THE UNIVERSITY OF CALGARY

FORCE-FREE MAGNETIC FIELDS

by

HATEM ZAGHLOUL

A DISSERTATION

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHYSICS AND ASTRONOMY

CALGARY, ALBERTA

DECEMBER 1993

ISBN 0-315-94012-3

Name Hatem Zaghloul

Dissertation Abstracts International is arranged by broad, general subject categories. Please select the one subject which most nearly describes the content of your dissertation. Enter the corresponding four-digit code in the spaces provided.

Physics; General

SUBJECT TERM

0	6	0	5

SUBJECT CODE

U·M·I

Subject Categories

THE HUMANITIES AND SOCIAL SCIENCES

COMMUNICATIONS AND THE ARTS
Architecture 0729
Art History 0377
Cinema 0900
Dance 0378
Fine Arts 0357
Information Science 0723
Journalism 0391
Library Science 0399
Mass Communications 0708
Music 0413
Speech Communication 0459
Theater 0465

EDUCATION
General 0515
Administration 0514
Adult and Continuing 0516
Agricultural 0517
Art 0273
Bilingual and Multicultural 0282
Business 0688
Community College 0275
Curriculum and Instruction 0727
Early Childhood 0518
Elementary 0524
Finance 0277
Guidance and Counseling 0519
Health 0680
Higher 0745
History of 0520
Home Economics 0278
Industrial 0521
Language and Literature 0279
Mathematics 0280
Music 0522
Philosophy of 0998
Physical 0523
Psychology 0525
Reading 0535
Religious 0527
Sciences 0714
Secondary 0533
Social Sciences 0534
Sociology of 0340
Special 0529
Teacher Training 0530
Technology 0710
Tests and Measurements 0288
Vocational 0747

LANGUAGE, LITERATURE AND LINGUISTICS
Language
- General 0679
- Ancient 0289
- Linguistics 0290
- Modern 0291

Literature
- General 0401
- Classical 0294
- Comparative 0295
- Medieval 0297
- Modern 0298
- African 0316
- American 0591
- Asian 0305
- Canadian (English) 0352
- Canadian (French) 0355
- English 0593
- Germanic 0311
- Latin American 0312
- Middle Eastern 0315
- Romance 0313
- Slavic and East European 0314

PHILOSOPHY, RELIGION AND THEOLOGY
Philosophy 0422
Religion
- General 0318
- Biblical Studies 0321
- Clergy 0319
- History of 0320
- Philosophy of 0322

Theology 0469

SOCIAL SCIENCES
American Studies 0323
Anthropology
- Archaeology 0324
- Cultural 0326
- Physical 0327

Business Administration
- General 0310
- Accounting 0272
- Banking 0770
- Management 0454
- Marketing 0338

Canadian Studies 0385
Economics
- General 0501
- Agricultural 0503
- Commerce-Business 0505
- Finance 0508
- History 0509
- Labor 0510
- Theory 0511

Folklore 0358
Geography 0366
Gerontology 0351
History
- General 0578
- Ancient 0579
- Medieval 0581
- Modern 0582
- Black 0328
- African 0331
- Asia, Australia and Oceania 0332
- Canadian 0334
- European 0335
- Latin American 0336
- Middle Eastern 0333
- United States 0337

History of Science 0585
Law 0398
Political Science
- General 0615
- International Law and Relations 0616
- Public Administration 0617

Recreation 0814
Social Work 0452
Sociology
- General 0626
- Criminology and Penology 0627
- Demography 0938
- Ethnic and Racial Studies 0631
- Individual and Family Studies 0628
- Industrial and Labor Relations 0629
- Public and Social Welfare 0630
- Social Structure and Development 0700
- Theory and Methods 0344

Transportation 0709
Urban and Regional Planning 0999
Women's Studies 0453

THE SCIENCES AND ENGINEERING

BIOLOGICAL SCIENCES
Agriculture
- General 0473
- Agronomy 0285
- Animal Culture and Nutrition 0475
- Animal Pathology 0476
- Food Science and Technology 0359
- Forestry and Wildlife 0478
- Plant Culture 0479
- Plant Pathology 0480
- Plant Physiology 0817
- Range Management 0777
- Wood Technology 0746

Biology
- General 0306
- Anatomy 0287
- Biostatistics 0308
- Botany 0309
- Cell 0379
- Ecology 0329
- Entomology 0353
- Genetics 0369
- Limnology 0793
- Microbiology 0410
- Molecular 0307
- Neuroscience 0317
- Oceanography 0416
- Physiology 0433
- Radiation 0821
- Veterinary Science 0778
- Zoology 0472

Biophysics
- General 0786
- Medical 0760

EARTH SCIENCES
Biogeochemistry 0425
Geochemistry 0996
Geodesy 0370
Geology 0372
Geophysics 0373
Hydrology 0388
Mineralogy 0411
Paleobotany 0345
Paleoecology 0426
Paleontology 0418
Paleozoology 0985
Palynology 0427
Physical Geography 0368
Physical Oceanography 0415

HEALTH AND ENVIRONMENTAL SCIENCES
Environmental Sciences 0768
Health Sciences
- General 0566
- Audiology 0300
- Chemotherapy 0992
- Dentistry 0567
- Education 0350
- Hospital Management 0769
- Human Development 0758
- Immunology 0982
- Medicine and Surgery 0564
- Mental Health 0347
- Nursing 0569
- Nutrition 0570
- Obstetrics and Gynecology 0380
- Occupational Health and Therapy 0354
- Ophthalmology 0381
- Pathology 0571
- Pharmacology 0419
- Pharmacy 0572
- Physical Therapy 0382
- Public Health 0573
- Radiology 0574
- Recreation 0575
- Speech Pathology 0460
- Toxicology 0383

Home Economics 0386

PHYSICAL SCIENCES

Pure Sciences
Chemistry
- General 0485
- Agricultural 0749
- Analytical 0486
- Biochemistry 0487
- Inorganic 0488
- Nuclear 0738
- Organic 0490
- Pharmaceutical 0491
- Physical 0494
- Polymer 0495
- Radiation 0754

Mathematics 0405
Physics
- General 0605
- Acoustics 0986
- Astronomy and Astrophysics 0606
- Atmospheric Science 0608
- Atomic 0748
- Electronics and Electricity 0607
- Elementary Particles and High Energy 0798
- Fluid and Plasma 0759
- Molecular 0609
- Nuclear 0610
- Optics 0752
- Radiation 0756
- Solid State 0611

Statistics 0463

Applied Sciences
Applied Mechanics 0346
Computer Science 0984
Engineering
- General 0537
- Aerospace 0538
- Agricultural 0539
- Automotive 0540
- Biomedical 0541
- Chemical 0542
- Civil 0543
- Electronics and Electrical 0544
- Heat and Thermodynamics 0348
- Hydraulic 0545
- Industrial 0546
- Marine 0547
- Materials Science 0794
- Mechanical 0548
- Metallurgy 0743
- Mining 0551
- Nuclear 0552
- Packaging 0549
- Petroleum 0765
- Sanitary and Municipal 0554
- System Science 0790

Geotechnology 0428
Operations Research 0796
Plastics Technology 0795
Textile Technology 0994

PSYCHOLOGY
General 0621
Behavioral 0384
Clinical 0622
Developmental 0620
Experimental 0623
Industrial 0624
Personality 0625
Physiological 0989
Psychobiology 0349
Psychometrics 0632
Social 0451

Nom ____________________

Dissertation Abstracts International est organisé en catégories de sujets. Veuillez s.v.p. choisir le sujet qui décrit le mieux votre thèse et inscrivez le code numérique approprié dans l'espace réservé ci-dessous.

SUJET ____________________ CODE DE SUJET [ ][ ][ ][ ] U·M·I

Catégories par sujets

HUMANITÉS ET SCIENCES SOCIALES

COMMUNICATIONS ET LES ARTS

- Architecture 0729
- Beaux-arts 0357
- Bibliothéconomie 0399
- Cinéma 0900
- Communication verbale 0459
- Communications 0708
- Danse 0378
- Histoire de l'art 0377
- Journalisme 0391
- Musique 0413
- Sciences de l'information 0723
- Théâtre 0465

ÉDUCATION

- Généralités 515
- Administration 0514
- Art 0273
- Collèges communautaires 0275
- Commerce 0688
- Économie domestique 0278
- Éducation permanente 0516
- Éducation préscolaire 0518
- Éducation sanitaire 0680
- Enseignement agricole 0517
- Enseignement bilingue et multiculturel 0282
- Enseignement industriel 0521
- Enseignement primaire. 0524
- Enseignement professionnel 0747
- Enseignement religieux 0527
- Enseignement secondaire 0533
- Enseignement spécial 0529
- Enseignement supérieur 0745
- Évaluation 0288
- Finances 0277
- Formation des enseignants 0530
- Histoire de l'éducation 0520
- Langues et littérature 0279
- Lecture 0535
- Mathématiques 0280
- Musique 0522
- Orientation et consultation 0519
- Philosophie de l'éducation 0998
- Physique 0523
- Programmes d'études et enseignement 0727
- Psychologie 0525
- Sciences 0714
- Sciences sociales 0534
- Sociologie de l'éducation 0340
- Technologie 0710

LANGUE, LITTÉRATURE ET LINGUISTIQUE

- Langues
  - Généralités 0679
  - Anciennes 0289
  - Linguistique 0290
  - Modernes 0291
- Littérature
  - Généralités 0401
  - Anciennes 0294
  - Comparée 0295
  - Mediévale 0297
  - Moderne 0298
  - Africaine 0316
  - Américaine 0591
  - Anglaise 0593
  - Asiatique 0305
  - Canadienne (Anglaise) 0352
  - Canadienne (Française) 0355
  - Germanique 0311
  - Latino-américaine 0312
  - Moyen-orientale 0315
  - Romane 0313
  - Slave et est-européenne 0314

PHILOSOPHIE, RELIGION ET THÉOLOGIE

- Philosophie 0422
- Religion
  - Généralités 0318
  - Clergé 0319
  - Études bibliques 0321
  - Histoire des religions 0320
  - Philosophie de la religion 0322
- Théologie 0469

SCIENCES SOCIALES

- Anthropologie
  - Archéologie 0324
  - Culturelle 0326
  - Physique 0327
- Droit 0398
- Économie
  - Généralités 0501
  - Commerce-Affaires 0505
  - Économie agricole 0503
  - Économie du travail 0510
  - Finances 0508
  - Histoire 0509
  - Théorie 0511
- Études américaines 0323
- Études canadiennes 0385
- Études féministes 0453
- Folklore 0358
- Géographie 0366
- Gérontologie 0351
- Gestion des affaires
  - Généralités 0310
  - Administration 0454
  - Banques 0770
  - Comptabilité 0272
  - Marketing 0338
- Histoire
  - Histoire générale 0578
  - Ancienne 0579
  - Médiévale 0581
  - Moderne 0582
  - Histoire des noirs 0328
  - Africaine 0331
  - Canadienne 0334
  - États-Unis 0337
  - Européenne 0335
  - Moyen-orientale 0333
  - Latino-américaine 0336
  - Asie, Australie et Océanie 0332
- Histoire des sciences 0585
- Loisirs 0814
- Planification urbaine et régionale 0999
- Science politique
  - Généralités 0615
  - Administration publique 0617
  - Droit et relations internationales 0616
- Sociologie
  - Généralités 0626
  - Aide et bien-àtre social 0630
  - Criminologie et établissements pénitentiaires 0627
  - Démographie 0938
  - Études de l' individu et de la famille 0628
  - Études des relations interethniques et des relations raciales 0631
  - Structure et développement social 0700
  - Théorie et méthodes 0344
  - Travail et relations industrielles 0629
- Transports 0709
- Travail social 0452

SCIENCES ET INGÉNIERIE

SCIENCES BIOLOGIQUES

- Agriculture
  - Généralités 0473
  - Agronomie. 0285
  - Alimentation et technologie alimentaire 0359
  - Culture 0479
  - Élevage et alimentation 0475
  - Exploitation des péturages 0777
  - Pathologie animale 0476
  - Pathologie végétale 0480
  - Physiologie végétale 0817
  - Sylviculture et faune 0478
  - Technologie du bois 0746
- Biologie
  - Généralités 0306
  - Anatomie 0287
  - Biologie (Statistiques) 0308
  - Biologie moléculaire 0307
  - Botanique 0309
  - Cellule 0379
  - Écologie 0329
  - Entomologie 0353
  - Génétique 0369
  - Limnologie 0793
  - Microbiologie 0410
  - Neurologie 0317
  - Océanographie 0416
  - Physiologie 0433
  - Radiation 0821
  - Science vétérinaire 0778
  - Zoologie 0472
- Biophysique
  - Généralités 0786
  - Medicale 0760

SCIENCES DE LA TERRE

- Biogéochimie 0425
- Géochimie 0996
- Géodésie 0370
- Géographie physique 0368
- Géologie 0372
- Géophysique 0373
- Hydrologie 0388
- Minéralogie 0411
- Océanographie physique 0415
- Paléobotanique 0345
- Paléoécologie 0426
- Paléontologie 0418
- Paléozoologie 0985
- Palynologie 0427

SCIENCES DE LA SANTÉ ET DE L'ENVIRONNEMENT

- Économie domestique 0386
- Sciences de l'environnement 0768
- Sciences de la santé
  - Généralités 0566
  - Administration des hipitaux 0769
  - Alimentation et nutrition 0570
  - Audiologie 0300
  - Chimiothérapie 0992
  - Dentisterie 0567
  - Développement humain 0758
  - Enseignement 0350
  - Immunologie 0982
  - Loisirs 0575
  - Médecine du travail et thérapie 0354
  - Médecine et chirurgie 0564
  - Obstétrique et gynécologie 0380
  - Ophtalmologie 0381
  - Orthophonie 0460
  - Pathologie 0571
  - Pharmacie 0572
  - Pharmacologie 0419
  - Physiothérapie 0382
  - Radiologie 0574
  - Santé mentale 0347
  - Santé publique 0573
  - Soins infirmiers 0569
  - Toxicologie 0383

SCIENCES PHYSIQUES

Sciences Pures

- Chimie
  - Genéralités 0485
  - Biochimie 487
  - Chimie agricole 0749
  - Chimie analytique 0486
  - Chimie minérale 0488
  - Chimie nucléaire 0738
  - Chimie organique 0490
  - Chimie pharmaceutique 0491
  - Physique 0494
  - PolymÇres 0495
  - Radiation 0754
- Mathématiques 0405
- Physique
  - Généralités 0605
  - Acoustique 0986
  - Astronomie et astrophysique 0606
  - Electronique et électricité 0607
  - Fluides et plasma 0759
  - Météorologie 0608
  - Optique 0752
  - Particules (Physique nucléaire) 0798
  - Physique atomique 0748
  - Physique de l'état solide 0611
  - Physique moléculaire 0609
  - Physique nucléaire 0610
  - Radiation 0756
- Statistiques 0463

Sciences Appliqués Et Technologie

- Informatique 0984
- Ingénierie
  - Généralités 0537
  - Agricole 0539
  - Automobile 0540
  - Biomédicale 0541
  - Chaleur et ther modynamique 0348
  - Conditionnement (Emballage) 0549
  - Génie aérospatial 0538
  - Génie chimique 0542
  - Génie civil 0543
  - Génie électronique et électrique 0544
  - Génie industriel 0546
  - Génie mécanique 0548
  - Génie nucléaire 0552
  - Ingénierie des systämes 0790
  - Mécanique navale 0547
  - Métallurgie 0743
  - Science des matériaux 0794
  - Technique du pétrole 0765
  - Technique minière 0551
  - Techniques sanitaires et municipales 0554
  - Technologie hydraulique 0545
- Mécanique appliquée 0346
- Géotechnologie 0428
- Matières plastiques (Technologie) 0795
- Recherche opérationnelle 0796
- Textiles et tissus (Technologie) 0794

PSYCHOLOGIE

- Généralités 0621
- Personnalité 0625
- Psychobiologie 0349
- Psychologie clinique 0622
- Psychologie du comportement 0384
- Psychologie du développement 0620
- Psychologie expérimentale 0623
- Psychologie industrielle 0624
- Psychologie physiologique 0989
- Psychologie sociale 0451
- Psychométrie 0632

THE UNIVERSITY OF CALGARY
FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a dissertation entitled "Force-Free Magnetic Fields" submitted by Hatem Zaghloul in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

H.A. Buckmaster

Supervisor, Dr. H.A. Buckmaster, Physics and Astronomy

C.J. Bland

Supervisor, Dr. C.J. Bland, Physics and Astronomy

Dr. D. Hobill, Physics and Astronomy

Dr. J. Śniatycki, Mathematics and Statistics

R.E. Collin

External Examiner, Dr. R.E. Collin, Case Western University

C. Chu

External Examiner, Dr. C.C. Chu, Independent Consultant

Date: 23 December 1993

ABSTRACT

This dissertation reviews the subject of force-free (FF) magnetic fields and electromagnetic fields with $\mathbf{E} \| \mathbf{B}$. It provides a critical analysis of the FF condition which shows that it behaves properly under parity and time reversal transformations. It reviews the proof that it is impossible to have completely FF magnetic fields as well as the time development and the stability of FF magnetic fields and reviews the known FF magnetic field solutions in rectangular, cylindrical and spherical coordinate systems. It shows that there are solutions that cannot be included in the accepted "general" spherical coordinates solution. More general spherical coordinate solutions are then obtained. The dissertation reviews the application of FF magnetic fields into plasma confinement and introduces a new spheromak configuration that may prove to be more stable than currently accepted configurations.

This dissertation also examines the conditions under which transverse electromagnetic (TEM) waves exist in a sourceless medium. It shows that TEM waves can be classified according to whether their Poynting vector is identically zero or non-zero. The former are non-propagating TEM standing waves (TEMSW) with $\mathbf{E} \parallel \mathbf{B}$ and the latter are TEM traveling and standing waves with $\mathbf{E} \perp \mathbf{B}$ and $\mathbf{E} \not\parallel \mathbf{B}$. The general conditions under which $\mathbf{E} \parallel \mathbf{B}$ TEMSW exist are derived. Mathematical examples of these standing waves are given. An experimental configuration to realize one of these examples is revisited. The behavior of these waves under Lorentz transformations is discussed and it is shown that these waves are light-like and that their fields lead to well defined Lorentz invariants. The dissertation reviews the physical properties of $\mathbf{E} \parallel \mathbf{B}$ TEMSW and shows how the different Lorentz invariants can be used to classify TEM waves.

ACKNOWLEDGMENTS

I have to first praise Allah for guiding me and giving me the strength to complete this thesis.

I then would like to thank a number of dear people; my beloved mother for her continued encouragement, love and care, my dear wife, Sawzia, and my children, Ahmed and Fatima, for all their sacrifices that allowed me to do the needed work, Dr. H.A. Buckmaster, Dr. C.J. Bland, Dr. D. Hobill, and Dr. R. Taylor for their patience and their supervision of the work that led to this thesis, and Dr. K. Volk, Dr. T.H.T. Van Kalleveen, Dr. H. Laue, and Mr. O. Barajas for their input on various chapters.

I would like to thank the University of Calgary for providing the facilities and financial support and the Natural Sciences and Engineering Research Council of Canada, and the Killam foundation for Post-Graduate Scholarships.

I would like to mention my friends Dr. Orla Aaquist, Mr. Paul Milligan, Dr. M. Fattouche, Ms. A.L. Norman, Mr. Oscar Barajas, and Dr. A. Lakhtakia, for their continued support and encouragement. It meant a lot to me.

I finally would like to thank my dear administrative assistant, Cherie Rothfus, who typed and edited this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
FORCE FREE MAGNETIC FIELDS

1.1. INTRODUCTION

The origin of the hypothesis of force-free (FF) magnetic fields can be traced back to the work of London and London [1934] on superconductors. They were studied by Lundquist [1950] who also mentioned them in his work [1951] on the stability of magneto-hydrostatic fields. Lust and Schlute [1954] introduced them into stellar media to allow magnetic fields and large currents to exist simultaneously in stellar matter with vanishing Lorentz force (more exactly, the magnetic field exerts no force on the matter) [FERRARO AND PLUMPTON 1961]; i.e., $\mathbf{j}\times\mathbf{B}=0$, where $\mathbf{j}$ is the current density and $\mathbf{B}$ is the magnetic field. The Lorentz force vanishes if $\mathbf{j}$ is parallel to $\mathbf{B}$; i.e., if $\mathbf{j}=\alpha\mathbf{B}$, where α is a scalar coefficient of proportionality and may be a constant or a function of space and time. However, the case of constant α proved to be very interesting since it has exact analytical solutions.

Force-free magnetic fields have not been dealt with in texts on classical electrodynamics. However, their applications in specific fields have appeared in several books: Chandrasekhar [1961] and Parker [1979] discuss their applications in explaining stellar physics phenomena, Campbell and Evetts [1972] discuss their application to enhance the critical current in superconductors, and Zijlstra [1967] mentioned their application in the design of force-reduced coils.

This chapter and parts of chapter 2 are in essence the same as the work of Zaghloul and Barajas [1993]. The remainder of this chapter attempts to review the subject of FF magnetic fields. First the FF condition is analyzed critically and it is shown that it behaves properly under Lorentz, rotational, parity and time reversal transformations. The chapter reviews the proof that it is impossible to have completely FF magnetic fields. Finally, the chapter reviews how FF magnetic fields are used to explain solar flares, to increase the critical field in superconductors, in plasma confinement, and in the generation of large magnetic fields. Chapter 2 reviews the known magnetic field solutions in rectangular, cylindrical and spherical coordinate systems. It shows that there are solutions that cannot be included in the accepted "general" solution as given in spherical coordinates. More general spherical coordinate solutions are then obtained. Chapter 3 reviews the application of FF magnetic fields into plasma confinement and introduces a new spheromak configuration that may prove to be more stable than currently accepted configurations. Chapter 4 reviews the subject of $\mathbf{E} \parallel \mathbf{B}$ transverse electromagnetic (TEM) standing waves.

The subject of FF magnetic fields is a very broad subject and it is very difficult, and perhaps impossible, to review all the literature that was written on them in a single thesis. The author has attempted to review the literature and provide highlights of the major aspects of the field. Apologies are due if any work was not mentioned.

Natural units, where the speed of light, c, the permittivity and permeability are all unity, are used in this thesis except where practical applications make SI units more desirable. Vectors are denoted by bold letters. Roman subscripts signify the components of a vector in 3-dimensions. Greek sub- and super-scripts signify 4-dimensional tensors. Other notational definitions are given as required.

1.2. THE FORCE-FREE CONDITION

The equations governing electromagnetic fields are Maxwell's equations,

$$\nabla \cdot \mathbf{E} = \rho \quad , \quad \nabla \cdot \mathbf{B} = 0 \quad ,$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \;, \quad \nabla \times \mathbf{B} = \mathbf{j} + \frac{\partial \mathbf{E}}{\partial t} \tag{1.1}$$

where **E** is the electric field and ρ is the charge density, together with

$$\mathbf{j} = \sigma \mathbf{E} + \sigma \mathbf{v} \times \mathbf{B} + \rho \mathbf{v}, \tag{1.2}$$

where $\mathbf{v}$ is the flow velocity and σ is the conductivity. The conductivity is in general a tensor. For simplicity, we will treat it as a time-pseudo-scalar that is independent of space and time. This is equivalent to ignoring plasma oscillations and assuming that the medium is both isotropic and homogeneous. Force-free magnetic fields exert no force on the material, i.e., $\mathbf{j} \times \mathbf{B} = 0$. This can be achieved if (i) $\mathbf{j} = 0$, (ii) $\mathbf{B} = 0$, or (iii) $\mathbf{j} = \alpha \mathbf{B}$. The first two cases are straightforward to handle and may be considered to be special cases of the third case; namely the cases where $\alpha = 0$ and α is infinity, respectively.

In the development of the equations for FF magnetic fields, the term corresponding to the displacement current, $\partial \mathbf{E}/\partial t$, is ignored [FERRARO AND PLUMPTON 1961, PAI 1963]. In magnetohydrodynamics, this approximation is termed the magnetohydrodynamic approximation and is justified since the following three conditions are generally satisfied [PAI 1963] (i) the flow velocity, **v**, is much less than the speed of light, (ii) the electric field in the frame, where the current carrying conductors are stationary, is of the same order as the induced electric field $\mathbf{v} \times \mathbf{B}$ and (iii) very high frequencies are not considered. For FF fields, this approximation allows us to write

$$\nabla \times \mathbf{B} = \mathbf{j} = \alpha \mathbf{B} \tag{1.3}$$

or

$$\nabla \times \mathbf{B} = \alpha \mathbf{B} \tag{1.3'}$$

Equation (1.3') is called the FF condition. However, if any of the conditions listed above is not satisfied then this equation does not properly represent the FF condition. We will first derive a FF condition when the electric field and its time derivative are not negligible and v << c. In this case,

$$\mathbf{j} = \sigma\mathbf{E}. \tag{1.4}$$

Note that α is not necessarily a constant; it may be a function of space and time. We will study the case where α is constant separately since it is simple and interesting.

1.2.1. α is a Constant

If α is constant, then

$$\nabla \times \mathbf{E} = \nabla \times \frac{\mathbf{j}}{\sigma} = \frac{\alpha}{\sigma} \nabla \times \mathbf{B} = -\frac{\partial \mathbf{B}}{\partial t} \tag{1.5}$$

and

$$\nabla \times \mathbf{B} = \mathbf{j} + \frac{\partial \mathbf{E}}{\partial t} = \alpha\mathbf{B} + \frac{\alpha}{\sigma}\frac{\partial \mathbf{B}}{\partial t} \tag{1.6}$$

Substitution of (1.5) into (1.6) gives:

$$\frac{\sigma}{\alpha}\frac{\partial \mathbf{B}}{\partial t} = \alpha\mathbf{B} + \frac{\alpha}{\sigma}\frac{\partial \mathbf{B}}{\partial t}$$

$$-\left(\frac{\sigma}{\alpha} + \frac{\alpha}{\sigma}\right)\frac{\partial \mathbf{B}}{\partial t} = \alpha\mathbf{B} \tag{1.7}$$

which leads to

$$\mathbf{B}(\mathbf{r}, t) = \mathbf{B}_0(\mathbf{r}) e^{-\alpha^2 \sigma t/(\sigma^2 + \alpha^2)} \tag{1.8}$$

This solution reduces to that of Chandrasekhar and Kendall [1957] when $\sigma >> \alpha$. The solution (1.8) has the following two properties: (a) the vectors $\partial\mathbf{B}/\partial t$, $\partial\mathbf{E}/\partial t$, $\mathbf{B}$, and $\mathbf{j}$ are all parallel, and (b) if $\alpha = 0$ or $\sigma = 0$, the time dependence vanishes and the field is static.

It follows that FF magnetic fields with $\mathbf{j} = \alpha\mathbf{B}$, where α is a non-zero constant, also satisfy (1.3') with a modified constant. However, notice that if the conductivity is zero, the current density is zero and we have FF magnetic fields (for which α is zero) but $\mathbf{E}$ need no longer be parallel to $\mathbf{B}$. The relation between electromagnetic fields with $\mathbf{E} \parallel \mathbf{B}$ and FF magnetic fields was first demonstrated by Chu and Ohkawa [1982]. The idea of electromagnetic fields with $\mathbf{E} \parallel \mathbf{B}$ has stirred a controversy in literature over the past few years. They will be studied in more detail in Chapter 4.

1.2.2. α is a Function of Space and Time

Recent work on FF magnetic fields has concentrated mainly on the case where α is a function of spatial coordinates [For example ALY 1984]. This work assumes that FF fields satisfy $\nabla \times \mathbf{B} = \alpha\mathbf{B}$. However, this condition ignores the displacement current, and unlike the case of constant α, this does not merely change the constant. Substituting $\mathbf{j} = \alpha\mathbf{B} = \sigma\mathbf{E}$ into the two curl equations of (1.1) leads to

$$\nabla \times \mathbf{B} = \alpha\mathbf{B} + \frac{\partial \mathbf{E}}{\partial t} = \alpha\mathbf{B} + \frac{\alpha}{\sigma}\frac{\partial \mathbf{B}}{\partial t} + \frac{1}{\sigma}\frac{\partial \alpha}{\partial t}\mathbf{B}$$

$$= (\alpha + \frac{1}{\sigma}\frac{\partial \alpha}{\partial t})\,\mathbf{B} + \frac{\alpha}{\sigma}\frac{\partial \mathbf{B}}{\partial t}, \tag{1.9}$$

and

$$\nabla \times \mathbf{E} = \nabla \times \left(\frac{\alpha}{\sigma}\mathbf{B}\right) = \frac{\alpha}{\sigma}\nabla \times \mathbf{B} + \frac{1}{\sigma}\nabla\alpha \times \mathbf{B} = -\frac{\partial \mathbf{B}}{\partial t} \tag{1.10}$$

Substitution for $\partial\mathbf{B}/\partial t$ from (1.10) into (1.9) and rearranging lead to

$$\nabla \times \mathbf{B} = \left[\sigma\left(\alpha\sigma + \partial\alpha/\partial t\right)/\left(\alpha^2 + \sigma^2\right)\right]\mathbf{B} - \left[\alpha/\left(\alpha^2 + \sigma^2\right)\right]\nabla\alpha \times \mathbf{B} \tag{1.11}$$

This means that $\nabla \times \mathbf{B} \neq \alpha\mathbf{B}$ unless $\nabla\alpha \| \mathbf{B}$ or $\nabla\alpha = 0$. The latter is impossible if α is a function of spatial coordinates and the former is generally not the case.

1.2.3. Relativistic Force-Free Fields

We will now investigate the FF condition from a relativistic point of view. The covariant (electromagnetic) Lorentz force density is $f_0 = \mathbf{E} \cdot \mathbf{j}$ and $\mathbf{f} = (\nabla \cdot \mathbf{E})\mathbf{E} + \mathbf{j} \times \mathbf{B}$. If the electromagnetic field is FF in all frames then all four components of f_μ must vanish. f_0 vanishes if $\mathbf{E} \perp \mathbf{j}$ or if either $\mathbf{E}$ or $\mathbf{j}$ vanishes. If $\rho = \nabla \cdot \mathbf{E} \neq 0$, $\mathbf{E} = \mathbf{j} \times \mathbf{B}/\rho$ which means that: (i) $\mathbf{E} \perp \mathbf{j}$; i.e. $f_0 = 0$ and (ii) such fields cannot have $\mathbf{E} \parallel \mathbf{B}$. Note that the orthogonality of the current and electric field is not a common phenomenon. Evangelidis [1987] studied the compatibility between the relativistic and the non-relativistic FF condition. It is interesting to note that London and London [1934] used relativistic terms when they mentioned the absence of pondermotive volume forces inside the superconductor.

Stenzil and Urratia [1990] demonstrated a FF (vanishing covariant Lorentz force) electromagnetic field in a magnetized plasma.

The remainder of this thesis, however, uses the more restricted FF condition (1.3').

1.3. THE POSSIBILITY OF STRESS FREE CURRENT CARRYING COILS

1.3.1. Parker's Proof

Parker [1958] gave a proof that it is impossible to configure coils such that they are free from all magnetic forces. It is believed that Parker was the first to deliver a proof for this assertion. However, the theorem was known without proof before Parker published his proof. Furth et al. [1957] mentioned a proof similar to that due to Parker.

To prove his assertion, Parker [1958] considered the wires "as being composed of perfectly conducting, classical fluid, rather than rigid metal, and held in place by suitable volume forces, f_i, x_i of constraint." He set the total force per unit volume of the fluid, F_i, equal to $f_i + \frac{\partial M_{ij}}{\partial x_j}$ where $M_{ij} = B_i B_j - \frac{1}{2}\delta_{ij}B^2$ is Maxwell's stress tensor in the absence of an electric field. The wires have no internal kinetic energy since they are assumed to be classical fluids. For static equilibrium, the scalar virial equation [PARKER 1954] requires

$$\int_V x_i F_i dV = 0 \tag{1.12}$$

The Einstein summation convention is used throughout this section. Now, since

$$\frac{\partial x_i M_{ij}}{\partial x_j} = \delta_{ij} M_{ij} + x_i \frac{\partial M_{ij}}{\partial x_j} = M_{ii} + x_i \frac{\partial M_{ij}}{\partial x_j} \tag{1.13}$$

then substitution into (1.12) yields

$$\int_V \left[x_i f_i + \frac{\partial x_i M_{ij}}{\partial x_i} - M_{ii} \right] dV = 0 \tag{1.14}$$

We substitute $M_{ij} = -B^2/2$ and use the divergence theorem

$$\int_V x_i f_i dV + \int_S x_i M_{ij} dS_j + \int_V \frac{B^2}{2} dV = 0 \tag{1.15}$$

where dS_j is an element of the surface S bounding the volume V. Parker [1958] then explained that since far away from the currents the field must decrease as $1/r^3$ then the surface integral in (1.15) can be made negligible by choosing S sufficiently large. In this case,

$$\int_V \frac{B^2}{2} dV = -\int_V x_i f_i dV \tag{1.16}$$

which is equation (3) of Parker [1958]. Parker [1958] finally stated that since the left hand side is positive definite (vanishes if the field vanishes everywhere) then there must be a net inward force, f_i, exerted to keep the conductors in place. Note that Levy [1962] gave a similar proof but used $\partial S_{ij}/\partial x_j$, where S_{ij} is the mechanical stress tensor, rather than f_i.

1.3.2. Zaghloul's Refutation of Parker's Proof

Zaghloul [1989] attempted to refute Parker's [1958] and Levy's [1962] proofs. This subsection will review Zaghloul's work on the subject and clarify that Zaghloul erred by concluding that it is possible to configure current carrying coils in a FF configuration.

The electromagnetic force, which in the absence of an electric field is purely magnetic, will be termed L_i. This force is the Lorentz force $\mathbf{L} = \mathbf{j} \times \mathbf{B}$ and, as will be shown below, satisfies $L_i = \partial M_{ij}/\partial x_j$. To get a clearer view of the physical significance of each term, Cartesian coordinates are used and the summation convention is temporarily dropped,

$$\begin{aligned} L_x &= \frac{\partial M_{xx}}{\partial x} + \frac{\partial M_{xy}}{\partial y} + \frac{\partial M_{xz}}{\partial z} \\ &= -B_y \frac{\partial B_y}{\partial x} - B_z \frac{\partial B_z}{\partial x} + B_x \frac{\partial B_x}{\partial x} + B_y \frac{\partial B_x}{\partial y} + B_x \frac{\partial B_y}{\partial y} + B_z \frac{\partial B_x}{\partial z} + B_x \frac{\partial B_z}{\partial z} \\ &= B_x (\nabla \cdot B) + [(\nabla \times B) \times B]_x \end{aligned} \tag{1.17}$$

With similar expressions for the y- and z- components. Now $F_i = f_i + L_i$ which when substituted in (1.12) leads to the conclusion that $\int_V x_i F_i dV$ is dependent on the angle between $\mathbf{j}$ and $\mathbf{B}$ which contradicts the result of Parker [1958]. Note that (1.13) becomes

$$\frac{\partial x_i M_{ij}}{\partial x_j} = M_{ii} + x_i[\mathbf{j}\times\mathbf{B}]_i = -\frac{1}{2}B^2 + x_i[\mathbf{j}\times\mathbf{B}]_i \tag{1.18}$$

and if **j** || **B**, i.e., if the magnetic field is FF, then $\frac{\partial x_i M_{ij}}{\partial x_j} = -\frac{1}{2}B^2$.

Zaghloul [1989] interpreted this to mean that for FF magnetic fields, the magnetic energy density is actually the divergence of a vector, **N**, where $N_j = -x_i M_{ij}$. Furthermore, in this case, the surface integral in (1.15) that Parker [1958] made negligible by increasing the surface area is exactly equal to the total magnetic energy, U_m and hence is independent of the surface area. To determine the order of magnitude of the vector, **N**, use is made of the equivalent to Gauss' law in this case, $\int_S N_j dS_j = U_m$. By choosing Gauss' surface to be a sphere of radius R and considering that the vector, **N**, should be radial far away from the currents, Zaghloul [1989] obtained $N_r = \frac{U_m}{4\pi r^2}$.

Zaghloul [1989] concluded that he had refuted Parker's proof that it is impossible to configure current carrying conductors in a manner such that they are free of stresses and that he had also shown that the magnetic energy density of FF magnetic fields is the divergence of a vector, **N**, where $N_j = -x_i M_{ij}$. However, even though Zaghloul's [1989] analysis was mathematically correct, it lacked physical significance: stress-free current carrying coils have a magnetic field that drops at a rate lower than $1/r^3$. To be precise, the terms in M_{ij} have a $1/r^3$ dependence on the distance. This would only be possible if the coils extended over all space which is physically unrealizable since the lowest possible order of multipole expansion possible for magnetic fields is a dipole with $1/r^3$ dependence for the fields and $1/r^6$ for M_{ij}.

In conclusion, Parker [1958] and Levy [1962] had proven that it is impossible to position current carrying conductors in a force-free configuration.

1.4. STABILITY OF FORCE-FREE MAGNETIC FIELDS

The question of magnetic stability will be addressed through the variational principle in which one will be looking for magnetic fields that minimize an energy integral subject to appropriate constraints. One such constraint that has gained much popularity is the magnetic helicity.

1.4.1. On Magnetic Helicity

Woltjer [1958a] introduced the integral

$$K_0 = \int_V \mathbf{A} \cdot \mathbf{B} dV \tag{1.19}$$

and showed that if the volume, V, is surrounded by a perfect conductor ($\mathbf{B} \cdot \mathbf{dS} = 0$ at the boundary where dS is an element of area), if the system is a closed system not interacting with the outside world ($\partial \mathbf{A}/\partial t = 0$ at the boundary) and if the conductivity is infinite within the volume, then $\partial K_0/\partial t = 0$, which means that for the conditions he imposed (which really are applicable within a confined plasma), K_0 is constant during the development of the system. However, Woltjer's arguments did not consider the scalar potential. We will now show that: (i) Woltjer's conclusion is correct even when the scalar potential is considered and (ii) K_0 is gauge invariant [TAYLOR 1986] under the conditions proposed by Woltjer [1958a]. To prove (i) we follow Woltjer's steps except for the inclusion of the scalar potential, $\nabla\phi$. If the conductivity is infinitely large, the total electric force per unit charge, $\mathbf{E}_t = \mathbf{E} + \mathbf{v} \times \mathbf{B}$, must be infinitesimally small and in this case

$$\mathbf{E} = -\mathbf{v} \times \mathbf{B} = -\partial \mathbf{A}/\partial t - \nabla\phi \tag{1.20}$$

$$\frac{\partial \mathbf{A}}{\partial t} = \mathbf{v} \times \mathbf{B} - \nabla\phi \tag{1.21}$$

$$\frac{\partial \mathbf{A}}{\partial t} \cdot \mathbf{B} = -\nabla\phi \cdot \mathbf{B} \tag{1.22}$$

$$\begin{aligned}
\frac{\partial}{\partial t}\int_V \mathbf{A}\cdot\mathbf{B}dV &= \int_V \frac{\partial \mathbf{A}}{\partial t}\cdot\mathbf{B}dV + \int_V \mathbf{A}\cdot\frac{\partial \mathbf{B}}{\partial t}dV \\
&= \int_V \frac{\partial \mathbf{A}}{\partial t}\cdot\mathbf{B}dV + \int_V \mathbf{A}\cdot\frac{\partial \nabla\times\mathbf{A}}{\partial t}dV \\
&= \int_V \frac{\partial \mathbf{A}}{\partial t}\cdot\mathbf{B}dV + \int_V \nabla\cdot(\mathbf{A}\times\frac{\partial \mathbf{A}}{\partial t})dV - \int_V \frac{\partial \mathbf{A}}{\partial t}\cdot(\nabla\times\mathbf{A})dV \\
&= \int_V \nabla\cdot(\mathbf{A}\times\frac{\partial \mathbf{A}}{\partial t})dV \\
&= \int_S (\mathbf{A}\times\frac{\partial \mathbf{A}}{\partial t})\cdot d\mathbf{S} = 0
\end{aligned} \tag{1.23}$$

To prove (ii) we have to prove that

$$\int_V \mathbf{A}\cdot\mathbf{B}dV = \int_V \mathbf{A}'\cdot\mathbf{B}dV \tag{1.24}$$

where $\mathbf{A}' = \mathbf{A} + \nabla\psi$ and ψ is any scalar function.

$$\begin{aligned}
K_0 = \int_V \mathbf{A}\cdot\mathbf{B}dV &= \int_V \mathbf{A}'\cdot\mathbf{B}dV - \int_V \nabla\psi\cdot\mathbf{B}dV \\
&= \int_V \mathbf{A}'\cdot\mathbf{B}dV - \int_V \nabla\cdot(\psi\mathbf{B})dV + \int_V \psi\nabla\cdot\mathbf{B}dV \\
&= \int_V \mathbf{A}'\cdot\mathbf{B}dV - \int_S \psi\mathbf{B}\cdot d\mathbf{S} = \int_V \mathbf{A}'\cdot\mathbf{B}dV
\end{aligned} \tag{1.25}$$

The integral K_0 has been termed the magnetic helicity.

1.4.2. On the Stability of Force-free Magnetic Fields

It is obvious that if the magnetic energy, $U_m = \int_V \frac{B^2}{2} dV$, is to be minimized, the minimum will occur at zero magnetic field. Appropriate constraints have to be imposed on the variations of the magnetic field. In a simplistic approach, one can require that the magnetic helicity discussed above remain unchanged and add it to the magnetic energy through the use of a Lagrange multiplier, λ; it is then required to find the magnetic fields that minimize

$$F = U_m + \lambda K_0 \tag{1.26}$$

We start by finding the variation in F, δF, due to small variation, $\delta \mathbf{B}$, in the magnetic fields, $\mathbf{B}$ and then find the conditions on $\mathbf{B}$ such that $\delta F = 0$. We are again considering the same boundary conditions as those introduced in the discussion on the magnetic helicity.

$$\begin{aligned}
\delta F &= \int_V [2\mathbf{B}\cdot\delta\mathbf{B} + \lambda\mathbf{A}\cdot\delta\mathbf{B} + \lambda(\nabla\times\mathbf{A})\cdot\delta\mathbf{A}]dV \\
&= \int_V [2\mathbf{B}\cdot\delta\mathbf{B} + 2\lambda\mathbf{A}\cdot\delta\mathrm{B} + \lambda\nabla\cdot(\mathbf{A}\times\delta\mathbf{A})]dV \\
&= \int_V [2\mathbf{B}\cdot\delta\mathbf{B} + 2\lambda\mathrm{A}\cdot\delta\mathbf{B}]dV + \lambda\int_S \mathbf{A}\times\delta\mathbf{A}\cdot d\mathbf{S} \\
&= \int_V [2\mathbf{B} + 2\lambda\mathbf{A}]\cdot\delta\mathbf{B}dV
\end{aligned} \tag{1.27}$$

For an arbitrary $\delta\mathbf{B}$, δF vanishes if $\mathbf{B} = -\lambda\mathbf{A}$ which becomes $\nabla\times\mathbf{B} = \alpha\mathbf{B}$ if one sets $\alpha = -\lambda$ and takes the curl of both sides. This apparently means that the fields satisfying the FF condition are minimum energy states.

That δF vanishes for FF magnetic fields should be evident since if the magnetic force vanishes, there are no forces trying to change the state of the field. However, there is no guarantee that this is a minimum energy state. This was best summarized by

Woltjer [1959], "Obviously minimizing the energy, we obtain stable configurations. But the variational principle gives us true minima as well as other extrema, the latter corresponding to unstable equilibria. Thus stability occurs only for certain values of the Lagrange multipliers in our equations. Generally the nature of the extremum and thus its stability can only be established by determining the second variation of the energy. However, in some cases we can find the most stable configuration, i.e., the state for which the energy is an absolute minimum by simpler means." By "simpler means" Woltjer [1959] meant to compare the energy values with other energy values and check that it is indeed a minimum. Woltjer [1959] appealed to these "simpler means" in his proofs that constant α FF fields are energy minima.

1.4.3. On Criticism of the Stability of Force-Free Fields

Salingaros [1987] criticized the statement that FF magnetic fields represent the state of minimum magnetic energy subject to appropriate constraints on the variations of the field. Salingaros [1987] used the constraints

$$K_1 = \int_V |\nabla \times B|^2 dV \text{ and } K_2 = \int_V \mathbf{B} \cdot \nabla \times B dV \qquad (1.28)$$

and derived

$$\delta F = 0 \text{ if } 1/2\mathbf{B}_0 + \lambda_1 \nabla \times \mathbf{B}_0 + \lambda_2 \nabla \times \nabla \times \mathbf{B}_0 = 0 \qquad (1.29)$$

where $\mathbf{B}_0$ is a fixed initial field configuration, and λ_1 and λ_2 are the two Lagrange multipliers associated with K_1 and K_2 in the energy integral. Salingaros [1987] also stated that for a variation $\varepsilon\mathbf{h}$ around $\mathbf{B}_0$

$$\delta^2 F = \varepsilon^2 \int_V \left[\frac{1}{2}h^2 + \lambda_1 \mathbf{h} \cdot \nabla \times \mathbf{h} + \lambda_2 |\nabla \times \mathbf{h}|^2\right] dV \qquad (1.30)$$

However, Salingaros [1987] stated, "The condition (5a) above is satisfied by several distinct classes of vector fields. First, substitute the field $\nabla \times \mathbf{B} = k\mathbf{B}$ into (5a) to determine the condition for those fields to be an energy extrema. One has

$$\lambda_2 k^2 + \lambda_1 k + \frac{1}{2} = 0 \Rightarrow \delta^2 F = 0 \tag{6}$$

The class of vector fields corresponds to an energy extremum which is a saddle point, since $\mathbf{h}(r)$ satisfies the same identity as $\mathbf{B}(r)$ (6) so that $\delta^2 F = 0$ (5b)." By forcing the variations of $\mathbf{B}$ to satisfy the FF condition, Salingaros [1987] was only searching amongst FF fields for those which are energy extrema. He found that amongst the fields satisfying $\nabla \times \mathbf{B} = k\mathbf{B}$, those with k's satisfying equation (6) of Salingaros are energy extrema. Note that since his equation (6) involves Lagrange multipliers which can be arbitrarily chosen, it can be satisfied by any k. Another fact is that Salingaros [1987] interpreted the vanishing of the second derivative as the point is a saddle point. Following his argument (searching amongst FF fields), we notice that the dependence of the second variation on $\mathbf{h}$ is identical to the dependence of F on $\mathbf{B}$. This suggests that not only does the second variation vanish, all higher order variations vanish. This means that the energy integral is a constant. This is an interesting concept. We therefore conclude that the work of Salingaros [1987] was not of a general nature and that he misinterpreted the vanishing of the second integral.

Lundquist [1951] directly addressed the question of the stability of FF magnetic fields and investigated the stability of the magnetic field in a long cylinder and concluded that the FF field components

$$B_\rho = 0, \; B_\phi = B_0 J_1(\alpha\rho), \text{ and } B_z = B_0 J_0(\alpha\rho) \tag{1.31}$$

(where B_0 is the amplitude of the field, J_0 and J_1 are Bessel functions and ρ, ϕ, z are cylindrical coordinates) are stable against perturbations of the form

η = (c cos bx sin az, 0, c' sin bx cos az) (x, y, and z are Cartesian coordinates) (1.32)

Woltjer [1958b] has shown that general FF fields with constant α are stable against spherical expansion.

A second variation under rather general circumstances, based on the energy principle of Bernstein et al. [1958] was worked out by Voslanber and Callebaut [1962] and surprisingly they found that a wide class of FF magnetic fields with constant α is energy maxima states (the second variation is negative).

In conclusion of this section, FF magnetic fields represent states of equilibrium (if only magnetic forces are present). They are not always stable equilibria.

1.5. TIME DEVELOPMENT OF FORCE-FREE MAGNETIC FIELDS

The equations governing the time development of a FF magnetic field are

$$\nabla \times \mathbf{B} = \alpha\mathbf{B} = \mathbf{j} \;\Rightarrow\; \alpha = \frac{\mathbf{B}\cdot\nabla\times\mathbf{B}}{B^2} \tag{1.33}$$

$$\nabla\cdot\mathbf{B} = 0 \;\Rightarrow\; \mathbf{B}\cdot\nabla\alpha = 0 \tag{1.34}$$

together with Faraday's law

$$\nabla\times\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t} \tag{1.35}$$

and since $\mathbf{E} = -\mathbf{v}\times\mathbf{B} + 1/\sigma\mathbf{j}$, (1.35) reduces to

$$\nabla \times (\frac{1}{\sigma}\nabla \times \mathbf{B} - \mathbf{v}\times\mathbf{B}) = \frac{1}{\sigma}\nabla\times\alpha\mathbf{B} - \nabla\times\mathbf{v}\times\mathbf{B} + \frac{\partial\mathbf{B}}{\partial t} \tag{1.36}$$

and if the fluid is assumed to be stationary ($\mathbf{v} = 0$)

$$\partial\mathbf{B}/\partial t = \left(1/\sigma\,\mathbf{B}\times\nabla\alpha - \alpha^2\mathbf{B}\right) \tag{1.37}$$

Ferraro and Plumpton [1961] suggested that in order to solve these equations when α is a function of space and time, one has to find a solution, **B**, at a certain instant of time such that this solution satisfies (1.3') and $\mathbf{B}\times\nabla\times\mathbf{B}=0$ and find α through (1.33) and finally use (1.37) to define **B** at later times. Jette and Sreenivasan [1969] showed that this does not guarantee that at later times, **B** will satisfy the solenoidal condition nor $\mathbf{B}\times\nabla\times\mathbf{B}=0$. Now, if a magnetic field is once FF, does it always remain FF? The answer to this question has been given by Jette [1970]: only if α is constant in space and time will a magnetic field that is once FF always remain FF. Earlier, Jette and Sreenivasan [1969] had shown that α must be either a constant or a function of both space and time.

1.6. TRANSFORMATION PROPERTIES OF THE FORCE-FREE CONDITION

The FF condition and its magnetic field solutions have been criticized by Salingaros [1986] who attempted to show that the condition is not invariant under parity nor general spatial transformations and that its solutions are not gauge invariant. It was later shown to be gauge invariant by Maheswaran [1986] and invariant under parity by Collinson [1987]. This chapter reviews that controversy and shows that this equation is invariant under (i) parity, and (ii) time reversal and that the constant α must be a pseudoscalar. This section, also, clarifies that the vector differential equation is a special case of Ampere's law and that under Lorentz transformations, Ampere's law, and not necessarily $\nabla\times\mathbf{B}=\alpha\mathbf{B}$, is invariant.

1.6.1. Gauge Invariance

Salingaros [1986] suggested that solutions of (1.3') are not gauge invariant. He suggested that two vector potentials only different by the gradient of a scalar function should lead to the same magnetic field, i.e.,

if $\quad \mathbf{B} = \nabla \times \mathbf{A} \qquad (1.38)$

and $\quad \mathbf{A}' = \mathbf{A} + \nabla\lambda \qquad (1.39)$

then $\quad \mathbf{B}' = \mathbf{B} \qquad (1.40)$

Salingaros [1986] then suggested that an implication of (1.3') is that

$$\mathbf{B} = \alpha\mathbf{A} + \nabla\varphi \tag{1.41}$$

and $\quad \mathbf{B}' = \alpha\mathbf{A}' + \nabla\varphi = \alpha\mathbf{A} + \alpha\nabla\lambda + \nabla\varphi \qquad (1.42)$

and since $\nabla \cdot \mathbf{B}' = 0$, we get $\alpha\nabla^2\lambda = 0$ which is not true in general.

Maheswaran [1986] (and later Brownstein [1987a]) pointed out that the function φ above does not have to be the same in (1.41) and (1.42). This means that

$$\mathbf{B}' = \alpha\mathbf{A}' + \nabla\varphi' \tag{1.43}$$

with $\quad \nabla\varphi' = \nabla\varphi + \alpha\nabla\lambda \qquad (1.44)$

which would lead to that $\nabla^2\varphi = -\alpha\nabla\cdot\mathbf{A}$, $\nabla^2\varphi' = -\alpha\nabla\cdot\mathbf{A}'$, and

$$\nabla^2\lambda = \nabla\cdot(\mathbf{A} - \mathbf{A}') \tag{1.45}$$

1.6.2. Parity

Salingaros [1986] also criticized the magnetic field solutions of the FF equation (1.3') and argued that since the operator, ∇, is a vector operator that changes sign under reflection of the Cartesian coordinates, then (1.3') is not invariant under parity. Collinson [1987] pointed out that the quantity, α, is a pseudoscalar which means that it changes its sign under reflection of the Cartesian coordinates. This meant that (1.3') is invariant under parity.

Salingaros [1986], however, stated "A magnetic field **B** is known to be a parity eigenvector of even parity." This statement demonstrates a common confusion about parity; namely mixing intrinsic parity and spatial symmetry. The parity operator, $\boldsymbol{P}$, maps the coordinate $\mathbf{r}$ onto $-\mathbf{r}$ and maps an arbitrary vector function of spatial coordinates and time, $\mathbf{F}(\mathbf{r},t)$, onto a vector function $\mathbf{F}'(-\mathbf{r},t)$. It is customary to say that the vector $\mathbf{F}$ is a polar vector (or of odd intrinsic parity) if $\mathbf{F}'(-\mathbf{r},t) = -\mathbf{F}(-\mathbf{r},t)$ and axial (or of even intrinsic parity) if $\mathbf{F}'(-\mathbf{r},t) = \mathbf{F}(-\mathbf{r},t)$. Accordingly, $\mathbf{E}$ is polar and $\mathbf{B}$ is axial [ROHRLICH 1965]. This means that

$$\boldsymbol{P}\,[\mathbf{B}(\mathbf{r},t)] = \mathbf{B}(-\mathbf{r},t) \tag{1.46}$$

On the other hand, Salingaros' statement means that

$$\boldsymbol{P}\,[\mathbf{B}(\mathbf{r},t)] = \mathbf{B}(\mathbf{r},t) \tag{1.47}$$

Comparison of (1.46) and (1.47) shows that Salingaros' statement is valid if and only if

$$\mathbf{B}(\mathbf{r},t) = \mathbf{B}(-\mathbf{r},t) \tag{1.48}$$

which means that **B** is an even function of spatial coordinates. This is not always true since the spatial symmetry (if any) of **B** and **E** is dependent on the spatial symmetry (if any) of the charge density. It is important to stress that all physical quantities must have either even or odd intrinsic parities but do not have to be odd or even functions of spatial coordinates. This confusion of intrinsic parity and spatial symmetry was not addressed by any of the respondents to Salingaros [1986].

1.6.3. General Coordinate Transformations

Salingaros [1986] finally suggested that if a solution of the scalar Helmholtz equation, ψ, is a function of the z coordinate; e.g., $\psi = a\sin(\alpha z)$, then a solution that satisfies (1.3') is

$$\mathbf{B} = a(\sin(\alpha z), \cos(\alpha z), 0) \tag{1.49}$$

and under a transformation of coordinates

$$(y,z) \rightarrow (u(y,z), v(y,z))$$

the magnetic field, **B**, would transform to **B'**

$$\mathbf{B}' = (\psi, \frac{J}{\alpha}\frac{\partial \psi}{\partial v}, -\frac{J}{\alpha}\frac{\partial \psi}{\partial u}) \tag{1.50}$$

where $J = \frac{\partial u}{\partial y}\frac{\partial v}{\partial z} - \frac{\partial u}{\partial z}\frac{\partial v}{\partial y}$ is the Jacobian of the transformation. The solution (1.49) would only be form invariant if $J = 1$ which is valid only for translations and rotations and not any general coordinate transformation. Salingaros [1986] then suggested that the lack of form invariance of the solution invalidates the solution being a magnetic field and appealed to the argument that "The independence of the choice of reference system is essential for all differential equations which describe physical fields." Collinson [1987] pointed out the lack of form invariance of the solution under a general coordinate transformation does not mean that the equation is not covariant (independent of the choice of the reference system). Collinson also explained that the solution (1.49) is a special case of Hansen's [1935] method of solving the vector Helmholtz equation which will be described in the next chapter. Brownstein [1987a] stressed that solutions of the form of (1.47) are only intended for Cartesian coordinates and that one cannot expect such solutions to maintain this form under general coordinate transformations. His explanation of the relationship between (1.49) and Hansen's [1935] method of solution is more explicit than that of Collinson [1987].

1.6.4. Time-Reversal

We will now discuss the behavior of (1.3') under time reversal. Equation (1.3') is a special case of Ampére's law

$$\nabla \times \mathbf{B}(\mathbf{r},t) = \frac{\partial \mathbf{E}(\mathbf{r},t)}{\partial t} + \mathbf{j}(\mathbf{r},t) \tag{1.51}$$

when the magnetic field, **B**, is parallel to $\partial \mathbf{E}/\partial t + \mathbf{j}$, i.e.,

$$\alpha \mathbf{B} = \frac{\partial \mathbf{E}}{\partial t} + \mathbf{j} \tag{1.52}$$

Now, since **B** is an axial vector and both $\partial \mathbf{E}/\partial t$ and **j** are polar vectors [RORHLICH 1965], α must be a pseudoscalar and the FF condition (1.3') is invariant under parity. This is what was stated by Collinson [1987]. Similarly, since time reversal maps **B** onto - **B** and $\partial \mathbf{E}/\partial t + \mathbf{j}$ onto $-(\partial \mathbf{E}/\partial t + \mathbf{j})$, α must be a proper scalar under time reversal; i.e., time reversal maps α onto α, and (1.3') is invariant under time reversal.

1.6.5. Examples

The results in 1.6.2. and 1.6.4. can be applied to the example given by Chu and Ohkawa [1982] which is the same as that used by Salingaros [1986];

$$\mathbf{j} = 0 \tag{1.53a}$$

$$\mathbf{k} = (0, 0, k) \tag{1.53b}$$

$$\mathbf{B} = B_0 \,(\sin kz, \cos kz, 0) \cos \omega t \tag{1.53c}$$

$$\mathbf{E} = E_0 \,(\sin kz, \cos kz, 0) \sin \omega t \tag{1.53d}$$

Both **B** and **E** satisfy (1.3') and α in this case is the only non-zero component of the propagation vector, **k**, which is a polar vector, i.e., its components reverse sign under parity, and this means that $k = \alpha$ is a pseudoscalar in agreement with the results of 1.6.2. However, since the propagation vector is a vector of magnitude $2\pi/\lambda$ where λ is the wavelength of the wave and its direction is the direction of the propagation of the wave and since time reversal reverses the direction of propagation of the wave, it is concluded that time reversal maps **k** onto **-k** which means that k is a time pseudoscalar. Time reversal maps $\mathbf{B}(\mathbf{r},t)$ onto $\mathbf{B}_T(\mathbf{r},t) = -\mathbf{B}(\mathbf{r},-t)$ and maps $\mathbf{E}(\mathbf{r},t)$ onto $\mathbf{E}_T(\mathbf{r},t) = \mathbf{E}(\mathbf{r},-t)$. Applying these results to (1.53c) and (1.53d) gives;

$$\mathbf{B}_T(\mathbf{r},t) = -B_0(\sin kz, \cos kz, 0)\cos(-\omega t) \tag{1.53c'}$$

$$\mathbf{E}_T(\mathbf{r},t) = E_0(\sin kz, \cos kz, 0)\sin(-\omega t) \tag{1.53d'}$$

Both $\mathbf{B}_T$ and $\mathbf{E}_T$ satisfy (1.3'). This demonstrates how the FF condition (1.3') is invariant under both parity and time reversal.

Note the two curl equations of Maxwell's equations in a sourceless medium

$$\nabla \times \mathbf{B} = \frac{\partial \mathbf{E}}{\partial t} \tag{1.54}$$

and

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{1.55}$$

lead, for a monochromatic electromagnetic plane wave, to

$$\nabla \times \mathbf{E} = -j\omega\mathbf{B} \tag{1.56}$$

and

$$\nabla \times \mathbf{B} = j\omega\mathbf{E} \tag{1.57}$$

Simple substitution shows that the vector $\mathbf{E} + j\mathbf{B}$ is a solution for the FF condition. Varadan et al. [1987] reported a similar observation.

The electric and magnetic fields of a circularly polarized monochromatic wave in vacuum

$$\mathbf{E} = \cos(\omega t - kz)\hat{\mathbf{x}} - \sin(\omega t - kz)\hat{\mathbf{y}} \tag{1.58a}$$

and

$$\mathbf{B} = \sin(\omega t - kz)\hat{\mathbf{x}} + \cos(\omega t - kz)\hat{\mathbf{y}} \tag{1.58b}$$

satisfy the FF condition (1.3') with α replaced with -k. Circularly polarized waves are valid, physically significant, and realizable solutions for Maxwell's equations. Note that the wave described by (1.58) is a right-handed wave since **E**, **B** and **k** describe a right handed coordinate system; i.e., $\mathbf{k} = \mathbf{E} \times \mathbf{B}$. The electric and magnetic fields of a left-handed circularly polarized wave (with $\mathbf{k} = -\mathbf{E} \times \mathbf{B}$) satisfy the FF condition (1.3') with α replaced with k. This led Brownstein [1987a] to suggest that the FF condition (1.3') should be replaced with a new condition

$$\nabla \times \mathbf{B} = \xi k \mathbf{B} \tag{1.59}$$

where ξ is a pseudoscalar (with respect to both time-reversal and parity) equal to 1 for left-handed circularly polarized wave and equal to -1 for a right-handed circularly polarized wave. Brownstein [1987a] was motivated to preserve the physical significance of k as the propagation constant. However, if α is viewed as the coefficient of proportionality between the current and the magnetic field, there may be no need to preserve its positive sign. For this reason we will not adopt (1.59) as the FF condition.

1.6.6. Lorentz Transformations

Finally, it is important to stress that the FF condition is a special case of Ampere's law and that under Lorentz transformations, Ampere's law and not necessarily (1.3') is covariant.

1.6.6.1. Sourceless Media

The second rank electromagnetic field tensor

$$F_{\mu\nu} = \begin{bmatrix} 0 & B_z & -B_y & -E_x \\ -B_z & 0 & B_x & -E_y \\ B_y & -B_x & 0 & -E_z \\ E_x & E_y & E_z & 0 \end{bmatrix} \tag{1.60}$$

when multiplied by the four vector potential, A^ν, leads to

$$F_{\mu\nu}A^{\nu} = \left(\mathbf{A}\times\mathbf{B} + \phi\mathbf{E},\ -\mathbf{E}\cdot\mathbf{A}\right) \tag{1.61}$$

and for magnetic fields satisfying the FF condition (1.3') with constant α, it is possible to set $\mathbf{A}$ parallel to $\mathbf{B}$ through a proper choice of the scalar function ψ where $\mathbf{B} = \mathbf{A} + \nabla\psi$ and to set $\phi = 0$ through a proper gauge transformation; two possibilities arise

(i) the Lorentz invariant $\mathbf{E}\cdot\mathbf{B} = 0 \Rightarrow \mathbf{E}\cdot\mathbf{A} = 0$ and $F_{\mu\nu}A^{\nu} = 0$; consequently the FF condition is covariant in all Lorentz frames; and

(ii) the Lorentz invariant $\mathbf{E}\cdot\mathbf{B} \neq 0 \Rightarrow \mathbf{E}\cdot\mathbf{A} \neq 0$ and in the frame where $\mathbf{B}$ satisfies (1.3'), $F_{\mu\nu}A^{\nu} = (0,\ 0,\ 0,\ -\mathbf{E}\cdot\mathbf{A})$ and in any other frame $\mathbf{A}\times\mathbf{B} \neq 0$. Which means that the FF condition is not Lorentz covariant.

An example of case (i) above is the circularly polarized wave of (1.58). If this wave is transformed to a Lorentz frame, K', moving in the positive z-direction at a speed v with respect to the frame in which (1.58) was expressed. Since, $\mathbf{E}$ and $\mathbf{B}$ have zero components in the direction of the relative motion between the two frames,

$$\mathbf{E}' = \gamma(\mathbf{E} + \mathbf{v}\times\mathbf{B}) = \gamma(1-\mathrm{v})\mathbf{E} \tag{1.62a}$$

$$\mathbf{B}' = \gamma(\mathbf{B} - \mathbf{v}\times\mathbf{E}) = \gamma(1-\mathrm{v})\mathbf{B} \tag{1.62b}$$

with primed quantities representing quantities in the frame K' and $\gamma = 1/\sqrt{1-\mathrm{v}^2}$. Substitution shows that $\mathbf{E}'$ and $\mathbf{B}'$ satisfy

$$\nabla\times\mathbf{B}' = -k'\,\mathbf{B}' \tag{1.63}$$

with $k' = \gamma(k - \mathrm{v}\omega)$. This sheds further light on the fact that α in (1.3') is not a universal constant; it may change upon transformation. If the frame K' was moving in the positive y-direction, we get

$$\mathbf{B}' = \gamma[\sin(\omega t - kz)\hat{\mathbf{x}} + \cos(\omega t - kz)\hat{\mathbf{y}} + v\cos(\omega t - kz)\hat{\mathbf{z}}] \tag{1.64a}$$

and $$\mathbf{E}' = \gamma[\cos(\omega t - kz)\hat{\mathbf{x}} - \sin(\omega t - kz)\hat{\mathbf{y}} - v\sin(\omega t - kz)\hat{\mathbf{z}}] \tag{1.64b}$$

and again substitution shows that **E'** and **B'** satisfy

$$\nabla \times \mathbf{B}' = -k'\,\mathbf{B}' \tag{1.65}$$

with $k' = k$.

An example of case (ii) above is the wave of Chu and Ohkawa [1982] given by Eqs. (1.53).

1.6.6.2. Media with Sources

The effect of Lorentz transformations on the FF condition in a medium with sources will be demonstrated by an example. Consider the FF field obtained by Ferraro and Plumpton [1961]

$$\mathbf{B} = B_0[0,\ \sin(\alpha x),\ \cos(\alpha x)] \tag{1.66}$$

This field is time independent and hence is associated with the current density

$$\mathbf{j} = \frac{B_0}{\alpha}[0,\ \sin(\alpha x),\ \cos(\alpha x)] \tag{1.67}$$

Note that in this example, α has no physical significance other than being a constant of proportionality between two parallel vectors. In a frame K' moving with a speed v in the positive z-direction, $\alpha' = \alpha$, $x' = x$,

$$\mathbf{B}' = B_0[0,\ \gamma\sin(\alpha x),\ \cos(\alpha x)] \tag{1.68}$$

$$\mathbf{j}' = \frac{B_0}{\alpha}[0,\ \sin(\alpha x),\ \gamma\cos(\alpha x)] \tag{1.69}$$

$$\mathbf{E}' = B_0[-\gamma v\sin(\alpha x),\ 0\ ,0] \tag{1.70}$$

and unless $\gamma = 1$ ($v = 0$), $\mathbf{B}'$ and $\mathbf{j}'$ are not parallel.

In conclusion, it has been shown that the FF condition (1.3') is invariant under parity and time reversal and that Ampere's law, the parent equation of (1.3'), is invariant under Lorentz transformations although (1.3') is not necessarily invariant under Lorentz transformations.

1.7. APPLICATIONS OF FORCE-FREE MAGNETIC FIELDS

1.7.1. Solar Flares

As explained in the introduction to this chapter, FF magnetic fields were introduced to allow both large currents and large magnetic fields to exist in stellar material. For this reason, it is not surprising that solar physics is the field where FF fields have been most extensively used. Low [1977] showed that FF magnetic field with non-constant α can become a more complex field which may annihilate explosively to produce a solar flare. Solar flares are complicated since they exhibit transitions from conditions of slow passive evolution to fast active evolution. FF magnetic fields with constant α and with α varying from region to region have been successfully applied to explain the details of such processes [see VAN HOVEN 1981].

1.7.2. Superconductors

Achieving and maintaining a FF configuration in a superconductor leads to a large increase in the limit of its current carrying capacity. This section will explain why FF configurations are required if larger magnetic fields are to be generated from the available superconductors. In the simple derivation of the relation between the current density and the magnetic field in superconductors, the superconductor is assumed to be a slab in the yz- plane and the external magnetic field is assumed to be uniform and parallel to the z-axis. Using the London equation, the magnetic field inside the slab satisfies

$$\mathbf{B} = B_0 e^{-x/\lambda} \hat{z} \tag{1.71}$$

where λ is the penetration depth. Since there is no electric field inside the metal

$$\nabla \times \mathbf{B} = -\frac{\partial B_z}{\partial x}\hat{y} = j_y \hat{y} \tag{1.72}$$

i.e.,

$$\mathbf{j} = (B_0/\lambda)e^{-x/\lambda}\hat{y} \tag{1.73}$$

The Lorentz force, **L**, is

$$\mathbf{L} = \mathbf{j} \times \mathbf{B} = \left(B_0^2/\lambda\right)e^{-2x/\lambda}\hat{\mathbf{x}} \tag{1.74}$$

as shown schematically in Figure 1.1.

This force, **L**, acts on the super-current carriers, pushing them towards the inside of the superconductor, and this has the effect of destroying the force for a flat planar superconducting slab. The analysis of more complicated configurations leads to the same conclusion. The elimination or reduction of the Lorentz force, without reducing the field or the current, is desirable. This would require a special configuration of the external magnetic field. Josephson [1966] has explained that the force-balance equation $\mathbf{j} \times \mathbf{B} + \mathbf{P} = 0$, where **P** is the pinning force, suggests that with a finite pinning force it is possible to increase the critical current by changing the angle between **j** and **B**. This has been confirmed experimentally. The ideal configuration is a FF configuration. This configuration is impossible to achieve. Matsushita [1981] has shown that the FF current is unstable unless pinning centers stabilize it.

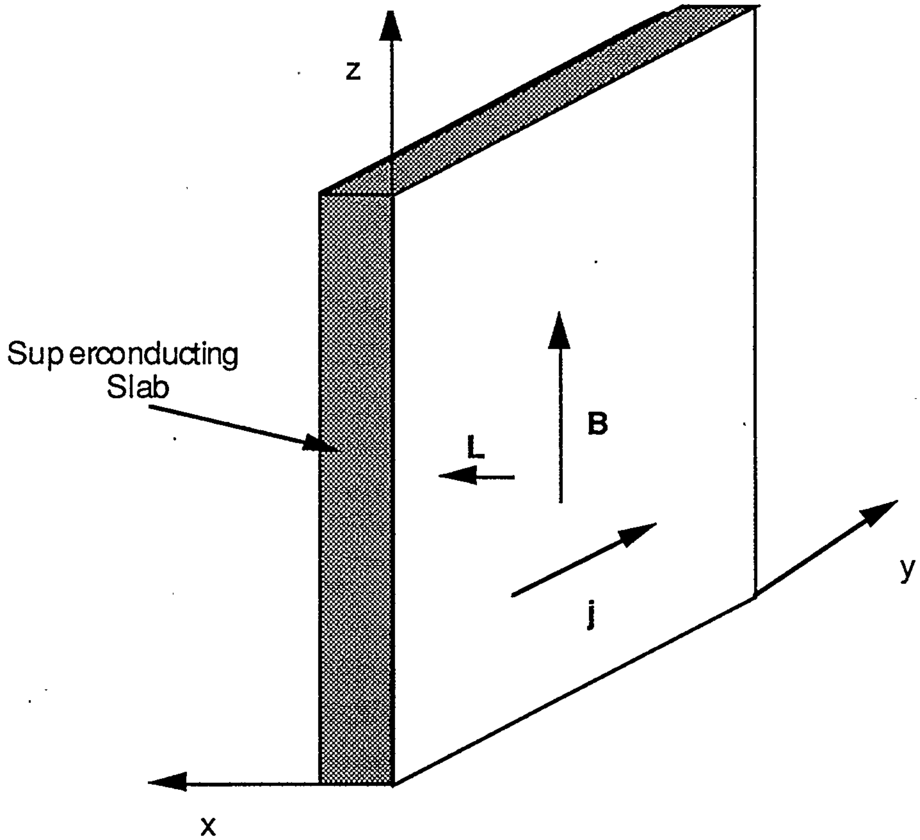

Figure 1.1 A schematic diagram showing the relationship between the magnetic field, the current and the Lorentz force.

1.7.3 Generation of Large Magnetic Fields

It has been stated earlier that it is impossible to construct current carrying coils of finite dimensions such that they are totally free of magnetic stresses. Current carrying coils must have some support to counter the magnetic stresses. This should be obvious when it is desired to generate a large magnetic field in some region of space for use in different applications: e.g., nuclear magnetic resonance (NMR) or electron paramagnetic resonance (EPR). In such applications, it is desirable to have the magnetic field in a region free of conductors. This means that in that region, the field is non-zero while no current is present. This application is in contrast with the plasma confinement case where it is desirable to enhance the FF configuration in the plasma and no free space magnetic field is desired. It is often the case in conventional coils generating large magnetic fields that the magnetic stresses exceed the tensile strength of the coil material [ZIJLSTRA 1967]. This problem is further complicated by the fact that large magnetic fields require large currents which in turn are associated with large ohmic losses and require substantial cooling. It is advantageous for such coils to shift the magnetic stresses to more convenient regions.

Furth et al. [1957] showed the design for an approximately FF configuration of coils. Yangfang and Luguang [1983] have demonstrated an implementation of this design using superconductors. Wells and Mill [1962] reported on their implementation of a force-reduced toroidal coil system. They included photographs showing that a conventionally wound torus failed at a field of 2.7 T. In the same set of photographs, the force-reduced design did not show appreciable deflection at 2.7 T.

1.7.4. Plasma Confinement

Taylor [1974] showed, assuming that the plasma internal energy is negligible compared with the magnetic energy, that the final state of relaxation will be the state of minimum energy subject only to the single invariant

$$K_0 = \int_V \mathbf{A} \cdot \mathbf{B} dV \tag{1.19}$$

where V is the total volume of the system. This state has been determined by Woltjer [1958a] and is a force-free magnetic field. Robinson [1969] reported that the longitudinal field and current are reversed near the wall in the Zeta Tokamak if the pinch parameter θ is close to 1.2. Taylor [1974] treated the plasma magnetic field as force-free and used the second set of cylindrical solutions of Ferraro and Plumpton [1961] to show that the field would be reversed at $\alpha\rho = 2.404$ corresponding to $\theta = 1.202$, which is in good agreement with Robinson's observations [ROBINSON 1969].

CHAPTER 2
MAGNETIC FIELD SOLUTIONS OF THE FORCE-FREE CONDITION

2.1. INTRODUCTION

This chapter will attempt to find the solution of the FF condition

$$\nabla \times \mathbf{B} = \alpha \mathbf{B}, \tag{1.3'}$$

subject to the constraint

$$\nabla \cdot \mathbf{B} = 0 \tag{2.1}$$

in rectangular, cylindrical and spherical coordinates. It will first review the known solutions and then derive a new solution in spherical coordinates that does not fit in the "general" solution provided by Chandrasekhar and Kendall [1957].

Taking the curl of (1.3') and invoking (2.1), shows that the magnetic field must satisfy the vector Helmholtz equation

$$\nabla^2 \mathbf{B} + \alpha^2 \mathbf{B} = 0 \tag{2.2}$$

Equation (2.2) can be solved if the method of Hansen [1935] is exploited; a good review of this method is given by Collin [1991]; if a scalar function, ψ, satisfies the scalar Helmholtz equation

$$\nabla^2 \psi + \alpha^2 \psi = 0 \tag{2.3}$$

then three and only three independent vectors can be formed from ψ such that they satisfy the vector Helmholtz equation. These three vectors are: (i) $\mathbf{L} = \nabla\psi$; (ii) $\mathbf{P} = \nabla \times \psi\mathbf{r}$, and

(iii) $\mathbf{T} = \nabla \times \nabla \times \psi\mathbf{r}$. Another such set may be formed by using **a** in place of **r**, where **a** is a constant vector. Note that **r** is not a constant vector. The general solution of (2.2) is then obtained by: (i) finding a complete set of solutions for the scalar wave equation, (ii) finding the three vector solutions for each member of this set and (iii) adding all these vectors. Now a general solution corresponding to a single function is

$$\mathbf{B} = a\mathbf{L} + b\mathbf{P} + c\mathbf{T} = a\nabla\psi + b\nabla \times \psi\mathbf{r} + c\nabla \times \nabla \times \psi\mathbf{r} \tag{2.4}$$

where a, b and c are constants.

We manipulated (1.3') and (2.1) to obtain (2.2). However, not every solution of (2.2) is necessarily a solution of (1.3') and (2.1). The two conditions; $\nabla \cdot \mathbf{B} = 0$ (implies $a = 0$), and $\nabla \times \mathbf{B} = \alpha\mathbf{B}$ (implies $b = \alpha c$):

$$\begin{aligned}
\nabla \times \mathbf{B} &= b\nabla \times \nabla \times \psi\mathbf{r} + c\nabla \times \nabla \times \nabla \times \psi\mathbf{r} \\
&= b\nabla \times \nabla \times \psi\mathbf{r} - c\nabla^2 \times \nabla \times \psi\mathbf{r} \\
&= b\nabla \times \nabla \times \psi\mathbf{r} + \alpha^2 c\nabla \times \psi\mathbf{r} \\
&= \alpha\mathbf{B} = \alpha b\nabla \times \psi\mathbf{r} + \alpha c\nabla \times \nabla \times \psi\mathbf{r}
\end{aligned} \tag{2.5}$$

Thus we look for solutions of the form

$$\mathbf{B} = B_0(\alpha\nabla \times \psi\mathbf{r} + \nabla \times \nabla \times \psi\mathbf{r}), \tag{2.6}$$

2.2. RECTANGULAR COORDINATES

Nakagawa and Raadu [1972] obtained general FF magnetic fields in rectangular coordinates:

$$B_x = \sum_{k \neq 0} \frac{i}{k^2}\left[ak_y - k_x\left(k^2 - a^2\right)^{1/2}\right] B_k e^{ik\cdot r - (k^2 - a^2)^{1/2} z},$$

$$B_y = \sum_{k\neq 0} \frac{-i}{k^2}\left[ak_x - k_y\left(k^2 - a^2\right)^{1/2}\right]B_k x e^{ik\cdot r - \left(k^2 - a^2\right)^{1/2} z},$$

$$B_z = \sum_{k\neq 0} B_k e^{ik\cdot r - \left(k^2 - a^2\right)^{1/2} z}, \tag{2.7}$$

where $\mathbf{k} = k_x\hat{\mathbf{x}} + k_y\hat{\mathbf{y}}$, $k_z^2 k^2 = k_x^2 + k_y^2$, $\mathbf{r} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}}$, the summation is over all possible values of both k_x and k_y excluding $k_x = k_y = 0$, and B_k are the Fourier coefficients of $B_z(x,\ y,\ z = 0)$. Ferraro and Plumpton [1961] obtained the single solution

$$\mathbf{B} = B_0[\sin 0,\ \sin(\alpha x),\ \cos(\alpha x)] \tag{2.8}$$

Note that in rectangular coordinates, (2.2) is actually three independent equations of the form (2.3). The first step in finding the most general solution to (2.2) is to find a complete set of solutions, ϕ, to (2.3). Each component of the solution of (2.2) is a sum of such functions.

2.3. CYLINDRICAL COORDINATES

Lust and Schlute [1954] gave the differential equation for an axially symmetric FF magnetic field in cylindrical coordinates and found some special analytical as well as numerical solutions. Chandrasekhar [1956] also developed the equations governing FF axially symmetric magnetic fields in cylindrical coordinates but did not give the solutions. Lundquist [1950] found particular solutions with radial and azimuthal components only

$$B_\rho(\mathbf{z}) = B_0 \sin(\alpha z),\ B_\phi(z) = B_0 \cos(\alpha z),\ B_z = 0 \tag{2.9}$$

and solutions with azimuthal and axial components only:

$$B_\rho = 0,\ B_\phi(\rho) = B_0 J_1(\alpha\rho),\ B_z(\rho) = B_0 J_0(\alpha\rho) \tag{2.10}$$

These last solutions have been used to explain a phenomenon in plasma confinement as has been explained in Chapter 1.

2.4. SPHERICAL COORDINATES

Chandrasekhar and Kendall [1957] (C&K) derived general FF magnetic fields in spherical coordinates.

$$B_r(r,\theta,\phi) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \frac{n(n+1)}{\alpha r} Z_n(\alpha r) P_n^m(x) \times \left[A_{n,m}\cos(m\phi) + B_{n,m}\sin(m\phi)\right],$$

$$B_\theta(r,\theta,\phi) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \Bigg(\frac{1}{\alpha r}\frac{d}{dr}[rZ_n(\alpha r)]\frac{\partial}{\partial\theta}P_n^m(x) \times \left[B_{n,m}\cos(m\phi) + A_{n,m}\sin(m\phi)\right]$$

$$+(m/\sin\phi) Z_n(\alpha r) P_n^m(x) \left[A_{n,m}\cos(m\phi) - B_{n,m}\sin(m\phi)\right] \Bigg)$$

$$B_\phi(r,\theta,\phi) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \Bigg(\frac{1}{\alpha r\sin\theta}\frac{d}{dr}[rZ_n(\alpha r)]\frac{\partial}{\partial\theta}P_n^m(x)\left[-A_{n,m}\cos(m\phi) + B_{n,m}\sin(m\phi)\right]$$

$$-Z_n(\alpha r) P_n^m(x)\left[B_{n,m}\cos(m\phi) + A_{n,m}\sin(m\phi)\right] \Bigg) \tag{2.11}$$

where $x = \cos\theta$, $Z_n(\alpha r) = (\pi/2\alpha r)^{1/2}\, J_{n+1/2}(\alpha r)$, $J_{n+1/2}(\alpha r)$ is a linear combination of the Bessel functions $J_{n+1/2}$ and $J_{-n-1/2}$, P_n^m is the associated Legendre polynomial of order m and degree n, and $A_{n,m}$ and $B_{n,m}$ are coefficients to be determined from boundary conditions. Other workers have attempted to obtain FF magnetic fields in spherical coordinates. Chandrasekhar [1957] earlier obtained axially symmetric FF magnetic fields in spherical coordinates. Nakagawa [1973] obtained the same solutions as those of C&K [1957] but with explicit coefficients. Freire [1966] obtained a field [Eqs.(46) of his paper, with the sign of B_ϕ corrected, and with **H** replaced by **B**]:

$$B_r = 0,\ B_\theta = (A/r)\sin(\alpha r),\ B_\phi = (A/r)\cos(\alpha r) \tag{2.12}$$

However, direct substitution shows that (2.12) does not satisfy the solenoidal condition

$$\nabla \cdot \mathbf{B} = 0, \tag{2.1}$$

and hence is not a magnetic field [BROWNSTEIN 1987b]. It appears that Freire [1966] assumed that if each field component does not depend on its own coordinate then the solution satisfies (2.1). However, this is not the case for spherical coordinates since in this system of coordinates, the divergence operator is given by

$$\nabla \equiv \frac{1}{r^2}\frac{\partial}{\partial r} r^2 \mathbf{u}_r + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\,\mathbf{u}_\theta + \frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\mathbf{u}_\phi, \tag{2.13}$$

where $\mathbf{u}_r$, $\mathbf{u}_\theta$ and $\mathbf{u}_\phi$ are unit vectors. The solution that Freire [1966] was seeking is interesting because it had a vanishing radial component and cannot be included in the "general" solution of C&K given in (2.11).

Zaghloul and Barajas [1990] (Z&B) obtained FF magnetic fields without a radial component, $\mathbf{B} = [0,\ B_\theta(r,\ \theta,\ \phi),\ B_\phi(r,\ \theta,\ \phi)]$ by assuming that these solutions are separable; $B_\theta(r,\ \theta,\ \phi) = B_\theta^r B_\theta^\theta B_\theta^\phi$ and $B_\phi(r,\ \theta,\ \phi) = B_\phi^r B_\phi^\theta B_\phi^\phi$, where the superscripts are labels indicating the only coordinate on which each function depends. The dependence of these solutions on r is similar to that derived by Freire [1966]. Freire solved the differential equations

$$\frac{d^2}{dr^2}(rB_\theta^r) = -\alpha^2 r B_\theta^r \tag{2.14a}$$

and
$$\frac{d^2}{dr^2}(rB_\phi^r) = -\alpha^2 r B_\phi^r \tag{2.14b}$$

to obtain

$$B_\theta^r = \left(\frac{1}{r}\right)\left(A_+ e^{i\alpha r} + A_- e^{-i\alpha r}\right), \tag{2.14a'}$$

$$B_\phi^r = \left(\frac{K}{\alpha r}\right)\left(\frac{d}{dr}\right)\left(r B_\theta^r\right)$$

$$= \left(\frac{i}{r}\right) K \left(A_+ e^{i\alpha r} - A_- e^{-i\alpha r}\right), \tag{2.14b'}$$

where A_+, A_-, and K are constants. Equation (2.1) leads to

$$\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\left(B_\phi \sin\theta\right) = -\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}B_\phi, \tag{2.15}$$

and (1.3') leads to the solutions given by (2.14) and to

$$\frac{1}{r^2\sin\theta}\left(r\frac{\partial}{\partial\theta}\left(B_\phi \sin\theta\right) - r\frac{\partial}{\partial\phi}B_\theta\right) = 0. \tag{2.16}$$

Equations (2.15) and (2.16) when multiplied by r $\sin\theta$ become

$$\frac{\partial}{\partial\theta}\left(B_\theta \sin\theta\right) = -\frac{\partial}{\partial\phi}B_\phi, \tag{2.15'}$$

$$\frac{\partial}{\partial\theta}\left(B_\phi \sin\theta\right) = \frac{\partial}{\partial\phi}B_\theta. \tag{2.16'}$$

Equations (2.15') and (2.16') are not valid if $r = 0$ or if $\theta = 0$ or π. It is obvious from the similarity of (2.15') and (2.16') that B_θ and B_ϕ have similar dependence on the angular coordinates, θ and ϕ. For this reason, only the solutions of (2.16') will be derived and those of (2.15') will be inferred from them. Algebraic and differential manipulation as well as using the separation of variables mentioned earlier lead to

$$\frac{d}{d\theta}\left(\sin\theta\frac{d}{d\theta}\left(B_\theta^\theta \sin\theta\right)\right) = m^2 B_\theta^\theta \tag{2.17}$$

and

$$\frac{d^2}{d\phi^2}B_\theta^\phi = -m^2 B_\theta^\phi, \tag{2.18}$$

where m is a real constant. Equation (2.18) has the solution

$$B_\theta^\phi = B_{0+}^\phi e^{i(m\phi+\delta_{m+})} + B_{0-}^\phi e^{-i(m\phi-\delta_{m-})} \tag{2.18'}$$

where B_{0+}^ϕ, B_{0-}^ϕ, δ_{m+}, and δ_{m-} are real constants. The requirement of single valuedness of the function B_θ^ϕ forces m to be an integer. The solution of (2.17) can be brought about by using a variable γ in place of θ, where $\gamma = \ln[\tan(\theta/2)]$; i.e., $d\gamma = (1/\sin\theta)d\theta$. Consequently,

$$\frac{d^2}{d\gamma^2}B_\theta'^\theta = m^2 B_\theta'^\theta, \tag{2.19}$$

where $B_\theta'^\phi = B_\theta^\phi \sin\theta$. Now

$$B_\theta'^\theta = B^\theta{}_+ e^{m\gamma} + B^\theta{}_- e^{-m\gamma}, \tag{2.20}$$

where B_+^θ and B_-^θ are constants. Reintroducing θ in (2.20) gives

$$B_\theta^\theta = \left\{B_+^\theta[\tan(\theta/2)]^m + B_-^\theta[\cot(\theta/2)]^m\right\}\csc\theta. \tag{2.21}$$

The real solutions are

$$B_\theta(r,\ \theta,\ \phi) = \frac{B_{m+}}{r\sin\theta}[\tan(\theta/2)]^m \sin(\alpha r + m\phi + \delta_{m+})$$

$$+\frac{B_{m-}}{r\sin\theta}[\cot(\theta/2)]^m \sin(\alpha r - m\phi + \delta_{m-}), \tag{2.22}$$

$$B_\phi(r,\ \theta,\ \phi) = \frac{B_{m+}}{r\sin\theta}[\tan(\theta/2)]^m \cos(\alpha r + m\phi + \delta_{m+})$$

$$+\frac{B_{m-}}{r\sin\theta}[\cot(\theta/2)]^m \cos(\alpha r - m\phi + \delta_{m-}), \tag{2.23}$$

where B_{m+} and B_{m-} are coefficients related to the earlier coefficients. The more general compact solution is

$$B_\theta(r,\ \theta,\ \phi) = \sum_{m=-\infty}^{m=+\infty} \frac{B_m}{r\sin\theta}\left(\tan\frac{\theta}{2}\right)^m \sin(\alpha r + m\phi + \delta_m) \tag{2.24}$$

$$B_\phi(r,\ \theta,\ \phi) = \sum_{m=-\infty}^{m=+\infty} \frac{B_m}{r\sin\theta}\left(\tan\frac{\theta}{2}\right)^m \cos(\alpha r + m\phi + \delta_m) \tag{2.25}$$

These solutions:

(1) satisfy both the solenoidal and the FF conditions,

(2) are singular at $\theta = 0$ or π. However, they are not valid on the axis since it was excluded in the derivation,

(3) are singular at $r = 0$ which is also excluded from the derivation,

(4) for $m > 0$, the solutions are defined at $\theta = 0$ and actually both components vanish at $\theta = 0$ for $m>1$. The $m=1$ solution is interesting since it must be independent of ϕ when $\theta = 0$. This is the case since

$$\mathbf{u}_\theta = \cos\phi\hat{\mathbf{x}} + \sin\phi\hat{\mathbf{y}} \quad \text{and} \quad \mathbf{u}_\phi = -\sin\phi\hat{\mathbf{x}} + \cos\phi\hat{\mathbf{y}} \tag{2.26}$$

leading to

$$\mathbf{B} = \frac{1}{r}\left[\sin(\alpha r+\delta_1)\,\hat{\mathbf{x}} - \cos(\alpha r+\delta_1)\,\hat{\mathbf{y}}\right] \tag{2.27}$$

which is independent of ϕ;

(5) have, if only a single solution with a particular m is considered, a magnitude that is independent of ϕ

$$B = \frac{B_m}{r\sin\theta}\tan\left(\frac{\theta}{2}\right)^m \tag{2.28}$$

which for m = 1 reduces to

$$B = \frac{B_1}{r\sin\theta}\tan\left(\frac{\theta}{2}\right) = \frac{B_1}{2r\cos\left(\frac{\theta}{2}\right)^2} \tag{2.29}$$

and

(6) have the coefficients B_m which are to be determined from boundary conditions and may be time dependent.

Evangelidis [1975] obtained spherical coordinates solutions with no radial component as well as solutions with only a radial component. However, they were obtained for the case of non-constant α.

If $B_\theta^\theta B_\theta^\phi \sin\theta$ is considered one function, then we obtain equation (35) of Schelkunoff [1938]

$$B_\theta^\theta B_\theta^\phi \sin\theta = f_1\left(e^{i\phi}\tan\frac{\theta}{2}\right) + f_2\left(e^{-i\phi}\tan\frac{\theta}{2}\right) \tag{2.30}$$

where f_1 and f_2 are arbitrary functions. However, note that the solution of Z&B is a series expansion of (2.30). Schelkunoff [1938] obtained these solutions in his search for transverse electric spherical waves; i.e., waves with no electric radial component. This is not surprising since Z&B obtained their solutions in their search for magnetic fields with no radial component. Schelkunoff [1952] later found that typical functions of the complex variable $z = e^{i\phi}\tan(\theta/2)$ are TEM solutions to the Helmholtz equation in spherical coordinates. Schelkunoff [1952] also derived the solution of Z & B as a solution for the associated Legendre equation for the case n = 0. Schelkunoff identified that in this case the radial components of both the electric and magnetic fields vanish and that the waves are TEM waves.

The general solutions for B_θ and B_ϕ, in compact form largely due to Schelkunoff [1938], are

$$B_\theta = \frac{1}{r\sin\theta}\left[A_+ e^{i\alpha r} f_1\left(e^{i\phi}\tan\frac{\theta}{2}\right) + A_- e^{-i\alpha r} f_2\left(e^{-i\phi}\tan\frac{\theta}{2}\right)\right] \tag{2.31}$$

$$B_\phi = \frac{1}{r\sin\theta}\left[A_+ e^{i\alpha r} f_1\left(e^{i\phi}\tan\frac{\theta}{2}\right) - A_- e^{-i\alpha r} f^2\left(e^{-i\phi}\tan\frac{\theta}{2}\right)\right] \tag{2.32}$$

The real solutions will depend on the functions f_1 and f_2.

It is concluded that the solutions of C&K [1957] and those of Nakagawa [1973] are not the most general. The most general solution of the FF condition in spherical coordinates will be investigated and derived in the next section.

2.5. THE GENERAL FORCE-FREE MAGNETIC FIELD IN SPHERICAL COORDINATES

In Section 2.4., it was shown that the spherical FF field without a radial component

$$B_\theta(r,\ \theta,\ \phi) = \sum_{m=-\infty}^{m=+\infty} \frac{B_m}{r\sin\theta}\left(\tan\frac{\theta}{2}\right)^m \sin(\alpha r + m\phi + \delta_m) \tag{2.33}$$

$$B_\phi(r,\ \theta,\ \phi) = \sum_{m=-\infty}^{m=+\infty} \frac{B_m}{r\sin\theta}\left(\tan\frac{\theta}{2}\right)^m \cos(\alpha r + m\phi + \delta_m) \tag{2.34}$$

cannot be included in the "general" FF of C&K. This raises the question: what is the most general FF field in spherical coordinates? This will be addressed through (i) checking whether the solution of Z&B fits in Hansen's form, (ii) finding the scalar function that generates the solution of Z&B, (iii) finding the general ϕ that is a solution to the Helmholtz equation, and (iv) deriving the general solution of the FF condition in spherical coordinates.

We will first find the solution of Z&B using the method of Hansen [1935]. Letting

$$\mathbf{B} = \nabla \times r\varphi \mathbf{u}_r + \nabla \times \nabla \times r\varphi \mathbf{u}_r \tag{2.35}$$

and setting B_r to zero, we get

$$\frac{d}{d\theta}\left(\sin\theta \frac{d\varphi}{d\theta}\right) + \frac{1}{\sin\theta}\frac{d^2\varphi}{d\phi^2} = 0 \tag{2.36}$$

where φ is a solution of the SHE. Equation (2.36) is, through simple manipulation, identical to (2.17) and has the solution

$$\varphi = \sum_{m=-\infty}^{\infty} \frac{B_m}{r}\left(\tan\frac{\theta}{2}\right)^m \sin(\alpha r + m\phi + \delta_m) \tag{2.37}$$

This suggests that Hansen's method is valid and that the limitation on the "general" solution of C&K may have been introduced through the particular solution of the SHE used by Hansen and later by C&K. We will now attempt to find the general solution of the VHE. We first notice that (2.14a) is the zeroth order spherical Bessel differential equation which, in agreement with the solution of Z&B, has a solution $\sin(\alpha r)/r$. We also notice that equation (2.17) is the zero order Legendre equation (with m = 0). This means that the solutions of Schelkunoff and Z&B are solutions of the zero order Legendre equation. The general solution of Legendre's equation is

$$\psi_n^m = p_n^m P_n^m + q_n^m Q_n^m \tag{2.38}$$

where p_n^m and q_n^m are constants and $Q_n^m(x)$ are Legendre's polynomials of the second-kind. The solution, $Q_n^m(x)$, is generally ignored because of its singularity at $x = \pm 1$.

This singularity can be avoided if the $\theta = 0$ and π axes are not included in the solution. It may be appropriate to include $Q_n^m(x)$ in the general solution once the singularities are removed from the usable region of the solution. However, $Q_0(x) = \frac{x+1}{x-1}$ cannot be used to construct the $(\tan(\theta/2))^m$ of Z&B. The relationship between the $(\tan(\theta/2))^m$ of Z&B and the known solutions of the Helmholtz equation in spherical coordinates will now be derived.

2.5.1. Solutions of the Scalar Helmholtz Equation in Spherical Coordinates

The scalar Helmholtz equation in spherical coordinates can be separated into a radial part and an angular part. The solution of the angular part is the well known spherical wave functions that are formed using the associated Legendre polynomials of the first and second kinds and the solutions of the radial part are the Hankel functions of the first and second kinds. We will now briefly review the associated Legendre polynomials of the first and second kinds.

2.5.1.1. Associated Legendre Polynomials of the First and Second Kinds

We first search for the general solution, P, for the associated Legendre equation (the Legendre equation with $n \neq 0$)

$$(1-x^2)\frac{d^2P}{dx^2} - 2x\frac{dP}{dx} + [n(n+1)(1-x^2) - m^2]P = 0 \tag{2.39}$$

The method of Frobenius where the solution is assumed to be expandable as a Taylor series

$$P = \sum_{i=0}^{\infty} a_i x^i \tag{2.40}$$

Substitution leads to the following recursion relationship

$$a_{i+2} = \frac{m^2 - n(n+1) + 2i^2}{i^2 + 3i + 2} a^i - \frac{i^2 - 3i + 2 - n(n+1)}{i^2 + 3i + 2} a_{i-2} \tag{2.41}$$

This recursion relationship leads to two independent solutions (series): one with even powers of x and one with odd powers of x. These series have some properties:

(1) For a given n and m where n≥m, one of the two series will truncate at some power of x whereas the other series will be an infinite series. The series that truncates is labeled P_n^m and the other is labeled Q_n^m. These series are known as the associated Legendre polynomials of the first and second kind respectively.

(2) For a given n, the odd series truncate if $-n \leq m \leq n$ and m is odd; whereas the even series truncates if $-n \leq m \leq n$ and m is even. If m>n or m<-n, both odd and even series are infinite. It is customary to maintain the convention that for odd m, the odd series is P_n^m and for even m, the even series is P_n^m.

(3) The infinite series diverges at $x = \pm 1$.

(4) The P_n^m polynomials are the common Legendre polynomials and form a complete set if $-n \leq m \leq n$. This completeness means that any square integrable function of θ can be expanded in terms of a sum of these polynomials.

(5) Q_n^m and P_n^m are orthogonal to each other for each integer n and m since one is an even series and the other is an odd series.

2.5.2. The Solutions of Z&B in Terms of Spherical Solutions of the Scalar Helmholtz Equation

2.5.2.1. The Angular Part

In section 2.4., it was stated that Schelkunoff [1952] derived the solutions (2.24) and (2.25) which were later derived by Z&B. Schelkunoff derived (2.24) and (2.25) as solutions of the associated Legendre equation for the case n = 0. In this subsection, we will show that tan(θ/2) is a solution of the associated Legendre equation for the case n = 0 and m = 1. To show this one has to remember that

$$\tan(\frac{\theta}{2}) = \frac{\sin(\frac{\theta}{2})}{\cos(\frac{\theta}{2})} = \frac{2\sin(\frac{\theta}{2})\cos(\frac{\theta}{2})}{2\cos^2(\frac{\theta}{2})} = \frac{\sin\theta}{1+\cos\theta} = \frac{\sqrt{1-x^2}}{1+x}$$
$$= (1+x)^{-1}(1-x^2)^{1/2} = (1-x+x^2-x^3+x^4+\cdots)(1-\frac{x^2}{2}-\frac{x^4}{8}-\frac{x^6}{16}-\cdots)$$
$$= 1-x+\frac{x^2}{2}-\frac{x^3}{2}+\frac{3x^4}{8}-\frac{3x^5}{8}+\cdots$$
$$= (1-x)(1+\frac{x^2}{2}+\frac{3x^4}{8}+\cdots) \tag{2.42}$$

Given the recursion relationship (2.41), Setting n = 0 and m = 1 leads to

$$P = (a_0 + a_1 x)\left(1+\frac{x^2}{2}+\frac{3x^4}{8}+\ldots\ldots\right) \tag{2.43}$$

Further setting $a_0 = -a_1$ leads to

$$P = a_0 \tan(\frac{\theta}{2}) \tag{2.44}$$

In a similar manner, one can show that

$$P^m = L\tan(\theta/2)^m \tag{2.45}$$

where L is a constant.

It can now be concluded that the angular dependence of the solutions of Z&B can be included in the angular dependence of the solutions of C&K if the latter solutions are expanded to include the Legendre polynomials of the second kind. However it is important to stress, that

(1) the presence of $\sin\theta$ in the denominator of the solutions of C&K suggests that the solutions are singular at $\theta = 0$. This is not the case because for $m \neq 0$, the Legendre polynomials of the first kind always have a $\sin\theta$ term;

(2) even if the singularity at $\theta = 0$ can be removed by using appropriate boundaries for the magnetic material, the solution Q_n^m cannot be expanded in terms of P_n^m. This should be obvious since the two solutions have different symmetries around the $\theta = \pi/2$ plane; and

(3) Even though P_0^1 and Q_0^1 are singular at $\theta = 0$, the sum of the two polynomials is not singular at $\theta = 0$.

2.5.2.2. The Radial Part

The spherical Bessel differential equation is

$$\frac{dR^2}{dr^2} + \frac{2}{r}\frac{dR}{dr} + \left(1 - \frac{n}{r^2}\right)R = 0 \tag{2.46}$$

With n = 0, (2.46) reduces to

$$\frac{dR^2}{dr^2} + \frac{2}{r}\frac{dR}{dr} + r = 0 \tag{2.47}$$

which reduces to (2.14a) with simple manipulation and if αr is substituted for r. Equation (2.47) has known solutions which are the spherical Bessel functions, j_n, and the spherical Neumann functions, n_n. These solutions are generally combined to form the Hankel functions of the first- and second-kind which are defined as

$$h_n^{(1)} = j_n + in_n \quad \text{and} \quad h_n^{(2)} = j_n - in_n \tag{2.48}$$

Since $j_0 = \dfrac{\sin r}{r}$ and $n_0 = \dfrac{\cos r}{r}$, substitution shows that (2.14a') is a sum of $h_0^{(1)}$ and $h_0^{(2)}$. This means that the radial dependence in the solutions of Z&B is contained in the general solution of the spherical Bessel differential equation.

2.5.3. The General Force-Free Magnetic Field in Spherical Coordinates

The analysis in subsections 2.5.1. and 2.5.2. showed that the solutions of Z&B can be included in the solutions of C&K if the associated Legendre polynomials of the second kind and the spherical Neumann functions are added to the associated Legendre polynomials of the first kind and the spherical Bessel functions respectively. This means that the most general solution for the SHE in spherical coordinates is

$$\psi = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \left(a_n h_n^{(1)} + b_n h_n^{(2)}\right)\left(c_{nm} P_m^n + d_{nm} Q_m^n\right) e^{im\phi} \tag{2.49}$$

where a_n, b_n, c_{nm} and d_{nm} are constants to be determined from the boundary conditions. Accordingly, the most general solution of the vector Helmholtz equation is

$$\mathbf{C} = d\nabla\psi + e\nabla \times \psi\mathbf{r} + f\nabla \times \nabla \times \psi\mathbf{r} \tag{2.50}$$

or

$$\mathbf{C} = g\nabla\psi + h\nabla \times \psi\mathbf{a} + l\nabla \times \nabla \times \psi\mathbf{a} \tag{2.51}$$

where d, e, f, g, h and l are constants to be determined from the boundary conditions and **a** is a constant vector. The general constant vector in spherical coordinates is

$$\mathbf{a} = (\lambda \sin\theta \cos\phi + \mu \sin\theta \sin\phi)\mathbf{u}_r + (\lambda \cos\theta \cos\phi + \mu \cos\theta \sin\phi)\mathbf{u}_\theta$$
$$+ (\mu \cos\phi - \lambda \sin\phi)\, \mathbf{u}_\phi \tag{2.52}$$

where λ, μ and ν are constants. Note that Hansen [1935] originally used the radial vector, $\mathbf{r}$, in his solution. C&K suggested that a constant vector should be used but actually used $\mathbf{r}$. Stratton [1941] has shown that although $\mathbf{r}$ is not constant, it is a valid choice. Flammer [1957] stated that the only known vectors that, combined with ψ as in (2.50) and (2.51), lead to solutions of the VHE are $\mathbf{r}$ and $\mathbf{a}$. The rules that such vectors obey are not stated in literature and are currently under investigation by the author.

The most general known magnetic fields satisfying $\nabla\times\mathbf{B}=\alpha\mathbf{B}$ in spherical coordinates are

$$\mathbf{B} = \nabla\times\psi\mathbf{r} + \nabla\times\nabla\times(\psi/\alpha)\mathbf{r} \tag{2.53}$$

and
$$\mathbf{B} = \nabla\times\psi\mathbf{a} + \nabla\times\nabla\times(\psi/\alpha)\mathbf{a} \tag{2.54}$$

Equation (2.53) leads to fields (the form is largely due to Rosenbluth and Bussac [1979])

$$\begin{aligned}
B(r,\,\theta,\,\phi) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\ & (c_m e^{im\phi} + d_{-m}e^{-im\phi}) \times \\
& \Big\{\Big[\frac{n(n+1)}{\alpha r}(a_n j_n(\alpha r) + b_n n_n(\alpha r))(p_n^m P_n^m + q_n^m Q_n^m)\Big]\mathbf{u}_r \\
& + \Big[\frac{im}{\sin\theta}(p_n^m P_n^m + q_n^m Q_n^m)(a_n j_n + b_n n_n) \\
& \quad - \frac{d}{d\theta}(p_n^m P_n^m + q_n^m Q_n^m)\frac{1}{\alpha r}\frac{d}{dr}(a_n r j_n + b_n r n_n)\Big]\mathbf{u}_\theta \\
& + \Big[\frac{im}{\sin\theta}(p_n^m P_n^m + q_n^m Q_n^m)\frac{1}{\alpha r}\frac{d}{dr}(a_n r j_n + b_n r n_n) \\
& \quad + \frac{d}{d\theta}(p_n^m P_n^m + q_n^m Q_n^m)(a_n j_n + b_n n_n)\Big]\mathbf{u}_\phi\Big\}
\end{aligned} \tag{2.55}$$

In conclusion, this chapter has derived the most general FF magnetic field in spherical coordinates. The next chapter will attempt to find an application where the new general solution leads to a different result than the "general" solution of C&K.

CHAPTER 3
APPLICATION OF FORCE-FREE MAGNETIC FIELDS TO PLASMA CONFINEMENT

3.1. NUCLEAR FUSION AS AN ALTERNATE ENERGY SOURCE

It is estimated that the current annual energy consumption is about 0.3 Q $\left(1\ \mathrm{Q} \equiv 1.05 \times 10^{21}\ \mathrm{J}\right)$ and that the there is only 11 Q of oil reserves and an estimated 100 Q of coal that can be mined economically. Given these estimates together with the fact that more and more countries are becoming industrialized which could bring the annual energy consumption to 0.7 Q, it is natural to speculate that fossil fuel reserves of the earth cannot supply the energy needed to support the continued consumption of advanced countries for more than a century. This speculation has encouraged scientists to search for an alternate fuel source. This search has led to the use of nuclear fission where one heavy nucleus breaks into smaller and more stable nuclei. An example of such a fission is the reaction

$$ {}^{235}_{92}\mathrm{U} + {}^{1}_{0}\mathrm{n} = 2 \times {}^{118}_{50}\mathrm{Sn} + 8\mathrm{e} + 236\mathrm{Mev} \tag{3.1} $$

However, in this case, the fuel is the uranium isotope ${}^{235}_{92}\mathrm{U}$ and it is estimated that there is 2.5×10^{10} kg uranium on earth; of which only 0.7% is ${}^{235}_{92}\mathrm{U}$ and since the reactor efficiency is 20%, it is estimated that all the ${}^{235}_{92}\mathrm{U}$ in the world can only supply 7 Q. A more efficient reaction is the fast breeder reaction (used in the super Phoenix in France)

$$ {}^{238}_{92}\mathrm{U} + {}^{1}_{0}\mathrm{n} \Rightarrow {}^{239}_{92}\mathrm{U} \Rightarrow {}^{239}_{93}\mathrm{Np} \Rightarrow {}^{239}_{94}\mathrm{Pu} \tag{3.2} $$

The world supply of ${}^{238}_{92}U$ would generate 300 Q. This makes fast breeder reactors seem like the choice for the immediate future. However, ${}^{239}_{94}Pu$ is radio-active. The waste of a fission reactor poses a serious limitation on the usefulness of such reactors.

An alternate source of energy is nuclear fusion where two light nuclei fuse together forming a heavier and more stable nucleus. A host of such reactions are possible; a promising candidate is

$$ {}^{2}_{1}D + {}^{3}_{1}T = {}^{4}_{2}He + {}^{1}_{0}n + 17.6\,Mev \tag{3.3} $$

where ${}^{3}_{1}T$ is the hydrogen isotope tritium, ${}^{2}_{1}D$ is the hydrogen isotope deuterium. Tritium is a radio-active material with a half life of 12 years; there is no natural tritium. Tritium can be produced through the nuclear reactions

$$ {}^{6}_{3}Li + {}^{1}_{0}n = {}^{4}_{2}He + {}^{3}_{1}T + 4.8\,Mev \tag{3.4} $$

$$ {}^{7}_{3}Li + {}^{1}_{0}n = {}^{4}_{2}He + {}^{3}_{1}T + {}^{1}_{0}n - 2.47\,Mev \tag{3.5} $$

Given the world supply of lithium and deuterium, it is estimated that the three reactions (3.3), (3.4) and (3.5) can generate 6×10^{6} Q. However, even though these reactions are free of radio-active waste, water containing ${}^{3}_{1}T$ is dangerous since it can harm the reproductive organs through its β-decay if present in humans.

3.2. THE CONFINEMENT PROBLEM

Reaction (3.3), as a prototype of nuclear reactions involving at least two species of nuclei, is promising for energy production. However, to release this energy, enough energy must be delivered to the nuclei so that they can overcome the Coulomb barrier

$$ U_0 = \frac{Z_1 Z_2 e^2}{4\pi\varepsilon_0 r_0} \tag{3.6} $$

where Z_1 and Z_2 are the number of charges on the two nuclei, e is the charge of an electron and r_0 is the radius of a nucleus (approximately 5×10^{-15} m). For deuterium and tritium ($Z_1 = Z_2 = 1$) this barrier is about 0.29 MeV. Overcoming the Coulomb barrier is generally achieved through heating a plasma rich with the two species. It is possible to assume that all gases become plasma above 10^6 K. The temperature needed to overcome the Coulomb barrier can be derived from the fact that the kinetic energy, K.E., of a particle at a temperature T is

$$\mathrm{K.E.} = \frac{1}{2}mv^2 = \frac{3}{2}kT \tag{3.7}$$

where k is Boltzmann's constant ($K = 1.3709 \times 10^{-23}$ J/K). The temperature required to overcome the 0.29 MeV barrier is

$$T = \frac{2 \times 0.29 \times 10^6 \times 1.602 \times 10^{-19}}{3 \times 1.3709 \times 10^{-23}} = 2.25 \times 10^9 \mathrm{K} \tag{3.8}$$

This temperature is realized through external heating that is applied to the plasma. The value of this temperature raises a serious problem since most known metals melt if T > 3500 K. Consequently a method must be found to prevent the plasma from coming into contact with the walls of the container. Magnetic and inertial confinement are two tested methods of plasma confinement [NIU 1988].

3.3. MAGNETIC CONFINEMENT

A single particle of charge e traveling at a constant velocity **v** in a uniform magnetic field **B** will experience a Lorentz force $e\mathbf{v} \times \mathbf{B}$. This force is perpendicular to the velocity of the particle and does not change the magnitude of this velocity; it only introduces curvature to the velocity. This means that the kinetic energy of a charged particle is unaltered by its presence in a uniform magnetic field. If $\mathbf{v} \perp \mathbf{B}$ then the path of the particle would describe a circle whose axis is the magnetic field direction and whose radius is $r_c = \frac{mv}{eb}$ where m is the mass of the particle. The angular velocity of this motion $\omega_c = \frac{eB}{m}$ is called the Larmor frequency. The Lorentz force vanishes if $\mathbf{v} \parallel \mathbf{B}$, This means that an arbitrary constant velocity would have a resultant whose component parallel to the field, $v_{\parallel}$, is unaltered by the presence of the field and whose component perpendicular to the field, $v_{\perp}$, describes circular motion about the field; the particle cannot drift away from the field lines. However, the particle is free to drift along the field lines.

3.3.1. Magnetic Mirrors

If the magnetic field was parallel to the axis of a cylinder containing the charged particles and if the field increased towards the edges of the cylinder, then according to $\nabla \cdot \mathbf{B} = 0$, the magnetic field must have a radial component with

$$\frac{1}{r}\frac{\partial}{\partial r}(rB_r) + \frac{\partial B_z}{\partial z} = 0 \tag{3.9}$$

If $\frac{\partial B_z}{\partial z}$ is constant, then

$$B_r = -\frac{1}{2} r \frac{\partial B_z}{\partial z} \tag{3.10}$$

and the z-component of the Lorentz force

$$m\frac{dv_z}{dt} = ev_\phi B_r = -\frac{1}{2} ev_\phi r \frac{\partial B_z}{\partial z} \tag{3.11}$$

which means that the particle will be subject to a force sending it towards the weaker magnetic field. If both ends of the cylinder have stronger fields as may be realized by winding a coil as shown in Figure 3.1, then the particles will tend to stay inside the cylinder away from the walls provided that

(i) the field increases slowly as a function of the distance along the axis of the cylinder, z; and

(ii) the radius of the cylinder is larger than the Larmor radius.

This process describes the concept of magnetic mirrors which are a means of magnetic plasma confinement [REITZ AND MILFORD 1967]. Note that

(i) particles of opposite charges travel in opposite directions, hence all particles, in the described case, will experience a force pushing them towards the weaker field;

(ii) particles traveling with large $v_{||}$, will not be reflected back by the magnetic mirror; and

(iii) the particles are not allowed to cross magnetic field lines and thus travel on the surface of flux tubes.

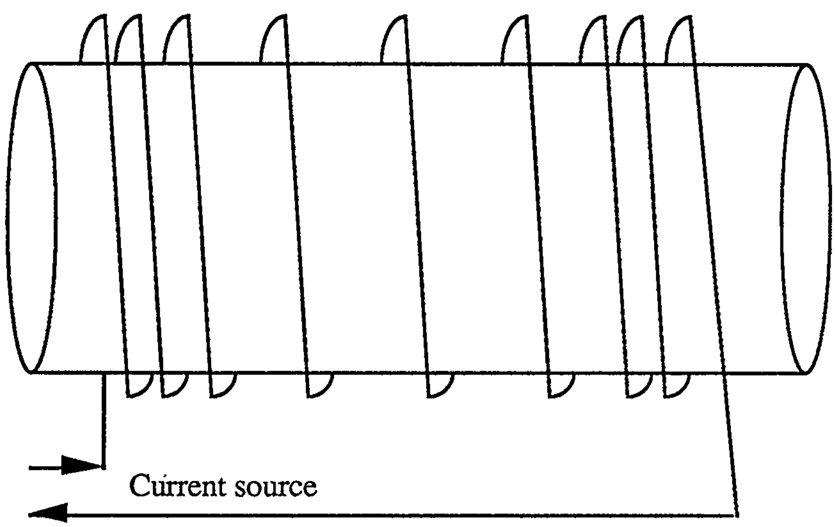

Figure 3.1 Schematic diagram showing how a coil could be wound to make a magnetic mirror.

3.3.2. Toroidal Confinement

Another method of plasma confinement is by bending the magnetic field lines into circles of radius R. This way, the particles will be confined to the interior of a torus whose major radius is the radius of the field circle and whose minor radius is the Larmor radius. However, this confinement is not perfect since bending the field into circles leads to that the particles traveling along the field lines will experience a centripetal acceleration $\frac{v_{\parallel}^2}{R}$. This acceleration is accounted for by a Lorentz force due to a component of the velocity that is perpendicular to the magnetic field, v_D. This means that

$$ev_D B = m\frac{v_{\parallel}^2}{R} \tag{3.12}$$

or
$$v_D = \frac{mv_{\parallel}^2}{eBR} \tag{3.13}$$

Since this speed depends on the square of $v_{\parallel}$ then its direction will be dictated by the sign of the charge; positive charges will tend to flow downwards and negative charges will tend to flow upwards. This charge separation introduces an electric field which is perpendicular to the magnetic field and $\mathbf{E} \times \mathbf{B}$ leads to a drift motion. This drift motion can be made to cancel the charge separation by making the field lines not close on themselves; rather after a complete rotation around the torus the field line shifts by an angle $\iota = 2\pi/n$ (n is a real number) called the rotational transfer angle. This shift in magnetic field lines makes the direction of the $\mathbf{E} \times \mathbf{B}$ drift motion isotropic [HAGLER AND KRISTIANSEN 1977].

3.3.3. The Magnetic Pinch Effect

The confinement of a current filament by its own self-magnetic field is known as the pinch effect [SPITZER 1962]. A plasma confined by its own magnetic field is thus said to be self-pinched. Thus, the pinch effect is a phenomenon that allows a magnetic field to apply a pressure on a plasma that could counter the plasma pressure. This pinch results when a magnetic field rapidly builds up and is not able to penetrate into the plasma due to the freezing of magnetic field lines in a plasma of infinite conductivity. The pinch effect is a useful phenomenon in plasma confinement.

To understand the pinch effect, first, the freezing of magnetic field lines in a plasma of infinite conductivity is explained. The total flux crossing an area A is given by

$$\phi = \int \mathbf{B} \cdot d\mathbf{A} \tag{3.14}$$

where dA is an element of the area; the flux has a time derivative

$$\frac{d\phi}{dt} = \int \frac{d\mathbf{B}}{dt} \cdot d\mathbf{A} + \int \mathbf{B} \cdot d\frac{d\mathbf{A}}{dt} \tag{3.15}$$

At every point on the boundary of the area, the rate of change of the area is proportional to the component of the velocity perpendicular to the boundary at that point, **dl**. This means that

$$d\frac{d\mathbf{A}}{dt} = \mathbf{v} \times d\mathbf{l} \tag{3.16}$$

also if the conductivity is infinite, the total electric force per unit charge, $\mathbf{E} + \mathbf{v} \times \mathbf{B}$, must tend to zero and from Maxwell's equation

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times (\mathbf{v} \times \mathbf{B}) \tag{3.17}$$

Equations (3.16) and (3.17) when substituted into (3.15) lead to

$$\frac{d\phi}{dt} = \int \nabla \times (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{A} + \int \mathbf{B} \cdot (\mathbf{v} \times d\mathbf{l}) \tag{3.18}$$

applying Stoke's theorem to the first integral on the right hand side yields

$$\frac{d\phi}{dt} = \int (\mathbf{v}\times\mathbf{B})\cdot d\mathbf{l} + \int \mathbf{B}\cdot(\mathbf{v}\times d\mathbf{l}) \tag{3.19}$$

which means that the flux does not change and that the magnetic field lines are frozen in the plasma [NICHOLSON 1983].

The magnetic pressure is related to the magnetic portion of the Lorentz force. It is customary to express the force acting on a plasma as

$$\mathbf{F} = \mathbf{j}\times\mathbf{B} - \nabla p \tag{3.20}$$

where p is the plasma pressure. Combining this with $\nabla\times\mathbf{B} = \mathbf{j}$ and the vector equality

$$\nabla\left(\frac{B^2}{2}\right) = \mathbf{B}\times\nabla\times\mathbf{B} + (\mathbf{B}\cdot\nabla)\mathbf{B} \tag{3.21}$$

one obtains,

$$\mathbf{F} = -\nabla\left(\frac{B^2}{2}\right) - \nabla p + (\mathbf{B}\cdot\nabla)\mathbf{B} \tag{3.22}$$

If $(\mathbf{B}\cdot\nabla)\mathbf{B}=0$, then

$$\mathbf{F} = -\nabla\left(\frac{B^2}{2}\right) - \nabla p \tag{3.23}$$

This suggests that the magnetic energy density, $\frac{B^2}{2}$, plays the role of a magnetic pressure.

Now, for a current traveling in the z direction of a cylindrical coordinate system, the magnetic field is in the azimuthal direction and can be calculated from Ampére's law

$$B(r) = \frac{1}{r}\int_0^r r'\, j(r')dr' \tag{3.24}$$

and in this case

$$\frac{\partial B(r)}{\partial r} = -\frac{1}{r^2}\int_0^r r' j(r')dr' + j(r) = -\frac{1}{r}B(r) + j(r) \tag{3.25}$$

The magnetic Lorentz force is in the radial direction and has the value

$$F = -B\frac{\partial B(r)}{\partial r} - \frac{B^2}{r} = \frac{\partial p}{\partial r} \tag{3.26}$$

which suggests that

$$p = \frac{B^2}{2} - \int_0^r \frac{B^2}{r'}dr' \tag{3.27}$$

and if the magnetic field increased suddenly, then by the freezing of the magnetic field lines, the field cannot penetrate deep into the plasma and we have a small surface region where there is magnetic pressure. This pressure is capable of balancing the plasma pressure and keeping the plasma away from the walls.

3.3.4. Mathematical Theory of Relaxed States

Taylor [1974] showed, assuming that the plasma internal energy is negligible compared with the magnetic energy, that the final state of relaxation will be the state of minimum energy subject only to the invariance of the magnetic helicity

$$K_0 = \int_V \mathbf{A} \cdot \mathbf{B} dV \tag{1.19}$$

where V is the total volume of the system. This state has been determined through the variational principle in Chapter 1.4. and is a constant α FF magnetic field [WOLTJER 1958a]. Robinson [1969] reported that the longitudinal field and current are reversed near the wall in the Zeta Tokamak if the pinch parameter $\theta \approx 1.2$. Taylor [1974] treated the plasma magnetic field as FF and used the second set of cylindrical solutions of Ferraro and Plumpton [1961] to show that the field would be reversed at $\alpha\rho = 2.404$ corresponding to $\theta = 1.202$, in agreement with the observations of Robinson [1969].

Bodenson et al. [1981] showed that the FF condition

$$\nabla \times \mathbf{B} = \alpha \mathbf{B} \tag{1.3'}$$

is an eigenvalue problem with an operator ∇, an eigenvector $\mathbf{B}$, and an eigenvalue α. Equation (1.3') is not a Hermitian equation. However, it can be repressed as

$$\nabla \times \nabla \times \mathbf{A}_i = \lambda_i \nabla \times \mathbf{A}_i \tag{3.28}$$

which is Hermitian and in that case, all eigenvalues are real, eigenvectors corresponding to distinct eigenvalues are orthogonal in the sense

$$\int \mathbf{A}_i \cdot \mathbf{A}_j dV = 0 \tag{3.29}$$

if $\lambda_i \neq \lambda_j$ and if more than one eigenvector correspond to the a single eigenvalue, these eigenvectors can be orthogonalized according to the Gramm-Schmidt orthogonalization. Note that (3.29) is an integration over all space. However, it can be satisfied if the volume is the interior of a perfectly conducting shell provided $\mathbf{A} \cdot \mathbf{n} = 0$ on the surface, where $\mathbf{n}$ is a vector perpendicular to the surface. This can be achieved since if $\mathbf{B}$ satisfies (1.3') then

$$\mathbf{B} = \alpha \mathbf{A} + \nabla \varphi \tag{1.39}$$

and a special choice of the scalar function (φ = constant) would lead to $\mathbf{A} \cdot \mathbf{n} = 0$ on the surface.

3.3.5. The Tokamak

The tokamak is a magnetic plasma confinement device that utilizes the fact that the plasma relaxes to a FF state if allowed to do so by the boundary conditions. The device is shown in Figure 3.2 [NIU 1988] and basically consists of a torus with a minor radius a and a major radius b as shown in Figure 3.3. The minor circle is surrounded by toroidal field coils that generate a magnetic field that has the direction of the major circle of the torus. The torus is surrounded by a transformer with primary windings supplied with an external current whereas the secondary windings of the transformer is the plasma. This transformer is the mechanism for generating a current along the major circle of the torus which in turn produces a poloidal field. The torus thus described may be approximated by a cylinder of radius a and of height b whose two ends are joined together. In treating the field generated within the torus cylindrical coordinates and cylindrical FF fields are used.

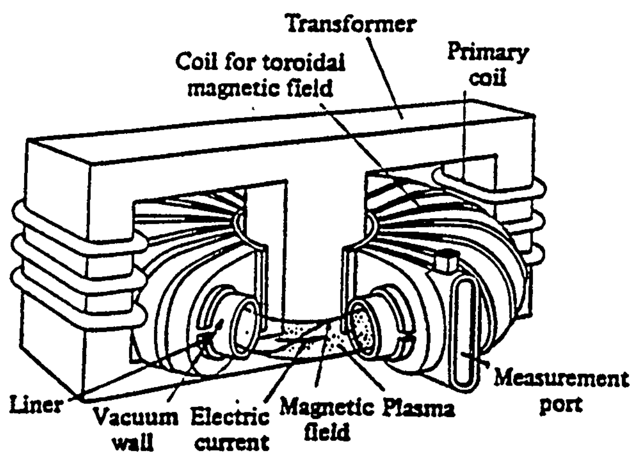

Figure 3.2 A schematic diagram showing the basic structure of a Tokamak [NIU 1988]

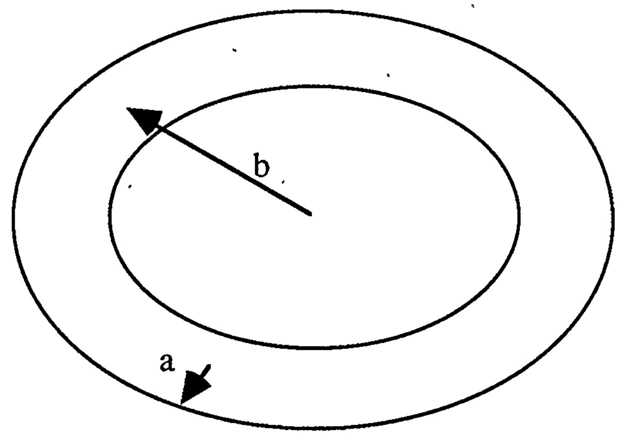

Figure 3.3 A schematic diagram showing a torus with a major radius b and a minor radius a

One of the biggest tokamak problems in the past was in the production of the toroidal field. This was due to the fact that only superconductors were capable of generating the required magnetic field (5-10 T). However, the reactor produces neutrons and other forms of radiation which could damage the superconducting wires and more importantly, the reactor operates at temperatures near 10^8 K whereas the superconducting wires have to operate below 20 K. However, the second of these problems is starting to be relaxed with the development of high-temperature superconductors. Note that the further the coils are from the plasma, the more the field they had to produce.

3.3.6. The Spheromak

The spheromak [ROSENBLUTH AND BUSSAC 1979] (R&B) is another magnetic plasma confinement device that utilizes the fact that the plasma relaxes to a FF state if allowed to do so by the boundary conditions. A typical spheromak has a conducting spherical shell surrounding the plasma. R&B showed that, for stability purposes, some ellipticity must be introduced to the spherical shell. A spheromak has a number of advantages over a tokamak:

(1) relative compactness;

(2) a toroidal blanket is not necessary;

(3) an external toroidal field is not required; and

(4) the volume of the plasma is simply connected which leads to simplifications in the mathematical analysis of the problem.

3.3.7. Force-Free Magnetic Fields in a Spheromak

R&B analyzed the magnetohydrodynamic stability of a spheromak assuming that relaxed states will follow Taylor's theory. They obtained a general solution similar to that originally given by C&K. They expressed their solution as a sum over different modes

$$\begin{aligned}\mathbf{B}_m = \sum_{n=m}^{\infty} b_n^m e^{im\phi}\Bigg\{&\left[n(n+1)P_n^m \frac{j_n(\alpha r)}{\alpha r}\right]\mathbf{u}_r \\ &+\left[\frac{im}{\sin\theta}P_n^m j_n(\alpha r) - \sin\theta P'^m_n \frac{1}{\alpha r}\frac{d}{dr}(r j_n(\alpha r))\right]\mathbf{u}_\theta \\ &+\left[\sin\theta P'^m_n j_n(\alpha r) + \frac{in}{\sin\theta}P_n^m \frac{1}{\alpha r}\frac{d}{dr}(r j_n(\alpha r))\right]\mathbf{u}_\phi\Bigg\}\end{aligned} \tag{3.30}$$

where $P'^m_n = \frac{d}{dx}P_n^m(x)$ and the b_n^m are coefficients to be determined from the boundary conditions. R&B noted that n = 0 and m = 0 are not allowed since they represent magnetic monopoles at the origin of the coordinate system. They stated that the classical spheromak equilibrium will be with n = 1 and m = 0. For simplicity, they assumed that $b_1^0 = 1$. If it is assumed that the plasma is contained within a conducting shell of radius R, then the normal component of the magnetic field at the shell's surface must vanish. This means that α must satisfy

$$j_1(\alpha R) = \frac{\sin(\alpha R)}{(\alpha R)^2} - \frac{\cos(\alpha R)}{\alpha R} = 0 \tag{3.31}$$

The first zero of $j_1(\alpha R)$ is at $\alpha R = 4.493$. They took the first zero of the spherical Bessel function because it leads to the smallest α and hence the lowest energy. This is evident from the proportionality of the total magnetic energy and the magnetic helicity

$$U_m = \int_V \frac{B^2}{2} dV = \int_V \frac{\alpha}{2} \mathbf{A} \cdot \mathbf{B} dV + \int \frac{1}{2} \nabla\varphi \cdot \mathbf{B} dV = \frac{\alpha}{2} K_0 + 0 \tag{3.32}$$

and if K_0 is a constant of the motion, then the smallest α corresponds to the lowest energy.

R&B studied the stability of oblate and prolate spheromaks whose surface can be given by

$$C(r, \theta) = r - R(1 + \varepsilon \cos^2\theta) = 0 \tag{3.33}$$

by finding the ratio $\delta\alpha/\alpha$ where $\delta\alpha = \alpha - \alpha_0$ where $\alpha_0 R = 4.493$. A particular configuration would be stable against higher order modes if $\delta\alpha/\alpha$ is positive since a positive ratio means that other modes lead to larger α's and consequently larger magnetic energies. The non-vanishing gradient of the function describing the surface, ∇C, at any point, (r_0, θ_0, ϕ_0), is a vector in the direction perpendicular to the surface at that point. Then, the condition for the vanishing of the magnetic field component perpendicular to the surface is given by

$$\mathbf{B} \cdot \nabla C = 0 \tag{3.34}$$

provided $\nabla C \neq 0$. Since

$$\nabla C = \mathbf{u}_r + 2\varepsilon \sin\theta \cos\theta \mathbf{u}_\theta \tag{3.35}$$

it is obvious that the radial component of ∇C can never vanish and in this case the condition for the vanishing of the magnetic field component perpendicular to the surface is

$$\mathbf{B}\cdot\nabla C = B_r + 2\varepsilon\sin\theta\cos\theta B_\theta = 0 \qquad (3.36)$$

A Taylor expansion of the spherical Bessel functions in (3.30) that includes terms of first order in $\delta\alpha$ and ε is

$$f(\alpha r) = f(\alpha_0 R) + (R\delta\alpha + \alpha_0 R\varepsilon\cos^2\theta)\frac{d}{d\alpha r}f(\alpha r)\big|_{\alpha_0 R} \qquad (3.37)$$

Setting $\frac{d}{d(\alpha r)}f(\alpha r) = f'(\alpha r)$ and substituting (3.30) and (3.37) into (3.36) leads to

$$b_1^m e^{im\phi}\left[2P_1^m(R\delta\alpha + \alpha R\varepsilon\cos^2\theta)\frac{j_1'(\alpha_0 R)}{\alpha_0 R} - 2\varepsilon\sin\theta\cos\theta\sin\theta\, P_1^{m\prime}\, j_1'(\alpha_0 R)\right]$$

$$+12 b_3^m e^{im\phi} P_3^m \frac{j_3'(\alpha_0 R)}{\alpha_0^2 R} + \cdots = 0 \qquad (3.38)$$

Defining the function I(m) as

$$I(m) = \int_{-1}^{1} P_1^m(\mu)\ \mathbf{B}\cdot\nabla C\ d\mu, \qquad (3.39)$$

one gets

$$I(0) = b_1^0 j_1'(\alpha_0 R)\left[\frac{4}{3}\frac{\delta\alpha}{\alpha} + \frac{4}{15}\varepsilon\right] = 0 \qquad (3.40)$$

and $$I(1) = b_1^1 j_1'(\alpha_0 R)\left[\frac{8}{3}\frac{\delta\alpha}{\alpha} + \frac{16}{15}\varepsilon\right] = 0 \qquad (3.41)$$

leading to

$$\frac{\delta\alpha}{\alpha}\bigg|_{m=0} = -\frac{\varepsilon}{5} \qquad (3.42)$$

and $$\left.\frac{\delta\alpha}{\alpha}\right|_{m=1} = -\frac{2\varepsilon}{5} \tag{3.43}$$

This led R&B to conclude that the prolate spheromak ($\varepsilon > 0$) is unstable whereas the oblate spheromak ($\varepsilon < 0$) is stable; flattening the spheromak is essential for stability.

3.3.8. A Novel Spheromak Design

A novel spheromak design is now proposed. This design is achieved through confining the plasma between two perfectly conducting spherical shells of radii R_0 and R_1 as shown in Figure 3.4. The presence of the perfectly conducting shells implies that the magnetic field must be tangential at both boundaries; i.e., the radial component of the magnetic field, B_r, must vanish at both boundaries. This, in turn, implies that αR_0 and αR_1 are both zeros of the spherical Bessel function which reduces the number of applicable solutions; the smaller set of solutions enhances the stability of the solution.

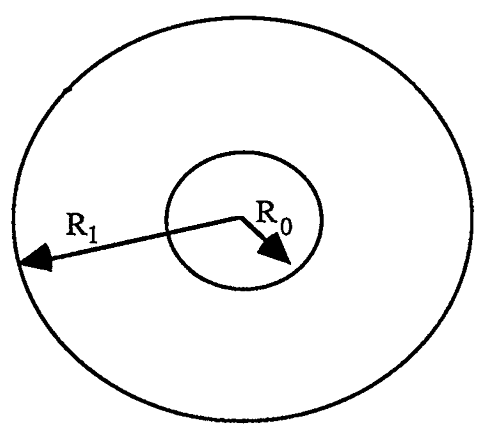

Figure 3.4 A cross section of the proposed spheromak

The interior shell excludes the origin from the plasma magnetic field and thus, in principle, allows the presence of the fields that belong to the general fields proposed in the previous chapter. This means that the applicable general spherical coordinates FF field is

$$\mathbf{B}_m = \sum_{n=m}^{\infty} e^{im\phi}\left\{\left[n(n+1)p_n^m P_n^m \frac{a_n j_n(\alpha r) + d_n n_n(\alpha r)}{\alpha r}\right]\mathbf{u}_r\right.$$

$$+\left[-\sin\theta\ p_n^m P'^m_n \frac{d}{dr}[a_n r j_n(\alpha r) + d_n r n_n(\alpha r)]\right.$$

$$\left.+\frac{im}{\sin\theta} p_n^m P_n^m [a_n j_n(\alpha r) + d_n n_n(\alpha r)]\right]\mathbf{u}_\theta$$

$$+\left[-\sin\theta\ p_n^m P'^m_n [a_n j_n(\alpha r) + d_n n_n(\alpha r)]\right.$$

$$\left.\left.+\frac{in}{\sin\theta} p_n^m P_n^m \frac{1}{\alpha r}\frac{d}{dr}[a_n r j_n(\alpha r) + d_n r n_n(\alpha r)]\right]\mathbf{u}_\phi\right\} \tag{3.44}$$

It is of interest to determine the effect of ellipticity on the stability of the proposed spheromak. If the internal shell has the surface

$$C_0(r,\ \theta) = r - R_0\left(1 + \varepsilon_0 \cos^2\theta\right) = 0 \tag{3.45}$$

and the outer shell has the surface

$$C_1(r,\ \theta) = r - R_1\left(1 + \varepsilon_1 \cos^2\theta\right) = 0 \tag{3.46}$$

The conditions of the vanishing of the magnetic field component perpendicular to the surfaces are $\mathbf{B} \cdot \nabla C_k = 0$ where $k = 0, 1$. Substitution of (3.44) into (3.36) leads to

$$\mathbf{B}_m \cdot \nabla C = \sum_{n=m}^{\infty} e^{im\phi}\left\{\left[n(n+1)p_n^m P_n^m \frac{a_n J_n(\alpha r) + d_n n_n(\alpha r)}{\alpha r}\right]\right.$$

$$+2\varepsilon \sin\theta\cos\theta\left[-\sin\theta p_n^m P'^m_n \frac{d}{dr}[a_n r j_n(\alpha r) + d_n r n_n(\alpha r)]\right.$$

$$\left.\left.+\frac{im}{\sin\theta} p_n^m P_n^m [a_n j_n(\alpha r) + d_n n_n(\alpha r)]\right]\right\} \tag{3.47}$$

Expanding $j_n(\alpha r)$ and $n_n(\alpha r)$ as in (3.37)

$$j_n(\alpha r) = j_n(\alpha_0 R) + \left(R\delta\alpha + \alpha_0 R\varepsilon\cos^2\theta\right)j_n'(\alpha_0 R) \tag{3.48}$$

$$n_n(\alpha r) = n_n(\alpha_0 R) + \left(R\delta\alpha + \alpha_0 R\varepsilon\cos^2\theta\right)n_n'(\alpha_0 R) \tag{3.49}$$

Substitution of (3.48) and (3.49) into (3.47) keeping only first order terms of $\delta\alpha$ and ε leads to

$$\begin{aligned}
B_m \cdot \nabla C = p_1^m e^{im\phi}\Bigg[& 2P_1^m\left[\frac{a_1 j_1(\alpha_0 R) + d_1 n_1(\alpha_0 R)}{\alpha_0 R}\right] \\
& + \frac{\left(R\delta\alpha + \alpha_0 R\varepsilon\cos^2\theta\right)\left[a_1 j_n'(\alpha_0 R) + d_1 n_1'(\alpha_0 R)\right]}{\alpha_0 R} \\
& - 2\varepsilon\sin^2\theta\cos\theta P_1'^m\left[R\left[a_1 j_1'(\alpha_0 R) + d_1 n_1'(\alpha_0 R)\right] + \left[a_1 j_1(\alpha R) + d_1 n_1(\alpha R)\right]\right] \\
& + 2im\varepsilon\cos\theta P_1^m\left[a_1 j_1(\alpha_0 R) + a_1 n_1(\alpha_0 R)\right]\Bigg] \\
& + 12 p_3^m e^{im\phi} P_3^m\left[\frac{a_3 j_3(\alpha_0 R) + d_3 n_3(\alpha_0 R)}{\alpha_0 R}\right] + \cdots = 0
\end{aligned} \tag{3.50}$$

However, the boundary condition $B_r(\alpha_0 R)\Big|_{\theta=\frac{\pi}{2}} = 0$ translates to

$$a_n j_n(\alpha_0 R) + d_n n_n(\alpha_0 R) = 0.$$

Equation (3.50) reduces to

$$\begin{aligned}
2p_1^m P_1^m e^{im\phi}\Bigg[& \frac{\left(R\delta\alpha + \alpha_0 R\varepsilon\cos^2\theta\right)\left(a_1 j_1'(\alpha_0 R) + d_1 n_1'(\alpha_0 R)\right)}{\alpha_0 R}\Bigg] \\
& - \varepsilon\sin^2\theta\cos\theta R\left[a_1 j_1'(\alpha_0 R) + d_1 n_1'(\alpha_0 R)\right] \\
& + 12 p_3^m P_3^m e^{im\phi}\cos\theta\left[\frac{a_3 j_3(\alpha_0 R) + d_3 n_3(\alpha_0 R)}{\alpha_0 R}\right] + \cdots = 0
\end{aligned} \tag{3.51}$$

Using definition (3.39) of I(m)

$$I(0) = 2P_1^0\left[a_1 j_1'(\alpha_0 R) + d_1 n_1'(\alpha_0 R)\right]\left[\frac{4}{3}\frac{\delta\alpha}{\alpha} + \frac{4}{15}\varepsilon\right] = 0 \tag{3.52}$$

$$I(1) = P_1'\left[a_1 j_1'(\alpha_0 R) + d_1 n_1'(\alpha_0 R)\right]\left[\frac{8}{3}\frac{\delta\alpha}{\alpha} + \frac{16}{15}\varepsilon\right] = 0 \tag{3.53}$$

Leading to

$$\left.\frac{\delta\alpha}{\alpha}\right|_{m=0} = -\frac{\varepsilon}{5} \tag{3.54}$$

$$\left.\frac{\delta\alpha}{\alpha}\right|_{m=1} = -\frac{2\varepsilon}{5} \tag{3.55}$$

Which are identical results to those for the conventional spheromak. It is concluded that the prolate spheromak is unstable whereas the oblate spheromak is stable. At the precision analysis presented in this thesis, it seems that having one shell prolate (with ε_p) and the other oblate (with ε_0) will lead to a stable spheromak provided that $\varepsilon_p < \varepsilon_0$.

The condition $B_r(\alpha R_n)\Big|_{\theta=\frac{\pi}{2}} = 0$ implies

$$a_n j_n(\alpha R_0) + d_n n_n(\alpha R_o) = 0 \tag{3.56a}$$

$$a_n j_n(\alpha R_1) + d_n n_n(\alpha R_1) = 0 \tag{3.56b}$$

Equation (3.56a) leads to

$$a_n = -\frac{a_n n_n(\alpha R_0)}{d_n(\alpha R_0)} \tag{3.57}$$

Substitution of (3.57) into (3.56b) leads to

$$n_n(\alpha R_0) j_n(\alpha R_1) - n_n(\alpha R_1) j_n(\alpha R_0) = 0 \tag{3.58}$$

Choice of R_0 and R_1 restricts the value of α. The radii R_0 and R_1 were set to 4.493 and 7.725, the lowest non zero values for α satisfying (3.58) for $0 \leq n \leq 8$ were determined numerically and is shown in figure 3.5.

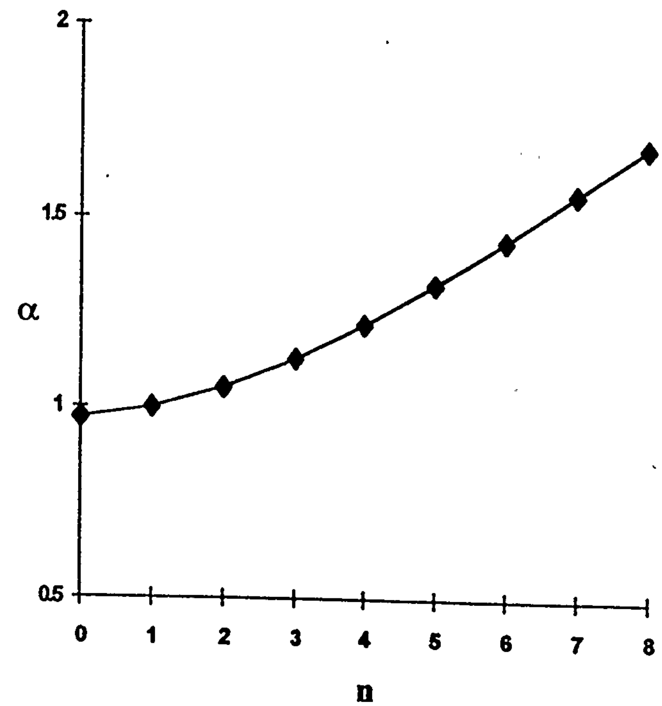

Figure 3.5 A graph showing the lowest non-zero values of α such that
$n_n(4.493\alpha)j_n(7.725\alpha) - n_n(7.725\alpha)j_n(4.493\alpha) = 0$
for values of n between 0 and 8

The function $j_1(\alpha r)$ has its first zero at 4.493 and its second zero at 7.725 The function $n_1(\alpha r)$ has its first zero at 2.798 and its second zero at 6.121. This means that for a given value of α, a configuration supporting $j_1(\alpha r)$ (with an external shell to internal shell radii ratio of 1.669) will have approximately 1.6 the volume of a configuration supporting $n_1(\alpha r)$ (with an external shell to internal shell radii ratio of 2.188). Numerical integration showed that if α is assumed to be unity, then the total magnetic energy in the configuration supporting $j_1(\alpha r)$ is approximately 0.247 of the total magnetic energy in the configuration supporting $n_1(\alpha r)$. Although these values can be used only as indicators, one is led to conclude that the configuration supporting $j_1(\alpha r)$ is more desirable than a configuration supporting $n_1(\alpha r)$. A spheromak supporting $j_1(\alpha r)$ has approximately 3.651 the volume of a conventional spheromak (with the same α).

Obvious problems with the proposed spheromak configuration are

(i) a mechanism is required to fix the interior shell with respect to the outer shell. A possible solution is a dielectric that does not perturb the magnetic field. However, this dielectric will face the full temperature and radiation of the plasma; and

(ii) the difficulty of positioning field coils inside the interior shell.

CHAPTER 4

TRANSVERSE ELECTROMAGNETIC WAVES WITH E || B

4.1. INTRODUCTION

The solutions of Maxwell's equations of common interest are concerned with the propagation of electromagnetic (EM) energy in the form of transverse electromagnetic (TEM) waves in free space, material media, and transmission lines as well as transverse electric and magnetic waves in waveguides. Such solutions are characterized by the orthogonality of the electric field **E** and the magnetic flux density, $\mathbf{E} \perp \mathbf{B}$. An examination of many senior undergraduate and graduate electrodynamics textbooks [STRATTON 1941, SMYTHE 1950, PANOFSKY AND PHILLIPS 1955, PURCELL 1965, REITZ AND MILFORD 1967, LORRAIN AND CORSON 1970, COOK 1975, JACKSON 1976, PORTIS 1978, MARION AND HEALD 1980] reveals that there is very little discussion of the conditions under which TEM standing waves exist. Textbooks with a more engineering emphasis [RAMO ET AL. 1965, JORDAN AND BALMAIN 1968, BEKEFI AND BARRETT 1968 and SHADOWITZ 1975] generally discuss standing waves in the context of the measurement of the voltage standing wave ratio for transmission lines. Consequently, it is not surprising that the recent demonstration by Chu and Ohkawa [1982] that a class of TEM waves with $\mathbf{E} \parallel \mathbf{B}$ exists, provoked considerable critical reaction [LEE 1983, SALINGAROS 1985 and SALINGAROS 1986] as well as a rebuttal [CHU 1983]. Unfortunately, most of this discussion appears to be due to the failure of the various authors to define their terminology and explain their assumptions. It is intuitive to assume that all TEM waves propagate. This seems to be the reaction of the respondents [LEE 1983, SALINGAROS 1985 and SALINGAROS 1986] and probably many other readers. Chu and Ohkawa [1982] appear to have realized that their proposed class consisted only of standing waves because their example was of this class and Chu [1983] was explicit in his rebuttal. Recently, Zaghloul et al. [1987] derived the general conditions under which TEM waves with **E** || **B** exist. They showed that it is useful to classify TEM waves according to whether their Poynting vector is identically zero or non-zero. The former are **E** || **B** standing waves and the latter are traveling and standing TEM waves with **E** |/| **B**. Zaghloul and Buckmaster [1988] gave a pedagogic review of the subject and studied the physical properties of TEM standing waves (TEMSW) with **E** || **B**. The approach of Zaghloul et al. [1987] was to find the conditions on the vector potential of **E** || **B** TEMSW. Shimoda et al. [1990] found the conditions on the EM energy density of **E** || **B** TEMSW. Baylis and Jones [1989] derived the conditions of **E** || **B** TEMSW using the Pauli algebra approach. Uehara et al. [1989] obtained **E** || **B** non-transverse solutions.

It is interesting to note that although a controversy surrounded **E** || **B** TEMSW, the fact that the superposition of circularly polarized waves of the same helicity, amplitude and frequency, traveling in opposite directions leads to **E** || **B** TEMSW has been in literature for approximately three decades [HARRINGTON 1961]. **E** || **B** TEMSW should not be confused with **E** || **B** static fields. **E** || **B** static fields have been the subject of various exercises on relativistic electrodynamics [LANDAU AND LIFSHITZ 1975 and JACKSON 1976] **E** || **B** static fields are easy to obtain (at least within a finite region of space); place a charged spherical object on the axis of a magnetic dipole and, on the axis, the electric and magnetic fields are parallel. Note, however, that Misner et al. [1973] and Pathria [1974] had exercises that dealt with **E** || **B** waves.

This chapter is, in essence, the same as the work of Zaghloul and Buckmaster [1993]. The chapter reviews **E** ‖ **B** TEMSW. It reviews TEMSW in general, derives the conditions under which **E** ‖ **B** TEMSW exist in a sourceless medium, shows how the examples introduced in literature previously can be obtained by solving the general conditions, and finally reviews the physical properties of **E** ‖ **B** TEMSW. The chapter concludes by introducing some applications of these standing waves.

4.2. SOLUTIONS OF MAXWELL'S EQUATIONS IN A SOURCELESS MEDIUM

Maxwell's equations in a uniform sourceless medium, assumed to be vacuum for simplicity, are [BEKEFI AND BARRETT 1977]

$$\nabla \times \mathbf{E} - \frac{\partial \mathbf{B}}{\partial t} = 0 \tag{4.1a}$$

$$\nabla \times \mathbf{B} + \frac{\partial \mathbf{E}}{\partial t} = 0 \tag{4.1b}$$

$$\nabla \cdot \mathbf{E} = 0 \tag{4.1c}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{4.1d}$$

The solution of (4.1a-d) can be obtained by introducing scalar ϕ and vector **A** potentials [JACKSON 1976] such that

$$\mathbf{E} = - \nabla\phi - \frac{\partial \mathbf{A}}{\partial t} \tag{4.2a}$$

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{4.2b}$$

Substitution of (4.2a) and (4.2b) into (4.1a) and (4.1c) yields;

$$\nabla\times\nabla\times\mathbf{A}+\frac{\partial}{\partial t}\left[\nabla\phi+\frac{\partial\mathbf{A}}{\partial t}\right]=0 \tag{4.3a}$$

$$\nabla^2\phi+\frac{\partial}{\partial t}(\nabla\cdot\mathbf{A})=0 \tag{4.3b}$$

The vector Laplacian $\Delta\mathbf{A}\equiv\nabla\ (\nabla\cdot\mathbf{A})-\nabla\times\nabla\times\mathbf{A}$, so (4.3a) becomes

$$\Delta\mathbf{A}-\frac{\partial^2\mathbf{A}}{\partial t^2}-\nabla\left[\nabla\cdot\mathbf{A}+\frac{\partial\phi}{\partial t}\right]=0 \tag{4.3c}$$

Notice that $\Delta\mathbf{A}\equiv\nabla^2\mathbf{A}$ in Cartesian rectangular coordinates. Equations (4.3b) and (4.3c) can be separated by choosing the Lorentz gauge [JACKSON 1976];

$$\nabla\cdot\mathbf{A}+\frac{\partial\phi}{\partial t}=0 \tag{4.4}$$

Substitution of (4.4) into (4.3b) and (4.3c) yields

$$\nabla^2\phi-\frac{\partial^2\phi}{\partial t^2}=0 \tag{4.5a}$$

$$\Delta\mathbf{A}-\frac{\partial^2\mathbf{A}}{\partial t^2}=0 \tag{4.5b}$$

A general TEM solution of (4.5b) can be expressed in the form due to d'Alembert [SMYTHE 1950 and MORSE AND FESHBACH 1953]

$$\mathbf{A}(\mathbf{r},t)=\mathbf{A}_+(\eta)+\mathbf{A}_-(\xi) \tag{4.6a}$$

where $\mathbf{A}_+(\eta)=\sum_{i=1}^{3}\mathbf{A}_+(\eta)_i 1_i$, $\mathbf{A}_-(\xi)=\sum_{i=1}^{3}\mathbf{A}_-(\xi)_i\mathbf{u}_i$, $\eta\equiv\mathbf{k}\cdot\mathbf{r}-\omega t$, $\xi\equiv\mathbf{k}\cdot\mathbf{r}+\omega t$ and $\mathbf{u}_i$ are mutually orthogonal unit vectors in an arbitrary orthogonal coordinate system. Notice that the most general TEM solution of (4.5b) is a sum of solutions of the form

(4.6a) over all possible propagation vectors, **k.** **k** and ω, the angular frequency, are related by

$$k^2 = \omega^2 \tag{4.6b}$$

4.3. TRAVELING AND STANDING TEM WAVES

The two solutions of (4.5b) given by (4.6a) can be considered as the vector potentials of two TEM waves traveling in opposite directions, i.e., energy propagation takes place in two opposite directions. In general, there is more energy flow in one direction than in the other so that there is a net energy flow in one direction. Such a TEM wave is called a traveling wave. In the special case where the energy propagated in one direction is equal to that propagated in the opposite direction, there is no net energy flow in the medium and the sum of the two TEM waves forms what is generally known as a standing wave. Mathematically, the amount of energy density propagated is proportional to the magnitude of the Poynting vector [MARION AND HEALD 1980] **S** where

$$\mathbf{S} \propto \mathbf{E} \times \mathbf{B} \tag{4.7}$$

The condition for a standing wave is that the time average of **S** vanishes [RAO 1977]. This can be achieved if (i) **S** is zero all the time everywhere in the region of space under consideration, i.e.. $\mathbf{S}(\mathbf{r}, t) = 0$ or (ii) **S** changes sign with time and has a zero time average, i.e. $\mathbf{S}(\mathbf{r}, t)_{av} = 0$. Examination of (4.7) shows that $\mathbf{S} = 0$ if either **E** or **B** is zero (the cases of electro- and magneto-statics) or if $\mathbf{E} \parallel \mathbf{B}$. This means that for time varying fields, (i) leads to TEMSW with $\mathbf{E} \parallel \mathbf{B}$ whereas (ii) leads to the more familiar $\mathbf{E} \perp \mathbf{B}$ TEMSW. The fact that $\mathbf{S}(\mathbf{r}, t) = 0$, for time varying fields, leads to $\mathbf{E} \parallel \mathbf{B}$ TEMSW will be used in Sec. 4.4. to derive the general conditions for the existence of these standing waves in a sourceless medium. Note, however, that although all standing waves have a Poynting vector whose time average is zero, not all waves with $\mathbf{S} = 0$ are standing waves, in the strict sense, as was shown by Shimoda et al. [1990]. However, it can be stated that if only a single frequency is present, then the vanishing of the Poynting vector (or its time average) is an indication that the wave is a standing wave.

4.4. E || B TEMSW IN A SOURCELESS MEDIUM

4.4.1 The Work of Zaghloul et al. [1987]

The derivation of the general conditions for the existence of TEMSW with **E** || **B** was given first in a condensed form by Zaghloul et al. [1987]. The derivation given here is more complete and details the necessary assumptions [ZAGHLOUL AND BUCKMASTER 1988]. This derivation can be simplified by assuming that $\nabla \cdot A = 0$, $\partial\phi/\partial t = 0$ and $\nabla\phi = 0$ This is equivalent to choosing the Coulomb gauge[1] which is consistent with the earlier choice of the Lorentz gauge. However, it should be emphasized that different vector and scalar potentials that are related by a gauge

[1] The Coulomb gauge is generally given as the condition of transversality of the vector potential with no restriction on the scalar potential. However the extra restrictions imposed on the scalar potential do not conflict with the condition imposed on the vector potential. It is natural that such extra restrictions remove some of the arbitrariness in the potentials. i.e. the scalar potential is determined to within an additive constant and the vector potential is determined to within the gradient of a function of the coordinates that satisfies Laplace's equation.

transformation lead to the same electric and magnetic fields [MARION AND HEALD 1980]. Now

$$\mathbf{B}(\mathbf{r},t) = \nabla \times \mathbf{A} = \mathbf{k} \times (\mathbf{A}'_+ + \mathbf{A}'_-) \tag{4.8}$$

where $\mathbf{A}'_+ \equiv \dfrac{d\mathbf{A}_+(\eta)}{d\eta}$ and $\mathbf{A}'_- \equiv \dfrac{d\mathbf{A}_-(\xi)}{d\xi}$ so **B** is always transverse since $\mathbf{k} \perp \mathbf{B}$ and

$$\mathbf{E}(\mathbf{r},\ t) = \frac{\partial \mathbf{A}}{\partial t} = \omega(\mathbf{A}'_+ - \mathbf{A}'_-) \tag{4.9}$$

since $\nabla\phi = 0$. Notice that **E** is not necessarily transverse. A necessary condition for $\mathbf{E} \perp \mathbf{k}$, which is a requirement for a TEM wave, is $\mathbf{k}\cdot(\mathbf{A}'_+ - \mathbf{A}'_-) = 0$. Solutions of (4.5a) and (4.5b) for which the Poynting vector $\mathbf{S}(\mathbf{r},\ t) = 0$ can be obtained by substituting (4.8) and (4.9) into (4.7)

$$\begin{aligned} \mathbf{S}\ (\mathbf{r},\ t) &\propto \omega(\mathbf{A}'_+ - \mathbf{A}_-) \times \mathbf{k} \times (\mathbf{A}'_+ - \mathbf{A}'_-) \\ &= \omega[(\mathbf{A}'_+ - \mathbf{A}'_-)\cdot(\mathbf{A}'_+ + \mathbf{A}'_-)]\mathbf{k} - \omega[\mathrm{k}\cdot(\mathbf{A}'_+ - \mathbf{A}'_-)](\mathbf{A}'_+ + \mathbf{A}'_-) \\ &= \omega[(\mathbf{A}'_+ - \mathbf{A}'_-)\cdot(\mathbf{A}'_+ + \mathbf{A}'_-)]\mathbf{k} - 0 \end{aligned} \tag{4.10}$$

since $\mathbf{k}\cdot(\mathbf{A}'_+ - \mathbf{A}_-) = 0$ for TEM waves. Consequently, the condition for **S** (**r**, t) = 0 is

$$(\mathbf{A}'_+ + \mathbf{A}'_-)\cdot(\mathbf{A}'_+ - \mathbf{A}'_-) = 0 \tag{4.11}$$

It is easy to interpret this condition vectorially using Fig. 4.1 since

$$(\mathbf{A}'_+ + \mathbf{A}'_-) \perp (\mathbf{A}'_+ - \mathbf{A}'_-)$$

Analytically,

$$(\mathbf{A}'_+ + \mathbf{A}'_-)\cdot(\mathbf{A}'_+ - \mathbf{A}'_-) = |\mathbf{A}'_+|^2 - |\mathbf{A}'_-|^2 = 0 \tag{4.12}$$

i.e. $|\mathbf{A'}_+| = |\mathbf{A'}_-|$. It was mentioned in Sec. 4.3. that $\mathbf{S}(\mathbf{r}, t) = 0$ if (i) $\mathbf{E} = 0$, (ii) $\mathbf{B} = 0$ or (iii) $\mathbf{E} \parallel \mathbf{B}$ so the conditions for non-trivial $\mathbf{E} \parallel \mathbf{B}$ TEMSW are $|\mathbf{A'}_+| = |\mathbf{A'}_-|$, $\mathbf{A'}_+ \neq \pm\mathbf{A'}_-$, $\nabla\cdot\mathbf{A}_+ = 0$ and $\nabla\cdot\mathbf{A}_- = 0$.

Recently, Gray [1992] criticized the condition obtained by Zaghloul et al. [1987] claiming, without proof, that it is over-restrictive compared with the original work of Chu and Ohkawa [1982]. His claim is unfounded as it related to the work of Chu and Ohkawa. However, the claim of Zaghloul et al. [1987] that their condition is the general condition for $\mathbf{E} \parallel \mathbf{B}$ TEMSW is correct only for monochromatic TEM waves, the solutions of Chu and Ohkawa [1982] belong to this class.

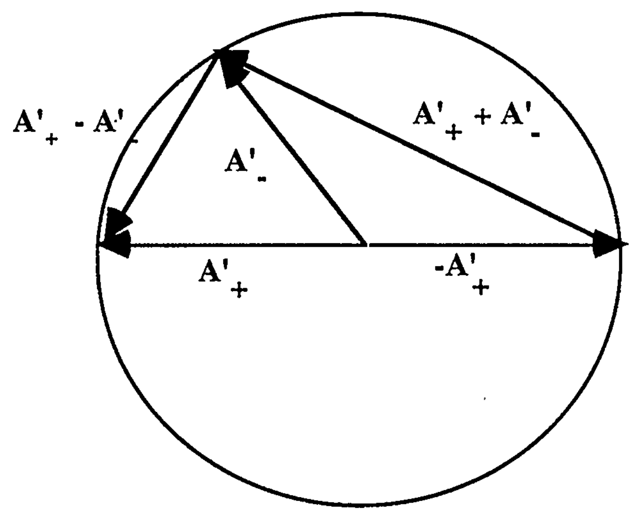

Figure 4.1 A schematic diagram showing the vectorial relationship between the vectors $\mathbf{A'}_+$ and $\mathbf{A'}_-$.

4.4.2. The Work of Shimoda et al. [1990]

Recently, Shimoda et al. [1990] derived a new condition for $\mathbf{E} \parallel \mathbf{B}$ TEMSW. The basic point of their derivation is the same as that of Zaghloul et al. [1987]: $\mathbf{E} \parallel \mathbf{B}$

TEMSW must satisfy $\mathbf{E} \times \mathbf{B} = 0$. They started from Maxwell's equations in free space (4.1a-d) and looked for homogeneous plane waves with $E_z = B_z = 0$ and $\mathbf{E} \times \mathbf{B} = 0$. If the wavefront of the wave is the xy-plane, then **E** and **B** do not depend upon either x or y. Then (4.1a) and (4.1b) become

$$\frac{\partial E_x}{\partial t} = -\frac{\partial B_y}{\partial z}, \ \frac{\partial E_y}{\partial z} = \frac{\partial B_x}{\partial t} \tag{4.1a'}$$

$$\frac{\partial B_x}{\partial t} = \frac{\partial E_y}{\partial z}, \ \frac{\partial B_y}{\partial z} = -\frac{\partial E_x}{\partial t} \tag{4.1b'}$$

The vanishing of the Poynting vector implies

$$E_x B_y = E_y B_x \tag{4.13}$$

substitution of (4.1a') and (4.1b') into the differentiation of (4.13) with respect to time and z yields that the energy density, $\varepsilon = (E^2 + B^2)/2$ is independent of both z and t:

$$\frac{\partial \varepsilon}{\partial t} = \frac{\partial \varepsilon}{\partial z} = 0 \tag{4.15}$$

This is the condition for $\mathbf{E} \parallel \mathbf{B}$ TEMSW. The vanishing of the time derivative of the energy density is a consequence of the vanishing of the Poynting vector as should be obvious from direct substitution in Poynting's theorem

$$\frac{\partial \varepsilon}{\partial t} = -\nabla \cdot \mathbf{S}$$

4.5. CLASSIFICATION OF EM WAVES USING LORENTZ INVARIANTS

It will now be shown how Lorentz invariants can be used to classify TEM waves [ZAGHLOUL AND BUCKMASTER 1990]. The invariant $I_2 = \mathbf{E} \cdot \mathbf{B}$ divides EM waves into two basic classes; the first class has $I_2 = 0$ and $\mathbf{E} \perp \mathbf{B}$ in all Lorentz frames and the second class has $I_2 = 0$ and $\mathbf{E} \perp \mathbf{B}$ in all Lorentz frames. The fields of the second class have been called parallelizable by Salingaros [1985]. However, note that the superposition of a wave with $\mathbf{E} \perp \mathbf{B}$ and a wave with $\mathbf{E} \parallel \mathbf{B}$ results in an example where it is impossible to find a frame where the $\mathbf{E} \perp \mathbf{B}$ component is parallelized.

The invariant ε_{sw} given by [MISNER ET AL. 1973 and ZAGHLOUL ET AL. 1988]

$$\varepsilon_{sw}^2 = I_1^2 + I_2^2 = \varepsilon^2 - S^2 \tag{4.16}$$

where S is the magnitude of the Poynting vector and $I_1 = E^2 - B^2$ is a Lorentz invariant [BARUT 1964], is a measure of the standing wave component of an EM wave. Equation (4.16) shows that if $\varepsilon_{sw} = 0$, then $\varepsilon = S$, i.e., all the EM energy density within any volume propagates. This is the case for pure traveling waves (waves with no standing wave component). Equation (4.16) also shows that $\varepsilon_{sw} = 0$ if and only if both I_1 and I_2 vanish. This means that pure traveling monochromatic waves can be characterized by the vanishing of the invariants I_1 and I_2. Such waves are referred to in literature [MISNER ET AL. 1973] as null or pure radiation fields.

The Poynting vector, **S**, is not invariant but is a useful tool in classifying EM waves into standing waves with $\mathbf{E} \parallel \mathbf{B}$ for which $\mathbf{S} = 0$ and traveling and standing waves with $\mathbf{E} \nparallel \mathbf{B}$ for which $S \neq 0$.

It will be shown in (ii) of Sec. 4.8. that time-varying waves with **E** || **B** cannot have $I_1 = 0$. This means that $I_1 \neq 0$ and $I_2 \neq 0$ for parallelizable waves. Moreover, $\mathbf{E} \perp \mathbf{B}$ standing waves have $\varepsilon_{sw} \neq 0$ and $I_2 = 0$ and consequently cannot have $I_1 = 0$. This means that the invariant $I_1 = E^2 - B^2$ divides EM waves into two classes; the first class has $I_1 = 0$ and contains waves with no standing wave component. These are waves for which the electric energy density, $E^2/2$, and the magnetic energy density, $B^2/2$, are equal in all Lorentz frames. The second class has $I_1 \neq 0$ and contains all waves with a standing wave component. This suggest that pure monochromatic traveling waves can be characterized by the vanishing of the invariant I_1 alone. Since such waves also have $I_2 = 0$ it is concluded that, for time varying fields, the vanishing of I_1 is necessary and sufficient for the vanishing of I_2.

Note that both the Lorentz invariants [BARUT 1964] $I_2 = E^2 - B^2$ and $I_2 = \mathbf{E} \cdot \mathbf{B}$ are non-zero for **E** || **B** TEMSW. Recall that for monochromatic TEM waves

$$\mathbf{E} = \omega\left[\frac{dA_+(\eta)}{d\eta} - \frac{dA_-(\xi)}{d\xi}\right] \text{ and } \mathbf{B} = \mathbf{k} \times \left[\frac{dA_+(\eta)}{d\eta} + \frac{dA_-(\xi)}{d\xi}\right] , \tag{4.17}$$

$$E^2 = \omega^2\left[\left|\frac{dA_+(\eta)}{d\eta}\right|^2 + \left|\frac{dA_-(\xi)}{d\xi}\right|^2 - 2\left[\frac{dA_+(\eta)}{d\eta} \cdot \frac{dA_-(\xi)}{d\xi}\right]\right] , \tag{4.18a}$$

$$B^2 = k^2\left[\left|\frac{dA_+(\eta)}{d\eta}\right|^2 + \left|\frac{dA_-(\xi)}{d\xi}\right|^2 + 2\left[\frac{dA_+(\eta)}{d\eta} \cdot \frac{dA_-(\xi)}{d\xi}\right]\right] , \tag{4.18b}$$

Subtracting (4.18b) from (4.18a) yields

$$I_1 = E^2 - B^2 = -4\omega^2\left[\frac{dA_+(\eta)}{d\eta} \cdot \frac{dA_-(\xi)}{d\xi}\right] \tag{4.19}$$

and the product of the square roots of (4.18a) and (4.18b) yields

$$I_2 = \mathbf{E}\cdot\mathbf{B} = \pm EB = \pm 2\left[\varepsilon^2 - \omega^2\left[\frac{dA_+(\eta)}{d\eta}\cdot\frac{dA_-(\xi)}{d\xi}\right]\right]^{1/2} \tag{4.20}$$

It is conjectured that I_1 and I_2 are zero on the average (both time and space). Brownstein [1987a] has shown that the integration of each invariant over all space and time must vanish if it has a "well-defined meaning."

Equation (4.19) suggests that $I_1 = 0$ only if $\mathbf{A}'_+ \perp \mathbf{A}'_-$. $\mathbf{A}'_+ \perp \mathbf{A}'_-$ implies that $I_2 = \pm 4K\omega|\mathbf{A}'_+||\mathbf{A}'_-|$. In which case, $\mathbf{E} \perp \mathbf{B}$ if and only if $|\mathbf{A}'_+|$ or $|\mathbf{A}'_-| = 0$. This means that for monochromatic waves, $I_1 = 0$ and $I_2 = 0$ only for waves with no standing wave component. It will be shown later that monochromatic $\mathbf{E} \parallel \mathbf{B}$ TEMSW cannot have $I_1 = 0$. Table 4.1 summarizes this classification of EM waves.

It is instructive to give examples for each of the classes given in Table 4.1 An example of pure traveling $\mathbf{E} \perp \mathbf{B}$ TEM waves is a linearly polarized wave propagating in the positive z-direction:

$$\mathbf{E} = E_0\,(\cos[\omega t - kz], 0, 0) \tag{4.21a}$$

$$\mathbf{B} = B_0\left(0,\ \cos[\omega t - kz],\ 0\right) \tag{4.21b}$$

Note that in the units chosen in this thesis $E_0 = B_0$ and $k = \omega$. An example of an $\mathbf{E} \perp \mathbf{B}$ TEMSW is the sum of this example with another of the same amplitude, frequency and helicity propagating in the negative z-direction:

$$\mathbf{E} = E_0\,(\cos[\omega t - kz], 0, 0) + E_0\,(\cos[\omega t + kz], 0, 0) = E_0\,(\cos\omega t \cos kz, 0, 0) \tag{4.22a}$$

$$\mathbf{B} = B_0\,(0, \cos[\omega t - kz], 0) + B_0\,(0, -\cos[\omega t + kz], 0) = B_0\,(0, \sin\omega t \sin kz, 0) \tag{4.22b}$$

	I_1 $E^2 - B^2$	I_2 $\mathbf{E} \cdot \mathbf{B}$	ε_{sw}	S
$\mathbf{E} \perp \mathbf{B}$ (pure traveling)	0	0	0	≠0
$\mathbf{E} \perp \mathbf{B}$ (standing)	≠0	0	≠0	≠0
$\mathbf{E} \parallel \mathbf{B}$ (standing)	≠0	≠0	≠0	0
E I/I B and $\mathbf{E} \perp \mathbf{B}$ (parallelizable)	0	≠0	≠0	≠0
E I/I B and $\mathbf{E} \perp \mathbf{B}$ (not parallelizable)	≠0	≠0	≠0	≠0

Table 4.1: Table classifying EM waves according to the vanishing (0) and non-vanishing (≠0) of some EM Lorentz invariants and of the magnitude of the Poynting vector (S) [ZAGHLOUL AND BUCKMASTER 1990]

An example of an **E** || **B** TEMSW that will be obtained below is [CHU AND OHKAWA 1982]:

$$\mathbf{E} = E_0(\sin kz, \cos kz, 0)\sin\omega t \tag{4.23a}$$

$$\mathbf{B} = B_0(\sin kz, \cos kz, 0)\cos\omega t \tag{4.23b}$$

An example of a parallelizable TEM standing wave is:

$$\mathbf{E} = E'_0 (\sin kz \sin\omega t - \beta \cos kz \cos\omega t, \cos kz \sin\omega t + \beta \sin kz \cos\omega t, 0) \tag{4.24a}$$

$$\mathbf{B} = B'_0 (\sin kz \sin\omega t + \beta \cos kz \sin\omega t, \cos kz \cos\omega t - \beta \sin kz \sin\omega t, 0) \tag{4.24b}$$

where E'_0, B'_0 and β are coefficients. Note that this latter example may be obtained from the previous example of **E** || **B** TEM standing waves by a suitable Lorentz transformation to a frame moving in the positive z-direction.

An example of a wave that does not have **E** ⊥ **B**, does not have **E** || **B** and cannot be parallelized can be constructed from the superposition of the waves given by (4.21) and (4.23).

In conclusion, this section showed that Lorentz invariants of the EM waves can be used to classify EM waves according to whether they are orthogonal (**E** ⊥ **B**) or not orthogonal (**E** $\not\perp$ **B**) and according to whether they have a standing wave component or not.

4.6. MATHEMATICAL EXAMPLES OF TEMSW WITH E || B

4.6.1. E || B TEMSW of Chu and Ohkawa [1982]

Chu and Ohkawa [1982] proposed that if **F** is a vectorial function satisfying both (4.5b) and the Coulomb gauge and can be expressed in the form due to Zaghloul et al. [1987]

$$\mathbf{F}(\mathbf{r}, t) = \mathrm{Re}\,(e^{i\omega t})\,\mathrm{Re}\,(\mathbf{F}_k(r)) \tag{4.25}$$

then the vector potential

$$\mathbf{A}(\mathbf{r}, t) = \mathrm{Re}\left(e^{i\omega t}\right)\left[\mathbf{F}_k(r) + k^{-1}\nabla \times \mathbf{F}_k(\mathbf{r})\right] \tag{4.26}$$

will lead to **E** || **B** TEMSW. This approach defines a class of **E** || **B** TEMSW that may be characterized by the fact that the vector potential **A** satisfies the equation $\nabla \times \mathbf{A} = k\mathbf{A}$. However, Zaghloul et al. [1987] have shown that not all vector potentials satisfying $\nabla \times \mathbf{A} = k\mathbf{A}$ belong to this class. An example of a vector potential satisfying (4.25) and (4.26) is

$$\mathbf{A}(\mathbf{r}, t) = \mathbf{A}_0\,(\sin kz, \cos kz, 0)\cos\omega t \tag{4.23c}$$

which leads to the **E** || **B** fields given by (4.23a) and (4.23b)

Note that Harrington [1961] introduced this example of **E** || **B** TEMSW before Chu and Ohkawa [1982].

4.6.2. E || B TEMSW of Girotti et al.

Girotti et al. [see CHU AND OHKAWA 1987] proposed that if the vector potential can be expressed in the form

$$\mathbf{A}(\mathbf{r}, t) = f(\mathbf{r})\,\mathbf{G}(t) \tag{4.27}$$

then the resulting **E** and **B** will be parallel if

$$\frac{d\mathbf{G}(t)}{dt} \parallel \nabla f(\mathbf{r}) \times \mathbf{G}(t) \tag{4.28}$$

An example of a vector potential which is of the form given by (4.27) and satisfying (4.28) is

$$\mathbf{A}(\mathbf{r},\ t) = \mathbf{A}_0(\sin\omega t,\ \cos\omega t,\ 0)\cos kz \tag{4.29a}$$

which leads to the **E** || **B** fields

$$\mathbf{E}(\mathbf{r},\ t) = \omega\mathbf{A}_0(-\cos\omega t,\ \sin\omega t,\ 0)\cos kz \tag{4.29b}$$

$$\mathbf{B}(\mathbf{r},\ t) = k\mathbf{A}_0(\cos\omega t,\ -\sin\omega t,\ 0)\sin kz \tag{4.29c}$$

4.6.3. Solutions Generated Using The General Conditions of Zaghloul et al. [1987].

A general method for determining analytical expressions for those vector potentials which lead to **E** || **B** TEMSW has been outlined by Zaghloul and Buckmaster [1988]: if Cartesian coordinates are used and if the direction of the propagation vector (see iv of Sec. 4.8.) is assumed, without loss of generality, to be the z-direction; i.e. $\mathbf{k}$ = (0, 0, k), then it follows that if $A'_{\pm x} = (a\cos\theta_{\pm})$ and $A'_{\pm y} = (a\sin\theta_{\pm})$ where a is a constant and θ_{+} and θ_{-} are arbitrary functions of η and ξ respectively then the conditions for **E** || **B** TEMSW are satisfied provided θ_{+} and θ_{-} are chosen such that $\mathbf{A'}_{+} \neq \pm \mathbf{A'}_{-}$. In this case,

$$A_{+x}(\eta) = a\int \cos\theta_{+} d\eta, \quad A_{+y}(\eta) = a\int \cos\theta_{+} d\eta,$$

$$A_{-x}(\xi) = a\int \cos\theta_{-} d\xi \ \text{ and } \ A_{-y}(\xi) = a\int \sin\theta_{-} d\xi \tag{4.30}$$

The sign of a in (4.30) can be chosen arbitrarily to be positive or negative.

Examples:

(1) If $\theta_+ = \eta$ and $\theta_- = \xi$, then

$$A_{+x}(\eta) = \pm a \sin\eta\ ,\ A_{+y}(\eta) = a\cos\eta\ ,\ A_{-x}(\xi) = a\sin\xi \ \text{ and } \ A_{-y}(\xi) = a\cos\xi \tag{4.31}$$

The upper sign (+) leads to

$$A_x(\mathbf{r}, t) = A_{+x}(\eta) = A_{-x}(\xi) = a\sin(kz - \omega t) + a\sin(kz + \omega t) = 2a \sin kz \cos\omega t \tag{4.31a}$$

and

$$A_y(\mathbf{r},t) = A_{+y}(\eta) = A_{-y}(\xi) = a\cos(kz - \omega t) + a\cos(kz + \omega t) = 2a\cos kz \cos\omega t \tag{4.31b}$$

These components give the vector potential

$$\mathbf{A}(\mathbf{r},t) = 2a\,(\sin kz, \cos kz, 0)\cos\omega t \tag{4.23c}$$

which is the example given by Chu and Ohkawa [1982]. The lower sign (-) leads to

$$A_x = A_{+x}(\eta) + A_{-x}(\xi) = -a\sin(kz - \omega t) + a\sin(kz + \omega t) = 2a\cos kz \sin\omega t \tag{4.31c}$$

and

$$A_y = A_{+y}(\eta) + A_{-y}(\xi) = -a\cos(kz - \omega t) + a\cos(kz + \omega t) = 2a\cos kz\cos\omega t \tag{4.31d}$$

These components lead to the vector potential

$$\mathbf{A}(\mathbf{r},t) = 2a\,(\sin\omega t, \cos\omega t, 0)\cos kz \tag{4.26a}$$

which is the example given by Girotti et al. [see CHU AND OHKAWA 1987].

(2) If $\theta_+ = \ln\eta$ and $\theta_- = \xi$, then

$$A_x = a/2(\omega t - kz)\{\sin[\ln(\omega t - kz)] - \cos[\ln(\omega t - kz)]\} + a\cos(\omega t - kz) \tag{4.32a}$$

$$A_y = a/2(\omega t - kz)\{\sin[\ln(\omega t - kz)] + \cos[\ln(\omega t - kz)]\} + a\sin(\omega t - kz) \quad (4.32b)$$

$$A_z = 0 \quad (4.32c)$$

This vector potential leads to **E** || **B** fields since **S**(**r** ,t) = 0.

4.6.3. Solutions Generated Using The General Conditions of Shimoda et al. [1990].

Shimoda et al. [1990] proposed that since the EM energy density is a constant in space and time, it is possible to write

$$E_x = \cos(f)\cos(g) \quad , \quad E_y = \cos(f)\sin(g) \quad (4.33a)$$

$$B_x = \sin(f)\cos(h) \quad , \quad B_y = \sin(f)\sin(h) \quad (4.33b)$$

where amplitudes of **E** and **B** are taken to be unity and f, g and h are arbitrary functions of z and t that satisfy

$$h = g \text{ or } \pi, \qquad \frac{\partial f}{\partial t} = \frac{\partial g}{\partial z}, \qquad \frac{\partial g}{\partial t} = \frac{\partial f}{\partial z} \quad (4.34)$$

consequently f and g must satisfy the scalar Helmholtz equation and must have the general solutions

$$f = F(z+t) + G(z-t), \qquad g = F(z+t) - G(z-t) \quad (4.35)$$

where F and G are arbitrary functions. The fields are

$$E_x = \cos(F+G)\cos(F-G) \quad , \quad E_y = \cos(F+G)\sin(F-G) \quad (4.36a)$$

$$B_x = \sin(F+G)\cos(F-G) \quad , \quad B_y = \sin(F+G)\sin(F-G) \quad (4.36b)$$

Examples:

Shimoda et al. [1990] chose different F(z+t) and G(z-t) to obtain different examples of

E || B :

(1) $F = k(z+t)/2$ and $G = k(z-t)/2$ yield the circularly polarized wave example of Harrington [1961] also proposed by Chu and Ohkawa [1982]).

(2) $F = k(z+t)/2$ and $G = -k(z-t)/2$ yield the example of Girotti et al. [see CHU AND OHKAWA 1987].

(3) $F = k(z+t)/2$ and $G = 0$ yield

$$\mathbf{E} = (1+\cos[kz+\omega t],\sin[kz+\omega t],0) \tag{4.37a}$$

$$\mathbf{B} = (\sin[kz+\omega t],1-\cos[kz+\omega t],0) \tag{4.37b}$$

This example is interesting because it does not represent a standing wave. It consists of a circularly polarized propagating wave superimposed on a static field. This appears to contradict the previous statements that **E || B** waves must be standing waves. However, this apparent contradiction is removed since the statements made earlier were concerned with non static fields.

(4) $F = k_1(z+t)$ and $G = k_2(z-t)$ yield

$$E_x = \cos(kz+\Delta\omega t)\cos(\Delta kz+\omega t)\,,\ E_y = \cos(kz+\Delta\omega t)\sin(\Delta kz+\omega t) \qquad (4.38a)$$

$$B_x = \sin(kz+\Delta\omega t)\cos(\Delta kz+\omega t)\,,\ E_y = \sin(kz+\Delta\omega t)\sin(\Delta kz+\omega t) \qquad (4.38a)$$

where $\Delta k = \Delta\omega = k_1 - k_2$ and $k = \omega = k_1 + k_2$. Again, this is not a standing wave, in the strict sense of the term though it does have a vanishing Poynting vector. It is the case where two circularly polarized waves of the same amplitude but different frequency are traveling in opposite directions. The energy propagated by one wave in one direction is compensated for by the energy propagated by the other wave in the opposite direction. This may justify calling this solution a standing wave.

4.7. A PHYSICAL EXAMPLE

The vector potential given by Girotti et al. [see CHU AND OHKAWA 1987]

$$\mathbf{A}(\mathbf{r},t) = A_0\left(\sin\omega t,\ \cos\omega t,\ 0\right)\cos kz \tag{4.29a}$$

This vector potential can be constructed by the superposition of two circularly polarized waves of the same magnitude with similar helicity [JACKSON 1976] traveling in opposite directions [CHU 1983];

$$\mathbf{A}(\mathbf{r},t) = \frac{A_0}{2}\left[(-\sin[kz-\omega t]),\cos[kz-\omega t],0\right] + (\sin[kz+\omega t],\cos[kz+\omega t],0) \tag{4.29a'}$$

$$\mathbf{E}(\mathbf{r},t) = \frac{\omega A_0}{2}\left[(-\cos[kz-\omega t],-\sin[kz-\omega t],0) + (-\cos[kz+\omega t],-\sin[kz+\omega t],0)\right] \tag{4.29b'}$$

$$\mathbf{B}(\mathbf{r},t) = \frac{kA_0}{2}\left[(\sin[kz-\omega t],-\cos[kz-\omega t],0) + (\sin[kz+\omega t]),\cos[kz+\omega t],0\right] \tag{4.29c'}$$

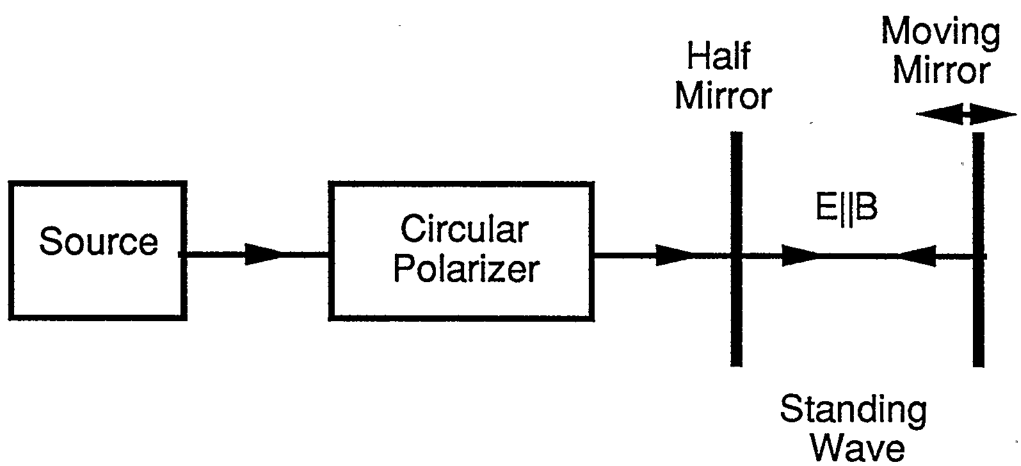

Figure 4.2 A schematic diagram of an experimental configuration that generates an E || B TEM standing wave pattern.

Figure 4.2 is a schematic diagram of an experimental configuration that generates the TEMSW pattern given by (4.29b') and (4.29c') with $\mathbf{E} \parallel \mathbf{B}$. Zaghloul and Buckmaster [1988] showed a thought-experiment realization of example of Harrington [1961] and

Chu and Ohkawa [1982] through a similar configuration. Brownstein [1991] pointed out that the electric field in the proposed set-up did not (and could not) vanish at the mirror. Subsection 4.9.1. will show how the example of Harrington [1961] and Chu and Ohkawa [1982] can be realized.

4.8. PHYSICAL PROPERTIES OF E || B TEMSW [ZAGHLOUL AND BUCKMASTER 1990]

(i) **E** || **B** TEMSW cannot be linearly polarized:

If the electric and magnetic fields of an **E** || **B** TEMSW are linearly polarized in, for example, the x-direction then $\mathbf{E} = (E_x, 0, 0)$ and $\mathbf{B} = (B_x, 0, 0)$ and

$$\nabla \times \mathbf{E} = \frac{\partial E_x}{\partial z}\hat{\mathbf{y}} - \frac{\partial E_x}{\partial y}\hat{\mathbf{z}} = \frac{\partial B_x}{\partial t}\hat{\mathbf{x}} \tag{4.39}$$

$$\nabla \times \mathbf{B} = \frac{\partial B_x}{\partial z}\hat{\mathbf{y}} - \frac{\partial B_x}{\partial y}\hat{\mathbf{z}} = \frac{\partial E_x}{\partial t}\hat{\mathbf{x}} \tag{4.40}$$

It is concluded from (4.39) and (4.40) that $E_x = E_x(x)$ and $B_x = B_x(x)$. However,

$$\nabla \cdot \mathbf{B} = \frac{\partial B_x}{\partial x} = 0 \tag{4.41}$$

so B_x is a constant and cannot be the only non-vanishing component of the magnetic field of a TEMSW and hence the electric and magnetic fields of an **E** || **B** TEMSW cannot be linearly polarized. Consequently, if **E** || **B** then both **E** and **B** must have more than one non-zero Cartesian component and, hence, may be decomposed into the superposition of two waves each with **E** || **B**.

A corollary is that the Lorentz invariant [JACKSON 1976] $I_1 = E^2 - B^2$ cannot be zero everywhere for **E** || **B** TEMSW.

ii) If **E** || **B** then $\mathbf{E} \neq \alpha\,\mathbf{B}$ where α is a constant:

If $\mathbf{E} = \alpha\,\mathbf{B}$ where α is a constant then from Maxwell's equations

$$\nabla \times \mathbf{E} = \alpha \nabla \times \mathbf{B} = \frac{\partial \mathbf{B}}{\partial t} \tag{4.42}$$

$$\nabla \times \mathbf{B} = \frac{\partial \mathbf{E}}{\partial t} = \alpha \frac{\partial \mathbf{B}}{\partial t} \tag{4.43}$$

$\alpha^2 = -1$ so **E** and **B** cannot be simultaneously real. Consequently, for **E** || **B** TEMSW,

$$\mathbf{E}(\mathbf{r}, t) = \alpha(\mathbf{r}, t)\, \mathbf{B}(\mathbf{r}, t)$$

where $\alpha(\mathbf{r}, t)$ may be a function of **r**, t or both. Substitution into Maxwell's equations shows that $\alpha(\mathbf{r}, t)$ must satisfy

$$\mathbf{B} \cdot \nabla \alpha = 0 \tag{4.44}$$

and

$$(1+\alpha^2)\nabla \times \mathbf{B} + \alpha \nabla \alpha \times \mathbf{B} = \frac{\partial \alpha}{\partial t}\mathbf{B} \tag{4.45}$$

which upon scalar multiplication by **B** yields

$$B^2 \frac{\partial \alpha}{\partial t} = \left(1+\alpha^2\right)\mathbf{B} \cdot (\nabla \times \mathbf{B}) \tag{4.46}$$

Equation (4.44) means that at any given instant of time, α is constant along any magnetic field line whereas (4.46) suggests that $\partial\alpha/\partial t$ is a measure of the curling of **B** in the direction of **B**. Note that if α is a function of t only, then **B** is parallel to $\partial B/\partial t$ and to **A** and satisfies

$$\nabla \times \mathbf{B} = \frac{1}{\left(1+\alpha^2\right)} \frac{\partial \alpha}{\partial t} \mathbf{B} \tag{4.47}$$

Also, if α is a function of $\mathbf{r}$ only, then $\nabla \times \mathbf{B}$ and $\partial \mathbf{B}/\partial t$ must be orthogonal to $\mathbf{B}$.

iii) If $\mathbf{E} \parallel \mathbf{B}$ in one Lorentz frame then no Lorentz frame exists where $\mathbf{E} \perp \mathrm{B}$:

This is a direct consequence of the invariance of the Lorentz invariant $I_2 = \mathbf{E} \cdot \mathbf{B}$.

iv) $\mathbf{E} \parallel \mathbf{B}$ TEMSW have uniform energy density:

Shimoda et al. [1990] have shown in their derivation of the condition for $\mathbf{E} \parallel \mathbf{B}$ waves in free space that such waves may be characterized by the fact that they have an energy density that is independent of both space and time. This is an interesting property that can be useful in obtaining uniform energy density resonant cavities as will be discussed in Subsection 4.9.1.

v) Direction of the propagation vector:

The direction of the propagation vector, $\mathbf{k}$, associated with $\mathbf{E} \parallel \mathbf{B}$ TEMSW, cannot be uniquely determined by investigating the direction of $\mathbf{E}$ and $\mathbf{B}$ at any point in space. The direction of $\mathbf{k}$ can be inferred from a consideration of the directions of the fields at different points. The two fields $\mathbf{E}$ and $\mathbf{B}$ remain parallel to each other at each point but their direction will change from point to point such that all of these directions lie in parallel planes, the transverse planes. This provides a unique definition of the direction of the propagation vector. It is probably more appropriate to use the term "wave vector" rather than "propagation vector" for these standing waves since they do not propagate.

vi) Transformation properties of $\mathbf{E} \parallel \mathbf{B}$ TEMSW:

a) Proper Lorentz transformations: If $\mathbf{E} \parallel \mathbf{B}$ and neither **E** nor **B** is static in one Lorentz frame then in any other frame $\mathbf{E} \nparallel \mathbf{B}$. This is obvious since if the Poynting vector vanishes in one frame it will not vanish in any other frame unless the wave is linearly polarized which is not the case for $\mathbf{E} \parallel \mathbf{B}$ TEMSW. This means that the Lorentz frame in which $\mathbf{E} \parallel \mathbf{B}$ is unique. This led Salingaros [1985] to state; "parallelizable fields for which a Lorentz frame exists where $\mathbf{E} \parallel \mathbf{B}$, and light-like fields, which are orthogonal and equal in magnitude, are mutually exclusive." Zaghloul [1988] noted that the term "light-like" in this definition is inappropriate since the term "light-like" originated in relativistic dynamics to describe 4-vectors of zero magnitude because the propagation 4-vector of light (as well as all TEM waves including $\mathbf{E} \parallel \mathbf{B}$ TEMSW) is of zero magnitude. Torrence [1991] pointed out that some waves may have $\mathbf{E} \nparallel \mathbf{B}$ and $\mathbf{E} \not\perp \mathrm{B}$ in all frames. This casts doubt on the appropriateness of the term "parallelizable." TEM waves with $I_1 = I_2 = 0$ are known as null or pure radiation fields [BARUT 1964].

b) Parity: If the charge density is a proper scalar, **E** is a polar vector and **B** is an axial vector. From (ii) for $\mathbf{E} \parallel \mathbf{B}$ TEMSW, $\mathbf{E} = \alpha\,(\mathbf{r},t)\,\mathbf{B}$ and $\alpha\,(\mathbf{r},t)$ must be a pseudoscalar.

c) Time-reversal: If the charge density is even under time-reversal, then **E** is even, **B** is odd and $\alpha(\mathbf{r},t)$ is odd under time-reversal.

d) Gauge transformations: **E**, **B** and hence $\mathbf{E} \parallel \mathbf{B}$ are invariant under gauge transformations [BARUT 1964].

vii) Magnetic Helicity:

Chu and Ohkawa [1982] showed that their example of **E** || **B** TEMSW possesses magnetic helicity, $\int \mathbf{A}\cdot\mathbf{B}dV$, with a time average of $2\pi\varepsilon/k$. All **E** || **B** TEMSW with α not dependent on **r** have a non-zero magnetic helicity value. This is obvious since for such waves **B** || **A** and the constant of proportionality between the two vectors is not a function of **r**. In particular, each member of Chu and Ohkawa's [1982] class of **E** || **B** TEMSW satisfies $\nabla\times\mathbf{B} = k\mathbf{B}$ and $\mathbf{B}=k\mathbf{A}$ and has a non-zero magnetic helicity since

$$\int \mathbf{A}\cdot\mathbf{B}dV = \int \frac{1}{k}B^2 dv = \frac{2}{k}U_{magnetic} \tag{4.48}$$

where $U_{magnetic}$ is the total magnetic energy whose time average is non-vanishing. Note that all **E** || **B** TEMSW with α not a function of **r** have a non-zero magnetic helicity. However, the example of Girotti et al. [see CHU AND OHKAWA 1987], which belongs to the class with α a function of **r** only, has zero magnetic helicity since $\mathbf{A}\cdot\mathbf{B} = 0$.

viii) Force-Free Nature of **E** || **B** TEMSW:

From the previous chapters, we know that force-free fields are magnetic fields that exhibit zero Lorentz force. This is achieved if the current density, **j**, is parallel to the magnetic field. i.e., $\mathbf{j} = b\mathbf{B}$ where b is a coefficient that may be a function of **r**. If the conductivity is scalar, $\mathbf{j} = \sigma\mathbf{E} = b\mathbf{B}$ which implies that **E** || **B**. This fact is not discussed in the literature since, as was shown in the previous chapters, the time dependence of such fields was either ignored or suppressed. Note that the case where both b and σ are constants is not in contradiction with (ii) above since (ii) was derived for a sourceless region of space. Note also that Chu and Ohkawa [1982] introduced their class of **E** || **B**

TEMSW in connection with force-free fields. However, not all **E** || **B** TEMSW are force-free.

ix) $k_\mu F^{\mu\nu}$ for **E** || **B** TEMSW:

E ⊥ **B** TEM waves with no standing wave component may be characterized by

$$h^\eta = k_\mu F^{\mu\eta} = (\mathbf{k}\cdot\mathbf{E}, -\omega\mathbf{E} + \mathbf{B}\times\mathbf{k}) = 0 \tag{4.51}$$

However, $h^\eta \neq 0$ for **E** || **B** TEMSW as direct substitution verifies. However, this does not characterize these waves since $h_\eta \neq 0$ for **E** |/| **B** TEM waves such as **E** ⊥ **B** TEMSW.

4.9. APPLICATIONS OF E || B TEMSW

4.9.1. Uniform Energy Cavities

Ruby lasers generally oscillate in several modes simultaneously. It was believed that the reason is that different parts of the laser are at different energy levels. To check this, it would be necessary to design a ruby laser with uniform energy density throughout. Evtuhov and Siegman [1965] outlined a technique to obtain this uniform energy density. The objective of this technique is to obtain a standing wave pattern of circularly polarized waves. The authors did not mention the observation that their proposed pattern has the property that **E** || **B.** They proposed that this standing wave pattern can be realized by placing two birefringent sections in the laser cavity before the mirrors as shown in Figure 4.3. The generated **E** || **B** TEMSW is identical to the example of Harrington [1961] and the total EM energy density, the electric energy density and the magnetic energy density are uniform throughout the volume.

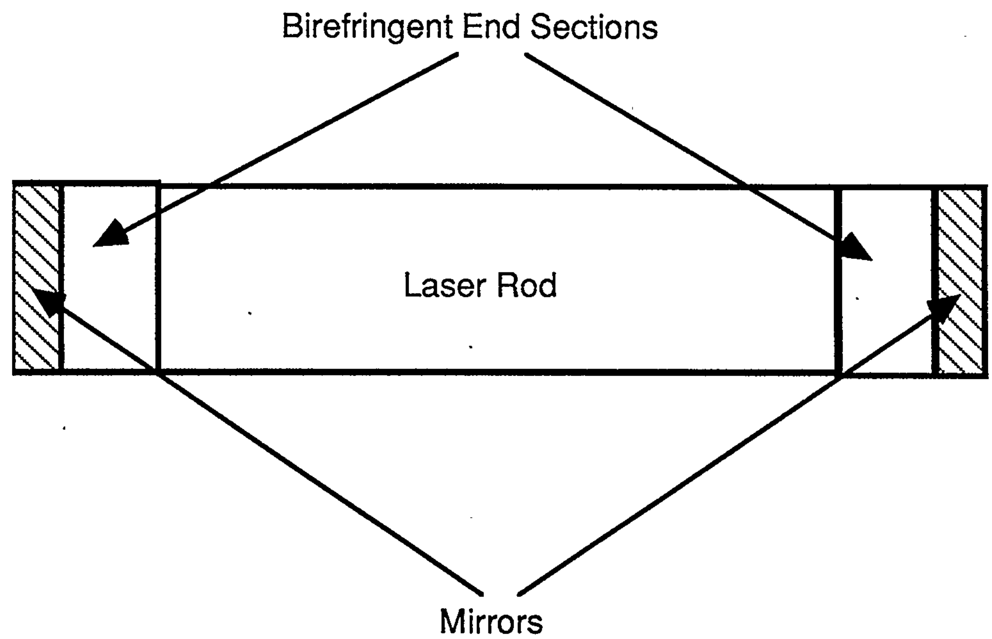

Figure 4.3 The configuration of a ruby laser proposed by Evtuhov and Siegman [1965] which realizes **E** || **B** TEMSW.

In electron paramagnetic resonance (EPR), a sample is placed in a homogeneous static magnetic field inside a cavity. The cavity is excited with microwaves, such that the photons have an energy that matches the energy gap generated by the static magnetic field. In EPR spectroscopy, it is desirable to determine the paramagnetic species in the sample. In EPR imaging, it is desirable to determine the spatial distribution of the paramagnetic species. This is achieved by placing the sample in a static magnetic field that has a different intensity at each point in the sample and then sweeping the magnetic field for a fixed microwave field. Marshall et al. [1988] have performed an imaging experiment where they filled the cavity with oxygen, a paramagnetic gas, and imaged the microwave magnetic energy density inside the cavity. This means that in EPR imaging, the distribution of the microwave magnetic energy can distort the resulting image. This can be avoided by having a uniform microwave magnetic energy density inside the cavity. If the cavity is excited with two orthogonal linearly polarized TEM waves simultaneously, a uniform energy density **E** || **B** TEMSW maybe set up in the cavity.

4.9.2. Reexamination of the EM Model of Matter

Equation (4.16)

$$\varepsilon_{sw}^2 = I_1^2 + I_2^2 = \varepsilon^2 - S^2 \tag{4.16}$$

can be compared with the expression for the invariant squared magnitude of the energy-momentum tensor p_μ of a particle [BARUT 1964]

$$c^2 p_\mu p^\mu = E^2 - c^2 p^2 = m_0{}^2 c^4 \tag{4.52}$$

where E=m c^2 is the self energy, p is the magnitude of the momentum, m_0 is the rest mass and m is the mass of the particle. The similarity between (4.16) and (4.52) suggests that somehow ε_{sw} plays the same role in electrodynamics that the rest energy of a particle plays in relativistic dynamics. The weakness of this similarity is that ε_{sw} is part of a rank-2 tensor whereas E is part of a rank-1 tensor. However, the similarity suggests that it may be possible to attribute the non-zero mass to waves similar in nature to **E** || **B** TEMSW.

4.9.3. EM Energy Storage

The generation of **E** || **B** TEMSW may allow the localization of EM energy. This result for **E** || **B** TEMSW differs from **E** ⊥ **B** TEMSW since the former has a vanishing Poynting vector and for the latter, there exists no Lorentz frame where the energy appears to be stationary. **E** || **B** TEMSW only require the direction of the electric and magnetic field vectors to change continuously which may be interpreted as that the energy does not move from any given location. This localization of EM energy may open the door for interesting applications such as storage ("bottling") of EM energy using, for example, a laser beam to make a laser battery. It is realized that, currently, the only way to generate either form of TEMSW requires reflectors and that losses occur upon reflection. This puts a practical limit on any possible utilization of this apparent localization in the immediate future.

4.10. CONCLUSIONS

This chapter reviewed the conditions under which TEM standing waves with **E** || **B** may exist in a sourceless medium as well as solutions of these conditions. It showed that these waves fit well within the theory of Lorentz invariants and that these waves have properties that distinguish them from other TEM standing waves. Applications of these waves have been proposed.

CHAPTER 5

CONCLUSIONS AND FURTHER RESEARCH

This thesis has reviewed the subject of FF magnetic fields. It has given a critical analysis of the FF condition and has shown that this condition behaves properly under parity and time reversal transformations. The proof that it is impossible to have completely FF magnetic fields has been reviewed. It has been also shown how general solutions of the FF condition in rectangular, cylindrical and spherical coordinate systems are obtained and explained how these fields can be applied both theoretically and experimentally to different disciplines of physics. This thesis has also shown that the "general" solution of the force-free condition in spherical coordinates only contained solutions with no-singularities. A more general solution was then obtained. The thesis then showed an example in magnetic plasma confinement where the more general solution can lead to experimental results different than those predicted by the previously accepted general solution. The proposed example included a new spheromak configuration. Spheromaks are more desirable than tokamaks because an external toroidal field is not required. The new spheromak configuration allows only a finite set of solutions to satisfy the boundary conditions set by the conducting shell. This enhances the stability of the spheromak against higher order modes.

Future work could concentrate on finding a practical method to contain the internal conducting spherical shell in place and to actually perform an experiment with this configuration.

The thesis reviewed the subject of transverse electromagnetic standing waves (TEMSW) with **E** || **B**. This subject has been controversial for the past few years. The thesis highlighted the fact that the controversy surrounding **E** || **B** TEMSW was predominantly due to lack of proper definition of terms and that when all terms used are clearly defined, the controversy disappears. An example of lack of definition of terms is the common belief that there are no time varying solutions to Maxwell's equation with **E** || **B**. This proved not to be the case. There are no *purely traveling, monochromatic*, time varying solutions of Maxwell's equation with **E** || **B**.

The thesis showed how **E** || **B** TEMSW fit in the larger picture of electromagnetics. It showed that these waves lead to proper Lorentz invariants. It derived the general conditions that these waves must satisfy and demonstrated how these conditions may be solved to find examples of **E** || **B** TEMSW. It showed how an example of these waves can be realized in an experiment. The thesis finally concluded by giving some possible applications of **E** || **B** TEMSW. The first of these applications is of interest in a number of physics disciplines: lasers and EPR and NMR imaging being examples. The laser example has been implemented in the past without labeling the resultant waves as **E** || **B** TEMSW. The EPR and NMR equivalents would be interesting and future research should concentrate on solving some of the technical questions such as whether birefringent sections be used or whether the physical example proposed in the thesis be implemented. The remaining proposed applications are very crude in nature and much more work has to be done in order to develop any form of theory around these applications.

REFERENCES

Aly, J.J., 1984, Astrophys. J., **283**, 349.

Barut, A.O., 1964, Electrodynamics and Classical Theory of Fields and Particles, (Dover, New York).

Baylis, W.E., and Jones, G., 1989, J. Phys. A.: Math. Gen., **A22**, 17.

Bekefi, G. and Barrett, A.H., 1977, Electromagnetic Vibrations, Waves, and Radiation, (MIT Press, Cambridge).

Bernstein, I.B., Frieman, E.A., Kruskal, M.D., and Kulsrud, R.M., 1958, Proc. Roy. Soc. (Lond.), **A244**, 17.

Bodenson, A., Marklin, G., An, Z.G., Chen, H.H., Lee, Y.C., and Liu, C.S., 1981, Phys. Fluids, **24**, 1682.

Brownstein, K.R., 1987a, Phys. Rev., **35**, 4856.

Brownstein, K.R., 1987b, Private communication.

Brownstein, K.R., 1991, Private communication.

Campbell, A.M., and Evetts, J.E., 1972, Critical Currents in Superconductors, (Taylor and Francis, London).

Chandrasekhar, S., 1956, Astrophys. J., **124**, 232.

Chandrasekhar, S. and Kendall, P.C., 1957, Astrophys. J., **126**, 457.

Chandrasekhar, S., 1957, Proc. Natl. Acad. Sci. U.S., **42**, 1.

Chandrasekhar, S., 1961, Hydrodynamic and Hydromagnetic Stability, (Dover, New York).

Chu, C. and Ohkawa, T., 1982, Phys. Rev. Lett., **48**, 837.

Chu, C., 1983, Phys. Rev. Lett. **50**, 139 .

Chu, C. and Ohkawa, T., 1987, Phys. Rev. Lett., **58**, 424.

Collin, R.E., 1991, Field Theory of Guided Waves, 2nd Ed., (IEEE Press, Pistacaway).

Collinson, C.D., 1987, J. Phys. A: Math. Gen., **20**, L193.

Cook, D.M., 1975, The Theory of the Electromagnetic Field, (Prentice-Hall, Englewood Cliffs).

Evangelidis, E.A., 1975, "Force Free Fields in Plasmas," published in Athens. Private Communication, October 1990.

Evangelidis, E.A., 1987, Astrophysics and Space Science, **143**, 113.

Evtuhov, V., and Siegman, A.E., 1965, Applied Optics, **4**, 142.

Ferraro, V.C.A. and Plumpton, C., 1961, An Introduction to Magneto-Fluid Mechanics, (Oxford U. P., London).

Flammer, C., 1957, Spheroidal Wave Functions, (Stanford U.P., Stanford, CA).

Freire, G., 1966, Am. J. Phys. **34**, 567.

Furth, H.P., Levine, M.A., and Waniek, R.W., 1957, Rev. Sci. Instrum., **28**, 949.

Gray, J.E., 1992, J. Phys. A: Math. Gen., **25**, 5373.

Hagler, M.O, and Kristiansen, M., 1977, An Introduction to Controlled Thermonuclear Fusion, (Lexington Books, Lexington).

Hansen, W.W., 1935, Phys. Rev., **47**, 139.

Harrington, 1961, Time-Harmonic Electromagnetic Fields, (McGraw-Hill, New York).

Jackson, J.D., 1976, Classical Electrodynamics, 2nd Ed. (Wiley, New York).

Jette, A.D, 1970, J. Mathematical Anal. Appl., **29**, 109 .

Jette, A.D. and Sreenivasan, S.R., 1969, Phys. Fluids, **12**, 2544.

Jordan, E.C., and Balmain, K.G, 1968, Electromagnetic Waves and Radiating Systems, 2nd Ed., (Prentice-Hall, Englewood Cliffs).

Josephson, B.D., 1966, Phys. Rev., **152**, 211.

Lee, K.K., 1983, Phys. Rev. Lett. **50**, 138.

Levy, R.H., 1962, A.R.S. J., **32**, 787 .

London, F., and London, H., 1934, Proc. Roy. Soc. (Lond.), **A149**, 71.

Lorrain, P., and Corson, D., 1970, Electromagnetic Fields and Waves, 2nd Ed., (Freeman, San Francisco).

Low, B.C., 1977, Astrophys. J., **212**, 234.

Lundquist, S., 1950, Ark. Fys., **2**, 35. (Was unavailable to the author)

Lundquist, S., 1951, Phys. Rev., **83**, 307.

Lust, K. and Schlute, A., 1954, Zeitschrift für Astrophysik, **34**, 263.

Maheswaran, M., 1986, J. Phys. A: Math. Gen., **19**, L761.

Marion, J.B., and Heald, M.H., 1980, Classical Electromagnetic Radiation, 2nd Ed., (Academic, New York).

Marshall, A., Suits, B.H., and Ulmor, T., 1988, J. Magn. Reson., **76**, 494.

Matsushita, T.,1981, Phys. Lett., **86**, 123.

Misner, C.W., Thorne, K.S. and Wheeler, J.A., 1973, Gravitation (Freeman, San Francisco) pp.480-1.

Morse, P.M., and Feshbach, H., 1953, Methods of Theoretical Physics Part I, (McGraw-Hill, New York), p.844.

Nakagawa, Y. and Raadu, M.A., 1972, Solar Physics, **25**, 127.

Nakagawa, Y., 1973, Astron. Astrophys., **27**, 95.

Nicholson, D.R., 1983, An Introduction to Plasma Physics, (John Wiley, New York).

Niu, K., 1988, Nuclear Fusion, (Campridge U.P., Cambridge, UK).

Pai, S-I., 1963, Magnetohydrodynamics and Plasma Physics, (Springer-Verlag, Vienna) p.17.

Panofsky, W.K.H., and Phillips, M., 1955, Classical Electricity and Magnetism, (Addison-Wesley, Reading).

Parker, E.N.,1954, Phys. Rev., **96**, 1686.

Parker, E.N.,1958, Phys. Rev., **109**, 1440 .

Parker, E.N., 1979, Cosmical Magnetic Fields, (Clarendon Press, Oxford).

Pathria, R.K., 1974, The Theory of Relativity, 2nd Ed., (Hindustan Pub. Corp., Delhi).

Portis, A.M., 1978, Electromagnetic Fields-Sources and Media, (Wiley, New York).

Purcell, E.M., 1965, Electricity and Magnetism (McGraw-Hill, New York).

Ramo, S., Whinnery, J.R. and Van Duzer, T., 1965, Fields and Waves in Communication Electronics, (Wiley, New York).

Rao, N.N., 1977, Elements of Engineering Electromagnetics, (Prentice-Hall, Englewood Cliffs).

Reitz, J.R., and Milford, F.J., 1967, Foundations of Electromagnetic Theory, (Addison-Wesley, Reading).

Robinson, D.C., 1969, Plasma Phys., **11**, 893.

Rohrlich, F., 1965, Classical Charged Particles, (Addison-Wesley, Reading).

Rosenbluth, M.N., and Bussac, M.N., 1979, Nuclear Fusion, **19**, 489.

Salingaros, N.A., 1985, Am. J. Phys. **53**, 361.

Salingaros, N.A., 1986, J. Phys. A: Math. Gen. **19**, L101 and L705.

Salingaros, N.A., 1987, Hadronic J., **10**, 109.

Schelkunoff, S.A., 1938, AIEE Trans., **57**, 744.

Schelkunoff, S.A., 1952, Advanced Antenna Theory, (John Wiley, New York).

Shadowitz, A., 1975, The Electromagnetic Field, (McGraw Hill, New York).

Shimoda, K., Kawai, T., and Uehara, K., 1990, Am. J. Phys., **58**, 394.

Smythe, W.R., 1950, Static and Dynamic Electricity, 2nd Ed., (McGraw-Hill, New York).

Spitzer, L., 1962, Physics of Fully Ionized Gases, (Interscience, New York).

Stenzil, R.L., and Urrutia, J.M., 1990, Phys. Rev. Lett., **65**, 2011.

Stratton, W., 1941, Electromagnetic Fields, (McGraw-Hill, New York).

Taylor, J.B., 1974, Phys. Rev. Lett.,**33**, 1139.

Taylor, J.B., 1986, Rev. Mod. Phys., **58**, 741.

Torrence, R., 1991, Private Communication.

Uehara, K., Kawai, T., and Shimoda, K., 1989, J. Phys. Soc. Jpn., **58**, 3570.

Van Hoven, G., 1981, "Simple-Loop Flares: Magnetic Instabilities," pp.217-275 and the two articles by J.Birn and K.Schindler, "Two-Ribbon Flares: Magnetostatic Equilibria" and "Two-Ribbon Flares: Post-Flare Loops," pp.337-428 in Solar Flare Magnetohydrodynamics, E. R. Priest, Ed., (Gordon and Breach, New York) and the references therein.

Varadan, V.K., Lakhtakia, A., Varadan, V.V., 1987, J. Phys. A: Math. Gen., **20**, 2649.

Voslanber, D., and Callebaut, D.K., 1962, Phys. Rev., **128**, 2016.

Wells, D.R., and Mill, R.G., 1962, in High Magnetic Fields, (MIT, Cambridge) pp.44-47.

Woltjer, L., 1958a, Proc. Nat. Acad. Sci., **44**, 489.

Woltjer, L., 1958b, Astrophys. J., **128**, 384.

Woltjer, L., 1959, Proc. Nat. Acad. Sci., **45**, 769.

Yangfang, B. and Luguang, Y., 1983, IEEE Trans. Mag., **Mag-19**, 324.

Zaghloul, H.,Volk, K. and Buckmaster, H.A., 1987, Phys. Rev. Lett. **58**, 423 Attention is drawn to a serious typographic error: the condition for $\mathbf{E} \parallel \mathbf{B}$ TEM waves should be $dA_+(\eta)/d\eta \not\parallel dA_-(\xi)/d\xi$ rather than $dA_+(\eta)/d\eta \parallel dA_-(\xi)/d\xi$. Also note that TEM waves with $\mathbf{S}(\mathbf{r},t)\neq 0$ are $\mathbf{E} \not\parallel \mathbf{B}$ traveling and standing waves rather than $\mathbf{E} \perp \mathbf{B}$ traveling and standing waves and ref.4 should be J. Phys. **A19**, L101 (1986).

Zaghloul, H., 1988, Am. J. Phys., **56**,.

Zaghloul, H., Buckmaster, H.A., and Volk, K., 1988, Am. J. Phys., **56**, 274.

Zaghloul, H., and Buckmaster, H.A., 1988, Am J. Phys., **56**, 801.

Zaghloul, H., 1989, Phys. Lett. A, **140**, 95.

Zaghloul, H., and Barajas, O., 1990, Am. J. Phys., **58**, 783.

Zaghloul, H., and Buckmaster, H.A., 1990, "Properties of TEM Standing Waves with **E** || **B**", Ann. Phys. Fr., 21-18.

Zaghloul, H., and Barajas, O., 1993, "Transverse Electromagnetic Waves with **E** || **B**," in Essays on the Formal Aspects of Electromagnetic Theory, A.Lakhtakia, Ed., (World Scientific, Singapore) pp.207-227.

Zaghloul, H., and Buckmaster, H.A., 1993, "Force-Free Magnetic Fields," in Essays on the Formal Aspects of Electromagnetic Theory, A.Lakhtakia, Ed., (World Scientific, Singapore) pp.183-206.

Zijlstra, H., 1967, Experimental Methods in Magnetism, Vol-1, (North Holland, Amsterdam).

APPENDIX A:

PARITY AND TIME REVERSAL OF HERMITIAN IRREDUCIBLE TENSOR OPERATORS

ABSTRACT

This appendix discusses irreducible spherical tensor operators (ISTOs) and shows that the common practice of ignoring the basis of such tensor operators can lead to confusion when their Hermitian properties are analyzed. The importance of the Hermitian properties of ISTOs arises from the fact that all operators with real observables in quantum mechanics must be Hermitian and that the application of the time reversal operator to ISTOs is related to Hermitian conjugation. The definitions of parity and time reversal of ISTOs are also discussed. The appendix concludes by reviewing some of the applications of parity and time reversal to ISTOs. These applications include a proof that the magnetic quantum number, m, can have only integer values, restrictions to the relation between the spatial parity of electric and magnetic multipole fields, and the development of selection rules. The last of these applications is important since it includes a derivation that odd order terms cannot be present in the non-relativistic spin Hamiltonian which describes the energy levels of a paramagnetic ion.

A.1. TENSORS

Tensors, like vectors, are mathematical entities that have components obeying a specific law under coordinate transformations. The components x_i of a vector where the subscript i takes on any value between 0 and n, the dimension of the vector space, transform according to

$$x'_j = \frac{\partial x'_j}{\partial x_i} x_i \qquad \text{(A.1)}$$

where the Einstein summation convention is used. A vector is a rank-1 tensor. A form of (A.1) that is more generalizable to tensors of higher rank is

$$x'_j = A^i_j x_i \qquad \text{(A.1a)}$$

A tensor of rank-2 has the transformation law

$$T_{ij} = A^k_i A^l_j T_{kl} \qquad \text{(A.2)}$$

with obvious generalization to tensors of higher ranks. A rank-2 tensor has n^2 components where n throughout this appendix is the dimension of the rank-1 vector space. These components are usually thought of in tensor calculus as the elements of an $n \times n$ matrix. It should be clear that there is no reason for this association other than mathematical convenience; it is equally appropriate to express these n^2 components as a column vector. Such a vector will have a transformation law similar to (A.1) but the indices will run from 1 to n^2. The transformation matrix of this single column rank-2 tensor is the result of the direct product of the matrix A^i_j with itself. This suggests that a rank-2 tensor can be treated as a rank-1 tensor but in a space of a larger dimension (=

n^2). In this appendix, all tensors will be expressed as column vectors. An example of a rank-2 tensor is the dyadic product of two rank-1 tensors (vectors) A_i and B_i;

$$T_{ij} = A_i B_j \tag{A.3}$$

Note that (A.3) is an expression for the components of the tensor. A basis is required to express the tensor; in Cartesian coordinates, the tensor whose components are given by (A.3) can be written as

$$T = AB = A_x B_x \mathbf{ii} + A_x B_y \mathbf{ij} + \cdots\cdots + A_z B_z \mathbf{kk} \tag{A.4}$$

where **i**, **j**, and **k** are unit vectors along the coordinate axes. The combinations **ii**, **ij**, ..., **kk** are called the bases (unit vectors) of the representation. This example will be used repeatedly throughout this appendix.

It is important to define both proper tensors and pseudo-tensors. Jackson [1976] stated; "If a tensor of rank N transforms under spatial inversion with a factor $(-1)^N$, it is called a true tensor or just a tensor, while if the factor is $(-1)^{N+1}$, it is called a pseudo-tensor." e.g., a proper tensor of rank-N, A, gets mapped under spatial inversion onto a new tensor, A', such that

$$A'(\mathbf{r}) = (-1)^N A(\mathbf{r}) \tag{A.5}$$

where $\mathbf{r}'$ is the result of the spatial inversion of $\mathbf{r}$; $\mathbf{r}' = -\mathbf{r}$. It is suggested that the existence of pseudo-tensors is an unfortunate inheritance of the twentieth century physicists. Vectors were known and well established before the advent of tensors and consequently, any mathematical quantity with a preferred direction was called a vector. For example in electromagnetics both the electric field and the magnetic field are vectors although the magnetic field is defined via the cross product of a vector operator with a vector which implies that it does not reverse its sign under spatial inversion. A partial remedy was to label these vectors as pseudo-vectors or axial vectors [JACKSON 1976, WIGNER 1959] as opposed to proper or polar vectors that reverse sign under spatial inversion. Terms such as pseudo- or proper can be avoided by realizing that pseudo-tensors are proper tensors of a different rank. In electromagnetics, the components of magnetic and electric fields of an electromagnetic wave are the components of a proper tensor of rank-2.

A.2. IRREDUCIBLE TENSORS

A rank-2 tensor can be expressed as a column vector with n^2 components that will transform under a coordinate transformation according to (A.1a). It is possible to perform some unitary transformation that makes it possible to divide the new tensor T' into n vectors with 1, 3, 5, ..., 2k+1 components. For example, when n = 3, the 3 vectors have 1, 3 and 5 components. Notice that calling them vectors implies that each of them obeys a transformation law similar to (A.1). These vectors are called irreducible tensors. Each of these vectors belongs to an invariant subspace of the n^2 dimensional linear space of which the rank-2 tensor is an element. The irreducible part arises from the fact that decomposing a space into its invariant subspaces is termed reduction. When it is not possible to reduce the space any further, the resulting decomposition is called irreducible. The matrix A^i_j is then block diagonal; its non-zero elements are in 1×1, 3×3 and 5×5 matrices along the diagonal of the matrix A^i_j. It is customary to state that T is decomposed into its irreducible tensors and write

$$T = T^{(0)} + T^{(1)} + T^{(2)} \tag{A.6}$$

and to call the number in brackets in the superscript the degree of the irreducible tensor. An irreducible tensor of degree-k has 2k+1 components. The irreducible tensors of the example given by (A.4) are [WIGNER 1959]

$$\mathrm{T}^{(0)}=\frac{1}{3}\left[(\mathbf{ii}+\mathbf{jj}+\mathbf{kk})\left(\mathrm{T}_{xx}+\mathrm{T}_{yy}+\mathrm{T}_{zz}\right)\right] \tag{A.7a}$$

$$\mathrm{T}^{(1)}=\frac{1}{2}\left[(\mathbf{jk}-\mathbf{kj})\left(\mathrm{T}_{yz}-\mathrm{T}_{zy}\right)+(\mathbf{ki}-\mathbf{ik})\left(\mathrm{T}_{zx}-\mathrm{T}_{xz}\right)+(\mathbf{ij}-\mathbf{ji})\left(\mathrm{T}_{xy}-\mathrm{T}_{zy}\right)\right] \tag{A.7b}$$

$$\begin{aligned}\mathrm{T}^{(2)}=\frac{1}{2}\Big[&\tfrac{1}{3}(2\mathbf{kk}-\mathbf{ii}-\mathbf{jj})\left(2\mathrm{T}_{zz}-\mathrm{T}_{xx}-\mathrm{T}_{yy}\right)+(\mathbf{ii}-\mathbf{jj})\left(\mathrm{T}_{xx}-\mathrm{T}_{yy}\right)+(\mathbf{jk}+\mathbf{kj})\left(\mathrm{T}_{yz}+\mathrm{T}_{zy}\right)\\&+(\mathbf{ki}+\mathbf{ik})\left(\mathrm{T}_{zx}+\mathrm{T}_{xz}\right)+(\mathbf{ij}+\mathbf{ji})\left(\mathrm{T}_{xy}+\mathrm{T}_{yx}\right)\Big]\end{aligned} \tag{A.7c}$$

Equations (A.7) give both the components and the unit vectors. The form of the unit vectors shows that each irreducible tensor is actually a Cartesian tensor of rank-2. It is important to realize that although the irreducible tensor of degree-1 has three components, it does not belong to the linear space, V_3^1, of tensors of rank-1. There exists an isomorphism (a one-to-one and onto relation) between V_3 and the linear space V_3^2 whose elements are irreducible tensors of degree-1 that result from the reduction of a rank-2 tensor. Notice that the same difference exists between the linear spaces of irreducible tensors of degree-1 that result from the reduction of tensors of different ranks. This difference plays a very important role when the parity of such irreducible tensors is considered.

A.3. IRREDUCIBLE SPHERICAL TENSORS

An irreducible tensor is a tensor that obeys the transformation law (A.1a) with the elements A^i_j replaced by the Wigner [1959] coefficients; a tensor that transforms under a rotation of the coordinate axes according

$$T'_{m'} = D^j_{mm'} T_m \tag{A.8}$$

is called an irreducible spherical (standard) tensor (IST). This definition is due to Wigner [1959] who used the term irreducible tensors in place of ISTs. The source of the term spherical is the fact that (A.8) is also the transformation law for spherical harmonics. The term standard arises from the fact that these tensors were the first form of irreducible tensors studied. ISTs are related to irreducible Cartesian tensors (ICT) by simple unitary transformations. The IST of degree-1 is related to the ICT of degree-1 (x,y,z) by [SILVER 1976 p.52]

$$T^1_{-1} = \frac{1}{\sqrt{2}}(x - iy)$$

$$T^1_0 = z$$

$$T^1_1 = \frac{-1}{\sqrt{2}}(x + iy) \tag{A.9}$$

Condon and Shortley [1935] realized that a degree-1 IST (without labeling it as such) satisfies the same commutation relations with the angular momentum operators as the spherical harmonics. Racah [1942] defined ISTs using these commutation relations. According to Racah [1942], an IST of degree-k is a tensor with 2k+1 components satisfying the same commutation relations with the angular momentum operators as the spherical harmonics.

$$\left[J_+, T_q^k\right] = \sqrt{(k-q)(k+q+1)}\,T_{q+1}^k$$

$$\left[J_z, T_q^k\right] = qT_q^k$$

$$\left[J_-, T_q^k\right] = \sqrt{(k+q)(k-q+1)}\,T_{q-1}^k \tag{A.10}$$

The two definitions are equivalent.

A.4. HERMITIAN IRREDUCIBLE SPHERICAL TENSOR OPERATORS

An operator, A, is said to be Hermitian if it is equal to its adjoint, A^+; i.e., $A^+ = A$ [FEYNMAN ET AL. 1965]. This definition is applicable to tensor operators. A tensor operator, T, is said to be Hermitian if it is equal to its adjoint [SCHWINGER 1952], T^+, i.e., $T = T^+$. The components of the adjoint of an ISTO satisfy the following commutation relations with the angular momentum operators:

$$\left[J_z, T_q^{k+}\right] = -qT_q^{k+} \tag{A.11a}$$

$$\left[J_-, T_q^{k+}\right] = -\sqrt{(k+q)(k-q+1)}\,T_{q+1}^{k+} \tag{A.11b}$$

$$\left[j_+, T_q^{k+}\right] = -\sqrt{(k-q)(k+q+1)}\,T_{q-1}^{k+} \tag{A.11c}$$

Investigation of (A.11) leads to the conclusion that T_q^{k+} behaves under spatial rotations similar to $(-1)^q T_{-q}^{k+}$. This means that the adjoint of a component of an ISTO is a component of an ISTO. i.e.,

$$T_q^{k+} = (-1)^q S_{-q}^k \tag{A.12}$$

where S is an ISTO. Schwinger [1952] makes a clear distinction between the adjoint of a component of an ISTO, T_q^{k+}, and a component of the adjoint of an ISTO, $T^{+k}_{\ q}$, and writes (using a different labeling convention)

$$T^{+k}_{\ q} = (-1)^q T_q^{k+} \tag{A.13}$$

This distinction will be used in later applications. The adjoint of each component of an ISTO behaves under rotations similar to a different component (except for the component with q = 0) and no two different components of an ISTO are the same. This becomes self evident if each component is associated with the direction of its unit vector. This means that only the component with q = 0 of an ISTO may be Hermitian [BRINK AND SATCHLER 1967] and that for an ISTO to be Hermitian, $T = T^+$, there must be an exact equality between the adjoint of each component (with its unit vector) and the component (with its unit vector) it rotates similar to. i.e.,

$$T_q^{k+}\mathbf{e}_q^{k+} = T_{-q}^{k}\mathbf{e}_{-q}^{k} \tag{A.14}$$

and if $T_q^{k+} = (-1)^q T_{-q}^k$ then $\mathbf{e}_q^{k+} = (-1)^q \mathbf{e}_{-q}^k$. This is best clarified by considering the following example:

Consider the matrix elements obtained when the components of the degree-1 angular momentum ISTO, $j_{+1} = \frac{-1}{\sqrt{2}}\left(j_x + ij_y\right)$, $j_0 = j_z$, $j_{-1} = \frac{1}{\sqrt{2}}\left(j_x - ij_y\right)$, are evaluated in the basis set $|1,+1\rangle$, $|1,0\rangle$, $|1,-1\rangle$ [SILVER 1976 p.211]:

$$T_1^1 = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$T_0^1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

$$T_{-1}^1 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{A.15}$$

where the matrix has the same label as the component. Notice that only the matrix T_0^1 is Hermitian. The matrices T_1^1 and T_{-1}^1 are not Hermitian (even their sum is not Hermitian) but they satisfy $T_q^{k+} = (-1)^q T_{-q}^k$. Since the expectation values of the angular momentum operators form a Hermitian matrix, it can be concluded that something must be missing from the above argument. The unit vectors, $\mathbf{e}_1^1 = \frac{-1}{\sqrt{2}}(\mathbf{i} + i\mathbf{j})$, $\mathbf{e}_0^1 = \mathbf{k}$, and $\mathbf{e}_{-1}^1 = \frac{1}{\sqrt{2}}(\mathbf{i} - i\mathbf{j})$ are missing. If every element of each matrix is associated with the corresponding unit vector then the matrix which results when these matrices are added together (which is the matrix of expectation values of the angular momentum ISTO) is Hermitian.

It is also stated in literature that if the components of an ISTO satisfy

$$T_q^{k+} = (-1)^{p-q} T_{-q}^k \tag{A.16}$$

where the choice of p is arbitrary [BRINK AND SATCHLER 1967], then the ISTO is said to be Hermitian. Schwinger [1952] followed the convention used initially by Racah [1942] and set p = 0. Edmonds [1957] followed the convention used by Fano and Racah [1959] and set p = k. The relation between the two notations is the same as the relation between the two notations of spherical harmonics: the spherical harmonics of Condon and Shortley [1935]

$$Y^l_m(\pi-\theta,\phi+\pi)=e^{im(\phi+\pi)}P^l_m(-\cos\theta)=(-1)^l e^{im\phi}P^l_m(-\cos\theta)=(-1)^l Y^l_m(\theta,\phi) \quad (A.17)$$

and the spherical harmonics of Biedenharn and Rose [1954] which were later used and labeled by Huby [1954], $\mathcal{Y}^l_m = i^l Y^l_m$. It is important to stress that the relation (A.16) is for the components and that the unit vectors must satisfy a similar relation.

A.5. PARITY AND TIME REVERSAL OF IRREDUCIBLE TENSOR OPERATORS

A.5.1. Parity

The parity of a tensor is its behavior under the operation of space inversion at a point, for simplicity, taken to be the origin of the coordinate system [LUDERS 1957 p.3]. This operation can be established by simply reversing the signs of the coordinate axes (passive inversion). It can also be realized by inverting every point of space about the origin (active inversion) [WOLF 1969 and HERMANN 1975]. It is important to realize that it is irrelevant which method of inversion is chosen to study a particular problem provided that the two methods are not mixed.

It is important to distinguish between two kinds of parity; intrinsic parity, ξ, [ROMAN 1960 p.427] which is also termed internal parity by Beresteteskii et al. [1971 Sec.13] and is defined to be odd or even depending on whether the tensor transforms under parity with a factor = -1 or 1 respectively and spatial symmetry around the origin which is termed both orbital parity [DE-SHALIT AND TALMI 1963 p.153 and BERESTETESKII ET AL. 1971 Sec.13] and spatial parity [SILVER 1976 p.211]. Examples of spatial symmetry are even functions for which f(**r**) = f(-**r**) and odd functions for which $f(\mathbf{r}) = -f(-\mathbf{r})$. The product of the intrinsic parity and the orbital parity forms what may be called total parity [DE-SHALIT AND TALMI 1963 p.153 and BERESTETSKII ET AL. 1971 Sec. 13] or space parity by Roman [1960 p.251].

A.5.1.1. Intrinsic Parity

A pseudo-tensor of some rank is a proper tensor of a higher rank. So, all tensors have the same intrinsic parity provided they are all expressed in the maximum possible rank. However, for irreducible tensors, the rank of the tensor is somehow suppressed in the notation. Also, it was stated earlier that it is possible to deduce the rank of an irreducible tensor from the rank of its unit vectors but this is also not straight forward if the general convention in literature of presenting the components without their unit vectors is followed. This difficulty in determining the rank of the tensor when combined with the usage of pseudo-tensors which renders the rank rather useless in determining the intrinsic parity of any given tensor, explains why the rare literature on the parity of irreducible tensors is rather ambiguous. In Sec. A.5.3, the case of polar and axial ISTOs will be discussed.

A.5.1.2. Orbital Parity

Cohen-Tannoudji et al. [1977] introduce even and odd operators: under the operation of parity, **P**, an operator, B, is mapped onto a new operator, B', where

$$B' = \mathbf{P}^{-1} B \mathbf{P} \tag{A.18}$$

If B' = B then the operator is said to be even and if B' = -B then the operator is said to be odd. Furthermore an even operator commutes with **P** and an odd operator anti-commutes with **P**. According to this classification, the position and the momentum operators are odd operators and the Hamiltonian is an even operator.

In simple words, an odd operator transforms an even function into an odd function and an odd function into an even function and vice versa. We can see that Cohen-Tannoudji et al. [1977] have given us a handy tool to determine whether a particular tensor is even or odd. This might not be suitable for double tensor operators. The parity of such operators is the product of the parities of its constituents. This is illustrated by the following example:

> The orbital angular momentum operator, **L**, is the vector product of the position, **r**, and the momentum, **p**, operators. The position and momentum operators are odd ICTs of degree-1 and are related by similarity transformations to ISTs of degree-1. Their dyadic product, **rp**, is an even Cartesian tensor operator of rank-2 which may be decomposed into 3 ISTs of degrees 0, 1 and 2. The IST of degree-1 that results from this reduction is **L** and carries the same (even) parity as the reducible Cartesian tensor. This demonstrates that the degree of an ISTO does not restrict its parity. The parity of an ISTO is independent of its degree.

A spherical harmonic (the prototype for ISTOs) of degree l in the Condon and Shortley [1935] phase convention, has even orbital parity if l is even and odd orbital parity if l is odd. However, it is not clearly explained in literature whether all ISTOs have a similar rule for their orbital parity or not.

A.5.2. Time Reversal of Irreducible Tensor Operators

An operator, $\hat{T}$, acting on a eigenfunction, $|\psi\rangle$, leads to a new eigenfunction, $|\psi'\rangle$, that may be expressed as an expansion of the eigenfunctions of the Hamiltonian, $|\phi_n\rangle$. i.e.,

$$\hat{T}|\psi\rangle = |\psi'\rangle = \sum c_n |\phi_n\rangle \tag{A.19}$$

where $c_n = \langle \phi_n | \psi' \rangle$.

The effect of the time reversal (inversion or reflection) operator on quantum mechanical states was first studied by Wigner [1959]. He gave the following definition of the time reversal operator: "It transforms a state ϕ into the state $\theta\phi$ in which all velocities (including the "spinning" of the electrons) have opposite directions to those in ϕ. (Hence, "reversal of the direction of motion" is perhaps more felicitous, though longer, expression than "time inversion.")" Brink and Satchler [1967], Edmonds [1957] and Wigner [1959] give a clear discussion of the effect of time reversal on quantum mechanical states:

> The time reversal operator, θ, is an operator that maps t onto -t and accordingly leaves the position coordinates unaltered whilst it reverses the signs of the momentum coordinates. To determine the effect of time reversal on a wavefunction, we study its effect on Schrodinger's equation

$$\hat{H}\psi = i\hbar \frac{\partial \psi}{\partial t} \tag{A.20}$$

$$\theta \hat{H}\psi = \theta i\hbar \frac{\partial \psi}{\partial t} \tag{A.21}$$

and if it is assumed that the Hamiltonian is real and time independent, then $\theta\hat{H} = \hat{H}\theta$ and since $\theta\frac{\partial}{\partial t} = -\frac{\partial}{\partial t}$,

$$\hat{H}\theta\psi = -i\hbar\frac{\partial\psi}{\partial t} \tag{A.22}$$

The complex conjugate of (A.20) is given by

$$\hat{H}^*\theta\psi^* = -i\hbar\frac{\partial\psi^*}{\partial t} \tag{A.23}$$

If it is assumed that there is a unitary operator, $\hat{O}$, such that $\hat{O}\hat{H}^* = \hat{H}\hat{O}$ then (A.23) is reduced to

$$\hat{H}\hat{O}\psi^* = -i\hbar\frac{\partial\psi^*}{\partial t} \tag{A.24}$$

Comparison of (A.22) and (A.24) shows that $\theta\psi$ satisfies the same Schrodinger equation as $\hat{O}\psi^*$. This means that it is possible to express the time reversal operator θ as the product of a unitary operator, $\hat{O}$, times the complex conjugation operator, $\hat{K}$; i.e.,

$$\theta = \hat{O}\hat{K} \tag{A.25}$$

Applying (A.25) to (A.19) yields

$$\theta|\psi\rangle = \hat{O}\hat{K}|\psi\rangle = \hat{O}|\psi^*\rangle \tag{A.26}$$

This leads to

$$\theta\hat{T}|\psi\rangle = \theta\hat{T}\theta^{-1}\theta|\psi\rangle = \sum\theta c_n|\phi_n\rangle = \sum c_n^*\theta|\phi_n\rangle = \sum c_n^*\langle\phi_n| \tag{A.27}$$

Taking the complex conjugate of (A.19) gives

$$\langle\psi|\hat{T}^{*} = \langle\psi'| = \sum c_n^{*}\langle\phi_n| \tag{A.28}$$

Dirac [1956] used (A.28) to define the adjoint of an operator. Comparison of (A.27) and (A.28) shows that it is possible to make the identification

$$\theta\hat{T} = \hat{T}\theta^{*} \tag{A.29}$$

In analogy with the parity operator, it is said that an operator is even under time reversal if it commutes with the time reversal operator and odd under time reversal if it anti-commutes with the time reversal operator. Merzbacher [1970] gives a similar definition for time odd and even tensors; an operator is time odd (-) or time even (+) if it satisfies

$$\theta T_q^k \theta^{-1} = \pm(-1)^q T_{-q}^k \tag{A.30}$$

A.5.3. On Polar and Axial Tensor Operators

Chamberlain [1977] dealt with the subject by classifying irreducible tensors into Cartesian (polar) and angular momentum (axial) tensors in a manner similar to the classification of vectors:

A polar tensor operator is a tensor operator satisfying

$$\mathbf{P} C_q^k \mathbf{P}^{-1} = (-1)^{k-q} C_{-q}^k \tag{A.31a}$$

$$\theta C_q^k \theta^{-1} = (-1)^{k-q} C_{-q}^k \tag{A.31b}$$

and an axial tensor operator is a tensor operator satisfying

$$\mathbf{P} J_q^k \mathbf{P}^{-1} = (-1)^{q} J_{-q}^k \tag{A.32a}$$

$$\theta J_q^k \theta^{-1} = (-1)^{q} J_{-q}^k \tag{A.32b}$$

However the effect of parity on the polar spherical tensor given by Chamberlain [1977] is

$$\mathbf{P}T^{1}_{\pm 1}\mathbf{P}^{-1} = -T^{1}_{\mp 1} \tag{A.33}$$

This shows that a polar tensor operator does not satisfy the definition given by Chamberlain [1977]. The same inconsistency occurs with axial tensor operators.

Polar and axial tensor operators were re-introduced by Chatterjee et al. [1983] and their properties were discussed by Tuszynski et al. [1985], Buckmaster et al. [1985] and Tuszynski [1985]. The definitions used by these authors do not lead to the inconsistencies exhibited by the definitions of Chamberlain [1977]:

A polar irreducible tensor operator is a tensor operator satisfying

$$\mathbf{P}T^{k}_{q}(\mathbf{E})\mathbf{P}^{-1} = (-1)^{k}T^{k}_{q}(\mathbf{E}) \tag{A.34a}$$

$$\theta T^{k}_{q}(\mathbf{E})\theta^{-1} = (-1)^{-q}T^{k}_{-q}(\mathbf{E}) \tag{A.34b}$$

where **E** is a prototype for any polar vector (i.e., **E**, **D**, **r** and **p**) and an axial tensor operator is a tensor operator satisfying

$$\mathbf{P}T^{k}_{q}(\mathbf{B})\mathbf{P}^{-1} = T^{k}_{q}(\mathbf{B}) \tag{A.35a}$$

$$\theta T^{k}_{q}(\mathbf{B})\theta^{-1} = (-1)^{k-q}T^{k}_{-q}(\mathbf{B}) \tag{A.35b}$$

where **B** is a prototype for any axial vector (i.e., **B**, **H**, **J**, **L** and **S**).

Note that Chatterjee et al. [1983] as well as Chamberlain [1977] defined polar and axial tensor operators according to their behavior under both parity and time reversal whereas in the classification of vectors, only their behavior under parity was taken into consideration. These extra restrictions lead to that some tensor operators are neither polar nor axial. Examples of such unclassified tensor operators are the double tensor operators formed by the product of a polar tensor and an axial tensor [BUCKMASTER ET AL. 1985 and TUSZYNSKI 1985], the complex potential tensor operator (which is not a double tensor operator) and the linear momentum spherical tensor operator (which is a double tensor operator resulting from the product of a position tensor and an orbital angular momentum tensor operator). In their application of polar and axial tensor operators Tuszynski et al. [1985] suggested that phenomenological spin tensor operators are polar tensor operators but then Buckmaster et al. [1985 p.4003] showed that this suggestion (assumption) was not necessary to derive a symmetry adapted generalized spin Hamiltonian. Rudowicz and Bramley [1987] criticized the suggestion that phenomenological spin is a polar vector on the grounds that if that was the case then the Zeeman electronic spin Hamiltonian ($\alpha\ \mathbf{B}\cdot\mathbf{S}$) appears to be a pseudo-scalar.

It is important to note that the classification of tensor operators into polar and axial tensor operators is based on the intrinsic parity of these operators.

A.5.4. ON "A NOTE ON TIME REVERSAL OPERATOR" BY WONG [1969]

In applying time reversal to tensor operators, caution should be exercised to avoid confusing the two spherical harmonics conventions discussed earlier. An example of such confusion arises in the note by Wong [1969] who suggested including some form of parity into time reversal and then gave an example;

$$2x = (x + iy) + (x - iy) = T^1_{-1} - T^1_1 \qquad \text{(A.36)}$$

This is the notation used by Racah [1942] in his first paper on irreducible tensors and is one of the two notations of Edmonds [1957 eqs. 5.1.3 and 5.5.3]. When Wong [1969] applied time reversal to (A.36), he assumed that $\theta T_q^k \theta^{-1} = (-1)^{k+q} T_{-q}^k$ which is the second notation used by Edmonds [1957 eq. 5.5.3] and Fano and Racah [1959]. This meant that the left hand side of (A.36) was time even whilst the right hand side was time odd. Wong [1969] concluded that some coefficient was missing. Actually according to Fano and Racah [1959]

$$2ix = (ix + y) + (ix - y) = -T_1^1 + T_{-1}^1 \tag{A.37}$$

and both sides of (A.37) are time odd. Also according to Racah [1942]

$$\theta T_q^k \theta^{-1} = (-1)^q T_{-q}^k \tag{A.38}$$

and both sides of (A.36) are time even. Note that Judd and Runciman [1976] gave a proper argument for introducing some form of parity into time reversal of total angular momentum eigenfunctions. However, it should be noted that parity is introduced into time reversal if the Condon and Shortley spherical harmonics phase convention is used and not if the spherical harmonics phase convention of Huby [1954] is used. This means that later works [BUCKMASTER ET AL. 1985 and ARMSTRONG 1971] that are based on the conclusions of Wong [1969] may be valid. Note, however, that the parity that Judd and Runciman [1976] had in mind is the orbital parity which is determined by the degree of the irreducible spherical tensor operator.

A.6. APPLICATIONS OF PARITY AND TIME REVERSAL OF ISTO'S

A.6.1. The Magnetic Quantum number, m, Can Only Assume Integer Values

The fact that the magnetic quantum number, m, can assume only integer values is generally proved in textbooks using the argument that the wavefunctions must be single valued at every point in space and since two points in space with the same **r** and θ and with their angles ϕ differing by integer multiples of 2π are the same point, then $e^{im\phi} = e^{im(\phi+2\pi)}$ and m must be an integer. Ross [1980] showed that this result can be obtained by using the invariance of the orbital angular momentum, **L**, under parity:

> Since they have no degeneracy, the eigenfunctions of L_z must possess definite parity and since all of them are of the form $e^{im\phi}$, it follows that the change ϕ to $\phi+\pi$ (the parity transformation) must give
>
> $$e^{im(\phi+\pi)} = \pm e^{im\phi} \tag{A.39}$$
>
> and hence m must be an integer.

A.6.2. The Spatial Parity of the Electric and Magnetic Multipoles

It is well known in electromagnetism that electromagnetic fields in source free regions can be expanded into spherical harmonics [SLATER AND FRANK 1947 Ch.III, BLATT AND WEISSKOPF 1952 Appendix-B, SILVER 1976 Ch.29, EDMONDS

1957 p.83, DE-SHALIT AND TALMI 1963 Ch.16 and ROSE 1957 Sec.27]. Such an expansion coincides with the multipole expansion of the fields; the spherical harmonic of degree-l corresponds to a l-pole. The electric scalar potential of an electric 2^l-pole is [SLATER AND FRANK 1947 p.34]

$$\psi(\mathbf{r}) = \sum_{m}\left(Ar^{l} + \frac{B}{r^{l}}\right)P_{m}^{l}(\cos\theta)(C\sin m\phi + D\cos m\phi) \tag{A.40}$$

where A, B, C and D are constants to be determined from the boundary conditions and P_m^l are associated Legendre polynomials. Since the electric field $\mathbf{E} = -\nabla\psi$, it is immediately obvious that a constant electric field results from an electric dipole for which $l = 1$. However, a constant electric field may also result from a magnetic monopole (if it exists) for which $l = 0$ and similarly for a constant magnetic fields, if we replace every electric with magnetic. It is straight forward to see that the electric field produced by a single electric charge (an electric monopole with $l = 0$) is an odd function of position (or simply odd), provided that the position of the charge is taken as the origin of coordinates. The electric field produced by two equal and opposite electric charges separated by a small distance (an electric dipole with $l = 1$) is an even function of the coordinates (or simply even), provided the center of the dipole is the origin of the coordinate system. This can be generalized to state that the electric (magnetic) field of an electric (magnetic) 2^l-pole is even or odd depending on whether l is odd or even respectively. It is known [ROSE 1957] that if the electric field is even then the magnetic field must be odd and vice versa; this means that the magnetic (electric) field of an electric (magnetic) 2^l-pole is even or odd depending on whether l is even or odd respectively. To verify this result, consider a constant electric field. A constant electric field can be produced from either an electric dipole, $l = 1$, and is even or from a magnetic monopole with $l = 0$ and is also even. These results agree with the fact that a constant electric field is even. It is customary in literature [BLATT AND WEISSKOPF 1952, DE SHALIT AND TALMI 1963] to state that the parity of the electric 2^l-pole is $(-1)^l$ and the parity of a magnetic 2^l-pole is $(-1)^{l+1}$. Two points have to be clarified; (i) in this statement; only spatial parity is considered; i.e., intrinsic parity is not considered and (ii) the parity assigned to a particular multipole is that of its magnetic field; this is only a matter of convention. The spatial parity of electromagnetic fields is of importance in the derivation of selection rules.

A.6.3. Selection Rules

By selection rules is meant those rules that must be satisfied such that a matrix element, $A_{jmj'm'}$, is non-zero

$$A_{jmj'm'} = \langle j', m' | \hat{A} | j, m \rangle \neq 0 \tag{A.41}$$

It appears to the author that Pauli [1926] was the first to use the expression "selection rules" and Dirac [1956 p.159] discussed the selection rules for the angular momentum operator. For an arbitrary ISTO, (A.41) is an expression for the matrix element of a component, i.e.,

$$A_{jmj'm'} = \langle j', m' | T_q^k | j, m \rangle \neq 0 \tag{A.42}$$

and selection rules for such matrix elements are based on a number of considerations.

A.6.3.1. The Superselection Rule

A very strong selection rule, generally called the superselection rule is that only states with integral or half-integral j values may be superposed. This selection rule is a result of the fact that states with half-integral j values change their sign after a rotation of 2π whilst states with integral j values do not change their sign after such a rotation and it is inadmissible to mix such states since all states in a superposition of states should either change their sign or not [MERZBACHER 1970 p.379]. The application of this superselection rule shows that any matrix element with $j-j' = -\frac{2k+1}{2}$, where k is an integer, vanishes.

A.6.3.2. Only Integer Rank Hermitian Irreducible Tensor Operators Exist

Schwinger [1952 p.261] showed that Hermitian tensors, $T^{+} = T$, exist only for integer degree:

> Using Racah's [1942] phase convention, the adjoint of a tensor operator of degree-k is a new tensor operator of degree-k whose components are defined to be
>
> $$T^{+k}_{\ q} = (-1)^{q} T^{k+}_{-q} = i^{2q} T^{k+}_{-q} \tag{A.43}$$
>
> and the adjoint of this new tensor is
>
> $$T^{++k}_{\ \ q} = i^{2q} T^{k+}_{-q} = i^{2q}\left(i^{2q}\left(i^{-2q} T^{k+}_{q}\right)\right)^{+} = (-1)^{2q} T^{k}_{q} \tag{A.44}$$
>
> Now, since k can be only integer or half-integer and if k is integer then q is integer and vice versa then it is legitimate to write (A.44) as

$$T^{++} = (-1)^{2k} T \tag{A.45}$$

It is obvious that if T is Hermitian, $T^{+} = T$, then $T^{++} = T$ and k must be integer.

It is often stated in literature that irreducible tensor operators of half-integer degree cannot exist and an argument similar to that mentioned above is given as a proof. The difference between the two arguments is that whilst Schwinger [1952 p.261] appealed to the definition of Hermitian tensor operators, other authors appeal to the fact that the adjoint of the adjoint of a linear operator is the linear operator itself [DIRAC 1956 p.27]. However if half-integer degree ISTO did exist, they cannot be Hermitian and therefore cannot have a definite time reversal behavior. i.e., if k is half-integer

$$\theta T_q^k \theta^{-1} \neq \pm T_{-q}^{k+} \tag{A.46}$$

rather

$$\theta T_q^k \theta^{-1} = \pm i T_{-q}^{k+} \tag{A.47}$$

Moreover, such a tensor operator would couple particles of integer spin with particles of half-integer spin. The fact that such a coupling has never been observed physically justifies the exclusion of tensor operators of half-integer degree.

A.6.3.3. The Selection Rule of Multipole Radiation

Blatt and Weisskopf [1952 p.588] derived that electric dipole radiation with $l = 1$ can only be emitted from a particle if the final and initial states of the particle are of different parity.

"The transition probability T_{ab} between the states ψ_a and ψ_b is proportional to the matrix element H'_{ab}:

$$H'_{ab} = \frac{1}{2}\int \psi_b^*(\mathbf{j}\cdot\mathbf{A})\psi_a d\tau$$

where **j** is the current density operator and **A** is the vector potential of the emitted radiation. The vector potential **A** is proportional to the electric field **E** in a light wave of definite frequency. Hence **A** has the same parity as **E**. The current operator **j** has parity -1, since it is an ordinary (polar) vector and hence changes sign under a reflection of the coordinates. (More precisely: because the operator for **j** anti-commutes with the parity operator **P**: **jP** = -**Pj**.) The parity of (**j**·**A**) therefore is the opposite of the parity of **E**; since **E** and **H** always have opposite parity in multipole radiation, we conclude that (**j**·**A**) has the same parity as the magnetic field in the emitted wave. This was the reason for using **H** in the parity classification of the multipole fields."

From this paragraph, we can arrive at the following selection rule:

The electromagnetic radiation emitted by a particle in its transition from a state ψ_a to a state ψ_b must satisfy

$$p_a + p_b + p_{radiation} = 2n \qquad \text{(A.48)}$$

where p_a, p_b and $p_{radiation}$ are the spatial parities of the states ψ_a and ψ_b and the emitted radiation respectively and n is an integer. This means that the emitted radiation must be even if the two states have the same spatial parity and must be odd if the two states have opposite spatial parities.

It should be noted that

(i) The parity mentioned in the above quotation is spatial parity,

(ii) It is incorrect that the spatial parity of the current operator **j** is -1 because it is a polar vector; that **j** is a polar vector is related to its intrinsic parity and is not relevant to the above argument. For a single particle [BRINK AND SATCHLER 1967]

$$\mathbf{j} = \frac{eg_L}{m}\mathbf{P} = -\frac{i\hbar eg_L}{m}\nabla \tag{A.49}$$

where **P** is the momentum operator, m is the particle's mass, g_L is its orbital g factor (eg_L is the particle's charge). **j** is an odd operator since **P** is an odd operator. This is what Blatt and Weisskopf [1952] meant in the footnote on p.588. This lack of distinction between intrinsic parity and spatial parity is evident in the works of de-Shalit and Talmi [1963 p.152 eq.16.33] and Rose [1957 p.138] who stated that **j** is an odd operator because it is a polar vector.

(iii) The point of the proof above is that if a function is integrated over all space then the result of integration must be zero if the function is odd. This was explained in a complicated form by de-Shalit and Talmi [1963 p.152].

(iv) This selection rule has been invoked to show that matrix elements corresponding to odd rank terms in the multipole expansion of the crystal field in EPR are zero. This argument was used by Bleaney and Stevens [1953 p.128] where $p_a = l_a$ and $p_b = l_b$ and l_a and l_b are the orbital angular momentum quantum numbers of the two states respectively, the initial and final states are assumed to have the same l, and the multipoles under consideration are electric multipoles.

(v) This selection rule does not apply to transitions between spin wavefunctions since the evaluation of matrix elements involving spin eigenfunctions involves a sum over all possible spin coordinates rather than an integration over all space. Transitions between spin states are of interest in EPR.

A.6.3.5. The Selection Rules of Tuszynski [1985]

Tuszynski [1985] gave the following selection rules which he derived from consideration of the parity and time reversal of ISTOs:

For an axial tensor operator defined by eqs. (A.35),

$$\langle s,l,j,m|\theta^{-1}\left(\theta T_q^k(\mathbf{B})\theta^{-1}\right)\theta|s,l',j',m'\rangle = (-1)^{j-m+k-q+j'-m'}\langle s,l,j,-m|T_{-q}^k(\mathbf{B})|s,l',j',-m'\rangle \tag{A.50}$$

since $\theta|l,j,m\rangle = (-1)^{j-m}|s,l,j,m\rangle$ and $\langle s,l,j,m|\theta^{-1} = (-1)^{j-m}\langle s,l,j,m|$ [BIEDENHARN AND ROSE 1953 p.736]. The 3-j symbols have the following symmetry

$$\begin{pmatrix} j & k & j' \\ m & -q & -m' \end{pmatrix} = (-1)^{j+k+j'}\begin{pmatrix} j & k & j' \\ -m & q & m' \end{pmatrix} \tag{A.51}$$

since the Wigner-Eckart theorem relates the matrix elements to a reduced matrix element, $\langle\cdots\|\cdots\|\cdots\rangle$, through

$$\langle s,l,j,m|T_q^k|s,l',j',m'\rangle = (-1)^{j-m}\begin{pmatrix} j & k & j' \\ m & -q & -m' \end{pmatrix}\langle s,1,j\,\|T^k\|s,l',j'\rangle \tag{A.52}$$

we find that

$$\langle s,l,j,m|T_q^k|s,l',j',m'\rangle = (-1)^{j+k+j'}\langle s,l,j,-m|T_{-q}^k|s,l',j',-m'\rangle \tag{A.53}$$

substitution of (A.53) into (A.50) yields

$$\langle s,l,j,m|T_q^k(\mathbf{B})|s,l',j',m'\rangle = (-1)^{2(j+k+j')+m-1-m'}\langle s,l,j,m|T_q^k(\mathbf{B})|s,l',j',m'\rangle$$

(A.54)

and this is an identity since 2(j+k+j') is an even integer (Tuszynski [1985] did not mention the evenness of this integer) due to the superselection rule (actually the superselection rule leads to that j+j' must be an integer; that k must be an integer has been shown above) and m-q-m' = 0 due to the conservation of angular momentum. Note that $l+l'$ which appeared in the exponent of (-1) in eq. 26 of Tuszynski [1985] should be j+j' for polar ISTOs. It is concluded that time reversal produces no constraints on the matrix elements of all ISTOs.

Similarly, the parity of axial ISTOs leads to

$$\langle s,l,j,m|\mathbf{P}^{-1}\mathbf{P}T_q^k(\mathbf{B})\mathbf{P}^{-1}\mathbf{P}|s,l',j',m'\rangle = (-1)^{l+l'}\langle s,l,j,m|T_q^k(\mathbf{B})|s,l',j',m'\rangle \qquad \text{(A.55)}$$

which requires that $(l+l')$ must be an even integer for a non-zero matrix element of an axial ISTO and the parity of polar ISTOs leads to

$$\langle s,l,j,m|\mathbf{P}^{-1}\left(\mathbf{P}T_q^k(\mathbf{E})\mathbf{P}^{-1}\right)\mathbf{P}|s,l',j',m'\rangle = (-1)^{l+l'+k}\langle s,l,j,m|T_q^k(\mathbf{E})|s,l',j',m'\rangle \qquad \text{(A.56)}$$

which requires that $(l+l'+k)$ must be an even integer.

This classification is incomplete and incorrect for the following reasons:

(i) as was mentioned earlier, some ISTOs are neither axial nor polar (the complex potential as well as the linear momentum ISTOs),

(ii) it considered only the spatial parity of states and the intrinsic parity of ISTOs, and

(iii) consideration of the intrinsic parity alone should not lead to any selection rule.

However, these selection rules should be understood in exactly the same way as the selection rules of multipole fields. i.e. if the spatial parity of an ISTO is determined by (A.34) then its selection rule is "$l+l'$+k must be even" and if the spatial parity of the ISTO is determined by (A.35) then its selection rule is "$l+l'$ must be even".

A.7. CONCLUSIONS

This appendix has reviewed the definition of Hermitian irreducible tensor operators (ISTO) and the effect of the operations of parity and time reversal on ISTOs. It showed that the common practice of ignoring the basis of such tensors can lead to a confusion when one is analyzing their Hermitian properties. Examples of this confusion were presented. The appendix also reviewed the definitions of parity and time reversal of irreducible spherical tensor operators. This appendix also showed that there are two forms of parity commonly used: intrinsic parity and spatial parity. These two parities are different operations that lead to different results. Care must be taken when applying parity to state which parity is really meant.

This appendix has reviewed some of the applications of parity and time reversal to irreducible spherical tensor operators. These applications include a proof that the magnetic quantum number, m, can only have integer values, restrictions to the relation between the spatial parity of electric and magnetic multipole fields, and the development of selection rules.

REFERENCES

Armstrong, L., 1971, Theory of the Hyperfine Structure of Free Atoms, (Wiley, New York), p.151.

Beresteteskii, V.B., Lifshitz E.M. and Pitaevskii, L.P., 1971, Relativistic Quantum Theory, (Addison-Wesley, Reading).

Biedenharn, L.C., and Rose, M.E., 1953, Rev. Mod. Phys., **25**, 729.

Blatt, J.M., and Weisskopf, V.F., 1952, Theoretical Nuclear Physics, (Wiley, New York) p.526.

Bleaney, B., and Stevens, K.W.H., 1953, Rep. Prog. Phys., **16**, 108, (p.128).

Brink, D.M., and Satchler, G.R., 1967, Angular Momentum, (Clarendon Press, Oxford).

Buckmaster, H.A., Chatterjee, R., and Tuszynski, J.A., 1985, J. Chem. Phys., **83**, 4001.

Chamberlain, J.R., 1977, Physica B, **89B**, 107, (pp.111-2).

Chatterjee, R., Tuszynski, J.A., and Buckmaster, H.A., 1983, Can. J. Phys., **61**, 1613.

Cohen-Tannoudji, C., Diu, B., and Laloe, F., 1977, Quantum Mechanics, Vol. I, (Hermann, Paris).

Condon, E.U., and Shortley, G.H., 1935, The Theory of Atomic Spectra, (Campbridge U.P., Cambridge).

De-Shalit, A., and Talmi, I., 1963, Nuclear Shell Theory, (Academic Press, New York).

Dirac, P.A.M., 1956, The Principles of Quantum Mechanics, (Clarendon Press, Oxford).

Edmonds, A.R., 1957, Angular Momentum in Quantum Mechanics, (Princeton U.P., New Jersey).

Fano, U., and Racah, G., 1959, Irreducible Tensorial Sets, (Academic Press, New York).

Feynman, R.P., Leighton, R.B., and Sands, M., 1965, The Feynman Lectures in Physics, Vol.III, (Addison-Wesley, Reading), p. III-20-3.

Herman, R., 1975, Ricci and Levi-Civita's Tensor Analysis Paper, (Math. Sci., Brookline).

Huby, 1954, Proc. Phys. Soc. (Lond), **67**, 1103.

Jackson, J.D., 1975, Classical Electrodynamics 2nd Ed. (Wiley, New York)

Judd, B.R., and Runciman, W.A., Proc. Roy. Soc. Lond. A., **352**, 91.

Landau, L.D., and Lifshitz, E.M., 1975, The Classical Theory of Fields, Pergamon, London, pp.10, 62, 63 and 65.

Luders, G., 1957, Ann. Physics, **2**, 1.

Merzbacher, E., 1970, Quantum Mechanics, (Wiley, New York).

Pauli, W., 1926, Zeit. Phy., **36**, 336.

Racah, G., 1942, Phys. Rev., **62**, 438.

Roman, P., 1960, Theory of Elementary Particles, (North Holland, Amsterdam).

Rose, M.E., 1957, Elementary Theory of Angular Momentum, (Wiley, New York) p.131.

Ross, A.W., 1980, Am. J. Phys., **48**, 413.

Rudowicz, C., and Bramley, R., 1987, J. Phys. C: Solid State Phys., **20**, L77.

Schwinger, J., 1952, "On Angular Momentum," in Quantum Theory of Angular Momentum, 229, (p.261).

Silver, B.L., 1976, Irreducible Tensorial Methods, (Academic Press, New York)

Slater, J.C., and Frank, N.H., 1947, Electromagnetism, Ch. III, (McGraw-Hill, New York).

Tuszynski, J.A., Buckmaster, H.A., Chatterjee, R., and Boteler, J.M., 1985, J. Magn. Reson., **63**, 241.

Tuszynski, J.A., 1985, Physica A, **131A**, 289.

Wigner, E.P., 1959, Group Theory and its Applications to the Quantum Mechanics of Atomic Spectra, (Academic Press, New York).

Wolf, A.A., 1969, Am. J. Phys., **37**, 531.

Wong, E.Y., 1969, J. Chem. Phys. **50**, 4120.

APPENDIX B:

A SIMPLE METHOD FOR THE EVALUATION OF THE NOISE PERFORMANCE OF MICROWAVE MIXER AND DETECTOR DIODES

ABSTRACT

This appendix describes a simple method for the evaluation of the noise performance of the various microwave (microwave) mixer diodes which can be used in electron paramagnetic resonance (EPR) spectrometers. The advantage of this method over other methods is that it optimizes the performance of the diodes for EPR applications by determining the optimum operating conditions for each microwave diode. This appendix describes the instrumentation and the experimental procedure that have been developed and reports the results of evaluating various types of microwave diodes. The analysis of the measurements reported suggests that

1) the noise spectrum of microwave diodes can be divided into three frequency intervals where the spectral power density satisfies: (i) $1/f^2$, (ii) $1/f^{\alpha}$ where $0 < \alpha < 1$ depending on the microwave bias power level and, (iii) f independent i.e. thermal
2) the "1/f" corner frequency increases as the microwave bias power level increases, and

3) the DC load has almost no effect on the noise performance of the microwave diodes.

B.1. INTRODUCTION

Most EPR spectroscopists are motivated to maximize the sensitivity of their spectrometers so that they can observe new phenomena as well as make measurements with improved signal-to-noise ratio (SNR). Consequently, they are interested in using microwave diodes that optimize the sensitivity of their EPR spectrometers. It is unfortunate that microwave diode manufacturers specifications are not appropriate for EPR spectrometers. Most microwave mixer diodes are designed for use in radar equipment and, hence, are specified by their conversion loss and noise figure at 30 MHz. Detector diodes are specified by their broad band tangential sensitivity at frequencies up to approximately 1 MHz and their typical DC bias current. EPR spectrometers use magnetic field modulation frequencies up to 100 kHz and synchronous demodulation at the modulation frequency so the "1/f" corner frequency of mixer diodes is of interest [VAN DER ZIEL 1954]. The term "1/f" arises from the observation that the dominant noise spectral component at low frequencies is approximately inversely proportional with the frequency at which it is measured. This low frequency has been called flicker noise. There is no accepted theory for the 1/f dependence of flicker noise because it is still not fully understood. Recent work suggests that it is related to condition of the surface of the semiconductor [KING 1978 p.81]. Most EPR spectroscopists resort to a trial and error approach to evaluate diodes by determining the spectrometer sensitivity using the SNR of a standard sample. This approach does not ensure that the diode is evaluated under optimum operating conditions. A few systematic studies using a variety of noise measurement techniques have been attempted for this purpose [BUCKMASTER AND DERING 1965, BUCKMASTER AND RATHIE 1971, BUCKMASTER AND COHN-SFETCU 1973].

Recently Zaghloul et al. [1990a and 1990b] described a simple method for diode evaluation to assist in the design of a 9 GHz EPR spectrometer using magnetic field modulation below 1 kHz. This appendix is an expansion of this work. It first reviews the concept of noise in electric devices, introduces the common microwave mixer/detector diodes, and gives the parameters that characterize the performance of these diodes. It then gives a brief review of the methods that have been used to measure the noise performance of microwave mixer/detector diodes with emphasis on their use in EPR spectrometers. It, finally, introduces the configuration to measure this performance and reports the results of performing measurements on different diodes.

B.2. NOISE [BREWSTER 1987]

This section concerned with the electric noise present in electric circuits. Noise is a term that generally refers to unwanted signals. The Cambridge Dictionary of Science and Technology gives the following definition of noise as used in the telecommunications industry: "(1) In general, any unwanted disturbance superimposed on a useful signal and tending to obscure its information content. (2) Any undesired disturbance within a useful frequency band. Those caused by parallel services may be termed interference." These unwanted signals maybe man-made or natural. A common example of man-made noise is the 60 Hz noise that is generated by the brushes of AC electric motors. Examples of natural noise are lightening and the magnetic phenomena associated with solar activity. Noise is unpredictable (random) in nature. Noise in electronic circuits and devices has a number of sources. A simple resistor produces noise due to collisions of the finitely charged electrons with the molecules. Since the mobility of electrons depends on the temperature of the material, it is expected that the noise produced depends on the temperature; at absolute zero temperature, the noise vanishes. This noise is called thermal or Johnson noise and for a resistor, R, its mean square level, V_n^2, is:

$$V_n^2 = 4kTBR \tag{B.1}$$

where T is the absolute temperature (Kelvin), B is the bandwidth over which the noise is measured (noise bandwidth) and k is Boltzmann's constant (1.38×10^{-23} J/K). The electron

pulses produced by molecular collisions are extremely short. Thus, the power spectral density of thermal noise is constant over a wide frequency range and, therefore, is often referred to as white noise. In electric circuits, the maximum power is delivered from a source when the load impedance is equal to the source impedance (R_S for the single resistor). Accordingly, the maximum noise power that the resistor can deliver is

$$P_n = kTB \tag{B.2}$$

Equation (B.2) defines the noise temperature, T_S, for all sources of white noise

$$T_s = \frac{\text{noise power delivered by source in bandwidth B}}{\text{kB}} \tag{B.3}$$

The noise temperature is often used to quantitatively characterize the noise performance of devices. Another quantity performing a similar function is the noise figure (factor), F, of a two-port network (an electric circuit with two ports: an input port and an output port)

$$F = \frac{\text{available noise power at two}-\text{port output}}{\text{available noise power if two}-\text{port were noiseless}} \tag{B.4}$$

with the source at the standard ambient temperature, T_o = 290 K. Note that if the noise generating device is an amplifier of a power gain, G, then its noise figure is

$$F = \frac{P_o}{GkT_oB_n} \tag{B.5}$$

where N_o is the available noise power out of the amplifier, B_n is the bandwidth of the amplifier and T_o = 290 K. The noise power available at the output of the amplifier is

$$N_o = GN_i + GkT_AB_n = GkB_n(T_s + T_A) \tag{B.6}$$

where T_A is the noise temperature of the amplifier and N_i is the noise power input to the amplifier, and

$$\frac{S_o}{N_o} = \frac{GS_i}{GkB_n(T_s + T_A)} = \frac{T_s}{T_s + T_A}\left[\frac{S_i}{N_i}\right] \tag{B.7}$$

where S_i and S_o are powers of the input and output signals of the amplifier, respectively. The noise figure and the noise temperature are interchangeable concepts since they are related through

$$F = \frac{1}{G}\left[\frac{N_o}{N_i}\right]_{T_s=T_o} = \frac{T_o + T_A}{T_o}. \tag{B.8}$$

Combining (B.8) and (B.7) yields

$$F = \left(\frac{S_i}{N_i}\right) \Big/ \left(\frac{S_o}{N_o}\right) \tag{B.9}$$

which provides an alternate definition of the noise figure. The noise figure and noise temperature as defined above are sufficient for the characterization of a single device.

When M devices are cascaded, the effective noise temperature, T_e, and effective noise figure, F_e, can be calculated from the individual devices noise temperatures, T_1, T_2, ..., T_M or noise figures F_1, F_2, ..., F_M, power gains, G_1, G_2, ..., G_M and the overall system bandwidth, B_n.

$$N_o = G_1G_2\cdots\cdots G_M kT_sB_n + G_1G_2\cdots\cdots G_M kT_1B_n + G_2G_3\cdots\cdots G_M kT_2B_n + \cdots\cdots + G_M kT_M B_n \tag{B.10}$$

$$F_e = \frac{1}{G_1G_2\cdots\cdots G_M}\frac{N_o}{N_i}\bigg|_{T_s=T_o} = \frac{T_o+T_1}{T_o} + \frac{T_2}{G_1T_o} + \frac{T_3}{G_1G_2T_o} + \cdots\cdots + \frac{T_M}{G_1G_2\cdots\cdots G_M T_o}$$

$$= F_1 + \frac{F_2-1}{G_1} + \frac{F_3-1}{G_1G_2} + \cdots\cdots + \frac{F_M}{G_1G_2\cdots\cdots G_{M-1}} \tag{B.11}$$

This formula is known as the Friis equation.

B.3. MICROWAVE DETECTOR DIODES

Diodes are devices that ideally allow electric charge to flow in only one direction. Ideal diodes do not exist. Actual diodes may be described as resistors having a small resistance in one direction (few ohms) and a large resistance in the other direction (few kilo ohms). There are both vacuum tube and semiconductor diodes. Semiconductors are made of quadrivalent elements doped with either trivalent elements (p-type) or pentavalent elements (n-type). This doping leads to the presence of excess electrons in n-type semiconductors and a lack of electrons (holes) in p-type semiconductors. Placing an n-type semiconductor next and a p-type semiconductor leads to the migration of electrons from the n-type to the p-type and holes from the p-type to the n-type. This migration leads to (i) an electrostatic potential between the two semiconductors with the n-type being at the higher potential and (ii) a region near the junction surface which is depleted of both electrons and holes and is called the depletion region. This wafer of the two semiconductors forms a semiconductor diode (often called p-n diode) and allows a current if the external source supplies electrons to the n-type semiconductor. Diodes connected in this manner are said to be forward biased otherwise they are reverse biased. Reverse biased diodes retard the flow of charge unless the potential across the diode exceeds a critical voltage called the breakdown voltage. Voltages larger than the breakdown voltage cause the diode to appear as if it is shorted.

When an ordinary p-n diode is forward biased, an excess of charge is stored in it. If the voltage is reversed rapidly, charge carriers near the external connections are swept away giving rise to a large reverse current and charge carriers near the p-n junction disappear due to the recombination of oppositely charged carriers. Only when this recombination process is completed does the diode recover its blocking capability [MAZDA 1981]. The time it takes a diode to recover its blocking capability is called recovery time and may be shortened by introducing impurities such as gold into the junction. These impurities help trap the free charge carriers. Other methods to shorten the recovery time include the use a single type of charge carriers (point contact and Schottky diodes) and increasing the doping density thus reducing the size of the depletion region (tunnel diodes) [MAZDA 1981]. Typical I-V characteristics of point contact, Schottky and tunnel diodes are given in Fig. B.1 [KING 1979].

B.3.1. Point Contact Diodes [GINZTON 1957]

Point contact diodes have a pointed metal whisker which is in contact with an n-type semiconductor wafer. If the work function of the semiconductor is lower than that of the whisker then a potential barrier is generated which impedes the motion of electrons from the metal to the semiconductor. This barrier is increased when the semiconductor is externally made more negative than the metal and vice versa. This means that the whisker-semiconductor contact (point contact) provides a rectifying junction. The I-V characteristics of point contact diodes are identical to those of p-n diodes.

B.3.2. Schottky Diodes [BEEFHORTH AND GOLDSMID 1970]

Similar to a point contact diode, a simple Schottky diode is formed by a metal-semiconductor (n-type) junction rather than a semiconductor-semiconductor junction as in a p-n diode. The difference between point contact diodes and Schottky diodes is that the metal whisker is replaced by metal deposited on the semiconductor in the latter. The metal is separated from the semiconductor by a thin insulating film. The insulating film is kept as thin as possible. The inevitable oxide film forming on the surface of the semiconductor acts as the insulating film. The theory of operation of Schottky diodes is that if the semiconductor has a lower work function than the metal then electrons will flow from the semiconductor to the metal and reside on the metal-insulator interface.

Schottky diodes have similar I-V characteristics to p-n diodes. The major difference between the characteristics of the two diodes is that Schottky diodes can operate as rectifiers up to very high frequencies since only one species of charge carriers is available (electrons) and they have an extremely short relaxation time.

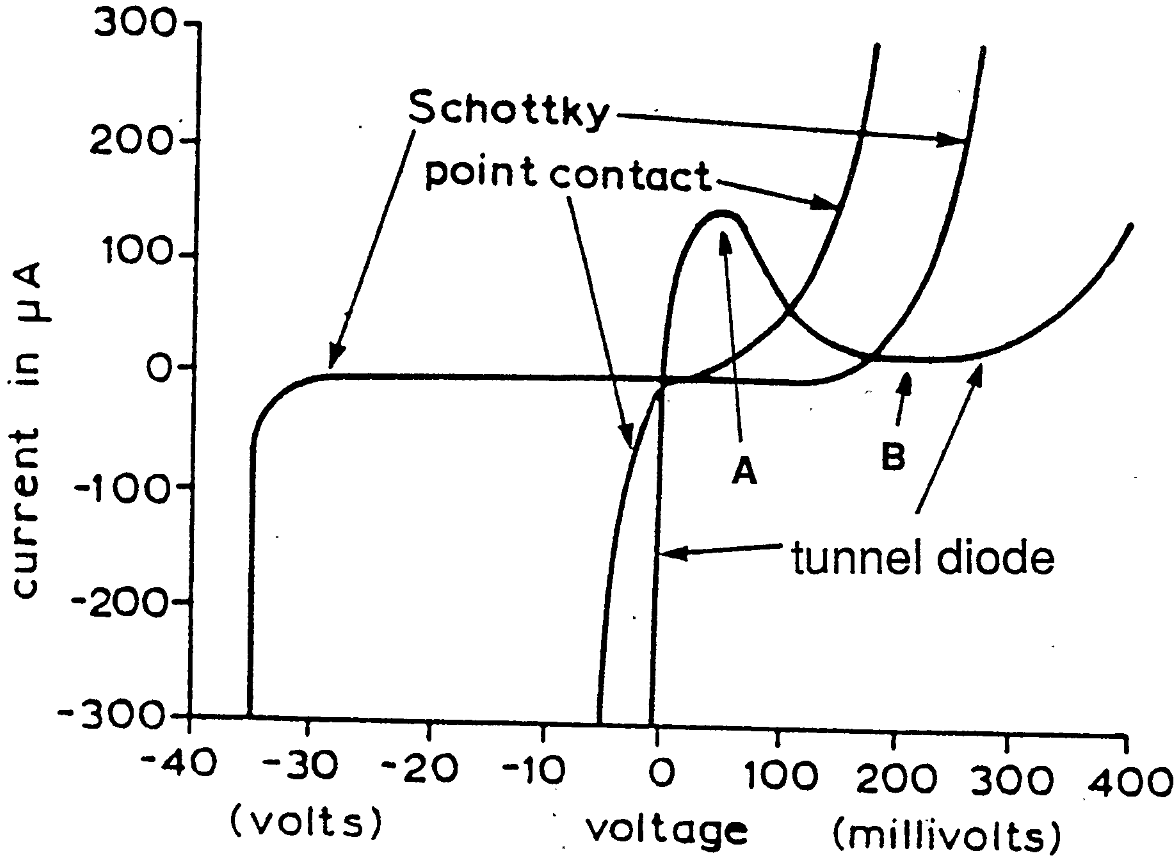

Figure B.1 I-V characteristics of typical point contact, Schottky and tunnel diodes [KING 1979].

B.3.3. Tunnel Diodes [MAZDA 1981]

A tunnel diode is a densely doped p-n diode. Increasing the doping density has the effect of narrowing the depletion region which permits charge transfer by quantum mechanical tunneling. Tunneling leads to the desirable feature of extremely short recovery times. Another byproduct of tunneling is the negative resistance region of the diode shown between points A and B on Fig. B.1. There is a current in reverse biased tunnel diodes because of tunneling by electrons from the p-type to the n-type. When forward biased, current rapidly increases by tunneling of electrons from the n-type to the p-type up to point A. This leads to a shift in the electron distributions on both sides of the junction and a fall in the tunneling current up to point B. After this point, the current increases with voltage as in ordinary p-n diodes. Tunnel diodes have a very small breakdown voltage and are restricted to small signal applications. It is important to note that tunnel diodes are often termed "back diodes." The two terms are used interchangeably. In this appendix, the term "tunnel diode" will be used.

B.3.4. Waveguide Diode Mounts

The design of the waveguide crystal mounts for microwave diodes is complicated. In principle, the diode is positioned where there is a maximum of a standing wave pattern set up by a proper termination inside the mount. The microwaves impinging on the diode are rectified and fed through, typically, a dumb bell transformer from the diode to a coaxial connection. The inductance and capacitance of this transformer (to the body of the waveguide) act as a low pass filter and attenuate the high frequency after mixing (or detection).

This section has reviewed some of the diodes used for detection and mixing of microwave signals. Many other diodes that are used for various purposed in microwaves circuits including the generation of microwave signals exist. The analysis of such diodes is beyond the scope of this thesis. However, before concluding this section it is appropriate to mention p-i-n diodes since they are used in the experimental configuration to evaluate the noise performance of microwave diodes. Adding an intrinsic layer of a semiconductor material between the p-type and n-type semiconductors increases the width of the depletion region and leads to PIN diodes that can act as microwave switches.

B.4. FACTORS CHARACTERIZING DETECTOR DIODE PERFORMANCE

Tangential Sensitivity: the power necessary to produce a specific SNR at the output.

DC Bias Current: the current required for a diode to operate with forward bias.

IF Impedance: output impedance of the diode at the intermediate frequency (IF). It is a function of the bias level and the design of its mount.

Noise Temperature: the ratio of the noise power of a diode to that of a resistor whose resistance is equal to the diode IF impedance at $T_o = 290K$.

Conversion or Insertion Loss: qualitatively, it is a measure of the loss of signal due to the presence of the diode. Quantitatively, it is defined as the ratio of the output intermediate frequency signal power to the input RF power in the signal to be tested.

Noise Figure: the ratio between the noise power at two-port output and the available noise power if two-port were noiseless with the source at the standard ambient temperature, $T_0 = 290$ K.

1/f Corner Frequency: The "1/f" corner frequency is the frequency at which the flicker noise becomes smaller than the thermal noise (noise caused by the random motion of electrons in a conductor) and the shot noise (noise generated by the random passage of charges through barriers). This corner frequency is generally determined by fitting a straight line (with a slope of approximately -1) to the low frequency portion of the noise curve (plotted on log-log scale), plotting a straight line (slope = 0) to the high frequency portion of the same curve and determining the frequency where the two lines intersect.

Equivalent Noise Resistance (R_A): the resistance that will produce thermal noise of the same magnitude as the noise generated by the diode.

Figure-of-Merit (M): a measure of the amount of power lost in the diode. It is equivalent to the insertion loss of mixer diodes. Quantitatively it is defined to be

$$M = \frac{\beta R_o}{\sqrt{R_o + R_A}} \tag{B.20}$$

where R_0 is the diode impedance, R_A is the equivalent noise resistance of the diode and β is the ratio of the output current to the input power. The minimum detectable power is inversely proportional to M.

B.5. DIODES NOISE MEASUREMENTS

Measurement of the noise of microwave detector and mixer diodes is equivalent to determining the noise that the diode adds to the measured signal. A good measure of this noise is the noise figure described above. However, different factors can be used if the noise is measured for a specific application. For example, in EPR, it suffices to measure the sensitivity of an EPR spectrometer given actual operating conditions. This has generally been the approach taken by EPR spectroscopists. This method does not guarantee that the performance of the diode is optimized.

Buckmaster and Dering [1965] developed a simple experimental configuration to measure the conversion gain and noise temperature of microwave mixer diodes. However, they did not ensure that the diodes were matched at both the microwave and intermediate frequencies. Buckmaster and Rathie [1971 I] developed a system to measure the noise spectrum form 10 Hz to 1 MHz using a tunable switching radiometer. This system was utilized by Buckmaster and Rathie [1971 II] to measure the noise performance of microwaves detector diodes. The advantage of this system was that it allowed correct matching of the diodes. Buckmaster and Cohn-Sfetcu [1973] developed an experimental configuration to measure microwave diodes noise. The configuration was that of a synchrodyne spectrometer with a reflection cavity. Their configuration had a stabilized source delivering microwaves to two arms: a reference arm where the reference signal is attenuated and phase shifted, and a signal arm where the signal is attenuated and reflected from a reflection cavity through a magic-T. The reference arm is connected to the H-port of a second magic-T and the signal arm is connected to its E-port. The sum and difference ports of this second magic-T were connected to two diode mounts where the diodes under test were mounted. The outputs of the diodes were added through a matching transformer that limited the spectrum analyzed to the range 0.1-20 kHz. The amplified output of the transformer was recorded on a tape recorder, played back, filtered through a low pass anti aliasing filter, sampled, digitized and analyzed using a Fourier Analyzer. The Fourier Analyzer allowed the accumulation of the noise power spectral density over a wide bandwidth in a short period of time. This novel feature of their configuration was modified and re-introduced in the configuration developed for the noise measurements described in this appendix. However, the shortcoming of this configuration is that it retained components of the spectrometer that are not essential to characterize the diodes. In particular, the cavity was present during the measurements which made it difficult to isolate the effects of the diode from those of the cavity. The study was performed on tunnel and point contact diodes. Cohn-Sfetcu and Buckmaster [1974] later removed the cavity from their configuration and used a PIN modulator and studied the noise performance of diodes near 9.3 GHz.

Jaggard and King [KING 1978] developed a configuration similar to that of Cohn-Sfetcu and Buckmaster [1974] for use in a comprehensive study of noise in microwave diodes at 5 GHz. This configuration had a source delivering microwaves to the H-port of a magic-T whose E-port is terminated through a matched load. The sum and difference ports of this magic-T were connected to two arms: a signal arm where the signal could be modulated with a ferrite modulator and precisely attenuated and phase shifted using a precision coaxial attenuator and a slotted line respectively, and a reference arm where the signal was attenuated. The reference arm is connected to the sum port of a second magic-T and the signal arm is connected to its difference port. The E-port was terminated through a matched load and the H-port port was connected to the diode mount. The diode had an external bias circuit. The output of the diode was amplified and recorded using a narrow band voltmeter. The best feature of this configuration is its simplicity. Its most obvious shortcoming is using the narrow-band voltmeter rather than measuring the entire low frequency spectrum and taking the Fourier transform. The study was performed on tunnel, Schottky and point contact diodes.

The remainder of this appendix describes a simple method for microwave diode evaluation which was developed to assist in the design of an X-band EPR spectrometer using magnetic field modulation below 1 kHz. The configuration is similar to that of Jaggard and King [KING 1979] except that

(i) the operating frequency was 9.4 GHz rather than 5 GHz,

(ii) the fast Fourier transform (FFT) of wideband low frequency measurements was used rather than recording narrow band voltage measurements,

(iii) self bias rather than DC external bias was used,

(iv) the option to change the loading of the diodes was introduced, and

(v) a PIN diode rather than a ferrite modulator was used so that many different signals could be introduced simultaneously.

Measurements were performed on various microwave diodes: point contact, Schottky, tunnel and Gallium Arsenide. The output of the microwave diode was recorded under various operating conditions: video load, microwave bias power and modulation frequency. Measurements were performed to determine the effect of the preamplifier that followed the microwave diode on the SNR. The recorded spectra were used to determine the SNR, the noise floor and the "1/f" corner frequency. Comparison of these factors for the different types of microwave diodes has shown that some

Schottky-barrier diodes have noise figures at 1 kHz that are as low as those for tunnel diodes.

B.6. INSTRUMENTATION

Figure B.2 is a block diagram of the experimental configuration used to make the measurements reported in this appendix. The microwave power was generated using a VA232 klystron with an HP716B klystron power supply. The frequency was stabilized using a Curry, McLaughlin and Len, (MOS1) microwave oscillator stabilizer. The maximum microwave power delivered using this source was 200 mW. The microwave frequency was measured using an HP5246L electronic counter with an HP5255A frequency converter. The microwave signal was divided equally between two arms using a magic-tee with the difference arm terminated. The reference or bias arm had an isolator, a precision rotary vane attenuator, a precision rotary vane phase shifter and an isolator. The reference arm power was monitored using a 20 dB directional coupler and a microwave power meter. The signal arm had an isolator, a precision rotary vane attenuator, an HP8735B PIN modulator, a rotary vane precision attenuator and an isolator. The microwave power in the signal arm was attenuated by 100 dB: 30 dB before and 70 dB after the PIN modulator to isolate the VSWR of this modulator and to ensure linear operation. The PIN modulator was driven using a function generator. This method of modulation was chosen to generate repeatable low level modulation signals. The microwave power in these two arms was summed using another magic-tee whose difference arm was terminated. The sum of these two signals traveled through a bi-directional coupler and an on-off switch into a slide screw tuner diode mount De Mornay-Bonardi (DBG-319). The attenuated outputs of the bi-directional coupler were

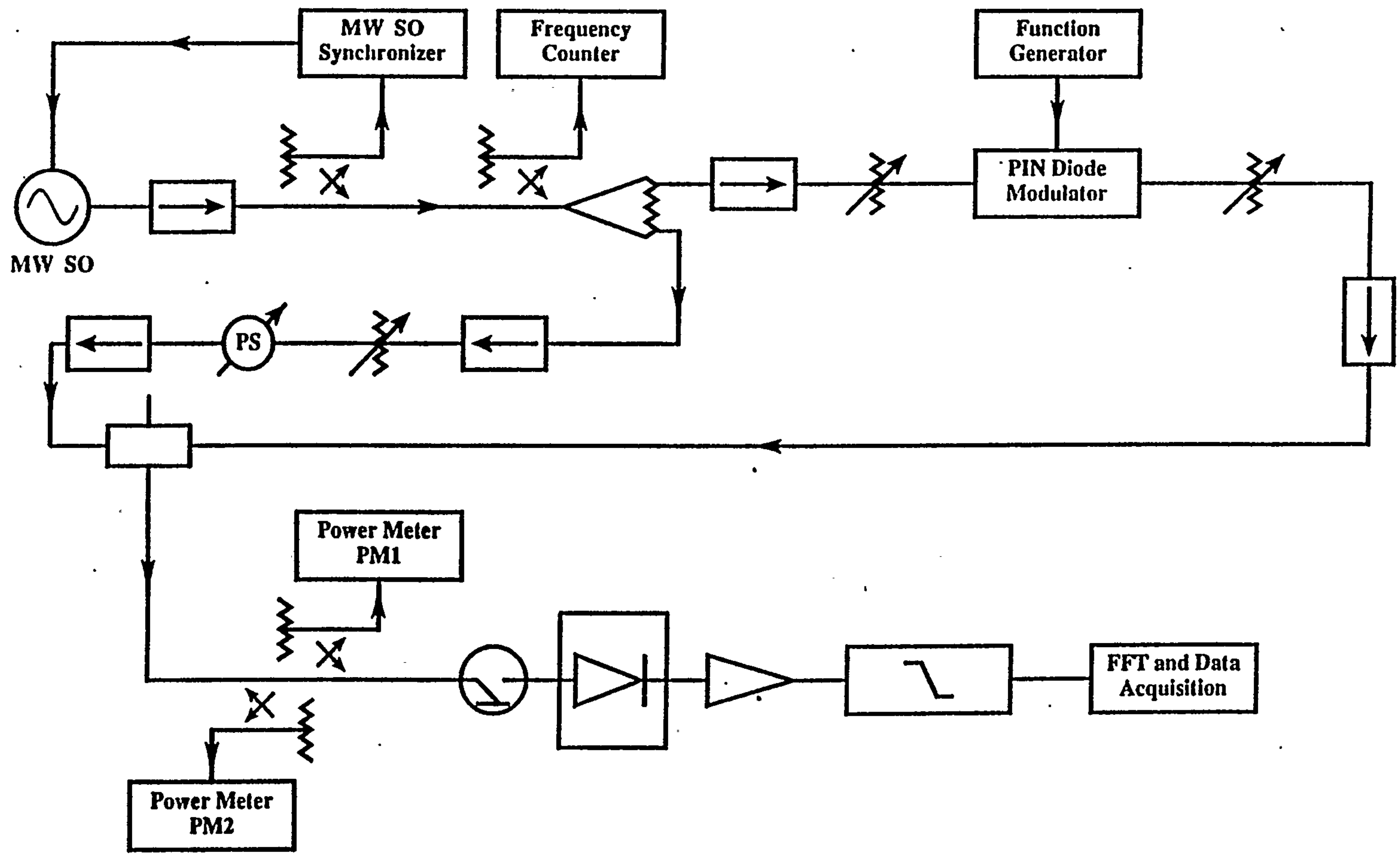

Figure B-2 Block diagram of the experimental configuration developed to evaluate the performance of microwave mixer diodes connected to power meters connected to monitor the incident and reflected power levels. The output of the diode was amplified using a Stanford Research SR552 preamplifier and processed using an antialiasing low pass Butterworth filter (Rockland 1042F using both of its filters in series with 0 dB gain) and a Data Precision DP6000 digital oscilloscope. The input to the SR552 preamplifier was shunted using a parallel pair of 1 H, 48 Ω inductances in series with a variable resistor since this amplifier did not provide the necessary DC return path to the diode. This resistance, combined with the resistance of the inductances, provided the DC load for the diode. The inductances were chosen to maximize the reactance and minimize the total resistance of this load. The DC voltage drop across the load resistance was monitored using a digital multimeter.

B.7. THEORY OF MEASUREMENTS

The analysis presented here follows that of Collin [1992] except for that in his analysis the local oscillator (LO) and the received radio frequency (RF) signal were of different frequencies. Consequently, he had three diode equivalent circuits:

1) the equivalent circuit for the RF signal,
2) the equivalent circuit for the LO signal, and
3) the equivalent circuit for the signal at the intermediate frequency (IF) i.e. the difference between the RF frequency and the LO frequency.

In the present analysis the LO is replaced by the reference arm signal, and the RF is replaced by the signal arm signal. Both of these signals operated at the same frequency leading to a zero IF. Accordingly, only two equivalent circuits need to be considered: an RF+LO equivalent circuit and a base band (BB) equivalent circuit.

A microwave diode can be represented by the circuit shown in Fig. B.3 where c_d is the diode junction capacitance, R_s is the series resistance, L_s is the series inductance and c_j is the package capacitance. Elements other than the ideal diode can be ignored in the analysis provided that the microwave frequencies involved are low. The I-V characteristic of an ideal diode is described by

$$i_d = I_s[\exp(v_d/v_c) - 1] \tag{B.12}$$

where I_s is the diode reverse saturation current, the diode characteristic voltage $v_c = \frac{nkT}{e}$, T is the diode temperature, n is a diode dependent parameter and is generally between 1 and 1.5, and v_d is the voltage incident on the diode. In practice, v_d is approximately 0.025V.

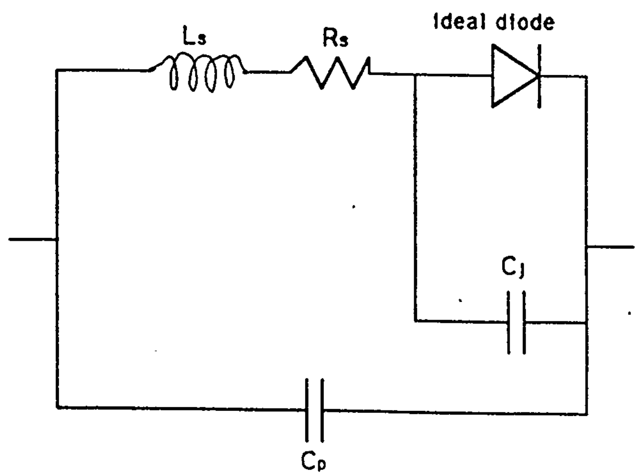

Figure B.3
Equivalent circuit for a microwave dioden [COLLIN 1992]

Further assumptions needed to develop simplified equivalent circuits are:

1) the RF+LO impedance is essentially zero at BB frequencies,
2) the BB impedance, Z_b, is essentially zero at RF+LO frequencies, and
3) the network in what the diode is embedded has zero impedance at all frequencies except those in the vicinity of RF+LO and BB frequencies. This assumption does

not allow voltages other than those at the RF+LO and BB frequencies to exist across the diode.

With these assumptions, the equivalent circuits for the different signals in the diode are shown in Figs. B.4a and B.4b. In Fig. B.4a, Z_s and Z_p are series and parallel impedances that are dependent on the components in the reference and signal arms. In Fig. B.4b, Z_{sh} is a shunt impedance providing the DC return path and Z_a is the BB amplifier impedance.

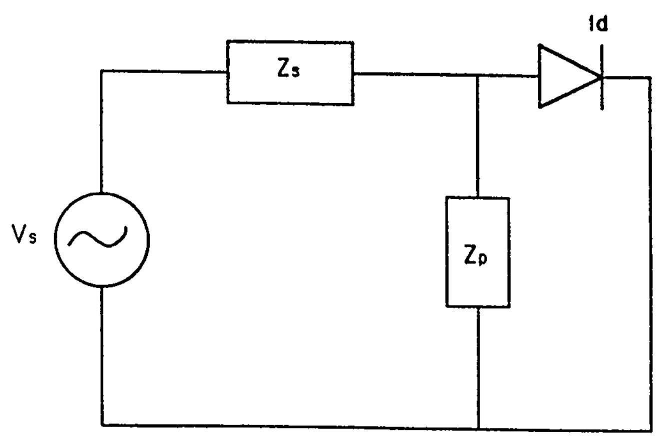

Figure B.4a
Equivalent circuit for RF and LO signals in the diode

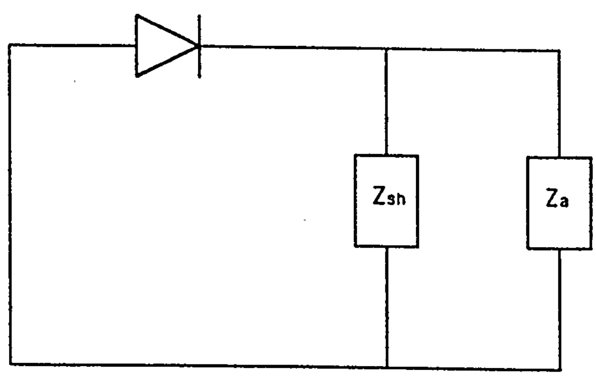

Figure B.4b
Equivalent circuit for base band signals

Expanding the exponential function in (B.12) in a Taylor series

$$i_d = I_s\left(\frac{v_d}{v_c} + \frac{v_d^2}{2v_c^2} + \frac{v_d^3}{6v_c^3} + \cdots\cdots\right) \qquad \text{(B.13)}$$

Collin [1992] highlighted an interesting fact; a number of published works on mixer diodes set $v_d = v_{LO}+v_{RF}$. This is incorrect since an ideal diode is a nonlinear element that introduces a relationship (B.12) between the voltage drop across it and the current passing through it. If the circuit that the diode is in can support voltages at the IF then it is more appropriate to set $v_d = v_{LO} + v_{RF} - v_{IF}$. In the present analysis,

$$v_d = v_s + v_r - v_b \tag{B.14}$$

where $v_r = v_{ro}\cos(\omega_o t+\phi_r)$, v_{ro} is the amplitude of the reference arm signal, ω_o is the klystron frequency, ϕ_r is the phase shift introduced into the reference signal, v_s is the signal arm signal

$$v_s = \sum_{k=0}^{\infty} v_{sk}\cos(\omega_o t + k\omega_m t + \phi_{sk}) \tag{B.15}$$

ω_m is the frequency of the signal driving the PIN modulator, v_{sk} are the Fourier coefficients of the square wave output from the PIN modulator, ϕ_{sk} are the phase shifts introduced into the signal at the frequency $\omega_o + k\omega_m$, and v_b is the BB voltage resulting from the mixing of RF and LO signals. Substitution of (B.14) into (B.13) yields

$$i_d = I_s\left[\frac{1}{v_c}(v_r + v_s - v_b) + \frac{1}{2v_c^2}\left(v_r^2 + v_s^2 + v_b^2 + 2v_r v_s - 2v_r v_b - 2v_s v_b\right)\right.$$
$$\left. + \frac{1}{6v_c^3}\left(v_r^3 + v_s^3 - v_b^3 + 3v_r^2 v_s + 3v_r v_s^2 - 3v_r^2 v_b + 3v_r v_b^2 - 3v_s^2 v_b + 3v_s v_b^2\right) + \cdots\cdots\right] \tag{B.16}$$

The even order terms $v_r^2,\ v_s^2,\ v_b^2,\ v_r^4,\ldots$ contribute a DC current

$$I_o = I_s\left(\frac{v_{ro}^2}{4v_c^2} + \sum_{k=0}^{\infty}\frac{v_{sk}^2}{4v_c^2} + \cdots\cdots\right) \tag{B.17}$$

This DC current was used to bias the diodes in the experiments performed.

The terms of (B.16) having appreciable components within the bandwidth of the BB amplifier are v_b, $v_r v_s$, and $v_r^2 v_b$ leading to a BB current

$$V_{bk} = \frac{\frac{1}{2v_c^2} V_r Z_b}{1 + \frac{I_s Z_b}{v_c} + \frac{I_s Z_b V_r^2}{4v_c^3}} V_{sk} \tag{B.18}$$

in phasor notation for a particular frequency $k\omega_m$, (B.18) becomes

$$I_{bk} = I_s\left(-\frac{V_{bk}}{v_c} + \frac{V_r V_{sk}}{2v_c^2} - \frac{v_{ro}^2 v_{bk}}{4v_c^2} \right) = \frac{V_{bk}}{Z_b} \tag{B.19}$$

Rearrangement of (B.19) leads to

$$V_{bk} = \frac{\frac{1}{2v_c^2} V_r Z_b}{1 + \frac{I_s Z_b}{v_c} + \frac{I_s Z_b V_r^2}{4v_c^3}} V_{sk} \tag{B.20}$$

which shows that, under the assumptions made, there is a linear relationship between the BB voltage at any particular frequency $k\omega_m$ and the signal voltage at the frequency $\omega_o + k\omega_m$.

B.8. EXPERIMENTAL PROCEDURE

The PIN modulator was used to square wave modulate the microwave power in the signal arm. This produces all the odd harmonics of its driving signal repetition frequency which enabled the simultaneous measurement of the SNR of all the odd harmonics of the fundamental component of the square wave. The bi-directional coupler was used to measure the voltage standing wave ratio (VSWR) of the diode and its mount. In this VSWR measurement, only the reflected power is affected by tuning the diode and hence it was only necessary to minimize the reflected power to tune the diode. The function of the VSWR arm was checked by first closing S-2. A 1 dB difference between the incident and the reflected power levels was observed. A 17.7 dB difference between these two power levels was observed when the diode mount was replaced by a 0.15 standard reflection (HP X916D). This agrees with the theoretical value: 0.15 reflection leads to -16.4 dB and there is a loss of 1 dB in the waveguide. An 11.4 dB difference between the two power levels was recorded when the diode mount was replaced by a 0.1 standard reflection (HP X916C) which was also in agreement with the theoretical value. It was found that the diode could also be matched to the microwave circuit by maximizing the DC current in the load resistor since both methods yielded the same optimum operating conditions.

The spectrum studied for each diode consisted of four frequency decades which were measured separately. Each decade was obtained by modulating the PIN modulator with a frequency near the low frequency end of that decade (i.e., 20 Hz, 200 Hz, 5 kHz and 50 kHz) and recording the FFT of the amplified output of the diode. The highest frequency decade was recorded using a Tektronix 7L13 spectrum analyzer in a Tektronix

7834 storage oscilloscope whereas the three lower decades were obtained using the FFT function on the DP6000 digital oscilloscope. The effect of the preamplifier and the low pass filter on the spectrum was determined by recording their swept frequency transfer functions using a Stanford Research lock-in amplifier (SR510). Figure B.5a shows the transfer function of the SR552 preamplifier with its shunting input inductance. Note the shape of the transfer function at frequencies below 100 Hz due to the high pass filter at the input of the amplifier. Figure B.5b shows the characteristic shape of the Butterworth low pass filter in one frequency decade. The shape of the transfer function of the Butterworth filter near the cut-off frequency sets an upper limit on the usable portion of each frequency range.

The following experimental procedure was used to study the microwave diodes:

1) The diode was placed in its mount,
2) The diode load was adjusted to the manufacturer's recommended value,
3) The reference arm signal power was set to the desired level using the attenuator (A1) and PM1,
4) The diode was tuned in the mount by minimizing the VSWR of the incident microwave power,
5) The frequency of the modulation signal generator was set to 5000 Hz (200 Hz) and the low pass antialiasing filter was set to 45 kHz (4.5 kHz),
6) The DP6000 was triggered manually and a frame of the amplified output of the diode was recorded, The DP6000 was programmed to record the fast Fourier transform (FFT) (real only, no windowing, forward FFT and magnitude coefficient) of each frame, and

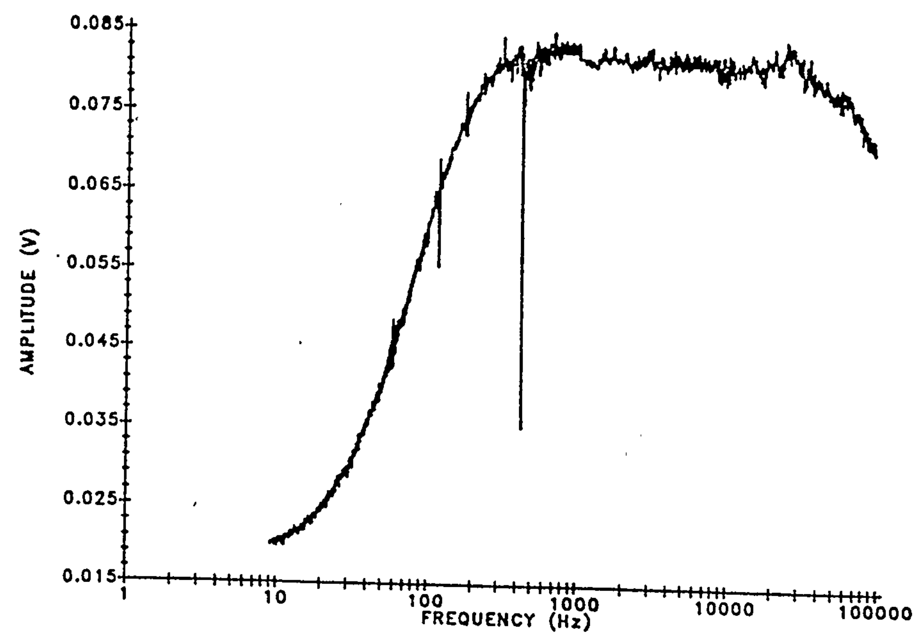

Figure B.5a The transfer function of the SR552 preamplifier with its shunting impedence.

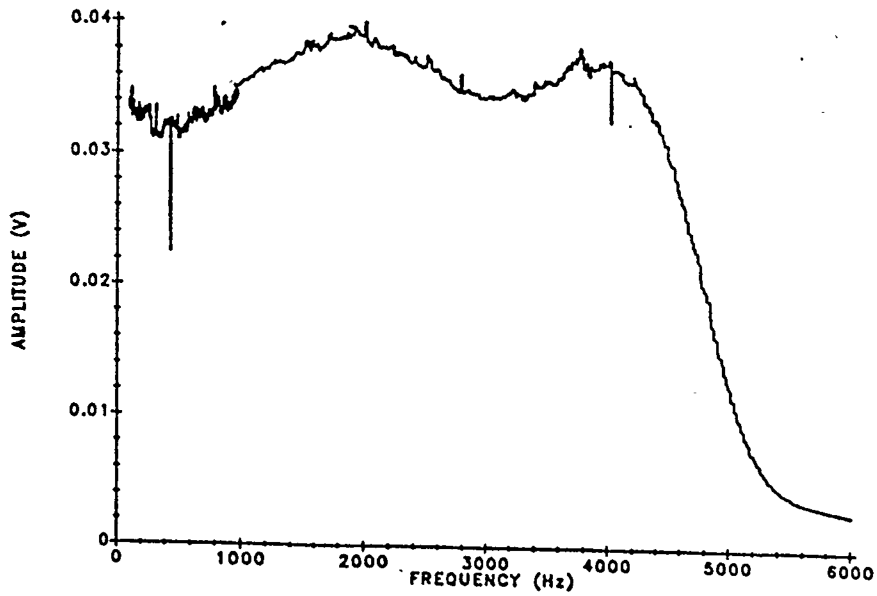

Figure B.5b The transfer function of the low pass Butterworth filter in the middle frequency band.

7) The measured spectrum of each frequency decade was normalized to a 1 Hz bandwidth and the three decades were merged.

The microwave diodes to be studied were screened. Two to seven of each model, depending on the available number of diodes, were chosen. The microwave power level, the diode load and the RF matching suggested by manufacturers were used for these initial measurements. These were made so as to assess the variability of diodes of the same model and to choose representative diodes. Another advantage for this set of runs was to determine matched pairs of diodes to be used later in the microwave balanced mixer of an EPR spectrometer. A 50 kHz modulation frequency was used. The amplified output of each diode was displayed using a Tektronix 7L13 spectrum analyzer and recorded on the DP6000 digital oscilloscope (all buffers except buffer 1 off, 2000 points, 1 ms period, external trigger using the 'positive gate out' signal from the Tektronix 7834, trigger level of 3V, filter out, input set to 5V and the signal into buffer 1 came from the 'vertical signal out' of the Tektronix 7834). A theoretical explanation of the resulting spectra may be found in the next section.

One typical microwave diode of each model was selected from this screening for further assessment. Typical here means that its SNR and noise floor were close to the measured mean SNR and noise floor for that model. The data sets were collected using a procedure similar to that described above. The microwave bias power level was varied and an optimal power level, at the recommended DC load, was determined. Bias power levels of 0.02, 0.05, 0.1, 0.2, 0.5, 1.0, 2.0, 5.0 mW were used. DC load values of 24, 51, 124, 244, 524, 1024 Ω were used. Recall that the highest frequency decade was obtained using the spectrum analyzer rather than the digital oscilloscope.

B.9. DATA ANALYSIS RESULTS

Figures B.5a-5c give the FFT [20 log(FFT of DP6000 input normalized to 1 Hz BW)] for selected a) point contact (-7 dBm microwave bias power and 122 Ω load resistance), b) Schottky (-7 dBm microwave bias power and 122 Ω load resistance) and c) tunnel (-7 dBm microwave bias power and 22 Ω load resistance) diodes. Each graph is the combination of two measurement decades. Note that no windowing was needed with the FFT since the square wave signal was sampled synchronously. The obvious peaks on the graphs are the amplitudes of the fundamental component and its harmonics. The difference between these peaks and the noise level gives the SNR. However, the amplitude of the peaks decrease as 1/f as expected for the FFT of a square wave and should not be misinterpreted as an increase in the insertion loss. The deviation from this 1/f relationship at the beginning of each decade is due to the introduction of a new fundamental frequency for each decade. However, minor variations from the expected 1/f could be attributed to variation of the insertion loss with frequency. The insertion loss is calculated by subtracting the amplitude of the largest peak from that of input local oscillator (reference
arm) power level.

The presence of the signal peaks made it difficult to analyze the measured spectra using a computer. Instead, all spectra were plotted and then analyzed manually. This analysis was performed, using weighted moving 5 point averages. The data was analyzed by

(i) ignoring the signal peaks,

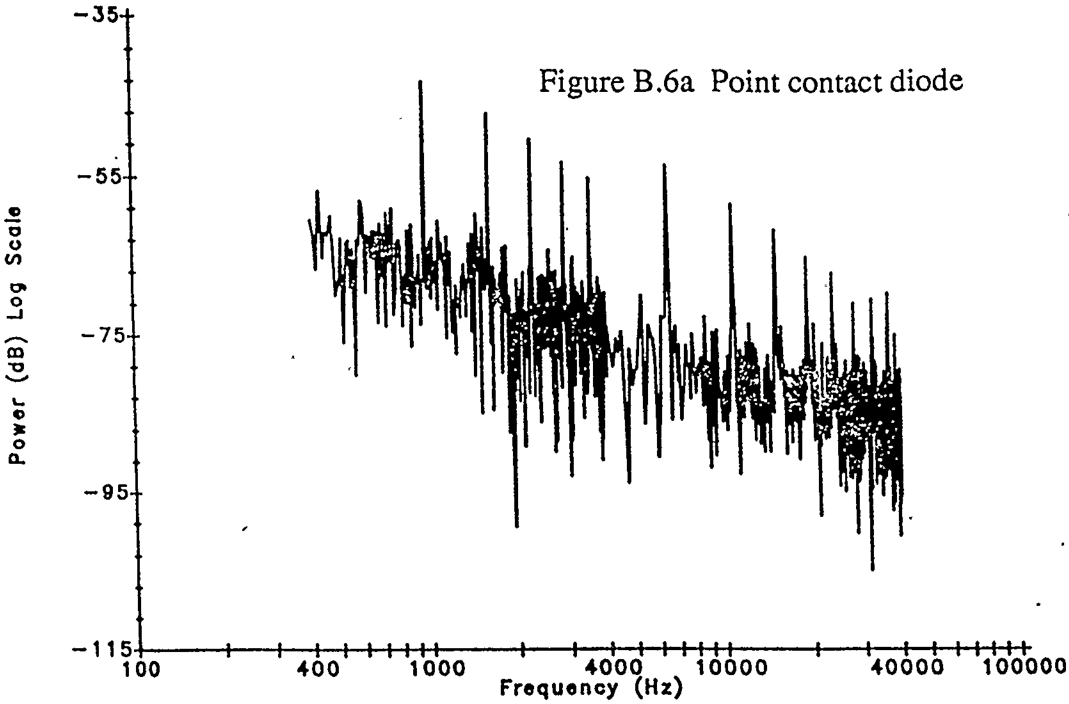

Figure B.6a Point contact diode

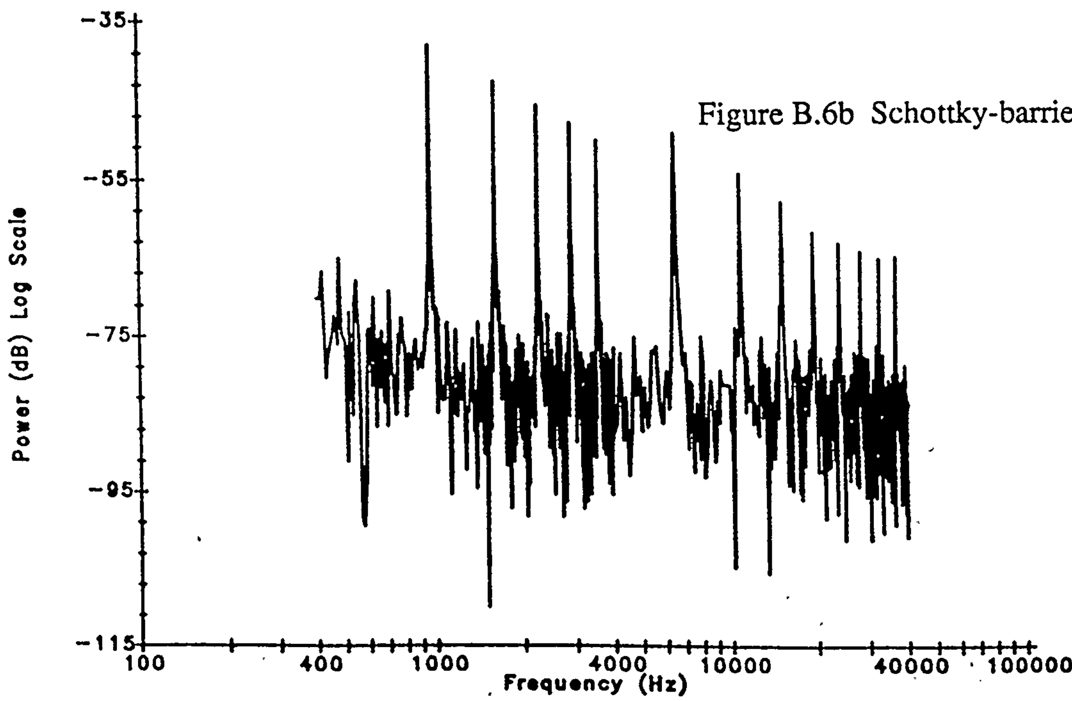

Figure B.6b Schottky-barrier diode

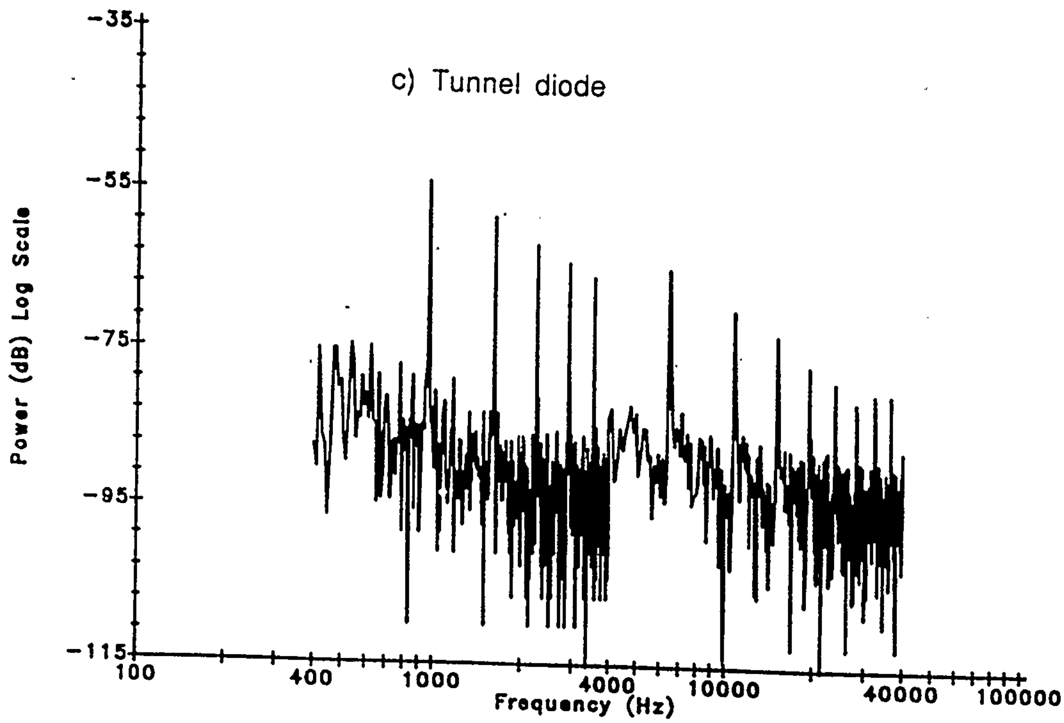

Figure B.6c The FFT of the amplified output of the microwave mixer diodes: a) point contact, b) Schottky, and c) tunnel.

(ii) finding the thermal noise level which is the high frequency noise level (does not depend on frequency),

(iii) whenever possible a straight line was fitted to the low frequency noise,

(iv) slope of this line was determined, and

(v) the frequency where this line intersects the thermal noise level was determined and was taken to be the "1/f" corner frequency.

Most noise spectra demonstrated strong curvature at low frequencies and it was not possible to obtain an accurate fit a single straight line to the data. In such cases two lines with different slopes where fitted to different segments of the low frequency data and the intersection of the line fitting the higher frequency range with the thermal noise level was taken as the "1/f" corner frequency. This is shown on Fig. B.8a.

Analysis showed that the "1/f" corner frequency for the point contact diode is above the maximum measurement frequency of 45 kHz in agreement with the measurements of Buckmaster and Dering [1965] and Jaggard [KING 1978]. The "1/f" corner frequency for the Schottky and the tunnel diodes was approximately 4 kHz in agreement with the measurements of Jaggard [KING 1978].

Anderson and van der Ziel [1952], Fonger et al. [1958] and Watkins [1959] have reported that, under some conditions the, flicker ("1/f") noise could be zero at a particular current in a junction diode. This observation has been confirmed in this work. For example, the PFM-422 and the MA-40075 diodes were found to exhibit no "1/f" noise at microwave bias power levels of -10 dBm and -13 dBm respectively. King [9] suggested that the flicker noise in tunnel diodes follows a $1/f^{\alpha}$ law where $1.35 < \alpha < 1.5$

rather than $\alpha = 1$. This work suggests that $\alpha = 1$ is an approximation and, in fact, α can vary between 0 and 2 depending on the microwave bias power level. Figures B.7a-7g are the FFT of the output of the MA40075 diode at different power levels. Figures B.8a-8g are the FFT of the output of the 1N23WE diode at different power levels. It is easy to see that there is no "1/f" noise at a microwave bias power level of -13 dBm, that $\alpha = 2$ for 7 dBm and that for intermediate microwave bias power levels, $0 < \alpha < 2$. Inspection of these and other spectra showed that spectral power density always has a $1/f^{\alpha}$ frequency dependency and that at the higher microwave bias power levels, $\alpha = 2$ at the low frequencies, $0 < \alpha < 2$ at transition frequencies and $\alpha = 0$ at high frequencies. The frequencies at the boundaries between these ranges are a decreasing function of the microwave bias power level. These results are in agreement with the observations of the above mentioned researches.

Table B.1 gives the thermal noise floor of each microwave diode at different reference arm power levels. Table B.2 gives the second corner frequency between the transition range and the thermal noise range for each microwave diode at different reference arm power levels. It was found that changing the DC load for the microwave diodes had a negligible effect on the output of the diodes. This was the case for the values of load resistances used in these experiments. This result is rather surprising, particularly for tunnel diodes, since they have always been assumed to perform optimally with minimum load resistance.

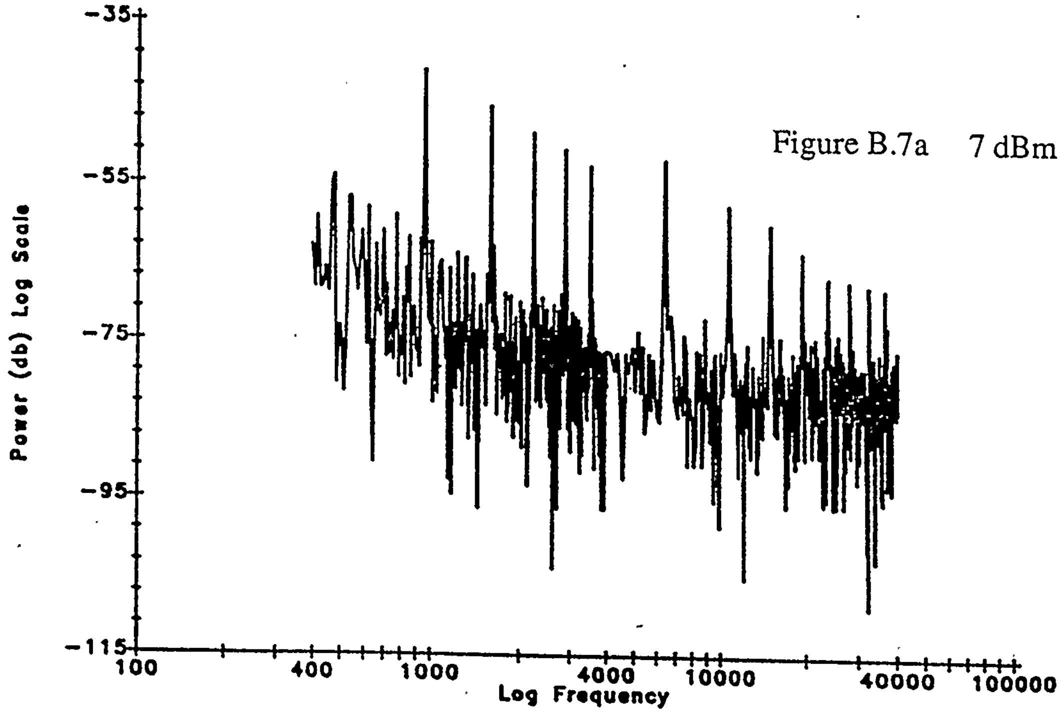

Figure B.7a 7 dBm

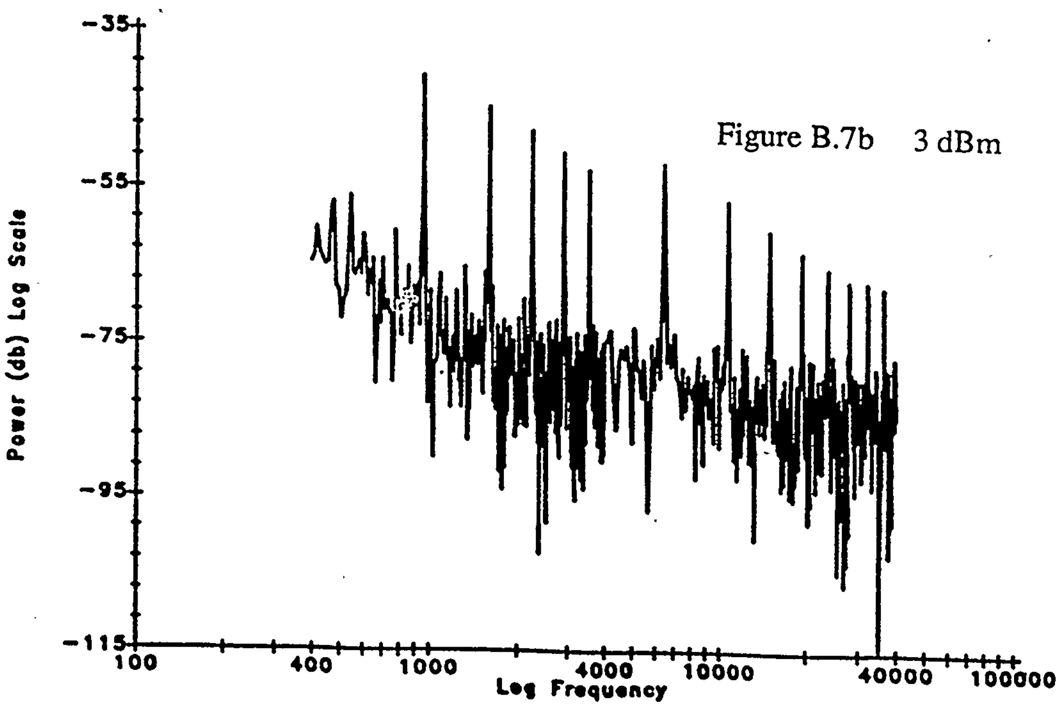

Figure B.7b 3 dBm

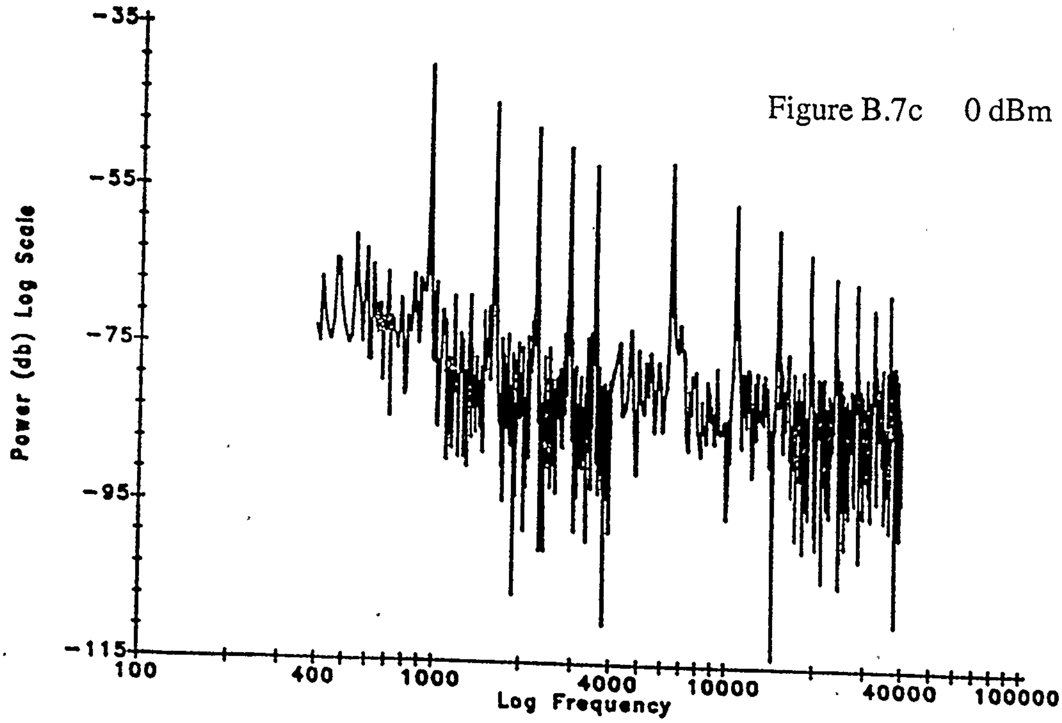

Figure B.7c 0 dBm

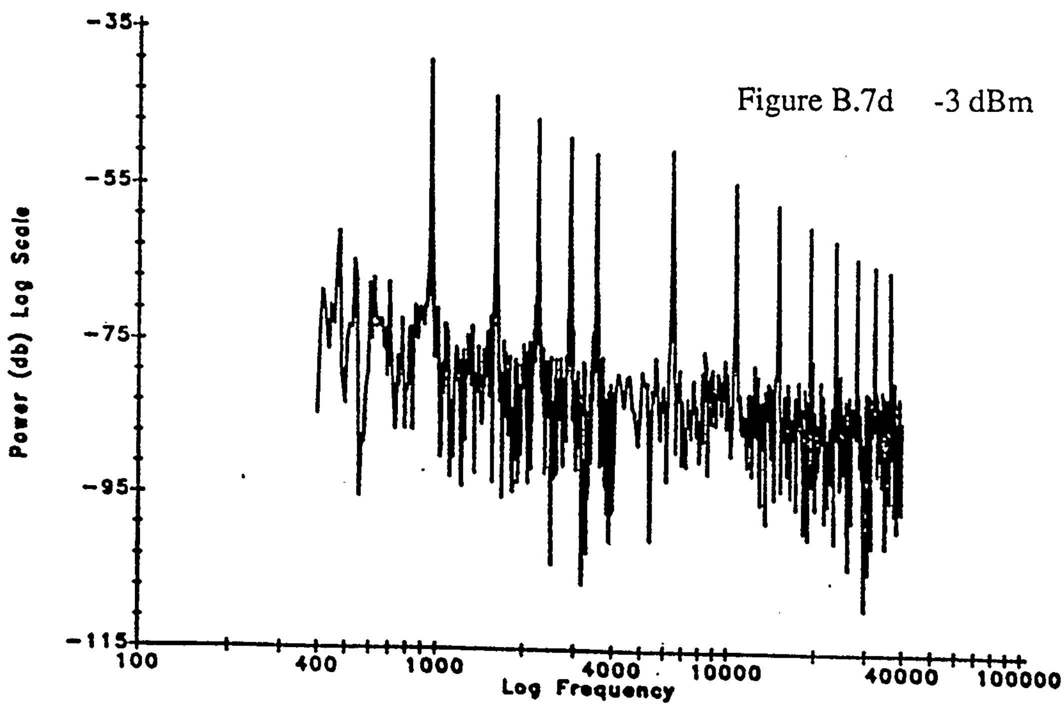

Figure B.7d -3 dBm

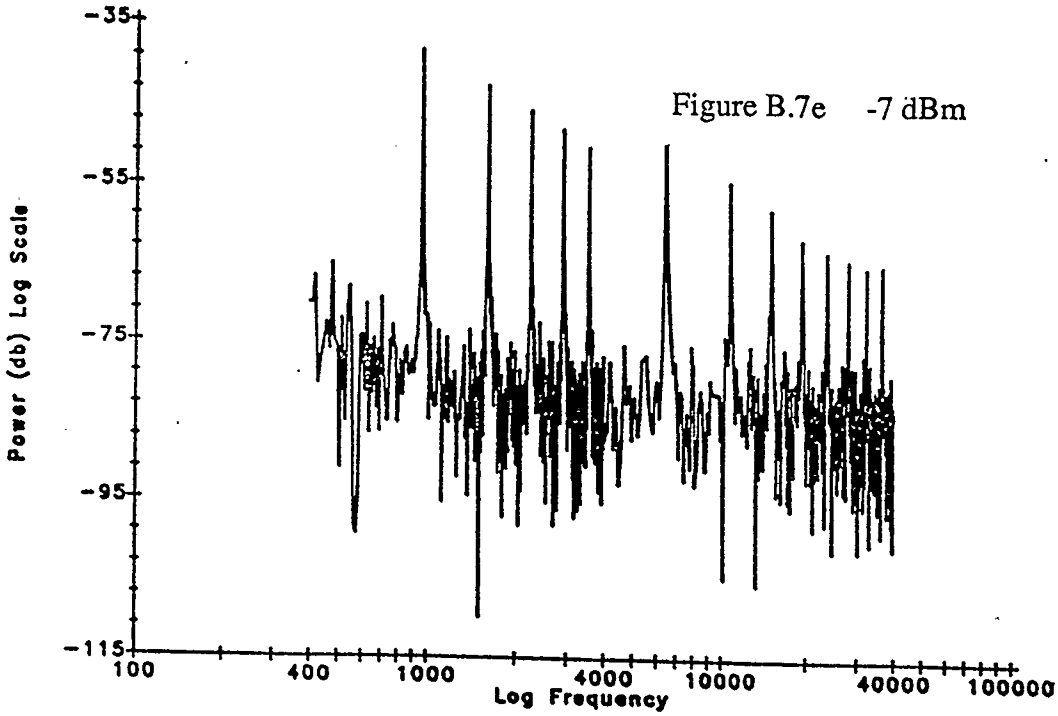
Figure B.7e -7 dBm
-35
-55
-75
-95
-115
Power (db) Log Scale
100
400
1000
4000
10000
40000
100000
Log Frequency

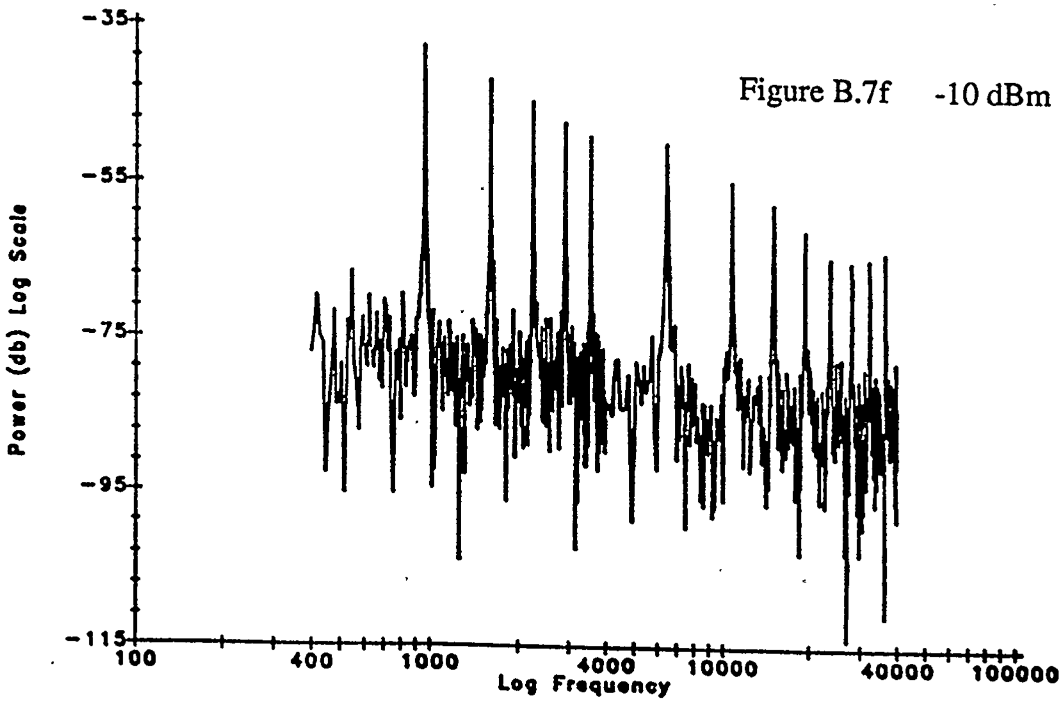
Figure B.7f -10 dBm
-35
-55
-75
-95
-115
Power (db) Log Scale
100
400
1000
4000
10000
40000
100000
Log Frequency

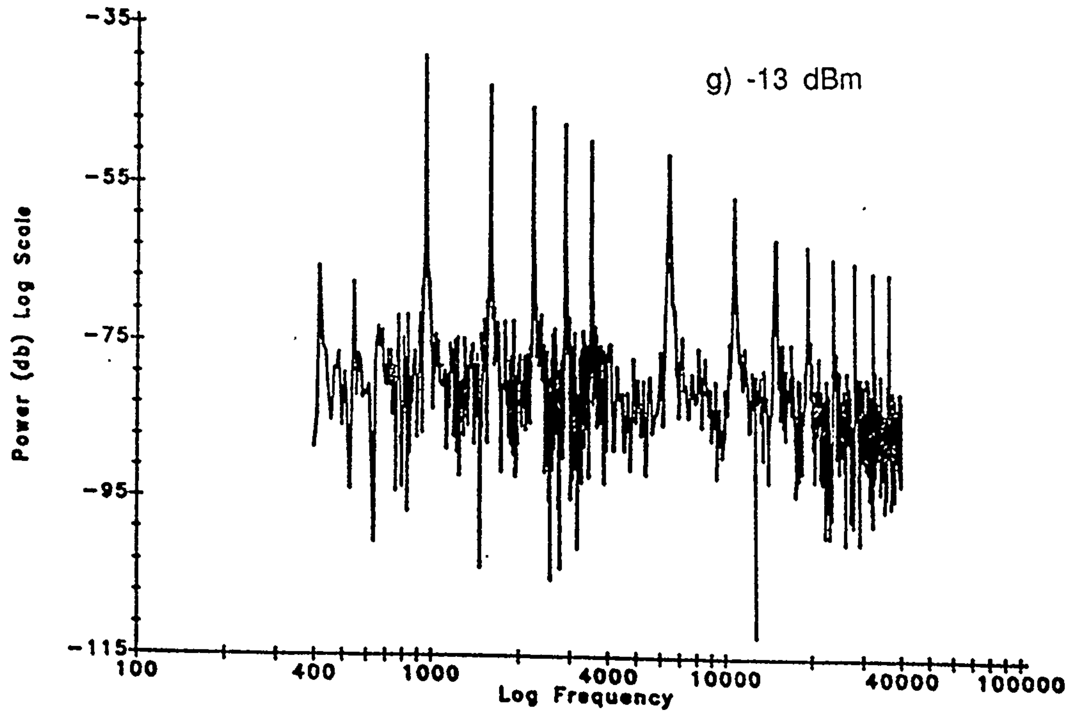

Figure B.7g The FFT of the amplified output of the MA40075 diode at different microwave bias power levels: a) 7 dBm, b) 3 dBm, c) 0 dBm, d) -3 dBm e) -7 dBm, f) -10 dBm and g) -13 dBm.

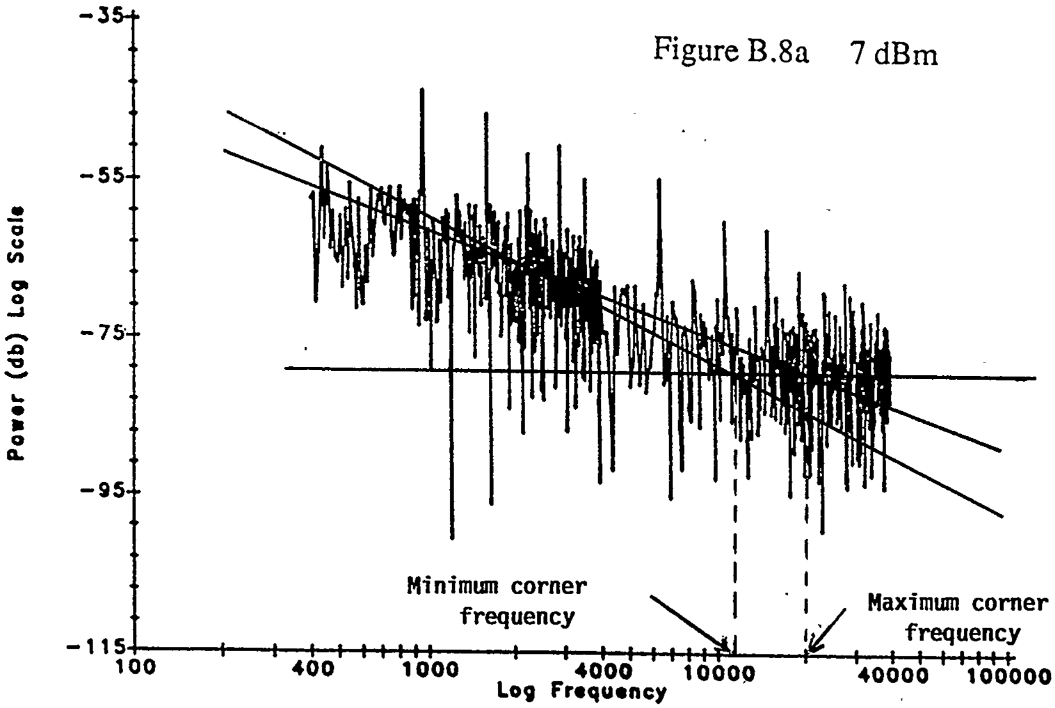
Figure B.8a 7 dBm
Power (db) Log Scale
-35
-55
-75
-95
-115
100
400
1000
4000
10000
40000
100000
Log Frequency
Minimum corner frequency
Maximum corner frequency

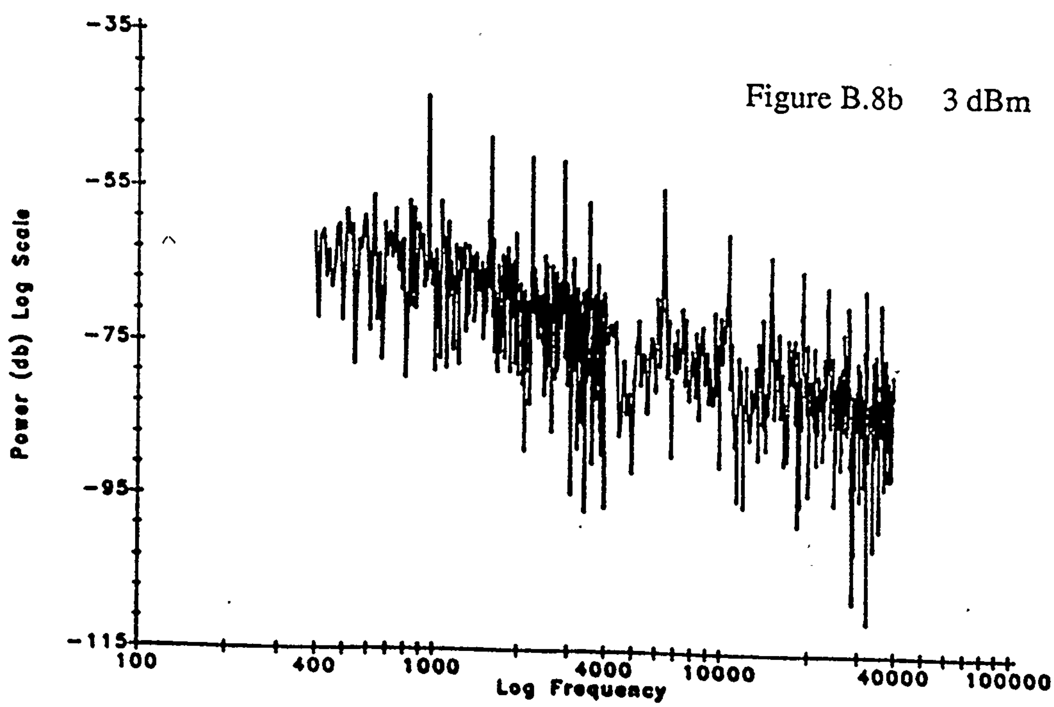
Figure B.8b 3 dBm
Power (db) Log Scale
-35
-55
-75
-95
-115
100
400
1000
4000
10000
40000
100000
Log Frequency

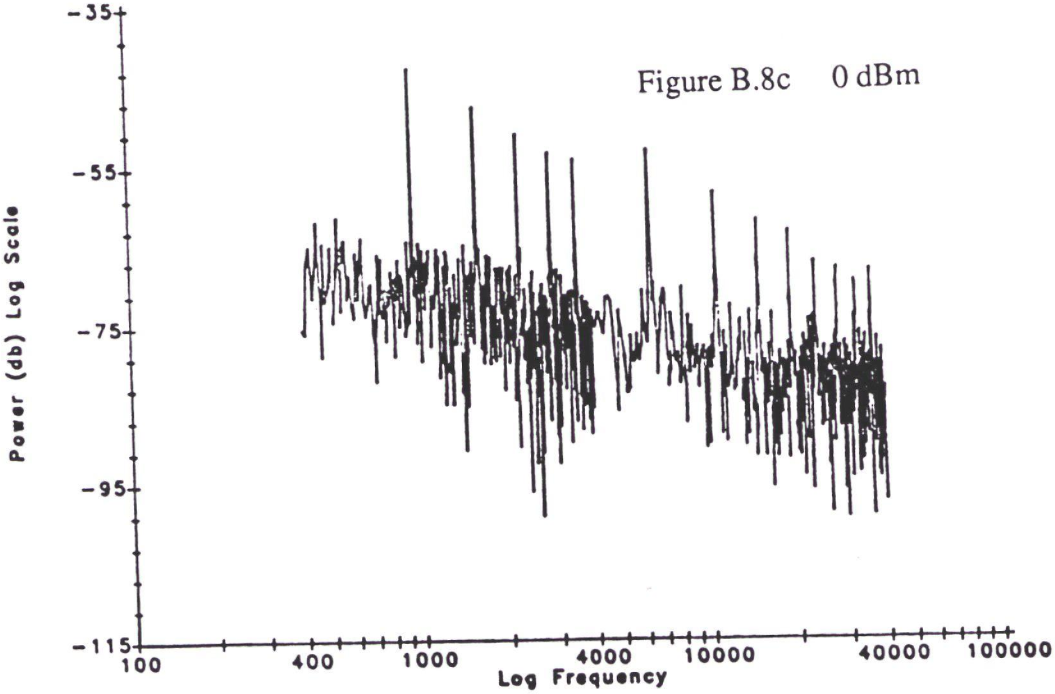
Figure B.8c 0 dBm
Power (db) Log Scale
-35
-55
-75
-95
-115
100
400
1000
4000
10000
40000
100000
Log Frequency

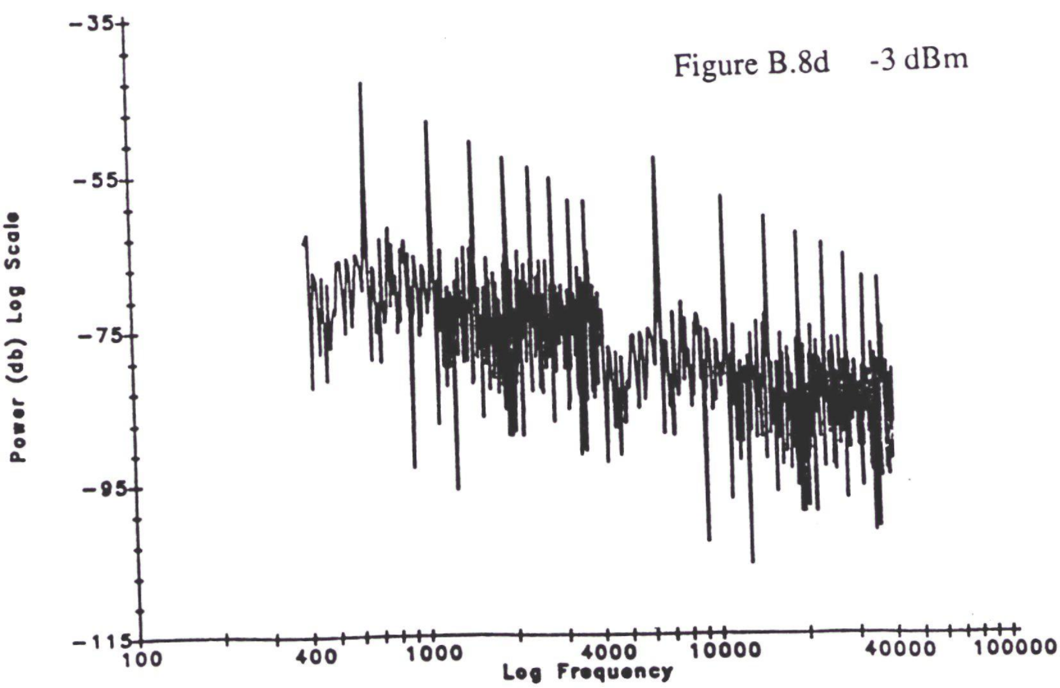
Figure B.8d -3 dBm
Power (db) Log Scale
-35
-55
-75
-95
-115
100
400
1000
4000
10000
40000
100000
Log Frequency

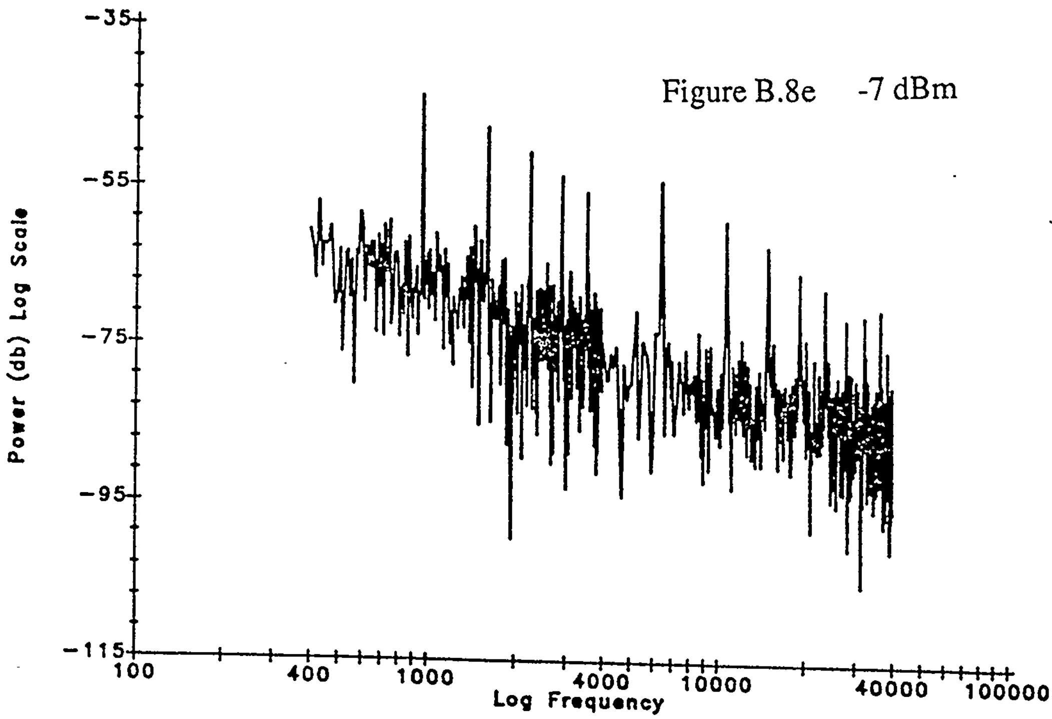
Figure B.8e -7 dBm
-35
-55
-75
-95
-115
Power (db) Log Scale
100
400
1000
4000
10000
40000
100000
Log Frequency

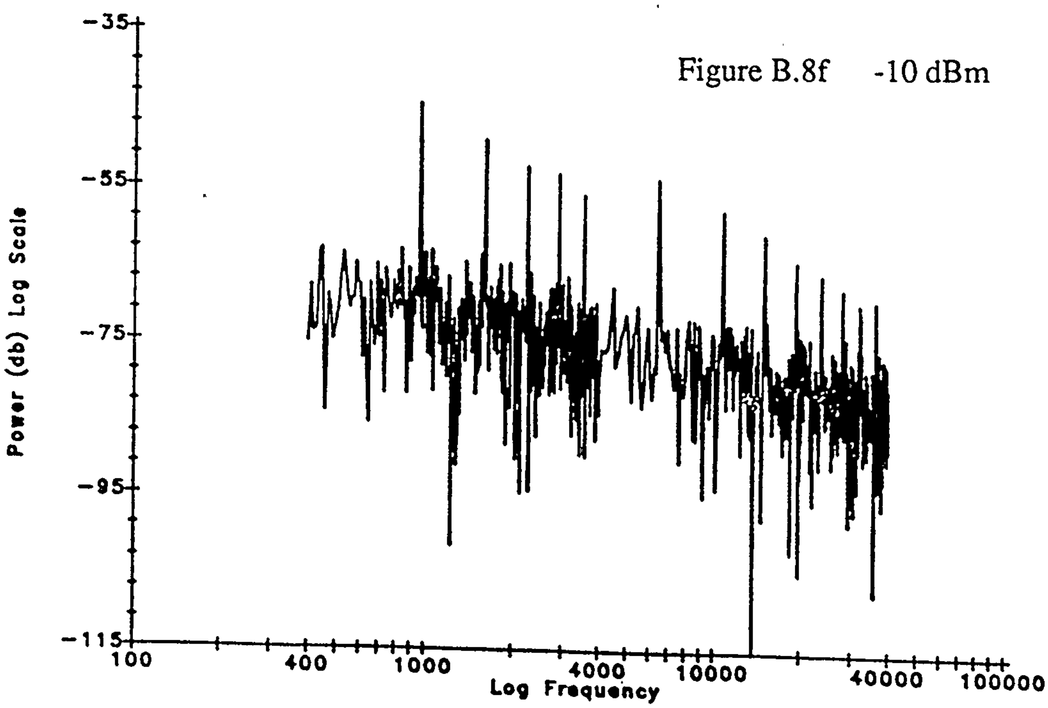
Figure B.8f -10 dBm
-35
-55
-75
-95
-115
Power (db) Log Scale
100
400
1000
4000
10000
40000
100000
Log Frequency

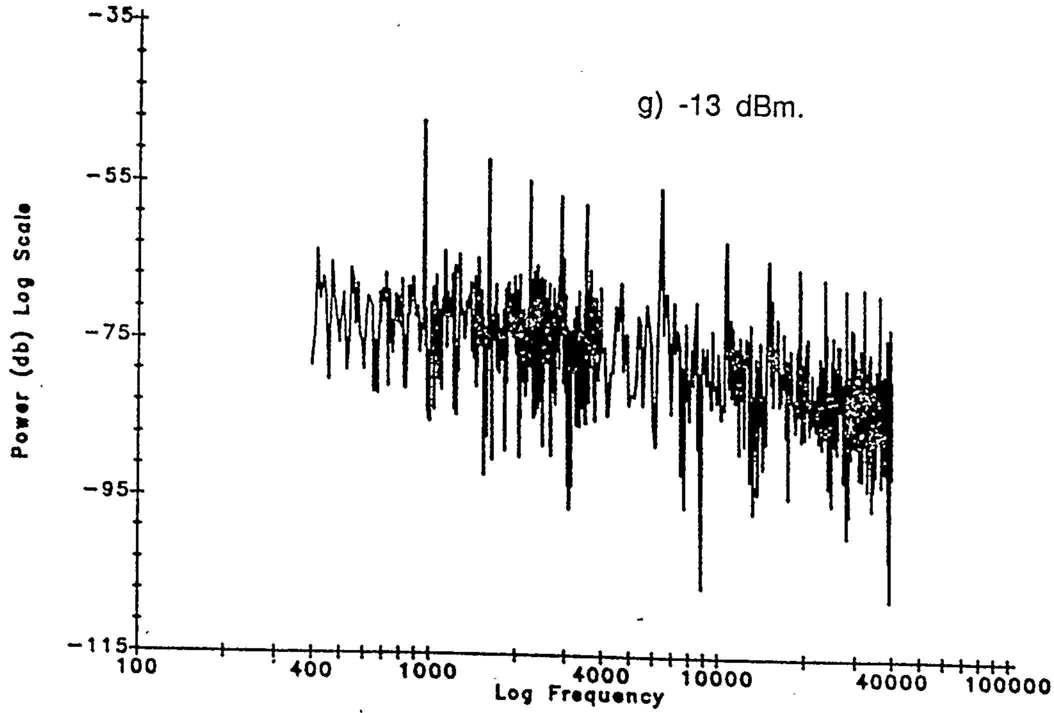

Figure B.8g The FFT of the amplified output of the 1N23WE diode at different microwave bias power levels: a) 7 dBm, b) 3 dBm, c) 0 dBm, d) -3 dBm e) -7 dBm, f) -10 dBm, and g) -13 dBm.

B.10. CONCLUSIONS

This appendix described a simple method for the evaluation of the various types of microwave diodes which can be used in EPR spectrometers. The advantages of this method are:

1) it ensures that the optimum operating conditions for each diode are determined,
2) it allows a wide frequency band to be measured simultaneously, and
3) it generates square wave amplitude modulation thus enabling the conversion gain of the diode to be observed at more than one frequency simultaneously.

This method has been used to study a selection of microwave diodes. It has been found that:

1) the noise spectrum of microwave diodes can be divided into three ranges: $1/f^2$ range, a transition range where the spectral power density follows a $1/f^\alpha$ law where $0 < \alpha < 2$ depending on the microwave bias power level and a thermal range where the spectral density is constant over a wide range of frequency,
2) the "1/f" corner frequency increases with the microwave bias power level, and
3) the DC load has almost no effect on the noise performance of the diodes.

It is suggested that, for future work, measurements of the noise level reported here are performed without using any modulation on the microwave power and to lower frequencies. It is believed that such measurements will reveal that the flicker noise has a $1/f^\alpha$ dependency on the frequency and that α changes with frequency and with microwave bias power level.

Table B.1 Table giving the thermal and shot noise floor (± 3dB) of selected microwave diodes at different microwave bias power levels

Diode	Diode	Power Level (dBm)						
Model	Type	-13	-10	-7	-3	0	3	7
1N23WE	Point	<-80	-83	-85	-84	-82	-81	-79
MA4130A	Point	<-85	-82	-80	<-80	<-76	-76	
MA40050-3	Schottky			-84	-85	-84	-84	-82
MA40071E	Schottky	-80	-85	-81	-85	-82	-82	-83
MA40090	Schottky	-85	-86	-88	-84	-87	-81	-83
MA40194-3	Schottky	-90	-90	-90	-90	-87	-81	-83
HPA2627	Schottky			-88	<-88	-87	-85	-80
PKD100M-42	Schottky	-85	-88	-92	-88	-86	-84	-84
27100bd-UOC	Tunnel	-95	-94		-99	-96	-96	
PHILCOL4164	Tunnel	-95	-92		-97	-90	-94	
PFM422	GaAs				-80	-81	-80	-80

Table B.2 Table giving the maximum value of the corner frequency (kHz) for selected microwave diodes at different microwave bias power levels.

Diode	Diode	Power Level (dBm)						
Model	Type	-13	-10	-7	-3	0	3	7
1N23WE	Point	>40	32	24	23	22	20	20
MA4130A	Point	>40	>40	15	10	>40	20	15
MA40050-3	Schottky			7.5	8	9	10	7
MA40071E	Schottky	>40	>40	15	12	7.5	5	5.5
MA40090	Schottky	5	8	20	10	14	10	12
MA40194-3	Schottky	1.8	1.8	5.5	5.5	15	35	38
HPA2627	Schottky	30		30	>40	30	>30	30
PKD100M-42	Schottky	4	3.5	6.5	4	4.2	4.2	3.3
27100bd-UOC	Tunnel	1.3	2.2		5.5	4.6	3.5	
PHILCOL4164	Tunnel	3.3	2.5		5.5	4.6	3.5	
PFM422	GaAs				6	4	3.5	3.5

Table B.3 Table giving the minimum value of the corner frequency (kHz) for selected microwave diodes at different microwave bias power levels.

Diode	Diode	Power Level (dBm)						
Model	Type	-13	-10	-7	-3	0	3	7
1N23WE	Point	>40	32	17	14	10	15	12
MA4130A	Point	>40	>40	2.6	4.5	>40	20	2.8
MA40050-3	Schottky			2.3			1.9	
MA40071E	Schottky	>40	>40	2.5	8.5	3.9	4	4.5
MA40090	Schottky	1.5	2.7	3.9	3.5			
MA40194-3	Schottky	1.2		2	2.3	2.7	10	
HPA2627	Schottky	30		30	>40	30	>30	30
PKD100M-42	Schottky							
27100bd-UOC	Tunnel					2.2	3.3	
PHILCOL4164	Tunnel		1.5	3	1.7	2.5	2.5	1.8
PFM422	GaAs							

REFERENCES

Anderson, R.L., and van der Ziel, A., 1952, IRE Trans. Electron Devices, **ED-1**, 20.

Beefhorth, T.H., and Goldsmid, H.J., 1970, Physics of Solid State Devices, (Pion, London).

Brewster, R.L., 1987, Telecommunications Technology, (Ellis Horwood, Chichester, U.K.).

Buckmaster, H.A., and Dering, J.C., 1965, Can. J. Phys., **43**, 1088.

Buckmaster, H.A., and Rathie, R.S., 1971, Can. J. Phys., **49**, I. 849, II. 853,

Buckmaster, H.A., and Cohn-Sfetcu, S., 1973, in Magnetic Resonance and Related Phenomena, V.Hovi, Ed. Proc. XVII Colloque Ampere, Turku, Finland, I. 276, II. 280, (North Holland, Amsterdam).

Cohn-Sfetcu, S., and Buckmaster, H.A., 1974, IEEE Trans. Instrum. and Meas., **IM-23**, 102.

Collin, R.E., 1992, Foundations for Microwave Engineering, 2nd Ed., (McGraw-Hill, New York), Chapter 12.

Fonger, W.H., Loferski, J.J., and Rappaport, P., 1958, J. Appl. Phys., **29**, 558.

Ginzton, E.L., 1957, Microwave Measurements, (McGraw Hill, New York).

Gorelkinskii, Y.V., and Kim, A.A., 1985, Izv. Akad. Nauk. Kaz. SSR, Ser. Fiz.-Mat., **24** (Chem. Abst. **103**:114780e).

Karthe, W., and Wehrsdorfer, E., 1979, J. Magn. Reso., **33**, 107.

King, R.J., 1978, Microwave Homodyne Systems, (Peter Peregrinus, Stevenage).

Mazda, F.F., 1981, Discrete Electronic Components, (Cambridge, U.P. New York).

Ondria, J.G., 1968, IEEE Trans Microwave Theory and Technique, **MTT-16**, 767.

Van der Ziel, A., 1979, Advances Electronics And Electron Phys., **47**, 225.

Von Hippel, A.R., 1954, Dielectrics and Waves, (MIT Press, Cambridge).

Watkins, T.B., 1959, Proc. Phys. Soc. (Lond), pt. 1, **73**, 59.

Zaghloul, H., van Kalleveen, T.H.T., Hansen, C.H., and Buckmaster, H.A., 1990a, Proc. of the IEEE Instrumentation and Measurement Technologies Conference, IMTC-90, Emerging Measurement Techniques IM90-3A1, pp.104, San Jose, CA, February 13-15, 1990.

Zaghloul, H., van Kalleveen, T.H.T., Hansen, C.H., and Buckmaster, H.A., 1990b, IEEE Trans. Instrum. and Meas., **IM-39**, 928.

APPENDIX C:

HAIR COLOUR DETERMINATION USING EPR SPECTROSCOPY

ABSTRACT

The characterization of intact human hair colour (pigmentation) using CW-EPR is described. It is found that the g-factor classifies this hair into three rather than two different groups. The g-factors of eumelanin in black hair is 2.0036(1), phaeomelanin in yellow hair 2.0043(1) and trichsiderin in red hair 2.0046(1).

C.1. INTRODUCTION

Dr. Buckmaster and a co-worker, Dr. T.H.T. van Kalleveen, drew the author's attention the use of EPR in studying the metal ion/wool systems [KOKOT 1972]. A subsequent literature search revealed that EPR has been used to study and characterize human hair. This research left a number of questions unanswered: is there a difference in the EPR spectra of curly and straight hair?, red and yellow hair? and light blonde and white hair? In attempting to answer these questions, the present appendix resulted.

This appendix details the work of Zaghloul et al. [1989] drawing attention to an interesting result concerning the characterization of human hair colour (pigmentation) using CW-EPR spectroscopy that appears to have escaped notice in the extensive literature concerning natural and synthetic melanins. Except for a few early exploratory EPR studies of various biological materials including intact human hair [COMMONER ET AL. 1954, EDWARDS ET AL. 1959 and 1961, MASON ET AL, 1960, BLOIS ET AL. 1964], all studies have concentrated upon the identification of the free radical present in eumelanin and phaeomelanin which were considered responsible for the pigmentation in black and "red" hair respectively. Consequently, these studies have focused on the photochemical [CHEKEDAL ET. AL. 1977 and 1978, SEALY ET AL. 1980] and chemical properties [SEALY ET AL. 1980, 1982a and 1982b] of natural and synthetic melanins. Mammalian forms of eumelanin and phaeomelanin are interrelated and both are amorphous, heterogeneous polymers [NICOLAUS 1982]. There does not appear to be any reference in the EPR literature to measurements of trichsiderin which is a third human pigmentation that has been associated with red ("auburn") hair [SWIFT 1977]. It would appear that most EPR workers have failed to distinguish between the two types of "red" hair: yellowish red due to phaeomelanin [SWIFT 1977] and auburn red due to trichsiderin [SWIFT 1977].

C.2. EXPERIMENTAL PROCEDURE

This appendix describes 9 GHz CW-EPR measurements performed on thirty eight intact hair samples. The samples were various shades of black, yellow and red. These were the natural colours of the samples; none of these samples were dyed or treated in any way that alters the hair colour. The samples were prepared by placing 0.200(5) gm of hair into a Norell XR-55 pyrex sample tube of 0.4 cm inner diameter and 17.5 cm length. The hair was packed at the bottom of each tube. A glass rod of 13.0 cm length was used to push the hair towards the bottom of the tube to ensure similar packing of all samples. The tubes were then inserted in a TE_{102} rectangular sample cavity of a non-commercial 9 GHz synchrodyne CW-EPR spectrometer. All spectra were recorded as the first Fourier absorption coefficient, a_1, of the absorption, a_0, under identical spectrometer settings. The spectra were obtained using an SR510 lock-in amplifier and a modified version of its data acquisition software. The spectra were analyzed on a SUN-3 workstation to obtain the EPR spectral parameters. The parameters used to characterize the different samples were the g-factor and the relative spin concentration. The latter quantity is obtained by double numerical integration of a_1 and is a measure of the number of unpaired spins in the sample.

C.3. RESULTS

Figure C.1 shows typical 9 GHz CW-EPR spectra for a red (solid) and a black (dashed) hair sample. This figure illustrates the different g-factors for these samples. Figure C.2 is a histogram of the g-factor values that were measured. It can be seen that the g-factors are grouped around three distinct values: 2.0036(1), 2.0043(1) and 2.0046(1). Visual inspection of the hair colour showed that different shades of black hair were associated with the peak centered at 2.0036(1), yellow hair with the peak centered at 2.0043 and red hair with the peak centered at 2.0046. The g-factor of black hair is in agreement with the value of 2.0037(1) reported by Blois et. al (5) and of 2.003(1) reported by Mason et. al. [1960] and others [SEALY ET AL. 1980]. Based on Swift's [1977] association between pigmentation and hair colour, it is reasonable to assume that the reported g-factors are characteristic of eumelanin, phaeomelanin and trichsiderin, respectively. The relative spin concentration was correlated with the shade of the hair colour in agreement with the measurements of Mason et. al. [1960]. One white hair sample produced an EPR spectrum with an SNR of approximately 2 whereas dark black hair produced signals of SNR exceeding 1000. This was confirmed by bleaching two black samples with hydrogen peroxide. This bleaching reduced the relative spin concentration by factors of ten and thirty without altering the g-factor.

It should be observed that the fifteen yellow and red intact hair samples used in this study could not be distinguished visually before the EPR measurements had been performed. However, it was obvious after these measurements that these samples did belong to two different colour groups: yellow (including yellowish red) and red. This remark, combined with the small difference in the g-factors of the two groups, provides a

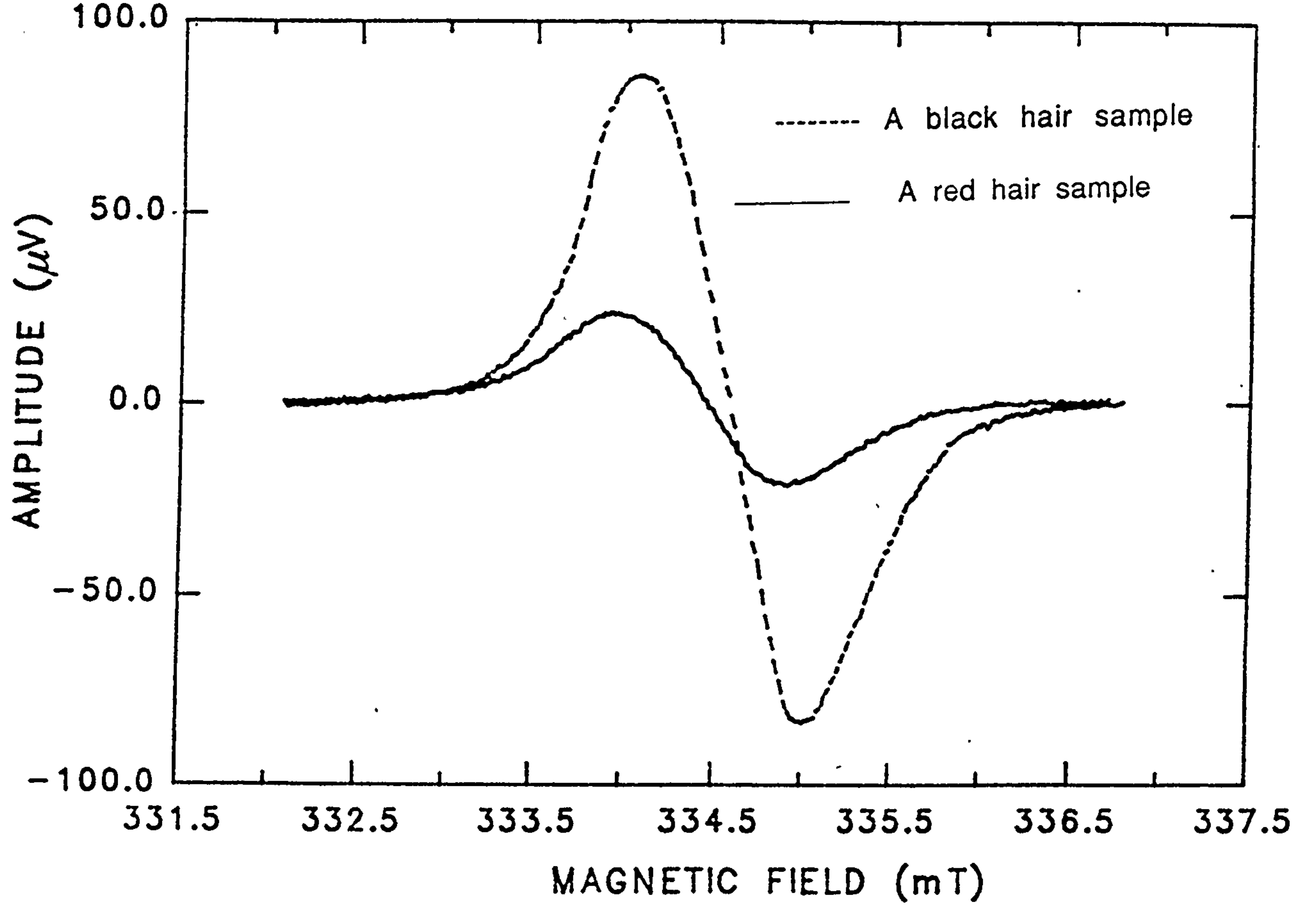

Figure C.1 Typical 9GHz CW-EPR spectra for a red (solid) and a black (dashed) hair sample.

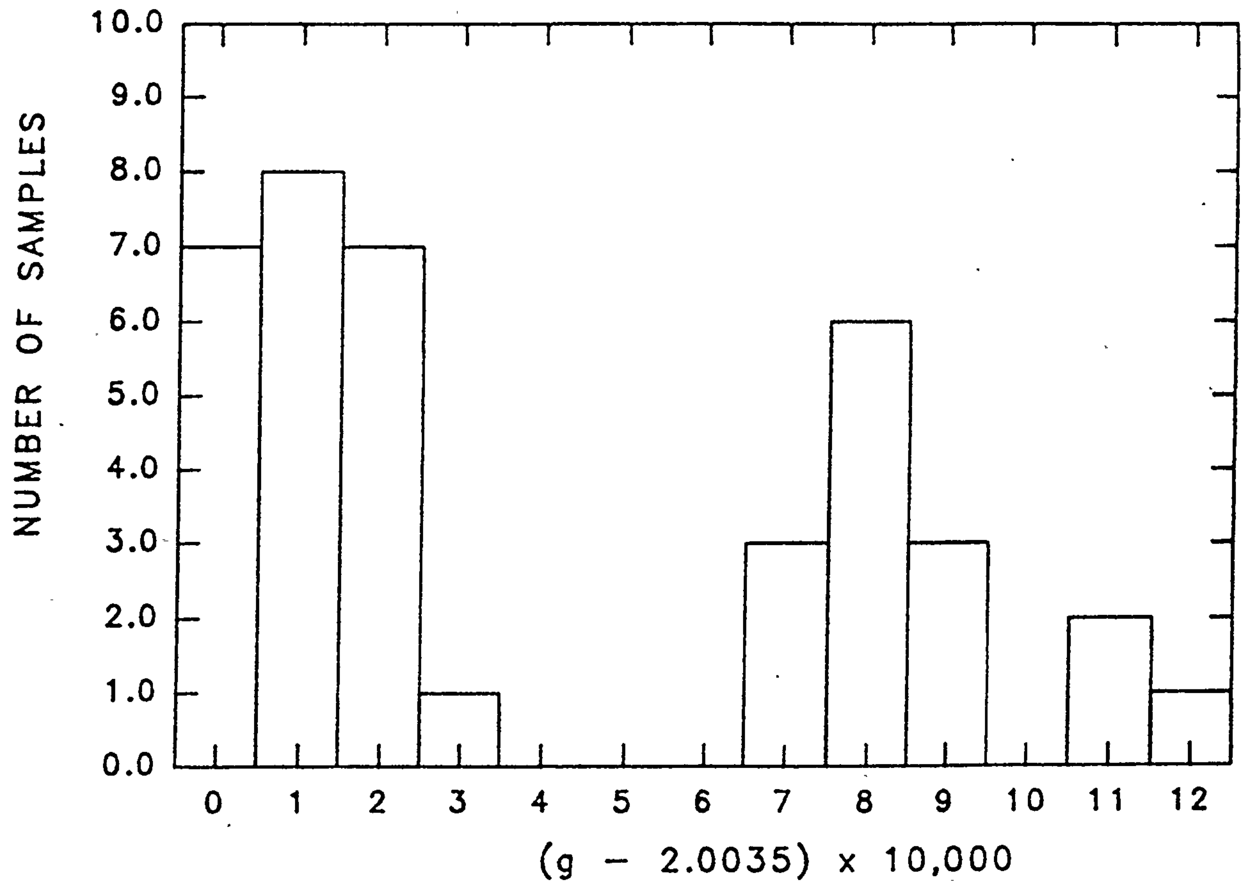

Figure C.2 A histogram of the measured g-factor values.

probable explanation for the failure of earlier workers to observe the results reported in this appendix.

If the method described in this appendix is to be used by different researchers who wish to share their results, it is suggested that the above experiments be performed with powdered hair samples in order to reduce the variability of the measured g-factors. This will ensure that all samples are identically packed. The number of samples should be increased to establish an improved statistical correlation between the EPR spectral line parameters (g-factor, relative spin concentration and the linewidth) and the hair colour.

C.4. CONCLUSIONS

This appendix described characterization of intact human hair colour (pigmentation) using CW-EPR. It is found that the g-factor classifies this hair into three rather than two different groups. The g-factors of eumelanin in black hair is 2.0036(1), phaeomelanin in yellow hair 2.0043(1) and trichsiderin in red hair 2.0046(1).

REFERENCES

Blois, M.S., Zahlan, A.B., and Maling, J.E., 1964, Biophys. J., **4**, 471.

Chedekel, M.R., Post, P.W., Deibel, R.M., and Kalus, M., 1977, Photochem, Photobio., **26**, 651.

Chedekel, M.R., Smith, S.K., Post, P.W., Pokora, A, and Vessel, D.L., 1978, Proc. Natl. Acad. Sci. USA, **75**, 5395.

Commoner, B., Townsend, J., and Pake, G., 1954, Nature, **174**, 689.

Edwards, M.1., Kerkut, G.A., Munday, K.A., Ingram, D.J.E., Saxena, M.C., and Leach, K., 1959, Naure, **184**, 1402.

Edwards, M.L., Kerkut, G.A.., Munday, K.A., and Leach, K., 1961, Nature, **190**, 1005.

Kokot, S., 1972, "An Investigation of Transition Metal Ions in Keratin", Ph.D. thesis, School of Textile and Technology, University of New South Wales, Australia.

Mason, H.S., Ingram, D.J.E., and Allen, B.J., 1960, Arch. Biochem. Biophys., **86**, 225.

Nicolaus, R., 1982, Melanin, (Herman, Paris).

Sealy, R.C., Felix, C.C., Hyde, J.S., and Swartz, H.M., 1980, in Vol. 4 of Free Radicals in Biology, Pryor, W.A., Ed., pp 209, (Academic Press, New York).

Sealy, R.C., Hyde, J.S., Felix, C.C., Menon, I.A., and Prota, G., 1982a, Science, **217**, 545.

Sealy, R.C., Hyde, J.S., Felix, C.C., Menon, I.A., Prota, G., Swartz, H.M., Persad, S. and Haberman, H.F., 1982b, Proc. Natl. Acad. Sci. USA, **79**, 2885.

Swift, J.A., 1977, "The Histology of Keratin Fibers", in Chemistry of Natural Protein Fibres, Asquith, R.S., Ed. pp. 81-146, (Plenum, New York).

Zaghloul, H., van Kalleveen, T.H.T., and Buckmaster, H.A., 1989, Abstract No. PE 29, CAP Congress, Guelph, ON, June 26-28, 1989, and Abstract No. 144, 12th International EPR Symposium, Denver, CO, USA, July 30- August 3, 1989.