

Embedded Element Pattern Prediction in Mutually Coupled Array Antennas Using Isolated Element Far-Field Data

André S. Conradie 20899068

Report submitted in partial fulfilment of the requirements of the module Project (E) 448 for the degree Baccalaureus in Engineering in the Department of Electrical and Electronic Engineering at Stellenbosch University.

Supervisor: Prof. D. I. L. de Villiers

November 2020

Acknowledgements

I would like to express my sincere thanks to the following people:

Prof. D.I.L. de Villiers for his guidance, encouragement and insight throughout this project.

My parents for loving and supporting me and for raising me to be curious.

The friends I have made over the last four years for making sure that I spend enough time away from my textbooks.



Plagiaatverklaring / Plagiarism Declaration

- 1. Plagiaat is die oorneem en gebruik van die idees, materiaal en ander intellektuele eiendom van ander persone asof dit jou eie werk is.
 - Plagiarism is the use of ideas, material and other intellectual property of another's work and to present is as my own.
- 2. Ek erken dat die pleeg van plagiaat 'n strafbare oortreding is aangesien dit 'n vorm van diefstal is.
 - I agree that plagiarism is a punishable offence because it constitutes theft.
- 3. Ek verstaan ook dat direkte vertalings plagiaat is.

 I also understand that direct translations are plagiarism.
- 4. Dienooreenkomstig is alle aanhalings en bydraes vanuit enige bron (ingesluit die internet) volledig verwys (erken). Ek erken dat die woordelikse aanhaal van teks sonder aanhalingstekens (selfs al word die bron volledig erken) plagiaat is.

 Accordingly all quotations and contributions from any source whatsoever (including the internet) have been cited fully. I understand that the reproduction of text without quotation marks (even when the source is cited) is plagiarism
- 5. Ek verklaar dat die werk in hierdie skryfstuk vervat, behalwe waar anders aangedui, my eie oorspronklike werk is en dat ek dit nie vantevore in die geheel of gedeeltelik ingehandig het vir bepunting in hierdie module/werkstuk of 'n ander module/werkstuk nie.
 - I declare that the work contained in this assignment, except where otherwise stated, is my original work and that I have not previously (in its entirety or in part) submitted it for grading in this module/assignment or another module/assignment.

20899068	-A Stoneda
Studentenommer / Student number	Handtekening / Signature
Conradie, A.S.	9 November 2020
Voorletters en van / Initials and surname	Datum / Date

Abstract

English

Antenna arrays have the potential to meet the directivity requirements in modern communication and radar systems while reducing the cost and complexity of these systems. The analysis of antenna arrays can be time-consuming, particularly during the synthesis of sparse antenna arrays where several configurations must be analysed. This project considers the implementation of the Spherical-Wave Plane-Wave method for the fast prediction of embedded element patterns in $\lambda/2$ -dipole antenna arrays and its application to the synthesis of sparse arrays. This method uses sampled far-field radiation and scattering data from an isolated element to model mutual coupling in antenna arrays through a multi-scattering approach. The accuracy of the method is evaluated by comparing its results for several regular and irregular array configurations to method of moment-based full-wave simulations. The implementation in this paper is restricted to dipole antennas, but other work has shown that it may be generalised to more complex radiating structures.

Afrikaans

Antennasamestellings het die potensiaal om aan die uitstralingsverseistes in moderne kommunikasie- en radarstelsels te voldoen, terwyl dit koste verminder en die stelsel se implementering vereenvoudig. Die analise van sulke samestellings kan tydrowend wees, veral wanneer verskeie oneweredige samestellings geanaliseer moet word. Die Sferiese Golf-Platvlakgolfmetode vir die voorspelling van uitstralingspatrone word vir $\lambda/2$ -dipoolantennasamestellings geïmplementeer en aangewend vir die ontwerp van yl samestellings. Hierdie metode gebruik die uitstralings- en verstooiingsinligting van geïsoleerde elemente om die koppeling tussen verskeie elemente in 'n samestelling deur 'n meervoudigeverstrooiingsproses te modelleer. Die akkuraatheid van hierdie metode word ge-evalueer deur die voorspelde uitstralingspatrone met momentmetode-gebaseerde volgolf-analise resultate te vergelyk. Die huidige implementering is beperk tot dipoolantennas, maar ander werk het reeds gewys dat die metode veralgemeen kan word om meer ingewikkelde antennasamestellings te analiseer.

Contents

De	eclara	tion		ii
Αŀ	ostrac	t		iii
Lis	st of	Figures	S	vi
Lis	st of	Tables		viii
No	omen	clature		ix
1.	Intro	oductio	on	1
	1.1.	Backg	round	. 1
	1.2.	Projec	et Description	. 2
2.	The	oretica	l formulation	4
	2.1.	Introd	luction	. 4
	2.2.	Backg	round	. 4
		2.2.1.	Maxwell's equations	. 4
		2.2.2.	Full-wave analysis	. 4
		2.2.3.	The method of moments	. 5
	2.3.	The S	W-PW method	. 7
		2.3.1.	Intuitive description	. 7
		2.3.2.	Mathematical formulation	. 8
	2.4.	Spheri	ical wave expansion	. 11
	2.5.	Conclu	usion	. 17
3.	Imp	lement	ation and results	18
	3.1.	Introd	luction	. 18
	3.2.	Imple	mentation	. 18
		3.2.1.	The input data	. 19
		3.2.2.	The antenna type	. 19
		3.2.3.	Error metrics	. 19
		3.2.4.	Convergence	. 20
		3.2.5.	Software implementation and dependencies	. 21
	9 9	Dagul		21

Contents v

		3.3.1.	Spherical wave expansion	21
		3.3.2.	Two-element test arrays	23
		3.3.3.	Uniformly spaced linear arrays	23
	3.4.	Infinite	e ground plane model	25
	3.5.	Conclu	sion	26
4.	Арр	lication		28
	4.1.	Introd	uction	28
	4.2.	The de	esign process	28
		4.2.1.	Synthesis excluding mutual coupling	28
		4.2.2.	Synthesis including mutual coupling	30
	4.3.	Conve	x optimisation	31
		4.3.1.	Interior point methods	32
		4.3.2.	Newton's method	34
	4.4.	Array	synthesis	35
		4.4.1.	Dipole array antenna with 10λ aperture	35
		4.4.2.	Dipole array antenna with 120λ aperture	37
	4.5.	Conclu	sion	37
5.	Sum	mary a	and Conclusion	39
Bil	bliogi	aphy		41
Α.	Proj	ect Pla	nning Schedule	43
В.	Out	comes	Compliance	44

List of Figures

1.1.	The presence of a passive conducting structure has a significant impact on the radiation characteristics of the active antenna	2
2.1.	A single active element induces scattered fields at adjacent elements. Each of these scattered fields induce a subsequent set of scattered fields. Iterating these interactions yields the active element's EEP. This is similar to Figure	
	1 in [5]	8
2.2.	The Legendre polynomials represent the radial variation in SWE, but the unnormalised associated Legendre polynomials pose a risk of numerical	
	errors due to the rapid increase in the functional values as m and n increase.	14
2.3.	The spherical Hankel function of the second kind is a complex-valued	
	function which represents outward traveling waves	15
3.1.	Convergence plots for the Jacobi and Gauss-Seidel iterative schemes for an eqi-spaced linear array of nine elements. The logarithm of the relative	
	difference, ρ , is plotted against the iteration number, k , when calculating	
	the EEP of the centre element for various inter-element spacings shown in	
	the legend in terms of λ . The convergence rate and robustness is improved	
	by adopting the Gauss-Seidel scheme	21
3.2.	Plots of the normalised H-plane directivity of the EEP of an active element	
	with a single terminated passive element placed at different separation distances. The EEPs determined for $N=3$ and $N=12$ using the SW-PW	
	method are compared to those exported from FEKO	23
3.3.	Plots depicting the error in the prediction of EEPs in two-element arrays	24
3.4.	Plots comparing the MATLAB and FEKO results for a three-element array	
	with constant spacings of 0.35λ	24
3.5.	Plots comparing the MATLAB and FEKO results for a nine-element array	
	with constant spacings of 0.35λ	25
3.6.	The nine-element, $\lambda/2$ uniformly spaced array, $\lambda/4$ in front of an infinite	
	ground plane	26
3.7.	Plots of the embedded element normalised H-plane directivity of a nine-	
	element array in front of an infinite ground plane	26

4.1.	Block diagram of the iterative approach including intermediate step em-	
	ploying SW-PW-determined EEPs. This is similar to Figure 1 in $[2]$	31
4.2.	The log-barrier approximation reduces inter-point method solutions to a	
	sequence of Newton's method solutions	34
4.3.	Result of the 10λ aperture array, synthesized in the absence of mutual	
	coupling	35
4.4.	Plots depicting the results of the 10λ array synthesis process	37
4.5.	Normalised H-plane directivity of the results for the 120λ aperture array	
	when utilising IEPs and SW-PW-determined EEPs	38

List of Tables

3.1.	SWE reconstructed near-field values for different numbers of modes com-					
	pared to near-field values determined in FEKO. Reported errors are for					
	magnitude reconstruction	22				
4.1.	Element positions and complex excitations (in Volt) after each iteration	36				

Nomenclature

Variables and functions

 λ Wavelength

B Magnetic flux density

 μ Permeability

H Magnetic field intensity

D Electric flux density

 ϵ Permittivity, Small number

E Electric field intensity

 ∇ Del operator

M Magnetic current density, TE-modes

J Electric current density, MoM current density vector

 ρ_e Electric charge density

 ρ_m Magnetic charge density

Z MoM matrix

V MoM excitation vector

M Number of array elements

 N_m Number of mesh segments

 \mathbf{Z}_p^{far} Matrix mapping a known current distribution to its far-field pattern

f Far-field pattern

 $\hat{\boldsymbol{n}}$ Outward pointing unit normal vector

 \mathbf{E}^{s} Matrix mapping incident E-field to its induced scattered E-field

 \mathbf{w}_{pm} Weight of field radiated by element m at element p

 \hat{r} Spherical coordinate system radial unit vector

 $\hat{\boldsymbol{\theta}}$ Spherical coordinate system azimuthal unit vector

 $\hat{\phi}$ Spherical coordinate system elevation unit vector

f^s Matrix mapping incident far-field pattern to its induced scattered

far-field pattern

Nomenclature x

 \mathbf{r}_m Array element reference vector

 ψ Solution to the homogeneous scalar wave equation

F Vector potential

A Vector potential

N TM-modes

R Radial component of ψ

 Θ Elevation component of ψ

 Φ Azimuthal component of ψ

 P_n Regular Legendre polynomial of degree n

 P_n^m Associated Legendre Polynomial of degree n and order m

 \bar{P}_n^m Normalised associated Legendre polynomial

 z_n Spherical Bessel function of order n

 Z_n Ordinary Bessel function of order n

 $h_n^{(2)}$ Spherical Hankel function of the second kind and order n

N Truncation index for SWE

 \mathbf{m}_{nm} Generating function for TE-modes in SWE

 \mathbf{n}_{nm} Generating function for TM-modes in SWE

 c_{nm} Scaling factor for SWE modes

 ζ Intrinsic impedance of free space

k Wavenumber, Iteration index

 Q_{smn} SWE coefficients

 \hat{z} Rectangular coordinate system z-directed unit vector

 Z_0 Characteristic impedance

 N_{Ω} Number of angular samples in far-field representation

 ρ Relative difference

W Matrix of weights of radiated fields at other elements

 δ Stopping condition for SW-PW iterations and Newton's method

w Excitation Vector

AF Array factor

 $||\cdot||_{l_n}$ l_n -norm

 $M(\hat{r})$ Target radiation pattern mask

Nomenclature xi

$Z^{(i)}$	Convex optimisation weight matrix at the i^th iteration
$z_p^{(i)}$	p^{th} entry in $Z^{(i)}$
$g(\cdot)$	General function
$I_{_}$	Indicator function
$\hat{I}_{_}$	Log-barrier approximation function

Acronyms and abbreviations

CEM Computational Electromagnetics

EEP Embedded Element Pattern

EM Electromagnetic

EMACS Electromagnetics and Microwave, Antenna and Computational Systems

IEP Isolated Element Pattern

MoM Method of Moments

PWS Plane Wave Spectrum

RMSE Root Mean Square Error

SLL Sidelobe Level

SWE Spherical Wave Expansion

TE Transverse Electric

TM Transverse Magnetic

Chapter 1

Introduction

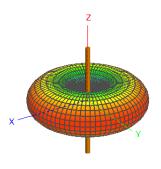
1.1. Background

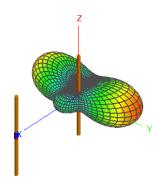
In order to obtain the large directivity required by many modern communication and radar systems, antennas with apertures which are large in terms of wavelength must be used. This may be achieved by means of array antennas or reflector antennas. Although array antennas are often more expensive to design and implement, they offer attractive advantages over reflector antennas. Such advantages include their being made flat so as to occupy less space than reflector antennas, and their offer of electronic beam-steering capabilities which renders arrays more versatile than reflectors [1].

The analysis of antenna arrays can be time-consuming, particularly during the design process when several full-wave simulations must be performed [2]. Much time has been devoted to developing methods to simplify the synthesis of antenna arrays, however, many of these methods utilise isolated or isotropic elements which neglect the effects of mutual coupling and lead to sub-optimal practical implementations [3]. An optimisation framework which aims to minimise the number of full-wave simulations that need to be performed during synthesis of maximally sparse antenna arrays in the presence of coupling effects, is proposed in [2].

The effects of mutual coupling may be observed through the difference between embedded and isolated element patterns [4]. Isolated element patterns (IEP) display the radiation characteristics of a single antenna when its environment is free of other radiating or scattering structures. Conversely, embedded element patterns (EEP) display the radiation characteristics of an element in the *presence* of other radiating or scattering structures by combining their contributions into a single radiation pattern. Figure 1.1a depicts the radiation pattern of a single active $\lambda/2$ -dipole antenna placed at the origin, while Figure 1.1b depicts the radiation pattern of the same active element with an identical passive element placed $\lambda/2$ away.

A fast method to predict embedded element patterns was proposed in [4] and fully developed in [5]. This method follows a multi-scattering approach and utilises only the radiation and scattering characteristics of a single isolated element to determine the effects of mutual coupling in array antennas. This is a post-processing step which eliminates





(a) Isolated element pattern.

(b) Embedded element pattern.

Figure 1.1: The presence of a passive conducting structure has a significant impact on the radiation characteristics of the active antenna.

the need for multiple full-wave simulations to be performed when a number of array configurations must be analysed, for example during array synthesis.

1.2. Project Description

The main aim of this project is to implement the Spherical Wave-Plane Wave (SW-PW) method from [4] for $\lambda/2$ -dipole antennas, thus enabling the fast analysis of arrays utilising these antennas. The secondary aim is to incorporate this implementation into the array synthesis framework in [2].

Chapter 2

In this chapter an exposition of the SW-PW method is provided and a representation of Spherical Wave Expansion (SWE) is derived. This provides a full description of the theoretical background required for the implementation of the SW-PW method.

Chapter 3

In this chapter the implementation of the SW-PW method for $\lambda/2$ -dipole array antennas is discussed. This includes generating the necessary input data and integrating with the code base provided by the Electromagnetics and Microwave, Antenna and Computational Systems (EMACS) research group at Stellenbosch University. The convergence of the SW-PW method is evaluated numerically, before the accuracy of the reconstruction of near-field data from far-field data using spherical wave expansion is examined by comparing these results to the near-field results from FEKO. Finally, the accuracy of the SW-PW method is evaluated by comparing the root mean square error (RMSE) over angle and using figures to compare the predicted embedded element and array pattern results to

those calculated in FEKO.

Chapter 4

In this chapter the SW-PW method is incorporated into the framework for the synthesis of maximally sparse antenna arrays in [2] as an intermediate step to further reduce the number of full-wave simulations required during array synthesis. Two benchmark examples therein are repeated to obtain comparable results and convex optimisation methods which are key to this procedure are discussed.

Chapter 5

This report is concluded by reflecting on the results obtained in earlier chapters and considering future extensions of this work.

Chapter 2

Theoretical formulation

2.1. Introduction

The SW-PW method is demonstrated here by means of a brief survey of the background against which the method was proposed together with a qualitative description and mathematical formulation. SWE plays an integral role in the propagation step of the SW-PW method and its representation for outward propagating waves in homogeneous free space is derived.

2.2. Background

2.2.1. Maxwell's equations

Electromagnetics is the study of the effects of electrically charged particles at rest and in motion. The fundamental mathematical description of these interactions is provided by Maxwell's equations in (2.1)-(2.4) and the accompanying constitutive relations $\mathbf{B} = \mu \mathbf{H}$ and $\mathbf{D} = \epsilon \mathbf{E}$.

$$\nabla \times \mathbf{E} = -\frac{\partial}{\partial t} \mathbf{B} - \mathbf{M} \tag{2.1}$$

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial}{\partial t} \mathbf{D} \tag{2.2}$$

$$\nabla \cdot \mathbf{D} = \rho_e \tag{2.3}$$

$$\nabla \cdot \mathbf{B} = \rho_m \tag{2.4}$$

2.2.2. Full-wave analysis

Full-wave computational electromagnetic (CEM) methods approximate Maxwell's equations numerically. Some unknown electromagnetic property is discretised by subdividing the structure into a large number of small segments (this is commonly referred to as meshing). This unknown property is projected onto a set of known basis functions by assuming a functional dependence of the spacial variation of the unknown in each segment. A

CEM method is applied to determine the basis function magnitudes, and the result is an approximation of the unknown property in each segment [6].

In high frequency applications it is often challenging to accurately determine the phase of rapidly oscillating sinusoidal sources. Hence, the mesh density is determined by the required phase accuracy, rather than the need to accurately resolve the geometry of a physical structure. The finer the mesh, the more accurate the results, however, as the mesh density increases, the number of unknowns and the computational cost of finding a solution also increases [6].

2.2.3. The method of moments

The method of moments (MoM) is one of the most popular CEM methods and is implemented in various commercial full-wave simulation software packages such as FEKO and GRASP. The MoM is typically applied in the frequency domain, assuming sinusoidal excitation and using the $e^{j\omega t}$ phasor convention [6].

In the MoM, the radiating or scattering structure is replaced by equivalent surface currents which are discretised during the meshing procedure. A matrix equation representing the coupling effect of each segment on every other segment is then derived by applying a Green's function. The relevant electromagnetic boundary conditions are applied, resulting in a set of linear equations. The solution of this set of equations yields the approximate current in each segment [6]. The matrix dimensions may range from hundreds to thousands of entries, with the upper bound determined by the available computational resources [6].

As to the mathematical description of an iterative solution scheme for the MoM, [7] is instructive. The matrix equation described above typically assumes the form in

$$\mathbf{ZJ} = \mathbf{V},\tag{2.5}$$

where V is a known excitation vector, Z is a matrix describing the inter-segment coupling and J is re-assigned as the vector of unknown current densities to maintain continuity with [7]. The matrix entries are usually complex values.

When a finite antenna array of M identical disjoint elements, each discretised using N_M distinct basis functions, is considered, (2.5) is expanded to

$$\begin{bmatrix} \mathbf{Z}_{11} & \mathbf{Z}_{12} & \cdots & \mathbf{Z}_{1M} \\ \mathbf{Z}_{21} & \mathbf{Z}_{22} & \cdots & \mathbf{Z}_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{Z}_{M1} & \cdots & \cdots & \mathbf{Z}_{MM} \end{bmatrix} \begin{bmatrix} \mathbf{J}_1 \\ \mathbf{J}_2 \\ \vdots \\ \mathbf{J}_M \end{bmatrix} = \begin{bmatrix} \mathbf{V}_1 \\ \mathbf{V}_2 \\ \vdots \\ \mathbf{V}_M \end{bmatrix}.$$
(2.6)

The diagonal entries of the block-partitioned **Z**-matrix represent the intra-element coupling and the off-diagonal entries represent the inter-element coupling. A single block matrix

 \mathbf{Z}_{ij} may have hundreds or thousands of entries, so it is readily appreciated how large the entire \mathbf{Z} -matrix may become.

The large number of entries in these matrices renders direct solution by matrix inversion a computationally expensive task. Instead, an iterative block based solution scheme based on the Jacobi method may be utilised. When this method is applied to (2.6),

$$\begin{bmatrix} \mathbf{J}_{1} \\ \mathbf{J}_{2} \\ \vdots \\ \mathbf{J}_{M} \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_{11}^{-1} (\mathbf{V}_{1} - \Sigma_{m=1,m\neq1}^{M} \mathbf{Z}_{1m} \mathbf{J}_{m}) \\ \mathbf{Z}_{22}^{-1} (\mathbf{V}_{2} - \Sigma_{m=1,m\neq2}^{M} \mathbf{Z}_{2m} \mathbf{J}_{m}) \\ \vdots \\ \mathbf{Z}_{MM}^{-1} (\mathbf{V}_{M} - \Sigma_{m=1,m\neqM}^{M} \mathbf{Z}_{Mm} \mathbf{J}_{m}) \end{bmatrix}$$

$$(2.7)$$

is found. Therein, $\mathbf{J}_1, ..., \mathbf{J}_M$ are the full-wave method of moment current vectors for elements 1 to M. The current on the p^{th} element of the array may then be written as

$$\mathbf{J}_p = \mathbf{Z}_{pp}^{-1} (\mathbf{V}_p - \Sigma_{m=1, m \neq p}^M \mathbf{Z}_{pm} \mathbf{J}_m). \tag{2.8}$$

 $\mathbf{J}_{p}^{(0)} = \mathbf{Z}_{pp}^{-1} \mathbf{V}_{p}$ is the current on element p when it is viewed in isolation, then (2.8) can be transformed into the iterative solution scheme given by

$$\mathbf{J}_{p}^{(k)} = \mathbf{J}_{p}^{(0)} - \sum_{m=1, m \neq p}^{M} \mathbf{Z}_{pp}^{-1} \mathbf{Z}_{pm} \mathbf{J}_{m}^{(k-1)}.$$
 (2.9)

Herein, $\mathbf{J}_p^{(k)}$ is the current on element p calculated at iteration k. The second term on the right hand side accounts for the inter-element coupling effects by adding the currents $\mathbf{Z}_{pp}^{-1}\mathbf{Z}_{pm}\mathbf{J}_m^{(k-1)}$ induced on element p by the fields radiated by adjacent elements $m=1,2,\ldots,M$ with $m\neq p$. The elements are first viewed in isolation and the currents are then updated according to (2.9) until convergence is reached.

This iteration scheme may be expressed in words as follows: The current on each element p is equal to the current on element p when it is viewed in isolation plus the current induced on element p by the fields radiated by adjacent elements. In each step, the iteration scheme determines the additional effect that currents induced on elements m in the previous step have on the currents induced on element p in this step. The scheme terminates when the additional effects disappear or become negligibly small.

Iterative schemes such as (2.9) are considered in attempts to avoid the high computational costs of solving (2.6) directly through matrix inversion. Direct solutions scale as $O((M \times N_M)^3)$, while iterative schemes attempt to solve multiple smaller problems which scale as $O(N_M^2)$ [7]. In both cases, it remains clear that the computational time increases with both the number of basis functions and the number of antennas in the array. The SW-PW method attempts to rapidly approximate these full-wave solutions.

2.3. The SW-PW method

Finding the far-field pattern of a radiating element becomes a trivial task when its current distribution is known. From this relationship between far-field patterns and currents, one might presume that there exists an approximate method similar to (2.9), but directly utilising far-field patterns rather than currents. Such a method has the potential to simplify antenna array synthesis by reducing the number of time-consuming full-wave simulations that must be performed during the design process.

The following discussion considers an array of M identical antennas with the aim of formulating an iterative solution scheme to determine the far-field pattern of the array using only the far-field radiation and scattering data of a single isolated element.

2.3.1. Intuitive description

This discussion follows [4], where a proof of concept was given, closely.

In (2.9), the current distribution vectors are unknown. When assuming that they are known, the existence of a matrix \mathbf{Z}_p^{far} which maps these known current distributions to their radiated far-field patterns may also be assumed. Multiplying (2.9) by this matrix yields

$$\mathbf{Z}_{p}^{far}\mathbf{J}_{p}^{(k)} = \mathbf{Z}_{p}^{far}\mathbf{J}_{p}^{(0)} - \Sigma_{m=1, m \neq p}^{M}\mathbf{Z}_{p}^{far}\mathbf{Z}_{pp}^{-1}\mathbf{Z}_{pm}\mathbf{J}_{m}^{(k-1)}.$$
(2.10)

By this definition, $\mathbf{Z}_p^{far}\mathbf{J}_p^{(0)}$ becomes the far-field radiated by the element p in isolation, denoted by $\mathbf{f}_p^{(0)}$. Similarly, $\mathbf{Z}_p^{far}\mathbf{Z}_{pp}^{-1}\mathbf{Z}_{pm}\mathbf{J}_m^{(k-1)}$ is the far-field scattered by element p due to the incident field from element m. The far-field pattern of element m from the previous iteration is denoted by $\mathbf{f}_m^{(k-1)}$ and the function mapping this field incident on element p to the field scattered by p, by \mathbf{f}_{pm} . Finally, $\mathbf{Z}_p^{far}\mathbf{J}_p^{(k)}$ becomes $\mathbf{f}_p^{(k)}$, the field radiated by element p and is calculated as the sum of the field due to its own excitation and the scattered fields due to incident fields from adjacent elements. The result is

$$\mathbf{f}_p^{(k)} = \mathbf{f}_p^{(0)} - \sum_{m=1, m \neq p}^M \mathbf{f}_{pm} \mathbf{f}_m^{(k-1)}.$$
 (2.11)

The assumption of known current distribution vectors is validated by considering the full-wave solution of a single active element, p, as the staring point. The M-1 other elements are passive terminated elements. Iterating (2.11) until it converges then yields p's EEP, which is its radiation pattern due to both its own excitation and the mutual coupling between all array elements. After repeating this process for all M elements, the array pattern may be determined as the superposition of the EEPs. Since the elements are identical, only a single isolated element's radiation and scattering characteristics is required to determine the array pattern in the presence of mutual coupling.

This multi-scattering approach is illustrated in Figure 2.1, where the radiated field

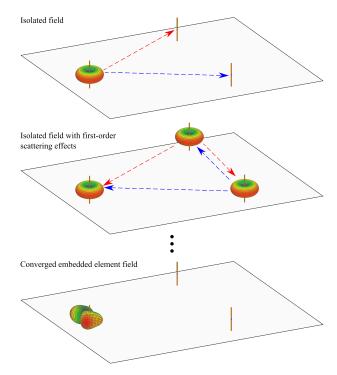


Figure 2.1: A single active element induces scattered fields at adjacent elements. Each of these scattered fields induce a subsequent set of scattered fields. Iterating these interactions yields the active element's EEP. This is similar to Figure 1 in [5].

from the excited element induces scattered fields at each other element. Each scattered field then induces it's own subsequent set of scattered fields and the result of the converged iterative scheme is the active element's EEP.

2.3.2. Mathematical formulation

This discussion follows [5] closely.

By the equivalence principle, the source distribution inside a region can be replaced by equivalent sources outside the region which would produce the same fields [8]. Consider a radiation source inside a region S, with free space both inside and outside that region, and known fields on the (imaginary) surface of S. Equivalent surface currents on S may be calculated as $\mathbf{J}_s = \hat{\boldsymbol{n}} \times \mathbf{H}$ and $\mathbf{M}_s = \mathbf{E} \times \hat{\boldsymbol{n}}$. \mathbf{J}_s and \mathbf{M}_s are the electric and magnetic surface currents, respectively, $\hat{\boldsymbol{n}}$ is the outward pointing unit normal vector on the surface of S, and \mathbf{E} and \mathbf{H} are the known fields on the surface of S. The fields at any point outside S may then be determined by integrating over these current distributions, and by the uniqueness principle, these are the originally postulated fields.

As to the implementation of the SW-PW method, the effect is that the angular components on a minimum sphere containing all radiating parts of an element provide a full description of the field radiated or scattered by the element.

Equation (2.9) considers currents, multiplying it by \mathbf{Z}_{pp} yields an iterative scheme

considering voltages. This description is generalised by considering E-fields in

$$\mathbf{E}_{p}^{(k)} = \mathbf{E}_{p}^{(0)} - \sum_{m=1, m \neq p}^{M} \mathbf{E}_{pm}^{s} \mathbf{E}_{m}^{(k-1)}.$$
 (2.12)

Herein, $\mathbf{E}_p^{(0)}$ is the E-field of element p in isolation and \mathbf{E}_{pm}^s is a function which maps the field radiated by element m to that scattered by element p.

Local spherical coordinate systems are defined at elements p and m and denoted the unit vectors by $\{\hat{\boldsymbol{r}}_p, \hat{\boldsymbol{\theta}}_p, \hat{\boldsymbol{\phi}}_p\}$ and $\{\hat{\boldsymbol{r}}_m, \hat{\boldsymbol{\theta}}_m, \hat{\boldsymbol{\phi}}_m\}$. Following the discussion above, the angular components of the E-field radiated by element m may be restricted to its own minimum sphere and this is written as $\mathbf{E}_m^{(k-1)} \mid_{\theta_m,\phi_m}$, where \mid_{θ_m,ϕ_m} indicates the restriction. \mathbf{E}_{pm}^s then takes the form

$$\mathbf{E}_{pm}^{s} = \begin{bmatrix} E_{pm_{\theta_{p},\theta_{m}}}^{s} & E_{pm_{\theta_{p},\phi_{m}}}^{s} \\ E_{pm_{\phi_{p},\theta_{m}}}^{s} & E_{pm_{\phi_{p},\phi_{m}}}^{s} \\ E_{pm_{r_{p},\theta_{m}}}^{s} & E_{pm_{r_{p},\phi_{m}}}^{s} \end{bmatrix}, \tag{2.13}$$

where the columns map the $\hat{\boldsymbol{\theta}}_{m}$ - and $\hat{\boldsymbol{\phi}}_{m}$ -directed components of the field radiated by m to the fields scattered by element p.

The effect that radiated fields have on adjacent elements is central to the multiple scattering approach in the SW-PW method. Rather than restricting the field radiated by element m to its own minimum sphere, it is propagated outward and restricted it to its tangential components on the minimum sphere of element p. This is called the weight of m at p. $\mathbf{E}_{m}^{(k-1)}$ is replaced by $\mathbf{w}_{pm}^{(k-1)}$ to emphasize this propagation step. The existence of a transformation matrix \mathbf{E}_{pm}^{s} is established by the equivalence principle.

When the inter-element separation approaches the far-field distance, the process simplifies. The propagating waves may be approximated as transverse electromagnetic waves which have no radially directed components, and the $\frac{1}{r}e^{jkr}$ free-space propagation factor may be applied. The far-field pattern and the propagation factor are separable, so the far-field pattern provides the full description of the E-field's directional dependence. As a result, \mathbf{E}_{pm}^s becomes a 2×2 transformation and the E-field notation in (2.12) may be replaced with the equivalent far-field notation in

$$\mathbf{f}_{p}^{(k)} = \mathbf{f}_{p}^{(0)} - \sum_{m=1, m \neq p}^{M} \mathbf{f}_{pm}^{s} \mathbf{w}_{pm}^{(k-1)}.$$
 (2.14)

The notation utilising the weight of the fields from elements m at element p is retained to emphasize the propagation step.

The array far-field is determined by applying the superposition principle. This result at iteration k is given by

$$\mathbf{f}^{(k)} = \sum_{m=1}^{M} \mathbf{f}_m^{(k)} e^{jk\hat{\mathbf{r}}\cdot\mathbf{r}_m},\tag{2.15}$$

where k is the wavenumber, \hat{r} is the directional unit vector from the array reference point, r_m is the reference point of element m and the $e^{jk\hat{r}\cdot r_m}$ term shifts the phase reference of

element m to the array reference point.

Equation (2.14) represents the foundation of the SW-PW method. The following practical considerations are yet to be addressed: (i) the availability of far-field radiation data; (ii) the propagation of the field radiated by element m to the minimum sphere of element p; (iii) how to apply (2.14), which is defined for elements in each other's far-fields, to elements in coupled arrays; and (iv) how to obtain scattered field data.

As to (i) above, the far-field radiation data is readily obtainable from any commercial full-wave simulation package.

As to (ii) above, in each step, the E-field in (2.12), or its far-field equivalent in (2.14), must be propagated from elements m to element p. The field must be evaluated at the minimum sphere of element p, and then it must be restricted to its tangential components in the coordinate system of element p. To evaluate the field of element m at a set of points in space distinct from those where it is known, the field must be projected onto a finite set of analytically known modes such that it may be reconstructed at these other points. In [4], both the spherical wave expansion (SWE) and the plane wave spectrum (PWS) are considered and it is determined that the SWE is the superior choice for this application.

As to (iii) above, the accuracy of such an expansion is dependent both on the number of modes included in the finite series description, and the data that is used to determine the expansion coefficients. If the coefficients are determined by using the sampled E-field data close to the minimum sphere, the expansion will provide accurate results for any point in the near field outside the sphere used to determine the coefficients. On the other hand, if a larger sphere or far-field data is used, the expansion should not produce accurate near-field results, because the data that was used to determine the coefficients did not include the evanescent field-effects close to the radiating element. However, as will be shown show in Chapter 3, the near field may be reconstructed a distances resembling the inter-element spacing in antenna arrays when far-field data is used to determine the coefficients. This allows for (2.14) to be applied to elements in one another's near-fields even though far-field data is being used.

Finally, as to (iv) above, the field incident to element p from element m must be restricted to its tangential components in the coordinate system of element p and the scattered field must be determined. A rigorous manner in which to do this is to determine the restriction directly from the SWE of the incident field. A simpler method is to turn the incident fields at element p into a single plane wave with orthogonal polarisations, using the SWE results. The field scattered by element p is then estimated from a set of scattering data imported from a full-wave simulation package where it was determined by impinging plane waves with unit amplitude on the isolated element p from multiple directions. In [4], it was determined that this simpler process exhibits results comparable to those of the more rigorous direct computation when applied to dipole antennas of varying lengths and inter-element spacings. Therein, the name "SW-PW method" was

coined seeing as the SWE is used to determine the amplitude and phase of incident plane waves used to model the inter-element coupling effects.

2.4. Spherical wave expansion

This section considers the derivation of SWE which is integral to the propagation step in the SW-PW method. The starting point is the homogeneous scalar wave equation,

$$\nabla^2 \psi + k^2 \psi = 0, \tag{2.16}$$

which is rewritten in spherical coordinates as

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\psi^2} + k^2\psi = 0 \tag{2.17}$$

The vector wave equation in spherical coordinates does not reduce to three simple uncoupled scalar wave equations as it does in rectangular coordinates [9]. Instead, it is noted that any EM wave may be represented as the superposition of transverse electric (TE) and transverse magnetic (TM) modes, thereafter [9] is illustrative. Therein, it is shown that the TE modes can always be found by letting the vector potential \mathbf{F} have only a radial component, $\mathbf{F} = F_r \hat{\mathbf{r}}$, setting the vector potential \mathbf{A} equal to zero and solving the the scalar wave equation for $r\psi = F_r$. The TE modes are then recovered from the definition of the vector potential. Similarly, the TM modes are found by assuming only a radial dependence in \mathbf{A} and setting \mathbf{F} equal to zero. The vectors \mathbf{F} and \mathbf{A} are defined in (2.18)-(2.19) and (2.20)-(2.21) respectively, where ϕ_m and ϕ_e represent arbitrary magnetic and electric scalar potentials which are not to be confused with the coordinate direction ϕ .

$$-\frac{1}{\epsilon}\nabla \times \mathbf{F} = \mathbf{E}_F \tag{2.18}$$

$$\nabla \cdot \mathbf{F} = -j\omega\mu\epsilon\phi_m \tag{2.19}$$

$$-\frac{1}{\mu}\nabla \times \mathbf{F} = \mathbf{H}_A \tag{2.20}$$

$$\nabla \cdot \mathbf{A} = -j\omega \mu \epsilon \phi_e \tag{2.21}$$

Further, when (2.17) is solved as specified above, the result is the same for both **A** and **F**. In [10], the TE and TM modes are represented by **M** and **N** respectively, and it is shown that the field components may be recovered from

$$\mathbf{M} = \nabla \times \mathbf{F} \tag{2.22}$$

and

$$\mathbf{N} = \frac{1}{k} \nabla \times \mathbf{M}.\tag{2.23}$$

M is reused for continuity with [10], but neither (2.1), nor (2.22) will be written again, so it should not cause confusion.

The aim of the present section is restated as solving (2.17) for the vector potential over distance, $\psi = F_r/r$. The strategy involves decomposing the coupled differential equation into three standard form differential equations with known solutions. The method of separation of variables is invoked to assume a solution of the form $\psi(r, \theta, \phi) = R(r)\Theta(\theta)\Phi(\phi)$:

$$\Theta\Phi \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R}{\partial r} \right) + R\Phi \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \Theta}{\partial \theta} \right) + R\Theta \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \Phi}{\partial \psi^2} + k^2 R\Theta \Phi = 0. \quad (2.24)$$

To decompose (2.24), similar variables are grouped together and replaced by constants. To this end, (2.24) is divided by $R\Theta\Phi$ and multiplied by $r^2\sin^2\theta$. The partial derivatives may be replaced by full derivatives since the argument of each derivative contains only a single variable. This yields

$$\frac{\sin^2 \theta}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \frac{\sin \theta}{\Theta} \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) + \frac{1}{\Phi} \frac{d^2 \Phi}{d\psi^2} + (kr \sin \theta)^2 = 0.$$
 (2.25)

The third term contains only the variable ϕ , so the substitution $\frac{1}{\Phi} \frac{d^2 \Phi}{d\psi^2} = -m^2$ may be applied. This is a standard second order differential equation with solution (2.26). Physically, the solution must be periodic in both θ and ϕ , hence m must be an integer.

$$\Phi = e^{-jm\phi} \tag{2.26}$$

Substituting this result and dividing by $\sin^2 \theta$ yields

$$\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + \frac{1}{\Theta\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right) - \left(\frac{m}{\sin\theta}\right)^2 + (kr)^2 = 0. \tag{2.27}$$

The θ -dependence is restricted to the second and third terms and these terms may also be replaced by a constant, but the form of the constant is not as clear as before. For now, it is denoted by c, with $\frac{1}{\Theta \sin \theta} \frac{d}{d\theta} (\sin \theta \frac{d\Theta}{d\theta}) - (\frac{m}{\sin \theta})^2 = -c$. Equation (2.24) has been reduced to the following three differential equations:

$$\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + \left[(kr)^2 - c\right]R = 0 \tag{2.28}$$

$$\frac{1}{\sin\theta} \frac{d}{d\theta} \left(\sin\theta \frac{d\Theta}{d\theta} \right) + \left[c - \left(\frac{m}{\sin\theta} \right)^2 \right] \Theta = 0 \tag{2.29}$$

$$\frac{d^2\Phi}{d\psi^2} + \Phi m^2 = 0 \tag{2.30}$$

The derivation is interrupted to introduce two well-known differential equations. Equation (2.31) is the standard form of Bessel's equation of order α , where α is an arbitrary complex number. The cases $\alpha^2 = n(n+1)$ and $\alpha^2 = (n+1/2)^2$ are called Bessel's equations of integer and half-integer order respectively, and are of particular interest in various physical systems. The solutions of Bessel's equation are called Bessel functions.

$$x^{2}\frac{d^{2}y}{dx^{2}} + x\frac{dy}{dx} + (x^{2} - \alpha^{2})y = 0$$
 (2.31)

Another well-known differential equation is shown in (2.32) and is called *Legendre's* equation, where γ and ν may be complex and are called the degree and order of the function, respectively. When both γ and ν are integers, the solutions of (2.32) are called the associated Legendre polynomials, while the solutions for $\nu = 0$ are called the ordinary Legendre polynomials.

$$(1 - x^2)\frac{d^2y}{dx^2} - 2x\frac{dy}{dx} + \left[\gamma(\gamma + 1) - \frac{\nu^2}{1 - x^2}\right]y = 0$$
 (2.32)

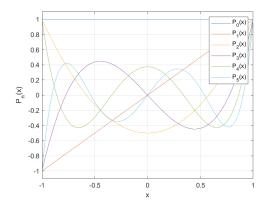
It would be convenient for the unsolved differential equations in (2.28) and (2.29) to be of the above forms with known solutions. To that end, the transformation of variables $\eta = \cos \theta$, and consequently $d\theta = -d\eta/\sin \theta$ and $\sin^2 \theta = 1 - \eta^2$ is applied to (2.29) to obtain

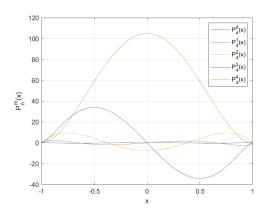
$$(1 - \eta^2)\frac{d^2\Theta}{d\eta^2} - 2\eta\frac{d\Theta}{d\theta} + \left[c - \frac{m^2}{1 - \eta^2}\right]\Theta = 0.$$
 (2.33)

This is precisely in the form of (2.32) if c takes the in the form $\gamma(\gamma + 1)$. The set of solutions to Legendre's equation which exhibit the periodic characteristics required by the physical setup of EM field propagation are called the associated Legendre polynomials, with both γ and ν integers. Hence c = n(n+1) is defined where n is an integer:

$$(1 - \eta^2) \frac{d^2 \Theta}{d\eta^2} - 2\eta \frac{d\Theta}{d\theta} + \left[n(n+1) - \frac{m^2}{1 - \eta^2} \right] \Theta = 0.$$
 (2.34)

The solutions to (2.34) are the associated Legendre polynomials of the first kind, denoted by P_n^m , with n being any positive integer and $|m| \le n$. The associated Legendre polynomials are related to the regular Legendre polynomials, P_n in (2.35), by (2.36), where P_n are the solutions to (2.34) for m = 0. For m = 0, the associated Legendre polynomials reduce to the regular Legendre Polynomials and for m > n, $P_n^m = 0$. Legendre functions for m < 0 are related to those for m > 0 by (2.37). Figure 2.2a depicts the regular Legendre polynomials for n = 0, 1, 2, 3, 4.





- (a) The regular Legendre polynomials for n = 0, 1, 2, 3, 4.
- (b) The associate Legendre polynomials for n = 4 and m = 0, 1, 2, 3, 4.

Figure 2.2: The Legendre polynomials represent the radial variation in SWE, but the unnormalised associated Legendre polynomials pose a risk of numerical errors due to the rapid increase in the functional values as m and n increase.

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} x^n$$
 (2.35)

$$P_n^m(x) = (-1)^m (1 - x^2)^{m/2} \frac{d^m}{dx^m} P_n(x)$$
(2.36)

$$P_n^{-m}(x) = (-1)^m \frac{(n-m)!}{(n+m)!} P_n^m(x)$$
(2.37)

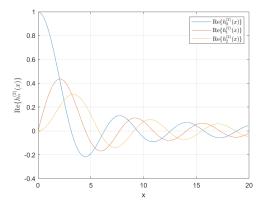
The solution of (2.17) is not unique, which allows two further restrictions to be placed on the solutions of (2.28)-(2.30), [11]. The inclusion of negative values of m would require additional scaling and unnecessarily increase the complexity of the implementation. Hence the associated Legendre polynomials are restricted to those with positive order m by replacing P_n^m with $P_n^{|m|}$. Figure 2.2b shows that the values of the associated Legendre polynomials quickly become very large as n increases. Hence these functions are normalised as in (2.38) to avoid numerical errors. Finally, a solution to (2.29) is given in (2.39).

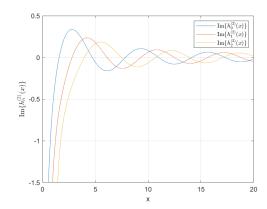
$$\bar{P}_n^m = (-1)^m \sqrt{\frac{2n+1}{2} \frac{(n-m)!}{(n+m)!}} P_n^m(\cos \theta)$$
 (2.38)

$$\Theta(\theta) = \bar{P}_n^{|m|}(\cos \theta) \tag{2.39}$$

The form of constant c was determined as c = n(n+1), with n an integer. This is included in (2.28) and rewritten as

$$\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + \left[(kr)^2 - n(n+1)\right]R = 0,$$
(2.40)





(a) The real part of the spherical Hankel function of the second kind for n = 0, 1, 2.

(b) The imaginary part of the spherical Hankel function of the second kind for n = 0, 1, 2.

Figure 2.3: The spherical Hankel function of the second kind is a complex-valued function which represents outward traveling waves.

and applying the chain-rule for differentiation yields

$$r^{2}\frac{d^{2}R}{dr^{2}} + 2r\frac{dR}{dr} + \left[(kr)^{2} - n(n+1) \right] R = 0.$$
 (2.41)

Further, division by r^2 and the substitution $R(r) = r^{-1/2}\tilde{R}(r)$ from [12] yields

$$\frac{d^2\tilde{R}}{dr^2} + \frac{1}{r}\frac{d\tilde{R}}{dr} + \left[k^2 - \frac{(n + \frac{1}{2})^2}{r^2}\right]\tilde{R} = 0.$$
 (2.42)

A change of variables r = kr and multiplication by $(kr)^2$ yields

$$(kr)^2 \frac{d^2 \tilde{R}(kr)}{d(kr)^2} + (kr) \frac{d\tilde{R}(kr)}{d(kr)} + \left[(kr)^2 - \left(n + \frac{1}{2} \right)^2 \right] \tilde{R}(kr) = 0$$
 (2.43)

and renders the equation recognisable as Bessel's equation of half-integer order in (2.31). The solutions to (2.42) are given in [12] as

$$\tilde{R}(r) = Ar^{1/2}z_n(kr) = BZ_{n+1/2}(kr),$$
(2.44)

where $z_n(kr)$ are the spherical Bessel functions which are related to the ordinary Bessel functions $Z_{n+1/2}$ by

$$z_n(kr) = \sqrt{\frac{\pi}{2kr}} Z_{n+1/2}(kr). \tag{2.45}$$

There are four types of spherical Bessel functions. Those of the first and second kind represent standing waves, while the third kind represents inward traveling waves and the fourth kind represents outward traveling waves. Those of the third and fourth kind are also called spherical Hankel functions of the first and second kind. Clearly the spherical Bessel function of the fourth kind (spherical Hankel function of the second kind) is applicable to

radiating elements. Figure 2.3 depicts the real and imaginary parts of the spherical hankel functions of the second kind for n = 0, 1, 2, and the solution to (2.28) is

$$R(r) = z_n^{(4)}(kr) = h_n^{(2)}(kr). (2.46)$$

The product of (2.26), (2.39) and (2.46) is

$$\psi_{nm}(r,\theta,\phi) = R(r)\Theta(\theta)\Phi(\phi) = h_n^{(2)}(kr)\bar{P}_n^{|m|}(\cos\theta)e^{-jm\phi}, \qquad (2.47)$$

which is a solution for (2.17). This solution is in a the form of a generating function which produces a unique function for each nm-pair, and is called a mode of the SWE.

Practically, the infinite series of possible modes must be truncated. Commercial implementations such as GRASP [13] define the upper bound, N, for n as

$$N = kr_0 + \max\{10, 3.6\sqrt[3]{kr_0}\},\tag{2.48}$$

where r_0 is the radius of the minimum sphere entirely enclosing the radiating element.

Equations (2.18)-(2.21) and the accompanying discussion stated that the ψ was in fact not a generating function for the E-field, but a generating function for the potential field over distance. The application of (2.22) and (2.23) to (2.47) yields

$$\mathbf{m}_{nm} = -h_n^{(2)}(kr) \frac{jm\bar{P}_n^{|m|}(\cos\theta)}{\sin\theta} e^{-jm\phi} \hat{\boldsymbol{\theta}}$$

$$-h_n^{(2)}(kr) \frac{d\bar{P}_n^{|m|}(\cos\theta)}{d\theta} e^{-jm\phi} \hat{\boldsymbol{\phi}}$$
(2.49)

and

$$\mathbf{n}_{nm} = \frac{n(n+1)}{kr} h_n^{(2)}(kr) \bar{P}_n^{|m|}(\cos\theta) e^{-jm\phi} \hat{\boldsymbol{r}}$$

$$+ \frac{1}{kr} \frac{d}{d(kr)} \{kr h_n^{(2)}(kr)\} \frac{d\bar{P}_n^{|m|}(\cos\theta)}{d\theta} e^{-jm\phi} \hat{\boldsymbol{\theta}}$$

$$- \frac{1}{kr} \frac{d}{d(kr)} \{kr h_n^{(2)}(kr)\} \frac{jm\bar{P}_n^{|m|}(\cos\theta)}{\sin\theta} e^{-jm\phi} \hat{\boldsymbol{\phi}}$$
(2.50)

which represent generating functions for the TE and TM modes respectively.

From [13], there are two popular representations of the SWE functions, called the Q-modes and the ab-modes. The Q-modes are preferred here, and the expansion functions are written as

$$\mathbf{F}_{1mn} = c_{mn} \mathbf{m}_{mn} \tag{2.51}$$

and

$$\mathbf{F}_{2mn} = c_{mn} \mathbf{n}_{mn},\tag{2.52}$$

where,

$$c_{mn} = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{n(n+1)}} \left(-\frac{m}{|m|} \right)^m \tag{2.53}$$

is a scaling factor which accounts for the normalisation of the associated Legendre functions. The E-field may then be written as

$$\mathbf{E}(r,\theta,\phi) = \mathbf{E}^{TE}(r,\theta,\phi) + \mathbf{E}^{TM}(r,\theta,\phi)$$

$$= k\sqrt{\zeta} \sum_{n=1}^{n=\infty} \sum_{m=-n}^{m=n} \left[Q_{1mn} \mathbf{F}_{1mn}(r,\theta,\phi) + Q_{2mn} \mathbf{F}_{2mn}(r,\theta,\phi) \right],$$
(2.54)

where Q_{1mn} and Q_{2mn} are the coefficients which must be determined during the expansion process, and $\zeta \approx 377\Omega$ is the intrinsic impedance of free space. The product $k\sqrt{\zeta}$ transforms the unit of the right-hand-side of (2.54) to [V/m]. The set of functions \mathbf{F}_{1mn} and \mathbf{F}_{2mn} for all possible mn-combinations forms a complete orthogonal basis for propagating waves in 3-D space. Any propagating wave has a unique expression of the form in (2.54).

In order to determine the values of the SWE coefficients, the E-field must be known on the surface of some sphere encompassing all radiating parts of the antenna. Further, the number of modes included in the SWE are truncated according to (2.48). If the result of this truncation is x nm-modes, the E-field must be known at a minimum of x points of the sphere surrounding the radiating element and the values of the SWE basis functions must be computed at these x points. This combination of known basis function and E-field values at x points yields a system of linear equations which is easily solvable to produce the SWE coefficients.

2.5. Conclusion

This chapter began with a brief discussion of computational electromagenetics and used (2.9) as the starting point for the derivation of the SW-PW method, which is written in general from in (2.12) and adjusted for far-fields in (2.14). A generating function for the basis functions of the SWE representation of propagating waves was derived from the homogeneous wave equation in spherical coordinates in (2.17) and written in its final form in (2.49) and (2.50). The theoretical description in this chapter is used to implement the SW-PW method in Chapter 3.

Chapter 3

Implementation and results

3.1. Introduction

The practical implementation of the SW-PW method to determine EEPs in linear arrays of $\lambda/2$ -dipole antennas is discussed in this chapter. The predicted EEPs for two-element arrays and uniformly spaced arrays with a larger number of elements are validated by comparing them to full-wave simulation results obtained from FEKO. Without loss of generality, all antennas are \hat{z} -directed and oriented along the x-axis in the xy-plane and all simulations are conducted at a frequency of 1 GHz.

The input far-field radiation and scattering data is obtained from FEKO. The SW-PW method may be implemented in any scripting language, and MATLAB is selected for use here.

3.2. Implementation

The mathematical formulation of the SW-PW method requires near-field values to be used to determine the coupling and scattering effects between adjacent antennas. Far away from the radiation element, the radiation pattern and the propagation factor become separable, as in $\mathbf{E}(\hat{r}) = \frac{1}{r}e^{-jkr}\mathbf{f}(\hat{r})$. It is convenient to use far-field data to represent radiation characteristics, because the absence of radially dependent field-components means that that only angular variation in field strength must be represented and hence it can be represented by a smaller data set. The far-field data does not exhibit all the evanescent mode information of the near-field, but it will be shown it contains enough of this information to recover near-field values at distances close to the minimum sphere for $\lambda/2$ -dipole antennas. As a result, only arrays with inter-element separations larger than this smallest distance at which the near-field can be reconstructed, can be evaluated.

Far-field data is used to model the coupling between antenna elements, which is done by projecting the far-field data onto a set of mutually orthogonal basis functions which may be used to reconstruct the field at (almost) any point in space. To determine the field scattered by an element m due to a field radiated by an element p, the radiated field must be reconstructed on the minimum sphere of element m. This reconstruction may be used to determine the field scattered by element m, as detailed in Section 2.3.2. This was done in [4] and two conclusions were reached: Firstly, that this process becomes intractable for large arrays. Secondly, that similar results may be achieved by reconstructing the field only at the feed point of element m and then using simulated scattering data to obtain the field scattered by m. This scattered field data is obtained from a full-wave simulation package by impinging plane waves with unit amplitude on the element from various directions. From the linearity of Maxwell's equations, this scattering data may be scaled by the amplitude and phase of the reconstructed E-field at the feed point of element m to determine it's scattered field due to the field incident from element p. This computationally more efficient method is used in the present implementation.

3.2.1. The input data

The required input data includes the radiation and scattering data from each element. Since the arrays consist of identical elements, this constitutes a single radiation pattern and a set of scattering patterns due to incident unit amplitude plane waves from different directions. The isolated element is loaded with $Z_0 = 50~\Omega$ when it is active and terminated in the same impedance when it is inactive. FEKO may be used to determine the following results: (i) The radiated far-field for an element excited by a $1 \angle 0^{\circ}V$ source; and (ii) The set of scattered far-fields for a terminated element excited by $1 \angle 0^{\circ}V/m$, $\hat{\theta}_{m}$ - and $\hat{\phi}_{m}$ -polarised incident plane waves from multiple directions.

3.2.2. The antenna type

All antennas are \hat{z} -directed $\lambda/2$ -dipole antennas oriented along the x-axis in the xy-plane. This choice of antenna type simplifies the implementation in two ways. Firstly, an incoming wave which is polarised perpendicular to the orientation of the antenna will induce zero current on the antenna and hence not have a scattering effect, which means that only \hat{z} -directed E-field components need to be considered when determining scattered fields. Secondly, the radiation from dipole antennas is isotropic in the H-plane and their structure is rotationally symmetric, which means that the angle of incidence has no effect on the antenna's scattering pattern. Consequently, the set of scattered far-fields reduces to a single scattered far-field.

3.2.3. Error metrics

The RMSE over angle is used to compare the H-plane directicity of the SW-PW predicted EEPs to those calculated via the MoM in FEKO. In

$$RMSE = \sqrt{\frac{1}{N_{\Omega}} \sum_{n=1}^{N_{\Omega}} |\mathbf{f}_D^{SW-PW} - \mathbf{f}_D^{FEKO}|^2}$$
(3.1)

 \mathbf{f}_D^{SW-PW} is the SW-PW solution \mathbf{f}_D^{FEKO} is the FEKO solution and N_{Ω} is the number of angular points in the representation.

Further, the predicted and reference directivities are visually compared to check for inaccurately modeled isolated spikes or sidelobes which may be misrepresented by the RMSE metric due to the division term. The accuracy of the SWE is analysed by directly comparing the magnitude and phase components of the reconstructed E-field to the near-field values exported from FEKO.

3.2.4. Convergence

When the Jacobi method is applied to the MoM-matrix, its convergence is analysed in a straightforward manner [5]. In the SW-PW method, theoretical convergence analysis is not so simple and the convergence analysis herein follows [5]. The weights of incident field components at each element, written as \mathbf{w}_{pm} in (2.14), are updated iteratively and assembled into a matrix \mathbf{W} . Convergence is reached when the relative difference between \mathbf{W} in the current and previous iterations becomes negligible. This is written as

$$\rho^{(k)} = \frac{||\mathbf{W}^{(k)} - \mathbf{W}^{(k-1)}||_2}{||\mathbf{W}^{(k)}||_2} < \delta, \tag{3.2}$$

where $\rho^{(k)}$ is the relative difference at iteration k and $\delta = 0.001$ is used as the stopping criterion. A maximum number of 100 iterations is imposed on each EEP calculation and calculations typically required in the range of 15 iterations to reach convergence.

The EEP calculation for a nine-element array with elements spaced 0.5λ apart diverges when the Jacobi iterative scheme is used, but the solution converges when the Gauss-Seidel iterative scheme is used. This is shown in Figures 3.1a and 3.1b, where the yellow line in Figure 3.1a shows the divergent case. The relative difference on the y-axis is plotted on a logarithmic scale, so the plots show exponential convergence rates. The convergence rate of the Gauss-Seidel method is superior to that of the Jacobi method.

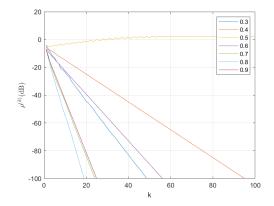
In the Jacobi scheme, all calculations in iteration k are done using the results in iteration k-1. The Gauss-Seidel scheme is adopted instead, wherein the already-obtained iteration k results use used in the remaining calculations of iteration k. This is written as

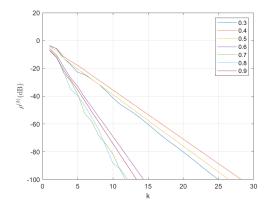
$$\mathbf{f}_{p}^{(k)} = \mathbf{f}_{p}^{(0)} - \sum_{m=1, m \neq p}^{M} \mathbf{f}_{pm}^{s} \mathbf{w}_{pm}^{(k_s)}, \tag{3.3}$$

where

$$k_s = \begin{cases} k, & m p. \end{cases}$$

$$(3.4)$$





- (a) Jacobi iterative scheme (2.14). Diverges for 0.5λ spacing.
- **(b)** Gauss-Seidel iterative scheme (3.3). Exhibits an improved convergence rate.

Figure 3.1: Convergence plots for the Jacobi and Gauss-Seidel iterative schemes for an eqi-spaced linear array of nine elements. The logarithm of the relative difference, ρ , is plotted against the iteration number, k, when calculating the EEP of the centre element for various inter-element spacings shown in the legend in terms of λ . The convergence rate and robustness is improved by adopting the Gauss-Seidel scheme.

3.2.5. Software implementation and dependencies

The present SW-PW immplementation relies on two MATLAB classes provided by the EMACS research group at Stellenbosch University. The first class enables the representation and manipulation of far-fields as objects in MATLAB, and the second class performs the SWE and reconstruction of far-fields.

The SW-PW function receives a matrix containing antenna positions and excitations and returns a cell array containing all embedded element patterns and the corresponding array pattern. The only other initial information is the set of imported radiated and scattered far-fields.

For each active element p, the radiated field is expanded into its SWE modes and reconstructed at every other element m and the imported scattering information is then used to find the field scattered by each element m. These scattered fields are in turn reconstructed at every other other element to determine subsequent scattering effects. This process is repeated until the stopping criterion in (3.2) is met and the result is the active element's EEP. This process is repeated for every element in the array and the array pattern is determined as the superposition of scaled versions of these EEPs.

3.3. Results

3.3.1. Spherical wave expansion

When the SWE coefficients are determined from sampled near-field data on a sphere slightly larger than the minimum sphere, the accurate reconstruction of near-field data at

any point outside the sampled sphere may be expected. Herein, the SWE coefficients are determined from far-field rather than near-field data.

The use of far-field data provides a simplification in the implementation of SWE. The far-field is formally defined at $r \to \infty$ which allows for the spherical Hankel function to be replaced by its asymptotic approximation,

$$h^{(2)}(kr) \to j^{n+1} \frac{e^{-jkr}}{kr}.$$
 (3.5)

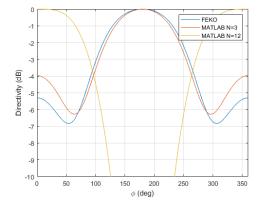
Since the far-field data does not exhibit all evanescent field information, it cannot reasonably be expected to accurately reconstruct the near-field close to the radiating element. The accuracy of the near-field reconstruction from far-field data is analysed in Table 3.1 by comparing the reconstructed values to those determined in FEKO. These results consider a single dipole antenna driven by a 1V source. The percentage error reported is for the magnitude of the reconstruction.

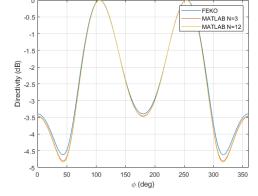
Table 3.1: SWE reconstructed near-field values for different numbers of modes compared to near-field values determined in FEKO. Reported errors are for magnitude reconstruction.

	FE	КО	MATLAB $N=3$			MATLAB $N = 12$		
Dist $[\lambda]$	$ E_z \left[\frac{V}{m}\right]$	$\angle E_z$ [°]	$ E_z \left[\frac{V}{m}\right]$	$\angle E_z[^{\circ}]$	error [%]	$ E_z \left[\frac{V}{m}\right]$	$\angle E_z$ [°]	error $[\%]$
0.25	4.306	121.1	3.835	135.3	10.9	8.448	101.5	96.2
0.30	3.885	107.2	3.627	111.3	6.7	4.195	105.3	8.0
0.35	3.525	92.6	3.398	93.8	3.6	3.562	92.5	1.1
0.40	3.215	77.4	3.152	77.7	1.9	3.221	77.4	0.2
0.45	2.947	61.8	2.916	61.8	1.1	2.949	61.8	0.04
0.50	2.716	45.8	2.701	45.7	0.6	2.716	45.8	0.01

The physical significance of larger numbers of m and n is that they represent faster angular oscillations in the basis functions. In the theoretical development of the SWE, an upper bound for the number of modes to include in the truncated series is given in (2.48), however, this suggestion is made with the view of determining the SWE coefficients from sampled near-field data where the fast-oscillating evanescent field effects are prominent. When far-field data is used to determine the coefficients and higher-order modes are included, is has the effect trying to model fast-oscillations from slowly oscillating data. For this reason, the near-field reconstruction improves close to the minimum sphere when a smaller than suggested maximum index N is used. Table 3.1 also shows that the reconstruction accuracy barely decreases further away from the minimum sphere for this smaller N. Figure 3.2 shows that better approximations of EEPs for elements placed close to each other's minimum spheres may be obtained for N = 3 than for N = 12, and a small price is paid in terms of accuracy at larger distances from the minimum sphere.

Further, the SWE and near-field reconstruction calculations are the most computationally expensive parts of the SW-PW method and by using a smaller number of modes, the speed of the implementation is increased significantly.





- (a) The terminated element is placed 0.25λ from the active element. At this small distance, better results are obtained for a smaller value of N.
- (b) The terminated element is placed 0.5λ from the active element. The accuracy sacrifice to use a smaller N is negligible.

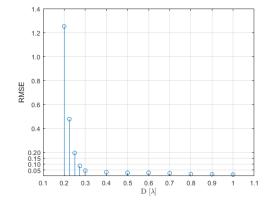
Figure 3.2: Plots of the normalised H-plane directivity of the EEP of an active element with a single terminated passive element placed at different separation distances. The EEPs determined for N=3 and N=12 using the SW-PW method are compared to those exported from FEKO.

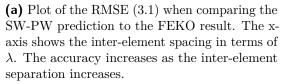
3.3.2. Two-element test arrays

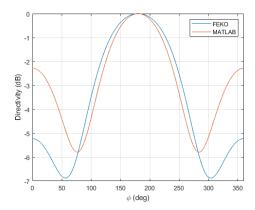
The SW-PW method is used to predict the EEPs in two-element arrays with varying inter-element spacings to confirm that the present implementation works. The labels "MATLAB" and "FEKO" are used to denote the EEPs determined using the SW-PW method in MATLAB and those determined in FEKO, respectively, and all plots depict the normalised H-plane directivity. Figure 3.3a shows the RMSE when comparing the EEPs determined in MATLAB to those imported from FEKO. The error decreases rapidly outside the minimum sphere and the results become almost identical as the inter-element separation grows. This was expected from the accuracy of the SWE reconstruction of the near-field in Table 3.1. Figure 3.3b shows an example of an error which would be unacceptable for the application in Chapter 4. Figure 3.2b depicts an example of an almost perfectly predicted EEP.

3.3.3. Uniformly spaced linear arrays

The results in Section 3.3.2 impose a clear restriction on the application of the SW-PW method in that the results become less reliable as the inter-element spacings approach the minimum sphere. This section considers linear uniformly spaced arrays of a larger number





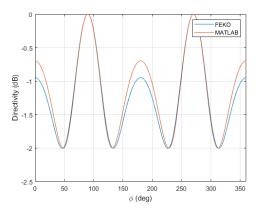


(b) Plot comparing the SW-PW prediction and FEKO results of the EEP of a two-element array with inter-element spacing of 0.225λ . This is an example of an error that is unacceptable for our applications.

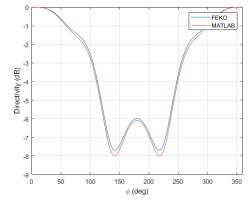
Figure 3.3: Plots depicting the error in the prediction of EEPs in two-element arrays.

of elements. The prediction accuracy is expected to decrease at least slightly due to the increase in mutual coupling effects.

For two-element arrays, EEPs may be predicted to within a RMSE of 0.05 when the elements are 0.3λ or further apart. To account for the increase in mutual coupling, the separation is increased to 0.35λ and it is observed that the accuracy does not decrease when the number of elements in the array increases. Figure 3.4 depicts the EEPs for a three-element array with inter-element spacings of 0.35λ . The MATLAB prediction of the EEP of the centre element is less than 0.3 dB out in some of the turning points, and accurate everywhere else. The left-most EEP is almost perfectly predicted.



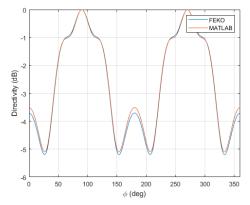
(a) Plot of the H-plane directivity for the centre element of the array, with RMSE=0.0468. This is only marginally worse than the two-element array of similar spacing.

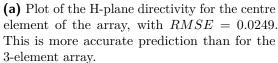


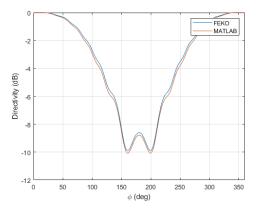
(b) Plot of the H-plane directity of the left-most element of the array, with RMSE=0.0396. The pattern turning points are accurately predicted in MATLAB.

Figure 3.4: Plots comparing the MATLAB and FEKO results for a three-element array with constant spacings of 0.35λ .

Figure 3.5 shows the EEPs for the centre and left-most elements of a 9-element array







(b) Plot of the H-plane directity of the left-most element of the array, with RMSE=0.0393. The MATLAB and FEKO results are almost identical.

Figure 3.5: Plots comparing the MATLAB and FEKO results for a nine-element array with constant spacings of 0.35λ .

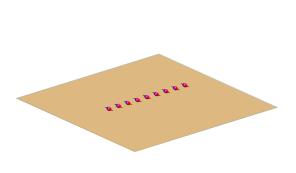
with 0.35λ inter-element spacings. The largest difference between the MATLAB and FEKO results is seen in the turning points of the central element, and this difference is less than 0.3 dB. The EEP of the left-most element is almost perfectly predicted.

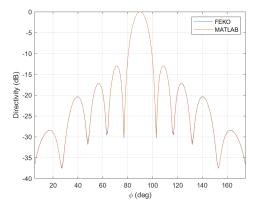
The SW-PW method is not meant to replace full-wave analysis, but rather to supplement it. It is clear that it method may be used to quickly estimate an EEP, but that it cannot make perfect predictions. This section considered inter-element spacings of 0.35λ , and from Sections 3.3.1 and 3.3.2, it may conclude that the predictions become more accurate as the these spacings grow. These predicted EEPs may be useful in certain cases, such as the examples in Chapter 4, but its suitability to other applications must be evaluated according to the requirements of those applications.

3.4. Infinite ground plane model

An infinite ground plane is modelled using image theory, i.e. the structure is analysed without the ground plane and by adding the mirror image of the antennas above the ground plane, below the ground plane [5]. If the ground plane is at y=0 and the array elements are in the positive half-space, the results for $0 \le \phi \le \pi$ are kept and those for $\pi \le \phi \le 2\pi$ are discarded. Practically, this doubles the number of array elements before the SW-PW method is executed and transforms it from a linear to a planar array. Figure 3.6a depicts the FEKO model of a nine-element array of \hat{z} -directed $\lambda/2$ -dipole antennas which are uniformly spaced at $\lambda/2$ apart and $\lambda/4$ in front of the ground plane at y=0, and Figure 3.6b depicts the normalised H-plane directivity for this array when it is uniformly excited with 1V sources. The SW-PW prediction follows the FEKO result closely.

Figure 3.7 depicts the central and right-most elements in the above array's normalised



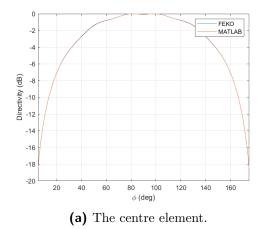


(b) The SW-PW and FEKO H-plane directivity results are almost identical.

(a) The array model from FEKO.

Figure 3.6: The nine-element, $\lambda/2$ uniformly spaced array, $\lambda/4$ in front of an infinite ground plane.

H-plane directivity. The number of elements are doubled before the SW-PW method is applied, and subsequently the fields of each element and its respective mirror image must be combined to obtain the element's radiation pattern in the presence of the other antennas and the ground plane. The SW-PW results show good correspondence with the FEKO simulation, wherein the array was modelled in the presence of a ground plane.



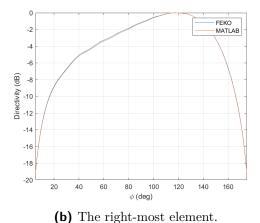


Figure 3.7: Plots of the embedded element normalised H-plane directivity of a nine-element array in front of an infinite ground plane.

3.5. Conclusion

This chapter illuminated all practical aspects regarding implementation of the SW-PW method to $\lambda/2$ -dipole antenna arrays. Figure 3.1 showed that the convergence rate and robustness is improved by adopting the Gauss-Seidel method in place of the Jacobi method. Table 3.1 showed that the the near-field can be reconstructed to within an accuracy of 5% and 1% at distances of 0.35λ and 0.5λ , respectively when using an upper bound of N=3

in the SWE. Figure 3.3a confirmed that the present implementation "works". Figure 3.3b showed that EEPs cannot be accurately predicted when the elements are placed in each other's minimum spheres, and Figure 3.2 showed how the prediction accuracy increases as the elements are placed further apart. Figures 3.4 and 3.5 showed that the prediction accuracy does not decrease significantly when the number of elements in the array increases. Finally, Figure 3.7 showed how the SW-PW method may also be applied to model arrays in the presence of a ground plane. Chapter 4 considers the application of the SW-PW method to the synthesis of array antennas.

Chapter 4

Application

4.1. Introduction

The potential benefit of using a minimum number of antennas to synthesize arbitrary farfield patterns is clear as it can decrease the complexity and cost of communication systems. A framework which combines iterative l_1 -norm minimisation and full-wave simulations to design maximally sparse arrays is proposed in [2]. This chapter considers the addition of an intermediate step into this framework, wherein the SW-PW method is used to compute the EEPs, with the aim of reducing the number of full-wave simulations needed to achieve a maximally spare array. Convex optimisation methods which are required compute the l_1 -norm minimisation in this process, are also discussed.

4.2. The design process

The method in [2] is followed closely in applying a thinned array approach. A densely populated array aperture is initialised with the aim of switching off (removing) as many of the elements as possible while meeting some predefined far-field pattern specifications.

4.2.1. Synthesis excluding mutual coupling

Consider N arbitrarily placed identical elements at positions $\{\mathbf{r}_n\}_{n=1}^N$ and let $\hat{\boldsymbol{r}} = \sin\theta\cos\phi\hat{\boldsymbol{x}} + \sin\theta\sin\phi\hat{\boldsymbol{y}} + \cos\theta\hat{\boldsymbol{z}}$ be the directional unit vector. If the effects of mutual coupling are disregarded and the elements are assumed to be isolated or isotropic, the array far-field pattern may be written as

$$\mathbf{f}(\hat{\mathbf{r}}) = \mathbf{f}_{el}(\hat{\mathbf{r}})AF(\hat{\mathbf{r}}). \tag{4.1}$$

Since the elements all have the same radiation patterns, $\mathbf{f}_{el}(\hat{\boldsymbol{r}})$, the array pattern $\mathbf{f}(\hat{\boldsymbol{r}})$ becomes the product of the element pattern and the array factor given by $AF(\hat{\boldsymbol{r}}) = \sum_{n=1}^{N} w_n e^{jkr_n \cdot \hat{\boldsymbol{r}}}$, where w_n is excitation of element n. The design process then simplifies to the manipulation of the closed-form expression for the array factor.

The effects of mutual coupling are seldom negligible and a more realistic representation of the array pattern is given by

$$\mathbf{f}(\hat{\boldsymbol{r}}) = \sum_{n=1}^{N} w_n \mathbf{f}_n(\hat{\boldsymbol{r}}), \tag{4.2}$$

where $\mathbf{f}_n(\hat{\boldsymbol{r}}) = \mathbf{f}_n^0(\hat{\boldsymbol{r}})e^{jk\boldsymbol{r}_n\cdot\hat{\boldsymbol{r}}}$ is the embedded element pattern of element n, with its phase reference point shifted to the array reference point.

Let $\mathbf{w} = [w_1, w_2, \dots, w_N]^T$, with T indicating the vector transpose, be the N-dimensional excitation vector. The symbol \mathbf{w} is reused here (without subscripts) for continuity with [2]. The far-field \mathbf{f} may be decomposed into its co- and cross-polar components as $\mathbf{f} = f_{co}\hat{\mathbf{co}} + f_{xp}\hat{\mathbf{xp}}$, and the far-field pattern may then be written in vector notation as

$$\mathbf{f}(\hat{\mathbf{r}}) = \left[\mathbf{w}^T \mathbf{f}_{co}(\hat{\mathbf{r}})\right] \hat{\mathbf{co}} + \left[\mathbf{w}^T \mathbf{f}_{xp}(\hat{\mathbf{r}})\right] \hat{\mathbf{xp}}.$$
 (4.3)

Herein, \mathbf{f}_{co} and \mathbf{f}_{xp} are N-dimensional vectors containing the co-and cross-polar components respectively of each of the N elements.

Equation (4.3) represents a convenient matrix expression for the far-field pattern of the entire array. The plan remains to initialise an array with a large number of elements and to eliminate as many of these elements as possible to arrive at a maximally sparse array which fulfills the array pattern specifications. An element may be eliminated when it does not make a contribution to the array pattern. This happens when the element's excitation is zero, i.e. the element is switched off. This aim may be written as

$$\arg\min ||\mathbf{w}||_{l_0},\tag{4.4}$$

where the l_0 -norm gives the number of non-zero entries in its argument vector.

Informally, a non-convex optimisation problem is one which has a global optimal value as well as a locally optimal value or values. This means that ordinary optimisation methods are unlikely to ever find the globally optimal solution. The optimisation of the l_0 -norm is an example of such a problem, which means that global optimisation methods must be invoked in the quest for a solution. These methods are inefficient and become intractable even for moderate arrays.

The minimization of the l_0 -norm may be approximated by an iterative weighted l_1 -norm minimisation, which is a convex optimisation problem that may be solved by conventional methods. If \hat{r}_s is the unit vector in the desired scanning direction, $M(\hat{r})$ is a mask representing the desired SLL and the co-polar far-field component is dominant, then the

 i^{th} iteration of the l_1 -norm minimisation may be written as

arg min
$$||\mathbf{Z}^{(i)}\mathbf{w}^{(i)}||_{l_1}$$

subject to
$$\begin{cases} f_{co}(\hat{\boldsymbol{r}}_s) = 1, \\ |f_{co}(\hat{\boldsymbol{r}})|^2 \leq M(\hat{\boldsymbol{r}}), \hat{\boldsymbol{r}} \in mask \end{cases}$$

$$(4.5)$$

where the p^{th} element of the diagonal matrix $\mathbf{Z}^{(i)}$ is given by

$$z_p^{(i)} = \frac{1}{\left| w_p^{(i-1)} \right| + \epsilon}. (4.6)$$

By multiplying each excitation by this weighting factor based on its value in the previous iteration, the apparent contribution, i.e. the size, of each small value in \mathbf{w} is magnified in the following iteration, leading to its value being suppressed further in the following minimisation iteration. The parameter ϵ is included to avoid numerical errors should a weight in \mathbf{w} be equal to zero and to enable elements which are off in one iteration to be turned on again in the next, if that improves the solution. This process is repeated until a set of active elements is clearly identifiable.

The inactive elements typically have excitations in the range of -200 dB and they may be discarded without affecting the array pattern. The result is a set of element positions and excitations which yields an array pattern conforming to the pattern conditions (beam scanning directing and SLL) as closely as possible.

4.2.2. Synthesis including mutual coupling

In the method above, the excitation vector \mathbf{w} is the optimisation variable and the element patterns \mathbf{f}_{co} are fixed. This means that when the elements are "moved around" the effects of mutual coupling must be re-evaluated and that the determination of the excitation vector and the determination of the mutual coupling effect are two separate steps.

The synthesis process is intialised by attributing to each element its IEP. The procedure above is then followed to determine the positions of an *initial* optimally sparse solution. Once this set of initial active elements has been identified, their positions are used to determine the effects of mutual coupling. In [2], the active element EEPs are determined using full-wave simulations and the passive element EEPs are estimated as phase-shifted versions of the closest active EEPs. The excitation vector minimisation is then repeated using these EEPs, and if a new set of active elements is determined, the EEP calculation and substitution is repeated. If the set of active elements does not change from one iteration to the next, an optimally sparse solution has been obtained. In [2], this method typically converged in the third iteration utilising EEPs. Herein, an intermediate step wherein the SW-PW method is used to determine the EEPs is included with the aim

of further reducing the number of full-wave simulations required during array synthesis. Figure 4.1 shows a block diagram of this process and is similar to Figure 1 in [2], but also includes the intermediate step utilising the SW-PW method.

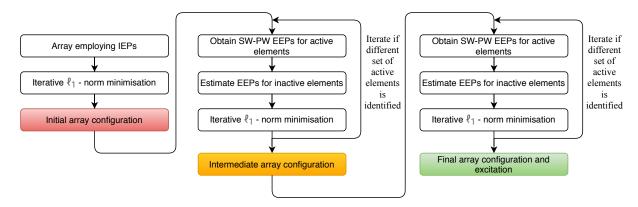


Figure 4.1: Block diagram of the iterative approach including intermediate step employing SW-PW-determined EEPs. This is similar to Figure 1 in [2].

This method allows for the placement of array elements at (almost) arbitrary positions in the array aperture. The resulting array may have elements at any of the quasi-continuous positions chosen in the initialisation step. In [2], $\lambda/100$ inter-element spacings are used, and $\lambda/20$ is used in Section 4.4. The elements are not really being moved around: in reality, elements have been placed at a large number of positions within the array aperture and by keeping the active and removing the inactive elements, the effects of moving them around to find their optimal positions is simulated.

4.3. Convex optimisation

This description of convex optimisation follows [14] closely.

General mathematical optimisation problems are of the form

arg min
$$g_0(\mathbf{x})$$

subject to $g_i(\mathbf{x}) \le b_i, i = 1, \dots, m.$ (4.7)

The vector $\mathbf{x} = [x_1, x_2, \dots, x_N]^T$ is the optimisation variable, g_0 is the objective function, g_i are the constraint functions and constants b_i are the bounds of the constraint functions. An optimal solution of the problem is the vector \mathbf{x} which produces the smallest value for the objective function while satisfying all the constraints.

Optimisation problems and their associated solution algorithms are grouped into different classes. The simplest class of optimisation problems are called linear programs. This includes all problems for which the objective and constraint functions are linear, i.e they satisfy

$$g_i(a\mathbf{x} + b\mathbf{y}) = ag_i(\mathbf{x}) + bg_i(\mathbf{y}) \tag{4.8}$$

for all N-dimensional vectors \mathbf{x} and \mathbf{y} , and all real numbers a and b.

Convex optimisation problems represent a more general description and includes all problems for which the objective and constraint functions satisfy

$$g_i(a\mathbf{x} + b\mathbf{y}) \le ag_i(\mathbf{x}) + bg_i(\mathbf{y}) \tag{4.9}$$

for all N-dimensional vectors \mathbf{x} and \mathbf{y} , and all real numbers a and b with $a+b=1, a\geq 0$ and $b\geq 0$. All linear problems are convex, but not all convex problems are linear.

There are efficient algorithms to solve linear and convex optimisation problems. However, if a problem is non-convex, such as (4.4), the search for a solution must resort to inefficient global optimisation methods or be restricted to locally optimal searches which require an initial *guess* and then searches for an optimal solution in the neighbourhood of this guess. Some non-convex problems may be reformulated as relaxed forms of convex problems, such as the reformulation of (4.4) in (4.5).

The solution algorithms for optimisation problems can be structured as a hierarchy. Quadratic optimisation problems are at the bottom of the hierarchy and are solved by solving a system of linear equations. Unconstrained or equality constrained problems are in the middle of the hierarchy and are solved using Newton's method by reducing them to a sequence of quadratic optimisation problems. Inequality constrained problems are at the top of the hierarchy and require the use of interior points methods to find a solution. These problems are solved by reducing them to a series of unconstrained or equality constrained problems. The problem in (4.5) requires the use of interior point methods.

4.3.1. Interior point methods

A brief discussion of interior point methods is provided here. For a full mathematical treatment, see [14]. Consider the inequality constrained problem in (4.7) which may be written as

arg min
$$g_0(\mathbf{x})$$

subject to $g_i(\mathbf{x}) \le 0, i = 1, ..., m$ (4.10)
 $\mathbf{A}\mathbf{x} = \mathbf{b}$

by incorporating the constraint boundaries b_i into the constraint functions g_i and assuming that this problem has a solution \mathbf{x} that satisfies $\mathbf{A}\mathbf{x} = \mathbf{b}$, where A is an $N \times N$ transformation matrix and \mathbf{b} is the $N \times 1$ constraint vector. The aim is to reformulate this problem as an equality constrained problem to which Newton's method may be applied. To this end, define the indicator function

$$I_{-}(u) = \begin{cases} 0 & u \le 0 \\ \infty & u > 0 \end{cases} \tag{4.11}$$

which allows for the problem to be rewritten as

arg min
$$g_0(\mathbf{x}) + \sum_{i=1}^m I_{-}(g_i(\mathbf{x}))$$

subject to $\mathbf{A}\mathbf{x} = \mathbf{b}$ (4.12)

where the constrains have been included in the objective function. The indicator function I_{-} goes to infinity for values not in the feasible set, i.e. those which do not satisfy the constraint equations $g_i(x) \leq 0$, and hence implicitly includes the constraints in the objective function. However, the objective function is not differentiable because the indicator function is not differentiable and Newton's method are not applicable yet. To apply Newton's method, the objective function must be made differentiable. To this end the indicator function may be approximated by

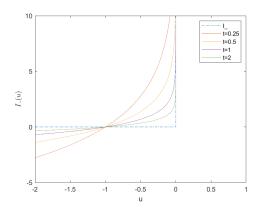
$$\hat{I}_{-}(u) = -\frac{1}{t}log(-u), \tag{4.13}$$

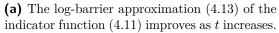
where the value of t > 0 sets the accuracy of the approximation and $\hat{I}_{-}(u)$ is convex, non-decreasing, differentiable and goes to infinity for arguments larger than u. This is called the log-barrier approximation and it is shown in Figure 4.2a for increasing values of t. The accuracy of the log-barrier approximation improves as the value of t increases. This approximation guarantees that the solution of (4.12) remains inside the feasible region by making the objective function value very large for solutions outside the feasible region. From the differentiability of $\hat{I}_{-}(u)$, Newton's method may be applied and the problem may be written as

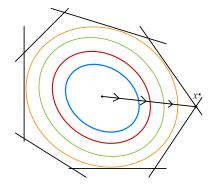
arg min
$$g_0(\mathbf{x}) + \sum_{i=1}^m -\frac{1}{t}\hat{I}_{-}(g_i(x))$$

subject to $\mathbf{A}\mathbf{x} = \mathbf{b}$. (4.14)

This problem is difficult to solve by Newton's method close to the boundary of the feasible region. Hence the problem does not reduce to a single problem solved by Newton's method, but a sequence of these problems. By starting with a small t, the solution is pushed to the centre of the feasible region. This solution is then used as a seed value the following iteration where the value of t is increased to allow the exploration of a larger part of the feasible region. Figure 4.2b shows how increasing parts of the feasible set are explored with each iteration and hence how a globally optimal solution is approached through a series of simplified steps. The black lines indicate the boundary of the feasible region and the coloured contours show the increasing regions of exploration. The line with arrows shows how the solution improves as larger parts of the feasible region are investigated and the optimal solution x^* is finally reached.







(b) The contours show how the log-barrier approximation allows for sequential exploration of the feasible set.

Figure 4.2: The log-barrier approximation reduces inter-point method solutions to a sequence of Newton's method solutions.

4.3.2. Newton's method

In the interest of brevity, only a qualitative description of Newton's method is provided here, whereafter the algorithm which is applied to the sequence of equality constrained optimisation problems obtained in the previous section, is stated. For a full mathematical treatment of this method, see [14].

Newton's method for the optimisation of unconstrained problems is similar to Newton's method for finding the roots of a polynomial, and can be extended to equality constrained optimisation problems.

To solve a continuously differentiable equation of the form f(x) = 0, some starting point is picked and f is approximated by a linear function at this point. The root of this approximation then becomes the next estimated point and this process is repeated until a point x for which the f(x) is sufficiently close to zero, is found. The minimum value of some twice continuously differentiable function coincides with the point where its first derivative is zero. Hence the process of finding this minimum value is the same as that of finding the root of f above, except that the objective is to find the root of the function's first derivative and that this may be achieved through approximations based on the function's second derivative. Since the objective function is convex, this solution will be the global optimal value.

Without further discussion, the algorithm from [14] for Newton's method for equality constrained minimisation is stated: given a starting point \mathbf{x} in the feasible region of the objective function and a tolerance δ (the symbol is re-used for continuity with [14]), repeat the following until the stopping condition is met:

- 1. Compute the Newton step Δx_{nt} and the decrement $\lambda(x)$;
- 2. Stop if $\lambda^2/2 \leq \delta$;

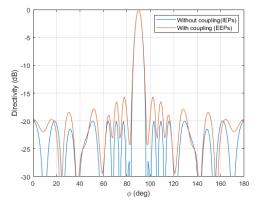
- 3. Choose the step size k by backtracking line search; and
- 4. Update the optimal point $x := x + k\Delta x_{nt}$.

4.4. Array synthesis

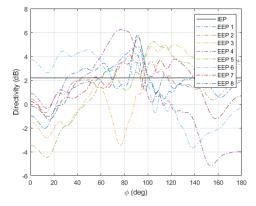
The two design examples in [2] are repeated here with the target SLL increased by 2 dB to compensate for the absence of the PEC ground plane. The degradation of IEP synthesis results is determined using the SW-PW method and the degradation of SW-PW EEP results is determined using FEKO. The convex optimisation MATLAB package from [15] is used to solve (4.5).

4.4.1. Dipole array antenna with 10λ aperture

The design objective is to synthesise a symmetric broadside scanned array with a main beam width of 12°, SLL of -20 dB, and a maximum aperture size of 10λ . The array aperture is initialised by placing antennas with isolated element patterns at 0.05λ intervals along the x-axis and performing the l_1 -norm minimisation on the element excitations. The result of this step is a linear array which satisfies the design requirements in the absence of mutual coupling. Figure 4.3a shows the normalised H-plane directivity of the initial synthesised array pattern both including and excluding the effects of mutual coupling; the mutual coupling in this figure is calculated using the SW-PW method. Sidelobe degradation of 4.4 dB is observed when the mutual coupling effects are included. Figure 4.3b shows a comparison of the IEP used in this iteration as a solid line and the SW-PW EEPs as dashed lines, emphasizing the non-negligible nature of mutual coupling.



(a) The H-plane directivities with and without coupling. The pattern exhibits SLL degradation of 4.4 dB when coupling is included.



(b) Plot comparing the IEP used to determine the design and the EEPs of the active elements. Mutual coupling is non-negligable.

Figure 4.3: Result of the 10λ aperture array, synthesized in the absence of mutual coupling.

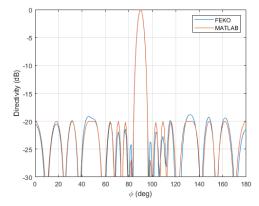
In the second iteration, the IEPs of the active elements are replaced by their respective EEPs calculated using the SW-PW method, and the IEPs of the inactive elements with phase-shifted versions of the nearest active element's EEP. The l_1 -norm minimisation is then repeated and the same set of active elements with further optimised excitations is identified. Since the same set of active elements has been identified for a second consecutive step, an optimally sparse solution has been found.

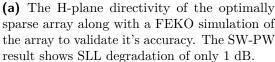
Figure 4.4a shows the array pattern designed using SW-PW EEPs and predicted using the SW-PW method, in orange, and the same array simulated in FEKO, in blue. Minor sidelobe degradation of less that 1 dB is observed when comparing the SW-PW array pattern to the FEKO validation. This is to be expected since the SW-PW is an approximation method and is intended to supplement full-wave simulations rather than replace them. Once an intermediate solution has been obtained in this way, the SW-PW EEPs may be replaced by those calculated using full-wave simulations, to complete the design process. Figure 4.4b shows the array pattern achieved when the SW-PW EEPs are replaced by their FEKO counterparts and a further l_1 -norm minimisation is performed. This result satisfies all design requirements.

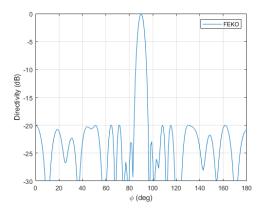
Table 4.1 shows how the active element excitations changed from the first to the final iterations. In this case, the set of active elements did not change from the first iteration, and therefore no time was saved by employing the SW-PW method as an intermediate step. However, from the examples in [2], it is reasonable to expect the set of active elements to change under more demanding SLL or main beam specifications, or if larger apertures are considered. Further, the SW-PW design in Figure 4.4a with minor sidelobe degradation may have been acceptable under certain circumstances, in which case subsequent iteration utilising full-wave EEPs may not have been performed. The sidelobe degradation may have been anticipated a target SLL of -21 dB may have been used in (4.5) to obtain an array configuration with a -20 dB SLL using only the SW-PW method to determine EEPS.

Table 4.1: Element	positions and	complex	excitations	(in Vo	olt) after	each iteration.

	Positions $[\lambda]$							
Iteration	0.4	1.35	2.05	2.4	2.95	3.3	4.2	5
Iter 1	$\begin{array}{ c c c c c c c c c c c c c c c c c c c$	0.239∠− 67.1°	$\begin{vmatrix} 0.114\angle - \\ 67.1^{\circ} \end{vmatrix}$		0.071∠− 67.1°			
Iter 2	$\begin{array}{ c c c c c c c c c c c c c c c c c c c$	0.190∠− 48.7°	$\begin{vmatrix} 0.069 \angle - \\ 82.6^{\circ} \end{vmatrix}$		$\begin{vmatrix} 0.068 \angle - \\ 97.1^{\circ} \end{vmatrix}$			0.066∠− 66.3°
Iter 3	0.189∠- 40.0°	0.191∠− 50.9°	0.096∠− 90.6°		0.054∠− 102.0°			0.063∠− 61.3°







(b) This array pattern is found when performing a subsequent iteration and replacing the SW-PW EEPs with full-wave EEPs. This satisfies all design specifications.

Figure 4.4: Plots depicting the results of the 10λ array synthesis process.

4.4.2. Dipole array antenna with 120λ aperture

For satellite applications, narrow beamwidths and strict sidelobe constraints over the earth's surface are often required. These applications may require antenna apertures with sizes in the range of 100λ . The present design specifications include a beamwidth of 1°, and SLL of -25 dB for 16° from broadside. This is the earth subtending angle as seen from the satellite, and radiation outside this angular range misses the earth.

The same process as before is followed, and Figure 4.5 depicts the initial and intermediate synthesis results. Using EEPs increased the number of active elements from 104 to 106. SLL degradation of 5 dB was observed when the SW-PW method was used to model the coupling in the initial design utilising IEPs. The intermediate design reached a SLL of -24.5 dB using the SW-PW EEPs. Sidelobe degradation may be expected if this array is to be analysed using full-wave simulations, but the purpose of this example is only to show that the SW-PW method may be used to improve the starting point from where full-wave analysis is used in the synthesis process. The SW-PW method predicts the SLL of the initial and intermediate arrays as -20 dB and -24.5 dB respectively. The intermediate result remains an estimation, but it is an estimation which is significantly closer to the end goal than that of the initial solution, hence it has the potential to reduce the number of full-wave simulations needed to achieve a final solution. This intermediate step seems to be more useful in arrays of this size as opposed to the 10λ aperture considered previously.

4.5. Conclusion

A detailed discussion of the framework for the synthesis of maximally sparse antenna arrays in the presence of mutual coupling from [2] was given in Section 4.2 and convex

38

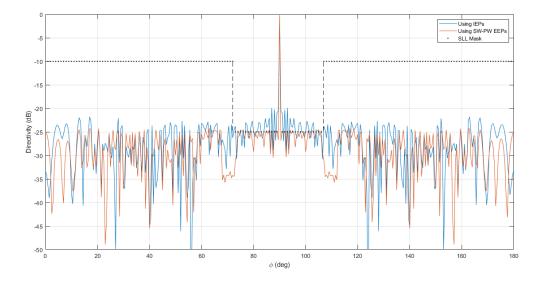


Figure 4.5: Normalised H-plane directivity of the results for the 120λ aperture array when utilising IEPs and SW-PW-determined EEPs.

optimisation was discussed in Section 4.3. Section 4.4 considered the incorporation of the SW-PW method as an intermediate step in the synthesis process outlined previously. The intermediate step provided improved synthesis results as compared to the inital result utilising IEPs, and that these results exhibit only minor SLL degradation when compared to full-wave simulations (see Figure 4.4a). The use of the SW-PW method fits seamlessly into the synthesis procedure in [2] and incorporating it as an intermediate step may further reduce the number of full-wave simulations required during array synthesis.

Chapter 5

Summary and Conclusion

An exposition of the SW-PW method and a derivation of SWE was provided in Chapter 2. Specifically, (2.14) describes the SW-PW method, and (2.49) and (2.50) are the generating functions for the TE and TM modes, respectively, in SWE. This provided the theoretical description needed to implement the SW-PW method in Chapter 3.

The results in Chapter 3, specifically Figures 3.3a, 3.4 and 3.5, and Table 3.1 showed that the SW-PW method was successfully implemented to predict embedded element patterns in coupled antenna arrays and that the use of far-field data for the reconstruction of the near-field is justified in this case. Figure 3.7 showed how the SW-PW method may be applied to arrays in the presence of a ground plane. This implementation does not present a significant decrease in computation time when compared to commercial full-wave simulation packages, due to the simplicity and small number of mesh segments of dipole antennas. However, the computational complexity of the SW-PW method does not increase significantly for more complex antenna structures and time the difference between SW-PW and full-wave calculations will be apparent for more complex antennas such as the log-periodic antennas in [5].

The results in Chapter 4, specifically Figures 4.3a and 4.4 showed that antenna arrays with small apertures may be synthesized using the SW-PW method to model the coupling effects. The use of the SW-PW method as an intermediate step in the synthesis process provides significantly improved results over the initial step using IEPs. This is further illustrated by Figure 4.5 which shows that the starting point from where full-wave simulations are used may be improved by incorporating this intermediate step. This has the potential to further reduce the number of full-wave simulations required during array synthesis, and hence to simplify this involved process.

A more complete version of the SW-PW method was implemented in [5] wherein it is applied to more complex radiating structures such as planar inverted-F and log-periodic antennas. Said implementation also characterised the S-parameters of array elements as opposed to only far-field radiation patterns considered in this project. Inclusion of these elements would render the implementation in the project of greater practical use.

Further improvements suggested in [5] include using near-field data and adding the radial component of the radiated fields to enable the modeling of arrays with smaller inter-element spacings.

In terms of the application of the SW-PW method to practical problems, this project showed that it has great potential when incorporated into synthesis methods. This application would also benefit greatly from the above-mentioned extensions since it would enable its use in the synthesis of a wider range of practical systems. Further work is also required to better characterise the potential reduction in the number of required full-wave simulations should the SW-PW method be incorporated as an intermediate step in the design framework in [2].

Bibliography

- [1] P.-S. Kildal, Foundations of Antenna Engineering. Gothenburg, Sweden: Kildal Antenn AB, 2015, pp. 299–333.
- [2] C. Bencivenni, M. V. Ivashina, R. Maaskant, and J. Wettergren, "Design of Maximally Sparse Antenna Arrays in the Presence of Mutual Coupling," *IEEE Antennas and Wireless Propagation Letters*, vol. 14, pp. 159–162, 2015.
- [3] C. Zhang, X. Fu, S. Peng, and Y. Wang, "Linear unequally spaced array synthesis for sidelobe suppression with different aperture constraints using whale optimization algorithm," in 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), Wuhan, China, June 2018, pp. 69–73.
- [4] D. I. L. de Villiers and R. Maaskant, "Element pattern prediction in mutually-coupled arrays through isolated antenna characterization," in 2017 International Symposium on Antennas and Propagation (ISAP), Phuket, Thailand, October 2017, pp. 1–2.
- [5] T. Marinović, D. I. L. de Villiers, D. J. Bekers, M. N. Johansson, A. Stjernman, R. Maaskant, and G. A. E. Vandenbosch, "Fast Characterization of Mutually-Coupled Array Antennas Using Isolated Antenna Far-Field Data," *IEEE Transactions on Antennas and Propagation*, pp. 1–1, 2020.
- [6] D. B. Davidson, Computational Electromagnetics for RF and Microwave Engineering, 2nd ed. Cambridge: Cambridge University Press, 2011, pp. 1–10.
- [7] D. J. Ludick, M. M. Botha, R. Maaskant, and D. B. Davidson, "Comparison of the iterative jacobi method and the iterative Domain Green's Function Method for finite array analysis," in 2016 10th European Conference on Antennas and Propagation, Davos, Switzerland, April 2016, pp. 1–5.
- [8] R. F. Harrington, *Time-Harmonic Electromagnetic Fields*. Hoboken, New Jersey: John Wiley Sons, Inc., 1961, pp. 106.
- [9] C. A. Balanis, *Advanced Engineering Electromagnetics*. Hoboken, New Jersey: John Wiley Sons, Inc., 1989, pp. 104–125,543–552.
- [10] J. A. Stratton, Electromagnetic Theory. New York: McGraw-Hill, 1941, pp. 392–420.

- [11] T. Marinović, "Vector Spherical Harmonic Expansion," 2019, [Online], Available: https://altairuniversity.com/40633-vector-spherical-harmonic-expansion-tomislav-marinovic/.
- [12] M. N. O. Sadiku, *Numerical Techniques in Electromagnetics*, 2nd ed. Boca Raton, Florida: CRC Press LLC, 2001, pp. 72–73.
- [13] TICRA, TICRA Tools User's Manual, 2020, pp. 333–338.
- [14] S. Boyd and L. Vandenberghe, *Convex Optimisation*. Cambridge: Cambridge University Press, 2004, pp. 1–16, 521–623.
- [15] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," 2014, [Online], Available: http://cvxr.com/cvx.

Appendix A

Project Planning Schedule

Date	Activity
28 Jul - 16 Aug	Background of theory and software (SW-PW method,
	SWE, MATLAB, FEKO).
$17~\mathrm{Aug}$ - $06~\mathrm{Sept}$	Implementation and testing of software.
14 Sept - $20 Sept$	Software Integration and testing.
21 Sept - 11 Oct	Review and application of design methodology.
12 Oct - 02 Nov	Report writing.
09 Nov	Report deadline.
16 Nov - 20 Nov	Skripsie oral.
25 Nov	Skripsie open day.

Appendix B

Outcomes Compliance

ELO	Description	Compliance	Sec. Ref.
1	Problem solving: Identify, for-	This is shown in the theoretical	2.3, 3.2-3
	mulate, analyse and solve complex	description and the implementa-	
	engineering problems creatively	tion of the SW-PW method in	
	and innovatively	Chapters 2 and 3 respectively, and	
		the overcoming of the barriers be-	
		tween the theoretical and practical	
		descriptions.	
2	Application of scientific and	This is shown in the derivation	$2.4,\ 3.2 - 3,\ 4.4$
	engineering knowledge: Apply	of SWE, and the implementation	
	knowledge of mathematics, nat-	of the SW-PW method and array	
	ural sciences, engineering funda-	synthesis techniques.	
	mentals and an engineering spe-		
	ciality to solve complex engineer-		
	ing problems.		
3	Engineering Design: Perform	This is shown in the implemen-	3.2 - 3.3, 4.4
	creative, procedural and non-	tation of the SW-PW method in	
	procedural design and synthesis of	Chapter 3 and in array synthesis	
	components, systems, engineering	in Chapter 4.	
	works, products or processes.		
4	Investigations, experiments	This is shown in the analysis of	3.3, 4.4
	and data analysis: Demon-	SW-PW and array synthesis re-	
	strate competence to design and	sults.	
	conduct investigations and experi-		
	ments.		

Appendix 45

\mathbf{ELO} Description Compliance Sec. Ref. 5 Engineering methods, skills This included using MATLAB and 3.2 and tools, including Informa-FEKO, and the use of approximation Technology: Demonstrate tions to the theoretical description competence to use appropriate of the SW-PW method to practiengineering methods, skills and cally implement it. tools, including those based on information technology. 6 Professional and technical This outcome is satisfied through communication: Demonstrate the report, the presentation and competence to communicate effeccommunication with my supervitively, both orally and in writing, sor. with engineering audiences and the community at large. 8 Individual work: Demonstrate All tests and result analysis was 3.3-4, 4.4-5 competence to work effectively as conducted alone. an individual. 9 Independent Learning Abil-This is shown in the derivation of 2.4, 4.3ity: Demonstrate competence to the SWE and its adaptation to use far-field data to reconstruct engage in independent learning through well-developed learning the near-field, and the analysis of skills. convex optimisation problems.