NATIONAL TECHNICAL UNIVERSITY OF ATHENS
School of Electrical and Computer Engineering
Division of Electromagnetics, Electrooptics and Electronic Materials

Oblique scattering/hybrid wave propagation in elliptical homogeneous configurations & spectrum/preconditioning of the singular domain integral equation

DOCTORAL DISSERTATION

GRIGORIOS P. ZOUROS
National Technical University of Athens

SUPERVISOR: JOHN A. ROUMELIOTIS
Professor / National Technical University of Athens

CO-SUPERVISOR: NEIL V. BUDKO
Assistant professor / Delft University of Technology

ATHENS, GREECE
2012

NATIONAL TECHNICAL UNIVERSITY OF ATHENS
School of Electrical and Computer Engineering
Division of Electromagnetics, Electrooptics and Electronic Materials

Oblique scattering/hybrid wave propagation in elliptical homogeneous configurations & spectrum/preconditioning of the singular domain integral equation

DOCTORAL DISSERTATION

GRIGORIOS P. ZOUROS
National Technical University of Athens

Three-member doctoral scientific committee:
John A. Roumeliotis
Neil V. Budko
John L. Tsalamengas

Approved by the seven-member doctoral scientific committee. Athens, 7th of June 2012.

John A. Roumeliotis *Professor / N.T.U.A.*	Neil V. Budko *Assistant professor / T.U.D.*	John L. Tsalamengas *Professor / N.T.U.A.*
Elias N. Glytsis *Professor / N.T.U.A.*	John L. Vomvoridis *Professor / N.T.U.A.*	Kyriakos K. Hizanidis *Professor / N.T.U.A.*
George J. Fikioris *Assistant professor / N.T.U.A.*		

GRIGORIOS P. ZOUROS

Electrical and Computer Engineering degree,
National Technical University of Athens, Greece.

Educational and Technological Electronic Engineering degree,
School of Educational and Technological Engineering of Athens, Greece.

This dissertation was set in Times® Ten
and Times® Ten Greek (Linotype Originals),
and in *MathTıme*™ *Professional* by Grigorios P. Zouros.
E-mail address: zouros@mail.ntua.gr, zouros@ieee.org
Website: https://sites.google.com/site/grigorioszouros

Times is a trademark of Linotype Corp.
Linotype is a trademark of Linotype GmbH.
MathTıme is a trademark of Publish or Perish, Inc.

All numerical results in this dissertation were obtained with the use of MATLAB®.
MATLAB is a trademark of The MathWorks, Inc.

An online version of this dissertation can be found on the website listed above.

This dissertation is printed on acid-free paper. ♾

The General Assembly of the 24th of January 2012 of the School of E.C.E. of N.T.U.A.,
approved the writing of this dissertation in english.

Preface

Most methods applied in electromagnetics nowadays are mainly based on numerics, since the classical analytical techniques are unable to cover complex configurations. However, the latter are respected from the analytical point of view, and they constitute reference points for the validation of the numerical methods. In Duffy's book [19] there is a nice table in his introduction, depicting the relationship between the various techniques. This doctoral dissertation covers three out of four areas: separation of variables, asymptotic analysis, and numerical techniques. The first two were completed during my thesis at National Technical University of Athens, while the third was completed when I was a visiting researcher at Delft Institute of Applied Mathematics, at Delft University of Technology.

The content of this doctoral dissertation is validated by the journal publications [84, 85, 86].

Abstract

The topic of this doctoral dissertation is oriented on two directions: the first consists of exact techniques whose power is based on analytical expressions. More specifically, two classical problems – one in electromagnetic scattering and one in electromagnetic wave propagation – are treated with the use of elliptical wavefunctions. The second direction is oriented toward numerics, where the singular domain integral equation is studied. Its great advantage relies on understanding the physical point of view, with applications like photonics, nanophysics, biological imaging, fiber optics, etc.

The electromagnetic scattering by an arbitrarily oriented elliptical cylinder having different constitutive parameters to those of the background medium, is treated in the first chapter. The separation of variables method is used to solve this prob-

lem, but, due to the oblique incidence of the source fields, hybrid waves for the scattered and the induced fields are generated, thus making the formulation complicated. Moreover, because of the different wavenumbers between the scatterer and the background medium, the orthogonality relations for Mathieu functions do not hold, leading to more complicated systems, compared to those of normal incidence, that should be solved in order to get the solution for the scattered or the induced field. The validation of the results reveals the high accuracy of the implementation, even for electrically large scatterers. Both polarizations are considered and numerical results are given for various values of the parameters. The method is exact and can be used for reference as an alternative validation for future methods involving scattering problems.

In the second chapter the cutoff wavenumbers of the elliptical dielectric waveguide are calculated exactly and analytically. Two separate methods are used to solve this problem. The first method is based on the separation of variables technique using Mathieu functions and gives the exact cutoff wavenumbers. The system matrices of which the roots of their determinant should be determined are complicated because of the nonexistence of orthogonality relations for Mathieu functions, due to the different constitutive parameters between the core and the cladding of the fiber. In the second method the cutoff wavenumbers are obtained through analytical expressions, when the eccentricity h of the elliptical core is specialized to small values. In the latter case, analytical, closed-form algebraic expressions, free of Mathieu functions, are obtained for the expansion coefficients $g_{sn}^{(2)}$ in the resulting relation $x_{1,sn}(h) = x_{1,sn}(0)\left[1 + g_{sn}^{(2)}h^2 + \mathcal{O}(h^4)\right]$ for the cutoff wavenumbers, where $x_{1,sn}(0)$ are the normalized cutoff wavenumbers of the circular dielectric waveguide. These expressions are valid for every different value of s and n, corresponding to every higher-order hybrid mode. Numerical results are given for various higher-order modes, as well as a comparison with the exact solution.

Finally, in the third chapter the domain integral equation method is studied which, with its FFT-based matrix-vector products, is a viable alternative to local methods in free-space scattering problems. However, it often suffers from the extremely slow convergence of iterative methods, especially in the transverse electric (TE) case with large or negative permittivity. We identify very dense line segments in the spectrum as partly responsible for this behavior, and the main reason why a normally efficient deflating preconditioner does not work. We solve this problem by applying an explicit multiplicative regularizing operator, which on the operator level transforms the system to the form "identity plus compact". On the matrix level this regularization reduces the length of the dense spectral segmets roughly by a factor of four, while preserving the ability to calculate the matrix-vector products using the FFT algorithm. Such a regularized system is then further preconditioned by deflating an apparently stable set of eigenvalues with largest magnitudes, which results in a robust acceleration of the restarted GMRES under constraint memory conditions.

Abstract in greek

Το αντικείμενο αυτής της διδακτορικής διατριβής προσανατολίζεται σε δύο κατευθύνσεις: η πρώτη περιλαμβάνει ακριβείς τεχνικές των οποίων η ισχύς βασίζεται σε αναλυτικές εκφράσεις. Πιο συγκεκριμένα μελετάται ένα κλασικό πρόβλημα της σκέδασης ηλεκτρομαγνητικών κυμάτων, με χρήση ελλειπτικών κυματικών συναρτήσεων. Η δεύτερη κατεύθυνση προσανατολίζεται στις αριθμητικές τεχνικές, όπου και μελετάται η ιδιόμορφη χωρική ολοκληρωτική εξίσωση. Το κύριο πλεονέκτημά της έγκειται στην κατανόηση φυσικών φαινομένων, όπως για παράδειγμα φω-

τονικές εφαρμογές, εφαρμογές νανοφυσικής, βιολογικές απεικονίσεις, οπτικές ίνες κ.λπ.

Στο πρώτο κεφάλαιο μελετάται η σκέδαση ηλεκτρομαγνητικών κυμάτων από έναν απέραντο ελλειπτικό κύλινδρο τυχαίου προσανατολισμού, ο οποίος έχει διαφορετικές συντακτικές παραμέτρους από τον περιβάλλοντα χώρο. Η επίλυση του προβλήματος γίνεται με τη μέθοδο χωρισμού των μεταβλητών, λόγω όμως της λοξής πρόσπτωσης δημιουργούνται υβριδικά κύματα για τα σκεδαζόμενα και τα επαγόμενα πεδία, γεγονός που περιπλέκει τον φορμαλισμό. Επιπλέον, λόγω των διαφορετικών κυματαρίθμων του σκεδαστή και του περιβάλλοντα χώρου, οι σχέσεις ορθογωνιότητας των συναρτήσεων Μάθιου δεν είναι δυνατόν να εφαρμοστούν. Έτσι οδηγούμαστε σε ακόμη πιο πολύπλοκα συστήματα, συγκρινόμενα με αυτά της κάθετης πρόσπτωσης. Η επίλυση αυτών των σύνθετων συστημάτων θα οδηγήσει στην εύρεση των σκεδαζόμενων και των επαγόμενων πεδίων. Η επικύρωση των αποτελεσμάτων αποκαλύπτει την υψηλή ακρίβεια της μεθόδου, ακόμη και για ηλεκτρικά μεγάλους σκεδαστές. Μελετώνται και οι δύο πολώσεις και δίνονται αριθμητικά αποτελέσματα για διάφορες τιμές των παραμέτρων. Η μέθοδος είναι ακριβής και μπορεί να χρησιμοποιηθεί ως αναφορά για την ορθότητα μελλοντικών τεχνικών ηλεκτρομαγνητικής σκέδασης.

Η ηλεκτρομαγνητική σκέδαση από διάφορα αντικείμενα είναι μεγάλου ενδιαφέροντος στον εφαρμοσμένο ηλεκτρομαγνητισμό. Για παράδειγμα, διαγνωστικές μεθοδολογίες ή ο ρυθμός απορρόφησης ακτινοβολίας, μπορεί να υπολογιστεί σε ελλειπτικά βιολογικά μοντέλα ιστών. Η μελέτη της σκέδασης από κυκλικούς κυλίνδρους, έχει συζητηθεί διεξοδικά στη βιβλιογραφία. Τα τελευταία χρόνια υπάρχει μια αυξανόμενη μελέτη της σκέδασης από την πιο δύσκολη ελλειπτική γεωμετρία, δηλ. κύλινδροι με διηλεκτρικό που αποτελούνται από ελλειπτικές διατομές, πολυστρωματικοί διηλεκτρικοί ελλειπτικοί κύλνδροι, ή ακόμη και ομοιογενείς ανισοτροπικοί ελλειπτικοί κύλινδροι.

Η πιο ευρέως χρησιμοποιούμενη μέθοδος για ελλειπτικούς σκεδαστές βασίζεται στο χωρισμό των μεταβλητών με ανάπτυξη των πεδίων με χρήση των ελλειπτικών κυματοσυναρτήσεων, δηλαδή συναρτήσεων Μάθιου. Είναι γνωστό ότι, όταν δύο μέσα – ο σκεδαστής και ο περιβάλλων χώρος – έχουν τους ίδιους κυματάριθμους, τότε ισχύει η ορθογωνιότητα των συναρτήσεων Μάθιου. Από την άλλη πλευρά, όταν ο σκεδαστής έχει δείκτη διάθλασης διαφορετικό από εκείνο του περιβάλλοντα χώρου, η ορθογωνιότητα δεν ισχύει, κάτι που οδηγεί σε περίπλοκες εκφράσεις, όταν κάποιος προσπαθήσει να ικανοποιήσει τις οριακές συνθήκες. Επιπλέον, μια πρόσθετη δυσκολία προκύπτει όταν η πρόσπτωση του επίπεδου κύματος δεν είναι κάθετη, αλλά πλάγια, με αποτέλεσμα έναν συνδυασμό TM και TE κυμάτων – υβριδικά κύματα – για τα επαγόμενα και τα σκεδαζόμενα πεδία.

Τα περισσότερα προβλήματα σκέδασης ασχολούνται με την κάθετη πρόσπτωση. Αρχικά, η μελέτη ξεκίνησε με τον υπολογισμό της οπίσθιας διατομής σκέδασης από έναν διηλεκτρικό ελλειπτικό κύλινδρο υπό κάθετη πρόσπτωση [79]. Ακριβείς εκφράσεις κλειστής μορφής έχουν ληφθεί [70] για τις διατομές σκέδασης από έναν ελλειπτικό μεταλλικό κύλινδρο που έχει μικρή εκκεντρότητα, χρησιμοποιώντας ασυμπτωτική ανάλυση, ενώ στην [72], η ίδια διαδικασία έχει ακολουθηθεί για έναν διηλεκτρικό ελλειπτικό κύλινδρο, επίσης με μικρή εκκεντρότητα. Σύνθετες ελλειπτικές γεωμετρίες έχουν επίσης μελετηθεί, π.χ. στην [60] για την περίπτωση ενός ελλειπτικού κυλίνδρου επικαλυμμένου με ομοιογενές υλικό. Η μέθοδος που χρησιμοποιήθηκε είναι ο χωρισμός των μεταβλητών. Τεχνικές για πολυστρωματικές δομές ελλειπτικών κυλίνδρων έχουν επίσης προταθεί. Στην [10] οι συγγραφείς έχουν μελετήσει την περίπτωση TM πρόσπτωσης για σκέδαση από πολυστρωματικό ελλειπτικό κύλινδρο. Έχουν κατασκευάσει μια αναδρομική δια-

δικασία για τον υπολογισμό του πεδίου, με βάση τις συναρτήσεις Μάθιου, και πάλι για κάθετη πρόσπτωση. Στην [9], έχει προταθεί μια αναλυτική λύση για έναν πολυστρωματικό ελλειπτικό κύλινδρο, στην περίπτωση που ο κύλινδρος και το μέσον που τον περιβάλλει έχουν τον ίδιο δείκτη διάθλασης, ή αλλιώς, τον ίδιο κυματάριθμο, ικανοποιώντας έτσι τις ιδιότητες ορθογωνιότητας των συναρτήσεσν Μάθιου. Τελικά, στην [42], η περίπτωση ενός ομοιογενούς ανισοτροπικού ελλειπτικού κυλίνδρου έχει επίσης μελετηθεί.

Η σκέδαση της λοξής πρόσπτωσης είναι μια πιο γενική περίπτωση σε όσα συζητήθηκαν μέχρι στιγμής. Πολλοί συγγραφείς έχουν συμβάλλει για την περίπτωση των κυκλικών κυλίνδρων, όπου διαφορετικές μέθοδοι έχουν χρησιμοποιηθεί. Στην [40], τα κυλινδρικά κυματαδιανύσματα χρησιμοποιούνται από τον συγγραφέα για τη μελέτη της σκέδασης από διηλεκτρικό κυκλικό κυλίνδρο. Η μέθοδος των βοηθητικών πηγών έχει χρησιμοποιηθεί στην [69] για τη διερεύνηση της σκέδασης από έναν τέλεια αγώγιμο κυκλικό κύλινδρο, αν και στην περίπτωση αυτή δεν εμφανίζονται υβριδικά κύματα. Μια εναλλακτική μέθοδος με βάση γενικευμένες σειρές παρουσιάζεται στην [39], και οι συγγραφείς επιβεβαιώνουν τα αποτελέσματα από την [40]. Η περίπτωση των ελλειπτικών κυλίνδρων ξεκίνησε στην [78], με το φορμαλισμό του προβλήματος μέχρι ένα συγκεκριμένο σημείο, αλλά δεν υπάρχουν αριθμητικά αποτελέσματα. Στην [68] ο συγγραφέας χρησιμοποιεί ολοκληροδιαφορικές εξισώσεις για να μελετήσει την σκέδαση από διηλεκτρικούς κυλίνδρους, αυθαίρετης διατομής, και επικυρώνει τα αποτελέσματά του με την εξέταση ελλειπτικών κυλίνδρων. Επίσης, στην [45], μελετάται η σκέδαση ηλεκτρομαγνητικών κυμάτων με χρήση του τανυστή Γκρήν, και δίνονται παραδείγματα με ελλειπτικές γεωμετρίες.

Είναι προφανές ότι η χρήση των τεχνικών που χρησιμοποιούνται για τους ελλειπτικούς κυλίνδρους, βασίζεται κυρίως στην κάθετη πρόσπτωση και στη χρήση υλικών με ίδιους δείκτες διάθλασης. Υπάρχει μια έλλειψη του συνδυασμού της λοξής πρόσπτωσης με υλικά που έχουν διαφορετικές συντακτικές παραμέτρους. Η μέθοδος που παρουσιάζεται εδώ είναι ακριβής και εναλλακτική εκείνων που βασίζονται στις ολοκληρωτικές εξισώσεις [68, 43]. Επιπλέον, αν και η επικύρωση για πολλές νέες τεχνικές βασίζεται σε κυκλικούς κυλίνδρους, με βάση τις γνωστές λύσεις των συναρτήσεων Μπέσελ, το πρόβλημα σκέδασης από ελλειπτικούς κυλίνδρους είναι σε πολλές περιπτώσεις μια ισχυρή απόδειξη της ορθότητας των αποτελεσμάτων. Ως εκ τούτου, τα αποτελέσματα της παρούσας μεθόδου μπορούν να χρησιμοποιηθούν ως αναφορά για την επικύρωση μελλοντικών μεθόδων που επιλύουν προβλήματα ηλεκτρομαγνητικής σκέδασης.

Στο δεύτερο κεφάλαιο υπολογίζονται οι κυματάριθμοι αποκοπής για έναν διηλεκτρικό ελλειπτικό κυματοδηγό. Ο υπολογισμός γίνεται με ακριβή καθώς και με αναλυτικό τρόπο. Για την επίλυση αυτού του προβλήματος χρησιμοποιήθηκαν δύο διαφορετικοί τρόποι. Ο πρώτος τρόπος βασίζεται στο χωρισμό μεταβλητών με χρήση των συναρτήσεων Μάθιου, ο οποίος δίνει τους ακριβείς κυματαρίθμους αποκοπής. Η μέθοδος αυτή οδηγεί σε σύνθετα γραμμικά ομογενή συστήματα των οποίων πρέπει να υπολογίσουμε τις ρίζες της ορίζουσας των αντίστοιχων πινάκων. Η πολυπλοκότητα των πινάκων αυτών οφείλεται στη μη ύπαρξη ορθογωνιότητας των συναρτήσεων Μάθιου λόγω των διαφορετικών συντακτικών παραμέτρων μεταξύ του πυρήνα και του περιβλήματος της ίνας. Η δεύτερη μέθοδος υπολογίζει τους κυματαρίθμους αποκοπής αναλυτικά, χρησιμοποιώντας κλειστούς τύπους. Οι εκφράσεις αυτές ισχύουν όταν η εκκεντρότητα του ελλειπτικού πυρήνα της ίνας περιορίζεται σε μικρές τιμές. Στην περίπτωση αυτή οι κυματάριθμοι αποκοπής έχουν τη μορφή $x_{1,sn}(h) = x_{1,sn}(0)\left[1 + g_{sn}^{(2)}h^2 + \mathcal{O}(h^4)\right]$, όπου οι συντελεστές ανάπτυξης $g_{sn}^{(2)}$ δίνονται από αναλυτικούς, κλειστούς αλγε-

βρικούς τύπους, χωρίς την παρουσία συναρτήσεων Μάθιου, ενώ $x_{1,sn}(0)$ είναι οι κυματάριθμοι αποκοπής του διηλεκτρικού κυκλικού κυματοδηγού. Οι αναλυτικές αυτές εκφράσεις ισχύουν για κάθε διαφορετική τιμή των δεικτών s και n, που αντιστοιχούν σε υβριδικούς ρυθμούς υψηλότερης τάξης. Δίνονται αριθμητικά αποτελέσματα για διάφορους ρυθμούς υψηλότερης τάξης και γίνεται σύγκριση με την ακριβή λύση.

Το προαναφερθέν εσωτερικό πρόβλημα έχει πολύ μεγάλο ενδιαφέρον και εφαρμογές, όπως στη χρήση των οπτικών κυματοδηγών σε δίκτυα διανομής σήματος σε κεραίες, σε συστήματα κινητής τηλεφωνίας, και ιδιαίτερα στη χρήση ως γραμμών μεταφοράς μεγάλης απόστασης για οπτικά δίκτυα. Η μελέτη των ρυθμών κυματοδήγησης μέσα σε έναν κυκλικό διηλεκτρικό κυματοδηγό είναι γνωστή εδώ και πέντε δεκαετίες. Στην [58] οι συγγραφείς εξάγουν πρώτοι τη χαρακτηριστική εξίσωση για την εύρεση των κυματαρίθμων αποκοπής για ανώτερης τάξης υβριδικούς ρυθμούς. Μια πιο λεπτομερής και εμπεριστατωμένη ανάλυση σχετικά με το θέμα των υψηλότερης τάξης ρυθμών δίνεται στην [44], όπου ο συγγραφέας παρουσιάζει τις χαρακτηριστικές εξισώσεις για εγκάρσια μαγνητικούς (TM), εγκάρσια ηλεκτρικούς (TE) και υβριδικούς (EH ή HE) ρυθμούς. Ασυμπτωτικές εκφράσεις παρουσιάζονται στην [63], οι οποίες ισχύουν για όλους τους ρυθμούς ανώτερης τάξης. Μια ολοκληρωμένη μελέτη για τους ρυθμούς της κυκλικής οπτικής ίνας δίνεται στην [81], μαζί με μια ωραία περιγραφή για τους ρυθμούς. Η μελέτη της διάδοσης σε ελλειπτικούς διηλεκτρικούς κυματοδηγούς είναι πολύ πιο δύσκολη και παρουσιάζει μεγάλο ενδιαφέρον. Ωστόσο, οι μελέτες αυτές είναι λίγες λόγω των δυσκολιών που προκύπτουν από την ελλειπτική γεωμετρία, και δεν υπάρχουν καθαρά κλειστής μορφής αναλυτικοί τύποι για τους υβριδικούς ρυθμούς κυματοδήγησης.

Οι ερευνητές έχουν χρησιμοποιήσει διαφορετικές μεθόδους, κατά τα τελευταία χρόνια, για να μελετήσουν τα χαρακτηριστικά των ρυθμών για οπτικές ίνες ελλειπτικής γεωμετρίας. Η μέθοδος της ανάπτυξης των πεδίων σε σειρές συναρτήσεων Μάθιου χρησιμοποιείται στην [77], όπου εξετάζεται μόνο ο θεμελιώδης ρυθμός. Στη συνέχεια, στην [80], χρησιμοποιούνται και πάλι οι συναρτήσεις Μάθιου, ο συγγραφέας όμως απλοποιεί το φορμαλισμό κάνοντας χρήση της συνθήκης της χαλαρής κυματοδήγησης, δηλαδή όταν οι δείκτες διάθλασης του πυρήνα και του περιβλήματος της ίνας είναι σχεδόν όμοιοι. Η ίδια προσέγγιση χρησιμοποιείται στην [64] για να μελετηθούν οι γραμμικά πολωμένοι ρυθμοί μέσα σε οπτικούς κυματοδηγούς που έχουν αυθαίρετη διατομή. Η τροποποιημένη μέθοδος των πεπερασμένων στοιχείων εφαρμόζεται στην [82] για να μελετηθούν οι ρυθμοί σε έναν κυκλικό οπτικό κυματοδηγό με ανομοιογενή πυρήνα, και η μέθοδος εφαρμόζεται και για την ελλειπτική περίπτωση. Η περίπτωση των διηλεκτρικών κυματοδηγών αυθαίρετης διατομής αντιμετωπίζεται στην [24], όπου οι συγγραφείς μελετούν την περίπτωση της ασθενής κυματοδήγησης. Στην [38] δίνεται μια πλήρης ανάλυση των ρυθμών, με βάση τις συναρτήσεις Μάθιου, και οι συγγραφείς δίνουν μια ποικιλία αποτελεσμάτων για τη διασπορά, για τους ανώτερης τάξης υβριδικούς ρυθμούς, χρησιμοποιώντας την ακριβή χαρακτηριστική εξίσωση. Η εργασία τους επεκτείνεται στην [52], όπου μελετώνται διάφορα χαρακτηριστικά διάδοσης, όπως η ταχύτητα ομάδας και η εξασθένηση των διαφόρων ανώτερης τάξης ρυθμών. Στην [51] υπολογίζεται ο κυματάριθμος αποκοπής του πρώτου ανώτερης τάξης ρυθμού, χρησιμοποιώντας το φορμαλισμό της [77]. Η μέθοδος των πεπερασμένων στοιχείων έχει χρησιμοποιηθεί στην [14], προκειμένου να υπολογιστούν οι συχνότητες αποκοπής για τους ανώτερους τάξης ρυθμούς, και υπό το πρίσμα της ασθενούς κυματοδήγησης, σε ίνες αυθαίρετης διατομής, και δίνονται αριθμητικά αποτελέσματα για ίνες με ελλειπτικό πυρήνα για διάφορες εκκεντρότητες. Στην [50], γίνεται

σύγκριση μεταξύ της μεθόδου των συναρτήσεων Μάθιου, της μεθόδου που χρησιμοποιήθηκε στην [56], και ερμηνεύονται διάφορες αποκλίσεις. Μια πειραματική προσέγγιση παρουσιάζεται στην [20], για τον υπολογισμό της αποκοπής κάποιων ρυθμών υψηλότερης τάξης σε διηλεκτρικούς ελλειπτικούς κυματοδηγούς. Η προσέγγιση της ασθενούς κυματοδήγησης εφαρμόζεται στην [61], για τον υπολογισμό των ρυθμών από την επίλυση της εξίσωσης Μάθιου με τη χρήση των χαρακτηριστικών αριθμών τους. Η ίδια προσέγγιση χρησιμοποιείται επίσης στην [74], διατυπώνοντας το πρόβλημα με χρήση ελλειπτικών σειρών Φουριέ. Τέλος, στην [27] υπολογίζονται οι ρυθμοί στη χαλαρή κυματοδήγηση, χρησιμοποιώντας μια απλοποιημένη χαρακτηριστική εξίσωση με βάση την ανάπτυξη σε συναρτήσεις Μάθιου.

Είναι προφανές ότι, παρόλο που το μεγαλύτερο μέρος της έρευνας γίνεται για την περίπτωση της ασθενούς κυματοδήγησης, όπου ο φορμαλισμός απλοποιείται σημαντικά, η γενική περίπτωση έχει επίσης μελετηθεί, ωστόσο, οι ακριβείς χαρακτηριστικές εξισώσεις που εμφανίζονται στις [77] και [38] χρειάζονται περισσότερη ανάπτυξη για να καταλήξουμε σε μια νέα χαρακτηριστική εξίσωση που να ισχύει στην αποκοπή. Επιπλέον, σε σύγκριση με τον κυκλικό διηλεκτρικό κυματοδηγό όπου έχουμε ξεχωριστές αναλυτικές, κλειστού τύπου εξισώσεις για τον άμεσο υπολογισμό των κυματαρίθμων αποκοπής [44], δεν υπάρχουν οι αντίστοιχες αναλυτικές, κλειστής μορφής χαρακτηριστικές εξισώσεις, απευθείας στην αποκοπή, για τον ελλειπτικό διηλεκτρικό κυματοδηγό.

Τέλος, στο τρίτο κεφάλαιο μελετάται η μέθοδος της χωρικής ολοκληρωτικής εξίσωσης. Η υλοποίηση της μεθόδου βασίζεται στον ταχύ μετασχηματισμό Φουριέ για τον ταχύ υπολογισμό των γινομένων μεταξύ πινάκων και διανυσμάτων, και αποτελεί μια εναλλακτική μέθοδο για προβλήματα σκέδασης, σε σχέση με τις αντίστοιχες «τοπικές» μεθόδους. Παρ' όλα αυτά, ένα αρνητικό σημείο της μεθόδου είναι η πολύ αργή σύγκλιση των επαναληπτικών μεθόδων που χρησιμοποιούνται, ιδιαίτερα στην περίπτωση της εγκάρσιας ηλεκτρικής πόλωσης με υψηλή ή αρνητική τιμή της επιτρεπτότητας. Διαπιστώνουμε ότι το φάσμα αποτελείται από πολύ πυκνά τμήματα στα οποία οφείλεται αυτή η συμπεριφορά της αργής σύγκλισης, και επιπλέον είναι ένας ακόμη λόγος για τον οποίο μέθοδοι βελτιστοποίησης οι οποίες βασίζονται σε απλή μετακίνηση των ιδιοτιμών σε πιο πρόσφορα σημεία, δεν επιτυγχάνουν. Το πρόβλημα αυτό της βελτιστοποίησης ξεπερνιέται εφαρμόζοντας ένα τελεστή ομαλοποίησης, ο οποίος, σε επίπεδο τελεστών μετασχηματίζει το σύστημα σε μορφή «μοναδιαίο και συμπαγές». Σε επίπεδο πινάκων, αυτή η ομαλοποίηση ελαττώνει κατά τέσσερις φορές το μήκος των πυκνών φασματικών τμημάτων, ενώ διατηρείται η δυνατότητα του υπολογισμού των γινομένων μεταξύ πινάκων και διανυσμάτων, με χρήση του ταχύ μετασχηματισμού Φουριέ. Το ομαλοποιημένο σύστημα βελτιστοποιείται περαιτέρω με μετακίνηση ευσταθών ιδιοτιμών με μεγαλύτερο μέτρο, κάτι που οδηγεί σε αξιόλογη επιτάχυνση της επαναληπτικής γενικευμένης μεθόδου ελαχίστου υπολοίπου (GMRES), και μάλιστα κάτω από περιορισμένες συνθήκες υπολογιστικής μνήμης.

Γενικά, οι μέθοδοι μπορούν να διαιρεθούν σε δύο κατηγορίες – τοπικές και ολικές – ανάλογα με τις εξισώσεις. Οι τοπικές μεθόδοι βασίζονται στις διαφορικές εξισώσεις Μάξγουελ [22, 23, 29, 30], είναι πιο δημοφιλείς λόγω των αραιών πινάκων που προκύπτουν και την ευκολία του προγραμματισμού τους. Παράλληλα με τις τοπικές μεθόδους, οι ολικές μεθόδοι βασίζονται στις ισοδύναμες ολοκληρωτικές εξισώσεις που χρησιμοποιούνται συχνά σε προβλήματα σκέδασης ελευθέρου χώρου. Εάν ένα αντικείμενο είναι μεγάλο και ομογενές, ή είναι τέλεια αγώγιμο, τότε το πρόβλημα συνήθως εκφράζεται με χρήση συνοριακών ολοκληρωτικών εξισώσεων, με θεμελιώδεις αγνώστους τα πεδία (ρεύματα) στις διεπαφές [28]. Εάν

ένα αντικείμενο είναι ανομοιογενές ή σύνθετο και αποτελείται από πολλά μικρά διαφορετικά μέρη, η πλέον ενδεδειγμένη ολική μέθοδος είναι αυτή της χωρικής ολοκληρωτικής εξίσωσης, σε δύο ή σε τρεις διαστάσεις [31, 32, 34, 41, 49, 53, 65, 67, 73, 33, 17]. Παρόλο που οι ολικές μέθοδοι παράγουν πυκνούς πίνακες, είναι γενικά πιο σταθερές, όσο αφορά τη διακριτοποίηση, και οι συνελικτικού τύπου ολοκληρωτικοί τελεστές επιτρέπουν μερικές φορές τον υπολογισμό των γινομένων πίνακα-διανύσματος με την ταχύτητα του ταχύ μετασχηματισμού Φουριέ. Αυτά τα χαρακτηριστικά καθιστούν τη μέθοδο της χωρικής ολοκληρωτικής εξίσωσης μια εναλλακτική λύση για προβλήματα σκέδασης ελεύθερου χώρου.

Οι κυριότερες δυσκολίες της παρούσας μεθόδου είναι η μη κανονικότητα, τόσο του τελεστή, όσο και του πίνακα που προκύπτει, και η εξαιρετικά αργή σύγκλιση των λίγων επαναληπτικών μεθόδων που μπορούν να εφαρμοστούν σε τέτοιους πίνακες. Αριθμητικά πειράματα δείχνουν σταθερά ότι η επαναληπτική γενικευμένη μέθοδος ελαχίστου υπολοίπου είναι η καλύτερη επιλογή [25]. Ως ένας πεπερασμένος αλγόριθμος, η πλήρης (χωρίς επανεκκίνηση) αυτή μέθοδος, τελικά θα συγκλίνει με οποιοδήποτε φυσικά λογικών διαστάσεων σκεδαστή [57]. Ωστόσο, ο αριθμός των αγνώστων είναι απαγορευτικός για τη χρήση της πλήρους μεθόδου, και συνεπώς την χρησιμοποιούμε με επανεκκίνηση. Δυστυχώς, μερικές φορές, ειδικά στα προβλήματα της ΤΕ σκέδασης με υψηλές τιμές επιτρεπτότητας ή διαστάσεων του σκεδαστή, η επαναληπτική γενικευμένη μέθοδος ελαχίστου υπολοίπου με επανεκκίνηση, καθώς και άλλες γνωστές επαναληπτικές μέθοδοι, συγκλίνουν πολύ αργά, ή δε συγκλίνουν καθόλου.

Πολλές επιτυχημένες στρατηγικές βελτισοποίησης έχουν προταθεί για τις τοπικές μεθόδους, βλέπε π.χ. [22, 23], καθώς και για τις συνοριακές ολοκληρωτικές εξισώσεις [2, 3, 15, 16, 66]. Η μέθοδος της χωρικής ολοκληρωτικής εξίσωσης υστερεί στον τομέα αυτό [12, 25, 26, 83], και όποιες αναφορές βελτιστοποίησης υπάρχουν, γίνονται κυρίως για χαμηλές τιμές επιτρεπτότητας ή για μικρούς σκεδαστές. Μια από τις δυσκολίες στο σχεδιασμό μιας κατάλληλης τεχνικής βελτιστοποίησης για τη μέθοδο αυτή, είναι ότι πρέπει οι πίνακες να είναι αραιοί, ώστε να είναι σε θέση να ανταγωνιστούν με τις τοπικές μεθόδους όσον αφορά τη μνήμη και την ταχύτητα. Πρόσφατα, μια τεχνική βελτιστοποίησης που βασίζεται σε μετακίνηση ιδιοτιμών, έχει προταθεί στην [62], για την επιτάχυνση της λύσης στην περίπτωση της ΤΜ σκέδασης. Μια λεπτομερής φασματική ανάλυση του πίνακα του συστήματος που αναφέρθηκε στην [62], δείχνει ότι οι απομακρυσμένες ιδιοτιμές (οι μεγαλύτερες σε μέτρο) είναι υπαίτιες για την στασιμότητα της πλήρης επαναληπτικής γενικευμένης μεθόδου ελαχίστου υπολοίπου. Παρόλα αυτά, όπως δείχνουμε εδώ, αυτή τη στρατηγική δε λειτουργεί στην περίπτωση της ΤΕ σκέδασης.

Αναλύουμε, συνεπώς, τη διαφορά ανάμεσα στα φάσματα της ΤΕ και ΤΜ σκέδασης και δείχνουμε ότι η διακριτοποιημένη έκδοση της μεθόδου οδηγεί σε μια πολύ πυκνή ομάδα ιδιοτιμών, καθιστώντας τη διαδικασία μετακίνησης ανέφικτη. Εναλλακτικά, χρησιμοποιούμε έναν ομαλοποιητή, ο οποίος είναι ένας συνελικτικού τύπου ολοκληρωτικός τελεστής, κάτι που επιτρέπει να υπολογίζουμε τα γινόμενα πινάκων-διανυσμάτων με τον αλγόριθμο του ταχύ μετασχηματισμού Φουριέ. Η εφαρμογή του ομαλοποιητή συμπιέζει τις πυκνές ομάδες ιδιοτιμών, και εν συνεχεία μπορούμε να μετακινήσουμε τις μεγαλύτερες σε μέτρο ιδιοτιμές. Έχουμε, εν συνεχεία, εκτελέσει μια σειρά αριθμητικών πειραμάτων τα οποία έχουν επανειλημμένα δείξει ότι, αφιερώνοντας ένα σταθερό ποσό της μνήμης του υπολογιστή για την μετακίνηση των ιδιοτιμών και, εν συνεχεία, χρήση της επαναληπτικής γενικευμένης μεθόδου ελαχίστου υπολοίπου, οδηγεί σε πιο αποτελεσματικό εξαγόμενο από το να αφιερώσουμε όλη τη μνήμη του υπολογιστή για τη μεγιστοποίηση της διάστασης του εσωτερικού υπόχωρου για χρήση της πλήρους μεθόδου.

Αν και η μέθοδος της χωρικής ολοκληρωτικής εξίσωσης – τόσο η ΤΜ όσο και η ΤΕ περίπτωση – είναι γνωστή στην επιστημονική κοινότητα εδώ και πολλά χρόνια [31, 32, 34, 41, 49, 53, 65, 67, 73], η πλήρης θεωρητική ανάλυση στις δύο διαστάσεις εξακολουθεί να λείπει. Ο λόγος μπορεί να είναι η παρουσία των ειδικών συναρτήσεων στον πυρήνα του ολοκληρωτικού τελεστή, κάτι που περιπλέκει τους υπολογισμούς, περισσότερο απ' ότι στην τρισδιάστατη περίπτωση [5, 57, 33, 17]. Ως εκ τούτου, πριν εισέλθουμε στις λεπτομέρειες των αριθμητικών τεχνικών, αφιερώνουμε κάποιο χώρο για την θεωρητική ανάλυση των δύο διαστάσεων. Τα αποτελέσματα από το θεωρητικό μέρος της εργασίας μας είναι: το σύμβολο (symbol) του ολοκληρωτικού τελεστή, ο ομαλοποιητής (regularizer), και η ισοδυναμία του φάσματος μεταξύ δύο και τριών διαστάσεων, για την ΤΕ περίπτωση.

Acknowledgements

I express my deepest gratitude to my family for their sacrifices. I am greatly indebted to my supervisor Dr. John Roumeliotis and to my co-supervisor Dr. Neil Budko, for their outstanding support. I thank Yota for helping me broaden my horizons. Special thanks are due to Mr. Georgios Tsogkas of National Technical University of Athens for the useful discussions, and to Mr. Jianing Wang of Delft University of Technology for the bibliographic support.

Introduction

Electromagnetic scattering from various objects is of great interest in applied electromagnetics. For instance, diagnostic methodologies in biological tissues can be deduced or the specific absorption rate (SAR) can be computed inside elliptic biological tissue models [11]. The study of the scattering by circular cylinders has been discussed thoroughly in the literature. In the recent years there is a growing study of scattering by the more difficult elliptical geometries, i.e. dielectric cylinders which are composed by elliptical cross sections, multilayered dielectric elliptical cylinders, or even homogeneous anisotropic elliptical cylinders.

The most widely used formulation for elliptic scatterers is based on the separation of variables method (SVM) by expanding the fields in terms of elliptical wavefunctions, i.e. Mathieu functions. It is known that when two media – the scatterer and the background medium – have the same wavenumber, then there exists the orthogonality relation for Mathieu functions. This also applies when the scatterer is perfectly conducting. On the other hand, when the scatterer has a refractive index different than that of the background medium, the orthogonality does not hold, leading to complicated expressions when one tries to satisfy the boundary conditions. Moreover, an additional difficulty arrives when the incidence on a dielectric cylinder is not normal but oblique, leading to cross polarization of the fields, namely a combination of TM and TE waves (or hybrid waves) for the scattered and induced fields.

Most scattering problems deal with normal incidence. Initially, the study was started in [79] by obtaining the back scattering cross section of a dielectric elliptical cylinder under normal incidence. In [59] the authors have considered the scattering by elliptic cylinders on various configurations, i.e. homogeneous cylinders and dielectric coated conducting cylinders, based on expanding the fields in terms of elliptical wavefunctions. Exact, closed form expressions have been obtained in [70] for the

scattering cross sections of the scattering by an elliptic metallic cylinder having small eccentricity, by using asymptotic analysis, while in [72], the same procedure has been followed to obtain results for a dielectric elliptic cylinder, also with small eccentricity. Composite elliptical bodies have also been studied, e.g. in [60] for the case of an impedance elliptic cylinder coated with homogeneous material. The method used there was the separation of variables with the impendance boundary condition (IBC). Techniques for multilayer elliptic structures have also been proposed. In [10] the authors have studied the TM illumination case for scattering by a multilayer elliptic cylinder. They have built a recursive procedure for the field computation, based on Mathieu functions, again for normal incidence. In [9], an analytical solution has been obtained for a multilayer elliptic cylinder in the case of isorefractive materials, i.e. when the two media have the same refractive index, or else, the same wavenumber, thus satisfying the orthogonality properties of Mathieu functions. The case of the normal scattering by multilayer elliptic cylinders is also reviewed in [48]. Ultimately, in [42], the case of a homogeneous anisotropic elliptic cylinder has been considered. The solution was obtained by expanding the fields using Mathieu functions as eigenfunctions, together with Fourier series. Their formulation includes media having different refractive indices, but still is under normal incidence.

Scattering of obliquely incident waves, sometimes being referred as 2.5D scattering or 3D scattering by an infinite cylinder, is a more general case to what discussed so far. Many authors have contributed for the case of circular cylinders where different methods have been employed. In [40], the circular cylindrical wavevector functions have been used by the author to study the scattering by dielectric circular cylinders. The method of auxiliary sources (MAS) has been used in [69] to investigate the scattering by a perfectly conducting circular cylinder, although in this case hybrid waves do not appear. An alternative method based on the generalised Debye series expansion (GDSE) is presented in [39] and the authors confirm the results from [40]. The version of elliptical cylinders was started in [78] by modelling the problem up to a specific point, but not providing any numerical results. Then, in [55], the problem for the case of a perfectly conducting elliptic cylinder embedded in an uniaxially anisotropic medium has been formulated. Impedance elliptic cylinders have also been studied, i.e. in [76], by the use of two alternative methods. One was the manipulation of the elliptical boundary it terms of polar coordinates and the other was the direct implementation of the finite difference method (FDM). In [68] the author uses integral-integrodifferential equations, which are solved via Nyström method, to study dielectric cylinders having arbitrarily shaped smooth boundary, and he validates his results by considering elliptic cylinders. Also, in [45], the Green's tensor technique is applied to study 2D and 3D electromagnetic scattering and examples are provided with elliptical geometries. Finally, in [43], the authors continue their work started in [42] by introducing oblique incidence. They use Galerkin's method as well as the FDTD method to validate the results.

It is evident that the formulation of the most techniques used for elliptical cylinders supports mostly normal incidence or isorefractive materials. There is a lack of combination of oblique incidence on materials having different constitutive parameters, and its corresponding investigation. The method presented here is exact and alternative to those that can support the aforementioned combination, i.e. methods that are based on integral equations [68, 43]. Moreover, although the validation for many new techniques relies on the circular geometry with the well known solution based on Bessel functions, the scattering problem from elliptical cross section cylinders is in many cases a challenging way for proving the correctness of the results. Therefore, this method can be used for reference as an alternative

validation for future methods involving scattering problems.

There is also a steady interest in the numerical simulation of the electromagnetic field in inhomogeneous media. The methods can be roughly divided into two categories – local and global – in accordance with the governing equations. The local methods based on the differential Maxwell's equations [22, 23, 29, 30] are generally more popular due to the sparse nature of their matrices and the ease of programming. Alongside the local methods, global methods based on an equivalent integral equation formulation are also often employed, especially in free-space scattering. If an object is large and homogeneous or has a perfectly conducting boundary then the problem is usually reduced to a boundary integral equation with the fields (currents) at the interfaces being the fundamental unknowns [28]. If an object is continuously inhomogeneous or is a composite consisting of many small different parts, the most appropriate global method is the domain integral equation (DIE), in two dimensions, or the volume integral equation, in three dimensions [31, 32, 34, 41, 49, 53, 65, 67, 73, 33, 17]. Although the global methods produce dense matrices, they are generally more stable with respect to discretization than the local ones, and the convolution-type integral operators sometimes allow to compute matrix-vector products at the FFT speed. These properties make the DIE method a viable alternative to local methods for certain free-space scattering problems.

The main difficulties with the DIE method are the non-normality of both the operator and the resulting system matrix, inherent to frequency-domain electromagnetic scattering, and the extremely slow convergence of the few iterative methods that can be applied with such matrices. Numerical experiments consistently show that the GMRES algorithm is the best choice [25]. Being a finite algorithm, the full (un-restarted) GMRES will eventually converge with any physically meaningfull scatterer and incident field [57]. Most of the practically interesting problems, however, involve the number of unknowns prohibitive for the full GMRES, and its restarted version is used. Unfortunately, sometimes, especially in TE scattering with large permittivities/dimensions, restarted GMRES as well Bi-CGSTAB converge very slowly or do not converge at all.

Several successful preconditioning strategies have been proposed for local methods, see e.g. [22, 23], and boundary integral formulations [2, 3, 15, 16, 66]. Whereas the domain integral equation method is lagging behind in this respect [12, 25, 26, 83], reporting convergence and acceleration thereof mainly for low-contrast and/or small objects. One of the difficulties in designing a suitable multiplicative preconditioner for this method is that it needs to be either sparse or have a block-Toeplitz form to be able to compete with local methods in terms of memory and speed. Recently a deflation-based preconditioner has been proposed in [62] for accelerating the DIE solver in the TM case. A thorough spectral analysis of the system matrix reported in [62] showed that the outlying eigenvalues (largest in magnitude) are responsible for a period of stagnation of the full GMRES, and that their deflation accelerates the iterative process. As we show here this strategy, however, does not work in the TE case. Moreover, using the deflation on top of a full GMRES puts an even higher strain on the memory.

To further extend the applicability of the DIE method towards objects with higher contrasts/sizes, in this dissertation we propose a two-stage preconditioner based on the regularization of the pertaining singular integral operator and subsequent deflation of the largest eigenvalues. To derive a regularizer we employ the symbol calculus, which also provides us with the essential spectrum of the original operator. Our regularizer may be viewed as a generalization of the Calderón identity recently employed for preconditioning of the boundary integral equation method

[2, 3, 15]. As an illustration and a proof of the concept we compute the complete spectrum of the system matrix for a few typical objects of resonant size. We analyze the difference in spectra between the TE and TM polarizations and demonstrate that the discretized version of the regularizer indeed contracts a very dense group of eigenvalues making the regularized system amenable to deflation. Unlike the previously mentioned boundary integral equation approach [2] the DIE method permits a straightforward discretization of the continuous regularizer without any additional precautions. As our regularizer is a convolution-type integral operator, all matrix-vector products can be carried out with the FFT algorithm. The subsequent deflation of the largest eigenvalues can be achieved with the standard Matlab's EIGS routine. Most importantly, these largest eigenvalues and the associated eigenvectors turn out to be rather stable, so that one can considerably accelerate the EIGS algorithm by setting a modest tolerance. We have performed a number of numerical experiments which have consistently demonstrated that given a fixed amount of computer memory it is better to spend some part of it on deflation than all of it on maximizing the dimension of the inner Krylov subspace of the restarted GMRES. Keeping the RESTART parameter of the GMRES in the order or larger than the dimension of the deflation subspace gives a speed-up which grows with the object size and permittivity.

Although the DIE method (both TM and TE versions of it) has been around in engineering community for many years [31, 32, 34, 41, 49, 53, 65, 67, 73], the full theoretical analysis of the two-dimensional case (in the strongly singular form) is still missing. The reason might be the presence of special functions in the kernel of the integral operator, which makes calculations more tedious than in the three-dimensional case [5, 57, 33, 17]. Therefore, before going into the numerical details, we devote some space here to the analysis of the two-dimensional singular integral operator. In this theoretical part we restrict ourselves to objects described by the constitutive parameters Hölder-continuous in $\mathbb{R}^2$. The analysis requires re-writing the problem as a standard singular integral equation with its strongly and weakly singular parts explicitly separated. The results of the theoretical part of our work are: the operator symbol, the explicit regularizer, and the apparent (spectral) equivalence of the two-dimensional transverse case and the more general three-dimensional problem. The subsequent numerical analysis of the problem utilizes an equivalent but visually different form of the equation, since there is no need to separate the strongly and weakly singular parts while compiling the system matrix.

Apart from the aforementioned scattering problems, the corresponding inner problem, i.e. the electromagnetic wave propagation in dielectric waveguides, is of great interest in engineering since it has many applications, such as their use in signal distribution networks for phased-array antennas, for cellular phone systems, and especially the elliptical core fibers as long-distance transmission lines for optical networks. The study of the modes propagating inside a circular dielectric waveguide is well known since five decades ago. In [58] the authors first extracted the equation for obtaining the cutoff wavenumbers for higher-order hybrid modes. A more detailed, rigorous analysis on the aforementioned theme for the higher-order modes is given in [44], where the author presents the characteristic equations for transverse magnetic (TM), transverse electric (TE) and hybrid (EH or HE) modes. Analytical asymptotic expressions are presented in [63], which are valid for all frequencies and higher-order modes. Then, a comprehensive study for all modes in circular optical fibers is given in [81], along with a nice description for the modes and their designation. The study of the propagation in elliptical dielectric waveguides, apparently much more difficult, is also of great interest. However, such studies are cumbersome due to the difficulties arising by the elliptical geometry, and pure analytical formulas for hybrid modes of

any order do not exist.

The researchers have employed different methods, during the last years, to study the characteristics of the modes for the elliptical fibers. The expansion in series of Mathieu functions is used in [77] where only the fundamental modes $_{e}HE_{11}$ and $_{o}HE_{11}$ are examined. Then, in [80], using again Mathieu series expansions, the author simplifies the formulation under the weakly guiding approximation, i.e. when the refractive indices of the core and the cladding of the fiber are almost similar, and he expands his study to some higher-order modes. The same approximation is used in [64] to study the linearly polarized (LP) modes in optical waveguides having arbitrary cross section, but the formulation now is based on the scalar wave equations. The modified finite element method is applied in [82] to study single mode circular optical waveguides having inhomogeneous cores, and they further apply it to the elliptical case. The case of arbitrary cross sectional shape dielectric waveguides is treated in [24], where the authors use integral representations for the fields, and they study the weakly guiding case as well as the general case of completely different refractive indices between the core and the cladding. In [38] a complete analysis on the modes, based on Mathieu functions, is given, and the authors give a variety of dispersion curves for higher-order hybrid modes, from the exact characteristic equations. Their work is extended in [52] where the propagation characeristics, such as the group velocity and the attenuation for various higher-order modes, are studied. In [51] the cutoff parameter of the first higher-order mode, i.e. $_{o}EH_{01}$, is computed from the formulation obtained by Yeh in [77] and discrepancies with previous results are resolved. The finite element method is then again used in [14] to obtain the cutoff frequencies for the higher-order LP modes in fibers of arbitrary cross section and numerical results are given for circular as well as for elliptical cores, for various eccentricities. An alternative method to that of Mathieu series expansion, namely the point-matching numerical method, is used in [56] for the calculation of some higher-order hybrid modes. Then, in [50], a comparison between the Mathieu series expansion and the point-matching method employed in [56] is made, and discrepancies are resolved. An experimental approach is presented in [20] for the cutoff calculation of some first higher-order modes in elliptical dielectric waveguides. The weakly guiding approach is again employed in [61] to determine the LP modes, by solving the Mathieu's equation with the use of the corresponding characteristic numbers. The same approach is also used in [74] by formulating the problem using the elliptical Fourier series expansion. Finally, the intensity distributions and the LP modes are calculated in [27], using a simplified characeristic equation based on Mathieu series expansion.

It is apparent that, although most of the research is made for the weakly guiding case where the formulation is considerably simplified, the general case was also studied, however, the exact characteristic equations obtained in [77] and [38] need more development to obtain a new characteristic equation valid at cutoff. Furthermore, compared with the circular case where we have two discrete analytical equations for the direct calculation of the cutoff wavenumbers [44], the corresponding analytical, closed-form characteristic equations, directly at cutoff, for the elliptical case do not exist.

Contents

Chapter 1

Oblique scattering by elliptical homogeneous dielectric cylinders

The electromagnetic scattering by an arbitrarily oriented elliptical cylinder having different constitutive parameters to those of the background medium, is treated in this chapter. The separation of variables method is used to solve this problem, but, due to the oblique incidence of the source fields, hybrid waves for the scattered and the induced fields are generated, thus making the formulation complicated. Moreover, because of the different wavenumbers between the scatterer and the background medium, the orthogonality relations for Mathieu functions do not hold, leading to more complicated systems, compared to those of normal incidence, that should be solved in order to get the solution for the scattered or the induced field. The validation of the results reveals the high accuracy of the implementation, even for electrically large scatterers. Both polarizations are considered and numerical results are given for various values of the parameters. The method is exact and can be used for reference as an alternative validation in future methods involving scattering problems.

1.1 Introduction

In this chapter the electromagnetic scattering of an obliquely incident plane wave by an elliptic dielectric cylinder is examined. The geometry of the scatterer is depicted in Figure 1.1 where the cross section corresponds to an ellipse having a major axis of length $2c_0$, a minor axis of length $2b_0$ and a focal distance of length $2a$. The incident wave is impinging with an angle ψ with respect to Ox semiaxis and with an angle θ with respect to Oz semiaxis. The permittivity and permeability of the scatterer is ϵ_2 and μ_2 respectively, while for the background medium is ϵ_1 and μ_1. All materials are lossless. Both polarizations are considered for oblique incidence, namely the case of

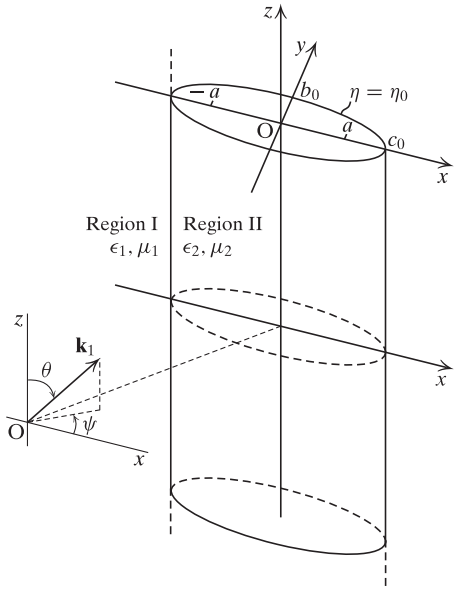

Figure 1.1 Geometry of the configuration.

a TM incident wave (or E-wave polarization) and the case of a TE incident wave (or H-wave polarization).

The problem is solved using the separation of variables based on the elliptical coordinate system. By expanding the fields in terms of radial and angular elliptical wave functions, there result four infinite sets of linear nonhomogeneous equations for the scattered electromagnetic field. Each of the four sets is actually built up from two equations involving both TM and TE waves, thus leading to a 2×2 linear system. Then the systems are solved numerically, by truncation, resulting to non-symmetric complex system matrices. Truncation of the systems should be appropriate in order not to increase the corresponding condition numbers. In the present case, due to the cross polarization of the fields, quantities including integrals of angular elliptical functions and their derivatives, should be carefully handled since non diagonal terms are present, unlike the case of isorefractive materials, or perfectly conducting boundaries, where only diagonal terms exist.

1.2 Scattering by a transverse magnetic incident wave

The case of an obliquely TM plane wave will be first considered. This has the form

$$\begin{aligned} E_z^{\mathrm{i}} &= \exp(-j\mathbf{k}_1 \cdot \mathbf{r}) \\ &= \exp[-jh_1(\cosh\eta\cos\varphi\cos\psi + \sinh\eta\sin\varphi\sin\psi)]\exp(-jk_1 z\cos\theta), \end{aligned} \tag{1.1}$$

where we have expressed the position vector $\mathbf{r}$ in terms of the elliptical coordinates (η, φ, z), by using the relations $x = a\cosh\eta\cos\varphi$ and $y = a\sinh\eta\sin\varphi$. In (1.1), $h_1 = \kappa_1 a = k_1 a\sin\theta = k_1 c_0 h\sin\theta$ with $h = a/c_0$ being the eccentricity of the ellipse and k_1 being the wavenumber of region I. The time dependence is of the form $\exp(j\omega t)$ and it is suppressed throughout. When $h = 0$ the ellipse degenerates into a circle of radius c_0 while, when $h = 1$, it degenerates into a strip of length $2a$. We

expand the incident wave of (1.1) in series of elliptical wave functions [47], namely

$$E_z^{\mathrm{i}} = \sqrt{8\pi}e^{-j\beta z}\Bigg[\sum_{m=0}^{\infty} j^{-m}\frac{Se_m(\kappa a,\cos\psi)}{M_m^{\mathrm{e}}(h_1)}Je_m(h_1,\cosh\eta)Se_m(h_1,\cos\varphi) + \sum_{m=1}^{\infty} j^{-m}\frac{So_m(\kappa a,\cos\psi)}{M_m^{\mathrm{o}}(h_1)}Jo_m(h_1,\cosh\eta)So_m(h_1,\cos\varphi)\Bigg], \tag{1.2}$$

where Je_m and Jo_m are the even and odd radial Mathieu functions of the first kind, Se_m and So_m are the even and odd angular Mathieu functions, and M_m^{e}, M_m^{o} are the normalization constants [47] and are given in Appendix 1.A. All quantities that relate integrals of angular Mathieu functions and the way to be calculated, are gathered in Appendix 1.A. In (1.2) we have defined $\beta = k_1\cos\theta$.

The scattered field in region I has both electric and magnetic field components, otherwise the boundary conditions cannot be satisfied. The z-component of the electric field is expressed as

$$E_z^{\mathrm{sc}} = e^{-j\beta z}\Bigg[\sum_{m=0}^{\infty} A_m^{\mathrm{e}}He_m^{(2)}(h_1,\cosh\eta)Se_m(h_1,\cos\varphi) + \sum_{m=1}^{\infty} A_m^{\mathrm{o}}Ho_m^{(2)}(h_1,\cosh\eta)So_m(h_1,\cos\varphi)\Bigg], \tag{1.3}$$

where $He_m^{(2)}$ and $Ho_m^{(2)}$ are the even and odd radial Mathieu functions of the fourth kind, sometimes also referred as Mathieu-Hankel functions. From now on, the superscript (2) will be ommited for simplicity. The corresponding magnetic scattered field H_z^{sc} in region I, has the same expression as E_z^{sc} in (1.3), but with different expansion coefficients, denoting them by D_m^{e} for the even part, and by D_m^{o} for the odd part.

The other components of the scattered electric and magnetic field are given by

$$E_\eta^{\mathrm{sc}} = -\frac{j\beta}{h_\eta\kappa_1^2}\frac{\partial E_z^{\mathrm{sc}}}{\partial\eta} - \frac{j\omega\mu_1}{h_\eta\kappa_1^2}\frac{\partial H_z^{\mathrm{sc}}}{\partial\varphi}, \tag{1.4}$$

$$E_\varphi^{\mathrm{sc}} = -\frac{j\beta}{h_\eta\kappa_1^2}\frac{\partial E_z^{\mathrm{sc}}}{\partial\varphi} + \frac{j\omega\mu_1}{h_\eta\kappa_1^2}\frac{\partial H_z^{\mathrm{sc}}}{\partial\eta}, \tag{1.5}$$

$$H_\eta^{\mathrm{sc}} = \frac{j\omega\epsilon_1}{h_\eta\kappa_1^2}\frac{\partial E_z^{\mathrm{sc}}}{\partial\varphi} - \frac{j\beta}{h_\eta\kappa_1^2}\frac{\partial H_z^{\mathrm{sc}}}{\partial\eta}, \tag{1.6}$$

$$H_\varphi^{\mathrm{sc}} = -\frac{j\omega\epsilon_1}{h_\eta\kappa_1^2}\frac{\partial E_z^{\mathrm{sc}}}{\partial\eta} - \frac{j\beta}{h_\eta\kappa_1^2}\frac{\partial H_z^{\mathrm{sc}}}{\partial\varphi}, \tag{1.7}$$

where $h_\eta = a\sqrt{\cosh^2\eta - \cos^2\varphi}$ is the scale factor of the elliptical coordinate system.

The induced field in region II has again both electric and magnetic field components. The induced electric field is expressed as

$$E_z^2 = e^{-j\beta z}\Bigg[\sum_{m=0}^{\infty} C_m^{\mathrm{e}}Je_m(h_2,\cosh\eta)Se_m(h_2,\cos\varphi) + \sum_{m=1}^{\infty} C_m^{\mathrm{o}}Jo_m(h_2,\cosh\eta)So_m(h_2,\cos\varphi)\Bigg], \tag{1.8}$$

while the induced magnetic field H_z^2 has the same expresseion as E_z^2 in (1.8), but with the expansion coefficients F_m^{e} and F_m^{o} for the even and odd part, respectively. Since (1.8) describes the fields in region II, it includes the quantity $h_2 = \kappa_2 a = k_2 a \sin\theta$, with k_2 being the wavenumber of region II. Relating the inner wavenumber to the outer wavenumber by $k_2 = \sqrt{\epsilon_2/\epsilon_1}\sqrt{\mu_2/\mu_1}k_1$, we express h_2 in terms of h_1 by $h_2 = \sqrt{\epsilon_2/\epsilon_1}\sqrt{\mu_2/\mu_1}h_1$.

The boundary condition for the z-component of the electric field on the elliptical boundary $\eta = \eta_0$ is

$$E_z^{\mathrm{i}} + E_z^{\mathrm{sc}} = E_z^2. \tag{1.9}$$

Upon satisfying it, and after using the orthogonality properties for the angular Mathieu functions, one time for the even function and one time for the odd function, we express the inner expansion coefficient in terms of the outer expansion coefficient by the formula

$$\begin{aligned} C_i^{\mathrm{e}} = \Bigg[& \sqrt{8\pi} \sum_{m=0}^{\infty} j^{-m} \frac{\mathrm{Se}_m(\kappa a, \cos\psi)}{M_m^{\mathrm{e}}(h_1)} \mathrm{Je}_m(\kappa a, \cosh\eta_0) M_{mi}^{\mathrm{e}}(h_1, h_2) \\ & + \sum_{m=0}^{\infty} A_m^{\mathrm{e}} \mathrm{He}_m(\kappa a, \cosh\eta_0) M_{mi}^{\mathrm{e}}(h_1, h_2) \Bigg] \\ & / \big[\mathrm{Je}_i(\gamma a, \cosh\eta_0) M_i^{\mathrm{e}}(h_2)\big], \quad i = 0, 1, 2, \ldots . \end{aligned} \tag{1.10}$$

A similar expression is obtained for the C_i^{o} coefficient versus A_m^{o} by simply changing "e" with "o" in (1.10), and carefully do the summation by starting m from 1 to ∞. The quantity M_{mi}^{e} appearing in (1.10) and its corresponding M_{mi}^{o} are defined in Appendix 1.A.

A similar procedure is followed to satisfy the boundary condition for the magnetic field intensity, namely

$$H_z^{\mathrm{sc}} = H_z^2, \tag{1.11}$$

and leads to the following relation

$$\begin{aligned} F_i^{\mathrm{e}} = & \sum_{m=0}^{\infty} D_m^{\mathrm{e}} \mathrm{He}_m(\kappa a, \cosh\eta_0) M_{mi}^{\mathrm{e}}(h_1, h_2) \\ & / \big[\mathrm{Je}_i(\gamma a, \cosh\eta_0) M_i^{\mathrm{e}}(h_2)\big], \quad i = 0, 1, 2, \ldots , \end{aligned} \tag{1.12}$$

where we have again expressed the inner expansion coefficient in terms of the outer one. Similarly as before, by changing "e" with "o" in (1.12), and interpreting the summation index m from 1 to ∞, we can express F_i^{o} in terms of D_i^{o}.

Substituting the expansion coefficients C_i^{e} and C_i^{o} in (1.8) by their equivalent from (1.10), we obtain the electric field intensity in region II given by a double series and expressed via the only unknowns now, namely, the expansion coefficients of region I. The same is done for the magnetic field H_z^2 by substituting F_i^{e} and F_i^{o} by their equivalents from (1.12).

Next, we proceed with the boundary condition at $\eta = \eta_0$ for the φ-components of the electric field, namely

$$-\frac{\beta}{\kappa_1^2}\frac{\partial E_z^{\mathrm{i}}}{\partial\varphi} - \frac{\beta}{\kappa_1^2}\frac{\partial E_z^{\mathrm{sc}}}{\partial\varphi} + \frac{\omega\mu_1}{\kappa_1^2}\frac{\partial H_z^{\mathrm{sc}}}{\partial\eta} = -\frac{\beta}{\kappa_2^2}\frac{\partial E_z^2}{\partial\varphi} + \frac{\omega\mu_2}{\kappa_2^2}\frac{\partial H_z^2}{\partial\eta}. \tag{1.13}$$

Multiplying both sides of (1.13) with $\mathrm{Se}_s(h_1, \cos\varphi)$ and integrating from $\varphi = 0$ to 2π,

there results the set of linear nonhomogeneous equations

$$
\begin{aligned}
&-\cos\theta\sum_{m=1}^{\infty} Ho_m(\kappa a,\cosh\eta_0)M_{ms}^{o'e}(h_1)A_m^o \\
&+\cos\theta\frac{\epsilon_1}{\epsilon_2}\frac{\mu_1}{\mu_2}\sum_{m=1}^{\infty}\sum_{i=1}^{\infty}A_m^o Ho_m(\kappa a,\cosh\eta_0)\frac{M_{mi}^o(h_1,h_2)M_{is}^{o'e}(h_2,h_1)}{M_i^o(h_2)} \\
&+\sqrt{\frac{\mu_1}{\epsilon_1}}He_s'(\kappa a,\cosh\eta_0)M_s^e(h_1)D_s^e \\
&-\sqrt{\frac{\mu_1}{\epsilon_1}}\frac{\epsilon_1}{\epsilon_2}\sum_{m=0}^{\infty}\sum_{i=0}^{\infty}D_m^e He_m(\kappa a,\cosh\eta_0)\frac{Je_i'(\gamma a,\cosh\eta_0)}{Je_i(\gamma a,\cosh\eta_0)}\frac{M_{mi}^e(h_1,h_2)M_{is}^e(h_2,h_1)}{M_i^e(h_2)} \\
&=\cos\theta\sqrt{8\pi}\sum_{m=1}^{\infty}j^{-m}So_m(\kappa a,\cos\psi)Jo_m(\kappa a,\cosh\eta_0)\frac{M_{ms}^{o'e}(h_1)}{M_m^o(h_1)} \\
&-\cos\theta\frac{\epsilon_1}{\epsilon_2}\frac{\mu_1}{\mu_2}\sqrt{8\pi}\sum_{m=1}^{\infty}\sum_{i=1}^{\infty}j^{-m}\frac{So_m(\kappa a,\cos\psi)}{M_m^o(h_1)}Jo_m(\kappa a,\cosh\eta_0) \\
&\times\frac{M_{mi}^o(h_1,h_2)M_{is}^{o'e}(h_2,h_1)}{M_i^o(h_2)},\quad s=0,1,2,\ldots.
\end{aligned}
\tag{1.14}
$$

In eqs. (1.14), the primed versions of Mathieu functions denote the derivative with respect to η. Moreover, the various quantities $M_{ms}^{o'e}$ appearing, are defined in Appendix 1.A. The set of (1.14) relates the odd part expansion coefficient of the scattered electric field with the even part expansion coefficient of the scattered magnetic field, hence leading to a cross relation. Since we have two unknowns here, another set is necessary, and thus a 2×2 linear system is composed. This extra set is obtained if we satisfy the final boundary condition for the φ-components of the magnetic field, namely

$$
-\frac{\omega\epsilon_1}{\kappa_1^2}\frac{\partial E_z^i}{\partial\eta}-\frac{\omega\epsilon_1}{\kappa_1^2}\frac{\partial E_z^{sc}}{\partial\eta}-\frac{\beta}{\kappa_1^2}\frac{\partial H_z^{sc}}{\partial\varphi}=-\frac{\omega\epsilon_2}{\kappa_2^2}\frac{\partial E_z^2}{\partial\eta}-\frac{\beta}{\kappa_2^2}\frac{\partial H_z^2}{\partial\varphi}. \tag{1.15}
$$

Next we multiply both sides of (1.15) with $So_s(h_1,\cos\varphi)$ and integrate from $\varphi=0$ to 2π. The resulting set has the same structure for LHS (left hand side) as that of eqs. (1.14), but different RHS (right hand side). To obtain the LHS we should make the following substitutions; for the part relating A_m^o: $A_m^o\to D_m^e$, "o"→"e", "o′e"→"e′o" and $\mu_i\leftrightarrow\epsilon_i, i=1,2$. For the part relating D_m^e: $D_m^e\to -A_m^o$, "e"→"o" and $\mu_i\leftrightarrow\epsilon_i, i=1,2$. The RHS of this new set would be given by

$$
\begin{aligned}
\text{RHS}=&\sqrt{\frac{\epsilon_1}{\mu_1}}\sqrt{8\pi}j^{-s}So_s(\kappa a,\cos\psi)Jo_s'(\kappa a,\cosh\eta_0) \\
&-\sqrt{\frac{\epsilon_1}{\mu_1}}\frac{\mu_1}{\mu_2}\sqrt{8\pi}\sum_{m=1}^{\infty}\sum_{i=1}^{\infty}j^{-m}\frac{So_m(\kappa a,\cos\psi)}{M_m^o(h_1)}Jo_m(\kappa a,\cosh\eta_0)\frac{Jo_i'(\gamma a,\cosh\eta_0)}{Jo_i(\gamma a,\cosh\eta_0)} \\
&\times\frac{M_{mi}^o(h_1,h_2)M_{is}^o(h_2,h_1)}{M_i^o(h_2)},\quad s=1,2,3,\ldots.
\end{aligned}
\tag{1.16}
$$

Instead of the aforementioned sets, there is also another couple which composes a new 2×2 linear system. This new group relates the remaining expansion coefficients A_m^e and D_m^o. To obtain it, we should multiply both parts of (1.13) with $So_s(h_1,\cos\varphi)$

and integrate from $\varphi = 0$ to 2π, and do the same with (1.15) with $Se_s(h_1, \cos\varphi)$. Alternatively, the last one can be directly obtained from the previous sets by using the following substitutions: $A_m^{\mathrm{o}} \to D_m^{\mathrm{o}}$, $D_m^{\mathrm{e}} \to -A_m^{\mathrm{e}}$ and $\mu_i \leftrightarrow \epsilon_i, i = 1, 2$.

1.3 Scattering by a transverse electric incident wave

To derive the case for a TE incident wave, a complete duality scheme should be applied to all formulas of Section 1.2. Therefore, the following changes are necessary: $E \to H$, $H \to -E$ and $\mu_i \leftrightarrow \epsilon_i, i = 1, 2$.

1.4 Scattering cross sections

Yeh [79] has obtained the back scattering cross section for a dielectric elliptical cylinder but this was for normal incidence. In the present case, due to the cross polarization, the full definition should be used to derive an expression that includes the contribution of both electric and magnetic field components. The bistatic cross section is given by [4]

$$\sigma = \lim_{\rho\to\infty} 2\pi\rho \frac{|\mathbf{E}^{\mathrm{sc}}|^2}{|\mathbf{E}^{\mathrm{i}}|^2}, \tag{1.17}$$

where ρ is the distance from the origin in polar coordinates. Then, using the asymptotic expansion for the radial Mathieu function of the fourth kind [8, 43] and the corresponding one for its derivative [43]

$$H_{\substack{\mathrm{e}\\ \mathrm{o}}m}(h_1, \cosh\eta) \underset{\eta\to\infty}{\sim} \sqrt{\frac{1}{\kappa_1 a \cosh\eta}}\, \exp[-j(\kappa_1 a \cosh\eta - (2m+1)\pi/4)], \tag{1.18}$$

$$H'_{\substack{\mathrm{e}\\ \mathrm{o}}m}(h_1, \cosh\eta) \underset{\eta\to\infty}{\sim} \sqrt{\frac{1}{\kappa_1 a \cosh\eta}}\, \exp[-j(\kappa_1 a \cosh\eta - (2m+1)\pi/4)] \times \Big(-\frac{1}{2} - j\kappa_1 a \cosh\eta\Big), \tag{1.19}$$

in (1.3), in the corresponding one for the H_z^{sc}, and finally in eqs. (1.4)–(1.7), we obtain the scattered far field expressions. Next, using the asymptotic formula $a\cosh\eta \sim_{\eta\to\infty} \rho$ and keeping only the dominant terms in the fields, i.e. those with $1/\sqrt{\rho}$ compared with $1/\rho^{3/2}$ behaviour, we finally obtain from (1.17) the back or radar ($k_1\sigma_{\mathrm{b}}$), the forward ($k_1\sigma_{\mathrm{f}}$), and the total ($k_1 Q_{\mathrm{t}}$) normalized scattering cross sections. For the back and forward scattering cross sections we have

$$k_1\sigma_{\mathrm{b}} = \frac{2\pi}{\sin\theta}\left[|G^E(\psi+\pi)|^2 + \frac{\mu_1}{\epsilon_1}|G^H(\psi+\pi)|^2\right], \tag{1.20}$$

$$k_1\sigma_{\mathrm{f}} = \frac{2\pi}{\sin\theta}\left[|G^E(\psi)|^2 + \frac{\mu_1}{\epsilon_1}|G^H(\psi)|^2\right], \tag{1.21}$$

where

$$\begin{bmatrix} G^E(\varphi) \\ G^H(\varphi) \end{bmatrix} = \sum_{m=0}^{\infty} \begin{bmatrix} A_m^{\mathrm{e}} \\ D_m^{\mathrm{e}} \end{bmatrix} j^m Se_m(h_1, \cos\varphi) + \sum_{m=1}^{\infty} \begin{bmatrix} A_m^{\mathrm{o}} \\ D_m^{\mathrm{o}} \end{bmatrix} j^m So_m(h_1, \cos\varphi). \tag{1.22}$$

Compared with the expressions in [70, 72], it is obvious that we have two contributions due to the hybrid fields, and also, the additional factor $1/\sin\theta$ that appears in the

oblique incidence case. For the total scattering cross section, defined in [54], we have

$$k_1 Q_{\rm t} = \frac{1}{\sin\theta}\left\{ \sum_{m=0}^{\infty} |A_m^{\rm e}|^2 M_m^{\rm e}(h_1) + \sum_{m=1}^{\infty} |A_m^{\rm o}|^2 M_m^{\rm o}(h_1) + \frac{\mu_1}{\epsilon_1}\left[\sum_{m=0}^{\infty} |D_m^{\rm e}|^2 M_m^{\rm e}(h_1) + \sum_{m=1}^{\infty} |D_m^{\rm o}|^2 M_m^{\rm o}(h_1) \right]\right\}. \tag{1.23}$$

1.5 Numerical results

A part of this section is devoted to the validation and then, various examples are presented and discussed. A complete vectorized Matlab package was developed for the calculation of Mathieu functions. Mathieu functions used here correspond to the ones used in [47]. For accurate calculation, the implementation on the computer was based on [7]. The linear systems, see for instance eqs. (1.14), are solved numerically, by truncation. What is important here is the appropriate choice of the truncation order so not to add terms that will cause the system to become ill-posed. So, we first check the projection of all RHS and we select the order of truncation to be such that terms of the order 10^{-16} and smaller not to be included. Then the system is solved using the Gauss elimination technique with partial pivoting.

1.5.1 Validation

The verification of the code was mainly preferred to be done by comparison with other, different techniques from the literature. In [72] the authors provide a table for the calculation of the back, forward and total scattering cross sections, under normal incidence. In Table 1.1 a comparison is made between the current work and theirs. In the special case where $\epsilon_2/\epsilon_1 \to \infty$ and $\mu_2/\mu_1 \to 0$, the results of this work conclude to the ones of [70] where the cylinder is a perfect conductor, and again under normal incidence. The agreement is shown in Table 1.1 where the extreme ratios $\epsilon_2/\epsilon_1 = 10^{15}$ and $\mu_2/\mu_1 = 10^{-15}$ have been used. From the values presented, it is evident that an agreement has been achieved between 4 and 7 significant digits. However this comparison is not sufficient to conclude about the accuracy of the present technique, since their algorithm is based on a series expansion of the eccentricity h, up to the order h^4.

Next, in Table 1.2, a comparison is made with [68] where the author has obtained highly accurate results with the use of the exponentially converging Nyström method. The author refers to an elliptic dielectric cylinder which is isorefractive compared to the background medium, and he presents results for the total tangential field components in the case where both TM and TE incident waves are present. The results are found to agree within 8 and 14 significant figures, thus revealing the high accuracy of the implementation of the work presented here. We further check for the correctness for an electrically large scatterer having $k_1 c_0 \simeq 25.5\pi$. As can be also seen from Table 1.2, even for this difficult case, the present algorithm yields an astonishing agreement up to 13 significant digits. The truncation order for the case of $k_1 c_0 \simeq 0.7\pi$ is 7 for TM and 9 for TE wave, respectively. For the case of $k_1 c_0 \simeq 25.5\pi$ is 39 terms for TM and 43 terms for TE wave, respectively. It is obvious that when $k_1 c_0$ factor increases, more terms are required if one wants to achieve high accuracy.

A self validation was also investigated with the satisfaction of the boundary conditions across the elliptic interface, which were verified to an excellent accuracy, as can be seen from Table 1.3. This constitutes a further check for the correctness of the present solution.

Table 1.1 Comparison with the closed form relations in [70, 72] based on asymptotic analysis.*

Method	$k_1\sigma_b$	$k_1\sigma_f$	k_1Q_t
TM incident wave: $\epsilon_2/\epsilon_1 = 2.54$, $\mu_2/\mu_1 = 1$, $k_1c_0 = \pi$, $\psi = 0°$			
[72]	6.18706994547200	184.9067048022120	27.12969561215999
This work	6.18703078788406	184.9062767761573	27.12969292771986
TE incident wave: $\epsilon_2/\epsilon_1 = 2.54$, $\mu_2/\mu_1 = 1$, $k_1c_0 = \pi$, $\psi = 45°$			
[72]	0.115168127410000	150.5303809198020	24.3357001800960
This work	0.115130859315949	150.5304186551822	24.3357297357245
TM incident wave: $\epsilon_2/\epsilon_1 = 10^{15}$, $\mu_2/\mu_1 = 10^{-15}$, $k_1c_0 = 1.4\pi$, $\psi = 18°$			
[70]	14.05514176996800	115.7218114169368	20.7438342450944
This work	14.05518224090304	115.7217521256232	20.7438734217104
TE incident wave: $\epsilon_2/\epsilon_1 = 10^{15}$, $\mu_2/\mu_1 = 10^{-15}$, $k_1c_0 = 1.4\pi$, $\psi = 63°$			
[70]	14.71419260987070	55.49594687890344	14.3503668802168
This work	14.71418407607783	55.49594751732805	14.3503702205521

*The values of the parameters used are $h = 0.1$ and $\theta = 90°$. The rest parameters are shown in the subcategories.

Finally, in Table 1.4 we see how the series of (1.3) for the magnetic field converge, together with the accompanied accuracy. For this case, whose values of parameters are shown also in Table 1.4, 21 terms are needed to catch an accuracy of 16 significant digits. The number of terms is increased when k_1c_0 is increased, as can be seen from Table 1.5, where now 75 terms are needed for the series of (1.3) for the electric field. In companion with Tables 1.4 and 1.5 are Figures 1.2 and 1.3 which illustrate the convergence for each case. It should be also noted that a verification on the convergence, like the one presented here, were also done for the series which accrue after differentiation, i.e. for the φ–components of the fields obtained from (1.5) and (1.7).

1.5.2 Applications

The presentation of the results is organized as follows: figures 1.4–1.10 have double vertical axes. The blue curves should be read with respect to the left axis while the red curves with respect to the right axis. We correspond to the left axis the results obtained for the case of a TM incident wave, while, to the right axis, the results from a TE incident wave. The values of the parameters are gathered in the captions.

In Figure 1.4 we plot the back scattering cross section versus the incident angle ψ for various tilt angles θ. The plot is in the range from 0 to 90 degrees, since the results are symmetric about the 90 degrees, as it is imposed by the geometry of the scatterer. We observe a highly sensitive behaviour for the back scattering cross section to the change of ψ, and this also depends on the change of the obliquely incident angle. When the incidence is very oblique (solid curves with $\theta = 10°$) the variation of $k_1\sigma_b$ is very smooth, then continuing on an intermediate value of θ (dashed curves with $\theta = 30°$), the variation starts to become important, and finally, when we approach the normal incidence (dash-dot curves with $\theta = 80°$), the variation becomes very large. The higher variation of $k_1\sigma_b$ in the case of the TE incident wave, is also evident.

The same remarks can be concluded from Figure 1.5 where we plot the forward scattering cross section. Again, when the tilt angle becomes larger, the sensitivity

Table 1.2 Comparison with the highly accurate results from [68].*

Method	$\lvert E_z(c_0,0)\rvert$	$\lvert E_\varphi(c_0,0)\rvert \times 10^{-3}$	$\lvert H_z(c_0,0)\rvert$	$\lvert H_\varphi(c_0,0)\rvert$
		$h = 0.6, k_1 c_0 = \sqrt{2}/2\pi \simeq 0.7\pi$		
[68]	0.776688259004	0.2868380432785	1.1811246931802	0.9340099224582
This work	0.776688259004474	0.286838044996280	1.181124693180238	0.934009922481920
Agreement†	12	8	14	10
		$h = \sqrt{1-(5/9)^2} \simeq 0.83, k_1 c_0 = 18\sqrt{2}\pi \simeq 25.5\pi$		
[68]	0.875644527	0.367480904914	1.126681679992	0.97588563775
This work	0.875644527300915	0.367480907116272	1.126681679992637	0.975885637775567
Agreement†	9	8	13	10

*The values of the parameters used here to match the ones in [68] are $\epsilon_1 = \sqrt{2}\epsilon_0$, $\mu_1 = \sqrt{2}\mu_0$, $\epsilon_2 = 2\epsilon_0$, $\mu_2 = \mu_0$, $\psi = 180° + 45°$ and $\theta = 180° - 40°$. The rest parameters are shown in the subcategories.
†In significant digits, compared with the results from [68].

Table 1.3 Satisfaction of the boundary conditions for all tangential field components.*

Total field	$\lvert E_z(c_0,0)\rvert$	$\lvert E_\varphi(c_0,0)\rvert$	$\lvert H_z(c_0,0)\rvert$	$\lvert H_\varphi(c_0,0)\rvert$
		TM incident wave		
Region I	0.8756445273012264	0.6943660706634107	$2.471228083329778 \times 10^{-18}$	0.00325649435176776
Region II	0.8756445273012283	0.6943660706632128	$2.471228083329779 \times 10^{-18}$	0.00325649435176778
Agreement†	14	12	15	14
		TE incident wave		
Region I	$3.289348154170203 \times 10^{-13}$	366.7869354856975	1.126681679992637	0.9791420617943808
Region II	$3.289348154170201 \times 10^{-13}$	366.7869354856993	1.126681679992639	0.9791420617943756
Agreement†	15	14	15	13

*The values of the parameters used are $\epsilon_1 = \sqrt{2}\epsilon_0$, $\mu_1 = \sqrt{2}\mu_0$, $\epsilon_2 = 2\epsilon_0$, $\mu_2 = \mu_0$, $\psi = 180° + 45°$, $\theta = 180° - 40°$, $h = \sqrt{1-(5/9)^2} \simeq 0.83$ and $k_1 c_0 = 18\sqrt{2}\pi \simeq 25.5\pi$.
†In significant digits.

Table 1.4 Convergence history for the series of (1.3), for the total magnetic field $|H_z(c_0, 0)|$ and for the case of a TM incident wave.*

NOT†	Partial sums $\times 10^{-3}$§	Accuracy¶
1	0.02762426190699955	0
2	0.264942507449316	0
3	0.722386463238474	0
4	0.7952206859739518	1
5	0.7891377483696527	1
6	1.356086513119823	0
7	1.315115632795634	2
8	1.28279671507106	1
9	1.283629533523272	3
10	1.283928540306621	4
11	1.283937221266295	5
12	1.283938171065654	6
13	1.283938130469819	8
14	1.283938203254267	7
15	1.283938199378802	7
16	1.283938198295074	9
17	1.283938198309434	10
18	1.283938198307139	12
19	1.28393819830737	13
20	1.283938198307413	13
21	1.283938198307413	16
22	1.283938198307413	16
23	1.283938198307413	16

*The values of the parameters used are $\epsilon_1 = \epsilon_0$, $\mu_1 = \mu_0$, $\epsilon_2 = 4\epsilon_0$, $\mu_2 = \mu_0$, $\psi = 45°$, $\theta = 30°$, $h = 0.4$ and $k_1 c_0 = 2\pi$.
†Number of terms.
§When the 15th decimal digit is missing, it is understood to be zero.
¶In significant digits, compared with the previous term.

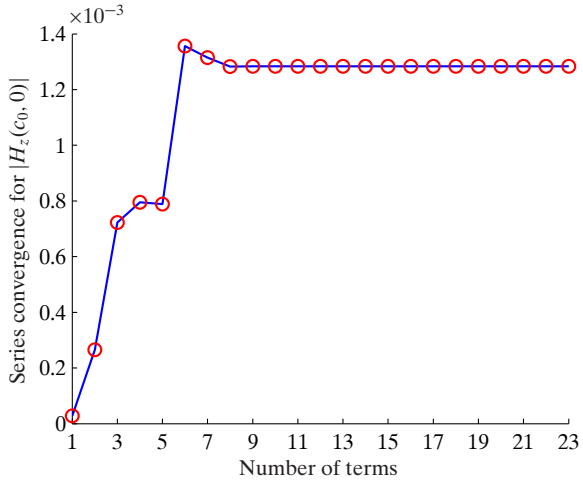

Figure 1.2 (color online). Illustration of the convergence for the data in Table 1.4.

Table 1.5 Convergence history for the series of (1.3), for the total electric field $|E_z(c_0, 0)|$ and for the case of a TE incident wave.*

NOT†	Partial sums ×10	Accuracy§
60	4.456426094495426	8
61	4.456426108333447	7
62	4.456426103445187	9
63	4.456426102262043	9
64	4.456426102687145	10
65	4.456426102779209	10
66	4.456426102745657	11
67	4.456426102739103	11
68	4.456426102741522	11
69	4.456426102741951	13
70	4.456426102741792	13
71	4.456426102741765	14
72	4.456426102741775	14
73	4.456426102741776	15
74	4.456426102741775	15
75	4.456426102741775	16
76	4.456426102741775	16
77	4.456426102741775	16

*The values of the parameters used are $\epsilon_1 = \epsilon_0$, $\mu_1 = \mu_0$, $\epsilon_2 = 4\epsilon_0$, $\mu_2 = \mu_0$, $\psi = 45°$, $\theta = 60°$, $h = 0.6$ and $k_1 c_0 = 10\pi$.
†Number of terms.
§In significant digits, compared with the previous term.

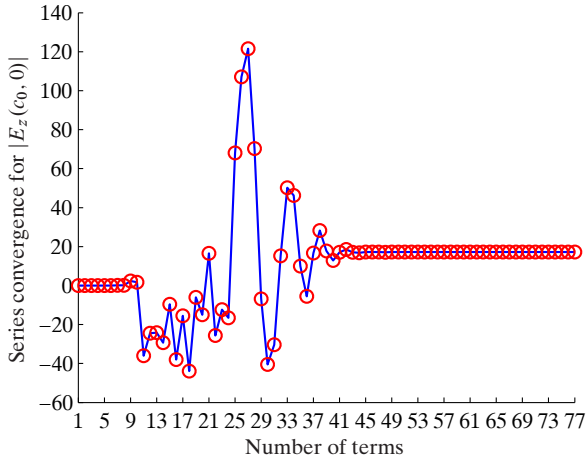

Figure 1.3 (color online). Illustration of the convergence for the data in Table 1.5.

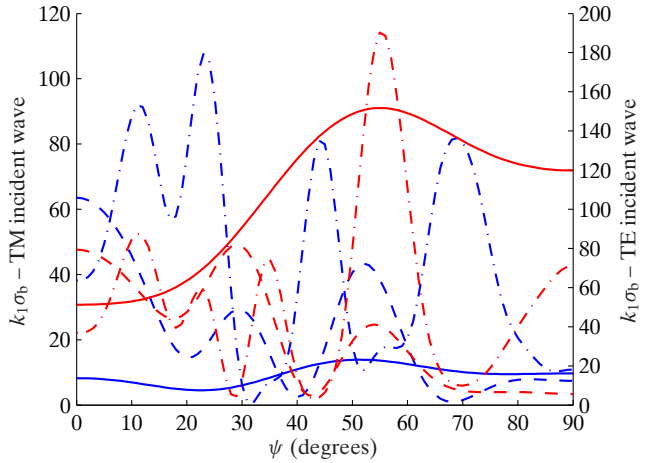

Figure 1.4 (color online). Back scattering cross section versus the incident angle ψ. Solid curves: $\theta = 10°$, dashed curves: $\theta = 30°$, dash-dot curves: $\theta = 80°$. The rest parameters are $\epsilon_2/\epsilon_1 = 2.54$, $\mu_2/\mu_1 = 1$, $h = 0.6$ and $k_1 c_0 = 10\pi$.

begins to appear, however, it is not so high compared to that of $k_1\sigma_b$. Moreover, we observe the same behaviour for both the two cases of TM and TE incident wave, which means that $k_1\sigma_f$ has almost the same magnitude for both TM and TE cases to the change of ψ. This is not true for the case of $k_1\sigma_b$.

As far as the total scattering cross section is concerned, see Figure 1.6, the same comments apply like the ones for the forward scattering cross section, although the magnitude of $k_1 Q_t$ is much smaller compared to that of $k_1\sigma_f$.

The sensitivity also depends on the size of the scatterer. In Figure 1.7 we illustrate this change for obliquely incident TM and TE waves at $\theta = 60°$, for two different values of $k_1 c_0 = 2\pi$ and $k_1 c_0 = 20\pi$, by plotting the back scattering

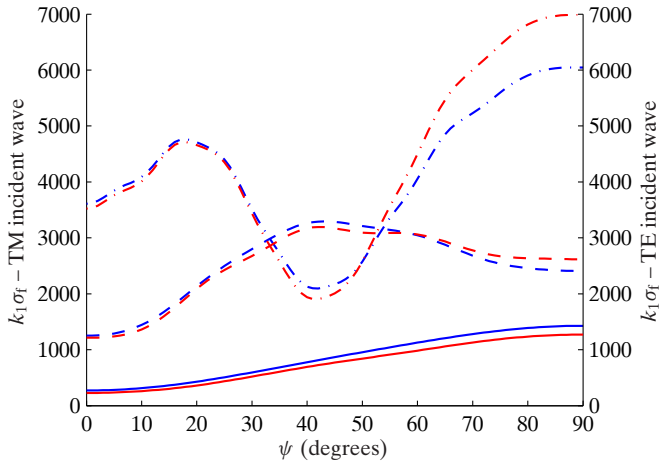

Figure 1.5 (color online). Forward scattering cross section versus the incident angle ψ. Solid curves: $\theta = 10°$, dashed curves: $\theta = 30°$, dash-dot curves: $\theta = 80°$. The other values of the parameters are the same as in Figure 1.4.

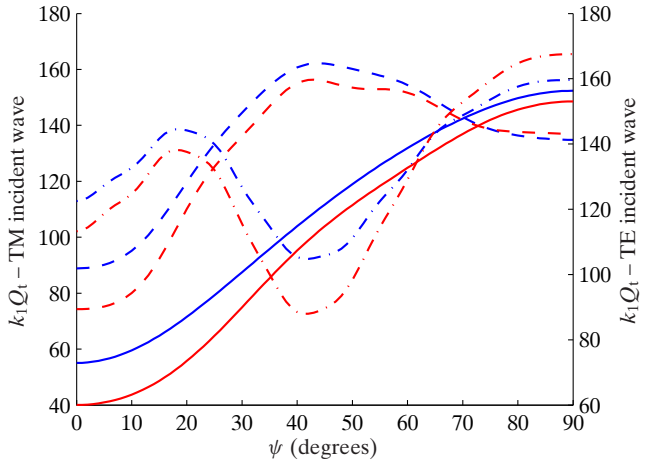

Figure 1.6 (color online). Total scattering cross section versus the incident angle ψ. Solid curves: $\theta = 10°$, dashed curves: $\theta = 30°$, dash-dot curves: $\theta = 80°$. The other values of the parameters are the same as in Figure 1.4.

cross sections in dB $[10\log(k_1\sigma_b)]$. Alternatively, these values of k_1c_0 correspond to $c_0/\lambda_1 = 1$ and $c_0/\lambda_1 = 10$, respectively, so we have a scatterer having size of the order of the incident wavelength, and one ten times bigger. The much higher sensitivity and variation of $k_1\sigma_b$ to the change of ψ, when the size of the scatterer is electrically increased, is evident.

Figure 1.8 shows the results for the back scattering cross section when we increase the eccentricity to the value $h = 0.8$. By comparing the results between Figure 1.8 and Figure 1.4 that correspond to very oblique incident angles (solid curves), we see that, when the ellipse approaches the strip, $k_1\sigma_b$ becomes more

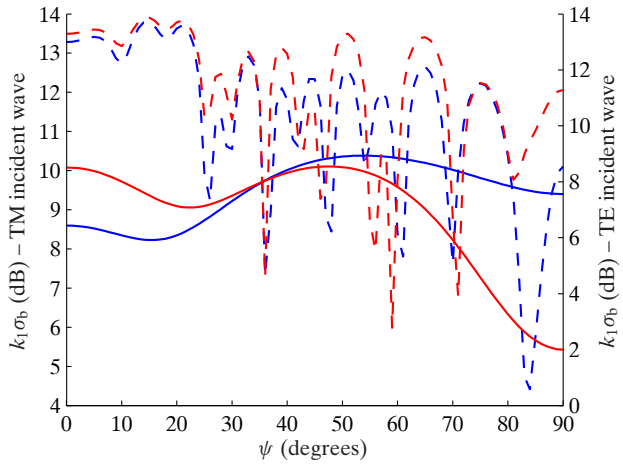

Figure 1.7 (color online). Back scattering cross section in dB versus the incident angle ψ, and for $\theta = 60°$. Solid curves: $k_1c_0 = 2\pi$, dashed curves: $k_1c_0 = 20\pi$. The other values of the parameters are the same as in Figure 1.4.

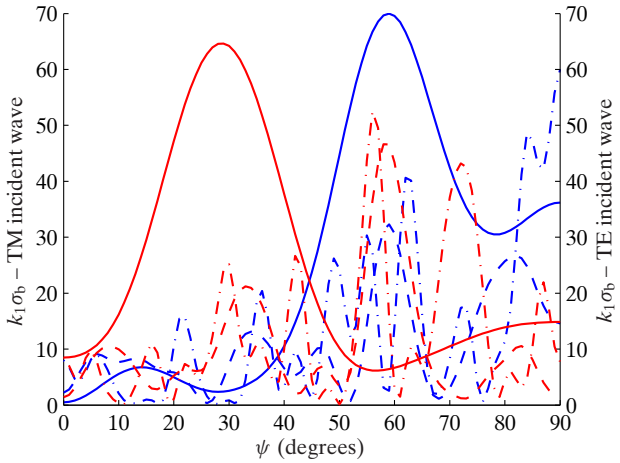

Figure 1.8 (color online). Back scattering cross section versus the incident angle ψ, but for $h = 0.8$. Solid curves: $\theta = 10°$, dashed curves: $\theta = 30°$, dash-dot curves: $\theta = 80°$. The other values of the parameters are the same as in Figure 1.4.

sensitive. For larger tilt angles, it continues to have the high sensitivity, but now the variation is larger for smaller tilt angles.

The back scattering cross section in dB (same definition as in Figure 1.7) versus the electrical size of the scatterer, is shown in Figure 1.9 for $\theta = 10°$, and in Figure 1.10 using $\theta = 70°$. The peaks and troughs observed in both Figures 1.9 and 1.10 correspond to resonant frequencies. The variation and the sensitivity becomes higher when the oblique incident angle increases, as we have already noticed from Figure 1.4. Moreover, in the case of $\theta = 70°$, the variations are greater for the TM incident wave.

Finally, in Table 1.6 we have gathered some results for the computed absolute

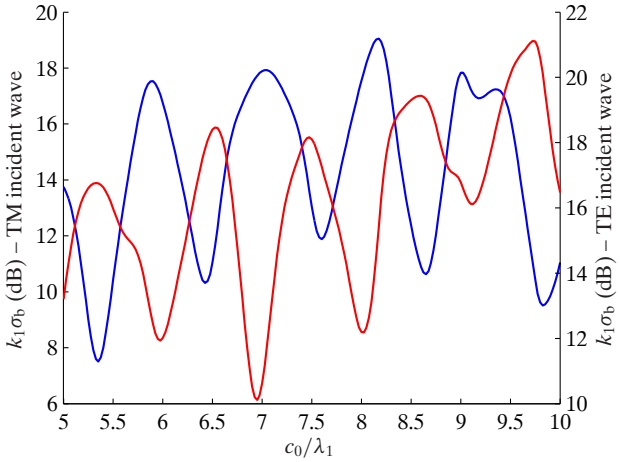

Figure 1.9 (color online). Back scattering cross section in dB versus c_0/λ_1 for $\psi = 45°$, $\theta = 10°$ and $h = 0.8$. The other values of the parameters are the same as in Figure 1.4.

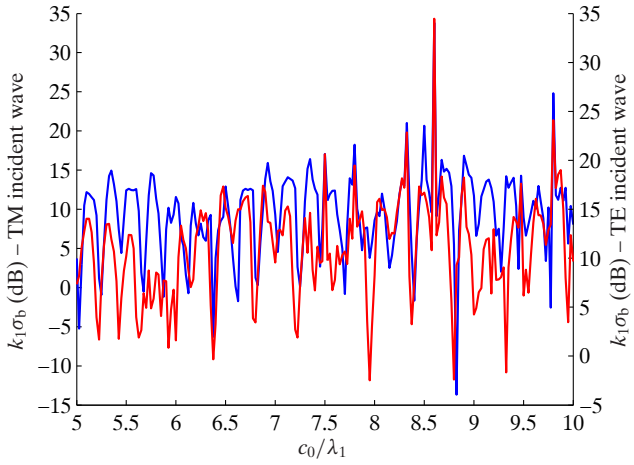

Figure 1.10 (color online). Back scattering cross section in dB versus c_0/λ_1 for $\psi = 45°$, $\theta = 70°$ and $h = 0.8$. The other values of the parameters are the same as in Figure 1.4.

value of all tangential field components. The values depicted are for the total fields, i.e. the scattered fields plus the incident fields, and on the elliptic boundary at the location $(\eta, \varphi) = (\eta_0, 0)$. This would be an extra comparison for validation in future methods.

1.6 Concluding remarks

The electromagnetic scattering of obliquely incident plane waves by an elliptical dielectric cylinder was studied by using the SVM. The formulation supports different constitutive parameters between the scatterer and the background medium, thus leading to nonorthogonal relations for angular Mathieu functions. Moreover, the existence of hybrid waves makes the formulation challenging, since the linear system from which the solution is retrieved is composed by coupled equations, which include both the contribution of TM and TE waves. The validation of the correctness of the method was achieved by comparing it with exact, closed-form relations, and additionally, with a highly accurate formulation. The results agree up to 13 significant figures, even for electrically large scatterers, thus revealing the highly accurate implementation. Finally, we observe and discuss the sensitivity of the scattering cross sections upon the change of various parameters. This method is exact, is alternative to the ones based on integral equations, and can be used for reference as an alternative validation for future methods involving scattering problems.

1.A Appendix: Integrals of angular Mathieu functions

In eqs. (1.2)–(1.16) integrals of various combinations of angular Mathieu functions have been used. The normalization constants M_m^{e} and M_m^{o} given in [47] can be calculated by substituting the definition series of angular Mathieu functions and

Table 1.6 Absolute value of all tangential total field components on the elliptic boundary.*

$\lvert E_z(c_0,0)\rvert$	$\lvert E_\varphi(c_0,0)\rvert$	$\lvert H_z(c_0,0)\rvert \times 10^{-4}$	$\lvert H_\varphi(c_0,0)\rvert \times 10^{-4}$
TM incident wave: $\epsilon_2/\epsilon_1 = 2.54$, $\mu_2/\mu_1 = 1$, $h = 0.1$, $k_1c_0 = \pi$, $\psi = 0°$, $\theta = 90°$			
2.288195081322319	0	0	38.94375263319
TM incident wave: $\epsilon_2/\epsilon_1 = 4$, $\mu_2/\mu_1 = 1$, $h = 0.4$, $k_1c_0 = \pi$, $\psi = 45°$, $\theta = 90°$			
1.803748227000343	0	0	35.38444329392
TM incident wave: $\epsilon_2/\epsilon_1 = 4$, $\mu_2/\mu_1 = 1$, $h = 0.4$, $k_1c_0 = 2\pi$, $\psi = 45°$, $\theta = 30°$			
0.209999529274052	0.195835248999691	12.83938198307	6.033213282836437
TM incident wave: $\epsilon_2/\epsilon_1 = 4$, $\mu_2/\mu_1 = 1$, $h = 0.6$, $k_1c_0 = 10\pi$, $\psi = 45°$, $\theta = 60°$			
0.223189903589569	0.196366158675060	6.225945875074653	7.885613920065429
$\lvert E_z(c_0,0)\rvert \times 10^2$	$\lvert E_\varphi(c_0,0)\rvert \times 10^2$	$\lvert H_z(c_0,0)\rvert$	$\lvert H_\varphi(c_0,0)\rvert$
TE incident wave: $\epsilon_2/\epsilon_1 = 2.54$, $\mu_2/\mu_1 = 1$, $h = 0.1$, $k_1c_0 = \pi$, $\psi = 45°$, $\theta = 90°$			
0	1.173910610000002	0.669332490754157	0
TE incident wave: $\epsilon_2/\epsilon_1 = 4$, $\mu_2/\mu_1 = 1$, $h = 0.4$, $k_1c_0 = \pi$, $\psi = 45°$, $\theta = 90°$			
0	1.395819481293371	1.180533876627618	0
TE incident wave: $\epsilon_2/\epsilon_1 = 4$, $\mu_2/\mu_1 = 1$, $h = 0.4$, $k_1c_0 = 2\pi$, $\psi = 45°$, $\theta = 30°$			
1.652547067361899	3.570317971553199	0.515151791537351	0.343462144368244
TE incident wave: $\epsilon_2/\epsilon_1 = 4$, $\mu_2/\mu_1 = 1$, $h = 0.6$, $k_1c_0 = 10\pi$, $\psi = 45°$, $\theta = 60°$			
0.44564261027417750	1.295488454706039	0.086477523480504	0.163519279817156

*The values of the parameters are shown in the subcategories.

using the orthogonality relations for cosines and sines. Therefore we have

$$\int_0^{2\pi} S_{\substack{\mathrm{e}\\\mathrm{o}}m}(h,\cos\varphi)S_{\substack{\mathrm{e}\\\mathrm{o}}v}(h,\cos\varphi)d\varphi = M^{\substack{\mathrm{e}\\\mathrm{o}}}_{mv}(h)\delta_{mv}$$
$$= \delta_{mv}2\pi\sum_{\substack{n=0\\1}}^{\infty}\frac{1}{\varepsilon_n}B^{\substack{\mathrm{e}\\\mathrm{o}}}_n(h,m)B^{\substack{\mathrm{e}\\\mathrm{o}}}_n(h,v), \qquad (1.A.1)$$

with $m, v = 0, 1, 2\ldots$. Stacked notation have been used to include both even and odd functions at once. In (1.A.1), δ_{mv} is Kronecker's delta while $\varepsilon_0 = 1$, $\varepsilon_n = 2, (n \geq 1)$ is Neumann factor and $B^{\substack{\mathrm{e}\\\mathrm{o}}}_n(h,l)$ are the expansion coefficients for Mathieu functions [47]. l and n should be both even or both odd. All these types of integrals we encounter here, actually define a $M \times V$ matrix whose elements are defined for every different value of m and v, e.g. $m = 0, 1, 2, \ldots, M$ and $v = 0, 1, 2, \ldots, V$. In the case of (1.A.1), only the diagonal terms are present, while the rest are all zero.

The inner products $M^{\mathrm{e}}_{mv}(h_1,h_2)$ and $M^{\mathrm{o}}_{mv}(h_1,h_2)$ are not orthogonal and are given by [72]

$$\int_0^{2\pi} S_{\substack{\mathrm{e}\\\mathrm{o}}m}(h_1,\cos\varphi)S_{\substack{\mathrm{e}\\\mathrm{o}}v}(h_2,\cos\varphi)d\varphi = M^{\substack{\mathrm{e}\\\mathrm{o}}}_{mv}(h_1,h_2)$$
$$= 2\pi\sum_{\substack{n=0\\1}}^{\infty}\frac{1}{\varepsilon_n}B^{\substack{\mathrm{e}\\\mathrm{o}}}_n(h_1,m)B^{\substack{\mathrm{e}\\\mathrm{o}}}_n(h_2,v), \qquad (1.A.2)$$

with $m, v = 0, 1, 2\ldots$. These constitute a full $M \times V$ matrix. When $h_1 = h_2$, (1.A.2) degenerates to (1.A.1).

Combinations between angular functions and their derivatives are also present, namely

$$\int_0^{2\pi} S'_{\substack{\mathrm{e}\\\mathrm{o}}m}(h,\cos\varphi)S_{\substack{\mathrm{o}\\\mathrm{e}}v}(h,\cos\varphi)d\varphi = M^{\substack{\mathrm{e}'\mathrm{o}\\\mathrm{o}'\mathrm{e}}}_{mv}(h)$$
$$= \mp\pi\sum_{n=1}^{\infty}nB^{\substack{\mathrm{e}\\\mathrm{o}}}_n(h,m)B^{\substack{\mathrm{o}\\\mathrm{e}}}_n(h,v), \qquad (1.A.3)$$

where the primed angular functions denote the derivative with respect to φ. In this case, even if we have the same h, these inner products are not orthogonal, such as that of (1.A.1). Moreover, due to the existence of one even type and one odd type function, indices m and v are not the same, as it is also in (1.A.1). Therefore, these quantities constitute a full $M \times V$ matrix. When we speak for even indices, the one should begin from 0 and the other from 2. When we speak for odd indices, both should begin from 1. For $M^{\mathrm{e}'\mathrm{o}}_m$, $m = 0, 1, 2, \ldots$ and $v = 1, 2, 3, \ldots$ while for $M^{\mathrm{o}'\mathrm{e}}_m$, $m = 1, 2, 3, \ldots$ and $v = 0, 1, 2, \ldots$. This fact results in a combination of existence of both diagonal and non diagonal terms when setting up the systems. For instance, in the set of eqs. (1.14), the matrix that corresponds to D^{e}_m consists of full matrix elements plus diagonal terms, while, the matrix that corresponds to A^{o}_m, consists of two full matrix elements.

The more general inner product

$$\int_0^{2\pi} S'_{\substack{\mathrm{e}\\\mathrm{o}}m}(h_1,\cos\varphi)S_{\substack{\mathrm{o}\\\mathrm{e}}v}(h_2,\cos\varphi)d\varphi = M^{\substack{\mathrm{e}'\mathrm{o}\\\mathrm{o}'\mathrm{e}}}_{mv}(h_1,h_2)$$
$$= \mp\pi\sum_{n=1}^{\infty}nB^{\substack{\mathrm{e}\\\mathrm{o}}}_n(h_1,m)B^{\substack{\mathrm{o}\\\mathrm{e}}}_n(h_2,v), \qquad (1.A.4)$$

also appears. All the aforementioned commentaries are valid. When $h_1 = h_2$, (1.A.4) also degenerates to (1.A.3).

Chapter 2

Hybrid wave propagation in elliptical dielectric waveguides

The cutoff wavenumbers of the elliptical dielectric waveguide are calculated exactly and analytically. Two separate methods are used to solve this problem. The first method is based on the separation of variables technique using Mathieu functions and gives the exact cutoff wavenumbers. The system matrices of which the roots of their determinant should be determined are complicated because of the nonexistence of orthogonality relations for Mathieu functions, due to the different constitutive parameters between the core and the cladding of the fiber. In the second method the cutoff wavenumbers are obtained through analytical expressions, when the eccentricity h of the elliptical core is specialized to small values. In the latter case, analytical, closed-form algebraic expressions, free of Mathieu functions, are obtained for the expansion coefficients $g_{sn}^{(2)}$ in the resulting relation $x_{1,sn}(h) = x_{1,sn}(0)\left[1 + g_{sn}^{(2)}h^2 + \mathcal{O}(h^4)\right]$ for the cutoff wavenumbers, where $x_{1,sn}(0)$ are the normalized cutoff wavenumbers of the circular dielectric waveguide. These expressions are valid for every different value of s and n, corresponding to every higher-order hybrid mode. Numerical results are given for various higher-order modes, as well as a comparison with the exact solution.

2.1 Introduction

The geometry of the waveguide is depicted in Figure 2.1, where the core corresponds to an ellipse having a major axis of length $2c_0$, a minor axis of length $2b_0$, and an interfocal distance of length $2a$. The permittivity and the permeability of the core are ϵ_1 and μ_1, respectively, while, for the cladding, they are ϵ_2 and μ_2. For our calculations, the core and the cladding materials are non magnetic, therefore implying that $\mu_1 = \mu_2 = \mu_0$, the permeability of free space.

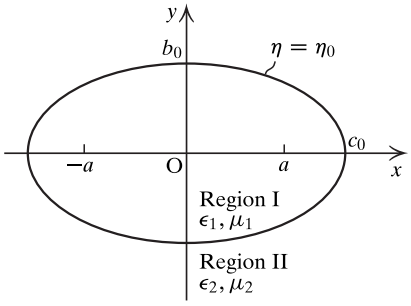

Figure 2.1 Geometry of the dielectric waveguide.

We follow two alternative procedures to obtain the solution of the problem. The first one, based on the separation of variables in terms of Mathieu functions, leads us to homogeneous linear systems at cutoff that, compared with the metallic elliptical waveguide, are much more complicated due to the different constitutive parameters in the core and the cladding. This fact leads to nonorthogonal relations between angular Mathieu functions and their derivatives. A similar complication is obtained in the scattering by obliquely incident waves – as we have seen in Chapter 1 – however, in the present case, the situation is more challenging since the roots of the determinant of this complex system are required, and not simply the solution of the system, which is the case when dealing with the scattering problem. Setting this determinant equal to zero, it defines the characteristic equation of the problem the roots of which are the exact cutoff wavenumbers for the corresponding hybrid modes. Once the system is set up, the roots are obtained using numerical techniques that will be described later.

When the solution is specialized to small values of the eccentricity $h = a/c_0$, i.e. for $h \ll 1$, analytical, closed-form expressions are obtained for the normalized cutoff wavenumbers in the form $x_{1,sn}(h) = x_{1,sn}(0)\left[1 + g_{sn}^{(2)}h^2 + \mathcal{O}(h^4)\right]$. The expansion coefficients $g_{sn}^{(2)}$ for all higher-order modes are given by analytical, closed-form expressions, independent of h, while $x_{1,sn}(0)$ corresponds to the circular dielectric waveguide having radius c_0. Once a new cutoff wavenumber is of interest, the first procedure will require repetition, from the beginning, for each different h or for each different mode order. On the other hand, once the coefficients $g_{sn}^{(2)}$ are known, the same cutoff wavenumber is immediately obtained by a quick "back-of-the envelope" calculation, for each small value of h and for every order.

The independent exact solution described above is required to validate the correctness of the results of our analytical, closed-form formulas, and to conclude about the restriction of how much small values of h should be used in order to keep low errors in our analytical formulas.

2.2 Mode designation

The elliptical dielectric waveguide maintains only hybrid modes, due to the lack of the symmetry of the configuration. As can be also seen from [38, 51, 56], these modes are split in two groups, the even modes and the odd modes. According to the definition given by Yeh in [77], the even group of modes corresponds to the equations that describe the even part of the axial component of the magnetic field, while, the odd group of modes corresponds to the equations that describe the odd part of the

axial magnetic field component. Each group has two different types of hybrid modes, i.e. the EH and HE. So we speak about ${}_{\mathrm{e}}\mathrm{EH}_{sn}$ and ${}_{\mathrm{e}}\mathrm{HE}_{sn}$ modes for the even group and for ${}_{\mathrm{o}}\mathrm{EH}_{sn}$ and ${}_{\mathrm{o}}\mathrm{HE}_{sn}$ modes for the odd group.

The circular dielectric waveguide has four different types of modes, namely the TM_{0n}, TE_{0n}, EH_{sn} and HE_{sn}. The characteristic equations at cutoff for the aforementioned modes, which will give the corresponding normalized cutoff wavenumbers $x_{1,sn}(0) = x^0_{1,sn}$, are [44]

$$\mathrm{TM}_{0n}, \mathrm{TE}_{0n} : J_0(x^0_{1,0n}) = 0, \quad n \geqslant 1 \tag{2.1}$$

$$\mathrm{EH}_{sn} : J_s(x^0_{1,sn}) = 0, \quad x_{1,sn} \neq 0,\ s \geqslant 1,\ n \geqslant 1 \tag{2.2}$$

$$\mathrm{HE}_{1n} : J_1(x^0_{1,1n}) = 0, \quad n \geqslant 1 \tag{2.3}$$

$$\mathrm{HE}_{sn} : J_{s-1}(x^0_{1,sn}) \left(\frac{\epsilon_1}{\epsilon_2} + 1\right) - \frac{x^0_{1,sn} J_s(x^0_{1,sn})}{s-1} = 0, \quad s \geqslant 2,\ n \geqslant 1 \tag{2.4}$$

When the eccentricity of the ellipse becomes zero, the core of the fiber becomes a circle with radius c_0. In this case, the ${}_{\mathrm{o}}\mathrm{EH}_{0n}$ and ${}_{\mathrm{e}}\mathrm{HE}_{0n}$ modes degenerate to TM_{0n} and TE_{0n} modes, respectively, while ${}_{\mathrm{e}}\mathrm{EH}_{sn}$ and ${}_{\mathrm{o}}\mathrm{EH}_{sn}$ ($s \geqslant 1$) degenerate to EH_{sn}, and finally, ${}_{\mathrm{e}}\mathrm{HE}_{sn}$ and ${}_{\mathrm{o}}\mathrm{HE}_{sn}$ ($s \geqslant 1$) degenerate to HE_{sn}.

2.3 Solution of the problem

The longitudinal component E_z^{I} in the region I of the waveguide is expressed as

$$E_z^{\mathrm{I}}(\eta, \varphi, z) = e^{-j\beta z}\left[\sum_{m=0}^{\infty} A_m^{\mathrm{e}} \mathrm{Je}_m(\kappa a, \cosh\eta)\mathrm{Se}_m(\kappa a, \cos\varphi) + \sum_{m=1}^{\infty} A_m^{\mathrm{o}} \mathrm{Jo}_m(\kappa a, \cosh\eta)\mathrm{So}_m(\kappa a, \cos\varphi)\right] \tag{2.5}$$

where η, φ are the transverse elliptical-cylindrical coordinates with respect to $x\mathrm{O}y$, Je_m and Jo_m are the even and odd radial Mathieu functions of the first kind, Se_m and So_m are the even and odd angular Mathieu functions, β is the propagation constant, and $\kappa = (k_1^2 - \beta^2)^{1/2}$ with $k_1 = \omega(\epsilon_1\mu_1)^{1/2}$ being the wavenumber of region I. The time dependence $\exp(j\omega t)$ is suppressed throughout. A similar expression as in (2.5) also holds for the axial component H_z^{I} with the only difference that the expansion coefficients A_m^{e} and A_m^{o} are replaced by D_m^{e} and D_m^{o}, respectively.

In region II, the longitudinal component E_z^{II} is expressed as

$$E_z^{\mathrm{II}}(\eta, \varphi, z) = e^{-j\beta z}\left[\sum_{m=0}^{\infty} C_m^{\mathrm{e}} \mathrm{Ke}_m(\gamma a, \cosh\eta)\mathrm{Se}_m(-j\gamma a, \cos\varphi) + \sum_{m=1}^{\infty} C_m^{\mathrm{o}} \mathrm{Ko}_m(\gamma a, \cosh\eta)\mathrm{So}_m(-j\gamma a, \cos\varphi)\right] \tag{2.6}$$

where Ke_m and Ko_m are the even and odd modified radial Mathieu functions of the second kind, defined in [1] by the relation $K_{\substack{\mathrm{e}\\\mathrm{o}}m}(\gamma a, \cosh\eta) = \pi/2 \exp[-j(m+1)\pi/2] H^{(2)}_{\substack{\mathrm{e}\\\mathrm{o}}m}(-j\gamma a, \cosh\eta)$, with $H^{(2)}_{\substack{\mathrm{e}\\\mathrm{o}}m}$ being the even/odd radial Mathieu functions of the fourth kind. Moreover, in (2.6) $\gamma = (\beta^2 - k_2^2)^{1/2}$ with $k_2 = \omega(\epsilon_2\mu_2)^{1/2}$ being the

wavenumber of region II. Again, a similar expression to that in (2.6) holds also for the longitudinal component H_z^{II} with the only difference that the expansion coefficients C_m^{e} and C_m^{o} are replaced by F_m^{e} and F_m^{o}, respectively.

Satisfying the boundary conditions for the longitudinal components at the elliptical interface $\eta = \eta_0$, namely $E_z^{\mathrm{I}} = E_z^{\mathrm{II}}$ and $H_z^{\mathrm{I}} = H_z^{\mathrm{II}}$, and using the orthogonality properties for the angular Mathieu functions, we can express the expansion coefficients of the fields in region II in terms of the expansion coefficients of the fields in region I, i.e. $C^{\mathring{\mathrm{o}}}{}_m$ in terms of $A^{\mathring{\mathrm{o}}}{}_m$ and $F^{\mathring{\mathrm{o}}}{}_m$ in terms of $D^{\mathring{\mathrm{o}}}{}_m$.

Next, we need to satisfy the remaining two boundary conditions at $\eta = \eta_0$ for the tangential components, namely $E_\varphi^{\mathrm{I}} = E_\varphi^{\mathrm{II}}$ and $H_\varphi^{\mathrm{I}} = H_\varphi^{\mathrm{II}}$. The procedure is similar to the one described in Chapter 1, so the details will not be repeated here for brevity. Expressing the tangential component in terms of the longitudinal one, we end up with the following two sets of linear homogeneous equations. The first set is

$$\begin{aligned}
&\frac{-\beta}{\kappa^2}\sum_{m=0}^{\infty} \mathrm{Je}_m(\kappa a, \cosh\eta_0) M_{ms}^{\mathrm{e'o}}(\kappa a) A_m^{\mathrm{e}} \\
&- \frac{\beta}{\gamma^2}\sum_{m=0}^{\infty}\sum_{i=0}^{\infty} A_m^{\mathrm{e}} \mathrm{Je}_m(\kappa a, \cosh\eta_0) M_{mi}^{\mathrm{e}}(\kappa a, -j\gamma a) \frac{M_{is}^{\mathrm{e'o}}(-j\gamma a, \kappa a)}{M_i^{\mathrm{e}}(-j\gamma a)} \\
&+ \frac{\omega\mu_1}{\kappa^2} \mathrm{Jo}_s'(\kappa a, \cosh\eta_0) M_s^{\mathrm{o}}(\kappa a) D_s^{\mathrm{o}} \\
&+ \frac{\omega\mu_2}{\gamma^2}\sum_{m=1}^{\infty}\sum_{i=1}^{\infty} D_m^{\mathrm{o}} \mathrm{Jo}_m(\kappa a, \cosh\eta_0) \frac{\mathrm{Ko}_i'(\gamma a, \cosh\eta_0)}{\mathrm{Ko}_i(\gamma a, \cosh\eta_0)} \\
&\times \frac{M_{mi}^{\mathrm{o}}(\kappa a, -j\gamma a) M_{is}^{\mathrm{o}}(-j\gamma a, \kappa a)}{M_i^{\mathrm{o}}(-j\gamma a)} = 0, \quad s \geqslant 1
\end{aligned} \tag{2.7}$$

while the second one is resulting from (2.7) using the following substitutions: $A_m^{\mathrm{e}} \to D_m^{\mathrm{o}}$, "e"→"o", "e′o"→"o′e", and $D_m^{\mathrm{o}} \to -A_m^{\mathrm{e}}$, "o"→"e", $\mu_{1,2} \to \epsilon_{1,2}$. Moreover, the index s should be interpreted from 0 to ∞. In (2.7), the primed versions of Mathieu functions denote their derivative with respect to η, while the various inner products $M_{ms}^{\mathrm{e'o}}$ etc, are all gathered in the Appendix 1.A. It is worthy to note the difference between the inner products of this problem and the ones appeared in Chapter 1, where here some contain imaginary arguments.

These two sets described above, define a 2×2 homogeneous system with unknowns the expansion coefficients for the even part of E_z^{I} and for the odd part of H_z^{I}. Thus, according to the description given in Section 2.2, the roots of the determinant of this system matrix will describe the ${}_{\mathrm{o}}\mathrm{EH}_{sn}$ and ${}_{\mathrm{o}}\mathrm{HE}_{sn}$ modes. Apart from the aforementioned homogeneous system, there exists its dual one, which occurs if we simply swap $A_m^{\mathrm{e}} \to D_m^{\mathrm{e}}$, $D_m^{\mathrm{o}} \to -A_m^{\mathrm{o}}$, and $\epsilon_{1,2} \leftrightarrow \mu_{1,2}$. The roots of the determinant of this dual system matrix will describe the ${}_{\mathrm{e}}\mathrm{EH}_{sn}$ and ${}_{\mathrm{e}}\mathrm{HE}_{sn}$ modes. Further details about the indices s and n of the modes will be given in the Appendix 2.A, where we further develop these sets to reach the cutoff condition. The roots obtained from these determinants will give us the exact cutoff wavenumbers, and will be used to conclude about the accuracy of our analytical formulas.

Now we proceed with the expansion of (2.7) and its pair, by keeping terms up to the order of h^2. The expansions for small eccentricities of the Mathieu functions appearing in (2.7) are given in [35], the expansions for small eccentricities of the various inner products M^{e} and M^{o} are given in the Appendix of [72], while, the expansions for small eccentricities of $M^{\mathrm{e'o}}$ and of $M^{\mathrm{o'e}}$, are given in the Appendix 2.B. Substituting these expansions in (2.7), we conclude to the following infinite linear set

of homogeneous equations for the expansion coefficients A_s^{e} and D_s^{o}:

$$
\begin{aligned}
&-\beta\left(\frac{1}{\kappa^2}\alpha_{s-2,s}+\frac{1}{\gamma^2}a_{s-2,s}\right)A_{s-2}^{\mathrm{e}}+\omega\frac{\mu_2}{\gamma^2}b_{s-2,s}D_{s-2}^{\mathrm{o}}\\
&-\beta\left(\frac{1}{\kappa^2}\alpha_{ss}+\frac{1}{\gamma^2}a_{ss}\right)A_s^{\mathrm{e}}+\omega\left(\frac{\mu_1}{\kappa^2}\delta_{ss}+\frac{\mu_2}{\gamma^2}b_{ss}\right)D_s^{\mathrm{o}}\\
&-\beta\left(\frac{1}{\kappa^2}\alpha_{s+2,s}+\frac{1}{\gamma^2}a_{s+2,s}\right)A_{s+2}^{\mathrm{e}}+\omega\frac{\mu_2}{\gamma^2}b_{s+2,s}D_{s+2}^{\mathrm{o}}=0,\quad s=1,2,\ldots
\end{aligned}
\tag{2.8}
$$

The quantities α, a, δ and b appearing in (2.8) are defined in the Appendix 2.B. Following a similar procedure for the pair of (2.7), we conclude to a second infinite linear set of homogeneous equations, namely:

$$
\begin{aligned}
&-\omega\frac{\epsilon_2}{\gamma^2}a'_{s-2,s}A_{s-2}^{\mathrm{e}}-\beta\left(\frac{1}{\kappa^2}\delta'_{s-2,s}+\frac{1}{\gamma^2}b'_{s-2,s}\right)D_{s-2}^{\mathrm{o}}\\
&-\omega\left(\frac{\epsilon_1}{\kappa^2}\alpha'_{ss}+\frac{\epsilon_2}{\gamma^2}a'_{ss}\right)A_s^{\mathrm{e}}-\beta\left(\frac{1}{\kappa^2}\delta'_{ss}+\frac{1}{\gamma^2}b'_{ss}\right)D_s^{\mathrm{o}}\\
&-\omega\frac{\epsilon_2}{\gamma^2}a'_{s+2,s}A_{s+2}^{\mathrm{e}}-\beta\left(\frac{1}{\kappa^2}\delta'_{s+2,s}+\frac{1}{\gamma^2}b'_{s+2,s}\right)D_{s+2}^{\mathrm{o}}=0,\quad s=1,2,\ldots
\end{aligned}
\tag{2.9}
$$

The various primed quantities are again gathered in the Appendix 2.B. Although the index s in (2.9) should start from zero, the zero value has been extracted so that both sets (2.8) and (2.9) start from $s=1$. These two sets constitute a 2×2 homogeneous system from which the ${}_{\mathrm{o}}\mathrm{EH}_{sn}$ and ${}_{\mathrm{o}}\mathrm{HE}_{sn}$, $s\geqslant 1$ modes are described. The aforementioned special case where $s=0$ in (2.9) contributes to ${}_{\mathrm{o}}\mathrm{EH}_{0n}$ modes, and the corresponding linear homogeneous equation, valid only for the expansion coefficient A_0^{e}, is given by

$$
-\omega\left(\frac{\epsilon_1}{\kappa^2}\alpha'_{00}+\frac{\epsilon_2}{\gamma^2}a'_{00}\right)A_0^{\mathrm{e}}-\omega\frac{\epsilon_2}{\gamma^2}a'_{2,0}A_2^{\mathrm{e}}=0 \tag{2.10}
$$

The linear system defined by (2.8) and (2.9) has the same form with the one appearing in [37], although in the present case it is homogeneous. As it is described in [37], in order to analytically express its determinant, these two sets must be transformed beforehand so that the off diagonal terms in the determinant be free of zero order terms in their expansions. Before we proceed with the transformation, it is more convenient to rewrite (2.8) and (2.9) in a compact form by introducing the quantities

$$
\zeta_{ss}^{(\ell)}=\frac{1}{\kappa^2}\alpha_{ss}^{(\ell)}+\frac{1}{\gamma^2}a_{ss}^{(\ell)},\quad \ell=0,2 \tag{2.11}
$$

$$
\xi_{ss}^{(\ell)}=\frac{\mu_1}{\kappa^2}\delta_{ss}^{(\ell)}+\frac{\mu_2}{\gamma^2}b_{ss}^{(\ell)},\quad \ell=0,2 \tag{2.12}
$$

$$
\zeta_{ss}^{\prime(\ell)}=\frac{\epsilon_1}{\kappa^2}\alpha_{ss}^{\prime(\ell)}+\frac{\epsilon_2}{\gamma^2}a_{ss}^{\prime(\ell)},\quad \ell=0,2 \tag{2.13}
$$

$$
\xi_{ss}^{\prime(\ell)}=\frac{1}{\kappa^2}\delta_{ss}^{\prime(\ell)}+\frac{1}{\gamma^2}b_{ss}^{\prime(\ell)},\quad \ell=0,2 \tag{2.14}
$$

which correspond to the expressions inside the brackets of the coefficients A_s^{e} and D_s^{o} in (2.8) and (2.9). So, for instance, the expression inside the brackets of A_s^{e} in (2.8) is written compactly as $\zeta_{ss}=\zeta_{ss}^{(0)}+\zeta_{ss}^{(2)}h^2$. To achieve the aforementioned transformation, we multiply (2.8) with $\beta/\xi_{ss}^{(0)}$ and (2.9) with $\omega/\xi_{ss}^{\prime(0)}$ and then we add

their corresponding terms. We do the same by multiplying (2.8) with $\omega/\zeta_{ss}^{(0)}$ and (2.9) with $-\beta/\zeta_{ss}'^{(0)}$ and then we add their corresponding terms. In the two new resulting sets, we substitute ω by $k_2/(\epsilon_2\mu_2)^{1/2}$ and the ratio k_2^2/β^2 by $(\gamma^2+\kappa^2)/[\epsilon_1\mu_1/(\epsilon_2\mu_2)\gamma^2+\kappa^2]$ and finally we arrive at two sets having the form

$$t_{s-2,s}A_{s-2}^{\mathrm{e}} + u_{s-2,s}D_{s-2}^{\mathrm{o}} + t_{ss}A_s^{\mathrm{e}} + u_{ss}D_s^{\mathrm{o}} + t_{s+2,s}A_{s+2}^{\mathrm{e}} + u_{s+2,s}D_{s+2}^{\mathrm{o}} = 0, \quad s = 1, 2, \ldots \tag{2.15}$$

$$t'_{s-2,s}A_{s-2}^{\mathrm{e}} + u'_{s-2,s}D_{s-2}^{\mathrm{o}} + t'_{ss}A_s^{\mathrm{e}} + u'_{ss}D_s^{\mathrm{o}} + t'_{s+2,s}A_{s+2}^{\mathrm{e}} + u'_{s+2,s}D_{s+2}^{\mathrm{o}} = 0, \quad s = 1, 2, \ldots \tag{2.16}$$

In (2.15) and (2.16), $t_{ss} = T_{ss}^{(0)} + T_{ss}^{(2)}h^2 + \mathcal{O}(h^4)$ and $u'_{ss} = U_{ss}'^{(0)} + U_{ss}'^{(2)}h^2 + \mathcal{O}(h^4)$ where

$$T_{ss}^{(\ell)} = -\frac{\zeta_{ss}^{(\ell)}}{\xi_{ss}^{(0)}} - \frac{1}{\epsilon_2\mu_2}\frac{\gamma^2+\kappa^2}{\frac{\epsilon_1\mu_1}{\epsilon_2\mu_2}\gamma^2+\kappa^2}\frac{\zeta_{ss}'^{(\ell)}}{\xi_{ss}'^{(0)}}, \quad \ell = 0, 2 \tag{2.17}$$

$$U_{ss}'^{(\ell)} = \frac{\xi_{ss}'^{(\ell)}}{\zeta_{ss}'^{(0)}} + \frac{1}{\epsilon_2\mu_2}\frac{\gamma^2+\kappa^2}{\frac{\epsilon_1\mu_1}{\epsilon_2\mu_2}\gamma^2+\kappa^2}\frac{\xi_{ss}^{(\ell)}}{\zeta_{ss}^{(0)}}, \quad \ell = 0, 2 \tag{2.18}$$

while $t'_{ss} = T_{ss}'^{(2)}h^2 + \mathcal{O}(h^4)$ and $u_{ss} = U_{ss}^{(2)}h^2 + \mathcal{O}(h^4)$. The quantities $t_{s\pm2,2}$, $u_{s\pm2,2}$, $t'_{s\pm2,2}$ and $u'_{s\pm2,2}$ lead to $\mathcal{O}(h^4)$ terms and do not contribute in the development of the determinant up to $\mathcal{O}(h^2)$ studied here. Since $T_{ss}'^{(0)} = U_{ss}^{(0)} = 0$, the aforementioned requirement for the off diagonal terms of the determinant is now fulfilled. Therefore, the expansion of the determinant Δ is [37]

$$\Delta = P(t_{qq})P(u'_{qq})\Bigg[1 - \sum_{\substack{w=0\\1}}^{\infty}\frac{t_{w,w+2}t_{w+2,w}}{t_{ww}t_{w+2,w+2}} - \sum_{\substack{w=0\\1}}^{\infty}\frac{u_{ww}t'_{ww}}{t_{ww}u'_{ww}} - \sum_{\substack{w=0\\1}}^{\infty}\frac{u_{w,w+2}t'_{w+2,w}}{t_{ww}u'_{w+2,w+2}} - \sum_{\substack{w=0\\1}}^{\infty}\frac{t'_{w,w+2}u_{w+2,w}}{u'_{ww}t_{w+2,w+2}} - \sum_{\substack{w=0\\1}}^{\infty}\frac{u'_{w,w+2}u'_{w+2,w}}{u'_{ww}u'_{w+2,w+2}}\Bigg] \tag{2.19}$$

with the quantity $P(t_{qq})$ given in [36]

$$P(t_{qq}) = P(T_{qq}^{(0)})\Bigg[1 + h^2\sum_{\substack{w=0\\1}}^{\infty}\frac{T_{ww}^{(2)}}{T_{ww}^{(0)}} + \mathcal{O}(h^4)\Bigg] \tag{2.20}$$

$$P(T_{qq}^{(0)}) = T_{ww}^{(0)}T_{w+2,w+2}^{(0)}\cdots, \quad w = 0, 1 \tag{2.21}$$

Stacked notation has been used to include both even and odd indices at once. Same expressions as (2.20) and (2.21) hold for $P(u'_{qq})$ as well by substituting "T" with "U'". By inspection, all the summations in (2.19) are $\mathcal{O}(h^4)$, so the expansion of the determinant by keeping terms up to the order h^2 is

$$\Delta = P(T_{qq}^{(0)})P(U_{qq}'^{(0)})\Bigg\{1 + h^2\sum_{\substack{w=0\\1}}^{\infty}\left[\frac{T_{ww}^{(2)}}{T_{ww}^{(0)}} + \frac{U_{ww}'^{(2)}}{U_{ww}'^{(0)}}\right] + \mathcal{O}(h^4)\Bigg\} \tag{2.22}$$

As it is also explained in [36], the quantities $T_{ww}^{(\ell)}$ and $U_{ww}'^{(\ell)}$ depend on the normalized cutoff wavenumbers $x_1 = x_1(h) = \kappa c_0$ of the elliptical dielectric waveguide that have values near $x_1(0)$, i.e. the cutoff wavenumbers of the circular dielectric waveguide.

This means that $x_1(h)$ is expanded around $h = 0$ as $x_1(h) = x_1^{(0)} + x_1^{(2)}h^2 + \mathcal{O}(h^4)$. In the remaining text we use x_1^0 instead of $x_1^{(0)}$ for simplicity. Therefore, these values are retrieved by setting the quantity inside the curly brackets in (2.22) equal to zero. Expanding $T_{ss}^{(\ell)}(x_1(h))$ in Maclaurin series in terms of h, we get

$$\begin{aligned} T_{ss}^{(\ell)}(x_1(h)) &= T_{ss}^{(\ell)}(x_1^0) + x_1^{(2)}\frac{\mathrm{d}T_{ss}^{(\ell)}(x_1^0)}{\mathrm{d}x_1}h^2 + \mathcal{O}(h^4) \\ &= [T_{ss}^{(\ell)}]^{(0)} + [T_{ss}^{(\ell)}]^{(2)}h^2 + \mathcal{O}(h^4), \quad \ell = 0, 2 \end{aligned} \tag{2.23}$$

Same expansion holds for $U_{ss}^{\prime(\ell)}(x_1(h))$. Setting $\Delta = 0$ and retaining only the large terms up to the order h^2, yields

$$1 + \left[\frac{T_{ss}^{(2)}(x_1(h))}{T_{ss}^{(0)}(x_1(h))} + \frac{U_{ss}^{\prime(2)}(x_1(h))}{U_{ss}^{\prime(0)}(x_1(h))}\right]h^2 = 0 \tag{2.24}$$

Multiplying (2.24) with $T_{ss}^{(0)}(x_1(h)) \neq 0$ and using (2.23) as well as the corresponding expansion for $U_{ss}^{\prime(\ell)}(x_1(h))$, we get

$$[T_{ss}^{(0)}]^{(0)} + \left\{[T_{ss}^{(0)}]^{(2)} + [T_{ss}^{(2)}]^{(0)} + \frac{[T_{ss}^{(0)}]^{(0)}}{[U_{ss}^{\prime(0)}]^{(0)}}[U_{ss}^{\prime(2)}]^{(0)}\right\} h^2 = 0 \tag{2.25}$$

As can be seen, but not shown here for brevity, after laborious calculations using (2.23), (2.17), (2.18) and (2.11)–(2.14), as well as (2.B.6), (2.B.9), (2.B.12), (2.B.15), (2.B.17), (2.B.19), (2.B.21) and (2.B.23), these two expressions reduce to the well known characteristic equation for the circular optical fiber, i.e. equation (8.2–49) of [44]. Equating separately the zero order term as well as the coefficient of h^2 to zero in (2.25), we get

$$T_{ss}^{(0)}(x_1^0) = 0 \tag{2.26}$$

$$x_1^{(2)} = -\frac{T_{ss}^{(2)}(x_1^0) + U_{ss}^{\prime(2)}(x_1^0)}{\mathrm{d}T_{ss}^{(0)}(x_1^0)/\mathrm{d}x_1} \tag{2.27}$$

These two expressions are valid at cutoff. Equation (2.26) corresponds to the cutoff characteristic equations of the circular dielectric waveguide, i.e. (2.2)–(2.4), and so x_1^0 are the roots of these equations. Once x_1^0 is known for the desired mode, $x_1^{(2)}$ is calculated from (2.27) in terms of x_1^0, and therefore the desired cutoff wavenumber for the elliptical dielectric waveguide is given by $x_1 = x_1^0 + x_1^{(2)}h^2$, or equivalently by $x_{1,sn} = x_{1,sn}(0)\left[1 + g_{sn}^{(2)}h^2\right]$.

The denominator of (2.27) is obtained by simply differentiating the appropriate equation of Section 2.2 for the mode of interest. In order to get the expression for $T_{ss}^{(2)}(x_1^0)$ in (2.27), we use (2.17), (2.11)–(2.14), as well as (2.B.5)–(2.B.24). We note that the second order expansions of (2.B.16) and (2.B.20) contain the second kind Hankel functions with various indices, namely H_s and $H_{s\pm2}$, as well as their corresponding derivatives. The asymptotic expansions for small arguments of Hankel functions and their derivatives are [1]

$$H_0(z) \underset{|z|\to 0}{\sim} -\frac{2j}{\pi}\ln\frac{j\Gamma z}{2} \tag{2.28}$$

$$H_n(z) \underset{|z|\to 0}{\sim} \frac{j\Gamma(n)}{\pi}\left(\frac{2}{z}\right)^n, \quad n \geqslant 1 \tag{2.29}$$

$$H_0'(z) \underset{|z|\to 0}{\sim} -\frac{2j}{\pi}\frac{1}{z} \tag{2.30}$$

$$H_n'(z) \underset{|z|\to 0}{\sim} -\frac{j\Gamma(n+1)}{\pi}\frac{2^n}{z^{n+1}}, \quad n \geqslant 1 \tag{2.31}$$

In (2.28), $\Gamma = \exp(\gamma)$ with γ the Euler-Mascheroni constant, while in (2.29) and (2.31) $\Gamma(n)$ is the Gamma function. Since the expression for $T_{ss}^{(2)}(x_1^0)$ must be calculated at cutoff, i.e. when $x_2 = \gamma c_0 \to 0$, appropriate substitution should be done in (2.B.16) and (2.B.20) for the Hankel functions since they have different asymptotic behaviour when $s = 0$ and when $s \geqslant 1$, as can be seen from (2.28)–(2.31). It should be reminded that all the expansions in Appendix 2.B are originating from the expansions for Mathieu functions and so all the indices appearing must be nonnegative. So, for instance, when we examine the case $s = 1$, H_s and H_{s+2} in (2.B.16) and (2.B.20) must be substituted from (2.29) while H_{s-2} should disappear. Same applies for the derivatives. A similar procedure also holds when evaluating $U_{ss}'^{(2)}(x_1^0)$ in (2.27).

The wavenumbers obtained from the aforementioned procedure correspond to ${}_{\mathrm{o}}\mathrm{EH}_{sn}$ and ${}_{\mathrm{o}}\mathrm{HE}_{sn}$ ($s \geqslant 1$, $n \geqslant 1$) modes. If we repeat the same process but with the dual sets of (2.8) and (2.9), we will get expansions giving the cutoff wavenumbers for ${}_{\mathrm{e}}\mathrm{EH}_{sn}$ and ${}_{\mathrm{e}}\mathrm{HE}_{sn}$ ($s \geqslant 1$, $n \geqslant 1$) modes. Precisely at cutoff ($x_2 = 0$), however, these expansions are the same, and just above cutoff ($x_2 \ll 1$) they start to separate, given by prohibitively lengthy expressions. This means that exactly at cutoff, both even and odd modes have the same cutoff wavenumbers, which are given by the same expansions. Concluding, we have the following results for $x_1^{(2)}$ in (2.27) (the various α's and g's appearing in the following expressions as well as in the expressions of the Appendix 2.B are given in equations (A.6)–(A.12) in the Appendix of [70]):

$$\begin{aligned} x_1^{(2)}\Big|_{{}_{\mathrm{e}}^{\mathrm{o}}\mathrm{EH}_{1n},{}_{\mathrm{o}}^{\mathrm{e}}\mathrm{HE}_{1n}} &= x_1^0\{2 + [g_2^{\mathrm{e}}(1) + g_2^{\mathrm{e}}(1)g_0^{\mathrm{o}}(1) - 2g_2^{\mathrm{o}}(1)](x_1^0)^2 \\ &\quad - 16\alpha_{20}^{\mathrm{e}+}(1) + 48\alpha_{20}^{\mathrm{o}+}(1)\}/[x_1^0 J_0(x_1^0) - J_1(x_1^0)], \quad n \geqslant 2 \end{aligned} \tag{2.32}$$

$$\begin{aligned} x_1^{(2)}\Big|_{{}_{\mathrm{o}}^{\mathrm{e}}\mathrm{EH}_{2n}} &= x_1^0\Big\{3 + 4\{g_2^{\mathrm{e}}(2)[2 + g_0^{\mathrm{o}}(2)] - 5g_2^{\mathrm{o}}(2)\}(x_1^0)^2 \\ &\quad + 96[-2\alpha_{20}^{\mathrm{e}+}(2) + \alpha_{20}^{\mathrm{o}+}(2)]\Big\}/[4x_1^0 J_1(x_1^0) - 8J_2(x_1^0)], \quad n \geqslant 1 \end{aligned} \tag{2.33}$$

$$\begin{aligned} x_1^{(2)}\Big|_{{}_{\mathrm{e}}^{\mathrm{o}}\mathrm{HE}_{2n}} &= x_1^0\Big\{3 + 4\{g_2^{\mathrm{e}}(2)[2 + g_0^{\mathrm{o}}(2)] - 5g_2^{\mathrm{o}}(2)\}(x_1^0)^2 \\ &\quad + 96[-2\alpha_{20}^{\mathrm{e}+}(2) + \alpha_{20}^{\mathrm{o}+}(2)]\Big\}/\Big\{4x_1^0\epsilon_1\mu_1/(\epsilon_2\mu_2)J_0(x_1^0) \\ &\quad - 4\{\epsilon_1\mu_1/(\epsilon_2\mu_2) + [(x_1^0)^2 - 1]\}J_1(x_1^0)\Big\}, \quad n \geqslant 1 \end{aligned} \tag{2.34}$$

$$\begin{aligned} x_1^{(2)}\Big|_{{}_{\mathrm{o}}^{\mathrm{e}}\mathrm{EH}_{sn}} &= x_1^0\Big\{1 + s + s^2\{g_2^{\mathrm{e}}(s)[g_0^{\mathrm{o}}(s) + s] \\ &\quad - g_2^{\mathrm{o}}(s)(s^2+1)\}(x_1^0)^2 + 8(s+1)[-s^3\alpha_{20}^{\mathrm{e}+}(s) \\ &\quad + (s+2)\alpha_{20}^{\mathrm{o}+}(s)]\Big\}/\{s^2[x_1^0 J_{s-1}(x_1^0) - sJ_s(x_1^0)]\}, \quad s \geqslant 3, \quad n \geqslant 1 \end{aligned} \tag{2.35}$$

$$x_1^{(2)}\Big|_{\substack{\mathrm{e}\\\mathrm{o}}\mathrm{HE}_{sn}} = (s-1)x_1^0\Big\{1+s+s^2\big\{g_2^{\mathrm{e}}(s)\big[g_0^{\mathrm{o}}(s)+s\big]-g_2^{\mathrm{o}}(s)(s^2+1)\big\}(x_1^0)^2 + 8(s+1)\big[-s^3\alpha_{20}^{\mathrm{e}+}(s)+(s+2)\alpha_{20}^{\mathrm{o}+}(s)\big]\Big\} / \Big\{s^2\big\{(s-1)x_1^0\epsilon_1\mu_1/(\epsilon_2\mu_2)J_{s-2}(x_1^0)-\big[(s-1)^2\epsilon_1\mu_2/(\epsilon_2\mu_2) + (x_1^0)^2-(s-1)^2\big]J_{s-1}(x_1^0)\big\}\Big\}, \quad s\geqslant 3, \quad n\geqslant 1 \tag{2.36}$$

The dominant modes ${}_{\mathrm{o}}^{\mathrm{e}}\mathrm{HE}_{11}$ have $x_1^{(2)}=0$ and are always cutoff at zero.

In the special case where $s=0$, we no longer have two coupled sets, but only one equation given by (2.10). This case is similar to [36] and the cutoff wavenumbers are given again from (2.26)–(2.27) but setting $U_{ss}'^{(2)}(x_1^0)=0$. However $T_{00}^{(\ell)}$ is not given by (2.17) but from the equation

$$T_{00}^{(\ell)} = -\frac{1}{\sqrt{\epsilon_2\mu_2}}\frac{\sqrt{\gamma^2+\kappa^2}}{\sqrt{\frac{\epsilon_1\mu_1}{\epsilon_2\mu_2}\gamma^2+\kappa^2}}\left[\frac{\epsilon_1}{\kappa^2}\alpha_{00}'^{(\ell)}+\frac{\epsilon_2}{\gamma^2}a_{00}'^{(\ell)}\right], \quad \ell=0,2. \tag{2.37}$$

$T_{00}^{(0)}(x_1^0)$ now corresponds to (2.1) for TM and TE modes of the circular waveguide and the cutoff wavenumbers obtained from this special case correspond to ${}_{\mathrm{o}}\mathrm{EH}_{0n}$ and ${}_{\mathrm{e}}\mathrm{HE}_{0n}$ ($n\geqslant 1$) modes of the elliptical waveguide, with $x_1^{(2)}$ given by

$$x_1^{(2)}\Big|_{{}_{\mathrm{o}}\mathrm{EH}_{0n},{}_{\mathrm{e}}\mathrm{HE}_{0n}} = \big\{3g_2^{\mathrm{e}}(0)(x_1^0)^2J_0(x_1^0) - \big[4J_0(x_1^0)+(x_1^0)^2J_2(x_1^0)\big]\alpha_{20}^{\mathrm{e}+}(0)\big\}/J_1(x_1^0), \quad n\geqslant 1 \tag{2.38}$$

2.4 Numerical results and discussion

This section is organized in three subsections: first, we check the validity of the exact method by comparing it with results from other works in the literature. Then our analytical method is checked versus our exact method for various eccentricities, and finally, a variety of numerical results and a practical application relevant to microwave frequencies is presented and discussed.

2.4.1 Validation of the exact method

The exact cutoff wavenumbers are obtained by setting the determinant of (2.A.1), or its dual, equal to zero. The routines for Mathieu functions used to build up the system matrix are the same with the ones used in Chapter 1. Then the system matrix is truncated and the roots are found by using a combination of the bisection and the inverse quadratic interpolation method. The truncation order used for the determinant was 12, however even 8 terms could be used to converge to the correct root. So, the computational performance is high and the result is exact, but the algorithm must be repeated, from the beginning, each time a new root is of interest, each time the materials of the two regions of the waveguide are changed, or each time the eccentricity of the elliptical core is changed.

First, we have validated the results for the special case when the eccentricity $h=0$. In this case the results should match the corresponding ones for the circular dielectric waveguide, whose cutoff wavenumbers are given by the exact formulas (2.1)–(2.4). The 3rd column in Tables 2.3 and 2.4 – which we will introduce in Subsection 2.4.3 – depict the exact cutoff wavenumbers for the circular waveguide,

Table 2.1 Comparison of the normalized cutoff wavenumber κb_0 for the first higher-order mode ${}_{\mathrm{o}}\mathrm{EH}_{01}$, of the current work with other methods from the literature.

c_0/b_0	h	This work	[74]	[51]
1.1	0.42	2.24	2.240	2.237
1.5	0.75	1.791	1.791	1.778
2	0.87	1.468	1.468	1.451
4	0.97	0.942	0.942	0.925
10	0.995	0.5544	0.554	0.547
20	0.999	0.3814	0.381	0.379

obtained by (2.1)–(2.4), for two different relative permittivity values, namely $\epsilon_1/\epsilon_2 = 2.56$ and $\epsilon_1/\epsilon_2 = 1.0404$. We have verified these results to all 5 significant digits shown in both Tables, for all modes, by setting the eccentricity of our exact method equal to zero. A table depicting this comparison is not provided, since the agreement is perfect.

Next we compare our results with the results from the published works [74] and [51], where the authors have used the refractive indices $n_1 = 1.46$ and $n_2 = 1.34$, for the first higher-order mode ${}_{\mathrm{o}}\mathrm{EH}_{01}$. In [74, 51], the authors give results for the normalized cutoff wavenumber $V = \kappa b_0$, while our method gives results for the normalized cutoff wavenumber κc_0. So, we need to multiply our results with the ratio b_0/c_0 and then compare them. The comparison is shown in Table 2.1 for various values of the eccentricity. The agreement is evident from both works [74, 51].

Furthermore, we compare our results with the results from the published works [14, 74, 61, 24]. The methods used in these works are described in the Introduction, and are based on the weakly guiding approximation, when the difference of the refractive indices between the core and the cladding of the waveguide is small. Authors use different nomenclature for the modes, however, in our comparison we will adopt the nomenclature given in [27], namely the linearly polarized modes ${}_{\substack{\mathrm{e}\\\mathrm{o}}}\mathrm{LP}_{sn}$, $s \geqslant 0$, $n \geqslant 1$. It should be noted that our exact method is full and that the linearly polarized modes are provided here merely for illustration. In Subsection 2.4.3 numerical results are given and discussed for the complete hybrid modes. The values of the refractive indices used to match the ones in the aforementioned published works are $n_1 = 1.485$ and $n_2 = 1.47$. Again, as it was the case for Table 2.1, in all these works the authors give results for the normalized cutoff wavenumber $V = \kappa b_0$. Table 2.2 shows this comparison for 3 different cases – namely 3 different eccentricities – for the first 12 successive modes. The agreement is evident, especially with [14] and [74], and in some cases in all 4 significant figures. Even for the case of high eccentricity $h \simeq 0.87$, some of the results agree in all significant digits.

2.4.2 Validation of the analytical formulas

The approximated cutoff wavenumbers are obtained using the formula $x^0_{1,sn}(1 + x^{(2)}_{1,sn}/x^0_{1,sn}h^2)$ where $x^0_{1,sn}$ is obtained by (2.1)–(2.4) and $x^{(2)}_{1,sn}$ by (2.32)–(2.36) and (2.38), depending on the mode of interest. This is the result of daunting cumbersome calculations, and the desired cutoff wavenumber is obtained immediately, using these closed-form analytical expressions, for every higher-order mode, for different materials, and for different eccentricities. The limitation stems from the fact that these formulas require small eccentricity.

The validity of our analytical formulas is verified by comparison with the exact

Table 2.2 Comparison of the normalized cutoff wavenumbers κb_0 of the current work with other methods from the literature.

1st case: $c_0/b_0 = 1.2$ or $h \simeq 0.55$						
No	Mode	This work	[14]	[74]	[61]	[24]
1	${}_{e}LP_{01}$	0	0	0	0	0
2	${}_{e}LP_{11}$	2.105	2.102	2.101	2.18	2.14
3	${}_{o}LP_{11}$	2.305	2.303	2.302	2.35	2.32
4	${}_{e}LP_{21}$	3.304	3.286	3.285	3.40	3.36
5	${}_{o}LP_{21}$	3.514	3.509	3.510	3.54	3.52
6	${}_{e}LP_{02}$	3.756	3.739	3.737	3.84	3.83
7	${}_{e}LP_{31}$	4.555	4.552	4.550	4.62	4.59
8	${}_{o}LP_{31}$	4.645	4.643	4.640	4.66	4.65
9	${}_{e}LP_{12}$	4.987	4.989	4.985	5.00	5.00
10	${}_{o}LP_{12}$	5.357	5.356	5.356	5.43	5.40
11	${}_{e}LP_{41}$	5.72	5.721	5.714	6.20	5.72
12	${}_{o}LP_{41}$	5.76	5.759	5.754	5.80	5.75
2nd case: $c_0/b_0 = 1.5$ or $h \simeq 0.75$						
No	Mode	This work	[14]	[74]	[61]	[24]
1	${}_{e}LP_{01}$	0	0	0	0	0
2	${}_{e}LP_{11}$	1.793	1.791	1.791	1.81	1.82
3	${}_{o}LP_{11}$	2.193	2.193	2.193	2.26	2.26
4	${}_{e}LP_{21}$	2.753	2.736	2.736	2.84	2.80
5	${}_{o}LP_{21}$	3.185	3.182	3.182	3.19	3.17
6	${}_{e}LP_{02}$	3.654	3.638	3.639	3.73	3.71
7	${}_{e}LP_{31}$	3.785	3.782	3.782	3.83	3.83
8	${}_{o}LP_{31}$	4.098	4.097	4.095	4.10	4.08
9	${}_{e}LP_{12}$	4.631	4.631	4.629	4.64	4.65
10	${}_{e}LP_{41}$	4.764	4.763	4.757	5.35	4.78
11	${}_{o}LP_{41}$	5.016	5.017	5.012	5.10	5.18
12	${}_{o}LP_{12}$	5.238	5.237	5.237	5.30	5.28
3rd case: $c_0/b_0 = 2$ or $h \simeq 0.87$						
No	Mode	This work	[14]	[74]	[61]	[24]
1	${}_{e}LP_{01}$	0	0	0	0	0
2	${}_{e}LP_{11}$	1.47	1.467	1.468	1.51	1.51
3	${}_{o}LP_{11}$	2.077	2.077	2.076	2.10	2.10
4	${}_{e}LP_{21}$	2.196	2.180	2.181	2.26	2.26
5	${}_{o}LP_{21}$	2.848	2.846	2.846	2.89	2.89
6	${}_{e}LP_{31}$	2.983	2.981	2.981	3.00	3.01
7	${}_{e}LP_{02}$	3.535	3.522	3.520	3.63	3.61
8	${}_{o}LP_{31}$	3.539	3.537	3.537	3.52	3.55
9	${}_{e}LP_{41}$	3.738	3.738	3.734	4.64	3.77
10	${}_{o}LP_{41}$	4.234	4.238	4.232	4.2	4.24
11	${}_{e}LP_{12}$	4.293	4.295	4.293	4.4	4.3
12	${}_{e}LP_{51}$	4.49	4.496	4.487	—	—

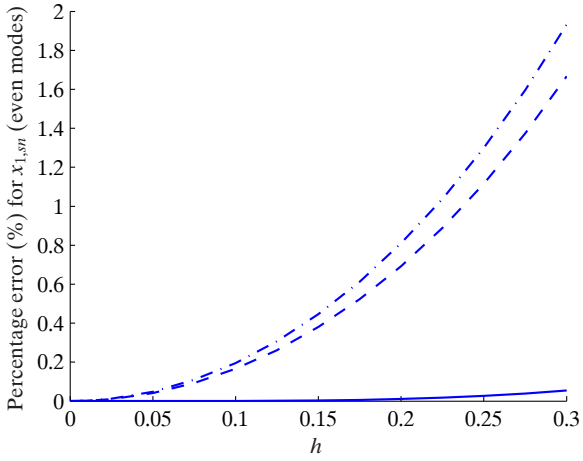

Figure 2.2 (color online). Percentage error for $x_{1,sn}$ for various higher-order even modes when $\epsilon_1/\epsilon_2 = 2.56$. Solid curve: ${}_{\mathrm{e}}\mathrm{HE}_{01}$; dashed curve: ${}_{\mathrm{e}}\mathrm{HE}_{21}$; dashed-dotted curve: ${}_{\mathrm{e}}\mathrm{HE}_{31}$.

cutoff wavenumbers obtained by setting the determinant of (2.A.1) equal to zero. The cutoff wavenumbers with even indices, for example ${}_{\mathrm{o}}\mathrm{HE}_{mn}$ with $m = 0, 2, \ldots, n \geqslant 1$, are extracted by building up the even part of (2.A.1). The cutoff wavenumbers with odd indices, those with $m = 1, 3, \ldots$, are extracted by taking the odd part of (2.A.1). This property was also reported in [38]. The same holds also for the even modes, for example ${}_{\mathrm{e}}\mathrm{HE}_{mn}$, but in this case the dual of (2.A.1) should be used. In Figure 2.2 we give the percentage relative error |(exact value − approximate value)/exact value| × 100 for the cutoff wavenumbers of various higher-order even modes. The details and the values of the parameters used are given in the captions. The approximate values used correspond to our analytical formulas. As it has been already pointed out in

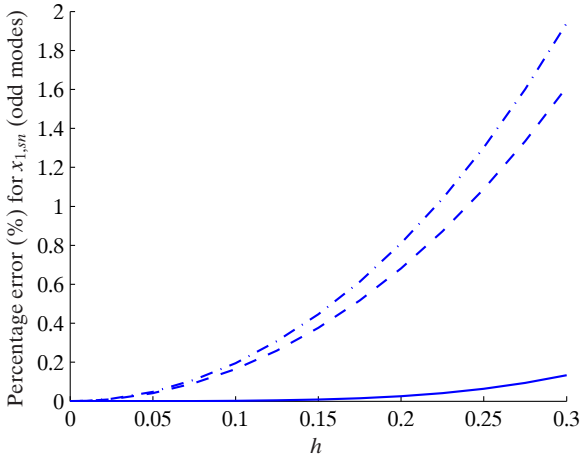

Figure 2.3 (color online). Percentage error for $x_{1,sn}$ for various higher-order odd modes when $\epsilon_1/\epsilon_2 = 2.56$. Solid curve: ${}_{\mathrm{o}}\mathrm{EH}_{01}$; dashed curve: ${}_{\mathrm{o}}\mathrm{HE}_{21}$; dashed-dotted curve: ${}_{\mathrm{o}}\mathrm{HE}_{31}$.

Table 2.3 Normalized cutoff wavenumbers $x_{1,sn}$ for an elliptical dielectric waveguide.*

No	Mode	$x_{1,sn}(0) = x^0_{1,sn}$	$g^{(2)}_{sn} = x^{(2)}_{1,sn}/x^0_{1,sn}$
1	${}^{e}_{o}\mathrm{HE}_{11}$	0	0
2	${}_{e}\mathrm{HE}_{01}$, ${}_{o}\mathrm{EH}_{01}$	2.4048	0.25
3	${}^{e}_{o}\mathrm{HE}_{21}$	2.8652	0.086284
4	${}^{e}_{o}\mathrm{HE}_{12}$, ${}^{e}_{o}\mathrm{EH}_{11}$	3.8317	0.54122
5	${}^{e}_{o}\mathrm{HE}_{31}$	4.3551	0.056829
6	${}^{e}_{o}\mathrm{EH}_{21}$	5.1356	2.4333
7	${}_{e}\mathrm{HE}_{02}$, ${}_{o}\mathrm{EH}_{02}$	5.5201	0.25
8	${}^{e}_{o}\mathrm{HE}_{41}$	5.6855	0.029099
9	${}^{e}_{o}\mathrm{HE}_{22}$	5.7780	−0.46638
10	${}^{e}_{o}\mathrm{EH}_{31}$	6.3802	1.3524
11	${}^{e}_{o}\mathrm{HE}_{51}$	6.9445	0.002074
12	${}^{e}_{o}\mathrm{HE}_{13}$, ${}^{e}_{o}\mathrm{EH}_{12}$	7.0156	−2.4471
13	${}^{e}_{o}\mathrm{HE}_{32}$	7.3868	−0.41643
14	${}^{e}_{o}\mathrm{EH}_{41}$	7.5883	0.95942
15	${}^{e}_{o}\mathrm{HE}_{61}$	8.1616	−0.024383

*The value of the parameter used is $\epsilon_1/\epsilon_2 = 2.56$.

Section 2.3, precisely at cutoff, our analytical formulas give the cutoff wavenumbers for both even and odd modes. So, the results depicted in Figure 2.2 are the analytical results compared with the ones obtained from the dual of (2.A.1). In Figure 2.3 we present the same comparison but for the odd modes, i.e. the analytical results compared with the ones obtained from (2.A.1) itself.

As it can be concluded from Figures 2.2 and 2.3, the errors remain low enough even for some relatively high values of h, although the expansions used are valid up to the order of h^2 only. For instance, for the ${}_{e}\mathrm{HE}_{01}$ mode we get an error less than 0.1%, and for the ${}_{o}\mathrm{EH}_{01}$ mode an error less than 0.14%, even up to $h = 0.3$. For the higher-order modes depicted in Figures 2.2 and 2.3, the errors are maintained below the level of 2%. From these figures, as well as from many other of our available numerical results not published here, the sprectrum of the modes maintains errors below 0.4% for eccentricities up to 0.1. Lower errors for higher eccentricities require more terms, $\mathcal{O}(h^4)$ and higher, in the expansion. Nevertheless, this will lead to prohibitively lengthy expressions for the expansion coefficients, as it can be concluded from the present analysis where, although only $\mathcal{O}(h^2)$ terms are considered, the expressions are still lengthy.

2.4.3 Discussion

In Table 2.3 we list the first fifteen successive cutoff wavenumbers $x_{1,sn}$ for the hybrid modes appearing in the elliptical dielectric waveguide. In this case, the core of the fiber is polystyrene having a relative refractive index $n_1/n_2 = 1.6$, or equivalently, a relative permittivity of $\epsilon_1/\epsilon_2 = 2.56$. The third column of the Table, namely $x_{1,sn}(0)$ corresponds to the circular dielectric waveguide, while the fourth column gives the values of the expansion coefficients $g^{(2)}_{sn}$ for the corresponding elliptical dielectric waveguide, having small eccentricity.

When the fractional refractive index defined by $\Delta = 1 - n_2/n_1$ is small, practically $\Delta < 0.05$, then we talk about the weakly guiding approximation (WGA) [13].

Table 2.4 Normalized cutoff wavenumbers $x_{1,sn}$ for an elliptical dielectric waveguide.*

No	Mode	$x_{1,sn}(0) = x^0_{1,sn}$	$g^{(2)}_{sn} = x^{(2)}_{1,sn}/x^0_{1,sn}$
1	${}^{e}_{o}HE_{11}$	0	0
2	${}_{e}HE_{01}$, ${}_{o}EH_{01}$	2.4048	0.25
3	${}^{e}_{o}HE_{21}$	2.4215	−0.0092828
4	${}^{e}_{o}HE_{12}$, ${}^{e}_{o}EH_{11}$	3.8317	0.54122
5	${}^{e}_{o}HE_{31}$	3.8525	−0.039927
6	${}^{e}_{o}EH_{21}$	5.1356	2.4333
7	${}^{e}_{o}HE_{41}$	5.1588	−0.07852
8	${}_{e}HE_{02}$, ${}_{o}EH_{02}$	5.5201	0.25
9	${}^{e}_{o}HE_{22}$	5.5274	−0.49137
10	${}^{e}_{o}EH_{31}$	6.3802	1.3524
11	${}^{e}_{o}HE_{51}$	6.405	−0.12158
12	${}^{e}_{o}HE_{13}$, ${}^{e}_{o}EH_{12}$	7.0156	−2.4471
13	${}^{e}_{o}HE_{32}$	7.0271	−0.46769
14	${}^{e}_{o}EH_{41}$	7.5883	0.95942
15	${}^{e}_{o}HE_{61}$	7.6145	−0.1672

*The value of the parameter used is $\epsilon_1/\epsilon_2 = 1.0404$.

This means that more than a single mode can have the same propagation constant β. In Table 2.4 we give the cutoff wavenumbers for this case where the above criterion for WGA is satisfied, by selecting $n_1/n_2 = 1.02$ (or $\epsilon_1/\epsilon_2 = 1.0404$), which is equivalent to $\Delta = 0.02$. If one uses the approximate equations for finding the cutoff wavenumbers, he will conclude that a group of modes have the same cutoff wavenumber. For instance, all three modes ${}_{e}HE_{01}$, ${}_{o}EH_{01}$ and ${}^{e}_{o}HE_{21}$ (No 2 and 3 in Table 2.4) have the same cutoff wavenumber, leading to the ${}^{e}_{o}LP_{11}$ mode. Rigorously speaking, however, the ${}^{e}_{o}HE_{21}$ mode is still separated by a small fraction. This is indeed the case from our results in Table 2.4, which can serve for a comparison with future methods studying the LP modes. It should be noticed that the modes ${}^{e}_{o}EH_{sn}$ ($s \geq 0$, $n \geq 1$) and ${}^{e}_{o}HE_{sn}$ ($s = 0, 1$, $n \geq 1$) have the same expansion coefficients, as can be concluded from both Tables, while the ${}^{e}_{o}HE_{sn}$ ($s \geq 2$, $n \geq 1$) modes do not. This comes from the fact that the ${}^{e}_{o}EH_{sn}$ ($s \geq 0$, $n \geq 1$) and ${}^{e}_{o}HE_{sn}$ ($s = 0, 1$, $n \geq 1$) modes are independent of the relative permittivity, as can be seen from (2.32)–(2.36) and (2.38), while the ${}^{e}_{o}HE_{sn}$ ($s \geq 2$, $n \geq 1$) modes do depend on it. This also leads to the fact that the cutoff wavenumbers are safely separated each other, at least by 0.2, in the case where $\epsilon_1/\epsilon_2 = 2.56$, since, the larger the contrast between the core and the cladding permittivity, the larger the separation between the modes. This is why we notice a change in the order between lines No 7 and 8 in Tables 2.3 and 2.4: the current value of permittivity used in Table 2.3 makes the ${}^{e}_{o}HE_{41}$ mode to overpass the ${}_{e}HE_{02}$ and ${}_{o}EH_{02}$ modes.

The expansion coefficients $g^{(2)}_{0n}$ for the ${}_{e}HE_{0n}$ and ${}_{o}EH_{0n}$ ($n \geq 1$) modes are all the same and equal to 0.25. This can be seen from both Tables 2.3 and 2.4 only for ${}_{e}HE_{01}$ and ${}_{o}EH_{01}$ modes, but it was verified to other higher-order modes as well. This behaviour appeared also in the TE_{e1n} and TE_{o1n} ($n \geq 1$) modes of the elliptical metallic waveguide [71].

A practical application with relevance at microwave frequencies is presented

Table 2.5 Normalized cutoff wavenumbers and cutoff frequencies for an elliptical dielectric waveguide.*

No	Approx. value $x_{1,sn}$	Exact value, even modes $x_{1,sn}$	Exact value, odd modes $x_{1,sn}$	Percentage error, even modes	Percentage error, odd modes
1	0	0	0	0	0
2	2.411	2.411	2.411	0.002042	0.002766
3	2.868	2.872	2.872	0.1664	0.1658
4	3.852	3.843	3.839	0.2412	0.3506
5	4.358	4.366	4.366	0.1966	0.1966
6	5.261	5.149	5.148	2.177	2.179
7	5.534	5.534	5.534	0.002166	0.002605
8	5.687	5.699	5.699	0.216	0.2159
9	5.751	5.793	5.793	0.7276	0.7248
10	6.466	6.396	6.396	1.1	1.1

No	Mode	Cutoff frequency in GHz for the circular waveguide	Cutoff frequency in GHz for the elliptical waveguide
1	${}_{\text{e}}^{\text{o}}\text{HE}_{11}$	0	0
2	${}_{\text{e}}\text{HE}_{01}, {}_{\text{o}}\text{EH}_{01}$	9.187	9.21
3	${}_{\text{e}}^{\text{o}}\text{HE}_{21}$	10.95	10.95
4	${}_{\text{e}}^{\text{o}}\text{HE}_{12}, {}_{\text{e}}^{\text{o}}\text{EH}_{11}$	14.64	14.72
5	${}_{\text{e}}^{\text{o}}\text{HE}_{31}$	16.64	16.65
6	${}_{\text{e}}^{\text{o}}\text{EH}_{21}$	19.62	20.1
7	${}_{\text{e}}\text{HE}_{02}, {}_{\text{o}}\text{EH}_{02}$	21.09	21.14
8	${}_{\text{e}}^{\text{o}}\text{HE}_{41}$	21.72	21.73
9	${}_{\text{e}}^{\text{o}}\text{HE}_{22}$	22.07	21.97
10	${}_{\text{e}}^{\text{o}}\text{EH}_{31}$	24.37	24.7

*The values of the parameters used are $\epsilon_1/\epsilon_2 = 2.56$, $h = 0.1$, and $c_0 = 1$cm.

in Table 2.5 where $\epsilon_1/\epsilon_2 = 2.56$, $h = 0.1$, and $c_0 = 1$cm, for the first 10 successive hybrid modes of Table 2.3. In the upper part of Table 2.5 we depict the normalized cutoff wavenumbers, obtained by our analytical formulas and by our exact solution. Then we compare them by taking the percentage relative errors. As it can be seen from the last two columns, the percentage error is maintained at low levels, mainly below 0.73%, but in the case of the ${}_{\text{e}}^{\text{o}}\text{EH}_{21}$ and ${}_{\text{e}}^{\text{o}}\text{EH}_{31}$ modes, this error slightly exceeds the level of 2% and 1%, respectively, which is still small. In the lower part of Table 2.5 we depict the cutoff frequencies in GHz, obtained by the relation $f = x_{1,sn}/[2\pi c_0 \mu_0^{1/2}(\epsilon_1 - \epsilon_2)^{1/2}]$, using our analytical formulas, for both the circular case and the elliptical case. The range for the single mode operation is $0 < f < 9.21$ GHz for the elliptical waveguide. Comparing the cutoff frequencies between the circular and the elliptical case, we note a slightly increased value for the elliptical case. This means that the range of the single mode operation for an elliptical fiber is increased, a fact that was also observed by Yeh in [77]. This is also shown in Table 2.3 where the first few $g_{sn}^{(2)}$'s are positive, and thus lead to increased $x_{1,sn}$'s, compared with $x_{1,sn}(0)$'s. Some $g_{sn}^{(2)}$'s in both Tables 2.3 and 2.4 are negative, however the error compared with the exact value was less than 1%; thats the reason why in line No 9 of Table 2.5 the frequency is decreased.

2.5 Concluding remarks

In this work the cutoff wavenumbers of an elliptical dielectric waveguide were obtained exactly and analytically. The exact solution requires the algorithm to be repeated, from the beginning, each time a new mode, different materials of the waveguide, or a different eccentricity is needed. On the contrary, our analytical, closed-form formulas are valid for every small eccentricity, for every higher-order hybrid mode, and for different materials. The validation of our exact solution was performed by comparing with the case of the circular dielectric waveguide and with previously published results. The validation of our analytical formulas was done by comparison with our exact solution. The errors are maintained low enough, lower than 0.4% for eccentricities up to 0.1 for a big spectrum of higher-order modes, and in some cases this small error was even extended at higher values of eccentricities up to 0.3.

2.A Appendix: The exact characteristic equation at cutoff

The set appearing in (2.7) and its pair, define the homogeneous system with unknowns the expansion coefficients A_m^{e} and D_m^{o}. We multiply both sides of (2.7) by $\kappa^2\gamma^2$ and use $\omega = k_2/(\epsilon_2\mu_2)^{1/2}$. Then there appears the ratio k_2/β which equals to $k_2/\beta = (\gamma^2+\kappa^2)^{1/2}/[\epsilon_1\mu_1/(\epsilon_2\mu_2)\gamma^2+\kappa^2]^{1/2}$. At cutoff, $\gamma \to 0$. Expanding k_2/β in Maclaurin series, we get $k_2/\beta = 1 + [1-\epsilon_1\mu_1/(\epsilon_2\mu_2)]/(2\kappa^2)\gamma^2 + \mathcal{O}(\gamma^4)$. Now, the normalized wavenumbers $x_1 = \kappa c_0$ are obtained through $\kappa a = \kappa c_0 h = x_1 h$. Similarly holds $\gamma a = \gamma c_0 h = x_2 h$. So, by keeping terms in (2.7) and in its pair up to the order γ^2, and using the aforementioned substitutions for κa and γa, there results the following system matrix M at cutoff, namely

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \tag{2.A.1}$$

where

$$M_{11} = -x_2^2 A_1(x_1) - x_1^2 A_2(x_1, x_2) \tag{2.A.2}$$

$$M_{12} = x_2^2 \frac{\mu_1}{\mu_2}\sqrt{\frac{\mu_2}{\epsilon_2}}\mathrm{diag}(d(x_1)) + \left(x_1^2 + \frac{1-\frac{\epsilon_1\mu_1}{\epsilon_2\mu_2}}{2}x_2^2\right)\sqrt{\frac{\mu_2}{\epsilon_2}}D_3(x_1, x_2) \tag{2.A.3}$$

$$M_{21} = -x_2^2 \frac{\epsilon_1}{\epsilon_2}\sqrt{\frac{\epsilon_2}{\mu_2}}\mathrm{diag}(a(x_1)) - \left(x_1^2 + \frac{1-\frac{\epsilon_1\mu_1}{\epsilon_2\mu_2}}{2}x_2^2\right)\sqrt{\frac{\epsilon_2}{\mu_2}}A_3(x_1, x_2) \tag{2.A.4}$$

$$M_{22} = -x_2^2 D_1(x_1) - x_1^2 D_2(x_1, x_2) \tag{2.A.5}$$

If N is the truncation order, then M is of the order $2N \times 2N$, while the submatrices $A_{1,2,3}, D_{1,2,2}, \mathrm{diag}(a), \mathrm{diag}(d) \in \mathbb{R}^{N\times N}$. At cutoff, $x_2 \to 0$ and the field in region II detaches and does not decay. The $\mathcal{O}(x_2^2)$ terms appearing in (2.A.1) are necessary for the validity of all types of modes for the case when $h \to 0$, i.e. the circular dielectric waveguide. This means that (2.A.1) must be set up just above cutoff with $x_2 \ll 1$. On the contrary, if we dismiss these terms by setting (2.A.1) precisely at cutoff with $x_2 = 0$, then the roots that correspond to the HE modes when $h \to 0$, i.e. (2.4), do not appear. This fact was also observed by Saad in [56] when he was missing some modes due to unfavorable numerical conditions. So to conclude, by keeping $x_2 \ll 1$

and setting $h \to 0$, the roots of the equation $\det(M) = 0$ are identical to all the roots given by (2.1)–(2.4).

2.B Appendix: Expansions for the characteristic equation

The analytical expressions of the various quantities appearing in (2.8) are found after lengthy, laborious calculations, by using the expansions for the Mathieu functions from [35], the expansions for the various inner products M^{e} and M^{o} from the Appendix of [72], and finally, the expansions for the various inner products $M^{\mathrm{e}'\mathrm{o}}$ and $M^{\mathrm{o}'\mathrm{e}}$. The last are defined in the Appendix 1.A and their expansions are

$$[M_{ss}^{\mathrm{e}'\mathrm{o}}(-j\gamma a, \kappa a)]^{(0)} = -\pi s g_0^{\mathrm{o}}(s), \quad s \geqslant 1 \tag{2.B.1}$$

$$[M_{ss}^{\mathrm{e}'\mathrm{o}}(-j\gamma a, \kappa a)]^{(2)} = -\pi s \left[x_1^2 g_2^{\mathrm{o}}(s) - x_2^2 g_0^{\mathrm{o}}(s) g_2^{\mathrm{e}}(s)\right], \quad s \geqslant 1 \tag{2.B.2}$$

$$[M_{ss}^{\mathrm{o}'\mathrm{e}}(-j\gamma a, \kappa a)]^{(0)} = \pi s g_0^{\mathrm{o}}(s), \quad s \geqslant 1 \tag{2.B.3}$$

$$[M_{ss}^{\mathrm{o}'\mathrm{e}}(-j\gamma a, \kappa a)]^{(2)} = \pi s \left[x_1^2 g_0^{\mathrm{o}}(s) g_2^{\mathrm{e}}(s) - x_2^2 g_2^{\mathrm{o}}(s)\right], \quad s \geqslant 1 \tag{2.B.4}$$

The expansions with higher order indices like $[M_{s\pm2,s}^{\mathrm{e}'\mathrm{o}}]^{(2)}$ and $[M_{s\pm2,s}^{\mathrm{o}'\mathrm{e}}]^{(2)}$ lead to $\mathcal{O}(h^4)$ terms, so they are neglected. Therefore, the various quantities appearing in (2.8) are given in what follows:

$$\alpha_{ss} = \alpha_{ss}^{(0)} + \alpha_{ss}^{(2)} h^2 \tag{2.B.5}$$

$$\alpha_{ss}^{(0)} = -\frac{g_0^{\mathrm{o}}(s) s J_s(x_1) \pi^{3/2}}{\sqrt{2}} \tag{2.B.6}$$

$$\begin{aligned}\alpha_{ss}^{(2)} = \frac{\pi^{3/2} s x_1^2}{\sqrt{2}} \Big\{ &- \left[2 g_2^{\mathrm{e}}(s) g_0^{\mathrm{o}}(s) + g_2^{\mathrm{o}}(s)\right] J_s(x_1) \\ &+ g_0^{\mathrm{o}}(s) J_{s-2}(x_1) \alpha_{20}^{\mathrm{e}-}(s) + g_0^{\mathrm{o}}(s) J_{s+2}(x_1) \alpha_{20}^{\mathrm{e}+}(s) \Big\}\end{aligned} \tag{2.B.7}$$

$$a_{ss} = a_{ss}^{(0)} + a_{ss}^{(2)} h^2 \tag{2.B.8}$$

$$a_{ss}^{(0)} = \alpha_{ss}^{(0)} \tag{2.B.9}$$

$$\begin{aligned}a_{ss}^{(2)} = -\frac{\pi^{3/2} s}{\sqrt{2}} \Big\{ &\left[2 g_2^{\mathrm{e}}(s) g_0^{\mathrm{o}}(s) + g_2^{\mathrm{o}}(s)\right] J_s(x_1) x_1^2 - 4 g_2^{\mathrm{e}}(s) g_0^{\mathrm{o}}(s) J_s(x_1) x_2^2 \\ &- g_0^{\mathrm{o}}(s) x_1^2 \left[J_{s-2}(x_1) \alpha_{20}^{\mathrm{e}-}(s) + J_{s+2}(x_1) \alpha_{20}^{\mathrm{e}+}(s)\right] \Big\}\end{aligned} \tag{2.B.10}$$

$$\delta_{ss} = \delta_{ss}^{(0)} + \delta_{ss}^{(2)} h^2 \tag{2.B.11}$$

$$\delta_{ss}^{(0)} = \frac{[g_0^{\mathrm{o}}(s)]^2 J_s'(x_1) \pi^{3/2} x_1}{\sqrt{2}} \tag{2.B.12}$$

$$\begin{aligned}\delta_{ss}^{(2)} = \frac{g_0^{\mathrm{o}}(s) \pi^{3/2}}{\sqrt{2} s} \Big\{ &s \left\{3 g_2^{\mathrm{o}}(s) J_s'(x_1) x_1^3 + g_0^{\mathrm{o}}(s) \left[J_s(x_1) - J_s'(x_1) x_1\right]\right\} \\ &- g_0^{\mathrm{o}}(s) x_1^3 \left[J_{s-2}'(x_1)(s-2) \alpha_{20}^{\mathrm{o}-}(s) + J_{s+2}'(x_1)(s+2) \alpha_{20}^{\mathrm{o}+}(s)\right] \Big\}\end{aligned} \tag{2.B.13}$$

$$b_{ss} = b_{ss}^{(0)} + b_{ss}^{(2)} h^2 \tag{2.B.14}$$

$$b_{ss}^{(0)} = -\frac{j[g_0^{\mathrm{o}}(s)]^2 H_s'(-jx_2) J_s(x_1)\pi^{3/2}x_2}{\sqrt{2}H_s(-jx_2)} \tag{2.B.15}$$

$$\begin{aligned} b_{ss}^{(2)} = {} & \frac{g_0^{\mathrm{o}}(s)\pi^{3/2}}{\sqrt{2}s[H_s(-jx_2)]^2}\Big\{H_s(-jx_2)J_s(x_1)s\big\{g_0^{\mathrm{o}}(s)H_s(-jx_2) \\ & + jH_s'(-jx_2)\left[g_0^{\mathrm{o}}(s) - 3g_2^{\mathrm{o}}(s)x_1^2\right]x_2 + j4g_2^{\mathrm{o}}(s)H_s'(-jx_2)x_2^3\big\} \\ & + jg_0^{\mathrm{o}}(s)(s-2)x_2\alpha_{20}^{\mathrm{o}-}(s)\big\{H_s(-jx_2)H_s'(-jx_2)J_{s-2}(x_1)x_1^2 \\ & + \left[H_{s-2}(-jx_2)H_s'(-jx_2) - H_s(-jx_2)H_{s-2}'(-jx_2)\right]J_s(x_1)x_2^2\big\} \\ & + jg_0^{\mathrm{o}}(s)(s+2)x_2\alpha_{20}^{\mathrm{o}+}(s)\big\{H_s(-jx_2)H_s'(-jx_2)J_{s+2}(x_1)x_1^2 \\ & - H_s(-jx_2)H_{s+2}'(-jx_2)J_s(x_1)x_2^2 + H_s'(-jx_2)H_{s+2}(-jx_2)J_s(x_1)x_2^2\big\}\Big\} \end{aligned} \tag{2.B.16}$$

In (2.B.5)–(2.B.16) J denotes the Bessel function, J' denotes the derivative of J with respect to its argument, H denotes the Hankel function of the second kind with the superscript (2) omitted for simplicity, and finally, H' denotes the derivative of H with respect to its argument. It is reminded that the arguments are $x_1 = \kappa c_0$ and $x_2 = \gamma c_0$. Finally, the index s is interprented as $s \geqslant 1$.

The primed quantities appearing in (2.9) have the same expansion forms as the unprimed ones, and are given in what follows:

$$\alpha_{ss}'^{(0)} = \frac{\sqrt{2}J_s'(x_1)\pi^{3/2}x_1\tanh\eta_0}{\varepsilon_s} \tag{2.B.17}$$

$$\begin{aligned} \alpha_{ss}'^{(2)} = {} & -\sqrt{2}\pi^{3/2}x_1^3\tanh\eta_0\big[-3g_2^{\mathrm{e}}(s)J_s'(x_1) + J_{s-2}'(x_1)\alpha_{20}^{\mathrm{e}-}(s) \\ & + J_{s+2}'(x_1)\alpha_{20}^{\mathrm{e}+}(s)\big]/\varepsilon_s \end{aligned} \tag{2.B.18}$$

$$a_{ss}'^{(0)} = -\frac{j\sqrt{2}H_s'(-jx_2)J_s(x_1)\pi^{3/2}x_2\tanh\eta_0}{H_s(-jx_2)\varepsilon_s} \tag{2.B.19}$$

$$\begin{aligned} a_{ss}'^{(2)} = {} & j\sqrt{2}\pi^{3/2}x_2\tanh\eta_0\Big\{g_2^{\mathrm{e}}(s)H_s(-jx_2)H_s'(-jx_2)J_s(x_1)(-3x_1^2+4x_2^2) \\ & + \big\{H_s(-jx_2)H_s'(-jx_2)J_{s-2}(x_1)x_1^2 + \big[H_{s-2}(-jx_2)H_s'(-jx_2) \\ & - H_s(-jx_2)H_{s-2}'(-jx_2)\big]J_s(x_1)x_2^2\big\}\alpha_{20}^{\mathrm{e}-}(s) \\ & + \big[H_s(-jx_2)H_s'(-jx_2)J_{s+2}(x_1)x_1^2 - H_s(-jx_2)H_{s+2}'(-jx_2)J_s(x_1)x_2^2 \\ & + H_s'(-jx_2)H_{s+2}(-jx_2)J_s(x_1)x_2^2\big]\alpha_{20}^{\mathrm{e}+}(s)\Big\}/\big\{[H_s(-jx_2)]^2\varepsilon_s\big\} \end{aligned} \tag{2.B.20}$$

$$\delta_{ss}'^{(0)} = \frac{\sqrt{2}[g_0^{\mathrm{o}}(s)]^2 J_s(x_1)\pi^{3/2}s^2\tanh\eta_0}{\varepsilon_s} \tag{2.B.21}$$

$$\begin{aligned} \delta_{ss}'^{(2)} = {} & g_0^{\mathrm{o}}(s)\pi^{3/2}sx_1^2\tanh\eta_0\big\{sJ_s(x_1)\big[g_2^{\mathrm{e}}(s)g_0^{\mathrm{o}}(s)\varepsilon_s + g_2^{\mathrm{o}}(s)(2+\varepsilon_s)\big] \\ & - 2g_0^{\mathrm{o}}(s)J_{s-2}(x_1)(s-2)\alpha_{20}^{\mathrm{o}-}(s) \\ & - 2g_0^{\mathrm{o}}(s)J_{s+2}(x_1)(s+2)\alpha_{20}^{\mathrm{o}+}(s)\big\}/\big(\sqrt{2}\varepsilon_s\big) \end{aligned} \tag{2.B.22}$$

$$b_{ss}'^{(0)} = \frac{\sqrt{2}[g_0^{\mathrm{o}}(s)]^2 J_s(x_1)\pi^{3/2}s^2\tanh\eta_0}{\varepsilon_s} \tag{2.B.23}$$

$$b'^{(2)}_{ss} = \sqrt{2}g^{\mathrm{o}}_0(s)\pi^{3/2}s\tanh\eta_0\Big\{sJ_s(x_1)\{\big[g^{\mathrm{e}}_2(s)g^{\mathrm{o}}_0(s)+2g^{\mathrm{o}}_2(s)\big]x_1^2 - 4g^{\mathrm{o}}_2(s)x_2^2\} - g^{\mathrm{o}}_0(s)x_1^2\big[J_{s-2}(x_1)(s-2)\alpha^{\mathrm{o}-}_{20}(s) + J_{s+2}(x_1)(s+2)\alpha^{\mathrm{o}+}_{20}(s)\big]\Big\}/\varepsilon_s \tag{2.B.24}$$

In (2.B.17)–(2.B.24) $\varepsilon_0 = 1$, $\varepsilon_n = 2$, $(n \geqslant 1)$ is the Neumann factor, and the index s is now interpreted as $s \geqslant 0$ for α'_{ss} and a'_{ss}, and as $s \geqslant 1$ for δ'_{ss} and b'_{ss}.

Chapter 3

Spectrum & preconditioning of the singular domain integral equation

The domain integral equation method with its FFT-based matrix-vector products is a viable alternative to local methods in free-space scattering problems. However, it often suffers from the extremely slow convergence of iterative methods, especially in the transverse electric (TE) case with large or negative permittivity. We identify very dense line segments in the spectrum as partly responsible for this behavior, and the main reason why a normally efficient deflating preconditioner does not work. We solve this problem by applying an explicit multiplicative regularizing operator, which on the operator level transforms the system to the form "identity plus compact". On the matrix level this regularization reduces the length of the dense spectral segmets roughly by a factor of four while preserving the ability to calculate the matrix-vector products using the FFT algorithm. Such a regularized system is then further preconditioned by deflating an apparently stable set of eigenvalues with largest magnitudes, which results in a robust acceleration of the restarted GMRES under constraint memory conditions.

3.1 Singular domain integral equation

We consider the two-dimensional frequency-domain electromagnetic scattering problem on an inhomogeneous object of finite spatial support exhibiting contrast in both electric and magnetic properties. Figure 3.1 depicts this configuration where the inhomogeneous object occupies the domain D. We assume that the object (a cylinder) as well the source of the incident field (a plane wave, a current carrying line, etc.) are invariant in the x_3-direction of the spatial Cartesian system. The Maxwell equations describing the total electromagnetic field in the (x_1, x_2)-plane in

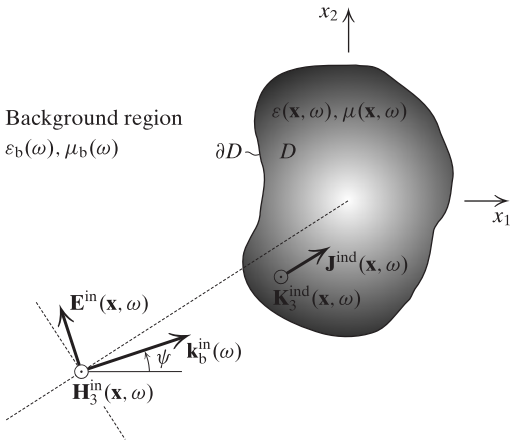

Figure 3.1 The configuration of the problem.

the presence of an inhomogeneous magneto-dielectric scatterer can be written as

$$-\partial_2 H_3(\mathbf{x},\omega) - i\omega\varepsilon(\mathbf{x},\omega)E_1(\mathbf{x},\omega) = -J_1(\mathbf{x},\omega), \tag{3.1.a}$$

$$\partial_1 H_3(\mathbf{x},\omega) - i\omega\varepsilon(\mathbf{x},\omega)E_2(\mathbf{x},\omega) = -J_2(\mathbf{x},\omega), \tag{3.1.b}$$

$$-\partial_2 E_1(\mathbf{x},\omega) + \partial_1 E_2(\mathbf{x},\omega) - i\omega\mu(\mathbf{x},\omega)H_3(\mathbf{x},\omega) = 0, \tag{3.1.c}$$

and

$$\partial_2 H_1(\mathbf{x},\omega) - \partial_1 H_2(\mathbf{x},\omega) - i\omega\varepsilon(\mathbf{x},\omega)E_3(\mathbf{x},\omega) = -J_3(\mathbf{x},\omega), \tag{3.2.a}$$

$$\partial_2 E_3(\mathbf{x},\omega) - i\omega\mu(\mathbf{x},\omega)H_1(\mathbf{x},\omega) = 0, \tag{3.2.b}$$

$$-\partial_1 E_3(\mathbf{x},\omega) - i\omega\mu(\mathbf{x},\omega)H_2(\mathbf{x},\omega) = 0, \tag{3.2.c}$$

where we introduce the two-dimensional position vector $\mathbf{x} = (x_1, x_2)$; $\mathbf{J}$ is the electric current density; ε and μ are the possibly complex-valued dielectric permittivity and magnetic permeability; and ω is the angular frequency. It is easy to notice that the first three equations, (3.1), are decoupled from the last three, (3.2). Reflecting the fact that the electric field strength is transverse to the axis of invariance, equations (3.1) are said to describe the transverse electric or TE case (also somewhat confusingly called H-wave polarization). Analogously, the three equations (3.2) correspond to the transverse magnetic, or TM case (also known as E-wave polarization).

The integral formulation of the scattering problem is obtained by first introducing the so-called incident field ($\mathbf{E}^{\text{in}}$, $\mathbf{H}^{\text{in}}$), which has the same physical source as the total field, but satisfies some simpler version of the above equations. Usually, the incident field is considered in a homogeneous isotropic background medium with parameters $\varepsilon_b(\omega)$ and $\mu_b(\omega)$. Then, it is fairly easy to see that the scattered field, defined as the difference between the total and the incident fields

$$\mathbf{E}^{\text{sc}}(\mathbf{x},\omega) = \mathbf{E}(\mathbf{x},\omega) - \mathbf{E}^{\text{in}}(\mathbf{x},\omega), \quad \mathbf{H}^{\text{sc}}(\mathbf{x},\omega) = \mathbf{H}(\mathbf{x},\omega) - \mathbf{H}^{\text{in}}(\mathbf{x},\omega), \tag{3.3}$$

also satisfies Maxwell's equations in the same homogeneous medium, i.e.,

$$-\partial_2 H_3^{\text{sc}}(\mathbf{x}) - i\omega\varepsilon_b E_1^{\text{sc}}(\mathbf{x}) = -J_1^{\text{ind}}(\mathbf{x}), \tag{3.4.a}$$

$$\partial_1 H_3^{\text{sc}}(\mathbf{x}) - i\omega\varepsilon_b E_2^{\text{sc}}(\mathbf{x}) = -J_2^{\text{ind}}(\mathbf{x}), \tag{3.4.b}$$

$$-\partial_2 E_1^{\text{sc}}(\mathbf{x}) + \partial_1 E_2^{\text{sc}}(\mathbf{x}) - i\omega\mu_b H_3^{\text{sc}}(\mathbf{x}) = -K_3^{\text{ind}}(\mathbf{x}), \tag{3.4.c}$$

and

$$\partial_2 H_1^{\rm sc}(\mathbf{x}) - \partial_1 H_2^{\rm sc}(\mathbf{x}) - i\omega\varepsilon_{\rm b} E_3^{\rm sc}(\mathbf{x}) = -J_3^{\rm ind}(\mathbf{x}), \tag{3.5.a}$$

$$\partial_2 E_3^{\rm sc}(\mathbf{x}) - i\omega\mu_{\rm b} H_1^{\rm sc}(\mathbf{x}) = -K_1^{\rm ind}(\mathbf{x}), \tag{3.5.b}$$

$$-\partial_1 E_3^{\rm sc}(\mathbf{x}) - i\omega\mu_{\rm b} H_2^{\rm sc}(\mathbf{x}) = -K_2^{\rm ind}(\mathbf{x}), \tag{3.5.c}$$

where the implicit dependence on ω has been omitted for simplicity. The physical sources of the scattered field are the induced current densities, also known as contrast currents:

$$J_k^{\rm ind}(\mathbf{x}) = -i\omega\left[\varepsilon(\mathbf{x}) - \varepsilon_{\rm b}\right] E_k(\mathbf{x}), \quad k = 1,2,3; \tag{3.6.a}$$

$$K_m^{\rm ind}(\mathbf{x}) = -i\omega\left[\mu(\mathbf{x}) - \mu_{\rm b}\right] H_m(\mathbf{x}), \quad m = 1,2,3. \tag{3.6.b}$$

It is convenient to write the Maxwell equations in the following matrix form:

$$\begin{bmatrix} -i\omega\varepsilon_{\rm b} & 0 & -\partial_2 \\ 0 & -i\omega\varepsilon_{\rm b} & \partial_1 \\ -\partial_2 & \partial_1 & -i\omega\mu_{\rm b} \end{bmatrix} \begin{bmatrix} E_1^{\rm sc} \\ E_2^{\rm sc} \\ H_3^{\rm sc} \end{bmatrix} = \begin{bmatrix} -J_1^{\rm ind} \\ -J_2^{\rm ind} \\ -K_3^{\rm ind} \end{bmatrix}, \tag{3.7}$$

$$\begin{bmatrix} -i\omega\mu_{\rm b} & 0 & \partial_2 \\ 0 & -i\omega\mu_{\rm b} & -\partial_1 \\ \partial_2 & -\partial_1 & -i\omega\varepsilon_{\rm b} \end{bmatrix} \begin{bmatrix} H_1^{\rm sc} \\ H_2^{\rm sc} \\ E_3^{\rm sc} \end{bmatrix} = \begin{bmatrix} -K_1^{\rm ind} \\ -K_2^{\rm ind} \\ -J_3^{\rm ind} \end{bmatrix}. \tag{3.8}$$

The two-dimensional Fourier transform of (3.7) and (3.8) with respect to coordinates x_1 and x_2 takes us from the $(\mathbf{x},\omega)$ domain to the $(\mathbf{k},\omega)$ domain, whereby $\partial_n \to -ik_n$, $n = 1,2$. In this way we arrive at the following linear algebraic problem of the form $\mathbb{A}\tilde{\mathbf{F}} = -\tilde{\mathbf{S}}$ for the TE case (similar procedure is followed for the TM case which is ommited for brevity):

$$\begin{bmatrix} -i\omega\varepsilon_{\rm b} & 0 & ik_2 \\ 0 & -i\omega\varepsilon_{\rm b} & -ik_1 \\ ik_2 & -ik_1 & -i\omega\mu_{\rm b} \end{bmatrix} \begin{bmatrix} \tilde{E}_1^{\rm sc} \\ \tilde{E}_2^{\rm sc} \\ \tilde{H}_3^{\rm sc} \end{bmatrix} = - \begin{bmatrix} \tilde{J}_1^{\rm ind} \\ \tilde{J}_2^{\rm ind} \\ \tilde{K}_3^{\rm ind} \end{bmatrix}. \tag{3.9}$$

Solving (3.9) via matrix inversion, we obtain the scattered fields in terms of the induced currents:

$$\begin{bmatrix} \tilde{E}_1^{\rm sc} \\ \tilde{E}_2^{\rm sc} \\ \tilde{H}_3^{\rm sc} \end{bmatrix} = \begin{bmatrix} \frac{k_1^2 - \omega^2\varepsilon_{\rm b}\mu_{\rm b}}{i\omega\varepsilon_{\rm b}} & \frac{k_1 k_2}{i\omega\varepsilon_{\rm b}} & ik_2 \\ \frac{k_1 k_2}{i\omega\varepsilon_{\rm b}} & \frac{k_2^2 - \omega^2\varepsilon_{\rm b}\mu_{\rm b}}{i\omega\varepsilon_{\rm b}} & -ik_1 \\ ik_2 & -ik_1 & i\omega\varepsilon_{\rm b} \end{bmatrix} \begin{bmatrix} \tilde{A}_1 \\ \tilde{A}_2 \\ \tilde{F}_3 \end{bmatrix}, \tag{3.10}$$

where we have introduced the vector potentials

$$\tilde{A}_k = \frac{1}{k_1^2 + k_2^2 - \omega^2\varepsilon_{\rm b}\mu_{\rm b}} \tilde{J}_k^{\rm ind}, \quad k = 1,2,3; \tag{3.11.a}$$

$$\tilde{F}_m = \frac{1}{k_1^2 + k_2^2 - \omega^2\varepsilon_{\rm b}\mu_{\rm b}} \tilde{K}_m^{\rm ind}, \quad m = 1,2,3. \tag{3.11.b}$$

Transforming (3.10) back to the $(\mathbf{x},\omega)$ domain and recognizing the partial derivatives as $k_n \to i\partial_n$, $n = 1,2$, we get

$$\begin{bmatrix} E_1^{\rm sc} \\ E_2^{\rm sc} \\ H_3^{\rm sc} \end{bmatrix} = \begin{bmatrix} -\frac{\partial_1^2}{i\omega\varepsilon_{\rm b}} - \frac{\omega^2\varepsilon_{\rm b}\mu_{\rm b}}{i\omega\varepsilon_{\rm b}} & -\frac{\partial_1\partial_2}{i\omega\varepsilon_{\rm b}} & -\partial_2 \\ -\frac{\partial_1\partial_2}{i\omega\varepsilon_{\rm b}} & -\frac{\partial_2^2}{i\omega\varepsilon_{\rm b}} - \frac{\omega^2\varepsilon_{\rm b}\mu_{\rm b}}{i\omega\varepsilon_{\rm b}} & \partial_1 \\ -\partial_2 & \partial_1 & i\omega\varepsilon_{\rm b} \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ F_3 \end{bmatrix}. \tag{3.12}$$

The $(\mathbf{x}, \omega)$-domain vector potentials are the spatial convolutions of the induced currents with the scalar Green's function, namely

$$A_k(\mathbf{x}, \omega) = \int_{\mathbf{x}' \in \mathbb{R}^2} g(\mathbf{x} - \mathbf{x}', \omega) J_k^{\text{ind}}(\mathbf{x}', \omega) \mathrm{d}\mathbf{x}', \quad k = 1, 2, 3, \tag{3.13.a}$$

$$F_m(\mathbf{x}, \omega) = \int_{\mathbf{x}' \in \mathbb{R}^2} g(\mathbf{x} - \mathbf{x}', \omega) K_m^{\text{ind}}(\mathbf{x}', \omega) \mathrm{d}\mathbf{x}', \quad m = 1, 2, 3. \tag{3.13.b}$$

The Green's function $g(\mathbf{x}, \omega)$ is the two-dimensional inverse Fourier transform of the expression $1/(k_1^2 + k_2^2 - \omega^2 \varepsilon_\mathrm{b} \mu_\mathrm{b})$ appearing in (3.11). It is easy to verify that this Green's function satisfies the non-homogeneous Helmholtz equation with a line current (two-dimensional Dirac's delta function) as a source term, located at $\mathbf{x}'$ position. Since the scattered fields are supposed to satisfy the so-called radiation boundary condition (i.e. outgoing waves decaying at infinity), out of the two possible solutions of the said Helmholtz equation one choses

$$g(\mathbf{x} - \mathbf{x}', \omega) = \frac{i}{4} H_0^{(1)}(k_\mathrm{b} |\mathbf{x} - \mathbf{x}'|). \tag{3.14}$$

The other possible solution has the form of the Hankel function of the second kind, and is chosen when a different time convention is used, i.e. for the time-dependence of the form $e^{i\omega t}$.

Substituting the induced currents from (3.6) and expressing the scattered fields as $\mathbf{E}^{\text{sc}} = \mathbf{E} - \mathbf{E}^{\text{in}}$ and $\mathbf{H}^{\text{sc}} = \mathbf{H} - \mathbf{H}^{\text{in}}$, we arrive at the following integro-differential equations, for TE and TM cases respectively, with the total fields as the fundamental unknown:

$$\begin{bmatrix} E_1^{\text{in}} \\ E_2^{\text{in}} \\ H_3^{\text{in}} \end{bmatrix} = \begin{bmatrix} E_1 \\ E_2 \\ H_3 \end{bmatrix} - \begin{bmatrix} k_\mathrm{b}^2 + \partial_1^2 & \partial_1 \partial_2 & -i\omega\mu_\mathrm{b}(-\partial_2) \\ \partial_2 \partial_1 & k_\mathrm{b}^2 + \partial_2^2 & -i\omega\mu_\mathrm{b}\partial_1 \\ -i\omega\varepsilon_\mathrm{b}(-\partial_2) & -i\omega\varepsilon_\mathrm{b}\partial_1 & k_\mathrm{b}^2 \end{bmatrix} \begin{bmatrix} g * (\chi_\mathrm{e} E_1) \\ g * (\chi_\mathrm{e} E_2) \\ g * (\chi_\mathrm{m} H_3) \end{bmatrix}, \tag{3.15}$$

$$\begin{bmatrix} H_1^{\text{in}} \\ H_2^{\text{in}} \\ E_3^{\text{in}} \end{bmatrix} = \begin{bmatrix} H_1 \\ H_2 \\ E_3 \end{bmatrix} - \begin{bmatrix} k_\mathrm{b}^2 + \partial_1^2 & \partial_1 \partial_2 & -i\omega\varepsilon_\mathrm{b}\partial_2 \\ \partial_2 \partial_1 & k_\mathrm{b}^2 + \partial_2^2 & -i\omega\varepsilon_\mathrm{b}(-\partial_1) \\ -i\omega\mu_\mathrm{b}\partial_2 & -i\omega\mu_\mathrm{b}(-\partial_1) & k_\mathrm{b}^2 \end{bmatrix} \begin{bmatrix} g * (\chi_\mathrm{m} H_1) \\ g * (\chi_\mathrm{m} H_2) \\ g * (\chi_\mathrm{e} E_3) \end{bmatrix}, \tag{3.16}$$

where

$$\chi_\mathrm{e}(\mathbf{x}, \omega) = \frac{\varepsilon(\mathbf{x}, \omega)}{\varepsilon_\mathrm{b}} - 1, \tag{3.17}$$

$$\chi_\mathrm{m}(\mathbf{x}, \omega) = \frac{\mu(\mathbf{x}, \omega)}{\mu_\mathrm{b}} - 1, \tag{3.18}$$

are the normalized electric and magnetic contrast functions, respectively, and the star $(*)$ denotes the 2D convolution, for example

$$g * (\chi_\mathrm{e} E_1) = \int_{\mathbf{x}' \in \mathbb{R}^2} g(\mathbf{x} - \mathbf{x}', \omega) \chi_\mathrm{e}(\mathbf{x}') E_1(\mathbf{x}') \mathrm{d}\mathbf{x}'. \tag{3.19}$$

Although the integro-differential equations (3.15)–(3.16) can to a certain extent be analyzed directly on the pertaining Sobolev spaces [33, 17], we prefer to transform them into singular integral equations and work on the Hilbert space $L_2^{(3)}(D)$ of vector-valued functions with spatial support on D.

As the two problems (3.15) and (3.16) are mathematically identical, up to some constants and non-essential sign changes in the differential matrix operator, we shall concentrate on just one of them, say the TE case (3.15). A standard singular integral

equation is obtained from (3.15) by carrying out the spatial derivatives of the weakly singular integrals (3.19), and writing the result as

$$u^{\text{in}} = Du + GXu + KXu, \tag{3.20}$$

where D and X are "diagonal" multiplication operators, G is a principal-value singular integral operator, and K is a compact integral operator. In our case, things become a little more complicated due to the matrix-valued kernel and the involvement of special functions.

The system (3.15) contains nine scalar integro-differential operators. If the partial derivatives are carried out, then we arrive at nine "pure" integral operators whose kernels have different degrees of singularity. The weakly singular kernels result in compact operators, whereas the strongly singular kernels require some extra caution. The presence of the second-order partial derivatives in the upper left (2×2)-corner of the derivative matrix in (3.15) indicates that we have four scalar strongly singular kernels in the TE case. The corresponding operators are called singular integral operators.

To arrive at the standard form of the singular integral operator we first separate the domain of integration D in two sub-domains as $D = [D \setminus D(\epsilon)] \cup D(\epsilon)$ where $D(\epsilon)$ is a circular area around $\mathbf{x}$ with the radius ϵ. Let u_r be either of the electric field components E_1 or E_2, then the product between the operator matrix and the convolution vector in (3.15) can be written as

$$\begin{aligned} I_1 = &\lim_{\epsilon \to 0} \int_{\mathbf{x}' \in [D \setminus D(\epsilon)]} \partial_k \partial_r g(\mathbf{x} - \mathbf{x}') \chi_{\text{e}}(\mathbf{x}') u_r(\mathbf{x}') \mathrm{d}\mathbf{x}' \\ &+ \lim_{\epsilon \to 0} \partial_k \int_{\mathbf{x}' \in D(\epsilon)} \partial_r g(\mathbf{x} - \mathbf{x}') \chi_{\text{e}}(\mathbf{x}') u_r(\mathbf{x}') \mathrm{d}\mathbf{x}', \quad k, r = 1, 2. \end{aligned} \tag{3.21}$$

Here and in what follows we save some space by not showing the parametric dependence on ω. The first term in (3.21) is recognized as the principal value, while the second integral will be denoted by I_2. This second term incorporates the Green's function of (3.14). Utilizing asymptotic expansions for small arguments for the zero order Bessel and Neumann functions [1], namely

$$J_0(z) = 1 - \frac{z^2}{4} + O(z^4), \tag{3.22}$$

$$N_0(z) = \frac{2}{\pi} \gamma J_0(z) + \frac{2}{\pi} J_0(z) \ln \frac{z}{2} + O(z^2), \tag{3.23}$$

we replace the Hankel function in Green's function by its asymptotic expansion based on (3.22) and (3.23), arriving at

$$\begin{aligned} g(\mathbf{x} - \mathbf{x}') &= -\frac{\gamma}{2\pi} + \frac{i}{4} - \frac{1}{2\pi} \ln \frac{k_{\text{b}} |\mathbf{x} - \mathbf{x}'|}{2} + O(|\mathbf{x} - \mathbf{x}'|^2) \\ &= g_0(\mathbf{x} - \mathbf{x}') + g_1(\mathbf{x} - \mathbf{x}'), \end{aligned} \tag{3.24}$$

where we have defined

$$g_0(\mathbf{x} - \mathbf{x}') = g(\mathbf{x} - \mathbf{x}') - \left[-\frac{1}{2\pi} \ln \frac{k_{\text{b}} |\mathbf{x} - \mathbf{x}'|}{2} \right], \tag{3.25}$$

and

$$g_1(\mathbf{x} - \mathbf{x}') = -\frac{1}{2\pi} \ln \frac{k_{\text{b}} |\mathbf{x} - \mathbf{x}'|}{2}. \tag{3.26}$$

The function $g_0(\mathbf{x}-\mathbf{x}')$ is not singular and its contribution (even after differentiation) in I_2 will vanish in the limit $\epsilon \to 0$. The nonzero contribution to I_2 comes from $g_1(\mathbf{x}-\mathbf{x}')$. Taking the first partial derivative ∂_r of $g_1(\mathbf{x}-\mathbf{x}')$, we get

$$\partial_r g_1(\mathbf{x}-\mathbf{x}') = \frac{1}{2\pi}\frac{-\Theta_r}{|\mathbf{x}-\mathbf{x}'|}, \tag{3.27}$$

where $\Theta_r = (x_r - x_r')/|\mathbf{x}-\mathbf{x}'|$. A singularity of order one has appeared, which is still a weak singularity (the order of the singularity is smaller than the dimension of the manifold, which is two-dimensional in the present problem). Applying the second partial derivative ∂_k we obtain a strong second-order singularity:

$$\partial_k\partial_r g_1(\mathbf{x}-\mathbf{x}') = \frac{1}{2\pi}\partial_k\frac{-\Theta_r}{|\mathbf{x}-\mathbf{x}'|} = \frac{1}{2\pi}\frac{1}{|\mathbf{x}-\mathbf{x}'|^2}\left(2\Theta_k\Theta_r - \delta_{kr}\right). \tag{3.28}$$

Thus, omitting the g_0 part (since it disappears in the limit) we re-write the second term of (3.21) as

$$\begin{aligned}&\lim_{\epsilon\to 0}\partial_k\int_{\mathbf{x}'\in D(\epsilon)}\partial_r g(\mathbf{x}-\mathbf{x}')\chi_{\mathrm{e}}(\mathbf{x}')u_r(\mathbf{x}')\mathrm{d}\mathbf{x}'\\ &= \lim_{\epsilon\to 0}\int_{\mathbf{x}'\in D(\epsilon)}\frac{1}{2\pi}\partial_k\frac{-\Theta_r}{|\mathbf{x}-\mathbf{x}'|}\chi_{\mathrm{e}}(\mathbf{x}')u_r(\mathbf{x}')\mathrm{d}\mathbf{x}'.\end{aligned} \tag{3.29}$$

By adding and subtracting $\chi_{\mathrm{e}}(\mathbf{x})u_r(\mathbf{x})$ inside the integral of (3.29), we get

$$\begin{aligned}&\lim_{\epsilon\to 0}\partial_k\int_{\mathbf{x}'\in D(\epsilon)}\partial_r g(\mathbf{x}-\mathbf{x}')\chi_{\mathrm{e}}(\mathbf{x}')u_r(\mathbf{x}')\mathrm{d}\mathbf{x}'\\ &= \frac{1}{2\pi}\lim_{\epsilon\to 0}\int_{\mathbf{x}'\in D(\epsilon)}\partial_k\frac{-\Theta_r}{|\mathbf{x}-\mathbf{x}'|}\left[\chi_{\mathrm{e}}(\mathbf{x}')u_r(\mathbf{x}') - \chi_{\mathrm{e}}(\mathbf{x})u_r(\mathbf{x})\right]\mathrm{d}\mathbf{x}'\\ &+ \frac{1}{2\pi}\lim_{\epsilon\to 0}\chi_{\mathrm{e}}(\mathbf{x})u_r(\mathbf{x})\int_{\mathbf{x}'\in D(\epsilon)}\partial_k\frac{-\Theta_r}{|\mathbf{x}-\mathbf{x}'|}\mathrm{d}\mathbf{x}'.\end{aligned} \tag{3.30}$$

Assuming the Hölder continuity of the function $\chi_{\mathrm{e}}(\mathbf{x})u_r(\mathbf{x})$, i.e., assuming that there exist $\alpha, C > 0$, such that for all $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^2$,

$$|\chi_{\mathrm{e}}(\mathbf{x})u_r(\mathbf{x}) - \chi_{\mathrm{e}}(\mathbf{x}')u_r(\mathbf{x}')| \le C|\mathbf{x}-\mathbf{x}'|^{\alpha}, \tag{3.31}$$

we effectively lower the order of singularity in the first integral in the RHS of (3.30). Hence, in the limit $\epsilon \to 0$ this term is zero. Interchanging ∂_k with $-\partial_k'$ and applying the 2D divergence theorem in the remaining term of (3.30), we obtain

$$\begin{aligned}&\lim_{\epsilon\to 0}\partial_k\int_{\mathbf{x}'\in D(\epsilon)}\partial_r g(\mathbf{x}-\mathbf{x}')\chi_{\mathrm{e}}(\mathbf{x}')u_r(\mathbf{x}')\mathrm{d}\mathbf{x}'\\ &= \frac{1}{2\pi}\chi_{\mathrm{e}}(\mathbf{x})u_r(\mathbf{x})\lim_{\epsilon\to 0}\oint_{\mathbf{x}'\in\partial D(\epsilon)} -\nu_k\frac{-\Theta_r}{|\mathbf{x}-\mathbf{x}'|}\mathrm{d}\mathbf{x}',\end{aligned} \tag{3.32}$$

with $\nu_k = -\Theta_k$ being the projection of the normal unit vector of the polar coordinate system on the x_k axis. Finally, we arrive at the result

$$\begin{aligned}&\lim_{\epsilon\to 0}\partial_k\int_{\mathbf{x}'\in D(\epsilon)}\partial_r g(\mathbf{x}-\mathbf{x}')\chi_{\mathrm{e}}(\mathbf{x}')u_r(\mathbf{x}')\mathrm{d}\mathbf{x}'\\ &= \frac{1}{2\pi}\chi_{\mathrm{e}}(\mathbf{x})u_r(\mathbf{x})\lim_{\epsilon\to 0}\oint_{\mathbf{x}'\in\partial D(\epsilon)} -\frac{\Theta_k\Theta_r}{|\mathbf{x}-\mathbf{x}'|}\mathrm{d}\mathbf{x}'\\ &= -\frac{1}{2\pi}\chi_{\mathrm{e}}(\mathbf{x})u_r(\mathbf{x})\lim_{\epsilon\to 0}\oint_{\mathbf{x}'\in\partial D(1)}\Theta_k\Theta_r\mathrm{d}\mathbf{x}'\\ &= -\frac{1}{2}\chi_{\mathrm{e}}(\mathbf{x})u_r(\mathbf{x})\delta_{rk},\end{aligned} \tag{3.33}$$

where we notice that the last contour integral is over the unit circle. Hence, (3.21) can now be written as

$$I_1 = p.v. \int_{\mathbf{x}' \in D} \partial_k \partial_r g(\mathbf{x}-\mathbf{x}') \chi_{\mathrm{e}}(\mathbf{x}') u_r(\mathbf{x}') \mathrm{d}\mathbf{x}' - \frac{1}{2} \chi_{\mathrm{e}}(\mathbf{x}) u_r(\mathbf{x}) \delta_{rk}, \quad k, r = 1, 2. \tag{3.34}$$

Now, we can represent (3.15) in the standard form

$$\begin{bmatrix} E_1^{\mathrm{in}} \\ E_2^{\mathrm{in}} \\ H_3^{\mathrm{in}} \end{bmatrix} = \begin{bmatrix} S & 0 & 0 \\ 0 & S & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \\ H_3 \end{bmatrix} + p.v. \begin{bmatrix} G_{11} & G_{12} & 0 \\ G_{21} & G_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix} * \begin{bmatrix} X_{\mathrm{e}} E_1 \\ X_{\mathrm{e}} E_2 \\ X_{\mathrm{m}} H_3 \end{bmatrix} \tag{3.35}$$

$$+ \begin{bmatrix} K_{11} & K_{12} & K_{13} \\ K_{21} & K_{22} & K_{23} \\ K_{31} & K_{32} & K_{33} \end{bmatrix} * \begin{bmatrix} X_{\mathrm{e}} E_1 \\ X_{\mathrm{e}} E_2 \\ X_{\mathrm{m}} H_3 \end{bmatrix}, \tag{3.36}$$

with S denoting the operator of (pointwise) multiplication with the function $s(\mathbf{x}) = 1+1/2\chi_{\mathrm{e}}(\mathbf{x})$, and I – the identity operator. The kernels in the principal-value operator in (3.35) are easily recognized using (3.28) and (3.15), and are given by

$$G_{nm}(\mathbf{x}) = -\frac{1}{2\pi|\mathbf{x}|^2} \left[2\Theta_n \Theta_m - \delta_{nm}\right], \quad n, m = 1, 2. \tag{3.37}$$

The kernels in the compact operator in (3.35) are easily obtained if we first calculate the first order derivative of Green's function which is

$$\partial_r g(\mathbf{x}-\mathbf{x}') = -\frac{i}{4} k_{\mathrm{b}} \frac{x_r - x_r'}{|\mathbf{x}-\mathbf{x}'|} H_1^{(1)}(k_{\mathrm{b}}|\mathbf{x}-\mathbf{x}'|), \quad r = 1, 2. \tag{3.38}$$

as well as the mixed derivative

$$\partial_k \partial_r g(\mathbf{x}-\mathbf{x}') = \frac{i}{4} \Bigg[k_{\mathrm{b}} \frac{1}{|\mathbf{x}-\mathbf{x}'|} \left(2 \frac{x_r - x_r'}{|\mathbf{x}-\mathbf{x}'|} \frac{x_k - x_k'}{|\mathbf{x}-\mathbf{x}'|} - \delta_{kr} \right) H_1^{(1)}(k_{\mathrm{b}}|\mathbf{x}-\mathbf{x}'|) - k_{\mathrm{b}}^2 \frac{x_r - x_r'}{|\mathbf{x}-\mathbf{x}'|} \frac{x_k - x_k'}{|\mathbf{x}-\mathbf{x}'|} H_0^{(1)}(k_{\mathrm{b}}|\mathbf{x}-\mathbf{x}'|) \Bigg], \quad k, r = 1, 2, \tag{3.39}$$

where δ_{kr} is the Kronecker's delta. So, using (3.38), (3.39) and (3.15), the aforementioned compact kernels are given by

$$K_{nm} = \left[\frac{1}{2\pi|\mathbf{x}|^2} - \frac{ik_{\mathrm{b}}}{4|\mathbf{x}|} H_1^{(1)}(k_{\mathrm{b}}|\mathbf{x}|) \right] \left[2\Theta_n \Theta_m - \delta_{nm}\right] + \frac{ik_{\mathrm{b}}^2}{4} H_0^{(1)}(k_{\mathrm{b}}|\mathbf{x}|) \left[\Theta_n \Theta_m - \delta_{nm}\right], \quad n, m = 1, 2; \tag{3.40}$$

$$K_{13} = -\frac{\omega \mu_{\mathrm{b}} k_{\mathrm{b}} \Theta_2}{4} H_1^{(1)}(k_{\mathrm{b}}|\mathbf{x}|); \tag{3.41}$$

$$K_{31} = -\frac{\omega \varepsilon_{\mathrm{b}} k_{\mathrm{b}} \Theta_2}{4} H_1^{(1)}(k_{\mathrm{b}}|\mathbf{x}|); \tag{3.42}$$

$$K_{23} = \frac{\omega \mu_{\mathrm{b}} k_{\mathrm{b}} \Theta_1}{4} H_1^{(1)}(k_{\mathrm{b}}|\mathbf{x}|); \tag{3.43}$$

$$K_{32} = \frac{\omega \varepsilon_{\mathrm{b}} k_{\mathrm{b}} \Theta_1}{4} H_1^{(1)}(k_{\mathrm{b}}|\mathbf{x}|); \tag{3.44}$$

$$K_{33} = -\frac{ik_b^2}{4} H_0^{(1)}(k_b|\mathbf{x}|). \tag{3.45}$$

As can be deduced from asymptotic expansions of special functions, all K_{nm} kernels are weakly singular, i.e., their singularity is less than that of the factor $|\mathbf{x}-\mathbf{x}'|^{-2}$, and the corresponding integrals exist in the usual sense. The remaining strong singularity is explicitly contained within the G_{nm} kernel. The latter kernel also features a very special tensor $2\theta_n\theta_m - \delta_{nm}$, which guarantees the existence of the integral in the sense of a principal value. The standard assumption, sufficient to guarantee the boundedness of the matrix singular integral operator G and the equivalence of the integral equation (3.20) to the original Maxwell's equations, is that of Hölder continuity of the dielectric permittivity and of all components of the incident field vector for all $\mathbf{x} \in \mathbb{R}^2$.

The solution is sought on $L_2^{(3)}(D \subset \mathbb{R}^2)$, i.e., inside a finite spatial domain D. It is well-known that an integral operator with a weakly singular kernel is compact on $L_2^{(3)}(D)$, however, the same operator is, generally, not compact on $L_2^{(3)}(\mathbb{R}^2)$, since a weakly singular kernel often decays too slowly at infinity. On the other hand, analysis of the existence of the solution of integral equations with a strongly singular integral operator is usually performed on domains without an edge, i.e., on $L_2^{(3)}(\mathbb{R}^2)$, see [46]. A way around this apparent technical contradiction was proposed in [57] (Lemma 9.1, pp. 45–46), where the analysis of existence is performed on $L_2^{(3)}(\mathbb{R}^2)$ with the weakly singular integral operators replaced by equivalent compact operators. The procedure involves a weighting (pointwise multiplication) of the equations with a finite differentiable function (equal to one inside D). The existence of the solution of such a weighted integral equation established by standard methods on $L_2^{(3)}(\mathbb{R}^2)$ implies the existence of the solution of the original equation on $L_2^{(3)}(D)$.

3.2 Symbol & existence

At the core of the Mikhlin-Prössdorf approach, successfully applied in [5, 57] to the three-dimensional volume integral equation, is the symbol calculus. It allows to study the existence of a solution and shows a way to reduce the original singular integral operator to a manifestly Fredholm operator of the form "identity plus compact". The symbol of an integral operator in the present two-dimensional case is a matrix-valued function $\Phi_{nm}(\mathbf{x}, \mathbf{k})$, $n, m = 1, 2, 3$, $\mathbf{x} \in \mathbb{R}^2$, and $\mathbf{k} \in \mathbb{R}^2$. Construction of the symbol follows a few simple rules explained in [46, 57, 5]: the symbol of the sum/product of operators is the sum/product of symbols; the symbol of a compact operator is zero; the symbol of a multiplication operator is the multiplier function itself; the symbol of the strongly singular integral operator is the principal value of the Fourier transform of its kernel (that is where the second argument $\mathbf{k}$ – the dual of $\mathbf{x}$ – comes from). It follows, in particular, that the symbol of the matrix operator is a matrix-valued function. Applying these rules with the use of operator notation, (3.35) can be written as

$$(A\mathbf{E})(\mathbf{x}) = \left(\mathbb{E} + \frac{1}{2}\mathbb{M}\right)\mathbf{E}(\mathbf{x}) + \mathbb{A}^s\mathbb{M}\mathbf{E}(\mathbf{x}') + \mathbb{K}\mathbf{E}(\mathbf{x}'), \tag{3.46}$$

where $\mathbb{A}^s$ is a (matrix) singular integral operator, $\mathbb{K}$ is a (matrix) compact integral operator, $\mathbb{E}$ is a (matrix) identity operator, and $\mathbb{M}(\mathbf{x}) = \chi_e(\mathbf{x})\mathbb{E}_2$ is a (matrix) multiplication operator, where the lower-right element of $\mathbb{E}_2$ is set to zero. The singular

term of (3.46) is of the form

$$(A^{\mathrm{s}}\mathbb{M}\mathbf{E})(\mathbf{x}) = \int_{\mathbf{x}'\in D} \frac{\mathbb{F}(\boldsymbol{\Theta})}{|\mathbf{x}-\mathbf{x}'|^2}\chi_{\mathrm{e}}(\mathbf{x}')\mathbf{E}(x')\mathrm{d}\mathbf{x}'. \tag{3.47}$$

According to (3.28), the characteristic (matrix) function $\mathbb{F}(\boldsymbol{\Theta})$ has the form

$$\mathbb{F}(\boldsymbol{\Theta}) = -\frac{1}{2\pi}[2\mathbb{Q}(\mathbf{x}-\mathbf{x}') - \mathbb{I}_2], \tag{3.48}$$

where $\mathbb{I}_2$ is the (3×3) identity matrix with the lower-right element set to zero, and the tensor $\mathbb{Q}$ is given by

$$\mathbb{Q}(\mathbf{x}-\mathbf{x}') = \begin{bmatrix} \frac{(x_1-x_1')^2}{|\mathbf{x}-\mathbf{x}'|^2} & \frac{(x_1-x_1')(x_2-x_2')}{|\mathbf{x}-\mathbf{x}'|^2} & 0 \\ \frac{(x_1-x_1')(x_2-x_2')}{|\mathbf{x}-\mathbf{x}'|^2} & \frac{(x_2-x_2')^2}{|\mathbf{x}-\mathbf{x}'|^2} & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{3.49}$$

For $\chi_{\mathrm{e}}(\mathbf{x})$ Hölder-continuous on $\mathbb{R}^2$, the symbol of the compound singular integral operator of (3.46) can be computed as (see [46]):

$$\mathrm{Smb}(\mathbb{A}) = \mathbb{I} + \frac{1}{2}\chi_{\mathrm{e}}(\mathbf{x})\mathbb{I}_2 + \mathrm{Smb}(\mathbb{A}^{\mathrm{s}})\chi_{\mathrm{e}}(\mathbf{x}), \tag{3.50}$$

where $\mathbb{I}$ is the ordinary (3×3) identity matrix. Since the characteristics $\mathbb{F}$ depends only on $\mathbf{x}-\mathbf{x}'$, the symbol of the singular integral operator $\mathbb{A}^{\mathrm{s}}$ is the Fourier transform of its kernel $\mathbb{F}(\boldsymbol{\Theta})/|\mathbf{x}-\mathbf{x}'|^2$ with respect to the variable $\mathbf{y} = \mathbf{x}-\mathbf{x}'$, i.e. it is a $\mathbf{k}$-domain matrix-valued function $\tilde{\mathbb{F}}^{\mathrm{s}}(\mathbf{k})$. Computing this Fourier transform is a daunting task, and we shall use a shortcut proposed in [46].

Let A^{s} be a single component of our matrix-valued operator $\mathbb{A}^{\mathrm{s}}$ and let $\Phi^{\mathrm{s}}(\mathbf{k})$ denote its symbol, which is one of the components of the matrix-valued symbol function $\tilde{\mathbb{F}}^{\mathrm{s}}(\mathbf{k})$ we are trying to compute. The symbol $\mathrm{Smb}(A^{\mathrm{s}}) = \Phi^{\mathrm{s}}(\mathbf{k})$ of a scalar simple singular integral operator can be expanded in a series of 2D spherical functions of order p, that is, in Fourier series of sines and cosines [46], namely

$$\Phi^{\mathrm{s}}(\tilde{\boldsymbol{\Theta}}) = \sum_{p=0}^{\infty}\left[\gamma_{2,p}a_p^{(1)}Y_{p,2}^{(1)}(\tilde{\boldsymbol{\Theta}}) + \gamma_{2,p}a_p^{(2)}Y_{p,2}^{(2)}(\tilde{\boldsymbol{\Theta}})\right], \tag{3.51}$$

where $\tilde{\boldsymbol{\Theta}} = \mathbf{k}/|\mathbf{k}|$ is the unit vector in the $\mathbf{k}$-domain, $Y_{p,2}^{(1)} = \sin(p\tilde{\phi})$, and $Y_{p,2}^{(2)} = \cos(p\tilde{\phi})$ is the basis of the expansion, $a_p^{(1)}$ and $a_p^{(2)}$ are the expansion coefficients, and $\tilde{\phi}$ is the directional angle of the unit vector $\tilde{\boldsymbol{\Theta}}$. In (3.51), $\gamma_{2,p} = \pi i^p\Gamma(p/2)/\Gamma((2+p)/2)$ [46]. Since any component $f(\boldsymbol{\Theta})$ of the characteristic matrix-valued function $\mathbb{F}(\boldsymbol{\Theta})$ given by (3.48) depend only on $\boldsymbol{\Theta}$, we can expand each of them in a Fourier series as well

$$f(\boldsymbol{\Theta}) = \sum_{p=1}^{\infty}\left[a_p^{(1)}\sin(p\phi) + a_p^{(2)}\cos(p\phi)\right], \tag{3.52}$$

where ϕ is the directional angle of the unit vector $\boldsymbol{\Theta} = (\mathbf{x}-\mathbf{x}')/|\mathbf{x}-\mathbf{x}'|$. It was shown in [46] that the expansion coefficients in (3.52) are the same with those of (3.51). Using (3.48) and (3.49), we see that the components of the characteristic $\mathbb{F}$ are

$$\mathbb{F}(\boldsymbol{\Theta}) = -\frac{1}{2\pi}[2\mathbb{Q}(\mathbf{x}-\mathbf{x}') - \mathbb{I}_2] = -\frac{1}{2\pi}\begin{bmatrix} 2\cos^2\phi - 1 & 2\cos\phi\sin\phi & 0 \\ 2\cos\phi\sin\phi & 2\sin^2\phi - 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{3.53}$$

To calculate the expansion coefficients, we must equate each element $[\mathbb{F}(\boldsymbol{\Theta})]_{kr}$, $k, r = 1, 2$, from (3.53) with the series from (3.52). For example, for the element $[\mathbb{F}(\boldsymbol{\Theta})]_{11}$ we have

$$-\frac{1}{2\pi}[2\cos^2\phi - 1] = -\frac{1}{2\pi}\cos(2\phi) = \sum_{p=1}^{\infty}\Big[a_p^{(1)}\sin(p\phi) + a_p^{(2)}\cos(p\phi)\Big]. \tag{3.54}$$

From this expansion it is obvious that the only nonzero expansion coefficient is $a_2^{(2)} = -1/(2\pi)$, while the rest are all zero (i.e. $a_p^{(1)} = 0\ \forall p$ and $a_p^{(2)} = 0\ \forall p \neq 2$). The same procedure is followed for the rest of the components in (3.53). Then, we substitute the known expansion coefficients in (3.51) to get the elements $[\tilde{\mathbb{F}}^{\mathrm{s}}(\tilde{\boldsymbol{\Theta}})]_{kr}$, $k, r = 1, 2$. Continuing our example, the element $[\tilde{\mathbb{F}}^{\mathrm{s}}(\tilde{\boldsymbol{\Theta}})]_{11}$ is obtained through the substitutions $\gamma_{2,2} = \pi i^2\Gamma(1)/\Gamma(2) = -\pi$, $Y_{2,2}^{(2)} = \cos(2\tilde{\phi})$ and $a_2^{(2)} = -1/(2\pi)$, and therefore $[\tilde{\mathbb{F}}^{\mathrm{s}}(\tilde{\boldsymbol{\Theta}})]_{11} = 1/2\cos(2\tilde{\phi}) = \cos^2(\tilde{\phi}) - 1/2$. Following the same procedure for all components, we finally get

$$\mathrm{Smb}(\mathbb{A}^{\mathrm{s}}) = \tilde{\mathbb{F}}^{\mathrm{s}}(\tilde{\boldsymbol{\Theta}}) = \begin{bmatrix} \cos^2\tilde{\phi} - 1/2 & \sin\tilde{\phi}\cos\tilde{\phi} & 0 \\ \sin\tilde{\phi}\cos\tilde{\phi} & \sin^2\tilde{\phi} - 1/2 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \mathbb{Q}(\mathbf{k}) - \frac{1}{2}\mathbb{I}_2. \tag{3.55}$$

Substituting the result of (3.55) back in (3.50), we get the symbol of the complete operator as the following (3×3) matrix-valued function:

$$\mathrm{Smb}(\mathbb{A})(\mathbf{x}, \mathbf{k}) = \mathbb{I} + \chi_{\mathrm{e}}(\mathbf{x})\mathbb{Q}(\mathbf{k}). \tag{3.56}$$

Several important conclusions can be made already at this stage. First of all we notice an almost complete equivalence of the symbol in the two-dimensional TE case with the previously obtained symbol of the three-dimensional scattering operator [5, 57]. The necessary and sufficient condition for the existence of a solution is now readily obtained as the condition on the invertibility of Φ. The explicit inverse of the symbol matrix is found to be

$$\Phi^{-1}(\mathbf{x}, \tilde{\phi}) = \mathbb{I} + \chi'_{\mathrm{e}}(\mathbf{x})\mathbb{Q}, \tag{3.57}$$

where $\chi'_{\mathrm{e}}(\mathbf{x}, \omega) = \varepsilon_{\mathrm{b}}/\varepsilon(\mathbf{x}, \omega) - 1$. Hence, the existence condition is

$$\varepsilon(\mathbf{x}, \omega) \neq 0, \quad \mathbf{x} \in \mathbb{R}^2. \tag{3.58}$$

In the TM case (3.16) the symbol can be derived along the same lines and turns out to be $\mathbb{I} + \chi_{\mathrm{m}}(\mathbf{x})\mathbb{Q}$, with the existence condition being $\mu(\mathbf{x}, \omega) \neq 0$, $\mathbf{x} \in \mathbb{R}^2$. It is interesting to note the difference in these existence conditions with respect to the three-dimensional case, where both $\varepsilon(\mathbf{x})$ and $\mu(\mathbf{x})$ are required to be nonzero at the same time.

3.3 Regularizer, spectrum & uniqueness

The inverse of the symbol is the symbol of the regularizer – an operator, which reduces the original singular integral operator to the form "identity plus compact". Since Φ^{-1} has the same form as Φ, we conclude that our original operator with its electric contrast function changed into χ'_{e} may be employed as a regularizer. Thus, if $A(\chi_{\mathrm{e}})$ is the original operator from (3.20)–(3.35), and $A(\chi'_{\mathrm{e}})$ denotes the same operator with the contrast function replaced by χ'_{e}, then we have

$$A(\chi'_{\mathrm{e}})A(\chi_{\mathrm{e}}) = I + K, \tag{3.59}$$

where K is a generic compact operator. This looks strikingly similar to the Calderón identity [2, 3, 15].

It is easy to see that the application of the regularizer eliminates the essential spectrum – a part of the operator spectrum that is invariant under compact perturbations of the operator. Indeed, even if $A(\chi_e)$ has a nontrivial essential spectrum, densely spread over a part of the complex plane, the spectrum of $I + K$ consists of isolated points only. As can be deduced from the symbol (3.56), the essential spectrum of the operator $A(\chi_e)$ in the TE case is given by

$$\sigma_{ess} = \left\{ \varepsilon(\mathbf{x})/\varepsilon_b;\ \ \mathbf{x} \in \mathbb{R}^2 \right\}, \tag{3.60}$$

and by $\sigma_{ess} = \left\{ \mu(\mathbf{x})/\mu_b; \mathbf{x} \in \mathbb{R}^2 \right\}$ in the TM case.

A well-known sufficient condition for the uniqueness of the solution has the form $\operatorname{Im} \varepsilon(\mathbf{x})/\varepsilon_b > 0$, $\mathbf{x} \in D$, see e.g. [34]. Simple algebraic manipulations allow expressing the resolvent operator $[A(\chi_e) - \lambda I]^{-1}$ of the eigenvalue problem as the inverse of the original operator $[A(\chi'_e)]^{-1}$ with a contrast function χ'_e containing a new relative permittivity function $\varepsilon'(\mathbf{x}, \lambda)$, which depends on the parameter λ. Then, applying the opposite of the uniqueness condition, i.e. $\operatorname{Im} \varepsilon'(\mathbf{x}, \lambda) \leq 0$, $\mathbf{x} \in D$, one obtains a wedge-shaped bound on the location of discrete eigenvalues:

$$\begin{aligned} &\operatorname{Im} \varepsilon(\mathbf{x}) - [\operatorname{Im} \varepsilon(\mathbf{x}) + \operatorname{Im} \varepsilon_b] \operatorname{Re} \lambda \\ &+ [\operatorname{Re} \varepsilon(\mathbf{x}) - \operatorname{Re} \varepsilon_b] \operatorname{Im} \lambda + \operatorname{Im} \varepsilon_b |\lambda|^2 \leqslant 0, \quad \mathbf{x} \in D. \end{aligned} \tag{3.61}$$

3.4 Discretization & collocation

Equation (3.35) defines an algebraic system $Au = b$ for the numerical evaluation of the fields. To derive the matrix elements, its more convenient to use the equivalent integral form of (3.35).

The application of the matrix operator of (3.15) – the matrix that contains the partial derivatives – on Green's functions, give us the Green's tensor. In order to get the equivalent integral form of (3.35), we need to split the Green's tensor in two parts. The first part, denoted by $\mathbb{A}(\mathbf{x} - \mathbf{x}')$, corresponds to the second order or mixed derivatives only, i.e.

$$\mathbb{A}(\mathbf{x} - \mathbf{x}') = \left(k_b^2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} \partial_1^2 & \partial_1 \partial_2 & 0 \\ \partial_2 \partial_1 & \partial_2^2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right) g(\mathbf{x} - \mathbf{x}'). \tag{3.62}$$

The second part, denoted by $\mathbb{B}(\mathbf{x} - \mathbf{x}')$, corresponds to the first order derivatives only, i.e.

$$\mathbb{B}(\mathbf{x} - \mathbf{x}') = \begin{bmatrix} 0 & 0 & \partial_2 \\ 0 & 0 & -\partial_1 \\ -\partial_2 & \partial_1 & 0 \end{bmatrix} g(\mathbf{x} - \mathbf{x}'). \tag{3.63}$$

To obtain the explicit relations for $\mathbb{A}(\mathbf{x} - \mathbf{x}')$ and $\mathbb{B}(\mathbf{x} - \mathbf{x}')$, we use the first and mixed order derivatives of Green's function given by (3.38) and (3.39), respectively. So, the $\mathbb{A}$ tensor is given by

$$\begin{aligned} \mathbb{A}(\mathbf{x} - \mathbf{x}') = &\frac{i}{4} k_b \frac{1}{|\mathbf{x} - \mathbf{x}'|} H_1^{(1)}(k_b |\mathbf{x} - \mathbf{x}'|)[2\mathbb{Q}(\mathbf{x} - \mathbf{x}') - \mathbb{I}_2] \\ &- \frac{i}{4} k_b^2 H_0^{(1)}(k_b |\mathbf{x} - \mathbf{x}'|)[\mathbb{Q}(\mathbf{x} - \mathbf{x}') - \mathbb{I}]. \end{aligned} \tag{3.64}$$

The tensor $\mathbb{Q}$ was introduced in (3.49) while $\mathbb{I}_2$ is, as explained in Section 3.2, the (3×3) identity matrix with the lower-right element set to zero. The $\mathbb{B}$ tensor is obtained with the use of (3.38)

$$\mathbb{B}(\mathbf{x}-\mathbf{x}') = -\frac{i}{4}k_{\mathrm{b}}H_1^{(1)}(k_{\mathrm{b}}|\mathbf{x}-\mathbf{x}'|)\boldsymbol{\Theta}(\mathbf{x}-\mathbf{x}')\times, \tag{3.65}$$

where we have now introduced the $\boldsymbol{\Theta}\times$ tensor given by

$$\boldsymbol{\Theta}(\mathbf{x}-\mathbf{x}')\times = \begin{bmatrix} 0 & 0 & \frac{x_2-x_2'}{|\mathbf{x}-\mathbf{x}'|} \\ 0 & 0 & -\frac{x_1-x_1'}{|\mathbf{x}-\mathbf{x}'|} \\ -\frac{x_2-x_2'}{|\mathbf{x}-\mathbf{x}'|} & \frac{x_1-x_1'}{|\mathbf{x}-\mathbf{x}'|} & 0 \end{bmatrix}. \tag{3.66}$$

After this splitting, we can easily express (3.35) in the following integral form

$$\begin{aligned} \mathbf{E}^{\mathrm{in}}(\mathbf{x}) = &\left[1+\frac{1}{2}\chi_{\mathrm{e}}(\mathbf{x})\right]\mathbf{E}(\mathbf{x}) \\ &- p.v.\int_{\mathbf{x}'\in D}\mathbb{A}(\mathbf{x}-\mathbf{x}')\chi_{\mathrm{e}}(\mathbf{x}')\mathbf{E}(\mathbf{x}')\mathrm{d}\mathbf{x}' \\ &- i\omega\mu_{\mathrm{b}}\int_{\mathbf{x}'\in D}\mathbb{B}(\mathbf{x}-\mathbf{x}')\chi_{\mathrm{m}}(\mathbf{x}')\mathbf{H}(\mathbf{x}')\mathrm{d}\mathbf{x}', \end{aligned} \tag{3.67}$$

$$\begin{aligned} \mathbf{H}^{\mathrm{in}}(\mathbf{x}) = &\mathbf{H}(\mathbf{x}) - \int_{\mathbf{x}'\in D}\mathbb{A}(\mathbf{x}-\mathbf{x}')\chi_{\mathrm{m}}(\mathbf{x}')\mathbf{H}(\mathbf{x}')\mathrm{d}\mathbf{x}' \\ &+ i\omega\varepsilon_{\mathrm{b}}\int_{\mathbf{x}'\in D}\mathbb{B}(\mathbf{x}-\mathbf{x}')\chi_{\mathrm{e}}(\mathbf{x}')\mathbf{E}(\mathbf{x}')\mathrm{d}\mathbf{x}', \end{aligned} \tag{3.68}$$

where the vectors $\mathbf{E}(\mathbf{x}) = [E_1(\mathbf{x}), E_2(\mathbf{x}), 0]^T$ and $\mathbf{H}(\mathbf{x}) = [0, 0, H_3(\mathbf{x})]^T$. Same applies for $\mathbf{E}^{\mathrm{in}}(\mathbf{x})$ and $\mathbf{H}^{\mathrm{in}}(\mathbf{x})$.

Equation (3.67) can be rewritten as

$$\begin{aligned} \mathbf{E}^{\mathrm{in}}(\mathbf{x}) = &\left[1+\frac{1}{2}\chi_{\mathrm{e}}(\mathbf{x})\right]\mathbf{E}(\mathbf{x}) \\ &- p.v.\int_{\mathbf{x}'\in D_{\mathrm{s}}}\mathbb{A}(\mathbf{x}-\mathbf{x}')\chi_{\mathrm{e}}(\mathbf{x}')\mathbf{E}(\mathbf{x}')\mathrm{d}\mathbf{x}' \\ &- \int_{\mathbf{x}'\in D\setminus D_{\mathrm{s}}}\mathbb{A}(\mathbf{x}-\mathbf{x}')\chi_{\mathrm{e}}(\mathbf{x}')\mathbf{E}(\mathbf{x}')\mathrm{d}\mathbf{x}' \\ &- i\omega\mu_{\mathrm{b}}\int_{\mathbf{x}'\in D_{\mathrm{s}}}\mathbb{B}(\mathbf{x}-\mathbf{x}')\chi_{\mathrm{m}}(\mathbf{x}')\mathbf{H}(\mathbf{x}')\mathrm{d}\mathbf{x}' \\ &- i\omega\mu_{\mathrm{b}}\int_{\mathbf{x}'\in D\setminus D_{\mathrm{s}}}\mathbb{B}(\mathbf{x}-\mathbf{x}')\chi_{\mathrm{m}}(\mathbf{x}')\mathbf{H}(\mathbf{x}')\mathrm{d}\mathbf{x}', \end{aligned} \tag{3.69}$$

where we have separated the domain D into a domain $D \setminus D_{\mathrm{s}}$ "free" of singularity, and a domain D_{s} which encloses the position vector $\mathbf{x}$ and hence the singularity. In this way we have three integrals in the usual sense and one in the sense of principal value.

Let us introduce a uniform two-dimensional N-node grid over a rectangular computational domain D. For simplicity let the grid step h be the same in both directions. Typically the grid step is chosen in accordance with the following empirical rule: find out the highest local value of the refractive index $n = \max\{\sqrt{\varepsilon'(\mathbf{x})\mu'(\mathbf{x})]/\varepsilon_{\mathrm{b}}\mu_{\mathrm{b}}}\}$, $\mathbf{x} \in D$, where prime denotes the real part; choose an integer k – the number of points

per smallest wavelength; and, finally, compute the grid step as $h = \lambda_b/(kn)$, where λ_b is the wavelength in the background medium. We set our discretization here at the traditional level of $k = 15$ points, which, as we verified, gives a small global error (in the order of 10^{-5}) with respect to the analytical solution for a homogeneous circular cylinder. For rectangular grid-conforming boundaries the error with this discretization rule could be even smaller, since most of the error for a circular cylinder comes from a poor geometrical representation of the boundary and/or bad approximation of the area of the circular cross-section.

Applying a simple collocation technique with the mid-point rule, the usual sense integrals over the domain $D \setminus D_s$ would be given by

$$\begin{aligned} &- \int_{\mathbf{x}' \in D \setminus D_s} \mathbb{A}(\mathbf{x} - \mathbf{x}') \chi_e(\mathbf{x}') \mathbf{E}(\mathbf{x}') d\mathbf{x}' \\ &\approx - \sum_{\substack{m=1 \\ m \neq n}}^{N-1} \mathbb{A}(\mathbf{x}_n - \mathbf{x}_m) \chi_e(\mathbf{x}_m) \mathbf{E}(\mathbf{x}_m) S_m, \quad n = 1, 2, \ldots, N; \end{aligned} \tag{3.70}$$

$$\begin{aligned} &- i\omega\mu_b \int_{\mathbf{x}' \in D \setminus D_s} \mathbb{B}(\mathbf{x} - \mathbf{x}') \chi_m(\mathbf{x}') \mathbf{H}(\mathbf{x}') d\mathbf{x}' \\ &\approx -i\omega\mu_b \sum_{\substack{m=1 \\ m \neq n}}^{N-1} \mathbb{B}(\mathbf{x}_n - \mathbf{x}_m) \chi_m(\mathbf{x}_m) \mathbf{H}(\mathbf{x}_m) S_m, \quad n = 1, 2, \ldots, N; \end{aligned} \tag{3.71}$$

where $S_m = h^2$ is the surface of each elementary cell D_m of side h in the computational domain.

We now proceed to calculate the principal value integral in (3.69). We have

$$\begin{aligned} &p.v. \int_{\mathbf{x}' \in D_s} \mathbb{A}(\mathbf{x} - \mathbf{x}') \chi_e(\mathbf{x}') \mathbf{E}(\mathbf{x}') d\mathbf{x}' \\ &= \lim_{\epsilon \to 0} \int_{\mathbf{x}' \in D_s \setminus D(\epsilon)} \mathbb{A}(\mathbf{x} - \mathbf{x}') \chi_e(\mathbf{x}') \mathbf{E}(\mathbf{x}') d\mathbf{x}' \\ &\approx \chi_e(\mathbf{x}_n) \mathbf{E}(\mathbf{x}_n) \lim_{\epsilon \to 0} \int_{\mathbf{x}' \in D_s \setminus D(\epsilon)} \mathbb{A}(\mathbf{x} - \mathbf{x}') d\mathbf{x}', \end{aligned} \tag{3.72}$$

where the domain of integration in the last integral is over the singular cell but with an exception of a small circular neighborhood having radius ϵ. This is illustrated in Figure 3.2. We further define the tensor $\mathbb{L}$ as

$$\mathbb{L} = \lim_{\epsilon \to 0} \int_{\mathbf{x}' \in D_s \setminus D(\epsilon)} \mathbb{A}(\mathbf{x} - \mathbf{x}') d\mathbf{x}'. \tag{3.73}$$

Now we recall the form that $\mathbb{A}(\mathbf{x} - \mathbf{x}')$ has from (3.64). It is convenient to rewrite (3.64) in terms of $\mathbb{Q}(\mathbf{x} - \mathbf{x}')$ and $\mathbb{I}_2$, namely

$$\begin{aligned} \mathbb{A}(\mathbf{x} - \mathbf{x}') = &\frac{i}{4} \left[2k_b \frac{1}{|\mathbf{x} - \mathbf{x}'|} H_1^{(1)}(k_b|\mathbf{x} - \mathbf{x}'|) - k_b^2 H_0^{(1)}(k_b|\mathbf{x} - \mathbf{x}'|) \right] \mathbb{Q}(\mathbf{x} - \mathbf{x}') \\ &+ \frac{i}{4} \left[k_b^2 H_0^{(1)}(k_b|\mathbf{x} - \mathbf{x}'|) - k_b \frac{1}{|\mathbf{x} - \mathbf{x}'|} H_1^{(1)}(k_b|\mathbf{x} - \mathbf{x}'|) \right] \mathbb{I}_2. \end{aligned} \tag{3.74}$$

To compute (3.73), we transform the singular cell D_s from rectangular shape to a circular disk having center at $\mathbf{x}$ and radius a_n, as shown in Figure 3.3. The transformed

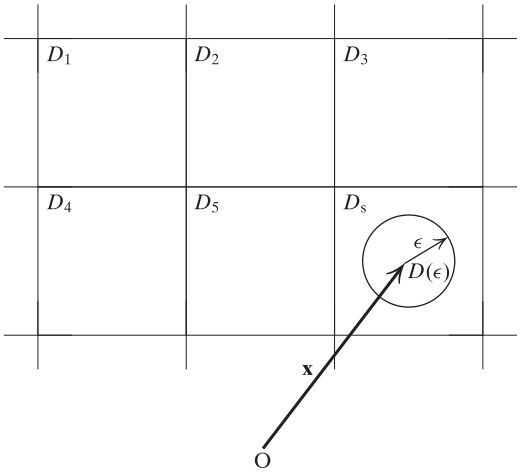

Figure 3.2 Some elementary cells of the domain D and the singular cell.

circular cell with the original one have the same surface $S = \pi a_n^2 = h^2$. Then, (3.74) is transformed into polar coordinates as

$$\begin{aligned}\mathbb{A}(\rho,\phi) =& \frac{i}{4}\left[2k_\mathrm{b}\frac{1}{\rho}H_1^{(1)}(k_\mathrm{b}\rho) - k_\mathrm{b}^2 H_0^{(1)}(k_\mathrm{b}\rho)\right]\mathbb{Q}(\phi) \\ &+ \frac{i}{4}\left[k_\mathrm{b}^2 H_0^{(1)}(k_\mathrm{b}\rho) - k_\mathrm{b}\frac{1}{\rho}H_1^{(1)}(k_\mathrm{b}\rho)\right]\mathbb{I}_2. \end{aligned} \tag{3.75}$$

Now, we calculate each element of the tensor $\mathbb{L}$ from

$$\begin{aligned}L_{pq} = \frac{i}{4}\Bigg\{ &\lim_{\epsilon\to 0}\int_{\rho=\epsilon}^{a_n}\left[2k_\mathrm{b}\frac{1}{\rho}H_1^{(1)}(k_\mathrm{b}\rho) - k_\mathrm{b}^2 H_0^{(1)}(k_\mathrm{b}\rho)\right]\rho\mathrm{d}\rho\int_{\phi=0}^{2\pi} Q_{pq}(\phi)\mathrm{d}\phi \\ &+ \int_{\rho=\epsilon}^{a_n}\left[k_\mathrm{b}^2 H_0^{(1)}(k_\mathrm{b}\rho) - k_\mathrm{b}\frac{1}{\rho}H_1^{(1)}(k_\mathrm{b}\rho)\right]\rho\mathrm{d}\rho\int_{\phi=0}^{2\pi}\delta_{pq}\mathrm{d}\phi\Bigg\}, \quad p,q = 1,2. \end{aligned} \tag{3.76}$$

The integration in (3.76) is carried out only three times since the tensor $\mathbb{Q}(\mathbf{x}-\mathbf{x}')$ is symmetric. The angular integral with integrand function $Q_{pq}(\phi)$ gives $\pi\delta_{pq}$, while the other with the integrand function δ_{pq} gives $2\pi\delta_{pq}$. The radial integrals are evaluated

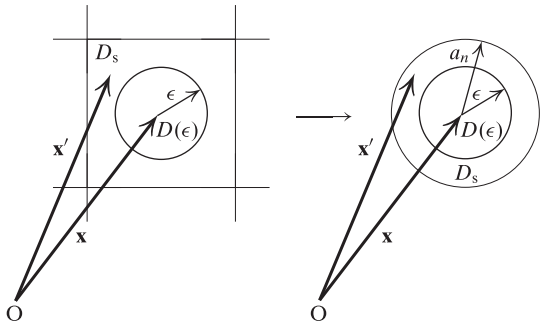

Figure 3.3 Transformation of the singular cell.

by changing variables $z = k_\mathrm{b}\rho$ and using the well known integral formulas for Bessel functions [75]

$$\int^z \zeta^{n+1} Z_n(\zeta)\mathrm{d}\zeta = z^{n+1} Z_{n+1}(z), \tag{3.77}$$

$$\int^z \zeta^{-n+1} Z_n(\zeta)\mathrm{d}\zeta = -z^{-n+1} Z_{n-1}(z). \tag{3.78}$$

The result is

$$L_{pq} = \delta_{pq}\left\{\frac{i\pi a_n}{4} k_\mathrm{b} H_1^{(1)}(k_\mathrm{b} a_n) - \frac{i\pi}{4} k_\mathrm{b} \lim_{\epsilon\to 0}\left[\epsilon H_1^{(1)}(k_\mathrm{b}\epsilon)\right]\right\}, \quad p,q = 1,2. \tag{3.79}$$

The involved limit is easily calculated giving $-i2/(\pi k_\mathrm{b})$. Therefore, substituting the results for the tensor $\mathbb{L}$ back to (3.72)

$$\begin{aligned} &p.v.\int_{\mathbf{x}'\in D_\mathrm{s}} \mathbb{A}(\mathbf{x}-\mathbf{x}')\chi_\mathrm{e}(\mathbf{x}')\mathbf{E}(\mathbf{x}')\mathrm{d}\mathbf{x}' \\ &\approx \chi_\mathrm{e}(\mathbf{x}_n)\left[-\frac{1}{2} + \frac{i\pi a_n}{4} k_\mathrm{b} H_1^{(1)}(k_\mathrm{b} a_n)\right]\mathbf{E}(\mathbf{x}_n), \quad \mathbf{x}_n \in D_\mathrm{s}. \end{aligned} \tag{3.80}$$

What remains is the calculation of the integral involving $\mathbb{B}(\mathbf{x}-\mathbf{x}')$ over the domain D_s in (3.69). We have

$$\begin{aligned} &-i\omega\mu_\mathrm{b}\int_{\mathbf{x}'\in D_\mathrm{s}} \mathbb{B}(\mathbf{x}-\mathbf{x}')\chi_\mathrm{m}(\mathbf{x}')\mathbf{H}(\mathbf{x}')\mathrm{d}\mathbf{x}' \\ &\approx -i\omega\mu_\mathrm{b}\chi_\mathrm{m}(\mathbf{x}_n)\mathbf{H}(\mathbf{x}_n)\int_{\mathbf{x}'\in D_\mathrm{s}} \mathbb{B}(\mathbf{x}-\mathbf{x}')\mathrm{d}\mathbf{x}'. \end{aligned} \tag{3.81}$$

Utilizing (3.65) and (3.66), and expressing $\mathbb{B}(\mathbf{x}-\mathbf{x}')$ in polar coordinates, in the same manner as we did for the $\mathbb{L}$ tensor before, we arrive at an angular integration from 0 to 2π for each component Θ_{pq} of $\boldsymbol{\Theta}(\mathbf{x}-\mathbf{x}')\times$ tensor. This integration is zero for each component Θ_{pq}, and hence this term does not contribute.

Putting together the results of (3.70), (3.71) and (3.80), we arrive at the linear equations

$$\begin{aligned} &\left\{1 - \chi_\mathrm{e}(\mathbf{x}_n)\left[\frac{i\pi k_\mathrm{b} h}{4\sqrt{\pi}} H_1^{(1)}(k_\mathrm{b} h/\sqrt{\pi}) - 1\right]\right\}\mathbf{E}(\mathbf{x}_n) \\ &- h^2 \sum_{\substack{m=1 \\ m\neq n}}^{N-1} \mathbb{A}(\mathbf{x}_n - \mathbf{x}_m)\chi_\mathrm{e}(\mathbf{x}_m)\mathbf{E}(\mathbf{x}_m) \\ &- i\omega\mu_\mathrm{b} h^2 \sum_{\substack{m=1 \\ m\neq n}}^{N-1} \mathbb{B}(\mathbf{x}_n - \mathbf{x}_m)\chi_\mathrm{m}(\mathbf{x}_m)\mathbf{H}(\mathbf{x}_m) = \mathbf{E}^\mathrm{in}(\mathbf{x}_n), \quad n = 1,2,\ldots,N. \end{aligned} \tag{3.82}$$

Following the same analysis for (3.68), we arrive at the linear equations

$$\left\{1-\chi_{\rm m}(\mathbf{x}_n)\left[\frac{i\pi k_{\rm b}h}{2\sqrt{\pi}}H_1^{(1)}(k_{\rm b}h/\sqrt{\pi})-1\right]\right\}\mathbf{H}(\mathbf{x}_n)$$
$$-h^2\sum_{\substack{m=1\\m\neq n}}^{N-1}\mathbb{A}(\mathbf{x}_n-\mathbf{x}_m)\chi_{\rm m}(\mathbf{x}_m)\mathbf{H}(\mathbf{x}_m)$$
$$+i\omega\varepsilon_{\rm b}h^2\sum_{\substack{m=1\\m\neq n}}^{N-1}\mathbb{B}(\mathbf{x}_n-\mathbf{x}_m)\chi_{\rm e}(\mathbf{x}_m)\mathbf{E}(\mathbf{x}_m)=\mathbf{H}^{\rm in}(\mathbf{x}_n),\quad n=1,2,\ldots,N. \tag{3.83}$$

Equations (3.82) and (3.83) compose the linear system for the evaluation of the unknown total fields, with the following structure

$$\begin{bmatrix}A_{11}&A_{12}&A_{13}\\A_{21}&A_{22}&A_{23}\\A_{31}&A_{32}&A_{33}\end{bmatrix}\begin{bmatrix}u_1\\u_2\\u_3\end{bmatrix}=\begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix}. \tag{3.84}$$

In (3.84), $[u_1,u_2,u_3]^T=[E_1(\mathbf{x}_n),E_2(\mathbf{x}_n),H_3(\mathbf{x}_n)]^T$, $n=1,2,\ldots,N$, are the unknown total field components on the grid, while $[b_1,b_2,b_3]^T=[E_1^{\rm in}(\mathbf{x}_n),E_2^{\rm in}(\mathbf{x}_n),H_3^{\rm in}(\mathbf{x}_n)]^T$ contains the grid values of the incident field. For example, the TE plane wave impinging to the object at an angle ψ with respect to x_1-axis is represented by: $[b_1]_n=\exp[ik_{\rm b}(x_{1,n}\cos\psi+x_{2,n}\sin\psi)]$, $[b_2]_n=-[b_1]_nk_{\rm b}\sin\psi/(\omega\varepsilon_{\rm b})$, and $[b_3]_n=[b_1]_nk_{\rm b}\cos\psi/(\omega\varepsilon_{\rm b})$, $n=1,\ldots,N$. The elements of the system matrix in (3.84) are now easily recognized with the use of (3.82), (3.83), (3.64), (3.65) and (3.14). The results are

$$[A_{\ell\ell}]_{nm}=-k_{\rm b}^2h^2\chi_{\rm e}(\mathbf{x}_m)\left\{\left[\frac{i}{2k_{\rm b}r_{nm}}H_1^{(1)}(k_{\rm b}r_{nm})-\frac{i}{4}H_0^{(1)}(k_{\rm b}r_{nm})\right]\theta_{\ell,nm}\theta_{\ell,nm}\right.$$
$$\left.+\left[\frac{i}{4}H_0^{(1)}(k_{\rm b}r_{nm})-\frac{i}{4k_{\rm b}r_{nm}}H_1^{(1)}(k_{\rm b}r_{nm})\right]\delta_{\ell\ell}\right\},\quad \ell=1,2,\quad m\neq n; \tag{3.85}$$

$$[A_{\ell\ell}]_{nn}=1+\left[1-\frac{i\pi k_{\rm b}h}{4\sqrt{\pi}}H_1^{(1)}(k_{\rm b}h/\sqrt{\pi})\right]\chi_{\rm e}(\mathbf{x}_n),\quad \ell=1,2; \tag{3.86}$$

$$[A_{\ell q}]_{nm}=-k_{\rm b}^2h^2\chi_{\rm e}(\mathbf{x}_m)\left[\frac{i}{2k_{\rm b}r_{nm}}H_1^{(1)}(k_{\rm b}r_{nm})-\frac{i}{4}H_0^{(1)}(k_{\rm b}r_{nm})\right]$$
$$\times\theta_{\ell,nm}\theta_{q,nm}(1-\delta_{nm}),\quad \ell,q=1,2,\quad \ell\neq q; \tag{3.87}$$

$$[A_{13}]_{nm}=i\omega\mu_{\rm b}h^2\chi_{\rm m}(\mathbf{x}_m)\frac{ik_{\rm b}}{4}H_1^{(1)}(k_{\rm b}r_{nm})\theta_{2,nm}(1-\delta_{nm}); \tag{3.88}$$

$$[A_{32}]_{nm}=-i\omega\varepsilon_{\rm b}h^2\chi_{\rm e}(\mathbf{x}_m)\frac{ik_{\rm b}}{4}H_1^{(1)}(k_{\rm b}r_{nm})\theta_{1,nm}(1-\delta_{nm}); \tag{3.89}$$

$$[A_{31}]_{nm}=i\omega\varepsilon_{\rm b}h^2\chi_{\rm e}(\mathbf{x}_m)\frac{ik_{\rm b}}{4}H_1^{(1)}(k_{\rm b}r_{nm})\theta_{2,nm}(1-\delta_{nm}); \tag{3.90}$$

$$[A_{23}]_{nm}=-i\omega\mu_{\rm b}h^2\chi_{\rm m}(\mathbf{x}_m)\frac{ik_{\rm b}}{4}H_1^{(1)}(k_{\rm b}r_{nm})\theta_{1,nm}(1-\delta_{nm}); \tag{3.91}$$

$$[A_{33}]_{nm}=-k_{\rm b}^2h^2\chi_{\rm m}(\mathbf{x}_m)\frac{i}{4}H_0^{(1)}(k_{\rm b}r_{nm}),\quad m\neq n; \tag{3.92}$$

$$[A_{33}]_{nn} = 1 + \left[1 - \frac{i\pi k_b h}{2\sqrt{\pi}} H_1^{(1)}(k_b h/\sqrt{\pi})\right] \chi_m(\mathbf{x}_n). \tag{3.93}$$

In (3.85)–(3.93) we have defined $r_{nm} = |\mathbf{x}_n - \mathbf{x}_m|$, $\theta_{\ell,nm} = (x_{\ell,n} - x_{\ell,m})/r_{nm}$, $\ell = 1, 2$, $n, m = 1, \ldots, N$, and $x_{\ell,n}$ denotes the Cartesian component of the 2D position vector $\mathbf{x}_n$, pointing at the nth node of the grid.

3.5 GMRES convergence

Both the system matrix and the original integral operator are neither symmetric nor normal. Therefore, the range of iterative methods applicable to the problem is extremely limited, with (restarted) GMRES typically showing the best performance. Figure 3.4 shows the GMRES convergence histories on the original system $Au = b$ for three physically distinct scatterers (all with $\chi_m = 0$) and two different polarizations, corresponding to the TE and TM cases. The three scatterers are: a lossless object with large positive real permittivity, $\varepsilon/\varepsilon_b = 16$, an object with small losses and a large negative real part of the permittivity, $\varepsilon/\varepsilon_b = -16 + i1.5$, and an inhomogeneous object consisting of two concentric layers with the inner part having large losses, $\varepsilon/\varepsilon_b = 2.5 + i20$, and the outer layer having a large positive real permittivity, $\varepsilon/\varepsilon_b = 16$. All objects are cylinders with square cross-sections with the side length $a = \lambda_b$. The inner core of the inhomogeneous object has also a square cross-section with the side length $a/2$. The RESTART parameter of the GMRES algorithm is equal to 40 (we shall study the influence of this parameter in more detail later). In Figure 3.4 (left) we illuminate the scatterers with a plane wave whose electric field vector is orthogonal to the axis of the cylinder (TE), whereas in Figure 3.4(right) the electric field vector of the incident plane wave is parallel to the axis of the cylinder (TM). Since all three objects do not exhibit any magnetic contrast with respect to the background medium, the integral operators are different for these TE and TM cases – see (3.15), (3.16). The latter contains now only the weakly singular part and is, therefore, of the form

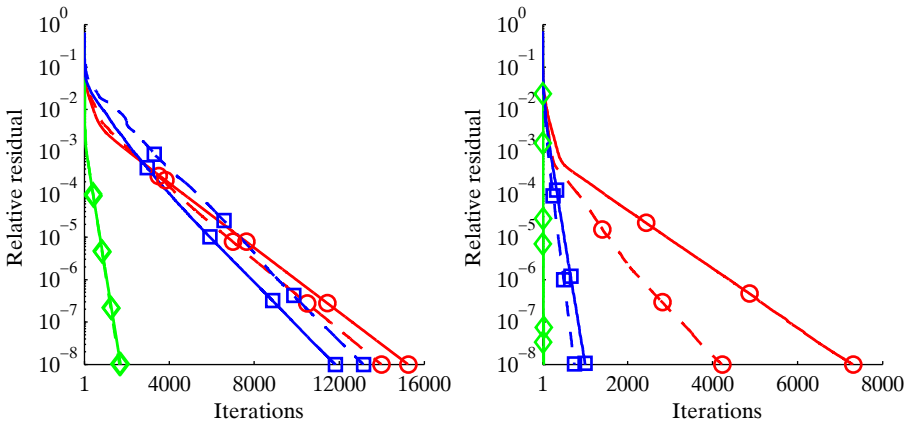

Figure 3.4 (color online). Convergence of restarted GMRES applied to the original system with (dashed) and without (solid) deflation of the largest eigenvalues for three physically distinct scatterers. Deflation accelerates GMRES in the TM case (right), but does not work in the TE case (left). Lossless object – red lines (with circles), object with small losses and negative real part of permittivity – green lines (with diamonds), and inhomogeneous object – blue lines (with squares).

"identity plus compact".

Next to the convergence plots of the original system in Figure 3.4 we present the convergence histories of the restarted GMRES with a deflation-based right preconditioner (dashed lines). As was suggested in [62], deflating the largest eigenvalues of A in the TM case may significantly accelerate the convergence of the full (un-restarted) GMRES. The preconditioner deflating r largest eigenvalues of A is constructed by first running the EIGS algorithm to retrieve these r eigenvalues and the corresponding eigenvectors. With the help of the modified Gram-Schmidt algorithm we further build an orthonormal basis for the retrieved set of eigenvectors. Let this basis be stored in a $3N \times r$ matrix V. Then, the right preconditioner [21] has the form

$$P^{-1} = I_{3N} + V\left[T^{-1} - I_r\right]V^*, \tag{3.94}$$

with

$$T = V^*AV. \tag{3.95}$$

As was observed in [62] for the TM case, although the system matrix is not normal, the matrix V has full column rank, and the matrix T has a decent condition number. We confirm this behavior in the present TE case. The term "deflation" refers to the fact that the spectrum of AP^{-1} looks like the spectrum of A with r largest eigenvalues moved to the point $1 + i0$ of the complex plane, and the corresponding eigen-subspace has been projected out. Here, even with the restarted GMRES, we see some improvement in the TM case Figure 3.4(right). The TE case, however, is not affected by deflation (it may even become worse).

Table 3.1 and Table 3.2 summarize the numerical results corresponding to the test objects in the TM and TE cases. The $3N \times r$ matrix V needs to be stored in the memory. For the sake of fair comparison, we divide our memory between the inner Krylov subspace of the restarted GMRES and V, i.e., if we set RESTART $= k$, $k > r$ with the original matrix A, then we use RESTART $= k - r$ with the preconditioned matrix AP^{-1}.

One can understand the peculiar lack of progress in the TE case by considering Figure 3.5, which shows the spectra of the system matrix for the TE and TM cases corresponding to the numerical experiments of Figure 3.4. As can be seen, the difference between the spectra is in the presence of dense line segments $[1, \varepsilon(\mathbf{x})/\varepsilon_b]$, $\mathbf{x} \in \mathbb{R}^2$ and some extra isolated eigenvalues inside a distribution remotely resembling a shifted circumference (the latter is much more pronounced with homogeneous objects at higher frequencies, see [62]). The dense line segments in the TE spectrum contain a very large number of closely spaced eigenvalues, and the deflation process simply stagnates when it reaches the outer ends of these lines, corresponding to $\varepsilon(\mathbf{x})/\varepsilon_b$, $\mathbf{x} \in D$. Indeed, in the TE case we have seen no change in the convergence with AP^{-1} when we tried to deflate any eigenvalues beyond those that were present outside a circle with the center at zero and the radius reaching the largest value of $|\varepsilon(\mathbf{x})/\varepsilon_b|$, $\mathbf{x} \in \mathbb{R}^2$ for each particular scattering configuration. At the same time in the TM case we observed consistent improvement in convergence when deflating all eigenvalues outside the unit circle, see Table 3.1.

3.6 Preconditioning

In the three-dimensional case it was conjectured [5] that the dense line segments, which we now observe in the TE spectrum as well, have something to do with the essential spectrum, $\lambda_{ess} = \varepsilon(\mathbf{x})/\varepsilon_b$, of the corresponding singular integral operator. Obviously, there is no such thing as essential spectrum with matrices, and its role and

Table 3.1 Restarted GMRES with fair memory usage for three different test scatterers, TM case. Original and deflated systems.

System	$\varepsilon/\varepsilon_b$	r, deflated eigenvalues	RESTART	CPU time, seconds	Iterations, tol. 10^{-8}	Speed–up
A	16	0	40	168	7311	1.0
AP^{-1}	16	28	12	78	4230	2.2
A	$-16 + i1.5$	0	40	0.7	27	1.0
AP^{-1}	$-16 + i1.5$	28	12	0.5	14	1.4
A	16 and $2.5 + i20$	0	40	23	1000	1.0
AP^{-1}	16 and $2.5 + i20$	28	12	14	749	1.6

Table 3.2 Restarted GMRES with fair memory usage for three different test scatterers, TE case. Original, deflated, regularized, and regularized-plus-deflated systems.

System	$\varepsilon/\varepsilon_b$	r, deflated eigenvalues	RESTART	CPU time, seconds	Iterations, tol. 10^{-8}	Speed–up
A	16	0	40	350	15258	1.0
AP^{-1}	16	6	34	306	14002	1.1
$A_R A$	16	0	40	142	4030	2.5
$A_R AP^{-1}$	16	14	26	117	3626	3
A	$-16 + i1.5$	0	40	39	1708	1.0
AP^{-1}	$-16 + i1.5$	7	33	36	1659	1.1
$A_R A$	$-16 + i1.5$	0	40	36	1021	1.1
$A_R AP^{-1}$	$-16 + i1.5$	30	10	16	534	2.4
A	16 and $2.5 + i20$	0	40	275	11823	1.0
AP^{-1}	16 and $2.5 + i20$	3	37	310	13156	0.9
$A_R A$	16 and $2.5 + i20$	0	40	107	3032	2.6
$A_R AP^{-1}$	16 and $2.5 + i20$	14	26	94	2911	3

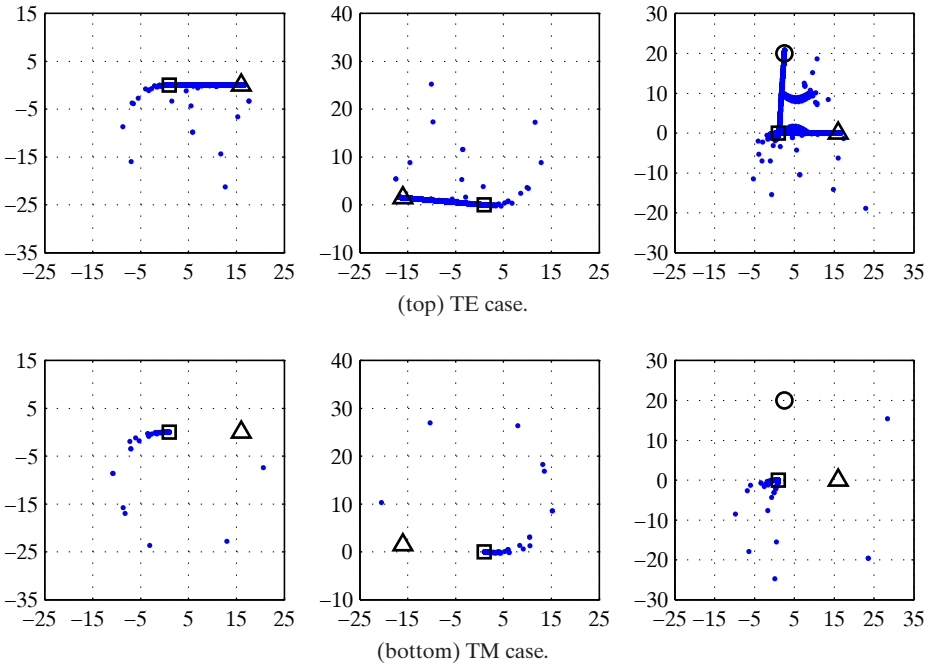

Figure 3.5 (color online). Eigenvalues of the system matrix A for three physically distinct scatterers in the TE (top) and TM (bottom) cases. Horizontal axis – real part, vertical axis – imaginary part. Also shown are the values of the relative permittivity of the object (triangles and circles), and the background medium (squares, always at one).

transformation in the process of discretization is so far unclear. Moreover, recently it was shown [18] that in the case of a discontinuous medium the essential spectrum will contain only the end and the middle points of the lines. Yet, it is also known [6] that the convex hull of the essential spectrum contains all other (discrete) eigenvalues of the problem in the low-frequency regime ($\omega \to 0$). Since the discrete eigenvalues of the operator correspond to "proper" eigenfunctions they do appear in the matrix spectrum upon discretization and should thus be accumulated inside the convex hull of the essential spectrum.

As the process of deflation is obviously hindered by the dense spectral segments one needs to reduce their extent prior to deflation. Numerical experiments and the above considerations suggest that we need to reduce the convex hull of the essential spectrum. On the operator level such a reduction can be achieved by regularization. We shall, therefore, apply the discretized regularizer A_{R}, i.e., a discrete version of the operator $A(\chi'_{\mathrm{e}})$, so that the problem becomes $A_{\mathrm{R}}Au = A_{\mathrm{R}}b$. Notice that no additional storage is required, since the original Green's tensor can be re-used. Matrix-vector products, however, do become twice more expensive to compute.

To see how this procedure affects the matrix spectrum we compute the eigenvalues of $A_{\mathrm{R}}A$ corresponding to the matrices analyzed in Figure 3.5 (top). The results are presented in Figure 3.6, where we see that the line segments have been substantially reduced. For a homogeneous object we can see what is going on also on the matrix level. The general form of our system matrix is $A = I - GX$, where I is a $3N \times 3N$ identity matrix, X is the diagonal matrix, containing the grid values

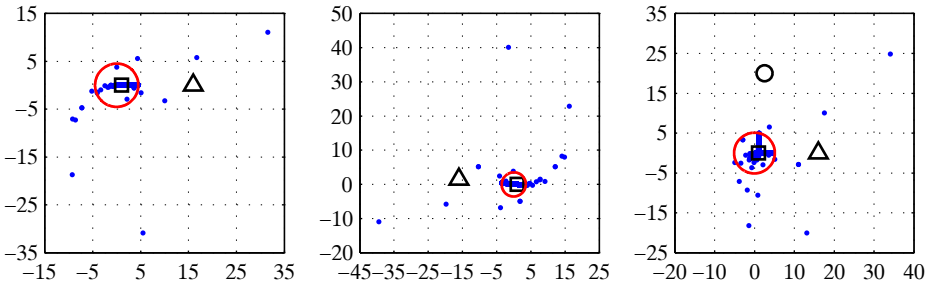

Figure 3.6 (color online). Eigenvalues of the regularized matrix $A_{\mathrm{R}}A$, corresponding to the spectra of A in the TE case, see Figure 3.5 (top). Large red circles show the estimated outer bounds of the dense cluster – deflation bounds. Also shown are the values of the relative permittivity of the object (triangles and circles), and the background medium (squares, always at one).

of the contrast function, and G is a dense matrix produced by the integral operator with the Green's tensor kernel. When the contrast is homogeneous and equal to a scalar complex value $\chi = \varepsilon/\varepsilon_{\mathrm{b}} - 1$ inside the scatterer, we can write our matrix as $A = I - \chi G$. Now, let ψ_n be an eigenvector of G, i.e. $G\psi_n = \beta_n \psi_n$. Then, the eigenvalues of A are:

$$\lambda_n = 1 - \beta_n \left(\frac{\varepsilon}{\varepsilon_{\mathrm{b}}} - 1 \right), \tag{3.96}$$

and the eigenvalues of $A_{\mathrm{R}}A$ are

$$\lambda'_n = \left[1 - \beta_n \left(\frac{\varepsilon_{\mathrm{b}}}{\varepsilon} - 1 \right) \right] \lambda_n. \tag{3.97}$$

Eliminating β_n we get

$$\lambda'_n = \left[1 + \frac{\varepsilon_{\mathrm{b}}}{\varepsilon} (1 - \lambda_n) \right] \lambda_n. \tag{3.98}$$

We have verified numerically that this transformation holds for the first two plots of Figure 3.5 (top) and Figure 3.6. Curiously, the seemingly quadratic amplification of the large eigenvalues in (3.98) is moderated by the inverse of the relative permittivity. That is why the regularized spectra are not as extended as one would fear. Unfortunately, such a simple proof is not possible with an inhomogeneous object, yet the effect of the regularization on the matrix spectrum in Figure 3.5 (top, right) and Figure 3.6 (right) is obviously similar to the homogeneous cases.

Table 3.2 shows that such a regularized system is already better than the original one in terms of convergence. The relative speed-up, however, is limited to approximately 2.5 times due to the extra work involved in computing the matrix-vector products with $A_{\mathrm{R}}A$. Moreover, no convergence improvement is observed with the negative permittivity object.

To further accelerate the convergence and make the method work with the negative permittivity objects as well (they are important in surface plasmon-polariton studies) we can now apply the deflation technique. For that we need to decide on the number of eigenvalues to be deflated. Obviously, it does not make sense to deflate beyond the first (from outside) dense cluster – deflation will only cost more time and memory, while the improvement will stagnate. After the regularization, the outer boundaries of the dense line segments have been all shifted inwards. However, due to nonlinear nature of the mapping (3.98) it is hard to tell exactly what an image

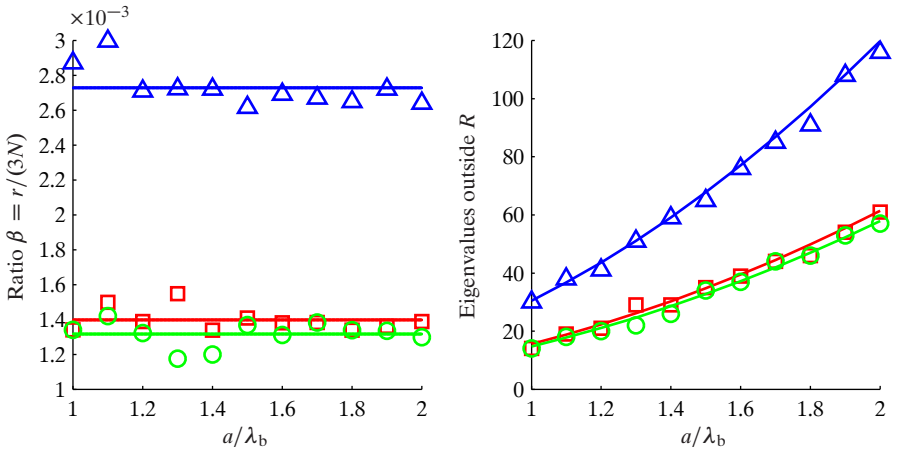

Figure 3.7 (color online). Dependence of the number of eigenvalues to be deflated on the wavelength for three test scatterers (same medium parameters as in previous figures). Left: ratio β of the number r of eigenvalues outside the deflation bound (circles with radius R, see Figure 3.6) to the total number $3N$ of eigenvalues for a classical discretization rule (fixed number of points per smallest medium wavelength). Right: actual number r of eigenvalues to be deflated (with magnitude greater than the deflation bound R) and our analytical prediction based on an average β coefficient estimated from the left plot.

of an arbitrary line segment would look like. Only for a lossless homogeneous scatterer are we able to predict that a real line segment $[1, \varepsilon/\varepsilon_b]$ will be mapped onto the real line segment $[1, 1 + (1 - \varepsilon/\varepsilon_b)^2/(4\varepsilon/\varepsilon_b)]$, which contracts the original segment by approximately four times. For a general line segment radially emerging from the point $1 + i0$ we have made a small numerical routine, which produces a discretized image according to (3.98) and finds the point with the largest absolute value – the radius of the circle beyond which all eigenvalues should be deflated. We have observed, by analyzing the spectra for various inhomogeneous objects, that the general regularization seems to transform each line segment according to (3.98). Therefore, in an inhomogeneous case, we can simply apply the aforementioned numerical mapping to each of the segments $[1, \varepsilon(\mathbf{x})/\varepsilon_b]$, $\mathbf{x} \in D$, and choose the point with the maximum absolute value as the deflation radius. The deflation bounds obtained in this way are shown in Figure 3.6 as solid red circles.

The number of eigenvalues outside the deflation bound depends on the physics of the problem. For example, we know that at low frequencies (large wavelengths) all discrete eigenvalues are located within the convex hull of the essential spectrum [6], and that at higher frequencies they tend to spread out. We have investigated this phenomenon in more detail and have found a pattern, similar to the one reported in [62]. As one can see from Figure 3.7 (left), the ratio of the number r of eigenvalues outside the deflation bound to the total number of eigenvalues $3N$ is roughly constant (for objects larger than λ_b), if at each wavelength we adjust the spatial discretization according to the classical rule (say, 15 points per smallest medium wavelength). Hence, with fixed medium parameters and a square computational domain with side a, the total number of unknowns and therefore eigenvalues grows as $3N = 3[\alpha^2(a/\lambda_b)^2 + 2\alpha(a/\lambda_b) + 1]$ with $\alpha = 15 \max\{\sqrt{(\varepsilon'/\varepsilon_b)}\}$. Thus, knowing $\beta = r/(3N)$ for some wavelength, we can estimate $r = 3N\beta$ at any other wavelength. This

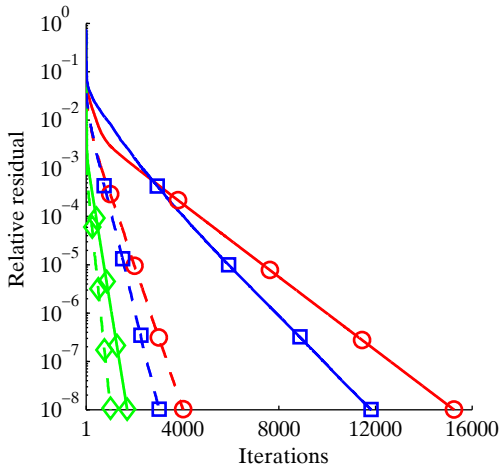

Figure 3.8 (color online). Convergence of restarted GMRES with the original $Au = b$ (solid) and preconditioned $A_{\mathrm{R}}AP^{-1}v = A_{\mathrm{R}}b$ (dashed) systems for the three physically distinct scatterers in the TE case. Red lines (with circles); $\varepsilon/\varepsilon_{\mathrm{b}} = 16$; Green lines (with diamonds): $\varepsilon/\varepsilon_{\mathrm{b}} = -16 + i1.5$; Blue lines (with squares): inhomogeneous object $\varepsilon/\varepsilon_{\mathrm{b}} = (16, 2.5 + i20)$.

empirical law comes very handy since solving a relatively low-frequency spectral problem can be really easy due to small system matrix dimensions. Otherwise, if there is no shortage of memory for the chosen wavelength, one can simply run the EIGS routine increasing the number of recovered largest eigenvalues until the deflation bound obtained above is reached.

Having a good estimate of the deflation parameter r we have applied the right preconditioner given by (3.94) to the regularized system for the same three objects as in all previous examples. The largest eigenvalues/eigenvectors of the regularized system turn out to be very stable, so that one can safely apply the EIGS routine with its tolerance set as high as 10^{-2}. Though, we have not seen any substantial acceleration beyond the 10^{-4} tolerance, which, in its turn, gave us a significant speed-up in the off-line work compared to the default machine precision tolerance. In all our experiments, the off-line time remained a small fraction of the total CPU time (for concrete numbers we refer to the example at the end of this Section).

The convergence histories for a restarted GMRES with this preconditioner are presented in Figure 3.8. This figure, and the iteration counts in general, are slightly misleading when one deals with an iterative algorithm like restarted GMRES. For example, we see in Table 3.2 that, apart from the negative permittivity case, the iteration counts after deflation are only slightly better than what we have already achieved with the regularized system. Hence, one could conclude that the deflation is not worth the effort. However, the GMRES algorithm does not scale linearly in time as a function of the RESTART parameter, i.e., while it may converge in less iterations for larger values of RESTART, the CPU time may not decrease as much. This has partly to do with the re-orthogonalization process, which takes more time for larger inner Krylov subspaces, and partly with the very bad spectral properties of our system matrix.

Thus, to decide upon the benefits of deflation we compare execution times for various values of RESTART parameter. In addition, we need to consider another

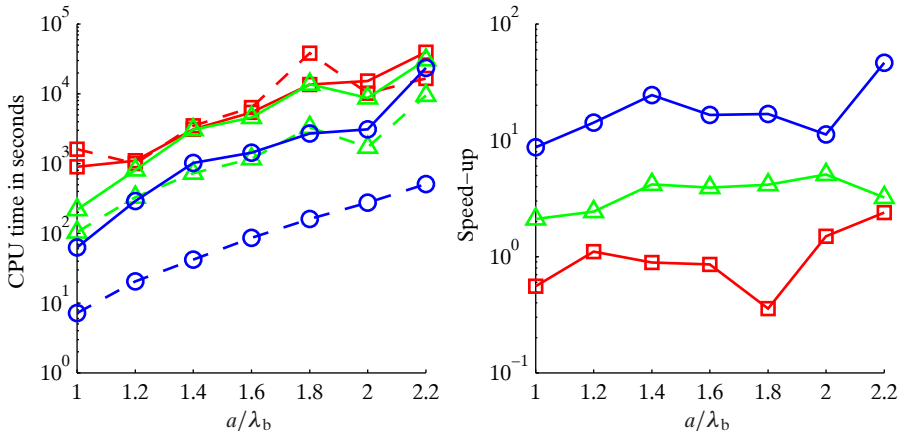

Figure 3.9 (color online). CPU time (left) and the achieved speed-up (right) as a function of wavelength (object size) for RESTART $= r/2$ (red, squares), RESTART $= 2r$ (green, triangles), and RESTART $= 10r$ (blue, circles). The permittivity of the scatterer is $\varepsilon/\varepsilon_b = 16$. Solid lines – original system $Au = b$; dashed lines – preconditioned system $A_R A P^{-1} v = A_R b$.

important constraint – the available memory. Since a fair comparison can only be achieved if we give the same amount of memory to both the original and the preconditioned systems, we give all the available memory to the GMRES algorithm in the former case, and split it between the RESTART and the deflation subspace of size r in the latter. Finally, we want to know if the efficiency of the proposed preconditioner depends on the physics of the problem, i.e., does it work better or worth for larger objects and contrasts? Of course, the system becomes larger and more difficult to solve for larger objects, and we have to be prepared to increase the value of RESTART for larger objects and contrasts to achieve any convergence at all. Hence, assuming that we have enough memory at least for deflation, we shall distinguish between the following three situations: limited memory, RESTART$(A_R A P^{-1}) < r$; moderate memory, RESTART$(A_R A P^{-1}) \sim r$; large memory, RESTART$(A_R A P^{-1}) > r$.

We observe from the plots of Figure 3.9 that with limited memory, where we choose RESTART$(A_R A P^{-1}) = r/2$ and RESTART$(A) = r$, the original system initially outperforms the preconditioned one. As the object size grows, so does the size r of the deflation subspace. Since in that case the RESTART$(A_R A P^{-1})$ parameter grows as well, at some point (for $a/\lambda = 1.9$ with this particular object) it reaches a value at which the preconditioned system breaks even and begins to outperform the original one for larger objects (maximum achieved speed-up is 2.4 times at $a/\lambda_b = 2.2$). With moderate memory, RESTART$(A_R A P^{-1}) = 2r$, RESTART$(A) = 3r$, the preconditioned system always performs better than the original one. Moreover, the relative speed-up Figure 3.9 (right) generally increases with the object size (maximum achieved speed-up is 5.1 times at $a/\lambda_b = 2$, minimum – 2.1 times at $a/\lambda_b = 1$). With large memory, i.e., for RESTART$(A_R A P^{-1}) = 10r$ and RESTART$(A) = 11r$, the speed-up may be as high as 46.6 times ($a/\lambda_b = 2.2$) (minimum speed-up of 8.7 times at $a/\lambda_b = 1$), and it also grows (on average) with the object size.

With objects of finite extent changing the wavelength of the incident field is not the same as changing the object permittivity. Hence, we repeat the same numerical experiments as in Figure 3.10, but now for an object of fixed size, $a/\lambda_b = 1$, and

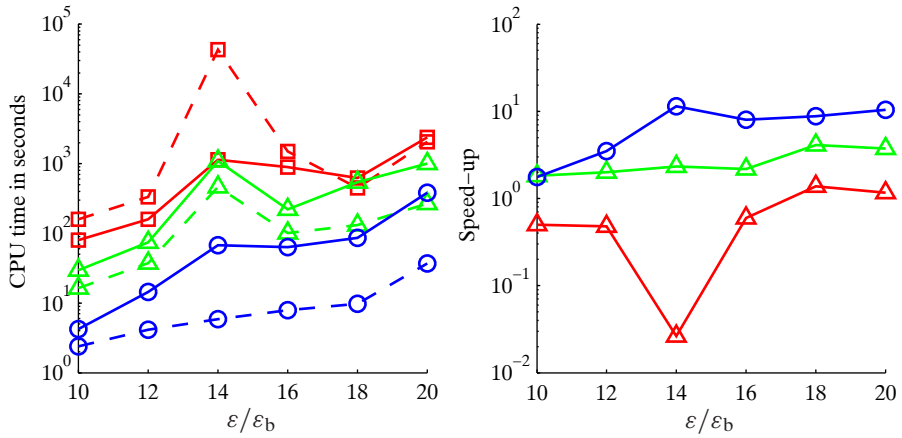

Figure 3.10 (color online). CPU time (left) and the achieved speed-up (right) as a function of permittivity for RESTART $= r/2$ (red, squares), RESTART $= 2r$ (green, triangles), and RESTART $= 10r$ (blue, circles). The scatterer is lossless with side length fixed at $a/\lambda_b = 1$. Solid lines – original system $Au = b$; dashed lines – preconditioned system $A_R AP^{-1}v = A_R b$.

varying permittivity. The deflation parameter r is estimated using the same algorithm as above and turns out to grow only slightly as a function of permittivity, climbing from 13 to 17 for the permittivity changing between 10 and 20. At the same time the total number of unknowns as well as eigenvalues grows very fast. Therefore, deflation did not accelerate the iterative process here as much as when we changed the wavelength, and most of the speed-up must be due to the regularization. For this relatively small scatterer we see some improvement in convergence only with relatively large amounts of memory devoted to the RESTART parameter, i.e., RESTART$(A_R AP^{-1}) \geq 2r$. Moreover, preconditioned GMRES with RESTART$(A_R AP^{-1}) = r/2$ did not converge for $\varepsilon/\varepsilon_b = 14$ at all (we set this point arbitrarily high in the plot). On the other hand, the maximum speed-up for RESTART $= 2r$ was 4.1 times (minimum 1.8 times), while for RESTART $= 10r$ the maximum speed-up was 11.4 times (minimum 1.6 times). From the previous experiments we may expect that the speed-up will be larger with larger objects.

As an example of a systematic application of the present preconditioner, we consider the problem, for which the DIE method is most suited, i.e., scattering on an object with continuously varying permittivity. The steps we suggest to follow are:

1. Determine the total available memory: M bytes
2. Fix the discretization rule at k points per smallest medium wavelength.
3. Estimate β from low-frequency matrix spectra.
4. Determine the maximal number of grid points $N_{\max}$:
   - If a – size of the computational domain D, and N – number of grid points, then $N = \alpha^2(a/\lambda_b)^2 + 2\alpha(a/\lambda_b) + 1$, where $\alpha = k \max\{\sqrt{(\varepsilon'(\mathbf{x})/\varepsilon_b)}\}, \mathbf{x} \in D$.
   - Memory needed for the system, right-hand side, and the unknown vector: $M_A = 720N$ bytes (one complex number – 16 bytes).
   - Memory required for deflation and GMRES: $M_{\text{Prec}} = 16r^2 + 48Nr + 48N\text{RESTART}$ bytes, where $r = 3N\beta$ and RESTART $= xr$, i.e., $M_{\text{Prec}} = 144(\beta^2 + \beta + x\beta)N^2$ bytes.

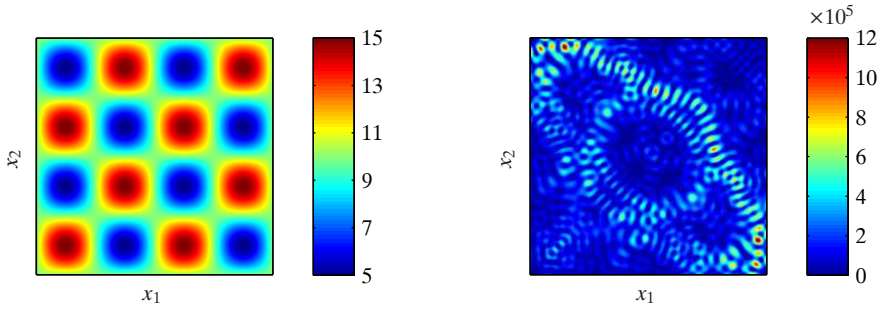

Figure 3.11 (color online). Scattering of a TE plane wave on an inhomogeneous object with continuously varying permittivity. Left – permittivity profile (side length $a/\lambda_b = 4$). Right – computed electric field intensity $|E_1|^2 + |E_2|^2$. Preconditioning gives 11.6 times speed-up for this problem.

- The largest affordable grid follows from $M_A + M_{\text{Prec}} \approx M$:
  $N_{\max} \approx [\sqrt{900 + M(\beta^2 + \beta + x\beta)} - 30]/[12(\beta^2 + \beta + x\beta)]$.

5. Determine the maximum affordable object size (smallest wavelength) via $[a/\lambda_b]_{\max} = (\sqrt{N_{\max}} - 1)/\alpha$.

The scatterer is depicted in Figure 3.11 (left). The relative permittivity function profile is given by

$$\varepsilon(\mathbf{x})/\varepsilon_b = 10 + 5\sin(4\pi x_1/a)\sin(4\pi x_2/a), \quad \mathbf{x} \in D, \tag{3.99}$$

where the origin of the coordinate system is in the middle of the computational domain, and $\varepsilon(\mathbf{x}) = \varepsilon_b$ for $\mathbf{x} \notin D$. The incident plane wave is impinging at an angle of 45 degrees with respect to the x_1-axis, so that both the object and the incident field are axially symmetric. We calculated β for this scatterer to be approximately 0.0011. Having $M = 4$ GB of memory available and choosing $k = 15$ and $\text{RESTART}(A_R A P^{-1}) = 8r$, from the above algorithm we determine the largest affordable size with this permittivity as $[a/\lambda_b]_{\max} = 4$. Hence, the total number of unknowns is $3N = 162867$ and the deflation parameter is $r = 180$. The GMRES was restarted at $\text{RESTART}(A_R A P^{-1}) = 1440$ and $\text{RESTART}(A) = 1620$. The results of this test are as follows: original system – 93020 seconds of CPU time, preconditioned system – 7999 seconds. The speed-up in pure "on-line" calculations is 11.6 times, as we would expect from our previous experiments with the homogeneous test object. The off-line work with the tolerance of EIGS set to 10^{-4} took only 745 seconds, so that even with this time included we get a 10.6 times speed-up. The computed electric field intensity, shown in Figure 3.11 (right), is axially symmetric as expected. The efficiency of the current preconditioner comes also in surface, not only when the object is electrically large, but also when we have to evaluate the fields for various incident angles, say for instance to compute the bistatic cross section for angles in the range from 0 to 2π. In some cases, changing just the incident angle, may lead the things to become even worse; so consider the same example as before with the only difference that the incident angle is now zero degrees. The results of this new test are as follows: original system – 162692 seconds/45.2 hours/1.9 days of CPU time, preconditioned system – 11620 seconds/3.2 hours. The speed-up in pure "on-line" calculations is 14 times.

The CPU timing is provided here merely for illustration and should not be used for direct comparison against the local methods, which was not the purpose

of this work either. For the particular problems considered above one could, for instance, substantially accelerate the calculations and free some valuable memory by computing the electric field only. On the other hand, the presented CPU timing data give us a good relative estimate of what to expect with objects having both electric and magnetic contrasts so that both electric and magnetic fields must indeed be computed at the same time.

Finally we should mention that, while we have been focusing here on the restarted GMRES, obviously, any other memory-efficient solver, which works with the original system $Au = b$, could also be applied to the preconditioned system $A_{\mathrm{R}}AP^{-1}v = A_{\mathrm{R}}b$. We have considered, for instance, Bi-CGSTAB, restarted GCR, and IDR. Sometimes we could achieve faster convergence by tuning the restart method of GCR, although, this did not happen with all types of scatterers (only with a homogeneous lossless one). Moreover, with some scatterers (inhomogeneous with losses) GCR and other algorithms would simply stagnate. Thus, for the moment, the restarted GMRES remains the most robust solver for the present problem, whereas other algorithms require more tuning and a more systematic study.

3.7 Concluding remarks

The analysis of the domain integral operator of the transverse electric scattering presented here helps to understand the reasons behind the extremely slow convergence of iterative methods often observed with this method. Unlike the TM case, which is equivalent to the Helmholtz equation and can be easily accelerated by deflating the largest eigenvalues, the TE case showed no improvement with this type of preconditioner. Direct numerical comparison of the TE and TM matrix spectra indicates that the TE problem features extended and very dense eigenvalue clusters accumulating around the convex hull of the essential spectrum of the pertaining singular integral operator. These clusters cause stagnation of the deflation process, especially with high-contrast objects. We have derived an analytical multiplicative regularizer and demonstrated on various test scatterers that its discretized version robustly reduces the extent of the troublesome dense clusters. Notably, the discretized regularizer does not require any extra computer memory and its action on a vector can be computed at FFT speed. Applying deflation of the largest-magnitude eigenvalues on the regularized system we could achieve up to 46.6 times acceleration of the restarted GMRES with relatively large memory, and up to 5.1 times with modest memory. The off-line work typically takes only a small fraction of the total time, since one can apply a very rough tolerance in the EIGS algorithm when recovering the largest-magnitude eigenvalues/eigenvectors. Somewhat surprisingly, this preconditioner showed a tendency to become more efficient (on average) with the increase in the object size and permittivity.

Bibliography

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions*. Dover, New York, 1972.

[2] F. P. Andriulli, K. Cools, H. Bağci, F. Olyslager, A. Buffa, S. Christiansen, and E. Michielssen. A multiplicative Calderón preconditioner for the electric field integral equation. *IEEE Trans. Ant. Propag.*, 56:2398–2412, 2008.

[3] H. Bağci, F. P. Andriulli, K. Cools, F. Olyslager, and E. Michielssen. A Calderón multiplicative preconditioner for coupled surface-volume electric field integral equations. *IEEE Trans. Ant. Propag.*, 58:2680–2690, 2010.

[4] C. A. Balanis. *Advanced Engineering Electromagnetics*. John Wiley & Sons, Inc., New York, 1989.

[5] N. V. Budko and A. B. Samokhin. Spectrum of the volume integral operator of electromagnetic scattering. *SIAM J. Sci. Comp.*, 28:628–700, 2006.

[6] N. V. Budko, A.B. Samokhin, and A.A. Samokhin. A generalized overrelaxation method for solving singular volume integral equations in low-frequency scattering problems. *Differential Equations*, 41:1262–1266, 2005.

[7] A. L. Van Buren and J. E. Boisvert. Accurate calculation of the modified Mathieu functions of integer order. *Quart. Appl. Math.*, LXV:1–23, 2007.

[8] W. J. Byun, J. W. Yu, and N. H. Myung. TM scattering from hollow and dielectric-filled semielliptic channels with arbitrary eccentricity in a perfectly conducting plane. *IEEE Trans. Microw. Th. Tech.*, 46:1336–1339, 1998.

[9] S. Caorsi and M. Pastorino. Scattering by multilayer isorefractive elliptic cylinders. *IEEE Trans. Ant. Propag.*, 52:189–196, 2004.

[10] S. Caorsi, M. Pastorino, and M. Raffetto. Electromagnetic scattering by a multilayer elliptic cylinder under transverse-magnetic illumination: Series solution

in terms of Mathieu functions. *IEEE Trans. Ant. Propag.*, 45:926–935, 1997.

[11] S. Caorsi, M. Pastorino, and M. Raffetto. Analytic SAR computation in a multilayer elliptic cylinder for bioelectromagnetic applications. *Bioelectromagn.*, 20: 365–371, 1999.

[12] M. Card, M. Bleszynskiz, and J. L. Volakis. A near-field preconditioner and its performance in conjunction with the BiCGstab(l) solver. *IEEE Antennas and Propagation Magazine*, 46:23–30, 2004.

[13] Allen H. Cherin. *An introduction to optical fibers*. McGraw-Hill, New York, 1983.

[14] K. S. Chiang. Finite element method for cutoff frequencies of weakly guiding fibres of arbitrary cross-section. *Opt. Quant. Electr.*, 16:487–493, 1984.

[15] S. H. Christiansen and J.-C. Nedelec. A preconditioner for the electric field integral equation based on Calderón formulas. *SIAM J. Numer. Anal.*, 40:1100–1135, 2003.

[16] R. Coifman, V. Rokhlin, and S. Wandzura. The fast multipole method for the wave equation: A pedestrian prescription. *IEEE Antennas and Propagation Magazine*, 35:7–12, 1993.

[17] M. Costabel, E. Darrigrand, and E. H. Koné. Volume and surface integral equations for electromagnetic scattering by a dielectric body. *J. Comput. Appl. Math.*, 234:1817–1825, 2010.

[18] M. Costabel, E. Darrigrand, and H. Sakly. The essential spectrum of the volume integral operator in electromagnetic scattering by a homogeneous body. *http://hal.archives-ouvertes.fr/hal-00646229*, November 2011. URL `http://hal.archives-ouvertes.fr/hal-00646229/en/`.

[19] D. G. Duffy. *Transform methods for solving partial differential equations*. CRC Press, Boca Raton, 2004.

[20] R. B. Dyott. Cut-off of the first higher order modes in elliptical dielectric waveguide: An experimental approach. *Electr. Lett.*, 26:1721–1723, 1990.

[21] J. Erhel, K. Burrage, and B. Pohl. Restarted GMRES preconditioned by deflation. *J. Com. App. Math.*, 69:303–318, 1996.

[22] Y. A. Erlangga. Advances in iterative methods and preconditioners for the Helmholtz equation. *Arch. Comput. Methods Eng.*, 15:37–66, 2008.

[23] Y. A. Erlangga, C. W. Oosterlee, and C. Vuik. A novel multigrid based preconditioner for heterogeneous Helmholtz problems. *SIAM J. Sci. Comp.*, 27: 1471–1492, 2006.

[24] L. Eyges, P. Gianino, and P. Wintersteiner. Modes of dielectric waveguides of arbitrary cross sectional shape. *J. Opt. Soc. Am.*, 69:1226–1235, 1979.

[25] Z. H. Fan, D. X. Wang, R. S. Chen, and E. K. N. Yung. The application of iterative solvers in discrete dipole approximation method for computing electromagnetic scattering. *Microwave Opt. Tech. Lett.*, 48:1741–1746, 2006.

[26] P. J. Flatau. Improvements in the discrete-dipole approximation method of computing scattering and absorption. *Optics Letters*, 22:1205–1207, 1997.

[27] I. Gómez-Castellanos and R. M. Rodríguez-Dagnino. Intensity distributions and cutoff frequencies of linearly polarized modes for a step-index elliptical optical fiber. *Opt. Eng.*, 46:045003, 2007.

[28] R. F. Harrington. *Field Computation by Moment Method*. IEEE press, Piscataway, NJ, 1993.

[29] F. Ihlenburg and I. Babuska. Dispersion analysis and error estimation of

Galerkin finite element methods for the helmholtz equation. *Int. J. Numer. Methods Eng.*, 38:3745–3774, 1995.

[30] F. Ihlenburg and I. Babuska. Finite element solution of the Helmholtz equation with high wave number. Part I: the h-version of the FEM. *Comput. Math. Appl.*, 30:9–37, 1995.

[31] A. S. Ilinski, A. B. Samokhin, and U. U. Kapustin. Mathematical modelling of 2D electromagnetic scattering. *Computers and Mathematics with Applications*, 40:1363–1373, 2000.

[32] N. Joachimowicz and C. Pichot. Comparison of three integral formulations for the 2-D TE scattering problem. *IEEE Trans. Microw. Th. Tech.*, 38:178–185, 1990.

[33] A. Kirsch. An integral equation approach and the interior tramsmission problem from maxwell's equations. *Inverse Problems and Imaging*, 1:107–127, 2007.

[34] R. E. Kleinman, G. F. Roach, and P. M. van den Berg. Convergent Born series for large refractive indices. *Journal of the Optical Society of America*, 7:890–897, 1990.

[35] G. C. Kokkorakis and J. A. Roumeliotis. Power series expansions for mathieu functions with small arguments. *Math. Comput.*, 70:1221–1235, 2001.

[36] G. C. Kokkorakis and J. A. Roumeliotis. Acoustic eigenfrequencies in concentric spheroidal-spherical cavities. *J. Soun. Vibr.*, 206:287–308, 1997.

[37] A. D. Kotsis and J. A. Roumeliotis. Acoustic scattering by a penetrable spheroid. *Acous. Phys.*, 54:153–167, 2008.

[38] J. E. Lewis and G. Deshpande. Modes on elliptical cross-section dielectric-tube waveguides. *IEE J. Microw. Opt. Acous.*, 3:147–155, 1979.

[39] R. Li, X. Han, and K. F. Ren. Generalized Debye series expansion of electromagnetic plane wave scattering by an infinite multilayered cylinder at oblique incidence. *Phys. Rev. E*, 79:036602, 2009.

[40] Kuo-Nan Liou. Electromagnetic scattering by arbitrarily oriented ice cylinders. *Applied Optics*, 11:667–674, 1972.

[41] D. E. Livesay and K.-M. Chen. Electromagnetic fields induced inside arbitrarilly shaped biological bodies. *IEEE Trans. Microw. Th. Tech.*, 22:1273–1280, 1974.

[42] S. C. Mao and Z. S. Wu. Scattering by an infinite homogenous anisotropic elliptic cylinder in terms of Mathieu functions and Fourier series. *J. Opt. Soc. Am. A*, 25:2925–2931, 2008.

[43] S. C. Mao, Z. S. Wu, and H. Y. Li. Three-dimensional scattering by an infinite homogeneous anisotropic elliptic cylinder in terms of Mathieu functions. *J. Opt. Soc. Am. A*, 26:2282–2291, 2009.

[44] D. Marcuse. *Light Transmission Optics*. Van Nostrand Reinhold Co., New York, 1972.

[45] O. J. F. Martin and N. B. Piller. Electromagnetic scattering in polarizable backgrounds. *Phys. Rev. E*, 58:3909–3915, 1998.

[46] S. G. Mikhlin and S. Prössdorf. *Singular Integral Operators*. Springer–Verlag, Berlin, 1986.

[47] P. M. Morse and H. Feshbach. *Methods of Theoretical Physics*. McGraw-Hill, New York, 1953.

[48] M. Pastorino. *Microwave Imaging*. John Wiley & Sons, Inc., Hoboken, 2010.

[49] A. F. Peterson and P. W. Klock. An improved MFIE formulation for TE-wave scattering from lossy, inhomogeneous dielectric cylinders. *IEEE Trans. Ant.*

Propag., 36:45–49, 1988.

[50] S. R. Rengarajan. On higher order mode cutoff frequencies in elliptical step index fibers. *IEEE Trans. Microw. Theory Tech.*, 37:1244–1248, 1989.

[51] S. R. Rengarajan and J. E. Lewis. First higher-mode cutoff in two-layer elliptical fibre waveguides. *Electr. Lett.*, 16:263–264, 1980.

[52] S. R. Rengarajan and J. E. Lewis. Propagation characteristics of elliptical dielectric-tube waveguides. *IEE Proc.*, 127:121–126, 1980.

[53] J. H. Richmond. TE-wave scattering by a dielectric cylinder of arbitrary cross-section shape. *IEEE Trans. Ant. Propag.*, 14:460–464, 1966.

[54] J. A. Roumeliotis, H. K. Manthopoulos, and V. K. Manthopoulos. Electromagnetic scattering from an infinite circular metallic cylinder coated by an elliptic dielectric one. *IEEE Trans. Microw. Th. Tech.*, 41:862–869, 1993.

[55] B. Rulf. Electromagnetic scattering by an elliptic cylinder embedded in a uniaxially anisotropic medium. *J. Opt. Soc. Am*, 56:595–601, 1966.

[56] S. M. Saad. On the higher order modes of elliptical optical fibers. *IEEE Trans. Microw. Theory Tech.*, MTT-33:1110–1113, 1985.

[57] A. B. Samokhin. *Integral Equations and Iteration Methods in Electromagnetic Scattering*. VSP, Utrecht, The Netherlands, 2001.

[58] S. P. Schlesinger, P. Diament, and A. Vigants. On higher-order hybrid modes of dielectric cylinders (Correspondence). *IRE Trans. Microw. Theory Tech.*, 8: 252–253, 1960.

[59] A. Sebak and L. Shafai. Generalized solutions for electromagnetic scattering by elliptical structures. *Comp. Phys. Commun.*, 68:315–330, 1991.

[60] A. R. Sebak. Scattering from dielectric-coated impedance elliptic cylinder. *IEEE Trans. Ant. Propag.*, 48:1574–1580, 2000.

[61] J. K. Shaw, W. M. Henry, and W. R. Winfrey. Weakly guiding analysis of elliptical core step index waveguides based on the characteristic numbers of Mathieu's equation. *J. Lightw. Techn.*, 13:2359–2371, 1995.

[62] J. Sifuentes. Preconditioning the integral formulation of the Helmholtz equation via deflation. Master's thesis, Applied Mathematics, Rice University, Houston, Texas, 2006. Online at: `http://scholarship.rice.edu/handle/1911/17917`.

[63] A. W. Snyder. Asymptotic expressions for eigenfunctions and eigenvalues of a dielectric or optical waveguide. *IEEE Trans. Microw. Theory Tech.*, MTT-17: 1130–1138, 1969.

[64] A. W. Snyder and W. R. Young. Modes of optical waveguides. *J. Opt. Soc. Am.*, 68:297–309, 1978.

[65] T. Søndergaard. Modeling of plasmonic nanostructures: Green's function integral equation methods. *Phys. Stat. Sol.*, 244:3448–3462, 2007.

[66] J. M. Song, C.-C. Lu, and W. C. Chew. Multilevel fast multipole algorithm for electromagnetic scattering by large complex objects. *IEEE Trans. Ant. Propag.*, 45:1488–1493, 1997.

[67] C.-C. Su. A simple evaluation of some principal value integrals for dyadic Green's function using symmetry property. *IEEE Trans. Ant. Propag.*, 35:1306–1307, 1987.

[68] J. L. Tsalamengas. Exponentially converging Nyström methods applied to the integral-integrodifferential equations of oblique scattering/hybrid wave propagation in presence of composite dielectric cylinders of arbitrary cross section. *IEEE Trans. Ant. Propag.*, 55:3239–3250, 2007.

[69] N. L. Tsitsas, E. G. Alivizatos, H. T. Anastassiu, and D. I. Kaklamani. Optimization of the method of auxiliary sources (MAS) for scattering by an infinite cylinder under oblique incidence. *Electromagn.*, 25:39–54, 2005.

[70] G. D. Tsogkas, J. A. Roumeliotis, and S. P. Savaidis. Scattering by an infinite elliptic metallic cylinder. *Electromagn.*, 27:159–182, 2007.

[71] G. D. Tsogkas, J. A. Roumeliotis, and S. P. Savaidis. Cutoff wavelengths of elliptical metallic waveguides. *IEEE Trans. Microw. Theory Tech.*, 57:2406–2415, 2009.

[72] G. D. Tsogkas, J. A. Roumeliotis, and Stylianos P. Savaidis. Electromagnetic scattering by an infinite elliptic dielectric cylinder with small eccentricity using perturbative analysis. *IEEE Trans. Ant. Propag.*, 58:107–121, 2010.

[73] J. L. Volakis. Alternative field representations and integral equations for modeling inhomogeneous dielectrics. *IEEE Trans. Microw. Th. Tech.*, 40:604–608, 1992.

[74] Y. H. Wang and X. Zhang. Elliptical Fourier series expansion method together with cutoff frequencies in elliptical optical waveguides. *J. Lightw. Techn.*, 16: 1933–1941, 1998.

[75] G. N. Watson. *A Treatise on the Theory of Bessel Functions*. Cambridge University Press, Cambridge, England, 1958.

[76] J. Yan, R. K. Gordon, and A. A. Kishk. Electromagnetic scattering from impedance elliptic cylinders using finite difference method (oblique incidence). *Electromagn.*, 15:157–173, 1995.

[77] C. Yeh. Elliptical dielectric waveguides. *J. Appl. Phys.*, 33:3235–3243, 1962.

[78] C. Yeh. Scattering of obliquely incident light waves by elliptical fibers. *J. Opt. Soc. Am.*, 54:1227–1231, 1964.

[79] C. Yeh. Backscattering cross section of a dielectric elliptical cylinder. *J. Opt. Soc. Am.*, 55:309–314, 1965.

[80] C. Yeh. Modes in weakly guiding elliptical optical fibres. *Opt. Quant. Electr.*, 8: 43–47, 1976.

[81] C. Yeh. Guided-wave modes in cylindrical optical fibers. *IEEE Trans. Educ.*, E-30:43–51, 1987.

[82] C. Yeh, K. Ha, S. B. Dong, and W. P. Brown. Single-mode optical waveguides. *Appl. Opt.*, 18:1490–1504, 1979.

[83] M. A. Yurkin and A.G. Hoekstra. The discrete dipole approximation: An overview and recent developments. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 106:558–589, 2007.

[84] G. P. Zouros. Electromagnetic plane wave scattering by arbitrarily oriented elliptical dielectric cylinders. *J. Opt. Soc. Am. A*, 28:2376–2384, 2011.

[85] G. P. Zouros and N. V. Budko. Transverse electric scattering on inhomogeneous objects: spectrum of integral operator and preconditioning. *SIAM J. Sci. Comp.*, 34:B226–B246, 2012.

[86] G. P. Zouros and J. A. Roumeliotis. Exact and closed-form cutoff wavenumbers of elliptical dielectric waveguides. *IEEE Trans. Microw. Th. Tech.*, 2012, submitted for publication.