FFTs διότι περιέχει δυο μόνο όρους σε αντίθεση με την (7.15) η οποία περιέχει τέσσερις όρους. Επομένως το γινόμενο του collocation πίνακα K με ένα διάνυσμα μπορεί να γραφεί ως εξής

$$\begin{aligned}
 \vec{y} &= K \vec{x} \Leftrightarrow \\
 \vec{y} &= (A(C_n \otimes I_N) + B(S_n \otimes I_N)) \vec{x} \Leftrightarrow
 \end{aligned}$$

$$\begin{cases} 1. \vec{y}_1 = (C_n \otimes I_N) \vec{x} \\ 2. \vec{y}_2 = (S_n \otimes I_N) \vec{x} \\ 3. \vec{y}_1 = A \vec{y}_1 \\ 4. \vec{y}_2 = B \vec{y}_2 \\ 5. \vec{y} = \vec{y}_1 + \vec{y}_2 \end{cases} \quad (7.18)$$

όπου στον πίνακα B έχουμε ενσωματώσει και τον πίνακα M (δηλαδή $B \leftarrow MB$). Τα βήματα 3 και 4 του παραπάνω τρόπου πολλαπλασιασμού μπορούν να υλοποιηθούν με την χρήση FFTs αφού και οι δυο πίνακες A και B είναι block-circulant. Ο πολλαπλασιασμός αυτός μπορεί να χρησιμοποιηθεί μόνο στις Krylov επαναληπτικές μεθόδους για την βελτίωση του χρόνου εκτελέσεως των μεθόδων. Άμεση μέθοδο, ανάλογη της FFTinv, όπως είχαμε στην περίπτωση των ίδιων συνοριακών συνθηκών δεν μπορούμε να έχουμε. Στον πίνακα (T14) βλέπουμε τα αποτελέσματα της επίλυσης του collocation συστήματος για διάφορα κανονικά πολύγωνα με διαφορετικές συνοριακές συνθήκες σε κάθε πλευρά χρησιμοποιώντας απευθείας πολλαπλασιασμό (Mvmul) και χρησιμοποιώντας την σχέση (7.18) με χρήση των FFTs (FFTmul). Οι συνοριακές συνθήκες σε κάθε πλευρά S_j δίνονται από τον αναδρομικό τύπο

$$\beta_j = \beta_{j-1} + \frac{\pi}{n}, \quad \beta_1 = \pi \quad (7.19)$$

Από τον πίνακα μπορούμε να συμπεράνουμε (T14) ότι

- Για μικρά πολύγωνα η χρήση των FFTs αυξάνει το χρόνο επίλυσης των γραμμικών συστημάτων σε όλες τις μεθόδους. Αυτό οφείλεται στο ότι ο πίνακας K έχει σχετικά μικρή διάσταση (π.χ. 4×128).
- Στα μέτριοι μεγέθους πολύγωνα (8-γωνο, 16-γωνο και 24-γωνο) οι προορυθμισμένες μέθοδοι είναι πάντα ταχύτερες με την χρήση FFTs ενώ οι μη-προορυθμισμένες είναι ταχύτερες για μεγάλα N .
- Στα μεγάλα πολύγωνα (32-γωνο και 40-γωνο) η χρήση των FFTs μειώνει το κόστος επίλυσης σε όλες τις περιπτώσεις. Στις περιπτώσεις που μας ενδιαφέρουν περισσότερο (μεγάλα N - μεγάλοι χρόνοι επίλυσης) το κέρδος από την χρήση των FFTs ξεπερνάει το 90 τοις εκατό για τις μη προορυθμισμένες μεθόδους. Παρατηρούμε ότι, για $N = 128$, το κέρδος στις μη προορυθμισμένες μεθόδους είναι πολλαπλάσιο από αυτό των προορυθμισμένων. Αυτό οφείλεται στο γεγονός ότι το κόστος των προορυθμισμένων μεθόδων εξαρτάται, εκτός από το πολλαπλασιασμό του collocation πίνακα με ένα διάνυσμα, και από την επίλυση κάποιων γραμμικών συστημάτων που σχετίζονται με την εφαρμογή του προορυθμιστή.

Εδώ πρέπει να σημειώσουμε ότι η συμπεριφορά των επαναληπτικών μεθόδων εξαρτάται ισχυρά από τις συνοριακές συνθήκες που χρησιμοποιούμε σε κάθε πλευρά. Πάντως το κέρδος από την χρήση των FFTs είναι δεδομένη, στις μεγάλες περιπτώσεις τουλάχιστον, αφού αυτό που μπορεί να γίνει με την αλλαγή των συνοριακών συνθηκών είναι να μειωθούν οι επαναλήψεις που απαιτούνται οπότε θα μειωθεί ο συνολικός χρόνος επίλυσης των γραμμικών αναλογικά και στις δυο περιπτώσεις (Mvmul και FFTmul).

Διαφορετικές Συνοριακές Συνθήκες σε κάθε Πλευρά (FFTmul vs Mvmtul)										
Κανονικό 4-γωνο						Κανονικό 24-γωνο				
	N = 16		N = 128				N = 16		N = 128	
	diff	%	diff	%			diff	%	diff	%
Bicgstab	0.004	94	-3.601	722		Bicgstab	-0.039	116	324.8	7
Gmres	-0.043	418	-1.841	599		Gmres	-0.011	107	187.5	7
Bicgstab (SGS)	-0.004	170	-0.189	142		Bicgstab (SGS)	0.305	34	77.0	66
Gmres (SGS)	-0.002	162	-0.141	140		Gmres (SGS)	0.705	36	184.0	65
Κανονικό 8-γωνο						Κανονικό 32-γωνο				
	N = 16		N = 128				N = 16		N = 128	
	diff	%	diff	%			diff	%	diff	%
Bicgstab	-0.047	290	9.430	46	Bicgstab	0.164	73	653.0	5	
Gmres	-0.029	208	5.440	47	Gmres	0.149	60	387.6	5	
Bicgstab (SGS)	0.006	57	0.370	89	Bicgstab (SGS)	1.632	22	379.0	61	
Gmres (SGS)	0.006	56	0.230	92	Gmres (SGS)	2.603	20	385.0	63	
Κανονικό 16-γωνο					Κανονικό 40-γωνο					
	N = 16		N = 128			N = 16		N = 128		
	diff	%	diff	%		diff	%	diff	%	
Bicgstab	-0.102	256	168.300	9	Bicgstab	0.548	47	825.1	5	
Gmres	-0.044	172	104.550	8	Gmres	0.270	54	496.1	5	
Bicgstab (SGS)	0.047	45	19.700	65	Bicgstab (SGS)	4.640	19	950.0	65	
Gmres (SGS)	0.087	38	32.900	64	Gmres (SGS)	3.400	28	485.0	65	

Πίνακας T14: *FFTmul vs Mvmtul*: Η διαφορά στον χρόνο επίλυσης του γραμμικού συστήματος με επαναληπτικές μεθόδους (α) χρησιμοποιώντας απευθείας πολλαπλασιασμό πίνακα επί διάνυμα (*Mvmtul*) και (β) χρησιμοποιώντας ταχείς μετασχηματισμούς Fourier (*FFTmul*) για τον πολλαπλασιασμό πίνακα επί διάνυμα στην περίπτωση που έχουμε διαφορετικές συνοριακές συνθήκες σε κάθε πλευρά. Σε κάθε πολύγωνο η πρώτη στήλη (*diff*) δείχνει την διαφορά του χρόνου επίλυσης με *Mvmtul* μείον τον χρόνο επίλυσης με *FFTmul* σε δευτερόλεπτα ενώ η δεύτερη στήλη τον λόγο του χρόνου επίλυσης με *FFTmul* προς τον χρόνο επίλυσης με *Mvmtul* εκφρασμένο ως ποσοστό επί τοις εκατό.

7.2 Συμπεράσματα

Από τα προηγούμενα φάνηκε ότι η ιδιαίτερη δομή του collocation πίνακα μπορεί, αν χρησιμοποιηθεί κατάλληλα, να μειώσει σημαντικά το κόστος επίλυσης της Dirichlet to Neumann απεικόνισης σε κανονικά πολύγωνα. Ειδικότερα μπορούμε να μειώσουμε το κόστος των Krylov επαναληπτικών μεθόδων στην περίπτωση των ίδιων αλλά και διαφορετικών συνοριακών συνθηκών σε κάθε πλευρά.

Στην περίπτωση των κανονικών πολυγώνων με τις ίδιες συνοριακές συνθήκες σε κάθε πλευρά, προτείνουμε την χρήση μια άμεσης μεθόδου (FFTinv) η οποία, για μεγάλες περιπτώσεις, οδηγεί σε μείωση του κόστους ως προς την άμεση μέθοδο του πακέτου Lapack πάνω από 99.5%. Η άμεση αυτή μέθοδος μπορεί να βελτιωθεί περαιτέρω αν μετατραπεί σε επαναληπτική (επιλύοντας επαναληπτικά κάποια μιγαδικά γραμμικά συστήματα).

Κεφάλαιο 8

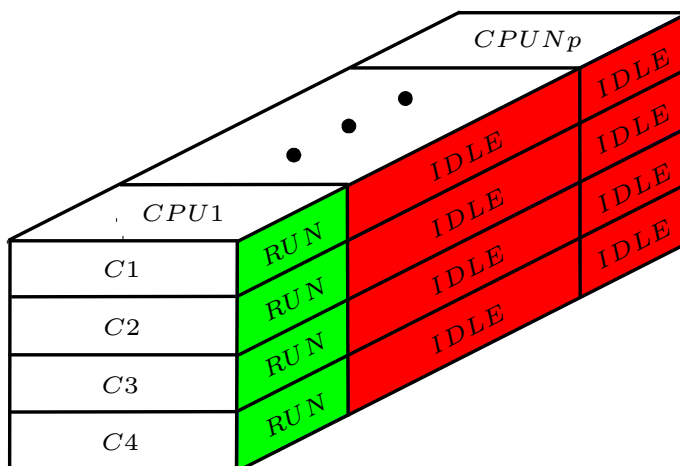
Παράλληλοι Υπολογισμοί / Υπολογισμοί Πλέγματος

Πριν μερικά χρόνια η ύπαρξη παράλληλων υπολογιστικών συστημάτων περιοριζόταν σε μεγάλα ερευνητικά ιδρύματα και συγκεκριμένους κυβερνητικούς οργανισμούς. Στις μέρες μας είναι πραγματικότητα η ύπαρξη παράλληλων υπολογιστών ακόμα και για προσωπική χρήση. Όλοι οι νέοι επεξεργαστές που κυκλοφορούν είναι πλέον πολυπύρρηνοι. Παρόλο την διαφαινόμενη αυτή εξέλιξη, όσο αφορά στον αριθμό των επεξεργαστών ανά υπολογιστικό σύστημα, η εξέλιξη του λογισμικού (software) για την εκμετάλλευση των δυνατοτήτων αυτών των συστημάτων δεν ακολουούθησε την εξέλιξη του υλικού (hardware). Το ζητούμενο στους επιστημονικούς υπολογισμούς δεν είναι η ταυτόχρονη εκτέλεση δυο ή περισσότερων διαφορετικών προγραμμάτων αλλά η επίλυση ενός προβλήματος με την ταυτόχρονη χρήση πολλών επεξεργαστών. Αν και έχουν αναπτυχθεί λογισμικά για τον αυτόματο διαχωρισμό των εργασιών σε πολλούς ταυτόχρονα επεξεργαστές (π.χ. η παράμετρος `-ap` στους MIPSpro Fortran 90 compilers για sgi μηχανές), η απόδοση τους δεν είναι η αναμενόμενη και έτσι είναι απαραίτητη, τουλάχιστον για τα σημερινά δεδομένα, η παρέμβαση του προγραμματιστή για την αποδοτική χρήση των παράλληλων συστημάτων. Κινούμενοι σε αυτή την κατεύθυνση θα υλοποιήσουμε ένα παράλληλο αλγόριθμο για την επίλυση του collocation συστήματος με την άμεση FFTinν μέθοδο.

8.1 Ο Παράλληλος Αλγόριθμος

Σκοπός μας είναι να παράξουμε ένα παράλληλο αλγόριθμο για την επίλυση ενός πραγματικού γραμμικού συστήματος $Ax = b$ όπου ο πίνακας των συντελεστών των αγνώστων είναι block-circulant $A = bcirc(A_1, A_2, \dots, A_n)$ και τα μπλοκ $A_1, A_2, \dots, A_n \in \mathbb{R}^{N \times N}$ είναι γενικοί πίνακες χωρίς κάποια ιδιαίτερη δομή. Θα θεωρήσουμε ότι οι διαθέσιμοι επεξεργαστές (έστω N_p) διαιρούν ακριβώς το πλήθος των μπλοκ του πίνακα A (n) και την διάσταση του κάθε μπλοκ (N). Δηλαδή απαιτούμε να ισχύει ότι

$$N_p \mid n \text{ και } N_p \mid N \quad (8.1)$$



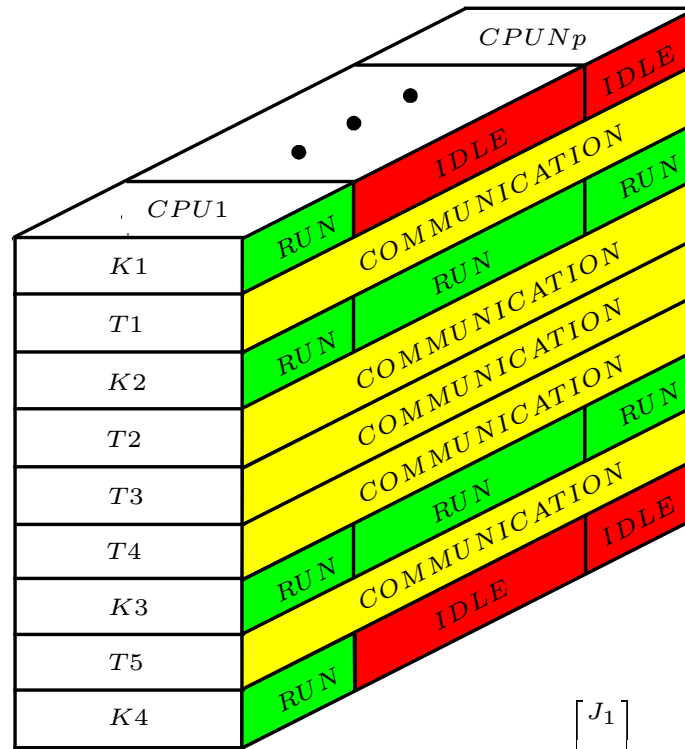
$$C1 : \quad \text{Υπολόγισε} \quad \vec{c} = \frac{1}{n} P^t \left(\bigoplus_{k=1}^N F_n^* \right) P \vec{b}$$

$$C2 : \quad \text{Υπολόγισε} \quad \begin{bmatrix} J_1 \\ \vdots \\ J_n \end{bmatrix} = P^t \left(\bigoplus_{k=1}^N F_n \right) P \begin{bmatrix} A_1 \\ \vdots \\ A_n \end{bmatrix}$$

$$C3 : \quad \text{Λύσε} \quad J_t \vec{y}_t = \vec{c}_t, \quad t = 1, \dots, n$$

$$C4 : \quad \text{Υπολόγισε} \quad \vec{x} = P^t \left(\bigoplus_{k=1}^N F_n \right) P \vec{y}$$

Σχήμα 8.1: Ο σειριακός αλγόριθμος της *FFTinv*.



$$J := \begin{bmatrix} J_1 \\ \vdots \\ J_n \end{bmatrix}$$

$$K1 : \quad \text{Υπολόγισε} \quad \vec{c} = \frac{1}{n} P^t \left(\bigoplus_{k=1}^N F_n^* \right) P \vec{b}$$

$$K2 : \quad \text{Υπολόγισε} \quad \begin{bmatrix} J_1 \\ \vdots \\ J_n \end{bmatrix} = P^t \left(\bigoplus_{k=1}^N F_n \right) P \begin{bmatrix} A_1 \\ \vdots \\ A_n \end{bmatrix}$$

$$K3 : \quad \text{Λύσε} \quad J_t \vec{y}_t = \vec{c}_t, \quad t = 1, \dots, n$$

$$K4 : \quad \text{Υπολόγισε} \quad \vec{x} = P^t \left(\bigoplus_{k=1}^N F_n \right) P \vec{y}$$

$T1 :$ Διαμοίρασε τον J (κατά στήλες)

$T2 :$ Συγκέντρωσε τον J στην $CPU1$

$T3 :$ Διαμοίρασε τον J (κατά γραμμές)

$T4 :$ Διαμοίρασε το $\vec{c}$

$T5 :$ Συγκέντρωσε το $\vec{y}$ στην $CPU1$

Σχήμα 8.2: Ο παράλληλος αλγόριθμος της FFT_{inv} .

Η βασική δομή των δύο αλγορίθμων που κατασκευάσαμε για την επίλυση του collocation συστήματος (σειριακός και παράλληλος) με την FFTinv μέθοδο φαίνεται στα σχήματα (8.1) και (8.2), αντίστοιχα. Για την καλύτερη κατανόηση των δυο αλγορίθμων επισημαίνουμε ότι

- Ο πίνακας συντελεστών του γραμμικού συστήματος είναι block-circulant και έτσι δεν είναι απαραίτητη η αποθήκευση ολόκληρου του πίνακα αλλά μόνο της πρώτης γραμμής (ή στήλης) των μπλοκ.
- Τα μπλοκ του πίνακα A δεν χαίρουν κάποιας ιδιαίτερης δομής και πρέπει να μεταφέρονται εξ' ολοκλήρου στον κάθε επεξεργαστή που τα χρειάζεται κάθε δεδομένη στιγμή.
- Τα μπλοκ $A_1, \dots, A_n$ δεν καταλαμβάνουν ξεχωριστό χώρο στην μνήμη αλλά έχουν φορτωθεί εξ' αρχής στον πίνακα J .
- Οι πίνακες $J_1, J_2, \dots, J_n$ είναι μιγαδικοί. Το ίδιο ισχύει και για τα διανύσματα $\vec{c}$ και $\vec{y}$. Επομένως η χρησιμοποιούμενη αριθμητική είναι αριθμητική μιγαδικών αριθμών (διπλής ακρίβειας).
- Η πράξη του πολλαπλασιασμού των πινάκων P και P^t με κάποιο διάνυσμα δεν εκτελείται. Αντ' αυτής γίνεται μια αναδιάταξη των στοιχείων του διανύσματος, αφού ο πίνακας P είναι πίνακας μετάθεσης [βήματα C1, C4, K1, K4]. Ανάλογα εκτελούμε και τον πολλαπλασιασμό των παραπάνω πινάκων με κάποιο πίνακα A θεωρώντας ότι ο πίνακας αποτελείται από διανύσματα στήλες [βήματα C2, K2].
- Η πράξη του πολλαπλασιασμού $\bigoplus_{k=1}^N F_n$ με ένα διάνυσμα (ή με κάποιο πίνακα θεωρούμενο ως σύνολο διανυσμάτων στήλης) μπορεί να αναλυθεί σε N διακριτούς μετασχηματισμούς Fourier τάξης n (DFT(n)) οι οποίοι μπορούν να υπολογιστούν αποδοτικά με N ταχείς μετασχηματισμούς Fourier τάξης n (FFT(n)) [βήματα C2, C4, K2, K4]. Ανάλογα ισχύουν και στην περίπτωση που αντί για F_n έχουμε F_n^* οπότε και οι μετασχηματισμοί Fourier είναι οι αντίστροφοι δηλαδή IDFT(n) και IFFT(n) [βήματα C1, K1].
- Για τον υπολογισμό των FFT χρησιμοποιούμε την έκδοση διπλής ακρίβειας του πακέτου FFTPACK(n) [FFTPACK].
- Η επίλυση των μιγαδικών γραμμικών συστημάτων γίνεται άμεσα με την ρουτίνα zgesv του πακέτου LAPACK [LAPACK] [βήματα C3, K3].
- Στον παράλληλο αλγόριθμο λέγοντας διαμοιρασμό μιας μεταβλητής (π.χ. ενός διανύσματος) εννοούμε ότι ο κάθε επεξεργαστής λαμβάνει ένα μέρος του διανύσματος και ότι όλοι οι επεξεργαστές λαμβάνουν διαφορετικά αλλά ίσου μεγέθους κομμάτια [βήματα T1, T3]. Αυτό μπορεί και συμβαίνει λόγω των περιορισμών (8.1).

- Στον παράλληλο αλγόριθμο μεταφέρουμε δυο φορές τον πίνακα J [βήματα $T1, T3$]. Αυτό συμβαίνει διότι στο βήμα $K2$ ο κάθε επεξεργαστής κάνει υπολογισμούς πάνω σε ολόκληρες στήλες του J (π.χ. χρειάζεται ολόκληρη την 1η στήλη του J) ενώ στο βήμα $K3$ τα συστήματα που καλείται να επιλύσει ο κάθε επεξεργαστής βρίσκονται αποθηκευμένα κατά γραμμές στον πίνακα J .
- Για την επίλυση του Dirichlet to Neumann Map ισχύει, εν γένει, ότι $n < N$.

8.2 Αριθμητικά Αποτελέσματα

Υλοποιήσαμε τους δυο αλγορίθμους (σειριακό και παράλληλο) χρησιμοποιώντας το πρότυπο MPI¹ για τρία παράλληλα υπολογιστικά συστήματα SGI - CM - CG. Τα χαρακτηριστικά αυτών των συστημάτων είναι

[SGI] Παράλληλος υπολογιστής SGI Origin 350 αποτελούμενος από 8 επεξεργαστές R16000@600MHz με 4MB L2 cache έκαστος. Η συνολική μνήμη του συστήματος είναι 4GB και το λειτουργικό σύστημα είναι IRIX version 6.5 με MipsPro compilers version 7.4.

[CM/CG] Συστάδα (cluster) τεσσάρων κόμβων SUN V240z διασυνδεδεμένων μέσω ethernet 100Mbps/1Gbps αντίστοιχα. Ο κάθε κόμβος αποτελείται από ένα διπλοπύρρηνο 1.5 GHz UltraSPARC IIIi επεξεργαστή με 1+1 L2 cache. Η συνολική μνήμη κάθε κόμβου είναι 2GB και το λειτουργικό σύστημα είναι Solaris 10.

Η γλώσσα προγραμματισμού που χρησιμοποιήσαμε είναι η Fortran και όλοι οι υπολογισμοί έγιναν με αριθμητική διπλής ακρίβειας. Για την μελέτη της απόδοσης του παράλληλου αλγόριθμου εκτελέσαμε πειράματα για μια πληθώρα κανονικών πολυγώνων $n = 8, 16, 24, 32, 40, 48, 56, 64, 72, 80$ με τις ίδιες συνοριακές συνθήκες σε κάθε πλευρά (Dirichlet) και για πλήθος συναρτήσεων βάσης

¹ Message Parsing Programming: Ένα πρότυπο το οποίο μας παρέχει την δυνατότητα ανταλλαγής μηνυμάτων (δεδομένων/εντολών) μεταξύ διασυνδεδεμένων διεργασιών (processes) και γενικότερα την δυνατότητα επικοινωνίας των διεργασιών. Η δυνατότητα αυτή μπορεί να αξιοποιηθεί για την υλοποίηση παράλληλων αλγορίθμων σε πολυεπεξεργαστικές μηχανές και συστάδες μηχανών. Όλες οι διεργασίες (επεξεργαστές στην δική μας περίπτωση) τρέχουν ένα αντίγραφο του ίδιου προγράμματος αλλά στον κάθε επεξεργαστή έχει ανατεθεί εξ' αρχής ένας αύξων αριθμός (rank) βάση του μπορούμε ξεχωρίζουμε τους επεξεργαστές και να δίνουμε διαφορετικές εντολές στον καθένα. Π.χ.

```
If (rank==0) then
  x = (-b+sqrt(D))/2
elseif (rank==1) then
  x = (-b-sqrt(D))/2
endif
```

Η ευθύνη της παραλληλοποίησης ανήκει καθαρά και μόνο στον προγραμματιστή.

$N = 8, 16, 32, 64, 128, 256, 512, 1024$. Για $N = 1024$ έχουμε αποτελέσματα μόνο για μικρά πολύγωνα ($n = 8, 16, 24$) λόγω περιορισμών μνήμης. Τα κανονικά πολύγωνα καθώς και η ακριβής λύση της εξίσωσης του Laplace παραμένουν ίδια όπως και στην παράγραφο (4.1).

Οι πίνακες/σχήματα που παραθέτουμε έχουν ως εξής:

Ο συνολικός χρόνος επίλυσης του γραμμικού συστήματος (πίνακες T1-T4) Ο χρόνος που απαιτείται για την πλήρη επίλυση του γραμμικού συστήματος. Όσο μικρότερος (για περισσότερους από ένα επεξεργαστές) τόσο καλύτερος ο παράλληλος αλγόριθμος.

Ο χρόνος εκτέλεσης των υπολογισμών (πίνακες T5-T8) Ο χρόνος αυτός περιλαμβάνει τον χρόνο των αριθμητικών πράξεων, τον χρόνο πρόσβασης στην μνήμη, τον χρόνο για το indexing των μεταβλητών και γενικότερα το χρόνο εκτέλεσης οποιασδήποτε διαδικασίας λαμβάνει χώρα στην ακολουθία των βημάτων για την επίλυση του συστήματος στο σύνολο των επεξεργαστών. Ο χρόνος αυτός μετριέται πάντα από τον πρώτο επεξεργαστή. Όσο μικρότερος (για περισσότερους από ένα επεξεργαστές) τόσο αποδοτικότερος ο παράλληλος αλγόριθμος.

Ο χρόνος επικοινωνίας (πίνακες T9-T12) Είναι ο χρόνος που απαιτείται για την ανταλλαγή δεδομένων μεταξύ των επεξεργαστών. Ο χρόνος αυτός μετριέται πάντα από τον πρώτο επεξεργαστή. Το άθροισμα των χρόνων εκτέλεσης των υπολογισμών και επικοινωνίας πρέπει να ισούται με τον συνολικό χρόνο επίλυσης. Όσο μικρότερος τόσο αποδοτικότερος ο παράλληλος αλγόριθμος.

Το speedup(επιτάχυνση) για την επίλυση του γραμμικού συστήματος (πίνακες T13-T16 και σχήματα 8.3-8.6) Το speedup για p επεξεργαστές (έστω $S(p)$) ορίζεται ως ο λόγος του χρόνου που απαιτείται για την επίλυση ενός προβλήματος σε ένα επεξεργαστή $T(1)$ προς το χρόνο που απαιτείται για την επίλυση του ίδιου προβλήματος σε p επεξεργαστές (έστω $T(p)$). Δηλαδή

$$S(N) = \frac{T(1)}{T(p)}.$$

Όσο μεγαλύτερο τόσο αποδοτικότερος ο παράλληλος αλγόριθμος.

Το speedup εκτέλεσης των υπολογισμών (πίνακες T17-T20 και σχήματα 8.3-8.6) Το speedup για την επίλυση του γραμμικού συστήματος χωρίς να λαμβάνονται υπόψιν οι χρόνοι επικοινωνίας. Όσο μεγαλύτερο τόσο αποδοτικότερος ο παράλληλος αλγόριθμος.

Λόγος χρόνου επικοινωνίας προς χρόνο υπολογισμών (πίνακες T21-T24) Το ποσοστό επί τοις εκατό του χρόνου επικοινωνίας προς το χρόνο εκτέλεσης των υπολογισμών. Όσο μικρότερος τόσο καλύτερος ο παράλληλος αλγόριθμος.

SGI - Ο Συνολικός Χρόνος Επίλυσης του Γραμμικού Συστήματος ($Ax=b$)										
$Np = 1$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.002	0.002	0.003	0.005	0.005	0.006	0.010	0.008	0.009	0.010
16	0.004	0.007	0.011	0.014	0.017	0.021	0.034	0.027	0.033	0.037
32	0.015	0.028	0.048	0.061	0.081	0.097	0.154	0.124	0.166	0.181
64	0.085	0.163	0.262	0.352	0.442	0.534	0.781	0.715	0.936	1.024
128	0.598	1.195	1.957	2.512	3.251	3.865	5.158	5.040	6.010	6.609
256	4.453	8.894	13.771	18.120	23.023	27.238	34.633	36.336	41.651	45.968
512	34.121	67.933	103.834	137.171	173.087	204.631	251.809	277.413	311.511	344.787
1024	260.977	523.725	840.509							
$Np = 2$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.002	0.002	0.003	0.003	0.004	0.005	0.007	0.006	0.007	0.007
16	0.003	0.005	0.008	0.010	0.012	0.014	0.022	0.019	0.022	0.024
32	0.010	0.018	0.031	0.039	0.050	0.060	0.091	0.078	0.103	0.111
64	0.050	0.095	0.154	0.207	0.263	0.314	0.450	0.429	0.552	0.605
128	0.330	0.674	1.110	1.385	1.836	2.214	2.918	2.854	3.389	3.735
256	2.441	4.775	7.687	9.664	12.710	15.074	19.065	19.360	22.338	24.640
512	17.589	35.113	55.305	70.704	94.934	108.974	133.524	148.872	161.570	178.722
1024	131.916	264.118	434.115							
$Np = 4$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.002	0.002	0.003	0.003	0.003	0.004	0.005	0.004	0.005	0.005
16	0.003	0.004	0.006	0.007	0.008	0.010	0.014	0.012	0.014	0.016
32	0.007	0.012	0.019	0.025	0.031	0.037	0.055	0.048	0.061	0.067
64	0.029	0.056	0.089	0.119	0.151	0.184	0.258	0.248	0.314	0.344
128	0.180	0.364	0.595	0.761	1.011	1.212	1.576	1.539	1.855	2.051
256	1.259	2.520	4.011	5.128	6.725	8.068	10.028	10.278	11.852	13.144
512	9.091	18.094	28.383	36.569	47.396	56.690	68.646	77.051	83.587	93.907
1024	67.189	134.927	250.684							
$Np = 8$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.002	0.002	0.002	0.003	0.003	0.004	0.014	0.004	0.005	0.006
16	0.003	0.004	0.005	0.006	0.009	0.008	0.012	0.010	0.012	0.013
32	0.006	0.010	0.015	0.018	0.023	0.027	0.038	0.034	0.048	0.045
64	0.021	0.039	0.058	0.078	0.097	0.125	0.163	0.160	0.195	0.219
128	0.118	0.219	0.349	0.457	0.592	0.708	0.930	0.942	1.086	1.203
256	0.744	1.392	2.216	2.935	3.721	4.496	5.583	5.760	6.624	7.371
512	5.123	10.185	15.299	19.529	25.334	30.483	37.195	41.147	44.951	49.902
1024	35.027	70.136	134.691							

Πίνακας T1: SGI - Ο συνολικός χρόνος επίλυσης του γραμμικού συστήματος, για 1,2,4 και 8 επεξεργαστές.

CM/CG - Ο Συνολικός Χρόνος Επίλυσης του Γραμμικού Συστήματος ($Ax=b$)										
$Np = 1$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.001	0.001	0.002	0.002	0.003	0.003	0.005	0.004	0.005	0.005
16	0.002	0.003	0.006	0.007	0.010	0.012	0.018	0.019	0.020	0.022
32	0.008	0.015	0.029	0.041	0.049	0.058	0.087	0.087	0.101	0.121
64	0.043	0.104	0.159	0.226	0.294	0.355	0.576	0.474	0.589	0.624
128	0.360	0.724	1.085	1.517	1.891	2.459	2.962	3.173	3.666	4.067
256	2.487	5.126	8.115	10.611	14.106	16.250	20.188	21.626	24.917	27.677
512	20.822	38.277	63.264	79.696	107.207	126.347	151.100	162.818	203.415	220.884
1024	174.190	354.891	498.594							
$Np = 2$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.001	0.001	0.002	0.002	0.002	0.002	0.004	0.003	0.004	0.004
16	0.002	0.003	0.004	0.005	0.007	0.008	0.012	0.012	0.013	0.014
32	0.005	0.010	0.019	0.024	0.034	0.038	0.051	0.052	0.060	0.070
64	0.026	0.060	0.090	0.127	0.166	0.203	0.260	0.268	0.338	0.356
128	0.199	0.387	0.608	0.815	1.017	1.331	1.583	1.727	1.949	2.188
256	1.352	2.699	4.144	5.583	7.192	8.242	10.250	11.340	12.730	14.060
512	10.040	19.780	30.700	41.240	52.290	61.610	73.650	84.280	100.600	107.400
1024	87.400	183.700	300.800							

Πίνακας T2: CM/CG - Ο συνολικός χρόνος επίλυσης του γραμμικού συστήματος, για 1 και 2 επεξεργαστές.

CM - Ο Συνολικός Χρόνος Επίλυσης του Γραμμικού Συστήματος ($Ax=b$)										
$Np = 4$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.003	0.004	0.005	0.006	0.008	0.009	0.011	0.012	0.013	0.015
16	0.006	0.011	0.016	0.021	0.027	0.030	0.037	0.040	0.045	0.050
32	0.021	0.040	0.058	0.076	0.095	0.115	0.138	0.155	0.179	0.196
64	0.078	0.158	0.239	0.327	0.409	0.495	0.593	0.664	0.762	0.837
128	0.361	0.797	1.088	1.485	1.829	2.244	2.644	2.965	3.376	3.775
256	1.725	3.496	5.288	7.062	8.946	10.790	12.650	14.220	16.060	17.810
512	9.634	18.940	29.040	38.940	48.800	57.820	68.600	78.310	91.420	99.280
1024	60.890	123.200	203.700							
$Np = 8$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.004	0.005	0.006	0.007	0.009	0.010	0.012	0.014	0.017	0.019
16	0.007	0.014	0.021	0.027	0.035	0.041	0.049	0.055	0.062	0.068
32	0.027	0.052	0.082	0.106	0.135	0.161	0.189	0.215	0.241	0.276
64	0.107	0.215	0.321	0.429	0.541	0.647	0.766	0.865	0.983	1.087
128	0.454	0.891	1.341	1.805	2.253	2.729	3.195	3.739	4.110	4.546
256	1.941	3.899	5.849	7.835	9.826	11.710	13.800	15.630	17.710	19.570
512	9.135	18.030	27.360	36.580	46.060	54.620	64.140	73.360	83.950	92.430
1024	46.600	95.710	145.400							

Πίνακας T3: CM - Ο συνολικός χρόνος επίλυσης του γραμμικού συστήματος, για 4 και 8 επεξεργαστές.

CG - Ο Συνολικός Χρόνος Επίλυσης του Γραμμικού Συστήματος ($Ax=b$)										
$Np = 4$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.002	0.002	0.002	0.003	0.004	0.004	0.005	0.004	0.005	0.005
16	0.003	0.004	0.006	0.007	0.008	0.010	0.012	0.013	0.015	0.016
32	0.007	0.012	0.020	0.024	0.030	0.035	0.042	0.045	0.052	0.062
64	0.027	0.050	0.070	0.096	0.122	0.145	0.184	0.194	0.234	0.253
128	0.135	0.254	0.391	0.531	0.653	0.833	1.001	1.086	1.242	1.376
256	0.790	1.610	2.478	3.315	4.278	4.894	6.172	6.711	7.544	8.414
512	5.887	11.430	17.650	23.630	30.040	35.300	42.650	48.140	57.850	61.620
1024	45.790	95.040	148.200							
$Np = 8$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.003	0.003	0.003	0.003	0.003	0.004	0.004	0.005	0.005	0.005
16	0.004	0.005	0.006	0.007	0.008	0.009	0.012	0.012	0.013	0.015
32	0.009	0.012	0.018	0.022	0.033	0.035	0.044	0.045	0.051	0.054
64	0.025	0.050	0.070	0.088	0.108	0.125	0.151	0.166	0.194	0.205
128	0.114	0.204	0.290	0.393	0.488	0.604	0.716	0.791	0.894	0.990
256	0.536	1.059	1.603	2.169	2.712	3.183	3.879	4.268	4.854	5.390
512	3.464	6.689	10.240	13.840	17.250	20.580	24.500	27.780	32.920	35.640
1024	23.900	51.340	78.910							

Πίνακας T4: CG - Ο συνολικός χρόνος επίλυσης του γραμμικού συστήματος, για 4 και 8 επεξεργαστές.

SGI - Ο Χρόνος Εκτέλεσης των Υπολογισμών										
$Np = 2$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.001	0.002	0.002	0.003	0.003	0.004	0.006	0.005	0.006	0.006
16	0.003	0.004	0.007	0.008	0.010	0.012	0.019	0.015	0.018	0.020
32	0.008	0.015	0.026	0.032	0.043	0.051	0.082	0.066	0.088	0.095
64	0.043	0.083	0.134	0.181	0.230	0.274	0.404	0.367	0.480	0.524
128	0.301	0.607	1.007	1.266	1.683	1.991	2.670	2.568	3.068	3.377
256	2.248	4.497	7.248	9.098	12.006	14.181	18.023	18.223	21.158	23.413
512	17.005	33.728	53.486	68.342	92.010	104.774	129.372	144.794	156.235	173.045
1024	129.678	260.211	426.186							
$Np = 4$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.001	0.001	0.002	0.002	0.002	0.003	0.004	0.003	0.004	0.004
16	0.002	0.002	0.004	0.005	0.006	0.007	0.011	0.009	0.011	0.012
32	0.005	0.008	0.014	0.018	0.023	0.027	0.044	0.036	0.047	0.051
64	0.022	0.042	0.069	0.093	0.117	0.139	0.207	0.190	0.245	0.268
128	0.152	0.303	0.501	0.633	0.845	0.988	1.343	1.273	1.546	1.704
256	1.105	2.248	3.589	4.562	6.002	6.974	9.027	9.147	10.582	11.719
512	8.453	16.956	26.683	34.269	44.455	51.386	64.672	72.527	78.424	87.260
1024	64.897	130.385	241.989							
$Np = 8$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.001	0.001	0.002	0.002	0.002	0.002	0.013	0.003	0.003	0.004
16	0.001	0.002	0.003	0.003	0.004	0.005	0.008	0.006	0.007	0.008
32	0.003	0.005	0.008	0.010	0.013	0.015	0.025	0.020	0.026	0.028
64	0.012	0.023	0.037	0.049	0.062	0.078	0.112	0.100	0.128	0.139
128	0.076	0.154	0.257	0.330	0.431	0.486	0.686	0.648	0.782	0.861
256	0.563	1.109	1.809	2.285	3.027	3.413	4.538	4.619	5.326	5.877
512	4.295	8.546	13.416	17.192	22.398	24.848	33.037	36.293	39.389	44.003
1024	32.471	65.148	124.676							

Πίνακας T5: SGI - Ο χρόνος εκτέλεσης των υπολογισμών, για 1,2,4 και 8 επεξεργαστές.

CM/CG - Ο Χρόνος Εκτέλεσης των Υπολογισμών										
$Np = 2$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.001	0.001	0.001	0.001	0.002	0.002	0.003	0.003	0.003	0.003
16	0.001	0.002	0.003	0.004	0.006	0.007	0.010	0.010	0.011	0.012
32	0.004	0.008	0.016	0.021	0.030	0.033	0.045	0.046	0.053	0.062
64	0.022	0.053	0.079	0.113	0.149	0.182	0.235	0.239	0.302	0.320
128	0.186	0.357	0.560	0.761	0.937	1.248	1.481	1.609	1.820	2.039
256	1.291	2.579	3.966	5.331	6.882	7.894	9.851	10.830	12.200	13.480
512	9.806	19.290	29.980	40.020	51.090	59.740	71.610	82.130	98.550	105.000
1024	85.030	181.800	282.000							

Πίνακας T6: CM/CG - Ο χρόνος εκτέλεσης των υπολογισμών, για 2 επεξεργαστές.

CM - Ο Χρόνος Εκτέλεσης των Υπολογισμών										
<i>Np</i> = 4	n = 8	16	24	32	40	48	56	64	72	80
N = 8	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.002	0.002	0.002
16	0.001	0.001	0.002	0.003	0.004	0.004	0.006	0.006	0.007	0.008
32	0.002	0.005	0.009	0.012	0.015	0.018	0.025	0.024	0.029	0.032
64	0.013	0.028	0.041	0.058	0.077	0.091	0.121	0.121	0.155	0.162
128	0.090	0.180	0.278	0.387	0.473	0.613	0.758	0.799	0.939	1.022
256	0.632	1.321	1.986	2.664	3.437	3.860	4.924	5.362	6.059	6.753
512	5.208	9.527	14.920	20.040	25.350	29.120	35.610	40.780	48.700	52.450
1024	42.000	88.600	132.000							
<i>Np</i> = 8	n = 8	16	24	32	40	48	56	64	72	80
N = 8	0.000	0.001	0.001	0.001	0.001	0.001	0.002	0.001	0.002	0.002
16	0.001	0.001	0.001	0.002	0.002	0.003	0.004	0.004	0.005	0.005
32	0.002	0.003	0.005	0.007	0.009	0.010	0.014	0.014	0.016	0.018
64	0.008	0.016	0.023	0.032	0.042	0.047	0.065	0.064	0.080	0.083
128	0.049	0.093	0.142	0.202	0.250	0.309	0.390	0.417	0.479	0.526
256	0.318	0.671	1.040	1.391	1.810	1.912	2.488	2.737	3.063	3.400
512	2.628	5.196	7.484	10.050	12.700	14.060	18.180	20.570	24.410	26.210
1024	19.650	44.060	68.430							

Πίνακας T7: CM - Ο χρόνος εκτέλεσης των υπολογισμών, για 4 και 8 επεξεργαστές.

CG - Ο Χρόνος Εκτέλεσης των Υπολογισμών										
<i>Np</i> = 4	n = 8	16	24	32	40	48	56	64	72	80
N = 8	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.002	0.002	0.002
16	0.001	0.001	0.002	0.003	0.004	0.004	0.007	0.006	0.006	0.007
32	0.003	0.005	0.008	0.012	0.015	0.018	0.024	0.025	0.029	0.033
64	0.012	0.028	0.041	0.058	0.076	0.090	0.121	0.123	0.155	0.165
128	0.091	0.179	0.283	0.389	0.474	0.613	0.761	0.799	0.938	1.021
256	0.620	1.323	2.012	2.669	3.446	3.865	5.057	5.406	6.087	6.748
512	5.201	9.528	14.980	20.060	25.440	29.300	35.930	40.820	49.030	52.480
1024	41.810	90.640	134.100							
<i>Np</i> = 8	n = 8	16	24	32	40	48	56	64	72	80
N = 8	0.000	0.001	0.001	0.001	0.001	0.001	0.002	0.001	0.002	0.002
16	0.001	0.001	0.001	0.002	0.002	0.003	0.004	0.004	0.004	0.005
32	0.002	0.003	0.005	0.008	0.009	0.010	0.014	0.014	0.016	0.018
64	0.007	0.015	0.022	0.032	0.041	0.046	0.064	0.063	0.081	0.085
128	0.049	0.093	0.143	0.204	0.247	0.309	0.386	0.418	0.481	0.528
256	0.316	0.676	1.039	1.440	1.806	1.908	2.504	2.715	3.089	3.396
512	2.624	5.196	7.481	10.110	12.780	14.290	17.960	20.620	24.670	26.470
1024	19.750	45.490	70.250							

Πίνακας T8: CG - Ο χρόνος εκτέλεσης των υπολογισμών, για 4 και 8 επεξεργαστές.

SGI - Ο Χρόνος Επικοινωνίας										
$Np = 2$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.000	0.000	0.000	0.000	0.001	0.001	0.001	0.001	0.001	0.001
16	0.001	0.001	0.001	0.002	0.002	0.002	0.003	0.003	0.003	0.003
32	0.002	0.003	0.005	0.006	0.007	0.009	0.009	0.011	0.015	0.016
64	0.007	0.012	0.020	0.026	0.033	0.040	0.046	0.062	0.072	0.081
128	0.029	0.066	0.103	0.119	0.153	0.224	0.248	0.286	0.322	0.359
256	0.193	0.278	0.439	0.567	0.704	0.893	1.042	1.137	1.179	1.227
512	0.584	1.385	1.818	2.362	2.924	4.200	4.153	4.078	5.335	5.677
1024	2.238	3.908	7.929							
$Np = 4$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.000	0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
16	0.001	0.001	0.002	0.002	0.002	0.003	0.003	0.003	0.004	0.004
32	0.002	0.004	0.005	0.007	0.008	0.010	0.011	0.012	0.014	0.016
64	0.007	0.014	0.019	0.026	0.034	0.044	0.052	0.058	0.069	0.076
128	0.028	0.061	0.094	0.128	0.166	0.224	0.234	0.267	0.310	0.347
256	0.154	0.272	0.422	0.566	0.723	1.094	1.001	1.131	1.270	1.424
512	0.638	1.138	1.700	2.300	2.941	5.304	3.973	4.524	5.163	6.647
1024	2.293	4.541	8.696							
$Np = 8$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.002
16	0.002	0.002	0.002	0.002	0.004	0.004	0.004	0.004	0.005	0.005
32	0.003	0.005	0.007	0.008	0.009	0.011	0.013	0.013	0.022	0.017
64	0.009	0.015	0.021	0.029	0.035	0.047	0.051	0.060	0.067	0.080
128	0.041	0.066	0.093	0.127	0.160	0.222	0.244	0.294	0.304	0.342
256	0.181	0.283	0.407	0.650	0.694	1.083	1.045	1.140	1.298	1.494
512	0.828	1.639	1.883	2.336	2.937	5.635	4.157	4.855	5.562	5.899
1024	2.556	4.989	10.015							

Πίνακας T9: SGI - Ο χρόνος επικοινωνίας, για 2,4 και 8 επεξεργαστές.

CM/CG - Ο Χρόνος Επικοινωνίας										
$Np = 2$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.001
16	0.000	0.000	0.001	0.001	0.001	0.001	0.001	0.002	0.002	0.002
32	0.001	0.002	0.003	0.003	0.004	0.005	0.006	0.007	0.007	0.008
64	0.004	0.006	0.011	0.014	0.017	0.021	0.026	0.029	0.036	0.036
128	0.013	0.030	0.049	0.053	0.080	0.083	0.102	0.118	0.129	0.150
256	0.061	0.120	0.178	0.252	0.310	0.348	0.399	0.516	0.533	0.574
512	0.234	0.488	0.721	1.221	1.197	1.877	2.042	2.150	2.018	2.389
1024	2.379	1.915	18.810							

Πίνακας T10: CM/CG - Ο χρόνος επικοινωνίας, για 2 επεξεργαστές.

CM - Ο Χρόνος Επικοινωνίας										
$Np = 4$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.002	0.003	0.004	0.005	0.006	0.008	0.009	0.010	0.011	0.013
16	0.005	0.009	0.014	0.018	0.023	0.026	0.031	0.034	0.038	0.042
32	0.018	0.035	0.049	0.064	0.080	0.097	0.114	0.131	0.151	0.164
64	0.065	0.129	0.198	0.269	0.332	0.404	0.472	0.542	0.607	0.675
128	0.271	0.618	0.810	1.098	1.356	1.631	1.886	2.166	2.437	2.753
256	1.093	2.175	3.303	4.398	5.508	6.925	7.729	8.859	9.996	11.060
512	4.426	9.408	14.130	18.890	23.450	28.700	32.990	37.530	42.720	46.820
1024	18.890	34.560	71.730							
$Np = 8$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.003	0.004	0.005	0.006	0.008	0.009	0.010	0.013	0.015	0.017
16	0.007	0.013	0.019	0.025	0.032	0.038	0.045	0.050	0.057	0.063
32	0.025	0.050	0.076	0.099	0.126	0.151	0.174	0.201	0.225	0.257
64	0.099	0.199	0.298	0.397	0.499	0.600	0.701	0.801	0.903	1.003
128	0.405	0.798	1.199	1.603	2.003	2.420	2.804	3.322	3.632	4.020
256	1.623	3.229	4.809	6.444	8.016	9.793	11.310	12.900	14.650	16.170
512	6.507	12.830	19.870	26.530	33.360	40.560	45.960	52.790	59.540	66.220
1024	26.950	51.640	77.010							

Πίνακας T11: CM - Ο χρόνος επικοινωνίας, για 4 και 8 επεξεργαστές.

CG - Ο Χρόνος Επικοινωνίας										
$Np = 4$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.001	0.002	0.001	0.002	0.002	0.003	0.003	0.002	0.003	0.003
16	0.002	0.002	0.003	0.004	0.004	0.006	0.005	0.007	0.009	0.009
32	0.004	0.007	0.011	0.012	0.015	0.018	0.018	0.020	0.023	0.028
64	0.015	0.022	0.029	0.038	0.045	0.054	0.063	0.071	0.079	0.088
128	0.044	0.074	0.108	0.142	0.178	0.220	0.240	0.287	0.305	0.355
256	0.169	0.286	0.466	0.646	0.832	1.028	1.115	1.304	1.457	1.666
512	0.686	1.901	2.666	3.566	4.599	5.998	6.723	7.315	8.819	9.140
1024	3.982	4.404	14.070							
$Np = 8$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.002	0.002	0.002	0.002	0.003	0.003	0.003	0.003	0.004	0.003
16	0.003	0.004	0.004	0.005	0.005	0.007	0.008	0.008	0.009	0.010
32	0.006	0.009	0.013	0.014	0.024	0.025	0.031	0.031	0.035	0.036
64	0.018	0.035	0.048	0.056	0.067	0.078	0.087	0.103	0.113	0.120
128	0.065	0.111	0.147	0.189	0.241	0.294	0.329	0.373	0.413	0.462
256	0.220	0.383	0.564	0.729	0.906	1.275	1.375	1.553	1.765	1.993
512	0.840	1.492	2.760	3.731	4.469	6.295	6.544	7.156	8.248	9.161
1024	4.144	5.850	8.658							

Πίνακας T12: CG - Ο χρόνος επικοινωνίας, για 4 και 8 επεξεργαστές.

SGI - Το Speedup Επίλυσης του Γραμμικού Συστήματος ($Ax=b$)										
$Np = 2$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.90	1.00	1.14	1.41	1.22	1.21	1.44	1.35	1.35	1.35
16	1.21	1.28	1.40	1.38	1.42	1.51	1.56	1.47	1.50	1.59
32	1.48	1.53	1.57	1.57	1.61	1.63	1.70	1.60	1.61	1.62
64	1.70	1.71	1.70	1.70	1.68	1.70	1.73	1.67	1.69	1.69
128	1.81	1.77	1.76	1.81	1.77	1.75	1.77	1.77	1.77	1.77
256	1.82	1.86	1.79	1.87	1.81	1.81	1.82	1.88	1.86	1.87
512	1.94	1.93	1.88	1.94	1.82	1.88	1.89	1.86	1.93	1.93
1024	1.98	1.98	1.94							
$Np = 4$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	1.02	1.33	1.28	1.75	1.64	1.65	1.93	1.89	1.68	1.82
16	1.55	1.82	1.87	2.07	2.09	2.09	2.37	2.18	2.29	2.38
32	2.02	2.36	2.53	2.46	2.59	2.63	2.79	2.58	2.73	2.71
64	2.88	2.90	2.95	2.95	2.92	2.91	3.02	2.88	2.98	2.97
128	3.32	3.28	3.29	3.30	3.21	3.19	3.27	3.27	3.24	3.22
256	3.54	3.53	3.43	3.53	3.42	3.38	3.45	3.54	3.51	3.50
512	3.75	3.75	3.66	3.75	3.65	3.61	3.67	3.60	3.73	3.67
1024	3.88	3.88	3.35							
$Np = 8$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.73	1.07	1.37	1.72	1.63	1.67	0.69	2.06	1.93	1.80
16	1.34	1.82	2.18	2.40	1.98	2.49	2.94	2.64	2.72	2.99
32	2.31	2.81	3.18	3.29	3.55	3.67	4.02	3.68	3.48	4.03
64	3.98	4.22	4.49	4.49	4.55	4.26	4.79	4.48	4.79	4.67
128	5.09	5.45	5.61	5.50	5.49	5.46	5.55	5.35	5.53	5.49
256	5.98	6.39	6.21	6.17	6.19	6.06	6.20	6.31	6.29	6.24
512	6.66	6.67	6.79	7.02	6.83	6.71	6.77	6.74	6.93	6.91
1024	7.45	7.47	6.24							

Πίνακας T13: SGI - Το speedup για την επίλυση του γραμμικού συστήματος, για 2,4 και 8 επεξεργαστές.

CM/CG - Το Speedup Επίλυσης του Γραμμικού Συστήματος ($Ax=b$)										
$Np = 2$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.78	0.99	1.10	1.21	1.28	1.30	1.37	1.29	1.32	1.23
16	1.02	1.25	1.37	1.43	1.44	1.48	1.57	1.63	1.52	1.51
32	1.47	1.53	1.55	1.67	1.46	1.52	1.71	1.66	1.68	1.72
64	1.65	1.74	1.76	1.77	1.77	1.74	2.21	1.77	1.74	1.75
128	1.81	1.87	1.78	1.86	1.86	1.85	1.87	1.84	1.88	1.86
256	1.84	1.90	1.96	1.90	1.96	1.97	1.97	1.91	1.96	1.97
512	2.07	1.94	2.06	1.93	2.05	2.05	2.05	1.93	2.02	2.06
1024	1.99	1.93	1.66							

Πίνακας T14: CM/CG - Το speedup για την επίλυση του γραμμικού συστήματος, για 2 επεξεργαστές.

CM - Το Speedup Επίλυσης του Γραμμικού Συστήματος ($Ax=b$)										
$Np = 4$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.24	0.29	0.34	0.35	0.36	0.35	0.44	0.34	0.37	0.36
16	0.31	0.31	0.34	0.35	0.36	0.39	0.49	0.47	0.44	0.44
32	0.37	0.38	0.50	0.53	0.52	0.51	0.63	0.56	0.56	0.62
64	0.55	0.66	0.67	0.69	0.72	0.72	0.97	0.71	0.77	0.75
128	1.00	0.91	1.00	1.02	1.03	1.10	1.12	1.07	1.09	1.08
256	1.44	1.47	1.53	1.50	1.58	1.51	1.60	1.52	1.55	1.55
512	2.16	2.02	2.18	2.05	2.20	2.19	2.20	2.08	2.23	2.22
1024	2.86	2.88	2.45							
$Np = 8$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.19	0.23	0.27	0.30	0.32	0.31	0.41	0.29	0.30	0.29
16	0.24	0.24	0.27	0.28	0.28	0.29	0.37	0.34	0.32	0.32
32	0.28	0.28	0.36	0.38	0.37	0.36	0.46	0.40	0.42	0.44
64	0.40	0.48	0.50	0.53	0.54	0.55	0.75	0.55	0.60	0.57
128	0.79	0.81	0.81	0.84	0.84	0.90	0.93	0.85	0.89	0.89
256	1.28	1.31	1.39	1.35	1.44	1.39	1.46	1.38	1.41	1.41
512	2.28	2.12	2.31	2.18	2.33	2.31	2.36	2.22	2.42	2.39
1024	3.74	3.71	3.43							

Πίνακας T15: CM: Το speedup για την επίλυση του γραμμικού συστήματος, για 4 και 8 επεξεργαστές.

CG- To Speedup Επίλυσης του Γραμμικού Συστήματος ($Ax=b$)										
$Np = 4$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.38	0.46	0.69	0.73	0.78	0.76	1.01	0.93	1.00	1.02
16	0.54	0.84	0.99	1.06	1.17	1.21	1.50	1.44	1.28	1.33
32	1.06	1.24	1.47	1.70	1.66	1.64	2.06	1.95	1.95	1.97
64	1.58	2.08	2.27	2.36	2.41	2.45	3.13	2.44	2.52	2.47
128	2.67	2.86	2.78	2.86	2.90	2.95	2.96	2.92	2.95	2.96
256	3.15	3.18	3.27	3.20	3.30	3.32	3.27	3.22	3.30	3.29
512	3.54	3.35	3.58	3.37	3.57	3.58	3.54	3.38	3.52	3.58
1024	3.80	3.73	3.36							
$Np = 8$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	0.26	0.36	0.53	0.64	0.79	0.85	1.13	0.93	0.95	1.03
16	0.44	0.68	0.95	1.08	1.24	1.24	1.49	1.56	1.49	1.51
32	0.89	1.23	1.64	1.88	1.49	1.66	1.96	1.95	1.99	2.25
64	1.72	2.06	2.26	2.56	2.71	2.85	3.81	2.85	3.04	3.05
128	3.15	3.55	3.74	3.86	3.87	4.07	4.14	4.01	4.10	4.11
256	4.64	4.84	5.06	4.89	5.20	5.11	5.20	5.07	5.13	5.13
512	6.01	5.72	6.18	5.76	6.21	6.14	6.17	5.86	6.18	6.20
1024	7.29	6.91	6.32							

Πίνακας T16: CG - To speedup για την επίλυση του γραμμικού συστήματος, για 4 και 8 επεξεργαστές.

SGI - To Speedup Εκτέλεσης των Υπολογισμών										
$Np = 2$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	1.06	1.16	1.35	1.64	1.52	1.44	1.65	1.58	1.61	1.59
16	1.56	1.59	1.71	1.68	1.73	1.79	1.80	1.78	1.79	1.86
32	1.81	1.86	1.86	1.87	1.88	1.90	1.88	1.87	1.88	1.90
64	1.96	1.96	1.95	1.95	1.92	1.95	1.93	1.95	1.95	1.95
128	1.98	1.97	1.94	1.98	1.93	1.94	1.93	1.96	1.96	1.96
256	1.98	1.98	1.90	1.99	1.92	1.92	1.99	1.99	1.97	1.96
512	2.01	2.01	1.94	2.01	1.88	1.95	1.95	1.92	1.99	1.99
1024	2.01	2.01	1.97							
$Np = 4$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	1.44	1.88	1.64	2.40	2.15	2.18	2.36	2.51	2.21	2.42
16	2.34	2.67	2.54	2.94	2.93	2.90	3.01	3.01	3.08	3.24
32	3.00	3.41	3.48	3.42	3.52	3.56	3.49	3.47	3.57	3.57
64	3.83	3.83	3.78	3.79	3.77	3.84	3.78	3.77	3.82	3.81
128	3.94	3.94	3.90	3.97	3.84	3.91	3.84	3.96	3.89	3.88
256	4.03	3.96	3.84	3.97	3.84	3.91	3.84	3.97	3.94	3.92
512	4.04	4.01	3.89	4.00	3.89	3.98	3.89	3.82	3.97	3.95
1024	4.02	4.02	3.47							
$Np = 8$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	1.27	1.65	2.11	2.61	2.55	2.46	0.77	3.04	2.85	2.51
16	2.78	3.49	3.71	4.14	4.02	4.38	4.51	4.60	4.53	4.99
32	4.93	5.59	5.90	5.87	6.07	6.34	6.08	6.11	6.41	6.43
64	7.06	7.06	7.07	7.20	7.10	6.84	6.99	7.18	7.31	7.34
128	7.82	7.78	7.63	7.61	7.54	7.96	7.52	7.77	7.68	7.67
256	7.91	8.02	7.61	7.93	7.61	7.98	7.63	7.87	7.82	7.82
512	7.95	7.95	7.74	7.98	7.73	8.24	7.62	7.64	7.91	7.84
1024	8.04	8.04	6.74							

Πίνακας T17: SGI - To speedup εκτέλεσης των υπολογισμών, για 2,4 και 8 επεξεργαστές.

CM/CG - To Speedup Εκτέλεσης των Υπολογισμών										
$Np = 2$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	1.03	1.28	1.40	1.52	1.55	1.55	1.61	1.59	1.60	1.59
16	1.26	1.53	1.65	1.72	1.70	1.73	1.77	1.88	1.79	1.77
32	1.77	1.83	1.80	1.91	1.65	1.74	1.93	1.91	1.91	1.95
64	1.93	1.95	2.00	1.99	1.97	1.95	2.46	1.98	1.95	1.95
128	1.94	2.03	1.94	1.99	2.02	1.97	2.00	1.97	2.01	1.99
256	1.93	1.99	2.05	1.99	2.05	2.06	2.05	2.00	2.04	2.05
512	2.12	1.98	2.11	1.99	2.10	2.11	2.11	1.98	2.06	2.10
1024	2.05	1.95	1.77							

Πίνακας T18: CM/CG - To speedup εκτέλεσης των υπολογισμών, για 2 επεξεργαστές.

CM - To Speedup Εκτέλεσης των Υπολογισμών										
$Np = 4$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	1.24	1.55	1.86	2.00	2.07	2.15	2.31	2.16	2.26	2.38
16	2.11	2.45	2.64	2.56	2.54	2.91	3.00	3.14	2.85	2.86
32	3.11	2.96	3.25	3.34	3.21	3.31	3.52	3.58	3.52	3.75
64	3.43	3.68	3.88	3.92	3.80	3.90	4.77	3.90	3.80	3.85
128	3.99	4.03	3.91	3.92	4.00	4.01	3.91	3.97	3.91	3.98
256	3.93	3.88	4.09	3.98	4.10	4.21	4.10	4.03	4.11	4.10
512	4.00	4.02	4.24	3.98	4.23	4.34	4.24	3.99	4.18	4.21
1024	4.15	4.01	3.78							
$Np = 8$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	1.43	1.86	2.32	2.30	2.86	2.87	3.13	2.98	3.04	3.23
16	2.91	3.34	3.92	3.86	4.07	3.96	4.65	4.39	4.32	4.25
32	4.46	5.37	5.73	6.03	5.78	5.86	6.07	6.24	6.40	6.66
64	5.62	6.62	6.82	6.98	7.04	7.50	8.88	7.37	7.33	7.48
128	7.34	7.76	7.62	7.52	7.56	7.96	7.59	7.61	7.66	7.74
256	7.82	7.64	7.80	7.63	7.79	8.50	8.11	7.90	8.13	8.14
512	7.92	7.37	8.45	7.93	8.44	8.99	8.31	7.92	8.33	8.43
1024	8.86	8.05	7.29							

Πίνακας T19: CM - Το speedup εκτέλεσης των υπολογισμών, για 4 και 8 επεξεργαστές.

CG - To Speedup Εκτέλεσης των Υπολογισμών										
$Np = 4$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	1.25	1.64	1.74	2.07	2.11	2.15	2.34	2.10	2.18	2.18
16	2.13	2.34	2.43	2.43	2.49	2.89	2.71	3.23	3.05	3.10
32	2.72	3.20	3.43	3.45	3.36	3.31	3.61	3.55	3.53	3.66
64	3.50	3.77	3.91	3.90	3.85	3.92	4.77	3.85	3.80	3.78
128	3.94	4.04	3.84	3.90	3.99	4.01	3.89	3.97	3.91	3.98
256	4.01	3.87	4.03	3.98	4.09	4.20	3.99	4.00	4.09	4.10
512	4.00	4.02	4.22	3.97	4.21	4.31	4.21	3.99	4.15	4.21
1024	4.17	3.92	3.72							
$Np = 8$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	1.42	1.85	2.26	2.46	2.88	2.60	2.94	2.96	3.05	3.13
16	2.45	3.51	3.73	3.95	4.15	4.45	4.59	4.91	4.77	4.64
32	3.40	5.35	5.79	5.03	5.59	5.81	6.36	6.42	6.39	6.88
64	6.48	6.83	7.27	7.12	7.18	7.69	8.94	7.50	7.28	7.37
128	7.38	7.81	7.60	7.43	7.65	7.95	7.67	7.59	7.63	7.70
256	7.87	7.58	7.81	7.37	7.81	8.52	8.06	7.97	8.07	8.15
512	7.94	7.37	8.46	7.88	8.39	8.84	8.41	7.90	8.25	8.34
1024	8.82	7.80	7.10							

Πίνακας T20: CG - Το speedup εκτέλεσης των υπολογισμών, για 4 και 8 επεξεργαστές.

SGI - Ο Λόγος του Χρόνου Επικοινωνίας προς τον Χρόνο Εκτέλεσης των Υπολογισμών (%)										
Np = 2	n = 8	16	24	32	40	48	56	64	72	80
N = 8	18	16	18	16	24	18	15	17	19	18
16	28	24	22	22	22	19	15	21	19	17
32	23	22	19	19	17	17	11	17	17	17
64	15	15	15	14	14	14	11	17	15	15
128	10	11	10	9	9	11	9	11	10	11
256	9	6	6	6	6	6	6	6	6	5
512	3	4	3	3	3	4	3	3	3	3
1024	2	2	2							
Np = 4	n = 8	16	24	32	40	48	56	64	72	80
N = 8	42	42	28	37	31	32	22	32	31	33
16	51	46	36	42	40	38	27	38	35	36
32	48	45	38	39	36	35	25	35	31	32
64	33	32	28	28	29	32	25	31	28	28
128	19	20	19	20	20	23	17	21	20	20
256	14	12	12	12	12	16	11	12	12	12
512	8	7	6	7	7	10	6	6	7	8
1024	4	3	4							
Np = 8	n = 8	16	24	32	40	48	56	64	72	80
N = 8	72	54	54	51	56	48	10	48	48	40
16	108	92	70	72	103	76	53	74	66	67
32	113	99	85	79	71	73	51	66	84	60
64	77	67	58	60	56	61	46	60	53	57
128	54	43	36	38	37	46	35	45	39	40
256	32	26	22	28	23	32	23	25	24	25
512	19	19	14	14	13	23	13	13	14	13
1024	8	8	8							

Πίνακας T21: SGI - Ο λόγος του χρόνου επικοινωνίας προς τον χρόνο εκτέλεσης των υπολογισμών (%), για 2,4 και 8 επεξεργαστές.

CM/CG - Ο Λόγος του Χρόνου Επικοινωνίας προς τον Χρόνο Εκτέλεσης των Υπολογισμών (%)										
Np = 2	n = 8	16	24	32	40	48	56	64	72	80
N = 8	27	26	24	23	19	18	16	22	20	28
16	22	21	20	19	17	17	12	15	17	16
32	20	19	16	14	13	14	12	15	14	13
64	17	12	14	12	12	12	11	12	12	11
128	7	8	9	7	9	7	7	7	7	7
256	5	5	4	5	5	4	4	5	4	4
512	2	3	2	3	2	3	3	3	2	2
1024	3	1	7							

Πίνακας T22: CM/CG - Ο λόγος του χρόνου επικοινωνίας προς τον χρόνο εκτέλεσης των υπολογισμών (%), για 2 επεξεργαστές.

CM - Ο Λόγος του Χρόνου Επικοινωνίας προς τον Χρόνο Εκτέλεσης των Υπολογισμών (%)										
$Np = 4$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	402	434	437	470	479	520	422	525	505	557
16	576	692	674	632	595	645	506	570	553	553
32	738	687	550	528	521	552	460	540	527	508
64	520	458	483	468	429	444	390	447	391	417
128	300	344	292	284	287	266	249	271	260	269
256	173	165	166	165	160	179	157	165	165	164
512	85	99	95	94	93	99	93	92	88	89
1024	45	39	54							
$Np = 8$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	655	697	747	670	796	812	652	921	921	1009
16	1094	1294	1342	1289	1365	1287	1151	1181	1249	1220
32	1475	1791	1511	1471	1483	1517	1217	1443	1433	1415
64	1293	1268	1277	1230	1195	1269	1081	1245	1124	1204
128	825	855	842	794	800	784	719	797	759	765
256	511	482	462	463	443	512	455	471	478	476
512	248	247	265	264	263	288	253	257	244	253
1024	137	117	113							

Πίνακας T23: CM - Ο λόγος του χρόνου επικοινωνίας προς τον χρόνο εκτέλεσης των υπολογισμών (%), για 4 και 8 επεξεργαστές.

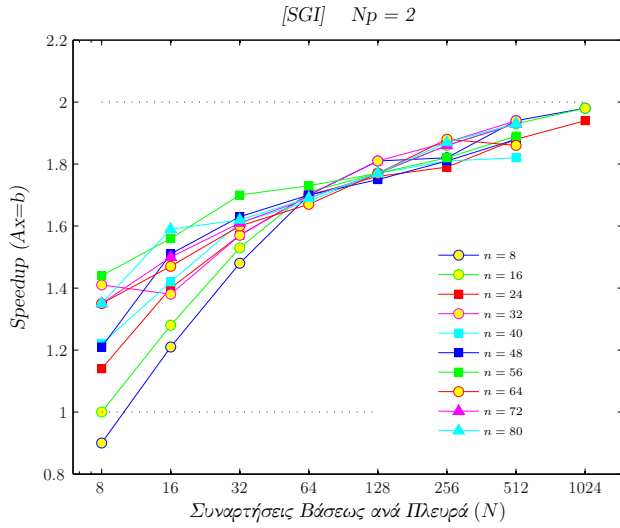
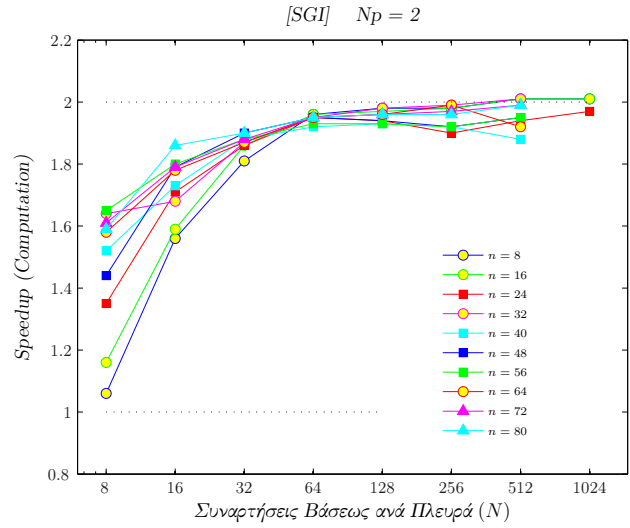
CG - Ο Λόγος του Χρόνου Επικοινωνίας προς τον Χρόνο Εκτέλεσης των Υπολογισμών (%)										
$Np = 4$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	223	252	144	180	167	181	130	124	116	112
16	288	175	145	127	111	138	81	123	138	133
32	154	157	132	103	102	101	75	82	81	86
64	121	81	72	65	59	60	52	58	51	53
128	48	41	38	36	38	36	32	36	32	35
256	27	22	23	24	24	27	22	24	24	25
512	13	20	18	18	18	20	19	18	18	17
1024	10	5	10							
$Np = 8$	$n = 8$	16	24	32	40	48	56	64	72	80
$N = 8$	431	402	322	275	260	203	158	217	218	201
16	448	408	286	260	233	258	206	213	220	207
32	279	334	253	168	275	249	224	229	221	205
64	276	231	221	178	165	170	135	163	140	142
128	134	120	103	93	97	95	85	89	86	87
256	70	57	54	51	50	67	55	57	57	59
512	32	29	37	37	35	44	36	35	33	35
1024	21	13	12							

Πίνακας T24: CG - Ο λόγος του χρόνου επικοινωνίας προς τον χρόνο εκτέλεσης των υπολογισμών (%), για 4 και 8 επεξεργαστές.

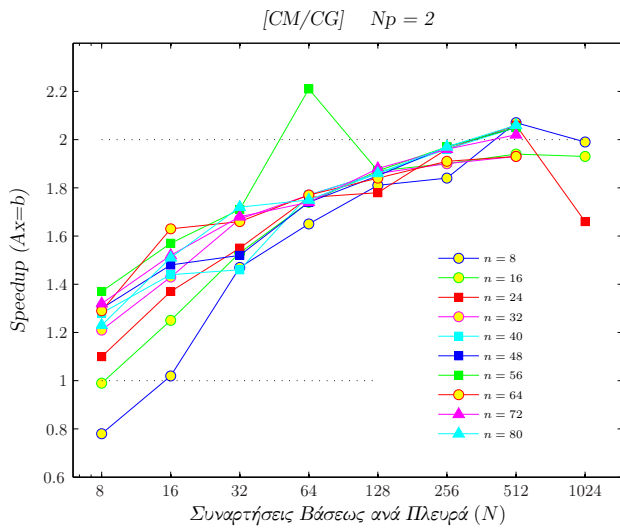
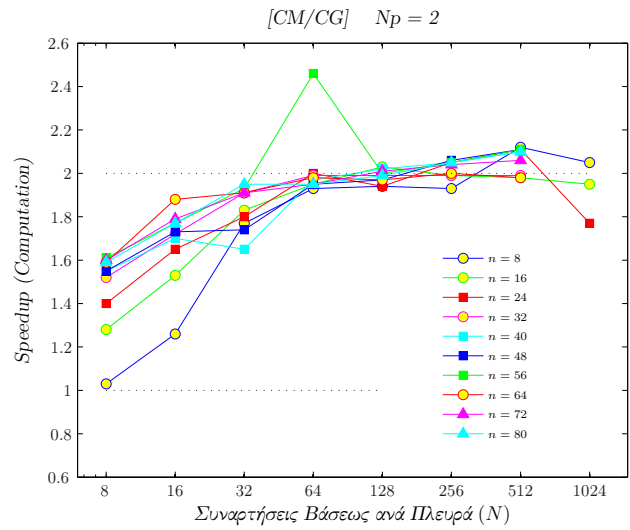
Μελετώντας τα αποτελέσματα που παρουσιάζονται στους πίνακες και τα σχήματα μπορούμε να παρατηρήσουμε τα εξής

- SGI: Το speedup για την επίλυση του συστήματος είναι πάντα μεγαλύτερο από την μονάδα εκτός από μερικές μικρές περιπτώσεις που εμφανίζονται μόνο για $N = 8$. Γενικά παρατηρούμε μια ανοδική πορεία του speedup για όλα τα πολύγωνα όσο αυξάνει το μέγεθος¹ των προβλημάτων εκτός από την μεγάλη περίπτωση ($n = 24, N = 1024$) για 2 και 4

¹ Λέγοντας μέγεθος του προβλήματος εννοούμε το γινόμενο των πλευρών του πολυγώνου με το πλήθος των συναρτήσεων βάσης ($n \times N$). Το μέγεθος του προβλήματος ταυτίζεται με την διάσταση του collocation πίνακα και όσο μεγαλύτερο είναι τόσο πιο χρονοβόρα περιμένουμε να είναι η λύση του συστήματος. Επίσης λέγοντας μικρή και μεγάλη περίπτωση (ή πρόβλημα) θα εννοούμε μικρό ή μεγάλο μέγεθος προβλήματος.

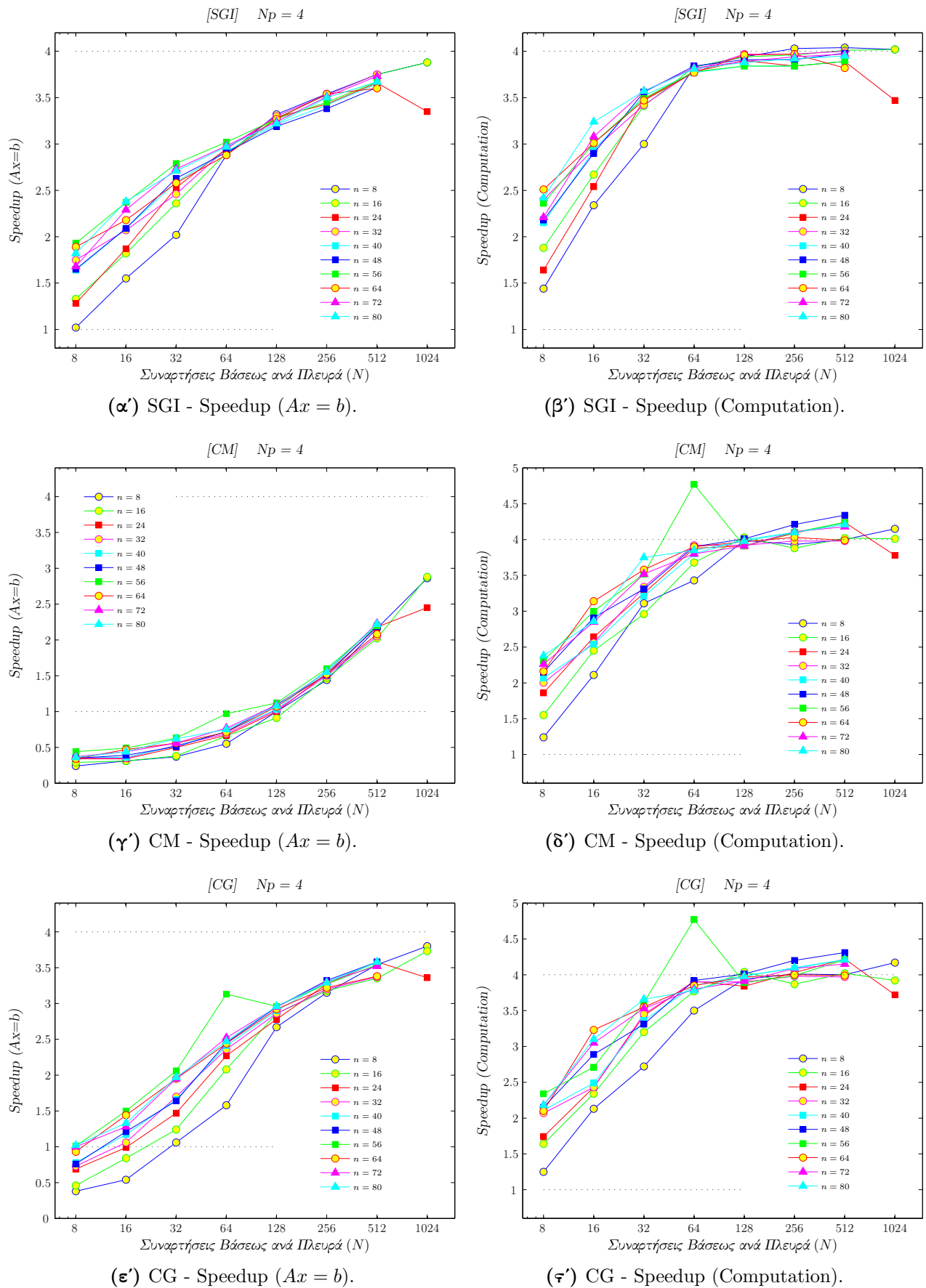
(α') SGI - Speedup ($Ax = b$).

(β') SGI - Speedup (Computation).

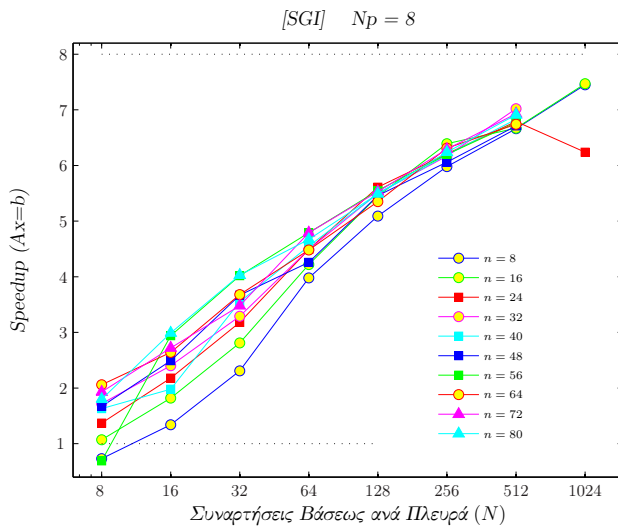
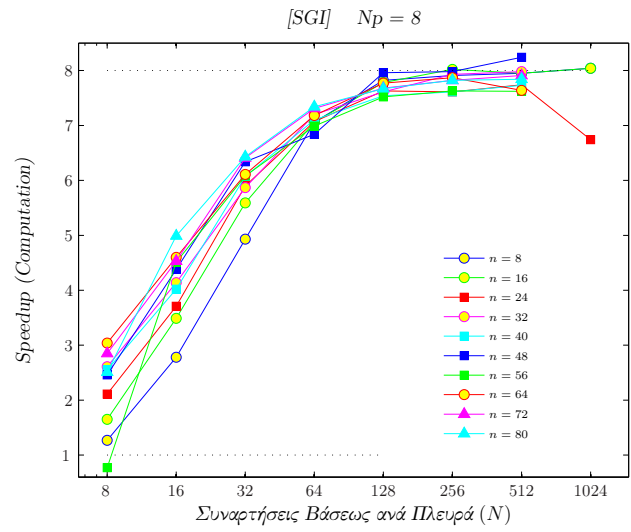
(γ') CM/CG - Speedup ($Ax = b$).

(δ') CM/CG - Speedup (Computation).

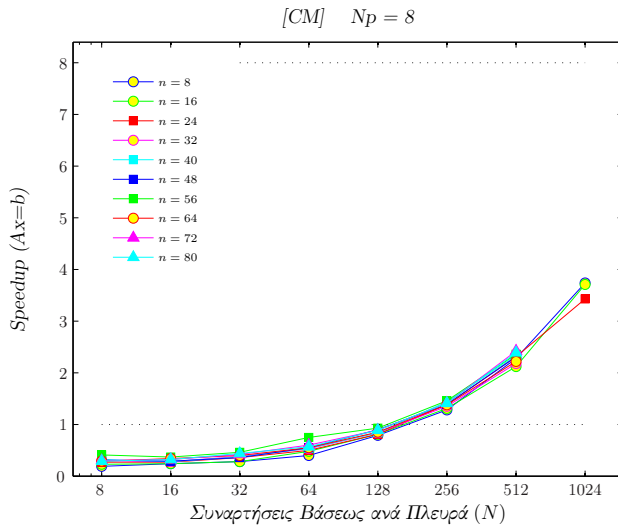
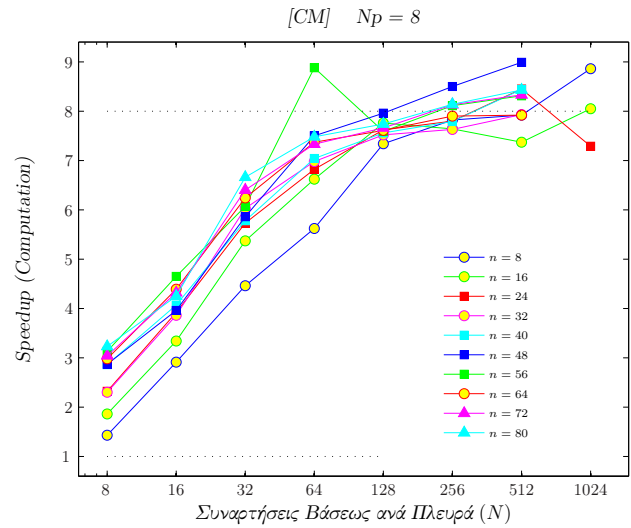
Σχήμα 8.3: Το speedup για την επίλυση του γραμμικού συστήματος (αριστερή στήλη) για 2 επεξεργαστές. Το speedup εκτέλεσης των υπολογισμών (δεξιά στήλη) για 2 επεξεργαστές.



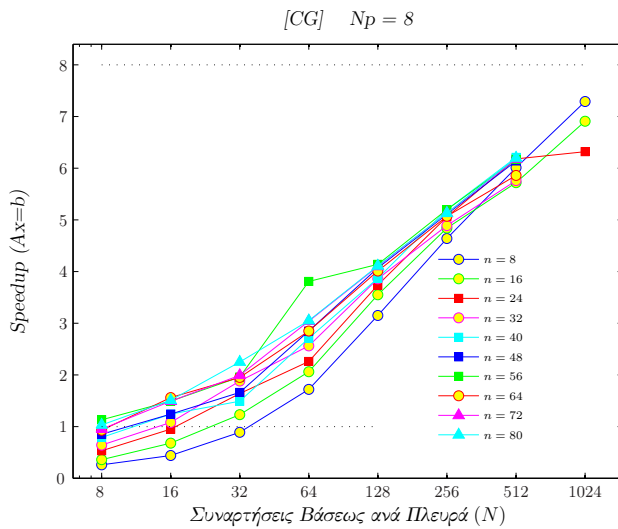
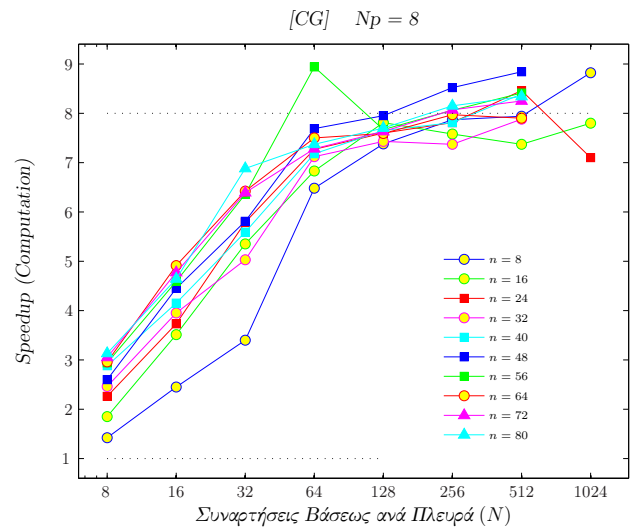
Σχήμα 8.4: Το speedup για την επίλυση του γραμμικού συστήματος (αριστερή στήλη) για 4 επεξεργαστές. Το speedup εκτέλεσης των υπολογισμών (δεξιά στήλη) για 4 επεξεργαστές.

(α') SGI - Speedup ($Ax = b$).

(β') SGI - Speedup (Computation).

(γ') CM - Speedup ($Ax = b$).

(δ') CM - Speedup (Computation).

(ε') CG - Speedup ($Ax = b$).

(ς') CG - Speedup (Computation).

Σχήμα 8.5: Το speedup για την επίλυση του γραμμικού συστήματος (αριστερή στήλη) για 8 επεξεργαστές. Το speedup εκτέλεσης των υπολογισμών (δεξιά στήλη) για 8 επεξεργαστές.

επεξεργαστές. Μάλιστα για τις άλλες δυο μεγάλες περιπτώσεις ($N = 1024$) για $n = 8$ και $n = 16$ τα speedup πρακτικά ταυτίζονται.

- CM/CG: Το speedup για την επίλυση του συστήματος για 2 επεξεργαστές είναι μεγαλύτερο της μονάδας εκτός από τις (μικρές) περιπτώσεις ($n = 8, 16, N = 8$). Όσο αυξάνονται οι επεξεργαστές τόσο αυξάνονται και οι περιπτώσεις για τις οποίες το speedup είναι μικρότερο της μονάδας στο CM (π.χ. για $Np = 4$ τα μισά προβλήματα έχουν $\text{speedup} < 1$ ενώ για $Np = 8$ τα 5/8 των προβλημάτων έχουν $\text{speedup} < 1$). Τα πράγματα βελτιώνονται στην περίπτωση του Gigabit δικτύου CG όπου για speedup μικρότερο της μονάδας παρατηρείται μόνο για περιπτώσεις όπου $N = 8, 16$. Γενικά παρατηρούμε, μια ανοδική τάση του speedup για όλα τα πολύγωνα όσο αυξάνει το N εκτός από την περίπτωση ($n = 24, N = 1024$) για δύο επεξεργαστές και τέσσερις επεξεργαστές (στο CG).
- CM/CG: Για δυο επεξεργαστές, σε μερικές περιπτώσεις (π.χ. στο 40-γωνο για $N = 512$) έχουμε υπεργραμμικό superlinear speedup.¹
- SGI / CM / CG: Το speedup εκτέλεσης των υπολογισμών φαίνεται να αυξάνει προοδευτικά όσο αυξάνει το N για όλα τα πολύγωνα εκτός από την περίπτωση ($n = 24, N = 1024$). Σε αρκετές περιπτώσεις εμφανίζεται υπεργραμμικό speedup ιδιαίτερα στην περίπτωση των cluster (CM/CG).
- SGI / CM / CG: Ο λόγος του χρόνου επικοινωνίας προς τον χρόνο εκτέλεσης των υπολογισμών και για κάθε πολύγωνο, αρχικά αυξάνει (μέχρι $N = 16$ ή $N = 32$) και κατόπιν μειώνεται. Η μείωση αυτή φτάνει και μέχρι μονοψήφια νούμερα για το SGI και το cluster στην περίπτωση των δυο επεξεργαστών.² Στην περίπτωση του CM η μείωση μετά βίας φτάνει το 50% για 4 επεξεργαστές ενώ για 8 επεξεργαστές ο λόγος διατηρείται πάντα πάνω από 100%. Στην περίπτωση τώρα του CG τα πράγματα είναι σαφώς καλύτερα αφού για 4 επεξεργαστές η μείωση φτάνει κάτω από το 20% για μεγάλες περιπτώσεις ενώ για 8 επεξεργαστές η μείωση ανέρχεται σε ποσοστά κάτω του 50% για $N = 512$ και κάτω του 22% για $N = 1024$.

Παρατήρηση 8.1 Σε μερικές περιπτώσεις ο χρόνος επίλυσης του *collocation* συστήματος για $n = 64$ είναι μικρότερος από τον χρόνο επίλυσης για $n = 56$ για τα ίδια N (π.χ. SGI για 1 επεξεργαστή και $N = 16$). Για την διερεύνηση αυτού του θέματος παραθέτουμε τον πίνακα (T25)

¹ Λέγοντας υπεργραμμικό speedup εννοούμε speedup μεγαλύτερο του αριθμού των επεξεργαστών. Αυτό αν και όχι τόσο συνηθισμένο απαντάται πολλές φορές σε παράλληλα προγράμματα και οφείλεται κυρίως στην καλύτερη οργάνωση των πράξεων που μπορεί να επιτύχει ο compiler (π.χ. λιγότερες αναφορές στην κύρια μνήμη) και όχι στην μείωση του συνολικού αριθμού των πράξεων κινητής υποδιαστολής. Ένα εξαιρετικό παράδειγμα αναφέρεται στο στο [PAR86].

² Ας υπενθυμίσουμε εδώ ότι ο κάθε κόμβος του cluster αποτελείται από δυο επεξεργαστές και επομένως όταν ένα πρόγραμμα εκτελείται σε δυο επεξεργαστές του cluster δεν υπάρχει ανάγκη επικοινωνίας μέσω του τοπικού δικτύου.

που περιέχει τον χρόνο εκτέλεσης των υπολογισμών για $n = 56$ και $n = 64$. Από αυτόν τον πίνακα βλέπουμε ότι για μικρά N ο χρόνος εκτέλεσης των υπολογισμών για $n = 64$ μπορεί να είναι μικρότερος από τον χρόνο για $n = 56$. Πώς μπορεί να συμβαίνει κάτι τέτοιο σε μία άμεση μέθοδο επίλυσης γραμμικών συστημάτων; Η απάντηση βρίσκεται στον FFT. Ο $FFT(n)$ μπορεί να θεωρηθεί ως ένας γρήγορος αλγόριθμος για τον πολλαπλασιασμό του πίνακα Fourier τάξεως n με ένα διάνυσμα n θέσεων. Το πόσο γρήγορος είναι όμως εξαρτάται άμεσα από το n . Αν το n είναι δύναμη του 2 τότε ο $FFT(n)$ πράγματι είναι της τάξεως $n \log_2 n$ αν όμως το n δεν είναι δύναμη του 2 τότε ο $FFT(n)$ δεν είναι, απαραίτητα, τάξεως $n \log_2 n$ ή είναι τάξεως $n \log_2 n$ αλλά με μεγαλύτερη σταθερά³. Στην περίπτωση μας κάναμε ένα πείραμα εκτελώντας τον $FFT(n)$ για διάφορες τιμές του n . Τα αποτελέσματα φαίνονται στο σχήμα (8.6). Από αυτό το σχήμα φαίνεται ξεκάθαρα γιατί αυτό το φαινόμενο παρουσιάζεται μόνο στο ζευγάρι ($n = 56, n = 64$). Γιατί όμως αυτό το φαινόμενο εμφανίζεται μόνο για μικρά N ; Διότι για μεγάλα N ο χρόνος επίλυσης των n γραμμικών συστημάτων [βήμα $C3/K3$] υπερβαίνει κατά πολύ τον χρόνο υπολογισμού του πίνακα J [βήμα $C2/K2$] (εκεί εμφανίζονται οι $N^2 FFT(n)$) με αποτέλεσμα να εξανεμίζεται η όποια διαφορά στην ταχύτητα των $FFT(64)$ σε σχέση με την ταχύτητα των $FFT(56)$.

Παρατήρηση 8.2 Για να μπορέσουμε να εκτιμήσουμε καλύτερα την απόδοση του αλγορίθμου, θα ήταν χρήσιμο να ξέρουμε τους επιμέρους χρόνους του κάθε βήματος (για τον σειριακό και τον παράλληλο αλγόριθμο). Για δυο περιπτώσεις ($n = 8, N = 512$) και ($n = 80, N = 512$) αυτοί οι επιμέρους χρόνοι φαίνονται στον πίνακα (T26). Έυκολα μπορούμε να παρατηρήσουμε τα εξής:

- Τα βήματα $C1/K1$ και $C4/K4$ καταναλώνουν πολύ λίγο χρόνο (λιγότερο από το 1% του συνολικού χρόνου). Αυτό σημαίνει ότι το βήμα αυτό δεν αξίζει να παραλληλοποιηθεί - το κόστος για την ανταλλαγή δεδομένων πιθανότατα θα εκμηδένιζε κάθε κέρδος από την ταχύτερη εκτέλεση των πράξεων (βλέπε κόστος βήματος $T4$).
- Τα πιο δαπανηρά βήματα είναι το $C3/K3$ δηλαδή η επίλυση των n γραμμικών συστημάτων.
- Στο CM, όταν χρησιμοποιούμε 4 επεξεργαστές τα βήματα $T1, T2$ και $T3$ είναι τόσο δαπανηρά όσο και το βήμα $K3$. Αν χρησιμοποιήσουμε και τους 8 επεξεργαστές (δηλαδή και τους τέσσερις υπολογιστές της συστάδας) τα βήματα $T1, T2$ και $T3$ είναι σχεδόν τρεις φορές πιο ακριβά από το βήμα $K3$, με αποτέλεσμα ο περισσότερος χρόνος της επίλυσης να αναλίσκεται στην ανταλλαγή δεδομένων και όχι στους καθαρούς υπολογισμούς. Στην περίπτωση του CG το κόστος των βημάτων $T1, T2$ και $T3$ είναι μικρότερο από το ένα τρίτο του χρόνου του βήματος $K3$. Εδώ φαίνεται ξεκάθαρα το κέρδος που μπορεί να έχουμε από την χρήση δικτύων υψηλών ταχυτήτων για την διασύνδεση των κόμβων ενός cluster υπολογιστών.

Παρατήρηση 8.3 Το βήμα $C2$ του (σειριακού) αλγορίθμου μπορεί να θεωρηθεί ως δυο ξεχωριστά βήματα: το βήμα $C2p$ (οι μεταθέσεις) και το βήμα $C2f$ (οι FFTs). Κάποιος θα περίμενε οι

³ Δυο αλγόριθμοι μπορεί να είναι της ίδιας τάξεως αλλά ο χρόνος για την εκτέλεση τους να διαφέρει. Π.χ. ο χρόνος εκτέλεσης ενός αλγορίθμου που απαιτεί $2N$ πράξεις είναι, θεωρητικά, διπλάσιος από τον χρόνο ενός άλλου αλγορίθμου που απαιτεί N πράξεις παρόλο που και οι δύο αλγόριθμοι είναι $O(N)$.

*FFT*s να είναι αρκετά πιο δαπανηροί από τις μεταθέσεις. Αυτό όμως δεν συμβαίνει, όπως φαίνεται από τον παρακάτω πίνακα.

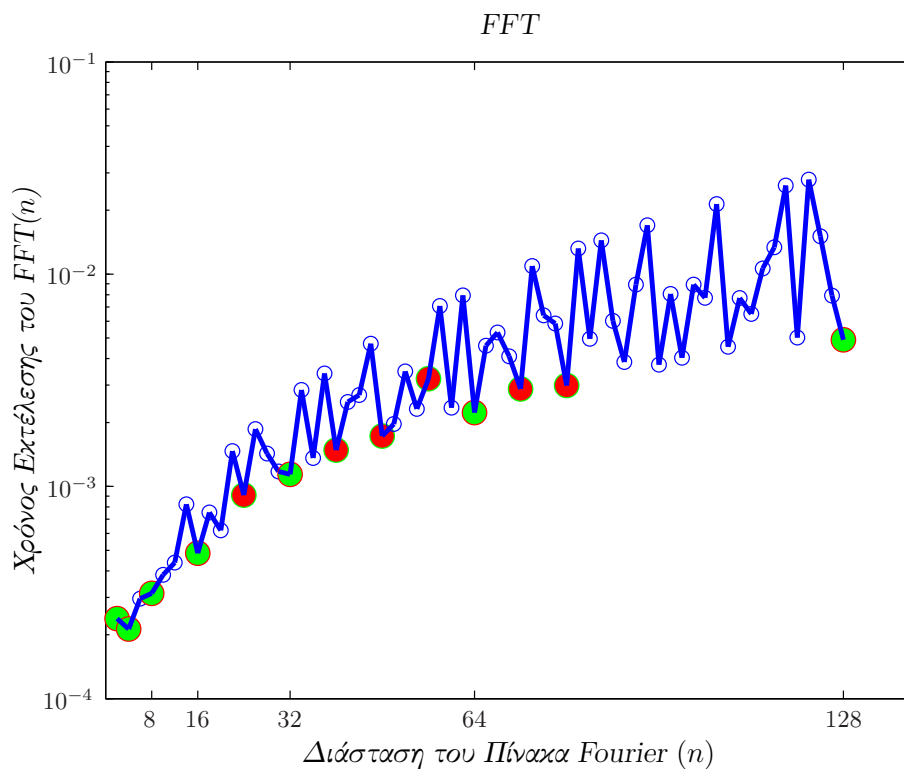
Χρόνος Βημάτων <i>C2p</i> και <i>C2f</i>		
$(n = 48, N = 128)$	<i>C2p</i>	<i>C2f</i>
<i>Origin</i> ($Np=1$)	0.154	0.149
<i>Andromeda</i> ($Np=1$)	0.064	0.063

Η παραπάνω συμπεριφορά είναι συνδυασμός του γεγονότος ότι ο *FFT* είναι πολύ "φθηνός" ενώ η μετάθεση είναι κάπως "περίπλοκη".

Παρατήρηση 8.4 Η μέθοδος *FFT* *in v* επιλύει γραμμικά συστήματα. Η πολυπλοκότητα όμως της μεθόδου δεν εξαρτάται αποκλειστικά από το μέγεθος του γραμμικού συστήματος αλλά από τις παραμέτρους n και N (παράγραφος 7.1.2). Αυτό φαίνεται πολύ καθαρά στο σχήμα (8.7a'). Το ενδιαφέρον είναι ότι αυτή η συμπεριφορά διατηρείται και στην περίπτωση του παράλληλου αλγόριθμου ακόμα και στην περίπτωση συστάδων υπολογιστών (σχήματα 8.7β' και 8.7γ'). Παρατηρήστε ότι μπορεί να υπάρξουν περιπτώσεις (π.χ. $(n = 8, N = 512), (n = 64, N = 128)$) όπου η επίλυση ενός συστήματος $2M \times 2M$ είναι κατά πολύ ταχύτερη από την επίλυση ενός συστήματος $M \times M$. Αυτή η συμπεριφορά δεν παρατηρείται στην *LU decomposition* και (για γενικούς πίνακες) ούτε στις επαναληπτικές μεθόδους.

(n=56,64) Χρόνος Εκτέλεσης των Υπολογισμών (σε δευτερόλεπτα)								
	Nr = 1		Nr = 2		Nr = 4		Nr = 8	
SGI	56	64	56	64	56	64	56	64
N = 8	0.010	0.008	0.006	0.005	0.004	0.003	0.013	0.003
16	0.034	0.027	0.019	0.015	0.011	0.009	0.008	0.006
32	0.154	0.124	0.082	0.066	0.044	0.036	0.025	0.020
64	0.781	0.715	0.404	0.367	0.207	0.190	0.112	0.100
128	5.158	5.040	2.670	2.568	1.343	1.273	0.686	0.648
256	34.633	36.336	18.023	18.223	9.027	9.147	4.538	4.619
512	251.809	277.413	129.372	144.794	64.672	72.527	33.037	36.293
CM/CG								
N = 8	0.005	0.004	0.003	0.003				
16	0.018	0.019	0.010	0.010				
32	0.087	0.087	0.045	0.046				
64	0.576	0.474	0.235	0.239				
128	2.962	3.173	1.481	1.609				
256	20.188	21.626	9.851	10.830				
512	151.100	162.818	71.610	82.130				
CM								
N = 8					0.002	0.002	0.002	0.001
16					0.006	0.006	0.004	0.004
32					0.025	0.024	0.014	0.014
64					0.121	0.121	0.065	0.064
128					0.758	0.799	0.390	0.417
256					4.924	5.362	2.488	2.737
512					35.610	40.780	18.180	20.570
CG								
N = 8					0.002	0.002	0.002	0.001
16					0.007	0.006	0.004	0.004
32					0.024	0.025	0.014	0.014
64					0.121	0.123	0.064	0.063
128					0.761	0.799	0.386	0.418
256					5.057	5.406	2.504	2.715
512					35.930	40.820	17.960	20.620

Πίνακας T25: SGI/CM/CG - Ο χρόνος εκτέλεσης των υπολογισμών για τις περιπτώσεις των πολυγώνων με 56 και 64 πλευρές και για τα τρία υπολογιστικά συστήματα.



Σχήμα 8.6: Ο χρόνος εκτέλεσης του FFT(n) της βιβλιοθήκης FFTPACK για διάφορες τιμές του n. Με έντονο πράσινο χρώμα έχουμε σημειώσει το χρόνο για δυνάμεις του 2 ενώ με κόκκινο το χρόνο για n που έχουμε χρησιμοποιήσει στα πειράματα.

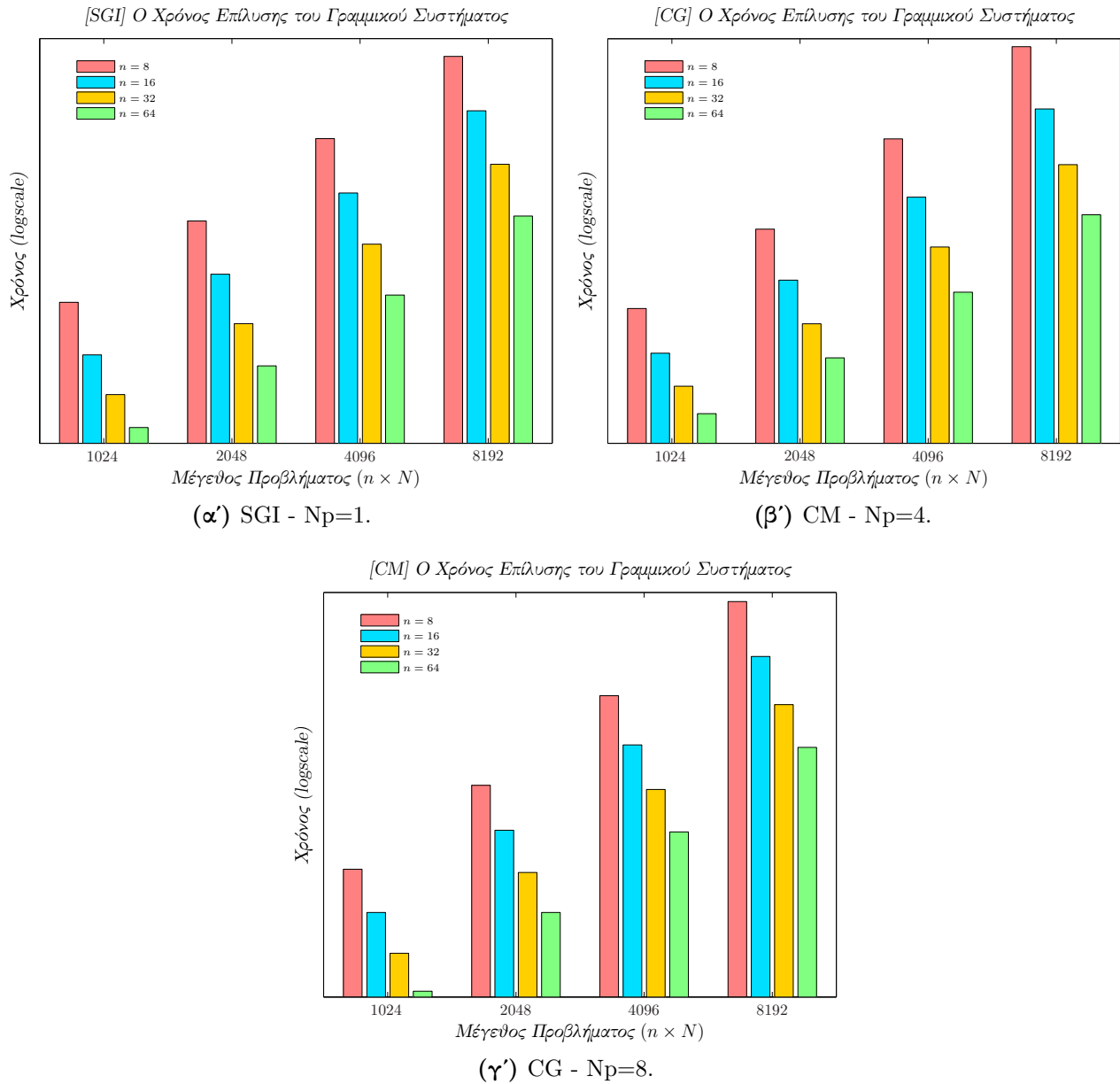
(n=8,N=512) Χρόνος των Βημάτων (σε δευτερόλεπτα)											
Np = 1			Np = 2			Np = 4			Np = 8		
Βήμα	SGI	CM/CG	Βήμα	SGI	CM/CG	SGI	CM	CG	SGI	CM	CG
C1	0.004	0.003	K1	0.003	0.002	0.003	0.002	0.002	0.003	0.002	0.002
C2	1.535	0.182	K2	0.768	0.595	0.386	0.286	0.291	0.229	0.145	0.145
C3	29.155	19.780	K3	14.858	9.197	7.452	4.871	4.876	3.875	2.454	2.462
C4	0.003	0.002	K4	0.003	0.002	0.003	0.002	0.002	0.003	0.002	0.002
			T1	0.210	0.125	0.206	1.481	0.228	0.223	2.145	0.296
			T2	0.177	0.056	0.169	1.426	0.321	0.172	2.126	0.225
			T3	0.173	0.057	0.161	1.413	0.153	0.148	2.291	0.212
			T4	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.001
			T5	0.003	0.000	0.007	0.178	0.139	0.030	0.119	0.098

(n=80,N=512) Χρόνος των Βημάτων (σε δευτερόλεπτα)											
Np = 1			Np = 2			Np = 4			Np = 8		
Βήμα	SGI	CM/CG	Βήμα	SGI	CM/CG	SGI	CM	CG	SGI	CM	CG
C1	0.041	0.043	K1	0.041	0.038	0.041	0.036	0.036	0.048	0.039	0.039
C2	18.398	2.963	K2	9.296	13.818	4.691	6.625	6.826	2.529	3.413	3.432
C3	302.811	194.104	K3	147.739	98.604	74.019	48.999	49.060	39.649	24.613	24.711
C4	0.037	0.025	K4	0.038	0.031	0.039	0.031	0.029	0.039	0.030	0.032
			T1	2.133	1.236	2.183	14.893	2.369	3.129	21.529	2.798
			T2	1.428	0.531	1.662	14.456	1.583	1.684	21.463	2.281
			T3	1.350	0.585	1.558	14.147	1.649	1.929	21.347	2.261
			T4	0.003	0.001	0.003	0.023	0.002	0.003	0.032	0.004
			T5	0.150	0.001	0.011	0.352	0.354	0.003	0.096	0.005

(n=8,N=512) Χρόνος των Βημάτων (επί % του συνολικού)											
Np = 1			Np = 2			Np = 4			Np = 8		
Βήμα	SGI	CM/CG	Βήμα	SGI	CM/CG	SGI	CM	CG	SGI	CM	CG
C1	0.01	0.01	K1	0.02	0.02	0.04	0.02	0.03	0.07	0.02	0.06
C2	5.00	0.91	K2	4.74	5.93	4.60	2.96	4.84	4.90	1.56	4.20
C3	94.98	99.07	K3	91.74	91.65	88.83	50.43	81.10	82.72	26.43	71.51
C4	0.01	0.01	K4	0.02	0.02	0.04	0.02	0.03	0.07	0.02	0.05
			T1	1.30	1.25	2.46	15.33	3.80	4.77	23.10	8.60
			T2	1.09	0.56	2.02	14.76	5.33	3.67	22.90	6.55
			T3	1.07	0.57	1.92	14.63	2.55	3.16	24.68	6.16
			T4	0.00	0.00	0.01	0.00	0.01	0.01	0.01	0.02
			T5	0.02	0.00	0.09	1.84	2.31	0.63	1.28	2.85

(n=80,N=512) Χρόνος των Βημάτων (επί % του συνολικού)											
Np = 1			Np = 2			Np = 4			Np = 8		
Βήμα	SGI	CM/CG	Βήμα	SGI	CM/CG	SGI	CM	CG	SGI	CM	CG
C1	0.01	0.02	K1	0.03	0.03	0.05	0.04	0.06	0.10	0.04	0.11
C2	5.73	1.50	K2	5.73	12.03	5.57	6.65	11.03	5.16	3.69	9.65
C3	94.25	98.46	K3	91.10	85.86	87.90	49.21	79.25	80.89	26.59	69.49
C4	0.01	0.01	K4	0.02	0.03	0.05	0.03	0.05	0.08	0.03	0.09
			T1	1.32	1.08	2.59	14.96	3.83	6.38	23.26	7.87
			T2	0.88	0.46	1.97	14.52	2.56	3.44	23.19	6.41
			T3	0.83	0.51	1.85	14.21	2.66	3.93	23.06	6.36
			T4	0.00	0.00	0.00	0.02	0.00	0.01	0.03	0.01
			T5	0.09	0.00	0.01	0.35	0.57	0.01	0.10	0.01

Πίνακας T26: SGI/CM/CG - Ο χρόνος που απαιτείται για την εκτέλεση των επιμέρους βημάτων των αλγορίθμων (8.1) και (8.2) για τις περιπτώσεις ($n = 8, N = 512$) και ($n = 80, N = 512$) και για τα τρία υπολογιστικά συστήματα. Ο πρώτος υποπίνακας παρουσιάζει τους απόλυτους χρόνους, σε δευτερόλεπτα, ενώ ο δεύτερος το ποσοστό του χρόνου του κάθε βήματος ως προς το συνολικό χρόνο του αλγορίθμου.



Σχήμα 8.7: Ο χρόνος επίλυσης του γραμμικού συστήματος για διάφορα μεγέθη προβλημάτων.

8.3 Συμπεράσματα

Σύμφωνα με όλη την παραπάνω ανάλυση μπορούμε να συμπεράνουμε ότι η άμεση μέθοδος FFTinν μπορεί να παραλληλοποιηθεί αποτελεσματικά όχι μόνο σε μηχανές κοινής μνήμης αλλά και σε συστάδες μηχανών. Παρατηρήσαμε ότι η επιτάχυνση των υπολογισμών, για μεγάλες περιπτώσεις φθάνει την βέλτιστη (τον αριθμό των επεξεργαστών) ενώ η επιτάχυνση για την λύση του συστήματος πλησιάζει αλλά δεν φθάνει, εν γένει, την βέλτιστη. Επίσης ο λόγος του χρόνου επικοινωνίας προς τον χρόνο εκτέλεσης των υπολογισμών, ιδιαίτερα στην περίπτωση στην περίπτωση της διασυνδεδεμένης με Mbps δίκτυο συστάδας, δεν είναι τόσο μικρός όσο θα θέλαμε.

Από τα αποτελέσματα που παρουσιάσαμε είναι προφανές ότι για την περαιτέρω βελτίωση του παράλληλου αλγορίθμου κανείς πρέπει να επιλύσει αποδοτικά το πρόβλημα μεταφοράς δεδομένων, πρόβλημα το οποίο θα αντιμετωπίσουμε στο μέλλον.

Κεφάλαιο 9

Υψηλής Τάξης Σύγκλιση

Η ημιτονοειδής βάση (3.4) είναι, κατά κάποια έννοια, η "φυσική" επιλογή βάσης για την αριθμητική προσέγγιση της Dirichlet to Neumann απεικόνισης λόγω της στενής σχέσης της με τους διακριτούς μετασχηματισμούς Fourier που εμφανίζονται στην (2.42). Από την μεριά όμως της θεωρίας προσεγγίσεως αυτή η βάση δεν είναι βέλτιστη. Το ερώτημα που τίθεται είναι αν η σύγκλιση της μεθόδου μπορεί να βελτιωθεί χρησιμοποιώντας μια διαφορετική επιλογή συναρτήσεων βάσης π.χ. πολυωνυμικές συναρτήσεις βάσης. Στο [FFX03] αναφέρεται η χρησιμοποίηση χαμηλού βαθμού πολυωνύμων (roof functions) για την μέθοδο FFX. Αυτή η επιλογή οδηγεί σε μεγαλύτερα σφάλματα από την ημιτονοειδή βάση. Στο παρών κεφάλαιο θα χρησιμοποιήσουμε ως συναρτήσεις βάσης πολυώνυμα προερχόμενα από τα Chebyshev πολυώνυμα πρώτου είδους [AST64, CRC03]. Θα δείξουμε ότι αυτή η επιλογή βάσης βελτιώνει κατά πολύ το ρυθμό σύγκλισης της μεθόδου [SFFS].

Για την επιλογή των collocation σημείων θα χρησιμοποιήσουμε την SFFS μέθοδο. Βέβαια τα διαγώνια μπλοκ του collocation πίνακα δεν θα είναι point διαγώνιοι πίνακες όπως ήταν στην περίπτωση της ημιτονοειδούς βάσης.

9.1 Chebyshev Βάση

Τα πολυώνυμα Chebyshev πρώτου είδους ορίζονται από την σχέση

$$T_n(x) = \cos(n \arccos x) \quad n = 0, 1, 2, \dots \quad (9.1)$$

Ο μόνος περιορισμός που έχουμε στην κατασκευή των συναρτήσεων βάσης είναι ο (2.40)

$$\phi_r^{(j)}(-\pi) = \phi_r^{(j)}(\pi) = 0,$$

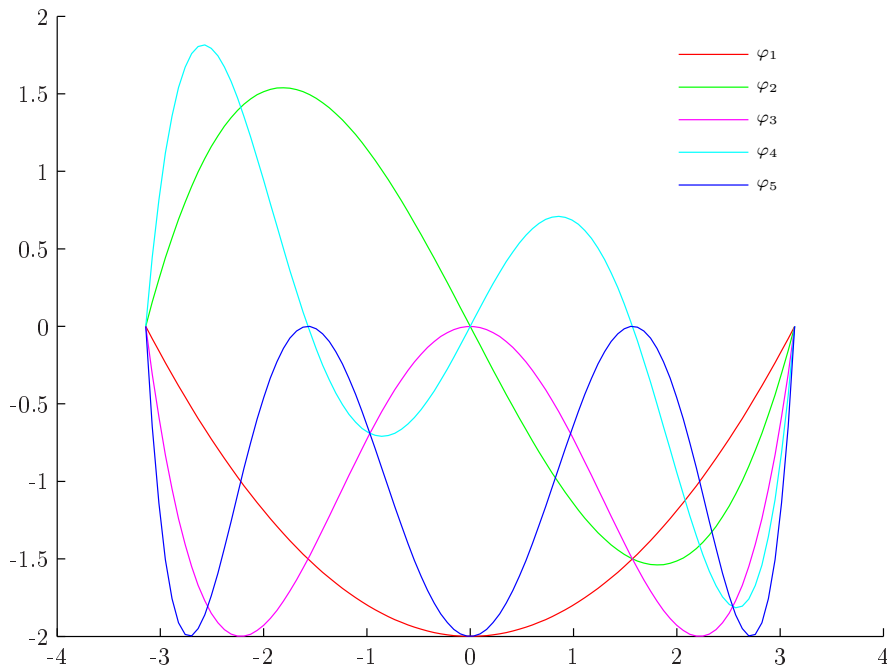
δηλαδή οι συναρτήσεις βάσης πρέπει να μηδενίζονται στα άκρα του διαστήματος $[-\pi, \pi]$. Μια οικογένεια συναρτήσεων βάσης που ικανοποιεί αυτόν τον περιορισμό είναι η

$$\varphi_r(s) = \begin{cases} T_{r+1}\left(\frac{s}{\pi}\right) - T_0\left(\frac{s}{\pi}\right) & , \quad r = 1, 3, 5, \dots \\ T_{r+1}\left(\frac{s}{\pi}\right) - T_1\left(\frac{s}{\pi}\right) & , \quad r = 2, 4, 6, \dots \end{cases}$$

ή ισοδύναμα

$$\varphi_r(s) = \begin{cases} T_{r+1}\left(\frac{s}{\pi}\right) - 1, & r = 1, 3, 5, \dots \\ T_{r+1}\left(\frac{s}{\pi}\right) - \frac{s}{\pi}, & r = 2, 4, 6, \dots \end{cases}$$

Το σχήμα (9.1) δείχνει τις 5 αρχικές ($r = 1, 2, 3, 4, 5$) συναρτήσεις βάσης.



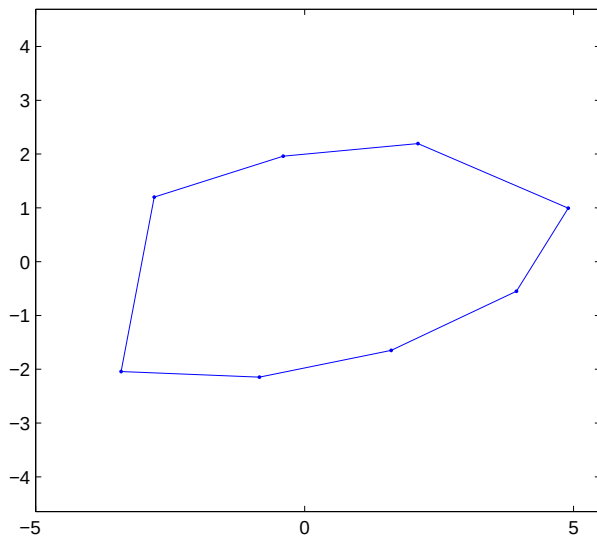
Σχήμα 9.1: Οι πολυωνυμικές συναρτήσεις βάσης που παράγονται από τα πολυώνυμα Chebyshev πρώτου είδους. Παρατηρήστε ότι όσο αυξάνει το r τόσο αυξάνονται οι "ταλαντώσεις" των συναρτήσεων.

9.2 Αριθμητικά Αποτελέσματα

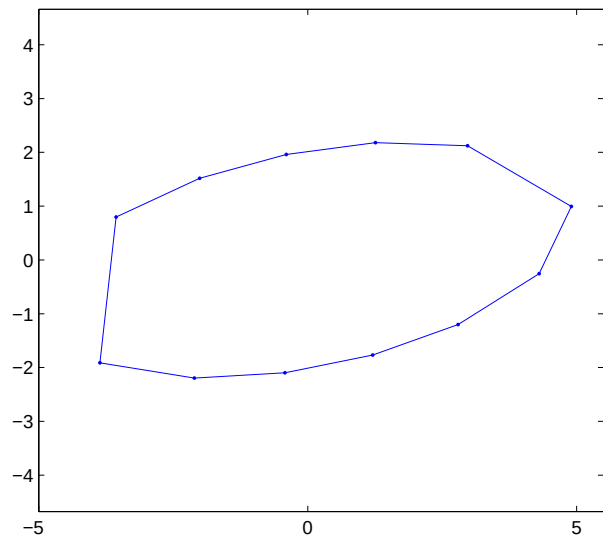
Για τον έλεγχο της απόδοσης της Chebyshev βάσης έχουμε επιλύσει την Dirichlet to Neumann απεικόνιση σε μια πληθώρα κανονικών και μη πολυγώνων και για πλήθος συναρτήσεων βάσης $N = 2, 4, 6, 8, 10, 12, 14, 16$ με τις ίδιες (Dirichlet) συνοριακές συνθήκες σε όλες τις πλευρές. Η αναλυτική λύση της διαφορικής εξίσωσης που επιλύουμε, η γεωμετρία των κανονικών πολυγώνων καθώς και το σχετικό σφάλμα που χρησιμοποιούμε είναι όπως στην παράγραφο (4.1). Οι κορυφές όλων των μη κανονικών πολυγώνων βρίσκονται πάνω στην έλλειψη

$$\frac{x^2}{5} + \frac{y^2}{2} = 1$$

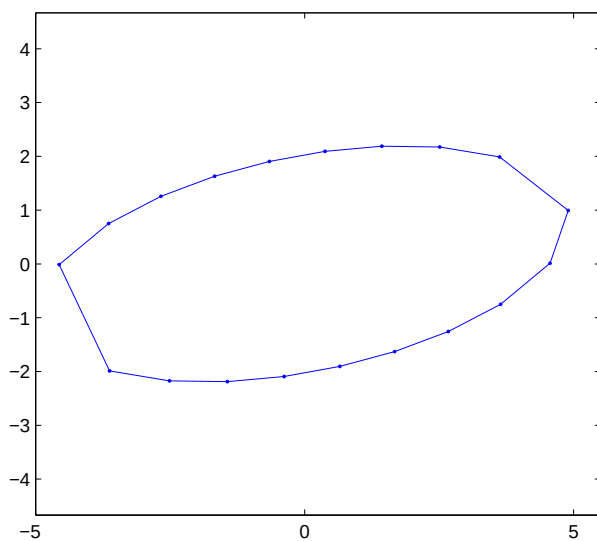
ενώ τα μή κανονικά πολύγωνα είναι στραμμένα κατά 0.2 (κατά την θετική φορά). Στο σχήμα (9.2) δείχνουμε, ενδεικτικά, τέσσερα από αυτά.



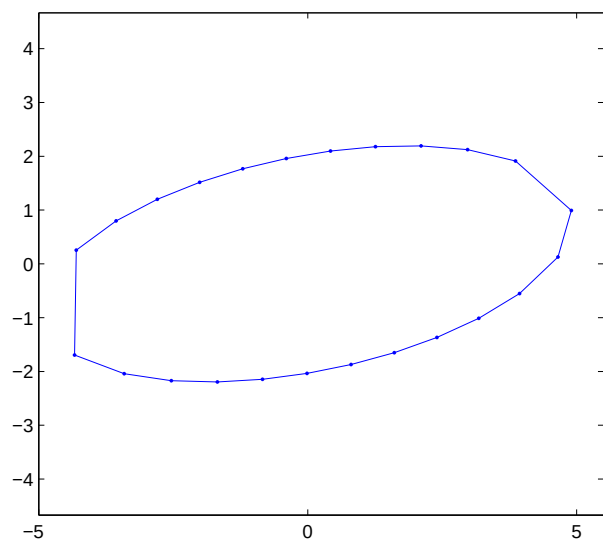
(α') Μη κανονικό πολύγωνο με 8 πλευρές.



(β') Μη κανονικό πολύγωνο με 12 πλευρές.



(γ') Μη κανονικό πολύγωνο με 19 πλευρές.



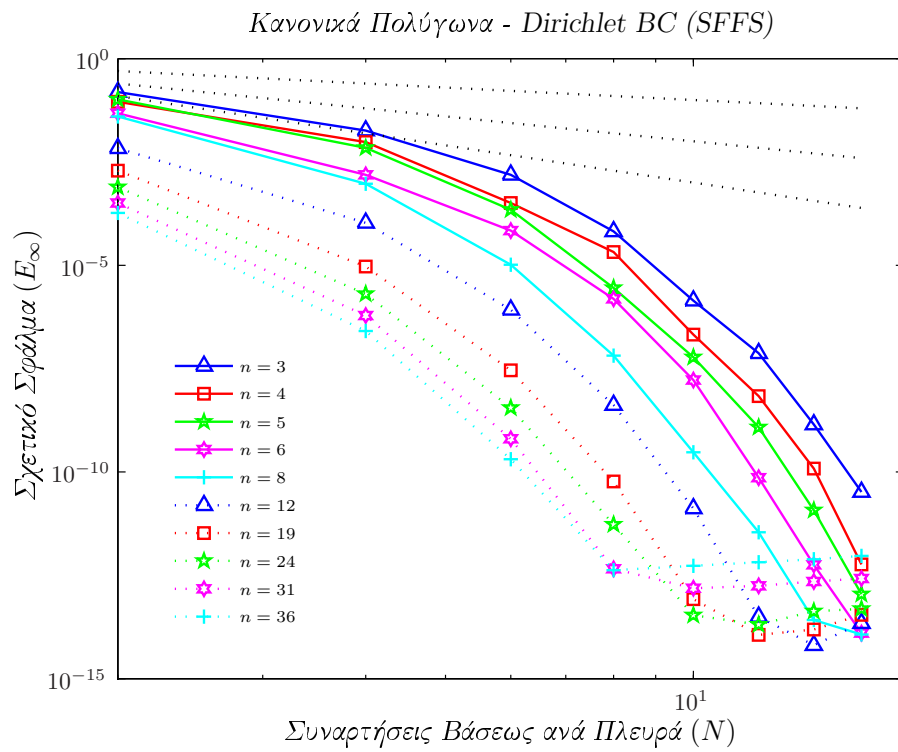
(δ') Μη κανονικό πολύγωνο με 24 πλευρές.

Σχήμα 9.2: Τα μη κανονικά πολύγωνα για την μελέτη της Chebyshev βάσης .

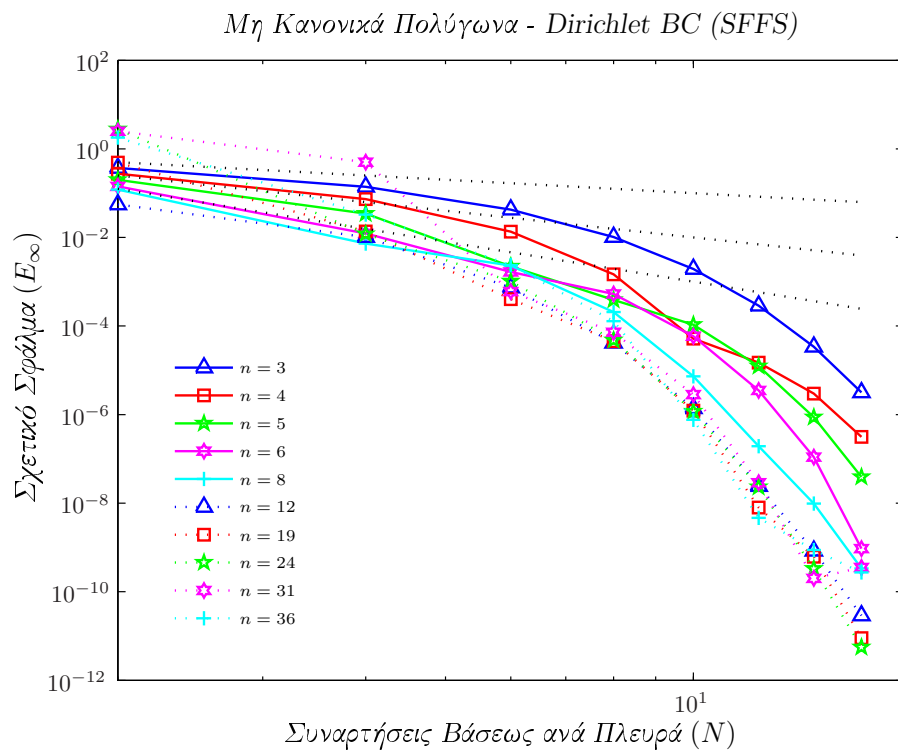
Στα σχήματα (9.3) και (9.4) φαίνεται η μείωση του σχετικού σφάλματος συναρτήσει του πλήθους των συναρτήσεων βάσης ανά πλευρά (N). Οι διακεκομμένες γραμμές στα σχήματα είναι οι ευθείες $y = \frac{1}{N}x$, $y = \left(\frac{1}{N}\right)^2 x$ και $y = \left(\frac{1}{N}\right)^3 x$.

Γενικά παρατηρούμε μια σημαντική βελτίωση του σφάλματος, σε σχέση με την ημιτονοειδή βάση. Μάλιστα στην περίπτωση των κανονικών πολυγώνων ο ρυθμός μείωσης του σφάλματος φαίνεται να είναι εκθετικός. Σε κάθε περίπτωση πάντως η τάξη σύγκλισης της μεθόδου είναι μεγαλύτερη από δυο. Στα κανονικά πολύγωνα και στις μεγάλες περιπτώσεις παρατηρούμε ένα breakdown της σύγκλισης. Αυτό είναι πιθανόν να οφείλεται σε σφάλματα στρογγύλευσης που υπεισέρχονται στους υπολογισμούς (παρατηρείστε ότι μιλάμε για σφάλματα της τάξεως του $1E-13$). Αυτό το breakdown δεν εμφανίζεται στα μη κανονικά πολύγωνα όπου η τάξη των σφαλμάτων είναι σαφώς μεγαλύτερη.

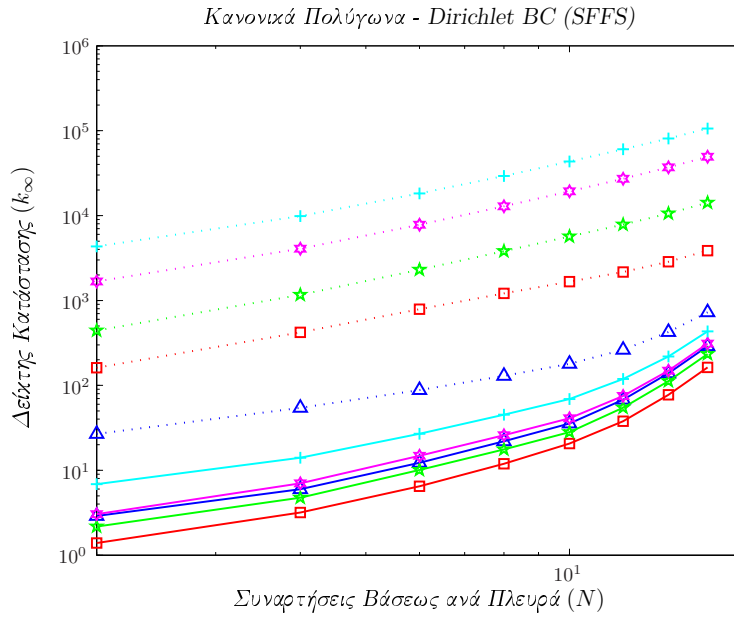
Ο δείκτης κατάστασης του collocation πίνακα (σχήματα 9.5 και 9.6) αυξάνει όσο αυξάνει και το πλήθος των συναρτήσεων βάσης. Για τα μικρά ($n < 12$) κανονικά πολύγωνα είναι πάντα μικρότερος από 500 ενώ για τα μικρά μη κανονικά πολύγωνα είναι πάντα μικρότερος από 1400. Για μεγαλύτερα πολύγωνα παίρνει αρκετά μεγάλες τιμές και ειδικά στα μή κανονικά πολύγωνα. Πάντως, στα μη κανονικά πολύγωνα, δεν φαίνεται να επηρεάζει την σύγκλιση της μεθόδου αφού, όπως φαίνεται από το σχήμα (9.4) η μέθοδος συγκλίνει ομαλά ακόμα και στα μεγάλα σε όλες τις περιπτώσεις.



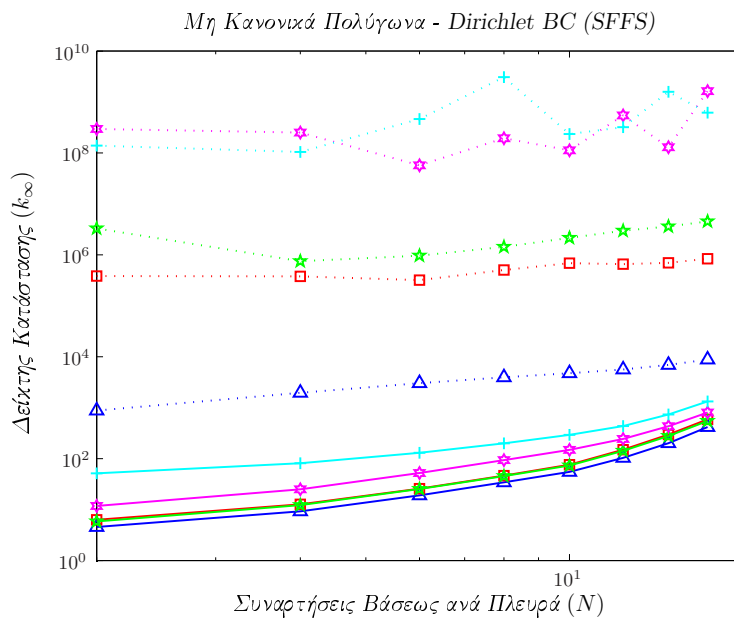
Σχήμα 9.3: Το σχετικό σφάλμα της μεθόδου SFFS για κανονικά πολύγωνα με Chebyshev συναρτήσεις βάσης και Dirichlet συνοριακές συνθήκες.



Σχήμα 9.4: Το σχετικό σφάλμα της μεθόδου SFFS για μη κανονικά πολύγωνα με Chebyshev συναρτήσεις βάσης και Dirichlet συνοριακές συνθήκες.



Σχήμα 9.5: Ο δείκτης κατάστασης του πίνακα των συντελεστών του γραμμικού συστήματος για τα κανονικά πολύγωνα με Chebyshev συναρτήσεις βάσης, υπολογισμένος με την εντολή `cond(A, inf)` της Matlab.



Σχήμα 9.6: Ο δείκτης κατάστασης του πίνακα των συντελεστών του γραμμικού συστήματος για τα μη κανονικά πολύγωνα με Chebyshev συναρτήσεις βάσης, υπολογισμένος με την εντολή `cond(A, inf)` της Matlab.

9.3 Συμπεράσματα

Το συμπέρασμα που βγαίνει από τα παραπάνω είναι ότι είναι δυνατή η χρησιμοποίηση και άλλων συναρτήσεων βάσης πέρα από την "φυσική" επιλογή που είναι οι ημιτονοειδείς συναρτήσεις βάσης. Μάλιστα η συγκεκριμένη επιλογή (Chebyshev) μας δίνει πολύ καλύτερη τάξη σύγκλισης της μεθόδου και στα κανονικά (σχεδόν εκθετική) και στα μη κανονικά (μεγαλύτερη από τετραγωνική) πολύγωνα. Η επιλογή κατάλληλων συναρτήσεων βάσης παραμένει ένα θέμα ανοιχτό για μελέτη με το οποίο θα ασχοληθούμε στο μέλλον.

Κεφάλαιο 10

Συμπεράσματα

Στην παρούσα διατριβή ασχοληθήκαμε με την νεοεισαχθήσα μέθοδο του Φωκά για την επίλυση γραμμικών και μη γραμμικών ολοκληρώσιμων Μερικών Διαφορικών Εξισώσεων. Σκοπός μας ήταν η μελέτη της αριθμητικής μεθόδου και σαν πρόβλημα μοντέλο χρησιμοποιήσαμε την εξίσωση του Laplace σε κλειστά κυρτά πολυγωνικά χωρία. Θεωρήσαμε τυχαίες συνοριακές συνθήκες στις πλευρές του πολυγώνου και διαμέσου της γενικευμένης Dirichlet to Neumann απεικόνισης υπολογίσαμε αριθμητικά τις "κάθετες" άγνωστες συνοριακές συνθήκες. Η γνώση των δυο αυτών "κάθετων" συνόλων συνοριακών συνθηκών είναι αρκετή για να μας δώσει την λύση της διαφορικής εξίσωσης στο σύνορο του πολυγώνου. Μέσω γνωστών ολοκληρωτικών μετασχηματισμών είναι δυνατή και η εύρεση της λύσης της διαφορικής εξίσωσης στο εσωτερικό του πολυγώνου.

Η αριθμητική μέθοδος έχει δυο βαθμούς ελευθερίας: την επιλογή των collocation σημείων καθώς και την επιλογή των συναρτήσεων βάσης. Εμείς προτείνουμε βελτιώσεις και για τους δυο αυτούς βαθμούς ελευθερίας. Επιλέξαμε ένα διαφορετικό σύνολο collocation σημείων (SFFS επιλογή) το οποίο οδηγεί σε, σταθερή, τετραγωνική σύγκλιση για κανονικά και μη πολύγωνα. Επίσης κατασκευάσαμε και ένα διαφορετικό σύνολο συναρτήσεων βάσης (βασισμένο στα Chebyshev πολυώνυμα 1ου είδους) η χρήση του οποίου οδηγεί σε σχεδόν εκθετική σύγκλιση της μεθόδου στα κανονικά πολύγωνα και πάντως μεγαλύτερη από τετραγωνική σε όλες τις περιπτώσεις.

Μελετώντας διάφορες επαναληπτικές μεθόδους για την επίλυση του collocation γραμμικού συστήματος καταλήξαμε στο συμπέρασμα ότι η ταχύτερη και ασφαλέστερη μέθοδος, στην περίπτωση που το πολυγωνικό χωρίο στο οποίο λύνουμε την διαφορική εξίσωση είναι κανονικό πολύγωνο και έχουμε ίδιου τύπου συνοριακές συνθήκες σε όλες τις πλευρές, είναι η μη προρυθμισμένη Bigstab. Στην περίπτωση των μη κανονικών πολυγώνων, λόγω των πολλών διαφορετικών γεωμετρικών μορφών ενός μη-κανονικού n -γώνου, είναι κάπως πρόωρο να βγάλουμε γενικά συμπεράσματα για την συμπεριφορά των επαναληπτικών μεθόδων.

Η προσεκτική μελέτη του collocation πίνακα αποκάλυψε μέρος της πλούσιας δομής του. Στην περίπτωση των κανονικών πολυγώνων με τις ίδιες συνοριακές συνθήκες σε όλες τις πλευρές ο

πίνακας είναι μπλόκ circulant ενώ στην περίπτωση των διαφορετικών συνοριακών συνθηκών ο πίνακας μπορεί να γραφεί ως συνδυασμός δυο μπλόκ circulant πινάκων με τις σταθερές να είναι διαγώνιοι πίνακες. Στην περίπτωση του τετραγώνου με ίδιες συνοριακές συνθήκες σε όλες τις πλευρές η μπλόκ circulant μορφή του collocation πίνακα μας επέτρεψε την εύρεση του φάσματος του πίνακα καθώς και του φάσματος των βασικών επαναληπτικών μεθόδων.

Η μπλόκ circulant μορφή του collocation πίνακα μας οδήγησε στην παραγωγή μιας άμεσης μεθόδου επίλυσης του (FFTin_v). Η μέθοδος αυτή οδήγησε σε σημαντική μείωση του χρόνου επίλυσης (σε μεγάλες περιπτώσεις αυτή η μείωση μπορεί να ανέλθει σε ποσοστά μεγαλύτερα του 99%).

Η αριθμητική επίλυση της γενικευμένης Dirichlet Neumann απεικόνισης σε πολύγωνα με μεγάλο πλήθος πλευρών μπορεί να οδηγήσει σε γραμμικά συστήματα των οποίων ο χρόνος επίλυσης είναι σχετικά μεγάλος. Σε αυτή την κατεύθυνση ασχοληθήκαμε με την κατασκευή ενός παράλληλου αλγόριθμου για την άμεση μέθοδο FFTin_v καταλήγοντας στο συμπέρασμα ότι η μέθοδος αυτή μπορεί να παραλληλοποιηθεί αποτελεσματικά όχι μόνο σε μηχανές κοινής μνήμης αλλά και σε δίκτυα υπολογιστών της τάξης του ενός Gbps.

Φυσικά υπάρχουν σημαντικά θέματα, όσο αφορά στην αριθμητική εφαρμογή της μεθόδου Φωκά, τα οποία χρήζουν περαιτέρω μελέτης και ευελπιστούμε ότι θα ασχοληθούμε με αυτά στο άμεσο μέλλον. Περιληπτικά αναφέρουμε:

- Αναλυτική μελέτη της συμπεριφοράς της Chebyshev βάσης.
- Εύρεση καλύτερων βάσεων.
- Αναλυτική μελέτη της συμπεριφοράς των επαναληπτικών μεθόδων για την περίπτωση των μη κανονικών πολυγώνων.
- Εύρεση του φάσματος του collocation πίνακα.
- Δομή και ιδιότητες του collocation πίνακα για κανονικά πολύγωνα με διαφορετικές συνοριακές συνθήκες σε κάθε πλευρά.
- Μελέτη και εύρεση αποδοτικών αλγορίθμων για την περίπτωση των μη κανονικών πολυγώνων - πιθανή κατηγοριοποίηση των μη κανονικών πολυγώνων.
- Βελτιστοποίηση του παράλληλου αλγόριθμου με περαιτέρω μείωση του κόστους επικοινωνίας.
- Επίλυση γενικότερων γραμμικών Ελλειπτικών Προβλημάτων Συνοριακών Τιμών.
- Εξέταση διαφορετικών τρόπων επίλυσης του collocation πίνακα λαμβάνοντας υπόψιν τις ασυμπτωτικές ιδιότητές του.

Βιβλιογραφία

- [AST64] M. Abramowitz and I. A. Stegun, **Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables**, *Dover, (Ninth Edition)*, 1964.
- [CIK95] S. L. Campbell, I. C. F. Ipsen, C. T. Kelley and C. D. Meyer, **Gmres and the Minimal Polynomial**, 1995.
- [CRC03] D. Zwillinger, **Standard Mathematical Tables and Formulae**, *Chapman and Hall, CRC Press, (31 Edition)*, 2003.
- [DAV79] P. J. Davis, **Circulant Matrices**, *Wiley-Interscience NY*, 1979.
- [DDS98] J. J. Dongarra, I. S. Duff, D. C. Sorensen and H. A. van der Vorst, **Numerical Linear Algebra for High Performance Computers**, SIAM, 1998.
- [DEM96] J. W. Demmel, **Applied Numerical Linear Algebra**, *Wiley-Interscience NY*, 1996.
- [DHE01] D. Day and M. A. Heroux, **Solving Complex-Valued Linear Systems via Equivalent Real Formulations**, *SIAM J. Sci. Comput.* **23**(2), 480-498, 2001.
- [EVA93] M. Eirmann and R. S. Varga, **Is the Optimal ω Best for the SOR Iteration Method**, *Lin. Alg. Appl. (182)*, 257-277, 1993.
- [FOK97] A. S. Fokas, **A unified transform method for solving linear and certain nonlinear PDEs**, *Proc. R. Soc. London A* **53**, 1411-1443, 1997.
- [FOK01] A. S. Fokas, **Two dimensional linear PDEs in a convex polygon**, *Proc. R. Soc. London A* **457**, 371-393, 2001.
- [FFX03] S. R. Fulton, A. S. Fokas and C. A. Xenophontos, **An analytical method for linear elliptic PDEs and its numerical implementation**, *J. Comp. Appl. Math.* preprint, 2003.
- [FRE93] R. W. Freud, **A Transpose Free quasi-minimal residual Algorithm for Non-Hermitian Linear Systems**, *SIAM J. Sci. Sta. Comp.* (13), 425-448, 1993.

- [GCL96] G. H. Golub, C. F. Van Loan, **Matrix Computations**, *The John Hopkins University Press, (Third Edition)*, 1996.
- [GPS96] A. Greenbaum and V. Pták and Z. Strakoš, **Any Nonincreasing Convergence Curve Is Possible for GMRES**, *SIAM Journal on Matrix Analysis and Applications* (17), 465-469, 1996.
- [GRE97] A. Greenbaum, **Iterative Methods for Solving Linear Systems**, *SIAM - Frontiers in Applied Mathematics* 17, (Philadelphia), 1997.
- [GST94] A. Greenbaum and Z. Strakoš, **Matrices that generate the same Krylov residual spaces**, *In Recent Advances in Iterative Methods*, Golub GH, Greenbaum A, Luskin M (eds). *IMA Volumes in Mathematics and its Applications* (60), 95-118, 1994.
- [HAD00] A. Hadjidimos, **Successive overrelaxation (SOR) and related methods**, *J. Comput. Appl. Math.* (123), 177-199, 2000.
- [HPS88] A. Hadjidimos, T. S. Papatheodorou and Y. G. Saridakis, **Optimal Block Iterative Schemes for Certain Large, Sparse and non Symmetric Linear Systems**, *Lin. Alg. Appl.* 110, 285-318, 1988.
- [HSA92] A. Hadjidimos and Y. G. Saridakis, **Modified Successive Overrelaxation (MSOR) and Equivalent 2-step Iterative Methods for Collocation Matrices**, *J. Comp. Appl. Math.* 42, 375-393, 1992.
- [HIG96] N. J. Higham, **Accuracy and Stability of Numerical Algorithms**, *SIAM*, 1996.
- [HYO81] L. A. Hageman and D. M. Young, **Applied iterative methods**, *Academic Press, New York*, (Second Edition), 1981.
- [KAN06] A. N. Kandili, **Επίλυση Ελλειπτικών Προβλημάτων σε Κανονικά Πολύγωνα με Χρήση Γνωστών Μεθόδων, Καθώς και Μεθόδων που Προκύπτουν από Νέες Μαθηματικές Αναλύσεις του Προβλήματος**, *Μεταπτυχιακή Διατριβή, Εργαστήριο Πληροφοριακών Συστημάτων Υψηλών Επιδόσεων, Τμήμα Μηχανικών Η/Υ και Πληροφορικής, Πολυτεχνική Σχολή, Πανεπιστήμιο Πατρών*, 2006.
- [KEL95] C. T. Kelley, **Iterative Methods for Linear and Nonlinear Equations**, *SIAM - Frontiers in Applied Mathematics* 16, (Philadelphia), 1995.
- [LOA92] C. F. Van Loan, **Computational Frameworks for the Fast Fourier Transform**, *SIAM - Frontiers in Applied Mathematics* 12, (Philadelphia), 1992.
- [MGE83] D. Mazancourt and D. Gerlic, **The inverse of a block-circulant matrix**, *IEEE Transactions on Antennas and Propagation* 31(5), 808-810, 1983.

- [MPS98] E. N. Mathioudakis, E. P. Papadopoulou and Y. G. Saridakis, **Non-Stationary Iterative Schemes for the Solution of Elliptic Collocation Systems**, *HERMA*, 1044-1051, 1989.
- [MPS06] E. N. Mathioudakis, E. P. Papadopoulou and Y. G. Saridakis, **Preconditioning for Solving Hermite Collocation by the BiCGSTAB**, *Trans. Math.*, 5, 811-816, 2006.
- [MRE00] M. Metcalf and J. Reid, **Fortran 90 Explained**, Oxford University Press, Inc., New York, NY, USA, 1990.
- [MTR92] J. Marsden and A. Tromba, **Διανυσματικός Λογισμός**, Πανεπιστημιακές Εκδόσεις Κρήτης (Μετάφραση Α. Γιαννόπουλος), 1992.
- [PAR86] D. Parkinson, **Parallel efficiency can be greater than unity**, *Parallel Computing*, Vol 3(3), 1986.
- [PGA99] T.C. Papanastasiou, G. Georgiou and A. Alexandrou, **Viscous Fluid Flow**, *CRC Press*, Boca Raton, FL, 1999.
- [SAA00] Y. Saad, **Iterative Methods for Sparse Linear Systems**, Second Edition, 2000.
- [SPA89] Y. G. Saridakis and T. S. Papatheodorou, **Optimal Block Accelerated Overrelaxation Method for a Large Class of 2-cyclic Matrices**, *Int. Jour. of Comp. Math.* 27, 223-242, 1989.
- [SFFS] A. G. Sifalakis, A. S. Fokas, S. Fulton and Y. G. Saridakis, **The Generalized Dirichlet-Neumann Map for Linear Elliptic PDEs and its Numerical Implementation**, *Journal of Computational and Applied Mathematics* (in press).
- [SFO93] G. L. G. Sleijpen and D. R. Fokkema, **BICGSTAB(l), for Linear Equations Involving Unsymmetric Matrices with Complex Spectrum**, *ETNA (1)*, 11-32, 1993.
- [SFPS] A. G. Sifalakis, S. Fulton, E. P. Papadopoulou and Y. G. Saridakis, **Direct and Iterative Solution of the Generalized Dirichlet-Neumann Map for Elliptic PDEs on Square Domains**, *J. Comp. and Appl. Math.* (Submitted).
- [SIF02] A. G. Sifalakis, **Επαναληπτικές Μέθοδοι Επίλυσης Γραμμικών Εξισώσεων Πεπερασμένων Στοιχείων**, *Μεταπτυχιακή Διατριβή, Εργαστήριο Εφαρμοσμένων και Υπολογιστικών Μαθηματικών, Γενικό Τμήμα, Πολυτεχνείο Κρήτης*, 2002.
- [SON89] P. Sonneveld, **CGS, a Fast Lanczos-Type Solver for Nonsymmetric Linear Systems**, *SIAM J. Sci. Sta. Comp.* 10(1), 36-52, 1989.

- [SPS] A. G. Sifalakis, E. P. Papadopoulou and Y. G. Saridakis, **Numerical Study of Iterative Methods for the Solution of the Dirichlet-Neumann Map for Linear Elliptic PDEs on Regular Polygon Domains**, *International Journal of Applied Mathematics and Computer Sciences* 4(3), 2007.
- [SSP] Y. G. Saridakis, A. G. Sifalakis and E. P. Papadopoulou, **Efficient Numerical Solution of the Generalized Dirichlet-Neumann Map for Linear Elliptic PDEs in Regular Polygon Domains**, *IMA Journal of Numerical Analysis (Submitted)*.
- [SSZ05] V. Simoncini and D. B. Szyld, **On the occurrence of superlinear convergence of exact and inexact Krylov subspace methods**, *SIAM Review* 2005 (47), 247-272, 2005.
- [SSZ07] V. Simoncini and D. B. Szyld, **Review Article: Recent computational developments in Krylov subspace methods for linear systems**, *Numerical Linear Algebra with Applications* (14), 1-59, 2007.
- [SUN94] **What Every Computer Scientist Should Know About Floating-Point Arithmetic**, ©Sun Microsystems, Inc., 1994.
- [TBA97] L. N. Trefethen and D. Bau III, **Numerical Linear Algebra**, SIAM, 1997.
- [TEM] R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. M. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine and H. Van der Vorst, **Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods**, *Electronic Version of SIAM Templates* , (Second Edition).
- [TIA00] Yongge Tian, **Rank Equalities Related to Generalized Inverses of Matrices and Their Applications**. 2000.
- [VAR00] R. S. Varga, **Matrix Iterative Analysis**, SPRINGER , (Second Edition), 2000.
- [VES97] R. Vescovo, **Inversion of Block-Circulant Matrices and Circular Array Approach**, *IEEE Transactions on Antennas and Propagation* 45(10), 1565-1567, 1997.
- [VOR92] H. A. van der Vorst, **Bi-CGSTAB, a Fast an Smoothly Converging Variant of Bi-CG for the Solution of Nonsymmetric Linear Systems**, *SIAM J. Sci. Sta. Comp.* (12), 631-634, 1992.
- [VOR03] H. Van Der Vorst, **Iterative Krylov Methods for Large Linear Systems**, *Cambridge Monographs on Applied and Computational Mathematics*, (First Edition), 2003.

- [YEI70] D. M. Young and H. E. Eidson, **On the Determination of the Optimum Relaxation Factor for the SOR Method when the Eigenvalues of the Jacobi Method are Complex**, *Report CNA-1 (Center of Numerical Analysis Univ. Texas Austin, Texas)*, 1970.
- [YOU54] D. M. Young, **Iterative Methods for Solving Partial Difference Equations of Elleiptic Type**, *Trans. Amer. Math Soc.* (76), 92-111, 1954.
- [YOU81] D. M. Young, **Iterative Solution of Large Linear Systems**, *Academic Press* New York, 1981.
- [WAT02] D. S. Watkins, **Fundamentals of Matrix Computations**, *Wiley-Interscience Series of Texts. Monographs and Tracts - Pure and Applied Mathematics (Second Edition)*, 2002.
- [FFTPACK] <http://www.netlib.org/fftpack/>
- [LAPACK] <http://www.netlib.org/lapack/>
- [QUADPACK] <http://www.netlib.org/quadpack/>

Παράρτημα



ELSEVIER



ScienceDirect

Journal of Computational and Applied Mathematics ■■■ (■■■) ■■■–■■■

JOURNAL OF
COMPUTATIONAL AND
APPLIED MATHEMATICSwww.elsevier.com/locate/cam

The generalized Dirichlet–Neumann map for linear elliptic PDEs and its numerical implementation[☆]

A.G. Sifalakis^a, A.S. Fokas^b, S.R. Fulton^c, Y.G. Saridakis^{a,*}^a*Applied Mathematics and Computers Lab, Department of Sciences, Technical University of Crete, 73100 Chania, Greece*^b*Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, UK*^c*Department of Mathematics and Computer Science, Clarkson University, Potsdam, NY 13699-5815, USA*

Received 30 October 2006; received in revised form 19 June 2007

Abstract

A new approach for analyzing boundary value problems for linear and for integrable nonlinear PDEs was introduced in Fokas [A unified transform method for solving linear and certain nonlinear PDEs, *Proc. Roy. Soc. London Ser. A* 53 (1997) 1411–1443]. For linear elliptic PDEs, an important aspect of this approach is the characterization of a generalized Dirichlet to Neumann map: given the derivative of the solution along a direction of an arbitrary angle to the boundary, the derivative of the solution perpendicularly to this direction is computed *without* solving on the interior of the domain. This is based on the analysis of the so-called *global relation*, an equation which couples known and unknown components of the derivative on the boundary and which is valid for all values of a complex parameter k . A collocation-type numerical method for solving the *global relation* for the Laplace equation in an arbitrary bounded convex polygon was introduced in Fulton et al. [An analytical method for linear elliptic PDEs and its numerical implementation, *J. Comput. Appl. Math.* 167 (2004) 465–483]. Here, by choosing a different set of the “*collocation points*” (values for k), we present a significant improvement of the results in Fulton et al. [An analytical method for linear elliptic PDEs and its numerical implementation, *J. Comput. Appl. Math.* 167 (2004) 465–483]. The new collocation points lead to well-conditioned collocation methods. Their combination with sine basis functions leads to a collocation matrix whose diagonal blocks are point diagonal matrices yielding efficient implementation of iterative methods; numerical experimentation suggests quadratic convergence. The choice of Chebyshev basis functions leads to higher order convergence, which for regular polygons appear to be exponential.

© 2007 Elsevier B.V. All rights reserved.

MSC: 35J25; 65N35; 64N99; 65F05; 65F10

Keywords: Elliptic PDEs; Dirichlet–Neumann map; Global relation; Collocation; Iterative methods

1. Introduction

A new approach for analyzing boundary value problems for linear and for integrable nonlinear PDEs in two dimensions was introduced in [5,6]. This approach was developed by several researchers, and has led to the emergence of a

[☆] This work was supported by the Greek Ministry's of Education *EPEAEK-Herakleitos* grant which is partially funded by the EU.

* Corresponding author. Tel.: +30 2821037740; fax: +30 2821037842.

E-mail address: yiannis@science.tuc.gr (Y.G. Saridakis).

new method for solving boundary value problems. This method involves two novel features:

- (a) It yields an analytic representation of the solution in the form of an integral in the complex k -plane.
- (b) It characterizes a generalized Dirichlet to Neumann map through the solution of the so-called *global relation*, which is an equation valid for all complex values of k , which couples specified and unknown values of the solution and its derivatives on the boundary.

For a large class of boundary value problems, the global relation can be solved analytically, and hence the generalized Dirichlet to Neumann map can be constructed in closed form. This includes linear evolution PDEs with spatial derivatives of arbitrary order on the half-line [7] and on a finite interval [10], the Laplace, the bi-harmonic and the modified Helmholtz equation in certain simple polygons [1–4] and the basic nonlinear integrable evolution PDEs on the half-line for certain simple boundary conditions [8,9]. However, for general boundary value problems, the global relation must be solved numerically.

The implementation of the new method to the case of the Laplace equation in an arbitrary bounded convex polygon was presented in [11], where:

- (a) It was shown that the global relation is not only a necessary but also a sufficient condition for existence. This reduces the problem of solving Laplace's equation, to the problem of solving the global relation.
- (b) A collocation-type numerical method was introduced for solving the global relation. This method is based on evaluating both the real and imaginary parts of the global relation at the following values of the complex parameter k ("*collocation points*"):

$$k = -\frac{l}{h_p}, \quad h_p = \frac{1}{2\pi}(z_{p+1} - z_p), \quad p = 1, 2, \dots, n, \quad l = 1, 2, \dots, \frac{N}{2}, \quad (1.1)$$

where n is the number of the sides of the polygon, $\{z_i\}_{i=1}^n$ are the corners of the polygon in the complex z -plane (with $z_{n+1} = z_1$) and N , which is chosen to be even, is the number of points used for the discrete approximation of the unknown boundary values. This choice has been motivated by the analytical integral representation in [6] (see also Remark 2.1).

In this paper, aiming to improve and stabilize the order of convergence and the associated conditioning number of the collocation method in [11], we use a different set of *collocation points* (the values of the complex parameter k): we use the values specified by Eq. (1.1) for the *imaginary part* of the global relation, but we choose the values

$$l = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots, \frac{N-1}{2}, \quad (1.2)$$

for the *real part* of the global relation.

The above choice, for sine basis functions, yields a linear algebraic system whose diagonal blocks of the associated coefficient matrix are point diagonal matrices. That is, the collocation coefficient matrix is by construction *block-Jacobi preconditioned*, pointing directly to the efficient implementation of iterative techniques. Numerical investigation suggests a quadratically convergent and well-conditioned collocation method. The classical Jacobi and Gauss–Seidel as well as the BiCGstab iterative techniques converge rapidly, yielding the discrete solution (for fixed N) after only a few iterations. Furthermore, the rate of convergence to the continuous solution as $N \rightarrow \infty$ can be improved by choosing a different set of basis functions. In particular, for Chebyshev basis functions, numerical experiments suggest that, for regular polygons, the convergence is exponential.

This paper is organized as follows: Section 2 outlines some of the analytical results of [11]. Section 3 presents the details of the new numerical solution of the global relation and explains the choice of the new collocation points. Section 4 presents numerical results for a variety of domains with different boundary conditions, and compares these results with the analogous results of [11]. Section 5 discusses further these results.

2. Overview

For elliptic PDEs in two dimensions x and y , it is convenient to replace the Cartesian coordinates x and y with the complex coordinates z and $\bar{z}$,

$$z = x + iy, \quad \bar{z} = x - iy. \quad (2.1)$$

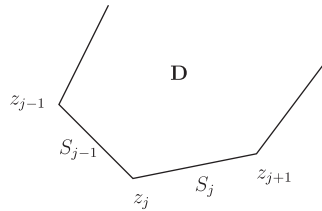


Fig. 1. Part of the bounded convex polygon with vertices z_j , sides S_j and interior D .

Using the equations

$$\partial_z = \frac{1}{2}(\partial_x - i\partial_y), \quad \partial_{\bar{z}} = \frac{1}{2}(\partial_x + i\partial_y),$$

it follows that the Laplace equation in the independent variable q can be rewritten in the form

$$\frac{\partial^2 q}{\partial z \partial \bar{z}} = 0. \quad (2.2)$$

This equation is equivalent to the equation

$$\frac{\partial}{\partial \bar{z}} \left(e^{-ikz} \frac{\partial q}{\partial z} \right) = 0, \quad (2.3)$$

for an arbitrary complex parameter k .

Suppose that the real-valued function $q(z, \bar{z})$ satisfies Laplace's equation (2.2) in a simply connected bounded domain D with boundary ∂D . Then, Eq. (2.3) together with the complex form of Green's theorem imply the equation

$$\int_{\partial D} e^{-ikz} \frac{\partial q}{\partial \bar{z}} dz = 0, \quad k \in \mathbb{C}. \quad (2.4)$$

Following [6] we shall refer to Eq. (2.4) as the *global relation* associated with Laplace's equation.

Suppose that $q(z, \bar{z})$ satisfies the Laplace equation in a convex bounded polygon with vertices $z_1, z_2, \dots, z_n$ (modulo n), which have indexed counterclockwise, see Fig. 1. Then the global relation (2.4) becomes

$$\sum_{j=1}^n \int_{S_j} e^{-ikz} \frac{\partial q}{\partial \bar{z}} dz = 0, \quad k \in \mathbb{C}, \quad (2.5)$$

where S_j denotes the side from z_j to z_{j+1} (not including the end points).

Proposition 2.1. *Let the real-valued function $q(z, \bar{z})$ satisfy the Laplace equation in the interior D of the polygon with corners $\{z_i\}_{i=1}^n$ depicted in Fig. 1. Let $g^{(j)}$ denote the derivative of the solution in the direction making an angle β_j , $0 \leq \beta_j \leq \pi$, with the side S_j , i.e.,*

$$\cos(\beta_j)q_s^{(j)} + \sin(\beta_j)q_n^{(j)} = g^{(j)}, \quad z \in S_j, \quad 1 \leq j \leq n, \quad (2.6)$$

where $q_s^{(j)}$ and $q_n^{(j)}$ denote the tangential and (outward) normal components of $\partial q / \partial \bar{z}$ along the side S_j . Let $f^{(j)}$ denote the derivative of the solution in the direction normal to the above direction, i.e.,

$$-\sin(\beta_j)q_s^{(j)} + \cos(\beta_j)q_n^{(j)} = f^{(j)}, \quad z \in S_j, \quad 1 \leq j \leq n. \quad (2.7)$$

The generalized Dirichlet–Neumann map, i.e., the relation between the sets $\{f^{(j)}\}_{j=1}^n$ and $\{g^{(j)}\}_{j=1}^n$, is characterized by the single equation

$$\sum_{j=1}^n |h_j| e^{i(\beta_j - km_j)} \int_{-\pi}^{\pi} e^{-ikh_j s} (f^{(j)}(s) - ig^{(j)}(s)) ds = 0, \quad k \in \mathbb{C}, \quad (2.8)$$

where

$$h_j := \frac{1}{2\pi}(z_{j+1} - z_j), \quad m_j := \frac{1}{2}(z_{j+1} + z_j), \quad j = 1, 2, \dots, n, \quad z_{n+1} = z_1. \quad (2.9)$$

Evaluating Eq. (2.8) on the following n -rays of the complex k -plane

$$k_p = -\frac{l}{h_p}, \quad l \in \mathbb{R}^+, \quad p = 1, \dots, n, \quad (2.10)$$

and multiplying the resulting equations by $\exp[-i(\beta_p + lm_p/h_p)]/|h_p|$, Eq. (2.8) yields the following set of n equations:

$$\sum_{j=1}^n \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-(il/h_p)(m_p - m_j)} \int_{-\pi}^{\pi} e^{il(h_j/h_p)s} (f^{(j)}(s) - ig^{(j)}(s)) ds = 0, \quad l \in \mathbb{R}^+, \quad p = 1, \dots, n. \quad (2.11)$$

Proof. Let a_j denote the angle of the side S_j with the horizontal axis, i.e.,

$$a_j = \arg(z_{j+1} - z_j), \quad h_j = |h_j|e^{ia_j}. \quad (2.12)$$

Eqs. (2.6) and (2.7) can be considered as two equations expressing $q_s^{(j)}$ and $q_n^{(j)}$ in terms of $g^{(j)}$ and $f^{(j)}$,

$$q_n^{(j)} = \sin(\beta_j)g^{(j)} + \cos(\beta_j)f^{(j)},$$

$$q_s^{(j)} = \cos(\beta_j)g^{(j)} - \sin(\beta_j)f^{(j)}.$$

Substituting these expressions in the identity

$$\frac{\partial q^{(j)}}{\partial z} = \frac{1}{2}e^{-ia_j}(q_s^{(j)} + iq_n^{(j)}),$$

we find

$$\frac{\partial q^{(j)}}{\partial z} = \frac{1}{2}e^{-i(a_j - \beta_j)}(g^{(j)} + if^{(j)}). \quad (2.13)$$

In order to evaluate the integral of the expression $\exp[-ikz]\partial q^{(j)}/\partial z$ along the side S_j , we introduce the following local coordinates: For $z \in S_j$, we let

$$z = \frac{1}{2}(z_j + z_{j+1}) + \frac{s}{2\pi}(z_{j+1} - z_j), \quad -\pi < s < \pi.$$

Note that for $s = -\pi$ and π , $z = z_i$ and z_{i+1} , respectively. Writing the above expression in terms of h_j and m_j , we find

$$z = m_j + sh_j. \quad (2.14)$$

Substituting the expressions for z and $\partial q^{(j)}/\partial z$ from Eqs. (2.14) and (2.13) into the global relation (2.5), and multiplying the resulting equation by $-i$ we find Eq. (2.8). $\square$

Remark 2.1. Suppose that the set $\{g^{(j)}\}_{j=1}^n$ is known. Then, Eq. (2.8) is a *single* equation for the n unknown functions $\{f^{(j)}(s)\}_{j=1}^n$. However, this equation is valid for *all* complex values of k . Thus, it is possible to generate a multitude of equations from the single equation (2.8), by prescribing a multitude of values for k . It is shown in [11] that a convenient choice for these values is the one defined in Eq. (2.10). This choice can be motivated as follows: suppose we only had one unknown $f^{(p)}$; then we would choose k in such a way that kh_p is real, i.e., $kh_p = -l$ or $kh_p = l$. The latter choice coincides with the values of k appearing in the analytical integral representation (see [6]). However, the integral representation involves e^{ikz} whereas the global relation involves e^{-ikz} , thus we now choose $kh_p = -l$.

Remark 2.2. In what follows we assume that $g^{(j)}$ are compatible in the corners and we seek a solution in the function class specified in [11].

Remark 2.3. The end values of the unknown functions $f^{(j)}$ can be calculated by the continuity requirements $q_z^{(j)}(z_j) = q_z^{(j-1)}(z_j)$. Namely, using expression (2.13) and setting $\delta_j = a_j - \beta_j$ we obtain (see also [11])

$$f^{(j)}(\pi) = \frac{\cos(\delta_{j+1} - \delta_j)g^{(j)}(\pi) - g^{(j+1)}(\pi)}{\sin(\delta_{j+1} - \delta_j)},$$

$$f^{(j)}(-\pi) = \frac{g^{(j-1)}(\pi) - \cos(\delta_j - \delta_{j-1})g^{(j)}(-\pi)}{\sin(\delta_j - \delta_{j-1})}.$$

3. The numerical method

Proposition 3.1. The generalized Dirichlet–Neumann map is defined by Proposition 2.1. Suppose that the set $\{g^{(j)}\}_{j=1}^n$ is given. Suppose that $f^{(j)}(s)$ is approximated by

$$f_N^{(j)}(s) = f_*^{(j)}(s) + \sum_{r=1}^N U_r^j \varphi_r(s), \quad j = 1, \dots, n, \quad N \text{ even integer}, \quad (3.1)$$

where $\varphi_r(s)$ are appropriate basis functions and

$$f_*^{(j)}(s) = \frac{1}{2\pi}[(s + \pi)f^{(j)}(\pi) - (s - \pi)f^{(j)}(-\pi)].$$

Then, the real coefficients U_r^j satisfy the Nn algebraic set of equations

$$\sum_{j=1}^n \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-i(l/h_p)(m_p - m_j)} \sum_{r=1}^N U_r^j F_r \left(\frac{lh_j}{h_p} \right) = G_p(l), \quad p = 1, 2, \dots, n, \quad (3.2)$$

where $G_p(l)$ denotes the known function

$$G_p(l) = i \sum_{j=1}^n \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-i(l/h_p)(m_p - m_j)} \int_{-\pi}^{\pi} e^{il(h_j/h_p)s} (g^{(j)}(s) + i f_*^{(j)}(s)) ds, \quad (3.3)$$

$F_r(l)$ denotes the integral

$$F_r(l) = \int_{-\pi}^{\pi} e^{ils} \varphi_r(s) ds, \quad r = 1, 2, \dots, N, \quad (3.4)$$

and l is chosen as follows: For the imaginary part of Eq. (3.2), $l = 1, 2, \dots, N/2$, whereas for the real part of Eq. (3.2), $l = \frac{1}{2}, \frac{3}{2}, \dots, N - 1/2$.

In the particular case that $\varphi_r(s)$ are the sine basis functions, that is,

$$\varphi_r(s) = \sin \left[r \left(\frac{\pi + s}{2} \right) \right], \quad r = 1, 2, \dots, N, \quad (3.5)$$

the term $j = p$ in the summation of the left-hand side of Eq. (3.2) is given by the following:

$$\begin{aligned} l = \frac{1}{2} : \pi U_1^p, \quad l = 1 : -\pi U_2^p, \\ l = \frac{3}{2} : -\pi U_3^p, \quad l = 2 : \pi U_4^p, \\ l = \frac{5}{2} : \pi U_5^p, \quad l = 3 : -\pi U_6^p, \\ \vdots \quad \quad \quad \vdots \end{aligned} \quad (3.6)$$

Proof. Substituting the representation (3.1) in Eqs. (2.11) we find Eqs. (3.2).

The term $j = p$ in the summation of the left-hand side of Eqs. (3.2) is given by

$$\sum_{r=1}^N U_r^j F_r(l) = \sum_{r=1}^N U_r^j R_r(l) + i \sum_{r=1}^N U_r^j I_r(l), \quad (3.7)$$

where $R_r(l)$ and $I_r(l)$ denote the real and imaginary parts of the integral $F_r(l)$ defined by Eq. (3.4). Let δ_{ab} denote the Kronecker function. Using

$$\begin{aligned} R_r\left(\frac{1}{2}\right) &= \pi\delta_{r1}, & R_r\left(\frac{3}{2}\right) &= -\pi\delta_{r3}, & R_r\left(\frac{5}{2}\right) &= \pi\delta_{r5}, & \dots \\ I_r(1) &= -\pi\delta_{r2}, & I_r(2) &= \pi\delta_{r4}, & I_r(4) &= -\pi\delta_{r8}, & \dots \end{aligned}$$

Eq. (3.7) yields the expressions (3.6). $\square$

Remark 3.1. To explain the choice of the new collocation points let us isolate one of the unknown integrals in the global relation (2.11) and denote by $H_p(k)$ the remaining terms,

$$\int_{-\pi}^{\pi} e^{i|k||h_p|s} f^{(p)}(s) ds = H_p(k), \quad k = |k|e^{i(\pi-\alpha_p)}, \quad p = 1, \dots, n. \quad (3.8)$$

Here we assume that the known linear function $f_*^{(p)}$ has been subtracted from $f^{(p)}$ as in (3.1) so that $f^{(p)}$ vanishes at $s = \pm\pi$. The complex conjugate, now, of Eq. (3.8) yields

$$\int_{-\pi}^{\pi} e^{-i|k||h_p|s} f^{(p)}(s) ds = \bar{H}_p(k), \quad k = |k|e^{i(\pi+\alpha_p)}, \quad p = 1, \dots, n. \quad (3.9)$$

The left-hand sides of Eqs. (3.8)–(3.9) involve both sines and cosines. However, since $f^{(p)}(s)$ vanishes at the end points we seek integral representations involving only sines. Thus, we let $s = 2t - \pi$ in Eqs. (3.8)–(3.9) and we find the following equations:

$$\begin{cases} 2 \int_0^{\pi} [\cos(2|k||h_p|t) + i \sin(2|k||h_p|t)] f^{(p)}(2t - \pi) dt = e^{i|k||h_p|\pi} H_p(-|k|e^{-i\alpha_p}), \\ 2 \int_0^{\pi} [\cos(2|k||h_p|t) - i \sin(2|k||h_p|t)] f^{(p)}(2t - \pi) dt = e^{-i|k||h_p|\pi} \bar{H}_p(-|k|e^{i\alpha_p}). \end{cases}$$

Subtracting these equations and letting

$$2|k||h_p| = l, \quad l = 1, 2, \dots$$

we find

$$4i \int_0^{\pi} \sin(lt) f^{(p)}(2t - \pi) dt = e^{il\pi/2} H_p\left(-\frac{l}{2h_p}\right) - e^{-il\pi/2} \bar{H}_p\left(-\frac{l}{2\bar{h}_p}\right), \quad (3.10)$$

for $l = 1, 2, \dots$, and $p = 1, \dots, n$. Taking $e^{il\pi/2}$ as a common factor and using the fact that $e^{-il\pi}$ equals -1 if l is odd and 1 if l is even, whereas $e^{il\pi/2}$ equals $(-1)^{l/2}$ if l is even Eq. (3.10) becomes

$$4 \int_0^{\pi} \sin(lt) f^{(p)}(2t - \pi) dt = \begin{cases} (-1)^{(l-1)/2} \left[H_p\left(-\frac{l}{2h_p}\right) + \bar{H}_p\left(-\frac{l}{2\bar{h}_p}\right) \right], & l \text{ odd}, \\ (-1)^{l/2} \left[H_p\left(-\frac{l}{2h_p}\right) - \bar{H}_p\left(-\frac{l}{2\bar{h}_p}\right) \right], & l \text{ even}. \end{cases}$$

4. Numerical results

Following directly Proposition 3.1, we construct the $Nn \times Nn$ linear system of equations, defined by relations (3.2)–(3.6), needed to determine the Nn unknown coefficients U_r^j in (3.1). Relation (3.2), which is the discrete analogue of the global relation (2.11), is a complex system of n equations. By choosing $\frac{N}{2}$ values of l for the real part of these equations and $\frac{N}{2}$ values of l for the imaginary part, we obtain a system of Nn real equations.

4.1. Comparison of SFFS and FFX methods for sine basis functions

The choice for the values of l specified in Proposition 3.1, in the case of sine basis functions, leads to a system whose coefficient matrix has the structure shown schematically in Fig. 2a, that is the $N \times N$ block diagonal submatrices are diagonal matrices. This should be compared to the block structure of the matrix obtained in [11] for the *physical space* case which is shown in Fig. 2b.

To illustrate the numerical implementation of the method, as well as to compare with the one presented in [11], we apply it to the Laplace equation for a variety of regular and irregular polygon domains with Dirichlet, Neumann and mixed ($\pi/3$) boundary conditions. In all cases the exact solution of the Laplace equation is given by

$$q(x, y) = \sinh(3x) \sin(3y), \quad (4.1)$$

and the corresponding boundary conditions as well as the corresponding functions $\{f^{(j)}(s)\}_{j=1}^n$ and $\{g^{(j)}(s)\}_{j=1}^n$ can be generated analytically. The relative error E_∞ , used to demonstrate the performance of the methods, is given by

$$E_\infty = \frac{\|f - f_N\|_\infty}{\|f\|_\infty}, \quad (4.2)$$

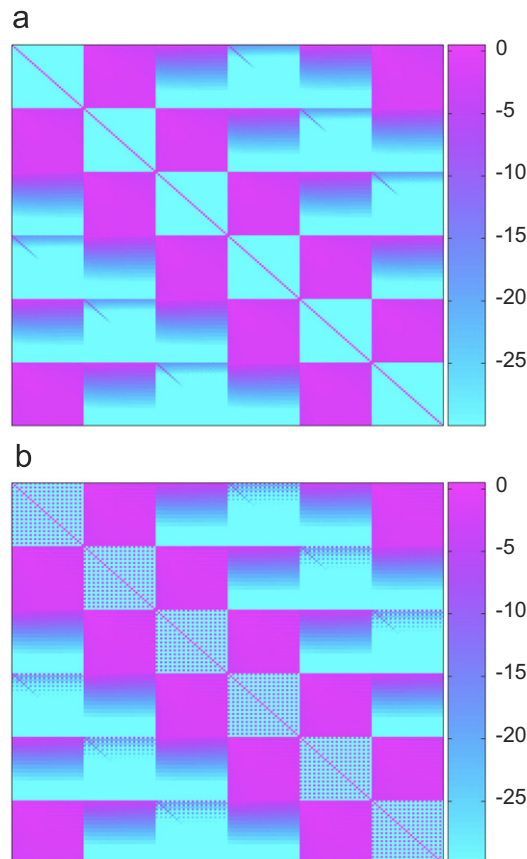


Fig. 2. The structures of the coefficient matrices produced (a) by Proposition 3.1 (SFFS method) and (b) in [11] (FFX method), for $n = 6$ and $N = 32$.

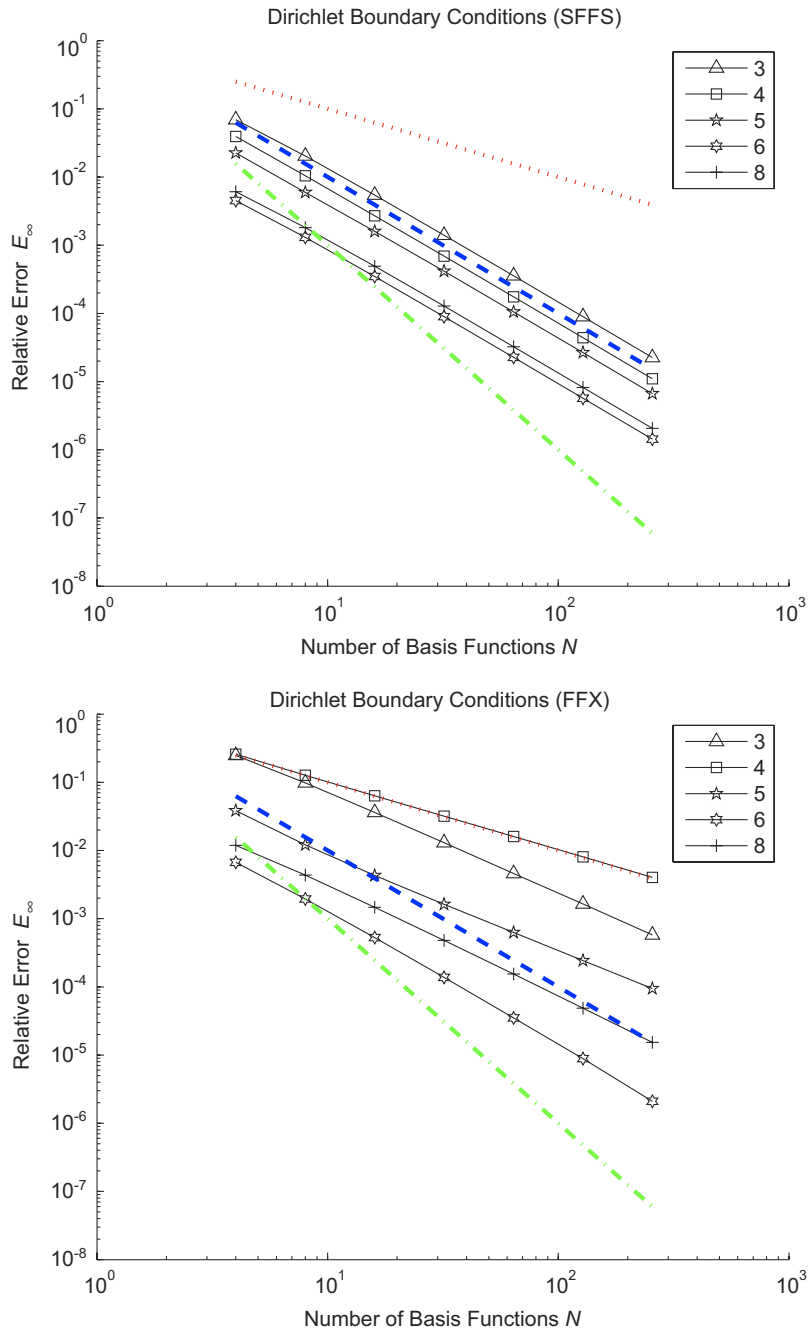


Fig. 3. Case of regular polygons and Dirichlet boundary conditions: E_∞ as a function of N for the method presented here (SFFS) and the method presented in [11] (FFX).

where

$$\|f\|_\infty = \max_{1 \leq j \leq n} \left\{ \max_{-\pi \leq s \leq \pi} |f^{(j)}(s)| \right\} \quad (4.3)$$

and

$$\|f - f_N\|_\infty = \max_{1 \leq j \leq n} \left\{ \max_{-\pi \leq s \leq \pi} |f^{(j)}(s) - f_N^{(j)}(s)| \right\}, \quad (4.4)$$

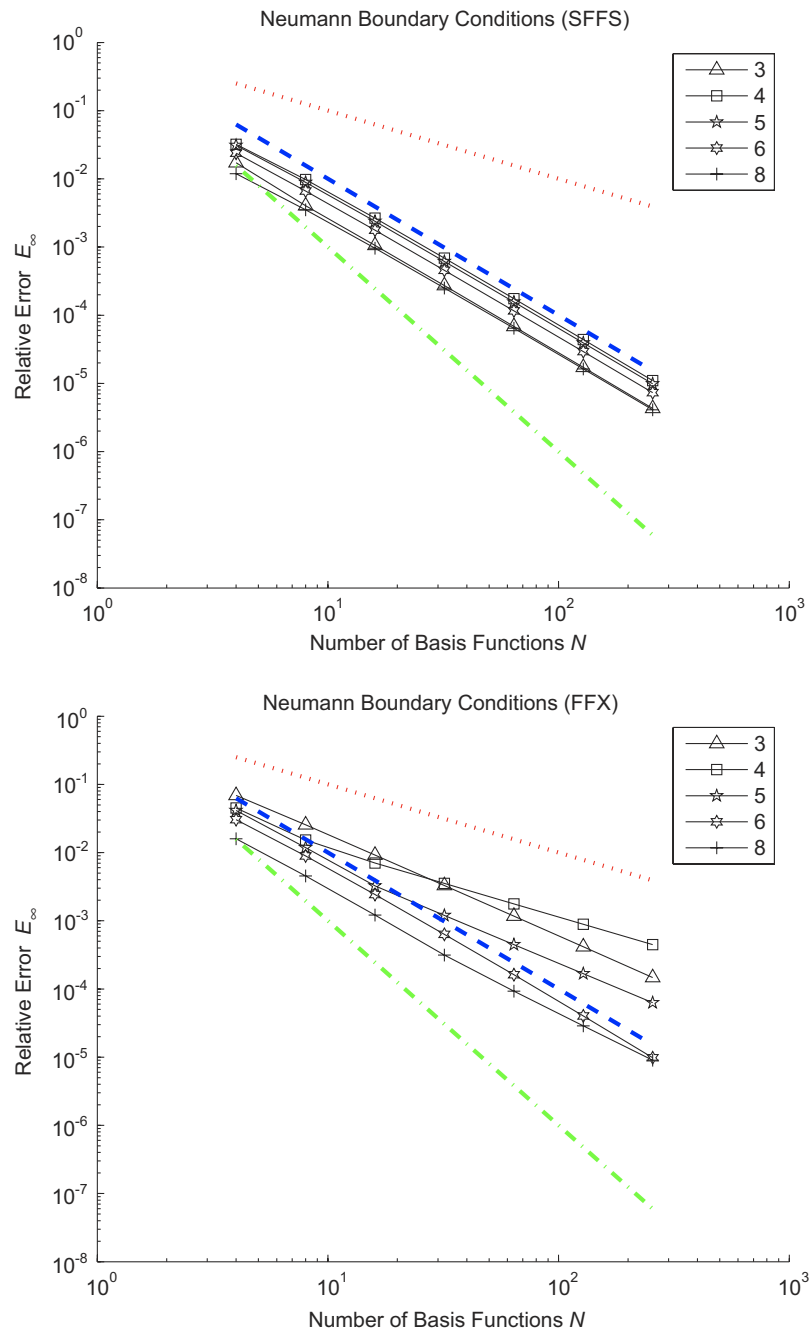


Fig. 4. Case of regular polygons and Neumann boundary conditions: E_∞ as a function of N for the method presented here (SFFS) and the method presented in [11] (FFX).

with $f_N^{(j)}$ as in (3.1), and the max over s is taken over a dense discretization of the interval $[-\pi, \pi]$. The number N of basis functions on each side remains the same in all experiments and takes successively the values $N = 4, 8, 16, 32, 64, 128, 256$. For the direct solution of the linear systems we have used the standard LAPACK routines, while for the computation of the right hand side vector we have used a routine (*dqawo*) from QUADPACK implementing the modified Clenshaw–Curtis technique.

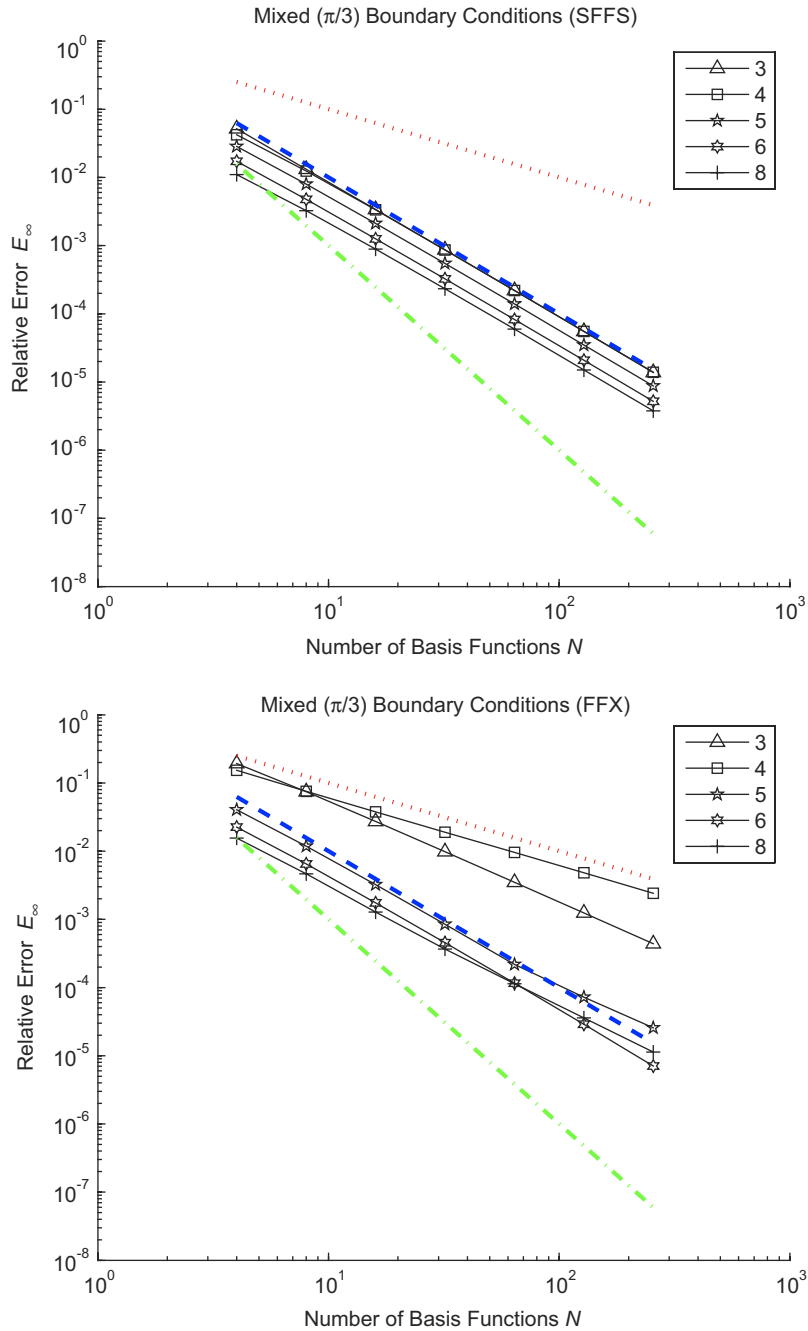


Fig. 5. Case of regular polygons and mixed boundary conditions: E_∞ as a function of N for the method presented here (SFFS) and the method presented in [11] (FFX).

Figs. 3–5 refer to the solution for *regular* polygons with Dirichlet ($\beta_j=0$), Neumann ($\beta_j=\pi/2$) and mixed ($\beta_j=\pi/3$) boundary conditions, respectively. The vertices z_j of the regular n -gons used in our experiments are lying on a circle centered at the origin with radius $\sqrt{2}$ and they are given by

$$z_j = \sqrt{2}e^{i[2(j-1)\pi/n-1/5]}, \quad j = 1, \dots, n.$$

Table 1
Regular triangle—Dirichlet BCs

N	SFFS method		FFX method	
	E_∞	O.o.C	E_∞	O.o.C
4	6.94E – 02	–	2.48E – 01	–
8	2.04E – 02	1.77	9.83E – 02	1.34
16	5.44E – 03	1.91	3.61E – 02	1.45
32	1.40E – 03	1.96	1.30E – 02	1.47
64	3.56E – 04	1.98	4.61E – 03	1.50
128	8.96E – 05	1.99	1.64E – 03	1.49
256	2.25E – 05	1.99	5.79E – 04	1.50

Table 2
Regular square—Neumann BCs

N	SFFS method		FFX method	
	E_∞	O.o.C	E_∞	O.o.C
4	3.20E – 02	–	4.51E – 02	–
8	9.78E – 03	1.71	1.53E – 02	1.56
16	2.66E – 03	1.88	7.04E – 03	1.12
32	6.88E – 04	1.95	3.53E – 03	1.00
64	1.75E – 04	1.98	1.77E – 03	1.00
128	4.40E – 05	1.99	8.90E – 04	0.99
256	1.10E – 05	2.00	4.46E – 04	1.00

Table 3
Regular pentagon—mixed BCs

N	SFFS method		FFX method	
	E_∞	O.o.C	E_∞	O.o.C
4	2.85E – 02	–	4.04E – 02	–
8	7.99E – 03	1.83	1.18E – 02	1.78
16	2.12E – 03	1.91	3.24E – 03	1.86
32	5.46E – 04	1.96	8.49E – 04	1.93
64	1.39E – 04	1.97	2.18E – 04	1.96
128	3.49E – 05	1.99	7.22E – 05	1.59
256	8.76E – 06	1.99	2.57E – 05	1.49

Table 4
Regular octagon—Dirichlet BCs

N	SFFS method		FFX method	
	E_∞	O.o.C	E_∞	O.o.C
4	6.09E – 03	–	1.19E – 02	–
8	1.81E – 03	1.75	4.37E – 03	1.45
16	4.93E – 04	1.88	1.47E – 03	1.57
32	1.28E – 04	1.95	4.79E – 04	1.62
64	3.25E – 05	1.98	1.54E – 04	1.64
128	8.21E – 06	1.98	4.88E – 05	1.66
256	2.06E – 06	1.99	1.54E – 05	1.66

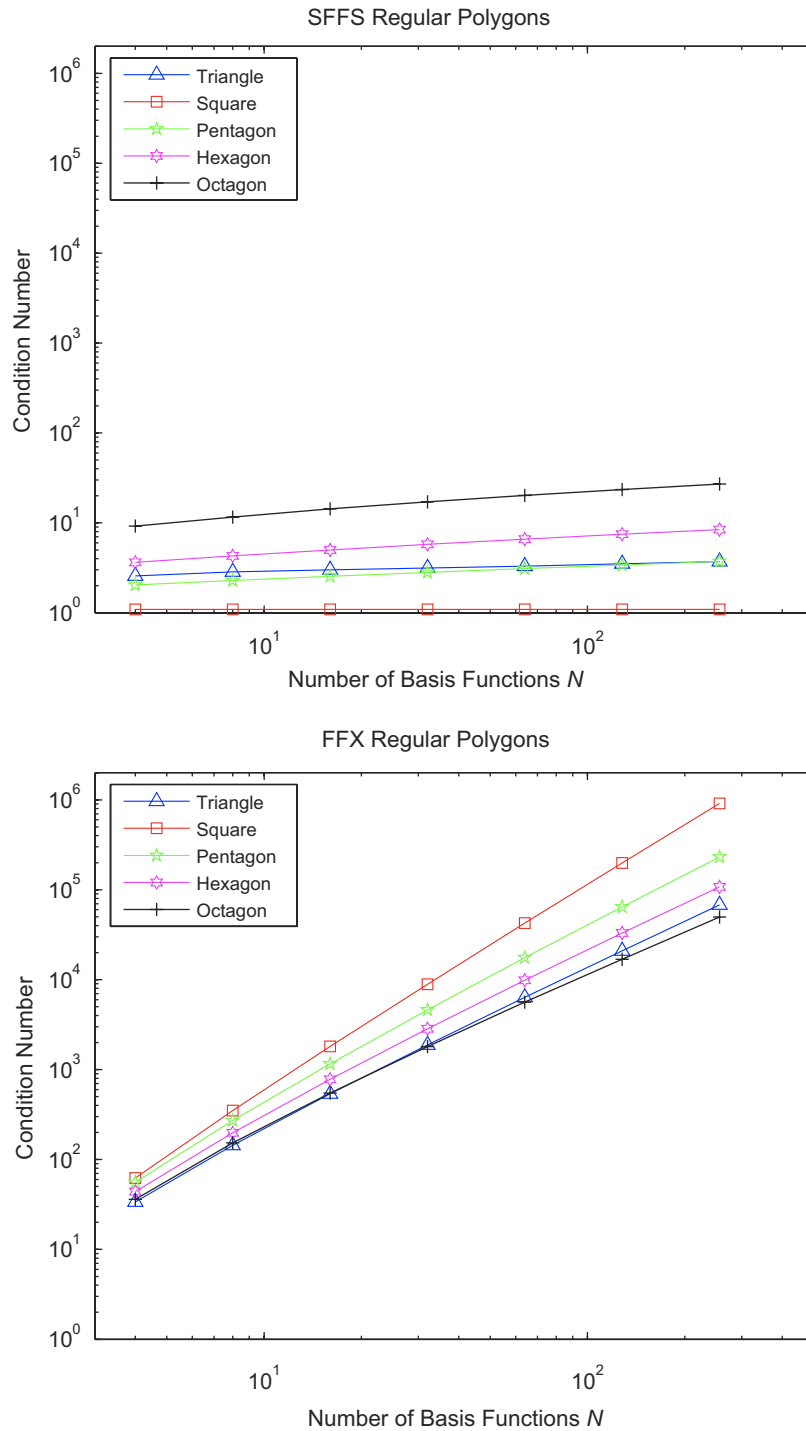


Fig. 6. Case of regular polygons: the condition numbers of the coefficient matrices as a function of N for the method presented here (SFFS) and the method presented in [11] (FFX).

The dotted, dashed and dash-dotted lines in the figures are the lines $1/N$, $1/N^2$ and $1/N^3$ indicating the slopes for convergence of order 1, 2 and 3, respectively. We note that in all cases both methods converge. However, the errors produced by the method described in Proposition 3.1 (SFFS) are significantly smaller.

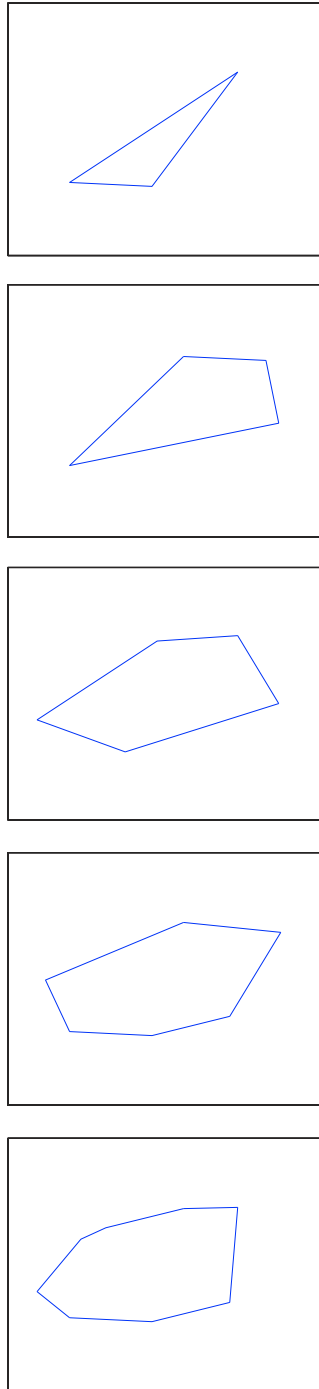


Fig. 7. The irregular polygons used in our experiments.

Furthermore, one may easily observe that the error lines for the SFFS method are parallel to the $1/N^2$ line indicating a quadratically convergent method with respect to the infinity error norm. To highlight this fact, the order of convergence (O.o.C) of both methods, for several cases of regular polygons, has been numerically estimated and included in Tables 1–4.

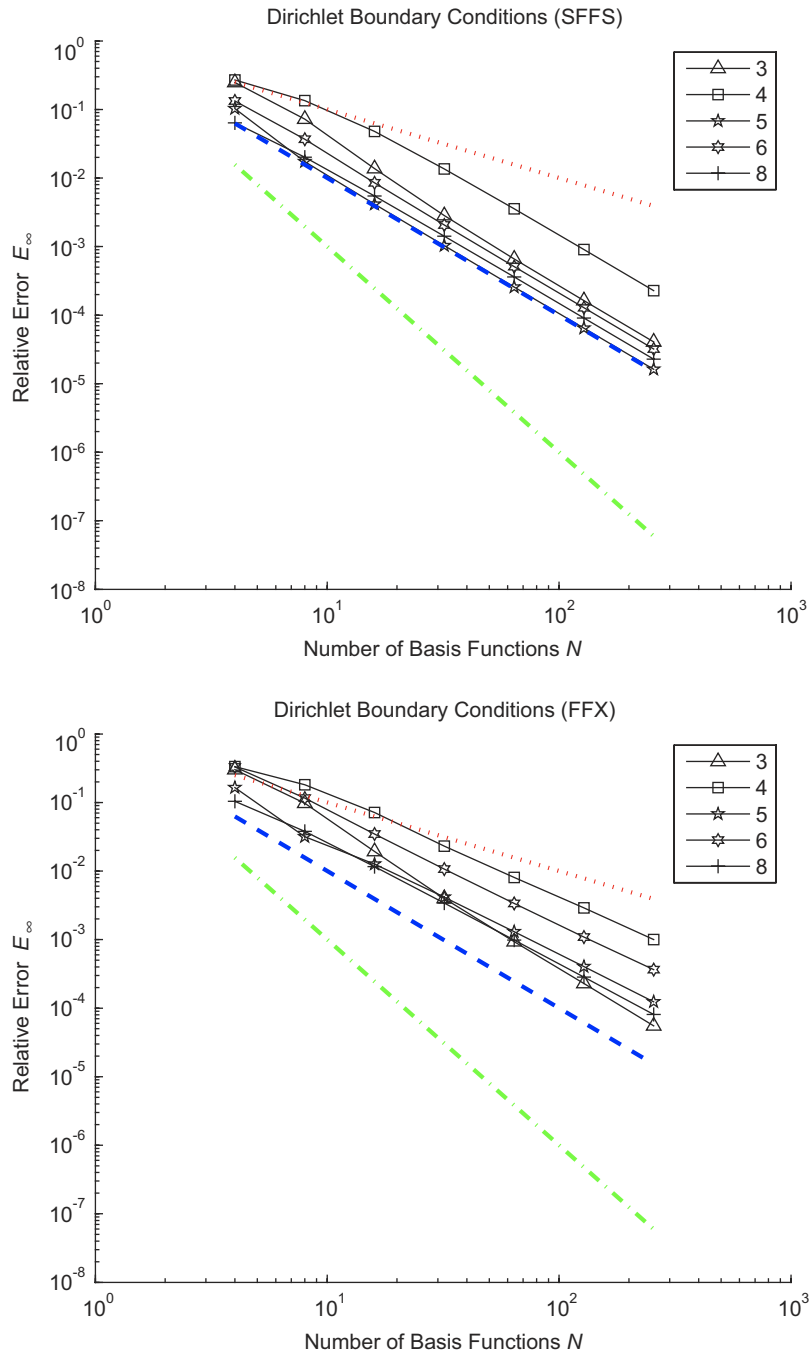


Fig. 8. Case of irregular polygons and Dirichlet boundary conditions: E_∞ as a function of N for the method presented here (SFFS) and the method presented in [11] (FFX).

As it pertains now to the *condition number* of the coefficient matrix associated with the SFFS method we point out the fact that it remains small and grows very slowly with N in contrary to the associated condition number with the FFX method (see Fig. 6).

In complete analogy to the above numerical treatment, we proceed now to the numerical study of the case of *irregular polygons*. The vertices z_j of the irregular n -gons (shown in Fig. 7) used in our experiments are lying on the ellipse

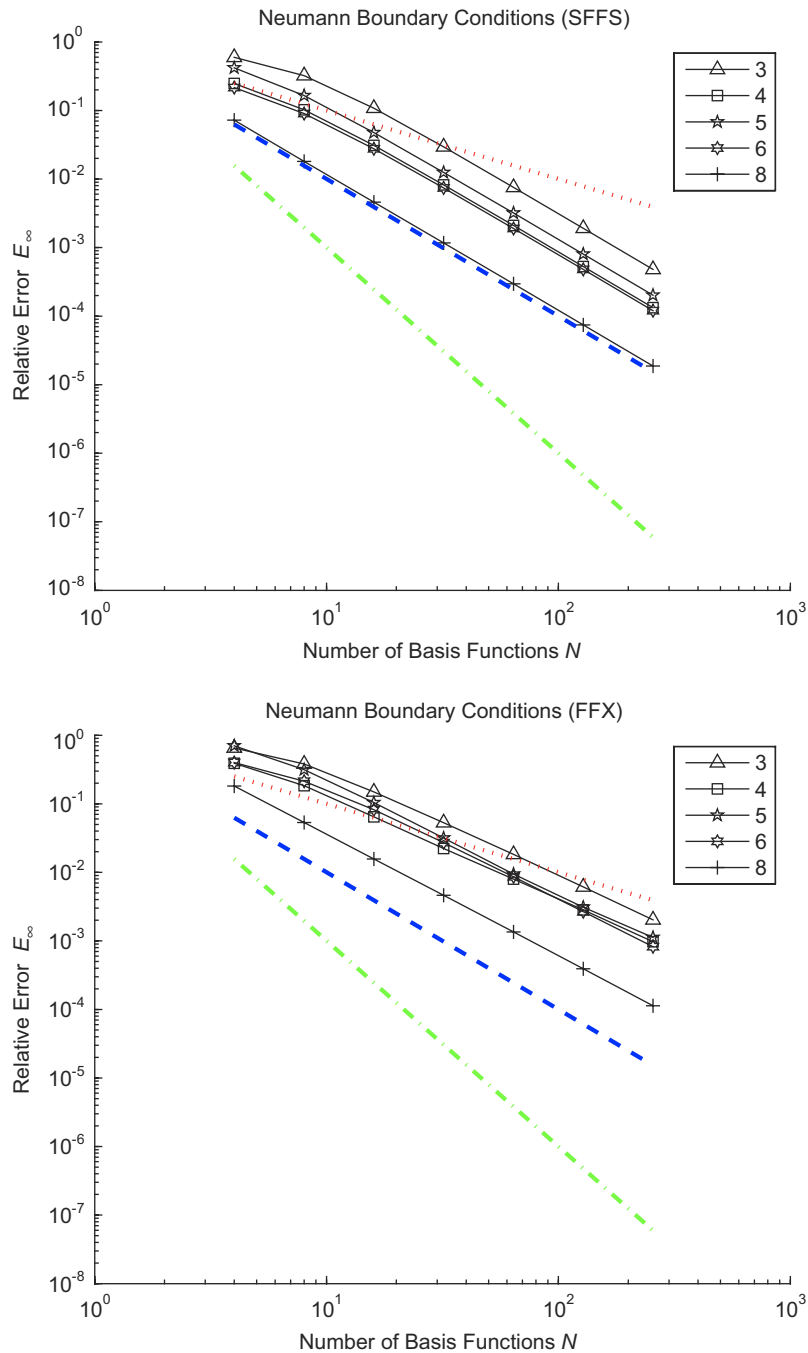


Fig. 9. Case of irregular polygons and Neumann boundary conditions: E_∞ as a function of N for the method presented here (SFBS) and the method presented in [11] (FFX).

$(x/5)^2 + (y/2)^2 = 1$ rotated by $\frac{1}{5}$. Their x -coordinates, before rotation, are given, in a counterclockwise motion, by

- *triangle*: $x = -4, -1, 3$;
- *quadrilateral*: $x = -4, 4, 4, 1$;
- *pentagon*: $x = -5, -2, 4, 3, 0$;

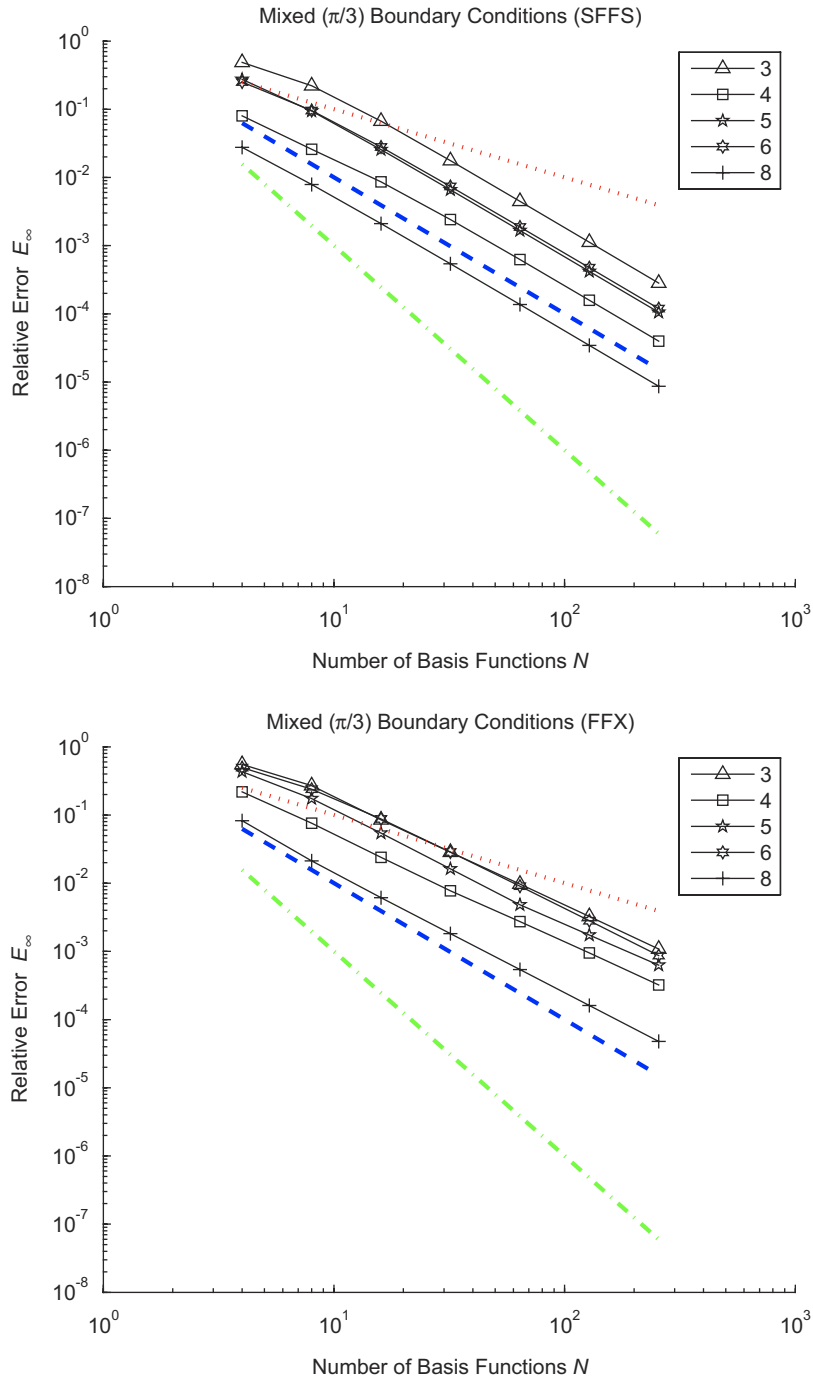


Fig. 10. Case of irregular polygons and mixed boundary conditions: E_∞ as a function of N for the method presented here (SFFS) and the method presented in [11] (FFX).

- *hexagon*: $x = -4, -1, 2, 4.5, 1, -4.5$;
- *octagon*: $x = -5, -4, -1, 2, 3, 1, -2, -3$.

Figs. 8–10 depict the relative error E_∞ as a function of N of both SFFS and FFX methods when they apply to the solution of the discrete global relation for the case of irregular polygons with Dirichlet, Neumann and mixed ($\beta_j = \pi/3$)

Table 5
Irregular quadrilateral—Dirichlet BCs

N	SFFS method		FFX method	
	E_∞	O.o.C	E_∞	O.o.C
4	2.70E – 01	–	3.33E – 01	–
8	1.35E – 01	1.00	1.81E – 01	0.88
16	4.79E – 02	1.49	7.13E – 02	1.34
32	1.36E – 02	1.82	2.31E – 02	1.63
64	3.56E – 03	1.93	8.05E – 03	1.52
128	9.05E – 04	1.98	2.90E – 03	1.47
256	2.28E – 04	1.99	1.00E – 03	1.54

Table 6
Irregular Pentagon—Neumann BCs

N	SFFS method		FFX method	
	E_∞	O.o.C	E_∞	O.o.C
4	4.19E – 01	–	7.01E – 01	–
8	1.65E – 01	1.34	3.15E – 01	1.15
16	4.77E – 02	1.79	1.04E – 01	1.60
32	1.25E – 02	1.93	3.16E – 02	1.72
64	3.19E – 03	1.97	9.40E – 03	1.75
128	8.03E – 04	1.99	3.09E – 03	1.61
256	2.02E – 04	1.99	1.12E – 03	1.46

Table 7
Irregular hexagon—mixed BCs

N	SFFS method		FFX method	
	E_∞	O.o.C	E_∞	O.o.C
4	2.51E – 01	–	4.96E – 01	–
8	9.56E – 02	1.39	2.39E – 01	1.05
16	2.80E – 02	1.77	8.74E – 02	1.45
32	7.37E – 03	1.93	2.86E – 02	1.61
64	1.88E – 03	1.97	9.04E – 03	1.66
128	4.74E – 04	1.99	2.81E – 03	1.69
256	1.19E – 04	1.99	8.69E – 04	1.69

Table 8
Irregular octagon—Dirichlet BCs

N	SFFS method		FFX method	
	E_∞	O.o.C	E_∞	O.o.C
4	6.38E – 02	–	1.04E – 01	–
8	2.01E – 02	1.67	3.79E – 02	1.46
16	5.48E – 03	1.87	1.16E – 02	1.71
32	1.42E – 03	1.95	3.41E – 03	1.77
64	3.60E – 04	1.98	9.85E – 04	1.79
128	9.07E – 05	1.99	2.83E – 04	1.80
256	2.28E – 05	1.99	8.07E – 05	1.81

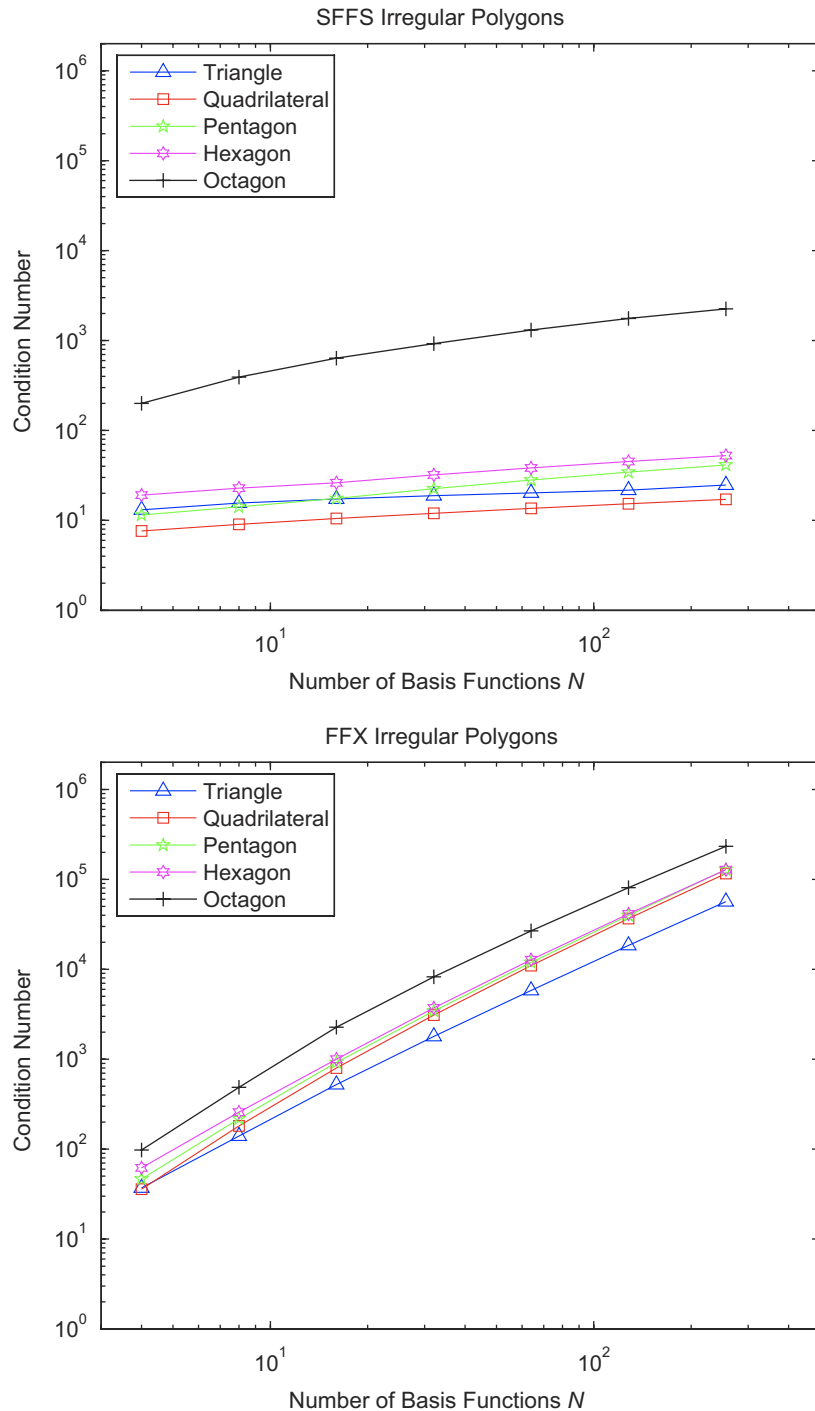


Fig. 11. Case of irregular polygons: the condition numbers of the coefficient matrices as a function of N for the method presented here (SFFS) and the method presented in [11] (FFX).

boundary conditions, respectively. Once more, both methods converge and the errors produced by the SFFS method are considerably smaller.

The quadratic convergence property of the SFFS method for the case of irregular polygons is being highlighted through Tables 5–8.

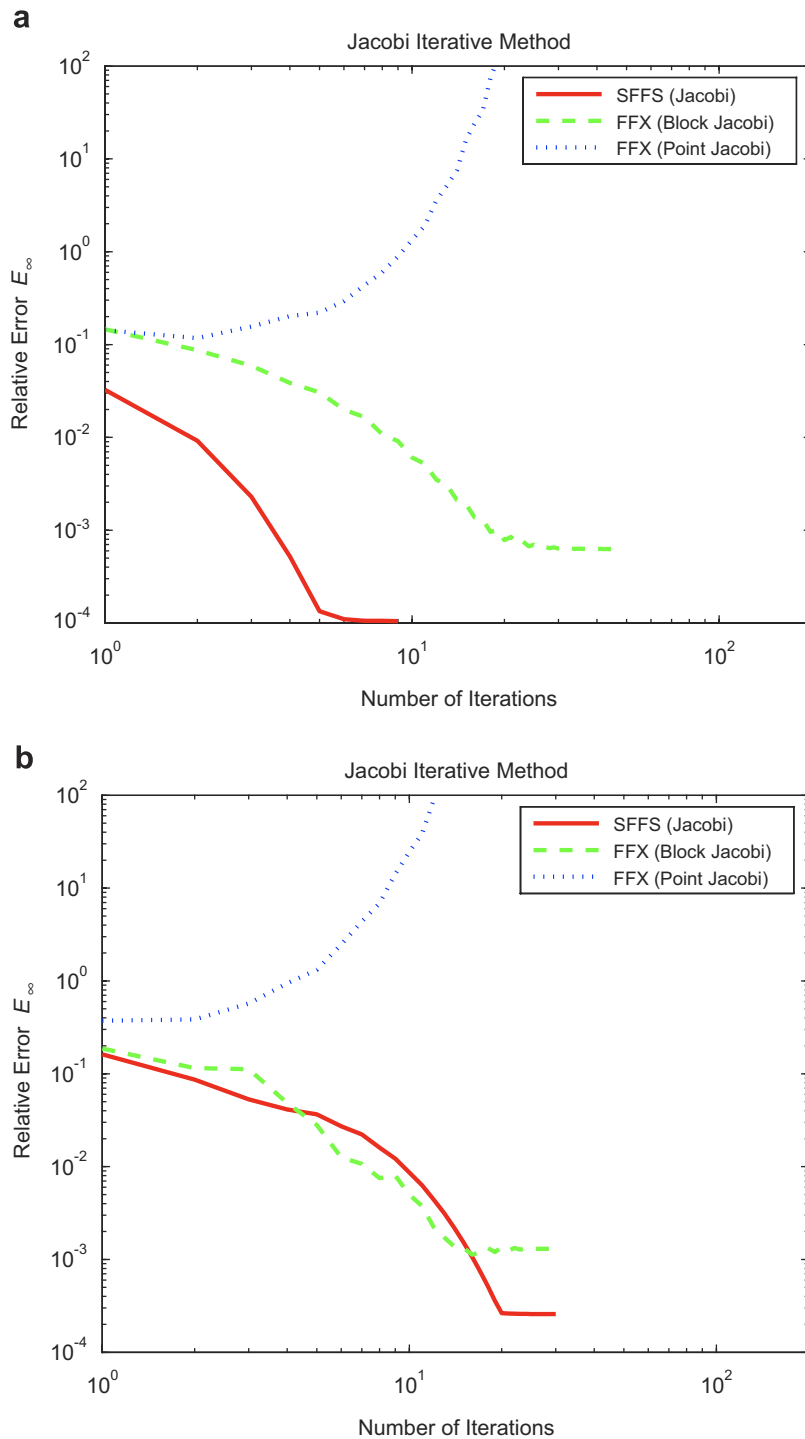


Fig. 12. The relative error E_∞ as a function of Jacobi iterations for (a) a regular and (b) an irregular pentagon ($N = 64$).

Finally, Fig. 11 refers to the behavior of the condition number of the coefficient matrices associated with the SFFS and FFX methods. Similar to the case of regular polygons, the condition number associated with the SFFS remains small and grows very slowly with N .

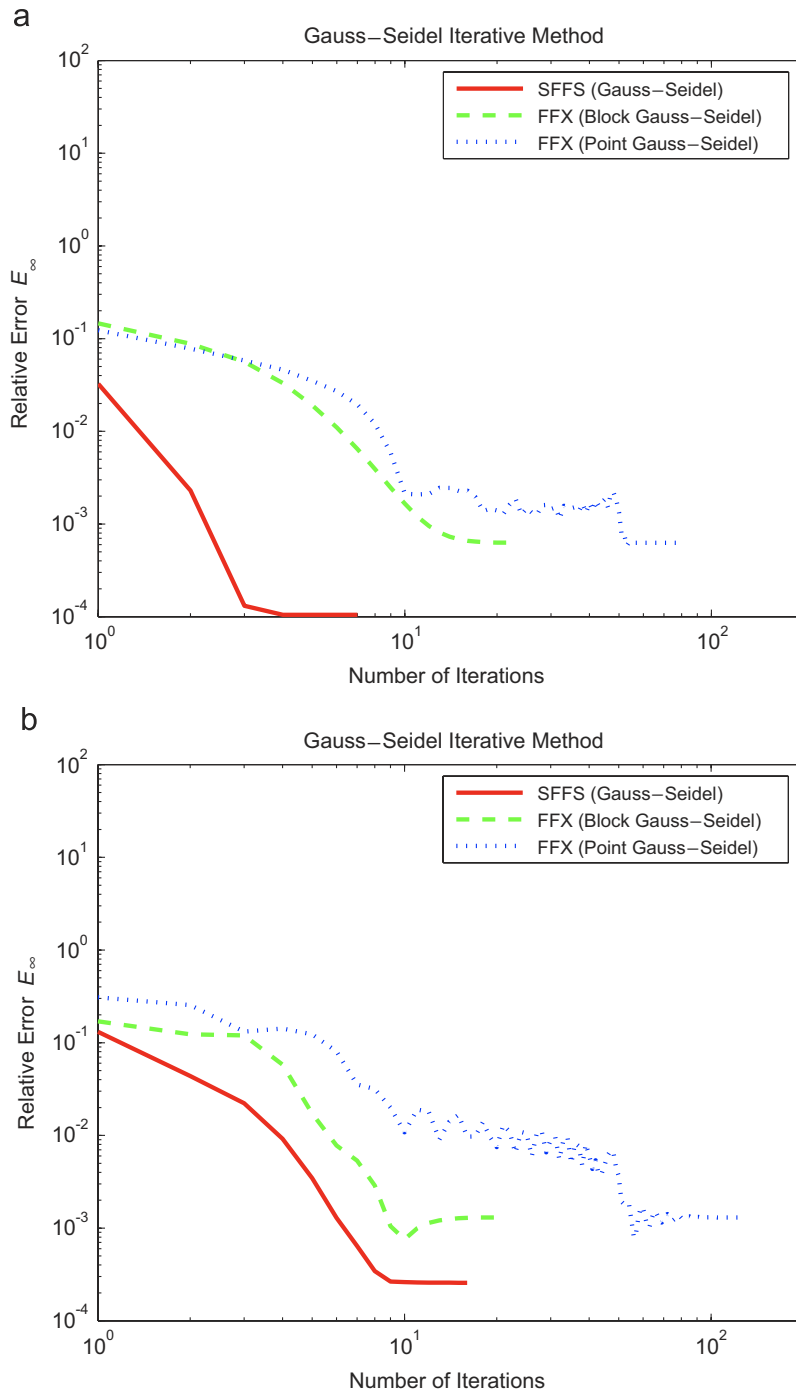


Fig. 13. The relative error E_∞ as a function of Gauss–Seidel iterations for (a) a regular and (b) an irregular pentagon ($N = 64$).

4.2. Performance of iterative solution methods for sine basis functions

Here we present some preliminary results for the performance of the iterative methods employed for the solution of the linear systems produced by the two methods (SFFS and FFX). A complete study for the behavior of iterative methods is outside the scope of the present work and will be presented elsewhere. Here we consider the classical *Jacobi*

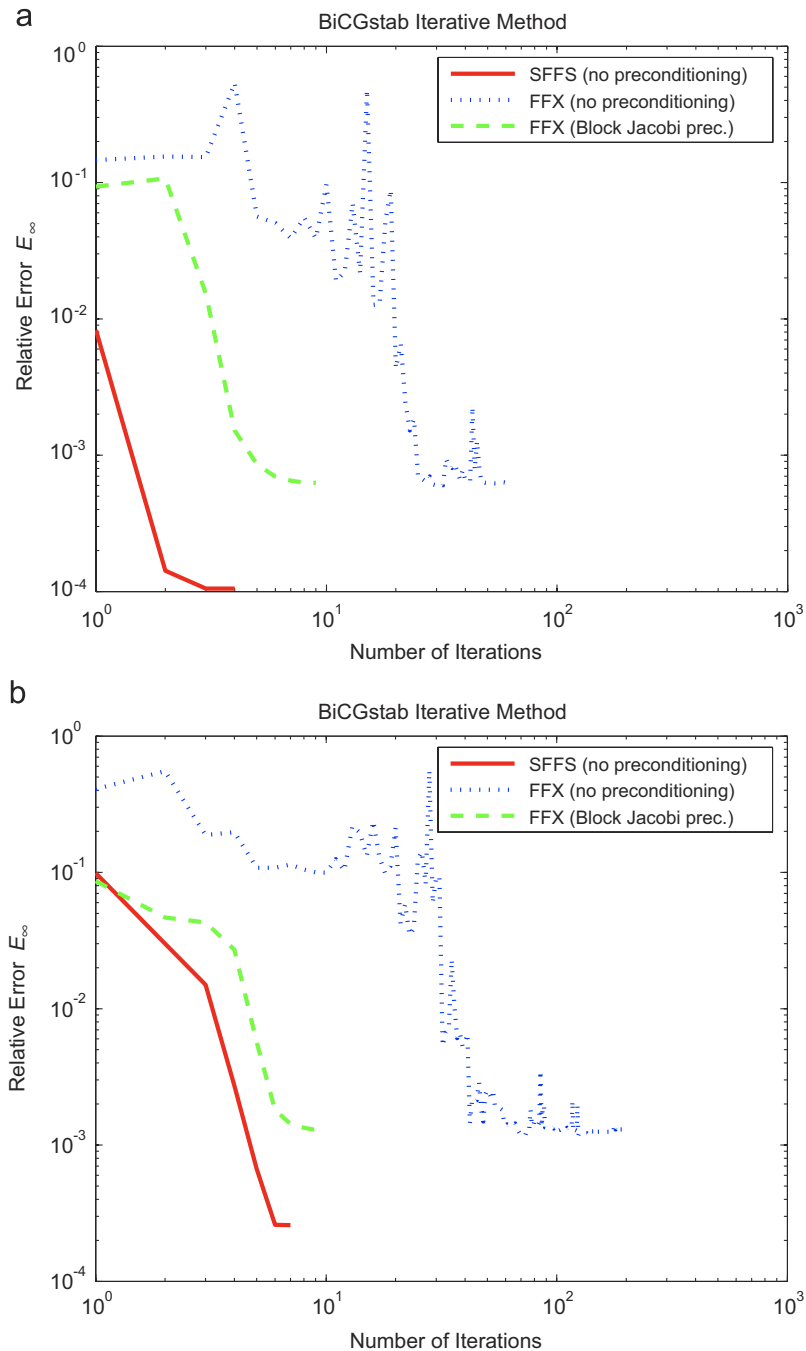


Fig. 14. The relative error E_∞ as a function of BiCGstab iterations for (a) a regular and (b) an irregular pentagon ($N = 64$).

and Gauss–Seidel as well as the *Bi-Conjugate Gradient stabilized* (BiCGstab) iterative methods. The results depicted in Figs. 12–14 refer to the reduction of the relative error E_∞ with respect to the number of iterations for the regular and irregular pentagons described earlier with the number of basis functions per side fixed at $N = 64$. To improve the convergence properties of the *point* Jacobi and Gauss–Seidel iterative methods for the linear system produced by the method in [11], we also considered their *block* analogs (the block size is fixed to N). For the same reason, together with the *unpreconditioned* BiCGstab we have also considered its *block Jacobi preconditioned* analog. In these cases, the computational cost increases due to the factorization involved. The maximum number of iterations allowed for all

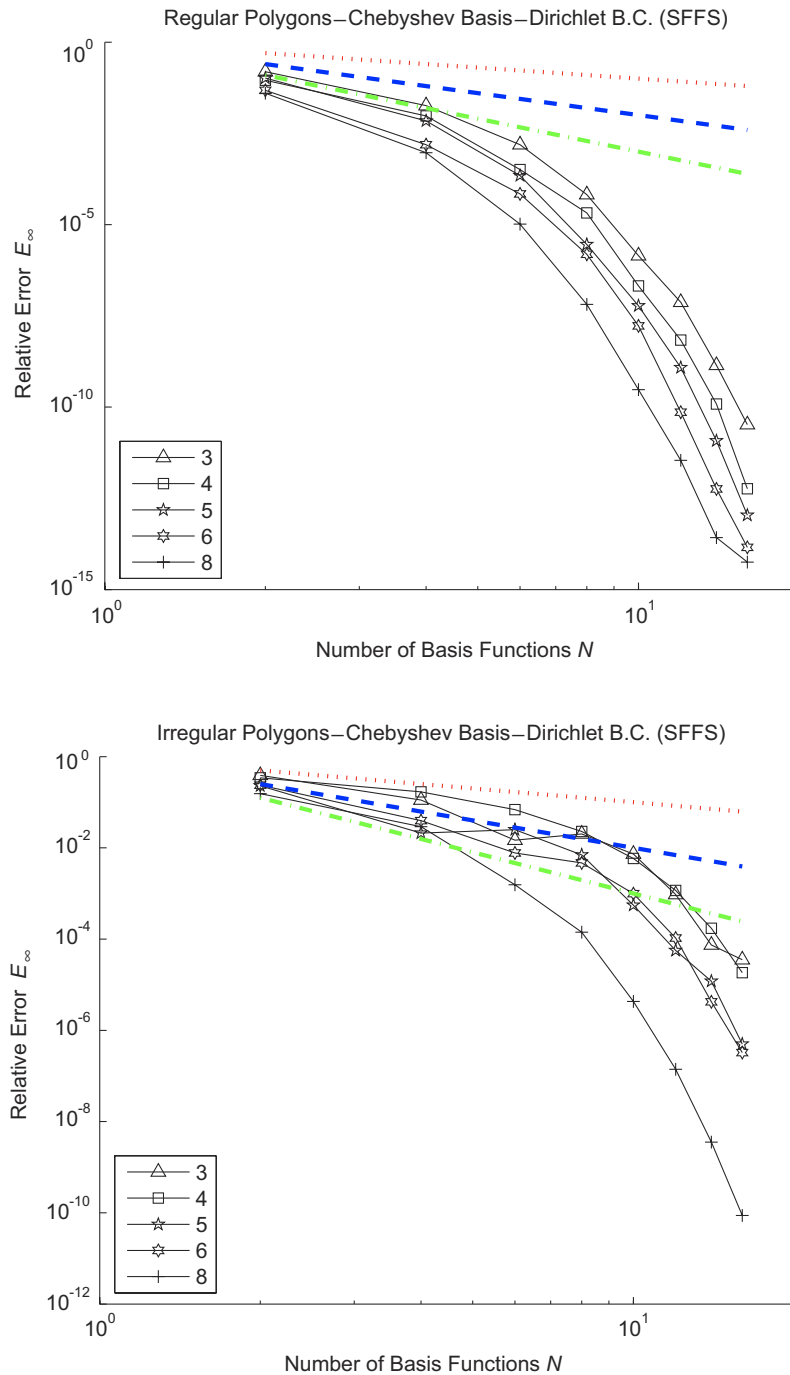


Fig. 15. E_∞ as a function of N for the SFFS method with Chebyshev basis functions for regular polygons (top panel) and irregular polygons (bottom panel) with Dirichlet boundary conditions.

methods to perform is set to 200 and the zero iterate $U^{(0)}$ is set to be equal to the right-hand side vector. The stopping criterion used for the Jacobi and Gauss–Seidel methods is

$$E_\infty^{(1)} = \frac{\|U^{(m+1)} - U^{(m)}\|_\infty}{\|U^{(m+1)}\|_\infty} \leq 10^{-6},$$

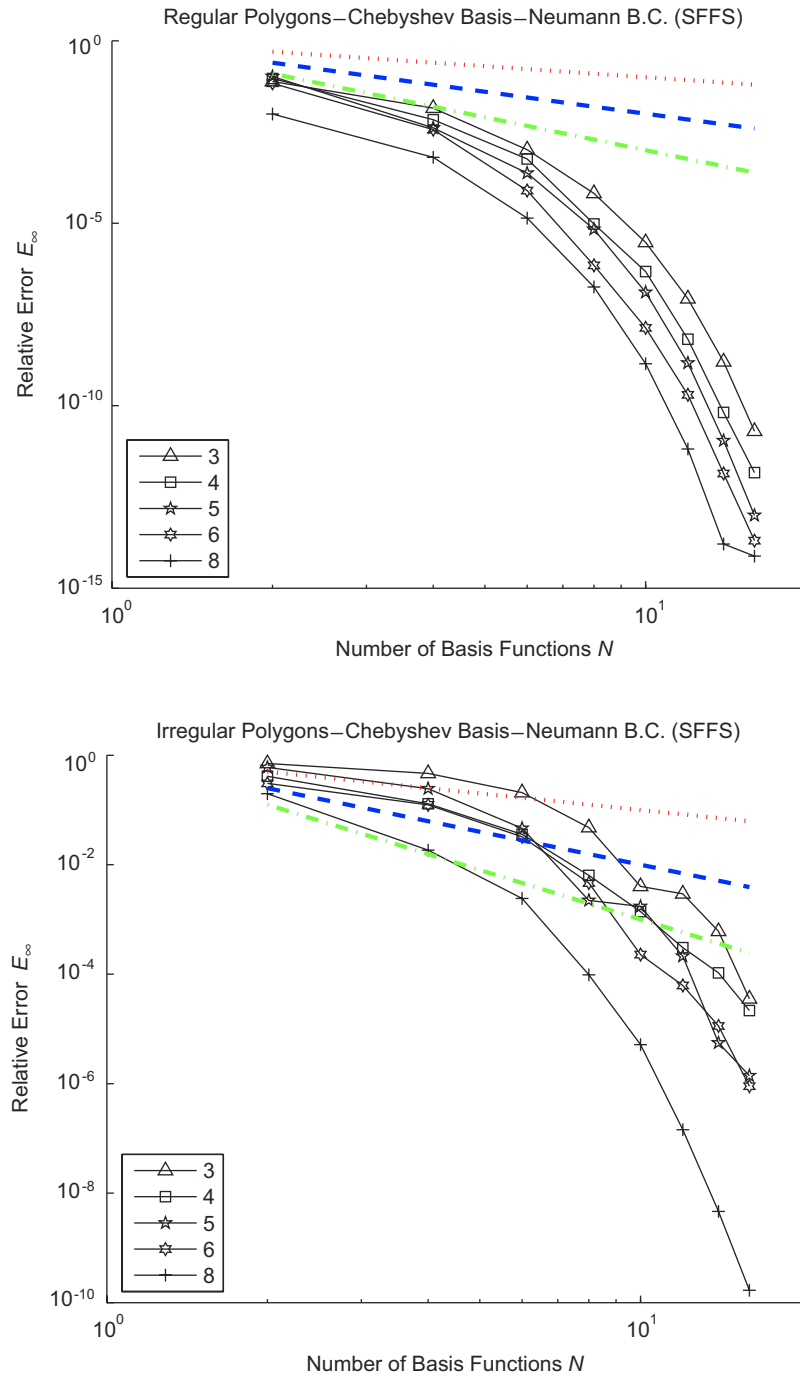


Fig. 16. E_∞ as a function of N for the SFSS method with Chebyshev basis functions for regular polygons (top panel) and irregular polygons (bottom panel) with Neumann boundary conditions.

while for the BiCGstab method is

$$E_\infty^{(2)} = \frac{\|b - AU^{(m)}\|_\infty}{\|b\|_\infty} \leq 10^{-6},$$

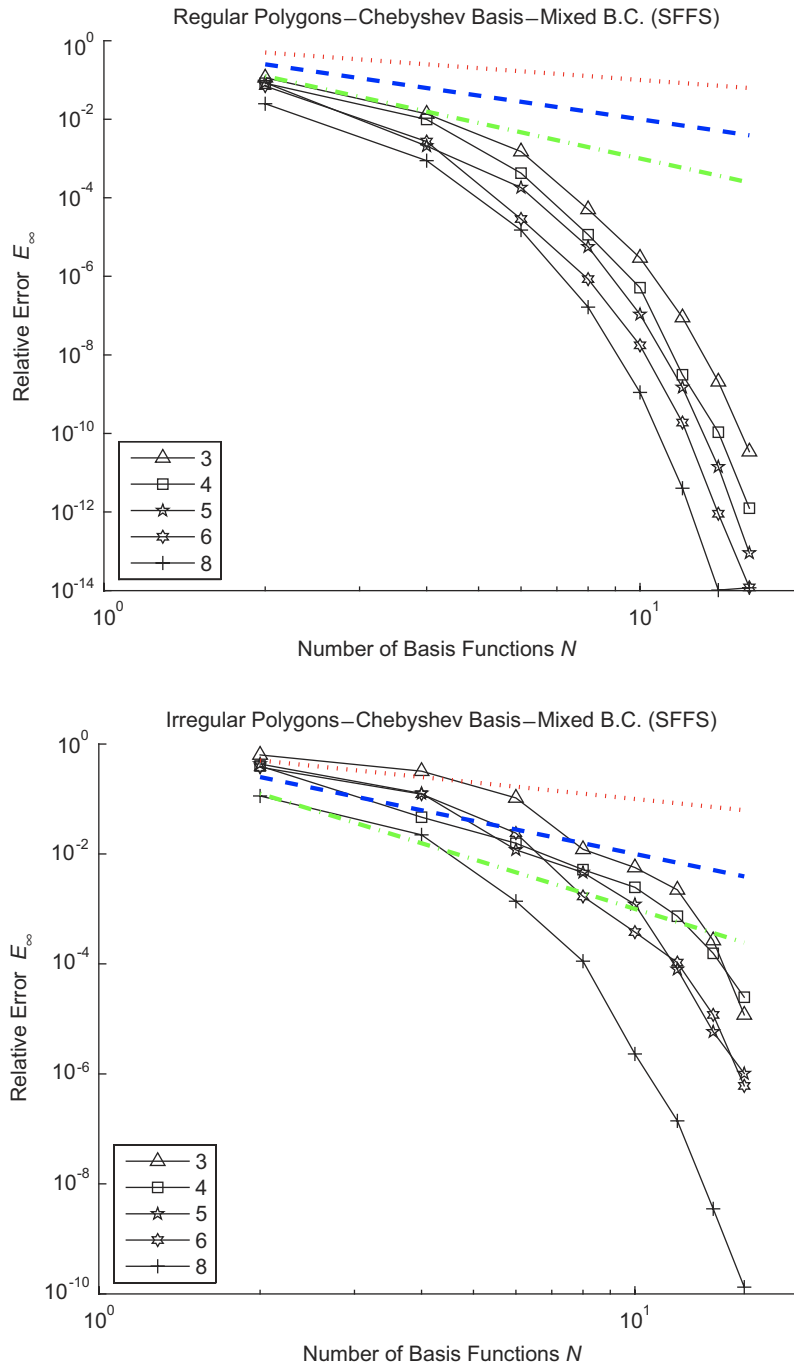


Fig. 17. E_∞ as a function of N for the SFFS method with Chebyshev basis functions for regular polygons (top panel) and irregular polygons (bottom panel) with mixed ($\pi/3$) boundary conditions.

where A , b and $U^{(m)}$ denote the coefficient matrix (with the appropriate modifications for the preconditioned cases), the right-hand side vector and the m th iterate, respectively.

Figs. 12–14 indicate the following:

- All iterative methods for the SFFS-algorithm converge rapidly in a few iterations. Specifically, for the case of the regular pentagon the Jacobi, Gauss–Seidel and BiCGstab iterative methods converge in 9, 7 and 4 iterations,

respectively. For the case of the irregular pentagon the corresponding numbers of iterations are 30, 16 and 7, respectively.

- The same iterative methods for the FFX-algorithm appear to have weak convergent properties. Namely, the Jacobi method diverges and the unpreconditioned BiCGstab method fails to reduce the relative error $E_\infty^{(2)}$ to a value less than 10^{-6} after 200 iterations for both cases of regular and irregular pentagons. The Gauss–Seidel method converges in 69 and 125 iterations for the case of regular and irregular pentagons, respectively. The situation improves by using the block Jacobi and Gauss–Seidel as well as the block Jacobi preconditioned BiCGstab methods, but this leads to extra computational cost due to the direct factorization involved. In these cases, the iterations needed for the convergence of the block-Jacobi, block-Gauss–Seidel and preconditioned-BiCGstab methods, respectively, are 46, 24 and 10 for the regular pentagon and 34, 21 and 10 for the irregular pentagon.

4.3. Higher-order convergence

The sine basis functions $\varphi_r(s)$ used in the approximation (3.1) in both the SFFS and FFX methods are natural in that they are closely connected to the underlying Fourier transform [cf. (3.4)]. Indeed, this is the reason that the diagonal blocks in the matrix are diagonal. However, from the point of view of approximation theory, these functions are not optimal. Thus, we can ask: Can the convergence be improved by using other basis functions? Here we explore only one choice, namely, Chebyshev polynomials. Since the endpoint values of the unknown functions $f^{(j)}$ have been removed (see Remark 2.3), the basis functions should vanish at the endpoints $s = \pm\pi$. To obtain reasonable conditioning, we construct polynomial basis functions from the Chebyshev polynomials $T_n(x) = \cos(n\cos^{-1}(x))$, defining $\varphi_r(s)$ for $r = 1, \dots, N$ by

$$\varphi_r(s) = \begin{cases} T_{r+1}\left(\frac{s}{\pi}\right) - T_0\left(\frac{s}{\pi}\right), & r \text{ odd,} \\ T_{r+1}\left(\frac{s}{\pi}\right) - T_1\left(\frac{s}{\pi}\right), & r \text{ even.} \end{cases}$$

The Fourier transforms of these functions needed for (3.4) can be computed using the recursion relation

$$2T_n(x) = \frac{T'_{n+1}(x)}{n+1} - \frac{T'_{n-1}(x)}{n-1}, \quad n > 1.$$

Figs. 15–17 show the error E_∞ as a function of N for the SFFS method using this Chebyshev basis for the regular and irregular polygon cases of Section 4.1. In each case the convergence is faster than quadratic; for the regular polygons, it appears to be exponential. The convergence is slower for the irregular polygons; the eventual breakdown of convergence may be due to roundoff error. In each case the error for a given value of N is smaller than that obtained with the sine basis. While the diagonal blocks in the linear system are no longer diagonal, the matrices are still well-conditioned, with condition numbers (not shown) typically less than 10 until the convergence breaks down.

5. Conclusions and remarks

We have introduced a revised implementation of the *collocation* numerical method introduced in [11] for solving linear elliptic PDEs with constant coefficients on arbitrary polygonal domains. The new version is based on a different choice of “*collocation points*”, as well as different choice of basis functions and produces significantly better approximations, especially for regular polygons. This is numerically demonstrated in Section 4; however, a rigorous proof for the convergence remains open. The associated condition number for polygons with relatively few number of sides remains confined and smaller than the corresponding one in [11].

For the case of sine basis functions, the computational advantage of our method is based on the fact that the block diagonal submatrices of the coefficient matrix of the linear system obtained from the discrete analog of the *global relation* (cf. [5,6]) are strictly diagonal. Thus, by construction and without introducing any computational cost from direct factorization, the coefficient matrix is “*block-Jacobi preconditioned*”. This property, besides the fact that it can be used to reduce the computational cost of direct factorization methods, points directly to the efficient implementation of the classical *SOR-like* and *CG-like* iterative methods. The rapid convergence properties of the Jacobi, Gauss–Seidel and BiCGstab iterative methods are numerically demonstrated in the previous section for regular and irregular pentagons.

A complete numerical study for the behavior of iterative methods pertaining to our algorithm, as well as their rigorous convergence analysis, is in progress.

For the case of Chebyshev basis functions, the block diagonal submatrices are no longer diagonal matrices, but the condition number is still small and numerical experiments indicate higher order of convergence. In particular, for regular polygons, it appears that the method converges exponentially.

This new method has been implemented recently to the modified Helmholtz and to the Helmholtz equations in the interior of a convex polygon.

References

- [1] Y.A. Antipov, A.S. Fokas, A transform method for the modified Helmholtz equation in a semi-strip, *Math. Proc. Cambridge Philos. Soc.* 138 (2005) 339–365.
- [2] D. ben-Avraham, A.S. Fokas, The modified Helmholtz equation in a triangular domain and an application to diffusion-limited coalescence, *Phys. Rev. E* 64 (2001) 016114–016116.
- [3] D. Crowdy, A.S. Fokas, Explicit integral solutions for the plane elastostatic semi-strip, *Proc. Roy. Soc. London* 460 (2004) 1285–1309.
- [4] G. Dassios, A.S. Fokas, The basic elliptic equations in an equilateral triangle, *Proc. Roy. Soc. London Ser. A* 461 (2005) 2721–2748.
- [5] A.S. Fokas, A unified transform method for solving linear and certain nonlinear PDEs, *Proc. Roy. Soc. London Ser. A* 53 (1997) 1411–1443.
- [6] A.S. Fokas, Two-dimensional linear PDEs in a convex polygon, *Proc. Roy. Soc. London Ser. A* 457 (2001) 371–393.
- [7] A.S. Fokas, A new transform method for evolution PDEs, *IMA J. Appl. Math.* 67 (2002) 559.
- [8] A.S. Fokas, Integrable nonlinear evolution equations on the half-line, *Comm. Math. Phys.* 230 (2002) 1–39.
- [9] A.S. Fokas, Linearizable initial boundary value problems for the sine-Gordon equation on the half-line, *Nonlinearity* 17 (2004) 1521–1534.
- [10] A.S. Fokas, B. Pelloni, A transform method for evolution PDEs on the interval, *IMA J. Appl. Math.* (2005) 1–24.
- [11] S. Fulton, A.S. Fokas, C. Xenophontos, An analytical method for linear elliptic PDEs and its numerical implementation, *J. Comput. Appl. Math.* 167 (2004) 465–483.

Numerical Study of Iterative Methods for the Solution of the Dirichlet-Neumann Map for Linear Elliptic PDEs on Regular Polygon Domains

A. G. Sifalakis and E. P. Papadopoulou and Y. G. Saridakis

Abstract—A generalized Dirichlet to Neumann map is one of the main aspects characterizing a recently introduced method for analyzing linear elliptic PDEs, through which it became possible to couple known and unknown components of the solution on the boundary of the domain without solving on its interior. For its numerical solution, a well conditioned quadratically convergent sine-Collocation method was developed, which yielded a linear system of equations with the diagonal blocks of its associated coefficient matrix being *point diagonal*. This structural property, among others, initiated interest for the employment of iterative methods for its solution. In this work we present a conclusive numerical study for the behavior of classical (Jacobi and Gauss-Seidel) and Krylov subspace (GMRES and Bi-CGSTAB) iterative methods when they are applied for the solution of the Dirichlet to Neumann map associated with the Laplace's equation on regular polygons with the same boundary conditions on all edges.

Keywords—Elliptic PDEs, Dirichlet to Neumann Map, Global Relation, Collocation, Iterative Methods, Jacobi, Gauss-Seidel, GMRES, Bi-CGSTAB.

I. INTRODUCTION

A New approach for analyzing boundary value problems for linear and for integrable nonlinear PDEs in two dimensions was introduced in [1] and [4]. This approach characterizes a generalized Dirichlet to Neumann map through the solution of the so-called *global relation*, namely, an equation, valid for all values

Manuscript received May 10, 2007. This work was supported by the Greek Ministry of Education "Herakleitos" EPEAEK Grant which is partially supported by the EU

A. G. Sifalakis is with the Department of Sciences, Technical University of Crete, University Campus, 73100 Chania, Greece (e-mail: sifal@science.tuc.gr).

E. P. Papadopoulou is with the Department of Sciences, Technical University of Crete, University Campus, 73100 Chania, Greece (e-mail: elena@science.tuc.gr)

Y. G. Saridakis is with the Department of Sciences, Technical University of Crete, University Campus, 73100 Chania, Greece (phone: +30-28210-37740; fax: +30-28210-37842; e-mail: yiannis@science.tuc.gr).

of a complex parameter k , coupling specified known and unknown values of the solution and its derivatives on the boundary.

For a large class of boundary value problems, the global relation can be solved analytically (see e.g. [4]-[6]), and hence the generalized Dirichlet to Neumann map can be constructed in closed form. However, for general boundary value problems, the global relation must be solved numerically.

The implementation of the new method to the case of the Laplace equation in an arbitrary bounded convex polygon was presented in [2], where a Sine-Collocation type numerical method was also introduced. In [3], we introduced a new set of *Collocation points* which yielded a well-conditioned quadratically convergent Sine-Collocation method. Moreover, the diagonal blocks of the associated coefficient matrix of the corresponding linear system were diagonal matrices. That is, the collocation coefficient matrix is by construction *block-Jacobi preconditioned*, pointing directly to the efficient implementation of iterative techniques. The first indicative numerical results, included in [3], relevant to the behavior of certain iterative methods, provided further support to this argument.

In the work herein, we present a conclusive numerical study for the behavior of the classical Jacobi and Gauss-Seidel as well as the GMRES[7] and the Bi-CGSTAB[8] iterative techniques applied to the solution of the model Dirichlet to Neumann map related to the Laplace's equation on several regular polygons with the same boundary conditions on all edges. We concluded that, while for polygons with small number of edges all methods converge rapidly, the Bi-CGSTAB method is the method of choice as it converges always faster in all cases of polygons independently of the number of edges, the number of basis functions and the boundary conditions.

The paper is organized as follows: Section 2 outlines some of the analytical results of [2] and [3] together

with the construction of the related Collocation linear system. Section 3 presents the numerical study for the behavior of the iterative methods considered.

II. OVERVIEW

To fix notation (see [2], [3]), let us consider the complex form of Laplace's equation in the independent variable q

$$\frac{\partial^2 q}{\partial z \partial \bar{z}} = 0. \quad (1)$$

The above equation is equivalent to the equation

$$\frac{\partial}{\partial \bar{z}} \left(e^{-ikz} \frac{\partial q}{\partial z} \right) = 0 \quad (2)$$

for an arbitrary complex parameter k . Suppose, now, that the real-valued function $q(z, \bar{z})$ satisfies Laplace's equation (1) in a simply connected bounded domain D with boundary ∂D . Then, equation (2), combined with the complex form of Green's theorem implies the equation

$$\int_{\partial D} e^{-ikz} \frac{\partial q}{\partial z} dz = 0, \quad k \in \mathbb{C}. \quad (3)$$

Equation (3) is referred to as the *global relation* associated with Laplace's equation. If we now consider the case of D being a convex bounded polygon with vertices $z_1, z_2, \dots, z_n$ (modulo n) have been indexed counter-clockwise (see Fig. 1), then the global relation (3) becomes

$$\sum_{j=1}^n \int_{S_j} e^{-ikz} \frac{\partial q}{\partial z} dz = 0, \quad k \in \mathbb{C}, \quad (4)$$

where S_j denotes the side from z_j to z_{j+1} (not including the end points).

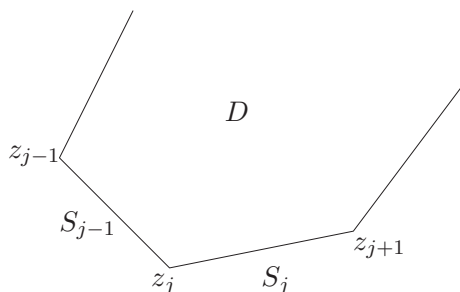


Fig. 1. Part of the bounded convex polygon with vertices z_j , sides S_j , and interior D .

Let, now, for $z \in S_j$, $1 \leq j \leq n$,

- $g^{(j)}$ denote the derivative of the solution in the direction making an angle β_j , $0 \leq \beta_j \leq \pi$ with the side S_j , i.e.

$$\cos(\beta_j) q_s^{(j)} + \sin(\beta_j) q_n^{(j)} = g^{(j)},$$

- $f^{(j)}$ denote the derivative of the solution in the direction normal to the above direction, i.e.

$$-\sin(\beta_j) q_s^{(j)} + \cos(\beta_j) q_n^{(j)} = f^{(j)},$$

where $q_s^{(j)}$ and $q_n^{(j)}$ denote the tangential and (outward) normal components of $\frac{\partial q}{\partial z}$ along the side S_j . For $z \in S_j$ and $-\pi < s < \pi$, with

$$z = \frac{1}{2}(z_j + z_{j+1}) + \frac{s}{2\pi}(z_{j+1} - z_j),$$

the *Generalized Dirichlet-Neumann map*, that is the relation between the sets $\{f^{(j)}(s)\}_{j=1}^n$ and $\{g^{(j)}(s)\}_{j=1}^n$, is characterized by the single equation

$$\sum_{j=1}^n |h_j| e^{i(\beta_j - km_j)} \cdot \int_{-\pi}^{\pi} e^{-ikh_j s} (f^{(j)} - ig^{(j)}) ds = 0, \quad k \in \mathbb{C} \quad (5)$$

where, for $j = 1, 2, \dots, n$,

$$h_j := \frac{1}{2\pi}(z_{j+1} - z_j), \quad (6)$$

$$m_j := \frac{1}{2}(z_{j+1} + z_j), \quad (7)$$

and $z_{n+1} = z_1$. Evaluating equation (5) on the following n -rays of the complex k -plane

$$k_p = -\frac{l}{h_p}, \quad l \in \mathbb{R}^+, \quad p = 1, \dots, n \quad (8)$$

and multiplying the resulting equations by $\exp[-i(\beta_p + lm_p/h_p)]/|h_p|$, equation (5) yields the following set of n equations (cf. [2]):

$$\sum_{j=1}^n \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-\frac{il}{h_p}(m_p - m_j)} \cdot \int_{-\pi}^{\pi} e^{il \frac{h_j}{h_p} s} (f^{(j)} - ig^{(j)}) ds = 0, \quad (9)$$

for $p = 1, \dots, n$.

Suppose that the set $\{g^{(j)}(s)\}_{j=1}^n$ is given, and that $\{f^{(j)}(s)\}_{j=1}^n$ is approximated by $\{f_N^{(j)}(s)\}_{j=1}^n$ where

$$f_N^{(j)}(s) = f_*^{(j)}(s) + \sum_{r=1}^N U_r^j \varphi_r(s) \quad (10)$$

with N being an even integer, and

$$2\pi f_*^{(j)}(s) = (s + \pi) f^{(j)}(\pi) - (s - \pi) f^{(j)}(-\pi). \quad (11)$$

Note that the values of $f^{(j)}(\pi)$ and $f^{(j)}(-\pi)$ can be computed by the continuity requirements at the vertices of the polygon. The set of functions $\{\varphi_r(s)\}_{r=1}^N$ are the *basis* functions which, for the purpose of this paper, are constructed from the Sine functions; namely

$$\varphi_r(s) = \sin(r \frac{\pi + s}{2}) \quad (12)$$

Then, the real coefficients U_r^j satisfy the $(Nn) \times (Nn)$ system of linear algebraic equations (cf. [3])

$$\sum_{j=1}^n \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-i \frac{l}{h_p} (m_p - m_j)} \cdot \sum_{r=1}^N U_r^j \int_{-\pi}^{\pi} e^{il \frac{h_j}{h_p} s} \varphi_r(s) ds = G_p(l) \quad (13)$$

where $G_p(l)$ denotes the known function

$$G_p(l) = i \sum_{j=1}^n \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-i \frac{l}{h_p} (m_p - m_j)} \cdot \int_{-\pi}^{\pi} U_r^j e^{il \frac{h_j}{h_p} s} \left(g^{(j)}(s) + i f_*^{(j)}(s) \right) ds \quad (14)$$

and l is chosen as follows:

- For the real part of equations (13) ,

$$l = \frac{1}{2}, \frac{3}{2}, \dots, \frac{N-1}{2}$$

- For the imaginary part of equations (13)

$$l = 1, 2, \dots, N/2.$$

It was the choice of the aforementioned collocation points (defined by the choice of the parameter l) that lead us to a well conditioned *point diagonal* collocation matrix, shown schematically in Fig. 2. The, numerically observed, convergence rate of the method is quadratic [3] and as the matrix is, by default, *Block Jacobi Preconditioned*, iterative methods appear to be a natural choice for the solution of the corresponding collocation system.

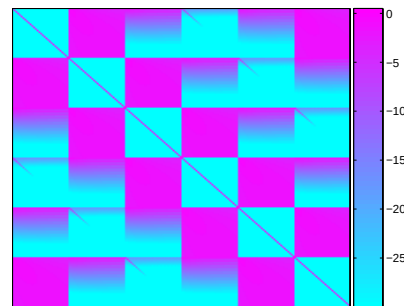


Fig. 2. The structure of the collocation coefficient matrix for $n = 6$ and $N = 32$

III. NUMERICAL STUDY OF ITERATIVE METHODS

In this section we have included results, from the numerical study conducted, pertaining to the behavior of iterative methods for the class of Laplace's problems having the same boundary conditions on all polygon's edges. For this class, following the analysis included in section II and particularly equation (13), which describes the linear system under consideration, one may easily verify that the Collocation coefficient matrix depends only on the choice and the number of the basis functions as well as on the number of the polygon's edges, while at the same time remains independent from the form of the boundary conditions and the form of the exact solution of Laplace's equation. Therefore, the numerical study focuses on the convergence behavior of iterative methods with respect to the number of vertices and the number of basis functions. However, for the sake of completeness, we have included experiments of all kind of boundary conditions for different polygon sizes. For simplicity and space economy, the results included here are related to a model Laplace equation with exact solution ([2], [3])

$$q(x, y) = \sinh(3x) \sin(3y). \quad (15)$$

We are pointing out once more, that all iterative methods behave similarly in all cases of problems of this class with Dirichlet ($\beta_j = 0$), Neumann ($\beta_j = \pi/2$) or Mixed ($\beta_j = \pi/3$) boundary conditions. The relative error E_∞ , used to demonstrate the convergence behavior of the iterative methods, is given by

$$E_\infty = \frac{\|f - f_N\|_\infty}{\|f\|_\infty}, \quad (16)$$

where

$$\|f\|_{\infty} = \max_{1 \leq j \leq n} \left\{ \max_{-\pi \leq s \leq \pi} |f^{(j)}(s)| \right\} \quad (17)$$

and

$$\|f - f_N\|_{\infty} = \max_{1 \leq j \leq n} \left\{ \max_{-\pi \leq s \leq \pi} |f^{(j)}(s) - f_N^{(j)}(s)| \right\}, \quad (18)$$

with $f_N^{(j)}$ as in (10), and the max over s is taken over a dense discretization of the interval $[-\pi, \pi]$. For the direct solution of the linear systems we have used the standard LAPACK routines, while for the computation of the right hand side vector we have used a routine (*dqawo*) from QUADPACK implementing the modified Clenshaw-Curtis technique. Apart from the classical Jacobi and Gauss-Seidel methods we have also considered the representative Krylov GMRES(10) and Bi-CGSTAB methods in two forms : (a) the unpreconditioned form or, equivalently, the block Jacobi preconditioned, due to the structure of the collocation matrix, and (b) the Symmetric Gauss-Seidel (SGS) preconditioned form, a two sided preconditioning method derived from the Symmetric SOR preconditioner for $\omega = 1$. The maximum number of iterations, allowed for all methods to perform, is set to 200 and the tolerance for the stopping criteria in all methods is set to 10^{-6} . The zero iterate $U^{(0)}$ is set to be equal to the right hand side vector. The results we have included refer to the representative cases of regular polygons with 8, 12, 19, 24, 31 and 36 edges. All polygons are constructed as in [3]. All experiments were conducted on a multiuser SUN V880 system using the Fortran-90 compiler.

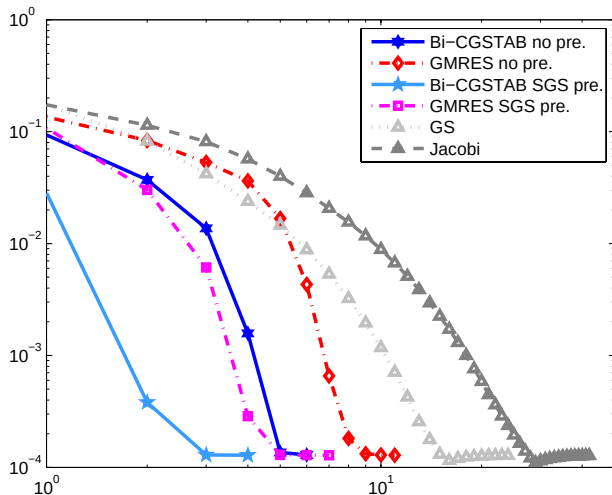


Fig. 1 : 8-gon/Dirichlet BC : Reduction of E_{∞} vs Iterations ($N = 32$)

TABLE I
REGULAR OCTAGON / DIRICHLET BC : NUMBER OF ITERATIONS

N	Jacobi	Gauss-Seidel	GMRES		Bi-CGSTAB	
			unpre	pre	unpre	pre
8	40	24	10	6	6	3
16	41	24	10	6	6	3
32	42	24	11	7	6	4
64	42	25	12	7	6	5

TABLE II
REGULAR OCTAGON / DIRICHLET BC : TIME (SEC $\times 10^{-2}$)

N	LU	Jacobi	Gauss-Seidel	GMRES		Bi-CGSTAB	
				unpre	pre	unpre	pre
8	0.6	2.1	1.0	0.4	0.6	0.5	0.5
16	4.5	4.7	2.6	1.3	1.7	1.4	1.7
32	40.6	13.9	8.5	5.8	9.3	4.85	12.4
64	414	105	34	31	30	24	39

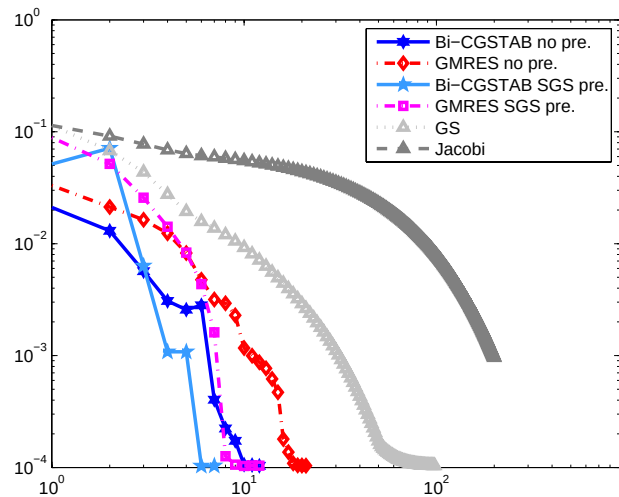


Fig. 2 : 12-gon/Dirichlet BC : Reduction of E_{∞} vs Iterations ($N = 32$)

TABLE III
REGULAR 12-GON / DIRICHLET BC : NUMBER OF ITERATIONS

N	Jacobi	Gauss-Seidel	GMRES		Bi-CGSTAB	
			unpre	pre	unpre	pre
8	200	87	19	12	10	7
16	200	94	21	12	10	7
32	200+	98	21	12	12	7
64	200+	100	21	13	11	7

TABLE IV
REGULAR 12-GON / DIRICHLET BC : TIME (SEC $\times 10^{-2}$)

N	LU	Jacobi	Gauss-Seidel	GMRES		Bi-CGSTAB	
				unpre	pre	unpre	pre
8	1.95	22.5	7.59	1.33	1.99	1.24	1.84
16	15.8	51.1	22.1	6.0	6.3	4.0	6.7
32	135	152	78	23	27	22	28
64	1450	1360	739	239	267	222	271

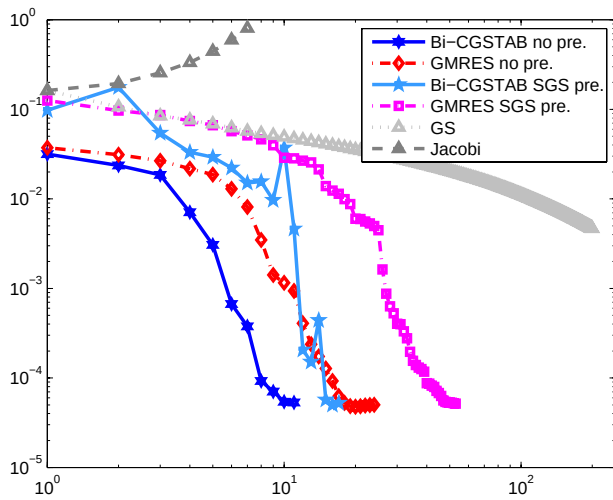
Fig. 3 : 19-gon/Neumann BC : Reduction of E_∞ vs Iterations ($N = 32$)

TABLE V

REGULAR 19-GON / NEUMANN BC : NUMBER OF ITERATIONS

N	Jacobi	Gauss-Seidel	GMRES		Bi-CGSTAB	
			unpre	pre	unpre	pre
8	div	200+	20	37	11	14
16	div	200+	23	53	11	16
32	div	200+	24	53	11	17
64	div	200+	24	52	11	19

TABLE VI

REGULAR 19-GON / NEUMANN BC : TIME (SEC)

N	LU	Jacobi	Gauss-Seidel	GMRES		Bi-CGSTAB	
				unpre	pre	unpre	pre
8	0.08	–	>0.44	0.03	0.12	0.03	0.08
16	0.66	–	>1.18	0.16	0.65	0.12	0.35
32	5.46	–	>4.33	0.76	2.86	0.67	1.68
64	159	–	>92.9	14.4	59.5	12.5	42.9

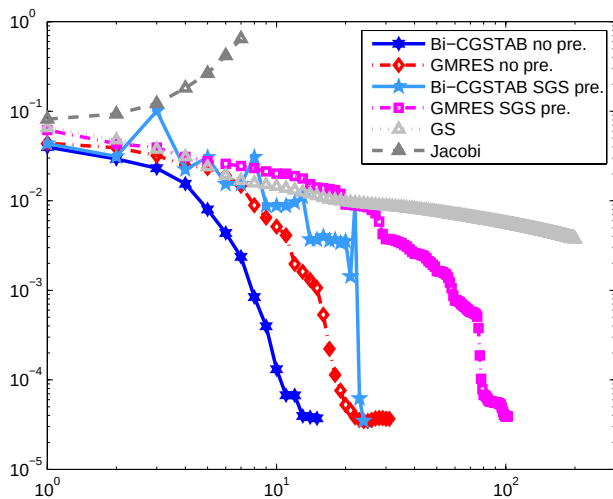
Fig. 4 : 24-gon/Dirichlet BC : Reduction of E_∞ vs Iterations ($N = 32$)

TABLE VII

REGULAR 24-GON / DIRICHLET BC : NUMBER OF ITERATIONS

N	Jacobi	Gauss-Seidel	GMRES		Bi-CGSTAB	
			unpre	pre	unpre	pre
8	div	200+	28	50	15	21
16	div	200+	29	78	14	25
32	div	200+	31	102	15	24
64	div	200+	32	114	16	25

TABLE VIII

REGULAR 24-GON / DIRICHLET BC : TIME (SEC)

N	LU	Jacobi	Gauss-Seidel	GMRES		Bi-CGSTAB	
				unpre	pre	unpre	pre
8	0.2	–	>0.72	0.07	0.25	0.06	0.19
16	1.37	–	>2.00	0.28	1.71	0.24	0.95
32	17.6	–	>16.8	3.67	21.6	3.61	10.7
64	323	–	>134	27.4	183	26.0	80.5

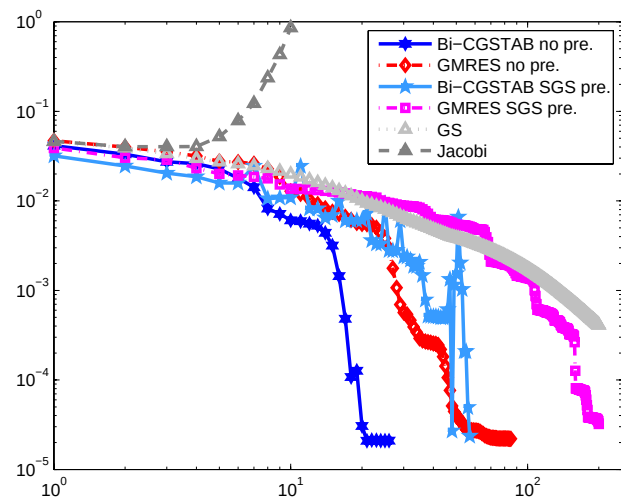
Fig. 5 : 31-gon / Mixed BC : Reduction of E_∞ vs Iterations ($N = 32$)

TABLE IX

REGULAR 31-GON / MIXED BC : NUMBER OF ITERATIONS

N	Jacobi	Gauss-Seidel	GMRES		Bi-CGSTAB	
			unpre	pre	unpre	pre
8	div	200+	91	103	23	44
16	div	200+	89	157	31	53
32	div	200+	85	200+	24	57
64	div	200+	87	200+	28	59

TABLE X

REGULAR 31-GON / MIXED BC : TIME (SEC)

N	LU	Jacobi	Gauss-Seidel	GMRES		Bi-CGSTAB	
				unpre	pre	unpre	pre
8	0.35	–	>1.17	0.36	0.83	0.14	0.64
16	2.96	–	>3.2	1.56	5.38	0.95	3.38
32	66.3	–	>61.1	33.0	>144	16.0	69.1
64	909	–	>271	144	>635	84.2	345

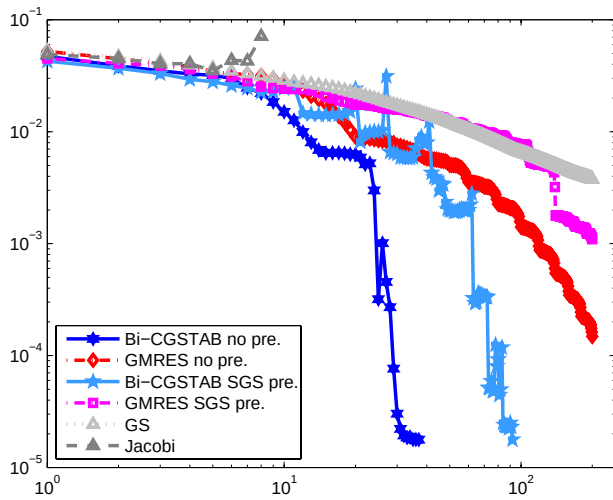
Fig. 6 : 36-gon/Dirichlet BC : Reduction of E_∞ vs Iterations ($N = 32$)

TABLE XI

REGULAR 36-GON / DIRICHLET BC : NUMBER OF ITERATIONS

N	Jacobi	Gauss-Seidel	GMRES		Bi-CGSTAB	
			unpre	pre	unpre	pre
8	div	200+	200+	200+	33	68
16	div	200+	200+	200+	35	76
32	div	200+	200+	200+	37	93
64	div	200+	200+	200+	37	95

TABLE XII

REGULAR 36-GON / DIRICHLET BC : TIME (SEC)

N	LU	Jacobi	Gauss-Seidel	GMRES		Bi-CGSTAB	
				unpre	pre	unpre	pre
8	0.6	—	>1.6	>1.1	>2.2	0.3	1.4
16	4.7	—	>4.7	>5.4	>10.4	1.6	7.0
32	133	—	>87	>105	>200	35	178
64	1300	—	>322	>389	>748	134	695

Figure 1 and Tables I and II summarize the results pertaining to the case of a Regular Octagon, a representative case for polygons with relatively *small* number of edges (triangles, squares, pentagons, hexagons, etc). Inspecting these data one may easily observe that :

- All methods converge smoothly in a few iterations
- The convergence rate of the Bi-CGSTAB method is faster than all other iterative methods
- Preconditioning improves the rate of convergence but increases the computational cost
- The convergence rate remains independent of N
- LU-factorization competes with the iterative methods only for small N .

As, now, the number of edges increases we observe that (see Fig. 2-6 and Tables III-XII):

- The Jacobi method diverges for polygons with medium and large number of edges
- The Gauss-Seidel method converges smoothly but very slowly
- The unpreconditioned versions of both Krylov methods converge smoothly
- The unpreconditioned version of the GMRES method fails to converge within 200 iterations for the case of the regular 36-gon
- The unpreconditioned Bi-CGSTAB converged in all experiments rapidly and its convergence rate remained faster than the unpreconditioned GMRES and the two classical iterative methods
- The SGS preconditioner for polygons with medium and large number of edges did not improve the convergence rate of both the Krylov methods and increased in all cases the computational cost
- The Krylov methods (except GMRES for the case of the 36-gon) are spectacularly more efficient than LU-factorization.

In conclusion, the unpreconditioned version of the Bi-CGSTAB is the most efficient and safe method to be employed for the solution of the linear system arising from the solution of the Dirichlet to Neumann map associated with the Laplace equation on regular polygons with the same Dirichlet, Neumann or Mixed Boundary Conditions on all edges. A spectral analysis for the theoretical justification of these results is under development.

- [1] A.S.Fokas, A unified transform method for solving linear and certain nonlinear PDEs, *Proc. R. Soc. London A* **53** (1997), 1411-1443.
- [2] S. Fulton, A.S. Fokas and C. Xenophontos, An Analytical Method for Linear Elliptic PDEs and its Numerical Implementation, *J. of CAM* **167** (2004), 465-483.
- [3] A. Sifalakis, A.S. Fokas, S. Fulton and Y.G. Saridakis, The Generalized Dirichlet-Neumann Map for Linear Elliptic PDEs and its Numerical Implementation, *J. of Comput. and Appl. Maths.* (in press)
- [4] A.S.Fokas, Two-dimensional linear PDEs in a convex polygon, *Proc. R. Soc. London A* **457** (2001), 371-393.
- [5] A.S. Fokas, A New Transform Method for Evolution PDEs, *IMA J. Appl. Math.* **67** (2002), 559.
- [6] G. Dassios and A.S. Fokas, The Basic Elliptic Equations in an Equilateral Triangle, *Proc. R. Soc. Lond. A* **461** (2005), 2721-2748.
- [7] Y. Saad and M. Schultz, GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems, *SIAM J. Sci. Statist. Comput.*, **7**, 1986, pp. 856-869.
- [8] H.A. Van Der Vorst, Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems, *SIAM J. Sci. Statist. Comput.*, **13**, 1992, pp. 631-644.

Direct and Iterative Solution of the Generalized Dirichlet-Neumann Map for Elliptic PDEs on Square Domains[★]

A.G.Sifalakis¹, S.R.Fulton²,
E.P. Papadopoulou¹ and Y.G.Saridakis^{1, *}

¹*Applied Mathematics and Computers Lab, Department of Sciences, Technical University of Crete, 73100 Chania, Greece*

²*Department of Mathematics and Computer Science, Clarkson University, Potsdam NY 13699-5815, USA*

Abstract

In this work we derive the structural properties of the Collocation coefficient matrix associated with the Dirichlet-Neumann map for Laplace's equation on a square domain. The analysis is independent of the choice of basis functions and includes the case of same type of boundary conditions on all sides, as well as the case of different boundary conditions on each side of the square domain. Taking advantage of said properties, we present efficient implementations of direct factorization and iterative methods, including classical SOR-type and Krylov subspace (Bi-CGSTAB and GMRES) methods appropriately preconditioned, for both Sine and Chebyshev basis functions. Numerical experimentation, to verify our results, is also included.

Key words : elliptic PDEs, Dirichlet-Neumann map, global relation, collocation, iterative methods, Jacobi, Gauss-Seidel, GMRES, Bi-CGSTAB

2000 MSC : 35J25, 65N35, 64N99, 65F05, 65F10

1 Introduction

Recently, Fokas[1,4] introduced a new unified approach for analyzing linear and integrable nonlinear PDEs. Central issue to this approach is a generalized Dirichlet to Neumann map, characterized through the solution of the so-called *global relation*, namely, an equation, valid for all values of a complex parameter k , coupling specified known and unknown values of the solution and its derivatives on the boundary. In

[★] This work was supported by the Greek Ministry's of Education *EPEAEK-Herakleitos* grant which is partially funded by the EU

^{*} Corresponding author. E-mail : yiannis@science.tuc.gr

particular, for the case of Laplace's equation, $q_{z\bar{z}} = 0$, in a convex bounded polygon D with vertices $z_1, z_2, \dots, z_n$ (modulo n) indexed counter-clockwise, the associated *Global Relation* takes the form (see also [2,3])

$$\sum_{j=1}^n \int_{S_j} e^{-ikz} q_z dz = 0, \quad k \in \mathbb{C}, \quad (1.1)$$

where S_j denotes the side from z_j to z_{j+1} (not including the end points). If, for $z \in S_j$, $1 \leq j \leq n$, we now let

- $g^{(j)}$ denote the derivative of the solution in the direction making an angle β_j , $0 \leq \beta_j \leq \pi$, with the side S_j , namely :

$$\cos(\beta_j) q_s^{(j)} + \sin(\beta_j) q_n^{(j)} = g^{(j)}, \quad (1.2)$$

- $f^{(j)}$ denote the derivative of the solution in the direction normal to the above direction, namely :

$$-\sin(\beta_j) q_s^{(j)} + \cos(\beta_j) q_n^{(j)} = f^{(j)}, \quad (1.3)$$

where $q_s^{(j)}$ and $q_n^{(j)}$ denote the tangential and (outward) normal components of q_z along the side S_j , then the *Generalized Dirichlet-Neumann map*, that is the relation between the sets $\{f^{(j)}(s)\}$ and $\{g^{(j)}(s)\}_{j=1}^n$, is characterized by the single equation

$$\sum_{j=1}^n |h_j| e^{i(\beta_j - km_j)} \int_{-\pi}^{\pi} e^{-ikh_j s} \left(f^{(j)} - ig^{(j)} \right) ds = 0, \quad k \in \mathbb{C} \quad (1.4)$$

where, for $j = 1, 2, \dots, n$, and $z_{n+1} = z_1$,

$$h_j := \frac{1}{2\pi} (z_{j+1} - z_j) \quad , \quad m_j := \frac{1}{2} (z_{j+1} + z_j) \quad , \quad s := \frac{z - m_j}{h_j} . \quad (1.5)$$

For the numerical solution of the Generalized Dirichlet-Neumann map in (1.4), a Collocation-type method has been developed (see [2] and [3]) : Suppose that the set $\{g^{(j)}(s)\}_{j=1}^n$ is given through the boundary conditions, and that $\{f^{(j)}(s)\}_{j=1}^n$ is approximated by $\{f_N^{(j)}(s)\}_{j=1}^n$ where

$$f_N^{(j)}(s) = f_*^{(j)}(s) + \sum_{r=1}^N U_r^j \varphi_r(s), \quad (1.6)$$

with N being an even integer, $2\pi f_*^{(j)}(s) := (s + \pi) f^{(j)}(\pi) - (s - \pi) f^{(j)}(-\pi)$ (the values of $f^{(j)}(\pi)$ and $f^{(j)}(-\pi)$ can be computed by the continuity requirements at the vertices of the polygon), and the set of real valued linearly independent functions $\{\varphi_r(s)\}_{r=1}^N$ being the *basis* functions. If we evaluate equation (1.4) on the following

n-rays of the complex k – plane: $k_p = -\frac{l}{h_p}$, $l \in \mathbb{R}^+$, $p = 1, \dots, n$, then the real coefficients U_r^j satisfy the system of linear algebraic equations

$$\sum_{j=1}^n \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-i\frac{l}{h_p}(m_p - m_j)} \sum_{r=1}^N U_r^j \int_{-\pi}^{\pi} e^{il\frac{h_j}{h_p}s} \varphi_r(s) ds = G_p(l) \quad (1.7)$$

where $G_p(l)$ denotes the known function

$$G_p(l) = i \sum_{j=1}^n \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-i\frac{l}{h_p}(m_p - m_j)} \int_{-\pi}^{\pi} e^{il\frac{h_j}{h_p}s} \left(g^{(j)}(s) + i f_*^{(j)}(s) \right) ds, \quad (1.8)$$

and l is chosen as follows: $l = \frac{1}{2}, \frac{3}{2}, \dots, \frac{N-1}{2}$ and $l = 1, 2, \dots, \frac{N}{2}$ for the real and imaginary part of equations (1.7), respectively, defining a set of *Collocation points*.

2 Collocation Matrix Structure for Square Domains

Consider, now, the square with vertices z_j and sides S_j , $j = 1, 2, 3, 4$ (modulo 4), indexed counter-clockwise, and interior $\mathbf{D}$, depicted in Fig. 2.1. Without any loss of the generality, we may assume that the square is centered at the origin, scaled and oriented so that one vertex (say z_1) is located at 1 so that

$$z_j = i^{j-1}, \quad j = 1, 2, 3, 4 \quad (2.1)$$

and the angle α_j of the side S_j from the real axis (measured counterclockwise) is given by

$$\alpha_j = \arg(z_{j+1} - z_j) = (2j + 1)\frac{\pi}{4}, \quad j = 1, 2, 3, 4. \quad (2.2)$$

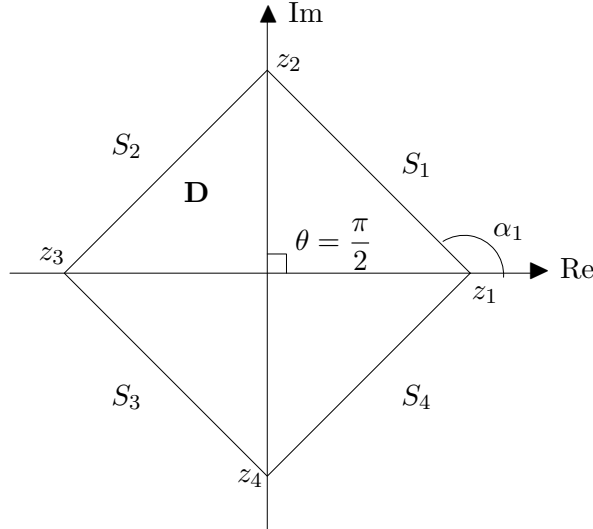


Figure 2.1 Square domain with vertices z_j , sides S_j and interior $\mathbf{D}$.

Case I : Same Boundary Conditions on all Sides

Assuming that the real-valued function $q(z, \bar{z})$ satisfies the Laplace's equation in the interior $\mathbf{D}$ of the square, described above, subject to the same type of Poincaré boundary conditions on all sides, that is

$$\cos(\beta) q_s^{(j)} + \sin(\beta) q_n^{(j)} = g^{(j)}, \quad z \in S_j, \quad 1 \leq j \leq 4, \quad (2.3)$$

and observing that the local coordinates of (1.5) take the form

$$m_j = \frac{1}{2} (z_j + z_{j+1}) = |m_j| e^{i(a_j - \frac{\pi}{2})} = \frac{1}{\sqrt{2}} e^{i(2j-1)\frac{\pi}{4}} = \frac{1}{\sqrt{2}} i^{(2j-1)/2}, \quad (2.4)$$

and

$$h_j = \frac{1}{2\pi} (z_{j+1} - z_j) = |h_j| e^{ia_j} = \frac{1}{\pi\sqrt{2}} e^{i(2j+1)\frac{\pi}{4}} = \frac{1}{\pi\sqrt{2}} i^{(2j+1)/2}, \quad (2.5)$$

we can easily obtain, from (1.4), that:

Lemma 2.1 *Let the real-valued function $q(z, \bar{z})$ satisfy the Laplace equation in the interior D of the square described above in this section. Let $g^{(j)}$ denote the derivative of the solution in the direction making an angle β , $0 \leq \beta \leq \pi$, with the side S_j (see (2.3)), and let $f^{(j)}$ denote the derivative of the solution in the direction normal to the above direction. The generalized Dirichlet-Neumann map is characterized by the equation*

$$\sum_{j=1}^4 e^{-kM_j} \int_{-\pi}^{\pi} e^{-kH_j s} \left(f^{(j)}(s) - i g^{(j)}(s) \right) ds = 0, \quad k \in \mathbb{C}, \quad (2.6)$$

where

$$M_j = im_j = \frac{1}{\sqrt{2}} i^{(2j+1)/2} \quad \text{and} \quad H_j = ih_j = \frac{1}{\pi\sqrt{2}} i^{(2j+3)/2}. \quad (2.7)$$

Proof. Upon simplification of the factors $|h_j|$ and $e^{i\beta_j}$, as $|h_j| = \frac{1}{2\pi}$ and $\beta_j = \beta$, from (1.4), the proof follows immediately. $\square$

Hence, upon evaluation of (2.6) on the following four rays of the complex k -plane

$$k_p = -\frac{l}{h_p}, \quad l \in \mathbb{R}^+, \quad p = 1, 2, 3, 4, \quad (2.8)$$

we obtain that:

Proposition 2.1 *Consider the generalized Dirichlet-Neumann map in Lemma 2.1. Suppose that the set $\{g^{(j)}\}_{j=1}^4$ is given through (2.3) and that the set $\{f^{(j)}\}_{j=1}^4$ is approximated by $\{f_N^{(j)}\}_{j=1}^4$ defined in (1.6). Then, the real coefficients U_r^j satisfy the $4N \times 4N$ linear system of equations*

$$\sum_{j=1}^4 e^{l\pi i^{j-p}} \sum_{r=1}^N U_r^j F_r (li^{j-p}) = G_p(l), \quad p = 1, 2, 3, 4, \quad (2.9)$$

where $G_p(l)$ denotes the known function

$$G_p(l) = i \sum_{j=1}^4 e^{l\pi i^{j-p}} \int_{-\pi}^{\pi} e^{ls i^{j-p+1}} (g^{(j)}(s) + i f_*^{(j)}(s)) ds, \quad (2.10)$$

$F_r(l)$ denotes the integral

$$F_r(l) = \int_{-\pi}^{\pi} e^{ils} \varphi_r(s) ds, \quad r = 1, 2, \dots, N, \quad (2.11)$$

and l is chosen as follows: For the real part of equations (2.9) $l = \frac{1}{2}, \frac{3}{2}, \dots, \frac{N-1}{2}$, whereas for the imaginary part of equations (2.9) $l = 1, 2, \dots, N/2$.

Proof. Observe that

$$\frac{M_j}{h_p} = \pi i^{j-p} \quad \text{and} \quad \frac{H_j}{h_p} = i^{j-p+1}. \quad (2.12)$$

Thus, evaluation of (2.6) at (2.8) yields the set of the four equations

$$\sum_{j=1}^4 e^{l\pi i^{j-p}} \int_{-\pi}^{\pi} e^{ls i^{j-p+1}} (f^{(j)}(s) - i g^{(j)}(s)) ds = 0, \quad l \in \mathbb{R}^+, \quad p = 1, 2, 3, 4, \quad (2.13)$$

hence, the proof follows immediately upon substitution of (1.6) into (2.13). $\square$

If we now let $A_{p,j} \in \mathbb{R}^{N,N}$ ($p, j = 1, 2, 3, 4$), to denote the $N \times N$ matrix with elements $a_{q,r}^{p,j}$ defined by

$$a_{q,r}^{p,j} = \begin{cases} \operatorname{Re} \left(e^{l\pi i^{j-p}} \int_{-\pi}^{\pi} e^{ls i^{j-p+1}} \varphi_r(s) ds \right), & l = \frac{1}{2}, \frac{3}{2}, \dots, \frac{N-1}{2} \\ \operatorname{Im} \left(e^{l\pi i^{j-p}} \int_{-\pi}^{\pi} e^{ls i^{j-p+1}} \varphi_r(s) ds \right), & l = 1, 2, \dots, N/2 \end{cases}, \quad (2.14)$$

for $q = 2l$ and $r = 1, 2, \dots, N$, then the collocation linear system, described in Proposition 2.1, may be written as

$$A_C \mathbf{U} = \mathbf{G}, \quad A_C \in \mathbb{R}^{4N, 4N}, \quad \mathbf{U}, \mathbf{G} \in \mathbb{R}^{4N}, \quad (2.15)$$

where

$$A_C = \begin{pmatrix} A_{1,1} & A_{1,2} & A_{1,3} & A_{1,4} \\ A_{2,1} & A_{2,2} & A_{2,3} & A_{2,4} \\ A_{3,1} & A_{3,2} & A_{3,3} & A_{3,4} \\ A_{4,1} & A_{4,2} & A_{4,3} & A_{4,4} \end{pmatrix}, \quad \mathbf{U} = \begin{pmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \\ \mathbf{U}_3 \\ \mathbf{U}_4 \end{pmatrix}, \quad \mathbf{G} = \begin{pmatrix} \mathbf{G}_1 \\ \mathbf{G}_2 \\ \mathbf{G}_3 \\ \mathbf{G}_4 \end{pmatrix} \quad (2.16)$$

and $\mathbf{U}_j \in \mathbb{R}^{N,1}$ and $\mathbf{G}_p \in \mathbb{R}^{N,1}$ denote the real vectors

$$\mathbf{U}_j = \{U_r^j\}_{r=1}^N = \left(U_1^j \ U_2^j \ \dots \ U_N^j \right)^T, \quad (2.17)$$

and

$$\mathbf{G}_p = \{G_q^p\}_{q=1}^N = \left(G_1^p \ G_2^p \ \dots \ G_N^p \right)^T, \quad (2.18)$$

with

$$G_q^p = \begin{cases} \operatorname{Re}(G_p(l)) , & l = \frac{1}{2}, \frac{3}{2}, \dots, \frac{N-1}{2}, \\ \operatorname{Im}(G_p(l)) , & l = 1, 2, \dots, N/2, \end{cases}, \quad q = 2l. \quad (2.19)$$

Following the notation above we prove:

Lemma 2.2 *The $N \times N$ real submatrices $A_{p,j} = \{a_{q,r}^{p,j}\}$, with $a_{q,r}^{p,j}$ being as defined in (2.14), satisfy*

$$A_{p,j} = E \begin{cases} A_0 & , \quad p = j \\ A_1 & , \quad |p - j| = 2 \\ O & , \quad |p - j| = 1, 3 \end{cases}, \quad (2.20)$$

where the elements of the matrix $A_0 = \{a_{q,r}\}_{q,r=1}^N$ are defined through the Finite Cosine/Sine Fourier Transform of the linear independent real valued basis functions $\phi_r(s)$, namely

$$a_{q,r} = \begin{cases} \int_{-\pi}^{\pi} \cos(\frac{q}{2}s) \phi_r(s) ds & , \quad q = \text{odd} \\ \int_{-\pi}^{\pi} \sin(\frac{q}{2}s) \phi_r(s) ds & , \quad q = \text{even} \end{cases}, \quad (2.21)$$

the matrix A_1 is defined by

$$A_1 = DA_0, \quad D = \operatorname{diag}(d_1, \dots, d_N), \quad d_q = (-1)^{q-1} e^{-q\pi}, \quad q = 1, \dots, N, \quad (2.22)$$

the matrix O denotes the null matrix and the diagonal matrix E is defined by

$$E = \operatorname{diag}(e_1, \dots, e_N), \quad e_q = e^{\frac{q}{2}\pi}, \quad q = 1, \dots, N. \quad (2.23)$$

Proof. Recall the definition of the elements $a_{q,r}^{p,j}$ from (2.14) and notice that, for $j = p$, there holds

$$a_{q,r}^{p,p} = e^{l\pi} \begin{cases} \operatorname{Re} \left(\int_{-\pi}^{\pi} e^{ils} \varphi_r(s) ds \right), & l = \frac{1}{2}, \frac{3}{2}, \dots, \frac{N-1}{2} \\ \operatorname{Im} \left(\int_{-\pi}^{\pi} e^{ils} \varphi_r(s) ds \right), & l = 1, 2, \dots, N/2 \end{cases}, \quad q = 2l.$$

Evidently, therefore,

$$a_{q,r}^{p,p} = e^{\frac{q}{2}\pi} a_{q,r} \quad (2.24)$$

where $a_{q,r}$ are as defined in (2.21), hence

$$A_{p,p} = EA_0 \quad , \quad p = 1, 2, 3, 4 \quad . \quad (2.25)$$

Similarly, as $i^{j-p} = -1$ for $|j - p| = 2$, there holds

$$a_{q,r}^{p,j} = e^{-l\pi} \begin{cases} \operatorname{Re} \left(\int_{-\pi}^{\pi} e^{-ils} \varphi_r(s) ds \right) = \int_{-\pi}^{\pi} \cos(ls) \phi_r(s) ds \quad , \quad l = \frac{1}{2}, \frac{3}{2}, \dots, \frac{N-1}{2} \\ \operatorname{Im} \left(\int_{-\pi}^{\pi} e^{-ils} \varphi_r(s) ds \right) = - \int_{-\pi}^{\pi} \sin(ls) \phi_r(s) ds \quad , \quad l = 1, 2, \dots, N/2 \end{cases} \quad ,$$

with $q = 2l$. Hence, for $|j - p| = 2$,

$$a_{q,r}^{p,j} = (-1)^{q-1} e^{-\frac{q}{2}\pi} a_{q,r} = e^{\frac{q}{2}\pi} ((-1)^{q-1} e^{-q\pi} a_{q,r}) \quad , \quad (2.26)$$

and therefore

$$A_{p,j} = EDA_0 = EA_1 \quad , \quad |p - j| = 2 \quad . \quad (2.27)$$

Finally, for $|j - p| = \text{odd}$, we have

$$a_{q,r}^{p,j} = \left(\int_{-\pi}^{\pi} e^{\pm ls} \varphi_r(s) ds \right) \begin{cases} \operatorname{Re} (e^{\pm il\pi}) = \cos(l\pi) = 0 \quad , \quad l = \frac{1}{2}, \frac{3}{2}, \dots, \frac{N-1}{2} \\ \operatorname{Im} (e^{\pm il\pi}) = \pm \sin(l\pi) = 0 \quad , \quad l = 1, 2, \dots, N/2 \end{cases} \quad ,$$

and, therefore,

$$A_{p,j} = O \quad , \quad |p - j| = \text{odd} \quad , \quad (2.28)$$

which completes the proof. $\square$

Therefore, it becomes apparent that

Proposition 2.2 *The Collocation linear system in (2.15) is equivalent to the system*

$$AU = (I_4 \otimes E^{-1})G \quad , \quad (2.29)$$

where $\otimes$ denotes the Kronecker (tensor) matrix product, A is defined by

$$A = \begin{pmatrix} A_0 & O & A_1 & O \\ O & A_0 & O & A_1 \\ A_1 & O & A_0 & O \\ O & A_1 & O & A_0 \end{pmatrix} = \begin{pmatrix} I & O & D & O \\ O & I & O & D \\ D & O & I & O \\ O & D & O & I \end{pmatrix} (I_4 \otimes A_0) \quad , \quad (2.30)$$

I_4 denotes the 4×4 identity matrix, and the matrices A_0 , A_1 , D and E are as defined in Lemma 2.2 above.

Remark 2.1 Notice that, as the basis functions $\varphi_r(s)$ are appropriately chosen real valued linearly independent functions, A_0 is nonsingular. Nonsingular is also the matrix B , defined by

$$B = \begin{pmatrix} I & O & D & O \\ O & I & O & D \\ D & O & I & O \\ O & D & O & I \end{pmatrix}, \quad (2.31)$$

as is apparently symmetric, strictly diagonally dominant and positive definite. Therefore, both matrices A in (2.30) and A_C in (2.16) are nonsingular too.

Remark 2.2 Observe that the matrix A in (2.30) is evidently *Block Circulant*. Naturally therefore, as $A_C = (I_4 \otimes E)A$, the collocation matrix A_C in (2.16) is *Block Circulant* too. In [5] it is shown that the Collocation coefficient matrix, although does not possess the special sparse structure of (2.30), remains *Block Circulant* for the case of general *Regular Polygons* with the same type of boundary conditions on all sides, allowing the deployment of FFT for the efficient solution of the corresponding collocation linear system.

Case II : Different Boundary Conditions on each Side

Let us now assume that the real-valued function $q(z, \bar{z})$ satisfies the Laplace's equation in the interior $\mathbf{D}$ of the square, described at the beginning of this section, subject to different type of Poincaré boundary conditions on each side, that is (see also equation (1.2))

$$\cos(\beta_j) q_s^{(j)} + \sin(\beta_j) q_n^{(j)} = g^{(j)}, \quad z \in S_j, \quad 1 \leq j \leq 4. \quad (2.32)$$

Then, the associated generalized Dirichlet-Neumann map is characterized by the equation

$$\sum_{j=1}^4 e^{i\beta_j} e^{-kM_j} \int_{-\pi}^{\pi} e^{-kH_j s} \left(f^{(j)}(s) - i g^{(j)}(s) \right) ds = 0, \quad k \in \mathbb{C}, \quad (2.33)$$

where M_j and H_j are as defined in Lemma 2.1, while Proposition 2.1 is being replaced by

Proposition 2.3 *Consider the generalized Dirichlet-Neumann map in (2.33). Suppose that the set $\{g^{(j)}\}_{j=1}^4$ is given through (2.32) and that the set $\{f^{(j)}\}_{j=1}^4$ is approximated by $\{f_N^{(j)}\}_{j=1}^4$ defined in (1.6). Then, the real coefficients U_r^j satisfy the $4N \times 4N$ linear system of equations*

$$\sum_{j=1}^4 e^{i\beta_j} e^{l\pi i^{j-p}} \sum_{r=1}^N U_r^j F_r(l i^{j-p}) = G_p(l), \quad p = 1, 2, 3, 4, \quad (2.34)$$

where $G_p(l)$ denotes the known function

$$G_p(l) = i \sum_{j=1}^4 e^{i\beta_j} e^{l\pi i^{j-p}} \int_{-\pi}^{\pi} e^{ls i^{j-p+1}} (g^{(j)}(s) + i f_*^{(j)}(s)) ds, \quad (2.35)$$

$F_r(l)$ is as in (2.11) and l is chosen as in Proposition 2.1.

The collocation linear system, described in Proposition 2.3 above, obviously is in the block partitioned form of (2.16) with the difference that the elements $\alpha_{q,r}^{p,j}$ of the submatrices $A_{p,j}$, used to defined the collocation matrix A_C in (2.16), are now defined by

$$\alpha_{q,r}^{p,j} = \begin{cases} \operatorname{Re} \left(e^{i\beta_j} e^{l\pi i^{j-p}} \int_{-\pi}^{\pi} e^{ls i^{j-p+1}} \varphi_r(s) ds \right), & l = \frac{1}{2}, \frac{3}{2}, \dots, \frac{N-1}{2} \\ \operatorname{Im} \left(e^{i\beta_j} e^{l\pi i^{j-p}} \int_{-\pi}^{\pi} e^{ls i^{j-p+1}} \varphi_r(s) ds \right), & l = 1, 2, \dots, N/2 \end{cases}, \quad (2.36)$$

and, of course, the vector $\mathbf{G}$ now refers to (2.35) instead of (2.10). It takes only a few simple algebraic manipulations to verify that

$$\alpha_{q,r}^{p,j} = a_{q,r}^{p,j} \cos(\beta_j) + \hat{a}_{q,r}^{p,j} \sin(\beta_j), \quad (2.37)$$

where $a_{q,r}^{p,j}$ is as defined in (2.14) and $\hat{a}_{q,r}^{p,j}$ is defined by

$$\hat{a}_{q,r}^{p,j} = \begin{cases} -\operatorname{Im} \left(e^{l\pi i^{j-p}} \int_{-\pi}^{\pi} e^{ls i^{j-p+1}} \varphi_r(s) ds \right), & l = \frac{1}{2}, \frac{3}{2}, \dots, \frac{N-1}{2} \\ \operatorname{Re} \left(e^{l\pi i^{j-p}} \int_{-\pi}^{\pi} e^{ls i^{j-p+1}} \varphi_r(s) ds \right), & l = 1, 2, \dots, N/2 \end{cases}, \quad (2.38)$$

with $q = 2l$ as always. Therefore, using also Proposition 2.2, the collocation coefficient matrix A_C now takes the form

$$A_C = (I_4 \otimes E) A (D_c \otimes I_N) + \hat{A} (D_s \otimes I_N), \quad (2.39)$$

where the matrices A and E are as defined in (2.30) and (2.23), respectively, the diagonal matrices D_c and D_s are defined by

$$D_c = \operatorname{diag}(\cos(\beta_1), \cos(\beta_2), \cos(\beta_3), \cos(\beta_4)) \quad (2.40)$$

and

$$D_s = \operatorname{diag}(\sin(\beta_1), \sin(\beta_2), \sin(\beta_3), \sin(\beta_4)) , \quad (2.41)$$

and the matrix $\hat{A} \in \mathbb{R}^{4N, 4N}$ is in the block partitioned form

$$\hat{A} = \begin{pmatrix} \hat{A}_{1,1} & \hat{A}_{1,2} & \hat{A}_{1,3} & \hat{A}_{1,4} \\ \hat{A}_{2,1} & \hat{A}_{2,2} & \hat{A}_{2,3} & \hat{A}_{2,4} \\ \hat{A}_{3,1} & \hat{A}_{3,2} & \hat{A}_{3,3} & \hat{A}_{3,4} \\ \hat{A}_{4,1} & \hat{A}_{4,2} & \hat{A}_{4,3} & \hat{A}_{4,4} \end{pmatrix}, \quad (2.42)$$

with the elements $\hat{a}_{q,r}^{p,j}$ of the submatrices $\hat{A}_{p,j} \in \mathbb{R}^{N,N}$ ($p, j = 1, 2, 3, 4,$) being defined in (2.38). With this notation we now prove that

Lemma 2.3 *The $N \times N$ real submatrices $\hat{A}_{p,j} = \left\{ \hat{a}_{q,r}^{p,j} \right\}$, with $\hat{a}_{q,r}^{p,j}$ being as defined in (2.38) satisfy*

$$\hat{A}_{p,j} = \begin{cases} E\hat{A}_0 & , \quad p = j \\ -ED\hat{A}_0 & , \quad |p - j| = 2 \\ \hat{D}\hat{A}_1 & , \quad p - j = -1, 3 \\ \hat{D}\hat{A}_2 & , \quad p - j = 1, -3 \end{cases} , \quad (2.43)$$

where the elements of the matrix $\hat{A}_0 = \left\{ \hat{a}_{q,r}^{(0)} \right\}_{q,r=1}^N$ are defined through the Finite Cosine/Sine Fourier Transform of the linear independent real valued basis functions $\phi_r(s)$, namely

$$\hat{a}_{q,r}^{(0)} = \begin{cases} -\int_{-\pi}^{\pi} \sin(\frac{q}{2}s) \phi_r(s) ds & , \quad q = \text{odd} \\ \int_{-\pi}^{\pi} \cos(\frac{q}{2}s) \phi_r(s) ds & , \quad q = \text{even} \end{cases} , \quad (2.44)$$

the elements of the matrix $\hat{A}_1 = \left\{ \hat{a}_{q,r}^{(1)} \right\}_{q,r=1}^N$ are defined by

$$\hat{a}_{q,r}^{(1)} = \int_{-\pi}^{\pi} e^{\frac{q}{2}s} \phi_r(s) ds , \quad (2.45)$$

the elements of the matrix $\hat{A}_2 = \left\{ \hat{a}_{q,r}^{(2)} \right\}_{q,r=1}^N$ are defined by

$$\hat{a}_{q,r}^{(2)} = (-1)^q \int_{-\pi}^{\pi} e^{-\frac{q}{2}s} \phi_r(s) ds , \quad (2.46)$$

the matrices D and E are as defined in Lemma 2.2 and the diagonal matrix $\hat{D}$ is defined by

$$\hat{D} = \text{diag} \left(\sin(\frac{\pi}{2}), \cos(2\frac{\pi}{2}), \dots, \sin((N-1)\frac{\pi}{2}), \cos(N\frac{\pi}{2}) \right) . \quad (2.47)$$

Proof. As in Lemma 2.2, recall the definition of the elements $\hat{a}_{q,r}^{p,j}$ from (2.38) and notice that, for $j = p$, there holds

$$\hat{a}_{q,r}^{p,p} = e^{l\pi} \begin{cases} -\text{Im} \left(\int_{-\pi}^{\pi} e^{ils} \varphi_r(s) ds \right) , & l = \frac{1}{2}, \frac{3}{2}, \dots, \frac{N-1}{2} \\ \text{Re} \left(\int_{-\pi}^{\pi} e^{ils} \varphi_r(s) ds \right) , & l = 1, 2, \dots, N/2 \end{cases} , \quad q = 2l .$$

Evidently, therefore,

$$\hat{a}_{q,r}^{p,p} = e^{\frac{q}{2}\pi} \hat{a}_{q,r}^{(0)} \quad (2.48)$$

where $\hat{a}_{q,r}^{(0)}$ are as defined in (2.44), hence

$$\hat{A}_{p,p} = E\hat{A}_0 \quad , \quad p = 1, 2, 3, 4 \quad . \quad (2.49)$$

Similarly, as $i^{j-p} = -1$ for $|j - p| = 2$, there holds

$$\hat{a}_{q,r}^{p,j} = e^{-l\pi} \begin{cases} -\text{Im} \left(\int_{-\pi}^{\pi} e^{-ils} \varphi_r(s) ds \right) = \int_{-\pi}^{\pi} \sin(ls) \phi_r(s) ds \quad , \quad l = \frac{1}{2}, \frac{3}{2}, \dots, \frac{N-1}{2} \\ \text{Re} \left(\int_{-\pi}^{\pi} e^{-ils} \varphi_r(s) ds \right) = \int_{-\pi}^{\pi} \cos(ls) \phi_r(s) ds \quad , \quad l = 1, 2, \dots, N/2 \end{cases} \quad ,$$

with $q = 2l$. Hence, for $|j - p| = 2$,

$$\hat{a}_{q,r}^{p,j} = (-1)^q e^{-\frac{q}{2}\pi} \hat{a}_{q,r}^{(0)} = -e^{\frac{q}{2}\pi} \left((-1)^{q-1} e^{-q\pi} a_{q,r}^{(0)} \right) \quad , \quad (2.50)$$

and therefore

$$\hat{A}_{p,j} = -ED\hat{A}_0 \quad , \quad |p - j| = 2 \quad . \quad (2.51)$$

Now, as $i^{j-p} = -i$ for $j - p = -1$ or $j - p = 3$, we have

$$\hat{a}_{q,r}^{p,j} = \left(\int_{-\pi}^{\pi} e^{ls} \varphi_r(s) ds \right) \begin{cases} -\text{Im} (e^{-il\pi}) = \sin(l\pi) \quad , \quad l = \frac{1}{2}, \frac{3}{2}, \dots, \frac{N-1}{2} \\ \text{Re} (e^{-il\pi}) = \cos(l\pi) \quad , \quad l = 1, 2, \dots, N/2 \end{cases} \quad ,$$

and, therefore,

$$\hat{A}_{p,j} = \hat{D}\hat{A}_1 \quad , \quad p - j = -1, 3 \quad . \quad (2.52)$$

Finally, as $i^{j-p} = i$ for $j - p = 1$ or $j - p = -3$, we have

$$\hat{a}_{q,r}^{p,j} = \left(\int_{-\pi}^{\pi} e^{-ls} \varphi_r(s) ds \right) \begin{cases} -\text{Im} (e^{il\pi}) = -\sin(l\pi) \quad , \quad l = \frac{1}{2}, \frac{3}{2}, \dots, \frac{N-1}{2} \\ \text{Re} (e^{il\pi}) = \cos(l\pi) \quad , \quad l = 1, 2, \dots, N/2 \end{cases} \quad ,$$

and, therefore,

$$\hat{A}_{p,j} = \hat{D}\hat{A}_2 \quad , \quad p - j = 1, -3 \quad , \quad (2.53)$$

which completes the proof. $\square$

Evidently, therefore, the matrix $\hat{A}$ in (2.42) can be expressed as

$$\hat{A} = (I_4 \otimes E)\tilde{A}_1 + (I_4 \otimes \hat{D})\tilde{A}_2 \quad , \quad (2.54)$$

where $\tilde{A}_1$ and $\tilde{A}_2$ denote the block circulant matrices

$$\tilde{A}_1 = \begin{pmatrix} \hat{A}_0 & O & -D\hat{A}_0 & O \\ O & \hat{A}_0 & O & -D\hat{A}_0 \\ -D\hat{A}_0 & O & \hat{A}_0 & O \\ O & -D\hat{A}_0 & O & \hat{A}_0 \end{pmatrix} \quad \text{and} \quad \tilde{A}_2 = \begin{pmatrix} O & \hat{A}_2 & O & \hat{A}_1 \\ \hat{A}_1 & O & \hat{A}_2 & O \\ O & \hat{A}_1 & O & \hat{A}_2 \\ \hat{A}_2 & O & \hat{A}_1 & O \end{pmatrix} \quad . \quad (2.55)$$

If we now let the matrix $\hat{B}$ to be defined by

$$\hat{B} = \begin{pmatrix} I & O & -D & O \\ O & I & O & -D \\ -D & O & I & O \\ O & -D & O & I \end{pmatrix}, \quad (2.56)$$

then, upon combination of the results above, we obtain

Proposition 2.4 *The Collocation coefficient matrix A_C , associated with the linear system described in Proposition 2.3, is expressed as*

$$A_C = (I_4 \otimes E) \left(B(I_4 \otimes A_0) (D_c \otimes I_N) + \hat{B}(I_4 \otimes \hat{A}_0) (D_s \otimes I_N) \right) + (I_4 \otimes \hat{D}) \tilde{A}_2 (D_s \otimes I_N). \quad (2.57)$$

where the diagonal matrix E and the matrix A_0 are defined in Lemma 2.2, the matrices B and $\hat{B}$ are as defined in (2.31) and (2.56) respectively, the diagonal matrices D_c and D_s are as defined in (2.40) and (2.41) respectively, the matrix $\hat{A}_0$ is defined in Lemma 2.3 and the matrix $\tilde{A}_2$ is as defined in (2.55).

Proof. Recall (2.55) and observe that $\tilde{A}_1 = \hat{B}(I_4 \otimes \hat{A}_0)$. This, combined with relations (2.30), (2.39) and (2.54) yields (2.57) and the proof follows. $\square$

3 Analysis and Implementation of Numerical Methods

Based on the structure, as well as the properties, of the Collocation coefficient matrix, in this Section we analyze and implement direct and iterative methods for determining the solution of the generalized Dirichlet-Neumann map associated to Laplace's equation on square domains. For the numerical experiments included, we considered the solution of the model Laplace's equation, with exact solution (cf. [2]-[3])

$$q(x, y) = \sinh(3x) \sin(3y). \quad (3.1)$$

The relative error E_∞ , used to demonstrate the convergence behavior of the direct and iterative methods considered, is given by

$$E_\infty = \frac{\|f - f_N\|_\infty}{\|f\|_\infty}, \quad (3.2)$$

where

$$\|f\|_\infty = \max_{1 \leq j \leq n} \left\{ \max_{-\pi \leq s \leq \pi} |f^{(j)}(s)| \right\} \quad (3.3)$$

and

$$\|f - f_N\|_\infty = \max_{1 \leq j \leq n} \left\{ \max_{-\pi \leq s \leq \pi} |f^{(j)}(s) - f_N^{(j)}(s)| \right\}, \quad (3.4)$$

with $f_N^{(j)}$ as in (1.6), and the max over s is taken over a dense discretization of the interval $[-\pi, \pi]$. For the direct solution of the linear systems we have used the standard LAPACK routines, while for the computation of the right hand side vector we have used a routine (*dqawo*) from QUADPACK implementing the modified Clenshaw-Curtis technique. As it pertains to the iterative methods, the maximum number of iterations, allowed for all methods to perform, is set to 200 and the zero iterate $U^{(0)}$ is set to be equal to the right hand side vector. All experiments were conducted on a multiuser SUN V240 system using the Fortran-90 compiler.

Case I : Same Boundary Conditions on all Sides

It is the special sparse structure, revealed in the previous section, of the collocation system, in (2.29), that allow us to efficiently and rapidly solve it.

Direct Solution

Taking advantage of the block structure of the matrix A in (2.30), and observing that the inverse of the matrix B in (2.31) is readily available by

$$B^{-1} = \hat{B}(I_4 \otimes C) , \quad (3.5)$$

where $\hat{B}$ is as defined in (2.56) and C is the diagonal matrix

$$C = \text{diag}(c_1, \dots, c_N) \quad , \quad c_q = \frac{1}{1 - d_q^2} = \frac{1}{1 - e^{-2q\pi}} \quad , \quad q = 1, \dots, N \quad , \quad (3.6)$$

with d_q denoting the diagonal elements of the matrix D in (2.22), it is evident that the collocation system (2.29) can be written as

$$(I_4 \otimes A_0)\mathbf{U} = \hat{B}(I_4 \otimes C)(I_4 \otimes E^{-1})\mathbf{G} , \quad (3.7)$$

or, equivalently, as

$$A_0\mathbf{U}_p = CE^{-1}(\mathbf{G}_p - D\mathbf{G}_{p'}) \quad , \quad p' = (p+2) \bmod 4 \quad , \quad p = 1, 2, 3, 4 \quad , \quad (3.8)$$

since the matrices C, D and E are diagonal and commute. The matrix A_0 , defined in Lemma 2.2, depends on the choice of basis functions $\varphi_r(s)$, as its elements are defined through their discrete cosine/sine Fourier transforms (see (2.21)). In [3] we considered the following two choices of basis functions :

(1) *Sine Basis Functions*

$$\varphi_r(s) = \sin\left(r\left(\frac{\pi + s}{2}\right)\right) \quad , \quad r = 1, \dots, N \quad . \quad (3.9)$$

(2) *Chebyshev Basis Functions*

$$\varphi_r(s) = \begin{cases} T_{r+1}\left(\frac{s}{\pi}\right) - T_0\left(\frac{s}{\pi}\right) , & r \text{ odd}, \\ T_{r+1}\left(\frac{s}{\pi}\right) - T_1\left(\frac{s}{\pi}\right) , & r \text{ even}. \end{cases} \quad , \quad r = 1, \dots, N \quad , \quad (3.10)$$

where $T_n(x) = \cos(n \cos^{-1}(x))$.

For the case of sine basis functions the matrix A_0 is point diagonal, hence the solution of (3.8) is readily available with computational cost of $\mathcal{O}(N)$. In general, though, including the case of Chebyshev basis functions, it is well known that the computational cost for solving the system (3.8) is $\mathcal{O}(N^3)$.

Iterative Solution

For an iterative analysis, independent from the choice of basis functions, one may take advantage of the *2-cyclic* (cf. [9]) nature of the matrix A in (2.30). Observing that its associated *weakly cyclic of index 2* (cf. [9]) block Jacobi iteration matrix T_0 can be expressed as

$$T_0 = (I_4 \otimes A_0^{-1})(I - B)(I_4 \otimes A_0) , \quad (3.11)$$

hence is similar to the matrix

$$I - B = - \begin{pmatrix} O & O & D & O \\ O & O & O & D \\ D & O & O & O \\ O & D & O & O \end{pmatrix} \quad (3.12)$$

where B is as defined in (2.31) and D is the diagonal matrix of (2.22), its spectrum $\sigma(T_0)$ satisfies

$$\sigma(T_0) = \{\pm e^{-q\pi} , \pm e^{-q\pi}\}_{q=1}^N , \quad (3.13)$$

and, obviously, its spectral radius $\varrho(T_0)$ is given by

$$\varrho(T_0) = e^{-\pi} \approx 0.0432 , \quad (3.14)$$

revealing a fast rate of convergence. Moreover, using well known results from the literature (e.g. cf. [9]), the spectral radii of the iteration matrices T_1 and $T_{\omega_{opt}}$, associated to the Gauss-Seidel and the optimal SOR iterative methods, respectively, satisfy

$$\varrho(T_1) = \varrho^2(T_0) = e^{-2\pi} \approx 0.0019 , \quad (3.15)$$

and

$$\varrho(T_{\omega_{opt}}) = \omega_{opt} - 1 = \frac{2}{1 + \sqrt{1 - e^{-2\pi}}} - 1 \approx 0.0005 , \quad (3.16)$$

revealing rapid convergence rates. However, we have to point out that, in view of (3.8), iterative methods may only compete direct factorization, since for all methods the main computational cost comes from the factorization of the matrix A_0 . To be more specific, with the change of variables

$$\mathbf{V} = A_0 \mathbf{U} , \quad (3.17)$$

the above iterative methods may be implemented through the following expressions:

- *Jacobi*

$$\mathbf{V}_p^{(m+1)} = -D\mathbf{V}_{p'}^{(m)} + E^{-1}\mathbf{G}_p , \quad p' = (p+2) \bmod 4 , \quad p = 1, 2, 3, 4 , \quad (3.18)$$

- *Gauss-Seidel*

$$\begin{cases} \mathbf{V}_p^{(m+1)} = -D\mathbf{V}_{p+2}^{(m)} + E^{-1}\mathbf{G}_p & , \quad p = 1, 2 \\ \mathbf{V}_{p+2}^{(m+1)} = -D\mathbf{V}_p^{(m+1)} + E^{-1}\mathbf{G}_p & , \quad p = 1, 2 \end{cases} \quad (3.19)$$

- *SOR*

$$\begin{cases} \mathbf{V}_p^{(m+1)} = (1 - \omega)\mathbf{V}_p^{(m)} - \omega D\mathbf{V}_{p+2}^{(m)} + \omega E^{-1}\mathbf{G}_p & , \quad p = 1, 2 \\ \mathbf{V}_{p+2}^{(m+1)} = (1 - \omega)\mathbf{V}_{p+2}^{(m)} - \omega D\mathbf{V}_p^{(m+1)} + \omega E^{-1}\mathbf{G}_p & , \quad p = 1, 2 \end{cases} \quad (3.20)$$

Consequently, by making also use of the fast convergence properties of the iterative methods considered, it is apparent that the computational cost, for the iterative solution, is $\mathcal{O}(N)$ for the case of sine basis functions, while, in general, including the case of Chebyshev basis functions, is $\mathcal{O}(N^3)$.

For completeness and uniformity (with the case of different boundary conditions) only purposes, we also consider two of the main representatives from the family of Krylov subspace iterative methods, namely the Bi-CGSTAB [6] and the GMRES [7] methods, for the solution of the preconditioned system

$$AM^{-1}\hat{\mathbf{U}} = (I_4 \otimes E^{-1})\mathbf{G} , \quad (3.21)$$

where, of course, $\hat{\mathbf{U}} = M\mathbf{U}$. Observing that both spectra $\sigma(T_0)$ and $\sigma(T_1) = \sigma^2(T_0)$ of the block Jacobi and block Gauss-Seidel iteration matrices, respectively, are real and *clustered* around zero, it is evident that if we choose the preconditioning matrix M to be the splitting matrix of the Jacobi or the Gauss-Seidel iterative methods, namely

$$M \equiv M_0 = I_4 \otimes A_0 \quad \text{or} \quad M \equiv M_1 = F(I_4 \otimes A_0) \quad (3.22)$$

where

$$F = \begin{pmatrix} I & O & O & O \\ O & I & O & O \\ D & O & I & O \\ O & D & O & I \end{pmatrix} , \quad (3.23)$$

then the spectrum of the preconditioned matrix AM^{-1} would satisfy

$$\sigma(AM_0^{-1}) = 1 - \sigma(T_0) \quad \text{or} \quad \sigma(AM_1^{-1}) = 1 - \sigma(T_1) , \quad (3.24)$$

since $T_0 = I - M_0^{-1}A$, $T_1 = I - M_1^{-1}A$ and the matrices $M^{-1}A$ and AM^{-1} are obviously similar. Therefore, the eigenvalues of the preconditioned matrices AM_0^{-1} and AM_1^{-1} are all real, located in the half complex plane with the origin being outside or towards the boundary of the the convex hull containing them, and clustered around unity. Hence, following [8], the Bi-CGSTAB is expected to have effective convergence properties.

To numerically demonstrate the above results we include Table 1 referring to the performance of all mentioned numerical methods when they apply to the model

problem, described at the beginning of this section, for the case of Chebyshev basis functions.

Table 1 Performance of Numerical Methods (Same BC — Chebyshev Basis Functions)

Method	Preconditioner	$N = 8$			$N = 16$		
		Error	Iter.	Time	Error	Iter.	Time
LU-factorization	—	2.09e-05	—	1.50e-04	5.78e-13	—	2.33e-04
Jacobi	—	2.09e-05	13	2.52e-04	5.78e-13	13	4.74e-04
Gauss-Seidel	—	2.09e-05	7	1.64e-04	5.78e-13	7	2.89e-04
SOR	—	2.09e-05	7	2.05e-04	5.78e-13	7	3.52e-04
Bi-CGSTAB	Jacobi	2.09e-05	2	7.27e-04	5.78e-13	2	8.43e-04
	Gauss-Seidel	2.09e-05	2	7.22e-04	5.78e-13	2	8.37e-04
GMRES(10)	Jacobi	2.09e-05	4	9.21e-04	5.78e-13	4	1.08e-03
	Gauss-Seidel	2.09e-05	3	8.71e-04	5.78e-13	3	1.02e-03

Case II : Different Boundary Conditions on each Side

The numerical treatment, for the case of different boundary conditions on each side of the square domain, largely depends on the boundary conditions used per se. Hence, the numerical results included for this case, are indicative and refer to the mixed boundary conditions (see (2.32)) obtained by making use of the following angles:

$$\beta_1 = \pi \quad , \quad \beta_2 = \frac{\pi}{4} \quad , \quad \beta_3 = \frac{\pi}{6} \quad , \quad \beta_4 = \frac{\pi}{3} \quad .$$

Recall, now, the associated, to the above boundary conditions, collocation matrix A_C , defined in Proposition 2.3, and observe that relation (2.57) or, equivalently, relation (2.39) combined with relation (2.54), contributes to its efficient construction, as it is written as a matrix combination of circulant matrices, one of which is the matrix A , defined in (2.30), associated to the case of same boundary conditions on all sides of the square.

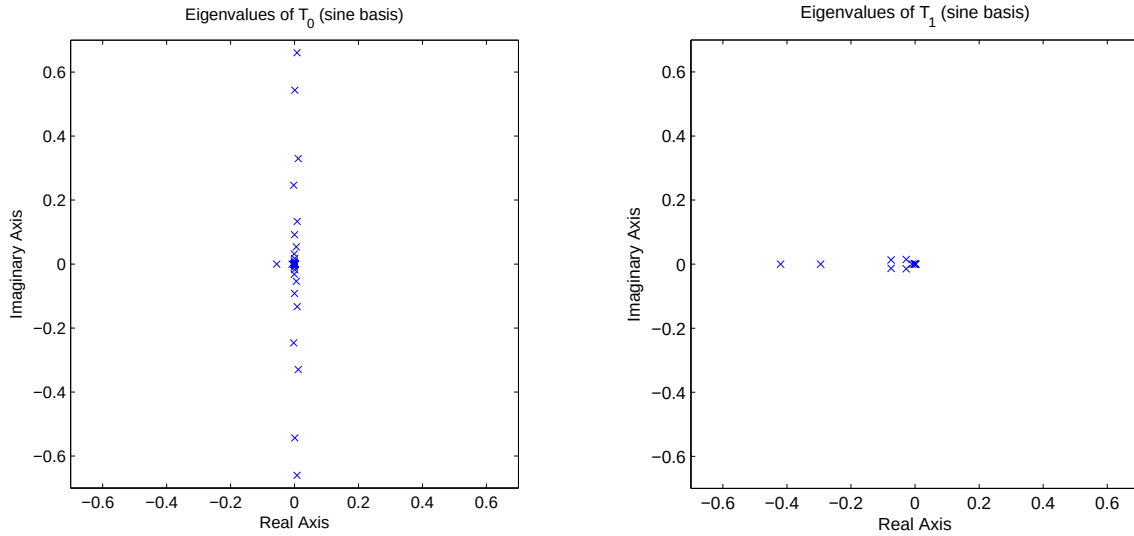


Fig. 1 : Eigenvalues of the block Jacobi and GS iteration matrices T_0 and T_1 for Sine Basis Functions ($N = 64$)

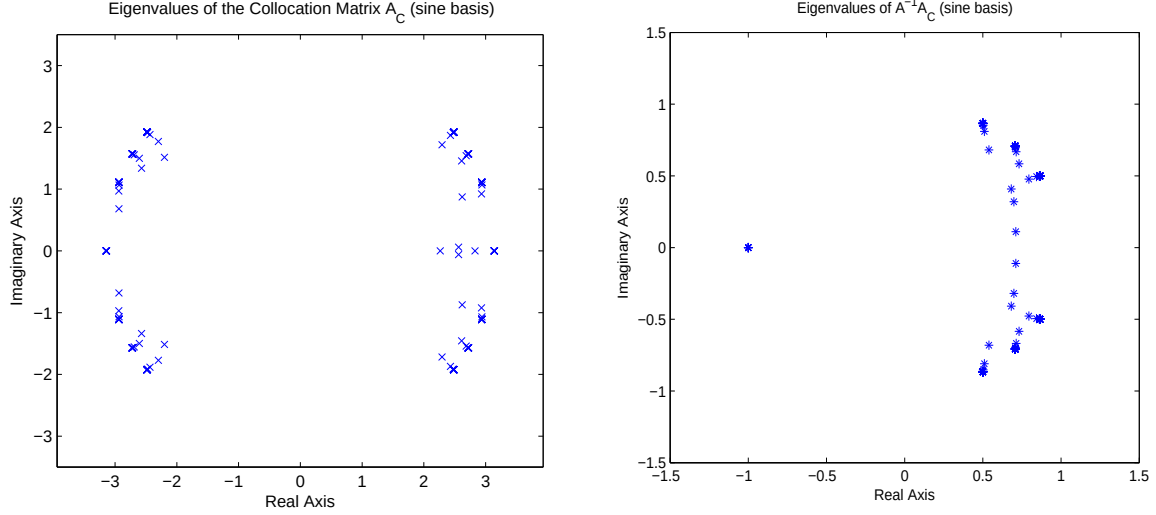


Fig. 2 : Eigenvalues of the Matrices A_C and $A^{-1}A_C$ for Sine Basis Functions ($N = 64$)

For sine basis functions, iterative methods are an effective alternative to direct factorization. And this because, as the collocation method, combined with the sine basis functions of (3.9), is quadratically convergent, it is necessary to use a sufficiently large number of basis functions (large N) to achieve a sufficiently small error norm.

To illustrate the convergence behavior of the classical block Jacobi and Gauss-Seidel (GS) methods, with iteration matrices $T_0 = M_0^{-1}N_0$ and $T_1 = M_1^{-1}N_1$ respectively, where

$$M_0 = \bigoplus_{p=1}^4 M_0^{(p)} \quad \text{with} \quad M_0^{(p)} = E \left(A_0 \cos(\beta_p) + \hat{A}_0 \sin(\beta_p) \right) \quad (3.25)$$

and M_1 defined analogously, we included Figure 1 depicting their eigenvalue distribution for a typical case ($N = 64$). Pertaining to the Krylov Bi-CGSTAB and GMRES methods, it is apparent that the use of the un-preconditioned versions is not suggested due to the A_C 's eigenvalue distribution depicted in Figure 2.

Table 2 Performance of Numerical Methods (Different BC — Sine Basis Functions)

Method	Preconditioner	$N = 32$			$N = 128$			$N = 512$		
		Error	Iter.	Time	Error	Iter.	Time	Error	Iter.	Time
LU-factor.	—	2.05e-03	—	2.29e-02	1.31e-04	—	1.51	7.69e-06	—	192.00
Jacobi	—	2.05e-03	35	2.53e-02	1.31e-04	43	0.76	7.67e-06	53	35.20
GS	—	2.05e-03	16	1.36e-02	1.31e-04	20	0.39	7.69e-06	24	19.30
Bi-CGSTAB	Jacobi	2.05e-03	8	1.48e-02	1.31e-04	9	0.51	7.70e-06	9	17.60
	GS	2.05e-03	4	1.08e-02	1.31e-04	5	0.40	7.69e-06	5	15.30
	A	2.05e-03	29	8.98e-03	1.31e-04	25	0.09	7.62e-06	32	15.00
GMRES(10)	Jacobi	2.05e-03	12	1.36e-02	1.31e-04	14	0.47	7.68e-06	16	17.20
	GS	2.05e-03	7	1.06e-02	1.31e-04	7	0.31	7.70e-06	7	12.70
	A	2.05e-03	37	8.44e-03	1.31e-04	35	0.07	7.67e-06	37	9.18

With respect to their preconditioned analogs, together with the block Jacobi and block GS preconditioning, we have also considered the case of using the block circulant matrix A of (2.30) as a preconditioner. And although the eigenvalue distribution of the preconditioned matrix $A^{-1}A_C$ (depicted in Figure 2) is not that encouraging, the fact that A^{-1} inverse is readily available combined with the large size of the matrices needed to be directly factored out, yields a very efficient preconditioning. In fact, the A -preconditioned GMRES method is significantly less time consuming, hence it is the method of preference. The performance results for all numerical methods considered for the case of sine basis functions have been included in Table 2 above.

For the case of Chebyshev basis functions the Collocation method appears to converge exponentially (cf. [3]). Therefore, one may achieve a small error norm with a few basis functions.

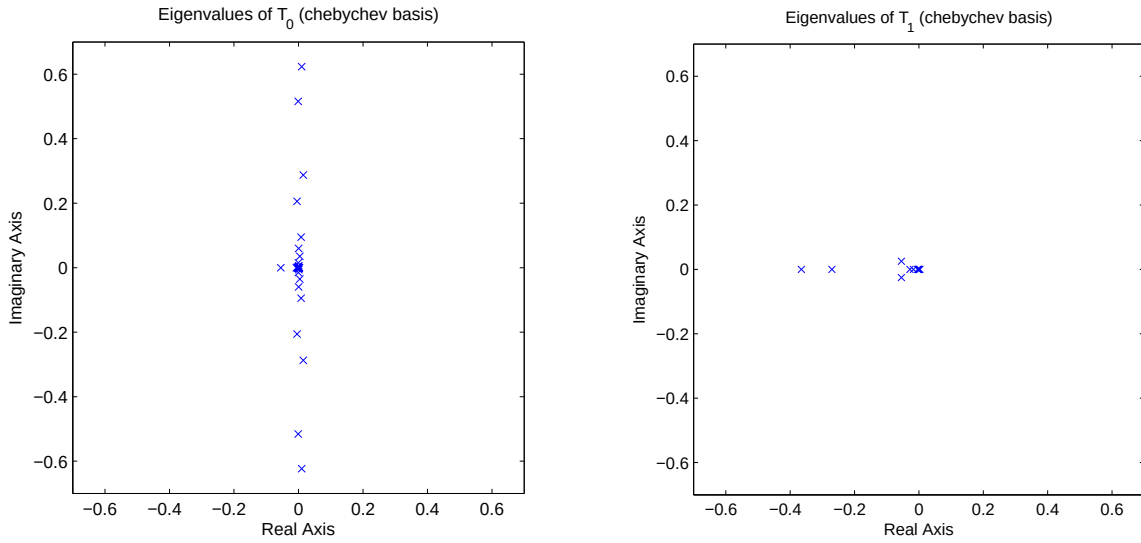


Fig. 3 : Eigenvalues of the block Jacobi and GS iteration matrices T_0 and T_1 for Chebyshev Basis Functions ($N = 16$)

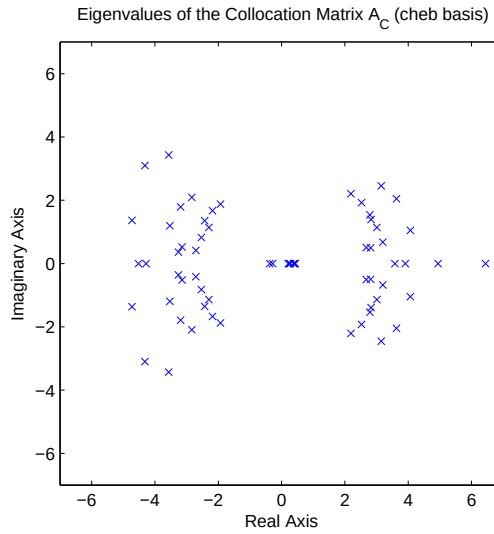


Fig. 4 : Eigenvalues of the Collocation Matrix A_C of (2.57) for Chebyshev Basis Functions ($N = 16$)

This fact leads to small size matrices and, therefore, direct factorization is more effective, than iterative methods, for their solution. Nevertheless, for comparison and demonstration purposes, together with the direct factorization method, we also consider the block Jacobi and GS methods, as well as their preconditioning analogs combined with the Bi-CGSTAB and GMRES methods. The eigenvalue distribution of the associated matrices T_0 , T_1 and A_C are depicted in Figures 3 and 4, while the performance results of all numerical methods considered are included in Table 3 below.

Table 3 Performance of Numerical Methods (Different BC — Chebyshev Basis Functions)

Method	Preconditioner	$N = 8$			$N = 12$			$N = 16$		
		Error	Iter.	Time	Error	Iter.	Time	Error	Iter.	Time
LU-factor.	—	4.38e-05	—	5.67e-04	1.45e-08	—	1.37e-03	1.15e-12	—	2.76e-03
Jacobi	—	4.38e-05	66	5.65e-03	1.45e-08	74	9.96e-03	1.16e-12	95	1.93e-02
GS	—	4.38e-05	30	3.16e-03	1.45e-08	34	5.23e-03	1.16e-12	36	8.28e-03
Bi-CGSTAB	Jacobi	4.38e-05	11	2.51e-03	1.45e-08	12	4.12e-03	1.16e-12	13	6.41e-03
	GS	4.38e-05	7	2.27e-03	1.45e-08	7	3.36e-03	1.15e-12	7	4.93e-03
GMRES(10)	Jacobi	4.38e-05	23	3.02e-03	1.45e-08	26	3.89e-03	1.16e-12	28	8.56e-03
	GS	4.38e-05	12	2.39e-03	1.45e-08	13	4.00e-03	1.16e-12	13	5.24e-03

References

- [1] A.S.Fokas, A unified transform method for solving linear and certain nonlinear PDEs, *Proc. R. Soc. London A***53** (1997), 1411-1443.
- [2] S. Fulton, A.S. Fokas and C. Xenophontos, An Analytical Method for Linear Elliptic PDEs and its Numerical Implementation, *J. of CAM* **167** (2004), 465-483.
- [3] A. Sifalakis, A.S. Fokas, S. Fulton and Y.G. Saridakis, The Generalized Dirichlet-Neumann Map for Linear Elliptic PDEs and its Numerical Implementation, *J. of Comput. and Appl. Maths.* (in press)
- [4] A.S.Fokas, Two-dimensional linear PDEs in a convex polygon, *Proc. R. Soc. London A***457** (2001), 371-393.
- [5] Y.G. Saridakis, A. Sifalakis and E.P. Papadopoulou, Efficient Solution of the Generalized Dirichlet-Neumann Map for Linear Elliptic PDEs in Regular Polygon Domains, (submitted)
- [6] H.A. Van Der Vorst, Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems, *SIAM J. Sci. Statist. Comput.*, 13,1992, pp. 631-644.
- [7] Y. Saad and M. Schultz, GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems, *SIAM J. Sci. Statist. Comput.*, 7,1986,pp. 856-869.

- [8] J. Dongarra, I. Duff, D. Sorensen, H. van der Vorst, *Numerical Linear Algebra for High-Performance Computers* , SIAM, 1998.
- [9] R.S. Varga, *Matrix Iterative Analysis* , Prentice-Hall, 1962.

Efficient Numerical Solution of the Generalized Dirichlet-Neumann Map for Linear Elliptic PDEs in Regular Polygon Domains

Y. G. SARIDAKIS, A. G. SIFALAKIS, E. P. PAPADOPOULOU

*Applied Math and Computers Lab, Dept of Sciences, Technical University of Crete,
73100 Chania, Greece*

[Received on 21 October 2007]

A new and novel approach for analyzing boundary value problems for linear and for integrable non-linear PDEs was recently introduced. For linear elliptic PDEs, an important aspect of this approach is the characterization of a generalized Dirichlet-Neumann map: given the derivative of the solution along a direction of an arbitrary angle to the boundary, the derivative of the solution perpendicularly to this direction is computed *without* solving on the interior of the domain. For this computation, a collocation-type numerical method has been recently developed. Here, we study the collocation's coefficient matrix properties. We prove that, for the Laplace's equation on regular polygon domains with the same type of boundary conditions on each side, the collocation matrix is Block Circulant, independently of the choice of basis functions. This leads to the deployment of the FFT for the solution of the associated collocation linear system, yielding significant computational savings. Numerical experiments are included to demonstrate the efficiency of the whole computation.

Keywords: Elliptic PDEs, Dirichlet-Neumann Map, Global Relation, Collocation, Block Circulant Matrices, FFT

1. Introduction

A new method for solving boundary value problems for linear and for integrable nonlinear PDEs in two dimensions was introduced by Fokas in 1 (1997), 2 (2001). This method involves two novel features:

- (a) It yields an analytic representation of the solution, in an integral form, in the complex k -plane.
- (b) It characterizes a generalized Dirichlet-Neumann map through the solution of the so-called *global relation*, an equation, valid for all complex values of k , which couples known and unknown components of the solution and its derivatives on the boundary.

For a large class of boundary value problems, the global relation can be solved analytically, and hence the generalized Dirichlet-Neumann map can be constructed in closed form. This includes linear evolution PDEs with spatial derivatives of arbitrary order on the half-line 4 (2002) and on a finite interval 5 (2005), the Laplace, the bi-harmonic and the modified Helmholtz equation in certain simple polygons 12 (2004), 6 (2001)-9 (2004), and the basic nonlinear integrable evolution PDEs on the half-line for certain simple boundary conditions 10 (2002)-11 (2004). However, for general boundary value problems, the global relation must be solved numerically.

In 12 (2004) and 13 (2007) a well conditioned and fast convergent collocation-type numerical method was developed and studied for the numerical solution of the Generalized Dirichlet-Neumann map associated to the generic model problem of Laplace's equation on an arbitrary convex polygon domain. The present work, which is a natural continuation of our work in 13 (2007), aims to the efficient solution of the associated collocation linear system. For the case of regular polygon domains, with the

same type of boundary conditions on each side, we prove, among other properties, that the Collocation coefficient matrix is *Block Circulant*, independently of the choice of basis functions. Evidently, therefore, by deploying the *Fourier Matrix*, the collocation matrix is transformed into a similar block diagonal matrix, the construction of which is implemented through the *Fast Fourier Transform* (FFT), yielding efficient implementation of direct and iterative methods for the solution of the collocation linear system.

This paper is organized as follows: Section 2 outlines some analytical results, as well as the collocation method of [13] (2007). Section 3 presents the structure and the properties of the collocation coefficient matrix for the case of regular polygons. And, finally, Section 4 presents the implementation through the FFT as well as numerical results, for a variety of boundary conditions, to demonstrate the efficiency of the whole computation.

2. Overview

For elliptic PDEs in two dimensions it is convenient to replace the Cartesian coordinates (x, y) with the complex coordinates $(z, \bar{z}) = (x + iy, x - iy)$. In doing so and using the equations

$$\partial_z = \frac{1}{2}(\partial_x - i\partial_y), \quad \partial_{\bar{z}} = \frac{1}{2}(\partial_x + i\partial_y),$$

Laplace's equation in the independent variable q can be written in the form

$$\frac{\partial^2 q}{\partial z \partial \bar{z}} = 0. \quad (2.1)$$

This equation is equivalent to the equation

$$\frac{\partial}{\partial \bar{z}} \left(e^{-ikz} \frac{\partial q}{\partial z} \right) = 0, \quad (2.2)$$

for an *arbitrary* complex parameter k .

The Global Relation

Suppose that the real-valued function $q(z, \bar{z})$ satisfies Laplace's equation (2.1) in a simply connected bounded domain D with boundary ∂D . Then, equation (2.2) together with the complex form of Green's theorem imply the equation

$$\oint_{\partial D} e^{-ikz} \frac{\partial q}{\partial z} dz = 0, \quad k \in \mathbb{C}, \quad (2.3)$$

which is referred to (cf. [2] (2001)) as the *global relation* associated with Laplace's equation. For the case of D being a convex bounded polygon with vertices $z_1, z_2, \dots, z_n$ (modulo n), which have indexed counter-clockwise, the global relation (2.3) becomes

$$\sum_{j=1}^n \int_{S_j} e^{-ikz} \frac{\partial q}{\partial z} dz = 0, \quad k \in \mathbb{C}, \quad (2.4)$$

where S_j denotes the side (z_j, z_{j+1}) .

The Generalized Dirichlet-Neumann Map

Using the identity

$$\frac{\partial q}{\partial z} = \frac{1}{2} e^{-i\alpha_j} \left(q_s^{(j)} + i q_n^{(j)} \right), \quad z \in S_j, \quad \alpha_j = \arg(z_{j+1} - z_j), \quad (2.5)$$

where $q_s^{(j)}$ and $q_n^{(j)}$ denote the tangential and (outward) normal components of $\frac{\partial q}{\partial z}$ along the side S_j , as well as the local coordinates

$$z = m_j + s h_j, \quad z_i < z < z_{i+1}, \quad -\pi < s < \pi, \quad (2.6)$$

with

$$m_j = \frac{1}{2} (z_j + z_{j+1}), \quad h_j = \frac{1}{2\pi} (z_{j+1} - z_j), \quad (2.7)$$

it was shown (e.g. 13 (2007)) that:

PROPOSITION 2.1 *Let the real-valued function $q(z, \bar{z})$ satisfy the Laplace equation in the interior D of the convex bounded polygon with corners $\{z_i\}_{i=1}^n$ described above. Let $g^{(j)}$ denote the derivative of the solution in the direction making an angle β_j , $0 \leq \beta_j \leq \pi$, with the side S_j , namely*

$$\cos(\beta_j) q_s^{(j)} + \sin(\beta_j) q_n^{(j)} = g^{(j)}, \quad z \in S_j, \quad 1 \leq j \leq n. \quad (2.8)$$

Let $f^{(j)}$ denote the derivative of the solution in the direction normal to the above direction, namely

$$-\sin(\beta_j) q_s^{(j)} + \cos(\beta_j) q_n^{(j)} = f^{(j)}, \quad z \in S_j, \quad 1 \leq j \leq n. \quad (2.9)$$

The generalized Dirichlet-Neumann map, that is the relation between the sets $\{f^{(j)}\}_{j=1}^n$ and $\{g^{(j)}\}_{j=1}^n$, is characterized by the single equation

$$\sum_{j=1}^n |h_j| e^{i(\beta_j - km_j)} \int_{-\pi}^{\pi} e^{-ikh_j s} \left(f^{(j)}(s) - i g^{(j)}(s) \right) ds = 0, \quad k \in \mathbb{C}. \quad (2.10)$$

Evaluating equation (2.10) on the following n -rays of the complex k -plane

$$k_p = -\frac{l}{h_p}, \quad l \in \mathbb{R}^+, \quad p = 1, \dots, n, \quad (2.11)$$

and multiplying the resulting equations by $\exp[-i(\beta_p + lm_p/h_p)] / |h_p|$, equation (2.10) yields the following set of n equations:

$$\sum_{j=1}^n \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-\frac{il}{h_p}(m_p - m_j)} \int_{-\pi}^{\pi} e^{il \frac{h_j}{h_p} s} \left(f^{(j)}(s) - i g^{(j)}(s) \right) ds = 0, \quad l \in \mathbb{R}^+, \quad p = 1, \dots, n. \quad (2.12)$$

Suppose, now, that the function set $\{g^{(j)}\}_{j=1}^n$ is known through appropriate boundary data. Then, it becomes apparent that the generalized Dirichlet-Neumann map, in its convenient form of (2.12), may

be used to determine the function set $\left\{f^{(j)}\right\}_{j=1}^n$. The end values of the unknown functions $f^{(j)}$ can be calculated by the continuity requirements $q_z^{(j)}(z_j) = q_z^{(j-1)}(z_j)$. Namely, rewriting equation (2.5) as

$$\frac{\partial q}{\partial z} = \frac{1}{2} e^{-i(\alpha_j - \beta_j)} \left(g^{(j)} + i f^{(j)} \right), \quad z \in S_j, \quad (2.13)$$

assuming that $g^{(j)}$ are compatible in the corners and setting $\delta_j = \alpha_j - \beta_j$ we obtain (cf 12 (2004))

$$f^{(j)}(\pi) = \frac{\cos(\delta_{j+1} - \delta_j) g^{(j)}(\pi) - g^{(j+1)}(\pi)}{\sin(\delta_{j+1} - \delta_j)},$$

$$f^{(j)}(-\pi) = \frac{g^{(j-1)}(\pi) - \cos(\delta_j - \delta_{j-1}) g^{(j)}(-\pi)}{\sin(\delta_j - \delta_{j-1})}.$$

The Numerical Method

The collocation-type numerical method, developed in 12 (2004)-13 (2007), for the determination of the function set $\left\{f^{(j)}\right\}_{j=1}^n$, is being described in the form of the following proposition (cf 13 (2007)):

PROPOSITION 2.2 *Consider the generalized Dirichlet-Neumann map in Proposition 2.1. Suppose that the set $\left\{g^{(j)}\right\}_{j=1}^n$ is given. Suppose that $f^{(j)}(s)$ is approximated by*

$$f_N^{(j)}(s) = f_*^{(j)}(s) + \sum_{r=1}^N U_r^j \varphi_r(s), \quad j = 1, \dots, n, \quad N \text{ even integer}, \quad (2.14)$$

where $\varphi_r(s)$ are appropriate basis functions, N is even and

$$f_*^{(j)}(s) = \frac{1}{2\pi} \left[(s + \pi) f^{(j)}(\pi) - (s - \pi) f^{(j)}(-\pi) \right].$$

Then, the real coefficients U_r^j satisfy the Nn algebraic set of equations

$$\sum_{j=1}^n \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-i \frac{l}{h_p} (m_p - m_j)} \sum_{r=1}^N U_r^j F_r \left(\frac{lh_j}{h_p} \right) = G_p(l), \quad p = 1, 2, \dots, n, \quad (2.15)$$

where $G_p(l)$ denotes the known function

$$G_p(l) = i \sum_{j=1}^n \frac{|h_j|}{|h_p|} e^{i(\beta_j - \beta_p)} e^{-i \frac{l}{h_p} (m_p - m_j)} \int_{-\pi}^{\pi} e^{il \frac{h_j}{h_p} s} (g^{(j)}(s) + i f_*^{(j)}(s)) ds, \quad (2.16)$$

$F_r(l)$ denotes the integral

$$F_r(l) = \int_{-\pi}^{\pi} e^{ils} \varphi_r(s) ds, \quad r = 1, 2, \dots, N, \quad (2.17)$$

and l is chosen as follows: For the imaginary part of equations (3.2) $l = 1, 2, \dots, N/2$, whereas for the real part of equations (3.2), $l = \frac{1}{2}, \frac{3}{2}, \dots, \frac{N-1}{2}$.

It is worthwhile to point out that different choices of the basis functions $\varphi_r(s)$ affect the convergence rate and the condition number of the method, as well as the structure of the coefficient matrix (cf 13 (2007) for a detailed treatment).

3. Matrix Properties for Regular Polygons

Let us now consider the regular polygon, depicted in Fig. 3.1, with vertices z_j and sides S_j , $j = 1, \dots, n$ (modulo n), indexed counter-clockwise, and interior $\mathbf{D}$. For notational simplicity and without any loss of the generality, we assume that the polygon is centered at the origin, scaled and oriented so that one vertex (say z_1) is located at 1, that is $z_1 = 1$. Setting $\theta = \frac{2\pi}{n}$, the vertices may be written as

$$z_j = \omega^{j-1}, \quad \omega = e^{i\theta}, \quad j = 1, \dots, n, \quad (3.1)$$

and the angle α_j of the side S_j from the real axis (measured counterclockwise) is given by

$$\alpha_j = \arg(z_{j+1} - z_j) = \frac{1}{2}[\pi + (2j-1)\theta] = \frac{\pi}{2} + (2j-1)\frac{\pi}{n}, \quad j = 1, \dots, n. \quad (3.2)$$

Suppose that the real-valued function $q(z, \bar{z})$ satisfies the Laplace's equation in the interior $\mathbf{D}$ of the regular n -gon, described above, subject to the same type of Poincaré boundary conditions on all sides, that is

$$\cos(\beta) q_s^{(j)} + \sin(\beta) q_n^{(j)} = g^{(j)}, \quad z \in S_j, \quad 1 \leq j \leq n. \quad (3.3)$$

Dirichlet and Neumann boundary conditions correspond to the special cases of $\beta = 0$ and $\beta = \frac{\pi}{2}$ respectively.

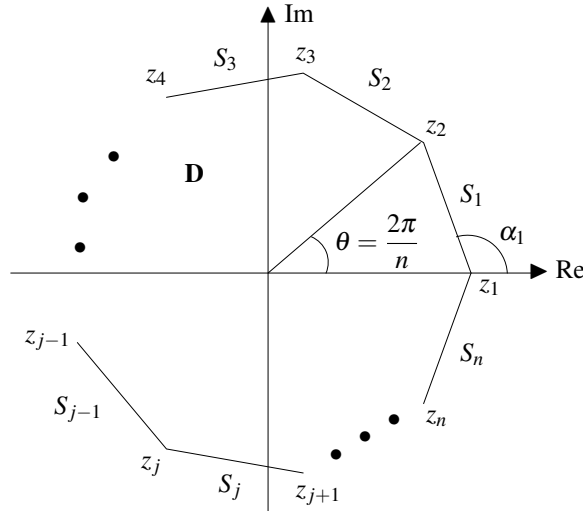


Figure 3.1 Regular n -gon with vertices z_j , sides S_j and interior $\mathbf{D}$.

Recalling, now, the local coordinates of (2.6) and observing that their parametrization of (2.7) is expressed as

$$m_j = \frac{1}{2} (z_j + z_{j+1}) = |m_j| e^{i(a_j - \frac{\pi}{2})} = \cos\left(\frac{\pi}{n}\right) e^{i(2j-1)\frac{\pi}{n}} = \cos\left(\frac{\pi}{n}\right) \omega^{(2j-1)/2}, \quad (3.4)$$

and

$$h_j = \frac{1}{2\pi} (z_{j+1} - z_j) = \frac{1}{2\pi} |h_j| e^{ia_j} = i \frac{1}{\pi} \sin\left(\frac{\pi}{n}\right) e^{i(2j-1)\frac{\pi}{n}} = i \frac{1}{\pi} \sin\left(\frac{\pi}{n}\right) \omega^{(2j-1)/2}, \quad (3.5)$$

it follows easily, from Proposition 2.1, that:

COROLLARY 3.1 *Let the real-valued function $q(z, \bar{z})$ satisfy the Laplace equation in the interior D of the regular n -gon with corners $\{z_i\}_{i=1}^n$ described above in this section. Let $g^{(j)}$, defined in (3.3), denote the derivative of the solution in the direction making an angle β , $0 \leq \beta \leq \pi$, with the side S_j , and let $f^{(j)}$ denote the derivative of the solution in the direction normal to the above direction. The generalized Dirichlet-Neumann map is characterized by the single equation*

$$\sum_{j=1}^n e^{-ikm_j} \int_{-\pi}^{\pi} e^{-ikh_j s} \left(f^{(j)}(s) - ig^{(j)}(s) \right) ds = 0, \quad k \in \mathbb{C}, \quad (3.6)$$

which, upon evaluation on the n -rays of the complex k -plane defined in (2.11), yields the following set of n equations:

$$\sum_{j=1}^n e^{l\phi\omega_{jp}} \int_{-\pi}^{\pi} e^{il\omega_{jp}s} \left(f^{(j)}(s) - ig^{(j)}(s) \right) ds = 0, \quad l \in \mathbb{R}^+, \quad p = 1, \dots, n \quad (3.7)$$

where

$$\phi = \pi \cot \frac{\pi}{n}, \quad \omega_{jp} = \omega^{j-p} \quad (3.8)$$

and $\omega = e^{i\theta}$ being as in (3.1).

Proof. The relations

$$\frac{m_p - m_j}{h_p} = i\phi(\omega_{jp} - 1) \quad \text{and} \quad \frac{h_j}{h_p} = \omega_{jp}, \quad (3.9)$$

together with the appropriate simplifications of constant quantities, relax equation (2.10) to (3.7) and the proof follows immediately. $\square$

As an immediate consequence of Corollary 3.1 and its combination with Proposition 2.2, one may readily obtain that:

COROLLARY 3.2 *Consider the generalized Dirichlet-Neumann map in Corollary 3.1. Suppose that the set $\{g^{(j)}\}_{j=1}^n$ is given through (3.3) and that the set $\{f^{(j)}\}_{j=1}^n$ is approximated by $\{f_N^{(j)}\}_{j=1}^n$ defined in (2.14). Then, the real coefficients U_r^j satisfy the Nn algebraic set of equations*

$$\sum_{j=1}^n e^{l\phi\omega_{jp}} \sum_{r=1}^N U_r^j F_r(l\omega_{jp}) = G_p(l), \quad p = 1, 2, \dots, n, \quad (3.10)$$

where $G_p(l)$ denotes the known function

$$G_p(l) = i \sum_{j=1}^n e^{l\phi\omega_{jp}} \int_{-\pi}^{\pi} e^{il\omega_{jp}s} (g^{(j)}(s) + if_*^{(j)}(s)) ds, \quad (3.11)$$

$F_r(l)$ denotes the integral

$$F_r(l) = \int_{-\pi}^{\pi} e^{ils} \varphi_r(s) ds, \quad r = 1, 2, \dots, N, \quad (3.12)$$

ϕ and ω_{jp} are as defined in (3.8), and l is chosen as in Proposition 2.2, namely: For the imaginary part of equations (3.10) $l = 1, 2, \dots, N/2$, whereas for the real part of equations (3.10) $l = \frac{1}{2}, \frac{3}{2}, \dots, \frac{N-1}{2}$.

System's Structure

Following Corollary 3.2, let $C(p, j, l, r)$ be defined by

$$C(p, j, l, r) = e^{l\phi\omega_{jp}} F_r(l\omega_{jp}) = e^{l\phi\omega_{jp}} \int_{-\pi}^{\pi} e^{il\omega_{jp}s} \varphi_r(s) ds, \quad (3.13)$$

with $C_R(p, j, l, r)$ and $C_I(p, j, l, r)$ to denote its real and imaginary parts respectively, that is

$$C_R(p, j, l, r) = \operatorname{Re}[C(p, j, l, r)] \quad \text{and} \quad C_I(p, j, l, r) = \operatorname{Im}[C(p, j, l, r)]. \quad (3.14)$$

Let, also, $C_{p,j} \in \mathbb{R}^{N,N}$ denote the real $N \times N$ matrix with elements

$$C_{p,j} = \{c_{2l,r}^{p,j}\}, \quad c_{2l,r}^{p,j} = \begin{cases} C_R(p, j, l, r) & , \quad l = \frac{1}{2}, \frac{3}{2}, \dots, \frac{N-1}{2} \\ C_I(p, j, l, r) & , \quad l = 1, 2, \dots, N/2 \end{cases}, \quad r = 1, 2, \dots, N. \quad (3.15)$$

If, now, $\mathbf{U}_j \in \mathbb{R}^{N,1}$ and $\mathbf{G}_p \in \mathbb{R}^{N,1}$ denote the real vectors

$$\mathbf{U}_j = \{U_r^j\}_{r=1}^N = [U_1^j \quad U_2^j \quad \dots \quad U_N^j]^T, \quad (3.16)$$

and

$$\mathbf{G}_j = \{G_\ell^p\}_{\ell=1}^N = [G_1^p \quad G_2^p \quad \dots \quad G_N^p]^T, \quad (3.17)$$

where

$$G_\ell^p = \begin{cases} G_R^p(l) & , \quad l = \frac{1}{2}, \frac{3}{2}, \dots, \frac{N-1}{2}, \\ G_I^p(l) & , \quad l = 1, 2, \dots, N/2, \end{cases}, \quad \ell = 2l, \quad (3.18)$$

with $G_R^p(l)$ and $G_I^p(l)$ to denote respectively the real and imaginary parts of $G_p(l)$ in (3.11), then it can be easily seen that:

PROPOSITION 3.1 *The linear system, described by equations (3.10)-(3.12) in Corollary 3.2, is given by*

$$\mathbf{C}\mathbf{U} = \mathbf{G}, \quad \mathbf{C} \in \mathbb{R}^{nN, nN}, \quad \mathbf{U}, \mathbf{G} \in \mathbb{R}^{nN}, \quad (3.19)$$

where

$$C = \begin{pmatrix} C_{1,1} & \cdots & C_{1,j} & \cdots & C_{1,n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ C_{p,1} & \cdots & C_{p,j} & \cdots & C_{p,n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ C_{n,1} & \cdots & C_{n,j} & \cdots & C_{n,n} \end{pmatrix}, \quad U = \begin{pmatrix} U_1 \\ \vdots \\ U_j \\ \vdots \\ U_n \end{pmatrix}, \quad G = \begin{pmatrix} G_1 \\ \vdots \\ G_p \\ \vdots \\ G_n \end{pmatrix}, \quad (3.20)$$

with $C_{p,j}$, U_j and G_p are as defined in (3.15)-(3.17).

Proof. Recall (3.13)-(3.14) and observe that the set of equations in (3.10) is written as

$$\sum_{j=1}^n \sum_{r=1}^N C(p, j, l, r) U_r^j = G_p(l), \quad p = 1, 2, \dots, n \quad (3.21)$$

or, equivalently, as

$$\sum_{j=1}^n \sum_{r=1}^N C_R(p, j, l, r) U_r^j = G_R^p(l) \quad \text{and} \quad \sum_{j=1}^n \sum_{r=1}^N C_I(p, j, l, r) U_r^j = G_I^p(l). \quad (3.22)$$

The above set of equations, by using (3.15) and (3.18), is expressed as

$$\sum_{j=1}^n \sum_{r=1}^N c_{2l,r}^{p,j} U_r^j = G_{2l}^p, \quad p = 1, 2, \dots, n \quad (3.23)$$

and, by letting $l = \frac{1}{2}, 1, \frac{3}{2}, 2, \dots, \frac{N-1}{2}, \frac{N}{2}$, as

$$\sum_{j=1}^n C_{p,j} U_j = G_p, \quad p = 1, 2, \dots, n, \quad (3.24)$$

which completes the proof. $\square$

Coefficient Matrix Properties

To reveal, now, the properties of the coefficient matrix C of (3.20) we first prove that :

PROPOSITION 3.2 *The coefficient matrix C , defined in relations (3.19)-(3.20) of Proposition 3.3, is Block Circulant. Namely,*

$$C = \text{bcirc}\{C_1, C_2, \dots, C_n\} = \begin{pmatrix} C_1 & C_2 & C_3 & \cdots & C_{n-1} & C_n \\ C_n & C_1 & C_2 & \cdots & C_{n-2} & C_{n-1} \\ C_{n-1} & C_n & C_1 & \cdots & C_{n-3} & C_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ C_3 & C_4 & C_5 & \cdots & C_1 & C_2 \\ C_2 & C_3 & C_4 & \cdots & C_n & C_1 \end{pmatrix}, \quad (3.25)$$

where

$$C_j = C_{1,j}, \quad j = 1, \dots, n \quad (3.26)$$

and $C_{1,j}$ are as defined in (3.15).

Proof. It suffices to prove that, for any $\mu = 1, \dots, n-1$, the matrices $C_{p,j}$ ($p, j = 1, \dots, n$) of (3.15) satisfy

$$C_{p,j} = C_{p',j'}, \quad (3.27)$$

where

$$p' = \begin{cases} p + \mu & , \quad 1 \leq \mu \leq n-p \\ p + \mu - n & , \quad n-p+1 \leq \mu \leq n-1 \end{cases} \quad (3.28)$$

and

$$j' = \begin{cases} j + \mu & , \quad 1 \leq \mu \leq n-j \\ j + \mu - n & , \quad n-j+1 \leq \mu \leq n-1 \end{cases}. \quad (3.29)$$

For this, recall $C(p, j, r, l)$ from (3.13) and observe that, for fixed r and l , there holds

$$C(p', j', r, l) = C(p, j, r, l) \quad (3.30)$$

since, from (3.1) and (3.8), $\omega^{\pm n} = 1$ and

$$\omega_{j'p'} = \omega^{j'-p'} = \omega^{j-p} = \omega_{jp}. \quad (3.31)$$

Evidently, therefore, for fixed r and l , we obtain, from (3.15), that

$$c_{2l,r}^{p,j} = c_{2l,r}^{p',j'}, \quad (3.32)$$

and the proof follows. $\square$

We know (e.g. 14 (1983), 15 (1979)) that the block circulant property of the coefficient matrix C allows its block diagonal factorization. To be specific, let $F_n \in \mathbb{C}^{n,n}$ denote the *Fourier Matrix*, that is

$$F_n = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega & \omega^2 & \dots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \dots & \omega^{2(n-1)} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \dots & \omega^{(n-1)(n-1)} \end{pmatrix}, \quad (3.33)$$

where $\omega = e^{i2\pi/n}$ is as in (3.1). Let also $P_n \in \mathbb{R}^{n,n}$ denote the cyclic permutation matrix

$$P_n = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \end{pmatrix} = \begin{pmatrix} O_{n-1,1} & I_{n-1} \\ 1 & O_{1,n-1} \end{pmatrix}, \quad (3.34)$$

where I_q denotes the $q \times q$ identity matrix and $O_{p,q}$ the $p \times q$ null matrix. It is, also, well known that, among other celebrated properties,

$$F_n^{-1} = \frac{1}{n} F_n^*, \quad (3.35)$$

where F_n^* denotes the *conjugate transpose*, while

$$P_n^n = I_n \quad \text{and} \quad P_n^\ell = \begin{pmatrix} O_{n-\ell, \ell} & I_{n-\ell} \\ I_\ell & O_{\ell, n-\ell} \end{pmatrix} \quad \text{for } \ell = 1, \dots, n-1, \quad (3.36)$$

and, furthermore,

$$P_n = F_n \Omega_n F_n^{-1}, \quad \Omega_n = \text{diag}(1, \omega, \omega^2, \dots, \omega^{n-1}). \quad (3.37)$$

If we now let $\otimes$ and $\oplus$ denote the Kronecker product and direct sum of matrices, respectively, then:

PROPOSITION 3.3 *The block circulant matrix C , defined in (3.25) of Proposition 3.2, is expressed as*

$$C = (F_n \otimes I_N) \left(\bigoplus_{\ell=1}^n \left(\sum_{j=1}^n \omega^{(\ell-1)(j-1)} C_j \right) \right) (F_n^{-1} \otimes I_N). \quad (3.38)$$

Proof. Observing that

$$\begin{aligned} C &= \sum_{j=1}^n (P_n^{j-1} \otimes I_N) \left(\bigoplus_{\ell=1}^n C_j \right) \\ &= \sum_{j=1}^n ((F_n \Omega_n^{j-1} F_n^{-1}) \otimes I_N) \left(\bigoplus_{\ell=1}^n C_j \right) \\ &= \sum_{j=1}^n (F_n \otimes I_N) (\Omega_n^{j-1} \otimes I_N) (F_n^{-1} \otimes I_N) \left(\bigoplus_{\ell=1}^n C_j \right) \\ &= (F_n \otimes I_N) \left(\sum_{j=1}^n (\Omega_n^{j-1} \otimes I_N) \left(\bigoplus_{\ell=1}^n C_j \right) \right) (F_n^{-1} \otimes I_N) \\ &= (F_n \otimes I_N) \left(\sum_{j=1}^n \left(\bigoplus_{\ell=1}^n \omega^{(\ell-1)(j-1)} C_j \right) \right) (F_n^{-1} \otimes I_N) \end{aligned}$$

the proof follows. □

At this point we would like to conclude this section with the following remarks:

REMARK 3.1 For the case of regular polygons with the same type boundary conditions on all sides, the coefficient matrix C of (3.25) is *independent* of the angle $\beta_j = \beta$ the derivative of the solution is making with the side S_j . Namely, C is *independent* of the type (i.e. Dirichlet, Neumann or Mixed) of *Boundary Conditions*.

REMARK 3.2 The coefficient matrix properties revealed in this section are *independent* of the choice of *basis functions* and the number of *polygon sides*. However, certain choices of basis functions may enrich the properties of the coefficient matrix. For example, for the natural choice of *sine* basis functions, i.e.

$$\varphi_r(s) = \sin \left[r \left(\frac{\pi + s}{2} \right) \right], \quad (3.39)$$

studied in 13 (2007), and in addition to the property that the block diagonal submatrices of C are point diagonal, that is (cf 13 (2007))

$$C_1 = \pi \text{diag}(1, -1, -1, 1, \dots, (-1)^{N-1}, (-1)^N), \quad (3.40)$$

it takes only a few algebraic manipulations to verify that (cf 16 (2007)), for regular polygons with the same type of boundary conditions on all sides, there also holds

$$C_{n-j+2} = DC_j D \quad , \quad D = \text{diag}(1, -1, \dots, 1, -1) \quad , \quad j = 2, \dots, \hat{n} \quad , \quad (3.41)$$

where $\hat{n} = n/2$ if $n = \text{even}$ while $\hat{n} = (n+1)/2$ if $n = \text{odd}$. Similarly, if one considers the case of a *square* domain, with the same type of boundary conditions on all sides, then, independently of the choice of basis functions, one may easily verify (cf 17 (2007)) that

$$C_2 = C_4 = O \quad \text{and} \quad C_3 = \hat{D}C_1 \quad , \quad \hat{D} = \text{diag}(d_1, \dots, d_N) \quad , \quad d_r = (-1)^{r-1} e^{-r\pi} \quad , \quad r = 1, \dots, N \quad . \quad (3.42)$$

4. Implementation & Numerical Verification

Having in mind that the cost for solving the linear system in (3.19), by the classical *LU*-factorization, is $\mathcal{O}(n^3 N^3)$, in this section we are aiming at the efficient implementation of the results in Propositions 3.2 and 3.3 and its numerical verification. For this purpose, recall the relation (3.38) and observe that, upon substitution in (3.19), the linear system at hand may be written as

$$A\mathbf{u} = \mathbf{g} \quad , \quad (4.1)$$

where

$$A = b\text{diag}(A_1, \dots, A_n) = \bigoplus_{\ell=1}^n A_\ell \quad , \quad A_\ell = \left(\sum_{j=1}^n \omega^{(\ell-1)(j-1)} C_j \right) \quad , \quad (4.2)$$

$$\mathbf{u} = (F_n^{-1} \otimes I_N) \mathbf{U} \quad , \quad (4.3)$$

and

$$\mathbf{g} = (F_n^{-1} \otimes I_N) \mathbf{G} \quad , \quad (4.4)$$

with C_j , $\mathbf{U}$ & $\mathbf{G}$ and F_n to be as defined in (3.26), (3.20) and (3.33), respectively. The cost of the solution, by the classical *LU*, of the equivalent block diagonal system in (4.1) has already been reduced to $\mathcal{O}(nN^3)$, implying substantial savings especially for regular polygons with a large number of sides. Of course, in addition to this cost, one has to consider also the cost to convert the system from its form in (3.19) into the form of (4.1). And in order to enjoy the said cost savings, this conversion has to be efficient. In this direction, aiming at the efficient deployment of the FFT, let us first observe that

$$\begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix} = (F_n \otimes I_N) \begin{pmatrix} C_1 \\ \vdots \\ C_n \end{pmatrix} \quad , \quad (4.5)$$

and prove that:

LEMMA 4.1 *Given the Fourier matrix in (3.33) there holds*

$$E(F_n \otimes I_N)E^T = \bigoplus_{r=1}^N F_n \quad , \quad (4.6)$$

where $E \in \mathbb{R}^{nN, nN}$ denotes the permutation matrix

$$E = \begin{pmatrix} E_1 \\ \vdots \\ E_r \\ \vdots \\ E_N \end{pmatrix}, \quad E_r = I_n \otimes \mathbf{e}_r^T \in \mathbb{R}^{n, nN}, \quad r = 1, \dots, N, \quad j = 1, \dots, n, \quad (4.7)$$

with $\mathbf{e}_r$ to denote the r -th unit vector of $\mathbb{R}^{N,1}$.

Proof. Simply observe that

$$\begin{aligned} E(F_n \otimes I_N)E^T &= \begin{pmatrix} I_n \otimes \mathbf{e}_1^T \\ \vdots \\ I_n \otimes \mathbf{e}_N^T \end{pmatrix} (F_n \otimes I_N) \begin{pmatrix} I_n \otimes \mathbf{e}_1 & \cdots & I_n \otimes \mathbf{e}_N \end{pmatrix} \\ &= \begin{pmatrix} I_n \otimes \mathbf{e}_1^T \\ \vdots \\ I_n \otimes \mathbf{e}_N^T \end{pmatrix} \begin{pmatrix} F_n \otimes \mathbf{e}_1 & \cdots & F_n \otimes \mathbf{e}_N \end{pmatrix} \\ &= \begin{pmatrix} F_n & O & \cdots & O \\ O & F_n & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & F_n \end{pmatrix}. \end{aligned}$$

□

Evidently, therefore, we can write

$$\begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix} = E^T \left(\bigoplus_{r=1}^N F_n \right) E \begin{pmatrix} C_1 \\ \vdots \\ C_n \end{pmatrix}, \quad (4.8)$$

$$\mathbf{g} = E^T \left(\bigoplus_{r=1}^N F_n^{-1} \right) E \mathbf{G}, \quad (4.9)$$

and

$$\mathbf{U} = E^T \left(\bigoplus_{r=1}^N F_n \right) E \mathbf{u}. \quad (4.10)$$

Hence, the cost of computing the vectors $\mathbf{g}$ or $\mathbf{U}$ is equivalent to the cost of applying N independent (parallel) IFFT/FFTs of order n , that is $\mathcal{O}(Nn \log n)$. Similarly, the cost of computing the matrices A_ℓ , $\ell = 1, \dots, n$ is equivalent to the cost of performing N^2 independent (parallel) FFTs of order n , that is $\mathcal{O}(N^2 n \log n)$. Thus, by following relations (4.8)-(4.10), the total cost of constructing and solving the system in (4.1) is $\mathcal{O}(N^3 n + N^2 n \log n)$.

To demonstrate the efficiency of our implementation we consider the solution of the model Laplace's equation, with exact solution (cf. 12 (2004)-13 (2007))

$$q(x, y) = \sinh(3x) \sin(3y), \quad (4.11)$$

by the direct *LU*-factorization method and, two main representatives of the Krylov subspace iterative methods, the BiCGSTAB and the GMRES(10) methods. The basis functions used are the sine functions defined in (3.39), and the relative error E_∞ , used to demonstrate the convergence behavior of the methods considered, is given by

$$E_\infty = \frac{\|f - f_N\|_\infty}{\|f\|_\infty}, \quad (4.12)$$

where

$$\|f\|_\infty = \max_{1 \leq j \leq n} \left\{ \max_{-\pi \leq s \leq \pi} |f^{(j)}(s)| \right\} \quad (4.13)$$

and

$$\|f - f_N\|_\infty = \max_{1 \leq j \leq n} \left\{ \max_{-\pi \leq s \leq \pi} |f^{(j)}(s) - f_N^{(j)}(s)| \right\}, \quad (4.14)$$

with $f_N^{(j)}$ as in (2.14), and the max over s is taken over a dense discretization of the interval $[-\pi, \pi]$. For the direct solution of the linear systems we have used the standard LAPACK routines, while for the computation of the right hand side vector we have used a routine (*dqawo*) from QUADPACK implementing the modified Clenshaw-Curtis technique. We have considered the un-preconditioned forms of the both BiCGSTAB and GMRES iterative methods. However, we point out that, as we are using the sine basis functions, the diagonal blocks C_1 , defined in (3.26), are point diagonal matrices (cf 13 (2007)), hence the coefficient matrix C of (3.25) may be considered as *block Jacobi preconditioned*. The maximum number of iterations, allowed for all iterative methods to perform, is set to 200 and the zero iterate $U^{(0)}$ is set to be equal to the right hand side vector. The results we have included refer to the representative cases of regular polygons with 8, 16, 24, 32, 40 and 48 vertices. All polygons are constructed as in 12 (2004). All experiments were conducted on a multiuser SUN V240 system using the Fortran-90 compiler.

Table 4.1 Regular 8-gon ($n = 8$)

Method	$N = 16$				$N = 128$			
	Regular		FFT		Regular		FFT	
	E_∞	Time	E_∞	Time	E_∞	Time	E_∞	Time
LU	4.93E-04	0.023	4.93E-04	0.002	8.22E-06	25	8.22E-06	0.345
BiCGSTAB	4.93E-04	0.004	4.93E-04	0.006	8.22E-06	0.226	8.15E-06	0.233
GMRES	4.92E-04	0.005	4.93E-04	0.006	8.74E-06	0.227	8.09E-06	0.269

Table 4.2 Regular 16-gon ($n = 16$)

Method	$N = 16$				$N = 128$			
	Regular		FFT		Regular		FFT	
	E_∞	Time	E_∞	Time	E_∞	Time	E_∞	Time
LU	2.71E-04	0.184	2.71E-04	0.004	4.46E-06	470	4.46E-06	0.742
BiCGSTAB	2.70E-04	0.013	2.71E-04	0.011	5.46E-06	0.744	3.91E-06	0.576
GMRES	2.70E-04	0.014	2.71E-04	0.014	5.82E-06	0.763	4.62E-06	0.837

Table 4.3 Regular 24-gon ($n = 24$)

Method	$N = 16$				$N = 128$			
	Regular		FFT		Regular		FFT	
	E_∞	Time	E_∞	Time	E_∞	Time	E_∞	Time
LU	1.40E-04	0.622	1.40E-04	0.006	2.23E-06	1540	2.23E-06	1.150
BiCGSTAB	1.41E-04	0.024	1.40E-04	0.018	5.72E-06	1.300	8.86E-06	0.970
GMRES	1.41E-04	0.032	1.40E-04	0.028	3.57E-06	1.510	2.23E-06	2.370

Table 4.4 Regular 32-gon ($n = 32$)

Method	$N = 16$				$N = 128$			
	Regular		FFT		Regular		FFT	
	E_∞	Time	E_∞	Time	E_∞	Time	E_∞	Time
LU	8.68E-05	1.490	8.68E-05	0.010	1.34E-06	4390	1.34E-06	1.570
BiCGSTAB	8.67E-05	0.060	8.69E-05	0.020	1.30E-06	3.160	2.08E-06	1.290
GMRES	8.74E-05	0.110	8.67E-05	0.080	1.88E-06	6.880	1.50E-06	4.980

Table 4.5 Regular 40-gon ($n = 40$)

Method	$N = 16$				$N = 128$			
	Regular		FFT		Regular		FFT	
	E_∞	Time	E_∞	Time	E_∞	Time	E_∞	Time
LU	6.05E-05	3.090	6.05E-05	0.010	8.73E-07	7560	8.73E-07	1.990
BiCGSTAB	6.02E-05	0.110	6.08E-05	0.030	8.62E-07	7.130	8.92E-07	1.860
GMRES	3.27E-04	0.310	6.08E-05	0.140	2.35E-03	12.50	1.97E-03	7.660

Table 4.6 Regular 48-gon ($n = 48$)

Method	$N = 16$				$N = 128$			
	Regular		FFT		Regular		FFT	
	E_∞	Time	E_∞	Time	E_∞	Time	E_∞	Time
LU	5.09E-05	9.180	5.09E-05	0.010	6.69E-07	13900	6.69E-07	2.610
BiCGSTAB	4.32E-05	0.190	5.09E-05	0.040	7.73E-07	12.80	7.34E-07	2.500
GMRES	1.06E-02	0.380	2.50E-03	0.220	1.27E-02	15.50	1.32E-02	10.70

Inspecting Tables 4.1-4.6 and Figs 1-3 it can be readily verified that:

- Applying the FFT formulation, all direct and iterative methods perform significantly faster, especially the direct solution method (*LU*-factorization, Fig 1).
- The unpreconditioned BiCGSTAB iterative method outperforms the upreconditioned GMRES(10), while GMRES fails to converge within 200 iterations for polygons with medium to large number of edges.
- Iterative methods compete with direct only for large number of basis functions (Fig 3)

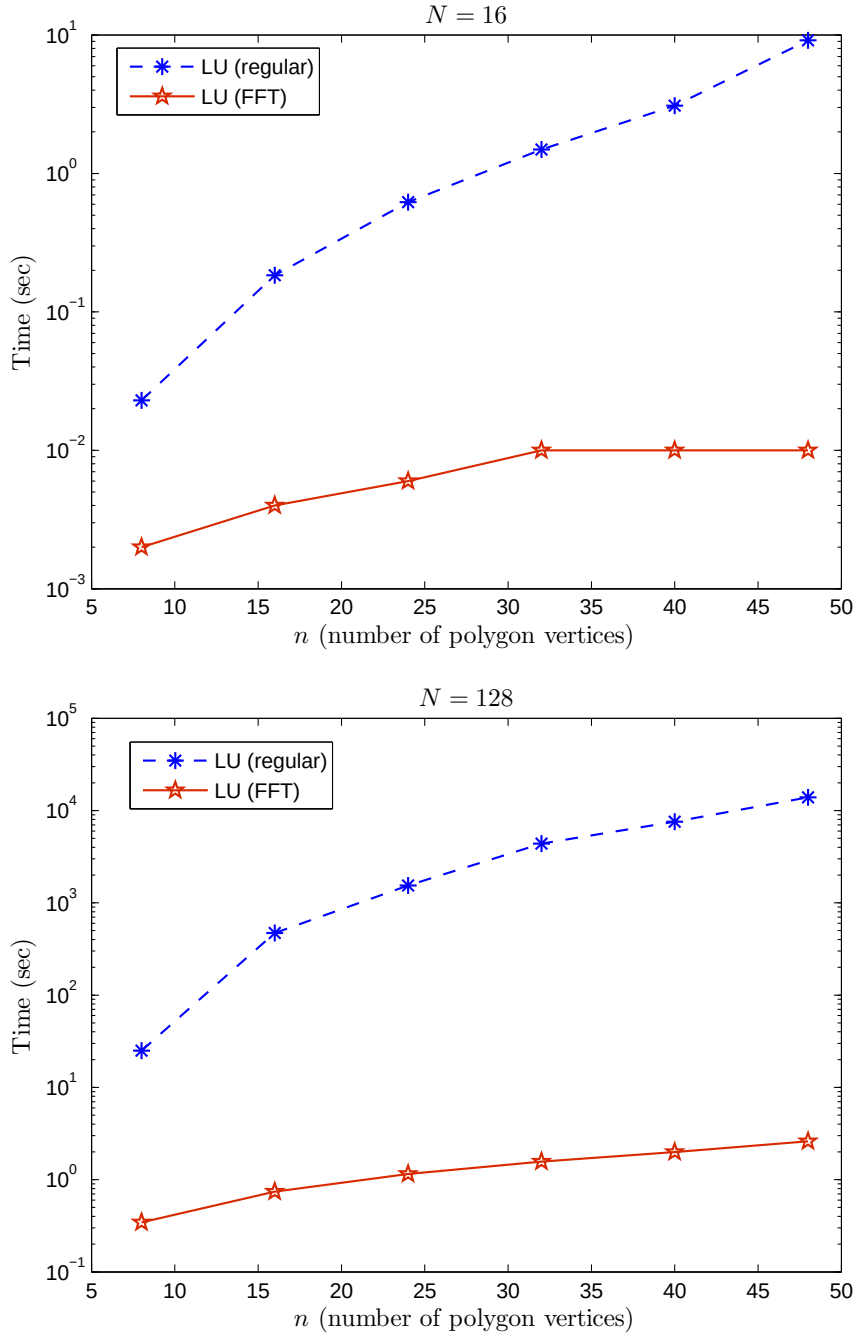


FIG. 1. Performance of the LU -factorization with (FFT) and without (regular) applying the FFT formulation (cases of $N = 16$ and $N = 128$).

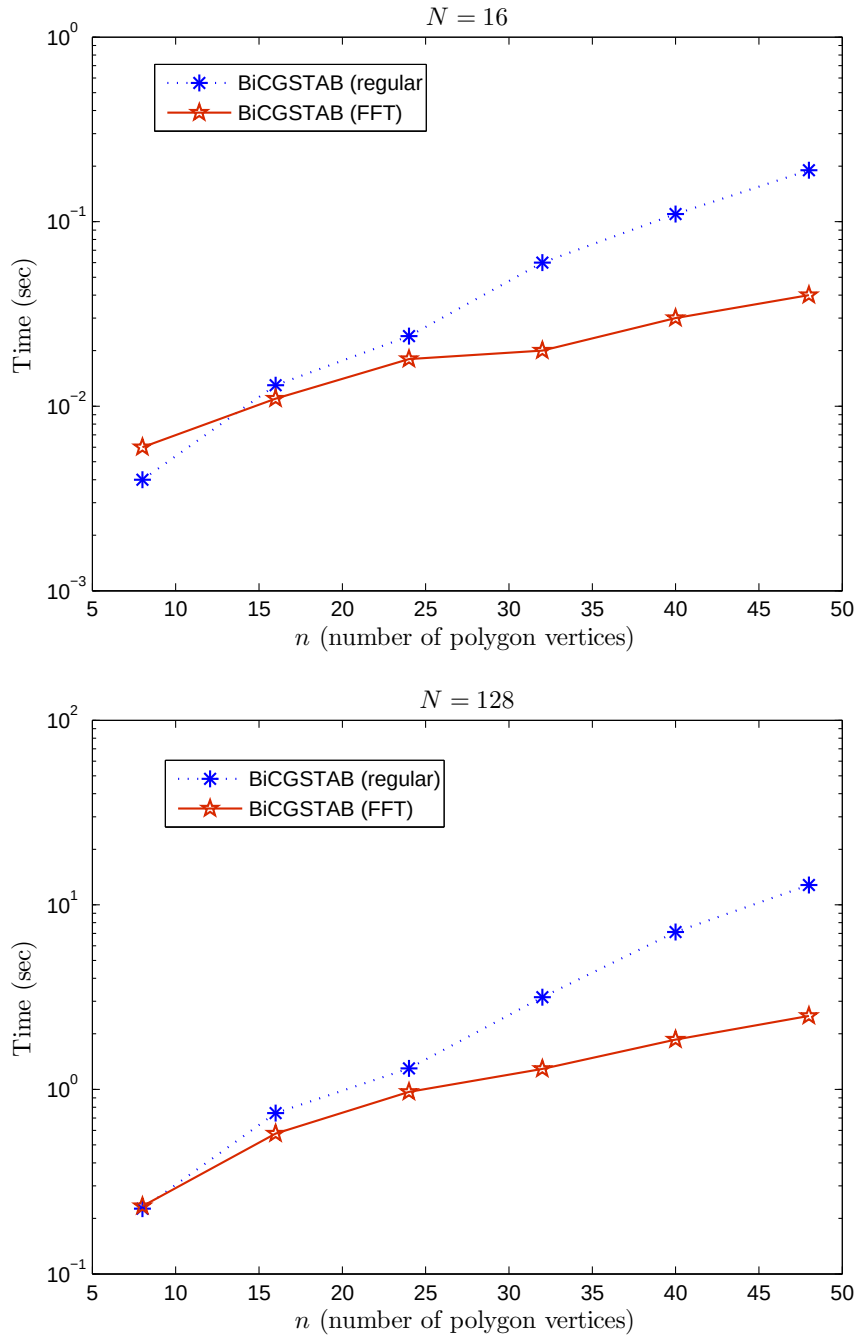
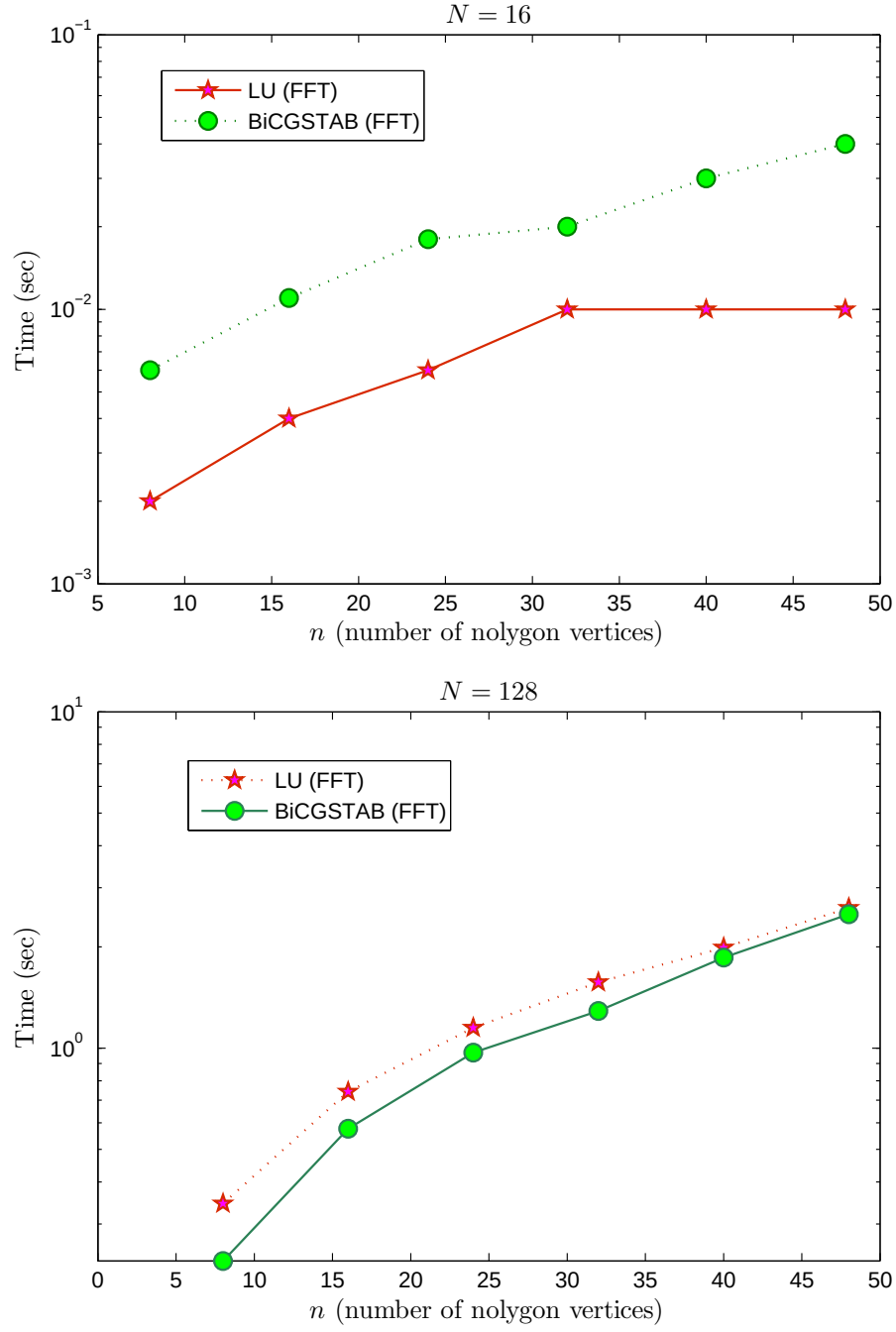


FIG. 2. Performance of the *BiCGSTAB* method with (FFT) and without (regular) applying the FFT formulation (cases of $N = 16$ and $N = 128$).

FIG. 3. Comparison of LU and BiCGSTAB methods, both using the FFT formulation, for the cases of $N = 16$ and $N = 128$.

5. Conclusions and Remarks

We have studied the properties of coefficient matrix of the collocation linear system associated with the Dirichlet-Neumann map for linear PDEs in regular polygon domains with the same type of boundary conditions. It was shown (Section 3) that, independently of the choice of basis functions, the matrix is *block circulant* and *independent of the boundary conditions*. These properties allowed the use of the FFT (Section 4), for the efficient solution of the collocation linear system, with significant improvement of the performance for both direct and iterative methods considered.

The set of matrix properties may be expanded if one considers certain choices of basis functions (see Remark 3.2). Finally, we note that an investigation for the matrix properties for other cases of regular and irregular polygons has already been undertaken.

ACKNOWLEDGEMENT : This work was supported by the Greek Ministry's of Education EPEAEK-Herakleitos grand which is partially funded by the EU. The authors wish to thank profs. A.S. Fokas and S.R. Fulton for many useful discussions.

REFERENCES

- FOKAS, A.S. (1997) A unified transform method for solving linear and certain nonlinear PDEs, *Proc. R. Soc. London A*, **53**, 1411–1443.
- FOKAS, A.S. (2001) Two-dimensional linear PDEs in a convex polygon, *Proc. R. Soc. London A* **457**, 371–393.
- FOKAS, A.S. AND KAPAEV, A.A. (2003) On a Transform Method for the Laplace Equation in a Polygon, *IMA J. Appl. Math.* **68**, 355–408.
- FOKAS, A.S. (2002) A New Transform Method for Evolution PDEs, *IMA J. Appl. Math.* **67**, 559–590.
- FOKAS, A.S. AND PELLONI, B. (2005) A Transform Method for Evolution PDEs on the Interval, *IMA J. Appl. Math.*, 1–24.
- BEN-AVRAHAM, D. AND FOKAS, A.S. (2001) The Modified Helmholtz Equation in a Triangular Domain and an Application to Diffusion-limited Coalescence, *Phys. Rev. E* **64**, 016114–6.
- ANTIPOV, Y.A. AND FOKAS, A.S. (2005) A Transform Method for the Modified Helmholtz Equation in a Semi-Strip, *Math. Proc. Camb. Phil. Soc.* **138**, 339–365.
- DASSIOS, G. AND FOKAS, A.S. (2005) The Basic Elliptic Equations in an Equilateral Triangle, *Proc. R. Soc. Lond. A* **461**, 2721–2748.
- CROWDY, D. AND FOKAS, A.S. (2004) Explicit Integral Solutions for the Plane Elastostatic Semi-Strip, *Proc. R. Soc. Lond.* **460**, 1285–1309.
- FOKAS, A.S. (2002) Integrable Nonlinear Evolution Equations on the Half-Line, *Comm. Math. Phys.* **230**, 1–39.
- FOKAS, A.S. (2004) Linearizable Initial Boundary Value Problems for the sine-Gordon Equation on the Half-Line, *Nonlinearity* **17**, 1521–1534.
- FULTON, S., FOKAS, A.S. AND XENOPHONTOS, C. (2004) An Analytical Method for Linear Elliptic PDEs and its Numerical Implementation, *J. of Comput. and Appl. Maths.* **167**, 465–483.
- SIFALAKIS, A.G., FOKAS, A.S., FULTON, S.R. AND SARIDAKIS, Y.G. The Generalized Dirichlet-Neumann Map for Linear Elliptic PDEs and its Numerical Implementation, *J. of Comput. and Appl. Maths.* (in press).
- DE MAZANCOURT, T. AND GERLIC, D. (1983) The inverse of a block circulant matrix, *IEEE Trans Ant. Prop.* **AP-31**, 808–810.
- DAVIS, P.J. (1979) Circulant Matrices, *Pure & Applied Mathematics*, John Wiley & Sons, Inc.
- SIFALAKIS, A.G. (2007) Iterative Methods and Grid Computing for the Innovative Solution of Elliptic PDEs, *PhD Thesis*, Technical University of Crete, (in Greek).
- SIFALAKIS, A.G., FULTON, S.R., PAPADOPOULOU, E.P. AND SARIDAKIS, Y.G. (2007) On the Iterative Analysis of the Generalized Dirichlet-Neumann Map for Linear Elliptic PDEs, *Proc. Int. Conf. in Num. Anal. (NUMAN07)*, 2007, Kalamata, Greece .