

Ford Lumban Gaol
Keshav Shrivastava
Jamil Akhtar *Editors*

Recent Trends in Physics of Material Science and Technology



Springer

Springer Series in Materials Science

Volume 204

Series editors

Robert Hull, Charlottesville, USA

Chennupati Jagadish, Canberra, Australia

Richard M. Osgood, New York, USA

Jürgen Parisi, Oldenburg, Germany

Tae-Yeon Seong, Seoul, Korea, Republic of (South Korea)

Shin-ichi Uchida, Tokyo, Japan

Zhiming M. Wang, Chengdu, China

The Springer Series in Materials Science covers the complete spectrum of materials physics, including fundamental principles, physical properties, materials theory and design. Recognizing the increasing importance of materials science in future device technologies, the book titles in this series reflect the state-of-the-art in understanding and controlling the structure and properties of all important classes of materials.

More information about this series at <http://www.springer.com/series/856>

Ford Lumban Gaol · Keshav Shrivastava
Jamil Akhtar
Editors

Recent Trends in Physics of Material Science and Technology

Editors

Ford Lumban Gaol
Bina Nusantara University
Jakarta
Indonesia

Keshav Shrivastava
University of Malaya
Kuala Lumpur
Malaysia

Jamil Akhtar
Sensors and Nano-Technology Group
CEERI Pilani
Pilani
India

ISSN 0933-033X ISSN 2196-2812 (electronic)
Springer Series in Materials Science
ISBN 978-981-287-127-5 ISBN 978-981-287-128-2 (eBook)
DOI 10.1007/978-981-287-128-2

Library of Congress Control Number: 2014956202

Springer Singapore Heidelberg New York Dordrecht London
© Springer Science+Business Media Singapore 2015

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

Springer Science+Business Media Singapore Pte Ltd. is part of Springer Science+Business Media (www.springer.com)

Contents

Green Synthesis of Silver Nanoparticles Using Latex Extract of <i>Thevetia Peruviana</i>: A Novel Approach Towards Poisonous Plant Utilization	1
Ni Nyoman Rupiasih, Avinash Aher, Suresh Gosavi and P.B. Vidyasagar	
1 Introduction	2
2 Materials and Methods	4
2.1 Preparation of Latex Extracts	4
2.2 Synthesis of Silver Nanoparticles	4
2.3 Characterization Techniques	5
3 Results and Discussion	5
4 Concluding Remarks	9
References	9
 A Study on the Power Functions of the Shewhart $\bar{X}$ Chart via Monte Carlo Simulation	11
M.B.C. Khoo	
1 Introduction	11
2 Shewhart $\bar{X}$ Chart	12
3 Shewhart $\bar{X}$ Chart's Power Functions	14
4 Shewhart $\bar{X}$ Chart's Probability Tables	16
4.1 Formulae and Computation of Probabilities for the Power Functions	16
4.2 Probability Tables	20
5 A Study on the Performance	20
6 Concluding Remarks	25
References	26

B-Spline Collocation with Domain Decomposition Method and Its Application for Singularly Perturbed Convection-Diffusion Problems.	27
Mas Irfan P. Hidayat, Bambang Ariwahjoedi and Setyamartana Parman	
1 Introduction	28
2 Global B-Spline Collocation.	29
3 Overlapping Domain Decomposition Methods	31
4 Numerical Examples	34
4.1 Poisson Equation on L-Domain	35
5 Numerical Simulations of Singularly Perturbed Convection-Diffusion Problems	41
5.1 Singularly Perturbed Problem.	41
5.2 Convection-Dominated Problem	42
6 Conclusions	45
References	46
Introducing the Dimensional Continuous Space–Time Theory	49
Luiz Cesar Martini	
1 Introduction	49
2 Embryo of the Primitive Continuous Space–Time	50
3 Introducing the Temporal Variation in the Length α	52
4 Identification of Light Speed as a Universal Constant	54
5 Determination of the Angular Frequency ω_0	55
6 Introduction of the Phase Variation	56
7 Alteration of Equilibrium Conditions.	57
8 Real Condition of the Invariability Before the Big-Bang	59
9 Characterization of the Dimensional Variation	60
10 The General Form of the Primitive Equation in Static and Dynamics Forms	61
11 Some Theoretical Results.	63
12 Conclusions	65
References	65
Application of Nikiforov-Uvarov Method for Non-central Potential System Solution.	67
C. Cari and A. Suparmi	
1 Introduction	67
2 Non-central Potential.	68
3 Nikiforov-Uvarov Method	70
4 Application of NU Method for Non-central Potential.	75

4.1	Energy Spectrum and Wave Function of 3D Oscillator Harmonics Plus Trigonometric Rosen-Morse Non-central Potential	75
4.2	Energy Spectrum and Wave Function of Eckart Plus Poschl-Teller Non-central Potential	85
References		92

Charge Transport Mechanism and Low Temperature Phase

Transitions in KIO_3		95
M.M. Abdel Kader, F. El-Kabbany, H.M. Naguib and W.M. Gamal		
1	Introduction	95
2	Experimental Procedures	99
2.1	Sample Preparation	99
2.2	Characterization Techniques	99
3	Results	101
3.1	Ferroelectricity	101
3.2	Dielectric Permittivity	101
3.3	Electrical Conductivity	103
4	Discussion	105
5	Conclusion	109
References		109

Low Power Transmitter for Capsule Endoscope 111

De Xing Lioe, Suhaidi Shafie, Harikrishnan Ramiah,
Nasri Sulaiman and Izhal Abdul Halin

1	Introduction	111
2	Imager Selection	113
3	Design Consideration	114
3.1	Carrier Frequency	114
3.2	Power Consumption	115
3.3	Imaging Quality	116
4	Proposed Transmitter	116
4.1	Mixer	117
4.2	Local Oscillator	118
5	Results	119
6	Conclusions	120
References		120

Romanovski Polynomials Method and Its Application

for Non-central Potential System 123

A. Suparmi and C. Cari

1	Introduction	123
2	Romanovski Polynomials	124

3	Application of Romanovski Polynomials for Energy Spectra and Wave Functions Analysis for Non-central Potential	127
3.1	Eckart Plus Poschl-Teller Non-central Potential	127
3.2	The Solution of Polar Schrodinger Equation for 3D Trigonometric Rosen Morse Non-central Potential	139
	References	146

Interpretation and Prediction of Diffusion Activation Energy from Thermodynamic Model

	P. Wu, Y. Zeng and H.M. Jin	
1	Introduction	149
2	Quasi-chemical Calculations	151
3	Results	152
4	Concluding Remarks	155
	References	156

Shear Strength of Single Lap Joint Aluminium-Thermoplastic Natural Rubber (Al-TPNR) Laminated Composite

	M.Z. Muzakkar, S. Ahmad, M.A. Yarmo, A. Jalar and M. Bijarimi	
1	Introduction	160
2	Experimental	161
2.1	Material	161
2.2	Preparation of Al-TPNR Laminated Composite	162
2.3	Characterization of Al-TPNR Laminated Composite	163
3	Result and Discussion	163
3.1	FTIR Spectroscopy Analysis	163
3.2	XPS Analysis	164
3.3	SEM Analysis	166
3.4	Lap Shear Strength	166
4	Conclusions	169
	References	169

A New Technique to Observe ENSO Activity via Ground-Based GPS Receivers

	Wayan Suparta, Ahmad Iskandar and Mandeep Singh Jit Singh	
1	Introduction	173
2	Methodology	174
2.1	PWV Retrieved from GPS	174
2.2	Data and Location	176
2.3	Data Processing	176
3	Results and Discussion	179

3.1	PWV and SSTa Variability	179
3.2	Monitoring of GPS PWV Variability During El Niño Event	180
3.3	Analysis of Rainfall During El Niño Event	182
4	Conclusion.	184
	References	185
Oligostilbenoids from <i>Vatica</i> Species and Bioactivities		187
A.S. Kamarozaman, N.F. Rajab and J. Latip		
1	Introduction	187
2	Oligostilbenoids from <i>Vatica</i>	189
2.1	Stilbene Monomer	189
2.2	Stilbene Dimmer	190
2.3	Stilbene Trimer	190
2.4	Stilbene Tetramer	191
2.5	Stilbene Hexamer	191
2.6	Stilbene Heptamer	191
2.7	Stilbene Octamer	192
3	Biological Activities	192
3.1	Chemopreventive	192
3.2	Antioxidant Activity	193
3.3	Anticancer.	193
3.4	Other Activities	194
4	Conclusions	195
	References	201
Investigation on Thermoacoustic Cooling Device with Variation in Stack Plate Size and Input Acoustic Energy		205
Nandy Putra and Dinni Agustina		
1	Introduction	206
2	Keypoints in Constructing the Thermoacoustic Refrigerator	207
2.1	Relation of the Resonance Frequency and Resonator.	208
2.2	Driver.	209
2.3	Stack	209
2.4	Drive Ratio	211
2.5	Working Gas	211
3	Procedure of the Experiments.	212
3.1	Experimental Set-Up.	212
4	Result and Discussion	214
4.1	Effect of Input Voltage Setting on Thermal Performance	214

4.2	The Effect of Plate Thickness on the Thermal Performance and Cooling Rate	217
5	Conclusions	219
	References	219

Characteristic of Coal Stockpile in Lowland and the Effect to Environment

	Rusdianasari, Susila Arita, Eddy Ibrahim and Ngudiantoro	
1	Description of Stockpile	221
1.1	Types of Stockpile Based on the System of Stockpiling	222
1.2	Technical Terms of Stockpiling	223
1.3	Coal Delivery	224
1.4	Potential Effects of Coal Stockpiling.	225
2	The Characteristics of Coal Stockpile in Lowlands	226
3	Environment Quality of Coal Stockpile	227
3.1	Air Quality in Stockpile	227
3.2	Water Quality	230
4	Environmental Impact of Coal Stockpile	232
4.1	Impact on Air Quality.	232
4.2	Impact to Water Quality	233
4.3	Impact to Soil Quality.	236
	References	243

Synthesis and Characterization of Calcium Iodate, Monohydrate Crystals Grown in Silica Gel

	S.J. Shitole	
1	Introduction	245
2	Experimental	246
3	Result and Discussion	247
4	Characterization	249
4.1	XRD (X-Ray Diffractometry).	249
4.2	FTIR Characterization.	251
4.3	Thermal Analysis	252
4.4	Powder SHG Measurements	254
5	Conclusions	255
	References	256

Reduction of Power Consumption for Centrifugation Process Using Separation Efficiency Model

	M.S. Salim, M.F. Abd Malek and Naseer Sabri	
1	Introduction	259
2	Separation Efficiency	261
3	Linear Velocity Profile Generator	262

4	Experimental Setup	262
5	Centrifugation Period Model	263
6	Centrifugation Power Consumption Model	264
7	Conclusions	268
	References	269

A Wide Range Measurement and Disclosure of Ultrasonic Attenuation for Very Low Concentrations of Solid in Liquids

Using Exponential Pulser Technique. 271

M.S. Salim, M.F. Abd Malek and Naseer Sabri

1	Introduction	271
2	Ultrasonic Pulser Design	274
3	Validation of PPD Design	276
4	Discussion	279
5	Conclusion.	280
	References	281

Wireless Sensor Network Wave Propagation in Vegetation 283

Naseer Sabri, S.A. Aljunid, M.S. Salim, S. Fouad and R. Kamaruddin

1	Introduction	283
2	Precision Agriculture Based WSN	284
3	Physical Aspect Challenges of Applying WSN in Agricultural Fields	285
4	Radio Wave Propagation Models in Simulation Tools	286
	4.1 Free Space Path Loss (FSPL).	286
	4.2 Two-Ray Ground Reflection Model	287
5	Wave Propagation Models in Vegetation	287
	5.1 Weissberger's Model	288
	5.2 Early ITU Vegetation Model	288
6	Total Path Loss in Vegetation	289
7	Simulation Setup of WSN in Agricultural Greenhouse	290
8	Results and Discussion	291
	8.1 Communication Coverage Simulation	293
	8.2 Network Connectivity Simulation	294
9	Conclusions	296
	References	297

Fiber Optic Sensors: Short Review and Applications. 299

Naseer Sabri, S.A. Aljunid, M.S. Salim and S. Fouad

1	Introduction	299
2	General Features of Fiber Optical Sensors	300
3	Fiber Optic Sensor Principles	302
4	Fiber Optic Sensors Category	303
5	Classification of Fibre Optical Sensors	304

6	Measurands of Optical Sensors.	305
6.1	Sensing Strategies.	305
6.2	Degree of Sensing	306
6.3	Role of Optical Fiber	306
7	Fiber Optical Sensor Types	306
8	Fiber Optical Sensing Applications	307
9	Conclusions	309
	References	310

Dynamic Effects of Piezoelectric Patch Actuators on Vibrational Response of Non-deterministic Structures:

	Modelling and Simulations.	313
	Azni N. Wahid, Asan G.A. Muthalif and Khairul A.M. Nor	
1	Introduction	313
1.1	Deterministic Structures Versus Non-deterministic Structures	313
1.2	Statistical Energy Analysis	314
2	Derivation of Equation of Motion	316
2.1	A Rectangular Plate with Distributed Point Masses Attached with a PZT Patch Actuator.	316
3	Power Delivered to a Non-deterministic Thin Rectangular Plate	321
4	Simulation Studies	321
4.1	Influence of PZT Patch Actuator Location and Size on the Average Power Delivered to the Structure	324
5	Conclusions	326
	References	335

Solubilities Prediction of Ginger Bioactive Compounds

	in Liquid Phase of Water by the COSMO-RS Method	337
	Syaripah Za'imah Syed Jaapar, Noor Azian Morad and Yoshio Iwai	
1	Introduction	338
2	COSMO-RS Calculation Procedures	340
3	Results and Discussion	341
3.1	Validation of Calculated Results.	341
3.2	Calculated Results for Solubilities of Ginger Bioactive Compounds in Water	347
4	Conclusion.	351
	References	351

Green Synthesis of Silver Nanoparticles Using Latex Extract of *Thevetia Peruviana*: A Novel Approach Towards Poisonous Plant Utilization

Ni Nyoman Rupiasih, Avinash Aher, Suresh Gosavi
and P.B. Vidyasagar

Abstract The development of green synthesis of nanoparticles has received increasing attention due to ease of preparation, less chemical handling and of eco-friendly nature. It is of considerable importance to expand their biological applications. Currently, a variety of inorganic nanoparticles with well-defined chemical composition, size and morphology have been synthesized by using different microorganisms, fungus, plant extracts and their applications in many cutting-edge technological areas have been explored. In the present study, silver nanoparticles (AgNPs) were successfully synthesized from AgNO_3 through a simple green synthesis route using the latex of *Thevetia peruviana* as reducing as well as capping agent. Synthesized silver nanoparticles were characterized using UV-Vis spectrophotometer, FTIR, SEM-EDS, XRD and HRTEM. UV-Vis spectra showed absorption at 570 nm with a shoulder at 395 nm. FTIR spectra analysis confirmed the functional groups involved in the silver nanoparticles formation. X-ray diffraction pattern of silver nanoparticles exhibited the diffraction angle at 34.48° , 40.12° and 45.39° which is corresponding to (222), (400) and (420) Bragg's reflection planes respectively with interplaner spacing of 2.59 Å obtained by HRTEM. TEM analysis showed the particles are spherical in nature with its size distribution between 10 and 30 nm.

N.N. Rupiasih (✉)

Department of Physics, Faculty of Mathematics and Natural Sciences,
Udayana University, Kampus Bukit Jimbaran, Badung, Bali 80362, Indonesia
e-mail: rupiasih69@yahoo.com

A. Aher · S. Gosavi · P.B. Vidyasagar

Department of Physics, University of Pune, Ganeshkhind, Pune 411007, India

© Springer Science+Business Media Singapore 2015

F.L. Gaol et al. (eds.), *Recent Trends in Physics of Material Science and Technology*,
Springer Series in Materials Science 204, DOI 10.1007/978-981-287-128-2_1

1 Introduction

Recently, the applications of nanoscale materials and structures have attracted great attention in nanoscience and nanotechnology. Nanoparticles are often referred to particles having one or more dimensions ranging from 1 to 100 nm and can be composed of different base materials. They do occur both in nature and as a result of industrial processes. They can exist in single, complex, aggregated or agglomerated forms with spherical, tubular and irregular shapes.

Nanoparticles exhibit completely new or improved properties compared to their bulk counterparts. These novel properties are derived due to the variation in specific characteristics such as size, distribution and morphology of the particles [1–3]. The two main reasons that materials at nano scale can have different properties are increased relative surface area and quantum size effects [4]. Nanoparticles have much greater surface area to volume ratio than their bulk forms, which can lead to greater chemical reactivity and affect their strength. Also at the nano scale, quantum effects can become much more important in determining the materials properties and characteristics, leading to novel optical, electrical and magnetic behaviours [4]. Technologically nanomaterials are resources designed at the molecular (nano meter) level to take advantage of their small size and novel properties which are generally not seen in their conventional, bulk counterparts.

The emergent properties of the nanoparticles have its potential for great impacts in electronics, medicine and other fields. In medicine, nanoparticles delivery systems had been used for tumor targeting [5]. The specific surface area is relevant for catalytic activity and other related properties such as antimicrobial activity [2]. As the specific surface area of nanoparticles is increased, their biological effectiveness can also increase on the account of a rise in surface energy. Gupta et al. reported that in cotton fabrics, silver nanoparticles show biocidal action against the bacteria *E. coli*, which is showing great potential to be used as an antiseptic dressing or bandage, which is in high demand for biomedical applications [6].

Nanoparticles with controlled size and composition are of fundamental and technological interest as they provide solutions to technological and environmental challenges in the areas of solar energy conversion, catalysis, medicine and water treatment. Global warming and climate change have induced a worldwide awareness and efforts to reduce generated hazardous wastes. Thus, “green” chemistry and chemical processes are progressively being integrated in science and industry for sustainable development [7, 8]. One of them is synthesis of nanoparticles by biological process which is implemented to develop safe, cost-effective and environmentally friendly technologies [7].

There are various methods, including physical and chemical which have been developed to prepare metal nanoparticles, such as chemical reduction [9], electrochemical, photochemical reduction, heat evaporation [10] and very recently via green chemistry route [10, 11]. Although physical and chemical methods are more popular in the synthesis of nanoparticles, the use of toxic chemicals are toxic enough to pollute the environment if large scale nanoparticles are produced.

It greatly limits their biomedical applications, in particular in clinical fields. Also, these methods are quite expensive due to low material conversions, high energy requirements, difficult and wasteful purifications [11]. Therefore, development of reliable, nontoxic and eco-friendly methods for synthesis of nanoparticles is of utmost importance to expand their biomedical applications as well as saving the environment. Thus, among biosynthetic methods employing biological things such as microorganisms or plant extracts have emerged as a simple, clean, eco-friendly, nontoxic and viable alternative to physical and chemical methods.

A great deal of effort has been put into the biosynthesis of inorganic material, especially metal nanoparticles using microorganisms and plants [1, 2, 12]. Biosynthesis methods are methods that use biological systems to synthesize nanoparticles. So far, different biological resources including herbal extracts [11–13], microalgae [14], fungus [3, 15] and bacteria [2, 16, 17] have been used for synthesis of metal nanoparticles. The use of plant extracts for the synthesis could be more advantageous, because it does not require elaborate processes such as intracellular synthesis and multiple purification steps or the maintenance of microbial cell cultures [7, 11, 12].

Up to now, metallic nanoparticles are mostly prepared from noble metals such as Ag, Pt, Au and Pd, among those; silver (Ag) is the metal of choice in the field of biological system, living organisms and medicine [18]. Among inorganic antibacterial agents, silver has been employed most extensively since ancient times to fight infections and control spoilage. The antibacterial and antiviral actions of silver, silver ion and silver compounds have been thoroughly investigated [19].

Gardea-Torresdey et al. first reported the formation of gold and silver nanoparticles by living plants [20, 21]. Their synthetic protocol by plant extract or biomass exemplifies the promising application of the green synthesis of metal nanoparticles. Recent research reported that silver nanoparticles have been synthesized using various natural products like green tea (*Camellia sinensis*), neem (*Azadirachta indica*) leaf broth, natural rubber, starch, aloe vera plant extract, lemongrass leaves extract, etc. [11, 13]. The shape and size of the nanoparticles synthesized using plants can be controlled and modulated by changing and controlling the pH of the solution [2]. The reduction of silver ions (Ag^+) in aqueous solution generally yields colloidal silver with particle diameters of several nanometers.

Studies have indicated that biomolecules like protein, phenols and flavonoids not only play a role in reducing the ions to the nanosize, but also play an important role in the capping of the nanoparticles [22]. The reduction of Ag^+ ions by combinations of biomolecules found in these extracts such as vitamins, enzymes/proteins, organic acids such as citrates, amino acids and polysaccharides [22–24] is environmentally benign, yet chemically complex.

Medicinal herbs are the local heritage with global importance. It has curative properties due to the presence of various complex chemical substances of different composition, which are found as secondary plant metabolite in one or more parts of the plants [7, 8, 11, 25]. These plant metabolites according to their composition are grouped as alkaloids, glycosides, corticosteroids, essential oils, etc.

Thevetia peruviana is an ever-green ornamental dicotyledonous plant that belongs to *Apocyanaceae* family. It is native of tropical America especially Mexico and West Indies and a close relative to *Nerium oleander*. It has naturalized in tropical regions [26, 27].

The whole plant contains a milky juice (white latex) containing a compound called thevetin that is used as a heart stimulant but in its natural form is extremely poisonous. It is poisonous owing to the presence of cardiac glycosides or cardiac toxins, which acted directly in the heart [26–28]. Ingestion of these plant parts could lead to death [26, 27]. The seed contains oil with high content of fatty acid. It is believed that *Thevetia* oil is rich in antioxidants with some specific components of vitamin E groups [27]. It is cultivated as ornamental tree standards in gardens and parks in temperate climates [29]. In the West Indies half a leaf is known as an emetic and purgative. In Java, Indian immigrants are reported to smoke using the dried leaves. Tincture of bark is also emetic and purgative and has been used as a febrifuge, but tests with bark and fruit extracts have proved inactive against avian malaria. In Indonesia it is reported being used as a fish-poison [30].

The present study demonstrated method for the synthesis of silver nanoparticles by the reduction of aqueous silver ions using latex from *Thevetia peruviana*, a poisonous as well as medicinal herbs plant [29]. It resulted in the extracellular formation of silver nanoparticles at room temperature which was further harvested by centrifugation and simple natural dried.

2 Materials and Methods

2.1 Preparation of Latex Extracts

Fresh fruits of *Thevetia peruviana* were collected, washed thoroughly 2–3 times with tap water, continuously rinse with distilled water, and used for the extraction. Milky white latex was obtained by cutting the green stems of the fruits. 100 ml of distilled water was added with 100 μ l of fresh latex extract, mixed gently and filtered through Whatman No.1 filter paper. The solution was stored at 4 °C for further used.

2.2 Synthesis of Silver Nanoparticles

A 50 μ l of aqueous solution of 1 M silver nitrate (AgNO_3) was added into 10 ml of latex extract solution and was kept at room temperature. The reaction of silver ions with the extract solution leads to the formation of silver nanoparticles (AgNPs). After around 6 h, the colour of the mixture turned into light purple and continuously changed to purple–brown indicating the formation of silver nanoparticles.

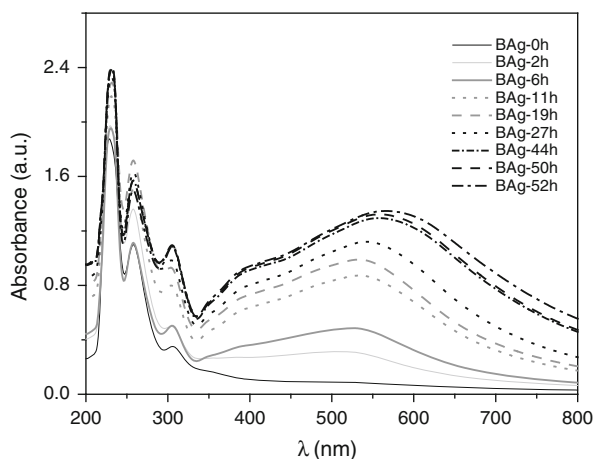
2.3 Characterization Techniques

UV-Vis absorption spectra were measured using JASCO V-670 UV-Vis spectrophotometer in the wavelength range of 200–800 nm. The progress of reaction between silver ions and the latex extract solution were monitored in extract solution at different reaction time between 0 and 52 h. Energy Dispersive elemental analysis is done with SEM-EDS (JEOL JSM 6360A). Crystalline nature of silver nanoparticles formed was examined by X-ray diffraction (XRD) analysis using X-ray diffractometer (BRUKER AXS D8) with a Cu-K α ($\lambda = 1.54 \text{ \AA}$) in the range of 20°–80°. Morphology and size of silver nanoparticles were investigated using transmission electron microscopy (TEM) using TECNAI G² 20U-TWIN instrument. FTIR spectrophotometer measurements were done using JASCO (FT/IR-6100) spectrophotometer in the range of 400–4,000 nm.

3 Results and Discussion

Figure 1 shows the UV-Vis absorption spectra of silver nanoparticles (AgNPs) synthesized using latex extract of *Thevetia peruviana* at different time interval. It shows the characteristic of surface plasmon resonance (SPR) absorption band at around 570 nm for silver nanoparticles (purple–brown solution). It is observed that the intensity of SPR bands increases as the reaction time progresses from 0 to 44 h and after 44 h of reaction time the SPR intensity is nearly constant indicating that the reaction process is almost completed. The observed absorption spectra are broad in nature with a tail at longer wavelength and a shoulder at around 395 nm indicating the formation of polydispers AgNPs. Figure 1 also shows the red shifted band from 510 to 570 nm which occurred with increasing reaction time from 0 to 52 h.

Fig. 1 UV-Vis absorption spectra of silver nanoparticles (AgNPs) synthesized at room temperature using latex extract of *Thevetia peruviana* at different time: 0–52 h



The strong absorption peak at around 230 nm observed indicated the presence of several organic compounds which are known to interact with silver ions. Also an absorption band observed at around 260 nm is attributed to the aromatic amino acids of proteins. It is well known that the absorption band at 260 nm arises due to electronic excitations of tryptophan and tyrosine residues in the proteins. This observation indicates the release of proteins into solution by latex of *Thevetia peruviana* and suggests a possible mechanism for the reduction of the silver ions present in the solution.

Figure 2 shows representative EDS profile for synthesized silver nanoparticles. The energy dispersive X-ray analysis reveals strong signal in the silver region and confirms the formation of silver nanoparticles. There is also a strong signal for Si, which due to Si (111) wafers which have been used as substrate to prepare thin film. Other elemental signals are recorded possibly due to elements from enzymes or proteins present within the latex of *Thevetia peruviana*.

A thorough study of the literature on *Thevetia peruviana* reveals that the major components of the plants are protein, phenols, oil, mome inositol, Ethyl palmitate, Ethyl linolenate and lesser amount of glycosides with the major components is thevetin [28]. It has been reported by Zhou et al. [23] that the carbonyl groups from the amino acid residues and peptides of proteins have a strong affinity to bind metals. So that protein can act as encapsulating agent and thus protect the nanoparticles from agglomeration. *Thevetia* oil contains high proportion of fatty acid and is known to be rich in antioxidants with some specific components of vitamin E groups [27]. This compound may also contribute to the reduction of Ag ions to silver nanosize. Ascorbic acid (vitamin E) is a reducing agent and can reduce, and thereby neutralize, reactive oxygen species leading to the formation of ascorbate radical and an electron. This free electron reduces the Ag^+ ion to AgO [22]. FTIR measurements were

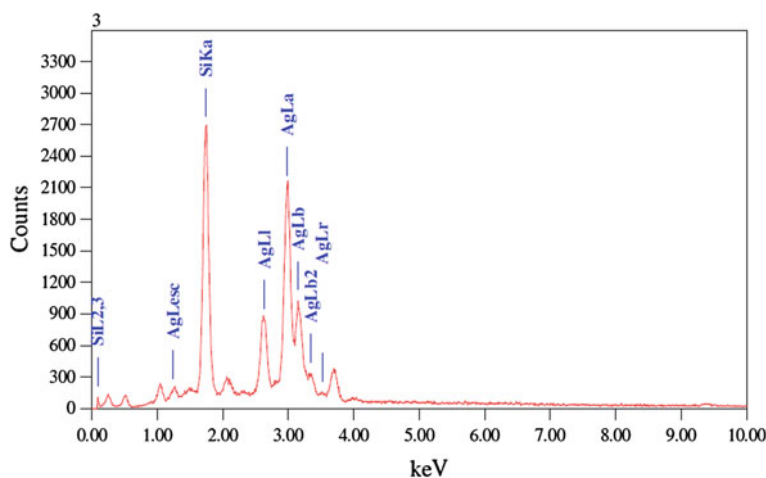


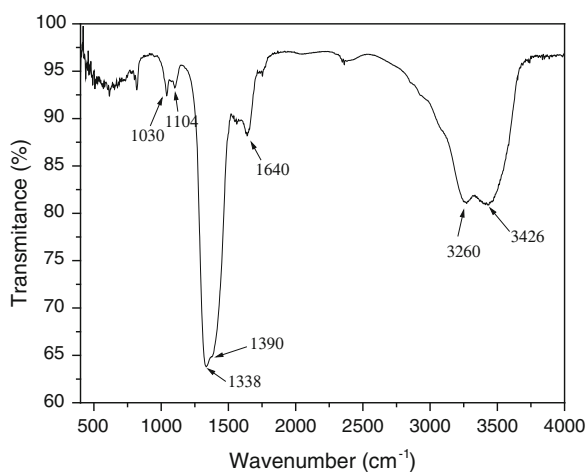
Fig. 2 EDS spectrum of silver nanoparticles (AgNPs) synthesized at room temperature using latex extract of *Thevetia peruviana*

carried out to identify the possible biomolecules responsible for the reduction of the Ag^+ ions and the capping of the bioreduced silver nanoparticles synthesized by the latex extract of *Thevetia peruviana*. The solution after complete reduction of Ag^+ was centrifuged at 15,000 rpm for 10 min to isolate the silver nanoparticles free from proteins or other compounds present in the solution. The representative spectra of nanoparticles obtained as shown in Fig. 3 manifest absorption bands located at 3426, 1640, 1390, 1338, 1104, and 1030 cm^{-1} . The absorption peak at around 1,030 cm^{-1} can be assigned as absorption peaks of $-\text{C}-\text{O}-\text{C}-$ or $-\text{C}-\text{O}-$. The peak at around 1,640 cm^{-1} is assigned to the amide I bonds of proteins. The bonds or functional groups such as $-\text{C}-\text{O}-\text{C}-$, $-\text{C}-\text{O}-$ and $-\text{C}=\text{C}-$ derived from heterocyclic compounds and the amide I bond derived from the proteins which present in the extract are the capping ligands of the nanoparticles [1, 11, 13, 22]. The band at 1,390 and 1,104 cm^{-1} , as well as a strong band at 1,338 cm^{-1} indicating presence of amine groups, as would be expected due to plant-origin of the samples. The peak at around 3,426 cm^{-1} was characteristic of $-\text{N}-\text{H}$ stretching of amide I band [22].

Figure 4 shows the XRD patterns of natural dried silver nanoparticles synthesized using latex of *Thevetia peruviana*. A number of Bragg reflections with 2θ values of 34.48°, 40.12° and 45.39° sets of lattice planes are observed which may be indexed to the (222), (400) and (420) faces of silver respectively. XRD pattern thus clearly illustrates that the silver nanoparticles formed in this present synthesis are crystalline in nature and having Face Centered Cubic crystal structure.

Figure 5 shows the TEM micrograph of the silver nanoparticles synthesized by latex of *Thevetia peruviana*. Figure 5a is the representative TEM image of synthesized silver nanoparticles. The morphology of the silver nanoparticles is found to be predominantly spherical. The nanoparticles have two dominant size distributions ranges such as 10–20 nm and 20–30 nm as shown in Fig. 5b, c shows HRTEM image of silver nanoparticles. The fringe spacing is found to be 2.59 Å. These results were in agreement with those reported by Kiruba Daniel et al. [1], Bar et al. [11] and Ahmad et al. [22].

Fig. 3 FTIR spectra of silver nanoparticles (AgNPs) synthesized at room temperature using latex extract of *Thevetia peruviana*



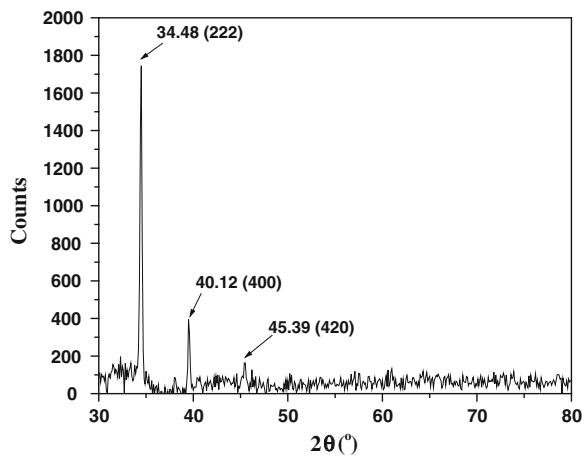


Fig. 4 XRD patterns of silver nanoparticles (AgNPs) synthesized at room temperature using latex extract of *Thevetia peruviana*

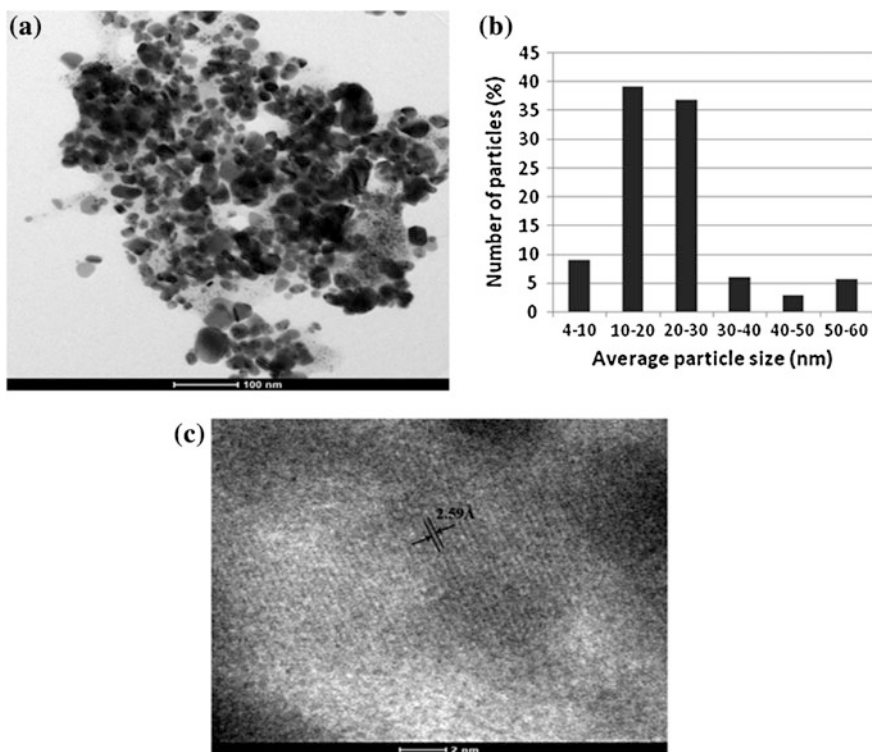


Fig. 5 **a** Representative TEM image of silver nanoparticles (AgNPs) synthesized at room temperature using latex extract of *Thevetia peruviana*. **b** Histogram of particle size distribution and **c** high resolution transmission electron microscopy (HRTEM) image of silver nanoparticles

4 Concluding Remarks

The rapid synthesis of silver nanoparticles (AgNPs) using latex extract of *Thevetia peruviana* has been demonstrated. UV-Vis absorption spectra showed that silver nanoparticles formed have characteristic of surface plasmon resonance (SPR) absorption band at around 570 nm with a shoulder at around 395 nm. Also observed, are strong absorption peaks at around 230 and 260 nm which indicate that compound may play a role in reducing the silver ions to nanosize and in the capping of the nanoparticles. This result corroborated with FTIR spectra obtained, confirmed that the functional groups are involved in silver nanoparticles formation. X-ray diffraction pattern of silver nanoparticles exhibited the diffraction angle at 34.48°, 40.12° and 45.39° which is corresponding to (222), (400) and (420) Bragg's reflection planes respectively, with interplaner spacing of 2.59 Å obtained by HRTEM image. TEM analysis showed the particles are spherical with size between 10 and 30 nm. The present green synthesis showed that the environment friendly and renewable latex of *Thevetia peruviana* could be used as an effective capping as well as reducing agent for the synthesis of silver nanoparticles. This eco-friendly route for the synthesis is a challenging alternative to chemical synthesis.

References

1. Kiruba Daniel SCG, Kumar R, Sathish V, Sivakumar M, Sunitha S, Anitha Sironmani T (2011) Green synthesis (*Ocimum tenuiflorum*) of silver nanoparticles and toxicity studies in Zebrafish (*Danio rerio*) model. *Int J Nanosci Nanotechnol* 2(2):103–117
2. Gurunathana S, Kalishwaralala K, Vaidyanathana R, Deepaka V, Pandiana SRK, Muniyandia J, Hariharana N, Eom SH (2009) Biosynthesis, purification and characterization of silver nanoparticles using *Escherichia coli*. *Colloids Surf B* 74:328–335
3. Li G, He D, Qian Y, Guan B, Gao S, Cui Y, Yokoyama K, Wang L (2012) Fungus-mediated green synthesis of silver nanoparticles using *Aspergillus terreus*. *Int J Mol Sci* 13:466–476
4. Buzea C, Pacheco CC, Robbie K (2007) Nanomaterials and nanoparticles: sources and toxicity. *Biointerphases* 2(4):MR17–MR172. <http://arxiv.org/ftp/arxiv/papers/0801/0801.3280.pdf>. Accessed 2 June 2013
5. Mohanraj VJ, Chen Y (2006) Nanoparticles—a review. *Trop J Pharm Res* 5(1):561–573
6. Gupta P, Bajpai M, Bajpai SK (2008) Investigation of antibacterial properties of silver nanoparticle-loaded poly(acrylamide-co-itaconic acid)-grafted cotton fabric. *J Cotton Sci* 12:280–286
7. Gavhanel AJ, Padmanabhan P, Kamble SP, Jangle SN (2012) Synthesis of silver nanoparticles using extract of neem leaf and triphala and evaluation of their antimicrobial activities. *Int J Pharm Bio Sci* 3(3):88–100
8. Geethalakshmi R, Sarada DVL (2010) Synthesis of plant-mediated silver nanoparticles using *Trianthema decandra* extract and evaluation of their anti microbial activities international. *J Eng Sci Technol* 2(5):970–975
9. Guzmán MG, Dille J, Godet S (2009) Synthesis of silver nanoparticles by chemical reduction method and their antibacterial activity. *Int J Chem Biol Eng* 2(3):104–111
10. Mukunthan KS, Elumalai EK, Patel TN, Murty VR (2011) *Catharanthus roseus*: a natural source for the synthesis of silver nanoparticles. *Asian Pac J Trop Biomed* 1:270–274

11. Bar H, Bhui DK, Gobinda PS, Sarkar P, De Sankar P, Misra A (2009) Green synthesis of silver nanoparticles using latex of *Jatropha curcas*. Colloids Surf A 339:134–139
12. Ponarulselvam S, Panneerselvam C, Murugan K, Aarthi N, Kalimuthu K, Thangamani S (2012) Synthesis of silver nanoparticles using leaves of *Catharanthus roseus* Linn. G. Don and their antiparasmodial activities. Asian Pac J Trop Biomed 2:574–580
13. Hudlikar M, Joglekar S, Kodam MDK (2012) Latex-mediated synthesis of ZnS nanoparticles: green synthesis approach. J Nanopart Res 14:865
14. Singaravelu G, Arockiamary JS, Ganesh Kumar V, Govindaraju K (2007) A novel extracellular synthesis of monodisperse gold nanoparticles using marine alga, *Sargassum wightii* Greville. Colloids Surf B 57:97–101
15. Mishra A, Tripathy SK, Wahab R, Jeong S-H, Hwang I, Yang Y-B, Kim Y-S, Shin H-S, Yun S-II (2011) Microbial synthesis of gold nanoparticles using the fungus *Penicillium brevicompactum* and their cytotoxic effects against mouse mayo blast cancer C₂C₁₂ cells. Appl Microbiol Biotechnol 92:617–630
16. Husseiny MI, Abd El-Aziz M, Badr Y, Mahmoud MA (2007) Biosynthesis of gold nanoparticles using *Pseudomonas aeruginosa*. Spectrochim Acta Part A 67:1003–1006
17. Natarajan K, Subbalaxmi Selvaraj V, Murty R (2010) Microbial production of silver nanoparticles. Dig J Nanomater Biostruct 5(1):135–140
18. Yamini SudhaLakshmi G (2011) Green synthesis of silver nanoparticles from *Cleome viscosa*: synthesis and antimicrobial activity. In: International conference on bioscience, biochemistry and bioinformatics IPCBEE, vol 5. IACSIT Press, Singapore, pp 334–337
19. Parameswari E, Udayasoorian C, Paul Sebastian S, Jayabalakrishnan RM (2010) The bactericidal potential of silver nanoparticles. Int Res J Biotechnol 1(3):044–049
20. Gardea-Torresdey JL, Parsons JG, Dokken K, Peralta-Videa J, Troiani HE, Santiago P, Jose-Yacamán M (2002) Formation and growth of Au nanoparticles inside live Alfalfa plants. Nano Lett 2:397–401
21. Gardea-Torresdey JL, Gomez E, Peralta-Videa E, Parsons JG, Troiani HE, Jose-Yacamán M (2003) Alfalfa sprouts: a natural source for the synthesis of silver nanoparticles. Langmuir 19:1357–1361
22. Ahmad N, Sharma S, Singh VN, Shamsi SF, Fatma A, Mehta BR (2011) Biosynthesis of silver nanoparticles from *Desmodium triflorum*: a novel approach towards weed utilization. Biotechnol Res Int 2011:8pp
23. Zhou Y, Lin W, Huang J, Wang W, Gao Y, Lin L, Li Q, Lin L, Mingming D (2010) Biosynthesis of gold nanoparticles by foliar broths: roles of biocompounds and other attributes of the extracts. Nanoscale Res Lett 5:1351–1359
24. Szczepanowicz K, Stefańska J, Socha RP, Warszyński P (2010) Preparation of silver nanoparticles via chemical reduction and their antimicrobial activity. Physicochem Probl Mineral Process 45:85–98
25. Deepak V, Kalishwaralal K, Pandian SRK, Gurunathan S (2011) An insight into the bacterial biogenesis of silver nanoparticles. In: Rai M, Duran N (eds) Industrial production and scale-up chapter 2 metal nanoparticles in microbiology. Springer, Berlin, pp 173–214
26. Usman LA, Oluwaniyi OO, Ibiyemi SA, Muhammad NO, Ameen OM (2009) The potential of oleander (*Thevetia peruviana*) in African agricultural and industrial development: a case study of Nigeria. J Appl Biosci 24:1477–1487
27. Mathuravalli K, Eswara Lakshmi R (2012) Analysis of phytochemical components and antimicrobial activity of the toxic plant—*Thevetia peruviana*. Indian J Innov Dev 1(2):97–101
28. Sangodare RSA, Agbaji AS, Dakare MA, Usman YO, Magomya A, Paul ED, Ibraheem AS, Etim NA, Ibrahim A, Aribido O, Shikeola O (2012) Investigation of the chemical constituent of extracts of *Thevetia peruviana* seed using GC-MS and FT-IR. Int J Food Nutr Saf 2(1):27–36
29. Zibbu G, Batra A (2011) *Thevetia peruviana* (Pers.) Schum: a plant with enormous therapeutic potential. J Pharm Res 4(12):4461–4464
30. Burkill HM (1985) Entry for *Thevetia nerifolia* Juss. (Family Apocynaceae) the useful plants of West Tropical Africa. Jstor Plant Sci 1. http://Plants.Jstor.Org/Upwta/1_411. Accessed 9 Oct 2012

A Study on the Power Functions of the Shewhart $\bar{X}$ Chart via Monte Carlo Simulation

M.B.C. Khoo

Abstract The Shewhart $\bar{X}$ control chart is used to monitor shifts in the process mean. However, it is less sensitive to small shifts. The Shewhart $\bar{X}$ chart's sensitivity can be enhanced by reducing the width of the control limits, increasing the subgroup size and using detection rules to enhance the chart's sensitivity. However, these actions will influence the power functions of the Shewhart $\bar{X}$ chart. A probability table providing the probabilities of detecting shifts in the mean, calculated using the formulae is recommended. However, the main setback is that the calculations of the probabilities using the formulae are complicated, laborious and time consuming. In this paper, a Monte Carlo simulation using the Statistical Analysis System (SAS) is conducted to compute these probabilities. The probabilities computed via Monte Carlo simulation are closed to that obtained using formulae. Therefore, the Monte Carlo simulation method is recommended as it provides savings, in terms of time and cost. In addition, the Monte Carlo simulation method is also more flexible in calculating the probabilities, for different combinations of the detection rules. The results obtained will enable practitioners to design and implement the Shewhart $\bar{X}$ chart more effectively.

1 Introduction

The control chart for monitoring process quality was introduced by Dr. Walter A. Shewhart in the 1920s. Shewhart defined the product attributes, types of product variations and proposed methods to collect, plot and analyze data [1].

The Shewhart $\bar{X}$ control chart is an important tool in Statistical Process Control (SPC). It detects assignable causes in process control so that process investigation and corrective actions can be made before many defective products are produced [2].

M.B.C. Khoo (✉)

School of Mathematical Sciences, Universiti Sains Malaysia, 11800 Minden,
Penang, Malaysia
e-mail: mkbc@usm.my

The main objective of applying the Shewhart $\bar{X}$ chart is to prevent failure so that production of low quality products will not occur. This enables cost savings and the production of high quality products.

The Shewhart $\bar{X}$ chart is powerful in detecting large process mean shifts but it is slow in detecting small shifts. Many researches have been made by quality experts to improve the Shewhart $\bar{X}$ chart's sensitivity. Some of these improvements include the synthetic $\bar{X}$ charts by [3–5], the sequential probability ratio test (SPRT) chart by [6], and the time weighted control charts (see [7–9], to name a few). However, more common approaches to increase the sensitivity of the $\bar{X}$ chart are by increasing the subgroup size, employing tighter control limits and applying sensitizing rules, such as those by [10–13]. All these approaches influence the power functions of the Shewhart $\bar{X}$ chart. Tables showing the probabilities of detecting shifts in the mean, computed using the formulae were presented by [14] to guide quality practitioners to construct the Shewhart $\bar{X}$ chart. In this work, a Monte Carlo simulation method is given, where similar results are obtained. This research is motivated by the fact that the Shewhart $\bar{X}$ chart is the most widely used control chart among practitioners, hence a good understanding of the power function of this chart enables the chart to be used more efficiently in process monitoring.

This chapter is organized as follows: The Shewhart $\bar{X}$ chart is discussed in Sect. 2. In Sect. 3, the power functions of the Shewhart $\bar{X}$ chart are presented. The probability table of the power functions of the Shewhart $\bar{X}$ chart is explained in Sect. 4. Section 5 compares the performance between the probabilities obtained via formulae and that computed using Monte Carlo simulation. Lastly, conclusions are drawn in Sect. 6.

2 Shewhart $\bar{X}$ Chart

The Shewhart $\bar{X}$ chart is a times series plot which provides assistance in identifying whether a process is in a state of statistical control. It is a variables control chart with no memory since it uses only the recent data in its control statistics. Having this property, the Shewhart $\bar{X}$ chart performs well when we are interested in the detection of large shifts. The Shewhart $\bar{X}$ contains three crucial decision lines, i.e., the center line (CL), upper control limit (UCL) and lower control limit (LCL). UCL and LCL are occasionally known as the “natural process limits” as they shows threshold at which the quality characteristic of a process being monitored is regarded statistically “unlikely”. The Shewhart $\bar{X}$ chart is constructed based on some statistical principles. As a common practice, the limits of the Shewhart $\bar{X}$ chart are taken as $\pm 3\sigma$ from the CL. That is, the UCL is drawn $+3\sigma$ above the CL whereas the LCL is drawn -3σ below the CL. The $\pm 3\sigma$ limits are chosen to strike a balance between the risk of the Type-I and Type-II errors. Here, the value of the mean of the statistic is represented by the CL. All the plotted sample points on the Shewhart $\bar{X}$ chart are connected so that the quality practitioner can have a better

view on how a process evolves over time. An out-of-control signal or action signal will be given when a plotted point falls outside the UCL or LCL limit. When a process has shifted, we wish to detect the assignable cause as soon as possible and on the other hand, we wish to have a minimum rate of false alarms when the process is in-control [8]. This is because a slow response to an out-of-control process can cause quality deterioration and increase quality costs while too high false alarm rates can give rise to unnecessary process adjustments and loss of confidence in the control charting methods.

Assume that a quality characteristic is normally distributed with mean μ and standard deviation σ . Here, both μ and σ are assumed known. If $X_1, X_2, \dots, X_n$ is a sample of size n , then the mean of this sample can be computed as [9]

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} \quad (1)$$

and we know that $\bar{X}$ is normally distributed with mean μ and standard deviation $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$. The probability that any sample mean will fall between

$$\mu + Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \quad (2a)$$

and

$$\mu - Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \quad (2b)$$

is $1 - \alpha$. Equations (2a) and (2b) can be taken as the UCL and LCL of the Shewhart $\bar{X}$ chart if μ and σ are known. To use the $\pm 3\sigma$ limits, we can just simply replace the $Z_{\alpha/2}$ by 3.

When the sample mean falls outside these limits, the process mean is considered as out-of-control. Corrective actions to search and eliminate the assignable causes are taken so that the process returns to its in-control state again. Note that the above results are approximately correct when the assumption of normality is violated due to the central limit theorem. The values of μ and σ are usually unknown in real situations and therefore, these parameters are required to be estimated from a set of in-control historical Phase-I data. When estimates are used in place of known parameters, at least 20–25 samples are required for generating better and reliable estimations. Note that a revision of the control limits periodically is necessary to ensure an effective use of the control chart.

3 Shewhart $\bar{X}$ Chart's Power Functions

In the literature of statistical quality control, hypothesis testing is always used to summarize an inference on the mean of a population which is given as:

$$\begin{aligned} H_0 : \mu &= \mu_0, \\ H_1 : \mu &\neq \mu_0. \end{aligned} \quad (3)$$

[9] defined the power as

$$\begin{aligned} \text{Power} &= 1 - \beta \\ &= P\{\text{reject } H_0 | H_0 \text{ is false}\}, \end{aligned} \quad (4)$$

where β is the Type-II error probability. In industries, quality users employ a control chart to prevent the production of defective outputs so that quality costs can be minimized [10]. Hence, an in-depth understanding of the method that helps to increase the power of a control chart in detecting process shifts is deemed indispensable.

The calculation of the control limits and sample standard deviation involves the subgroup size, n . So, n is one of the factors that can affect the power of a control chart. The n is proportional to the power of a control chart. As n increases, the power of a control chart increases and vice versa. However, using a large n is impractical in the industrial setting as it will inflate the quality cost. The control chart's sensitivity in detecting process mean shifts can be enhanced via the application of detection rules. Many authors have contributed to new methods on detection rules (also known as runs rules). Recent works on detection rules, were made by [11–15].

When more detection rules are used, the overall Type-I error probability is given by [9, 16]. The overall Type-I error probability is expressed as

$$\alpha = 1 - \prod_{i=1}^r (1 - \alpha_i), \quad (5)$$

where

α = probability of Type-I error,

r = number of detection rules used,

α_i = probability of Type-I error of the i th rule, for $i = 1, 2, \dots, r$.

The r detection rules in (5) are assumed to be independent of one another.

Based on the four detection rules given in Table 1, [14] had estimated the occurrence of a Type-I error probability in the first k subgroups. The following explains the four detection rules considered by [14]:

Rule 1 When one or more points fall beyond the $+3\sigma$ control limit, a process is declared as out-of-control.

Table 1 Type-I error probabilities for numerous combinations of detection Rules 1-4 [14]

Detection rules	Number of subgroups									
	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$	$k = 8$	$k = 9$	$k = 10$
1	0.003	0.005	0.008	0.011	0.013	0.016	0.019	0.021	0.024	0.027
1 and 2	0.003	0.006	0.011	0.015	0.020	0.024	0.028	0.034	0.039	0.043
1, 2 and 3	0.003	0.006	0.011	0.016	0.025	0.032	0.040	0.050	0.060	0.060
1, 2, 3 and 4	0.003	0.006	0.011	0.016	0.025	0.032	0.040	0.060	0.070	0.080

- Rule 2 When 2 of 3 successive points fall in the same zone and these points are between the $+2\sigma$ and $+3\sigma$ control limits, a process is declared as out-of-control.
- Rule 3 When 4 of 5 consecutive points fall in the same zone and these points are beyond the $+1\sigma$ control limit, a process is declared as out-of-control.
- Rule 4 When 8 consecutive points fall in the same zone, either above or below the CL, a process is declared as out-of-control.

The implementation of more detection rules will enhance the detection power of a control chart but at the expense of inflating the probability of the overall Type-I error.

4 Shewhart $\bar{X}$ Chart's Probability Tables

A detailed discussion regarding the formulae to compute the probabilities of several combinations of detection rules applied on the Shewhart $\bar{X}$ chart is provided in this section. We will also describe the probability table proposed by [14].

4.1 *Formulae and Computation of Probabilities for the Power Functions*

The following defines the symbols that will be used in this study:

k = number of subgroups.

$PDS(k)$ = probability of detecting an off-target signal within k subgroups.

p_k = probability of detecting an off-target signal at the k th subgroup.

Note that the subsequent discussions still consider the four detection rules employed by [14] as described in Sect. 3.

4.1.1 Using Detection Rule 1

Let “ e ” represents the probability that an $\bar{X}$ sample falls beyond the $+3\sigma$ limit (Fig. 1).

Wheeler [14] gave the following fundamental formula for the detection of an off-target signal within k subgroups:

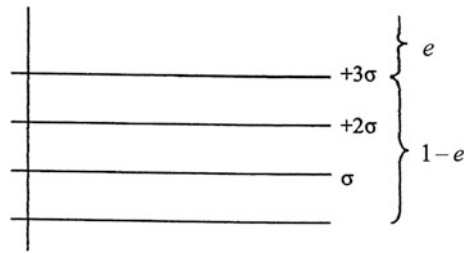


Fig. 1 Probabilities of an $\bar{X}$ sample on the Shewhart $\bar{X}$ chart, using detection Rule 1

$$\begin{aligned}
 PDS(k) &= \sum_{i=1}^k p_i \\
 &= e + e(1-e) + e(1-e)^2 + e(1-e)^3 + \cdots + e(1-e)^{k-1} \\
 &= \sum_{i=1}^k e(1-e)^{i-1} \\
 &= 1 - (1-e)^k.
 \end{aligned} \tag{6}$$

The case discussed here is for the upper sided control chart.

4.1.2 Using Detection Rules 1 and 2

Let

e = the probability that an $\bar{X}$ sample plots beyond the $+3\sigma$ limit,

f = the probability that an $\bar{X}$ sample plots between the $+2\sigma$ and the $+3\sigma$ limits,

g = the probability that an $\bar{X}$ sample plots between the CL and the $+2\sigma$ limit (Fig. 2).

At least 2 subgroups are needed in performing the analysis if detection Rule 2 is considered. We can use (6) to compute the probability, except for cases where $k = 2$ or more.

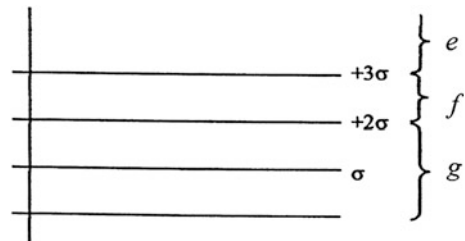


Fig. 2 Probabilities of an $\bar{X}$ sample on the Shewhart $\bar{X}$ chart, using detection Rules 1 and 2

For case $k = 2$, p_2 and $PDS(2)$ are [14]

$$p_2 = ge + fe + f^2 \quad (7)$$

and

$$\begin{aligned} PDS(2) &= p_1 + p_2 \\ &= e + ge + fe + f^2 \\ &= e + f^2 + (f + g)e. \end{aligned} \quad (8)$$

For case $k = 3$, p_3 and $PDS(3)$ can be expressed as [14]

$$p_3 = 2efg + 2f^2g + eg^2 \quad (9)$$

and

$$\begin{aligned} PDS(3) &= p_1 + p_2 + p_3 \\ &= e + f^2 + (f + g)e + 2efg + 2f^2g + eg^2 \\ &= e + f^2 + eg + ef(1 + 2g) + 2f^2g + eg^2. \end{aligned} \quad (10)$$

Using the same method, p_k , for $k = 4, 5, \dots, 10$ can be obtained using the following equations [14]:

$$p_4 = 3efg^2 + 2f^2g^2 + eg^3. \quad (11)$$

$$p_5 = ef^2g^2 + f^3g^2 + 4efg^3 + 2f^2g^3 + eg^4. \quad (12)$$

$$p_6 = 3ef^2g^3 + 3f^3g^3 + 5efg^4 + 2f^2g^4 + eg^5. \quad (13)$$

$$p_7 = 6ef^2g^4 + 5f^3g^4 + 6efg^5 + 2f^2g^5 + eg^6. \quad (14)$$

$$p_8 = ef^3g^4 + f^4g^4 + 10ef^2g^5 + 5f^3g^5 + 7efg^6 + 4f^2g^6 + eg^7. \quad (15)$$

$$p_9 = 4ef^3g^5 + 4f^4g^5 + 15ef^2g^6 + 9f^3g^6 + 8efg^7 + 2f^2g^7 + eg^8. \quad (16)$$

$$p_{10} = 11ef^3g^6 + 10f^4g^6 + 21ef^2g^7 + 10f^3g^7 + 8efg^8 + 2f^2g^8 + eg^9. \quad (17)$$

4.1.3 Using Detection Rules 1, 2 and 3

Let

e = the probability that an $\bar{X}$ sample plots beyond the $+3\sigma$ limit,

f = the probability that an $\bar{X}$ sample plots between the $+2\sigma$ and the $+3\sigma$ limits,

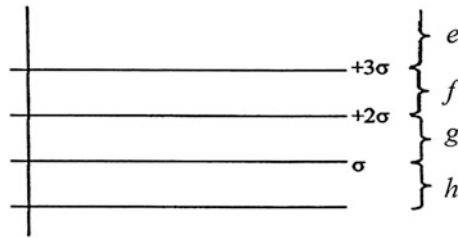


Fig. 3 Probabilities of an $\bar{X}$ sample on the Shewhart $\bar{X}$ chart, using detection Rules 1, 2 and 3

g = the probability that an $\bar{X}$ sample plots between the $+\sigma$ and the $+2\sigma$ limits,

h = the probability that an $\bar{X}$ sample plots between the CL and the $+\sigma$ limit (Fig. 3)

At least 4 subgroups are required in the analysis when using detection Rule 3. The probabilities obtained are the same as the equations given in Sect. 4.1.2, except for cases where $k = 4$ or more. When $k \geq 4$, p_k is obtained using the following equations [14]:

$$p_4 = 3efg^2 + 6efgh + 3efh^2 + 3eg^2h + 3egh^2 + eg^3 + eh^3 + 3f^2h^2 + 4f^2gh + 2f^2h^2 + g^4 + 4fh^3. \quad (18)$$

$$p_5 = 2ef^2gh + 2f^3gh + 12efg^2h + 14f^2g^2h + 16fg^3h + 4eg^3h + 4g^4h + ef^2h^2 + f^3h^2 + 12efgh^2 + 6f^2gh^2 + 6ef^2h^2 + 4efh^3 + 2f^2h^3 + 4egh^3 + eh^4. \quad (19)$$

$$p_6 = 9ef^2gh^2 + 9f^3gh^2 + 30efg^2h^2 + 20f^2g^2h^2 + 10eg^3h^2 + 16fg^3h^2 + 4g^4h^2 + 3ef^2h^3 + 3f^3h^3 + 20efgh^3 + 8f^2gh^3 + 10eg^2h^3 + 5efh^4 + 2f^2h^4 + 5egh^4 + eh^5. \quad (20)$$

$$p_7 = 19ef^2g^2h^2 + 18f^3g^2h^2 + 24efg^3h^2 + 25f^2g^3h^2 + 15fg^4h^2 + 6eg^4h^2 + 3g^5h^2 + 24ef^2gh^3 + 20f^3gh^3 + 60efg^2h^3 + 28f^2g^2h^3 + 20ef^3g^3 + 16fg^3h^3 + 4g^4h^3 + 6ef^2h^4 + 5f^3h^4 + 30efgh^4 + 10f^2gh^4 + 15eg^2h^4 + 6efh^5 + 2f^2h^5 + 6egh^5 + eh^6. \quad (21)$$

Due to the complexity, the Monte Carlo simulation is employed in the computation of the probabilities p_8 , p_9 and p_{10} .

4.1.4 Using Detection Rules 1, 2, 3 and 4

When detection Rule 4 is considered, at least 8 subgroups are called for in the analysis. Thus, the probabilities obtained are similar to that in the equations given in Sect. 4.1.3, except for cases where $k = 8$ or more. Owing to the complexity of the equations [14], reported that a total of 12,000 combinations are needed in the computation for detecting a shift. For this reason, the Monte Carlo simulation method is used in place of the formulae to calculate the probabilities for $k = 8, 9$ and 10.

4.2 Probability Tables

Wheeler [14] introduced a probability table obtained using the formulae. The probability table provides the probability of detecting sustained process mean shifts. The probability table includes four major characteristics defined as follows:

- Shift size:
The probability table provides a total of 11 sizes of shifts. They are 0.42σ , 0.67σ , 0.95σ , 1.25σ , 1.52σ , 1.72σ , 1.96σ , 2.16σ , 2.48σ , 2.75σ and 3.00σ .
- Number of subgroups, k :
The number of subgroups, $k = 1$ to $k = 10$ are shown in the probability table.
- Subgroup size, n :
Different subgroup sizes are indicated in the probability table of [14]. They are $n = 1$ to $n = 10$, $n = 12$, $n = 15$ and $n = 20$.
- Detection rules:
The necessary probabilities are computed based on several combinations of detection rules.

5 A Study on the Performance

The Statistical Analysis System (SAS) software is used in this study. The SAS computer programs are developed according to the framework of the probability table and the formulae (see also Sects. 4.1.1, 4.1.2, 4.1.3 and 4.1.4) given in [14], to compute the probabilities of detecting a sustained mean shift within particular number of subgroups following the shift. The results obtained are tabulated in Tables 2 and 3.

As can be observed from Tables 2 and 3,

- the probability of detecting a shift in the process mean becomes larger as the size of the shift increases, regardless of the values of n and k , and the combination of the detection rules being used.

- for all n , sizes of shift and combinations of the detection rules, the detection probability within k subgroups rises as k increases.
- The implementation of more detection rules can increase the sensitivity of a control chart, for any k , n and sizes of shifts.

A detailed comparison between Tables 2 and 3 shows that the probabilities in Table 2 are generally less than the corresponding probabilities in Table 3. This indicates that an increase in the subgroup size will result in an enhancement in the power of a control chart, irrespective of the value of k , shift size and detection rules being used.

The difference in terms of percentage is computed to enable us to make a more meaningful comparison between Wheeler's formulae method [14] and the Monte Carlo simulation method. Each probability table is compared in terms of percentages and the percentage difference can be computed as follows:

$$\text{Difference in terms of percentage} = \left| \frac{u - v}{v} \right| \times 100 \%, \quad (22)$$

where

u = probability obtained via Monte Carlo simulation,

v = probability calculated via formulae.

Tables 4 and 5 show the results for the percentage difference computed using (22). From Tables 4 and 5, we can observe that the probabilities computed by Monte Carlo simulation are quite closed to that obtained from the formulae approach. The existence of differences in the probabilities between the two methods are indicated by the boldfaced entries in Tables 4 and 5. It is very interesting to point out that the differences in the probabilities between the two methods are small. Indeed, for practical applications, the magnitude of these differences can be omitted.

6 Concluding Remarks

In SPC, the Shewhart $\bar{X}$ chart is a very popular process monitoring tool. The implementation of this chart is relatively simple. To enable a fast and effective design and implementation of the Shewhart $\bar{X}$ chart so that a more efficient system of process monitoring can be achieved, Tables 2, 3, 4 and 5 are provided. In fact, the approach by means of formulae is rather cumbersome. When detection Rule 4 is considered (see Sect. 4.1.4), the computation becomes very complicated which increases the chance of computation error. The findings of this study reveal that the Monte Carlo simulation method can favorably replace the formulae approach given by [14]. This is because the Monte Carlo simulation method allows the quality users to calculate the probabilities of detecting a sustained process mean shift within k subgroups, for the various detection rules more easily and quickly with the same accuracy.

References

1. Berger RW, Hart T (1986) Statistical process control, a guide for implementation. Marcel Dekker Inc., New York
2. Grant EL, Leavenworth RS (1980) Statistical quality control, 5th edn. McGraw-Hill Book Company, New York
3. Wu Z, Spedding TA (2000) A synthetic control chart for detecting small shifts in the process mean. *J Qual Technol* 32(1):32–38
4. Wu Z, Ou Y, Castagliola P, Khoo MBC (2010) A combined synthetic & $\bar{X}$ chart for monitoring the process mean. *Int J Prod Res* 48(24):7423–7436
5. Khoo MBC, Lee HC, Wu Z, Chen CH, Castagliola P (2011) A synthetic double sampling control chart for the process mean. *IIE Trans* 43(1):23–38
6. Stoumbos ZG, Reynolds MR Jr (2001) The SPRT control chart for the process mean with samples starting at fixed times. *Nonlinear Anal* 2(1):1–34
7. Roberts SW (1959) Control chart based on geometric moving averages. *Technometrics* 1(3):239–250
8. Khoo MBC, Wu Z, Chen CH, Yeong KW (2010) Using one EWMA chart to jointly monitor the process mean and variance. *Comput Stat* 25(2):299–316
9. Khoo MBC, Teh SY, Wu Z (2010) Monitoring process mean and variability with one double EWMA chart. *Commun Stat Theor M* 39(20):3678–3694
10. Lim TJ, Cho M (2009) Design of control charts with m-of-m runs rules. *Qual Reliab Eng Int* 25(8):1085–1101
11. Kim YB, Hong JS, Lie CH (2009) Economic-statistical design of 2-of-2 and 2-of-3 runs rule scheme. *Qual Reliab Eng Int* 25(2):215–228
12. Zarandi MHF, Alaeddini A, Turksen IB (2008) A hybrid fuzzy adaptive sampling—run rules for Shewhart control charts. *Inf Sci* 178(4):1152–1170
13. Antzoulakos DL, Rakitzis CA (2008) The modified r out of m control chart. *Commun Stat Simul Comput* 37(2):396–408
14. Wheeler DJ (1983) Detecting a shift in process average—tables of the power function for $\bar{X}$ charts. *J Qual Technol* 15(4):155–170
15. Ryan TP (1989) Statistical methods for quality improvement. Wiley, New York
16. Montgomery DC (2001) Introduction to statistical quality control, 4th edn. Wiley, New York

B-Spline Collocation with Domain Decomposition Method and Its Application for Singularly Perturbed Convection-Diffusion Problems

Mas Irfan P. Hidayat, Bambang Ariwahjoedi
and Setyamartana Parman

Abstract Global collocation method using B-spline basis functions is shown to be capable for solving elliptic partial differential equations in arbitrary complex domains. The global B-spline collocation approach is effectively alleviating difficulties commonly associated to B-spline based methods in handling such domains. Nonetheless, the global method, which is simply reduced to Bezier approximation of degree p with C^0 continuity, has led to the use of B-spline basis of high order in order to achieve high accuracy. The need for B-spline bases of high order is also more prominent in domains of large dimension. In addition, it may also lead to the ill-conditioning problem for combination of the use of B-spline bases of high order and increasing number of collocation points. In this chapter, global B-spline collocation scheme with domain decomposition techniques is introduced for solving Poisson equations in arbitrary complex domains. Overlapping Schwarz multiplicative and additive domain decomposition techniques are examined in this study. It is shown that the combination method produces higher accuracy with the B-spline basis of much lower order than that needed in implementation of the global method. The B-spline collocation with domain decomposition method hence improves the approximation stability of the global B-spline collocation method. Numerical simulations of singularly perturbed convection-diffusion problems are presented to further show the method efficacy and capability.

M.I.P. Hidayat (✉) · S. Parman

Departement of Mechanical Engineering, Universiti Teknologi PETRONAS,
Bandar Seri Iskandar, 31750 Tronoh, Perak Darul Ridzuan, Malaysia
e-mail: irfan@mat-eng.its.ac.id

S. Parman

e-mail: setyamartana@petronas.com.my

B. Ariwahjoedi

Department of Fundamental and Applied Science, Universiti Teknologi PETRONAS,
Bandar Seri Iskandar, 31750 Tronoh, Perak Darul Ridzuan, Malaysia
e-mail: bambang_ariwahjoedi@petronas.com.my

1 Introduction

The method of collocation is attractive for solving differential equations due to its ease of implementation and efficiency. The method applies directly to the original differential equations that describe the physics of the problem. In the last decades, collocation methods using B-spline functions for the numerical solutions of various types of partial differential equations have been paid attention and becoming an active research area. The spline collocation methods are interesting due to their high accuracies. There are many variants of spline collocation methods, which are mainly related to the locations of the collocation points, such as nodal, orthogonal and modified collocation (see the survey of spline collocation methods by Fairweather and Meade [1], Bialecki and Fairweather [2]). Other variants of spline collocation methods have been investigated as well: Botella [3], Greville [4, 5] and Moment [6] collocations. The accuracy of the spline collocation methods depends on the order of B-spline basis, regularity or continuity, and location of the collocation points [3–6]. Recently, B-spline collocation methods have been applied successfully in various areas including fluid dynamics [3, 4, 7–9], computational aero-acoustics [10], wave equation [11], non-linear analysis of laminated panels [12], vibration analysis [13, 14], and non-linear two-point boundary value problems [15].

However, there are serious shortcomings in the implementation of the B-spline collocation methods. The fact that basis needed for approximation is a tensor product operation from one-dimensional B-spline basis functions has restricted the method extension to applications in higher dimensions. Moreover, the tensor product structure that is essentially ‘rectangular’ in nature has also further limited the approximation to rectangular problem domain or other domains in which a rectangular structure can be envisaged. Obviously, this produces a quite severe limitation with respect to the flexibility and suitability of the method for applications in non-rectangular or irregular domains. In [16], it has been stated that the B-spline collocation methods have not yet reached the level of flexibility of the isoparametric formulation of the finite element method for arbitrary geometries.

Global collocation method using B-spline basis functions is shown to be capable for solving elliptic partial differential equations in arbitrary complex domains. The implementation of global collocation has been inspired by successful application of radial basis functions (RBF) collocation in non-uniform grids [17]. The global B-spline collocation approach is effectively alleviating difficulties commonly associated to B-spline based methods in handling such domains. Nonetheless, the global method, which is simply reduced to Bezier approximation of degree p with C^0 continuity, has led to the use of B-spline basis of high order in order to achieve high accuracy. The need for B-spline bases of high order is also more prominent in domains of large dimension. In addition, it may also lead to the ill-conditioning problem for combination of the use of B-spline bases of high order and increasing number of collocation points.

In this chapter, global B-spline collocation scheme with domain decomposition techniques is introduced for solving Poisson equations in arbitrary complex

domains. Overlapping Schwarz multiplicative and additive domain decomposition techniques are examined in this study [18]. It is aimed that the B-spline approximation can be more robust, but still preserving great flexibility of the global method in dealing with arbitrary complex domains. It is shown that the combination method produces higher accuracy with the B-spline basis of much lower order than that needed in implementation of the global method. The B-spline collocation with domain decomposition method hence improves the approximation stability of the global B-spline collocation method. Numerical applications of singularly perturbed convection-diffusion problems are given to further show the method efficacy and capability.

2 Global B-Spline Collocation

Consider the following boundary value problem:

$$\begin{aligned} Lu(\mathbf{x}) &= f(\mathbf{x}) & \forall \mathbf{x} \in \Omega \\ B^h u(\mathbf{x}) &= h(\mathbf{x}) & \forall \mathbf{x} \in \partial\Omega^h \\ B^g u(\mathbf{x}) &= g(\mathbf{x}) & \forall \mathbf{x} \in \partial\Omega^g \end{aligned} \quad (1)$$

where Ω is the problem domain, $\partial\Omega^h$ is the Dirichlet boundary, $\partial\Omega^g$ is the Neumann boundary, $\partial\Omega = \partial\Omega^h \cup \partial\Omega^g$ is the problem domain boundary, L is the Laplace operator (∇^2) in Ω , B^h is the differential operator on $\partial\Omega^h$, B^g is the differential operator on $\partial\Omega^g$, $\mathbf{x} \in R^d$.

Let also $\{CP_i = (\xi_i)\}_{i=1}^{NC}$ be NC collocation points in Ω , of which $\{(\xi_i)\}_{i=1}^{N_I}$ are interior points, $\{(\xi_i)\}_{i=1}^{N_h}$ are boundary points on $\partial\Omega^h$ and $\{(\xi_i)\}_{i=1}^{N_g}$ are boundary points on $\partial\Omega^g$, hence $NC = N_I + N_h + N_g$.

The approximate solution $\hat{u}(\mathbf{x})$ for the problem (1) is to be obtained with the collocation using B-splines. B-spline basis construction is briefly described in what follows.

Consider one-dimensional B-spline basis function construction given by the Cox-de Boor recursion formula [19, 20]:

$$N_{i,1}(t) = \begin{cases} 1 & \text{if } \tau_i < t < \tau_{i+1} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

$$N_{i,1}(t) = \frac{t - t_i}{t_{i+k-1} - t_i} N_{i,k-1}(t) + \frac{t_{i+k} - t}{t_{i+k} - t_{i+1}} N_{i+1,k-1}(t) \quad \text{for } 2 \leq k \leq n+1 \quad (3)$$

In the formula, the convention $0/0 = 0$ is used for the division calculation. Thus, $b = (N_{0,k}, \dots, N_{n,k})$ are the univariate B-spline bases for the space C of polynomial functions of degree less than or equal to $k-1$ on interval $I = [a_1, b_1] = [\tau_0, \tau_{n+k}]$.

It is also clear from (3) that for $k > 1$ the $N_{i,k}(t)$ are a linear combination of two $(k - 1)$ order basis functions.

The relationship between the number of knots in the knot vector $(\tau_i)_{i=0}^{n+k}(mk)$, control points $(n + 1)$ and B-spline order k is also given as follows:

$$mk = n + k + 1 \quad (4)$$

For higher dimensional problems, the basis functions are constructed by taking the product of the one-dimensional B-spline basis functions. For approximation in 2D, the basis functions are given by:

$$\mathbf{B} := (N_{i,k} \otimes M_{j,l})_{i=0,j=0}^{n,m} \quad (5)$$

where $\mathbf{B}$ are the tensor product B-spline bases for the space D of tensor product polynomial splines on the tensor product domain $[a_1, b_1] \times [a_2, b_2] = [\tau_0, \tau_{n+k}] \times [\varsigma_0, \varsigma_{m+l}]$. $N_{i,k}$ and $M_{j,l}$ are the B-spline basis functions of orders k and l .

Note that the tensor product B-spline bases have the partition of unity property:

$$\sum_{i=0,j=0}^{n,m} N_{i,k}(t) M_{j,l}(s) = 1 \quad (6)$$

In addition, continuity of a B-spline curve depends generally on its order. A B-spline curve of order k has in general C^{k-2} continuity. There is however a continuity constraint with respect to the multiplicity (mp) of a knot in the knot vector. It is called as knot continuity, where increasing the multiplicity (mp) of a knot reduces the continuity of the curve to C^{k-mp-1} at that knot. It also means only $p - mp$ derivatives exist for a knot of multiplicity mp , where p is the B-spline degree [21].

In this study a special class of B-spline approximation obtained by using an open-uniform knot vector with full multiplicity of its end knots (no interior knots) is used for the collocation. Such a B-spline approximation is chosen because it will produce global approximation in order to deal with arbitrary complex domains with arbitrary distribution of nodes. It is noted here the B-spline approximation is simply reduced into Bezier approximation of degree p with C^0 continuity [22].

Thus, we have here for the collocation (1) the following two-dimensional B-spline approximation:

$$\hat{u}(x, y) = \sum_{i=0,j=0}^{n,m} N_{i,k}(x) M_{j,l}(y) \alpha_{ij} \quad (7)$$

The B-spline approximation can be simplified to be much similar with shape functions in mesh-less and finite element methods [23] as:

$$\hat{u}(x, y) = \sum_{a=1}^{N_B} N^a(x, y) \alpha_a \quad (8)$$

where $N^a(x, y) = N_{i,k}(x) M_{j,l}(y)$, α_a are the coefficients of approximation related to $N^a(x, y)$ and N_B is the number of B-spline bases used in approximation.

3 Overlapping Domain Decomposition Methods

The basic idea of domain decomposition method (DDM) is to decompose one large global problem into many smaller sub-domain problems [18, 24, 25]. Only overlapping domain decomposition method with Schwarz multiplicative and additive techniques are focused on this work. Other DDM techniques, such as non-overlapping ones, are left as extensions of the present work, while the readers are also directed to a dedicated website on DDM [25] containing proceedings of the international conferences on DDM together with numerous references cited therein.

By overlapping domain decomposition method, a given computational domain is partitioned into smaller overlapping sub-domains. A typical option to accomplish such a decomposition task is by partitioning the computational domain Ω into K non-overlapping sub-domains and then extending from the shared boundaries of neighbouring sub-domains Ω_i and Ω_j an amount of overlapping portions o_{ij} to a larger domain Ω_i . The artificial interior boundaries are hence introduced here. This schematic is illustrated in Fig. 1.

Denote Ω_i , $\partial\Omega_i \setminus \Gamma_i$ and Γ_i the extended sub-domain, the natural boundary and the artificial interior boundary overlapped with other neighbouring sub-domains, respectively. Let also $\bar{\Omega}_i = \Omega_i \cup \partial\Omega_i \setminus \Gamma_i \cup \Gamma_i$ denote the closed sub-domain, S denote the operator of artificial interior boundary conditions and Λ_i be the artificial interior boundary value of sub-domain $\bar{\Omega}_i$ extracted from the neighbouring sub-domains. The original global problem is now reformulated as a series of sub-domain problems:

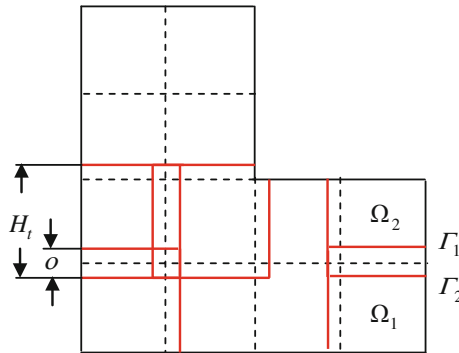


Fig. 1 Schematic of domain decomposition of Ω into smaller overlapping sub-domains Ω_i . The artificial boundaries Γ_i are part of the boundary of Ω_i that is interior of Ω

$$\begin{aligned}
Lu(\mathbf{x}) &= f_i && \text{in } \Omega_i \\
Bu(\mathbf{x}) &= b_i && \text{on } \partial\Omega_i \setminus \Gamma_i \\
Su(\mathbf{x}) &= \Lambda_i && \text{on } \Gamma_i
\end{aligned} \tag{9}$$

where B, S may specify Dirichlet, Neumann or mixed (Robin) boundary condition operator and $i = 1, 2, \dots, K$.

Communication and transmission of information among the sub-domains is carried out iteratively in the overlapping regions o_{ij} i.e. through the artificial interior boundaries Γ_i , which is described in the following.

Introduce the B-spline approximation in (8) as in the global B-spline collocation, the linear system for each sub-domain is obtained as follows:

$$\begin{bmatrix} \mathbf{A}_{\overline{\Omega}_i \setminus \Gamma_i} \\ \mathbf{A}_{\Gamma_i} \end{bmatrix} \boldsymbol{\alpha}_i = \begin{Bmatrix} \mathbf{F}_{\overline{\Omega}_i \setminus \Gamma_i} \\ \mathbf{F}_{\Gamma_i} \end{Bmatrix} \tag{10}$$

where $\mathbf{A}_{\overline{\Omega}_i \setminus \Gamma_i}$ are matrices representing the operators acting on $\overline{\Omega}_i$ excluding Γ_i , $\mathbf{F}_{\overline{\Omega}_i \setminus \Gamma_i}$ are the corresponding right-hand side vectors, in which the natural boundary conditions are imposed. The sub-matrices $\mathbf{A}_{\Gamma_i}$ and vectors $\mathbf{F}_{\Gamma_i}$ are produced from the artificial interior boundary conditions.

Solving further (10), it is obtained:

$$\boldsymbol{\alpha}_i = \mathbf{A}_i^{-1} \mathbf{F}_i \tag{11}$$

where $\mathbf{A}_i = [\mathbf{A}_{\overline{\Omega}_i \setminus \Gamma_i} \ \mathbf{A}_{\Gamma_i}]^T$, $\mathbf{F}_i = \{\mathbf{F}_{\overline{\Omega}_i \setminus \Gamma_i} \ \mathbf{F}_{\Gamma_i}\}^T$ and $\boldsymbol{\alpha}_i$ are the vector of coefficients of approximation of each sub-domain. Note that $\mathbf{F}_i$, containing the artificial interior boundary conditions, will continuously change with the updating results of its neighbouring sub-domains.

Recalling (8), the approximated values of the unknown function in sub-domain $\overline{\Omega}_i$ are obtained as:

$$\phi_i(\mathbf{x}) = G_i(\mathbf{x}) \boldsymbol{\alpha}_i \tag{12}$$

where $G_i(\mathbf{x}) = [N^a(\mathbf{x}_1), N^a(\mathbf{x}_2), \dots, N^a(\mathbf{x}_{Ns_i})]^T$ and Ns_i is the total number of collocation points in sub-domain $\overline{\Omega}_i$.

It can be seen that the approximated values of the unknown function in sub-domain $\overline{\Omega}_i$ are obtained in similar process as the approximation in global large domain, but involving update in the artificial interior boundary conditions. In addition, depending on how artificial interior boundary conditions $\mathbf{F}_i$ are updated, there are two types of overlapping domain decomposition methods i.e. additive Schwarz and multiplicative Schwarz methods. For clarity, both these overlapping domain decomposition techniques are described here.

In the overlapping additive Schwarz method, the artificial interior boundary values of each sub-domain problem at the $(n+1)$ th step are updated from the

results of the n th step. This is illustrated as follows: Let Ω be partitioned into two sub-domains Ω_1 and Ω_2 , where $\Omega_1 \cap \Omega_2 \neq \emptyset$. The additive Schwarz domain decomposition can be written as:

$$\begin{aligned} L\phi_1^n &= f && \text{in } \Omega_1 \\ B\phi_1^n &= b && \text{on } \partial\Omega_1 \setminus \Gamma_1 \\ \phi_1^n &= \phi_2^{n-1} && \text{on } \Gamma_1 \end{aligned} \quad (13)$$

$$\begin{aligned} L\phi_2^n &= f && \text{in } \Omega_2 \\ B\phi_2^n &= b && \text{on } \partial\Omega_2 \setminus \Gamma_2 \\ \phi_2^n &= \phi_1^{n-1} && \text{on } \Gamma_2 \end{aligned} \quad (14)$$

On the other hand, in the overlapping multiplicative Schwarz method, each sub-domain problem is solved sequentially or in series. Here, the artificial interior boundary values are soon updated after a sub-domain problem is solved (by (10) and (11)), where each sub-domain uses the most recent updates from its neighbouring sub-domains. The multiplicative Schwarz domain decomposition is then illustrated as:

$$\begin{aligned} L\phi_1^n &= f && \text{in } \Omega_1 \\ B\phi_1^n &= b && \text{on } \partial\Omega_1 \setminus \Gamma_1 \\ \phi_1^n &= \phi_2^{n-1} && \text{on } \Gamma_1 \end{aligned} \quad (15)$$

$$\begin{aligned} L\phi_2^n &= f && \text{in } \Omega_2 \\ B\phi_2^n &= b && \text{on } \partial\Omega_2 \setminus \Gamma_2 \\ \phi_2^n &= \phi_1^n && \text{on } \Gamma_2 \end{aligned} \quad (16)$$

Both the multiplicative and additive Schwarz domain decomposition can be expressed in terms of the Dirichlet artificial interior boundary conditions as follows [24]:

Algorithm (*Overlapping Domain Decomposition Method*)

$\phi \leftarrow$ Guess initial solution on the artificial boundaries

For $i = 1, 2, \dots$, until convergence

For $j = 1, 2, \dots, K$ Do:

Solve (4.3), get α_i

$\phi_i \leftarrow G_i \alpha_i$

If (Multiplicative method) Update $\mathbf{F}_i$ for each $\overline{\Omega}_i$ part of Ω from neighbouring domains

End Do

If (Additive method) Update all $\overline{\Omega}_i$ from previous steps

End Do

Note that the additive Schwarz algorithm lets a parallel implementation in the process of solving the sub-domain problems. The multiplicative Schwarz algorithm lacks of this parallel characteristic, however it should converge faster than the additive algorithm due to its direct update characteristic. Both the multiplicative and additive Schwarz methods are in fact similar to the classical block Gauss–Seidel iterative method and the block Jacobi iteration, respectively [26].

In implementations, index sets are used for clustering and numbering knot points in each sub-domain and the artificial interior boundaries. Colouring may also help the purpose. The index sets are also useful when local renumbering for knot points of sub-domains is considered. It provides a matrix or prolongation operator linking the local knot points to their global numbering/position in the original large domain.

It can be pointed out that using domain decomposition method, the B-spline approximation is to be applied for many smaller domains than the original domain. While the need for B-spline bases of high order in this global B-spline approximation would be more prominent in domains of large size or dimension, it may become apparent that the necessity would be less due to the reduced dimension for each sub-domain. This will be verified through the numerical experiments.

4 Numerical Examples

Numerical examples are presented here to examine the effectiveness of B-spline collocation method with domain decomposition techniques. To clearly assess the proposed method performance, B-spline approximations in applications of one global domain and decomposed domain are examined and the obtained numerical results are compared. Comparison is made in terms of *B-spline order*, *accuracy* and *elapsed time*. For measure of accuracies, the following L^2 error is used:

$$E_n = \sqrt{\frac{\sum_{i=1}^{NC} [u(\mathbf{x}_i) - \hat{u}(\mathbf{x}_i)]^2}{\sum_{i=1}^{NC} u(\mathbf{x}_i)^2}} \quad (17)$$

In the present study, B-spline basis functions are computed directly rather than symbolically such as in [27], thus the bases computation is faster. All numerical simulations are performed with MATLAB 2006 environment and ran on HP Compaq with OS Windows XP, processor of Intel Pentium 4 512 MB RAM. Such computer specifications are chosen in order to clearly observe gains obtained by the implementation of domain decomposition techniques. A fixed error tolerance 1×10^{-7} and zero initial guess on the artificial interior boundaries are used in all numerical simulations.

4.1 Poisson Equation on L-Domain

Consider the Poisson equation:

$$\begin{aligned} \nabla^2 u = & -\frac{751\pi^2}{144} \sin\left(\frac{\pi x}{6}\right) \sin\left(\frac{7\pi x}{4}\right) \sin\left(\frac{3\pi y}{4}\right) \sin\left(\frac{5\pi y}{4}\right) \\ & + \frac{7\pi^2}{12} \cos\left(\frac{\pi x}{6}\right) \cos\left(\frac{7\pi x}{4}\right) \sin\left(\frac{3\pi y}{4}\right) \sin\left(\frac{5\pi y}{4}\right) \\ & + \frac{15\pi^2}{8} \sin\left(\frac{\pi x}{6}\right) \sin\left(\frac{7\pi x}{4}\right) \cos\left(\frac{3\pi y}{4}\right) \cos\left(\frac{5\pi y}{4}\right) \end{aligned} \quad (18)$$

which is applied on the L-shaped domain as shown in Fig. 2 and has the exact solution as follows:

$$u(x, y) = \sin\left(\frac{\pi x}{6}\right) \sin\left(\frac{7\pi x}{4}\right) \sin\left(\frac{3\pi y}{4}\right) \sin\left(\frac{5\pi y}{4}\right) \quad (19)$$

For clarity, exact solution for the Poisson equation on L-domain is further depicted in Fig. 3. As can be seen, the problem shows ‘waving’ solution on the L-domain with peak values near the corner point (2, 2).

Note that the number of sub-domains of 3 and 4 are respectively used in the L-shaped domain for the implementations of domain decomposition method. Several numerical simulations are carried out to examine the influence of the number of sub-domains as well as that of the amount of overlapping.

Table 1 first describes the B-spline approximation on one global L-domain. The number of collocation points and the B-spline order are varied to assess numerical performances of the global B-spline collocation. The numerical performances are described in terms of simulation error and elapsed time.

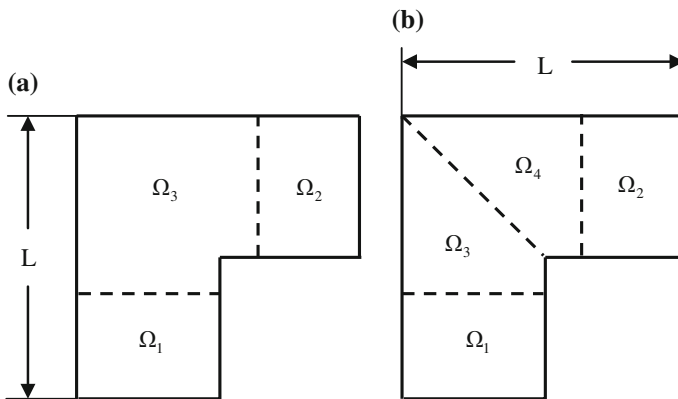


Fig. 2 L-shaped domain examined with: **a** three overlapping sub-domains, and **b** four overlapping sub-domains $L = 4$

Fig. 3 Exact solution of the Poisson equation on L-domain

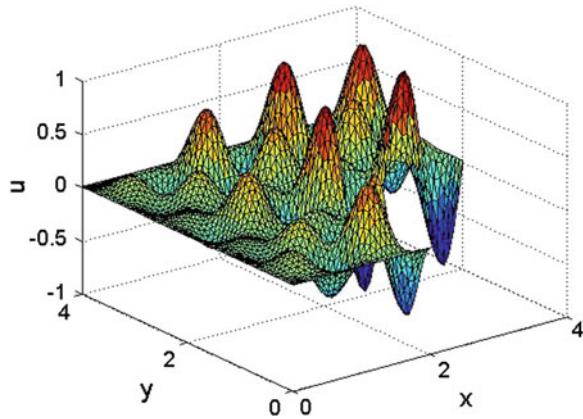


Table 1 Numerical accuracy for the Poisson equation on L-domain obtained by global B-spline collocation method

Number of collocation points	Number of domains	B-spline order k	L^2 error	Elapsed time (s)
721	1	23	5.086(-2)	2.60
		25	2.759(-2)	3.03
		27	4.634(-3)	3.40
1,791	1	23	2.250(-4)	5.59
		25	1.971(-5)	6.70
		27	1.757(-6)	7.91

In Table 1, it is shown that for the same B-spline order, the accuracy of B-spline approximation increases as the number of collocation points is increased. It is also obtained that for a fixed number of collocation points, the use of higher B-spline order will result in better accuracy in the approximation. The corresponding simulation elapsed times are also shown. It is observed that while the global B-spline collocation method effectively solves the PDE problems in irregular domain with arbitrary nodes, it leads to the use of B-spline bases of high orders. This may be explained from the general continuity posed by the global B-spline approximation which is C^0 only. It can be stressed out nonetheless that the global method does pose high accuracy and stable for solving Poisson equation in arbitrary domain.

Tables 2 and 3 describe the influence of the amount of overlapping in the global B-spline collocation with domain decomposition techniques for number of sub-domains 3 and 4, respectively.

Tables 4 and 5 describe the influence of the number of sub-domains in the global B-spline collocation with Schwarz multiplicative and additive decomposition techniques, respectively. Simulation results obtained with overlapping ratio 12 % are chosen.

Table 2 Influence of the amount of overlapping in B-spline collocation with Schwarz multiplicative and additive techniques

Number of nodes	Number of sub-domains	Overlapping ratio α/H_i (%)	Multiplicative			Additive		
			B-spline order k	L^2 error	Iteration steps	Elapsed time (s)	L^2 error	Iteration steps
721	3	6.3	16	8.593(-3)	36	3.61	8.583(-3)	50
			18	2.754(-3)	30	4.11	2.752(-3)	50
			20	1.716(-3)	34	5.46	1.712(-3)	50
		12	16	2.731(-2)	11	2.40	2.731(-2)	23
			18	8.289(-3)	12	2.74	8.288(-3)	25
			20	3.226(-3)	20	4.37	3.226(-3)	42
1,791	3	6.3	16	9.484(-4)	30	8.57	9.481(-4)	50
			18	1.917(-5)	27	11.15	1.946(-5)	44
			20	1.943(-6)	33	18.90	4.869(-6)	50
		12	16	7.400(-4)	12	5.35	7.429(-4)	24
			18	4.130(-5)	12	7.17	3.751(-5)	22
			20	2.330(-6)	14	10.69	2.361(-6)	24

Number of sub-domains is 3

Table 3 Influence of the amount of overlapping in B-spline collocation with Schwarz multiplicative and additive techniques

Number of nodes	Number of sub-domains	Overlapping ratio α/H_i (%)	Multiplicative			Additive		
			B-spline order k	L^2 error	Iteration steps	Elapsed time (s)	L^2 error	Iteration steps
721	4	6.3	16	4.426(-2)	34	3.60	4.425(-2)	50
			18	2.354(-2)	42	5.01	2.354(-2)	50
			20	1.904(-2)	47	6.61	1.914(-2)	50
		12	16	1.757(-2)	17	2.85	1.757(-2)	35
			18	7.902(-3)	19	3.51	7.902(-3)	39
			20	8.032(-3)	21	4.27	8.032(-3)	42
1,791	4	6.3	16	1.897(-4)	23	7.81	1.898(-4)	50
			18	1.460(-5)	29	11.87	1.609(-5)	50
			20	4.460(-6)	34	19.18	1.515(-5)	50
		12	16	1.928(-4)	23	6.07	3.447(-4)	26
			18	1.740(-5)	27	9.40	1.788(-5)	36
			20	5.283(-6)	33	13.15	2.552(-6)	31

Number of sub-domains is 4

Table 4 Influence of the number of sub-domains in B-spline collocation with domain Schwarz multiplicative decomposition technique

Number of collocation points	Number of sub-domains	Overlapping ratio α/H_t (%)	B-spline order k	L^2 error	Iteration steps	Elapsed time (s)
721	3	12	16	2.731(-2)	11	2.40
			18	8.289(-3)	12	2.74
			20	3.226(-3)	20	4.37
	4	12	16	1.757(-2)	17	2.85
			18	7.902(-3)	19	3.51
			20	8.032(-3)	21	4.27
1,791	3	12	16	7.400(-4)	12	5.35
			18	4.130(-5)	12	7.17
			20	2.330(-6)	14	10.69
	4	12	16	1.928(-4)	23	6.07
			18	1.740(-5)	27	9.40
			20	5.283(-6)	33	13.15

Table 5 Influence of the number of sub-domains in B-spline collocation with domain Schwarz additive decomposition technique

Number of collocation points	Number of sub-domains	Overlapping ratio α/H_t (%)	B-spline order k	L^2 error	Iteration steps	Elapsed time (s)
721	3	12	16	2.731(-2)	23	3.09
			18	8.288(-3)	25	3.94
			20	3.226(-3)	42	7.07
	4	12	16	1.757(-2)	35	3.87
			18	7.902(-3)	39	5.15
			20	8.032(-3)	42	6.57
1,791	3	12	16	7.429(-4)	24	7.93
			18	3.751(-5)	22	10.48
			20	2.361(-6)	24	15.49
	4	12	16	3.447(-4)	26	8.31
			18	1.788(-5)	36	15.70
			20	2.552(-6)	31	19.75

From Tables 1, 2, 3, 4 and 5, some important features of global B-spline collocation combined with domain decomposition techniques can be summarized as follows:

- (1) As the amount of overlapping increases, the number of iteration steps needed in the Schwarz multiplicative and additive techniques decreases. This may be understood that more overlapping requires more information exchange, thus allowing faster convergence.

- (2) It can be seen that compared with the global collocation method, the B-spline order is greatly reduced in the B-spline collocation with domain decomposition, thus avoiding ill-conditioning problem and producing increased accuracies. A decrease of the B-spline order up to seven was observed in the combination technique. For instant: B-spline collocation with domain decomposition with the number of collocation points 721, overlapping ratio 6.3 % and 3 sub-domains employed the B-spline order of 16, in comparison to the global B-spline collocation of order 23.
- (3) It is observed that in general more overlapping may lead to the increase of accuracy, in particular with respect to the use of higher number of collocation points or more number of sub-domains. In addition, more sub-domains can lead also to the increase of accuracy.
- (4) Schwarz additive decomposition technique takes longer time than Schwarz multiplicative technique in the present implementation. It is indeed observed that Schwarz multiplicative technique converges faster than the Schwarz additive algorithm due to its direct update characteristic. Nevertheless, Schwarz additive technique poses clear advantage in parallel computation, by which true potential of the additive technique can be seen and exploited.
- (5) It is observed that both Schwarz multiplicative and additive techniques produced similar accuracies in the present implementation of global B-spline collocation with domain decomposition techniques.
- (6) The use of more collocation points in the B-spline collocation combined with domain decomposition can increase the accuracy as that observed in the global B-spline collocation. Note that the use of more collocation points now becomes less costly as the B-spline order has been greatly reduced. Nevertheless, it should be pointed out also that the advantages of using domain decomposition techniques would become not too apparent when using much higher number of collocation points in particular with the use of small numbers of sub-domains. In such case, 'local' nature of domain decomposition may smear out and instead 'global' nature for each sub-domain becomes more apparent.

Showing high accuracies in solving Poisson equation in arbitrary domain, in the following section the B-spline collocation with domain decomposition techniques is applied for singularly perturbed convection-diffusion problems. Convection-diffusion equations may represent transient and steady heat conduction problems which may be problems of interest in many industrial and technological applications such as electronic cooling, heat transfer between pipe of vapour transport and fin of triangular or circular boundaries, encapsulation using functionally graded materials, and cryogenics [28].

5 Numerical Simulations of Singularly Perturbed Convection-Diffusion Problems

Singularly perturbed advection-diffusion-reaction problem is represented by:

$$\begin{aligned} -\varepsilon \nabla^2 u + \vec{b} \cdot \nabla u + cu &= f(x, y, \varepsilon) \quad \text{in } \Omega \\ u &= g \quad \text{on } \partial\Omega \\ \vec{b} &= (b_1, b_2), \quad b_1 \geq 0, b_2 \geq 0, c \geq 0 \quad \text{on } \partial\Omega \end{aligned} \quad (20)$$

where ε is a small diffusion parameter, b_1 , b_2 and c are constants, g is the Dirichlet boundary condition and $\partial\Omega$ is the boundary of Ω .

Such problem is also challenging to solve due to the presence of typical very thin boundary layers in the solution and thus it is involving multi-scales. Such sharp transitions and the multi-scale properties make the problem is often used as testing model for new numerical methods or algorithms.

5.1 Singularly Perturbed Problem

As the first application of the present method for the singularly convection-diffusion, the following problem is considered [24]:

$$\begin{aligned} -\varepsilon^2 \nabla^2 u + 2u &= f(x, y, \varepsilon) \quad \text{in } \Omega \\ u &= 0 \quad \text{on } \partial\Omega \end{aligned} \quad (21)$$

Here, f is chosen in such a way that the exact solution is given by:

$$u(x, y) = \left(1 - \frac{e^{-x/\varepsilon} + e^{-(1-x)/\varepsilon}}{1 + e^{-1/\varepsilon}}\right) \left(1 - \frac{e^{-y/\varepsilon} + e^{-(1-y)/\varepsilon}}{1 + e^{-1/\varepsilon}}\right) + x(1-x)y(1-y) \quad (22)$$

Problem domain is the L-domain as Fig. 2, but with $L = 1$. Three small overlapping sub-domains are employed to solve this problem. The value of ε is chosen as 0.01 and the iteration is considered to be convergent when the ratio of the difference of two iterations to the latest result is $< 1.0 \times 10^{-3}$.

Table 6 presents the obtained L^2 errors for this problem examined with 721 and 1,791 collocation points, respectively. All simulations are complete within three iterations.

Figure 4 depicts the computed solution and absolute or pointwise errors for the singularly perturbed problem obtained by using 721 collocation points with the B-spline order of 12. The maximum pointwise error value is 9.936E-2.

Table 6 Numerical accuracy for the singularly perturbed problem on L-domain by B-spline collocation with Schwarz multiplicative decomposition method

Number of collocation points	Number of sub-domains	B-spline order k	L^2 error
721	3	8	5.161(-2)
		10	2.144(-2)
		12	1.531(-2)
1,791	3	8	6.745(-2)
		10	3.248(-2)
		12	2.243(-2)

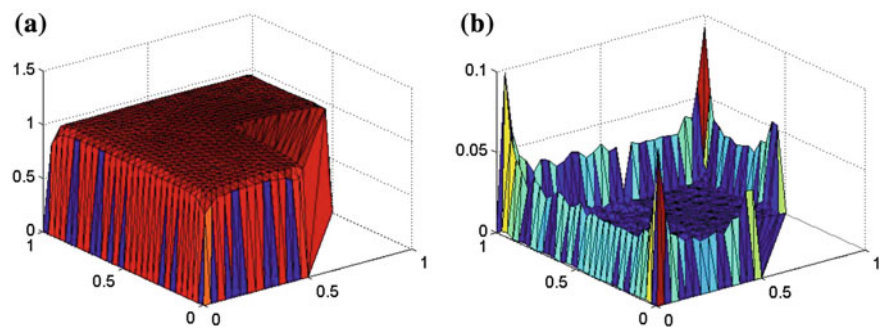


Fig. 4 **a** Computed solution and **b** pointwise errors for the singularly perturbed problem using 721 collocation points and the B-spline order of 12

The use of 4 sub-domains for numerical simulation of this problem produced similar accuracies. Note that the use of small numbers of sub-domains with small overlapped areas for the present B-spline collocation with domain decomposition techniques has shown sufficient good accuracies for this problem. The use of more sub-domains of eight or nine may be expected to produce further increase in accuracy.

The obtained results for the singularly perturbed problem hence show the capability and effectiveness of the present B-spline collocation with domain decomposition method.

5.2 Convection-Dominated Problem

The convection-diffusion problem with the convection-dominated term is considered as the next application, which is more challenging problem. The typical boundary layers will be clearly observed in the solutions of this problem.

$$\begin{aligned} -\varepsilon \nabla^2 u + \vec{b} \cdot \nabla u &= f(x, y, \varepsilon) \quad \text{in } \Omega \\ u &= g \quad \text{on } \partial\Omega \end{aligned}$$

(23)

Here, f is given by:

$$f(x,y) = (x+y)\left(1 - e^{-(1-x)/\varepsilon}e^{-(1-y)/\varepsilon}\right) + (x-y)\left(e^{-(1-y)/\varepsilon} - e^{-(1-x)/\varepsilon}\right) \tag{24}$$

This problem was presented in [29]. The exact solution is:

$$u(x,y) = xy\left(1 - e^{-(1-x)/\varepsilon}\right)\left(1 - e^{-(1-y)/\varepsilon}\right) \tag{25}$$

Two values ε of 0.1 and 0.01 are examined in the present study. $\vec{b} = (1, 1)$ are chosen. As in the previous problem 5.1, 3 small overlapping sub-domains are employed to solve this problem. The iteration is considered to be convergent also when the ratio of the difference of two iterations to the latest result is $<1.0 \times 10^{-3}$.

Figure 5 depicts the exact solutions of this convection-dominated problem for the two values parameter ε of 0.1 and 0.01. Note that sharper boundary layers are observed as the parameter ε becomes smaller i.e. convection term is more dominance.

Higher numbers of collocation points are required to solve this problem, thus it is examined with 1,791 and 2,757 collocation points. Tables 7 and 8 present the obtained L^2 errors and the iteration steps for the values ε of 0.1 and 0.01, respectively.

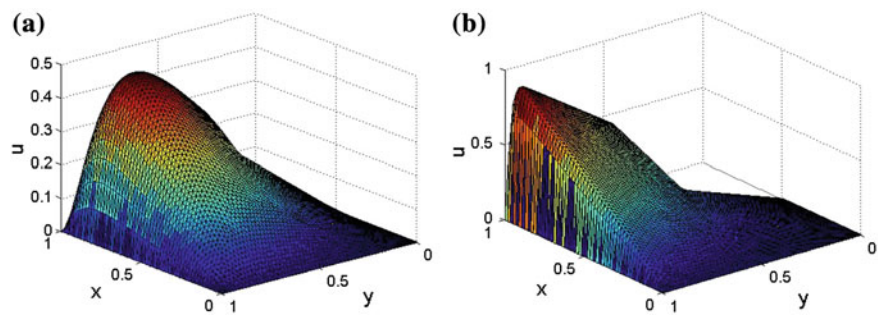


Fig. 5 Exact solutions of the convection-dominated problem 5.2 for two values ε of **a** 0.1, and **b** 0.01, respectively

Table 7 Numerical accuracy for the convection-dominated problem on L-domain by B-spline collocation with Schwarz multiplicative decomposition method ($\varepsilon = 0.1$)

Number of nodes	Number of sub-domains	B-spline order k	L^2 error	Iteration steps
1,791	3	8	9.529 (−3)	26
		10	1.467 (−4)	24
		12	2.486 (−6)	32
2,757	3	8	1.103 (−2)	17
		10	1.840 (−4)	22
		12	2.958 (−6)	32

Table 8 Numerical accuracy for the convection-dominated problem on L-domain by B-spline collocation with Schwarz multiplicative decomposition method ($\varepsilon = 0.01$)

Number of nodes	Number of sub-domains	B-spline order k	L^2 error	Iteration steps
1,791	3	16	8.589(−2)	10
		18	9.326(−3)	10
		20	3.957(−3)	15
2,757	3	16	1.471(−1)	10
		18	1.491(−2)	12
		20	3.236(−3)	14

Figures 6 and 7 depict further the computed solutions and pointwise errors for the convection-dominated problem for the chosen ε values with 2,757 collocation points.

Finally, we re-examine the convection dominated problem for $\varepsilon = 0.01$ by using six sub-domains of 2,757 collocation points, as shown in Fig. 8.

All parameters are kept the same as in the previous simulation, unless B-spline order of 22 is now employed. We can obtain further increase in accuracy. The

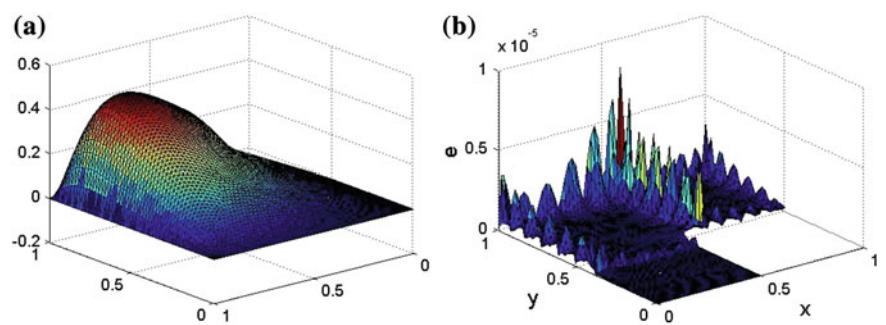


Fig. 6 **a** Computed solution and **b** pointwise errors for the convection-dominated problem ($\varepsilon = 0.1$, 2,757 collocation points and B-spline order 12)

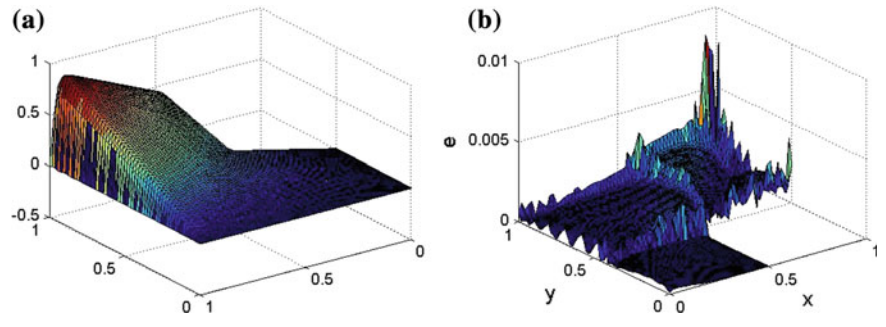


Fig. 7 **a** Computed solution and **b** pointwise errors for the convection-dominated problem ($\varepsilon = 0.01$, 2,757 collocation points and B-spline order 20)

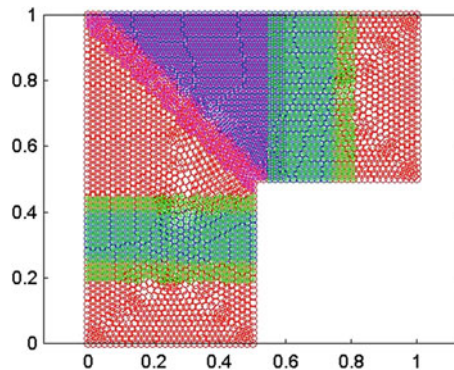


Fig. 8 L-domain decomposed into six overlapping sub-domains

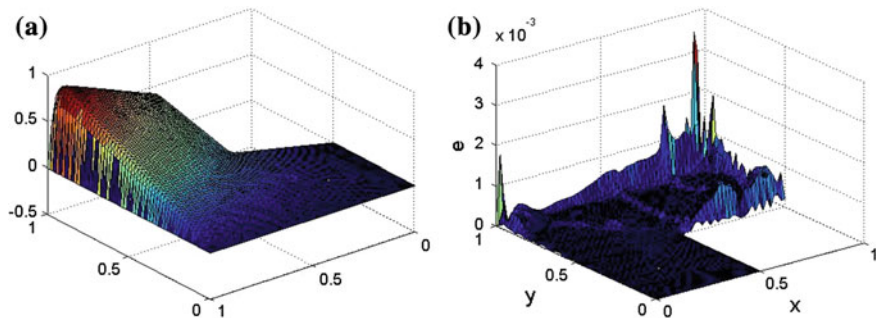


Fig. 9 **a** Computed solution and **b** pointwise errors for the convection-dominated problem with $\varepsilon = 0.01$ re-examined by using six sub-domains of 2,757 nodes. B-spline order 22

obtained L^2 error with six sub-domains is $6.981\text{E-}4$ in comparison the accuracy of $1.0490\text{E-}3$ which is obtained with three sub-domains.

Figure 9 presents the computed solutions and pointwise errors for the convection-dominated problem for $\varepsilon = 0.01$ re-examined by using six sub-domains of 2,757 collocation points and B-spline order 22.

6 Conclusions

In this study, overlapping domain decomposition methods of Schwarz multiplicative and additive have been combined with the global B-spline collocation method. It is shown that the combination method produced higher accuracy with the B-spline basis of much lower order than that needed in implementation of the initial method. The B-spline collocation with domain decomposition method hence improves the approximation stability of the global B-spline collocation method.

Furthermore, the combination method still retained great flexibility of the global B-spline collocation method in dealing with arbitrary complex domains. The efficacy and capability of the B-spline collocation with domain decomposition method are further shown through numerical simulations of singularly perturbed convection-diffusion problems.

Acknowledgments This research work is funded by a grant provided by Universiti Teknologi PETRONAS, which is gratefully acknowledged.

References

1. Fairweather G, Meade D (1989) A survey of spline collocation methods for the numerical solution of differential equations. In: Diaz JC (ed) *Mathematics for large scale computing*, Lecture notes in pure and applied mathematics, vol 120. Marcel Dekker, New York
2. Bialecki B, Fairweather G (2001) Orthogonal spline collocation methods for partial differential equations. *J Comput Appl Math* 128:55–82
3. Botella O (2002) On a collocation B-spline method for the solution of the Navier-Stokes equations. *Comput Fluids* 31:397–420
4. Johnson RW (2005) Higher order B-spline collocation at the Greville abscissae. *Appl Num Math* 52:63–75
5. Johnson RW (2005) A B-spline collocation method for solving the incompressible Navier-Stokes equations using an ad hoc method. *Comput Fluids* 34:121–149
6. Jator S, Sinkala Z (2007) A high order B-spline collocation method for linear boundary value problems. *Appl Math Comp* 191:100–116
7. Botella O (2003) A high-order mass-lumping procedure for B-spline collocation method with application to incompressible flow simulations. *Int J Numer Meth Eng* 41:1295–1318
8. Botella O, Shariff K (2003) B-spline methods in fluid dynamics. *Int J Comput Fluid Dyn* 17:133–149
9. Morinishi Y, Tamano S, Nakabayashi K (2003) A DNS algorithm using B-spline collocation method for compressible turbulent channel flow. *Comput Fluids* 32:751–776
10. Widjaja R, Ooi A, Chen L, Manasseh R (2005) Computational aeroacoustics using the B-spline collocation method. *C R Mecanique* 333:726–731
11. Saka B, Dag I (2007) Quartic B-spline collocation algorithms for numerical solution of the RLW equation. *Numer Methods Partial Differ* 23:731–751
12. Patlashenko I, Weller T (1995) Two-dimensional spline collocation method for nonlinear analysis of laminated panels. *Comput Struct* 57:131–139
13. Viswanathan KK, Navaneethakrishnan PV (2003) Free vibration study of layered cylindrical shells by collocation with splines. *J Sound Vib* 260:807–827
14. Hsu MH (2009) Vibration analysis of pre-twisted beams using the spline collocation method. *J Mar Sci Technol* 17(2):106–115
15. Rashidinia J, Ghasemi M (2011) B-spline collocation for solution of two-point boundary value problems. *J Comp App Math* 235:2325–2342
16. Botella O, Shariff K (2003) Solving the Navier-stokes equations using B-spline numerical methods. In: Hafez MM (ed) *Numerical simulations of incompressible flows*, World Scientific, Singapore, pp 277–287
17. Kansa EJ (1990) Multiquadric—a scattered data approximation scheme with applications to computational fluid dynamics II. *Comput Math Appl* 19:147–161
18. Schwarz HA (1870) Über einen Grenzübergang durch alternierendes Verfahren. *Ges Math Abh* 1:133–143

19. de Boor C (1972) On calculating with B-splines. *J Approx Theory* 6(1):50–62
20. Cox M (1972) The numerical evaluation of B-spline. *J Inst Math Appl* 10:134–149
21. Piegl L, Tiller W (1995) *The NURBS book*. Springer, New York
22. Farin G (2002) *Curves and surfaces for computer aided geometric design*. Academic Press, San Diego
23. Liu Y, Sun L, Xu F, Liu Y, Cen Z (2011) B spline-based method for 2-D large deformation analysis. *Eng Anal Bound Elem* 35:761–767
24. Zhou X, Hon YC, Li J (2003) Overlapping domain decomposition method by radial basis functions. *Appl Numer Math* 44:241–255
25. <http://www.ddm.org/>
26. Saad Y (2003) *Iterative methods for sparse linear systems*. SIAM Publications, Philadelphia
27. Hidayat MIP, Ariwahjoedi B, Parman S (2013) B-spline collocation with domain decomposition method. *J Phys Conf Ser* 423:012012
28. Ingber MS, Chen CS, Tanski JA (2004) A mesh free approach using radial basis functions and parallel domain decomposition for solving three-dimensional diffusion equations. *Int J Numer Meth Eng* 60:2183–2201
29. Zhang Z (2003) Finite element superconvergence on Shishkin mesh for 2-D convection-diffusion problems. *Math Comput* 72:1147–1177

Introducing the Dimensional Continuous Space–Time Theory

Luiz Cesar Martini

Abstract This article is an introduction to a new theory. The name of the theory is justified by the dimensional description of the continuous space-time of the matter, energy and empty space, that gathers all the real things that exists in the universe. The theory presents itself as the consolidation of the classical, quantum and relativity theories. A basic equation that describes the formation of the Universe, relating time, space, matter, energy and movement, is deduced. The four fundamentals physics constants, light speed in empty space, gravitational constant, Boltzmann's constant and Planck's constant and also the fundamentals particles mass, the electrical charges, the energies, the empty space and time are also obtained from this basic equation. This theory provides a new vision of the Big-Bang and how the galaxies, stars, black holes and planets were formed. Based on it, is possible to have a perfect comprehension of the duality between wave-particle, which is an intrinsic characteristic of the matter and energy. It will be possible to comprehend the formation of orbitals and get the equationing of atomics orbits. It presents a singular comprehension of the mass relativity, length and time. It is demonstrated that the continuous space-time is tridimensional, inelastic and temporally instantaneous, eliminating the possibility of spatial fold, slot space, worm hole, time travels and parallel universes. It is shown that many concepts, like dark matter and strong forces, that hypothetically keep the cohesion of the atomics nucleons, are without sense.

1 Introduction

In this article, the main equation of this theory is formally deduced, that is named the primitive equation of the continuous space-time. In order to give a first notion of the introduced concepts of this theory, imagine that any object occupying a specific

L.C. Martini (✉)

Electrical and Computing Engineering Faculty, University of Campinas, Campinas, Brazil
e-mail: martini@decom.fee.unicamp.br

space has dimensions that fill and define a volume in this space and this space cannot be occupied instantaneously by other object.

Imagine for example a stick in the space. This stick exists only because it has a potential energy that creates, defines and keeps its dimensions, determining a volume in the space.

In order to characterize the empiric concept of the existence of a volume occupying a determined space, it will be established the concept of a dimensional axis of potential energy, in which a elementary dimensional axis of potential energy occupying a space do not allows that other elementary axis of potential energy occupies the same space. This dimensional axis that constitutes an incremental axis, for example, of matter, that to exist physically needs a potential energy that creates and keeps it occupying a volume in the space. This first concept is introduced formally in the following [1].

2 Embryo of the Primitive Continuous Space–Time

In the last section, it will be established the rudiments of a theory that supposes the existence of a dimensional length that has a potential energy. This potential energy is stabilized from two opposed pairs of conjugated forces that have the same common origin [2].

It is defined a dimensional length α , whose end points are defined by the points P_1 and P_{-1} . At both points, there are two forces, one of contraction and the other of expansion, that are in equilibrium and maintain the potential at these points.

The conjugated forces of contraction f_{r_1} and $f_{r_{-1}}$ are named anti-space-time conjugated forces and the conjugated forces of expansion f_{s_1} and $f_{s_{-1}}$ are named space-time conjugated forces.

These two couples of conjugated forces $(f_{r_1}; f_{r_{-1}})$ and $(f_{s_1}; f_{s_{-1}})$ constitute one energy dimensional axis that sustains the elementary points P_1 and P_{-1} . This energy dimensional axis constituted by the elementary points P_1 and P_{-1} is the 0 dimensional basis of the space and it is going to be named as a elementary dimension of potential energy.

One elementary dimension of potential energy has a space field S , which energy is proportional to S^2 , i.e.

$$E = kS^2 \quad (1)$$

where k is a constant that depends on the dimensional length α . The space field S on the other hand should be dependent of the dimensional length α too. Imagine that the space field should be an exponential function of dimension length α , if it is differentiated or integrated it must result in the same space field. Suppose an origin

point in the space where many elementary axis of energy pass through. These axis and their energies are jointly independent, so they are mathematically orthogonal, that means that the total energy is the sum of the energies of each elementary axis. Then for two elementary axis of energy

$$E(P_{1,-1} + P_{2,-2}) = E(P_{1,-1}) + E(P_{2,-2})$$

where $P_{1,-1}$ is the elementary axis whose extreme points are P_1 and P_{-1} and $P_{2,-2}$ is the elementary axis whose extreme points are P_2 and P_{-2} . Equivalently,

$$\begin{aligned} E_1 &= E(P_{1,-1}) = k\vec{S}_1\vec{S}_1 = kS_1^2 \\ E_2 &= E(P_{2,-2}) = k\vec{S}_2\vec{S}_2 = kS_2^2 \\ E_{1,2} &= E(P_{1,-1} + P_{2,-2}) = kS_1^2 + kS_2^2 \end{aligned} \quad (2)$$

It is defined a volume V that contains two elementary dimensional axis of potential energy E_1 and E_2 , where the total energy is given by (2) and the space fields S_1 and S_2 are mathematically orthogonal. So, the volume V containing a set of m elementary dimensions is a euclidian space. As consequence, the addition of the energies E_1 and E_2 forms a one-dimensional mathematic space, that can be obtained from two zero-dimensional mathematic spaces, that has equal energies, $E_1 = E_2$.

In order to produce a two-dimensional mathematic space, it will start from two one-dimensional mathematic spaces. So it needs to add two more elementary dimension of potential energy. Now, considering a great number of elementary axis of potential energy that are uniformly distributed inside a volume V and constituted by elementary dimensions of length α , where the i -th elementary energy is given by $E_i = E_0$. Where E_0 is the elementary axis of potential energy, that is the same in all directions.

In order to establish an empiric formation law, it can firstly notice that the D -dimensional space of potential energy is formed by the addition of a power of two elementary energy dimensions.

Starting from a zero-dimensional mathematic space of energy $E(0)$, i.e. $E(0) = E_1 = E_0$.

The one-dimensional mathematical space of energy $E(1)$ is obtained by the addition of two zero-dimensional space of energy, i.e. $E(1) = E(0) + E(0) = E_1 + E_2 = 2E_0$.

The two-dimensional mathematical space of energy $E(2)$ is obtained by the addition of two one-dimensional mathematical space of energy, i.e. $E(2) = E(1) + E(1) = E_1 + E_2 + E_3 + E_4 = 4E_0$.

So, the D -dimensional mathematical space of energy $E(D)$ is given by:

$$E(D) = 2^D E_0 = e^{2\alpha_0 D} E_0 \quad (3)$$

So, D is defined as the dimension of a D -dimensional mathematical space of energy. In this way, it is found an elementary law that describes the total energy inside the volume V .

In the following, it is discovered that the total energy is dependent on the mathematical dimension D . Defining the space constant α_0 as

$$\alpha_0 = \frac{1}{2} \ln 2 \quad (4)$$

The energy $E(D)$ represents the energy of a dimensional space containing a set of elementary dimensions of potential energy.

The space of energy has dimension equal to D constructed from the bases of segments of one-dimensional axis of energy, given by

$$E_{1b} = E(D)^{\frac{1}{D}} = 2E_0^{\frac{1}{D}} = e^{2\alpha_0} E_0^{\frac{1}{D}} \quad (5)$$

So, the base zero-dimensional of one energy elementary dimension has energy, given by

$$E_{0b} = E_0^{\frac{1}{D}} \quad (6)$$

The (3) define an embryonal formula of the continuous space-time of the matter, energy and empty space that represents the totality of all real things in the Universe.

It can be demonstrated that

$$D = \infty \quad \text{for the time and for the empty space} \quad (7)$$

$$D = D_0 = 56 \quad \text{for the matter and for the energy} \quad (8)$$

On the other hand, conjugated forces of space-time f_s , do not exist for the empty space, there are only conjugated forces of anti-space-time, f_r .

3 Introducing the Temporal Variation in the Length α

Starting from the embryonal equations (1) and (3), it is possible to find an expression for E_0 . Since the elementary dimension of energy has a length α , setting $k = \alpha$ in (1). Since the space potential $S_{(0)}$ must be a function of α and an exponential law has to be obeyed, so doing.

$$E_0 = kS_0^2 = \alpha e^{2\alpha_0\alpha} \quad (9)$$

justified by the dimensional law 2^D , where by analogy it has been done $D = \alpha$ as elementary dimension.

In this theory, the universe is constituted only by dimensions whose unit of length is described in an absolute mathematical measurement system; that is perfectly characterized by the international measurement system (SI) [3]. It can be demonstrated that using the absolute mathematical measurement system, that

$$\alpha = \frac{1}{\pi} \quad (10)$$

Now appears the first question: How to introduce in the expression of E_0 space and anti-space representations and its temporal dependence? So, doing

$$\begin{aligned} E_0 &= \alpha S_0^2 = \alpha S_{0s} S_{0r} \\ E_0 &= \alpha e^{\alpha_0 \alpha} e^{\alpha_0 \alpha} = \alpha S_{0s} S_{0r} \end{aligned} \quad (11)$$

where S_{0s} is a space component and S_{0r} is an anti-space component of the elementary energy E_0 .

So it has

$$S_{0s} = e^{\alpha_0 D_{1s}} e^{\alpha_0 L_{1s}} = e^{\alpha_0 \frac{\alpha}{2}} S_{1s}, \quad (12)$$

where D_{1s} is a dimension of the stationary space given by $D_{1s} = L_{1s} = \frac{\alpha}{2}$ and L_{1s} is a dimensional length of the dimension D_{1s} of the stationary space.

As a consequence,

$$\begin{aligned} S_{0r} &= e^{\alpha_0 D_{1r}} e^{\alpha_0 L_{1r}} = e^{\alpha_0 \frac{\alpha}{2}} S_{1r} \\ D_{1r} &= L_{1r} = \frac{\alpha}{2} \end{aligned} \quad (13)$$

where S_{1s} is the elementary stationary component potential of the energy space and S_{1r} is the elementary stationary component potential of the anti space energy.

Now it is introduced a variation into the lengths L_{1s} and L_{1r} so it will be substituted respectively L_{1s} and L_{1r} by

$$\begin{aligned} L_{1s} &= \alpha_s(t_s) = f_s(t_s) + \frac{\alpha}{2} \\ L_{1r} &= \alpha_r(t_s) = f_r(t_s) + \frac{\alpha}{2} \end{aligned} \quad (14)$$

With the harmonic fluctuations $\alpha_s(t_s)$ and $\alpha_r(t_s)$ having a complete temporal dependence of t_s in a most simple form. It is designed t_s as the subjective time or the time that rises to the future like the common concept that we have about the time.

Otherwise, the real time is a dimension that oscillates as a pendulum with variable velocity. So it is defined a spatial fluctuation dependent of the time designed as space-time function given by $f_s(t_s) = (\alpha/2)\sin(\omega_0 t_s)$ where ω_0 must be determined.

4 Identification of Light Speed as a Universal Constant

On the other hand, before the Big-Bang E_0 can not have fluctuations because of the equilibrium between anti space and space time forces.

In this first moment because there is no correct notion of the time movement t_s , it is erroneously induced that $\alpha_r(t_s)$ is equal to $-\frac{\alpha}{2}\sin(\omega_0 t_s) + \frac{\alpha}{2}$.

So, before the Big Bang $E_0 = \alpha e^{z_0(D_{1s}+D_{1r})} e^{z_0 z_s(t_s)} e^{z_0 z_r(t_s)}$.

Now it must focus to the physical and mathematical nature of S_{1s} and S_{1r} .

The nature of the anti space and space forces are totally dynamics, in this sense they only can be represented by the dynamics components of $\alpha_s(t_s)$ and $\alpha_r(t_s)$.

Then it is defined the elementary embryonal equations components of dynamics space S_s and of dynamics anti space S_r in the same sense of S_{1s} and S_{1r} , so that

$$\begin{aligned} S_{1s} &= e^{z_0 D_{2s}} e^{z_0 f_s(t_s)} = e^{z_0 D_{2s}} S_s \\ S_{1r} &= e^{z_0 D_{2r}} e^{z_0 f_r(t_s)} = e^{z_0 D_{2r}} S_r \end{aligned} \quad (15)$$

Then $D_{2s} = D_{2r} = \alpha/2$. So they have a dynamics components of the energies given by

$$\begin{aligned} S_{0s} &= e^{z_0 D} e^{z_0(D_{1s}+D_{2s})} S_s \\ S_{0r} &= e^{z_0 D} e^{z_0(D_{1r}+D_{2r})} S_r \end{aligned}$$

with

$$\begin{aligned} D_s &= D_{1s} + D_{2s} = \alpha = \frac{1}{\pi} \\ D_r &= D_{1r} + D_{2r} = \alpha = \frac{1}{\pi} \end{aligned}$$

where $D_s = D_r = \alpha = 1/\pi$ are the dimensions of dynamics space and anti-space components. Its correspondent lengths L_s and L_r are equal to α that is the peak to peak variation of $f_s(t_s)$ and $f_r(t_s)$ respectively.

$$\begin{aligned} E(D) &= \alpha e^{2z_0 D} e^{z_0(D_s+D_r)} S_s S_r \\ &= \alpha e^{2z_0(D+\alpha)} S_s S_r \end{aligned} \quad (16)$$

Now the dimensional physical and mathematical space has dimension $D + \alpha$ and the quantities $\beta(D) = e^{2\alpha_0(D+\alpha)}$ are fundamental nature, dimensional function, that only depends of the dimension D .

This function $\beta(D)$ is physically and mathematically without dimension and to the mater and energy it has $D = D_0 = 56$. Then to the mater and energy, it is identified $\beta(D_0)$ with the square of the light velocity c in the empty space, expressed in meters per seconds.

In (16) by doing $D = D_0$, the value c numeric equal to the light velocity in the empty space, is the phase time velocity of the wave time into the mater and energy.

$$c = e^{\alpha_0(D_0 + \frac{1}{\pi})} \quad (17)$$

where $c = 2.997438563 \times 10^8$ m/s. The percentage difference between theoretical and practical values is $PD = -0.016$.

Now it is inquired that velocity has dimension, but this theory will proof that dimension time is exactly like space dimension, so meter by seconds is not dimensional.

This theory will proof that everything that exists into the universe is express in the unity of mathematical dimensional length α . So in a concise form it has

$$E(D_0) = \alpha c^2 S_s S_r \quad (18)$$

5 Determination of the Angular Frequency ω_0

Aiming to find a most complete equation that (16) represents, it will start admitting the principle that before the Big-Bang the continuous space time equation must not change with movement time. So the elementary cell of energy do not change with the time variation too.

The cell must be totally static, without manifestation of temperature, without mass like it is known, or electric charge. The cell must be only an elementary cell of potential energy that keeps its dimensions stables. So, in this way, the composition of the continuous space time equation must be present the dynamics equations S_s and S_r representatives of the forces of space and anti space that mutually one neutralizes the other from a derivative process.

As a start point principle, it will consider that the time was always present and it, in fact, made a continuous and indefinite process derivative in the energy at the elementary cells.

Meanwhile, before the Big-Bang the continuous temporal derivative in the energy, does not promote fluctuations in this static energy.

So, initially it will admit that the derivative of the energy before the Big-Bang was zero and it will be making

$$\frac{dE(D_0)}{dt_s} = 0$$

In this way, the primitive Universe was totally static in the zero absolute temperature. It was constituted only by elementary cells, completely stationaries, disposed side by side, totally occupying the empty space and All Universe Energy was stored in this elementary cells as potential energy.

Each elementary cell contains 2^{D_0} elementary dimensions of potential energy in a homogeneous spherical distribution of radius $\alpha/2$.

Because the elementary cells has this homogenous distribution it can be proven that the potential gravitational between the cells is always constant, independent of the volume of the cells that is considered.

It is going to be choose the referential space-time function $f_s(t_s)$ to have a sinusoidal form and the referential anti space-time function $f_r(t_s)$ to be a cosine, that it will be justified later and is given by

$$f_r(t_s) = -\frac{\alpha}{2} \cos(\omega_0 t_s)$$

It will be postulate that the equation of action into the continuous space time is a continued derivative equation. For a while it will be doing this action equation before the Big-Bang to be equal zero

$$\frac{dE(D_0)}{dt_s} = \omega_0 \alpha_0 \alpha^2 \frac{\sqrt{2}}{2} \cos\left(\omega_0 t_s - \frac{\pi}{4}\right) c^2 S_s S_r = 0$$

The unique form of this expression to be zero is done $\omega_0 = 0$. But so where will be the temporal variation?

6 Introduction of the Phase Variation

The question above will be solved by introducing a phase variation $\theta(t_s) = \omega_1 t_s$, where $\omega_1 = 2\pi \frac{v_f}{\lambda}$ and v_f is the phase velocity of the time wave, λ is the length of time wave produced by this phase velocity.

As in this moment the time will be considered into complete stationary space and not into movement matter, is possible to demonstrate that in this condition it must have $\frac{\lambda}{v_f} = 2\pi$.

When it will be considered the matter in movement with velocity v , it will be demonstrated that

$$\omega_1 = \sqrt{1 - \frac{v^2}{c^2}}$$

that is the known Lorentz factor [4].

In this moment it will be assumed that ω_1 is a constant value to be found.

Aiming not to disturb the equation before the Big-bang, it is introduced two opposite phases. So, before the Big-Bang, it is going to suppose initially that

$$f(t_s) = f_s(t_s) + f_r(t_s) = \alpha \frac{\sqrt{2}}{2} \sin(\omega_0 t_s - \frac{\pi}{4} + \theta(t_s) - \theta(t_s))$$

In this first moment it will be supposed that space and anti space forces will be represented by the phases $\theta(t_s)$ and $-\theta(t_s)$.

What does it mean having into the space a sinusoid with zero frequency and $\theta(t_s)$ phase changing with the time?

It means that the amplitude of the sinusoid will be instantaneously the same in all universe and this amplitude has a variation with the time which amplitude is defined by the phase. This equation depends on the time and is producing zero derivative.

7 Alteration of Equilibrium Conditions

After the Big-Bang the $\frac{dE(D_0)}{dt_s}$ equation has to be changed.

The start of the Big-Bang was caused by a unbalance in a single elementary cell. With finality to promote a transformation in the equation, in this first moment $-\theta(t_s)$ will be eliminated. The introduction of two opposite phase terms was resulting of a logical rationality, but it will be demonstrated that even before the Big-Bang, it was only necessary the phase $+\theta(t_s)$ and the phase $-\theta(t_s)$ was totally unnecessary.

So to promote the changing in a energy differential equation it will eliminate $-\theta(t_s)$, then the transformation will give

$$f(t_s) = \alpha \frac{\sqrt{2}}{2} \sin\left(\omega_0 t_s - \frac{\pi}{4} + \theta(t_s)\right)$$

Now the derivation of the energy equation is starting to produce a factor of energy reduction and dimensional contraction. The energy factor will form the mass and charge of the matter.

After the Big Bang it will postulate that to constitute the materia, the differential equation $\frac{dE(d_0)}{dt_s}$ will be separated in two components that are totally tied and produces combined action, but having independent actuation.

Considering the product $P(t_s) = S_s S_r = e^{\alpha_0 f(t_s)}$. The mathematical derivative of the product is

$$\begin{aligned} P'(t_s) &= \alpha_0 f'(t_s) P(t_s) = \alpha_0 f'(t_s) e^{\alpha_0 f(t_s)} \\ f'(t_s) &= \frac{df(t_s)}{dt_s} = \frac{\sqrt{2}}{2} \cos(\omega_1 t_s - \frac{\pi}{4}) \\ P'(t_s) &= \alpha_0 \frac{\sqrt{2}}{2} \cos(\omega_1 t_s - \frac{\pi}{4}) e^{\alpha_0 f(t_s)} \end{aligned}$$

Now it must attempt to the true physical derivative form of $P(t_s)$.

The time derivative actuation is produced over all functions in all equations and must separate space and anti space.

So it is postulated that the derivative form physically must produce:

$$P'_0 = \alpha_0 f'_s(t_s) e^{\alpha_0 f'_s(t_s)} + \alpha_0 f'_r(t_s) e^{\alpha_0 f'_r(t_s)}$$

Then it must have

$$E'(D_0) = \alpha c^2 P'_0 = \alpha c^2 \left[\alpha_0 \frac{\alpha}{2} \right] \cos(\omega_1 t_s) e^{\alpha_0 f'_s(t_s)} + \alpha c^2 \left[\alpha_0 \frac{\alpha}{2} \right] \sin(\omega_1 t_s) e^{\alpha_0 f'_r(t_s)}$$

The term $\alpha_0 \frac{\alpha}{2}$ will be proved in the Theory that it is an incremental contribution to the mass and charge of the elementary particles.

Then to the space and anti-space components it will have respectively

$$\begin{aligned} E'_s(d_0) &= \alpha c^2 [\alpha_0 \alpha / 2] \cos(\omega_1 t_s) e^{\alpha_0 f'_s(t_s)} \\ E'_r(d_0) &= \alpha c^2 [\alpha_0 \alpha / 2] \sin(\omega_1 t_s) e^{\alpha_0 f'_r(t_s)} \\ f'_s(t_s) &= \alpha / 2 \cos(\omega_1 t_s) \\ f'_r(t_s) &= \alpha / 2 \sin(\omega_1 t_s) \end{aligned}$$

Now it will try to reconstruct the primitive equation from its separation doing firstly

$$g(t_s) = E'_s(t_s) e^{\alpha_0 f'_s(t_s)} + E'_r(t_s) e^{\alpha_0 f'_r(t_s)} = \alpha c^2 [\alpha_0 \alpha / 2] \sqrt{2} \cos(\omega_1 t_s - \frac{\pi}{4}) e^{\alpha_0 f'(t_s)}$$

So in the equation above has the additive term from a sinuous and cosine given by

$$h(t_s) = \sqrt{2}\cos(\omega_1 t_s - \frac{\pi}{4})$$

This term is not reconstructed, and its represents an other separation term between $E'_s(D)$ and $E'_r(D)$ and it can be seen that

$$h(t_s) = \sqrt{2}\frac{e^{j(\omega_1 t_s - \frac{\pi}{4})}}{2} + \sqrt{2}\frac{e^{-j(\omega_1 t_s - \frac{\pi}{4})}}{2}$$

To restore the original form of this term its substitution will be postulated by

$$q(t_s) = \sqrt{2}\frac{e^{j(\omega_1 t_s - \frac{\pi}{4})}}{\sqrt{2}}\frac{e^{-j(\omega_1 t_s - \frac{\pi}{4})}}{\sqrt{2}}$$

In the reconstruction the term $[\alpha_0 \alpha/2]$ is obtained because it is resultant from the process of the separation by derivation.

Because it has introduced a complex time it must consider the time into $f_s(t_s)$ and $f_r(t_s)$ initially as a complex term, than it will have the reconstructed form

$$\begin{aligned} E'(D_0) &= \alpha c^2 \sqrt{2} \frac{e^{j(\omega_1 t_s - \frac{\pi}{4})}}{\sqrt{2}} \frac{e^{-j(\omega_1 t_s - \frac{\pi}{4})}}{\sqrt{2}} e^{\alpha_0 f'(t_s)} \\ E'(D_0) &= \alpha c^2 \sqrt{2} \frac{e^{j\omega_1 t_s}}{\sqrt{2}} \frac{e^{-j\omega_1 t_s}}{\sqrt{2}} e^{\alpha_0 \frac{\sqrt{2}}{2} \cos(j\omega_1 t_s - j\omega_1 t_s - \frac{\pi}{4})} \end{aligned}$$

The complex coordinates that appears will be explained when the Relativity Theory is introduced [5]. The complex coordinate that is found is the temporal coordinate which is necessary to be complex to obtain a real phase of the time into the space.

8 Real Condition of the Invariability Before the Big-Bang

Returning to consider again the derivative of the primitive energy equation.

It was made a preliminary analysis considering its mathematical derivative as been zero, however its derivative can not produce nullity of the continuous space time equation, but in sense that keep it unmodified.

So the continue derivative of the primitive continuous space time equation before the big bang has to produce itself, keeping its continuity.

So the real condition in a mathematical derivative definition before the Big-Bang must be $\frac{dE(D_0)}{d(jt_s)} = E(D_0)$, and do not zero.

To satisfy this condition it must eliminate one of two complex terms keeping only one of them. Studying the Relativity Theory the term $+j\omega_1 t_s$ must be conserved. So the following term will be eliminated. $\frac{e^{-j\omega_1 t_s}}{\sqrt{2}}$ and consequently $-j\omega_1 t_s$ in the $f'(t_s)$.

So in the primitive equation there is only a temporal complex coordinate $+j\omega_1 t_s$. Then it will have the primitive equation without derivative $E(D_0) = \alpha c^2 e^{j\omega_1 t_s} e^{\alpha_0 f(t_s)}$ where in this new condition $f(t_s) = j\frac{\alpha}{2} \sin(\omega_1 t_s) - j\frac{\alpha}{2} \cos(\omega_1 t_s)$ and its mathematical derivative equation with respect to complex temporal coordinate jt_s , before the Big-Bang, will be

$$\frac{dE(D_0)}{d(jt_s)} = \omega_1 \alpha c^2 e^{j\omega_1 t_s} e^{\alpha_0 f'(t_s)} + \alpha c^2 f'(t_s) e^{j\omega_1 t_s} e^{\alpha_0 f'(t_s)} \quad (19)$$

so that the derivative with respect to complex time coordinate jt_s produce the properly derivative energy of the continuous space time, so $\omega_1 = 1$ radian and $f'(t_s) = 0$ taking our attention to the term $f(t_s)$.

In this moment there is a question to solve: the continue derivation of $f_s(t_s)$ and $f_r(t_s)$ with respect to jt_s , must produce the nullity of the term $f'(t_s) = f'_s(t_s) + f'_r(t_s)$. How can it be possible?

9 Characterization of the Dimensional Variation

Now the above question will be solved! When the complex component is introduced and keeping the sinuous and cosine components, the complex component $e^{j\omega_1 t_s}$, is a description of the time. So the sinuous and cosine components in $f(t_s)$, can not be the time representation, but the spatial dimensional actuation of the time on the dimension. So inside of the sinuous and cosine terms the time is not the subjective time but the rotating amplitude constant vector $\vec{t}_0$ which amplitude is the constant equal to 1.

Otherwise, the time rotating vector actuation on the dimensional components in sinuous and cosine depends respectively of t_{ss} and t_{sr} and do not directly of t_s .

It can be seen that

$$\begin{aligned} t_{ss} &= jt_0 \sin(\omega_1 t_s) = j \sin(\omega_1 t_s) \\ t_{sr} &= jt_0 \cos(\omega_1 t_s) = j \cos(\omega_1 t_s) \end{aligned}$$

In the terms $f_s(t_s)$ and $f_r(t_s)$ it will be done respectively the transformations

$$j\omega_1 t_s = \omega_1 t_s \frac{t_{ss}}{\sin(\omega_1 t_s)} = \omega_1 t_s \frac{t_{sr}}{\cos(\omega_1 t_s)}$$

so it will have

$$\begin{aligned} f_s(t_{ss}) &= \frac{\alpha}{2} \sin \left(\omega_1 t_s \frac{t_{ss}}{\sin(\omega_1 t_s)} \right) \\ f_r(t_{sr}) &= -\frac{\alpha}{2} \cos \left(\omega_1 t_s \frac{t_{sr}}{\cos(\omega_1 t_s)} \right) \end{aligned}$$

that results

$$\begin{aligned} f_s(t_{ss}) &= \alpha/2 \ t_{ss} \\ f_r(t_{sr}) &= -\alpha/2 \ t_{sr} \end{aligned}$$

As seen above, $f_s(t_{ss})$ and $f_r(t_{sr})$ do not depends on ω_1 . Utilizing the Relativity Theory [6] it can be proven that t_{ss} and t_{sr} do not depends of an angular variant frequency ω_1 and in the expression of t_{ss} and t_{sr} , the angular frequency is $\omega_1 = -j$.

So in reality the correct forms are the following.

$$\begin{aligned} t_{ss} &= \sin(t_s) \\ t_{sr} &= \cos(t_s) \end{aligned} \tag{20}$$

$$\begin{aligned} f_s(t_{ss}) &= \alpha/2 \ t_{ss} = \alpha/2 \sin(t_s) \\ f_r(t_{sr}) &= -\alpha/2 \ t_{sr} = -\alpha/2 \cos(t_s) \end{aligned} \tag{21}$$

And the derivative of $f_s(t_{ss})$ and $f_r(t_{sr})$ is not with respect to jt_s but is with respect to components t_{ss} and t_{sr} respectively.

In this condition the sum of the continue derivation of $f_s(t_{ss})$ and $f_r(t_{sr})$ will be producing

$$f'(t_{ss}, t_{sr}) = \frac{df_s(t_{ss})}{dt_{ss}} + \frac{df_r(t_{sr})}{dt_{sr}} = 0 \tag{22}$$

10 The General Form of the Primitive Equation in Static and Dynamics Forms

So after the definition of time functions $f_s(t_{ss})$ and $f_r(t_{sr})$, now is presented the equations of the continuous space-time before the Big-Bang [7] in the static form:

$$E(D) = \alpha e^{2\alpha_0(D+\alpha)} e^{j\omega_1 t_s} e^{\alpha_0 f(t_{ss}, t_{sr})} \tag{23}$$

$$\begin{aligned}
\alpha_0 &= 1/2 \ln 2 \quad \text{space constant} \\
\alpha &= 1/\pi \quad \text{dimensional length} \\
\omega_1 &= 1 \quad \text{angular frequency of the time before the Big-Bang} \\
f(t_{ss}, t_{sr}) &= f_s(t_{ss}) + f_r(t_{sr}) \quad \text{static time function} \\
t_{ss} &= \sin(t_s) \quad \text{active time in the space component} \\
t_{sr} &= \cos(t_s) \quad \text{active time in the anti-space component} \\
t_s &= \text{subjective time} \\
f_s(t_{ss}) &= \alpha/2 \quad t_{ss} \quad \text{time function of the space component} \\
f_r(t_{sr}) &= -\alpha/2 \quad t_{sr} \quad \text{time function of the anti-space component} \\
v_f &= e^{\alpha_0(D+\alpha)} \quad \text{velocity phase of the time} \\
D &= \infty \quad \text{to the empty space and the time} \\
D &= D_0 = 56 \quad \text{to the matter and energy} \\
S_s &= E^{\alpha_0 f_s(t_{ss})} \quad \text{space component of the elementary static energy} \\
S_r &= E^{\alpha_0 f_r(t_{sr})} \quad \text{anti-space component of the elementary static energy}
\end{aligned}$$

Now is presented the equations of the continuous space-time before the Big-Bang [5] in the dynamic form:

$$E'(D) = \omega_1 \alpha e^{2\alpha_0(D+\alpha)} e^{j\omega_1 t_s} e^{\alpha_0 f'(t_{ss}, t_{sr})} \quad (24)$$

$$\begin{aligned}
f'(t_{ss}, t_{sr}) &= f'_s(t_{ss}) + f'_r(t_{sr}) \quad \text{dynamic time function} \\
f'_s(t_{ss}) &= \alpha/2 \quad \text{time function of the space component in continuous derivative} \\
f'_r(t_{sr}) &= -\alpha/2 \quad \text{time function of the anti-space component in continuous derivative} \\
S_{s'} &= E^{\alpha_0 f'_s(t_{ss})} \quad \text{space component of the elementary dynamic energy} \\
S_{r'} &= E^{\alpha_0 f'_r(t_{sr})} \quad \text{anti-space component of the elementary dynamic energy}
\end{aligned}$$

Then the primitive equation (24) of the continuous space time in its complete form in continuous derivation defined as Continuity equation of the continuous space time before the Big-Bang. After the Big-Bang everything starts to move and to the movement it has

$$\omega_1 = \sqrt{1 - \frac{v^2}{c^2}} \quad (25)$$

that is the angular frequency of the time in a matter moving with speed v .

So the elementary cells of potential energy of the primitive universe into the empty space totally stationary, without any movement or other form of energy are express by the (23) and (24) in static equilibrium. These equations are general and can be worked in many ways, obtaining surprising results. when the angular

frequency ω_1 is introduced, the number $D_0 = 56$ to the matter and energy, and $D = \infty$ to the time and empty space, the separation between space time and anti space time forces, the real time t instead of the subjective time t_s , and the basis of space and time.

11 Some Theoretical Results

Equation (23) can be conveniently separated in three terms like.

$$E(D) = e^{2\alpha_0(D+\alpha)} TS_s S_r \quad (26)$$

where is defined

$$\begin{aligned} T &= e^{i\omega_1 t_s} && \text{time equation} \\ S_s &= e^{\alpha_0 f_s(t_{ss})} && \text{space equation} \\ S_r &= e^{\alpha_0 f_r(t_{sr})} && \text{anti space equation} \end{aligned}$$

From the earlier equations, the space-time and the anti space-time equations are

$$C_s = e^{2\alpha_0(D+\alpha)} TS_s \quad \text{space-time equation} \quad (27)$$

$$C_r = e^{2\alpha_0(D+\alpha)} TS_r \quad \text{anti space-time equation} \quad (28)$$

After the Big-Bang the primitive equation of the continuous space time of the elementary cell of potential energy, was separated in two parts producing the equations of space time and anti space time spectrally tied and non separable into the matter, but in the energy there is only the space time component that is responsible by the wave propagation. The empty space has only the anti space component and the time has only the space component.

The empty space is characterized by its presence in distinct points. The time is characterized by the instantaneous actuation in all empty space. So only element of empty space is separated of another element of empty space, but the time is integralized on all empty space. The empty space, because it has only anti space component, it is physically totally static. Because the time has only space component and infinite space dimension it has an infinite phase velocity of propagation. So the time is a stationary wave and its instantaneous energy is simultaneous over all space.

The process of dimensional reduction where are engaged derivative and/or integral of the functions TS_s or TS_r will produce matter and energy, as resulting of this process. Now is presented some results of this theory:

$N = 7$ Number of atomics Shells

$$RT_1 = \frac{1}{2\pi\sqrt{2}} - \frac{1}{4\pi\sqrt{2}} \quad \text{Partial reduction of the time function}$$

$$RT_2 = \frac{1}{2\pi\sqrt{2}} \quad \text{Complete dimensional reduction of the time function}$$

$$M(0) = \frac{1}{2}(e^{\alpha_0 \frac{z}{2}} + e^{-\alpha_0 \frac{z}{2}}) \quad \text{Dimensional factor of correction}$$

$$Z(0) = \frac{1}{2} \left(M(0) + \frac{1}{M(0)} \right) \quad \text{Dimensional factor correction}$$

where the factor $M(0)$ is associated to the space and the $\frac{1}{M(0)}$ factor is associated to the anti-space. The factor $Z(0)$ is associated to the conjugated actuation between space and anti space.

Light speed into the empty space

$$c = e^{\alpha_0(D_0 + \frac{1}{\pi})} = 2.9974385630 \times 10^8 \text{ m/s} \quad (29)$$

$$\text{PD} = -0.016$$

where PD is the percentage difference.

Static proton mass

$$m'_p = \frac{\alpha_0}{\pi} (-\alpha_0^{D_0}) \alpha_0^{RT_1} \frac{1}{M(0)} \text{ kg} = 1.671795279 \times 10^{-27} \text{ kg} \quad (30)$$

$$\text{PD} = -0.049$$

Static neutron mass

$$m'_n = \frac{\alpha_0}{\pi} (-\alpha_0^{D_0}) \alpha_0^{RT_1} Z(0) \text{ kg} = 1.674341082 \times 10^{-27} \text{ kg} \quad (31)$$

$$\text{PD} = -0.034$$

Boltzmann constant [8]

$$k = \frac{1}{\alpha_0 \frac{\alpha_0}{\pi} \alpha_0^{2RT_1 - RT_2}} e^{-D_0} = 1.380299907 \times 10^{-23} \text{ J/K} \quad (32)$$

$$\text{PD} = -0.0262$$

Bohr Radius [8]

$$r_0 = \frac{1}{(m(8) - 1)c} = \frac{1}{63c} = 5.295526678 \times 10^{-11} \text{ m} \quad (33)$$

$$\text{PD} = 0.070$$

where

$$m(8) = \sum_{l=0}^7 (2l + 1) = 64$$

$$m(8) - 1 = D_0 + N$$

12 Conclusions

In this article, the equation called the primitive equation of the continuous space-time was formally deduced. This paper presents a completely new work that is a small part of a work quite extensive, resulting from years of intense research of the author. It was presented as an introduction, so it can be appreciated, analyzed and criticized. The conclusion will depend on future works all interconnected, and these conclusions may only be obtained after full publication of several other items that compound the theory.

References

1. Mott G (2006) A review of Einstein's theory of special relativity. *Int J Mod Phys* 15:755
2. Hummel RE (2001) *The wave-particle duality, electronic properties of materials*. Springer, New York, pp 6–13
3. Thompson A, Taylor B (2008) *Guide for the use of the International System of Units* NIST. Special publication 811. U.S. Government Printing Office, Washington DC
4. Hegerfeldt GC (1972) The Lorentz transformations: derivation of linearity and scale factor. *Società Italiana di Fisica* 10:257–267
5. Schröder UE (1990) Physical and conceptual foundations of the theory of special relativity. *Relativity* 2:11–58
6. Carmeli M (2008) Special relativity theory. *Relativity* 1:1–49
7. Trimble V (1996) Backgrounds and the Big Bang: some extracts from their history. *Int Astron Union* 168:9–16
8. Weisstein E (2012) Physical constants. <http://scienceworld.wolfram.com/physics/BohrRadius.html>

Application of Nikiforov-Uvarov Method for Non-central Potential System Solution

C. Cari and A. Suparmi

Abstract The energy eigenvalues and eigenfunctions of Schrodinger equation for a 3D harmonic oscillator potential plus Rosen-Morse non-central potential and Eckart plus trigonometric Poschl-Teller non-central potential are investigated using NU method. The bound state energy eigenvalues for both systems are given in a closed form and the corresponding radial wave functions are expressed in associated Laguerre polynomials for 3D harmonics oscillator while the radial and angular eigenfunctions are given in terms of Jacobi polynomials. The Rosen-Morse and Poschl-Teller potentials are considered to be the perturbation factors to the 3D harmonic oscillator and Eckart potentials that cause the decrease of angular momentum length but preserve the number of energy degeneracy.

1 Introduction

The Schrodinger equations of physical potentials have been studied intensively in recent years. Mostly methods used to obtain the exact solution of Schrodinger for physical potentials which consist of a class of shape invariant potentials are factorization method [1–3], super-symmetric quantum mechanics (SUSY QM) approach [4–8], Nikiforov-Uvarov (NU) method [9–13], and Romanovski polynomial [14–17]. Among these methods, there are some methods that interconnect to each other, SUSY QM with factorization method and WKB approach, NU method and Romanovski polynomial are developed based on hypergeometric differential equation. Shape invariant potentials is a class of one dimensional potentials (radial/central and angular functions potential) that obey to the properties proposed by Gendenshtein [8].

NU method, proposed by Nikiforov and Uvarov [18], has been widely used to solve second order linier differential equation without direct solution. The energy

C. Cari (✉) · A. Suparmi

Physics Department, Sebelas Maret University, Surakarta 57126, Indonesia
e-mail: carinln@yahoo.com

A. Suparmi
e-mail: suparmiuns@gmail.com

spectrum and the wave function of certain potential system are calculated using formulas derived from hypergeometric type equation which obtained by simple mathematical manipulation of the Schrodinger equation of the potential of interest. By suitable variable substitution, the Schrodinger equation for certain physical potential reduces to an generalized hypergeometric differential equation (GHDE), and by parameter and wave function substitutions to GHDE then GHDE reduces to hypergeometric/confluent hypergeometric differential equation that is called as hypergeometric type equation. As by product, a set of formulas, which is used to determine the energy spectrum and wave functions, are produced. Therefore the heart of NU method is determining the coefficients of the second, first, and zeroth derivatives from GHDE and plugging it into set of formulas.

In this chapter we analyze the energy spectra and wave functions of non-central potential: 3D oscillator harmonics plus trigonometric Rosen-Morse and Eckart plus Poschl-Teller non-central potentials using NU method. These two system potentials are separable potentials [19–21] therefore its Schrodinger equations solved using variable separation method. Three dimensional harmonics oscillator is one of exactly solvable potential that used to describe the nuclei, atomic or molecular vibration. Non-central potential composed of spherical harmonics oscillator with square of inverse potential together with ring-shaped non-central potential, or double ring shaped potential have been investigated intensively by some authors [22–25]. The Rosen-Morse potential is trigonometric potential which was proposed by Rosen-Morse [26] in 1935 and was used to describe the quark-gluon dynamics. The approximate bound state solution for trigonometric Rosen-Morse potential have also been studied for l -state solution [27–29], and Coulombic Rosen-Morse non-central potential, particularly for $\cot\theta$ part, has been investigated intensively [27]. For $l \neq 0$ and $r \ll 0$, the centrifugal term is approximated by hyperbolic, trigonometric or exponential functions [] and leads to the exact analytical solution of the radial Schrodinger equation. The exact analytical solutions of Schrodinger equations for some physical potentials are very essential since the knowledge of wave functions and energy contains all possible important information of the physical properties of quantum system.

2 Non-central Potential

The Schrodinger equations for central potentials, which are shape invariant potentials such as three dimensional harmonics oscillator, Kepler problems, Wood Saxon potential, Kratzer molecular potential, have been solved exactly using SUSY QM, WKB with Langer correction, and hypergeometric type equations particularly only for $l = 0$. Non-central potentials, which are separable ones, are also exactly solvable for $l = 0$. The separable non-central potential is given as

$$V(r, \theta, \varphi) = V(r) + \frac{V(\theta)}{r^2} + \frac{V(\varphi)}{r^2 \sin^2 \theta} \quad (1)$$

Schrodinger equation for non-central potential expressed in (1) is solved using variable separation method and is exactly solvable for $l = 0$ if $V(r)$, $V(\theta)$, and $V(\varphi)$ are shape invariants. The Schrodinger equation for non-central potential expressed in (1) are resolved into three second order linier differential equation: radial, polar and azimuthal Schrodinger equations. In the case of radial Schrodinger equation, for $l \neq 0$, however, is exactly solvable only if the contribution of centrifugal term, $\frac{l(l+1)}{r^2}$, for very small value of r , $r \ll 1$, is approximated by

$$\frac{l(l+1)}{r^2} \cong \frac{l(l+1)}{\sinh^2 r} \cong \frac{l(l+1)}{\sin^2 r}, \quad (2a)$$

Approximation expressed in (2a, 2b) was initially proposed by Greene and Aldrich [L] and newly improved [M], with $d_0 = 1/12$, given as

$$\frac{l(l+1)}{r^2} \cong \frac{l(l+1)(1+d_0)}{\sinh^2 r} \cong \frac{l(l+1)(1+d_0)}{\sin^2 r} \quad (2b)$$

The radial parts of non-central potential, $V(r)$, that produce the exact solution within r approximation either expressed in (2a) or (2b) are including radial Eckart potential, radial hyperbolic Poschl-Teller potential, radial trigonometric Poschl-Teller potential, radial hyperbolic Rosen-Morse potential, radial trigonometric Rosen-Morse potential, radial hyperbolic Scarf potential and radial trigonometric Scarf potential, as shown in Table 1.

On the other hand, the polar parts of non-central potentials that have exact solutions are listing in Table 2.

By using (1) the non-central potentials are achieved by combining the radial function potentials listed in Table 1 with polar potential listed in Table 2.

Table 1 Lists of radial potential that are solvable using NU method

No	Potential's name	Potential's equation
1	Eckart	$V(r) = \frac{\hbar^2}{2Ma^2} \left(V_0 \frac{e^{-r/a}}{(1-e^{-r/a})^2} - V_1 \frac{1+e^{-r/a}}{1-e^{-r/a}} \right)$
2	Poschl-Teller (trigonometric)	$V(r) = \frac{\hbar^2 x^2}{2M} \left(\frac{\kappa(\kappa-1)}{\sin^2 xr} + \frac{\eta(\eta-1)}{\cos^2 xr} \right)$
3	Modified Poschl-Teller/(hyperbolic)	$V(r) = \frac{\hbar^2 x^2}{2M} \left(\frac{\kappa(\kappa-1)}{\sinh^2 xr} - \frac{\eta(\eta+1)}{\cosh^2 xr} \right)$
4	Hyperbolic Rosen-Morse	$V(r) = \frac{\hbar^2 x^2}{2M} \left(\frac{v(v-1)}{\sinh^2 xr} - 2\mu \coth xr \right)$
5	Trigonometric Rosen-Morse	$V(r) = \frac{\hbar^2 x^2}{2M} \left(\frac{v(v+1)}{\sin^2 xr} - 2\mu \cot xr \right)$
6	Trigonometric scarf II	$V(r) = \frac{\hbar^2 x^2}{2M} \left(\frac{b^2+a(a-1)}{\sin^2 xr} - \frac{2b(a-\frac{1}{2}) \cos xr}{\sin^2 xr} \right)$
7	Hyperbolic scarf II	$V(r) = \frac{\hbar^2 x^2}{2M} \left(\frac{b^2+a(a+1)}{\sinh^2 xr} - \frac{2b(a+\frac{1}{2}) \cosh xr}{\sinh^2 xr} \right)$

Table 2 Lists of polar function potentials that are solvable using NU method

No	Polar potential	Potential's mathematical equation
1	Poschl-Teller	$V(\theta) = \frac{\hbar^2}{2M} \left(\frac{\kappa(\kappa-1)}{\sin^2 \theta} + \frac{\eta(\eta-1)}{\cos^2 \theta} \right)$
2	Rosen-Morse	$V(\theta) = \frac{\hbar^2}{2M} \left(\frac{v(v+1)}{\sin^2 \theta} - 2\mu \cot \theta \right)$
3	Scarf	$V(\theta) = \frac{\hbar^2}{2M} \left(\frac{b^2 + a(a-1)}{\sin^2 \theta} - \frac{2b(a-\frac{1}{2}) \cos \theta}{\sin^2 \theta} \right)$

In this chapter we will solve Schrodinger equation for non-central potential with $V(\varphi) = 1$,

$$V(r, \theta) = \frac{M\omega^2 r^2}{2} + \frac{\hbar^2}{2Mr^2} \left(\frac{v(v+1)}{\sin^2 \theta} - 2\mu \cot \theta \right) \quad (3)$$

and

$$V(r, \theta) = \frac{\hbar^2}{2Ma^2} \left(V_0 \frac{e^{-r/a}}{(1 - e^{-r/a})^2} - V_1 \frac{1 + e^{-r/a}}{1 - e^{-r/a}} \right) + \frac{\hbar^2}{2Mr^2} \left(\frac{\kappa(\kappa-1)}{\sin^2 \theta} + \frac{\eta(\eta-1)}{\cos^2 \theta} \right), \quad (4)$$

and the one dimensional Schrodinger equations are solved using Nikiforov-Uvarov method. Special for three dimensional harmonics oscillator we do not need the approximation value of r.

3 Nikiforov-Uvarov Method

Nikiforov-Uvarov (NU) method was developed based on the hypergeometric differential equation. In the following section the formulas used in NU method are derived from hypergeometric differential equation.

The hypergeometric differential equation expressed by Gau β [30] is given as

$$z(1-z) \frac{\partial^2 \Phi}{\partial z^2} + \{c - (a+b+1)z\} \frac{\partial \Phi}{\partial z} - ab\Phi = 0 \quad (5)$$

Equation (1) has three regular singular points at $z = 0$, $z = 1$, $z = \infty$. By using Frobenius method [30], the general solution at around point $z = 0$ is given as

$$\Phi(z) = A {}_2F_1(a, b; c; z) + B z^{1-c} {}_2F_1(a+1-c, b+1-c; 2-c; z) \quad (6)$$

where

$$\begin{aligned} {}_2F_1(a, b; c; z) &= \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{z^n}{n!} \\ &= 1 + \frac{ab}{c} \frac{z}{1!} + \frac{a(a+1)b(b+1)}{c(c+1)} \frac{z^2}{2!} \\ &\quad + \frac{a(a+1)(a+2)b(b+1)(b+2)}{c(c+1)(c+2)} \frac{z^3}{3!} + \dots \end{aligned} \quad (7)$$

$$(a)_n = a(a+1)(a+2)(a+3)\dots(a+n-1) \quad (8)$$

and c is not integer. When c is integer the second part of the solution becomes complicated and we have only the first part of the solution. For $a = -n$ or $b = -n$ the solution of (7) becomes finite. By substituting $z = \frac{x}{b}$ and for $b \rightarrow \infty$ then hypergeometric differential equation in (5) reduces to confluent hypergeometric equation given as

$$x \frac{\partial^2 \Phi}{\partial^2 z} + \{c - x\} \frac{\partial \Phi}{\partial z} - a\Phi = 0 \quad (9)$$

The solution of confluent hypergeometric equation in (9) at around regular singular point $z = 0$ is given as

$$\Phi(z) = A_1 F_1(a; c; z) + (Bz^{1-c} {}_1F_1(a+1-c; 2-c; z)) \quad (10)$$

with

$${}_1F_1(a; c; z) = \sum_{n=0}^{\infty} \frac{(a)_n}{(c)_n} \frac{z^n}{n!} = 1 + \frac{a}{c} \frac{z}{1!} + \frac{a(a+1)}{c(c+1)} \frac{z^2}{2!} + \frac{a(a+1)(a+2)}{c(c+1)(c+2)} \frac{z^3}{3!} + \dots \quad (11)$$

For $a = -n$ the polynomials in (11) becomes finite.

The one-dimensional Schrodinger equation of any shape invariant potential can be reduced into hypergeometric or confluent hypergeometric type differential equation, expressed in (5) or (9) by suitable variable transformation [12, 26]. The hypergeometric type differential equation, which is solved using Nikiforov-Uvarov method, is presented as

$$\frac{\partial^2 \Psi(s)}{\partial s^2} + \frac{\tilde{\tau}(s)}{\sigma(s)} \frac{\partial \Psi(s)}{\partial s} + \frac{\tilde{\sigma}(s)}{\sigma^2} \Psi(s) = 0 \quad (12)$$

where $\sigma(s)$ and $\tilde{\sigma}(s)$ are polynomials at most in the second order, and $\tilde{\tau}(s)$ is first order polynomial. Equation (12) is obtained from the Schrodinger equation of the

certain potential by suitable variable substitution. Equation (12) is solved by using separation of variable method which is expressed as

$$\Psi(s) = \phi(s)y(s) \quad (13)$$

for appropriate $\phi(s)$ function, and (12) reduces to

$$y''(s) + \left\{ 2 \frac{\phi'(s)}{\phi(s)} + \frac{\tilde{\tau}(s)}{\sigma(s)} \right\} y'(s) + \left\{ \frac{\phi''(s)}{\phi(s)} + \frac{\phi'(s)\tilde{\tau}(s)}{\phi(s)\sigma(s)} + \frac{\bar{\sigma}(s)}{\sigma^2(s)} \right\} y(s) = 0 \quad (14)$$

In order (14) is not more complex than (12), then the coefficient of $y'(s)$ in (14) has to be in the form of $\frac{\tau(s)}{\sigma(s)}$ that is

$$2 \frac{\phi'(s)}{\phi(s)} + \frac{\tilde{\tau}(s)}{\sigma(s)} = \frac{\tau(s)}{\sigma(s)} \quad (15)$$

By rewriting

$$\frac{\phi'(s)}{\phi(s)} = \frac{\pi(s)}{\sigma(s)} \quad (16)$$

then we obtain

$$2\pi(s) = (\tau(s) - \tilde{\tau}(s)) \rightarrow \tau(s) = \tilde{\tau}(s) + 2\pi(s) \quad (17)$$

where the new parameter $\pi(s)$ is the first order polynomial. By expressing the $\frac{\phi''(s)}{\phi(s)}$ term in (14) as

$$\frac{\phi''(s)}{\phi(s)} = \left[\frac{\phi'(s)}{\phi(s)} \right]' + \left[\frac{\phi'(s)}{\phi(s)} \right]^2 = \left[\frac{\pi(s)}{\sigma(s)} \right]' + \left[\frac{\pi(s)}{\sigma(s)} \right]^2 \quad (18)$$

and by setting the coefficient of $y(s)$ in (14) to be equal to $\frac{\bar{\sigma}(s)}{\sigma^2(s)}$ then from (14) and (18) we have

$$\left[\frac{\pi(s)}{\sigma(s)} \right]' + \left[\frac{\pi(s)}{\sigma(s)} \right]^2 + \frac{\phi'(s)\tilde{\tau}(s)}{\phi(s)\sigma(s)} + \frac{\bar{\sigma}(s)}{\sigma^2(s)} = \frac{\bar{\sigma}(s)}{\sigma^2(s)} \quad (19)$$

and (14) becomes

$$y''(s) + \frac{\tau(s)}{\sigma(s)} y'(s) + \frac{\bar{\sigma}(s)}{\sigma^2(s)} y(s) = 0 \quad (20a)$$

Equation (19) is rewritten as

$$\bar{\sigma}(s) = \tilde{\sigma}(s) + \pi^2(s) + \pi(s)(\tilde{\tau}(s) - \sigma'(s)) + \pi'(s)\sigma(s) \quad (20b)$$

In order the expression of (20a, 20b) as simple as possible then $\bar{\sigma}(s)$ in (20a, 20b) should be divisible by $\sigma(s)$ that is

$$\bar{\sigma}(s) = \lambda\sigma(s) \quad (21)$$

with λ is a constant, and thus (20a, 20b) reduces to

$$\sigma(s)y''(s) + \tau(s)y'(s) + \lambda y(s) = 0 \quad (22)$$

Equation (22) is called as hypergeometric type equation and its solutions as a functions of hypergeometric type, and (12) is called as generalized hypergeometric type equation [18]. The new parameter $\pi(s)$, which is the first order polynomial, is determined by using (20a, b) and (21) given as

$$\tilde{\sigma}(s) + \pi^2(s) + \pi(s)(\tilde{\tau}(s) - \sigma'(s)) - k\sigma(s) = 0 \quad (23)$$

with

$$\lambda - \pi'(s) = k \quad (24)$$

From (23) we have

$$\pi(s) = \frac{\sigma'(s) - \tilde{\tau}(s)}{2} \pm \sqrt{\left(\frac{\sigma'(s) - \tilde{\tau}(s)}{2}\right)^2 - \tilde{\sigma}(s) + k\sigma(s)} \quad (25)$$

Since the parameter $\pi(s)$ has the form of first order polynomial, then the expression under square root of (25) has to be a perfectly quadratic expression, that means the discriminate of the quadratic expression has to be zero, and so we obtain the value of k .

Before determining the solution of (22) it is necessary to show that the derivative of hypergeometric type differential equation is also a hypergeometric type differential equation. By setting $v_1(s) = y'(s) = \frac{\partial y(s)}{\partial s}$ in (22) that have been differentiated we have

$$\sigma(s)v_1''(s) + \tau_1(s)v_1'(s) + \mu_1 v_1(s) = 0 \quad (26)$$

with

$$\tau_1(s) = \tau(s) + \sigma'(s) \text{ and } \mu_1 = \lambda + \tau'(s) \quad (27)$$

Since $\tau_1(s)$ is the first polynomial and μ_1 is a parameter that s independence therefore (26) is the hypergeometric type differential equation. By repeating the step in obtaining (26) by substituting $v_2(s) = y''(s) = \frac{\partial^2 y(s)}{\partial s^2}$ in (26) that has been differentiated we get

$$\sigma(s)v_2''(s) + \tau_2(s)v_2'(s) + \mu_2v_2(s) = 0 \quad (28)$$

with

$$\tau_2(s) = \tau_1(s) + \sigma'(s) = \tau(s) + 2\sigma'(s) \quad (29a)$$

$$\mu_2 = \mu_1 + \tau_1'(s) = \lambda + 2\tau'(s) + \sigma''(s) \quad (29b)$$

By repeating the differentiation of (22) n times with $v_n(s) = y^{(n)}(s)$ such that we have

$$\sigma(s)v_n''(s) + \tau_n(s)v_n'(s) + \mu_nv_n(s) = 0 \quad (30)$$

and $\tau_n(s)$ dan μ_n yaitu,

$$\tau_n(s) = \tau(s) + n\sigma'(s) \quad (31)$$

$$\mu_n = \lambda + n\tau'(s) + \frac{n(n-1)}{2}\sigma''(s) \quad (32)$$

If $\mu_n = 0$, then from we obtain,

$$\lambda = \lambda_n = -n\tau'(s) - \frac{n(n-1)}{2}\sigma''(s) \quad n = 0, 1, 2, 3, \dots \quad (33)$$

The solution of (22) is obtained from condition that $y(s) = y_n(s)$, which is the nth order polynomial given as

$$y_n(s) = \frac{B_n}{\rho(s)} \frac{d^n}{ds^n} [\sigma^n(s)\rho(s)] \quad (34)$$

with B_n is normalization constant and $\rho(s)$ is weight function that satisfies the condition

$$\frac{d}{ds} [\sigma(s)\rho(s)] = \tau(s)\rho(s) \quad (35)$$

with $\sigma(s)$ and $\tilde{\sigma}(s)$ which are obtained from (12) are polynomials with mostly second order, and $\tilde{\tau}(s)$ is first order polynomial given as

$$\sigma(s) = as^2 + bs + c \text{ and } \tilde{\tau}(s) = fs + h \quad (36)$$

The useful formulas used to determine the energy spectra and the wave function of quantum system using NU method have been derived from hypergeometric differential equation.

4 Application of NU Method for Non-central Potential

By using (16), (17), (24), (25), (33), (34), and (35), the energy spectra and the corresponding wave functions of 3D oscillator harmonics plus trigonometric Rosen-Morse non-central potential and Eckart plus Poschl-Teller non-central potentials are calculated.

4.1 Energy Spectrum and Wave Function of 3D Oscillator Harmonics Plus Trigonometric Rosen-Morse Non-central Potential

The first part of this section discusses the solution of three dimensional Schrodinger equation for 3D HO potential with simultaneously the presence of trigonometric Rosen-Morse non-central potential whose potential is expressed as

$$\begin{aligned} & -\frac{\hbar^2}{2M} \left\{ \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right\} \Psi(r, \theta, \varphi) \\ & + \left\{ \frac{M\omega^2 r^2}{2} + \frac{\hbar^2}{2Mr^2} \left(\frac{v(v+1)}{\sin^2 \theta} - 2\mu \cot \theta \right) \right\} \Psi(r, \theta, \varphi) \} \\ & = E\Psi(r, \theta, \varphi) \end{aligned} \quad (37)$$

The non-central potential is separable one then (37) is solved using variable separation method by setting the wave function in (37) as $\psi(r, \theta, \varphi) = R(r)P(\theta)\phi(\varphi)$ and we obtain

$$\begin{aligned} \frac{1}{R} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) - \frac{r^2}{\hbar^2} M^2 \omega^2 r^2 + \frac{2Mr^2}{\hbar^2} E &= -\frac{1}{P \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) - \frac{1}{\phi \sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \\ &+ \left(\frac{v(v+1)}{\sin^2 \theta} - 2\mu \cot \theta \right) \\ &= \lambda = \ell(\ell+1) \end{aligned} \quad (38)$$

The azimuthal part obtained from the wave equation expressed in (38) is given as

$$\frac{1}{\phi} \frac{\partial^2}{\partial \varphi^2} = -m^2 \quad (39)$$

therefore the solution of azimuthal part of wave function as usual is

$$\phi = A_m e^{im\varphi} \quad (40)$$

From (38) we obtain the radial and angular parts of Schrodinger equation given as

$$\frac{1}{R} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) - \frac{r^2}{\hbar^2} M^2 \omega^2 r^2 + \frac{2Mr^2}{\hbar^2} E = l(l+1) \quad (41)$$

$$-\frac{1}{P \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{m^2}{\sin^2 \theta} + \left(\frac{\nu(\nu+1)}{\sin^2 \theta} - 2\mu \cot \theta \right) = l(l+1) \quad (42)$$

From the solution of (41) will be obtained the energy spectrum of 3D HO potential and the radial part of wave function while from (42) will be obtained the angular wave function and the value of l , orbital quantum number.

4.1.1 Solution of Radial Schrodinger Equation for 3D Oscillator Harmonics Plus Trigonometric Rosen-Morse Non-central Potential

The radial Schrodinger equation in (41) is rewritten as

$$\frac{\partial^2 R}{\partial r^2} + \frac{2}{r} \frac{\partial R}{\partial r} + \left(-\frac{M^2}{\hbar^2} \omega^2 r^2 + \frac{2\mu}{\hbar^2} E \right) R - \frac{l(l+1)}{r^2} R = 0 \quad (43)$$

By setting

$$\frac{2M}{\hbar^2} E = \varepsilon^2 \quad \frac{M^2 \omega^2}{\hbar^2} = \gamma^2 \quad \text{and} \quad R = \frac{\chi}{r} \quad (44)$$

and inserting it into (43) we get

$$\frac{\partial^2 \chi}{\partial r^2} + \left(-\gamma^2 r^2 + \varepsilon^2 - \frac{l(l+1)}{r^2} \right) \chi = 0 \quad (45)$$

Making a change variable $r^2 = x$ in (42) and change it into equation which is given as

$$x \frac{\partial^2 \chi}{\partial x^2} + \frac{1}{2} \frac{\partial \chi}{\partial x} - \left(\frac{\gamma^2 x^2 - \varepsilon^2 x + l(l+1)}{4x} \right) \chi = 0 \quad (46)$$

By comparing (12) and (46) we have

$$\sigma = x, \quad \tilde{\tau} = \frac{1}{2}, \quad \tilde{\sigma} = - \left\{ \frac{l(l+1) + \gamma^2 x^2 - \varepsilon^2 x}{4} \right\} \quad (47)$$

From (25) and (47) we get

$$\pi = \frac{1}{4} \pm \sqrt{\frac{1}{16} + \frac{l(l+1) + \gamma^2 x^2 - \varepsilon^2 x}{4} + kx} \quad (48)$$

The value of k is obtained from the condition that quadratic expression under the square root in (48) has to be completely square of first degree of polynomial therefore (48) is rewritten as

$$\pi = \frac{1}{4} \pm \frac{\gamma}{2} \left\{ x + 2 \frac{k - \frac{\varepsilon^2}{4}}{\gamma^2} \right\} \quad (49)$$

and the discriminate of the quadratic expression under the square root that has to be zero is given as

$$\left(k - \frac{\varepsilon^2}{4} \right)^2 - 4 \left(\frac{\gamma^2}{4} \right) \frac{\left(l + \frac{1}{2} \right)^2}{4} = 0 \quad (50)$$

From (50) we obtain the value of k as

$$k_1 = \frac{\varepsilon^2}{4} + \frac{\gamma}{2} \left(l + \frac{1}{2} \right) \quad \text{or} \quad k_2 = \frac{\varepsilon^2}{4} - \frac{\gamma}{2} \left(l + \frac{1}{2} \right) \quad (51)$$

By imposing the condition that $\tau' < 0$ then from (49) and (51) we get

$$\begin{aligned} \text{for } k_1 \quad \pi &= \frac{1}{4} - \frac{\gamma}{2} \left\{ x + 2 \frac{\frac{\gamma}{2} \left(l + \frac{1}{2} \right)}{\gamma^2} \right\} = -\frac{\gamma x}{2} - \frac{l}{2} \quad \text{or} \quad \pi = \frac{1}{4} - \frac{\gamma}{2} \left\{ x - 2 \frac{\frac{\gamma}{2} \left(l + \frac{1}{2} \right)}{\gamma^2} \right\} \\ &= -\frac{\gamma x}{2} + \frac{(l+1)}{2} \quad \text{for } k_2 \end{aligned} \quad (52)$$

By using (17), (24), (47), (51), and (52) we obtain

$$\begin{aligned} \text{for } k_1 : \lambda &= \frac{\varepsilon^2}{4} + \frac{\gamma}{2} \left(l + \frac{1}{2} \right) - \frac{\gamma}{2} = \frac{\varepsilon^2}{4} + \frac{\gamma}{2} \left(l - \frac{1}{2} \right) \quad \text{or} \\ \lambda &= \frac{\varepsilon^2}{4} - \frac{\gamma}{2} \left(l + \frac{3}{2} \right) \quad \text{for } k_2 \end{aligned} \quad (53)$$

$$\tau = \frac{1}{2} + 2 \left(-\frac{\gamma x}{2} - \frac{1}{2} \right) = -\gamma x - \left(l - \frac{1}{2} \right) \quad \text{for } k_1 \quad \text{or} \quad \tau = -\gamma x + \left(l + \frac{3}{2} \right) \quad \text{for } k_2 \quad (54)$$

From (33), (47) and (54) we obtain the same values of λ_n either for k_1 or k_2 , that is

$$\lambda = \lambda_n = -n_r(-\gamma) = \gamma n_r \quad (55)$$

The energy eigenvalue obtained by equating equations (53) and (55) is given as

$$\frac{\varepsilon^2}{4} = \gamma \left(n_r - \frac{(l - \frac{1}{2})}{2} \right) \quad \text{for } k_1 \quad \text{or} \quad \frac{\varepsilon^2}{4} = \gamma \left(n_r + \frac{(l + \frac{3}{2})}{2} \right) \quad \text{for } k_2 \quad (56)$$

To have physical meaning, the choice of the values of k , π , τ , λ and λ_n are all values for $k = k_2$ in (52)–(56) therefore the energy spectrum of 3D HO plus Rosen-Morse non-central potential is obtained from (44) which is given as

$$\varepsilon^2 = 2\gamma \left(2n_r + \left(l + \frac{3}{2} \right) \right) \rightarrow E = \hbar\omega \left(2n_r + l + \frac{3}{2} \right) \quad (57)$$

where n_r is radial quantum number, l is orbital quantum number and its values depend on the parameters of Rosen-Morse non-central potential. The orbital quantum number obtained from the solution of angular Schrodinger equation is expressed as

$$l = l' = \sqrt{\left(\sqrt{v(v+1) + m^2} + n_l + \frac{1}{2} \right)^2 - \frac{\mu^2}{\left(\sqrt{v(v+1) + m^2} + n_l + \frac{1}{2} \right)^2}} - \frac{1}{2} \quad (58a)$$

We can see from (58a) that for fixed values of principal quantum number, n_l , and radial quantum number n_r , the values of $l = l'$ is not fixed since it depends on the Rosen-Morse's parameter. Since the values of $l' \geq 0$ then from (58a) we obtain the condition that

$$\left(\sqrt{v(v+1)+m^2}+n_l+\frac{1}{2}\right)^4 - \left(\frac{\sqrt{v(v+1)+m^2}+n_l+\frac{1}{2}}{2}\right)^2 \geq \mu^2 \quad (58b)$$

The radial wave functions are calculated using (13), (16), (34), (35), (47), (52), and (54). The first part of the wave function obtained using (16), (47) and (52) is given as

$$\phi(r) = (x)^{\frac{(l'+1)}{2}} e^{-\frac{\gamma x}{2}} \quad (59)$$

and by using (35), (47), and (54) we get the weight function which is given as

$$\rho(x) = x^{l'+\frac{1}{2}} e^{-\gamma x} \quad (60)$$

The second part of radial wave function obtained using (34) and (60) is given as

$$y_{n_r}(x) = \frac{c_{n_r}}{\rho(x)} \frac{d^{n_r}}{dx^{n_r}} (\sigma^{n_r}(x) \rho(x)) = \frac{c_{n_r}}{x^{l'+\frac{1}{2}} e^{-\gamma x}} \frac{d^{n_r}}{dx^{n_r}} (x^{l'+\frac{1}{2}+n_r} e^{-\gamma x}) \quad (61)$$

From (61) we get the un-normalized first four of the second part of radial wave functions given as

$$y_0(x) = 1 \quad (62a)$$

$$y_1(x) = C_1(l' + 1.5 - \gamma x) \quad (62b)$$

$$y_2(x) = C_2((l' + 2.5)(l' + 1.5) - (2l' + 5)\gamma x + \gamma^2 x^2) \quad (62c)$$

$$y_3(x) = C_3((l' + 3.5)(l' + 2.5)(l' + 1.5) - (3l' + 10.5)(l' + 2.5)\gamma x + (3l' + 10.5)(\gamma x)^2 - (\gamma x)^3) \quad (62d)$$

The second part of radial wave function change into associated Laguerre polynomials by setting $\gamma x = z$ in (61), that is

$$y_{n_r}(x) = \frac{c_{n_r}(\gamma)^{l'+\frac{1}{2}}}{z^{l'+\frac{1}{2}} e^{-z}} \frac{d^{n_r}}{dz^{n_r}} (z^{l'+\frac{1}{2}+n_r} e^{-z}) = C_{n_r}(\gamma)^{l'+\frac{1}{2}} n_r! L_{n_r}^{l'+\frac{1}{2}}(z) \quad (63)$$

The un-normalized first four radial wave function obtained using (44), (59) and (62a)–(62d) are given as

$$R_0(x) = C_{n_r} r^{l'} e^{-\frac{\mu_0 r^2}{2\hbar}} \quad (65a)$$

$$R_1(x) = C_1 r^{l'} e^{-\frac{\gamma r^2}{2\hbar}} (l' + 1.5 - \gamma r^2) \quad (65b)$$

$$R_2 = C_1 r^{l'} e^{-\frac{\gamma r^2}{2}} ((l' + 2.5)(l' + 1.5) - (2l' + 5)\gamma r^2 + \gamma^2 r^4) \quad (65c)$$

$$R_3 = C_1 r^{l'} e^{-\frac{\gamma r^2}{2}} ((l' + 3.5)(l' + 2.5)(l' + 1.5) - ((3l' + 10.5)(l' + 2.5)\gamma r^2 + (3l' + 10.5)\gamma^2 r^4 - \gamma^3 r^6)) \quad (65d)$$

The effect of the presence of Rosen-Morse non-central potential to radial wave function is represented by the value of orbital momentum numbers, l' , that are not always be integer but it always be positive number. The normalization factor B_n in (51) can be obtained from the normalization condition of radial wave function which is expressed as

$$\int_0^\infty \chi_{n_r, l'}(r) \chi_{n_r, l'}(r) dr = \delta_{n_r, n_r} \quad (66a)$$

By inserting (51) into (54) we have

$$\int_0^\infty B_{n_r, l'} z^{\frac{(l'+1)}{2}} e^{-\frac{\gamma}{2} z} L_{n_r, l'}^{\frac{l'+1}{2}}(z) B_{n_r, l'} z^{\frac{(l'+1)}{2}} e^{-\frac{\gamma}{2} z} L_{n_r, l'}^{\frac{l'+1}{2}}(z) \frac{dz}{2\sqrt{\beta}z} = \delta_{n_r, n_r} \quad (66b)$$

The normalization condition for associated Laguerre polynomials is given as

$$\int_0^\infty z^{l'+\frac{1}{2}} e^{-z} L_{n_r, l'}^{\frac{l'+1}{2}}(z) L_{n_r, l'}^{\frac{l'+1}{2}}(z) dz = \frac{(n_r + l' + \frac{1}{2})!}{n_r!} \delta_{n_r, n_r} \quad (66c)$$

From (66a) and (66b) we get the normalization factor of radial wave function given as

$$B_{n_r} = \sqrt{\frac{2\sqrt{\gamma} n_r!}{(n_r + l' + \frac{1}{2})!}} \quad (66d)$$

The radial wave function of 3D HO plus Rosen-Morse non-central potential is expressed as associated Laguerre polynomials with the values of orbital quantum numbers are trigonometric Rosen-Morse's parameters dependence. By the absent of Rosen-Morse potential the radial wave function becomes the radial wave function of 3D HO. The effect of the presence of Rosen-Morse potential: the $\csc^2 \theta$ term causes the increase in wave amplitude and the wavelength as shown in Fig. 1, while the $\cot \theta$ term causes the decrease of the amplitude and the wavelength as shown in Fig. 2.

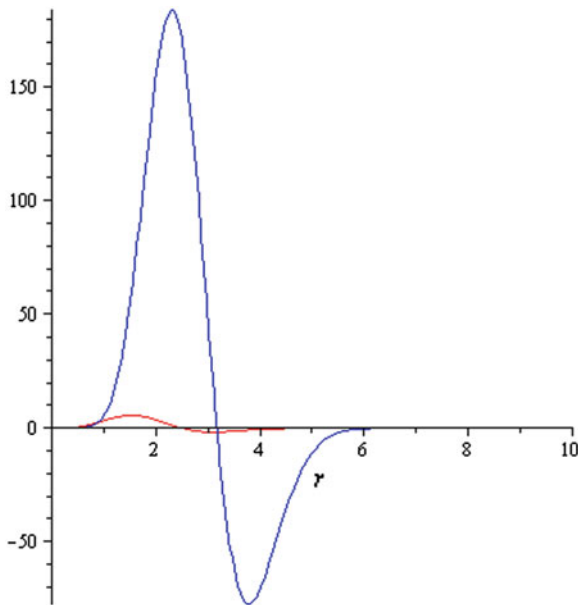


Fig. 1 The graph of 1st state of radial wave function for 3D HO-Rosen-Morse system for $\csc^2 \theta$

term — $R_{n_l m v \mu r}(r) = R_{21201}(r) = C_1 r^{3.65} e^{-\frac{r^2}{2}} (6.15 - \gamma r^2)$ — $R_{n_l m v \mu r}(r) = R_{21201}(r) = C_1 r^{3.65} e^{-\frac{r^2}{2}} (6.15 - \gamma r^2)$

4.1.2 The Solution of Angular Schrodinger Equation

The polar part of the Schrodinger equation expressed in (40) is rewritten as

$$\frac{\partial^2 P(\theta)}{\partial \theta^2} + \cot \theta \frac{\partial P(\theta)}{\partial \theta} - \left(\frac{v(v+1) + m^2}{\sin^2 \theta} - 2\mu \cot \theta \right) P(\theta) + l(l+1)P(\theta) = 0 \quad (67)$$

By making a change of variable, $\cot \theta = is$, in (67) we have

$$\frac{\partial}{\partial \theta} = i(1-s^2) \frac{\partial}{\partial s} \text{ and } \frac{\partial^2}{\partial \theta^2} = i(1-s^2) \frac{\partial}{\partial s} \left\{ i(1-s^2) \frac{\partial}{\partial s} \right\} \quad (68)$$

By inserting (68) into (67) we achieve

$$(1-s^2) \frac{\partial^2 P}{\partial s^2} - s \frac{\partial P}{\partial s} + \left\{ m^2 + v(v+1) - \frac{2\mu is + l(l+1)}{1-s^2} \right\} P = 0 \quad (69)$$

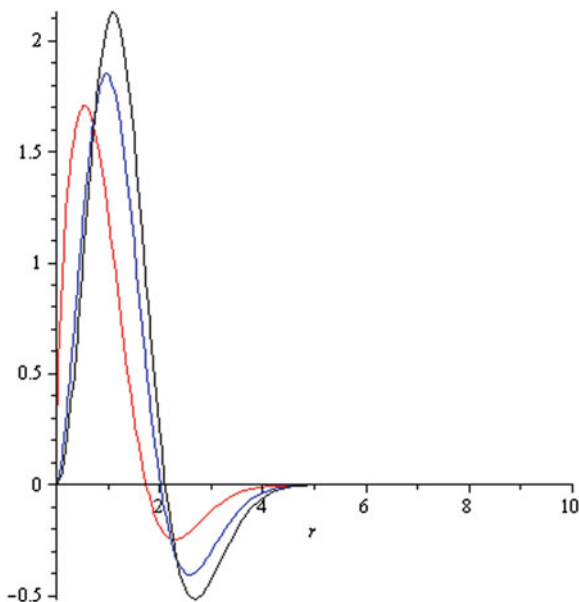


Fig. 2 The graph of 1st state of radial wave function for 3D HO-Rosen-Morse system for

cot term: — $R_{21061}(r) = C_1 r^{1.55} e^{-\frac{r^2}{2}} (4.05 - \gamma r^2)$ — $R_{n_l m v_l \mu r}(r) = R_{21021}(r) = C_1 r^{1.95} e^{-\frac{r^2}{2}} (4.45 - \gamma r^2)$

— $R_{210101}(r) = C_1 r^{0.52} e^{-\frac{r^2}{2}} (3.02 - \gamma r^2)$

$$\sigma = 1 - s^2, \quad \tilde{\tau} = -s, \quad \tilde{\sigma} = \{m^2 + v(v+1)\}(1 - s^2) - \{2\mu is + l(l+1)\} \quad (70)$$

By using (25) and (70) we get

$$\pi = -\frac{s}{2} \pm \sqrt{\frac{s^2}{4} - [m^2 + v(v+1) - l(l+1) - k] + 2\mu is + [m^2 + v(v+1) - k]s^2} \quad (71a)$$

$$\pi = -\frac{s}{2} \pm \sqrt{\left[m^2 + \left(v + \frac{1}{2}\right)^2 - k\right]} \left\{ s + \frac{2\mu i}{2\left\{m^2 + \left(v + \frac{1}{2}\right)^2 - k\right\}} \right\} \quad (71b)$$

The value of k in (71a) is obtained from the condition that quadratic expression under the square root in (71a) has to be perfectly square of first degree of polynomial therefore (71a) reduces to (71b), and the discriminate of the quadratic expression under the square root in (71a) has to be zero given as

$$\begin{aligned}
& -\mu^2 + k^2 - k \left\{ 2m^2 + 2 \left(v + \frac{1}{2} \right)^2 - \left(l + \frac{1}{2} \right)^2 \right\} \\
& + \left\{ m^2 + \left(v + \frac{1}{2} \right)^2 - \left(l + \frac{1}{2} \right)^2 \right\} \left\{ m^2 + \left(v + \frac{1}{2} \right)^2 \right\} \\
& = 0
\end{aligned} \tag{72}$$

The values of k obtained from (72)

$$k_1 = m^2 + \left(v + \frac{1}{2} \right)^2 - \frac{\left(l + \frac{1}{2} \right)^2 + \sqrt{\left(l + \frac{1}{2} \right)^4 + 4\mu^2}}{2} = m^2 + \left(v + \frac{1}{2} \right)^2 - p_1^2 \tag{73a}$$

$$k_2 = m^2 + \left(v + \frac{1}{2} \right)^2 - \frac{\left(l + \frac{1}{2} \right)^2 - \sqrt{\left(l + \frac{1}{2} \right)^4 + 4\mu^2}}{2} = m^2 + \left(v + \frac{1}{2} \right)^2 - p_2^2 \tag{73b}$$

with

$$p_1^2 = \frac{\left(l + \frac{1}{2} \right)^2 + \sqrt{\left(l + \frac{1}{2} \right)^4 + 4\mu^2}}{2} \text{ and } p_2^2 = \frac{\left(l + \frac{1}{2} \right)^2 - \sqrt{\left(l + \frac{1}{2} \right)^4 + 4\mu^2}}{2} \tag{74}$$

By inserting (73a, 73b) into (71b) we obtain the value of π that satisfies the condition $\tau' < 0$ given as

$$\pi = -\frac{s}{2} - p \left(s + \frac{\mu i}{p^2} \right) = -s \left(p + \frac{1}{2} \right) - \frac{\mu i}{p} \tag{75}$$

By inserting (70) and (75) into (17) we have

$$\tau = -s - s - 2p \left(s + \frac{\mu i}{p^2} \right) = -2s(1 + p) - \frac{2\mu i}{p} \tag{76}$$

The eigen values of the system are obtained by manipulating (24), (33), (70), (73a, 73b), (74), and (76) as follows. By using (24), (73a, 73b) and (74) we have

$$\lambda = m^2 + \left(v + \frac{1}{2} \right)^2 - p^2 - \left(p + \frac{1}{2} \right) \tag{77}$$

and by using (33), (70), and (76) we obtain

$$\lambda_n = -n(-2p - 2) + n(n - 1) = 2np + n + n^2 \tag{78}$$

By equating equation (77) and (78) we have

$$m^2 + \left(v + \frac{1}{2}\right)^2 - p^2 - \left(p + \frac{1}{2}\right) = 2np + n + n^2 \rightarrow m^2 + v(v+1) = \left(p + n + \frac{1}{2}\right)^2 \quad (79)$$

By inserting (74) into (79) we get

$$\pm \sqrt{m^2 + v(v+1)} - \left(n + \frac{1}{2}\right) = p = \pm \sqrt{\frac{\left(l + \frac{1}{2}\right)^2 \pm \sqrt{\left(l + \frac{1}{2}\right)^2 + 4\mu^2}}{2}} \quad (80)$$

To have physical meaning, from (80) we choose

$$\begin{aligned} -\sqrt{m^2 + v(v+1)} - \left(n + \frac{1}{2}\right) &= p = -\sqrt{\frac{\left(l + \frac{1}{2}\right)^2 + \sqrt{\left(l + \frac{1}{2}\right)^2 + 4\mu^2}}{2}} \\ &\rightarrow \left\{ \sqrt{m^2 + v(v+1)} + \left(n + \frac{1}{2}\right) \right\}^2 \\ &= \frac{\left(l + \frac{1}{2}\right)^2 + \sqrt{\left(l + \frac{1}{2}\right)^2 + 4\mu^2}}{2} \end{aligned} \quad (81)$$

that gives

$$\left(l + \frac{1}{2}\right) = \sqrt{\left(\sqrt{m^2 + v(v+1)} + n + \frac{1}{2}\right)^2 - \frac{\mu^2}{\left(\sqrt{m^2 + v(v+1)} + n + \frac{1}{2}\right)^2}} \quad (82)$$

Equation (82) shows that l , disturbed orbital momentum number, as a function of m , μ , v , and $n = n_l$ which is angular quantum number.

The first part of the polar wave function obtained from (16), (70) and (75) as follows:

$$\frac{\phi'}{\phi} = \frac{-s\left(p + \frac{1}{2}\right) - \frac{\mu i}{p}}{1 - s^2} = \frac{-s\left(p + \frac{1}{2}\right)}{1 - s^2} + \frac{-\frac{\mu i}{p}}{2(1 - s)} + \frac{-\frac{\mu i}{p}}{2(1 + s)} \quad (83)$$

$$\frac{d\phi}{\phi} = \frac{\frac{(p+\frac{1}{2})}{2}d(-s^2)}{1 - s^2} + \frac{\frac{\mu i}{p}d(-s)}{2(1 - s)} + \frac{-\frac{\mu i}{p}ds}{2(1 + s)} \rightarrow \phi = (1 - s)^{\frac{(p+\frac{1}{2})}{2} + \frac{\mu i}{2p}}(1 + s)^{\frac{(p+\frac{1}{2})}{2} - \frac{\mu i}{2p}} \quad (84)$$

The weight function of the second part of wave function is obtained from (35), (70), and (76) given as

$$\frac{\partial(\sigma\rho)}{\partial r} = \tau(r)\rho(r) - 2s\rho + (1-s^2)\rho' = \left(-2s(p+1) - \frac{2\mu i}{p}\right)\rho \quad (85)$$

$$\frac{\rho'}{\rho} = \frac{-2sp - \frac{2\mu i}{p}}{1-s^2} = \frac{-2sp}{1-s^2} + \frac{-\frac{2\mu i}{p}}{2(1-s)} + \frac{-\frac{2\mu i}{p}}{2(1+s)} \rightarrow \rho = (1-s)^{p+\frac{\mu i}{p}}(1+s)^{p-\frac{\mu i}{p}} \quad (86)$$

The second angular part obtained using (34), (70), and (86) is given as

$$y_n(s) = \frac{B_n}{(1-s)^{p+\frac{\mu i}{p}}(1+s)^{p-\frac{\mu i}{p}}} \frac{d^n}{ds^n} \left((1-s)^{p+\frac{\mu i}{p}+n} (1+s)^{p-\frac{\mu i}{p}+n} \right) \quad (87)$$

The polar wave function obtained from (13), (84) and (87) is

$$P(s) = (1-s)^{\frac{(p+\frac{1}{2})}{2}+\frac{\mu i}{2p}} (1+s)^{\frac{(p+\frac{1}{2})}{2}-\frac{\mu i}{2p}} y_n(s) \quad (88)$$

The total wave function of the system obtained from (40), (44), and (88) is

$$\Psi(r, s, \varphi) = B_{n_r} \gamma^{\frac{l}{2}} r^{\frac{l}{2}} e^{-\frac{l}{2} L_{n_r}} (\gamma r^2) (1-s)^{\frac{(p+\frac{1}{2})}{2}+\frac{\mu i}{2p}} (1+s)^{\frac{(p+\frac{1}{2})}{2}-\frac{\mu i}{2p}} y_n(s) e^{im\varphi} \quad (90)$$

with $s = -i \cot \theta$, and the energy spectrum is expressed in (57) and (58a).

4.2 Energy Spectrum and Wave Function of Eckart Plus Poschl-Teller Non-central Potential

The Schrodinger equation for for Eckart plus Poschl-Teller non-central potential is given as

$$\begin{aligned} & -\frac{\hbar^2}{2M} \left\{ \frac{1}{r^2} \frac{\partial}{\partial r} \left((r^2 \frac{\partial}{\partial r}) \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} (\sin \theta \frac{\partial}{\partial \theta}) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right\} \Psi(r, \theta, \varphi) \\ & + \frac{\hbar^2}{2Ma^2} \left(\frac{V_0 e^{\frac{-r}{a}}}{(1 - e^{\frac{-r}{a}})^2} - \frac{V_1 (1 + e^{\frac{-r}{a}})}{(1 - e^{\frac{-r}{a}})} \right) \Psi(r, \theta, \varphi) \} \\ & + \frac{\hbar^2}{2Mr^2} \left(\frac{\kappa(\kappa-1)}{\sin^2 \theta} + \frac{\eta(\eta-1)}{\sin^2 \theta} \right) \Psi(r, \theta, \varphi) \\ & = E \Psi(r, \theta, \varphi) \end{aligned} \quad (91)$$

The three dimensional Schrodinger equation expressed in (12) is solved using variable separation method by setting $\psi(r, \theta, \varphi) = R(r)P(\theta)\phi(\varphi)$ so we get

$$\begin{aligned} \frac{1}{R} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R}{\partial r} \right) - \frac{r^2}{a^2} \left(\frac{V_0 e^{\frac{-r}{a}}}{(1 - e^{\frac{-r}{a}})^2} - \frac{V_1 (1 + e^{\frac{-r}{a}})}{(1 - e^{\frac{-r}{a}})} \right) + \frac{2MEr^2}{\hbar^2} \\ = - \frac{1}{P \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial P}{\partial \theta} \right) - \frac{1}{\Phi \sin^2 \theta} \frac{\partial^2 \Phi}{\partial \varphi^2} + \frac{\kappa(\kappa - 1)}{\sin^2 \theta} + \frac{\eta(\eta - 1)}{\sin^2 \theta} = \lambda = l(l + 1) \end{aligned} \quad (92)$$

From (92) we get three differential equations with single variable as following:

$$\frac{1}{R} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R}{\partial r} \right) - \frac{r^2}{a^2} \left(\frac{V_0 e^{\frac{-r}{a}}}{(1 - e^{\frac{-r}{a}})^2} - \frac{V_1 (1 + e^{\frac{-r}{a}})}{(1 - e^{\frac{-r}{a}})} \right) + \frac{2MEr^2}{\hbar^2} = \lambda = l(l + 1) \quad (93a)$$

$$- \frac{1}{P \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial P}{\partial \theta} \right) + \frac{m^2}{\sin^2 \theta} + \frac{\kappa(\kappa - 1)}{\sin^2 \theta} + \frac{\eta(\eta - 1)}{\sin^2 \theta} = \lambda = l(l + 1) \quad (93b)$$

and

$$\frac{1}{\Phi} \frac{\partial^2 \Phi}{\partial \varphi^2} = -m^2 \rightarrow \Phi = A_m e^{im\varphi} \quad (94)$$

The radial part of Schrodinger equation is given as

By setting $\frac{2m}{\hbar^2} E = -\epsilon^2$, $R = \frac{\chi(r)}{r}$, applying an approximation for centrifugal term [14, 15], $\frac{1}{r^2} \cong \frac{1}{4a^2} \left(d_0 + \frac{1}{\sinh^2 \frac{r}{2a}} \right)$ for $\frac{r}{2a} \ll 1$ and $d_0 = \frac{1}{12}$ and changing the exponential term into hyperbolic function in (93a) we get

$$\frac{d^2 \chi(r)}{dr^2} - \frac{1}{a^2} \left\{ \frac{V_0 + l(l + 1)}{4 \sinh^2 \frac{r}{2a}} - V_1 \coth \frac{r}{2a} + \frac{l(l + 1)d_0 + 4a^2 \epsilon^2}{4} \right\} \chi(r) = 0 \quad (95)$$

By making a coordinate transformation, $r = f(s) = 2a \coth^{-1}(1 - 2s)$, in (95) we obtain

$$\begin{aligned} s(1 - s) \frac{\partial^2 \chi}{\partial s^2} + (1 - 2s) \frac{\partial \chi}{\partial s} \\ + \left\{ \frac{(V_0 + l(l + 1))s(1 - s)}{s(1 - s)} + \frac{V_1(1 - 2s)}{s(1 - s)} - \frac{\left(\frac{l(l + 1)d_0}{4a^2} + \epsilon^2 \right) a^2}{s(1 - s)} \right\} \chi \\ = 0 \end{aligned} \quad (96)$$

From (96) we get

$$\tilde{\tau} = 1 - 2s, \quad \sigma = s(1 - s) \quad (97a)$$

$$\tilde{\sigma} = (V_0 + l(l + 1))s(1 - s) + V_1(1 - 2s) - \left(\frac{l(l + 1)d_0}{4a^2} + \varepsilon^2 \right) a^2 \quad (97b)$$

Inserting (18) and (19) into (5) we have

$$\pi = \pm \sqrt{(V_0 + l(l + 1) - k)s^2 - (V_0 + l(l + 1) - 2V_1 - k)s - \left(V_1 - \left(\frac{l(l + 1)d_0}{4a^2} + \varepsilon^2 \right) a^2 \right)} \quad (98)$$

Due to the condition that the expression under the square root of (98) must be square of first degree polynomial, then (98) is rewritten as

$$\pi = \pm \sqrt{(V_0 + l(l + 1) - k)s^2} \left(s - \frac{V_0 + l(l + 1) - 2V_1 - k}{2(V_0 + l(l + 1) - k)} \right) \quad (99)$$

and the discriminate of the quadratic expression under the square root has to be zero, that is

$$\begin{aligned} & (V_0 + l(l + 1) - 2V_1 - k)^2 + 4(V_0 + l(l + 1) - k) \left(V_1 - \left(\frac{l(l + 1)d_0}{4a^2} + \varepsilon^2 \right) a^2 \right) \\ & = 0 \end{aligned} \quad (100)$$

From (100) we get

$$\begin{aligned} k &= (V_0 + l(l + 1) - 2\left(\frac{l(l + 1)d_0}{4a^2} + \varepsilon^2 \right) a^2) \pm 2\sqrt{\left(\frac{l(l + 1)d_0}{4a^2} + \varepsilon^2 \right)^2 a^4 - V_1^2} \\ &= A - C \end{aligned} \quad (101)$$

with

$$A = V_0 + l(l + 1) \text{ and } C = 2\left(\frac{l(l + 1)d_0}{4a^2} + \varepsilon^2 \right) a^2 \mp 2\sqrt{\left(\frac{l(l + 1)d_0}{4a^2} + \varepsilon^2 \right)^2 a^4 - V_1^2} \quad (102)$$

By imposing $\tau' < 0$ we have

$$\pi = -\sqrt{C} \left(s - \frac{C - 2V_1}{2C} \right) \text{ and } \tau = 1 - 2s - 2\sqrt{C} \left(s - \frac{C - 2V_1}{2C} \right) \quad (103)$$

Using (6) and (7) together with the values of k , π , τ , and σ we get

$$\lambda = k + \pi' = A - C - \sqrt{C} \quad (104)$$

and

$$\lambda_n = -n\tau' - \frac{n(n-1)}{2}\sigma'' = -n(-2 - 2\sqrt{C}) + n(n-1) \quad (105)$$

By equating equations (104) and (105) we have

$$A - C - \sqrt{C} = n + n^2 + 2n\sqrt{C} \rightarrow V_0 + (l + \frac{1}{2})^2 = (n + \sqrt{C} + \frac{1}{2})^2 \quad (106a)$$

$$\begin{aligned} \sqrt{C} &= \pm \sqrt{V_0 + (l + \frac{1}{2})^2 - (n + \frac{1}{2})} \\ &= \pm \sqrt{2(\frac{l(l+1)d_0}{4a^2} + \varepsilon^2)a^2} \mp 2\sqrt{(\frac{l(l+1)d_0}{4a^2} + \varepsilon^2)^2 a^4 - V_1^2} \end{aligned} \quad (106b)$$

The proper choice in (106b) that has physical meaning is

$$\begin{aligned} \sqrt{C} &= -\sqrt{V_0 + (l + \frac{1}{2})^2 - (n + \frac{1}{2})} \\ &= -\sqrt{2(\frac{l(l+1)d_0}{4a^2} + \varepsilon^2)a^2} + 2\sqrt{(\frac{l(l+1)d_0}{4a^2} + \varepsilon^2)^2 a^4 - V_1^2} \end{aligned} \quad (107)$$

and the energy spectrum produced is

$$E = -\frac{\hbar^2}{2M} \left\{ \frac{\left(\sqrt{V_0 + (l + \frac{1}{2})^2} + (n_r + \frac{1}{2}) \right)^2}{4a^2} + \frac{V_1^2}{a^2 \left(\sqrt{V_0 + (l + \frac{1}{2})^2} + (n_r + \frac{1}{2}) \right)^2} - \frac{l(l+1)d_0}{4a^2} \right\} \quad (108)$$

The energy spectrum of Eckart potential with the absent of Poschl-Teller potential is produced from (108).

The first part of the wave function is obtained by using (97a) and (103)

$$\begin{aligned} \frac{\phi'}{\phi} = \frac{\pi}{\sigma} &= \frac{-\sqrt{C}\left(s - \frac{C-2V_1}{2C}\right)}{s(1-s)} = \frac{-\sqrt{C}}{1-s} + \frac{C-2V_1}{s2\sqrt{C}} + \frac{C-2V_1}{(1-s)2\sqrt{C}} \rightarrow \phi(s) \\ &= s^{\frac{C-2V_1}{2\sqrt{C}}} (1-s)^{\frac{C+2V_1}{2\sqrt{C}}} \end{aligned} \quad (109)$$

The weight function of the radial part of the system is obtained using (35), (97a), and (103)

$$\begin{aligned} \frac{\partial(\sigma\rho)}{\partial s} &= \tau(s)\rho(s) \rightarrow (1-2s)\rho + s(1-s)\rho' \\ &= \left\{1 - 2s - 2\sqrt{C}\left(s - \frac{C-2V_1}{2C}\right)\right\}\rho \end{aligned}$$

that gives

$$\rho(s) = s^{\frac{C-2V_1}{\sqrt{C}}} (1-s)^{\frac{C+2V_1}{\sqrt{C}}} \quad (110)$$

The second part of the wave function is derived from Rodrigues relation expressed in (34). By inserting (97a) and (110) into (34) and by setting $\sqrt{C} = p$, C is (107) we get

$$y_n(s) = \frac{B_n}{s^{p-\frac{2V_1}{p}}(1-s)^{p+\frac{2V_1}{p}}} \frac{d^n}{ds^n} \left\{ s^{p-\frac{2V_1}{p}+n} (1-s)^{p+\frac{2V_1}{p}+n} \right\} \quad (111)$$

We finally obtain the complete wave functions from (16), (109) and (111) and with $\coth(r/2a) = 1 - 2s$ as

$$\chi(s) = s^{\frac{C-2V_1}{2\sqrt{C}}} (1-s)^{\frac{C+2V_1}{2\sqrt{C}}} y_n(s), \quad (112)$$

4.2.1 The Polar Schrodinger Equation Solution

By making a variable transformation $\cos 2\theta = s$ in (93b) we get

$$\begin{aligned} (1-s^2) \frac{\partial^2 P}{\partial s^2} - \left(\frac{1}{2} + \frac{3}{2}s\right) \frac{\partial P}{\partial s} \\ - \left(\frac{2[\kappa(\kappa-1) + m^2](1+s)}{4(1-s^2)} + \frac{2\eta(\eta-1)(1+s)}{4(1-s^2)} - \frac{l(l+1)(1-s^2)}{4(1-s^2)}\right) P \\ = 0 \end{aligned} \quad (114)$$

The form of (114) is similar to the (12). The orbital momentum number and the polar wave function are achieved from (114) by applying (12), (13), (16), (17), (22),

(24), (25), (33)–(35), with the solution steps similar to steps in Sect. 4.1.2 or 4.2.1. From (114) we have

$$\sigma = 1 - s^2 \tilde{\tau} = -\left(\frac{1}{2} + \frac{3}{2}s\right) \quad (115a)$$

$$\tilde{\sigma} = -\left\{\frac{[2(\kappa(\kappa-1) + m^2) + 2\eta(\eta-1) - l(l+1)]}{4} + \frac{[2(\kappa(\kappa-1) + m^2) - 2\eta(\eta-1)]s}{4} + \frac{l(l+1)s^2}{4}\right\} \quad (115b)$$

By imposing that $\tau' < 0$ and applying the condition of (25), then from (25) and (115a, 115b) we obtain

$$\pi = \frac{1-s}{4} - \sqrt{\left(\frac{l(l+1)}{4} - k + \frac{1}{16}\right)} \left(s + \frac{\frac{[2(\kappa(\kappa-1) + m^2) - 2\eta(\eta-1)]}{4} - \frac{1}{8}}{2\left(\frac{l(l+1)}{4} - k + \frac{1}{16}\right)}\right) \quad (116)$$

and

$$\begin{aligned} & \left\{\frac{2(\kappa(\kappa-1) + m^2) - 2\eta(\eta-1)}{4} - \frac{1}{8}\right\}^2 \\ & - 4\left\{\frac{l(l+1)}{4} - k + \frac{1}{16}\right\} \left\{\frac{[2(\kappa(\kappa-1) + m^2) + 2\eta(\eta-1) - l(l+1)]}{4} + \frac{1}{16}\right\} \\ & = 0 \end{aligned} \quad (117)$$

The value of k obtained from (117) is

$$\begin{aligned} k &= \left(\frac{l+1/2}{2}\right)^2 - \left(\frac{\sqrt{\kappa(\kappa-1) + m^2} \pm (\eta - 1/2)}{4}\right)^2 \\ &= \left(\frac{l+1/2}{2}\right)^2 - \frac{(\sqrt{o} \pm \sqrt{t})^2}{2} \end{aligned} \quad (118)$$

with

$$o = \frac{2(\kappa(\kappa-1) + m^2)}{4} \text{ and } t = \frac{2\eta(\eta-1)}{4} + \frac{1}{8} = \frac{(\eta-1/2)^2}{2} \quad (119)$$

By inserting (118) into (116) we have

$$\pi = -s \left(\frac{\sqrt{o} \pm \sqrt{t}}{\sqrt{2}} + \frac{1}{4} \right) - \frac{\sqrt{o} \mp \sqrt{t}}{\sqrt{2}} + \frac{1}{4} \quad (120)$$

By using (17), (115a) and (120) we get

$$\tau = -2s \left(\frac{\sqrt{o} \pm \sqrt{t}}{\sqrt{2}} + 1 \right) - 2 \frac{\sqrt{o} \mp \sqrt{t}}{\sqrt{2}} \quad (121)$$

By using (24), (33), (115a), (118), (120), and (121) we obtain

$$\lambda = \frac{(l + 1/2)^2}{4} - \left(\frac{\sqrt{o} \pm \sqrt{t}}{\sqrt{2}} \right)^2 - \left(\frac{\sqrt{o} \pm \sqrt{t}}{\sqrt{2}} + \frac{1}{4} \right) \quad (122)$$

and

$$\lambda_n = 2n \left(\frac{\sqrt{o} \pm \sqrt{t}}{\sqrt{2}} + 1 \right) + n(n + 1) \quad (123)$$

By equating (122) and (123) and also together with (119) we get l and the proper choice of l given as

$$l = \sqrt{\kappa(\kappa - 1) + m^2} + (\eta - 1/2) + 2n + 1 - 1/2 = \sqrt{\kappa(\kappa - 1) + m^2} + \eta + 2n_l \quad (124)$$

The values of π and τ corresponding to the proper choice of l are

$$\pi = -s \left(\frac{\sqrt{o} + \sqrt{t}}{\sqrt{2}} + \frac{1}{4} \right) - \frac{\sqrt{o} - \sqrt{t}}{\sqrt{2}} + \frac{1}{4} \quad (120a)$$

$$\tau = -2s \left(\frac{\sqrt{o} + \sqrt{t}}{\sqrt{2}} + 1 \right) - 2 \frac{\sqrt{o} - \sqrt{t}}{\sqrt{2}} \quad (121a)$$

The first part of the wave function obtained by using (16), (115a) and (120) is

$$\phi(s) = (1 - s)^{\sqrt{\frac{\kappa}{2}}}(1 + s)^{\sqrt{\frac{\kappa}{2}} + \frac{1}{4}} \quad (125)$$

The weight function for the second part of the wave function obtained by using (35) (115a) and (121a) is

$$\rho(s) = (1-s)^{2\sqrt{\frac{g}{2}}}(1+s)^{2\sqrt{\frac{l}{2}}} \quad (126)$$

By using (34), (115a) and (126) we obtain the second part of the polar wave function which is expressed in term of Jacobi polynomials given as

$$y_n(s) = \frac{B_n}{(1-s)^{2\sqrt{\frac{g}{2}}}(1+s)^{2\sqrt{\frac{l}{2}}}} \frac{d^n}{ds^n} \left\{ (1-s)^{2\sqrt{\frac{g}{2}}+n} (1+s)^{2\sqrt{\frac{l}{2}}+n} \right\} \quad (127)$$

The total polar wave function achieved from (13), (125) and (127) is given as

$$\Psi(s) = B_n (1 - \cos 2\theta)^{-\sqrt{\frac{g}{2}}} (1 + \cos 2\theta)^{-\sqrt{\frac{l}{2}+1}} \frac{d^n}{d(\cos 2\theta)^n} \left\{ (1 - \cos 2\theta)^{2\sqrt{\frac{g}{2}}+n} (1 + \cos 2\theta)^{2\sqrt{\frac{l}{2}}+n} \right\} \quad (128)$$

The total wave function of Eckart plus Poschl-Teller non-central potential found from (112) and (128) and the corresponding energy spectrum is expressed in (108).

The NU method is method developed based on hypergeometric differential equation but the application is wider since NU method is also applicable for problems that usually solved by confluent hypergeometric differential equation as for 3D harmonics oscillation.

Acknowledgments This work is partially supported by Hibah Pascasarjana Sebelas Maret University grant No: 2340/UN 27.10/PG/2012.

References

1. Sadeghi J (2007) Act Phys Pol A **112**(1):23
2. Sadeghi J, Pourhassan B (2008) EJTP **5**(17):193
3. Amani AR, Moghrimoazzen MA, Ghorbanpour H, Barzegaran S (2011) Af J Math Phys **10**:31–37
4. Salehi H (2011) App Math **2**:999–1004. doi:[10.4236/am.2011.28138](https://doi.org/10.4236/am.2011.28138)
5. Sheng JC, Ying Z, Lin ZX, Tian SL (2001) Commun Theor Phys **36**:641–646
6. Dutt R, Khare A, Sukhatme UP (1988) Am J Phys **56**:163
7. Chun-Seng J, Ying Z, Xiang-Lin Z, Liang Tian S (2001) Commun Theor Phys **36**(6):641–646
8. Gendenshtein L (1983) JETP Lett **38**:356
9. Ikhdair SM (2012) ISRN Math Phys **2012**(201525):20. doi:[10.5402/2012/201525](https://doi.org/10.5402/2012/201525)
10. Ikot AN, Akpabio LE, Obu JA (2011) JVR **6**(1):1–13
11. Halberg AS (2011) Int J Maths Math Sci **1**–9. doi:[10.1155/2011/358198](https://doi.org/10.1155/2011/358198)
12. Aktas M (2007) [arXiv:0701063v1](https://arxiv.org/abs/0701063) [quant-ph], pp 1–14

13. Cari C, Suparmi A (2012) IOSR-JAP 2(3):13–23. ISSN:2278-4861, doi:109790/4861-0231323
14. Movahedi M, Hamzavi M (2001) Int J Phys Sci 6(4):891–896
15. Alvarez-Castillo DE, Compean CB, Kirchbach M (2011) [arXiv:1105.1354v1](#) [quant-ph], pp 1–10
16. Alvarez –Castillo DE, Kirbach M (2007) Rev Mex Fis E53(2):143–154
17. Suparmi A, Cari C, Handhika J, Yanuarief, Marini H (2012) IOSR-JAP 2(2):43–51. ISSN:2278-4861, doi:109790/4861-0224351
18. Nikiforov AF, Uvarov UB (1988) Special function in mathematical physics. Birkhausa, Basel
19. Xian-Quan H, Guang L, Zhi-Min W, Lian-Bin N, Yan M (2010) Commun Theor Phys 53 (2):242–246
20. Salehi H (2011) Ap Math 2:999–1004. doi:10.4236/am.2011.28138
21. Aktas M (2007) [arXiv:quant-ph/0701063v1](#)
22. Zhang MC (2009) Int J Theor Phys 48:2625–2632
23. Aktas M (2007) [arXiv:0701063v1](#) [quant-ph], pp 1–14
24. Quan HX, Guang, Min WZ, Bin NL, Yan MA (2010) Commun Theor Phys 53:242–246
25. Shojaei MR, Rajabi AA (2011) Int J Phys Sci 6(33):7441–7446. doi:10.5897/IJPS11.653
26. Rosen N, Morse PM (1932) Phys Rev 42:2
27. Alvarez-Castillo DE, Compean CB, Kirchbach M (2011) [arXiv:1105.1354v1](#) [quant-ph], pp 1–10
28. Agboola D (2011) Pram J Phys 76(6):875–885
29. Compean CB, Kirchbach M (2005) [arXiv:quant-ph/0509055v2](#), pp 1–15
30. Qiang WC, Dong SH (2007) Phys Lett A 368(1–2):13–17

Charge Transport Mechanism and Low Temperature Phase Transitions in KIO_3

M.M. Abdel Kader, F. El-Kabbany, H.M. Naguib
and W.M. Gamal

Abstract Our report deals with the measurement of some electrical properties, namely the ac conductivity $\sigma(\omega, T)$ and the complex dielectric permittivity $\varepsilon^*(\omega, T)$ in the temperature interval $95 \text{ K} < T < 280 \text{ K}$ and at some selected frequencies (0.7–20 kHz) for polycrystalline samples of potassium iodate KIO_3 using a computerized RLC meter. The improper character of the ferroelectricity over the mentioned temperature range has been achieved by recording the ferroelectric hysteresis loops. The temperature dependence of each electrical parameter reveals that the compound undergoes two phase transitions at $T \approx 258 \text{ K}$ and at $T \approx 110 \text{ K}$. The frequency dependent conductivity seems to be in accordance with the power law $\sigma(\omega, T) \propto \omega^{s(T)}$ and the trend of temperature dependence of the frequency exponent s ($0 < s < 1$) suggests that the quantum mechanical tunneling (QMT) model is the main mechanism of the charge transport. Comparison with the behavior of the NH_4IO_3 in the same temperature range was outlined.

1 Introduction

Potassium iodate KIO_3 is a member of the monovalent metal iodate series of the general molecular formula MIO_3 where M stands for Li, Cs, Na, K... and/or NH_4 . This series of compounds attracted the attention of investigators due to their interesting properties. For example, the improper ferroelectric hydrogen-bonded NH_4IO_3 undergoes a ferroelectric phase transition at $\approx 368 \text{ K}$ [1–3]. Moreover, since the compound behaves as a proton conductor, therefore it has some practical applications such as chemical sensors, electrochromic displays and supercapacitors [3, 4]. Furthermore, LiIO_3 undergoes successive phase transitions and the high temperature phase behaves as a superionic conductor [5, 6]. Properties such as

M.M. Abdel Kader (✉) · F. El-Kabbany · H.M. Naguib · W.M. Gamal
Physics Department, Faculty of Science, Cairo University, Giza, Egypt
e-mail: m_m_abdelkader@yahoo.com

piezoelectric, pyroelectric and non-linear optical effects are common for some members of this series [7–11].

For KIO_3 , which is the subject of this article, it is known that most of the physical properties of the crystals of this compound such as, (dc) dielectric constant [12] piezoelectric and pyroelectric effects [13, 14], electro-optical behavior [13], non-linear optical properties [7, 14], ferroelectricity [15, 16] in addition to nuclear quadruple resonance (NQR) [12] and Raman scattering experiment [17], have been investigated previously. Later on, some of these properties were discussed [18]. Due to the increasing interest of these properties and the expected applications of such a material in laser technology, it is logical to re-examine these properties using more advanced techniques. For example, a somewhat recent work dealing with the temperature dependence for piezoelectric and ferroelastic properties, using resonance—antiresonance method [19], has been reported. The data indicate that the compound undergoes successive five phase transitions (I–II, II–III, III–IV, IV–V, V–VI) at the transition temperatures: $T_1 = 485$ K, $T_2 = 345.5$ K, $T_3 = 258$ K, $T_4 = 113$ K and $T_5 = 33$ K respectively. Similar behavior was observed earlier [12] with somewhat slight different in the value of the transition temperature for some phases. The anomaly that has been observed around ≈ 150 °C (423 K) by Crane [14], Herlach [12], Maeda et al. [19] and more recently at (428 ± 2) K by the present authors [20] has been already explained as a change in electrical conduction between extrinsic and intrinsic mechanisms [20] and hence it is not related to any phase transition [19, 20].

Among other things, KIO_3 single crystal exhibits noble non-linear optical properties in all phases [7]. Furthermore, its non-linear optical coefficient in the room temperature phase (phase III), is the largest between monovalent metal iodates (MIO_3) [7, 21] and is also greater than those of KH_2PO_4 (KDP), BaB_2O_4 (BBO) and LiBrO_5 (LBO), which are known as double frequency materials in the ultra-violet wavelength [22]. Although the presence of ferroelectric domains and twins at room temperature (phase III) prevents or at least hinders practical applications, yet using the poling technique and specially after detwenning and domain removal, it can be used for fabrication of non-linear optical devices [22].

Another point of interest is that the static piezoelectric constants of the single crystals of KIO_3 in phase III are about 50 times greater than those of α -quartz [23]. Due to this large piezoelectric effect, KIO_3 is a promising candidate for the fabrication of piezoelectric sensors [24]. Furthermore, the thermal characterization including specific heat and thermal decomposition in the temperature interval $260 \text{ K} < T < 600 \text{ K}$ were investigated [25] using Perkin—Elmer DSC.

To clarify the dynamic properties of the high temperature phase transition of KIO_3 crystals, the temperature dependence of low frequency (soft) optical modes was measured by Raman scattering [24]. More recently, detailed polarized Raman study were performed on KIO_3 single crystals over 78.5–553 K [26]. The improper character of ferroelectricity of the compound was suggested by Ahart et al. [24], supported by Liu et al. [26] and confirmed by Abdel-Kader et al. [20].

From the accurate measurements of some mentioned properties as a function of temperature [15, 17, 19, 20, 22–25], it has been reported that, the KIO_3 undergoes a

ferroelectric phase transition around 486 K and several phase transitions as summarized in Table 1. The table shows also the symmetry of some phases of KIO_3

Regarding the crystal structure, it is known that the structure of phase III using the single crystal X-ray diffraction technique, was determined by Crane [14] who also summarized the results of studies on KIO_3 until 1971 [14]. Furthermore, based on X-ray powder diffraction data, Hamid [27] suggested the space group of phase I

Table 1 Summary of phase transitions in KIO_3

Phase transition	Symmetry	Transition temperature (TK)	Technique or property	References
I–II	I Rhombohedral ($R\bar{3}m$, $Z = 1$) [20] Rhombohedral ($R\bar{3}$, $Z = 1$) [30] Rhombohedral ($R\bar{3}m$, $Z = 2$) [18] Rhombohedral ($R\bar{3}m$, $Z = 2$) [17]	485 K	Raman Spectroscopy	[24]
			Elastic and piezoelectric	[19]
			Detewenning and domain removing procedure	[22]
			Ferroelectricity and optical properties.	[15]
		(486 $\pm$ 1) K	Ac conductivity	[20]
II–III	II Monoclinic (Cm , $Z = 8$) [18] Monoclinic (Pm , $Z = 4$) [17]	345 K	Raman Spectroscopy	[24]
			Ac conductivity	[20]
			Thermal properties	[25]
			Ferroelectricity and optical properties.	[15]
		345.5 K	Elastic and piezoelectric	[19]
		345.6 K	Elastic properties	[23]
		343 K	Detewenning and domain removing procedure	[22]
III–IV	III Triclinic (PI , $Z = 4$) [31] Triclinic (PI , $Z = 4$) [17] Triclinic (PI , $Z = 4$) [18]	258 K	Ac conductivity	Present work
			Raman Spectroscopy	[17]
			Elastic and piezoelectric	[19]
IV–V	–	110 K	Ac conductivity	Present work
			Raman Spectroscopy	[17]
		113 K	Elastic and piezoelectric	[19]
V–VI	–	33 K	Elastic and piezoelectric	[19]

as R3 m which has not been yet confirmed. He also suggested the symmetry of phase II and determined the structure of Phase I from intensities taken by photographic method. It is known that, single crystals X-ray diffraction technique, has been faced with the some problems due to the existence of ferroelectric and/or ferroelastic domains [27].

Again single crystal technique has been used by Kalinin et al. [28] and has been faced with the same problems namely the tendency towards twinning and polymorphism [28].

To overcome all these problems, recently Kasatani et al. [29] determined, the crystal structure of phase I at 530 K using high energy X-ray technique and the MEM/Rietveld analysis. On the other hand, the high resolution neutron powder diffraction has been employed for determination of the crystal structure of phase I at 523 K [30], and also phase III at 300 K [31], at 100 and 10 K [32].

We have found that it is of interest to measure the ac conductivity and dielectric permittivity in the low temperature range $95\text{ K} < T < 280\text{ K}$ on polycrystalline samples. This technique is absent in literature for the title compound. Another reason for the use of this technique in the present work is the dependence of the physico—chemical properties of this compound on the method of its recrystallization. For example, different NQR¹²⁷ I spectra were found for crystals grown from water solution with or without HIO₃ [33]. Moreover, for crystallization of KIO₃ from an aqueous solution containing HIO₃ and KIO₃ with a ratio of HIO₃/KIO₃ greater than 7.4 %, crystals of KIO₃ and KH(IO₃)₂ are formed [23].

According to the best of our knowledge, no data on “electrical transport mechanism” in the investigated low temperature region was reported. Also the motivation behind the present work, is to compare our data with that of the very similar compound namely NH₄IO₃ [34].

In order to discuss the different models for conduction mechanism, the starting point is that, the total measured ac conductivity $\sigma_{ac(total)}$ at a given angular frequency (ω), can be separated into dc and ac components namely:

$$\sigma_{ac(total)} = \sigma_{dc} + \sigma_{(ac)} \quad (1)$$

The ac component $\sigma(\omega)$ is a function of both frequency (ω) and temperature (T) and may be written as $\sigma(\omega, T)$ whereas the dc component is a function of temperature only. According of Jonscher [35] $\sigma(\omega) \propto \omega^{s(T)}$, therefore

$$\sigma(\omega) = A(T)\omega^{s(T)} \quad (2)$$

This is the well known power law [35–43] and has been so widely observed for highly disordered materials and hence it has come to be known as a universal dynamic response.

The pre-factor A (T) is a constant and represents the degree of polarizability [38] and the frequency exponent s ($0 < s < 1$) describes the interaction (mainly of an electrostatic type) between the mobile ions with lattice around them or in more details, and according to the many body interaction model [35, 38], the interaction

between all dipoles participating in the polarization process is characterized by the parameter s . A unit value of s implies a pure Debye type interaction. The value of s decreases with the increase of interaction [38, 41]. The thermal behavior of the ac conductivity follows Arrhenius relation in most cases.

$$\sigma T = \sigma_0 \exp(-E/kT) \quad (3)$$

where E is the activation energy for the process, k is the Boltzmann's constant and σ_0 is constant.

Furthermore, it is known that the complex dielectric permittivity $\varepsilon^*(\omega, T)$ can be divided into the real part $\varepsilon'(\omega, T)$ and the imaginary part $\varepsilon''(\omega, T)$ according to the equation:

$$\varepsilon^*(\omega, T) = \varepsilon'(\omega, T) - j\varepsilon''(\omega, T) \quad (4)$$

(ε' and ε'' are frequency-dependent and represent the charging and loss currents respectively) [44].

2 Experimental Procedures

2.1 Sample Preparation

The material used in the present work was manufactured by the BDH Chemical Company Ltd. For electrical and/or ferroelectric measurements, a suitable amount of the fine polycrystalline powder of the compound, sufficient to prepare 5–6 pellets was further grained under 2 μm (micronized). The aim of this technique is to remove the effect of grain size and also to minimize the effect of porosity.

A microanalytical balance, type Sartorius, was used to achieve the equality of masses between all samples before compressing, (the mass of each sample was ≈ 0.6 gm). These samples were pressed under the same pressure (of about 2×10^8 N/m²) so we have practically identical pellets of thickness ≈ 2 mm and of diameter ≈ 1 cm.

Good electrical contact was attained by painting the opposite faces of each pellet with air drying conducting silver paste. Before any measurements, the samples were inserted into a dessicator over night to remove any humidity. A sample holder with brass electrodes was specially designed to fit the present measurements.

2.2 Characterization Techniques

To demonstrate and to check the presence of ferroelectricity at different temperatures, the circuit shown in Fig. 1a, was designed and manufactured in the Science Technical Center, Cairo University to fit the present measurement. The digital

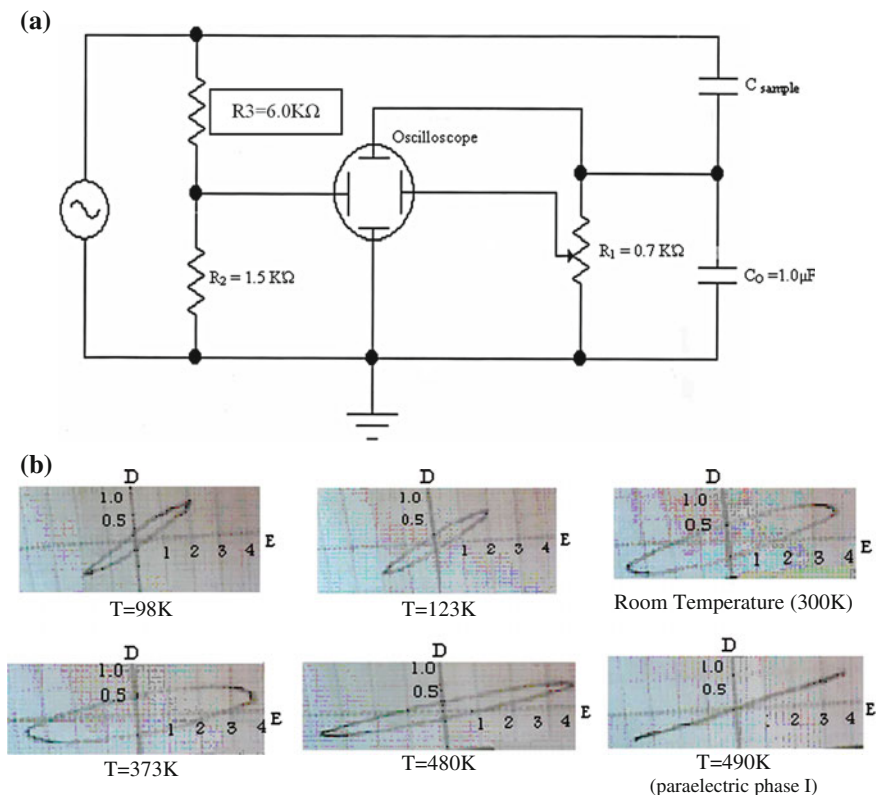


Fig. 1 **a** The designated circuit for detecting ferroelectric hysteresis loops. **b** The D-E hysteresis loops for KIO_3 (Arb. Scale)

storage oscilloscope type Instek GDS-820S, provided with a computerized camera as connected in the designated circuit shown was used to record ferroelectric hysteresis loops at different temperatures. Calibration, using triglycine sulphate (TGS), was carried out just before any actual run.

A computerized LCR bridge type Philips PM6304 was used to determine the ac conductivity and the dielectric permittivity at some selected frequencies. The bridge measures precisely the values of R , C , Z , Q ... However, in the present work the values of R and C are converted into conductivity and permittivity using the dimensions of the pellet and a simple program, since the bridge is interfaced to computer. The dc conductivity was measured by using an electrometer, type Keithly 614. Data were collected on at least three virgin samples and the results were found to be quite consistent and reproducible. All measurements were performed during heating run. The heating rate was 0.5 K/min and the experimental error of temperature is better than $+1 \%$. All measurements were performed at the thermal equilibrium. Calibration was done on a standard sample before any actual measurements.

3 Results

3.1 Ferroelectricity

The ferroelectric hysteresis loops (D-E loops) at some selected temperatures in the range $95\text{ K} < T < 280\text{ K}$ have been displayed and recorded on the monitor of the digital-storage oscilloscope. Two of these loops are shown in Fig. 1b, together with D-E loops recorded at high temperature up to 490 K where the compound is in the paraelectric phase. It seems likely that, the saturation in the D-E loops is not complete. This problem will be discussed later on (in the discussion section). The existence of this type of D-E loops is an indication that the compound behaves as an improper ferroelectric material.

3.2 Dielectric Permittivity

The temperature dependence ($95\text{ K} < T < 280\text{ K}$) of the real part ϵ' of the dielectric permittivity determined at some selected frequencies ($0.7\text{--}20\text{ kHz}$) is shown in Fig. 2.

In general, the value of ϵ' depends on both the frequency (f) and temperature (T). For a given frequency, each curve is characterized by the existence of two anomalies: a pronounced peak in the vicinity of $T_3 \approx 258\text{ K}$ and a small peak (anomaly) at $T_4 \approx 110\text{ K}$. Between the two transition regions, the value of ϵ' is almost temperature independent. It is known that, The compound undergoes a ferroelectric phase transition of an improper character (I–II) at $T_1 \sim (486 \pm 1)\text{ K}$, a change in the conduction mechanism at $(428 \pm 2)\text{ K}$ and a structural phase transition (II–III) at $T_2 \sim 345\text{ K}$.

The variation of ϵ'' , the imaginary part of the dielectric permittivity, as a function of temperature, $95\text{ K} < T < 280\text{ K}$ is shown in Fig. 3a. The frequency range ($1\text{--}10\text{ kHz}$) has been only considered for the purposes of clarity.

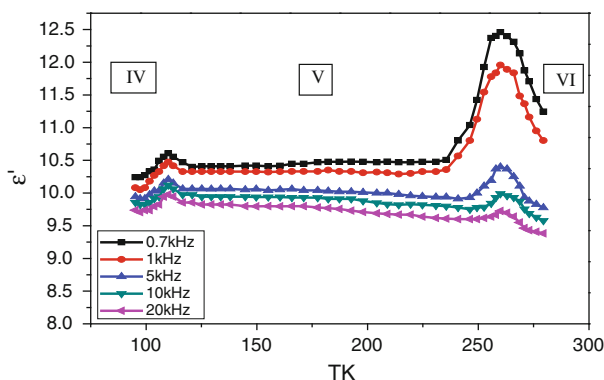


Fig. 2 Temperature dependence of the real part (ϵ') of the dielectric permittivity measured at different frequencies for KIO_3

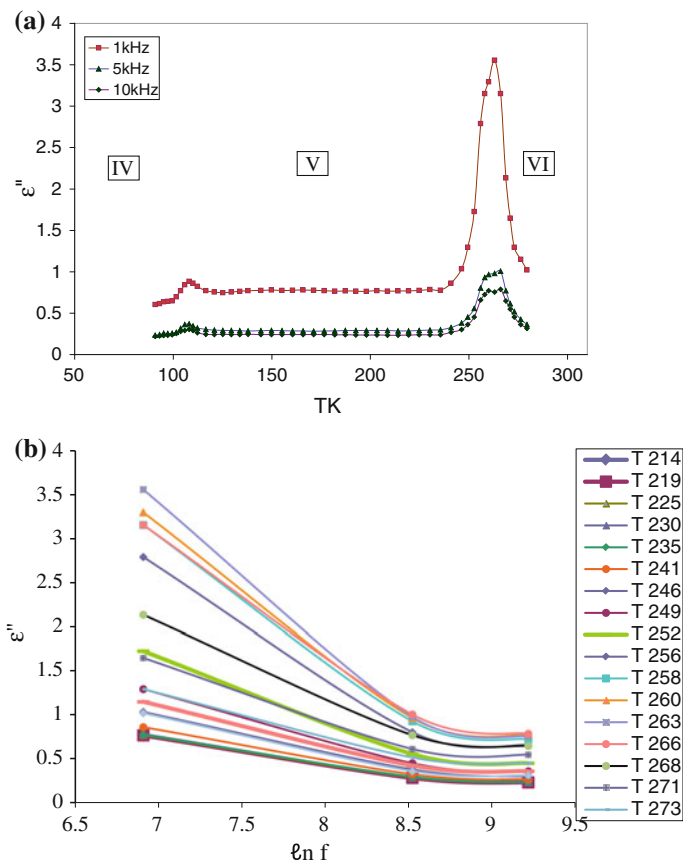


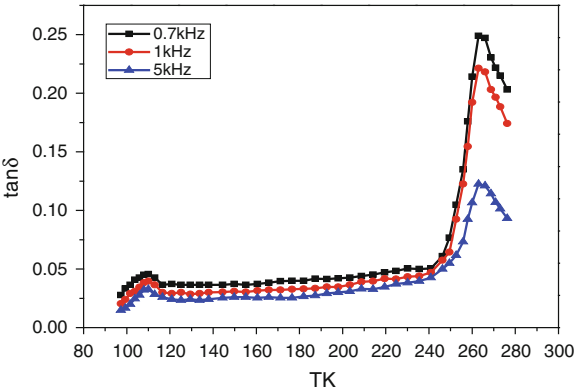
Fig. 3 **a** Temperature dependence of the imaginary part (ϵ'') of the dielectric permittivity measured at different frequencies for KIO_3 . **b** Relation between ϵ'' versus $\ln f$ at different temperatures for KIO_3

Apparently, the general trend of this plot is similar to that of ϵ' -T graph, Fig. 2, except that the value of ϵ'' at a given frequency and temperature is much reduced as compared with the corresponding value of ϵ' , Fig. 2.

Figure 3b shows the frequency dependence of the imaginary ϵ'' plotted as ϵ'' versus $\ln f$ at some selected temperatures. The plot indicates that the dispersion increases with increasing temperature and decreasing frequency which is in agreement with the typical behavior of the dielectric materials.

The variation of the (dielectric loss), $\tan \delta = \epsilon''/\epsilon'$ as a function of temperature ($95 \text{ K} < T < 280 \text{ K}$) measured at some selected frequencies is shown in Fig. 4. The general trend of the plot indicates also the existence of two phase transitions at the above mentioned temperatures and are reflected as two peaks.

Fig. 4 Temperature dependence of $\tan \delta$ measured at different frequencies for KIO_3



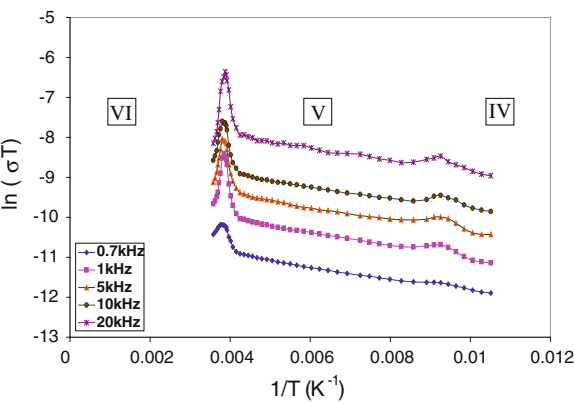
3.3 Electrical Conductivity

Obviously the conductivity depends on both, the temperature and the frequency.

3.3.1 Temperature Dependence

The ac conductivity $\sigma_{(\omega)}$ as function of temperature, $95 \text{ K} < T < 280 \text{ K}$ and frequency $0.7\text{--}20 \text{ kHz}$, plotted as $\ln(\sigma T)$ versus $1/T$ is shown in Fig. 5. The behavior of σ seems to be in accordance of Arrhenius relation. For a given frequency, and as the temperature increases from about 95 K , the first broad peak is observed at $\approx 258 \text{ K}$. The other phase transition is found at $\approx 110 \text{ K}$. Between the two transition regions the behavior of σ obeys Arrhenius relation.

Fig. 5 Variation of $(\ln \sigma T)$ vs. $1/T$ at different frequencies for KIO_3



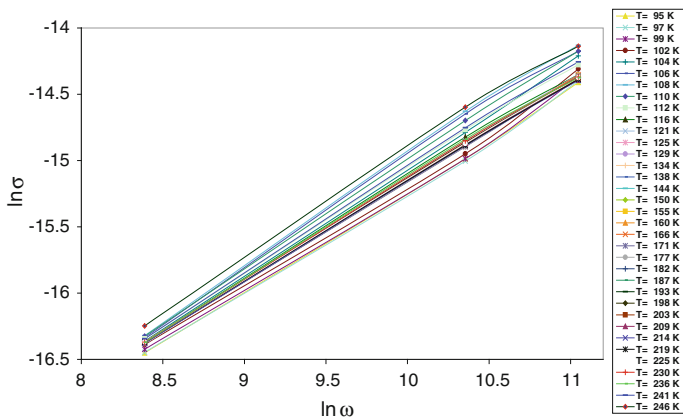


Fig. 6 Double logarithmic plot ($\ln \sigma$ vs. $\ln \omega$) for KIO_3

3.3.2 Frequency Dependence

The variation of $\ln \sigma$ versus $\ln \omega$ at different temperatures is shown in Fig. 6.

The plot consists of a series of straight lines of different slopes. The slope of each line gives the value of s , where $s = \left(\frac{\partial \ln \sigma}{\partial \ln \omega} \right)_{T=\text{const.}}$ at that temperature. The temperature dependence of s is shown in Fig. 7. Apparently, s is practically temperature independent except for in the phase transition regions.

Figure 8 shows the variation of the pre-exponential factor (A) with temperature, plotted as $\ln A$ versus T . The behavior of the plot indicates also the existence of two phase transitions at the same mentioned temperatures (258 and 110 K) which are reflected in the graph shown as “inverted peaks”.

Fig. 7 Temperature dependence of the frequency exponent (s) for KIO_3

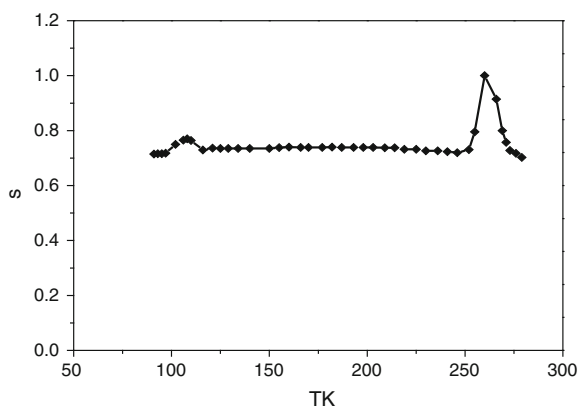
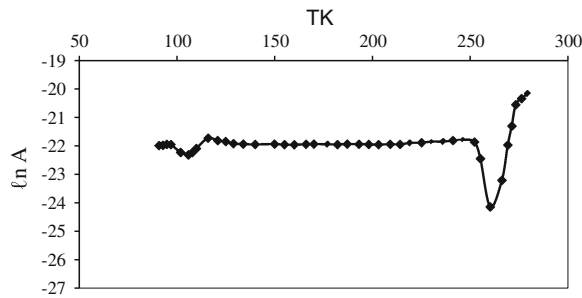


Fig. 8 Variation of $\ell n A$ versus T for KIO_3



4 Discussion

The values of the transition temperatures T_3 (III–IV) = 258 K and T_4 (IV–V) = 110 K of the present work was found to be consistent with those found by previous workers (Tables 1 and 2). For example, from his previous investigation of the ferroelectricity in KIO_3 single crystal, Salje [17] observed two phase transitions in the low-temperature region at $\approx -15^\circ\text{C}$ (258 K) and at -163°C (110 K).

Furthermore, Ivanov et al. [15] studied the temperature dependence of the optical axes angle $2v$, the birefringence in the z-cut $\delta n_{(z)}$ and the spontaneous polarization axes P_{sf} in a wide temperature range and they observed that at $T = -15^\circ\text{C}$ (258 K) all three curves [$2v(T)$, $\delta n_{(z)}(T)$ and $P_{sf}(T)$] exhibit only breaks and thus they suggested that this transition is to be of a second order. Moreover, although Ivanov et al. [15] did not investigate the phase transition (IV–V) in detail, yet they observed a regular quantitative change of the domain pattern which confirms the existence of a transition with a lowering of point symmetry. Moreover, they demonstrated the existence of ferroelectricity in the compound at the liquid nitrogen temperature -195°C (78 K) by observing the D-E hysteresis loops which was found to be clear at that temperature and thus they concluded that phase V behaves as a ferroelectric which is in a good agreement with the present work and also with the work of

Table 2 Summary of the crystal structure at some selected temperatures for the compound KIO_3

Temperature (K)	Crystal data	Technique	References
530	Rhombohedral, R3 m, Z = 1	High energy X-ray powder diffraction MEM/Reitveld analysis	[29]
523	Rhombohedral, R3, Z = 1	Neuton powder diffraction technique (Rietveld method)	[30]
Room temperature	Triclinic symmetry	Single crystal X-ray diffraction	[28]
523	Triclinic, P1, Z = 4	Neutron powder diffraction technique	[31]
300	Triclinic, P1, Z = 4	(Rietveld method)	[32]

Herlach [12], where the crystal remains ferroelectric below 485 K [12]. In somewhat recent work Maeda et al. [19] demonstrated the existence of low temperature phase transitions at $T_3 = -15\text{ }^\circ\text{C}$ (258 K) and $T_4 = -160\text{ }^\circ\text{C}$ (113 K). A similar behavior has been observed for NH_4IO_3 where the compound undergoes a ferroelectric phase transition at $\approx 368\text{ K}$ and behaves as a ferroelectric material below 368 K down to liquid nitrogen temperature [34].

Among other things, although some investigators have been reported on the low-temperature phase transitions in KIO_3 , through the study of the temperature dependence of some parameters, yet, the only crystal structure of the low temperature is that determined by Lucas [32]. The data do not show significant change, on the atomic level, relative to that of the room temperature, i.e. no strong evidence for low temperature structural phase transitions. As we mentioned above, the temperature dependence of some properties show detectable change at $\approx 258\text{ K}$ and at $\approx 110\text{ K}$. This inconsistency is not a major problem, since the structural phase transitions that involve only minor changes in atomic positions are not easy to detect [45]. Potassium chlorate KClO_3 represents an example of a crystal with a reported phase transition with very small structural change where the atomic arrangement and coordination above and below the phase transition are identical [45]. The same conclusion was also given by Brooker et al. [46]. From Raman studies of the phase transition in KClO_3 , they observed that the difference in the vibrational spectra of the two phases was very small and hence the transition (at 545 K) involved only minor rearrangement of the ClO_3 ion. Similar saturation may exist for the KIO_3 since the two compounds (KIO_3 and KClO_3) are members of halate series of the general molecular formula MXO_3 where ($\text{X} = \text{Cl}, \text{Br}$ and I).

From the crystal structure point of view [29–31], the I and O atoms exist in the crystals as IO_3 molecules rather than the IO_6 complex. In other words unlike the typical perovskite structure of ABO_3 ferroelectrics where six O atoms form O_6 octahedron surrounding, a B atom each I atom in KIO_3 has three nearest O atoms forming an IO_3 molecule [26].

At this stage, it is necessary to mention that although there is a good agreement between different authors that the compound undergoes two phase transitions in the temperature interval $95\text{ K} < T < 280\text{ K}$, yet there is a discrepancy concerning the nature and/or the exact mechanism of these transitions. Some authors consider these transitions to be associated with the orientational glass transitions [19]. Other authors [46] proposed that these transitions might be associated with the structural changes of domain which is generally independent on atomic arrangements.

The present work tries to construct a bridge or link between different visions. The motion and/or the change of the structure of ferroelectric domains with temperature was first suggested by Byrom and Lucas [30]. On the other hand, and in a similar way to the role of ClO_3^- as the main cause of the structural phase transition in the KClO_3 compound, one cannot neglect the effect of the reorientational motion of the IO_3^- , since, as we mentioned above, the spatial correlation between IO_3^- decreases with increasing temperature [29]. Thus the thermal energy at each transition temperature may be enough to cause the change of the structure of ferroelectric domains

and at the same time initiate (stimulate) the orientational motion of the IO_3^- and hence a very slight rearrangement of the molecule with a very minor change in the atomic position in a way similar to that observed in the KClO_3 [45, 46].

It is of interest to mention here that even for the structural phase transition of the KIO_3 that have been observed at ~ 487 K which is associated with the ferroelectric phase transition, the two X-ray diffraction patterns above and below the mentioned transition temperature are very similar [30]. In our vision, the exact mechanism of the low temperature phase transition is a future problem and requires the use of the high energy X-ray (synchrotron) powder (and not single) technique which is beyond the scope of the present work.

For the problems of unsaturated D-E hysteresis loops shown in Fig. 1, it is known that the compound behaves as an “improper ferroelectric material” [20, 24–26] and not as (usual) ferroelectrics. The characteristic features of the improper ferroelectric are [47, 48]:

- (i) a very slight change of dielectric constant near T_c
- (ii) a low spontaneous polarization

Logically, the spontaneous polarization is not the order parameter and not proportional to it as in the case of proper ferroelectrics. The factor (ii) is the most responsible for the unsaturation in the D-E hysteresis loops. The case of KIO_3 is very similar to the improper ferroelectric NH_4IO_3 where the D-E hysteresis loops in the ferroelectric phase are also unsaturated [34, 49].

Regarding the charge transport mechanism in different phases, it is known that the three main processes that contribute significantly to the ac conductivity [37, 50, 51] are the quantum mechanical tunneling QMT, the correlated barrier hopping (CBH), and the small polaron tunneling (SPT).

For QMT model, the value of s is given by

$$s = 1 - \frac{4}{\ln\left(\frac{\nu_{ph}}{\omega}\right)} = 1 - \frac{4}{\ln\left(\frac{1}{\omega\tau}\right)} \quad (5)$$

whereas for (CBH) model, we have

$$s = 1 - \frac{6kT}{E_0 - \ln\left(\frac{\nu_{ph}}{\omega}\right)} = 1 - \frac{6kT}{E_0 - \ln\left(\frac{1}{\omega\tau}\right)} \quad (6)$$

and finally for the (SPT),

$$s = 1 - \frac{4}{\ln\left(\frac{1}{\omega\tau_p}\right) - w_H/kT}. \quad (7)$$

The symbols in (5), (6) and (7) take their conventional meanings where k is Boltzmann's constant, E_0 is the optical energy band, ν is the frequency of the

phonon. τ_p is the relaxation time of the polaron (of the order of 10^{-13} s) and w_H is the activation energy involved in the electron transfer process between a pair of states. Moreover, the value of s in (5) is temperature independent, whereas in (6) s decreases with increasing the temperature and is less dependent on the frequency. For SPT, (7) predicts that, s increases as T increases.

Referring to Fig. 7, where s is plotted as a function of temperature, it is clear that the value of s is practically temperature independent i.e. in accordance with (5). Thus, the QMT model is the most likely one in this range of temperature. Moreover, if we put v_{ph} ($\sim 10^{-13}$ s $^{-1}$ and $\omega \sim 10^4$ rad/s) [37] into (5), then the value of $s \sim 0.76$, which is in a good agreement with our experimental value, Fig. 7.

Furthermore, for QMT model, the value of $\sigma(\omega)$ as given by Auston and Mott [52] and Efros [53] is

$$\sigma(\omega) = \frac{\eta}{3} e^2 kT [N(E_f(T))]^2 \alpha^{-5} \omega \left[\ln\left(\frac{v_{ph}}{\omega}\right) \right]^4 \quad (8)$$

where:

- e is the charge of electron,
- T is the absolute temperature,
- $N(E_f)$ is the density of energy states near the Fermi level,
- α is the electron wave function decay constant 1 Å [34],
- v_{ph} is the phonon frequency, and
- η is a constant.

Equation (8) predicted that, there is a linear relation between ω and $\frac{\sigma(\omega)}{[\ln(\frac{v_{ph}}{\omega})]^4}$ such a relation is presented in Fig. 9, which again confirm (8) where a series of almost straight lines of different slopes are observed. Thus the QMT model is the best one that describes the conduction mechanism in the investigated temperature region.

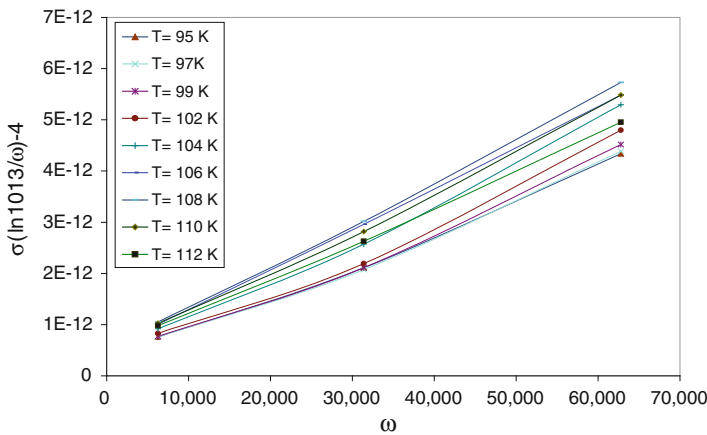


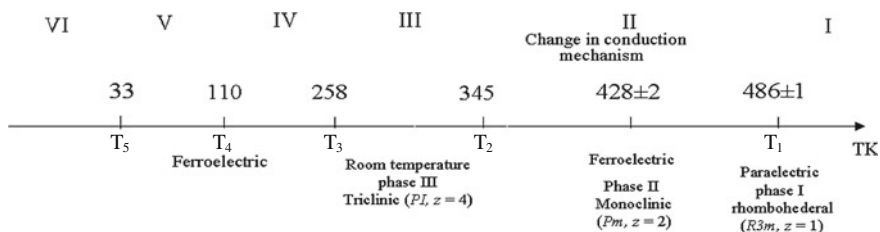
Fig. 9 Variation of $\sigma(\omega)/[\ln(v_{ph}/\omega)]^4$ versus ω at different temperatures for KIO₃

5 Conclusion

The present data indicate that in the investigated low temperature region $95\text{ K} < T < 280\text{ K}$, the compound remains ferroelectric (of an improper character) where the ferroelectric hysteresis loops, of quite insufficient saturation, are observed clearly over the mentioned temperature intervals. This is also the case for the similar compound NH_4IO_3 where it undergoes a ferroelectric phase transition of an improper character at $\approx 368\text{ K}$ and remains ferroelectric down to liquid nitrogen temperature. For the compound KIO_3 , the behavior of the temperature dependences of all electrical parameters confirms and supports each other and indicates clearly that compound undergoes two structural phase transitions at $T_3 \approx 258\text{ K}$ and at $T_4 \approx 110\text{ K}$ over the mentioned temperature range. It is known that the iodine and oxygen atoms in the present compound exist as IO_3 rather than as IO_6 complex and the spatial correlation between IO_3 becomes weak with increasing temperature. The thermal energy (kT) at the transition temperature is enough to cause the rearrangement of IO_3 molecules in a way similar to the case of KClO_3 . The thermal energy also causes an observable change in the structure of ferroelectric domain.

Regarding the electrical conduction, in the investigated low temperature region, the QMT model seems to be the main mechanism for the charge transport. This is to be expected, since this model usually exists in the low temperature range.

The low and high phase transitions as well as the change in conduction mechanism at $(428 \pm 2)\text{K}$ are shown in the following scheme. The values of the transition temperatures are taken from [20] and the present work except for transition at $T = 33\text{ K}$ is taken from [19].



References

1. Barabash A, Gavrilko T, Eshimov K, Puchkovskaya G, Shanchuk A (1999) J Mol Struct 145:511–512
2. Bismayer U, Salje E, Viswanathan K (1979) Phase Transitions 1:35
3. Abdel-Kader MM, El Kabbany F, Naguib H, Gamal WM (2008) Phase Transitions 81:83
4. Colombari P (1992) Proton conductors, (Solids, membranes and gels materials and devices) Part I P.I. Cambridge University Press, Cambridge
5. Cretz JM, Righi A, Galez C, Bourson P, Mareira RL (1998) Solid State Commun 105:481
6. Gaffar MA, Abu El-Fadl JA (1999) J Phys Chem Solid 60:1633

7. Bergman JG, Boyd GD Jr, Ashiken A, S-Kurtz (1969) *J Appl Phys* 40:2860
8. Keve ET, Abrahams SC, Berustein JL (1971) *J Chem Phys* 54:2556
9. Kurtz SK, Perry TT (1965) *Appl Phys Lett* 12:156
10. Crane GR, Bergman JC, Glass AM (1969) *J Am Ceram Soc* 52:655
11. Helg H, Gruncher H (1970) *J Phys Soc Jpn* 28(Supl):169
12. Herlach F, Helv (1961) *Helv Phys Acta* 34:305
13. Salje E (1972) *Z Kristallagr* 136:135
14. Crane GR (1972) *J Appl Cryst* 5:360
15. Ivanov NR, Shuvalov LA, Chikgladze OA (1973) *Phys Lett* 45A:437
16. Salje E (1971) *Z Kristallagr* 134:107
17. Salje E (1974) *Z Kristallagr* 139:317
18. Landolt Bornstein (1982) vol 111/16b. Springer, Berlin p 269
19. Maeda M, Takagi M, Suzuki I (2000) *J Phys Soc Jpn* 69:267
20. Abdel Kader MM, El-Kabbany F, Naguib HM, Gamal WM (2008) *Phase Transitions* 81:29
21. Filimonov AA, Survotov LG, Paklomov VS, Sonin AS (1965) *Sov Phys Cryst* 10:202
22. X- Yin, Lü MK (1992) *Appl Phys Lett* 60:2849
23. Haussühl S, Jyang W, Mengkai LÜ (1995) *Cryst Res Technol* 30:535
24. Ahart M, Kojam S, Takagi M, Maeda M, Suzuki I (1998) *Jpn J Appl Phys* 37:5687
25. Loiacono GM, Jacco JG (1985) *Mater Lett* 4:27
26. Liu L, Wu RQ, Ni ZH, Shen ZX, Feng YP (2006) *J Phys Conf Ser* 28:105 (and references there in)
27. Hamid SA (1973) *Z Kristallogr* 137:412
28. Kalinin VR, Llyukhin VV, Belov NV (1978) *Sov Phys Dokl* 12:1978
29. Kasatani H, Aoyagi S, Kuroiwa Y, Yagi K, Katayama R, Terauchi H (2003) *Nucl Instr Meth B* 199:49
30. Byrom PG, Lucas BW (1987) *Acta Crystallogr C* 43:1651
31. Lucas BW (1984) *Acta Crystallogr C* 40:1989
32. Lucas BW (1985) *Acta Crystallogr C* 41:1388
33. Bajsa DF, Barabash AI, Vertegel IG (1981) *Ukr Fiz Zhurnal* 10:1745
34. Abdel Kader MM, El-Kabbany F, Naguib HM, Gamal WM (2013) *Phase Transit* 36:947–958
35. Jonsher AK (1983) *Dielectric relaxation in solids*. Chelsea Dielectric press, London, p 223
36. Dyre JC (1988) *J Appl Phys* 64:2456
37. Elliott SR (1987) *Adv Phys* 36:135
38. Krthik C, Varma KBR (2006) *J Phys Chem Sol* 67:2437
39. Louati B, Gargouri M, Guidara K, Mhiri T (2005) *J Phys Chem Sol* 66:762
40. Dyre JC, Schroder TB (2000) *Rev Mod Phys* 72:873
41. Chen RH, Yen Chen-Chief, SHern CS, Fukami T (2006) *Sol State Ion* 177:2857
42. Lee WK, Liu JF, Nowick AS (1991) *Phys Rev Lett* 67:1559
43. Moawad HMM, Jain H, El-Mallawany R (2009) *J Phys Chem Sol* 70:224
44. Al Refaie SN, El-Rayyan HSB (1992) *J Mater Sci Lett* 11:958
45. Ramachandran N, Lonappan MA (1957) *Acta Cryst* 10:281
46. Brooker MH, Shapter JG (1989) *J Phys Chem Solids* 50:1087
47. Blinc R, Zekz B (1974) *Soft modes in ferroelectrics and antiferroelctrics*. North-Holand Publishing company, Amsterdam, pp 15, 45
48. Astrukov B, Levanyuk AP (1998) *Ferroelectric phenomena in crystals*. Springer, Berlin, p 74
49. Oka T, Mitsui T, Sawada S (1976) *J Phys Soc Jpn* 40:913
50. Gamati F, Fattoum A, Bohli N, Dhaoui W, Mohamed AB (2007) *J Phys: Condens Matter* 19:36203
51. Long AR (1982) *Adv Phys* 31:553
52. Austin IG, Mott NF (1969) *Adv Phys* 18:41
53. Efros AL (1981) *Phillos Mag* 43:829

Low Power Transmitter for Capsule Endoscope

De Xing Lioe, Suhaidi Shafie, Harikrishnan Ramiah, Nasri Sulaiman and Izhal Abdul Halin

Abstract This chapter presents the design technique of low power transmitter for the medical application of capsule endoscope. Considering the loss against frequency in a body wireless communication, ISM band of 434 MHz is employed in the design of the transmitter. This band has lower loss and relatively higher data rate compared to other standards. Inductorless architecture was adopted in the circuit design to reduce the circuit area, thus contribute to the reduction of capsule size. The core component of transmitter, the up-conversion mixer and ring oscillator is designed using CMOS 0.13 μm technology with voltage supply of 1.2 V. Both the mixer and ring oscillator consumes 1.57 mA of current, brings the dc power consumption of the transmitter to be 1.88 mW. Data rate of 3.5 Mbps ensure it can transmit high quality medical imaging. The proposed up-conversion and ring oscillator achieved low power and less area while still having the good performance. This achievement makes circuit integration for low power transmitter realizable.

1 Introduction

Conventional endoscopy is the process of diagnosing and examining gastrointestinal tract of the patients, which is conducted by either inserting a tube through the mouth (gastroscopy) or through rectum (colonoscopy). A camera is mounted at one end of

De Xing Lioe (✉) · S. Shafie · N. Sulaiman · I.A. Halin
Department of Electrical and Electronic Engineering, Faculty of Engineering,
Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
e-mail: dexiong@ieee.org

S. Shafie
e-mail: suhaidi@eng.upm.edu.my

H. Ramiah
Department of Electrical Engineering, Faculty of Engineering, University of Malaya,
50603 Kuala Lumpur, Malaysia

the endoscope and images are taken and sent to a display monitor. This is an effective way of diagnosis but the disadvantages are it causes pain and uncomfortable to the patients. Patients undergo gastroscopy may suffered from mild sore throat or feeling of distention from the insufflated air that was used during the procedure. On the other hand, colonoscopy causes discomfort and pain to the patients as the tube needs to be pushed into the intestine. Apart from the discomfort, conventional endoscopes have its' limitation, where it was not able to fully evaluate the disorder in the small intestine [1]. This is due to small intestine is very long and convoluted, therefore no available scopes are able to traverse the entire length of small intestine [2], resulted in some small intestine diseases to be hardly diagnosed (Fig. 1).

The evolution of endoscopy started around 10 years ago when Given Imaging produced the first ever capsule endoscope named M2A [3], thanks to advancement of the technology of CMOS image sensor, application specific integrated circuit (ASIC) devices, as well as white light emitting diode (LED) [4]. Generally a capsule endoscope is made up of a tiny camera, a lens, LEDs, battery and a wireless transmitter. A suspected gastrointestinal tract patient can now swallow a capsule size camera which takes images in the body and transmit to a receiver worn around

Fig. 1 Small intestine in human body

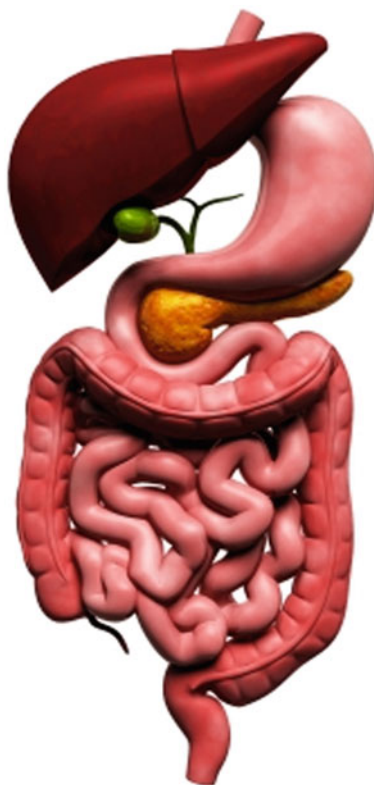


Image courtesy of dream designs / FreeDigitalPhotos.net

the waist, so to allow later evaluation by physician. The whole process usually takes about 6–8 h, but the patients are allowed to move freely during the process. Compared to conventional endoscopy, capsule endoscopy is a painless and non-invasive alternative. This method has proven to successfully diagnosed diseases which are unsuccessfully done using other diagnosis methods [5–7].

Few other companies have since introduced their own capsule endoscope such as EndoCapsule of Olympus, MicoCam of IntroMedic and OmOm from China. Despite the commercialization of these capsule endoscopes, still, there are many technical aspects that can be improved [8]. Many researches have been conducted to achieve better image resolution, higher frame rate, and longer process time, which are the three major properties lacking in current capsule endoscope, as well as some extra functions [9–12]. Poor quality of image may cause diagnosis error by physicians, or physicians might be unsure of the real situation causing a waste in capsule endoscopy. Certain part of the interested area in small intestine might be accidentally skipped if the frame rate is too low, therefore it is important for a higher frame rate.

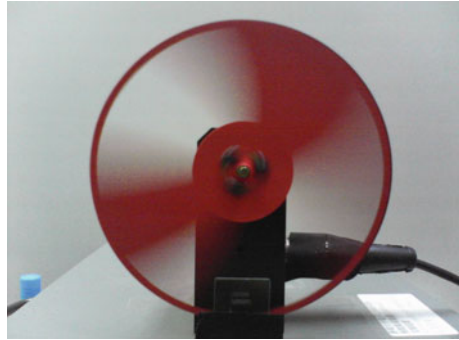
Most of the constraints on a capsule endoscope can be overcome if the battery life is sufficient. Since the available battery capacity for this particular application is very stringent due to its size limitation, a very low power consumption circuitry is highly essential. Prolonged battery life enabled the capsule endoscope to operate longer than average full process time in case of rare sequence where it takes much longer time for the capsule to travel through the entire body. Apart from that, improved functionalities may be added with having a longer battery life, such as microactuator for microbiopsy, two way communications to control the movement of capsule in the body, stopping mechanism to allow physician to monitor on gastrointestinal tract etc.

Transmitter of a capsule endoscope contributes the major power consumption towards the overall power usage, apart from image sensor. Power consumption of the capsule endoscope can be largely reduced if the transmitter is designed to be more efficient. Several researches show relatively high power consumption in the transmitter circuit, as well as commercially available capsule endoscopes provide only frame rate of two frames per second due to power constraint.

2 Imager Selection

The most important part in capsule endoscopy system is the image sensor chip which converts the reflected photon in term of voltage data. CMOS image sensor with global shutter [13, 14] is preferable for capsule endoscope since it promise low power consumption and can reduce the effect of the motion artifact. The motion artifact occurs when a moving object is captured by a camera with rolling shutter or when a moving camera with rolling shutter captures a still image. The produced image will be distorted as shown in Fig. 2. Then, the CMOS image sensor with wide dynamic range [15, 16] is also preferable in order to have the detail in the image of the captured object. However the processing must be on chip since the

Fig. 2 Distorted image captured by CMOS image sensor with rolling shutter [17]



number of output bits also is selected to be 8-bits in order to reduce the transmission data rates. The output data of selected image sensor is preferable to be in digital scheme rather than analog since it will be succeeded into transmitter block. The 8-bits data is then converted to serial data and succeeded to the transmitter. Serial data communication has been selected to reduce the transmission power and die space on designed transmitter. It also promise accurate data at receiver parts since the parallel data communication may suffer data mismatch and skew at the receiver.

3 Design Consideration

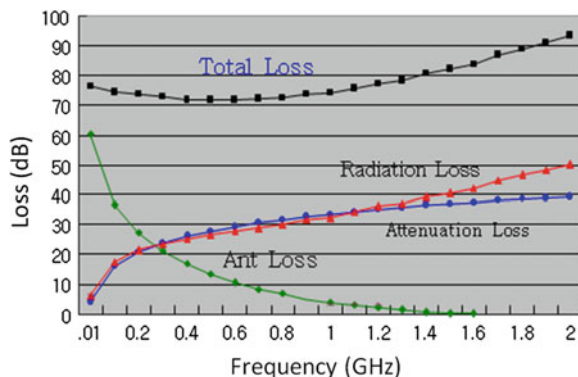
Designing the system aims at optimizing each and every one of the crucial characteristics, basically with the trade off of power consumption versus performance, where the performance includes image size, bit depth, frame rates, which sums up to be the data rate. The balance of both criteria of performance and power consumption has to be reached so that a quality capsule endoscope system is designed.

3.1 *Carrier Frequency*

There are a few commercially available capsule endoscopes and many researches been done on capsule endoscope. But up-to-date there is no standard available for the usage of capsule endoscope, which means that the creator or the researcher can determine their intended carrier frequency.

Federal Communications Commission (FCC) has allocated a frequency band of 402–405 MHz for Medical Implant Communications Service (MICS) in year 1999 [18]. After the introduction of this frequency band, it has been widely used in wireless implantable medical devices. MICS has very strict requirements to be followed, one of them is allowing transmission only up to 500 kbps, which is too low to address for the problem of low image quality and low frame rate of current capsule endoscope.

Fig. 3 Loss against frequency for in-body wireless communications [27]



Besides MICS, Wireless Medical Telemetry Services (WMTS) is another standard for wireless medical devices, with frequency range from 608 to 614 MHz. However, WMTS is heavily used and this band is only applicable to healthcare services in the United States [19]. The carrier frequency choices that were reported range from 20 MHz to 5 GHz, depending on the modulation scheme, such as frequency shift keying (FSK), phase shift keying (PSK), on-off keying (OOK), amplitude shift keying (ASK) etc., and transmission techniques [10, 20–24].

Little study has been conducted on the selection of carrier frequency for this specific application of wireless capsule endoscope. From one of the research conducted, it shows that for similar power consumption, lower frequency capsules tend to have better carrier signal-to-noise ratio (CSNR) on top of having smaller form factor [25]. Even though there is a trade-off between antenna size and wavelength, but carrier frequency below 500 MHz is favourable, moreover with doubt over the validity of this trade-off in the near-field transmission [26].

Figure 3 shows the relationship between radiation loss, attenuation loss, antenna loss, and total loss with frequency [27]. The graph shows that combination of all losses is minimal at the frequency range of 400–900 MHz. In this case, the proposed carrier frequency is taken to be 433.92 MHz, with range of 433.05–434.79 MHz, which is one of the Industrial, Scientific, Medical (ISM) band. This frequency fits in the outcome of the previous studies done, do not have limitation regarding to the allocation of transmission bandwidth, and no restriction on modulation method.

3.2 Power Consumption

The most essential aspect which gives the capsule endoscope to keep functioning is the power consumption. Considering it is a battery operated device, power management of the system must be well planned. Taking consideration into the size, two batteries of type SR927 W with dimensions of 9.5 mm × 2.73 mm is chosen.

The capacity for this type of battery is 60 mAh at 1.2 V. In order for the battery to last 8 h of continuous usage, maximum total power consumption has to be restricted not more than $(60 \text{ mAh}/8 \text{ h}) \times 1.2 \text{ V} \times 2$, equivalent to 18 mW. Dividing the power consumption into functional area of lighting, camera and other supporting circuits which took approximately 8.5 mW [23], the power consumption of the transmitter of the system has to be within 9.5 mW.

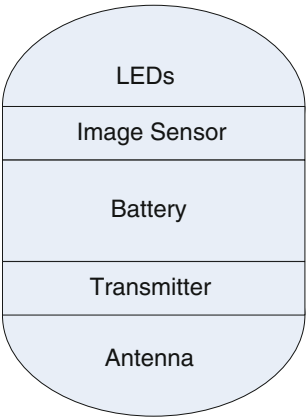
3.3 Imaging Quality

Resolution, bit depth, frame rates and data rates all combine to determine the quality of the capsule endoscopy imaging, which essentially is the main objective of diagnosis. Poor quality of imaging would bring to uncertainty of diagnosis and leads to time and financial loss. Resolution of the image should be large enough to be examined by physicians without doubt. The proposed resolution for the system is Quarter Video Graphics Array (QVGA) where the pixel size is 320×240 pixels. Frame rate of the camera is set to be five frames per second (fps) in order not to miss any intestinal area in concerned, contrast to 2 fps used in other work [20]. Flashes of lighting in synchronous with the frame rate is to be used in order to keep the power consumption low. Bit depth is another important attributes as difference in colour in intestine leads to discovery of diseases. 8 bits of colour is used in the design of the system to produce high quality imaging. Coming together with the mentioned characteristics, the data rate needed would be $320 \times 240 \times 5 \times 8$, which is equivalent to 3 Mbps. The transmitter of the system is designed to cater for data rate of 3.5 Mbps, compare to existing system of 1–2 Mbps [21, 26].

4 Proposed Transmitter

Wireless capsule endoscope architecture consists of LEDs for lighting, image sensor, battery, transmitter and antenna, as shown in Fig. 4. The building block in interest is the transmitter where it accounts for large percentage of the total power consumption and determines the quality of transmission. The proposed transmitter adopted direct conversion transmitter architecture which has simple signal path, thus making it attractive due to its low complexity circuit architecture and high level of integration. The signal from image sensor, baseband signal, is up-converted to RF signal directly in one step. RF signal can be directly connected to an antenna for transmission. This avoids the common problem of LO pulling caused by the power amplifier. The simplicity of the direct conversion transmitter has the benefit of high-integration, reduced chip area and cost, and low power consumption, especially without the need of power amplifier for such short distance transmission [20, 22]. This makes direct conversion transmitter a good choice when the requirement for a capsule endoscope transmitter is low power and small size.

Fig. 4 Simplified overall system of the capsule endoscope

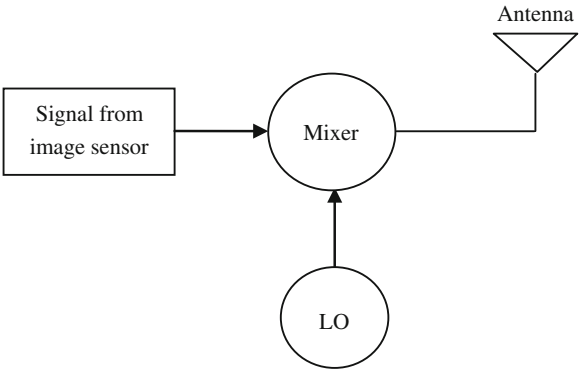


4.1 Mixer

Figure 5 shows the direct conversion transmitter architecture which consists of a mixer and local oscillator. Mixers can be classified as either an active mixer or a passive mixer. The difference is that active mixers in general have conversion gain while passive mixers have conversion loss. Active mixers are normally used in transmitter application due to higher conversion gain, which is one of the design considerations for transmitter. Other considerations for designing an up-conversion mixer are noise figure, linearity and dynamic range.

Since conversion gain is in consideration for the up-conversion mixer, an active mixer is favoured over passive mixer. Types of active mixer are single device, single-balanced and double-balanced. The simplest design of active mixer is the single device mixer, having the advantage of less component counts. The simplicity however causing the mixer to having a poor port-to-port feed-through performance [28]. This phenomenon is undesirable in a mixer. The concept of balanced mixers is introduced to perform desired signal cancellation of the component of output signal.

Fig. 5 Transmitter architecture



A single baseband signal is fed together with differential LO signal. The mixing of these signals resulted in the desired up-converted frequency as well as undesirable LO feed-through component. Baseband feed-through is cancelled out in the mixing process.

Signal cancellation concept is further advanced to form the double-balanced mixer. Double-balanced mixer offers both LO and baseband signal cancellation. The basic double-balanced mixer is the classic Gilbert cell mixer. The advantage of double-balanced mixer over the single-balanced counterpart is the isolation of all the ports in mixer, increases linearity and less susceptible to supply voltage noise. Since the architecture is twice as large as the single-balanced mixer, the power consumption is also increased, but this trade-off is worth in terms of mixer performance. The classic Gilbert cell mixer is however not suitable to achieve a good performance in low supply voltage and low power consumption due to the stacked transistors architecture.

4.2 Local Oscillator

An inductor paralleled with capacitor make up the resonator or the LC tank. LC tank alone is not a good oscillator as the amplitude of oscillation decays due to loss in the form of resistance heat. An active circuit is therefore needed in the circuit to compensate for this loss to ensure the oscillator will oscillate indefinitely and a steady-state oscillation is reached. The negative resistance introduced by the active circuit need to be equal to the resistance heat for the sustained oscillation. LC oscillators generally have the advantage of better phase noise compared to other oscillator such as ring oscillators. The advantage of the better phase noise however comes along with a setback of larger chip area for the implementation of LC oscillator due to the usage of inductors.

Another commonly used topology to implement a VCO is the ring oscillator. A ring oscillator can be implemented in single-ended architecture or differential architecture, consists of a number of gain stages. A single-ended architecture is an inverter based and constructed with a cascade of odd numbers of CMOS inverters while differential ring oscillator can be constructed using a differential pair with symmetric loads. The phase noise of differential ring oscillator is higher than that of the single-ended oscillator. Despite the disadvantage of having higher phase noise compare to single-ended ring oscillator, differential ring oscillator is preferred in most applications due to the lower sensitivity to substrate and supply noise, and also lower noise injection into other circuits on the same chip [29]. Phase noise requirement for this near-field targeted application is not essential due to low regulated power and short communication distance [30]. This further strengthens the advantage of differential ring oscillator towards single-ended ring oscillator. The less stringent requirement for phase noise also makes the differential ring oscillator a better choice than LC oscillator, in addition to the ability of ring oscillator to be highly integrated in a chip, where die area can be significantly reduced.

5 Results

The transmitter circuit has been designed and simulated using 0.13 μm CMOS technology. The layout of the active mixer and ring oscillator is as shown in Fig. 6 with die area within $1\text{ mm} \times 1\text{ mm}$. The small area is important for achieving highly integrated circuit with optimized performance. The choice of using mixer and ring oscillator which are without inductors greatly contributes to the minimization of die area.

The combination of mixer and ring oscillator consumes 1.57 mA of current with voltage supply of 1.2 V which translates to power consumption of 1.88 mW. The resulted power consumption is way below the targeted power performance of 9.5 mW. The achievement of low power consumption enables complete diagnosis without interruption of possible power failure. Apart from that, the reduction of power consumption can be utilized for other blocks of the system, which consequently able to boost the overall performance of the system. Figure 7 shows the simulated output waveform of the transmitter. Performance of the designed transmitter is shown in Table 1.

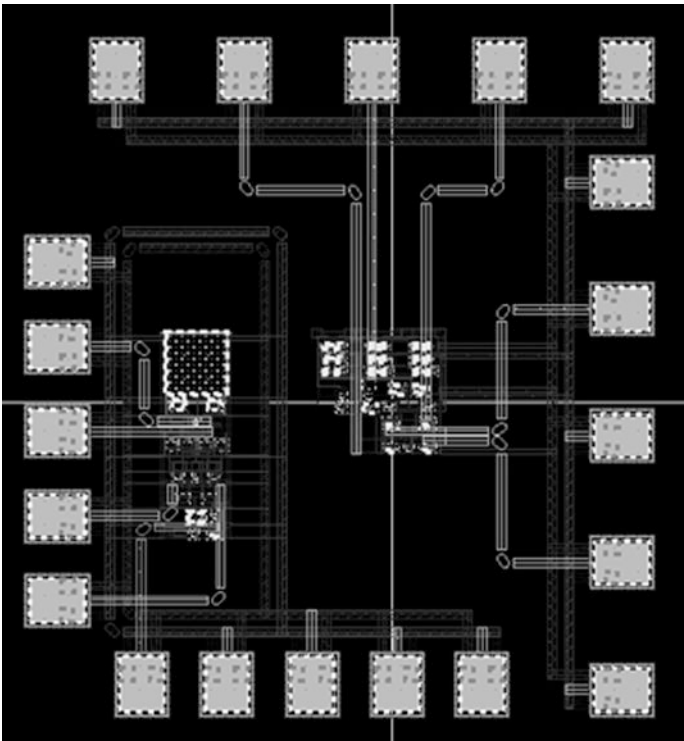


Fig. 6 Layout of the designed mixer and ring oscillator

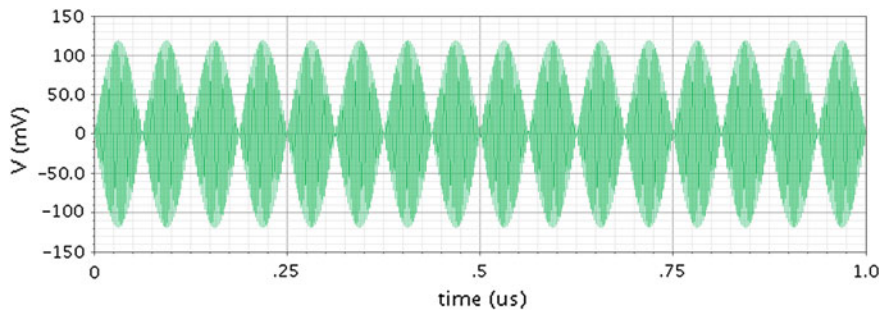


Fig. 7 Simulated output waveform of the transmitter

Table 1 Performance table of designed transmitter

Technology	CMOS 0.13 μm
Voltage supply	1.2 V
Current consumption	1.57 mA
Power consumption	1.88 mW
Carrier frequency	433.92 MHz
Data rate	3.5 Mbps
Die area	1 mm ²

6 Conclusions

Low power transmitter for wireless capsule endoscope has been designed which achieves the purpose of low power consumption and high data rate for high quality medical imaging. It offers higher frame rates by having higher data rates of 3.5 Mbps than other works done and better imaging quality which could improves the diagnosis accuracy and acceptance among surgeons and patients. The low power design of 1.88 mW ensures the capsule endoscope works without power failure for the whole diagnosis, as well as enables more voltage headroom to other parts of capsule endoscope system.

References

1. Leighton J, Sharma V et al (2006) Capsule endoscopy versus push enteroscopy for evaluation of obscure gastrointestinal bleeding with 1-year outcomes. *Dig Dis Sci* 51(5):891–899
2. Song B, Nam D et al (2010) Auto-focusing actuator and camera module including flexible diaphragm for mobile phone camera and wireless capsule endoscope. *Microsyst Technol* 16 (1):149–159
3. Given Imaging from <http://www.givenimaging.com>
4. Swain P (2003) Wireless capsule endoscopy. *Gut* 52(4):48–50

5. Leong AFPK (2003) Wireless capsule endoscopy: light at the end of the tunnel for obscure gastrointestinal bleeding. *J Singap Med Assoc* 44(10):2
6. Lee YY, Rahim AA et al (2007) Persistent gastrointestinal bleeding after jejunal resection of a bleeding jejunal angiodysplasia. *J Med Sci Res* (2):51–53
7. Tokuhara D, Watanabe K et al (2010) Wireless capsule endoscopy in pediatric patients: the first series from Japan. *J Gastroenterol* 45(7):683–691
8. Swain P (2010) At a watershed? Technical developments in wireless capsule endoscopy. *J Dig Dis* 11(5):259–265
9. Seibel EJ, Carroll RE et al (2008) Tethered capsule endoscopy, a low-cost and high-performance alternative technology for the screening of esophageal cancer and Barrett's esophagus. *IEEE Trans Biomed Eng* 55(3):1032–1042
10. Xinkai C, Xiaoyu Z et al (2009) A wireless capsule endoscope system with low-power controlling and processing ASIC. *IEEE Trans Biomed Circuits Syst* 3(1):11–22
11. Woo S, Kim T et al (2010) Stopping mechanism for capsule endoscope using electrical stimulus. *Med Biol Eng Comput* 48(1):97–102
12. Sang Heun L, Jaebok L et al (2011) A wideband spiral antenna for ingestible capsule endoscope systems: experimental results in a human phantom and a pig. *IEEE Trans Biomed Eng* 58(6):1734–1741
13. Yasutomi K, Itoh S, Kawahito S (2011) A two-stage charge transfer active pixel CMOS image sensor with low-noise global shuttering and a dual-shuttering mode. *IEEE Trans Electron Devices*, 58(3):740–747
14. Shafie S, Kawahito S, Itoh S (2008) A dynamic range expansion technique for CMOS image sensors with dual charge storage in a pixel and multiple sampling. *Sensors* 8:1915–1926
15. Shafie S, Kawahito S (2008) A wide dynamic range CMOS image sensor with dual charge storage in a pixel and a multiple sampling technique. In: *Proceedings of SPIE IS&T 2008*, vol 6816-5, San Jose, California, USA, Jan 2008, pp 1–9
16. Shafie S, Kawahito S, Abdul Halin I, Hasan WZW (2009) Non-linearity in wide dynamic range CMOS image sensors utilizing a partial charge transfer technique. *Sensors* 9:9452–9467 (Switzerland)
17. Shafie S (2008) A study on dynamic range expansion techniques with reduced motion blur for CMOS image sensor. Dissertation, Shizuoka University
18. FCC (2003) MICS band plan. <http://www.fcc.gov>
19. Patel M, Jianfeng W (2010) Applications, challenges, and prospective in emerging body area networking technologies. *IEEE Wireless Commun* 17(1):80–88
20. Itoh S, Kawahito S et al (2006) A 2.6 mW 2 fps QVGA CMOS one-chip wireless camera with digital image transmission function for capsule endoscopes. In: *Proceedings of IEEE international symposium on circuits and systems*
21. Baoyong C, Jinke Y et al (2007) Low-power transceiver analog front-end circuits for bidirectional high data rate wireless telemetry in medical endoscopy applications. *IEEE Trans Biomed Eng* 54(7):1291–1299
22. Jiang H, Li F, Chen X, Ning Y, Zhang X, Zhang B, Ma T, Wang Z (2010) A SoC with 3.9 mW 3 Mbps UHF transmitter and 240 μ W MCU for capsule endoscope with bidirectional communication. In: *Proceedings of IEEE Asian solid state circuits conference (A-SSCC)*, pp 1–4
23. Diao S, Zheng Y, Gao Y, Heng C-H, Je M (2010) A 7.2 mW 15 Mbps ASK CMOS transmitter for ingestible capsule endoscopy. In: *Proceedings of IEEE Asia Pacific conference on circuits and systems (APCCAS)*, pp 512–515
24. Yuan G, Yuanjin Z et al (2011) Low-power ultrawideband wireless telemetry transceiver for medical sensor applications. *IEEE Trans Biomed Eng* 58(3):768–772
25. Lei W, Drysdale TD et al (2007) In Situ characterization of two wireless transmission schemes for ingestible capsules. *IEEE Trans Biomed Eng* 54(11):2020–2027
26. Thoné J, Radiom S et al (2009) Design of a 2 Mbps FSK near-field transmitter for wireless capsule endoscopy. *Sens Actuators A* 156(1):43–48

27. Kim DK (2008) Considerations on high-data-rate in-body communications for WBAN. *IEEE* 802:15
28. Caverly R (2007) CMOS RFIC design principles. Artech House, Boston
29. Hajimiri A, Limotyrakis S et al (1999) Jitter and phase noise in ring oscillators. *IEEE J Solid-State Circuits* 34(6):790–804
30. Tekin A, Yuce MR et al (2008) Integrated VCOs for medical implant transceivers. *VLSI Des* 2008(912536):10

Romanovski Polynomials Method and Its Application for Non-central Potential System

A. Suparmi and C. Cari

Abstract The approximate analytical solution of Schrodinger equation for Eckart potential plus with trigonometric Poschl-Teller noncentral potential and trigonometric Rosen-Morse non-central potential systems are investigated using Romanovski polynomials. The approximate bound state energy eigenvalue of the first system is given in the close form and the corresponding approximate radial eigen functions is formulated in the form of Romanovski polynomials while the angular wave function is also expressed in Romanovski polynomials. The effect of the presence of trigonometric Poschl-Teller potential increases the angular wave function level. The presence of non-central potentials cause the orbital quantum numbers are mostly not integer.

1 Introduction

Schrodinger equations for a class of shape invariant potentials have been solved by using some methods such as SUSY WKB (SWKB) [1–4], SUSY operator and factorization method [5–9], NU method [10–14] and Romanovski polynomials [15–18]. Romanovski polynomials, which is a traditional method, consists of reducing Schrodinger equation by an appropriate change of the variable to the form of generalized hypergeometric equation [19]. The solution steps of Romanovski polynomials are rather similar to the steps applied in solution of Schrodinger equation by using NU method [20, 21], which is discussed in Chap. 5, and hypergeometric differential equation. The polynomial was discovered by Sir Routh [22] and rediscovered 45 years later by Romanovski [23]. Romanovski polynomial method is also called as finite Romanovski polynomial. The notion “finite” refers to the observation

A. Suparmi (✉) · C. Cari

Physics Department, Sebelas Maret University, Surakarta 57126, Indonesia
e-mail: suparmiuns@gmail.com

C. Cari

e-mail: carinln@yahoo.com

that, for any given set of parameters (i.e. in any potential) only a finite of polynomials appear orthogonal [24]. However the polynomial will not be finite if the certain condition is not satisfied and then leads to an infinite of polynomials that show orthogonal. It seems that NU and Romanovski polynomials methods are very similar in the way of variable substitution but they solve Schrodinger equation differently. NU method is applied wider than Romanovski polynomials since Romanovski polynomials could not be applied for potential solved using confluent hypergeometric differential equation.

In this chapter we discuss the Schrodinger equation solution for a particle which is in the field of Eckart potential with simultaneously presence of Poschl-Teller non-central potential and the polar Schrodinger equation for trigonometric Rosen-Morse potential using Romanovski polynomials. A non-central potential is potential as a function of radial and angular positions, it could be composed of radial function potentials and non-central potentials which are shown in Chap. 5. The simple choice of non-central potential is the separable potential [25–27]. The three dimensional Schrodinger equations of separable non-central potentials are exactly solvable as long the centrifugal term is approximated by hyperbolic function, trigonometric function or exponential function [28–32]. Due to the approximation of the centrifugal term, the energy spectra and the radial wave functions are approximately obtained for l -state solution and becomes exact solution for s-wave.

2 Romanovski Polynomials

One dimensional Schrodinger equation of potential of interest reduces to the differential equation of Romanovski polynomials by appropriate variable and wave function substitutions. The one dimensional Schrodinger equation is given as

$$-\frac{\hbar^2}{2M} \frac{\partial^2 \Psi(x)}{\partial x^2} + V(x) \Psi(x) = E \Psi(x) \quad (1)$$

where $V(x)$ is an effective potential which is mostly shape invariant potential. By suitable variable substitution $x = f(s)$ (1) changes into generalized hypergeometric type equation expressed as

$$\frac{\partial^2 \Psi(s)}{\partial s^2} + \frac{\tilde{\tau}(s)}{\sigma(s)} \frac{\partial \Psi(s)}{\partial s} + \frac{\tilde{\sigma}(s)}{\sigma^2(s)} \Psi(s) = 0 \quad (2)$$

with $\sigma(s)$ and $\tilde{\sigma}(s)$ are mostly polynomials of order two, $\tilde{\tau}(s)$ is polynomial of order one, of s , $\sigma(s)$, $\tilde{\sigma}(s)$, and $\tilde{\tau}(s)$ can have any real or complex values [21]. Equation (2) is solved by variable separation method. By introducing new wave function in (2) we obtain a hypergeometric type differential equation, which can be solved using finite Romanovski polynomials [16, 17, 21] which is expressed as

$$\sigma(s)y''(s) + \tau(s)y'(s) + \lambda y(s) = 0 \quad (3a)$$

with

$$\sigma(s) = as^2 + bs + c; \tau = fs + h \text{ and } -\{n(n-1) + 2n(1-p)\} = \lambda = \lambda_n \quad (3b)$$

Equation (3a) is described in the textbook by Nikiforov-Uvarov [21] where it is cast into self adjoint form and its weight function, $w(s)$, satisfies Pearson differential equation

$$\frac{d(\sigma(s)w(s))}{ds} = \tau(s)w(s) \quad (4)$$

The weight function, $w(s)$, is obtained by solving the Pearson differential equation expressed in (4) and by applying condition in (3b), that is

$$w(s) = \exp\left(\int \frac{(f-2a)s + (h-b)}{as^2 + bs + c} ds\right) \quad (5)$$

The corresponding polynomials are classified according to the weight function, and are built up from the Rodrigues representation which is presented as

$$y_n = \frac{B_n}{w(s)} \frac{d^n}{ds^n} \{(as^2 + bs + c)^n w(s)\} \quad (6)$$

with B_n is a normalization constant, and for $\sigma(s) > 0$ and $w(s) > 0$, $y_n(s)$'s are normalized polynomials and are orthogonal with respect to the weight function $w(s)$ within a given interval (s_1, s_2) , which is expressed as

$$\int_{-\infty}^{\infty} w(s)y_n(s)y_{n'}(s)ds = \delta_{nn'} \quad (7)$$

For Romanovski polynomials, the values of parameters in (3b) are:

$$a = 1, b = 0, c = 1, f = 2(1-p) \text{ and } h = q \text{ with } p > 0 \quad (8)$$

By inserting (8) into (5) we obtain the weight function as

$$w(s) = \exp\left(\int \frac{(f-2a)s + (c-b)}{as^2 + bs + c} ds\right) = \exp\left(\int \frac{(2-2p-2)s + q}{s^2 + 1} ds\right) \quad (9)$$

$$w(s) = (1+s^2)^{-p} e^{q \tan^{-1}(s)}$$

This weight function first reported by Routh [19] and then by Romanovski [23]. The polynomial associated with (9) are named after Romanovski and will be

denoted by $R_n^{(p,q)}(s)$. Due to the decrease of the weight function by s^{-2p} , integral of the type

$$\int_{-\infty}^{\infty} w^{(p,q)} R_n^{(p,q)}(s) R_{n'}^{(p,q)}(s) ds \quad (10)$$

will be convergent only if

$$n' + n < 2p - 1 \quad (11)$$

This means that only a finite number of Romanovski polynomials are orthogonal, and the orthogonality integral of the polynomial is expressed similar to the (7) where $y_n = R_n^{(p,q)}(s)$.

The differential equation satisfied by Romanovski Polynomial obtained by inserting (3b) and (8) into (3a) given as

$$\begin{aligned} (1+s^2) \frac{\partial^2 R_n^{(p,q)}}{\partial s^2} + \{2s(-p+1) + q\} \frac{\partial R_n^{(p,q)}(s)}{\partial s} \\ - \{n(n-1) + 2n(1-p)\} R_n^{(p,q)}(s) = 0 \end{aligned} \quad (12)$$

where $y_n = R_n^{(p,q)}(s)$. The heart of Romanovski polynomials method is in obtaining (12) from one dimensional Schrodinger equation. The Schrodinger equation of the potential of interest will be reduced into second order differential equation that is similar to (12) by an appropriate transformation of variable, for example, $r = f(s)$, and by introducing a new wave function which is given as

$$\Psi_n(r) = g_n(s) = (1+s^2)^{\frac{\beta}{2}} e^{\frac{-\alpha}{2} \tan^{-1} s} D_n^{(\beta,\alpha)}(s) \quad (13)$$

where $\Psi_n(s) = \Psi_n(x)$ is an eigen function of generalized hypergeometric equation in (2) which is the solution of Schrodinger equation for potential interest in (1), and

$$D_n^{(\beta,\alpha)}(s) = R_n^{(p,q)}(s) \quad (14)$$

From condition in (14) we get the relation between β with p , and α with q . The Romanovski polynomials obtained from Rodrigues formula expressed in (6) for the corresponding weight function in (9) is expressed as

$$R_n^{(p,q)}(s) = D_n^{(\beta,\alpha)}(s) = \frac{1}{(1+s^2)^{-p} e^{q \tan^{-1}(s)}} \frac{d^n}{ds^n} \left\{ (1+s^2)^n (1+s^2)^{-p} e^{q \tan^{-1}(s)} \right\} \quad (15)$$

If the wave function of the n th level in (13) is rewritten as

$$\Psi_n(r) = \frac{1}{\sqrt{\frac{df(s)}{ds}}} (1 + s^2)^{\frac{-p}{2}} e^{\frac{q}{2} \tan^{-1}(s)} R_n^{(p,q)}(s) \quad (16)$$

then the orthogonality integral of the wave functions expressed in (16) gives rise to orthogonality integral of the finite Romanovski polynomials, that is given as

$$\int_0^\infty \Psi_n(r) \Psi_{n'}(r) dr = \int_{-\infty}^\infty w^{(p,q)} R_n^{(p,q)}(s) R_{n'}^{(p,q)}(s) ds \quad (17)$$

In this case the values of p and q are not n -dependence where n is the degree of polynomials. However, if either (11) or (17) is not fulfilled then the Romanovski polynomials is infinity [16–18].

3 Application of Romanovski Polynomials for Energy Spectra and Wave Functions Analysis for Non-central Potential

Non-central potentials which consist are solvable by Romanovski polynomials. The non-central potentials that are solved using Romanovski polynomials are Eckart plus Poschl-Teller non-central potential system and polar Schrodinger equation of 3D oscillator harmonics plus trigonometric Rosen-Morse non-central potential system.

3.1 Eckart Plus Poschl-Teller Non-central Potential

The non-central potential is a potential of a function radial and angular simultaneously. The non-central potential which is constructed from Eckart potential and trigonometric Poschl-Teller non-central potential given as

$$V(r, \theta) = \frac{\hbar^2}{2Ma^2} \left(V_0 \frac{e^{-r/a}}{(1 - e^{-r/a})^2} - V_1 \frac{1 + e^{-r/a}}{1 - e^{-r/a}} \right) + \frac{\hbar^2}{2Mr^2} \left(\frac{\kappa(\kappa - 1)}{\sin^2 \theta} + \frac{\eta(\eta - 1)}{\cos^2 \theta} \right) \quad (18)$$

with V_0 and V_1 describe the depth of the potential well and are positives, $V_1 > V_0$, a is a positive parameter which to control the width of the potential well, M is the mass of the particle, and $0 < (r/a) < \infty$, $\kappa > 1$, $\eta > 1$. The non-central potential

expressed in (18) is separable ones therefore the Schrodinger equation of this potential is solved using variable separation method.

The three dimensional time-independent Schrodinger equation for Eckart potential combined with trigonometric Poschl-Teller non-central potential is

$$\begin{aligned}
 & -\frac{\hbar^2}{2M} \left\{ \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right\} \psi(r, \theta, \varphi) \\
 & + \frac{\hbar^2}{2Ma^2} \left(V_0 \frac{e^{-r/a}}{(1 - e^{-r/a})^2} - V_1 \frac{1 + e^{-r/a}}{1 - e^{-r/a}} \right) \psi(r, \theta, \varphi) \\
 & + \frac{\hbar^2}{2Mr^2} \left(\frac{\kappa(\kappa - 1)}{\sin^2 \theta} + \frac{\eta(\eta - 1)}{\cos^2 \theta} \right) \psi(r, \theta, \varphi) = E \psi(r, \theta, \varphi)
 \end{aligned} \tag{19}$$

By using variable separation method we get radial, polar and azimuthal parts of Schrodinger equation as following:

$$\frac{1}{R} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R}{\partial r} \right) - \frac{r^2}{a^2} \left(V_0 \frac{e^{-r/a}}{(1 - e^{-r/a})^2} - V_1 \frac{1 + e^{-r/a}}{1 - e^{-r/a}} \right) + \frac{2Mr^2}{\hbar^2} E = \lambda = l(l + 1) \tag{20}$$

$$\left\{ -\frac{1}{P \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial P}{\partial \theta} \right) - \frac{1}{\Phi \sin^2 \theta} \frac{\partial^2 \Phi}{\partial \varphi^2} \right\} + \left(\frac{\kappa(\kappa - 1)}{\sin^2 \theta} + \frac{\eta(\eta - 1)}{\cos^2 \theta} \right) = \lambda \tag{21}$$

$$\frac{1}{\Phi} \frac{\partial^2 \Phi}{\partial \varphi^2} = -m^2 \tag{22}$$

The azimuthal part of wave function obtained from (22) is given, as usual, as

$$\Phi = A_m e^{im\varphi} \tag{23}$$

The radial and polar parts of the Schrodinger equations are solved using Romanovski Polynomials.

3.1.1 Solution of Radial Part of Schrodinger Equation

By substitution $\frac{2M}{\hbar^2} E = -\varepsilon^2$ and $R = \frac{\chi(r)}{r}$ in (20) we get

$$\frac{\partial^2 \chi(r)}{\partial r^2} - \frac{1}{a^2} \left(V_0 \frac{e^{-r/a}}{(1 - e^{-r/a})^2} - V_1 \frac{1 + e^{-r/a}}{1 - e^{-r/a}} \right) \chi(r) - \varepsilon^2 \chi(r) - \frac{l(l + 1)}{r^2} \chi(r) = 0 \tag{24}$$

For $\frac{r}{a} < < 1$ the approximation of the centrifugal term in (24) [18, 19] is given as $\frac{1}{r^2} \cong \frac{1}{4a^2} \left(d_0 + \frac{1}{\sinh^2(r/2a)} \right)$ with $d_0 = 1/12$. In term of hyperbolic functions, (24) is rewritten as

$$\begin{aligned} & \frac{d^2\chi(r)}{dr^2} - \frac{l(l+1)}{4a^2} \left(d_0 + \frac{1}{\sinh^2(r/2a)} \right) \chi(r) \\ & - \frac{1}{a^2} \left(\frac{V_0}{4 \sinh^2(r/2a)} - V_1 \coth(r/2a) \right) \chi(r) - \varepsilon^2 \chi(r) \\ & = 0 \end{aligned} \quad (25)$$

and by making an appropriate change of variable, $r = f(x) = 2a \coth^{-1}(ix)$ in (25), we get

$$(1+x^2) \frac{d^2\chi}{dx^2} + 2x \frac{d\chi}{dx} - \left\{ V_0 + l(l+1) + \frac{4V_1 ix}{(1+x^2)} - \frac{l(l+1)d_0 + 4\varepsilon^2 a^2}{(1+x^2)} \right\} \chi = 0 \quad (26)$$

To solve (26) in terms of Romanovski polynomial, (13) suggests the substitution in (26) as [29]

$$\chi(f(x)) = g_n(x) = (1+x^2)^{\frac{\beta}{2}} e^{\frac{-\alpha}{2} \tan^{-1} x} D_n^{(\beta, \alpha)}(x) \quad (27)$$

where $1 < ix < \infty$.

By inserting (27) into (24) we obtain

$$\begin{aligned} & (1+x^2) \frac{\partial^2 D}{\partial x^2} + \{2x(\beta+1) - \alpha\} \frac{\partial D}{\partial x} \\ & - \left\{ \frac{\beta\alpha x - \frac{\alpha^2}{4} + \beta^2 + 4V_1 ix - (l(l+1)d_0 + 4\varepsilon^2 a^2)}{1+x^2} + V_0 + l(l+1) - \beta^2 - \beta \right\} D \\ & = 0 \end{aligned} \quad (28)$$

Equation (28) reduces to differential equation that satisfied by Romanovski polynomials if the coefficient of $\frac{1}{1+x^2}$ term in (28) is set to be zero, that are

$$-\frac{\alpha^2}{4} + \beta^2 - (l(l+1)d_0 + 4\varepsilon^2 a^2) = 0 \text{ and } \beta\alpha + 4V_1 i = 0 \quad (29)$$

and then (28) becomes

$$(1+x^2)\frac{\partial^2 D}{\partial x^2} + \{2x(\beta+1) - \alpha\}\frac{\partial D}{\partial x} - \{V_0 + l(l+1) - \beta^2 - \beta\}D = 0 \quad (30)$$

It is seen that the structure of (30) is similar to (12), and thus we assume that $D_n^{(\alpha,\beta)}(x) \cong R_n^{(p,q)}(x)$. By comparing the parameters between (12) and (30) we obtain the following relation:

$$\begin{aligned} V_0 + l(l+1) - \beta^2 - \beta &= n(n-1) + 2n(1-p); \\ 2(\beta+1) &= 2(-p+1) \text{ and } \alpha = -q \end{aligned} \quad (31)$$

From (31) we have $p = -\beta$, since $p > 0$ then the value of β obtained from (31) that has physical meaning is

$$\beta = \beta_n = -\sqrt{V_0 + \left(l + \frac{1}{2}\right)^2} - n - \frac{1}{2} \quad (32)$$

From (29) and (32) we obtain

$$\alpha^2 = -2(l(l+1)d_0 + 4\varepsilon^2 a^2) \pm 2\sqrt{\{l(l+1)d_0 + 4\varepsilon^2 a^2\}^2 - 16V_1^2} \quad (33)$$

$$\text{with } \alpha = \alpha_n = \frac{4V_1 i}{\sqrt{V_0 + \left(l + \frac{1}{2}\right)^2} + n + \frac{1}{2}} \quad (34)$$

Finally, the energy spectrum of the system achieved from (33) and (34) is given as

$$E_n = -\frac{\hbar^2}{2M} \left\{ \frac{\left(\sqrt{V_0 + \left(l + \frac{1}{2}\right)^2} + n + \frac{1}{2}\right)^2}{4a^2} + \frac{V_1^2}{a^2 \left(\sqrt{V_0 + \left(l + \frac{1}{2}\right)^2} + n + \frac{1}{2}\right)^2} - \frac{l(l+1)d_0}{4a^2} \right\} \quad (35)$$

The energy spectra of Eckart potential obtained using Romanovski polynomials in (35) is in agreement with the energy spectra obtained using NU method which is presented in Chap. 5.

To determine the wave function, (31), (32) and (34) are inserted into (9) and (15) so that we obtain the weight function $w(x)$ and the Romanovski polynomials $R_n^{(p,q)} = R_n^{(-\beta,-\alpha)}(x)$ as

$$w^{(-\beta,-\alpha)} = (1+x^2)^{-\left(\sqrt{V_0 + \left(l + \frac{1}{2}\right)^2} + n + \frac{1}{2}\right)} \exp\left(-\frac{4iV_1}{\sqrt{V_0 + \left(l + \frac{1}{2}\right)^2} + n + \frac{1}{2}} \tan^{-1}(x)\right) \quad (36a)$$

and

$$\begin{aligned} D_n^{(\beta, \alpha)}(x) &= R_n^{(p, q)}(x) = R_n^{(-\beta, -\alpha)}(x) \\ &= \frac{1}{(1+x^2)^{\beta_n} e^{-\alpha_n \tan^{-1}(x)}} \frac{d^n}{dx^n} \left\{ (1+x^2)^{\beta_n+n} e^{-\alpha_n \tan^{-1}(x)} \right\} \end{aligned} \quad (36b)$$

where β_n and α_n are expressed in (32) and (34). As a result, the wave function of the n th level is given by

$$\chi(f(x)) = g_n(x) = (1+x^2)^{\frac{\beta_n}{2}} e^{\frac{-\alpha_n}{2} \tan^{-1} x} R_n^{(-\beta, -\alpha)}(x) \quad (37)$$

By using trigonometric-hyperbolic functions relation

$$e^{-\frac{\alpha_n}{2} \tan^{-1} x} = e^{-\frac{\alpha_n}{2} \tan^{-1}(-i \coth(r/2a))} = \left(\frac{1 + \coth(r/2a)}{1 - \coth(r/2a)} \right)^{\frac{i\alpha_n}{4}} \quad (38)$$

then (37) becomes

$$\begin{aligned} \chi_n(r) &= g_n(r) \\ &= (1 - \coth(r/2a))^{\frac{\beta_n}{2} - \frac{i\alpha_n}{4}} (1 + \coth(r/2a))^{\frac{\beta_n}{2} + \frac{i\alpha_n}{4}} R_n^{(-\beta, -\alpha)}(-i \coth(r/2a)) \end{aligned} \quad (39)$$

$$\begin{aligned} R_n^{(-\beta, -\alpha)}(r) &= \frac{i^n}{(1 - \coth(r/2a))^{\beta_n - \frac{i\alpha_n}{2}} (1 + \coth(r/2a))^{\beta_n + \frac{i\alpha_n}{2}}} \\ &\quad \frac{d^n}{d(\coth(r/2a))^n} \left\{ (1 - \coth(r/2a))^{\beta_n - \frac{i\alpha_n}{2} + n} (1 + \coth(r/2a))^{\beta_n + \frac{i\alpha_n}{2} + n} \right\} \end{aligned} \quad (40)$$

The radial wave functions for Eckart plus Poschl-Teller non-central potential in (39) and (40) are in agreement with the result achieved by using NU method.

Since the β_n and α_n parameters, expressed in (32) and (34), are n -dependence then the orthogonality of the wave functions may not produce to the orthogonality integral of the polynomials [29], as shown in (41),

$$\int_0^\infty \chi_n(r) \chi_{n'}(r) dr = \delta_{nn'} \neq \int_1^\infty w^{(-\beta, -\alpha)} R_n^{(-\beta, -\alpha)}(x) R_{n'}^{(-\beta, -\alpha)}(x) dx \quad (41)$$

By carrying out the differentiations of (36b), we find the lowest four unnormalized Romanovski polynomials given as

$$R_0^{(-\beta_0, -\alpha_0)}(x) = 0 \quad (42)$$

$$R_1^{(-\beta_1, -\alpha_1)}(x) = (\beta_1 + 1)2x - \alpha_1 \quad (43)$$

$$R_2^{(-\beta_2, -\alpha_2)}(x) = 2(\beta_2 + 2)(2\beta_2 + 3)x^2 - 2\alpha_2(2\beta_2 + 3)x + \alpha_2^2 + 2\beta_2 + 4 \quad (44)$$

$$R_3^{(-\beta_3, -\alpha_3)}(x) = 4x^3(\beta_3 + 3)(2\beta_3 + 5)(\beta_3 + 2) - 6\alpha_3x^2(2\beta_3 + 5)(\beta_3 + 2) \\ + 2x(6\beta_3^2 + 3\alpha_3^2\beta_3 + 28\beta_3 + 6\alpha_3^2 + 34) - 2\alpha_3(2\beta_3 + 5) - \alpha(\alpha^2 + 2\beta + 6) \quad (45)$$

where β_n and α_n are expressed in (32) and (34). The lowest four degrees of un-normalized radial wave functions for arbitrary values of l are calculated by using (41) and (42)–(45).

3.1.2 The Solution of Polar Schrodinger Equation for Eckart Potential Combined with Non Central Poschl-Teller Potential

To solve the polar Schrodinger equation expressed in (19), we set the polar wave function as

$$P = \frac{Q(\theta)}{\sqrt{\sin \theta}}, \quad \frac{\partial P}{\partial \theta} = \frac{\frac{dQ}{d\theta}}{\sqrt{\sin \theta}} - \frac{1}{2} \frac{\cos \theta Q}{\sqrt{(\sin \theta)^3}} \quad (46)$$

where $Q(\theta)$ is the new polar wave function. By inserting (46) into (19) we obtain one dimensional polar Schrodinger equation as

$$\frac{d^2 Q}{d\theta^2} - \left\{ \frac{\kappa(\kappa - 1) + m^2 - (1/4)}{\sin^2 \theta} + \frac{\eta(\eta - 1)}{\cos^2 \theta} \right\} Q + (l(l + 1) + (1/4))Q = 0 \quad (47)$$

Equation (47) to be solved using Romanovski polynomials, therefore we substitute the variable θ and introduce a new wave function such that (47) reduces to generalized hypergeometric type equation expressed in (2) or into second order differential equation of Romanovski polynomials expressed in (10). By making a change of variable in (47), $\cos 2\theta = is$ then (47) becomes

$$(1 + s^2) \frac{\partial^2 Q}{\partial s^2} + s \frac{\partial Q}{\partial s} + \left[\frac{\kappa(\kappa - 1) + m^2 - \frac{1}{4} + \eta(\eta - 1)}{2(1 + s^2)} + \frac{\kappa(\kappa - 1) + m^2 - \frac{1}{4} - \eta(\eta - 1)}{2(1 + s^2)} is - \frac{(l + \frac{1}{2})^2}{4} \right] Q = 0 \quad (48)$$

Equation (48) will be reduced into differential equation of Romanovski polynomial by setting

$$Q(\theta) = g(s) = (1 + s^2)^{\frac{\beta}{2}} e^{\mp \tan^{-1} s} D_n^{(\alpha, \beta)}(s) \quad (49)$$

By inserting (49) into (48) we get

$$(1 + s^2) \frac{\partial^2 D}{\partial s^2} + \{s(2\beta + 1) - \alpha\} \frac{\partial D}{\partial s} - \left\{ \frac{2\beta s\alpha - s\alpha - \frac{\alpha^2}{2} + 2\beta^2 - 2\beta - \{\kappa(\kappa - 1) + m^2 - \frac{1}{4} - \eta(\eta - 1)\}is - \{\kappa(\kappa - 1) + m^2 - \frac{1}{4} + \eta(\eta - 1)\}}{2(1 + s^2)} + \frac{(l + \frac{1}{2})^2}{4} - \beta^2 \right\} D = 0 \quad (50a)$$

Equation (50a) reduces to the differential equation satisfied by Romanovski polynomials given as

$$(1 + s^2) \frac{\partial^2 D}{\partial s^2} + s(2\beta + 1) - \alpha \left\{ \frac{\partial D}{\partial s} - \left\{ \frac{(l + \frac{1}{2})^2}{4} - \beta^2 \right\} D \right\} = 0 \quad (50b)$$

when the coefficient of $\frac{1}{2(1+s^2)}$ is set to be zero, that are

$$-\left\{ \kappa(\kappa - 1) + m^2 - \frac{1}{4} + \eta(\eta - 1) \right\} - \frac{\alpha^2}{2} + 2\beta^2 - 2\beta = 0 \quad (51)$$

$$-\left\{ \kappa(\kappa - 1) + m^2 - \frac{1}{4} - \eta(\eta - 1) \right\} i + 2\beta\alpha - \alpha = 0 \quad (52)$$

By comparing (50b) and (10) we obtain

$$(2\beta + 1) = 2(-p + 1); \quad a = -q \quad (53)$$

$$\text{and } \frac{(l + \frac{1}{2})^2}{4} - \beta^2 = n(n - 1) + 2n(1 - p) \quad (54)$$

By using (51) and (52) and setting

$$\left(\kappa - \frac{1}{2} \right)^2 + m^2 - \frac{1}{4} = \kappa(\kappa - 1) + m^2 = F, \quad (\eta - \frac{1}{2})^2 = G \quad (55)$$

we obtain

$$\alpha = \pm i(\sqrt{F} \mp \sqrt{G}) \text{ and } \left(\beta - \frac{1}{2} \right) = \frac{(F - G)}{\pm 2i(\sqrt{F} \mp \sqrt{G})} i \quad (56)$$

To have physical meaning, the proper choice of the values of α and β in (56) are

$$\begin{aligned}\alpha &= i\left(\sqrt{F} - \sqrt{G}\right) = i\left(\sqrt{\kappa(\kappa-1) + m^2} - \left(\eta - \frac{1}{2}\right)\right) \text{ and} \\ \beta &= \frac{(\sqrt{F} + \sqrt{G}) + 1}{2} = \frac{\sqrt{\kappa(\kappa-1) + m^2} + (\eta + \frac{1}{2})}{2}\end{aligned}\quad (57)$$

From (54) and (57) we have

$$l = \left(\sqrt{\left(\kappa - \frac{1}{2}\right)^2 + m^2} - \frac{1}{4} + \eta + 2n \right) \quad (58)$$

Equation (58) shows that the values of l depend on the potential parameters, κ and η , and the degree of the Romanovski polynomial, $n = n_l$. The weight function obtained from (7), (54) and (57) is given as

$$w^{(p,q)} = w^{(-\beta+\frac{1}{2}, -\alpha)} = (1+s^2)^{\frac{\sqrt{F}+\sqrt{G}}{2}} e^{-i(\sqrt{F}-\sqrt{G})\tan^{-1}s} \quad (59)$$

The Romanovski polynomials are obtained by using (13) and (61) as

$$R_n^{(-\beta+\frac{1}{2}, -\alpha)}(s) = \frac{1}{(1+s^2)^{\frac{\sqrt{F}+\sqrt{G}}{2}} e^{-i(\sqrt{F}-\sqrt{G})\tan^{-1}s}} \frac{d^n}{ds^n} \left\{ \left((1+s^2)^{\frac{\sqrt{F}+\sqrt{G}}{2}+n} e^{-i(\sqrt{F}-\sqrt{G})\tan^{-1}s} \right) \right\} \quad (60)$$

and the polar wave functions obtained from (11) and (60) is given as

$$Q(\theta) = g_n(s) = (1+s^2)^{\frac{(\sqrt{F}+\sqrt{G})+1}{4}} e^{-\frac{i}{2}(\sqrt{F}-\sqrt{G})\tan^{-1}s} R_n^{(-\beta+\frac{1}{2}, -\alpha)}(s) \quad (61)$$

The polar eigen function obtained from (46) and (61) is given as

$$P_{nm\kappa\eta}(\theta) = \frac{Q_{nm\kappa\eta}}{\sqrt{\sin \theta}} = \frac{1}{\sqrt{\sin \theta}} \sqrt{(1+s^2)^{\frac{\sqrt{F}+\sqrt{G}+1}{2}} e^{-i(\sqrt{F}-\sqrt{G})\tan^{-1}s}} R_n^{(-\beta+\frac{1}{2}, -\alpha)}(s) \quad (62a)$$

$$\text{or } P_{nm\kappa\eta}(\theta) = \frac{1}{\sqrt{\sin \theta}} (1 - \cos 2\theta)^{\frac{\sqrt{F}}{2}+\frac{1}{4}} (1 + \cos 2\theta)^{\frac{\sqrt{G}}{2}+\frac{1}{4}} R_{n_l}^{(-\beta+\frac{1}{2}, -\alpha)}(-i \cos 2\theta) \quad (62b)$$

The orthogonality integral of the angular wave function obtained from (62a), (62b) is given as

$$\begin{aligned}
\int_0^{\frac{\pi}{2}} Q_n^*(\theta) Q_{n'}(\theta) d\theta &= \int_{-1}^1 (1+s^2)^{\frac{(\sqrt{F}+\sqrt{G})+1}{2}} e^{-\frac{i}{2}(\sqrt{F}-\sqrt{G})\tan^{-1}s} R_n^{(-\beta+\frac{1}{2},-\alpha)}(s) R_{n'}^{(-\beta+\frac{1}{2},-\alpha)}(s) \frac{\partial\theta}{\partial(is)} d(is) \\
&= -\frac{1}{2} \int_{-1}^1 (1+s^2)^{\frac{(\sqrt{F}+\sqrt{G})}{2}} e^{-\frac{i}{2}(\sqrt{F}-\sqrt{G})\tan^{-1}s} R_n^{(-\beta+\frac{1}{2},-\alpha)}(s) R_{n'}^{(-\beta+\frac{1}{2},-\alpha)}(s) d(is)
\end{aligned} \tag{63a}$$

where $\frac{\partial\theta}{\partial(is)} = -\frac{1}{2\sin 2\theta}$, $(1+s^2)^{1/2} = \sin 2\theta$ and from (63a) we have

$$\int_0^{\frac{\pi}{2}} Q_n^*(\theta) Q_{n'}(\theta) d\theta = -\frac{1}{2} \int_{-1}^1 w^{(-\beta+\frac{1}{2},-\alpha)} R_n^{(-\beta+\frac{1}{2},-\alpha)}(s) R_{n'}^{(-\beta+\frac{1}{2},-\alpha)}(s) d(is) \tag{63b}$$

Equation (63b) shows that the orthogonality of Romanovski polynomials is produced from the orthogonality of wave function but (63a), (63b) is not convergent [29, 30] since

$$n + n' < 2p - 1 \text{ or } n < -\frac{\sqrt{(\kappa - \frac{1}{2})^2 + m^2 - \frac{1}{4} + \eta + \frac{1}{2}}}{2} \tag{64}$$

and the interval of the variable is not in $-\infty < s < \infty$ interval.

The first four unnormalized Romanovski polynomials obtained from (60) are given as:

$$R_0^{(-\beta+\frac{1}{2},-\alpha)}(s) = 1 \tag{65}$$

$$R_1^{(-\beta+\frac{1}{2},-\alpha)}(s) = (\sqrt{F} + \sqrt{G} + 2)s - i(\sqrt{F} - \sqrt{G}) \tag{66}$$

$$R_2^{(-\beta+\frac{1}{2},-\alpha)}(s) = \left[(\sqrt{F} + \sqrt{G} + 4)(\sqrt{F} + \sqrt{G} + 3)s^2 - 2is(\sqrt{F} - \sqrt{G})(\sqrt{F} + \sqrt{G} + 3) - (\sqrt{F} - \sqrt{G})^2 + (\sqrt{F} + \sqrt{G} + 4) \right] \tag{67}$$

$$R_3^{(-\beta+\frac{1}{2},-\alpha)}(s) = \left[\left\{ (\sqrt{F} + \sqrt{G} + 6)(\sqrt{F} + \sqrt{G} + 5)((\sqrt{F} + \sqrt{G} + 4)s^3 - 3is^2(\sqrt{F} - \sqrt{G})(\sqrt{F} + \sqrt{G} + 5)(\sqrt{F} + \sqrt{G} + 4) \right. \right. \\
\left. \left. - 3s(\sqrt{F} + \sqrt{G} + 4)\{(\sqrt{F} - \sqrt{G})^2 - (\sqrt{F} + \sqrt{G} + 6)\} + i\{(\sqrt{F} - \sqrt{G})^2 - (3\sqrt{F} + 3\sqrt{G} + 16)\}(\sqrt{F} - \sqrt{G}) \right\} \right] \tag{68}$$

The solution of the first four of un-normalized polar wave functions are obtained from (57), (62a), (62b) and (65)–(68). The polar wave function (62b) is in agreement to the wave function obtained using NU method [31], but only for even numbers of polynomial degrees (n_l).

If there is no the presence of trigonometric Poschl-Teller potential, where $\kappa = 0$ and $\eta = 0$ then the polar wave function reduces to associated Legendre polynomials

and the orbital quantum number expressed in (36a), (36b) becomes $l = m + 2n_l$, with n_l is the degree of polynomial. However the associated Legendre polynomial obtained from this non-central potential are only those polynomials whose values of l and m are differed by even numbers since $l = m + 2n_l$. The effect of the presence of Poschl-Teller non-central potential to spherical harmonics is illustrated by using the three dimensional representation and the polar diagram of the absolute value of un-normalized angular wave functions obtained from (55), (62a), (62b), and (68) for $n = n_l = 3$. The 3D representations and polar diagram of $|Y_l^m|$ visualized using Mat Lab 7 are shown in Fig. 1 for $n_l = 3, \kappa = 0, \eta = 0, l = 6, m = 0$, Fig. 2 for $n_l = 3, \kappa = 2, \eta = 0, l = 7.4, m = 0$ and Fig. 3 for $n_l = 3, \kappa = 0, \eta = 2, l = 8, m = 0$. By comparing Figs. 2 and 3 with Fig. 1, it is concluded that there is a state change in angular wave function caused to the presence of Poschl-Teller potentials. Therefore it may be concluded that the number of the degeneracy of the system changes. By comparing Fig. 3 with Figs. 1 and 2 is concluded that the $\sec^2 \theta$ causes the change of the angular wave function state, while the effect of $\operatorname{cosec}^2 \theta$ term causes the absolute values of the angular wave function shifted to larger values of θ . Therefore the

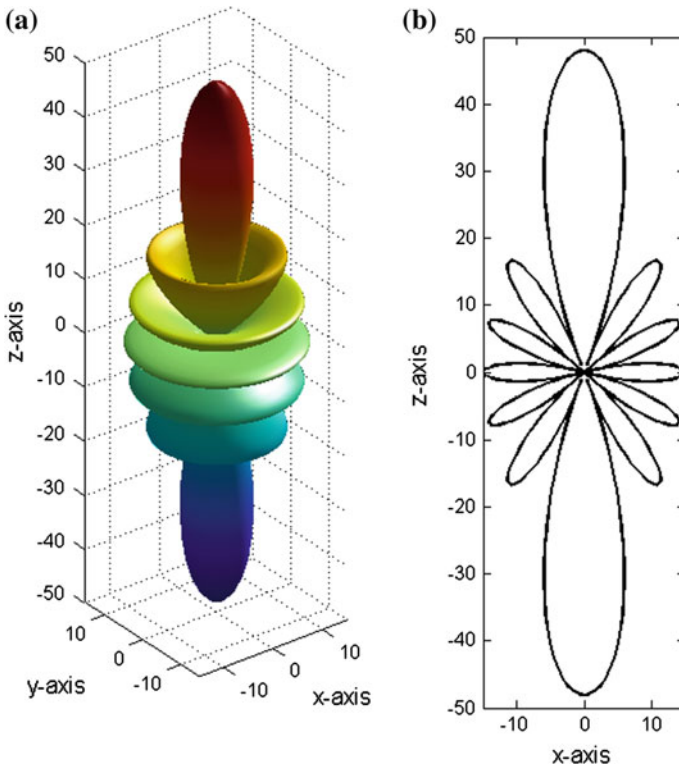


Fig. 1 **a** Three dimensional representation of $|Y_6^0|$ and **b** its polar diagram of $|Y_6^0| = \{86.63 \cos^3 2\theta + 23.63 \cos^2 2\theta - 55.13 \cos 2\theta - 7.13\}$

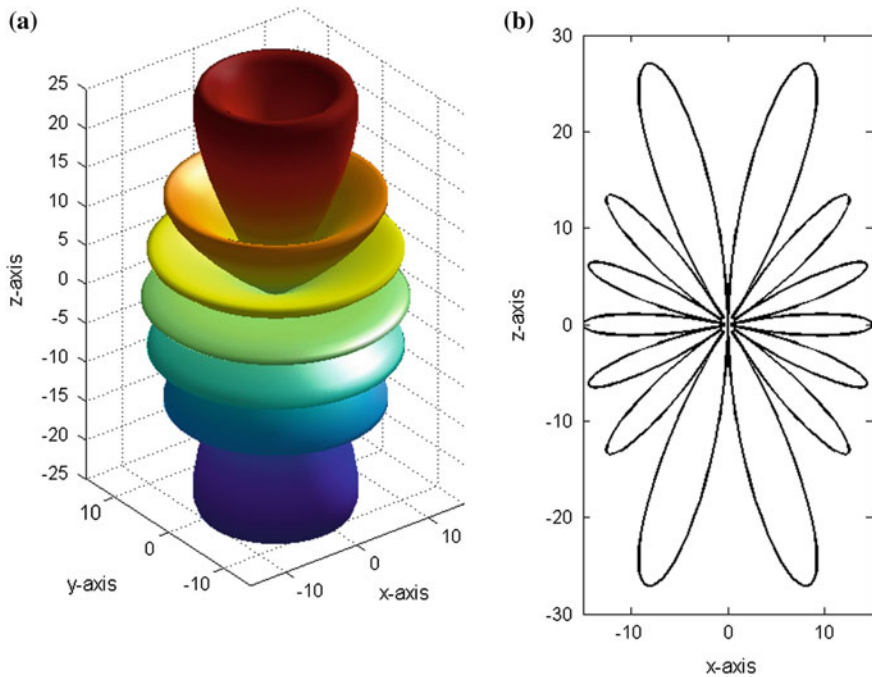


Fig. 2 **a** Three dimensional representation of $|Y_{7,4}^0|$ and **b** its polar diagram of $|Y_{7,4}^0| = \{199.48 \cos^3 2\theta + 164.79 \cos^2 2\theta - 48.36 \cos 2\theta - 28.67\} \sin^{1.4} \theta$

dominant effect of the presence of Poschl-Teller potential is coming from the $\sec^2 \theta$ term. The un-normalized angular wave functions illustrated in Figs. 1, 2 and 3 is in agreement with the result calculated using NU method [31]. By putting the new value of the orbital quantum number expressed in the energy of Eckart potential combined with trigonometric Poschl-Teller non-central potential is rewritten as

$$E_{n_r} = -\frac{\hbar^2}{2M} \left\{ \frac{V_1^2}{a^2 \left[\sqrt{V_o + (\sqrt{\kappa(\kappa-1)+m^2}) + \eta + 2n_l + \frac{1}{2}} \right]^2 + n_r + \frac{1}{2}} + \frac{\left[\sqrt{V_o + (\sqrt{\kappa(\kappa-1)+m^2}) + \eta + 2n_l + \frac{1}{2}} \right]^2}{4a^2} - \frac{l(l+1)d_0}{4a^2} \right\} \quad (69)$$

where n_l is a new polar quantum number and its values are non-negative integer, while n_r is radial quantum number and is nonnegative integer. From (69) we can calculate the energy for special case, for Eckart potential, we set $\kappa = \eta = m = n_l = 0$, therefore the energy spectrum of Eckart potential is

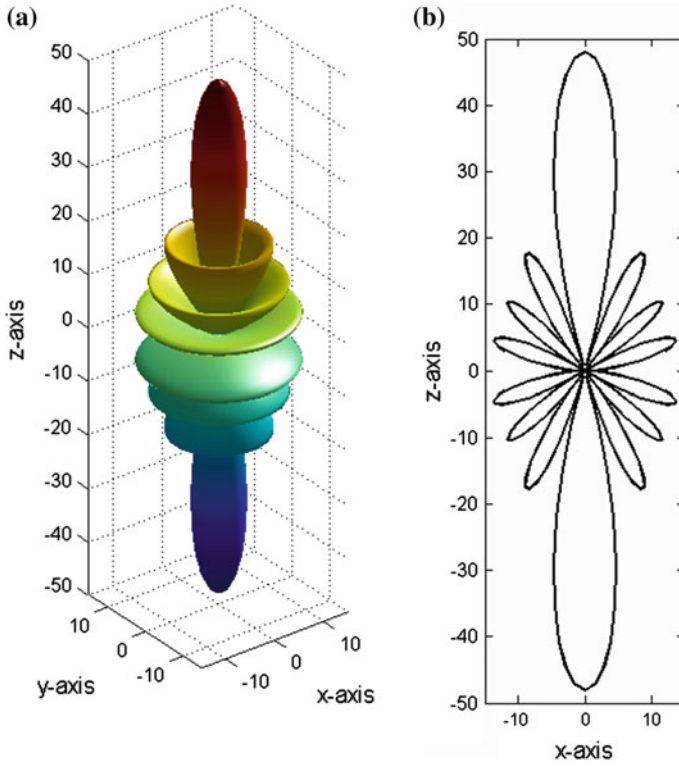


Fig. 3 **a** Three dimensional representation of $|Y_8^0|$ and **b** its polar diagram of $|Y_8^0| = \{268.13 \cos^3 2\theta - 160.88 \cos^2 2\theta - 86.63 \cos 2\theta + 27.38\} \cos^2 \theta$

$$E_{n_r} = -\frac{\hbar^2}{2M} \left\{ \frac{V_1^2}{a^2 \left[\sqrt{V_0 + \left(\frac{1}{2}\right)^2} + n_r + \frac{1}{2} \right]^2} + \frac{\left[\sqrt{V_0 + \left(\frac{1}{2}\right)^2} + n_r + \frac{1}{2} \right]^2}{4a^2} - \frac{l(l+1)d_0}{4a^2} \right\} \quad (70)$$

The total un-normalized wave function of the system obtained from (39) and (63a), (63b) is given as

$$\begin{aligned} \psi(r, \theta, \varphi) = & (1 - \coth(r/2a))^{\frac{\beta_{nr} - i\alpha_{nr}}{2} - \frac{1}{4}} (1 + \coth(r/2a))^{\frac{\beta_{nr} + i\alpha_{nr}}{2} - \frac{1}{4}} R_{n_r}^{(-\beta_{nr}, -\alpha_{nr})}(-i \coth(r/2a)) \\ & \times \frac{1}{\sqrt{\sin \theta}} (1 - \cos 2\theta)^{\frac{\sqrt{F} + \frac{1}{4}}{2}} (1 + \cos 2\theta)^{\frac{\sqrt{G} + \frac{1}{4}}{2}} R_{n_l}^{(-\beta + \frac{1}{2}, -\alpha)}(-i \cos 2\theta) e^{im\varphi} \end{aligned} \quad (71)$$

The wave function of the system in (68) reduces to the wave function of three dimensional Eckart potential by the absent of Poschl-Teller potential.

3.2 The Solution of Polar Schrodinger Equation for 3D Trigonometric Rosen Morse Non-central Potential

The polar part of the Schrodinger equation for trigonometric Rosen-Morse non-central potential is given as

$$\frac{\partial^2 P(\theta)}{\partial \theta^2} + \cot \theta \frac{\partial P(\theta)}{\partial \theta} - \left(\frac{v(v+1) + m^2}{\sin^2 \theta} - 2\mu \cot \theta \right) P(\theta) + l(l+1)P(\theta) = 0 \quad (72)$$

By setting $P = \frac{Q}{\sqrt{\sin \theta}}$ in (72) then (72) becomes

$$\frac{d^2 Q}{d\theta^2} - \left(\frac{v(v+1) + m^2 - \frac{1}{4}}{\sin^2 \theta} - 2\mu \cot \theta \right) Q + \left(l(l+1) + \frac{1}{4} \right) Q = 0 \quad (73)$$

To solve (73) we introduce a new variable $\cot \theta = s$ and (73) change into

$$(1+s^2) \frac{\partial^2 Q}{\partial s^2} + 2s \frac{\partial Q}{\partial s} - \left\{ \left(v(v+1) + m^2 - \frac{1}{4} \right) - \frac{2\mu s}{(1+s^2)} - \frac{l(l+1) + \frac{1}{4}}{(1+s^2)} \right\} Q = 0 \quad (74)$$

Equation (74) is solved in terms of Romanovski polynomial by setting

$$Q(\theta) = g_n(s) = (1+s^2)^{\frac{\beta}{2}} e^{-\frac{\alpha}{2} \tan^{-1} s} D_n^{(\beta, \alpha)}(s) \quad (75)$$

for $0 < s < \infty$

By inserting (75) into (74) we obtain

$$\begin{aligned} & (1+s^2) \frac{\partial^2 D}{\partial s^2} + \{2s(\beta+1) - \alpha\} \frac{\partial D}{\partial s} \\ & - \left\{ \frac{\beta s \alpha - \frac{\alpha}{4} + \beta^2 - 2\mu s - (l(l+1) + \frac{1}{4})}{1+s^2} + v(v+1) + m^2 - \frac{1}{4} - \beta^2 - \beta \right\} D \\ & = 0 \end{aligned} \quad (76)$$

Equation (76) reduces to differential equation satisfied by Romanovski polynomials

$$(1+s^2)\frac{\partial^2 D}{\partial s^2} + \{2s(\beta+1) - \alpha\}\frac{\partial D}{\partial s} - \left\{v(v+1) + m^2 - \frac{1}{4} - \beta^2 - \beta\right\}D = 0 \quad (77)$$

for

$$\beta s \alpha - \frac{\alpha^2}{4} + \beta^2 - 2\mu s - \left\{l(l+1) + \frac{1}{4}\right\} = 0 \quad (78)$$

By comparing (10) and (77) we obtain

$$\begin{aligned} (\beta+1) &= (-p+1); \alpha = -q \text{ and } v(v+1) + m^2 - \frac{1}{4} - \beta^2 - \beta \\ &= n(n-1) + 2n(1-p) \end{aligned} \quad (79)$$

From (78) we have

$$-\frac{\alpha^2}{4} + \beta^2 - \left\{l(l+1) + \frac{1}{4}\right\} = 0; \quad \beta\alpha - 2\mu = 0 \quad (80)$$

that give

$$\beta^2 = \frac{(l+\frac{1}{2})^2 \pm \sqrt{(l+\frac{1}{2})^4 + 4\mu^2}}{2} \quad (81)$$

and

$$\alpha^2 = \frac{8\mu^2}{(l+\frac{1}{2})^2 \pm \sqrt{(l+\frac{1}{2})^4 + 4\mu^2}} \quad (82)$$

Using (79) we obtain

$$p = -\beta; q = -\alpha \quad (83)$$

and

$$v(v+1) + m^2 = \left(\beta + n + \frac{1}{2}\right)^2 \quad (84)$$

Then from (84) we get

$$\beta = \sqrt{v(v+1) + m^2} - n - \frac{1}{2} \quad (85)$$

or

$$\beta = -\sqrt{v(v+1) + m^2} - n - \frac{1}{2} \quad (86)$$

By using (80) with (85) or with (86) we obtain

$$\alpha_n = \frac{2\mu}{\sqrt{v(v+1) + m^2} - n - \frac{1}{2}} \quad (87)$$

or

$$\alpha_n = -\frac{2\mu}{\sqrt{v(v+1) + m^2} + n + \frac{1}{2}} \quad (88)$$

From (85), (86), and (81) we obtain

$$\left(\sqrt{v(v+1) + m^2} - n - \frac{1}{2} \right)^2 = \frac{(l + \frac{1}{2})^2 \pm \sqrt{(l + \frac{1}{2})^4 + 4\mu^2}}{2} \quad (89)$$

$$\left(-\sqrt{v(v+1) + m^2} - n - \frac{1}{2} \right)^2 = \frac{(l + \frac{1}{2})^2 \pm \sqrt{(l + \frac{1}{2})^4 + 4\mu^2}}{2} \quad (90)$$

By imposing the condition that $p > 0$ then from (83) we have $-(\beta - \frac{1}{2}) > 0$ or $(\beta - \frac{1}{2}) < 0$, therefore the values of β and α that satisfy that condition are expressed in (86) and (88). The value of l that satisfies the system is obtained from (90), that is

$$l = l' = \sqrt{\left(\sqrt{v(v+1) + m^2} + n + \frac{1}{2} \right)^2 - \frac{\mu^2}{\left(\sqrt{v(v+1) + m^2} + n + \frac{1}{2} \right)^2}} - \frac{1}{2} \quad (91)$$

The weight function obtained from (7), (83), (86) and (88) is given as

$$w^{(-\beta, -\alpha)} = (1 + s^2)^{\beta_{n_l}} e^{-\alpha_{n_l} \tan^{-1} s} \quad (92)$$

Using (13) and (92) we obtain the Romanovski polynomials given as

$$R_n^{(-\beta, -\alpha)}(s) = \frac{1}{(1 + s^2)^{\beta_{n_l}} e^{-\alpha_{n_l} \tan^{-1} s}} \frac{d^{n_l}}{ds^{n_l}} \left\{ (1 + s^2)^{\beta_{n_l} + n_l} e^{-\alpha_{n_l} \tan^{-1} s} \right\} \quad (93)$$

The polar wave function obtained from (11) and (93) is given as

$$Q_{n_l}(\theta = \cot^{-1} s) = g_{n_l}(s) = (1 + s^2)^{\frac{\beta_{n_l}}{2}} e^{\frac{-\alpha_{n_l}}{2} \tan^{-1} s} R_{n_l}^{(-\beta, -\alpha)}(s) \quad (94)$$

The polar wave function obtained from (94) is given as

$$P_l^m(\theta) = \frac{Q_{n_l}(\theta)}{\sqrt{\sin \theta}} = (1 + s^2)^{\frac{\beta_{n_l}}{2} + \frac{1}{4}} e^{\frac{-\alpha_{n_l}}{2} \tan^{-1} s} R_{n_l}^{(-\beta, -\alpha)}(s) \quad (95)$$

Due to the condition that β_n and α_n are n-dependence, thus the Romanovski polynomials is infinity [16, 17] and the orthogonality of polynomials is not produced from orthogonality integral of the wave function, that is

$$\int \frac{Q_{n_l}}{\sqrt{\sin \theta}}(\theta) \frac{Q_{n_l'}}{\sqrt{\sin \theta}}(\theta) \sin \theta d\theta = \delta_{n_l n_l'} \neq \int (1 + s^2)^{\beta_{n_l}} e^{-\alpha_{n_l} \tan^{-1} s} R_{n_l}^{(-\beta, -\alpha)}(s) R_{n_l'}^{(-\beta, -\alpha)}(s) ds \quad (96)$$

Construction of Romanovski polynomial

The first four Romanovski polynomial are constructed using (93) are

$$R_0^{(-\beta + \frac{1}{2}, -\alpha)}(s) = 1 \quad (97)$$

$$R_1^{(-\beta + \frac{1}{2}, -\alpha)}(s) = 2(\beta_1 + 1)s - \alpha_1 \quad (98)$$

$$R_2^{(-\beta + \frac{1}{2}, -\alpha)}(s) = \{4(\beta_2 + 2)(\beta_2 + 1.5)s^2 - 4\alpha_2(\beta_2 + 1.5)s + \alpha_2^2 + 2(\beta_2 + 2)\} \quad (99)$$

$$R_3^{(-\beta + \frac{1}{2}, -\alpha)}(s) = 8(\beta_3 + 3)(\beta_3 + 2.5)(\beta_3 + 2)s^3 - 12\alpha_3(\beta_3 + 2.5)s^2 + 6\{\alpha_3^2 + 2(\beta_3 + 3)\}(\beta_3 + 2)s - \alpha_3\{\alpha_3^2 + 6\beta_3 + 16\} \quad (100)$$

The Romanovski polynomials expressed in (97–100) can be constructed manually or using computer programming with Mat Lab software. If Rosen-Morse noncentral potential is absent then

$$\beta_{n_l} = -m - n_l; \alpha_n = 0; l = n_l + m; \quad (101)$$

and the polar wave functions reduce to associated Legendre polynomials. The polar wave functions for $n_l = 3$ with different values of v and μ are shown in Table 1.

The third degree of Romanovski polynomials and the corresponding polar wave functions for $n_l = 3$, $m = 1$, $v = 0.2$ and $\mu = 0.2$ are calculated using (93), (95), (100) are listed in Table 1. The angular wave functions, $Y_{l'}^{m'}$, are obtained by multiplying the polar wave function listed in the last column at Table 1. with the

Table 1 The Romanovski polynomials and its corresponding un-normalized polar wave functions for Rosen-Morse potential for $n_l = 3$ and $m = 3$, with $\nu' + \frac{1}{2} = \sqrt{\nu(\nu + 1) + m^2} = m'$

No	n_l	m	ν	μ	$\nu' + \frac{1}{2}$	$l = l'$	$-\beta$	$R_{n_l}^{(-\beta, -\lambda)}(s)$	$P_{n_l m \nu \mu} = P_{\ell}^m(\theta)$
1	3	0	0	0	0	3	3.5	$-6 \cot^3 \theta + 9 \cot \theta$	$(\sin \theta)^3 (-6 \cot^3 \theta + 9 \cot \theta)$
2	3	1	0	0	1	4	4.5	$-60 \cot^5 \theta + 45 \cot^3 \theta$	$(\sin \theta)^4 (-60 \cot^5 \theta + 45 \cot^3 \theta)$
3	3	1	2	0	2.65	5.65	6.15	$-381.17 \cot^5 \theta + 156.87 \cot^3 \theta$	$(\sin \theta)^{5.65} \left(-381.17 \cot^5 \theta + 156.87 \cot^3 \theta \right)$
4	3	1	0	2	1	3.98	4.5	$-60 \cot^3 \theta + 54 \cot^2 \theta + 32.85 \cot \theta - 9.17$	$2(\sin \theta)^4 e^{0.45 \tan^{-1} \cot \theta + 4 \theta} \left(\frac{-60 \cos^3 \theta \sin \theta + 54 \cos^2 \theta \sin^2 \theta}{\sin^4 \theta} + 32.85 \cos \theta \sin^3 \theta - 9.17 \right)$
5	3	1	2	2	2.65	5.64	6.15	$-381.71 \cot^5 \theta + 118.15 \cot^3 \theta + 146.35 \cot \theta - 17.21$	$(\sin \theta)^{5.65} e^{0.33 \tan^{-1} \cot \theta} \left(\frac{-381.71 \cot^5 \theta + 118.15 \cot^3 \theta}{\sin^5 \theta} + 146.35 \cot \theta - 17.21 \right)$

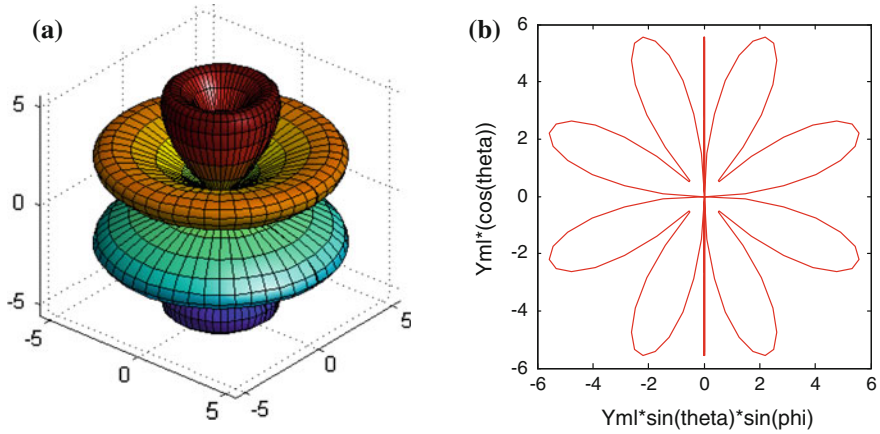


Fig. 4 **a** Three dimensional polar representation of absolute value of eigen function, $n_l = 3$, $m = 1$, $v = 0$ and $\mu = 0$ $Y_4^1 = 15(-4 \cos^3 \theta \sin \theta + 3 \sin^3 \theta \cos \theta)$. **b** Polar diagram of angular wave function $Y_4^1 = 15(-4 \cos^3 \theta \sin \theta + 3 \sin^3 \theta \cos \theta)$

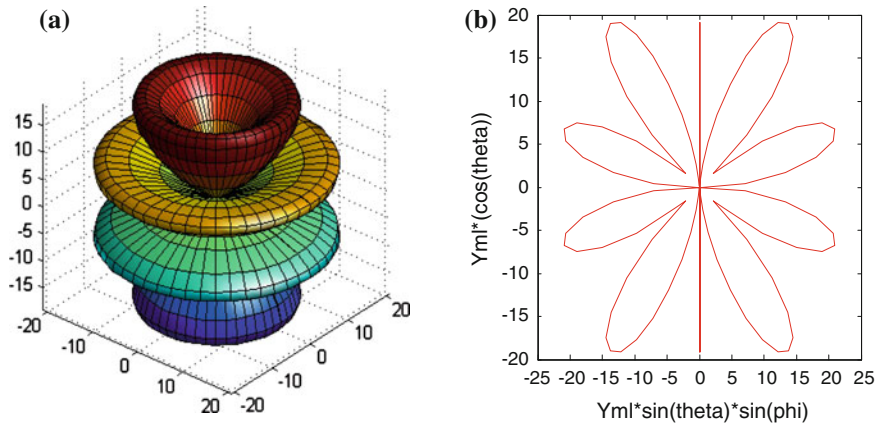


Fig. 5 **a** Three dimensional polar representation of absolute value of eigenfunction, $n_l = 3$, $m = 1$, $v = 2$ $\mu = 0$ $|Y_{5.65}^{2.65}| = (\sin \theta)^{5.65}(-381.17 \cot^3 \theta + 156.87 \cot \theta)$. **b** Polar diagram of angular wavefunction $|Y_{5.65}^{2.65}| = (\sin \theta)^{5.65} \times (-381.17 \cot^3 \theta + 156.87 \cot \theta)$

azimuthal wave function in (28). The polar diagram of orbital angular momentum eigen function, Y_l^m and three dimensional polar representation of the absolute value of the angular wave functions listed in Table 1 for all $m = l$ are graphed using Math Lab software shown in Figs. 4, 5, 6 and 7. By comparing Fig. (6a) and (7a) it can be

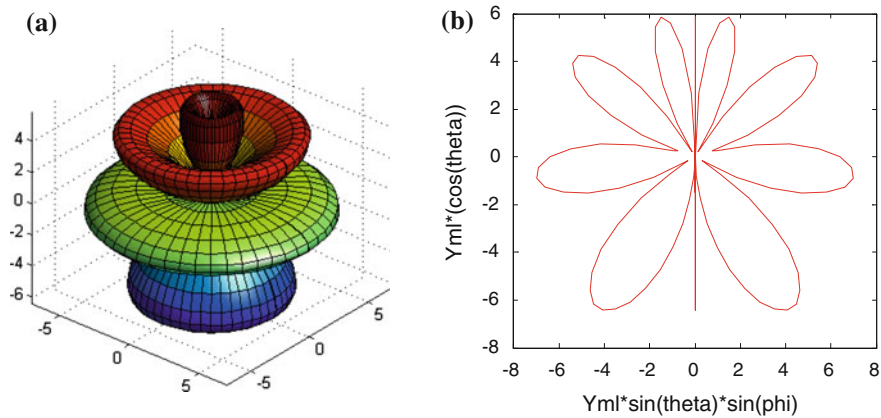


Fig. 6 **a** Three dimensional polar representation of absolute value of eigenfunction $n_l = 3, m = 1,$

$v = 0$ and $\mu = 2$ $|Y_{3,98}^1| = 2(\sin \theta)^4 e^{0.45 \tan^{-1} \cot \theta} \begin{pmatrix} -60 \cos^3 \theta \sin \theta \\ +54 \cos^2 \theta \sin^2 \theta \\ +32.85 \cos \theta \sin^3 \theta \\ -9.17 \end{pmatrix} \sin^4 \theta$. **b** Polar diagram

of angular wavefunction $|Y_{3,98}^1| = 2(\sin \theta)^4 e^{0.45 \tan^{-1} \cot \theta} \begin{pmatrix} -60 \cos^3 \theta \sin \theta \\ +54 \cos^2 \theta \sin^2 \theta \\ +32.85 \cos \theta \sin^3 \theta \\ -9.17 \end{pmatrix} \sin^4 \theta$

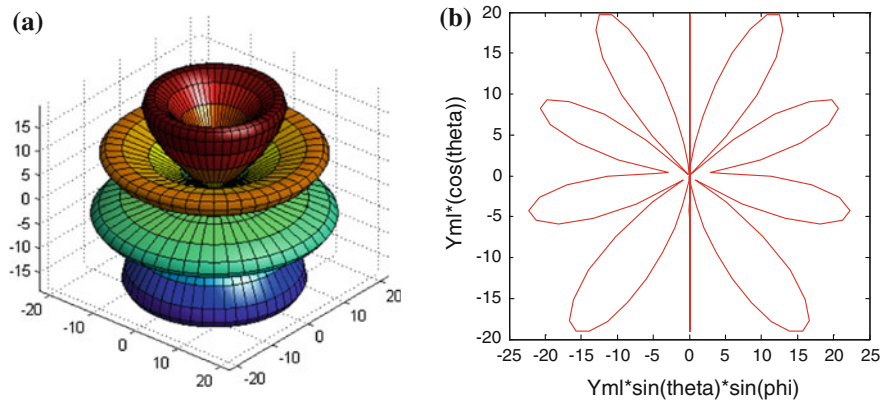


Fig. 7 **a** Three dimensional polar representation of absolute value of eigenfunction $n_l = 3,$

$m = 1, v = 2$ and $\mu = 2$ $Y_{5,64}^{2,65} = (\sin \theta)^{5.65} e^{0.33 \tan^{-1} \cot \theta} (-381.71 \cot^3 \theta + 118.15 \cot^2 \theta + 146.35 \cot \theta - 17.21)$. **b** Polar diagram of angular wavefunction $Y_{5,64}^{2,65} = (\sin \theta)^{5.65} e^{0.33 \tan^{-1} \cot \theta} (-381.71 \cot^3 \theta + 118.15 \cot^2 \theta + 146.35 \cot \theta - 17.21)$

shown that the effect of $\cot \theta$ is larger for lower level polar wave function. From Figs. (4a) and (5a) can be seen that $\csc^2 \theta$ causes the increase of the absolute value of the angular wave function, while $\cot \theta$ term causes the decrease of the absolute value of the angular wave function in the interval of $0 < \theta < \frac{\pi}{2}$ but causes the increase in the $\frac{\pi}{2} < \theta < \pi$ as shown in Figs. (6a) and (7a).

The total u-normalized wave function for the n level is given as

$$\psi(r, \theta, \varphi) = e^{im\varphi} \left\{ C_{n_r} r^{-(n_r+l'+\frac{1}{2})} e^{-\frac{\gamma r^2}{2}} \frac{d^{n_r}}{dr^{n_r}} (r^{2l'+1+2n_r} e^{-\gamma r^2}) \right\} \left[(\csc^2 \theta)^{-\frac{\beta_{n_l}}{2} + \frac{1}{4}} e^{\frac{\alpha_{n_l}}{2} \tan^{-1}(\cot \theta)} \frac{d^{n_l}}{d \cot \theta^{n_l}} \left\{ (\csc^2 \theta)^{\beta_{n_l} + n_l} e^{-\alpha_{n_l} \tan^{-1}(\cot \theta)} \right\} \right] \quad (102)$$

By the absent of Rosen-Morse potential the wave function in (102) reduces to the three dimensional spherical harmonics oscillator wave function.

The three dimensional Schrodinger equation for separable shape-invariant non-central potentials are solved using variable separation method. The 3D Schrodinger equation is separated into three one dimensional equations, radial and polar equations are solved using Romanovski polynomial while the azimuthal part is simple differential equation. The generalized hypergeometric type equation and so the Romanovski differential equations fall into two groups, first group, such as (26) and (74) have the same form, and so the (28) and (77). while the second group, such as (48) which is different to the form of (26) and (74) and so (50b) is different to the form of (28) and (77). Therefore by recognizing the form of the generalized hypergeometric type equation we can determine the form of differential equation of Romanovski polynomials. Even for complex variable, Romanovski polynomials method working very well in determining the energy spectra of the system but there is a limitation in producing the wave function, as in Poschl-Teller potential.

Acknowledgment This work is partially supported by Hibah Pascasarjana Sebelas Maret University grant No. 2340/UN27.10/PG/2012

References

1. Adhikari R, Dutt R, Sukhatme UP (1988) Phys Rev A 38(1988):1986
2. Comptet A, Bandrauk AD, Campbell DK (1985) Phys Lett B 150:159
3. Goudarzi H, Vahidi V (2011) Adv Stud Theor Phys 5(10):469
4. Inomata A, Suparmi A, Kurth S (1991) In: Dodonov VV, Man'ko VI (eds) Proceeding of 18th international Colloquium on group theoretical methods in physics. Springer, Berlin, p 399
5. Khare A, Bhaduriar RK (1993) [arXiv:hep-th/9310104v1](https://arxiv.org/abs/hep-th/9310104)
6. Chun-Seng J, Ying Z, Xiang-Lin Z, Liang Tian S (2001) Commun Theor Phys 36(6):641–646
7. Sadeghi J, Pourhassan B (2008) EJTP 5(17):193
8. Infeld L, Hull TD (1951) Rev Mod Phys 23:21–68
9. Gendenhtein LE (1983) JETP Lett 38:356
10. Ikhdair SM (2012) ISRN Math Phys 2012(201525):20. doi:[10.5402/2012/201525](https://doi.org/10.5402/2012/201525)

11. Ikot AN, Akpabio LE, Obu JA (2011) JVR 6(1):1–13
12. Halberg AS (2011) Int J Maths Math Sci 1–9 doi:[10.1155/2011/358198](https://doi.org/10.1155/2011/358198)
13. Aktas M (2007) [arXiv: 0701063v1](https://arxiv.org/abs/0701063v1) [quant-ph], pp 1–14
14. Cari C, Suparmi A (2012) IOSR-JAP 2(3):13–23. doi:[10.9790/4861-0231323](https://doi.org/10.9790/4861-0231323)
15. Movahedi M, Hamzavi M (2001) Int J Phys Sci 6(4):891–896
16. Alvarez-Castillo DE, Compean CB, Kirchbach M (2011) [arXiv: 1105.1354v1](https://arxiv.org/abs/1105.1354v1) [quant-ph] pp 1–10
17. Alvarez-Castillo DE, Kirbach M (2007) Rev Mex Fis E 53(2):143–154
18. Suparmi A et al (2012) IOSR-JAP 2(2):43–51. doi:[10.9790/4861-0224351](https://doi.org/10.9790/4861-0224351)
19. Koepf W, Masjed-Jamei M (2006) Integr Transform Spec Funct 17:559
20. Yasuk F, Berkdemir C, Berkdemir R (2005) J Phys A Math Gen 38:6579
21. Nikiforov AF, Uvarov UB (1988) Special function in mathematical physics. Birkhausa, Basel
22. Routh EJ (1884) Proc Lond Math Soc 16:245
23. Romanovski VI (1929) Compt Rend Acad Sci Paris 188:1023
24. Raposo A, Weber HJ, Alvarez-Castillo DE, Kirchbach M (2007) C Eur Phys J 5:253
25. Xian-Quan H et al (2010) Commun Theor Phys 53(2):242–246
26. Salehi H (2011) App Math 2:999–1004. doi:[10.4236/am.2011.28138](https://doi.org/10.4236/am.2011.28138)
27. Aktas M (2007) [arXiv:quant-ph/0701063v1](https://arxiv.org/abs/quant-ph/0701063v1)
28. Greene RI, Aldrich C (1976) Phys Rev A 14:2363
29. Jia CS, Chen T, Cui LG (2009) Phys Lett A 373:1621
30. Xu Y, He S, Jia CS (2010) Phys Scripta 81:045001
31. Qiang WC, Dong SH (2007) Phys Lett A 368(1–2):13–17
32. Ikhdair SM, Sever R (2008) Ann Phys 17(11):897–910

Interpretation and Prediction of Diffusion Activation Energy from Thermodynamic Model

P. Wu, Y. Zeng and H.M. Jin

Abstract Although self diffusion activation energy of metal is known to be strongly correlated to its melting point and heat of formation, prediction of diffusion activation energy remains a challenge and somehow an art. On the basis of the well developed quasi-chemical model, we demonstrate a general frame work to estimate surface/interface diffusion activation energy using the concentration and formation energy of the interface chemical bonds. Furthermore, this frame work reproduces all available diffusion activation energy data of Si on metals. We not only provide a practical method in predicting diffusion activation energy but also offer new physical insights about the interplay between chemical kinetics and thermodynamics.

1 Introduction

Diffusion activation energy data are frequently required in microstructure and phase transition research (e.g., in the glass formation process) where complicated chemical and structure parameters are involved [1–3]. However, diffusion activation energy data are scarce and the common techniques of experiment and computer simulation to derive these data are either not practical or time consuming due to wide range of changes of systems in terms of chemistry and microstructures [4, 5]. Therefore correlation methods were widely employed. Beke studied the dependence of diffusion to temperature and found that the diffusion coefficients usually obey the Arrhenius law [6]. Brown and Ashby showed the rate of diffusion at the melting

P. Wu (✉)

Singapore University of Technology and Design,
Singapore 138682, Singapore
e-mail: wuping@sutd.edu.sg

Y. Zeng · H.M. Jin

Institute of High Performance Computing, Agency for Science,
Technology and Research, 1 Fusionopolis Way, #16-16 Connexis,
Singapore 138632, Singapore

point is roughly a constant for a given crystal structure [7]. In addition, melting points show strong correlation to enthalpy of formation of alloys or compounds [8]. Any guidelines derived from these different approximation methods that leading to the prediction of diffusion activation energy will advance the diffusion research in general, although the nature of chemical bond formation is largely ignored in most of the previous correlation type of models. Recent advance in chemical thermodynamic calculations offers ready descriptions of detailed solution chemistry for systems of complicated chemistry as well as for multi-component and multiphase microstructures [9]. A direct link between the thermodynamics and diffusion activation will not only generate practical prediction method but also improve current theoretical understanding of chemical kinetics in general.

We demonstrated previously [10] that interface chemistry of metal and semiconductor can be qualified by quasi-chemical model (a solution chemistry model) [11, 12]; both the heat of formation of metal silicides and the eutectic temperatures of system consisted of the metal and semiconductors can be directly presented by the quasi-chemical model. Therefore, rather than a simple data correlation, our approach approximates the chemical bonding process, on the basis of the well developed quasi-chemical model, by calculating concentrations of different type of bonds based on the bond formation energy and solution chemical concentrations. This way, we maintain the rough physics of interface while developing correlation models.

In chemical catalysis research, factors that control chemical reaction paths and the corresponding potential energy surfaces [13, 14] are well studied and a simple linear correlation between the transition-state (TS) energy and the final-state (FS) energy of the respective elementary steps, namely, the Brønsted-Evans-Polanyi (BEP) relationship has been elucidated and applied in the analysis of surface elementary reactions. The BEP equation directly relates the change in activation energy of the reaction (a kinetic parameter) to the corresponding change of the reaction energy (a thermodynamic parameter). Liu and Hu [15] advanced the BEP method by using first principles calculations to systematically study a number of typical catalytic reactions on transition metal surfaces, and postulated that the reaction barrier can be decomposed using bond energies of all the associated reactant-surfaces and reactant-reactant pairs (or bonds). Their work directly linked the reaction barrier to the energetic of all sorts of involved chemical bonds and found that the dissociation reaction barrier is determined by the chemisorption energy at the final reaction state, in consistent with the BEP relationship.

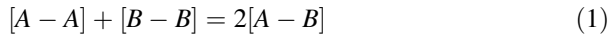
These studies suggest that the bond energies of reactants and products may play a critical role in determining the reaction barrier and diffusion activation energy. In statistical thermodynamics, all equilibrium properties are derived by using partition functions of all energetic micro-configurations, therefore, a transition state (on an energy path way for diffusion) has already been counted and discriminated by the partition functions. We aim to pick up the micro-configuration(s) that corresponding to the transition state by reverse engineering the thermodynamic solution. Therefore, the quasi-chemical (QC) model which describes the energy and concentration of bonds of all stable and metastable micro-configurations is applied in

this work. We first demonstrate such a reverse engineering technique (on QC) using the heat of formation and eutectic temperature data, and then diffusion activation energy. The resultant formulae can be used for prediction of activation energy merely from electro-negativities of the constituent elements.

2 Quasi-chemical Calculations

In this section, the relationship of reaction bond energy and mole fraction of the reactant pairs of a binary system is described by the quasi-chemical model [11, 12]. The numerical results of the relationship associated with various initial component compositions are presented.

An A-B binary system contains three types of nearest-neighbor pairs, namely, A-A, B-B, and A-B. The relative amounts of the three types of pairs are determined by the energy change associated with the formation of two A-B pairs from one A-A and one B-B pair, as described in reaction (1).



Let X_{AA} , X_{BB} and X_{AB} be the mole fractions for A-A, B-B and A-B pairs, and X_A and X_B are the initial mole fractions of components A and B, respectively. When the two components are mixed, A-B pairs are formed at the expense of A-A and B-B pairs. It follows from the mass balance that:

$$2X_A = 2X_{AA} + X_{AB} \quad (2)$$

$$2X_B = 2X_{BB} + X_{AB} \quad (3)$$

The molar enthalpy change of reaction (1) is denoted by ω , the net pair bond energy. The enthalpy of mixing, ΔH , is then given as following:

$$\Delta H = \left(\frac{X_{AB}}{2} \right) \omega \quad (4)$$

A “quasi-chemical equilibrium constant” for reaction (1) is obtained as

$$\frac{X_{AB}^2}{X_{AA}X_{BB}} = 4e^{\frac{-\omega}{RT}} \quad (5)$$

Substituting (2–3) into (5) yields (6) where it indicates that the mole fraction A-B bond, X_{AB} , is the function of the net pair bond energy ω , the initial component mole fraction X_A and the temperature T . R denotes the gas constant.

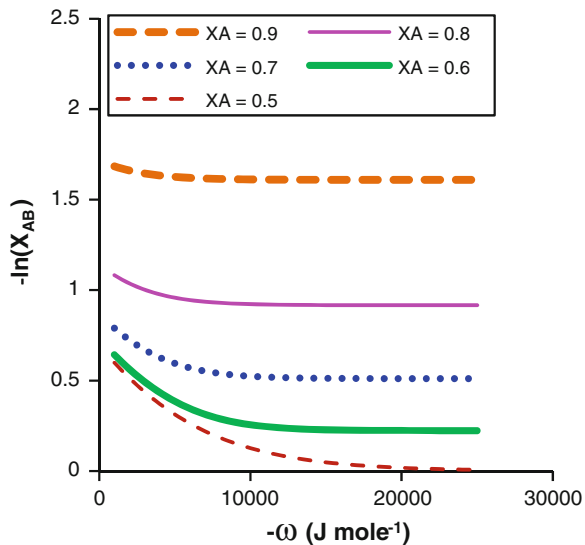
$$\frac{X_{AB}^2}{(2X_A - X_{AB})(2 - 2X_A - X_{AB})} = e^{\frac{-\omega}{RT}} \quad (6)$$

The analytical solution of X_{AB} cannot be obtained from (6). Thus we perform the numerical study by plotting $(-\omega)$ versus logarithm X_{AB} given $T = 300$ K and $\omega = -25 \sim 0$ kJ mole⁻¹. The results are shown in Fig. 1 with the initial component mole fraction $X_A = 0.5, 0.6, 0.7, 0.8$ and 0.9 , respectively. It is noted that due to the symmetry in the composition of component A and B, the results of $X_A = 0.1$ and $X_A = 0.9$ are equal. Similarly, the result of $X_A = 0.2$ is same as that of $X_A = 0.8$, and so forth.

3 Results

On the basis of quasi-chemical model, we examine four published data sets of formation heats, eutectic temperatures, and diffusion activation energies. Since the ω values for these systems are not available, we use the elemental electro-negativity (χ) to replace ω [10]. In particular, the squared difference of electro-negativities of two constituent elements $(\Delta\chi)^2$ is used. It is noted that $(\Delta\chi)^2$ is proportional to $-\omega'$ as derived from the well-known semi-empirical Miedema's model [16]. In this work, the Martynov-Batsanov electro-negativities (eV^{1/2}) [17] are used. In the study of the following four systems, $(\Delta\chi)^2$ is chosen as the abscissa, in consistent with the result of quasi-chemical model (Fig. 1).

Fig. 1 Dependence of net pair bond energy ω and the mole fraction of A-B bond X_{AB}



We first demonstrate the reverse engineering approach using the formation heats of silicides formed on the interfaces of transition metals and *n*-type silicon [10]. In order to obtain comparable result, the ordinate is formulated as the logarithmic ($\Delta H_f/\Delta\chi^2$) where ΔH_f is the formation heat. Since the temperature is the same for this set of data which leads to a same constant for each data point, the temperature is not taken into account. The resulting trend of $(\Delta\chi)^2$ versus $(\Delta H_f/\Delta\chi^2)$ is shown in Fig. 2. It is also found that the dependences of X_{AB} and ω generated from QC model with a range of initial composition X_A as shown Fig. 1 can be re-scaled to well match the trend of the experimental data. Beyond this composition range, it is impossible to reproduce the trend. As an illustration, the re-scaled dependence with $X_A = 0.6$ (bold line) is plotted in Fig. 2. In comparison with the results of quasi-chemical model (Fig. 1), $\ln(\Delta H_f/\Delta\chi^2)$ is equivalent to $-\ln(X_{AB})$, which implies a strong correlation between formation heat and the A-B bond properties (concentration and energy) at micro-configurations about $X_A = 0.6$.

The second system deals with the eutectic temperature of binary systems of transition metal silicides [10], since it is widely accepted for a strong correlation between diffusion and temperature [7]. In the same fashion of Fig. 2, the ordinate is formulated as the logarithmic ($T_{eu}/\Delta\chi^2$) where T_{eu} is the eutectic temperature. The result is shown in Fig. 3. Similarly as in Fig. 2, the QC dependence with $X_A = 0.6$ (bold line) is re-scaled to fit this set of experimental data. Furthermore, since $\ln(T_{eu}/\Delta\chi^2)$ is equivalent to $-\ln(X_{AB})$, it implies a strong correlation between the eutectic temperature and the A-B bond properties at micro-configurations about $X_A = 0.6$.

Inspired by the above success, we further study the diffusion of Si into metals by the reverse engineering technique. The diffusion activation energies and the corresponding average temperatures are extracted from [18] and summarized in Table 1. The ordinate is chosen as $\ln(\frac{Q}{RT(\Delta\chi)^2})$ where Q is the activation energy, R the gas constant, and T the average temperature. The resulting trend of $(\Delta\chi)^2$ versus $\ln(\frac{Q}{RT(\Delta\chi)^2})$ is shown in Fig. 4. Similarly, the QC dependence with $X_A = 0.6$ (bold line) is re-scaled to match this set of experimental data. Compared to Fig. 1, $\ln(\frac{Q}{RT(\Delta\chi)^2})$ is equivalent to $-\ln(X_{AB})$, implying strong correlation between the activation energy and the A-B bonding properties.

Fig. 2 Dependence of $\ln(\Delta H_f/\Delta\chi^2)$ and $(\Delta\chi)^2$ for the formation heats of transition-metal and silicon

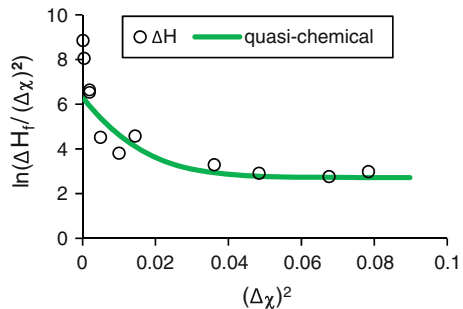


Fig. 3 Dependence of $\ln(T_{eu}/\Delta\chi^2)$ and $(\Delta\chi)^2$ for the eutectic temperatures of transition-metal silicides

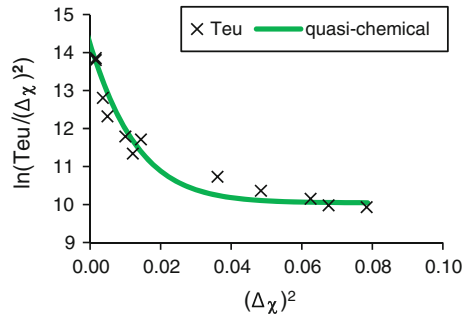
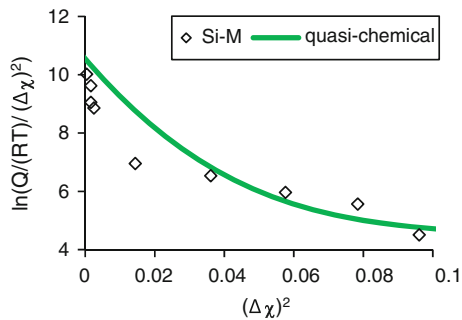


Table 1 Diffusion activation energies of Si into metals and the electro-negativities of metals

Metal	Average temp. °C	Activation energy cal mole ⁻¹	Metal electro-negativity eV ^{1/2}
Ti	1,050	39,700	1.86
Zr	1,100	55,750	1.7
V	1,100	61,200	2.22
Nb	1,100	48,000	2.03
Ta	1,000	34,600	1.94
Cr	1,000	22,760	2
Mo	1,350	78,000	1.94
W	1,000	63,000	1.79
Fe	1,000	22,030	1.67

Fig. 4 Dependence of $\ln(Q/RT(\Delta\chi^2))$ and $(\Delta\chi)^2$ for the diffusion activation energies of Si diffusion into metals

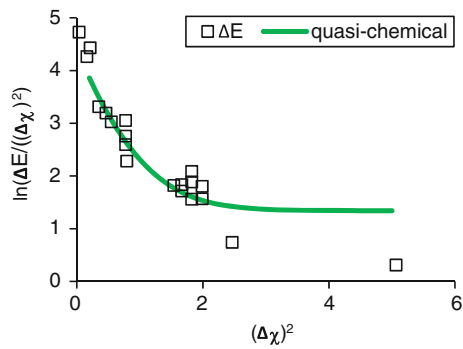


Finally, we study the reaction barrier (activation energy, ΔE) in the diffusion of O, S, N on (111) surfaces of various metals (Ni, Re, Ag, Pd, Pt, Ru, Cu) [13]. Using the data listed in Table 2, we calculate $\ln(\Delta E/\Delta\chi^2)$ and plot it against $(\Delta\chi)^2$ in Fig. 5. The term (RT) is omitted as the temperatures are the same for this set of data. Again, the QC dependence with $X_A = 0.6$ (bold line) is re-scaled to reproduce the experiment trend. Similarly, a strong correlation between the activation energy and the microscopic A-B bonding is demonstrated.

Table 2 Energies of initial and transition states in the diffusion of O, S, N on (111) surfaces of various metals (data read from [13]), and electro-negativities of each system

System	Energy, (eV)		Electro-negativity, (eV ^{1/2})	
	Initial state	Transition state	Non-metal	Metal
C/Ru	−6.69	−6.12	2.37	1.97
C/Pt	−6.27	−5.38	2.37	1.91
S/Re	−5.69	−5.21	2.65	2.06
S/Ru	−5.64	−5.08	2.65	1.97
N/Ru+	−5.55	−4.73	2.85	1.97
O/Ru+	−5.40	−4.80	3.32	1.97
O/Ru	−5.07	−4.34	3.32	1.97
N/Ru	−5.06	−4.54	2.85	1.97
S/Ni	−5.02	−4.64	2.65	1.76
O/Ru−	−4.79	−4.35	3.32	1.97
S/Pt	−4.74	−4.18	2.65	1.91
N/Ru−	−4.59	−3.98	2.85	1.97
C/Cu	−4.25	−3.73	2.37	1.08
S/Cu	−4.14	−3.89	2.65	1.08
O/Pt+	−4.08	−3.47	3.32	1.91
C/Cu−	−4.02	−3.56	2.37	1.08
O/Pt	−3.63	−3.16	3.32	1.91
O/Pd	−3.54	−3.06	3.32	2.08
O/Ag	−3.06	−2.71	3.32	1.07
H/Pt	−2.58	−2.37	2.1	1.91

Fig. 5 Dependence of $\ln(\Delta E/\Delta\chi^2)$ and $(\Delta\chi)^2$ for the reaction barriers of diffusion of O, S, N on (111) surfaces of various metals



4 Concluding Remarks

We demonstrated, on the basis of quasi-chemical model, a general frame work to formulate diffusion activation energy from solution thermodynamics. Although in this study our approach reproduced diffusion activation energy for relative simple

systems, in principle similar calculations may be carried out for multi-component and multiphase systems because the needed bond fraction and bond energy data can be derived from quasi-chemical model, which complements the popular computing intensive density functional theory (DFT) approaches [19–22]. The method offers opportunities to further understand the interplay between chemical kinetics and thermodynamics in general, and was applied recently in chemical gas sensor development [23].

References

1. Gill EK, Morrison JA (1963) Kinetics of solids. *Annu Rev Phy Chem* 14:205–228
2. Kang M, Zhang MX, Liu F, Zhu M (2009) Kinetics and morphology of isothermal transformations at intermediate temperature in 15CrMnMoV steel. *Mater Trans* 50(1):123–129
3. Wang ZY, Li YH, Adams JB (2000) Kinetic lattice Monte Carlo simulation of facet growth rate. *Surf Sci* 450(1–2):51–63
4. Rouxel T (2011) Thermodynamics of viscous flow and elasticity of glass forming liquids in the glass transition range. *J Chem Phys* 135(18):184501
5. Bletskan DI (2006) Glass formation in binary and ternary chalcogenide systems. *Chalcogenide Lett* 3(11):81–119
6. Beke D, Geszti T, Erdelyi G (1977) Connection between pre-exponential factor and activation enthalpy for self-diffusion and impurity diffusion. *Z Metallkde* 68(6):444–447
7. Brown AM, Ashby MF (1980) Correlations for diffusion constants. *Acta Metall* 28(8):1085–1101
8. Murti YVGS, Popson GA, Laskar AL (1990) Correlation of melting temperature and formation enthalpy of Schottky defects in ionic solids. *Phys Status Solidi A* 120(2): K133–K138
9. Hecht U, Granasy L, Pusztai T et al (2004) Multiphase solidification in multicomponent alloys. *Mat Sci Eng R* 46(1–2):1–49
10. Wu P, Zeng Y (2010) Quantifying the relationship between interface chemistry and metal electronegativity of metal-semiconductor interfaces. *J Mater Chem* 20(46):10345–10350
11. Wu P (1992) Optimization and calculation of thermodynamic properties and phase diagrams of multi-component oxide systems. Dissertation, University of Montreal
12. Wu P, Pelton AD (1992) Coupled thermodynamic phase-diagram assessment of the rare-earth-oxide aluminum-oxide binary-systems. *J Alloy Compd* 179:259–287
13. Nilekar AU, Greeley J, Mavrikakis M (2006) A simple rule of thumb for diffusion on transition-metal surfaces. *Angew Chem Int Ed* 45(42):7046–7049
14. van Santen RA, Neurock M, Shetty SG (2010) Reactivity theory of transition-metal surfaces: a Bronsted-Evans-Polanyi linear activation energy-free-energy analysis. *Chem Rev* 110(4):2005–2048
15. Liu ZP, Hu P (2001) General trends in CO dissociation on transition metal surfaces. *J Chem Phys* 114(19):8244–8247
16. Miedema AR (1992) Energy effects and charge-transfer in metal physics—modeling in real space. *Phys B* 182(1):1–17
17. Villars P, Hulliger F (1987) Structural stability domains for single-coordination intermetallic phases. *J Less-Common Met* 132(2):289–315
18. Samsonov GV, Vinitskii IM (1980) Handbook of refractory compounds (trans: Shaw K). IFI/Plenum, New York, p 220
19. Zhang J, Loh KP, Yang SW et al (2005) Exohedral doping of single-walled boron nitride nanotube by atomic chemisorption. *Appl Phys Lett* 87(24):243105

20. Lim FCH, Zhang J, Jin H et al (2013) A density functional theory study of CO oxidation on Pd-Ni alloy with sandwich structure. *Appl Cat A* 451:79–85
21. Zhang J, Cao XM, Hu P et al (2011) Density functional theory studies of ethanol decomposition on Rh(211). *J Phys Chem C* 115(45):22429–22437
22. Zhang J, Jin H, Sullivan MB et al (2009) Study of Pd-Au bimetallic catalysts for CO oxidation reaction by DFT calculations. *Phys Chem Chem Phys* 11(9):1441–1446
23. Choi SW, Katoch A, Sun GJ et al (2013) NO₂-sensing performance of SnO₂ microrods by functionalization of Ag nanoparticles. *J Mater Chem C* 1(16):2834–2841

Shear Strength of Single Lap Joint Aluminium-Thermoplastic Natural Rubber (Al-TPNR) Laminated Composite

M.Z. Muzakkar, S. Ahmad, M.A. Yarmo, A. Jalar and M. Bijarimi

Abstract In this work, we studied the effect of surface treatment of the aluminium surface and a coupling agent to improve adhesion between aluminium with organic polymer. Thermoplastic natural rubber (TPNR) matrix was prepared by melt blending of natural rubber (NR), liquid natural rubber (LNR) as a compatibilizer, linear low density polyethylene (LLDPE) and polyethylene grafted maleic anhydride (PE-g-MAH) as a coupling agent. The PEgMAH concentration used was varied from 0 to 20 %. In addition, the aluminium surface was pre-treated with 3-glycidoxypropyl trimethoxysilane (3-GPS) to enhance the mechanical properties of laminated composite. It was found that the shear strength of the single lap joint Al-TPNR laminated composite showing an increasing trend as a function of PE-g-MAH contents for the 3-GPS surface treated aluminium. Moreover, the FTIR-ATR, XPS and SEM analysis revealed that the strength improvement was associated to the chemical state of the compound involved.

M.Z. Muzakkar (✉)

Department of Chemistry, Faculty of Mathematics and Science,
Universitas Haluoleo, Kendari 93232, Indonesia
e-mail: Zakir1703@yahoo.com

S. Ahmad · A. Jalar

Faculty of Science and Technology, School of Applied Physics,
Universiti Kebangsaan Malaysia, 43000 Bangi, Selangor, Malaysia
e-mail: sahrim@ukm.my

A. Jalar

e-mail: azmn@ukm.my

M.A. Yarmo

Faculty of Science and Technology, School of Chemical Sciences
and Food Technology, Universiti Kebangsaan Malaysia,
43000 Bangi, Selangor, Malaysia
e-mail: ambar@ukm.my

M. Bijarimi

Faculty of Chemical and Natural Resources Engineering,
Universiti Malaysia Pahang, 26300 Kuantan, Malaysia
e-mail: bijarimi@ump.edu.my

1 Introduction

Recently, a lot of attention has been focused upon the development of polymer matrix composite with metals (sandwich/laminate composite). These composites are used in a wide range of applications such as aerospace, automotive, implants, electronic, packaging, defense, nuclear, medical, etc. As compared to homogenous metal sheets, the metal-polymer laminated composite have many unique properties in terms of low density and weight, high bending strength as well as good sound, vibration and damping characteristics [7, 24, 28, 37, 38].

The preparation of suitable metal-polymer laminates requires many practical considerations. It has been established that one of the most important considerations is the pretreatment process of metal surface either by mechanical, chemical, electrochemical and coupling agents. The coupling agent is well-known to enhance the molecular bonding between polymer matrix and inorganic particulates. It strengthens the interfacial adhesion when one end of the molecule is tethered to the reinforcement surface and the functionality at the other end reacts with the polymer phase [3–5, 8, 14, 27, 34, 39, 46]. Typically, the coupling agents are classified into organic, inorganic, and organic-inorganic groups. The organic agents include isocyanates, anhydrides, amides, imides, acrylates, chlorotriazines, epoxides, organic acids, monomers, polymers, and copolymers. On the other hand, the inorganic agents such as silicates can be found in the study of nanocomposite (Polymer-layered silicate nanocomposites). The other types include the organic-inorganic agents, i.e. silane and titanate [25].

It is a tedious and time-consuming process to select the appropriate coupling agents experimentally. Silane is recognized as an efficient coupling agent that can be used in the composites and adhesive formulations [30]. Silane coupling agents have a generic chemical structure $R_{(4-n)}-Si-(R'X)_n$ ($n = 1, 2$) where R is alkoxy, X represents an organofunctionality, and R' is an alkyl bridge connecting the silicon atom and the organofunctionality. The organofunctionality for every silanes is very specific. Most of them were used during the past decade by the researchers are amino, APS [44], mercapto, MRPS [2], glycidoxo, GPS [16, 41], vinyl, VTS [20], alkyl, HDS [1], or methacryloxy group, MPS [33].

Usually polar materials with good strength are selected as a polymer matrix. Functional groups like $-C=O$, $-OH$, $-COOH$, $-NH_2$, etc., make a polar polymer with inherent surface strength better than the non-polar polymers [8, 14, 38]. The non-polar polymers can be surface modified by surface oxidizing [10], chemical etching [19], surface grafting [31], flame treatment [13], electron beam [18], microwave irradiation [40], electrical plasma discharge [23] and glow discharge [22] etc. All of these processes aim to create polar groups at the polymer surface to increase the value of the surface-free energy and adhesive properties of polymeric materials. In recent years, many studies are dedicated to obtain new materials by blending between polyolefins with natural materials that is rich in functional groups like $-OH$, $-CO$, $-CHO$ and etc. For example, the use of starch [6], wood [11], rice straw [42] etc., and natural rubber [17, 21]. Blending between polyolefin as a thermoplastic polyolefin

(TPO) with natural rubber is known as thermoplastic elastomer (TPE) but it is commonly called as a thermoplastic natural rubber (TPNR).

Apart from modifying the polymer surface, the adhesion between polymeric material with the primary metal structures also requires some attention. It is well-known that the polyethylene exhibits insufficient adhesive bond strength due to low surface energy. This property drawback has been solved by chemical, thermal, mechanical and electrical treatments. However, the dry treatment method is more suitable for industrial applications. Previous works have shown the effect of the maleic anhydride coupling agent [8, 14, 28] by imparting polarity to a polyolefin surface that enables an improved adhesive bonding with aluminium and pretreated aluminium surface [3, 27, 34]. Chen et al. [7] prepared an aluminum-polypropylene (Al-PP) laminated composite by using PP-g-MAH as a coupling agent and 3-APS was used to treat the surface of the aluminum. da Silva et al. [12] fabricated steel-epoxy composite laminates and steel surfaces coated with an organic phosphate, while Pereira et al. [26] fabricated Al-epoxy composite laminates and then aluminium surface treated with sodium dichromate-sulfuric acid etch (CSA). Putman and Vaida [28] studied of mechanisms of interfacial adhesion in metal-polymer composites. They have produced a steel-polymer laminate composite from their studies. Cold-rolled low carbon steel (AISI/SAE1018) was used. Polymer laminates are hard TPU Pearlthane D11T75 merquinsa, soft TPU Pearlthane D11T85E merquinsa, nylene Custom Resins 401D, PP, and PPMA as a coupling agent. Steel surface was coated with an adhesive nano fullerene epoxy resin CNT epoxy. Sokolova et al. [35] produced a composite laminate consisting of high-quality austenitic stainless steel (316L) sheets with nominal thickness of 0.5 mm. The core consisted of a polyolefin foil being a mixture of polypropylene (PP) and polyethylene (PE) polymers, talc $[Mg_3Si_4O_{10}(OH)_2]$, rutile (TiO_2) and barite ($BaSO_4$). Adhesive for bonding the layers such as epoxy resin Koratac FL 201 (Kommerling) was used. Recently, Zhang et al. [45] produced a composite laminate of aluminum alloy-epoxy-soft steel. The main ingredient was the matrix diphenol epichlorohydrin-4,4-isopropylidene with acrylic butadiene copolymer (CTBN) as a reinforcing agent.

In the present work, we investigated the potential use of polyethylene grafted maleic anhydride (PE-g-MAH) coupling agent in the thermoplastic natural rubber (TPNR) LLDPE/NR/LNR matrix for Al-TPNR laminated composite. In the addition, the 3-GPS used as pretreatment to the surface of ground aluminium sheet in order to improve the shear strength of single lap joint aluminium-TPNR laminated composite.

2 Experimental

2.1 Material

Natural rubber (NR) was obtained from the Rubber Research Institute of Malaysia (RRIM). Liquid natural rubber (LNR) was prepared in our laboratory by photochemical oxidation of natural rubber (NR). Natural rubber (1 kg) was cut

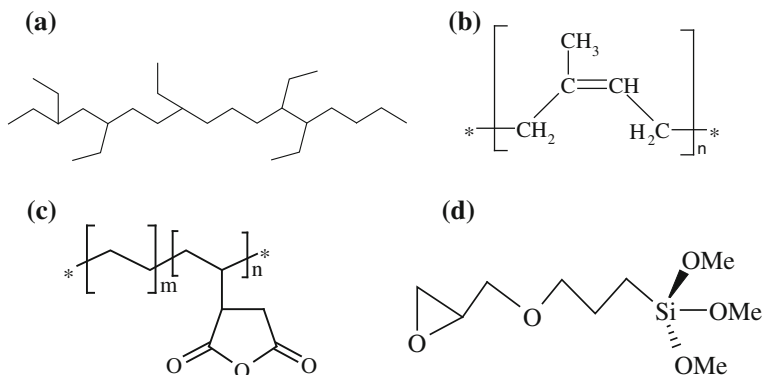


Fig. 1 Chemical structure of **a** LLDPE, **b** natural rubber (NR), **c** PE-g-MAH and **d** 3-GPS

approximately into $1 \text{ cm} \times 1 \text{ cm} \times 1 \text{ cm}$ cube. After that, it was mixed in toluene solvent with 0.2 g methylene blue, 0.26 g rose bengal and 10 mL methanol. Liquid natural rubber was formed when exposed under UV in a photoreactor after 12 days [32]. Linear low density polyethylene (LLDPE) from Exxon Mobil Chemical Manufacturing Singapore was used. The metal used is a pure aluminium sheet with specifications; thickness 1.5 mm, serial number AA1100 (Al = 99.00 %; Si and Fe = 0.05 %, Mn = 0.05 %; Zn = 0.10 %; Cu = 0.05–0.2 % and other content = 0.05 %). Other materials used are polyethylene-grafted-maleic anhydride, PE-g-MAH (melting point = 105°C ; $d = 0.925 \text{ g/mL}$) and silane coupling agent of 3-glycidyloxy-propyl-trimethoxy-silane (3-GPS), $[(\text{H}_2\text{COCH})-\text{CH}_2\text{O}(\text{CH}_2)_3\text{Si}(\text{OCH}_3)_3]$, both are manufactured in Aldrich Chemistry USA. The chemical structure of PE-g-MAH and 3-GPS are shown in Fig. 1.

2.2 Preparation of Al-TPNR Laminated Composite

Thermoplastic natural rubber (TPNR); LLDPE, NR, LNR (80:15:5) and PE-g-MAH were blended by using an internal mixer, ThermoHaake, at 140°C , rotor speed 50 r min^{-1} and mixing time of 12 min. Various concentrations of PE-g-MAH (0, 5, 10, 15, and 20 %) were incorporated into the compounds. The matrix is made of thin layers using a compression molding at a temperature of 160°C , pressure 0.5 MPa for 10 min. Aluminium sheets were cut to single lap shear tensile test, following to the standard of ASTM D 1002-72 for the sandwich structure. The aluminium sheets were boiled in distilled water for 30 min, cleaned with acetone at room temperature, and then ground with sandpaper of P600A. After that, the aluminium sheets were cleaned again with acetone at room temperature and then dried in the ambient air. Several sheets of aluminium were treated by dipping into a solution of 3-GPS 1 % (in ethanol), for 5 min, then stirring continuously for another

30 min. The sample matrix was then placed between two aluminium sheets and compressed to 10 min under the pressure of 0.5 MPa, at 160 °C temperature. The thickness of the matrix (TPNR) between two aluminium sheets are controlled to be about 0.05 mm by the applied pressure and moulding.

2.3 Characterization of Al-TPNR Laminated Composite

Tensile properties were measured using 10-tons Testometric M500-50CT tester with a crosshead speed of 1 mm/min. Every value of lap shear strength was averaged of five specimens include standard error. Fracture surface morphology of single lap shear tensile sample was carried out by using LEO 1450VP scanning electron microscope (SEM) with an acceleration voltage of 20 kV. Various elements, functional groups and bonds formed on the top surface of the aluminium were investigated using X-ray photoelectron spectroscopy (XPS) AXIS Ultra DLD with monochromatic Al K α radiation source (1486.6 eV), operating at a power of 150 W, was used to generate the emission of photoelectrons from the surface of the samples. Fourier transforms infrared (FTIR) spectroscopy spectrum 400 with attenuated total reflectance (ATR) method analysis to determinate functional group on the surface. For each spectrum, ATR ranges 4,000–650 cm⁻¹, 10 scans at a spectral 4 cm⁻¹ resolution were recorded.

3 Result and Discussion

3.1 FTIR Spectroscopy Analysis

Figure 2 shows FTIR spectra of LLDPE, PE-g-MAH, NR, TPNR (LLDPE/-NR/LNR) and TPNR plus PE-g-MAH (LLDPE/NR/LNR/PE-g-MAH). The functional groups that are related to the chemical state of Al-TPNR laminated composite interfaces such as C=O can be characterized in the frequency 1,709 cm⁻¹ (for PE-g-MAH), 1,645 cm⁻¹ (for system LLDPE/NR/LNR) and -OH in the frequency 3,394 cm⁻¹ respectively. The appearance of functional groups C=O and -OH on the matrix surface were believed to have formed during mixing as a result of material oxidation. These functional groups are important in the physical and chemical interactions in the bonding interface of Al-TPNR laminated composite [7, 28]. Chen et al. [8] reported that carboxylate complexes (Al-O-C=O) formed through interfacial interaction of MAH with metal substrates by Fourier transform infrared spectroscopy (FTIR) techniques and also proposed that carboxylate complexes formed by interfacial interaction between MAH grafted in PP and surfactant on the aluminium surface as a substrate. Yang et al. [41] reported that kinetics and reaction mechanism of γ -GPS on low carbon steel surfaces by FTIR-ATR, AFM,

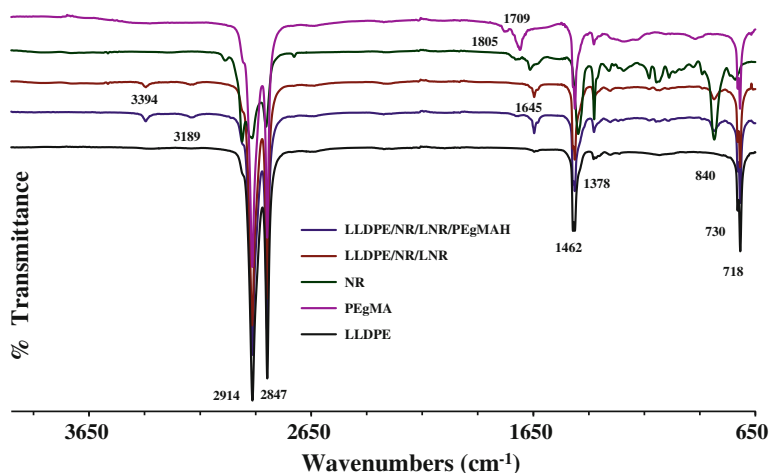


Fig. 2 FTIR spectra for LLDPE, PEGMAH, NR, LLDPE/NR/LNR and LLDPE/NR/LN-R/PEGMAH

NSS and theoretical calculation method. They reported that the reaction of γ -GPS on low carbon steel surfaces followed the conventional reaction mechanism, which can be described as a reaction (I) $(\text{Me (Metal)}-\text{OH} + \text{HO}-\text{Si} \rightarrow \text{Me}-\text{O}-\text{Si} + \text{H}_2\text{O})$ and reaction (II) $(\text{Si}-\text{OH} + \text{Si}-\text{OH} \rightarrow \text{Si}-\text{O}-\text{Si} (\text{siloxane}) + \text{H}_2\text{O})$.

3.2 XPS Analysis

An XPS spectrometer of the aluminium surface after the pretreatment procedure is shown in the Fig. 3a–c respectively. Al 2p spectra after pretreatment showed the formation of Al (72.25 eV), Al_2O_3 (74.45 eV) and 74.80 eV for Al–OOH and Al–Si bond (Al_2SiO_5), C 1s showed formation of C–C and C–H (285.6 eV), CO and COC (287.2 eV), COC and COH (286.3 eV) and COOH (288.5 eV), and O 1s also showed formation SiO_2 and MO (M = metal) (530.95 eV), O = (532.2 eV), and –O– (533.7 eV) at the surface of aluminum. The formation of OH or COOH functional groups on the surface of the aluminium, resulted in physical interaction, such as the van der Waals and hydrogen bonding interaction between TPNR or PE-g-MAH and surfactants (3-GPS) of the pretreated aluminium sheets at the interface. This has contributed to improve the interfacial adhesive strength and the chemical interactions at the interface [3, 5, 8, 28, 39].

Chen et al. [7] reported that the effect of the amount of maleic anhydride-grafted polypropylene (PP-g-MAH) in polypropylene (PP) and aluminium surface pretreated by 3-aminopropyltrimethoxysilane (3-APS) on lap shear strength of adhesive-bonded Al-PP laminated composite. For pure PP, the lap shear strengths were found to be weaker than composites. SEM analysis has shown that a fracture with

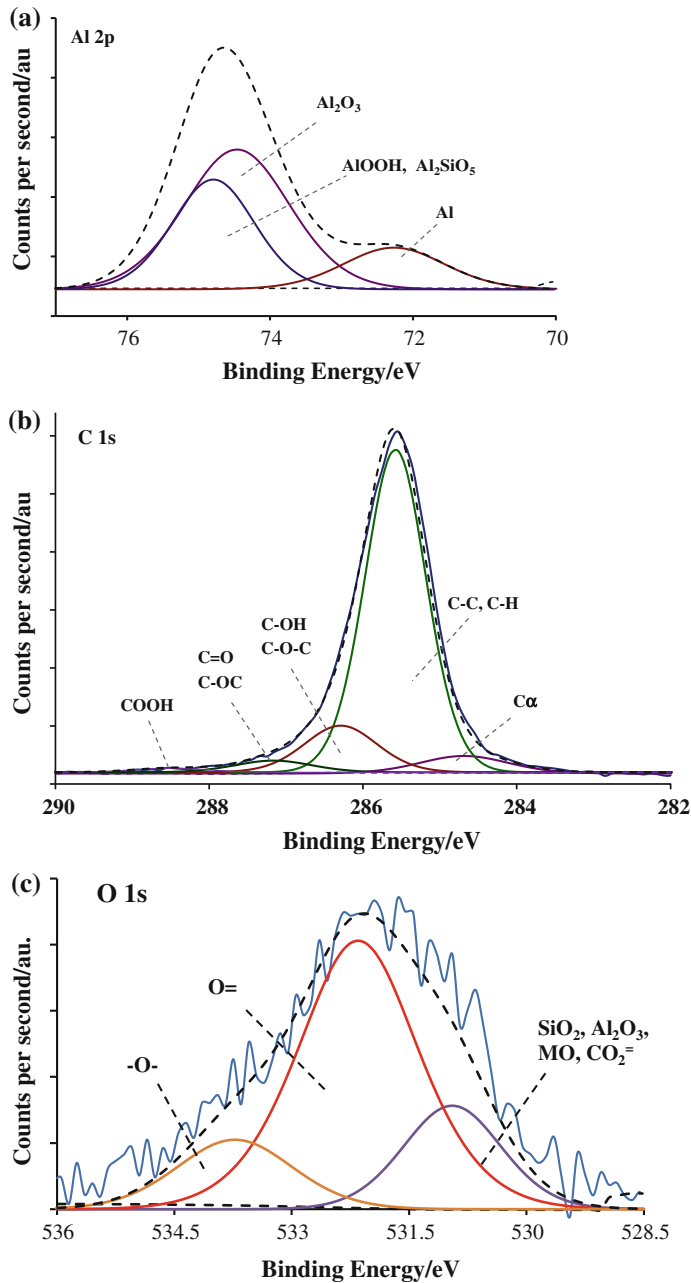


Fig. 3 XPS spectra of on the aluminium surface with 3-GPS, separated polymer surface Al-TPNR laminated composite; **a** Al 2p, **b** C 1s and **c** O 1s

almost no plastic deformation on the aluminium surface was observed. Modification of PP by the addition of 5–30 wt% amount of PP-g-MAH increased the lap shear strength significantly. It was due to the chemical interactions between OH , Al^{3+} or amino group NH_2 at the surface of the aluminium sheets and the polar functional anhydride groups and carboxylic groups COOH on PP-g-MAH at the interface (polar group investigated by XPS analysis).

3.3 SEM Analysis

Figure 4a, b shows the SEM micrographs of the lap shear fractured aluminium surfaces with and without PE-g-MAH in TPNR matrix and coupling agent, 3-GPS. It can be seen in Fig. 4a that almost all aluminum surfaces are clearly visible and almost no inherent TPNR, it is due to no interfacial physical and chemical interaction. In contrast, when the PE-g-MAH was added and aluminum surfaces pretreated with 3-GPS, it appears the fracture surface of aluminum remains as a strong matrix inherent as shown in Fig. 4b. Chen et al. [7] reported that pretreatment aluminum surface with the addition of coupling agent and maleic anhydride grafted into the plastic has enhanced the adhesion strength between polymer matrices with metal; hence an improved composite laminate could be obtained [8, 39].

3.4 Lap Shear Strength

The shear strength of the single lap joint of laminate composite was performed to analyze the effect of surface roughness, polymer type and metal surface treatment to the bond strength between metal and polymer (interface adhesion mechanism). It has been widely postulated that the creation of covalent bonds at the interface is sufficient for creating viable adhesive strength in adhesion-related applications [5, 7, 14].

Figure 5 shows the shear strength of the single lap joint Al-TPNR laminated composite with TPNR plus PE-g-MAH and the coupling agent, 3-GPS, which was added as a pretreatment on the surface of the aluminium sheet. With the addition of 15 % PE-g-MAH into the TPNR, the lap shear strength increased from 0.58 to 5.98 MPa as compared to TPNR system only. This remarkable increment of 931 % was further enhanced for the laminate composite treated with 3-GPS 1 % coupling agent. The lap shear strength was further increased to 13 % against the untreated laminate composite. This could be contributed by the physical interactions, such as the van der Waals interaction between TPNR and surfactants of the pretreated aluminium sheets as well as the mechanical interlocking between TPNR and surface roughness of the aluminium sheets at the interface. With the addition of grafting maleic anhydride into polyolefin and coupling agent on the aluminium surface, the increase in the lap shear strength should be ascribed to the contribution of chemical interactions at the interface [4, 7, 14, 24, 39].

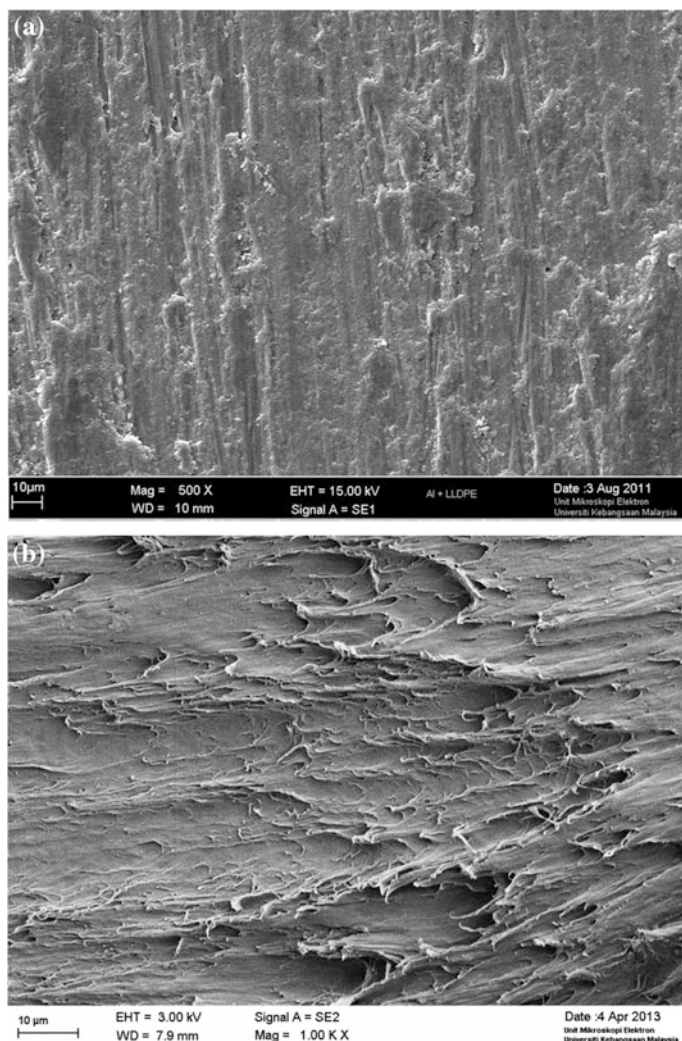
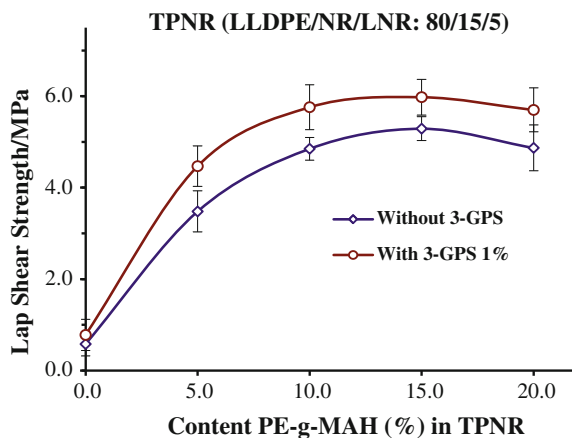


Fig. 4 SEM images of lap shear fractured surface; **a** of the adhering aluminum sheet without 3-GPS and PE-g-MAH, **b** adhered aluminium sheet with 3-GPS and PE-g-MAH 1 wt% added in TPNR matrix

The contribution in the lap shear strength laminated composite is not only due to the chemical interactions and physical strength at the interface, but also the procedure of curing. The physical and mechanical properties of adhesives, adhesive and bonding stiffness base metal interface also contributes the strength of the laminate composite. You et al. [43] proposed that the failure mode in the single lap joint may be changed with the curing procedure or conditions, physical and

Fig. 5 The effect of the PE-g-MAH content (%) on the shear strength of the single lap joint of Al-TPNR laminated composite: (—○—) with pretreatment 3-GPS 1 % solution, (—◇—) without pretreatment by 3-GPS solution



mechanical properties of the epoxy adhesive, and the steel base interface. By changing the extent of adhesive curing or interface properties, the stress distribution within the joint could be modified. Reis et al. [29] reported that the shear strength of lap joints was significantly influenced by stiffness, and the highest shear stresses were obtained by using high stiffness adherend materials. The superposition length influences the shear strength in different ways depending on the different adherend materials.

Figure 5 shows that after the addition of PE-g-MAH at 15 % concentrations, the shear strength decreases. This phenomenon occurs due to capacity (concentration) matrix interaction with the surface. Generally, the interface bonding surface strength increases monotonically with the number of interactions. Bonding strength can be determined by capacity surface bonding and interface mechanism. Song et al. [36] reported that adhesion and frictional performances of the molecule or functional groups textured surfaces strongly correlated to the wetting property of the textured surfaces that is the larger the water contact angle (WCA), the better the adhesion and friction performances of the textured surfaces. When the concentration of PE-g-MAH has reached a critical concentration, the excess concentration of PE-g-MAH is no longer interacting with the metal surface but interacts with the nonpolar from the other side of the PE-g-MAH. The interaction force between $-\text{COOH}$ functional groups of PE-g-MAH with $-\text{CH}_3$ in polymer or PE-g-MAH is a weak interaction compared interaction $-\text{COOH}$ with $-\text{OH}$ on the surface or $-\text{COOH}$ group on PP [9]. Ku et al. [15] reported that the shear strength of aluminum-polypropylene laminated composite increased after surface oxidation of polypropylene by atmospheric pressure plasma method. Increased PP surface roughness and concentration of hydrophilic functional groups $-\text{CO}$ and $-\text{OH}$ generated after plasma, indicated the formation of the chemical and physical interactions at the surface of the PP with aluminium.

4 Conclusions

The shear strength of single lap joint Al-TPNR laminated composite increased by 931 % when the PE-g-MAH was added at 15 % into the TPNR. With the 3-GPS 1 % coupling agent on the treated aluminium surface, the lap shear strength was further increased by another 13 % as compared to the untreated laminate composite. These phenomena could be associated with the physical and chemical interaction between the AlOOH , COOH , CO , COC , COH and Al-Si (Al_2SiO_5) as confirm by the XPS, FTIR and SEM analysis.

Acknowledgments This work is partially supported by the Ministry of Education and Culture, Directorate Jenderal Higher Education (DIKTI) Republic of Indonesia and Universiti Kebangsaan Malaysia for sponsoring this research project (UKM-DLP-2012-029). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

References

1. Abdelmouleh M et al (2004) Modification of cellulosic fibres with functionalised silanes: development of surface properties. *Int J Adhes Adhes* 24:43–54
2. Abdelmouleh M et al (2007) Short natural-fibre reinforced polyethylene and natural rubber composites: effect of silane coupling agents and fibres loading. *Compos Sci Technol* 67:1627–1639
3. Abel M-L et al (2006) Effect of solvent nature on the interaction of g-glycidoxypropyl trimethoxy silane on oxidised aluminium surface: a study by solution chemistry and surface analysis. *Int J Adhes Adhes* 26:16–27
4. Bhowmik S et al (2004) Experimental investigation into the effect of DC glow discharge pretreatment of HDPE on tensile lap shear strength. *Int J Adhes Adhes* 24:461–470
5. Carradò A et al (2011) Metal/polymer/metal hybrid systems: towards potential formability applications. *Compos Struct* 93:715–721
6. Cerclé C, Sarazin P, Favis BD (2013) High performance polyethylene/thermoplastic starch blends through controlled emulsification phenomena. *Carbohydr Polym* 92:138–148
7. Chen MA, Li HZ, Zhang XM (2007) Improvement of shear strength of aluminium-polypropylene lap joints by grafting maleic anhydride onto polypropylene. *Int J Adhes Adhes* 27:175–187
8. Chen MA, Xie X, Zhang XM (2009) Interactions of BTESPT silane and maleic anhydride grafted polypropylene with epoxy and application to improve adhesive durability between epoxy and aluminium sheet. *Prog Org Coat* 66:40–51
9. Cheng H, Hu Y (2012) Influence of chain ordering on frictional properties of self-assembled monolayers (SAMs) in nano-lubrication. *Adv Colloid Interface Sci* 171–172:53–65
10. Choi HS et al (2006) Surface oxidation of polyethylene using an atmospheric pressure glow discharge with liquid electrolyte cathode. *J Colloid Interface Sci* 300:640–647
11. Clemons C (2010) Elastomer modified polypropylene–polyethylene blends as matrices for wood flour–plastic composites. *Composites A Appl Sci Manuf* 41:1559–1569
12. da Silva LFM et al (2009) Effect of material, geometry, surface treatment and environment on the shear strength of single lap joints. *Int J Adhes Adhes* 29:621–632
13. Farris S et al (2010) The fundamentals of flame treatment for the surface activation of polyolefin polymers—a review. *Polymer* 51:3591–3605

14. Gao H et al (2012) Grafting effects of polypropylene/polyethylene blends with maleic anhydride on the properties of the resulting wood–plastic composites. *Compos A Appl Sci Manuf* 43:150–157
15. Ku JH et al (2013) Atmospheric pressure plasma treatment of polypropylene to improve the bonding strength of polypropylene/aluminum composites. *Compos B Eng* 45:1282–1287
16. Matinlinna JP, Lassila LVJ, Vallittu PK (2007) The effect of five silane coupling agents on the bond strength of a luting cement to a silica-coated titanium. *Dent Mater* 23:1173–1180
17. Mondal M et al (2013) Polypropylene/natural rubber thermoplastic vulcanizates by eco-friendly and sustainable electron induced reactive processing. *Radiat Phys Chem* 88:74–81
18. Murray KA et al (2012) The effects of high energy electron beam irradiation on the thermal and structural properties of low density polyethylene. *Radiat Phys Chem* 81:962–966
19. Nabesawa H et al (2008) Polymer surface morphology control by reactive ion etching for microfluidic devices. *Sens Actuators B* 132:637–643
20. Nachtigall S, Cerveira G, Rosa S (2007) New polymeric-coupling agent for polypropylene/wood–floor composites. *Polym Test* 26:619–628
21. Nakason C et al (2006) Dynamic vulcanization of natural rubber/high-density polyethylene blends: effect of compatibilization, blend ratio and curing system. *Polym Test* 25:782–796
22. Navaneetha Pandiyaraj K et al (2008) Adhesive properties of polypropylene (PP) and polyethylene terephthalate (PET) film surfaces treated by DC glow discharge plasma. *Vacuum* 83:332–339
23. Novák I et al (2008) Polymer matrix of polyethylene porous films functionalized by electrical discharge plasma. *Eur Polym J* 44:2702–2707
24. Parsa MH, Ahkami SNA, Ettehad M (2010) Experimental and finite element study on the spring back of double curved aluminum/polypropylene/aluminum sandwich sheet. *Mater Des* 31:4174–4183
25. Pavlidou S, Papaspyrides CD (2008) A review on polymer-layered silicate nanocomposites. *Prog Polym Sci* 33:1119–1198
26. Pereira AM et al (2010) Analysis of manufacturing parameters on the shear strength of aluminium adhesive single-lap joints. *J Mater Process Technol* 210:610–617
27. Prolongo SG, Ureña A (2009) Effect of surface pre-treatment on the adhesive strength of epoxy-aluminium joints. *Int J Adhes Adhes* 29:23–31
28. Putman CO, Vaidya UK (2011) Mechanisms of interfacial adhesion in metal-polymer composites—effect of chemical treatment. *Compos A Appl Sci Manuf* 42:906–915
29. Reis PNB, Ferreira JAM, Antunes F (2011) Effect of adherend's rigidity on the shear strength of single lap adhesive joints. *Int J Adhes Adhes* 31:193–201
30. Rider AN, Arnott DR (2000) Boiling water and silane pre-treatment of aluminium alloys for durable adhesive bonding. *Int J Adhes Adhes* 20:209–220
31. Rimpelová S et al (2013) Plasma treated polyethylene grafted with adhesive molecules for enhanced adhesion and growth of fibroblasts. *Mater Sci Eng C* 33:1116–1124
32. Seng LY et al (2011) Effects of liquid natural rubber (LNR) on the mechanical properties of LNR toughened epoxy composite. *Sains Malaysiana* 40:679–683
33. Sideridou ID, Karabela MM (2009) Effect of the amount of 3-methacryloxypropyltrimethoxysilane coupling agent on physical properties of dental resin nanocomposites. *Dent Mater* 25:1315–1324
34. Sinmazgelik T et al (2011) A review: fibre metal laminates, background, bonding types and applied test methods. *Mater Des* 32:3671–3685
35. Sokolova OA, Kühn M, Palkowski H (2012) Deep drawing properties of lightweight steel/polymer/steel sandwich composites. *Arch Civil Mech Eng* 12:105–112
36. Song Y et al (2010) Adhesion and friction properties of micro/nano-engineered superhydrophobic/hydrophobic surfaces. *Thin Solid Films* 518:3801–3807
37. Weiss M et al (2007) The influence of temperature on the forming behavior of metal/polymer laminates in sheet metal forming. *J Eng Mater Technol* 129:530–537

38. Wolf RA (2008) Improving adhesion performance between low surface tension composite and dissimilar substrates. In: Topical conference on path to innovation: new technologies in decorating and assembly, pp 137–140
39. Xie Y et al (2010) Silane coupling agents used for natural fiber/polymer composites: a review. *Compos A Appl Sci Manuf* 41:806–819
40. Xie R et al (2011) Aligned carbon nanotube coating on polyethylene surface formed by microwave radiation. *Compos Sci Technol* 72:85–90
41. Yang L et al (2010) Experimental and computational study on hydrolysis and condensation kinetics of γ -glycidoxypyltrimethoxysilane (γ -GPS). *Appl Surf Sci* 257:990–996
42. Yao F et al (2008) Rice straw fiber-reinforced high-density polyethylene composite: effect of fiber type and loading. *Ind Crops Prod* 28:63–72
43. You M et al (2003) Effect of metal as part of fillet on the tensile shear strength of adhesively bonded single lap joints. *Int J Adhes Adhes* 23:365–369
44. Yuan H et al (2012) Effect of surface modification on carbon fiber and its reinforced phenolic matrix composite. *Appl Surf Sci* 259:288–293
45. Zhang F et al (2013) Durability of adhesively-bonded single lap-shear joints in accelerated hygrothermal exposure for automotive applications. *Int J Adhes Adhes* 44:130–137
46. Zhang H-P et al (2009) Molecular dynamics simulations on the interaction between polymers and hydroxyapatite with and without coupling agents. *Acta Biomater* 5:1169–1181

A New Technique to Observe ENSO Activity via Ground-Based GPS Receivers

Wayan Suparta, Ahmad Iskandar and Mandeep Singh Jit Singh

Abstract In an attempt to study the effects of global climate change in the tropics for improving global climate model, this paper aims to detect the ENSO events, especially El Niño phase by using ground-based GPS receivers. Precipitable water vapor (PWV) obtained from the Global Positioning System (GPS) Meteorology measurements in line with the sea surface temperature anomaly (SSTa) are used to connect their response to El Niño activity. The data gathered from four selected stations over the Southeast Asia, namely PIMO (Philippines), KUAL (Malaysia), NTUS (Singapore) and BAKO (Indonesia) for the year of 2009/2010 were processed. A strong correlation was observed for PIMO station with a correlation coefficient of -0.90 , significantly at the 99 % confidence level. In general, the relationship between GPS PWV and SSTa at all stations on a weekly basis showed with a negative correlation. The negative correlation indicates that during the El Niño event, the PWV variation was in decreased trend. Decreased trend of PWV value is caused by a dry season that affected the GPS signals in the ocean-atmospheric coupling. Based on these promising results, we can propose that the ground-based GPS receiver is capable used to monitor ENSO activity and this is a new prospective method that previously unexplored.

1 Introduction

Global Positioning System (GPS) is satellite constellation of geodetic system that has been employed for determining position on the earth's surface in 3D and time. During its development, GPS technique can be applied to observe the characteristics of the Earth's atmosphere, including weather conditions. The 'weather' here, covering the sun until the surface activities. Bevis et al. [1] was firstly introduced the technique to observe the weather as well as climate by proposing the ground-based GPS and

W. Suparta (✉) · A. Iskandar · M.S.J. Singh
Space Science Centre (ANGKASA), Institute of Climate Change,
Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor Darul Ehsan, Malaysia
e-mail: wayan@ukm.my

combined with meteorological sensors which are known as the GPS meteorology. GPS meteorology carried to characterize the diversity of climate behavior in the lower atmosphere (the troposphere) and the upper atmosphere layers (Ionosphere) [2]. This can be done by exploiting the delay of GPS satellites in the atmosphere by propagation of electromagnetic signals. The measurements of GPS signals are then used to detect variations that occur in the atmosphere. One major of the atmospheric component caused the signal delay is the precipitable water vapor (PWV).

PWV is a paramount parameter that has crucial roles in the weather and global climate. One of the factors that possibly affecting the PWV variation is due to El Niño-Southern Oscillation (ENSO) phenomenon. ENSO is a complex phenomenon that results from interaction between the ocean and the atmosphere in the tropical Pacific Ocean, which consisted of warm (El Niño) and cool (La Niña) phases. There are some indicators to detect ENSO phenomenon that has been employed, such as sea level pressure anomaly (SLPa), sea surface temperature anomaly (SSTa) and multivariate ENSO index (MEI) and the zonal gradient of precipitation [3–6]. In addition to employ of these parameters, the United States collaborated with the French in 1992 to launch altimetry satellite TOPEX/Poseidon (T/P) for ENSO monitoring [7]. However, the data from this monitoring is still ongoing, which expected for climate research.

Here, GPS measurements through PWV have been employed to study the ENSO activity, where La Niña is one of the phases of ENSO that causes increased of PWV value [8]. The use of this technique is effective-cost and the system no extra maintenance for a longer period, and the data acquired is also useful for space weather studies. Therefore, this study will focus on the analysis of ENSO, especially El Niño phase by using GPS meteorology over the Southeast Asia region. With this technique, an understanding of the characteristics and physical mechanism during El Niño phase can be provided.

2 Methodology

2.1 PWV Retrieved from GPS

The concept of measurements of PWV from GPS technique is depicted in Fig. 1. From the figure, the GPS signal propagate to a receiver on the ground that has been interacting with sea surface of the ocean. When the GPS signal propagates through the Earth's atmosphere, it is affected by the variability of the refractive index of the ionosphere and troposphere. The excess delay of the signal causes bending of the signal, and the total delay along the slant path can be determined [9]. In line with that, the total tropospheric delay (ZTD) in the neutral atmosphere, which comprised the zenith hydrostatic delay (ZHD) and zenith wet delay (ZWD) can be calculated based on the improved Modified Hopfield model [2, 10]. The ZHD was calculated using the Saastamoinen model [11]. A Vienna mapping function (VMF1) was employed to reduce the atmospheric bias in the ZTD estimation [12]. The ZWD was computed by subtracting ZHD from ZTD. The ZWD was then transformed into

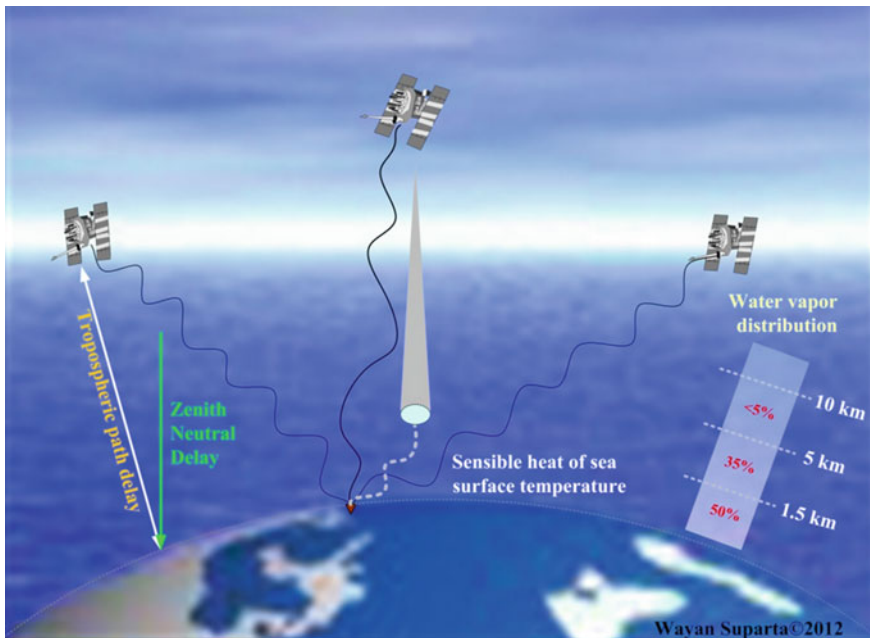


Fig. 1 Propagation of GPS signals to a receiver on the ground that covers the sea surface temperature influences adopted from Suparta et al. [8]

an estimate PWV by employing the surface temperature measured at a particular site. The PWV total (in mm) from a receiver position to the top of the atmosphere can be calculated. Detailed of PWV determination from GPS observations for this work can be found in the paper of Suparta et al. [2] and Suparta [13].

To calculate the PWV, the GPS data is combined with meteorological data (pressure (P in mbar), temperature (T in $^{\circ}\text{C}$) and relative humidity (H in percent)). The meteorological data obtained from the meteorology station should be collocated with the GPS station to obtain optimum accuracy of vertical water vapor. For the meteorological data that was not collocated with GPS station, interpolation of the meteorological data into the GPS station will be conducted by using the formula, as proposed by Klein Balting et al. [14]. The total PWV was estimated using the formula proposed by Bevis et al. [15]:

$$PWV = \pi(T_m)ZWD \quad (1)$$

where $\pi(T_m)$ is the conversion factor that varies with local climate (e.g., geographical condition, season and weather) and dependent on a weighted mean temperature (T_m), ZWD is Zenith Wet Delay derived from GPS observations. Based on the above formula, T_m is crucial parameter in the estimation of PWV value, and for the Western Pacific region with latitude range 20°N – 20°S and longitude 95°E – 156°E , the T_m was obtained as [16]

$$T_m = 0.83663 T_s + 48.103 \quad (2)$$

where T_m is in Kelvin and T_s is the surface temperature (in Kelvin). Note that the empiric T_m equation was obtained by a regression linear method for 15 selected radiosonde stations over the Western Pacific.

2.2 Data and Location

The GPS and meteorological data for this study were gathered from four selected stations over the Southeast Asia region. El Niño phase that occurred in 2009/2010 is selected, and the data was processed to verify the response of GPS PWV on ENSO phenomenon. The GPS station consists of the Philippine Island Manila Observatory (PIMO), Kuala Terengganu (KUAL), Nanyang Technological University Singapore (NTUS) and Bakosurtanal (BAKO). For the four selected stations, PIMO is located at Quezon City of Metro Manila, Philippines, while KUAL station is located on the East Coast of Peninsular Malaysia. Meanwhile, the NTUS station is located in the south-western part of Singapore, near the Jurong West Extension area, and slightly moved downward, BAKO is located in Cibinong, West Java, Indonesia. For the meteorological station, the data is collected from Manila airport weather station (Philippines), Kuala Terengganu (Malaysia), Nanyang Technological University (Singapore), and Soekarno-Hatta (Indonesia). The GPS data from KUAL station was obtained from the Department of Survey and Mapping Malaysia (DSMM), while another station was downloaded from the Scripps Orbit and Permanent Array Center (SOPAC) website. All the GPS data were supplied with resolution of a 30 s interval. The meteorological data were downloaded from the Weather Underground website.

The location of the study is depicted in Fig. 2. The specific of geographical coordinates for GPS stations and instruments set up of the GPS receivers is shown in Table 1, while the specific of geographical coordinates of meteorology station is compiled in Table 2. To relate the ENSO activity with PWV response, SSTa Oceanic Niño Index (ONI) in pathways of Niño 3.4 and Niño 4 regions is employed. The SSTa data were obtained from the National Oceanic and Atmospheric Administration (NOAA) which provided weekly.

2.3 Data Processing

To process and analyze all the parameters, a tropospheric water vapor program (*TroWav*) developed by Suparta et al. [2, 12], Suparta [13] was employed. The algorithms of the *TroWav* include satellite elevation angle, the ZTD, the ZHD, the ZWD and mapping function calculations. Figure 3 shows the flowchart of PWV determination. Then to study the relationship between GPS PWV and ENSO activity uses the correlative analysis. All the data are analyzed on a weekly basis

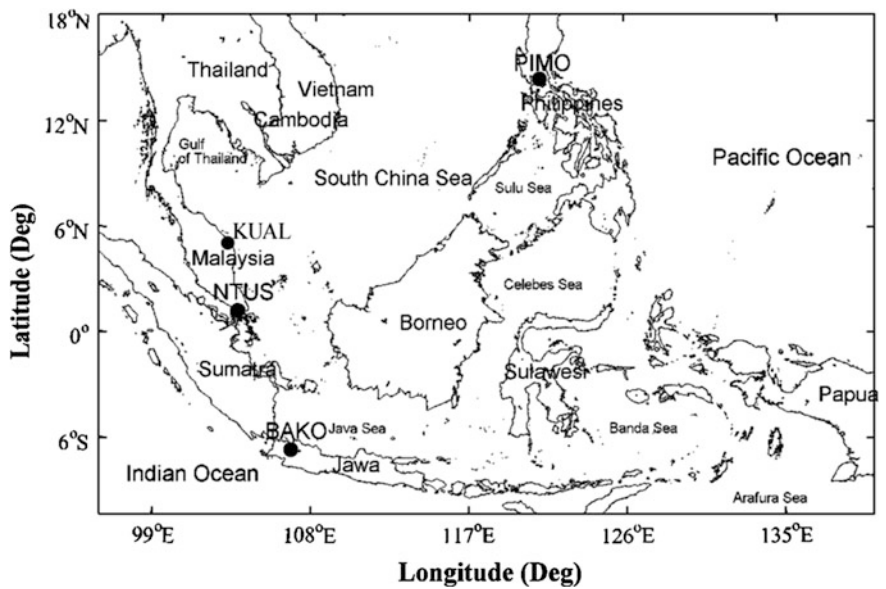


Fig. 2 Location of the study

Table 1 Geographical coordinates and instruments set up of GPS receivers

Station (country)	Latitude (degree)	Longitude (degree)	Height (m)	Type of GPS receivers
PIMO (Philippines)	14.64N	121.08E	95.53	ASHTech UZ-12
KUAL (Malaysia)	5.32N	103.14E	55.00	Trimble NetR5
NTUS (Singapore)	1.35N	103.68E	75.38	LEICA GRX1200GGPRO
BAKO (Indonesia)	6.49S	106.85E	158.20	LEICA GRX1200+GNSS

Table 2 Geographical coordinates of meteorological stations

Station (country)	Latitude (degree)	Longitude (degree)	Elevation (m)
Manila (Philippines)	14.50N	121.00E	21.00
Kuala Terengganu (Malaysia)	5.30N	103.10E	9.00
NTUS (Singapore)	1.35N	103.68E	75.38
Soekarno-Hatta (Indonesia)	6.10S	106.70E	8.00

because the SSTa data only available on a weekly. The weekly data processing is based on the GPS week from daily data. In the figure, GPS data for KUAL are available in binary format (*.dat), consisting of *.nav and *.obs. To translate, repair and testing the GPS data uses a Translate/Edit/Quality Check (TEQC) routine developed by UNAVCO (<http://www.unavco.org/>). The TEQC routine can convert *.dat file into the Receiver Independent Exchange Format (RINEX) files. Then split

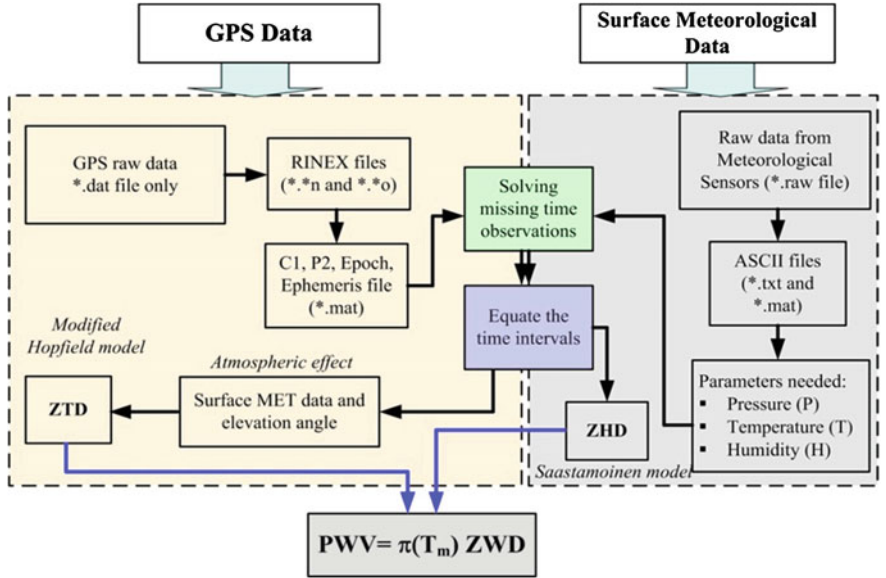


Fig. 3 Flowchart of PWV estimation adopted from Suparta et al. [8]

into individual files (L1, L2, C1, P2, epo, eph) for easy processing using Matlab™ program. For another station, all the RINEX observation file uses hatanaka (d-file) format and processed with 30 s interval (Fig. 4).

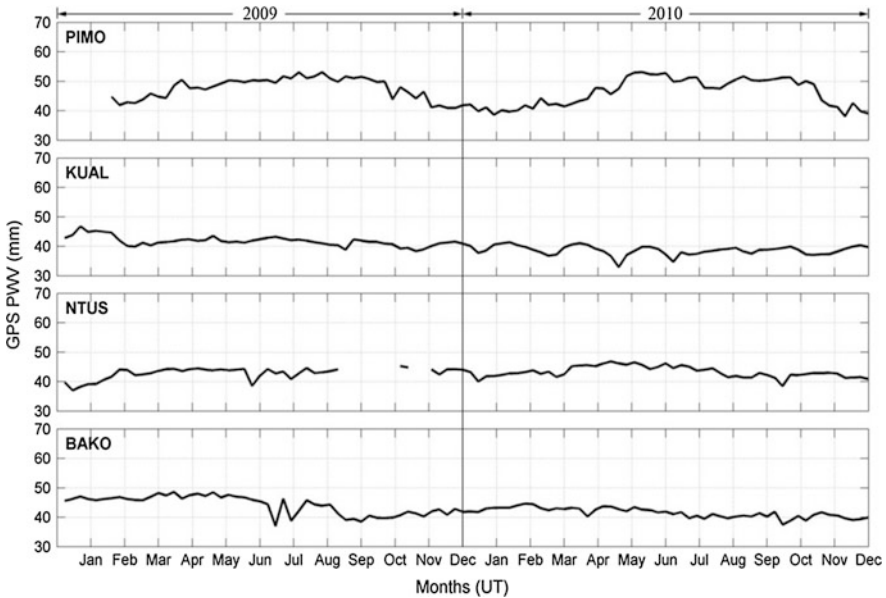


Fig. 4 The weekly average of PWV variation from GPS measurements for the year of 2009/2010

3 Results and Discussion

3.1 PWV and SSTa Variability

Figure 3 shows the weekly average of PWV variation from GPS measurements for the year of 2009/2010 at four selected stations. Based on the figure, the PWV pattern at each station is varied. For PIMO station, the range of PWV value was between 38 and 53.12 mm with an average of 46.90 mm. Another station like KUAL, NTUS and BAKO, the PWV value was lower than the PIMO station. Their averages of PWV were 40.20, 43.00 and 42.83 mm, respectively. Looked at the standard deviation (STD) value, the PWV for PIMO station was found higher of 4.33 mm than the other stations. The detail comparison of statistical analysis of PWV for all stations is presented in Table 3. From the figure, PIMO station showed a clear annual pattern of PWV, which is higher during the summer and lower during the winter. This trend almost followed by PWV at NTUS station, although the data is not completed. For BAKO and KUAL, the PWV trend is difficult to characterize although they appeared increased during the first inter-monsoon (April–May) and second inter-monsoon (October–November). The ripples peak of PWV occurred during June–July of 2009 at BAKO and NTUS stations shows that ENSO activity starting to disturb the atmosphere of both stations.

ENSO plays a crucial role in generating extreme conditions across the Pacific Ocean, in Indonesia and in the Borneo region. The more serious ENSO events with infamous their direct effects on communal health and livelihood were taken place in 1982/1983 and 1997/1998. For that reason, the ability to understand the ENSO pattern and predict its occurrences will allow for improved preparation to mitigate its effects and simultaneously support the preservation of the tropical rain forest in a targeted region. To observe their relevance to the water vapor changes during ENSO occurrences, two definitions to identify ENSO events can be applied. First, an ENSO event occurs when the 5-month running mean of the SST anomalies (SSTa) in the Niño 3.4 region is continuously higher than $\pm 0.4\text{ }^{\circ}\text{C}$ for 6 months [17]. The second definition was provided by the Japan Meteorological Agency (JMA) using a 1° meridionally narrower region of Niño 3 (4°N – 4°S and 150°W – 90°W), whereby an ENSO event is determined to have occurred when the 5-month running mean of the spatially averaged SSTa is larger than $\pm 0.5\text{ }^{\circ}\text{C}$ for six consecutive months [18]. The reason employed of 5-month running mean of SSTa is to average over the intra-seasonal variations in the tropical ocean. So that SSTa turns out to be a good indicator to detect ENSO activity.

Table 3 The statistical analysis of PWV

Station	Minimum (mm)	Mean (mm)	Maximum (mm)	STD (mm)
PIMO	38.06	46.90	53.12	4.33
KUAL	32.95	40.20	46.78	2.24
NTUS	36.99	43.00	46.83	1.92
BAKO	37.01	42.83	48.59	2.83

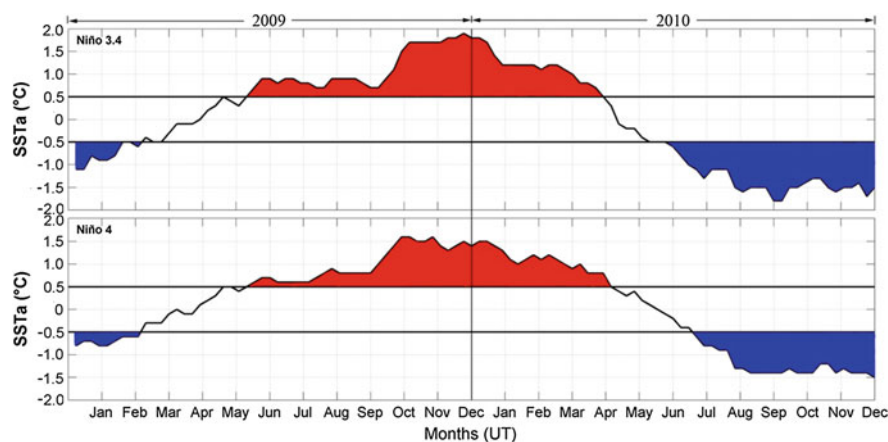


Fig. 5 The weekly SSTa variation at Niño 3.4 and Niño 4 regions for the year of 2009/2010

Figure 5 presents the ENSO identification according to JMA ($SST_a > 0.5^\circ\text{C}$) for eleven months or more is an ENSO warming (El Niño) event, while the shaded region with $SST_a < -0.5^\circ\text{C}$ (JMA) is an ENSO for cooling phase (La Niña event), and otherwise is neutral. From these two definitions, we confident to propose an ENSO event will occurs when $SST_a > \pm 0.5^\circ\text{C}$ at Niño 3 and Niño 3.4 regions. For the analysis, the weekly SSTa data throughout the year of 2009/2010 is presented. In general, from the SSTa changes at Niño 3.4 or Niño 4, the El Niño phase is occurred from the middle of June 2009 until April 2010. From the figure, it can also be seen that the maximum peak of El Niño was occurred in December 2009 for Niño 3.4 region with SSTa value of 1.9°C and October–November 2009 for Niño 4 region with SSTa value of 1.6°C . In our case, the selected site of study was located nearby the Niño 4 region.

3.2 Monitoring of GPS PWV Variability During El Niño Event

As mentioned previously, the correlative analysis is used to indicate the PWV response on El Niño event. During the event from June 2009 to April 2010, the relationship between GPS PWV and SSTa for Niño 3.4 and Niño 4 regions is varied at each station. Figure 6 shows the scatterplot between GPS PWV and SSTa for Niño 3.4 region for both increasing and decreasing phases. During the increased in intensity of El Niño in June–December of 2009 (Fig. 6a), a strong relationship was observed for PIMO station with a correlation coefficient of -0.90 , moderate relationship was obtained for KUAL station with a correlation coefficient of -0.56 and found no correlation for NTUS and BAKO stations, which are all significant at the 99 % confidence level. A low correlation during decreasing in intensity of El Niño

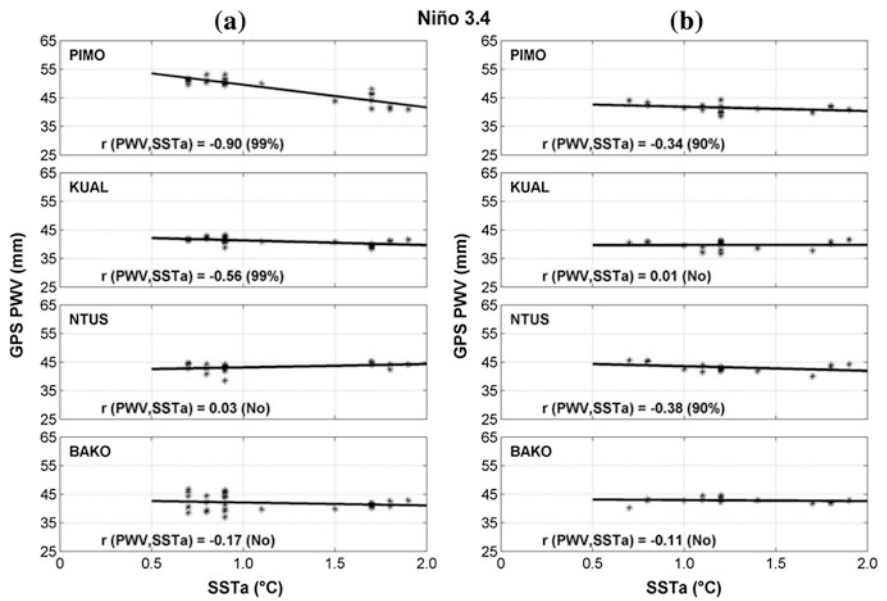


Fig. 6 Scatterplot between GPS PWV and SSTa for Niño 3.4 region during El Niño event for the year of 2009/2010 for **a** the increasing in intensity and **b** the decreasing in intensity

(December 2009–April 2010) was observed (Fig. 6b) for PIMO and NTUS stations with correlation coefficients are -0.34 and -0.38 , respectively, while for KUAL and BAKO stations are found no correlation (as indicated by **No** in the figure).

For the analyzing of relationship between GPS PWV and SSTa for Niño 4 region, the scatterplot during El Niño event 2009/2010 is presented in Fig. 7. As shown in Fig. 7a, the relationship between GPS PWV and SSTa during the increased in El Niño intensity for PIMO and KUAL stations were observed moderate, except for NTUS and BAKO with a low correlation. While during the decreasing in intensity of El Niño (Fig. 7b), PIMO and KUAL was not correlated. A low correlation was observed for NTUS and BAKO stations with correlation coefficients of -0.39 and -0.40 , respectively. From the figure, BAKO station was still with a low correlation, which likely low impacted by the El Niño event of 2009/2010. On the other hand, during increasing in intensity of El Niño, station located along the coast is more pronounced that the station in the land like NTUS and BAKO. The low effect of El Niño at specific site is probably due to season change from winter to summer.

The detail of the correlation coefficient value at all stations is compiled in Table 3. From the table, SSTa for Niño 3.4 shows the strong relationship on the GPS PWV. Moreover, the relationship between GPS PWV and SSTa at all stations during El Niño event is negative correlation, except for NTUS and BAKO during the increasing of El Niño intensity. The negative correlation indicated that the amount of GPS PWV is decreased during the event. This brings consequence that the delay of GPS signals from satellite to the receiver on the ground is larger. One

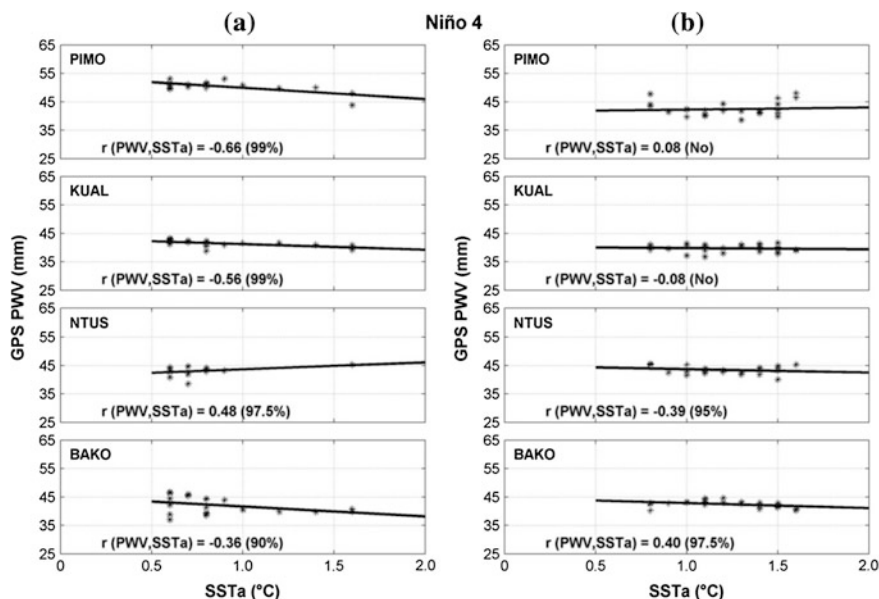


Fig. 7 Scatterplot between GPS PWV and SSTA from Niño 4 region during El Niño event for the year of 2009/2010 for **a** the increase in intensity and **b** the decrease in intensity

can be seen from Figs. 6 and 7 is that during increasing of El Niño intensity for El Niño 3.4 and El Niño 4 regions, the correlation trend is consistent for PIMO, KUAL and BAKO (decreasing trend) and NTUS (increasing trend). However, during decreasing of El Niño intensity at both Niño regions, only NTUS and BAKO was consistent in decreasing trend and PIMO and KUAL are opposite to each other. However, PIMO station experiences the highest El Niño effect than the other stations. This implies that the El Niño character in each year is apparently different which possibly due to the trade winds that across the region moved the water vapor over the ocean and occur too short or even occur too late (see Sect. 3.3 for detail comparison with precipitation). This depends on the latent heat of the sea surface temperature that will affect the distribution of water vapor in the atmosphere.

3.3 Analysis of Rainfall During El Niño Event

To strengthen the possible physical process of different response between PWV and SSTA variation, precipitation quantity from Tropical Rainfall Measuring Mission (TRMM) during a case of El Niño event is compared. Figure 8 shows the distribution of rainfall measured by TRMM-3B42.007 over the Southeast Asia region during El Niño from June 2009 to April 2010. From the figure, the most significant change in monthly average of rainfall was occurred in PIMO station (Philippines)

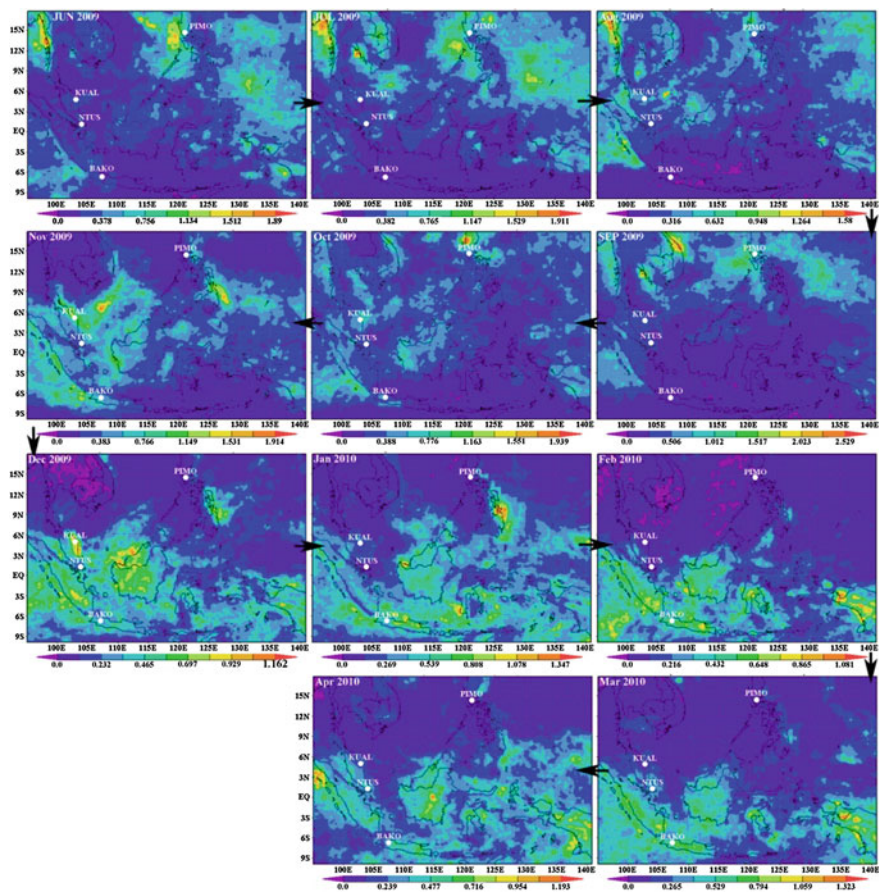


Fig. 8 Distribution of rainfall measured by TRMM-3B42.007 over the Southeast Asia region during El Niño event from June 2009 to April 2010

than the other stations. This is consistent with the PWV changes from GPS measurements. The beginning of El Niño event was from the month of June 2009 until September 2009, where the average of rainfall was higher, and the lower was occurred at KUAL, NTUS and BAKO stations. BAKO station is an area that has very low rainfall during the increasing of the intensity of El Niño from June to October of 2009. Along with the increasing intensity of the El Niño, high rainfall shifted towards the south. This causes an increase in rainfall for the station in KUAL, NTUS and BAKO. In addition to this condition, the drought occurred in the Philippines [19, 20] lead PIMO showed an opposite trend and gives a significant effect to the decreasing in agricultural production [21].

We here obviously obtained that the average of rainfall distribution is statistically comparable with PWV pattern as shown in Fig. 4. During increasing the intensity of El Niño, the monthly average of rainfall at PIMO station is increased

Table 4 The correlation coefficient between GPS PWV and SSTa during El Niño event for the case of 2009/2010

Station	The increase in intensity		The decrease in intensity	
	Niño 3.4	Niño 4	Niño 3.4	Niño 4
PIMO	−0.90	−0.66	−0.34	–
KUAL	−0.56	−0.56	–	–
NTUS	–	0.48	−0.38	−0.39
BAKO	–	−0.36	–	−0.40

Sign ‘–’ indicates no correlation

and the PWV pattern is decreased. This clearly understood that when the rainy occurred, the water vapor content in the atmosphere is fall become a water. The low or no correlation obtained between GPS PWV and SSTa in Table 4 can be seen through the contours of Fig. 8. Instead, the different thing was occurred at BAKO station where the rainfall is increased during the decreased intensity of El Niño, which is a similar trend with PWV pattern. The low of precipitation and small effect on El Niño was difficultly to differentiate the relationship trend.

4 Conclusion

This study uses a correlation analysis to determine the relationship between GPS PWV and SSTa variability during the case of El Niño event. The data from PIMO, KUAL, NTUS and BAKO stations during the year of 2009/2010 were selected to calculate PWV. The relationship between GPS PWV and SSTa was found negative correlation, although the different trend was observed at each station. The negative correlation during the El Niño event can be interpreted that the PWV trend is decreased. Decreasing of PWV is correlated with a dry season, where water vapor is minimal in the atmosphere. On the other hand, rainfall is minimal at the location of a strong El Niño. For the case of 2009/2010 during increasing of El Niño intensity, the station located along the coast such as PIMO showed more affected than the other stations. In contrast, increased of PWV along the ocean-atmosphere interactions caused the delay of GPS signals is higher, which possibly existing the atmospheric phenomena above the ocean.

From the correlation obtained between PWV and SSTa, the ground-based GPS receiver is capable proposed as an alternative technique to be used to detect ENSO activity, especially El Niño event, and this is a new prospective method that has not been utilized before. It should be noted here that the El Niño can affect the GPS signals through indirect effects of PWV measurements. To clarify the clear response of GPS PWV on the SSTa variation, a strong ENSO event will be considered for future work to distinguish the unwanted noise in the analysis. More GPS PWV data during El Niño event are also needed to better understand the ENSO mechanism, both during the occurrences of El Niño and El Niña phases.

Acknowledgments This research is partially supported by the Ministry of Higher Education Malaysia (MOHE) under LRGS/TD/2011/UKM/PG/01, UKM-LL-07-FGRS0211-2010 and Modal Insan grants. The authors would like to thank the Wyoming University for Radiosonde data, the Scripps Orbit and Permanent Array Center (SOPAC) for GPS data and Weather Underground for meteorological data.

References

1. Bevis M, Businger S, Herring TA et al (1992) GPS meteorology: remote sensing of atmospheric water vapor using the global positioning system. *J Geo Res* 97 (D14):15787–15801
2. Suparta W, Abdul Rashid ZA, Mohd Ali MA et al (2008) Observation of Antarctic precipitable water vapor and its response to the solar activity based on GPS sensing. *J Atmos Sol-Terr Phys* 70:1419–1447
3. Ropelewski CF, Jones PD (1987) An extension of the Tahiti-Darwin Southern oscillation index. *Mont Weat Rev* 115:2161–2165
4. Reynolds RW, Smith TM (1995) A high-resolution global sea surface temperature climatology. *J Clim* 8:1571–1583
5. Wolter K, Timlin MS (1998) Measuring the strength of ENSO events: how does 1997/98 rang? *Weather* 53(9):315–324
6. Curtis S, Adler R (2000) ENSO indices based on patterns of satellite-derived precipitation. *J Clim* 13:2786–2793
7. Chambers DP (1998) Monitoring El Niño with satellite altimetry El Niño workshop-online for education. Center for Space Research the University of Texas, Austin, p 16
8. Suparta W, Iskandar A, Singh MSJ et al (2013) A study of El Niño-Southern oscillation impacts to the South China Sea region using ground-based GPS receiver. *J Phys Conf Ser* 423:012043. doi:[10.1088/1742-6596/423/1/012043](https://doi.org/10.1088/1742-6596/423/1/012043)
9. Suparta W (2012) Detection of a katabatic wind event with GPS meteorology measurements at Scott Base Antarctica. *Meteorol Atmos Phys* 115(3–4):153–162
10. Hofmann-Wellenhof B, Lichtenegger H, Collins J (1994) Global positioning system theory and practice. Springer, New York, pp 113–115
11. Saastamoinen J (1972) Introduction to practical computation astronomical refraction. *Bul Geod* 106:383–397
12. Suparta W, Singh MSJ, Mohd Ali MA et al (2011) GPS water vapor monitoring and TroWav update for ENSO studies. Paper presented at the 2nd international conference on instrumentation, communications, information technology and biomedical engineering (ICICI-BME 2011), Institut Teknologi Bandung, Indonesia, 8–9 Nov 2011
13. Suparta W (2010) Using a global positioning system to estimate precipitable water vapor in Antarctica. *Polar Geogr* 33(1):63–79
14. Klein Baltink H, Derks HJP, Van Lammeren ACAP et al (1999) Water vapour from GPS tropospheric delay estimates, In: GPS water vapour meteorology, Beleids Commissie Remote Sensing (BCRS), chapter 2, pp 3–11
15. Bevis M, Businger S, Chiswel S et al (1994) GPS meteorology: mapping zenith wet delay onto precipitable water. *J App Met* 33:379–386
16. Suparta W, Iskandar A (2013) Modeling of weighted mean temperature over the Western Pacific Region to estimate GPS PWV. Paper presented at the 2013 IEEE international conference on space science and communication, Institute of Space Science, Universiti Kebangsaan Malaysia, Melaka, 1–3 July 2013
17. Trenberth KE (1997) The definition of El Niño. *Bull Amer Met Soc* 78:2771–2777
18. Bove MC, Elsner JB, Landsea CW et al (1998) Effect of El Niño on U.S. Landfalling hurricanes revisited. *Bull Amer Met Soc* 79(11):2477–2482

19. Yumul JGP, Dimalanta CB, Servando NT et al (2010) The 2009-2010 El Niño Southern oscillation in the context of climate uncertainty: the Philippine setting. *Philip J Sci* 139 (1):119–126
20. Holden WN (2013) Neoliberal mining amid El Niño induced drought in the Philippines. *J Geograp Geolog* 5(1):58–77
21. Roberts MG, Dawe D, Falcon W et al (2009) El Niño-Southern oscillation impacts on rice production in Luzon, the Philippines. *J App Met Clim* 48:1718–1724

Oligostilbenoids from *Vatica* Species and Bioactivities

A.S. Kamarozaman, N.F. Rajab and J. Latip

Abstract Reactive species (RS) which are generated from the pollution, deep fried and spicy foods, leakage of electrons from mitochondrial electron transport chains etc. may result in an oxidative damage in the body. The oxidative damage may lead to various diseases such as Alzheimer, atherosclerosis and cancer. In order to prevent such diseases, antioxidants play important roles in reducing the powerful oxidizing agents. *Vatica* species that belongs to the family of Dipterocarpaceae has been widely known to contain abundant source of oligostilbenoids which demonstrated interesting result in biological activities such as anticancer and antioxidant. This may lead to a development of drugs as well as natural antioxidants. In this chapter, we are highlighting the oligostilbenoids isolated from *Vatica* species from various researcher as well as the biological activities.

1 Introduction

Dipterocarpaceae is a large family comprising of 16 genera and 600 species [1] which are widely distributed in Africa, China, India, Indonesia, Malaysia, Myanmar, New Guinea, Philippines, Sri Lanka and Thailand [2–5]. This family is divided into three subfamilies namely Dipterocarpoideae, Monotoideae and Pakaraimoideae [6]. The first subfamily, Dipterocarpoideae is limited to tropical Asia while the

A.S. Kamarozaman (✉)

Centre of Foundation Studies, MARA University of Technology, Puncak Alam Campus,
42300 Bandar Puncak Alam, Selangor, Malaysia
e-mail: aisyahtsalihah@yahoo.co.uk

N.F. Rajab

School of Diagnostic and Applied Health Sciences, Faculty of Health Sciences, National
University of Malaysia, Jalan Raja Muda Abdul Aziz, 50300 Kuala Lumpur, Malaysia

J. Latip

School of Chemical Sciences and Food Technology, Faculty of Sciences and Technology,
National University of Malaysia, 43600 Bangi, Selangor, Malaysia

second subfamily is indigenous to Africa and the third is confined to tropical America [2]. Dipterocarpoideae which is the largest subfamily in Dipterocarpaceae is divided into four tribes namely *Dipterocarpeae*, *Dryobalanopseae*, *Shoreae* and *Vaticeae* [2, 6, 7]. The genera that belong to this largest subfamily are *Anisoptera*, *Cotylelobium*, *Dipterocarpus*, *Dryobalanops*, *Hopea*, *Neobalanocarpus*, *Parashorea*, *Shorea*, *Upuna*, *Vateria* and *Vatica* [1, 2, 4, 6, 8–10]. The largest genus found in this family is *Shorea* with 150 species followed by *Dipterocarpus* with 75 species [1].

The genus *Vatica* that belongs to the subfamily of Dipterocarpoideae is mainly found in Southeast Asia especially in Borneo and Malay Peninsula with 35 and 29 species were found respectively [4, 10, 11]. The species that have been recorded are *V. bancana*, *V. bella*, *V. cinerea*, *V. cuspidata*, *V. flavida*, *V. havilandii*, *V. heteroptera*, *V. lobata*, *V. lowii*, *V. maingayi*, *V. nitens*, *V. odorata*, *V. pallida*, *V. patula*, *V. pauciflora*, *V. perakensis*, *V. ridleyana*, *V. scortechinii*, *V. stapfiana* and *V. stipulata* [4]. *Vaticas* are usually small trees and sometimes shrubs. The trees can reach up to 160 ft in height and 38 in. in diameter. The species from this genus can be found in the lowland usually below 1,000 ft, hill-forest between 2,000 and 3,000 ft and peat swamp forests [4, 12].

Vaticas are small to medium-sized and rarely large trees [6, 11, 12]. Usually they have straight trunks without large buttresses. The crown is irregular, oblong, sympodial and non-emergent. The barks are usually in grey mottled and occasionally scroll-marked. The leaves are coriaceous, flowers are bisexual and fruits calyx are variables [6, 13]. The sap-wood is very pale and yellowish while the heart-wood is darkening to a dark-brown or reddish-brown after prolonged exposure to the air [12]. Certain species in this genus showed a considerable amount of sticky oil resembling ‘minyak keruing’. This species produced small amounts of greenish resin but it is not worked commercially [12].

There are two kinds of wood found in within the genus namely *resak* and *resak laru* [7, 12]. *Resak* is produced by certain species of *Vatica* and *Cotylelobium*. The wood is very hard, heavy and durable which is suitable for construction and in exposed situations. The species included in this kind of wood are *V. cinerea*, King, *V. curtisii*, King and *V. reticulata*, King. The other kind of wood is known as *resak laru* or *resak paya*. It is a common name applied to any species of *Vatica* which is used as *laru*, an ingredient added into toddy to retard the fermentation [4]. There is a large amount of sap-wood and the heart-wood which is much softer and lighter compared to *resak*. This kind of wood is rarely used as the wood obtained in a small-sizes and not durable. The species with this kind of wood are *V. bancana*, Scheff, *V. bella*, van Slooten, *V. stapfiana*, van Slooten and *V. wallichii*, Dyer [12].

Free radicals or also known as reactive species (RS) are constantly produced in cells through normal metabolic processes [14]. Free radicals are generated due to pollution, radiation, toxins as well as deep fried and spicy foods [15, 16]. Some are also produced by ‘accidents of chemistry’ such as the leakage of electrons from mitochondrial electron transport chains and direct reaction of autoxidisable molecules with molecular oxygen to produce superoxide free radical ($O_2^{\bullet-}$) [17]. Some of the examples of useful purposes from the generated RS are inter- and intra-cellular

of signaling [18] as well as defense against foreign organisms [19]. However, many RS are powerful oxidizing agents which capable in damaging DNA and other biomolecules [20].

Antioxidants are protection agents in reducing oxidative damage of human body from RS including reactive oxygen species (ROS) and retard the progress of many chronic diseases [21]. Thus, the free radicals can be removed from the body through antioxidant defense mechanism [22]. Oxidants or oxidative stress occurs when the balance of RS within the cells exceeds the level of antioxidants present. This situation may lead to the injury or killing of cells [23, 24], atherosclerosis, cancer, stroke, asthma, arthritis and other age related diseases [25, 26].

Hence, to overcome these diseases, many researchers have done studies in producing synthetic antioxidants such as butylated hydroxy anisole (BHA), butylated hydroxy toluene (BHT) and gallic acid esters. However, the synthesized antioxidants have contributed to the liver damage and carcinogenesis [27]. Therefore, studies in searching natural antioxidants have greatly increased to find the antioxidants which are safe to be used [22].

The chemical constituents found in *Vatica* such as oligostilbenoids have drawn much attention to be studied as they possessed various biological activities like antioxidant [28], anticancer [29] as well as chemopreventive [30] agents. In this chapter, we will focus on the oligostilbenoids isolated from the *Vatica* species and their biological activities.

2 Oligostilbenoids from *Vatica*

The family of Dipterocarpaceae is known to contain abundant source of oligostilbenoids [31] which formed through a condensation of resveratrol unit (3,5,4'-trihydroxystilbene) (**1**) [32]. Resveratrol unit is divided into two groups namely group A and B. Group A contains at least one five-membered oxygen heterocyclic ring, usually *trans*-2-aryl-2,3-dihydrobenzofuran (**2**) while Group B does not contain oxygen heterocyclic ring which the stilbene skeleton remain or has been modified [33, 34]. The examples of the oligostilbenoids in Group A are ϵ -viniferin (**3**), α -viniferin (**4**) and miyabenol A (**5**) while Group B are pallidol (**6**) and copalliphol B (**7**).

2.1 Stilbene Monomer

Resveratrol (3,5,4'-trihydroxystilbene) (**1**) which has been isolated from the stem of *V. rassak* is the first report of the occurring stilbene monomer in this family [35]. This stilbene monomer is rarely found in the family of Dipterocarpaceae instead, the constituent occur abundantly in glycoside form either *C*-glucoside or

O-glucoside. (*E*)-resveratrol 3-*O*- β -D-glucopyranoside or also known as piceid (**8**) was isolated from *V. diospyroides* [36], *V. rassak* [35], *V. pauciflora* [31], *V. albiramis* [37] and *V. odorata* [38].

2.2 Stilbene Dimmer

Stilbene dimmer in group A is classified into two subgroups, stilbene dimmer with stilbene skeleton and stilbene dimmer without stilbene skeleton. Stilbene dimmer that contain stilbene skeleton is ϵ -viniferin (**3**), which is a precursor to other oligostilbenoids [39] and was isolated from *V. affinis* [40], *V. rassak* [35], *V. pauciflora* [30, 31], *V. umbonata* [39], *V. albiramis* [37] and *V. odorata* [38]. The isomer, (*Z*)- ϵ -viniferin (**9**) was isolated from *V. albiramis* [37]. Other constituents in this subgroup are paucifloroside A (**10**) and vatalbinoside C (**11**), stilbene dimmers with glycoside isolated from *V. pauciflora* [31] and *V. albiramis* [37] respectively.

Laevifonol (**12**) that was isolated from *V. umbonata* [39] and *V. odorata* [38] is a stilbene dimmer without stilbene skeleton. It was formed by the condensation of ϵ -viniferin (**3**) and ascorbic acid [41]. Other oligostilbenoids fall into this subgroup is hemsleyanol A (**13**) and pauciflorol E (**14**) from *V. pauciflora* [31, 42]; albiraminol B (**15**), ampelopsin A (**16**), balanocarpol (**17**) and malibatol A (**18**) from *V. albiramis*. Stilbene dimmer glycosides in this subgroup are vatalbinosides D–F (**19–21**) from *V. albiramis* [37, 43].

Stilbene dimmer in group B is (+)-ampelopsin D (**22**) that was isolated from *V. pauciflora*. Biosynthesis of this compound might be from (+)- ϵ -viniferin as the stilbene skeleton is remained in the structure. Other constituents in this group are ampelopsin F (**23**) from *V. pauciflora* [31] and *V. umbonata* [39], isoampelopsin F (**24**) from *V. pauciflora* [31] as well as albiraminols C (**25**) and D (**26**) from *V. albiramis*. Vatalbinosides G–J (**27–30**) isolated from *V. albiramis* are stilbene dimmers glycoside in group B [44].

2.3 Stilbene Trimer

Biogenetically, stilbene trimer is formed through an oxidative coupling reaction between resveratrol (**1**) (stilbene monomer) and ϵ -viniferin (**3**) (stilbene dimmer). Stilbene trimers in group A are davidiol B (**31**) and pauciflorols A (**32**) and B (**33**) from *V. pauciflora*; stenophyllol B (**34**) from *V. pauciflora* [31] and *V. umbonata* [39]; vaticanols A (**35**) and E (**36**) from *V. rassak* [35, 45] and *V. pauciflora* [31]. In addition, pauciflorosides B (**37**) and C (**38**) isolated from *V. pauciflora* [31] as well as vaticaside A (**39**) from *V. rassak* [45] are stilbene trimers with glucosyl moiety.

Stilbene trimer without dihydrofuran ring is rarely found in this genus [46]. The constituents from this group are vaticanol G (**40**) isolated from *V. rassak* [47], *V. pauciflora* [30, 31] and *V. umbonata* [39] as well as its glycoside known as vaticaside D (**41**) isolated from *V. rassak* [47] and *V. pauciflora* [31].

2.4 Stilbene Tetramer

Stilbene tetramer is formed through condensation of two stilbene dimmers and classified into two groups. The first group is symmetrical stilbene and the second group is unsymmetrical stilbene. The constituents that fall in the first group are hopeaphenol (**42**) from *V. umbonata* [39], *V. albiramis* [37] and *V. odorata* [38]; hopeaphenol A (**43**) and isohopeaphenol A (**44**) from *V. oblongifolia* [48]; pauciflorol C (**45**) and vateriaphenol B (**46**) from *V. pauciflora* [31]; stenophyllol C (**47**) and vatalbinoside A (**48**) from *V. albiramis* [37] as well as vaticaffinol (**49**) from *V. affinis* [40].

The second group of stilbene tetramer is divided into two subgroups based on the ring in their structures. The first subgroup contains two units of unsymmetrical stilbene dimmers whereby the D unit is connected to the C2 ring through dihydrofuran ring [35]. The constituents in this subgroup are hemsleyanols C (**50**) and D (**51**) and isovaticanol B (**52**) from *V. pauciflora* [31]; vaticaphenol A (**53**) from *V. diospyroides* [36] and *V. oblongifolia* [48]; vaticanol B (**54**) from *V. rassak* [35], *V. pauciflora* [31], *V. umbonata* [39], *V. albiramis* [37] and *V. odorata* [38] as well as vaticanol F (**55**) from *V. rassak*. Furthermore, vaticasides B (**56**) and C (**57**) isolated from *V. rassak* [45] and *V. albiramis* and vatalbinoside B (**58**) from *V. albiramis* [37] are the stilbene tetramers in this subgroup with glucosyl moiety [45].

The second subgroup is stilbene tetramer with the two units of stilbene dimer (ϵ -viniferin) (**3**) combined to form dibenzobicyclo[3.2.1]octadiene system. Stilbene tetramers that belong to this subgroup are isovaticanol C (**59**) from *V. pauciflora* [31]; vatdiospyroidol (**60**) from *V. diospyroides* [36] and vaticanol C (**61**) from *V. rassak* [35], *V. pauciflora* [31] and *V. albiramis* [37].

2.5 Stilbene Hexamer

There are four stilbene hexamers isolated from *Vatica* species namely vaticanols D (**62**), H (**63**) and I (**64**) from *V. rassak* [35, 47, 49] as well as albiraminol A (**65**) from *V. albiramis* [43]. The structures of these constituents are composed from two different units of stilbene trimers.

2.6 Stilbene Heptamer

Stilbene heptamer is formed through oxidative coupling between stilbene tetramer and stilbene trimer. Vaticanol J (**66**) which was isolated from *V. rassak* is formed between vaticanol B (**54**) and a unit of stilbene trimer [47]. Pauciflorol D (**67**) isolated from *V. pauciflora* is formed through oxidative coupling between vaticaphenol A (**53**) (stilbene tetramer) and davidiol B (**31**) (stilbene trimer) [31].

2.7 Stilbene Octamer

Only one stilbene octamer has been isolated from the genus *Vatica*. Vatriaphenol A (**68**) was isolated from the stem of *V. albiramis* [37].

3 Biological Activities

Since 30 years ago, oligostilbenoids have shown a wide range of biological activity such as antibacterial [50–53], anti-HIV [54], anti-inflammatory [55, 56], chemopreventive [30] as well as cytotoxicity [57].

3.1 Chemopreventive

Resveratrol (**1**) modulates diverse biochemical processes such as inhibition of carcinogen activation, cellular proliferation and tumor promotion, blockade of tumor cell cycle progression, stimulation of carcinogen deactivation and induction of programmed cell death [58]. Studies done on the chemopreventive activity by resveratrol (**1**) revealed that the stilbene monomer significantly reduced the number of tumors per mouse in a two stage skin carcinogenesis model [59], inhibited *N*-methyl-*N*-nitrosourea-induced mammary tumorigenesis in female Sprague-Dawley rats [60] as well as inhibited *N*-nitrosomethylbenzylamine (NMBA)-induced esophageal tumorigenesis in F344 rats [61]. Furthermore, resveratrol (**1**) has the potential to inhibit the activation of NF- κ B and/or AP-1 in a cell or tissue-specific manner which may turn-off the genes regulating cell survival and stimulating proliferation as well as turn-on the genes initiating cancer cell death [58].

A stilbene dimer, ϵ -viniferin (**3**) exhibited a potential chemopreventive agent by protecting the Chang liver cells from oxidative damage induced by hydrogen peroxide, H₂O₂. Chang liver cells were treated with ϵ -viniferin (**3**) and vaticanol G (**40**) for 24 h and were exposed to H₂O₂ to induce the oxidative damage for 30 min. The percentage of cell viability before the treatment with ϵ -viniferin is 78.3 % while the viability after the treatment at 50 and 100 μ M is 106.9 and 111.0 % respectively. Meanwhile, the percentage of cells before treatment with vaticanol G (**40**) is 76.6 % while after treatment at 50 and 100 μ M is 85.4 and 84.4 % respectively. The proliferation of Chang liver cells had shown a significant increase after the treatment with ϵ -viniferin (**3**) while the treatment with vaticanol G (**40**) did not show any significant increase. This significant increase in Chang cells showed that ϵ -viniferin (**3**) has the ability to protect the cells from oxidative damage. High antioxidant capacity in ϵ -viniferin (**3**) is one of the factors that contribute to the significant protection of oxidative damage in Chang cells [30].

3.2 Antioxidant Activity

In vivo, resveratrol (**1**) exhibited an increase in plasma antioxidant capacity and a decrease in lipid peroxidation [62, 63] which is associated with coronary heart disease [64].

The antioxidant activity of ϵ -viniferin (**3**) by DPPH showed the IC_{50} value $52.6 \pm 1.1 \mu M$ [65]. Furthermore, ϵ -viniferin (**3**) displayed strong inhibitory against lipid peroxidation by Cu^{2+} -induced and peroxy radical AAPH-induced in HDL with the IC_{50} values 2.4 and 5.7 μM respectively. It indicated that the stilbene dimer has the potential in preventing lipoprotein peroxidation [66].

Studies on antidiabetogenic constituents found that vaticanols A (**35**) and G (**40**) demonstrated significant inhibition while vaticanol E (**36**) tended to inhibit the plasma glucose elevation which showed those compounds having antihyperglycemic effects. Furthermore, vaticanols A (**35**), E (**36**) and G (**40**) also inhibited plasma TG level after olive oil loading in mice, enzyme activity of porcine pancreatic lipase as well as demonstrated DPPH radical- and O_2^- radical scavenging activities. In addition, vaticanols A (**35**) and E (**36**) inhibited the enzyme activities of rat intestinal maltase and sucrose and of rat lens aldose reductase [67].

Meanwhile, hemsleyanol D (**51**) significantly inhibited plasma glucose level at doses of 100–200 mg/kg and moderately inhibited the rat lens aldose reductase. Balanocarpol (**17**) significantly displayed inhibitory effects on plasma glucose level at doses 100–200 mg/kg p.o. and suppressed the gastric emptying [68]. The unsymmetrical stilbene hexamer, vaticanol D (**62**) possessed a scavenging activity of super oxide at IC_{50} 7.4 μM [35].

3.3 Anticancer

Studies on the cytotoxicity of oligostilbenoids against the human promyelocytic leukemia HL-60 showed that resveratrol (**1**), ϵ -viniferin (**3**), vaticanol E (**36**), stenophyllol B (**34**), hopeaphenol (**42**), hemsleyanol D (**51**), vaticaside B (**56**), vaticanols B–D (**54**, **61**, **62**) and H–J (**63**, **64**, **66**) as well as vateriaphenol A (**68**) displayed the IC_{50} values 12.1 ± 0.17 [49], 44.2 ± 3.8 , 34.2 ± 1.9 , 25.0, 21.3 ± 1.9 , 12.3 ± 1.6 , 5.1 ± 0.3 , 4.8 ± 0.3 , 3.0 ± 0.2 , 9.8 ± 1.1 , 13.8 ± 0.8 , 7.8 ± 0.5 , 11.2 ± 0.7 and $9.7 \pm 0.6 \mu M$ respectively [29].

Furthermore, resveratrol (**1**), hopeaphenol (**42**), hemsleyanol D (**51**), vaticanols C (**61**), D (**62**), H–J (**63**, **64**, **66**) and vateriaphenol A (**68**) demonstrated the IC_{50} values 22.1 ± 1.8 , 28.6 ± 3.4 , 32.5 ± 2.1 , 3.2 ± 0.2 , 8.9 ± 0.9 , 15.1 ± 1.3 , 10.9 ± 0.8 , 21.5 ± 1.6 and $9.8 \pm 0.9 \mu M$ against colon cancer SW480 cells respectively. The growth suppression in the colon cancer by vaticanol C (**61**) is due to the apoptosis that was executed by the activation of caspase-3 [29].

In addition, ϵ -viniferin (**3**), hopeaphenol (**42**) and vaticanol B (**54**) exhibited the cytotoxicity against P-388 with the IC_{50} values 18.1 ± 0.7 [69], 3.2 [38] and

$46.4 \pm 1.5 \mu\text{M}$ [69] respectively. A novel stilbene tetramer, vatdiospyroidol (**60**) demonstrated the most potent cytotoxicity against KB (oral epidermoid carcinoma), with the EC_{50} value $1.0 \mu\text{g/mL}$ [36] while hopeaphenol (**42**), vaticanols D (**62**), H–J (**63**, **64**, **66**) and vateriaphenol A (**68**) displayed potent cytotoxicity against KB cells with ED_{50} values at $1.2 \mu\text{g/mL}$ [70], 11.8, 15.4, 11.0, 10.7 and $10.5 \mu\text{M}$ [44] respectively.

Vaticanol C (**61**) which is the unsymmetrical stilbene tetramer with dibenzobicyclo[3.2.1]octadiena system possessed very good biological activity especially in anticancer. This compound was found to induce considerable cytotoxicity in human cancer cell lines such as LNCaP (prostate), SH-SY5Y (neuroblastoma) and leukemia (U937). Vaticanols H (**63**) and I (**64**) showed 50 % inhibitory effect on the growth of prostate (LNCaP) and neuroblastoma (SH-SY5Y) cells when treated at the concentration of $10 \mu\text{M}$ while vaticanols D (**62**), J (**66**) and vateriaphenol A (**68**) displayed 50 % inhibitory effect on the growth of neuroblastoma cell (SH-SY5Y) only [29].

Furthermore, malibatol A (**18**) exhibited the cytotoxicity against host cells (CEM-SS) with the IC_{50} value $13 \mu\text{g/mL}$ [54]. Stenophyllol B (**34**) showed the cytotoxicity against HeLa cell at IC_{50} value $28.2 \mu\text{M}$ [71] while hopeaphenol (**42**) showed the cytotoxicity against human colon cancer (HCT-116; Caco-2) cell with IC_{50} 2–5 μM [72]. Vatdiospyroidol (**60**) demonstrated the most potent cytotoxicity against Col2 (colon cancer) as well as BC1 (breast cancer) with the EC_{50} values 1.9 and $3.8 \mu\text{g/mL}$ respectively [36].

In addition, vaticanol C (**61**) and hopeaphenol (**42**) displayed significant inhibitory effect on MMP-1 production from human dermal fibroblasts when tested at the concentration of $1 \mu\text{M}$ while stenophyllol C (**47**) and vaticanol B (**54**), vatalbinosides A (**48**) and B (**58**) showed blocking of $\text{IL-1}\beta$ -induced MMP-1 production when tested at $1 \mu\text{M}$ [37]. Meanwhile, hemsleyanol C (**50**) exhibited the most potent inhibitor when tested against Topoisomerase II [73].

3.4 Other Activities

Resveratrol (**1**) prevents platelet aggregation in vitro and in vivo [74, 75] and found to be potent in protecting against brain damage following cerebral ischemia [76]. The glucosylated-resveratrol, piceid (**8**) has shown a significant inhibitory effect on platelet aggregation induced by collagen (IC_{50} value of $69 \mu\text{M}$), adrenalin (IC_{50} value of $102 \mu\text{M}$), arachidonic acid (IC_{50} of $149 \mu\text{M}$) and ADP (IC_{50} of $218 \mu\text{M}$) when tested on human platelet-rich plasma [77].

ϵ -viniferin (**3**) which is a precursor for the oligostilbenoids [39] demonstrated a moderate inhibitory effect on murine tyrosinase activity with 24.6 ± 2.1 % inhibitions at $100 \mu\text{M}$ [78] as well as concentration-dependently inhibited the enzymatic activity of commercial (human recombinant) monoamine oxidase (MAO) isoform (MAO-A and MAO-B) activity which suggested that this stilbene dimmer might be used as a structural template in designing and developing new

antidepressant drugs [79]. Not only that, this compound showed the strongest antimutagenic effect against MNNG (*N*-methyl-*N'*-nitro-*N*-nitrosoguanidine) with the IC₅₀ value 27 µg/plate [80]. Meanwhile, the isomer of ε-viniferin (**3**), (*Z*)-ε-viniferin (**9**) displayed TNF-α inhibitory activity with the ratio 51.43 % [81] as well as antimutagenic effect [80]. Ampelopsin A (**16**) which is found to induce down-regulation of sphingosine kinase 1 expression and reduce DNA synthesis is a novel inhibitor of sphingosine kinase 1 activity [82]. Balanocarpol (**17**) demonstrated the inhibitory effects on plasma triglyceride level at a dose of 200 mg/kg p.o. The significant increase in the serum TG level indicated that this constituent has the potential to be used in preventing obesity [83]. This stilbene dimmer is a novel inhibitor against sphingosine kinase 1 activity which induced down-regulation of sphingosine kinase 1 expression, reduced DNA synthesis and stimulated PARP cleavage in breast MCF-7 cells [82]. Balanocarpol (**17**) showed very modest HIV-inhibitory activity with the ED₅₀ value 20 µg/mL [54] while (*Z*)-ε-viniferin (**9**) demonstrated antifungal activity [84].

Vaticanol E (**36**), pauciflorol B (**33**), hopeaphenol A (**43**), pauciflorol C (**45**), vaticaffinol (**49**) and hemsleyanol D (**51**) demonstrated antiviral activity against HSV-1 with IC₅₀ values 12.1 ± 0.5, 16.7 ± 0.5, 17.3 ± 0.5, 24.1 ± 0.6, 17.9 ± 0.3 and 9.1 ± 0.5 µM as well as HSV-2 with IC₅₀ values 4.5 ± 0.1, 3.2 ± 0.5, 3.7 ± 0.2, 3.3 ± 0.1, 3.2 ± 0.3 and 3.8 ± 0.2 µM respectively [85]. Vaticanols A (**35**), B (**54**), E (**36**) and G (**40**) possessed antifeedant activity against subterranean termite, *Reticulitermes speratus* [86]. In addition, vaticanols A (**35**) and G (**40**) exhibited complete inhibition on the growth of *Babesia gibsoni* in vitro at the concentration of 25 µg/mL while vaticanol B (**54**) at 50 µg/mL [87]. Hemsleyanol D (**51**) was found to be the most effective to inhibit the growth of MRSA with an MIC value at 2 µg/mL [88]. Furthermore, hopeaphenol A (**43**) and vaticaphenol A (**53**) demonstrated moderately active against MRSA with MIC values 100 and 50 µg/mL as well as *M. smegmatis* with MIC values 50 and 25 µg/mL respectively [48].

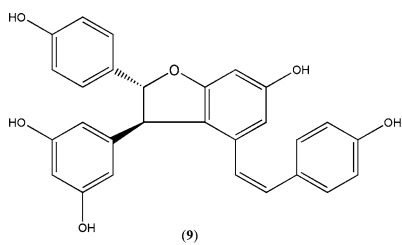
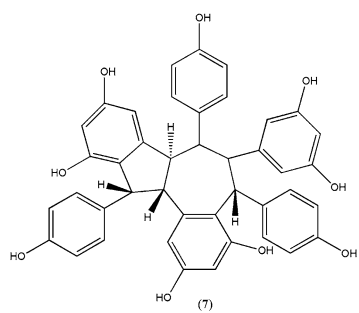
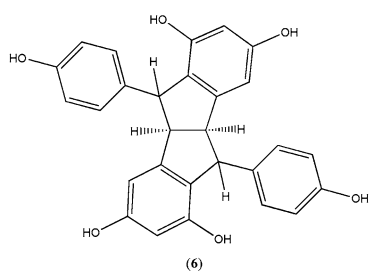
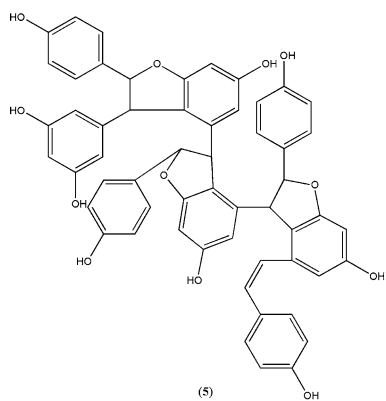
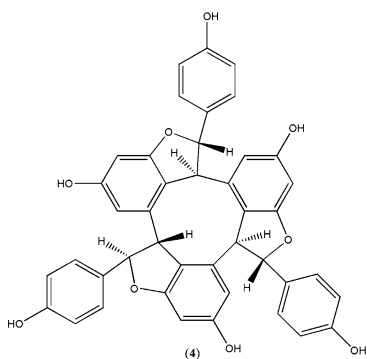
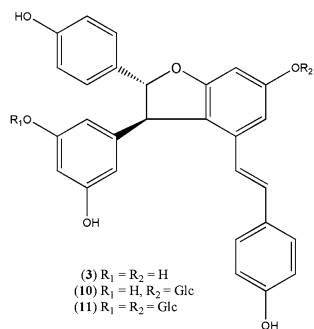
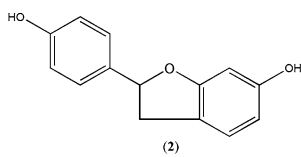
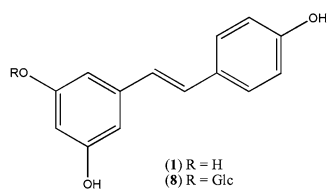
Vaticaffinol (**49**) exerts intensive immunosuppressive activity in vitro and in vivo. Based on this result, the stilbene tetramer may be a novel candidate compound used for T-cell-mediated autoimmune diseases [89].

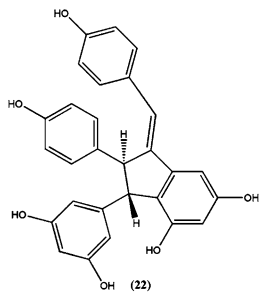
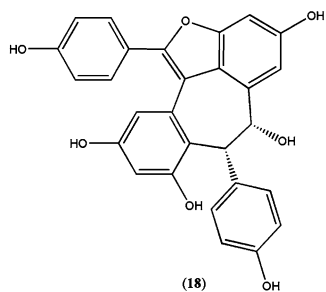
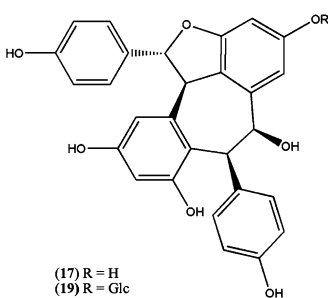
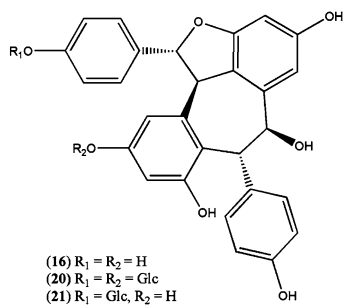
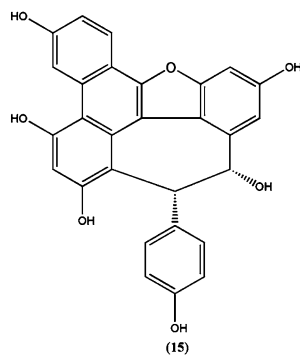
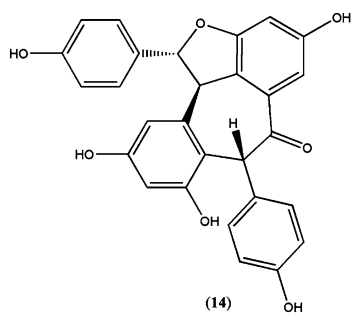
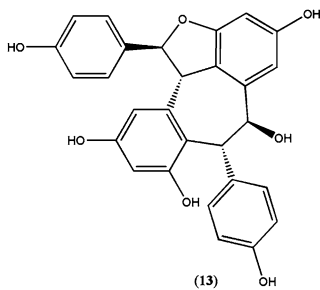
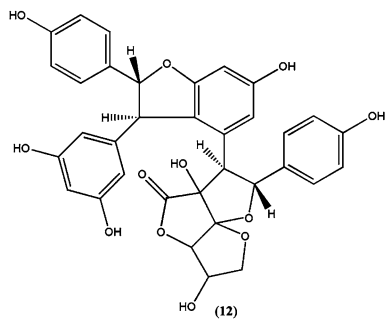
4 Conclusions

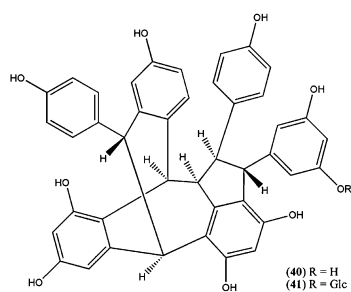
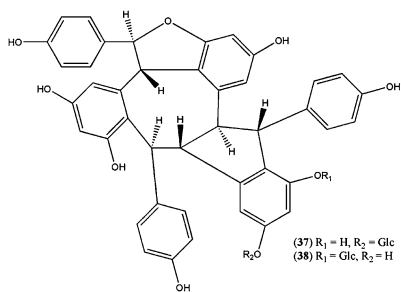
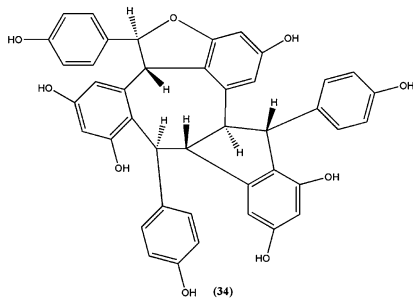
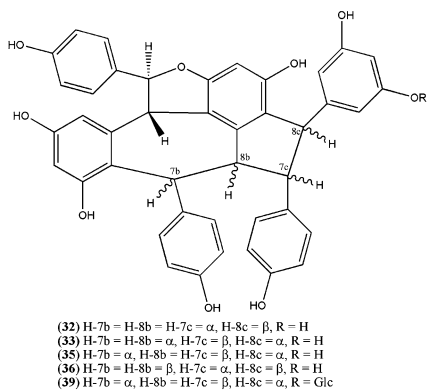
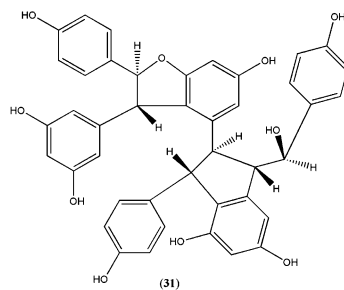
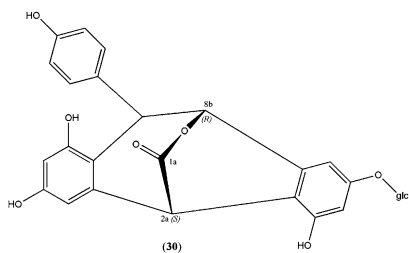
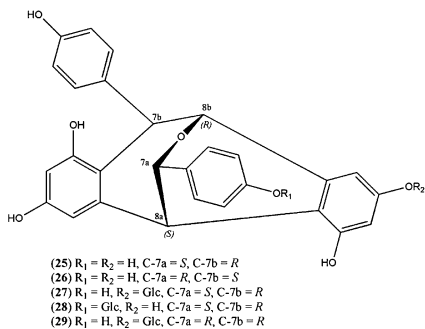
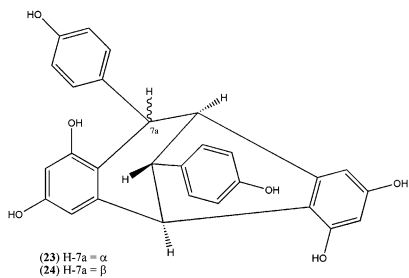
Species from the genus of *Vatica* were found to contain oligostilbenoids with various degree of polymerization, up to stilbene octamer. The oligostilbenoids showed significant biological activities especially in anticancer, antioxidant as well as chemopreventive activities. Perhaps these species can be used as a potential source of natural antioxidant in the future.

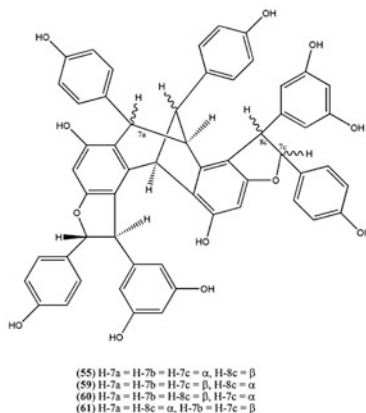
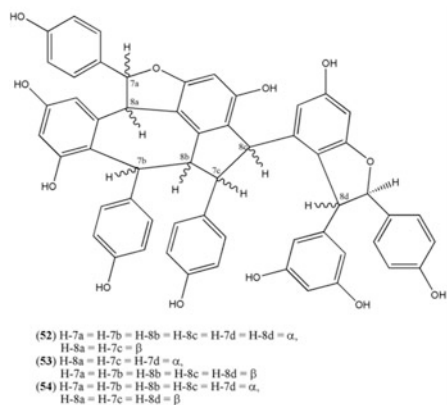
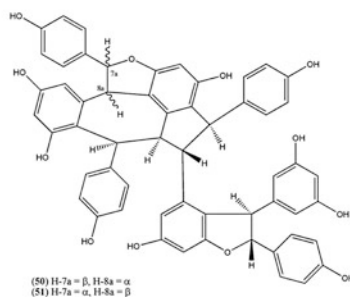
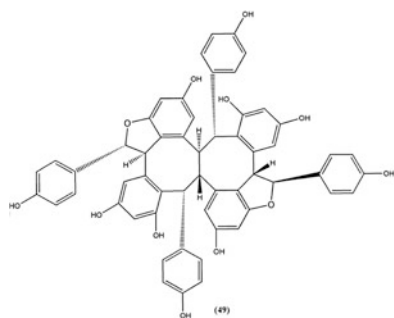
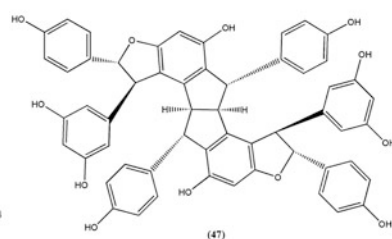
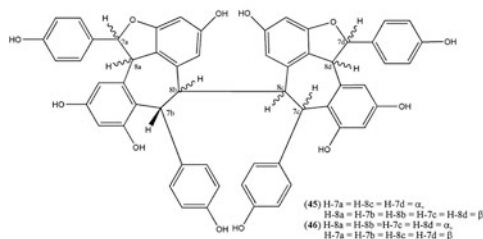
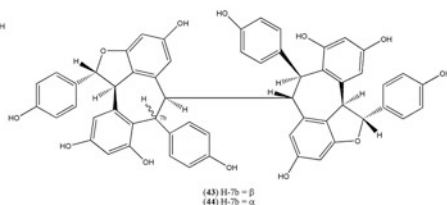
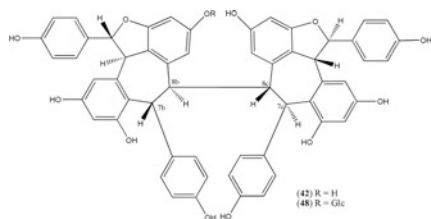
Acknowledgment We would like to express our greatest appreciation to the Ministry of Science, Technology and Innovation, Malaysia for the grant 02-01-02-SF0197 and Centre of Foundation Studies, UiTM Puncak Alam for the financial support.

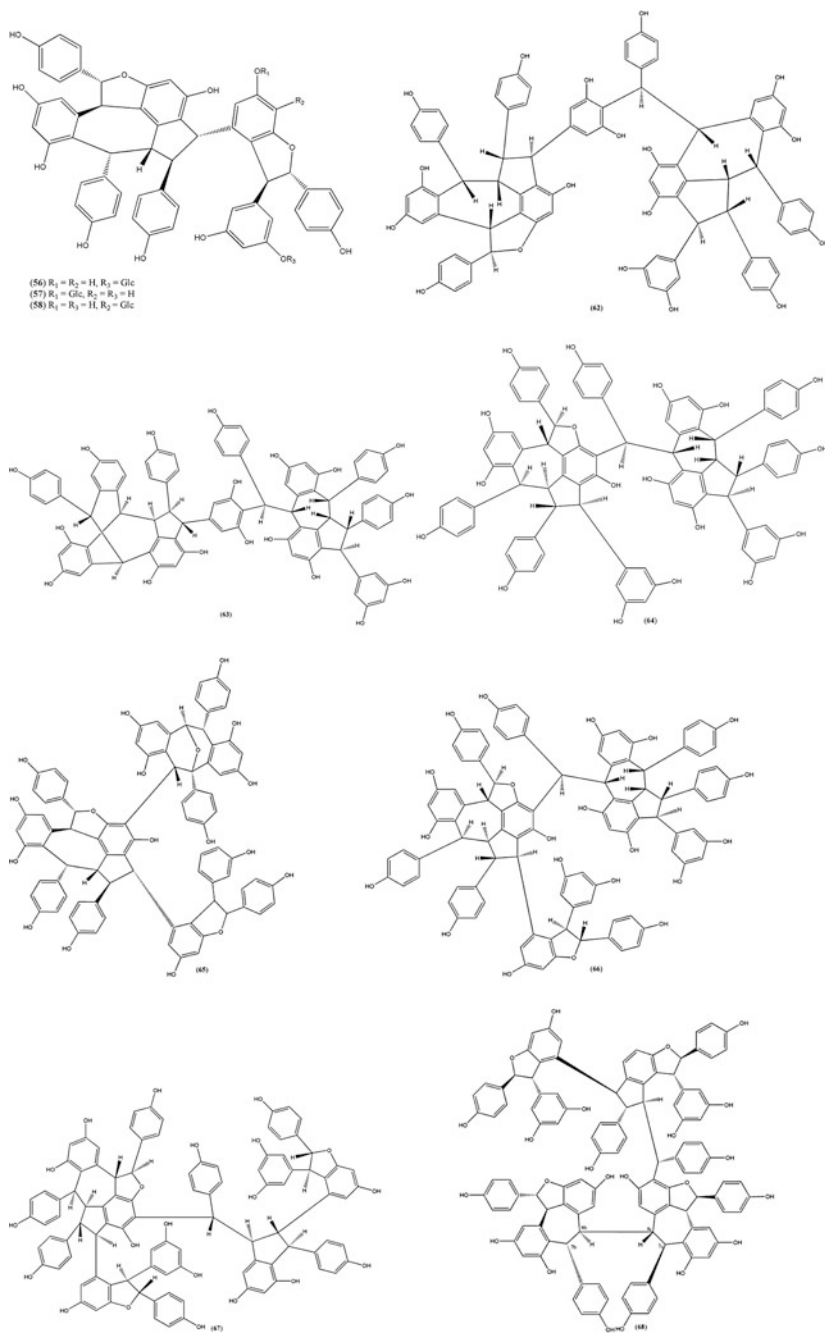
Appendix











References

1. Cronquist A (1981) An integrated system of classification of flowering plants. Columbia University Press, New York
2. Keng H (1983) Malayan seed plants, 3rd edn. Singapore University Press, Singapore
3. Newman MF, Burges PF, Whitmore TC (1999) Pedoman Identifikasi Pohon Dipterocarpaceae Pulau Kalimantan. Prosea, Bogor
4. Symington CF (1974) Foresters' manual of dipterocarps. Universiti Malaya, Kuala Lumpur
5. Usher G (1974) A dictionary of plants used by man, 1st edn. Constable and Company Ltd, London
6. Ashton PS, Arboretum A (1982) Dipterocarpaceae: Flora Malesiana 9, vol 1, 2nd edn. Dr. W. Junk Publishers, Hague
7. Foxworthy FW (1932) Dipterocarpaceae of the Malay Peninsula, Malayan forest records No. 10. Singapore
8. Bentham G (1965) Genera plantarum, vol 1. Stechert-Hafner Service Agency Inc., New York
9. Jong K (1976) Cytology of the Dipterocarpaceae: tropical trees. Academic Press, London
10. Kostermans AJG (1992) A handbook of the Dipterocarpaceae of Sri Lanka. WHT Publications, Colombo
11. Desch HE (1941) Dipterocarp timbers of the Malay Peninsula. In: Malayan forest records No. 14. Caxton Press Ltd, Kuala Lumpur
12. Burkill IH (1966) A dictionary of the economic products of the Malay Peninsula, vol 2. Art Printing Works, Kuala Lumpur
13. Ridley HN (1922) The Flora of the Malay Peninsula, vol 1, 1st edn. L. Reeve & Co. Ltd, London
14. Halliwell B (2001) Role of free radicals in the neurodegenerative diseases: therapeutic implications for antioxidant treatment. *Drugs Aging* 18(9):685–716
15. Halliwell B (1996) Antioxidants in human health and disease. *Ann Rev Nutr* 16:33–50
16. Halliwell B (2002) Effect of diet on cancer development: is oxidative DNA damage a biomarker? *Free Radic Biol Med* 32:968–974
17. Halliwell B, Gutteridge JMC (1999) Free radicals in biology and medicine, 3rd edn. University Press, Oxford
18. Sen CK, Sies H, Baeuerle PA et al (2000) Antioxidant and redox regulation of genes. Academic Press, San Diego
19. Bowie A, O'Neill LAJ (2000) Oxidative stress and nuclear factor- κ B activation. *Biochem Pharmacol* 59:13–23
20. Halliwell B (2007) Oxidative stress and cancer: have we moved forward? *Biochem J* 401:1–11
21. Pryor WA (1991) The antioxidant nutrient and disease prevention—what do we know and what do we need to find out? *Am J Clin Nutr* 53:391–393
22. Agrawal S, Kulkarni GT, Sharma VN (2011) A comparative study on the antioxidant activity of methanolic extracts of *Terminalia paniculata* and *Madhuca longifolia*. *Free Radic Antioxid* 1(4):62–68
23. Sawa T, Ohshima H (2006) Nitritative DNA damage in inflammation and its possible role in carcinogenesis. *Nitric Oxide* 14:91–100
24. Hofseth LJ, Saito S, Hussain SP et al (2003) Nitric oxide-induced cellular stress and p53 activation in chronic inflammation. *Proc Natl Acad Sci USA* 100:143–148
25. Cross CE, Halliwell B, Borish ET et al (1987) Oxygen radicals and human disease. *Ann Intern Med* 107:526–545
26. Diaz MN, Frei B, Vita JA et al (1997) Antioxidants and atherosclerotic heart disease. *N Engl J Med* 337:408–416
27. Wichi HP (1988) Enhanced tumor development by butylated hydroxyanisole (BHA) from the prospective of effect on forestomach and oesophageal squamous epithelium. *Food Chem Toxicol* 26:717–723

28. Nagmoti DM, Khatri DK, Juvekar PR et al (2012) Antioxidant activity and free radical-scavenging potential of *Pithecellobium dulce* Benth seed extracts. *Free Radic Antioxid* 2 (2):37–43
29. Ito T, Akao Y, Yi H et al (2003) Antitumor effect of resveratrol oligomers against human cancer cell lines and the molecular mechanism of apoptosis induced by vaticanol C. *Carcinogenesis* 24(9):1489–1497
30. Kamarozaman AS, Latip J, Syah YM et al (2013) Oligostilbenoids from *Vatica pauciflora* and the oxidative effect on Chang cells. In: *Journal of Physics: Conference Series*, 2013 international conference on science and engineering in mathematics, chemistry and physics (ScieTech 2013), January 2013, vol 423. IOP Publishing, Jakarta, p 012045. doi:[10.1088/1742-6596/423/1/012045](https://doi.org/10.1088/1742-6596/423/1/012045)
31. Ito T, Tanaka T, Iinuma M et al (2003) New resveratrol oligomers in the stem bark of *Vatica pauciflora*. *Tetrahedron* 59(28):5347–5363
32. Gorham J, Tori M, Asakawa Y (1995) *The biochemistry of the stilbenoids*. Chapman & Hall, London
33. Huang K-S, Lin M, Yu LN et al (2000) Four novel oligostilbenes from the roots of *Vitis amurensis*. *Tetrahedron* 56:1321–1329
34. Sotheeswaran S, Pasupathy V (1993) Distribution of resveratrol oligomer in plants. *Phytochemistry* 32(5):1983–1992
35. Tanaka T, Ito T, Nakaya K et al (2000) Oligostilbenoids in stem bark of *Vatica rassak*. *Phytochemistry* 54:63–69
36. Seo E-K, Chai H, Constant HL et al (1999) Resveratrol tetramers from *Vatica diospyroides*. *J Org Chem* 64:6976–6983
37. Abe N, Ito T, Ohguchi K et al (2010) Resveratrol oligomers from *Vatica albiramis*. *J Nat Prod* 73:1499–1506
38. Latip J, Zain WZWM, Ahmat N et al (2011) Cytotoxic oligostilbenoids from *Vatica odorata*. *Aus J Basic Appl Sci* 5(6):113–118
39. Atun S, Sjamsul AA, Emilio LG et al (2004) Oligostilbenoids from *Vatica umbonata* (Dipterocarpaceae). *Biochem Syst Ecol* 32:1051–1053
40. Sultanbawa MUS, Surendrakumar S, Wazeer MIM et al (1981) Novel resveratrol tetramer, vaticaffinol, from *Vatica affinis* Thw (Dipterocarpaceae). *J Chem Soc Chem Commun* 16:1204–1206
41. Hirano Y, Kondo R, Sakai K (2003) Novel stilbenoids isolated from the heartwood of *Shorea laevis*. *J Wood Sci* 49:53–58
42. Ito T, Tanaka T, Iinuma M et al (2004) Three new resveratrol oligomers from the stem bark of *Vatica pauciflora*. *J Nat Prod* 67(6):932–937
43. Abe N, Ito T, Oyama M et al (2011) Resveratrol derivatives from *Vatica albiramis*. *Chem Pharm Bull* 59(4):452–457
44. Abe N, Ito T, Oyama M et al (2011) Resveratrol dimers with an oxabicyclo ring in *Vatica albiramis*. *Heterocycles* 83(3):571–580
45. Ito T, Tanaka T, Ido Y et al (2001) Five new oligostilbenes with one or two dihydrofurans from the stem bark of *Vatica rassak*. *Heterocycles* 55(3):557–567
46. Ito T, Ali Z, Furusawa M et al (2005) Two novel trimeric resveratrol derivatives from *Cotylelobium lanceolatum*. *Chem Biodivers* 2:1200–1216
47. Ito T, Tanaka T, Nakaya K-I et al (2001) A novel bridged stilbenoid trimer and four highly condensed stilbenoid oligomers in *Vatica rassak*. *Tetrahedron* 57:7309–7321
48. Zgoda-Pols JR, Freyer AJ, Killmer LB et al (2002) Antimicrobial resveratrol tetramers from the stem bark of *Vatica oblongifolia* ssp. *oblongifolia*. *J Nat Prod* 65:1554–1559
49. Tanaka T, Ito T, Nakaya K et al (2000) Vaticanol D, a novel resveratrol hexamer isolated from *Vatica rassak*. *Tetrahedron Lett* 41:7929–7932
50. Chichewicz RH, Kouzi SA (2002) Resveratrol oligomers: structure, chemistry and biological activity. *Stud Nat Prod Chem* 26:507–579
51. Seo E-K, Douglas KA (2000) Bioactive constituents of the family Dipterocarpaceae. *Stud Nat Prod Chem* 23:531–561

52. Sultanbawa MUS, Surendrakumar S, Bladon P (1987) Distichol an antibacterial polyphenol from *Shorea disticha*. *Phytochemistry* 26(3):799–801
53. Zain WZWM, Ahmat N, Norizan NH et al (2011) The evaluation of antioxidant, antibacterial and structural identification activity of trimer resveratrol from Malaysia's Dipterocarpaceae. *Aus J Basic Appl Sci* 5(5):926–929
54. Dai JR, Hallock YF, Cardellina JH et al (1998) HIV-inhibitory and cytotoxic oligostilbenoids from the leaves of *Hopea malibato*. *J Nat Prod* 61:351–353
55. Huang K-S, Lin M, Cheng GF (2001) Anti-inflammatory tetramers of resveratrol from the roots of *Vitis amurensis* and the conformations of the seven-membered ring in some oligostilbenes. *Phytochemistry* 58:357–362
56. Kitanaka S, Ikezawa T, Yasukawa K et al (1990) (+)- α -viniferin, an anti-inflammatory compound from *Caragana chamlagu* root. *Chem Pharm Bull* 38:432–435
57. Haryoto S, Lia DJ, Yana MS et al (2008) Oligostilbenoids from *Shorea gibbosa* and their cytotoxic properties against P-388 cells. *J Nat Med* 62:195–198
58. Kundu JK, Surh YJ (2004) Molecular basis of chemoprevention by resveratrol: NF- κ B and AP-1 as potential targets. *Mutat Res* 555:65–80
59. Jang M, Cai L, Udeani GO et al (1997) Cancer chemopreventive activity of resveratrol, a natural product derived from grapes. *Science* 275:218–220
60. Bhat KPL, Lantvit D, Christov K et al (2001) Estrogenic and antiestrogenic properties of resveratrol in mammary tumor models. *Cancer Res* 61:7456–7463
61. Li ZG, Hong T, Shimada Y et al (2002) Suppression of *N*-nitrosomethylbenzylamine (NMBA)-induced esophageal tumorigenesis in F344 rats by resveratrol. *Carcinogenesis* 23:1531–1536
62. Wenzel E, Somoza V (2005) Metabolism and bioavailability of transresveratrol. *Mol Nutr Food Res* 49:472–481
63. Whitehead TP, Robinson D, Allaway S et al (1995) Effect of red wine ingestion on the antioxidant capacity of serum. *Clin Chem* 41(1):32–35
64. Holvoet P (2004) Oxidized LDL and coronary heart disease. *Acta Cardiol* 59(5):479–484
65. Liu W-B, Hu L, Hu Q et al (2013) New resveratrol oligomer derivatives from the roots of *Rheum lhasaense*. *Molecules* 18:7093–7102. doi:[10.3390/molecules18067093](https://doi.org/10.3390/molecules18067093)
66. Ngoc TM, Hung TM, Thuong PT et al (2008) Inhibition of human low density lipoprotein and high density lipoprotein oxidation by oligostilbenes from *Rhubarb*. *Biol Pharm Bull* 31(9):1809–1812
67. Matsuda H, Asao Y, Nakamura S et al (2009) Antidiabetogenic constituents from the Thai traditional medicine *Cotylelobium melanoxylon*. *Chem Pharm Bull* 57(5):487–494
68. Morikawa T, Chaipetch S, Matsuda H et al (2012) Antidiabetogenic oligostilbenoids and 3-ethyl-4-phenyl-3,4-dihydroisocoumarins from the bark of *Shorea roxburghii*. *Bioorg Med Chem* 20:832–840
69. Muhtadi Euis HH, Lia DJ et al (2006) Cytotoxic resveratrol oligomers from the tree bark of *Dipterocarpus hasseltii*. *Fitoterapia* 77:550–555
70. Ohyama M, Tanaka T, Ito T et al (1999) Antitumor agent 200.1 Cytotoxicity of naturally occurring resveratrol oligomers and their acetate derivatives. *Bioorg Med Chem Lett* 9:3057–3060
71. Nazri NAAM, Ahmat N, Abdullah M et al (2012) Antioxidant, antimicrobial and cytotoxic activities of resveratrol oligomers of *Shorea macroptera* Dyer. *Aus J Basic Appl Sci* 6(8):431–436
72. Niesen D, Gonzalez-Sarrias A, Yuan T et al (2012) Bioassay-guided isolation of cytotoxic constituents from *Carex vulpinoidea* seeds. In: 44th ACS national meeting and exposition, Philadelphia, PA, USA, August 2012
73. Yamada M, Hayashi K-I, Ikeda S et al (2006) Inhibitory activity of plant stilbene oligomers against DNA Topoisomerase II. *Biol Pharm Bull* 29(7):1504–1507
74. Bertelli AAE, Giovannini L, Giannessi D et al (1995) Antiplatelet activity of synthetic and natural resveratrol in red wine. *Int J Tissue React* 17(1):1–3

75. Wang Z, Huang Y, Zou J et al (2002) Effects of red wine and wine polyphenol resveratrol on platelet aggregation *in vivo* and *in vitro*. *Int J Mol Med* 9(1):77–79
76. Dong W, Li N, Gao D et al (2008) Resveratrol attenuates ischemic brain damage in the delayed phase after stroke and induces messenger RNA and protein express for angiogenic factors. *J Vasc Surg* 48(3):709–714
77. Orsini F, Pelizzoni F, Verotta L et al (1997) Isolation, synthesis and antiplatelet aggregation activity of resveratrol 3-*O*- β -D-glucopyranoside and related compounds. *J Nat Prod* 60:1082–1087
78. Ohguchi K, Tanaka T, Ito T et al (2003) Inhibitory effects of resveratrol derivatives from Dipterocarpaceae plants on tyrosinase activity. *Biosci Biotechnol Biochem* 67(7):1587–1589
79. Yanez M, Fraiz N, Cano E et al (2006) (-)-Trans-epsilon-viniferin, a polyphenol present in wines, is an inhibitor of noradrenaline and 5-hydroxytryptamine uptake and of monoamine oxidase activity. *Eur J Pharmacol* 542(1–3):54–60
80. Kim HJ, Chang EJ, Bae SJ et al (2002) Cytotoxic and antimutagenic stilbenes from seeds of *Paeonia lactiflora*. *Arch Pharmacol Res* 25:293–299
81. Yao C-S, Lin M, Wang Y-H et al (2004) Synthesis of the active stilbenoids by photooxidation reaction of trans- ϵ -viniferin. *Chin J Chem* 22(11):1350–1355
82. Lim KG, Gray AI, Pyne S et al (2012) Resveratrol dimers are novel sphingosine kinase 1 inhibitors and affect sphingosine kinase 1 expression and cancer cell growth and survival. *Br J Pharmacol* 166(5):1605–1616
83. Morikawa T, Chaipetch S, Matsuda H et al (2012) Anti-hyperlipidemic constituents from the bark of *Shorea roxburghii*. *J Nat Med* 66:516–524. doi:[10.1007/s11418-011-0619-6](https://doi.org/10.1007/s11418-011-0619-6)
84. Bala AEA, Kollmann A, Ducrot PH et al (2000) *Cis*- ϵ -viniferin: a new antifungal resveratrol dehydrodimer from *Cyphostemma crotalarioides* roots. *J Phytopathology* 148:29–32
85. Chen X, Qiao H, Liu T et al (2012) Inhibition of herpes simplex virus infection by oligomeric stilbenoids through ROS generation. *Antiviral Res* 95:30–36
86. Fukumoto H, Morimoto M, Fukuda Y et al (2007) Antifeedants in tropical Asian species, resak, against a subterranean termite, *Reticulitermes speratus* (Kolbe). *Kinki Daigaku Nogakubu Kiyo* 40:31–38
87. Subeki Nomura S, Matsuura H et al (2005) Anti-babesial activity of some Central Kalimantan plant extracts and active oligostilbenoids from *Shorea balangeran*. *Planta Med* 71(5):420–423
88. Nitta T, Arai T, Takamatsu H et al (2002) Antibacterial activity of extracts prepared from tropical and subtropical plants on Methicillin-Resistant *Staphylococcus aureus*. *J Health Sci* 48(3):273–276
89. Feng L-L, Wu X-F, Liu H-L et al (2013) Vaticaffinol, a resveratrol tetramer, exerts more preferable immunosuppressive activity than its precursor *in vitro* and *in vivo* through multiple aspects against activated T lymphocytes. *Toxicol Appl Pharmacol* 267:167–173

Investigation on Thermoacoustic Cooling Device with Variation in Stack Plate Size and Input Acoustic Energy

Nandy Putra and Dinni Agustina

Abstract A loudspeaker-driven thermoacoustic cooler device has been designed, constructed and tested to gain better understanding of its cooling performance. The influence of stack plate size to its performance was investigated. The plate was made of acrylic sheet in three different length variations, which were 6, 5 and 4. Each variation of experiment was conducted by varying plate thickness of the stack, 0.15, 0.5 and 1 mm, respectively. The experiments were conducted with various driver voltage input starting from setting 4–9 (Voltage peak-to-peak). The temperatures at the area of both ends of the parallel plate stack, which are cold side and hot side, were recorded. The results showed that thermoacoustic cooling effect occurred immediately and escalated rapidly in 2 min and showed a stable cooling temperature after 10 min. The experimental results confirmed that better thermal performance of the device and faster cooling rate yielded from higher voltage input. For each set of experiment, the input voltage setting, the operating frequency and other parameter of the stack were maintained the same. The thermal performance and cooling rate increased with the decrease of plate thickness. The largest temperature difference, 14.7 °C, was achieved with 0.15 mm plate thickness with 6 cm length at voltage setting 9. The experimental results showed that the effects of using different plate length were not the same for each thickness of stack plate. However, Stack plate size of 0.5 mm thickness and 6 cm length at the input voltage setting of 9, was arguably the optimum size in terms of consistent performance in cooling.

N. Putra (✉)

Applied Heat Transfer Laboratory, Department of Mechanical Engineering,
University of Indonesia, Kampus UI, Depok 16426, Indonesia
e-mail: nandyputra@eng.ui.ac.id

D. Agustina

Department of Mechanical Engineering, Syiah Kuala University, Darussalam,
Banda Aceh, Indonesia

1 Introduction

Refrigeration has been playing important role in almost every aspect of life to the level that living without it is impossible. It is even rated by the National Academy of Engineering as the 10th position in “Greatest engineering achievement in the 20th century” [1]. Despite the advantages, refrigeration’s high demand of energy and the use of harmful refrigerants have been some of major concerns in the obligatory need of greener technologies in the future. Therefore, the focus of efforts is not only the engineering aspects of aimed at improving engine efficiency but also reducing the use of fossil fuels and chlorofluorocarbons/hydrochlorofluorocarbons (HFCs/HCFCs) as refrigerant. The hazard rate of HFC/HCFC to the atmospheric layer is about 2,000–3,000 times greater than carbon dioxide [2].

To answer the challenges of the environmental impact of the use of conventional refrigerants, one alternative developed is thermoacoustic technology devices which convert heat energy to sound energy and vice versa. The former is known as thermoacoustic engines or prime movers and the latter is called thermoacoustic heat pumps or refrigerators. This alternative is an environmentally friendly because it uses harmless working gas. In addition, this machine can operate with relatively low maintenance costs are because of the lack of moving parts and can take advantage of high amplitude sound waves and gas waste heat as the driving force [3]. The two main reasons above make thermoacoustic cooling technology more compelling to investigate.

Since 1980s, various designs of thermoacoustic heat engines have been built and tested in order to increase their performance and reliability [4] and large thermoacoustic heat engines was introduced in 1992 [5]. The same endeavour on development of heat pumps (refrigerators) has also made a launch since 1986 [6]. In 1993, the first spacecraft thermoacoustic cryocooler was developed and used in the space shuttle [7] and a breaking through temperature of -65°C at cold side of a loudspeaker-driven thermoacoustic refrigerator was achieved in 2002 [8] by applying Rott’s theories [9] and the modification of the theory by Hoffler and Swift [3, 10]. Noted that, these devices above are attractive for safer operation using air or inert gas and simplified manufacturing with much less moving parts in the system [11].

Since then, a great number of attention in the existing literature that worked in an attempt to build and increase the performance of the refrigerator. Three main aspects to concern in designing a thermoacoustic refrigerator are included but not limited to the operating condition, working fluid and the geometry of the refrigerator parts [12]. Higher performance and efficiency were aimed by experimentally lowering the Prandtl number of the working fluid [13]. Numerical efforts were also developed by increasing the ratio of acoustic pressure to mean pressure (drive ratio) [14] and simulating the effects of blockage ratio, the ratio of plate thickness and plate spacing [15]. Building an empirical equation for resonance frequency of the device which often shifts from its theoretical number was also attempted [16].

The heart of a thermoacoustic refrigerator acted as a compressor in a conventional refrigerator is called a stack. The dimension of the stack plays an important role due to heat transfer occurs around it have great impact on the thermal performance of the device. Amplitude of sound wave input (in terms of acoustic energy) and drive ratio will determine the cooling capacity and the efficiency that can be achieved. Therefore, these three will be one of the first things to approach in constructing a thermoacoustic refrigerator.

So far, the application of thermoacoustic principles can be used for heat engine (or power plant), heat pumps (or refrigerator) or a combination of both. Some examples of applications that have been and are being developed are as follows:

- Liquefaction of natural gas. Acoustic energy to be used to pump heat to liquefy natural gas [17].
- Chip cooling. Piezo-electric element produces sound waves. A heat pump cools the chip termoakustik work [18].
- Electronics cooling on the ship. Loudspeakers produce sound waves and heat pump cools termoakustik electronic devices [7].
- Electricity from concentrated solar energy generating acoustic waves in a heated termoakustik devices. A linear motor generates electricity from these products [19].
- Cogeneration. Thermoacoustic heat engines generate acoustic energy from waste heat and linear motors convert acoustical energy into electricity.

2 Keypoints in Constructing the Thermoacoustic Refrigerator

The schematic presentation of $\lambda/4$ wavelength thermoacoustic refrigerator is shown in Fig. 1. It operates by using high amplitude sound waves and inert gases to produce cooling. The frequency of the driver, such as loudspeaker, and the

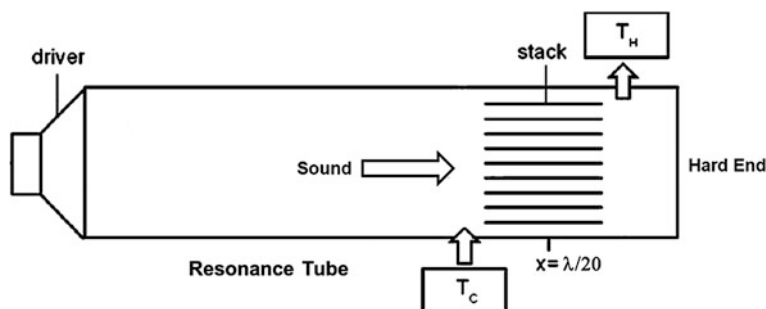


Fig. 1 The schematic of $\lambda/4$ wavelength thermoacoustic refrigerator

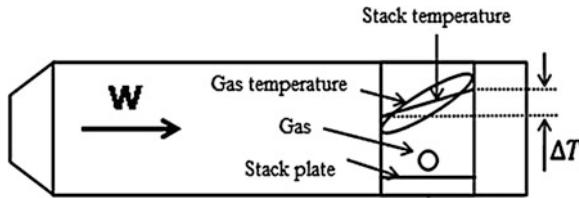


Fig. 2 The theoretical variations of stack and gas temperatures, and the temperatures at different locations inside and outside the stack, modified [20]

geometry of the resonator are calculated so as to get a standing sound wave in the resonator. A porous component called a stack is placed in the tube in such a way that a temperature difference due to thermoacoustic effect is created along the stack. One end of the stack starts to heat up while other end starts to cool down. By controlling temperature of hot side of stack (by means of a heat exchanger), the cold end of stack can provide lower temperature. There are five main components in a thermoacoustic refrigerator which are a resonator, a stack, a driver, working fluid and heat exchangers.

When there is no heat exchanger attached to the device, the fundamental study will be simplified. Such device is called thermoacoustic couples [4]. The theoretical variations of stack and gas temperatures is shown qualitatively in Fig. 2. ΔT is the difference in temperature of cold end and hot end of the stack. Acoustic power, W , is given by the speaker (driver) in the form of high amplitude sound waves. The length of the stack is assumed to be much less than the wavelength of the acoustic wave.

Main parts of the thermoacoustic couple including a sound generating device, resonance tubes, and stack. Besides greatly depending on the number of gas particles involved in the process, the system performance is also influenced by four main operational parameters (the drive ratio to excite thermoacoustic, operating frequency/resonance), the working fluid (Prandtl number), the geometry (blockage ratio, the diameter of the resonator, length of stack, stack position) and the material of stack and resonator [12].

2.1 Relation of the Resonance Frequency and Resonator

The design of the resonator tube having a major role to improve the efficiency of the thermoacoustic cooling device. Initial parameters are calculated based on the length of the resonator wavelength, λ . The length of resonator can be calculated by $\lambda/2$ or $\lambda/4$. The easiest way to decrease the acoustic power losses is to decrease the surface area of resonator wall; therefore the $\lambda/4$ wavelength resonator was chosen [3]. Theoretically, the resonator length, $L = \lambda/4$, giving ease of handling and dissipate only half the energy compared to the tube $\lambda/2$. Figure 3 shows some geometry of a resonator and resonator sections.

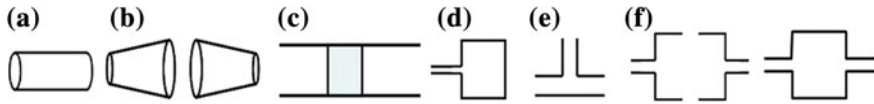


Fig. 3 Geometry of various resonators [21] **a** cylindrical or rectangular tube; **b** tapered tube; **c** capillary or tubule (regenerator); **d** Helmholtz resonator; **e** side tube; **f** reducer pipe

A smooth, linear cylindrical resonator without steps, misalignments and abrupt transition should be used to avoid unwanted eddying or non-linear pressure variations that would greatly complicated the analysis [17]. The length and diameter of resonator are 40 and 5.3 cm, respectively. The resonance frequency of sound wave is mainly determined by the length and the diameter of resonator [22].

$$f_n = \frac{nv}{4\left(L + \frac{14D}{23\pi}\right)}. \quad (1)$$

However, these theories must meet the result of the frequency test of the thermoacoustic device which was performed by inserting a condenser microphone into the resonator and connected it to the oscilloscope to get a relation between the frequency and its connecting voltage. The one with the highest voltage result was the resonance frequency which is best operating frequency for the device.

The material of resonator should be low in thermal conductivity and must endure vibration and pressure to the certain level. Acrylic tube was chosen because of these characteristics and its transparency help in order to arrange the stack and thermocouples and visualize them as well inside the resonator.

2.2 Driver

Thermoacoustic system is a thermodynamic cycle that is theoretically reversible. So that the driver of the heat pump can be either mechanical gas displacement system (such as speakers) or a heat engine. Loudspeakers are relatively easy to apply while the heat engine has a higher efficiency and has no moving parts [15].

2.3 Stack

The dimension of the stack should be around 10 % of the resonator volume and the spacing between the plates is determined by the thermal penetration depth of the stack which is 2–4 times the value of it [8]. In this experiment, acrylic sheet was cut into three different length which are 4, 5 and 6 cm with the wide adjusted to the inside diameter of the resonator. When the distance between the gas and the wall of

the stack is much larger than the thermal penetration depth, there will be an adiabatic heat transfer between the gas and the stack wall. Proposed optimum width is 2–4 δ_k [4]. 4 δ_k spacing was due to ease of fabrication [3], the distance 3 δ_k obtained from experimental results [8] while the distance 2.2 δ_k based on theoretical calculations and numerical [23]. The space between two plates ($2y_o$) is 0.55 mm (3 δ_k).

$$\delta_k = \sqrt{\frac{k}{\pi \rho f c_p}}. \quad (2)$$

Low thermal conductivity and relatively large heat capacity is among the characteristics of the stack material to minimize heat transfer along the x-axis and maximize heat storing while phasing the sound wave in order to produce cooling effect [13]. Widely available acrylic sheet was chosen to meet these criterions. The variations of stack plate thickness ($2l_o$) are 0.15, 0.5 and 1 mm.

Besides δ_k , stack performance is also determined by its location in the resonator. Stack must be close to the antinode pressure to obtain a higher temperature variation in each packet due to higher gas pressure amplitude in the region towards the pressure antinode [24]. The stack was located at the distance of $\lambda/20$ from the closed end of the resonator [8] as this position will bring the optimum performance to the device because a well-balanced of pressure and velocity of the sound wave is achieved [5].

Figure 4a shows a cross sectional view of parallel plate stack, the blockage ratio is defined as the ratio of area available to gas in the stack to the total area of the stack. It is expressed as

$$BR = \frac{2y_o}{2y_o + 2l_o}. \quad (3)$$

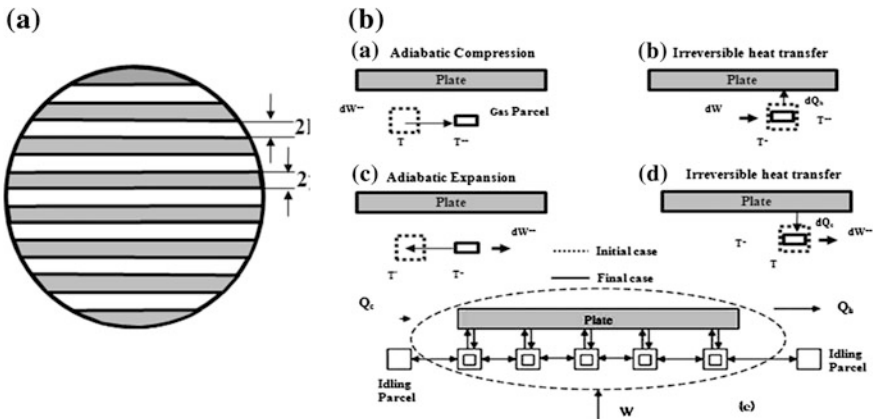


Fig. 4 a A cross sectional view of the stack. b Heat transfer between plate and gas parcels [25]

Figure 4b illustrates the gas parcels around the stack which experience displacement and temperature oscillation in association with the pressure variations with discrete steps as follows:

- When the gas is compressed, the temperature and the pressure will rise to T_{++} and P_{++} , the heat will move from gas to the stack.
- After the heat transfer, the gas temperature and pressure will drop to T_{+} and pressure will remain P_{++} . Gas will be expanded and moved back to its initial position with decreasing temperature.
- Gas temperature and pressure back to their initial value of T_{-} and P . Stack temperature is higher than the temperature of the gas.
- Heat moves from the stack to the gas so that the gas pressure and temperature back to the initial state.

This phenomena which is induced by sound energy continues and generates temperature difference along the stack. Occurrence of the temperature difference will gradually end if the critical temperature gradient (∇T_{crit}) across the stack is equal to the local temperature gradient [17], where

$$\nabla T_{crit} = \frac{p}{\xi \rho c_p}. \quad (4)$$

p is the acoustic pressure and ξ is the acoustic displacement amplitude. This temperature is important in determining the properties of a thermoacoustic device, since efficiency depends on a temperature differential caused by the sound waves that is larger than the critical temperature so that a large cooling effect is created [8].

2.4 Drive Ratio

Variation of the drive ratio was done by adjusting the input voltage to the loudspeaker. Voltage variations made to vary the amplitude of the sound wave which is the refrigerator's driver. The greater the input voltage, the higher the pressure (amplitude) of sound waves. Peak-to-peak voltage were varied ranging from 1 to 9 Vpp on the function generator. Confirmation of waveforms, Vpp value and the RMS value of the input voltage is done by connecting a microphone on the end of the resonator to an oscilloscope.

2.5 Working Gas

The thermal penetration depth in designing stack also put into consideration the properties of working gas, which in this experiment was air. Air was chosen mainly because of its availability and its Prandtl number falls into the acceptable category of working gas for thermoacoustic device [13].

3 Procedure of the Experiments

Components of the refrigerator was fabricated or chosen based on the literatures. The resonator was built from a 40 cm straight acrylic tube. The internal diameter of the tube is 5.3 and the wall thickness is 3 mm. The length of the resonator was set equal to quarter of wavelength of the acoustic wave based on (2). Air at atmospheric pressure was used as a working fluid, as mentioned in Sect. 2.

An 8 Ω loudspeaker with the maximum power of 30 W was used as acoustic driver and installed at one end of the resonator. Another end was closed tightly using the same material as the resonator. After determining the operating frequency experimentally, the square wave at frequency 180 Hz was set in built-in function generator of a National Instrument Elvis II and the sound wave was amplified.

The length and the thickness of each stack plate were 6 cm and 1 mm, respectively. The stack was made of acrylic sheet and cut manually to fit into a cylindrical stack holder. The spacing between the plates was realized by gluing 0.55 mm fishing line on to the surface of each plate. Then the plates were arranged and glued in a stack holder made of the same material as the plates. The stack holder was inserted into the resonator. The fabrication of the stack was done in the same procedures for other sets of experiments with plate thicknesses of 0.5 and 0.15 mm and plate lengths of 5 and 4 cm, respectively.

Five type K thermocouples with accuracy of 0.05 were used; one for ambient temperature and two of those were inserted at different position at each end area of the stack. The temperature data from all thermocouples were acquired simultaneously every second by using National Instrument data acquisition module 9211. Both thermocouples at hot side read the same temperature while those at the cold side showed difference in reading of 1–2 °C. The lowest reading was featured on the graphs.

3.1 Experimental Set-Up

Experimental procedures should meet several requirements to better ensure the accuracy of the data recorded. Before carrying out experiments on the thermal performance of the thermoacoustic cooling device, some things to consider such as the time of experiment, setting the ambient temperature, thermocouple placement as well as the suitability of the basic theory is applied in the experimental set-up.

Besides the necessity of accurate data, precision is crucial in experiment so that any variation of the data should be repeated under the same conditions to record the data. Repetition also needs to be done in the opposite direction of variation.

Figure 5 shows the experimental setup of the thermoacoustic refrigerator constructed at the Heat Transfer Laboratory, University of Indonesia. The stack was placed at 6 cm from the closed end. The frequency and wave amplitudes were measured by a condenser microphone which was located at pressure anti node of

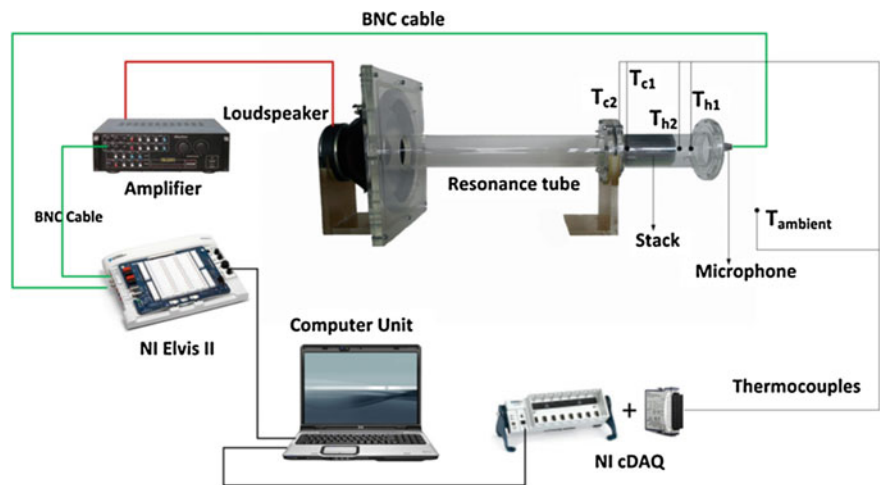


Fig. 5 The experimental set-up of the experiment

the resonator. The microphone was connected to a built-in oscilloscope of National Instrument Elvis II via Labview software. The oscilloscope confirmed the form and frequency of the wave input.

The experiment was conducted with variations in voltage peak to peak setting from 4 to 9 as the input voltage to the driver using a built in function generator in NI Elvis. The thickness of the plate stack was varied; 0.15, 0.5 and 1 mm. With each variation of the thickness, the length of the plate was also varied, 4, 5 and 6 cm. Figure 6 is the capture of one of the experiment that shows (a) the function generator used to input the sound wave and input voltage to the driver and (b) the oscilloscope monitor.

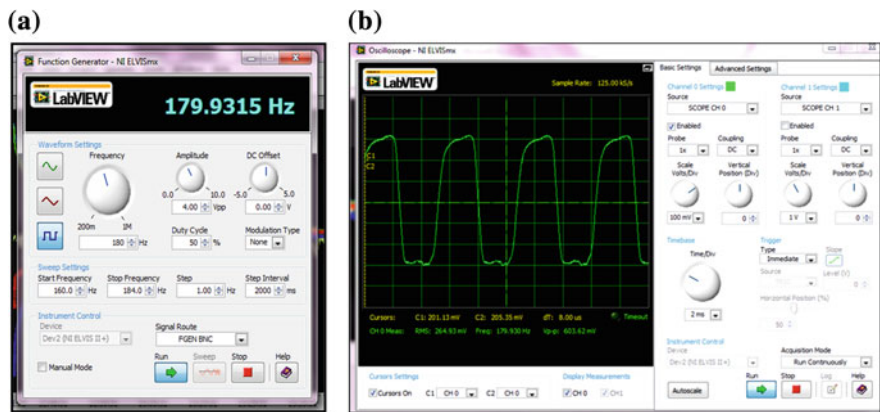


Fig. 6 Wave and voltage setting using national instrument Elvis II function generator (a) and b waveform and wave amplitude reading using NI Elvis II Oscilloscope

4 Result and Discussion

The theoretical resonance frequency for the resonator based on (2) is 209.6 Hz. The resonance frequency became around 180 Hz after installing the loudspeaker and placing the stack in the resonator. The existence of other structures and extra volume after the driver alter the resonance frequency [16]. 180 Hz is an approximation since the local temperature gradients established by the device alter the resonance frequency of the internal fluid [15].

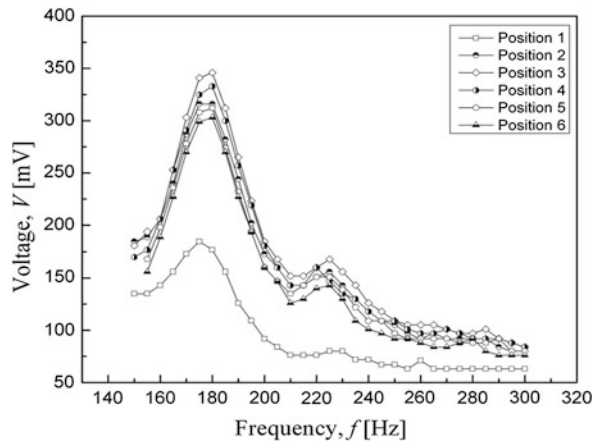
Figure 7 shows the tested frequencies and their connecting voltage result at different positions; positions 1–6 which are near the edge of the tube, $\lambda/20$ from the driver, $\lambda/8$, at the middle of tube length, $\lambda/20$ from the tube end and the last one is the left end. The frequency that resulted higher voltage was considered as the resonance frequency which was consistently around 180 Hz.

Types of sound waves which are used as the input affect the temperature difference between both ends of the stack, as shown in Fig. 8. Both ends of the stack temperature difference is of 7.8 and 5 °C respectively for the input square wave and sinusoidal wave input. Cooling is obtained by using a square wave remains stable up to 20 min, while cooling using sinusoidal wave decreases at about 10 min. Square wave has a higher energy than sinusoidal waves due to the strengthening of the amplitude.

4.1 Effect of Input Voltage Setting on Thermal Performance

To understand how sound wave amplitude affects the thermal performance of the resonator, sound wave amplitude was varied by setting the voltage peak to peak input from 4 to 9 with one increment for each experiment. The experiments were conducted for three different length of plate of 4, 5 and 6 cm, respectively. Figure 9

Fig. 7 Frequencies and their connecting voltage output at different position in resonator



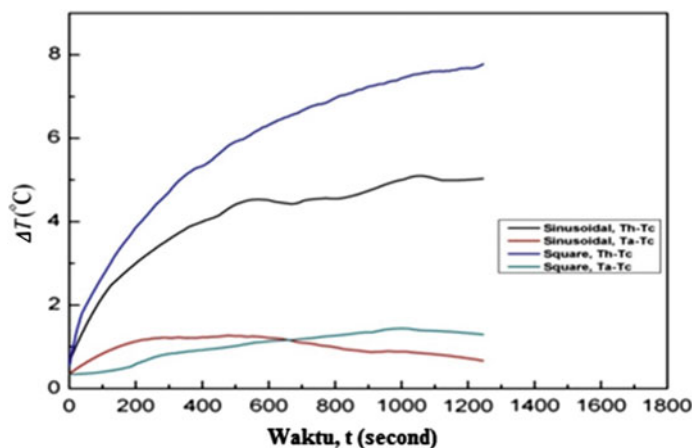


Fig. 8 Differences in temperature with different input wave

shows the performance for three variations on the thickness and the length of stack plate. Illustration of changes in ambient temperature difference and the cool side, the ambient temperature difference and the hot side and the cold side of the total temperature difference and heat each side indicated by $T_a - T_c$, $T_h - T_a$ and T_{total} . One block consists of four sections which represent data at 5, 10, 15 and 20 min duration.

The results showed that the length of 6 cm imply the largest temperature difference between cold and hot end of stack as shown in Fig. 9a–c. It is seen that the thermoacoustic phenomena inside the resonance tube yielded temperature difference between two extremities of the parallel plate stack. The result shown is the largest temperature difference, 14.8 °C achieved in 10 min time that occurred by using 0.15 mm plate thickness at input setting 9 Vpp.

Figure 10 presents the temperature difference across the stack yielded with various input voltage using 0.15 mm plate thickness. As expected the plot shows that higher input voltage will produce linearly higher temperature difference. The difference grows rapidly more than 50 % in 30 s and continues to around 80 % in 4 min and becomes generally constant in 10 min. These trends are in agreement to those shown in literature [26]. It is observed that starting from 10 min duration of experiment the local temperature gradient is relatively equal to the critical temperature gradient so that the acoustic energy have been fully used to overcome the heat dissipation [27].

The difference in temperature of the stack cold side (plate thickness of 0.15 mm) and the ambient temperature was plotted against time in Fig. 11. The data was recorded for 20 min with various input voltage to the loudspeaker. The figure shows that at 9 Vpp input to the driver the temperature of cold side decreased the most, 8.4 °C, in 90 s. The trend in Fig. 6 is similar to that found in the literature [28].

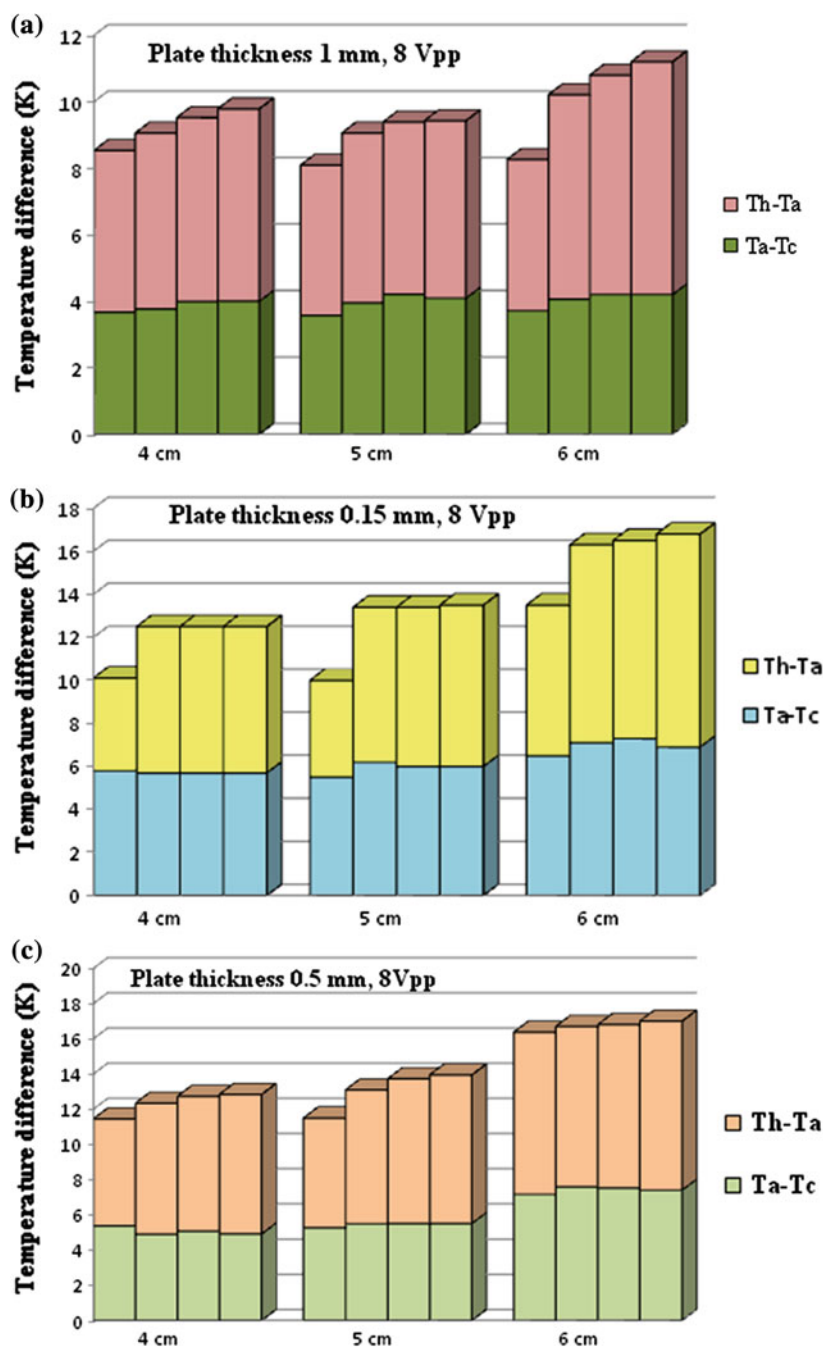


Fig. 9 Comparison of the results of the stack length 4–6 cm for **a** 1 mm thickness, **b** 0.15 mm thickness and **c** 0.5 mm thickness, on the temperature difference

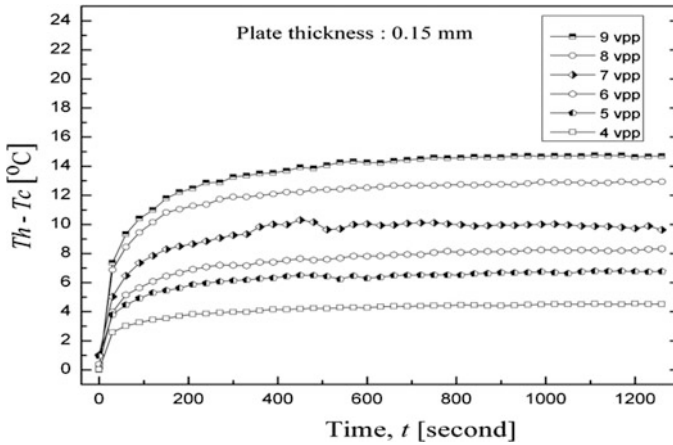


Fig. 10 The thermal performance of the resonator from 4 to 9 Vpp

The temperature then began fluctuating on the cold side of the stack due to drifting of thermocouples but it stayed lower than others. The figure also indicates that the cooling rate with input voltage 8 and 9 were faster than other input voltages although the trends are similar for all variations. The data was also recorded for different plate thickness, 0.5 and 1 mm with various input voltage. All measurements were recorded in a room with 26.5 °C ambient temperature (T_a).

4.2 The Effect of Plate Thickness on the Thermal Performance and Cooling Rate

Figure 9 depicts that using stack plate of 0.15 mm thickness produced the largest temperature difference across the stack and the fastest cooling rate compared to plate thickness of 0.5 and 1 mm. The input voltage to the driver remained constant during the experiment. At 9 Vpp, the largest temperature difference ($T_h - T_c$) for plate thicknesses of 0.15, 0.5 and 1 mm were 14.8, 13.1 and 9.2 °C respectively. It indicates that for the given resonator geometry, thinner plates means lower blockage ratio and produces higher thermal performance as found in literature [29].

Figure 9 also shows that at the same condition of input voltage, decrease in temperature at cold side of the stack ($T_c - T_a$) were up to 8.1, 7 and 4 °C for plate thicknesses of 0.15, 0.5 and 1 mm, respectively. It was observed from the experiment that thinner stack plate shows faster cooling rate at the cost more effort in fabricating the stack plate and deterioration of the stack plate after several set of experiment due to less rigidity in material used for the stack, as shown in Fig. 12.

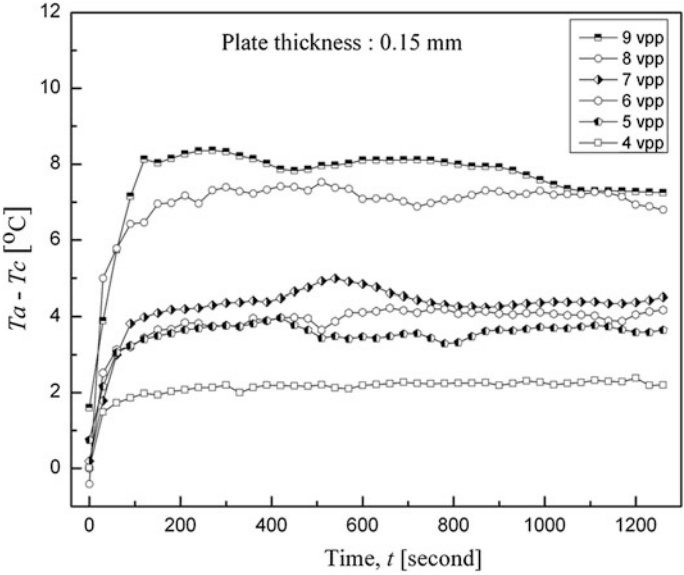


Fig. 11 Cooling rate of cold side of the stack, plate thickness of 0.15 mm

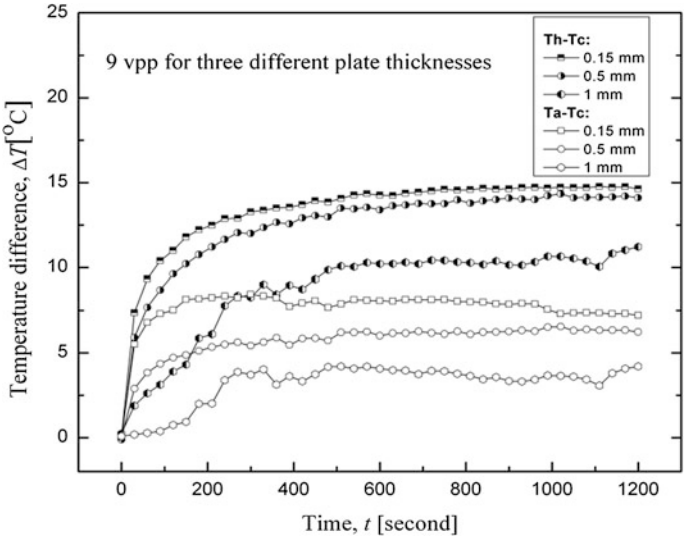


Fig. 12 Thermal performance and cooling rate with variation in stack plate thickness at 9 Vpp

5 Conclusions

The performance of the refrigerator in terms of temperature difference produced at the ends of the stack and the cooling rate were observed. The results showed that higher amplitude of sound wave (ranging from 4 to 9 Vpp setting) produce larger temperature difference, as expected. The experimental result showed that using plate with thickness of 0.15 mm and length of 6 cm at 9 Vpp yielded the largest temperature difference when compared to 0.5 and 1 mm plate for the given length and input voltage to the driver. Decrease in plate thickness leads to significant increase in cooling rate; which are for 0.15 mm plate thickness temperature dropped 8.5 °C below the ambient temperature in 2 min. For 0.5 and 1 mm plate thicknesses, the temperature drops 5 and 1 °C, respectively, below ambient in the same amount of time. Thinner plates which provide more area for heat transfer relative to the given resonator ensure faster cooling rate. However, whether the thinnest is the most optimum one still needs further study since by using plate of 0.5 mm thickness the cooling effect was more consistent and the stack structure was much more rigid compared to others.

References

1. Poesse ME (2012) Handbook of climate change mitigation. Springer, New York, pp 1821–1848
2. Lackner M, Chen W-Y, Suzuki T (2012) Introduction to climate change mitigation. Handbook of climate change mitigation. Springer, New York, pp 1–14
3. Swift G (2001) Thermoacoustics: a unifying perspective for some engines and refrigerators. *J Acoust Soc Am* 113(5):2379–2381
4. Wheatley J, Hofler T, Swift GW, Migliori A (1985) Understanding some simple phenomena in termoakustiks with applications to acoustical heat engines. *Am J Phys* 53:147–162
5. Swift GW (1992) Analysis and performance of a large thermoacoustic engine. *J Acoust Soc Am* 92:1551–1563
6. Hofler TJ (1986) Thermoacoustic refrigerator design and performance. PhD thesis, Physics Department, University of California, San Diego
7. Garret SI, Adeff JA, Hofler TJ (1993) Thermoacoustic refrigerator for space application. *J Thermophys Heat Transf* 7:595–599
8. Tijani MEH, Zeegers JCH, De Waele ATAM (2002) Construction and performance of a thermoacoustic refrigerator. *Cryogenics* 42(4):59–65
9. Rott N (1980) Thermoacoustics. *J Adv Appl Mech* 20:135–175
10. Swift GW (1988) Thermoacoustic engines. *J Acoust Soc Am* 84(4):1145–1180
11. Zink F, Viperman J, Schaefer L (2010) CFD simulation of thermoacoustic cooling. *Int J Heat Mass Transf* 53:3940–3946
12. Ghorbanian K, Hosseini H, Jafarholi M (2008) Design road-map for thermoacoustic refrigerators. *J Acoust Soc Am* 123(5):3546
13. Tijani MEH, Zeegers JCH, De Waele ATAM (2001) Prandtl number and thermoacoustic refrigerator. *J Acoust Soc Am* 112(1):134–143
14. Tasnim SH, Mahmud S, Fraser RA (2012) Effect of variation in working fluids and operating conditions on the performance of a thermoacoustic refrigerator. *Int Commun Heat Mass Transf* 39:762–768

15. Zontjens L (2008) Numerical investigations of the performance and effectiveness of thermoacoustic couples. Ph.D Dissertation, School of Mechanical Engineering, University of Adelaide. <http://digital.library.adelaide.edu.au/>
16. Ghazali NM, Ghazali AD, Ali IS, Rahman MAA (2012) Geometry effects on cooling in a standing wave cylindrical thermoacoustic resonator AIP Conf Proc 1440:1320
17. Wollan JJ, Swift GW, Backhauss SN, dan Gardner DL (2002) Development of a thermoacoustic natural gas liquefier. In: Proceedings of AIChE meeting, New Orleans
18. Symco O, Abdel Rahman E, Kwon Y, Behunin R (2004) Design and development of high-frequency thermoacoustic engines for thermal management in microelectronics. Microelectron J 35:185
19. Adeff JA, Hoffer TJ (2000) Design and construction of a solar powered, thermoacoustically driven thermoacoustic refrigerator. J Acoust Soc Am 107(6):37–42
20. Babaei H, Siddiqui K (2011) Modified theoretical model for thermoacoustic couples. Int J Therm Sci 50:206–213
21. Qing E, Feng W, Duan Yong L (2009) Thermoacoustic refrigeration device. <http://ieeexplore.ieee.org/iel5/4918025/4918026/04918931>
22. Moloney MJ, Hatten DL (2001) Acoustic quality factor and energy losses in cylindrical pipes. Am J Phys 69(3):311–314
23. Arnott WP, Bass HE, Raspet R (1991) General formulation of thermoacoustic for stack having arbitrarily shaped pore cross sections. J Acoust Soc Am 90:3228–3237
24. Wetzel M, dan Herman C (1997) Design optimization of thermoacoustics refrigerators. Int J Refrig 20:3–21
25. Ghazali NM, Aziz AA, Rajoo S (2006) Environmentally friendly refrigeration with Thermoacoustic research vote: 74166. University of Technology, Malaysia
26. Hariharan NM, Sivashanmugam P, Kasthuriangan S (2012) Influence of stack geometry and resonator length on the performance of thermoacoustic engine. Appl Acoust 73:1052–1058
27. Xiao JH (1995) Thermoacoustic heat transportation and energy transformation. Part 3: adiabatic wall thermoacoustic effect. Cryogenics 35:27
28. Russell DA, Weibull P (2002) Tabletop thermoacoustic refrigerator for demonstrations. Am J Phys 70(12)
29. Akhavanbazaz M, Siddiqui MHK, Bhat RB (2007) The impact of gas blockage on the performance of a thermoacoustic refrigerator. Exp Thermal Fluid Sci 32:231–239

Characteristic of Coal Stockpile in Lowland and the Effect to Environment

Rusdianasari, Susila Arita, Eddy Ibrahim and Ngudiantoro

Abstract Coal stockpile is a temporary shelter before they are sent to the consumer and is the operational center for a coal port. Stockpile is also used to mix coal in order to fit the needs of homogenization. Coal stockpile has wetlands in the flat topography lowlands affected by the tide. Influence the ups and downs of sea water is higher in this region compared to the influence of rainfall. Liquid waste from runoff and coal stockpile wetting is containing suspended solids and some dissolved substances. The existence of liquid waste from leachate can reduce the degree of acidity (pH) and increase the content of total suspended solids (TSS), iron (Fe) and manganese (Mn). Monitoring the effect to the environment on the coal stockpile is intended to perform an environmental assessment of the effect which arise due to the existence and operation of coal accumulation. Effect analysis based on the value of the effluent, air pollution (dust), soil and water by looking at the parameters of the coal waste water pH, TSS, metals Mn, and Fe, the total suspended particulates (TSP) parameters of air pollution and noise, and soil pollution in the form of physical and chemical characteristic of the soil.

1 Description of Stockpile

Coal stockpile is temporary shelter of coal before it is sent to consumer and center of operational for a Coal Port. Coal which is mined from the mine site will be brought into the port of coal by road and then shipped to the consumer. Stockpile serves as a buffer between the delivery and processing process, as well inventories, strategic and minimize short-term disruption or long term. It also serves as a mixing and distribution place by type of coal to conform with the request as required. Besides these purposes, the stockpile is also used to mix coal so the homogenization

Rusdianasari (✉) · S. Arita · E. Ibrahim · Ngudiantoro
State Polytechnic of Sriwijaya, Palembang, Indonesia
e-mail: diana-vsi@yahoo.com

process as needed. Homogenization aims to prepare the products from one type of material in which fluctuations in coal quality and size distribution are equated. In the homogenization process there are two types of blending and mixing [1].

Stockpile management is the process or procedure which consists of the quality setting of the coal in the stockpile build up procedure. This is done as an effort to make coal produced can be controlled, both quality and quantity. In addition stockpile management is also intended to reduce the losses that may appear from the handling process or handling of coal in the stockpile, such as rain, dust in the dry season, or burnt caused of coal burnt in the stockpile. Coal stockpiling setting is very important because it is related with the maintenance problems of quantity and quality of coal stacked in the stockpile [2].

1.1 Types of Stockpile Based on the System of Stockpiling

There are some types of stockpiles based on the system of stockpiling, as follows [3, 4]:

- Cone ply* is a system with a cone shape on one of end point until the required height is reached and continued based on long of stockpile. This system uses bulk tools, such as stacker reclaimer (Fig. 1).
- Chevron* is a system by stockpiling placement in one line material, along stockpile and stack frequently until reaches the required height. This system is good for bulk tools such as: belt conveyor or stacker reclaimer (Fig. 2).
- Chevcon* is a combination of stockpiling between cone ply system and chevron system (Fig. 3).

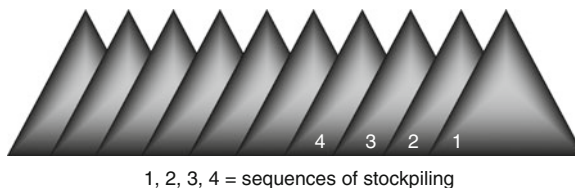


Fig. 1 Cone ply system

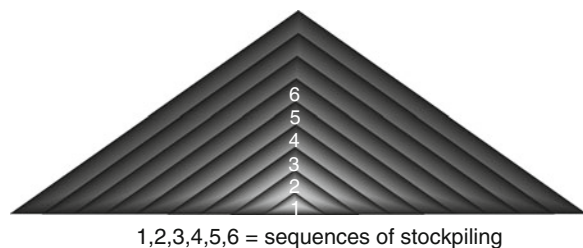
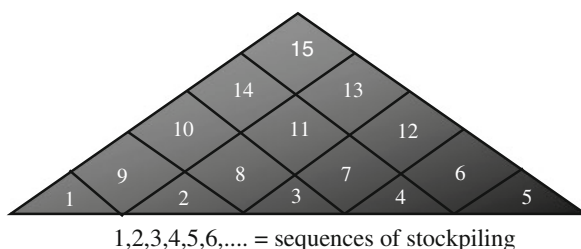


Fig. 2 Chevron system

Fig. 3 Windrow system

- d. *Windrow* is a system with stacks in parallel lines along the width of the stockpile and continued until the required height is reached. Commonly is used in tools such as: the backhoe, bulldozers, and loaders.

Coal stockpile where the study was conducted, Stockpile system is using chevron type because: contour shape, types of tools used, ease of handling, and to avoid self heating.

1.2 Technical Terms of Stockpiling

In the implementation of the stockpiling and demolition process must be regulated well. This is to avoid the stockpiling overcapacity. In this case technical Terms of stockpiling include [3]:

a. Coal quality

Coal as one of the technical terms of stockpiling should also be considered. Coal which is influential as follows:

(1) Coal which is stockpiled is one type

To avoid higher grade coal is burnt so for each stockpiling location is used similar coal (same grade and quality). That is because a lower grade coal is easier and faster to be burnt by itself, so heat from a lower grade coal accumulated and will influence higher grade coal to be burnt.

(2) Grain Size

Grain size has an influence to self heating indication, so in the coal stockpile handling should avoid coal production with the same size, because same grain size will make large enough space in stockpile and it's easy to happen the flow of air.

b. Basic surface design of stockpile

Basic surface of stockpile must be stabilized and bedding by using material which is strong enough to support the weight of coal stockpile. In addition basic surface of stockpile must be slightly convex to make drainage of stockpile fluently. It is intended to prevent puddle which is trapped in the middle of stockpile when it is rain. In stockpile with cone shape, the weight point will be

around the center of circle. This will make basic of stockpile will be decrease. If occurs decrease of basic of stockpile, it will make water is trapped in cavity which make difference humidity in stockpile which in the long term will make self heating or be accelerator when the temperature of top coal is increase [5].

c. Stockpile place condition

Stockpile place condition which influence to technical term of stockpile as follows:

(1) Clean area stockpile

Coal stockpile area must be free of all of the flammable material such as wood and waste. Besides it must be free of a pieces metals.

(2) Source of water with high pressure

Source of water with high pressure is very required if there is burning process around stockpile, example hidrant. Source of water with high pressure is required if there is burning around stockpile if it is not immediately extinguished then will influence temperature of stockpile will be increase and self heating process will be happened in stockpile.

(3) Making conduit surroundings stockpile

To drain water from coal stockpile is well which is from rain, and from water spraying surrounding stockpile area must be made conduit which is finally drain to settling pond. Water which is through coal stockpile will dissolve small coal from coal stockpile, so that small coal will be carried by water flow. Therefore before water is drained to river, it needs water treatment for water from that stockpile, or at least making settling pond. It will make small coal which is carried by water flow from stockpile, and it will not contaminated environment especially river. Besides settling pond, if the water from settling pond has $\text{pH} < 7$, it must be neutralized. Acid water neutralization with calcium carbonate and will be treated after through settling pond or before is drained to river or sea.

(4) Stockpile position

Stockpile Position must be considered to wind direction. Stockpile position is not facing wind direction, especially for long stockpile to avoid oxidation process in Stockpile [5].

1.3 Coal Delivery

Coal delivery is activity to load and unload coal which is in stockpile. Stockpile delivery has some system as follows [2]:

- a. LIFO (*Last In First Out*) System is a system where the last coal which is stockpiled will be taken for the first. In this system stockpiling process based on schedule but when delivery process the last coal which is stockpiled will be delivered in the first so this system may make coal is stockpiled for a long time.

- b. FIFO (*First In First Out*) System is a system where the first coal which is stockpiled will be taken for the first too. FIFO management in each stockpile for the Coal Company and in the end user must be done. Because it will prevent the risk of spontaneous burning in the stockpile. It because if coal is exposed into the air it will make coal is oxidate and it means will be happened self heating until spontaneous burning. Usually FIFO management is threatened with quality problem. Sometime coal which has stockpiled in the first in stockpile, it can not load and deliver with the reason that the quality is not required. But every FIFO management must be priority because besides as preventing some problem, FIFO system is cheapest.

1.4 Potential Effects of Coal Stockpiling

Potential effects of coal stockpiling is an effect or impact that could potentially appear from coal stockpiling. Potential effects of coal stockpiling varies in different types of coal, depending on the method of stockpiling [6]. Some potential effects of coal stockpiling that often occurs is as follows:

- a. Self heating and self heating factor of coal stockpiling

Self heating of coal stockpiling is a common thing and should get the attention especially on large coal stockpiling. Coal will oxidize when exposed on the surface during mining, so coal which is stockpiled will be oxidized continuously. The result of the oxidation reaction between oxygen and volatile gases from volatile matter in coal will produce heat [7].

When the oxidation reaction happens continuously, heat product will be increase, it also happens to coal stockpile. This temperature increase is also due to air circulation and heat in the stockpile is not fluently, so the temperature in the stockpile will accumulate and increase until it reaches the point of combustion temperature, which can eventually lead to self heating on the stockpile [8].

Some factors which cause coal self heating process, among others:

- (1) Stockpiling time

The longer coal will accumulate more heat which is stored in the stockpile, because the volume of air contained in the larger deposits, so the oxidation rate becomes higher.

- (2) Stockpiling methods

In a stockpile of coal needs to get compaction. With the compaction will be able to inhibit the occurrence of coal's self heating, because the space between the grains of the material between the coal is reduced. The tool used for compaction is track dozer.

- (3) Stockpiling conditions

Influence on the process conditions of self heating coal stockpiling, namely:

- **High of stockpile**
Height of the stockpile is too high will cause more heat is absorbed, it is because of the hypotenuse is formed will become longer, so the area is not compacted will be more extensive and will result in greater surface oxidized. Bituminous coal are stockpiled for more than 30 days should be the maximum height of stockpile are 6 m. As for the lignite coal stockpile are stockpiled for more than 14 days and the maximum height of stockpile are 4 m.
 - **Angle stockpiles**
Angle which is formed from a stack on the stockpile should be smaller than the angle of repose of coal mine. In general, coarse-sized material has a larger angle of repose than fine-sized material. The slope of the coal's stockpile are quite ideal is 38° .
 - **Grain size**
Basically the larger the surface area in direct contact with the outside air, the faster the combustion process itself happens. Conversely the larger the size of the lump of coal, the slower the swabakar process. Grain size of coal also affects the rate of the Oxidation Process. The more uniform grain size in a large stockpile of coal, the greater porosity and consequently the greater permeability of outside air to be circulated in the coal stockpile.
- (4) **Coal Parameter**
Parameters which is affecting self heating coal is as described at the beginning of this theory. Indication is affected the rate of oxidation increases as the decrease self heating coal rank.
- (5) **Self heating temperature**
All types of coal have indication to happen self heating process, but the time required and the temperature required for the self heating process is not same. For low rank coals have require a shorter time and lower temperature when compared to coal that has a high rank.

2 The Characteristics of Coal Stockpile in Lowlands

Coal stockpile at Muara Telang, Banyuasin Regency, South Sumatra, Indonesia, is one of the few existing stockpile at the Port of Tanjung Api Api region in the lowlands area which is affected by the tides. Stockpile location located at coordinates $02^\circ 31'00.22''\text{S}$ – $02^\circ 31'05.5''\text{S}$ $104^\circ 48'00.7''\text{E}$ and $02^\circ 30'37.6''\text{S}$ – $02^\circ 31'24.3''\text{S}$ $104^\circ 48'18.11''\text{E}$ – $104^\circ 48'20.6''\text{E}$, with an area ± 61.09 Ha [9].

Coal stockpile at Muara Telang has flat topografi which is influenced by tide and wet tropical climate. Stockpile dimensions with a maximum height of stockpile ± 15 m, with an angle of stockpile on conveyor belt, angle stockpile which is formed about 25° – 30° , while angle stockpile which is pushed by bulldozer is 38° – 44° ,

grain size about 1–30 cm for coal from mine and 50–100 mesh for coal blending. The quality of coal in stockpile contain total moisture 30.19 %, ash content 6.63 %, volatile matter 41.05 %, fixed carbon 40.28 %, total sulfur 0.67 % and calorific value 5,926 kcal/kg.

3 Environment Quality of Coal Stockpile

3.1 Air Quality in Stockpile

a. Temperature

Air temperature measurements are required, in which the gas content in the air is inversely proportional. At low temperatures, the concentration of gas pollutant in the air is considerably high (floating near the surface of the earth), increasing air temperature lowering the gas pollutant in air (gas rises in the atmosphere). The temperature of the air in coal stockpile locations ranged from 29.8 to 31.8 °C [9].

b. Carbon Monoxide (CO)

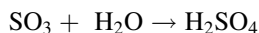
Carbon monoxide is a toxic gas that is colorless, odorless, and has no taste. This component has a weight of 96.5 % of the weight of the water and does not dissolve in water. Carbon monoxide is present in nature from incomplete combustion of the carbon or carbon-containing components, the reaction between carbon dioxide and carbon-containing components at high temperatures, the carbon dioxide will decompose into carbon monoxide and oxygen. Several studies have shown, the effect of CO on plant is usually not seen in practice. Effect of CO in humans at high concentrations can cause death, whereas contact with CO at relatively low concentrations (30,000 ppm or less) can be detrimental to health. Effect of CO on the human body is mainly caused by the reaction of CO with hemoglobin (Hb) in the blood.

The results of measurements taken at four sampling points below the level of CO was established standards 77–124 $\mu\text{g}/\text{Nm}^3$. For the content of CO in residential location, the content of CO is generally derived from the activity of motor vehicles and household activities [9].

c. Sulfur dioxide (SO₂)

Pollution by sulfur dioxide is mainly caused by two components of sulfur forms a colorless gas, namely sulfur dioxide (SO₂) and sulfur trioxide (SO₃), and both are called sulfur oxides (SO_x). Sulfur dioxide has a characteristic pungent odor and non-flammable in air, while the sulfur trioxide is a component that is not reactive. SO₂ in the air always produced in large quantities. The amount of SO₃ formed varies from 1–10 % of the total SO_x.

SO₃ in the air in the form of gas is possible only if the concentration of water vapor is very low. If the water vapor present in sufficient quantities, SO₃ and water vapor will soon combine to form droplets of sulfuric acid (H₂SO₄) with the following reaction:



The normal components contained in the air is H_2SO_4 not SO_3 . But the amount of H_2SO_4 in the atmosphere is more than SO_3 emissions, this suggests that the production of H_2SO_4 also from other mechanisms. After being in the atmosphere as SO_2 is converted into SO_3 (later to become H_2SO_4) by photolytic processes and catalytic. Total SO_2 oxidized to SO_3 is influenced by several factors including the water content, intensity, time, and distribution spectrum sunlight.

SO_x influence on plants can be affected by two factors, namely the effect of SO_2 concentration and contact time. Effects on humans and animals, SO_x at concentrations much higher than the concentration required is irritation to the respiratory system. Measurement results performed SO_2 on four sample locations as in Table 1 ($5.26\text{--}30.26 \mu\text{g}/\text{Nm}^3$), which is consist of still below the threshold standards ($900 \mu\text{g}/\text{Nm}^3$) [9].

d. Nitrogen oxides (NO_x)

Nitrogen oxides (NO_x) are a group of gases in the atmosphere of gas nitric oxide (NO) and nitrogen oxide (NO_2). Although other forms of nitrogen oxides exist, but these two gases are most commonly found as an air pollutant. Nitrogen oxide is a gas that is colorless and odorless, whereas nitrogen dioxide has a reddish brown color and pungent odor.

The sun's rays can lead to compounds react with nitrogen oxides that damage the ozone (O_3), which is a chemical compound that plays a role in the formation

Table 1 The results of measurements of air quality and noise levels in the coal stockpile

No.	Parameter	Unit	Measurement results				Threshold standard ^a
			U-1	U-2	U-3	U-4	
1	Air temperature	°C	29.8	30.6	31.8	30.5	–
2	CO	$\mu\text{g}/\text{Nm}^3$	77	124	92.4	90.2	30.000
3	SO_2	$\mu\text{g}/\text{Nm}^3$	5.26	30.26	10.16	10.10	900
4	NO_2	$\mu\text{g}/\text{Nm}^3$	2.08	21.73	6.55	6.50	400
5	HC	$\mu\text{g}/\text{Nm}^3$	0.6	2.5	0.3	0.2	160
6	TSP	$\mu\text{g}/\text{Nm}^3$	12	44	11	10	230
7	Noise	dBA	45.9	50.3	44.8	41.3	55 ^b
							70 ^c

U-1 Coal stockpile location, *U-2* Sri Tiga Village, *U-3* Telang riverside, *U-4* Karang Anyar Village

^a South Sumatra Governor Regulation No. 6 in 2012

^b For residence area

^c For industrial area

of nitrogen dioxide (NO_2), and formation of smog and acid rain. Some of the adverse effects caused by pollution is not caused by oxides NO_x , but because of its role in the formation of photochemical oxide as a critical component in smoke. Presence of NO_x in the atmosphere at a concentration of 3.5 ppm higher necrosis or damage woven leaf. NO_2 at a concentration of 5 ppm for 10 min inhaled by humans will result in a slight difficulty in breathing. The results of measurements on four sample locations, the actual NO_x ($2.08\text{--}21.73 \mu\text{g}/\text{Nm}^3$) is still below the threshold standard [9].

e. Total suspended particulate

Dust particles generally contain a variety of different chemical compounds with different sizes and different shapes, depending on where the source emission. Naturally particulate dust can be generated from dry dust carried by the wind or derived from a volcano vomit. Incomplete combustion of fuels containing carbon compounds will be pure or mixed with organic gases as well as the use of diesel engines that are not well maintained. Particulate dust drift is also produced from the incomplete combustion of coal produces the aerosol forming compound of tar beads. Compared to the burning of coal, oil and gas combustion generates fewer particulates floating dust. Motor vehicle density can increase the total emission of black smoke particulate dust.

Size of dust particulate solid or liquid form in the air depends on the size. Size of dust particulates that endanger health generally range between 0.1 and 10 microns. In general, the size of about 5 micron of particulate dust is airborne particulate matter that can go directly into the lungs and settles in the alveoli. This situation does not mean that the size of particulates greater than 5 micron are not dangerous, because the larger particulates may interfere with the upper respiratory tract and cause irritation. This situation will be worsen if the synergistic reaction with SO_2 gas contained in the air which comes from the activity of stone crusher and unloading coal. Furthermore particulate dust that floated and fluttered in the wind will cause irritation to the eyes and can impede penetration eye (visibility). The existence of spills of toxic metals contained in particulate dust in the air is the greatest danger to health. In general, the polluted air containing hazardous metals only about 0.01–3 % of all particulate dust in the air. However, these metals can be cumulative and possibly synergistic reaction occurs in the body's tissues.

The results of measurements taken at four sample locations, the content of particulate dust $10\text{--}44 \mu\text{g}/\text{Nm}^3$ and remained below established standards [9].

f. Noise

Noise in occupational health is defined as hearing voices that can reduce both quantitatively (increased hearing threshold) and qualitatively (narrowing of the spectrum of hearing) factors related to the intensity, frequency, duration and timing pattern. Noise is measured with a sound level meter. With the mechanism of action, if any object vibrates, it will lead to changes in air pressure that can be captured by these tools and will move the meter pointer.

Noise causes various disorders on society, labor and animals around him. Noise can cause disruption, such as impaired physiological, psychological, communication disorders and deafness, or disorders such as auditory disorders, such as disruption of hearing and non-auditory disturbances such as impaired communication, health hazards, decrease work performance, fatigue and stress. The results of conducted noise level measurements at four samples locations, the noise level is between 41.3 and 50.3 dBA, still below the threshold standard. The results of measurements taken at several locations in the stockpile area shown in Table 1. In Table 1 shows that the temperature has no significant difference, where the measurement takes place when the air temperature is quite bright [9].

3.2 *Water Quality*

Water quality can be seen from the parameters of physics and chemistry. Consequent changes in the physical parameters of the surface water are temperature and dissolved solids. The solids consist of organic and inorganic solids dissolved, and suspended sediment. This material will settle to the bottom of water causing specific siltation of water bodies. Another consequence of this solid growth of aquatic plants cause grow of water plant can be toxic to other creatures. The number indicates sludge solids contained in the water. Chemical parameters include the level of acidity (pH) and heavy metals.

a. Water temperature

Surface water quality standards (water bodies) set at normal temperatures. High surface water temperatures ($>45^{\circ}\text{C}$) will affect the speed of chemical reactions as well as life in the water. Temperature changes show activity of biological chemistry in solids and gases in water. Decaying os at high temperatures can lower the solubility of oxygen in the surface water. Therefore the aeration process will inhibit degradation of organic matter. Furthermore it will give affect to kill biota in water bodies and also the vegetation tropical. However, the averaged water temperature of 25.5°C is a good indicator and suitable to conditions (lower than average air temperature). The average value of the temperature measured was still met the quality standard [9].

b. Total suspended solids

The solids consist of organic and inorganic solids dissolved, and suspended sediment. This material will settle to the bottom of water over time causing siltation in particular on surface water bodies receiving. Another consequence is the solid substances pose particular the growth of aquatic plants and can be toxic to other creatures. The number indicates the amount of sludge solids contained in the water. From the analysis of dissolved solids, dissolved solid substances classified as very high compared to environmental standards, the concentration

of dissolved solids ranging from 249–355 mg/L, indicating that the location of water samples containing very low dissolved solids than the Environmental Quality Standard of 200 mg/L [9].

c. Acidity value (pH)

Normal water pH value is between 6 and 8, meanwhile the pH value is polluted, for example waste water varies depending on the type exhaust. The changes in acidity in wastewater either to the alkaline (higher pH) or to the acid (lower pH) would greatly disrupt the lives of fish and aquatic animals in the vicinity. Waste water with a low pH is highly corrosive to steel and will cause corrosion on iron pipes. Environmental Quality Standards for the parameters of the surface water pH is 6–9. Furthermore, water is highly acidic or alkaline will result in disruption of marine life and even the equipment used. From Table 2 shows that almost all the locations showed low pH values (tends to be acidic) and beyond the environmental quality standards [9].

d. Dissolved oxygen

Dissolved oxygen is a basic requirement for plant and animal life in the water. Living being in the water depends on the ability of water to maintain the minimum oxygen concentration required for life. Fish are aquatic creatures that require the highest oxygen, then invertebrates and the smallest is the bacteria need oxygen. Minimum dissolved oxygen concentration for microbial life can not be less than 5 ppm. The concentration of dissolved oxygen is too low can result in fish (neutron) and resulted in the rapid corrosion process because oxygen will bind hydrogen that coats the surface of the metal [9].

e. Heavy Metal and Toxic

Dissolved oxygen levels of the aquatic ecosystem in the vicinity of the study showed a relatively low value (less than 6 mg/L). Thus criteria, environmental quality aquatic ecosystems say less good or beyond the threshold standard.

Water is often contaminated with inorganic components, including heavy metals. Manganese (Mn) and iron are oxidized in water produced insoluble

Table 2 Analysis data of river water

No.	Parameter	Unit	Analysis results				Threshold standard ^a
			U-1	U-2	U-3	U-4	
1	Temperature	°C	26.5	26.5	26.5	26.5	–
2	pH	–	3.08	4.59	4.48	4.94	6–9
3	Total suspended solid	mg/L	249	310	355	289	200
4	Dissolved iron	mg/L	6.074	7.579	6.551	6.505	7
5	Dissolved manganese	mg/L	2.059	2.556	2.678	1.987	4

U-1 the main trench, U-2 trench 3, U-3 telang river, U-4 trench 4

^a South Sumatra Governor Regulation No. 8 in 2012 [16]

brownish color, which causes the water can not be used for domestic purposes and rocks containing compounds such as manganese and iron pyrite and hematite. In water bodies, iron (Fe) from the corrosion of heavy equipment and water pipes, metal materials as electrochemical reactions occurring on the surface. Water containing dissolved solids having electrically conductive properties to accelerate corrosion.

The results of laboratory tests on water samples taken showed levels of iron (Fe) in all locations are above the specified standards. This could be possible due to the oxidation of the pipes or objects that contain elements of iron along the river [9].

4 Environmental Impact of Coal Stockpile

4.1 Impact on Air Quality

a. Total Suspended Particulate (TSP)

Total Suspended Particulate is a very complex mixture of various organic and inorganic compounds in the air with the greatest diameter is very small, start from <1 to 500 microns. Dust particulate will be in the air within a relatively long time in suspended form in the air and into the human body through respiratory. In addition to negatively affect health, dust particle can also interfere with the power of invisibility eyes and also indicate chemical reactions, with different sizes and different shapes, depending on where the source emits [10]. Naturally particulate dust can be produced from the dry dust which is carried by the wind, oil and gas burning from the engine, and amount of motor vehicles. In the coal stockpile, when the dry season, the soil becomes dry so dust will scatter more, the operation of heavy equipment such as crusher stockpile also contributed dust, and truck traffic transporting coal from mines to coal stockpile and unloading coal contributes dust in the stockpile location.

Measurements have been done at some location. Content of particulate dust ($10\text{--}15\text{ }\mu\text{g}/\text{Nm}^3$) still below the quality standards established (threshold standard $230\text{ }\mu\text{g}/\text{Nm}^3$).

b. Noise

Noise in the stockpile location cause of the sounds of heavy equipment during loading and unloading process of coal and transportation vehicles of coal. Noise not only varies according to the sound pressure, but is also highly correlated with the frequency. There are several ways to reduce the influence of noise: reduce noise at the source, making the barrier to conductive media, and put earplugs. Noise reduction can be carried out by planting crops such as grasses, shrubs and trees. Plant species is effective to reduce the noise that has a thick canopy with shady leaves. Beside can reduce noise, while wind can produce sound [11].

Trees can reduce noise by absorbing soundwaves by the leaves, branches and twigs. Vegetation planting trees in the form of shelter belt, with a tight closure

and layered, can be substantial noise reduction up to 95 % of the source. To cope with the noise at the location of the stockpile is done by revegetation. Planted vegetation is expected to be able to reduce the impact of changes in air quality [12].

Noise level measurement results which is done at some stockpile location, the noise level (41.3–79 dBA). For measurements in a residential area of population is still below the quality standards which is established (55 dBA) while in the stockpile location the quality standards is exceeded the quality standards which is established (55 dBA).

4.2 *Impact to Water Quality*

Water conditions in the coal's stockpile location and stockpile activity is influenced by rainfall. The existence of coal stockpile activity will affect the quality of water for pH, TSS, and metals content [6].

To determine water quality in the stockpile location it will take sample of the water at several place/location, namely in the drainage, trenches 3 and 4, settling ponds and in the Telang river. Water samples were analyzed in the laboratory and compared with the water quality standards based on by peraturan Gubernur No. 8 Tahun 2012 is threshold standard to pH 6–9, TSS 200 ppm, metal content 9 ppm of Fe and metal content 4 ppm of Mn.

a. pH value

Normal pH value of water is approximately neutral is 6–8, while the pH value of the water is polluted, for example wastewater varies depending on debit. Acidity changing of the waste water either toward alkaline (pH up) or (pH down), would greatly interfere the lives of fish and other aquatic animals surrounding coal's stockpile location. In addition, waste water has a low pH is very corrosive to iron and steel and often cause corrosion on metal pipes [13]. Environmental Quality Standard for pH parameters on surface water area and waste water/water washing of coal is between 6 and 9 [14]. From the results of the laboratory analysis, pH values for water samples in the coal stockpile locations ranged from 3.08 to 5.87, and is still below the quality standard (threshold standard 6–9) and spatial distribution maps can be seen in Fig. 4.

b. Total Suspended Solids (TSS)

Solids consist of organic and inorganic solid materials are dissolved, settle, and suspended. Which is include total suspended solids is sludge, clay, metal oxides, sulfides, algae, bacteria and fungi. This material will settle to the bottom of water which gradually raises special silting on the surface of receiving water. Another impact this solid will indicate growth of certain aquatic plants can be toxic to other aquatic creature. Total solid will indicate that the contain sludge in the water. Total suspended solids are generally removed by flocculation and filtration.

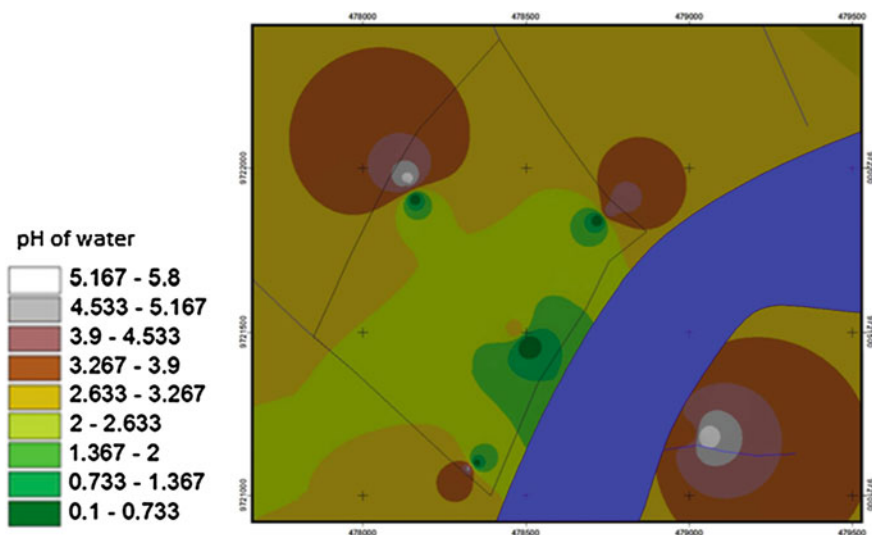


Fig. 4 Spatial distribution maps for pH of water in coal stockpile location

Total suspended solids contribute to turbidity (turbidity) by limiting light penetration for photosynthesis and visibility in the waters [15]. From the analysis of suspended solids, suspended solids content is high if compared to environmental quality standards, suspended solids concentrations ranged from 204–355 ppm, it shows that the water samples of location containing suspended solids higher than the Environmental Quality Standard of 200 ppm. Spatial distribution Map for pH values of water in the coal's stockpile locations can be seen in Fig. 5.

c. Heavy Metal

Water is often polluted by inorganic components, including heavy metals. Dangerous heavy metals are widely used in various purposes, therefore routinely produced on an industrial. Heavy metals in general, such as a mixture of iron (Fe), copper (Cu), total chromium (Cr), and aluminum (Al). Other metals were included in the heavy metal is manganese (Mn).

Manganese (Mn) and Iron (Fe) is oxidized in the brownish water and insoluble, which causes the water cannot be used for household and consumption of industrial water is so limited [14].

In the water surface, Fe levels is exceed of 1 ppm, but in groundwater, levels of Fe can be much higher. For the water which is poor of oxygen, such as groundwater, iron is present as Fe^{2+} fairly solid dissolved, while the river water flow and aeration occurs, oxidized Fe^{2+} is oxidized be Fe^{3+} which are difficult to dissolve at pH 6–8 (solubility only under a few $\mu\text{g/L}$), it also can be ferric hydroxide $\text{Fe}(\text{OH})_3$ or one type of oxide that is solid and can settle. In river water, iron is present as Fe^{2+} . Fe^{3+} dissolved, and Fe^{3+} in the form of a colloidal organic compounds. Iron is a main food source for the bacteria of iron

(centothrix, leptothrix, and gallionella) that can cause bad smell, dirty looks, and had a strange taste.

The results of laboratory tests on water samples which is taken at some point in the stockpile locations showed levels of iron (Fe) is exceed of the specified quality (threshold standard 7 ppm), while the levels of manganese (Mn) is still well below the standards (threshold standard 4 ppm). Spatial distribution map of Fe and Mn content in the water can be seen in Figs. 6 and 7.

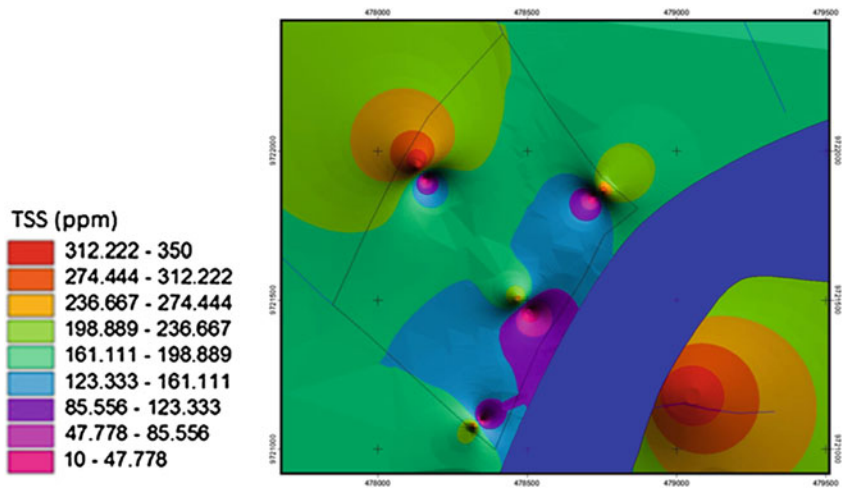


Fig. 5 Spatial distribution map for TSS of water in the coal stockpile locations

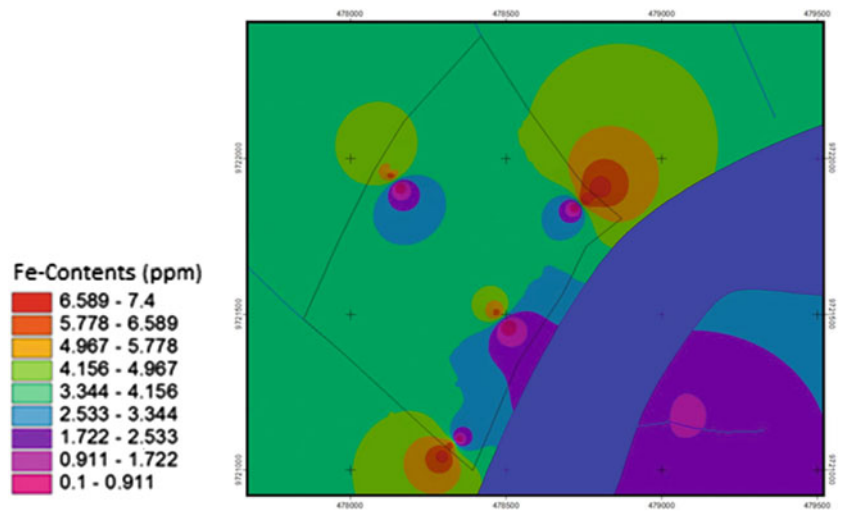


Fig. 6 Spatial distribution map of Fe content in the water

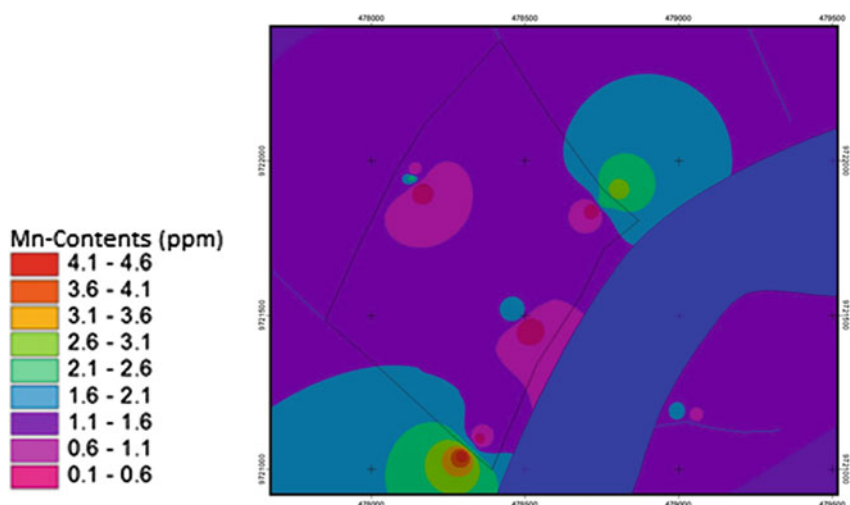


Fig. 7 Spatial distribution map of Mn content in the water

4.3 Impact to Soil Quality

The process of soil formation in the area of Muara Telang strongly influenced by water (alluvial processes), with the additional effect of a salt water seeping or flooding tide at surface of soil in some areas. Areas that are under the influence of flooding will form lands alluvial. In the areas which is not affected by salt water will form tropaquent soil. Existence is determined by the length of study in the most years, fluvial process is happened because of epipedon histik, sodium saturation which is more than 15 %, and levels of pyrite is high. When pyrite close to the soil surface, so that soil is called by sulfaquent.

Based on the results of field studies, the condition of the soil around the stockpile is thin peaty marshland. Soil type generally alluvial with shrub land cover with *Gelam* vegetation, palm, and marsh grass in the area of tidal marsh.

a. Chemical Characteristic of Soil

(1) pH of soil

Soil reaction shows acid or alkalinity of soil which is indicated by pH value. pH value indicates the number of the hydrogen ions (H^+) in the soil. The higher concentration of H^+ ions in the soil, the more acidic the soil. In the soil besides H^+ and other ions were also found OH^- ions, and the concentration is inversely proportional between H^+ and OH^- . While Acid soils the amount of H^+ ions is higher than OH^- . While alkali soil contains more OH^- than H^+ .

Alluvial or inceptisol generally have a very low pH of <4 , so type of this soil is difficult to use for cultivation. Soil acidity (pH) is the characteristic

of soil that influence absorption of nutrients by plants and is an indicator of toxic elements that influence the growth of the organism.

Field observation several points that represent the existing land units, namely the trench 3 and 4 and the center of the land. Composite sampling is done at depths up to 30 cm though. The results of the laboratory analysis showed that soil in stockpile locations have high acidity (low pH) with a range of 3.28–4.9 and its distribution map is shown in Fig. 8.

(2) C-organic content

The content of organic matter in the soil is one of the main factors in determining an aquaculture farm be success. It cause of organic content can increase chemical, physics and biology fertility of soil. Organic content is determined based on amount of C-Organic.

Organic content of soil determines interaction between abiotic and biotic component in soil ecosystem. In her research state that organic content with C-Organic form in the soil must be maintained >2 %, it will make organic content in the soil is not decrease by time cause of mineral decomposition so soil must be added with absolute organic content every year. Organic content is closely related with Capacity of Cation Exchanger and can increase Capacity of Cation exchanger of soil. Without adding organic content to soil, it will make soil gets chemical, physics, and biology degradation which can affect aggregate of soil and will happened soil compaction.

From the results of analysis with chemical method where soil has C-Organic content about 0.40–4.01 %. Distribution map for C-Organic content can be seen in Fig. 9.

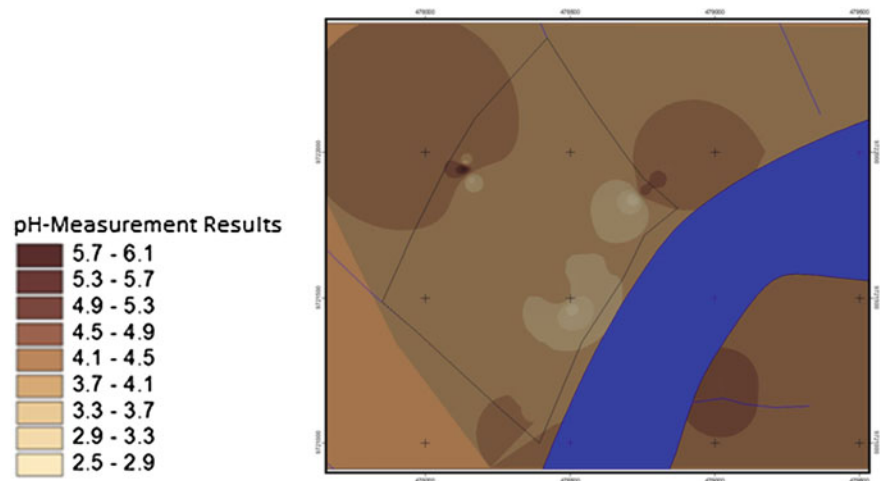


Fig. 8 Spatial distribution map for pH of soil

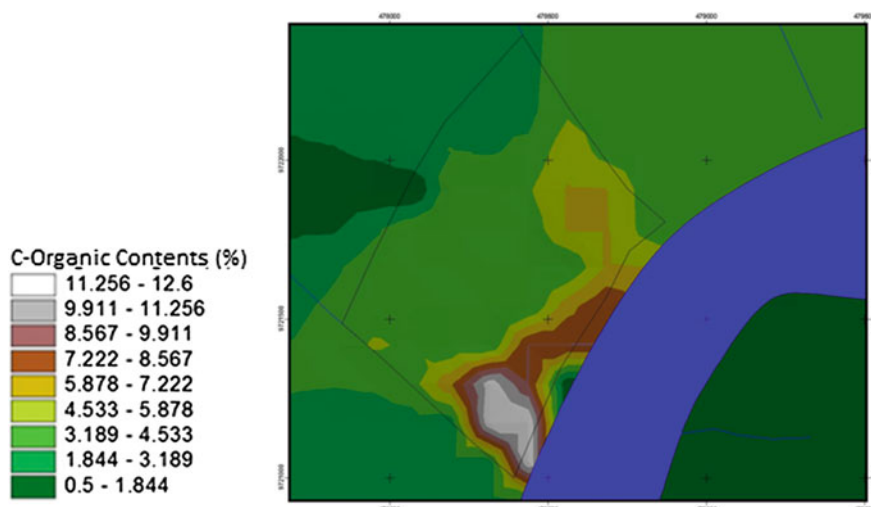


Fig. 9 Spatial distribution map for C-Organic content

(3) Nitrogen Content

Nitrogen is a macro essential nutrient, content about 15 % of the weight of plant and it uses in protein formation.

The source of N is atmosphere as primary sources, and others source are from activities in the soil as secondary sources. Symbiotic N Fixation is particularly presented in leguminosae (kind of plant) as certain bacteria. Organic content will miss N and other contents after decomposition process by microorganisms activities in the soil. N free from soil it cause of used by plant or microorganisms. Total N contents generally about between 2,000 until 4,000 kg/ha at 0–2 cm layer but it is provided for the plant are <3 % from total N contents. Function of Nitrogen is pushing the growth of plants in the vegetative phase and so important in the formation of chlorophyll, amino acids, fats, enzymes, and other compounds. Nitrogen content in the soil is in organic and inorganic forms. Organic form include NH_4 , NO_3 , NO_2 , N_2O and elements of N. Plants absorb these elements, especially in the form of NO_3 , but other forms are also able to absorb NH_4 , and urea ($\text{CO}(\text{N}_2))_2$ in the form of NO_3 . Furthermore, in the cycle, mineralization of organic Nitrogen will be happened in the soil while minerals get immobilization. Most N transported, mostly back as plant residues, lost to the atmosphere and back again, is lost through watering and increased again through fertilization. There are lost or gained cause of rain.

Relatively very low Nitrogen content ranged only from 0.05 to 0.22 %, spatial distribution map for the nitrogen content can be seen in Fig. 10.

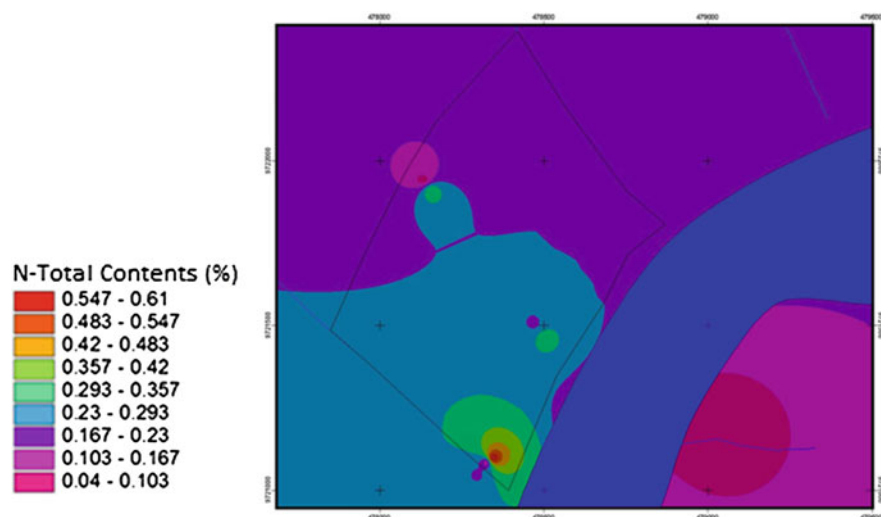


Fig. 10 Spatial distribution map for the nitrogen content

(4) Phosphorus Content (P-Bray)

Elements phosphorus (P) in soils derived from organic matter, fertilizers and minerals in the soil. Phosphorus is most readily absorbed by the plant at pH about 6–7. There are two types of phosphorus organic phosphorus and inorganic phosphorus. Organic phosphorus is usually found in the top layer is much richer in organic matter. If poor of phosphorus, inhibited cell division in plants and stunted growth. Phosphorus content of the soil of coal stockpile area ranged from 9.9 to 36.15 ppm and spatial distribution map for phosphorus content in the soil can be seen in Fig. 11.

(5) Content of Potassium

Potassium is the third nutrient after nitrogen and phosphorus is absorbed by plants in the form of K ions. Positive charge of the potassium will help neutralization process the electrical charge which is caused by the negative charge of Nitrates, Phosphates, or other elements. Potassium which is changed and absorbed by plants is dependent on the addition of external, fixation by the land itself and the addition of potassium itself.

Potassium soil is formed from weathering of rocks and minerals which is containing potassium. Through the process of decomposition of plant material and micro organisms will dissolve and then potassium back into the soil. Furthermore most of the potassium soil will be washed or eroded by soluble and process of loss will be accelerated again by the absorption process of plants and microorganisms. Some soil types have abundant potassium content. Potassium in the soil is found in weathered minerals

and release potassium ions. Adsorption ions on cation exchanged and absorbed quickly to plants. Organic soils contain a little of potassium. Potassium content in the soil of coal stockpile area is ranged from 0.19 to 1.28 me/100gr (Fig. 12).

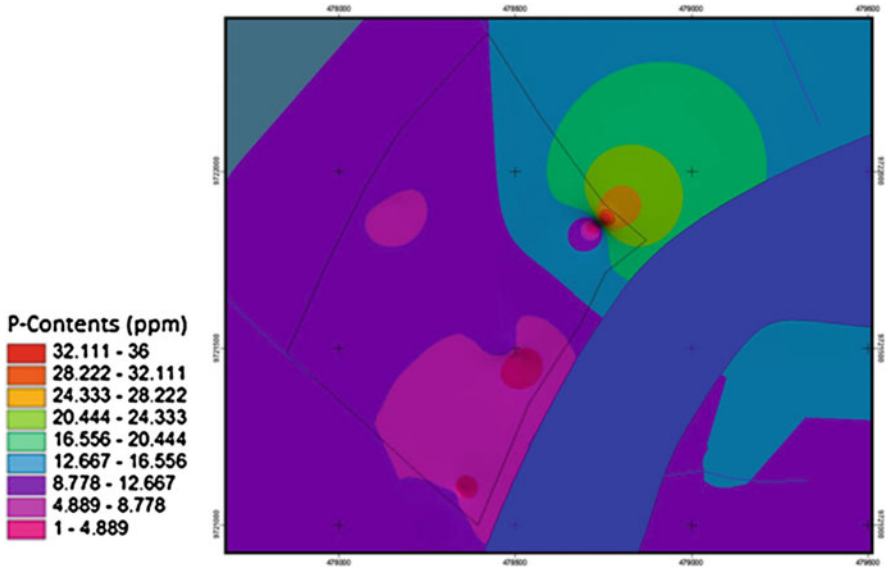


Fig. 11 Spatial distribution map for phosphorus content in the soil

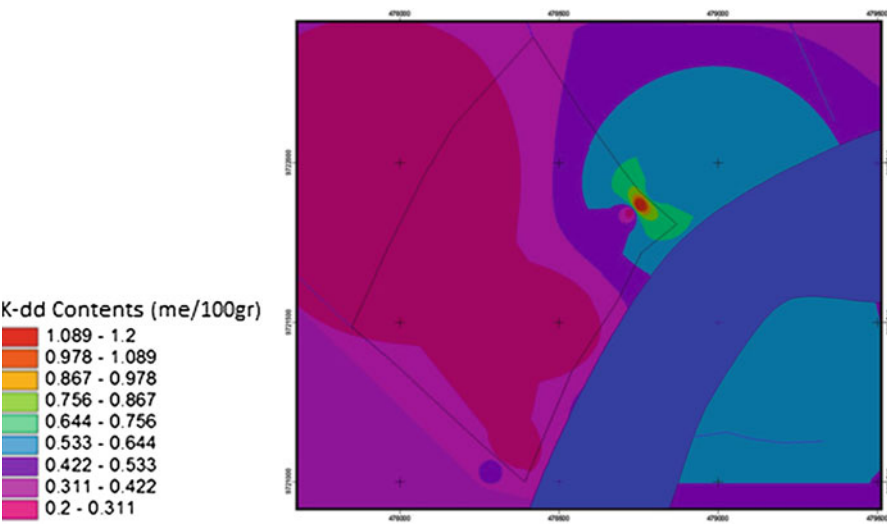


Fig. 12 Spatial distribution map for potassium content

b. Physical Characteristic of soil

(1) Porosity

Soil porosity is the ratio between the pore volume to the total volume of land, which shows the combination or arrangement of primary soil particles (sand, dust, and clay) for the secondary particles called aggregates [16]. Structure can change the texture effect by showing the relation of air humidity.

Aeration is one of the factors that determine the growth of plants. Soil aeration is a result of oxygen enter from the air through the pore spaces in the soil to the soil water to replace the oxygen which is used by plants and microorganism in the soil and carbon dioxide which is produce from soil to the air. Normal soil is soil which contains air and water in the sufficient quantities, balanced and steady. It is only found in the soil structure of the pore space, with the same ratio between macro and micro pores will stand from rain water. Porosity of the soil analysis results separately in the range 23–50 % of stockpile locations and spatial distribution map can be seen in Fig. 13.

(2) Permeability

Permeability is the ability of the soil to pass water. Soil with high permeability can increase the infiltration rate so it will decrease the rate of water run-off. On soil science, permeability is qualitatively defined as the reduction of gases, liquids or plant roots penetrate or pass. Soil permeability is a unit that includes soil infiltration and useful in soil treatment.

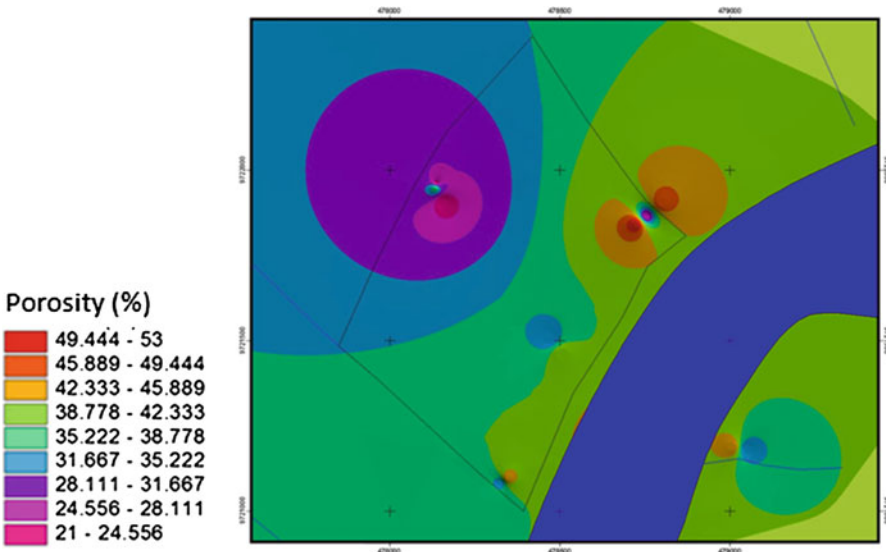


Fig. 13 Spatial distribution map for soil porosity

The top layer ranges from slow to moderately fast (0.20–9.46 cm/h), while in the bottom layer of relatively little slow to moderate (1.10–3.62 cm/h). Factors which is affecting the permeability is the texture, soil structure and porosity, viscosity and gravity, while the factors that influenced the permeability is drainage, infiltration, percolation, erosion, and evaporation. The results of soil analysis for permeability shows permeability of soil in coal stockpile has been rapid and very rapid (3.3–47.26 cm/h). Spatial distribution map for soil permeability can be seen in Fig. 14.

(3) Soil Texture

Soil texture is relative comparison between particles or fractions of primary soil, are sand, dust, clay and loam or well known in the field with a sense of roughness or smoothness of the soil. Particle/soil fraction is: sand <2–0.05 mm; dust <0.05–0.002 mm; clay <0.002 or <2 mm, which is known smooth clay <0.2 mm, while for colloidal material <0.001 mm. When it consists of particles/fraction of sand, dust and clay with the same ratio/proportional called clay. The smooth clay fraction are consist of smooth clay and colloidal materials, most of them are leached (leaching) to the bottom layer. The relation to plant growth, among which are:

- (a) If the good soil texture, such as there will be a sandy loam soil it will happened porosity and aeration which is good, This will facilitate penetration or permeation of plant root widely.
- (b) As the influence of extensive penetration of root, then the plant will have extensive root zone.

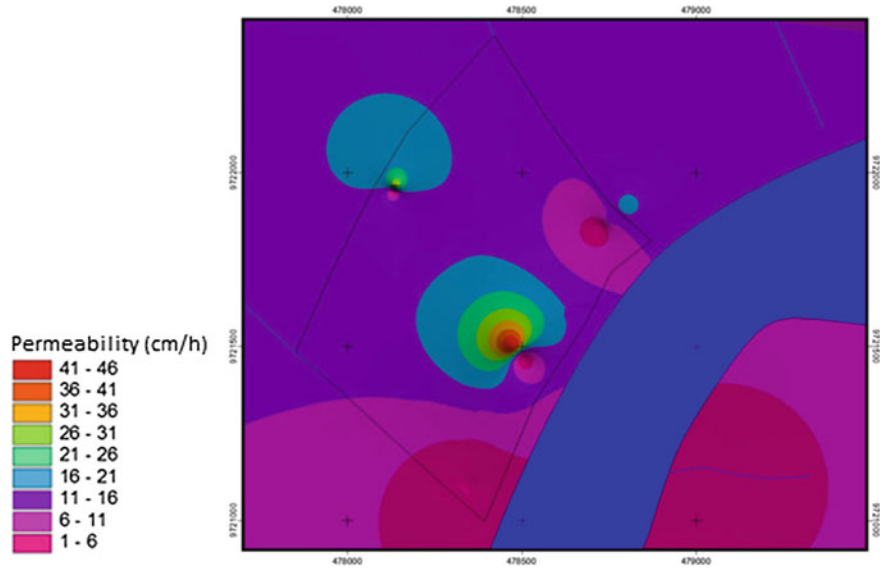


Fig. 14 Spatial distribution map for soil permeability

- (c) The extent of the root zone, will ensure high retrieval (absorption) nutrient elements of plant in the soil, so the plant will fertile.
- (d) The characteristic of good texture, soil structure will determine the formation of good soil.
- (e) Clay content, a measure colloid content of soil. The size of particle (colloid) will have a surface area and high pore space that also has a high absorption capacity and exchanging capacity of colloid. Absorption capability can be used for water and nutrients, so it will show soil fertility is increase. But if the clay content in the dominant composition or high or a less than ideal for the cultivation and soil treatment. High clay content causes percolation, infiltrasi, permeability, soil aeration, so makes it difficult to lower water and air circulation.

References

1. Badr T, Harion JL (2006) Effect of aggregate storage piles on dust emission. *Atmos Environ* 41:360–368
2. Carpenter AM (1999) Management of coal stockpile. IEA Coal Research, London
3. Hartman (2001) Introduction mining engineering. Wiley Interscience Publication, New York
4. Kravlen DWV (1981) Coal: typology-chemistry-physic-constitution. Elsevier Scientific Publishing Company, Amsterdam
5. Ejlali A (2009) A new criteria to design reactive coal stockpiles. *Int Commun Heat Mass Transf* 36:669–673
6. Wang J, Takarada T (2011) Environmental problem arising from coal handling and processing. Guma University, Japan
7. Ribeiro J, Ferreira da Silva E et al (2010) Burning of coal waste piles from Douro Coalfield (Portugal). *Int J Coal Geol* 81:359–372
8. Liu L, Zhou F (2010) A comprehensive hazard evaluation system for spontaneous combustion of coal in underground mining. *Int J Coal Geol* 82:27–36
9. Rusdianasari, Susila A, Eddy I et al (2013) Evaluation on environmental effect of coal stockpile in Muara Telang, Banyuasin, Indonesia. doi:[10.1088/1742-6596/423/1/012053](https://doi.org/10.1088/1742-6596/423/1/012053)
10. Koninklijke (1980) Dust concentration around coal stockpiles, Amsterdam
11. South Sumatra Governor Regulation No. 8 of 2012 on the Liquid waste quality standard for industrial activities, hotels, hospitals, domestic and coal mining
12. South Sumatra Governor Regulation No. 6 of 2012 on Quality standard emission stationary sources and limits of motor-vehicle exhaust emissions
13. Axcil A, Koldas S (2006) Acid mine drainage (AMD): causes, treatment and case studies. *J Clean Prod* 14:1139–1145
14. Fardiaz S (1992) Air and water pollution. Publisher Canicius, Jakarta
15. Ward CR, French D (2011) Element leachability from a coal stockpile and associated coastal sand deposits. *Fuel Process Technol* 92:817–824
16. Morashi ARA, Scallion J (2004) Porosity and hydrological changes in surface mine soil. In: 13th International soil conservation organization conference, ISCO, Brisbane, 2004

Synthesis and Characterization of Calcium Iodate, Monohydrate Crystals Grown in Silica Gel

S.J. Shitole

Abstract Simple gel technique was used to grow single crystals of calcium iodate, monohydrate, by single diffusion method. For the growth of crystals, the optimum conditions were established. Crystals having different morphologies like prismatic and prismatic pyramidal were obtained. Few crystals were opaque, some were translucent and some good quality transparent crystals were obtained. Doping of Cu^{+2} and Fe^{+3} was done, to study the effect on structure, thermal and nonlinear optical properties. Cell parameters were obtained from the X-ray diffractograms. Structural analysis was done by using FT-IR spectroscopy. TGA and DTA techniques were used to carry out thermal analysis. Nonlinear optical properties were studied using SHG measurements.

1 Introduction

The study of growth and characterization of single crystals is receiving increasing importance due to their various applications in solid-state technology and laser technology. With the absence of crystals, there would be no electronic industry and fiber optic communication. Growth of crystals by gel method is a promising technique for growing single crystals of substances which are sparingly soluble in water and decompose before their melting point [1]. A variety of crystals of high quality can be grown in a large range of solubility and temperature. In gel growth, crystals are mostly formed at ambient temperature and hence are free from strain often present in crystals prepared from melt or vapour. Method is inexpensive and within

S.J. Shitole (✉)

Department of Electronics, Z. B. Patil College, Dhule 424002, Maharashtra, India
e-mail: sjshitole@hotmail.com

the scope of small laboratories. Large scale movements like convection currents are almost completely suppressed, which otherwise could be harmful to the quality of crystal [2–7].

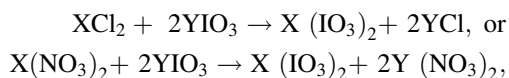
Single crystals of calcium iodate, which is non-centro-symmetric, exhibits prominent nonlinear optical property, and piezoelectric property [8]. In the present investigation, Single crystals of calcium iodate, monohydrate were grown by gel method. Since the compound decomposes before its melting point, conventional high temperature methods for its growth are not suitable. It is sparingly soluble in water; hence, gel method is the only alternative technique for growing the crystals of the size and quality as reported here at ambient temperature.

2 Experimental

To carry single and double diffusion experiments, test tubes and U-tubes were used as crystallizing vessels. Silica gel is used as growth medium. Acetic acid is used to acidify the gel. Proper chemical reaction is selected to get the desired product. The two soluble reactants are diffused into a gel, where they react to form an insoluble product. Dopants are incorporated either in gel or above the gel along with the supernatant. Optimum conditions were obtained by varying parameters such as density of gel, pH of gel, gel aging, concentration of reactants, concentration programming, concentration and amounts of dopants.

Various concentrations of acetic acid and those of sodium metasilicate were used to prepare gel. For this purpose, 5cc, 2N acetic acid was taken in a beaker, to which sodium metasilicate solution having different densities was added drop by drop with constant stirring by using magnetic stirrer. It avoids premature local gelling. To this mixture, 5cc of potassium or sodium iodate solution was added with constant stirring. The pH of the mixture was maintained at 4.2. Experiments were performed to optimize suitable pH value for growth of good quality crystals. This mixture was then transferred to the test tube and it was closed with cotton plug. The gel was allowed to set. It took nearly 12 days for setting. This set gel was aged for 5 days. Aging helps in nucleation control due to reduction in the diameter of the capillaries in gel. Calcium chloride or calcium nitrate was used as supernatant. Supernatants having different molarities were carefully poured over the set gels. Experiments were also performed by interchanging positions the reactants.

The chemical reactions inside the gel can be expressed as,



where X = Ca and Y = K or Na.

3 Result and Discussion

Parameters such as gel density, gel setting time, gel aging time, concentration of reactants, pH of gel, concentration of impurities have considerable effect on growth of crystals. Crystals having different morphologies were obtained. Prismatic crystals of size $3 \times 2 \times 2 \text{ mm}^3$ far away from gel interface and prismatic pyramidal crystals of size $5 \times 2 \times 2 \text{ mm}^3$ away from gel interface were obtained. It was observed that the number of crystals growing diminished with the increase in the distance from gel interface. It may be due to reduced rate of diffusion of supernatant. Second reason may be attributed to the aging of gel, since crystals in this region nucleate in a comparatively older gel.

High concentration of reactants results into either needle shaped crystals of size $8 \times 2 \times 2 \text{ mm}^3$ or hopper crystals of size $10 \times 3 \times 2 \text{ mm}^3$ near the gel interface. The reason may be high diffusion gradient near the gel interface. Less concentration of reactants leads to the formation of platy crystals due to slow diffusion rate and insufficient supply of reactants. Figure 1 shows few transparent needle shaped crystals, observed away from gel interface, at somewhat more concentration of reactants. Figure 2 shows hopper shaped crystals, grown inside the test tube near gel interface, due to higher concentration of reactants. Figure 3 shows different habits of calcium iodate crystals.

Color of the colored crystals is due to the inclusion of small amount of impurities or due to color centers [9]. Dake [10] studied rose red or pinkish variety of massive quartz. According to him, the color is due to the inclusion of a small amount of salts of titanium or manganese. Nassau [11] has studied the origin of color in minerals. He pointed out that color of amethyst, fluorite and smoky quartz is due to color centers. The effects that the impurities may have on acoustic, electric, magnetic and optical properties have led to deliberate attempts to include foreign ions in synthetic crystals [12].

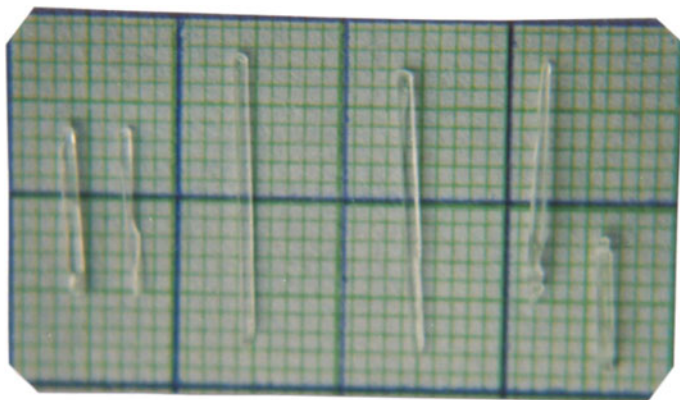


Fig. 1 Needle shaped crystals



Fig. 2 Hopper shaped crystals inside test tube

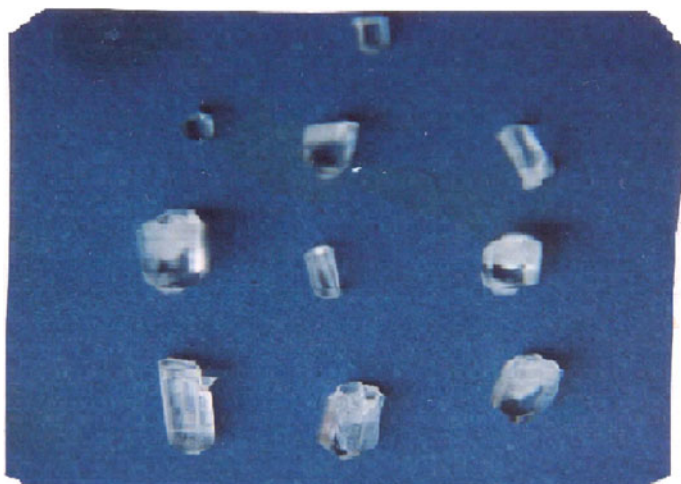


Fig. 3 Different morphologies of calcium iodate crystals

In the present investigation, dopants were incorporated either in gel or along with supernatant to study the effect of dopants on properties of the grown crystals of calcium iodate.

Very less amount and low concentration of impurity does not affect the growth and morphology of the crystals. High concentration of impurity, when added along

with the supernatant over the set gel is unable to penetrate the gel. On the other hand, an impurity solution, when incorporated in the gel, increases shining and improves morphology of the crystals. Cu^{+2} impurity induces light blue color while Fe^{+3} impurity induces light brown color in crystals. The effect of dopant on crystal growth is a complicated process. While selecting the dopant, the impurity ion substitution for Ca^{+} in crystal structure having smaller ionic radii than Ca^{+} should be favored. In this case, the ionic radius of Ca^{+} is 0.99 Å and that of Fe^{+3} and Cu^{+2} is 0.64 and 0.70 Å respectively. Very small amount of incorporation of impurities in calcium iodate crystals is confirmed by EDAX analysis.

4 Characterization

The characterization forms an indispensable part of the crystal growth activity. The physical and chemical assessment of a material as can be done in laboratory is defined as characterization. A crystal is said to be fully characterized when we know the identity, concentration, and position of all its constituent atoms. Characterization of the material is concerned with the determination of elemental composition, identification of crystalline phase, and quantification of impurities. Characterization of grown crystals involves understanding the relationship between the structure, perfection and chemical composition of the crystals with their structural, thermal and optical properties. For this purpose Calcium iodate crystals were characterized by XRD (X-ray diffractometry), FTIR (Infrared Spectroscopy), and Thermal Analysis [TGA (Thermo Gravimetric Analysis), DTA (Differential Thermal Analysis)] to study structure, and SHG Measurements to study NLO property.

4.1 XRD (X-Ray Diffractometry)

X-ray diffraction (XRD) is an excellent analysis technique for identifying unknown solid materials. The XRD pattern can act as a fingerprint to identify unknown materials in crystal and semi crystalline samples. The atoms are arranged in a regular pattern, and there is a smallest volume element, that by repetition in three dimensions describes the crystal. This smallest volume element is called a unit cell. The dimensions of the unit cell are described by three axes: a, b, c and the angles between them alpha, beta and gamma.

X-ray diffractograms were recorded using powder rotation photograph method. The specifications of diffractometer are, $\text{CuK}\alpha$ radiation $\lambda = 1.5418$ Å, 2θ range = 10 – 90° and scanning speed = $10^\circ/\text{min}$. Figures 4, 5, and 6 show diffractograms of undoped, Cu-doped and Fe-doped calcium iodate crystals respectively.

From diffractograms, it is clear that impurities are induced only in certain planes, but it does not change the structure of crystal. Impurities only cause a slight change in lattice parameters. Change in volume of unit cell of doped crystals may be

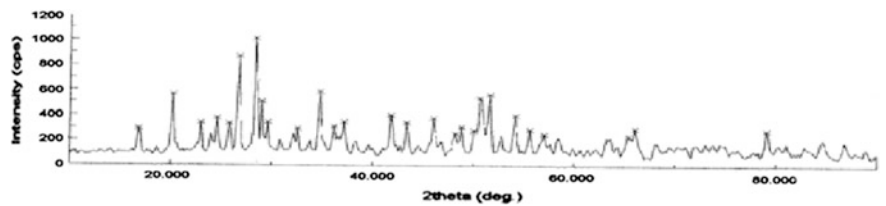


Fig. 4 X-ray diffractogram of undoped calcium iodate

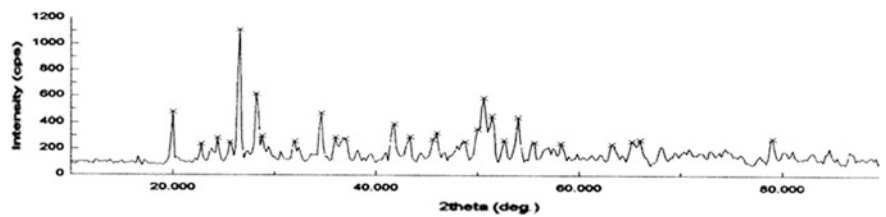


Fig. 5 X-ray diffractogram of Cu-doped calcium iodate

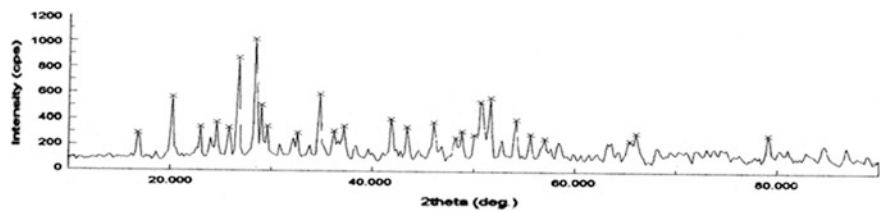


Fig. 6 X-ray diffractogram of Fe-doped calcium iodate

attributed to ionic radii of doped impurities (Cu^{+2} , Fe^{+3}). These crystals belong to monoclinic system with $a \neq b \neq c$ and $\alpha = 90^\circ \neq \beta$. The calculated values of a , b , c , and β are in good agreement with the reported ones [13]. The calculated values of lattice parameters, β , and volume of undoped and doped crystals with respect to reported values are represented in Table 1.

Table 1 Lattice parameters of calcium iodate, monohydrate

Lattice parameters	Undoped reported	Undoped observed	Cu-doped (observed)	Fe-doped (observed)
$a \text{ \AA}$	8.509	8.494 (1)	8.543 (7)	8.554 (4)
$b \text{ \AA}$	10.027	10.033 (2)	10.136 (7)	10.049 (2)
$c \text{ \AA}$	7.512	7.512 (2)	7.549 (6)	7.501 (3)
β°	95.27	95.00	94.00	95.00
$V(\text{\AA})^3$	638.213	637.457	651.673	642.349

4.2 FTIR Characterization

FTIR is an important technique for identification and characterization of a substance. Infrared Spectroscopy gives information on the vibrational and rotational modes of motion of a molecule.

In order for a vibrational mode in a molecule to be “IR active,” it must be associated with changes in the dipole. A permanent dipole is not necessary, as the rule requires only a change in dipole moment. A molecule can vibrate in many ways, and each way is called a vibration mode. For molecules with N atoms in them, linear molecules have $3N - 5$ degrees of vibrational modes, whereas non-linear molecules have $3N - 6$ degrees of vibrational modes, also called vibrational degrees of freedom.

Simple diatomic molecules have only one bond and only one vibrational band. If the molecule is symmetrical, the band is not observed in the IR spectrum, but only in the Raman spectrum. Asymmetrical diatomic molecules, absorb in the IR spectrum. More complex molecules have many bonds, and their vibrational spectra are correspondingly more complex.

FT-IR spectra of undoped and doped iodate crystals were scanned by using FT-IR spectrophotometer, Spectrum-2000, Perkin-Elmer model by placing sample KBr pellet in the sample beam in two ranges, 300–710 and 400–4,000 cm^{-1} . FT-IR spectra of undoped, Cu-doped, and Fe-doped calcium iodate crystals represented in Figs. 7, 8 and 9 respectively.

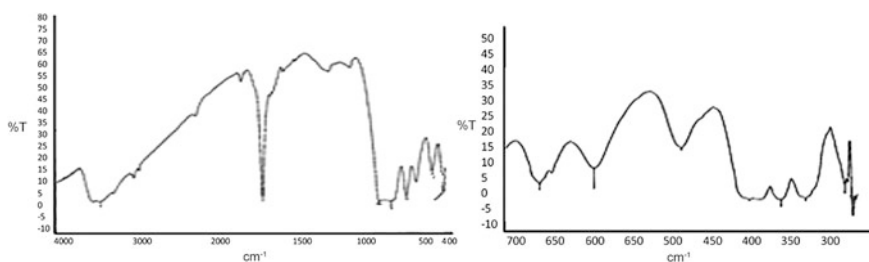


Fig. 7 FT-IR spectra of undoped calcium iodate in the range 4,000–400 and 710–300 cm^{-1}

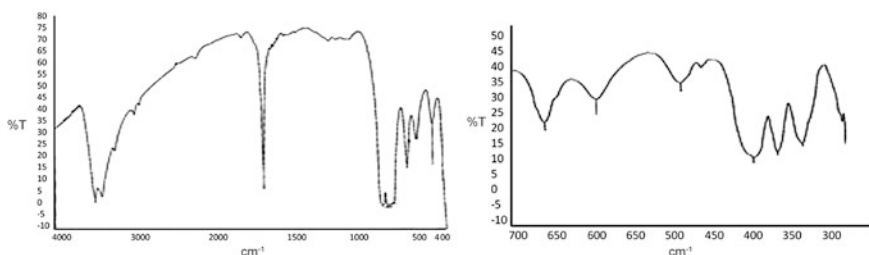


Fig. 8 FT-IR spectra of Cu-doped calcium iodate in the range 4,000–400 and 710–300 cm^{-1}

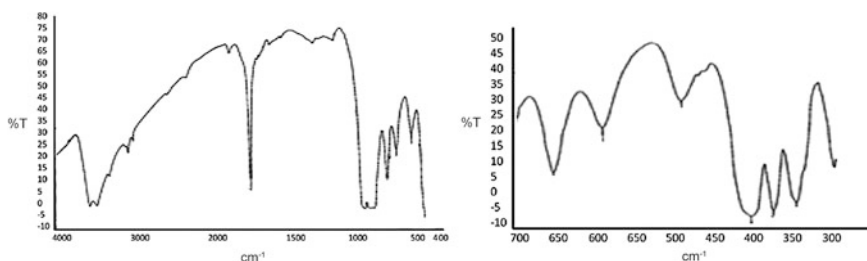


Fig. 9 FT-IR spectra of Fe-doped calcium iodate in the range 4,000–400 and 710–300 cm^{-1}

Table 2 Vibrational frequencies of undoped, Cu-doped, and Fe-doped calcium iodate crystals

Fundamental frequencies	Undoped (cm^{-1})	Cu-doped (cm^{-1})	Fe-doped (cm^{-1})
Symmetric stretching frequency, γ_1	759.15	759.15	759.15
Symmetric bending frequency, γ_2	396.77	396.74	396.72
Asymmetric stretching frequency, γ_3	818.83	815.86	818.85
Asymmetric bending frequency, γ_4	335.04	335.47	335.15

Fundamental frequencies of the pyramidal iodate ions as have been reported earlier by Nassau et al. [14], Dasent and Waddington [15], Balicheva and Petrova [16], Niquist and Kagal [17] Rocchiccioli [18], Dratovsky and Pacesova [19], and Nakamoto [20] are in good agreement with the observed values in the present investigation. Vibrational frequencies of undoped, Cu-doped, and Fe-doped calcium iodate crystals are given in Table 2.

4.3 Thermal Analysis

Thermal studies were carried out using Mettler Toledo Star system. Definite amount of sample was taken and heating was carried out from ambient to 900 $^{\circ}\text{C}$ for TGA and DTA at the rate of 5 $^{\circ}\text{C}/\text{min}$ in an air medium. Figure 10 shows TGA curve and Fig. 11 shows DTA curve of undoped calcium iodate. Figure 12 shows TGA curve and Fig. 13 shows DTA curve of Cu-doped calcium iodate. Figure 14 shows TGA curve and Fig. 15 shows DTA curve of Fe-doped calcium iodate crystals, respectively.

These crystals exhibit three steps explicitly. Table 3 shows kinetic data from TGA of these crystals. Results obtained are in good agreement with the reported one [21]. The compound is stable up to 220 $^{\circ}\text{C}$. The weight loss begins at 220 $^{\circ}\text{C}$. There is 3.2 % weight loss in the temperature range 220–280 $^{\circ}\text{C}$. This is the first step and is suggestive of loss of one water molecule from the crystal. This water loss step is not very sharp but extends over a wide temperature range of about 60 $^{\circ}\text{C}$

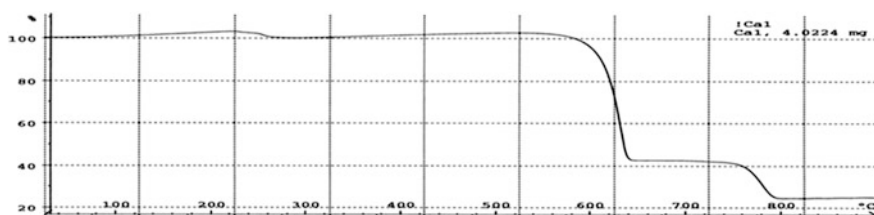


Fig. 10 TGA curve of undoped calcium iodate

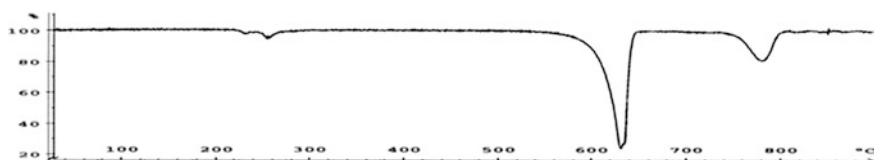


Fig. 11 DTA curve of undoped calcium iodate

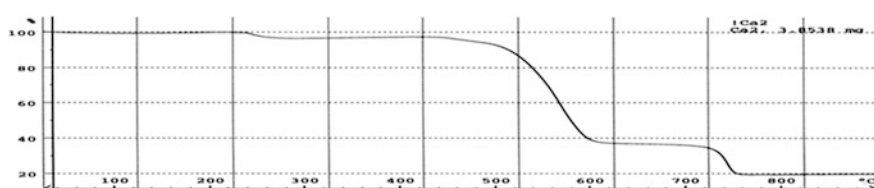


Fig. 12 TGA curve of Cu-doped calcium iodate

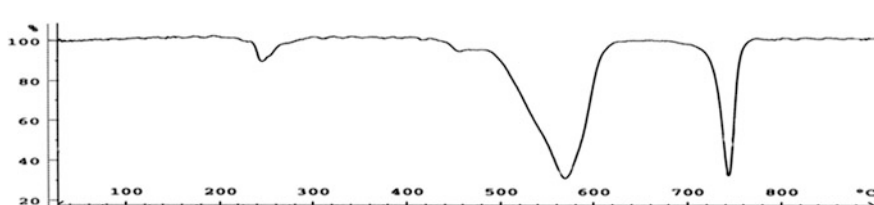


Fig. 13 DTA curve of Cu-doped calcium iodate

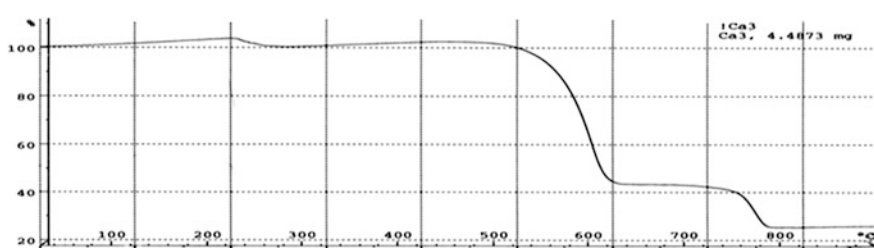


Fig. 14 TGA curve of Fe-doped calcium iodate

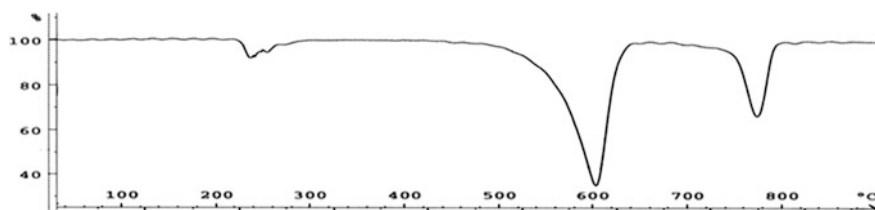


Fig. 15 DTA curve of Fe-doped calcium iodate

Table 3 Kinetic data from dynamic TGA of undoped, Cu-doped, and Fe-doped calcium iodate

Compound	Step	Temperature range (°C)	% weight loss	Probable product formed (solid)
Undoped calcium iodate, monohydrate	I	220–280	3.2	$\text{Ca}(\text{IO}_3)_2$
	II	580–640	58	$\text{Ca}_5(\text{IO}_6)_2$
	III	750–800	18	$\text{Ca}_5(\text{IO}_4)_2$
Cu-doped calcium iodate, monohydrate	I	240–280	2	$\text{Ca}(\text{IO}_3)_2$
	II	450–600	60	$\text{Ca}_5(\text{IO}_6)_2$
	III	720–750	18	$\text{Ca}_5(\text{IO}_4)_2$
Fe-doped calcium iodate, monohydrate	I	220–270	3.2	$\text{Ca}(\text{IO}_3)_2$
	II	500–620	58	$\text{Ca}_5(\text{IO}_6)_2$
	III	750–780	18	$\text{Ca}_5(\text{IO}_4)_2$

and indicates that lattice water is strongly hydrogen bonded with iodate group, showing that calcium iodate crystals are monohydrated.

For Cu-doped and Fe-doped calcium iodate crystals, almost similar results are obtained with slight changes in the temperature range of decomposition reactions and values of % weight loss. Thus doping has no effect on the structure of crystal. Kinetic data from dynamic TGA of undoped, Cu-doped, and Fe-doped calcium iodate is shown in Table 3.

4.4 Powder SHG Measurements

Powder SHG technique [22, 23] makes it possible to predict the magnitude of the nonlinear coefficients of the powder samples in a very simple manner, should provide a substantial increase in the number of new materials for use in nonlinear optic applications.

In SHG, an optical beam of frequency 2ω can be generated from the interaction of a high power laser beam of frequency ω with a suitable material either in the form of crystal or powder sample. It cannot occur in isotropic medium or in centrosymmetric crystals. Hence, a noncentrosymmetric crystal is a must to get SHG [24–26].

Table 4 Nonlinear coefficients of KDP and undoped calcium iodate crystals

Substance	Nonlinear coefficient (d's) m/v
KDP	6.3
Undoped calcium iodate	1.2535

For performing experiments, sample plates were prepared from powder samples of KDP, undoped, Cu-doped, and Fe-doped calcium iodate. All the samples were sieved by a 75 μm sieve, in order to maintain uniform particle size. Each sample plate was placed in front of Nd: YAG laser beam. Before sample plate, beam splitter splitted the input signal into reference signal, collected by photo detector and second signal on sample plate, which in turn generated second harmonic signal. Reference signal and second harmonic signal were both applied to digital storage oscilloscope. $V(\omega)$ of reference signal and $V(2\omega)$ of SHG signal were recorded as a function of arc lamp voltage. Conversion efficiency was calculated by taking ratio $V(\omega)/V(2\omega)$. Graphs of arc lamp voltage versus conversion efficiency were plotted and $\tan \theta$ has been found. Nonlinear coefficients were calculated by using the relation,

$$d^2 \text{ known}/d^2 \text{ unknown} = \tan \theta \text{ known}/\tan \theta \text{ unknown}$$

where,

$d \text{ known}$	nonlinear coefficient of KDP
$d \text{ unknown}$	nonlinear coefficient of the sample under test, to be found out
$\tan \theta \text{ known}$	slope of graph of arc lamp voltage versus conversion efficiency of reference sample (KDP)
$\tan \theta \text{ unknown}$	slope of graph of arc lamp voltage versus conversion efficiency of sample under test.

Table 4 gives the results of calculations of nonlinear coefficients (d's). Cu and Fe doped samples of calcium iodate do not show any nonlinear property. Output from these sample plates was completely stopped.

5 Conclusions

Gel technique can be successfully employed for the growth of calcium iodate crystals. In case of gel method, various habits can be obtained by changing various parameters. Well known Liesegang phenomenon has been observed in case of calcium iodate crystals. Calcium iodate crystals exhibit the phenomenon of efflorescence. Suitable value of pH for calcium iodate crystals is 4.2. Doping of impurities has no significant effect on the morphology of crystals. XRD results match very well with the standard JCPDS data. Dopants have been incorporated

only in certain planes. No major changes in structure have been observed. IR analysis confirms the presence of fundamental IR frequencies observed in all iodate compounds. Thermal analysis exhibits three steps explicitly on heating the samples. The first step involves dehydration at 250 °C, second step shows decomposition at 580 °C, and the third step involves again decomposition at 640 °C. Powder second harmonic generation experiments exhibit the nonlinear nature of the substance. But the intensity of SHG is very small. Cu^{+2} and Fe^{+3} doped samples have negative effect on SHG property. These samples do not show nonlinear effect.

References

1. Armington AF, O'Connor JJ et al (1968) Gel growth of cuprous halide crystals. *J Cryst Growth* 3–4:367–371
2. Shitole SJ, Saraf KB et al (2001) Growth and study of some gel grown group II single crystals of iodate. *Bull Mater Sci* 24(5):461–468
3. Shitole SJ, Saraf KB et al (2002) Growth, structural and microtopographical studies of calcium Iodate, monohydrate crystals grown by silica gel. *Cryst Res Technol* 37(5):440–445
4. Shitole SJ, Saraf KB et al (2007) Spherulitic growth of cadmium iodate in silica gel and its XRD and FTIR analysis. *J Adv Sci And Tech* 10(I & II):42–55
5. Joshi MS, Trivedi SG et al (1981) Growth of single crystals of barium iodate and strontium iodate in silica gels. *Cryst Res Technol* 16(1):19–26
6. Garud SL, Mahajan NK, Saraf KB et al (2009) Study of gel grown mixed crystals of $\text{Ba}_{1-x}\text{Ca}_x(\text{IO}_3)_4$. *Bull Mater Sci* 32(2):187–192
7. Garud SL, Saraf KB et al (2008) Growth and study of mixed crystals of Ca–Cd iodate. *Bull Mater Sci* 31(4):639–643
8. Dennis J, Henisch HK et al (1967) Nucleation and growth of crystals in gels. *J Electrochem Soc* 114:263–266
9. Dishovsky N, Boncheva – Maladenova Z et al (1981) Doping of gel-grown Ag_2SeO_4 single crystals. *J Cryst Growth* 51(1):147–148
10. Dake HC (1938) Quartz family minerals: a handbook for the mineral collector. Whittlesey House, New York
11. Nassau K et al (1978) The origins of color in minerals. *Am Mineral* 63:219–229
12. Saraf KB (1980) Ph. D. thesis, Sardar Patel University, Vallabh Vidyanagar, Gujrat, India
13. JCPDS card for X-ray diffraction data no. 26, p 1405
14. Nassau K, Shiever JW, Prescott BE et al (1973) Transition metal iodates. I. Preparation and characterization of the 3d iodates. *J Solid State Chem* 7:186–204
15. Dasent WE, Waddington TC et al (1960) Iodine–oxygen compounds. Part I. infrared spectra and structure of iodates. *J Chem Soc* 2429–2432
16. Balicheva TG, Petrova VA et al (1973) Vibrational spectra of the IO_3 -ion in alkali iodates. *J Struct Chem* 14(3):424–430
17. Nyquist RA, Kagal RD et al (1971) Infrared spectra of inorganic compounds. Academic Press, New York
18. Rocchiccioli C et al (1960) Spectroscopie Moléculaire – Etude Par Spectroscopie d' Absorption Infrarouge de Quelques Iodates Anhydres et Hydrates. *Compt Rend* 250 (1):1232–1234
19. Dratovsky M, Pacesova L (1968) *Russ Chem Rev* 37(4):243–254. doi:[10.1070/RC1968v037n04ABEH001631](https://doi.org/10.1070/RC1968v037n04ABEH001631)
20. Nakamoto K (1970) Infrared spectra of inorganic and coordination compounds, 2nd edn. Wiley-Interscience, New York

21. Sanyal GS, Nag K et al (1977) Thermal studies on the dimesoperiodates and iodates of Ca (II), Sr (II) and Ba (II) preparation and characterization of hexavalent iodates. *J Inorg Chem* 39:1127–1130
22. Chemla DS, Zyss J (1987) Nonlinear optical properties of organic molecules and crystals, vol 1–2. Academic Press, New York
23. Kurtz SK, Perry TT et al (1968) A powder technique for the evaluation of nonlinear optical materials. *J Appl Phys* 39(8):3798–3813
24. Bloembergen N, Armstrong JA, Ducuing J, Pershan PS et al (1962) Interactions between light waves in a nonlinear dielectric. *Phys Rev Lett* 127:1918–1939
25. Franken PA, Hill AE, Peters CW, Wernreich G et al (1961) Generation of optical harmonics. *Phy Rev Lett* 7:118–119
26. Morosin B, Bergman JG, Crane GR et al (1973) Crystal structure, linear and nonlinear optical properties of Ca (IO₃)₂·6H₂O. *Acta Cryst B* 29:1067–1072

Reduction of Power Consumption for Centrifugation Process Using Separation Efficiency Model

M.S. Salim, M.F. Abd Malek and Naseer Sabri

Abstract The power consumption, capacity, speed of rotation, separation precision and the centrifugation time are essential technical parameters of the centrifuge device, and hence their reliability may affect the system reliability and productivity. The modified period of velocity profile model (spinning period) which leads to decreasing the power consumption of laboratory centrifuge devices is derived. Based on the centrifugation time model, a fuzzy controller is proposed for the laboratory centrifuge device. The proposed controller design will produce high reliability by selecting various separation efficiencies, evaluating the separation efficiency period precisely, and reduction of separation power consumption. Based on the derived model, a low-cost controller modification leads to a shorter blood test time and lower power consumption while improving the separation efficiency to greater than 95 %.

1 Introduction

The centrifugal devices can use in a variety of medical and industrial applications. Typically, the power consumption, capacity, speed of rotation, separation precision and the centrifugation time are essential technical parameters of the centrifuge

M.S. Salim

School of Laser & Optoelectronics Engineering, AlNahrain University,
Baghdad 64074, Iraq

M.F. Abd Malek

School of Electrical System Engineering, University Malaysia Perlis,
01000 Kangar, Perlis, Malaysia

N. Sabri (✉)

School of Computer and Communication Engineering, University Malaysia Perlis,
01000 Kangar, Perlis, Malaysia
e-mail: muhsabri1967@yahoo.com

device, and hence their reliability may affect the system reliability and productivity. For this reason, centrifugation process is often obeying to developments and improvements continuously.

Continuing from our previous research [1], we derived a modified period of velocity profile model (spinning period) which leads to decreasing the power consumption of laboratory centrifuge devices. The new model has ability to estimate the power consumed for pre-defined separation efficiency.

Laboratory diagnosis of blood depends on separation efficiency which is done by means of centrifugal force and the spinning time for the blood sample. The current method of separating contents of a blood sample is not accurate in terms of centrifugation force and time. These variables are evaluated based on sedimentation theory, which calculates the sedimentation time based on higher particle density (red blood cells) and depth of sedimentation [2]. Therefore, about 0.5 ml volume of a 1 ml blood sample, the current centrifugation time is 5 min (also as recommended by the centrifuge device recommended manufacturer). To minimize or eliminate such uncertainties (due to damage of some red blood cells) from the process, we have developed a new technique that accurately predicts the required time for the separation then control the velocity profile period of centrifuge device. This technique is based on measuring the attenuation of the wave propagated through liquid its density varies with time of spinning [3, 4]. Separation efficiency of blood and plasma is evaluated empirically by percentage count for the red blood cells, white blood cells, HCT, and platelets in a sample using blood analyzer device (cell-dyne 1800). Experimental data was used to calculate the power consumption during the time taken for accurate separation. The schematic diagram of experimental attenuation measurements is shown in Fig. 1. Human blood was drawn from 84 healthy

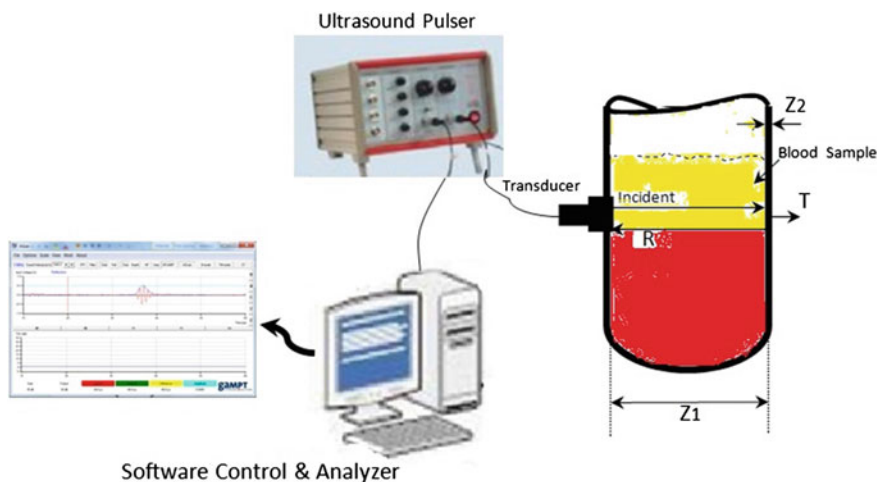


Fig. 1 Schematic diagram of experimental apparatus

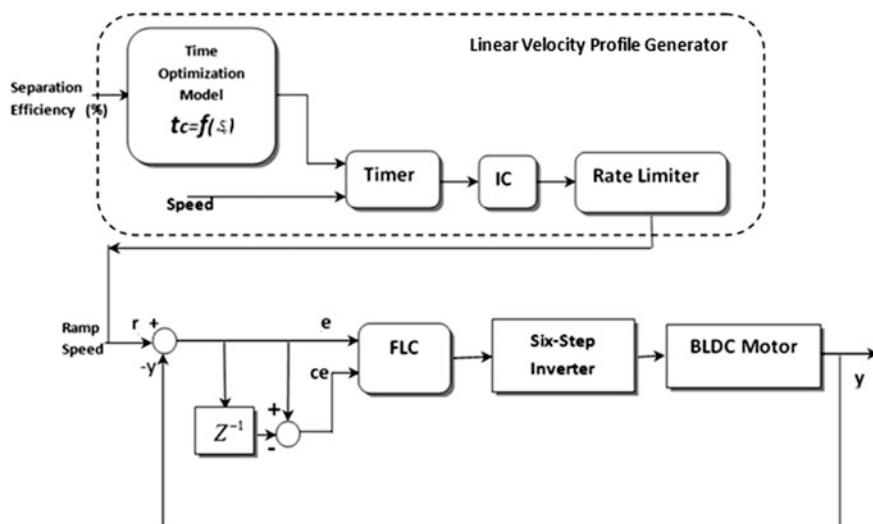


Fig. 2 Schematic diagram of proposed intelligent centrifugation fuzzy logic controller

volunteers, 23 women and 61 men; their HCT values ranged from 21 to 53.7 %; age range, 21–50 years. Samples of plasma were prepared using a serum separator tube (EDTA (K2); Demophorius Ltd, Cyprus, EU, UK). This method can also be applied to other purpose of centrifugal device, such as measurements of solid contents in liquid and sedimentation measurements [5]. Since the installation of additional measurement sensors may be costly and the sensors can also reduce the device reliability, they are not always the most cost-effective solution for the separation percentage monitoring of a centrifuge device. For this reason, the centrifuge velocity profile controller must be modified according to new mathematical model.

Based on the centrifugation time model, a fuzzy controller is proposed for the laboratory centrifuge device, Fig. 2. The proposed controller design is based on the separation efficiency and maximum spinning speed, which will produce high reliability by selecting various separation efficiencies, evaluating the separation efficiency period precisely, and reduction of separation power consumption. The design of the proposed fuzzy controller is divided into two parts, the linear velocity profile generator and the fuzzy logic controller, as illustrated in greater detail in the following sections.

2 Separation Efficiency

The various organs, tissues and cells throughout the body require oxygen from the blood. Other bodily elements also require carbon dioxide, which also depends on the cycling of blood effective techniques for separating; concentrating and accumulating

one or more of the components of blood will improve our ability to understand the properties of blood ascertain the state of an individual's health and treat bodily diseases. Therefore, accuracy in separating blood-plasma has advantages in various fields of scientific research, diagnostic testing and therapeutic treatment [6].

The separation efficiency S_e is a key factor for evaluating the performance of blood-plasma separation. It can be defined as

$$S_e = \frac{C_B - C_P}{C_B} \times 100 \% \quad (1)$$

where C_B is the cell concentration in the blood sample, and C_P is the cell concentration in the plasma, which is generated by the centrifugation process. For $C_P = 0$, the optimum separation process for plasma is present ($S_e = 100 \%$). In next sections, a relationship between power consumption of centrifuge device, separation period and separation efficiency are derived.

3 Linear Velocity Profile Generator

The proposed centrifugation velocity profile controller was designed based on the optimisation spinning time model. This model calculates the spinning time as a function of the separation efficiency. The mathematical model for the optimisation of the centrifugation time includes the acceleration and deceleration period because the blood sample is subjected to accelerated centrifugal force during the acceleration period and vice versa for the deceleration period [7].

These forces contribute the main centrifugal force (at 3,000 rpm) to separate the blood cells from the plasma. The need for a long acceleration and deceleration period (approximately 13 s) is observed because the blood cells encounter a high centrifugal force, which may cause damage. Therefore, to prevent this problem, the centrifugation speed gradually increases until it reaches a maximum (long settling time). The current periods of acceleration and deceleration are identical and equal to 13.6 s as a rapid mode operation. For this reason, as shown in Fig. 2, the FLC design begins with a ramp (linear) velocity profile generator, which is fed by the required spinning speed and separation efficiency percentage value. This generator consists of a Separation Efficiency (S_e) to Separation Period converter, a timer, an initial condition function and a speed limiter [8–10].

4 Experimental Setup

The Mathematical Model of the shorter separation process time and the separation efficiency greater than 95 % of 0.35 ml plasma, evaluated based on the mathematical model of ultrasound attenuation measurements in our previous research [1]. Utilize of

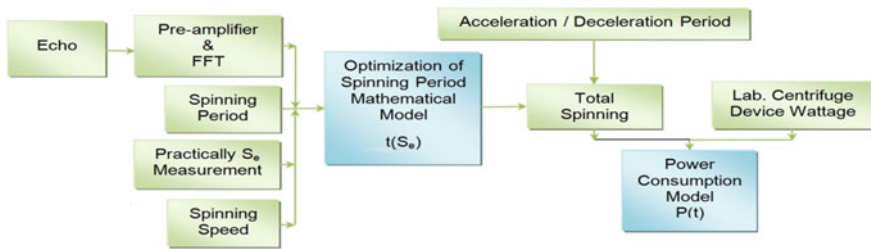


Fig. 3 Centrifugation power consumption model derivative sequence

attenuation measurements of wave propagate through a sample of blood during centrifugation process, is to calculate the separation efficiency of blood-plasma. In this research we divided the mathematical model into two models, first is the time duration of centrifugation as a function of separation efficiency model, while the second model is device power consumption as a function of time duration centrifugation technique, Fig. 3 summarize the derivative power consumption model.

The tools used in the experimental setup, as shown schematically in Fig. 1, consisted of a blood tube, the ultrasound pulser, Gampt-Echo Scan as a receiver, switch, and software analysis, while Fig. 4 shows the experimental procedure steps for evaluate the mathematical model. The transducer with center frequencies of 2 MHz was attached to the outside of the vessel. In the attenuation measurements reported here, data was obtained for the 2 MHz transducer, which gave the required detection depth and sufficient level of clarity of the separation efficiency measurements of the blood plasma. Table 1 was filled follow same procedure mentioned in our previous research for evaluating the optimization time for centrifugation process. Therefore the two part of mathematical model can be derived.

5 Centrifugation Period Model

The Separation Efficiency to Separation Period block is programmed with a Matlab package based on the optimisation centrifugation time model, which evaluates the time period required to achieve the predefined separation efficiency. The rest time in the velocity profile is determined by the time optimisation model (2) as a constant speed period. Figure 5 shows the velocity profile for 20 % separation efficiency, where the total period (T) of 20 % separation is approximately 55 s, calculated using (2) [11].

$$\begin{aligned}
 T &= t_c + t_{ACC} + t_{Dec} \\
 T &= (0.0324S_e - 0.194) + t_{Acc} + t_{Dec}
 \end{aligned}
 \tag{2}$$

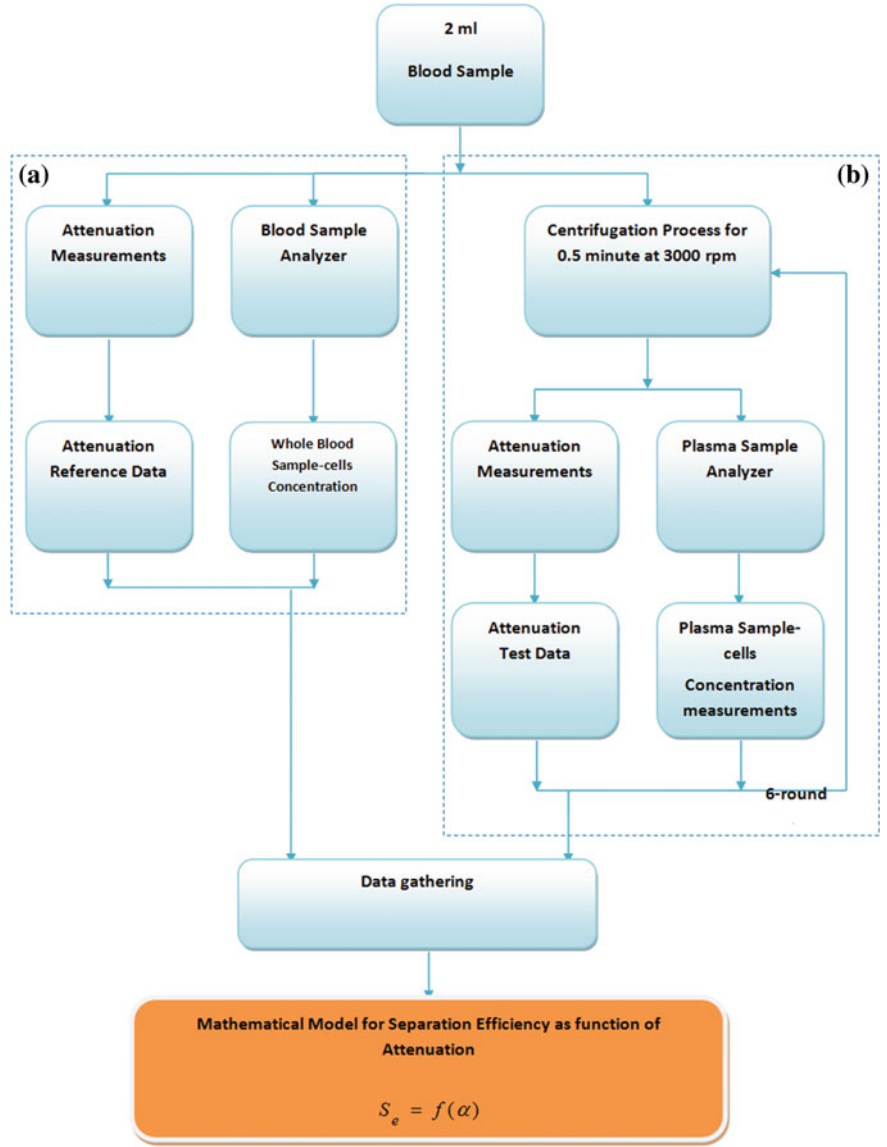


Fig. 4 Experimental procedure sequence for one blood sample test (a), Before centrifugation process, and after spinning operation (b)

6 Centrifugation Power Consumption Model

Power consumption is one of the most important measurement parameters for the design of a successful controller. The power consumption of the proposed controller can be evaluated based on the operation period of the centrifuge device.

Table 1 Evaluation of the attenuation, separation efficiency and power consumption for a discrete 5-min interval using a 180 W centrifuge device

Number of operation	Constant speed period (min)	Total spinning period (min)	Attenuation (dB/cm)	Plasma separation efficiency (%)	Power consumption (kW h)
1	0.5	0.953	−0.0257	20	0.002859
1	1	1.453	−0.0662	35	0.004359
1	1.5	1.953	−0.1202	55	0.005859
1	2	2.453	−0.1607	70	0.007359
1	2.5	2.953	−0.2012	85	0.008859
1	3	3.453	−0.2282	95	0.010359

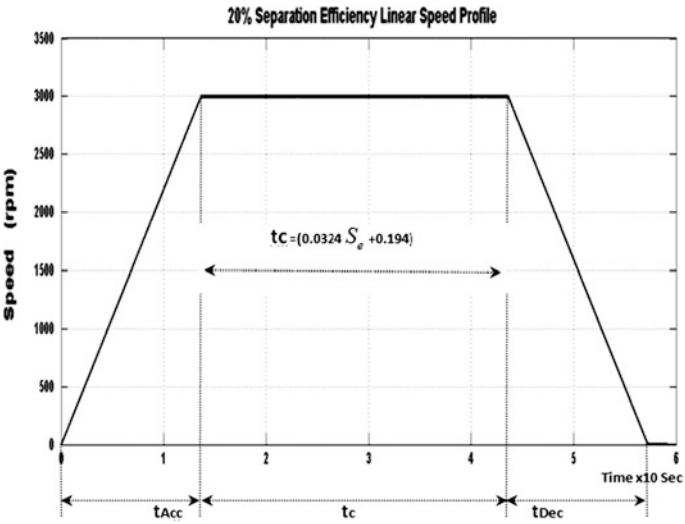


Fig. 5 Time period of linear velocity profile for 20 % separation efficiency

As mentioned in the previous section, the constant speed period (t_c) is limited by the separation efficiency value in (1). In addition, using rapid mode, the acceleration and deceleration periods are equal to 13.6 s. Therefore, the total spinning time changes based on predefined separation efficiency (t_c). Figure 6 shows the spinning period for various separation efficiencies. According to the experimental results, to obtain a volume of 0.35 ml with greater than 95 % plasma concentration, the appropriate time is 3 min. Therefore, the total centrifugation time is a 3.45 min. The electrical power consumption can be easily calculated using the following formula:

$$T_{PC}(\text{kW h}) = D_{PC}(\text{kW}) \times O_t(\text{h})$$

(3)

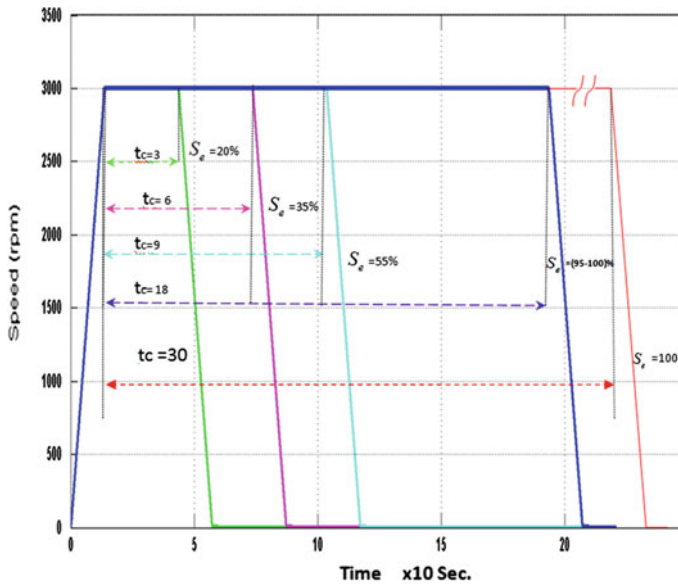


Fig. 6 Linear velocity profile period based on 20, 35, 55 and 95 % separation efficiency (S_e) using the proposed and classical methods

where T_{PC} is the total power consumption, D_{PC} is the device power consumption and O_t is the operation time period which can be formulate as:

$$O_t = T = (0.0324S_e - 0.194) + t_{Acc} + t_{Dec} \quad (4)$$

Thus, the power consumption can be rearranged as function of separation efficiency;

$$T_{PC}(\text{kWh}) = D_{PC}(\text{kWh}) \times (0.0324S_e - 0.194) + t_{Acc} + t_{Dec} \quad (5)$$

The power of the laboratory centrifuge device (Kubota corporation model 2420) used is 180 W. An evaluation of the separation efficiency and power consumption over a 5-min interval using a 180-W centrifuge device is shown in Table 1. The classical spinning time required to separate 0.35 ml of plasma from 1 ml blood with a separation efficiency greater than 95 % is 5 min according to the manufacturer's recommendations, while for proposed method, the time is 3.453 min, as shown in Fig. 6 [12].

Fig. 7 Power consumption versus 1 ml blood sample centrifugation time for 180 W centrifuge device

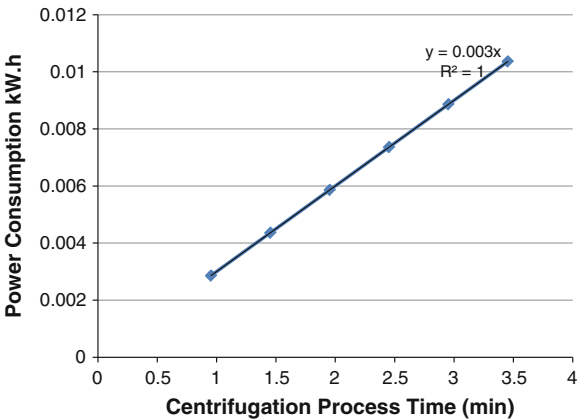


Fig. 8 The 180-W centrifuge device power consumption measurements for (20–100) % blood–plasma concentration

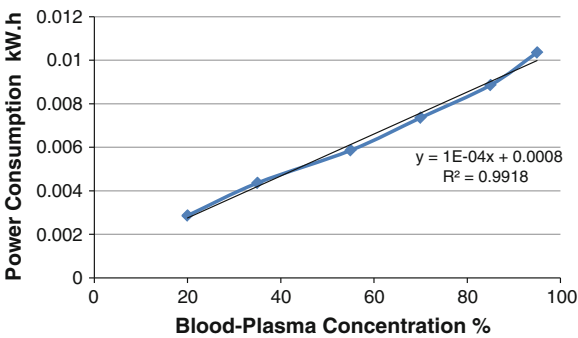


Table 2 Monthly electrical power consumption for a 180 W centrifuge device

Method type	Plasma separation efficiency (%)	Total spinning time $t_c + 2T_{Acc}$	Pc* for 100 time operating device daily (kW h)	Monthly PC* (kW h)
Classical	100	$5 + 2 \times 0.226$	1.6359	49.077
Proposed	95–100	$3 + 2 \times 0.226$	1.0359	31.077

PC* is power consumption

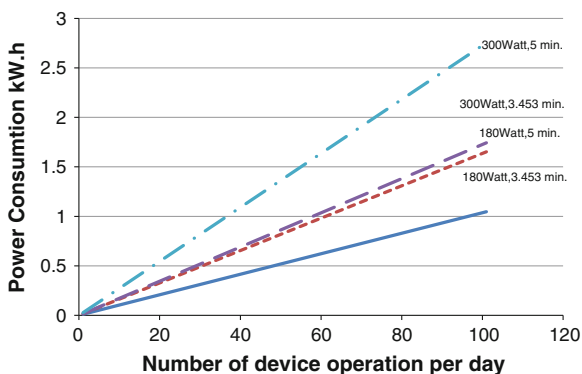


Fig. 9 Comparison between power consumption measurements results for daily operation device using classical method (5 min) and proposed method (3.453 min) for 300 W and 180-W laboratory centrifuge devices

7 Conclusions

Based on the mathematical model of the optimisation time for the centrifugation process, a mathematical of centrifugation process power consumption is derived. The centrifugation speed profile is linear, and it consists of the acceleration and deceleration period because the blood sample is subjected to an accelerated centrifugal force during the acceleration and deceleration period. These forces contribute the primary centrifugal force (at 3,000 rpm) for the precipitation of blood cells from plasma. Based on the derived model, a low-cost controller modification leads to a shorter blood test time and lower power consumption while improving the separation efficiency to greater than 95 %. In addition, increases the reliability of the centrifuge device for estimating the centrifugation period for predefined separation efficiency.

The relationships between the power consumption of the device and the centrifugation time and the plasma separation efficiency are linear, as shown in Figs. 7 and 8, respectively. As a result, the power consumption relative to the specific separation efficiency for any centrifuge device wattage can be estimated. The mathematical model of the centrifugation time derived based on attenuation measurements successfully conserved 18 kWh monthly when the device is used 100 times daily, Table 2. Figure 9 shows that the difference in the power consumption of the classical and proposed controller increases with the time of operation of the device per day in addition to the wattage of the device itself.

Additionally, the power consumption of a 180-W centrifuge device using the proposed controller is approximately equal to that of a 300-W centrifuge device using a classical controller.

References

1. Salim MS, Abd Malek MF, Heng RBW, Sabri SN, Juni KM (2011) A new measurement method of separation percentage for human blood plasma based on ultrasound attenuation. *IJPS* 6(30):6891–6898
2. Theory of sedimentation and centrifugation (2009) www.bbka.org.uk/local/iceni/bm~doc/pollensuspension-2.pdf
3. Greenwood MS, Adamson JD, Bamberger JA (2006) Long-path measurements of ultrasonic attenuation and velocity for very dilute slurries and liquids and detection of contaminants. *Ultrasonics* 44:461–466
4. Bain BJ (2006) *Blood cells: a practical guide*. Blackwell Publishing, Malden, 476p
5. Bamber JC (2004) Attenuation and absorption. In: Hill CR, Bamber JC, ter Haar GR (eds) *Physical principles of medical ultrasonics*. Wiley, Chichester, pp 93–166
6. Edward RT (2001) *An investigation of centrifugal blood-cell separation*. Ph.D. thesis, New Jersey Institute of Technology, Department of Chemical Engineering
7. Acarnley PP (1982) *Stepping motors guide to modern theory and practice*. Peter Peregrinus Ltd, London
8. Ramu K (1998) *Electronic control of machines*. ECpE 4324, A-1 Copies, Blacksburg (From the forthcoming book *Electronic control of machines*, Prentice-Hall, USA)
9. Leonard W (1985) *Control of electric drives*. Springer, Berlin
10. Chung S et al (1994) A robust speed control of brushless direct drive motor using integral variable structure control with sliding mode observer. Conference record of IAS annual meeting, vol 1, IEEE, Piscataway, NJ, pp 393–400
11. Salim MS, Abd Malek MF, Sabri N, Noaman NM, Juni KM (2013) Novel: time optimization model for centrifugation process: application in human blood-plasma separation. Measurement. <http://www.sciencedirect.com/science/article/pii/S0263224113002741>
12. Kubota Corporation–Laboratory Centrifuge Catalogue (2010) <http://www.centrifuge.jp/product/2420/index.html>

A Wide Range Measurement and Disclosure of Ultrasonic Attenuation for Very Low Concentrations of Solid in Liquids Using Exponential Pulser Technique

M.S. Salim, M.F. Abd Malek and Naseer Sabri

Abstract The low solid concentrations in solutions are discovered and measured based on measurement the attenuation that occurs with a wave propagating through a fluid. Currently methods (Pitch-Catch, Pulse-echo and Immersion method) have some of weak points related with range of detection and power consumption, therefore, to increase the measuring range to include the detection of very low concentrations in solution as well as not relying on the material that the container is made of, a new technique for the pulser was designed to overcome the disadvantages of the three currently methods. This technique has been thoroughly applied to dispensing with the use of a container or reflector with a high reflection coefficient and Expanding the measuring range to include very low concentrations. The new technique demonstrates the system's ability to distinguish the presence of particles in a fluid at a concentration lower than 10 %, which is below the limit of detection of the current method with the same device settings.

1 Introduction

To study the characteristics of fluids many researchers use common methods to measure the attenuation that occurs with a wave propagating through a fluid.

Frequently used containers or reflectors with a high reflection coefficient (H.R.C.), where the benefit of this material lies in strengthening the reflected signal to the

M.S. Salim

School of Laser and Optoelectronics Engineering, AlNahrain University, Baghdad 64074, Iraq

M.F. Abd Malek

School of Electrical System Engineering, University of Malaysia, Perlis, 01000 Perlis, Malaysia

N. Sabri (✉)

School of Computer and Communication Engineering, University of Malaysia, Perlis, 01000 Perlis, Malaysia

e-mail: muhsabri1967@yahoo.com

transducer or to generate multi echoes. Figure 1 show some of these methods which are currently used in the study of the characteristics of fluids or detecting small solid concentrations suspended in the fluid. Greenwood et al. used a container made of stainless steel to detect small concentrations of solids in a liquid using a pitch-catch method. The reason for using stainless steel is that it generates several feedback signals (echoes) resulting from transmitting a single pulse, as shown in Fig. 1a using a stainless steel container successfully to measure a concentration of less than 5 % silica (hereinafter referred to as method A) [3].

Secomski et al. developed a clinical method for noninvasive acoustic determination of hematocrit values in vivo. The value of hematocrit was determined initially in vitro from the pulse-echo measurements of acoustic attenuation. The attenuation was determined from the amplitude of echoes reflected from 3 mm diameter stainless steel affixed inside a wall opposite a transducer as shown in Fig. 1b. Using this method, we can increase the sensing of the reflected signal from the reflector which is used here to study the increase in liquid characteristics (hereinafter referred to as method B) [9]. In another method called the immersion method, it was used to immerse the transmitter and receiver in the liquid. This method does not require the use of a container with a high reflection coefficient, because of the direct measurement of the received signal from the transmitter as shown in Fig. 1c (hereinafter referred to as method C) [2, 10].

To sum up the aforementioned methods, the three methods have advantages and disadvantages. However, method A was useful in the accurate detection of small concentrations and also consumed little power; however, these features are satisfied based on using a container made of a high reflection coefficient (H.R.C.) material. Methods B, and C both used the same average transmitted power to detect very specific concentrations (as revealed concentrations depend on the power of the transmitted pulse); in addition method C depends on what the reflector material is made of.

In order to increase the measuring range to include the detection of very low concentrations in solution as well as not relying on the material that the container is made of, a new method was designed to overcome the disadvantages of the three abovementioned methods and continues to give accuracy and a wide measurement range. This new method, named Pulse Power Decay (PPD) technique (Exponential power pulser), has been thoroughly applied to the following:

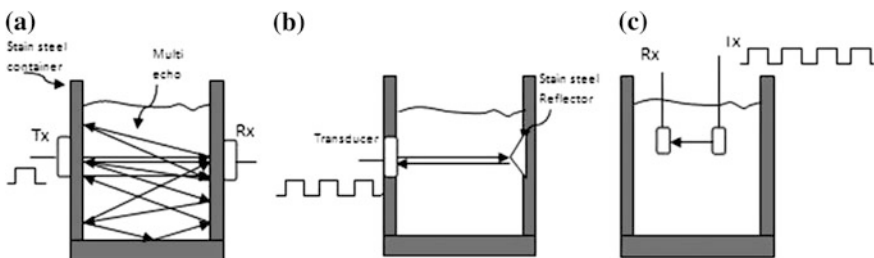


Fig. 1 Experimental method of fluid characteristics measurements and solid concentration content in liquid detection using **a** pitch-catch method, **b** pulse-echo method, and **c** immersion method

Table 1 Comparison between three currently concentration measurement methods and proposed method for transmission power, material of container is made of, measurement range

Method type	Transmitted pulses	Path (distance travelled)	Transmitted power	Material of reflector	Measurements range	PPD average power comparison	Measurements range depends on
A	Single pulse	Multi path (ND + 2W)	Constant (P_a)	H.R.C.	Wide	$P_a < P_{PPD}$	Transmitted pulse power and container type.
B	Multi pulse	Double path (2D)	Constant (P_b)	H.R.C.	Limited	$P_b > P_{PPD}$	Transmitted pulse power and reflector type
C	Multi pulse	Single path (D)	Constant (P_c)	None	Limited	$P_c > P_{PPD}$	Transmitted pulse power
PPD	Multi pulses	Single path (ND + 2W)	Decay (P_{PPD})	None	Wide	P_{PPD}	Decay power pulses

where P_a , P_b , P_c and P_{PPD} are the average power transmitted of method A, method B, method C, and method PPD respectively

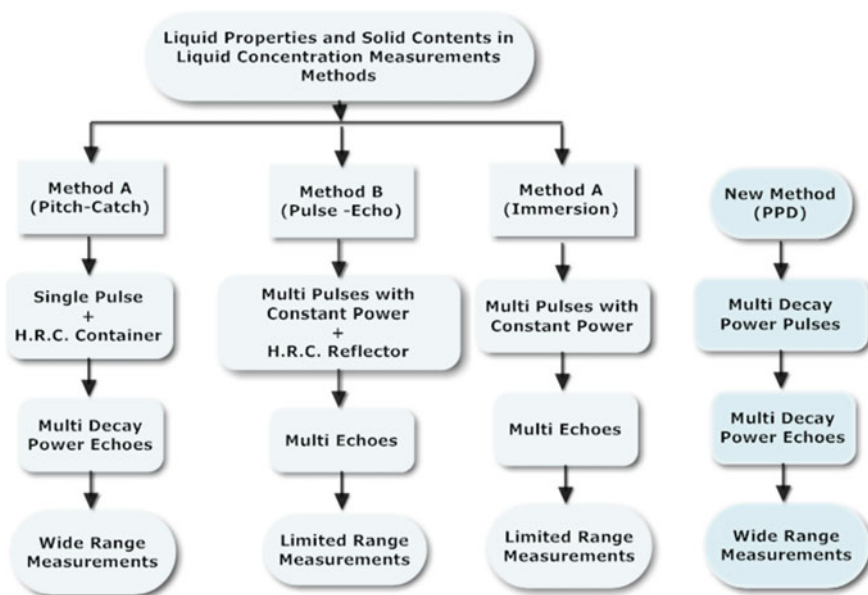


Fig. 2 Principles operation of A, B, C, and PPD method

- Dispensing with the use of a container or reflector with a high reflection coefficient (treatment method A and B).
- Expanding the measuring range to include very low concentrations (treatment method B and C).

Table 1 shows a comparison between currently three methods and the proposed method (PPD). Figure 2 illustrates the principle operations of the three mentioned methods and the proposed method. This new technique was tested using two containers made of stainless steel and Plexiglas, both containers filled in water and 5, 10, 20, 30, and 40 % Kaolin. Subsequent sections will give a detailed explanation of the new technology followed by a practical application.

2 Ultrasonic Pulser Design

The Power-Pulse Decay device (PPD) is a new technique that generates multi-decay power pulses. As mentioned above, the purpose of transmitting decay power pulses through a solution is to detect and measure various the concentrations of various particles.

A block diagram of the proposed ultrasonic pulser system (PPD) is shown in Fig. 3a. The PPD device consists of six hardware parts, panel control software, an I/O serial interface, the main controller, a high DC power supply, a pulse generator

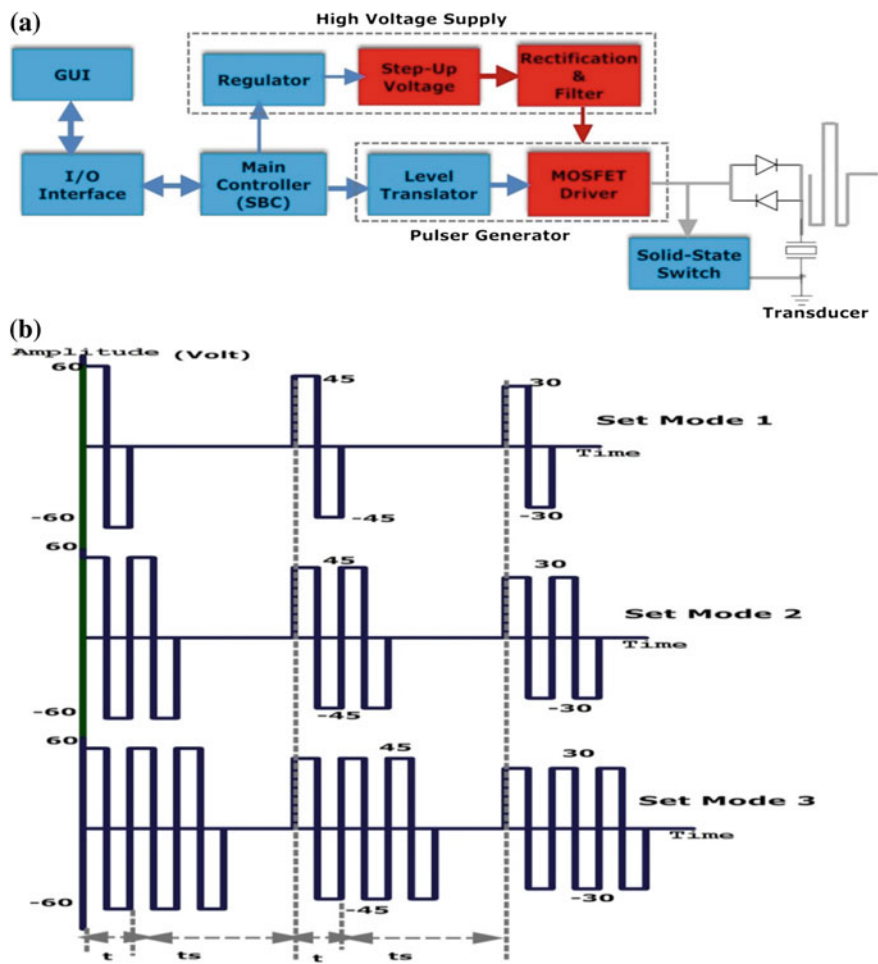


Fig. 3 A block diagram of the pulse-power decay ultrasonic pulser system (a), types of pulser sets mode (b)

and a solid-state switch. The transducer is driven by pulses from the decay power pulse generator, which delivers sets of high-voltage bipolar pulses. The advantage of using bipolar pulses is that the peak-to-peak pulse voltage can measure twice the voltage rating of the coaxial cable connecting the pulser and the transducer [11].

In addition, a bipolar pulse has a lower undesirable DC and a low-frequency component that may increase the leakage current compared to its unipolar counterpart. As a result, the size and cost of the cable can be substantially reduced, especially when a multi-element transducer array is used. The amplitude of the pulses in one set is determined by the output voltage of the high DC power supply, which converts the battery voltage of 12 V to $\pm 20\text{--}75$ VDC [1].

The amplitude of the received signal can be improved by increasing the pulse train amplitude, using a container with a low reflection coefficient and by driving the transducer at its resonance frequency. The frequency of the set pulses can be generated either by a programmable oscillator or by an external clock in which both methods are controlled by the main controller via control panel software. The pulse frequency, decay power factor, set size, time interval, set mode and start amplitude are programmed by the main controller with software control implemented by Visual Basic (Microsoft Inc.) [4].

Five modes of the pulse set are available with the PPD device; in the case of mode one, the pulse set transmits one pulse with the proper decay amplitude during one time interval, as shown in Fig. 3b. The schematic of the PPD device is shown in Fig. 3.

For the proposed technique with any experimental setup, the transducer records only the first individual signal (first echo), which corresponds to travel through the liquid. The first echo is the signal of interest (IS), and it occurs at the shortest time. To evaluate the performance of the PPD technique, a laboratory experiment in an industrial application field was conducted to detect and measure the concentration of kaolin in water. After verification of the performance of the PPD, this technique was adopted in our research. To obtain the attenuation measurements as a function of the frequency, the Fast Fourier transform (FFT) of the signal of the peak of interest (IS) for each pulse in the set was obtained for water as a reference and for the slurry [5, 8].

The data were obtained in ten steps, using the appropriate decay factor for the transmitted pulses and the receiver gains. Overlapping signals of interest (IS) were used to determine the effect of changing the amplitude of the pulse travelling through the *slurry*, and the data were normalised to a pulser voltage of ± 75 V. Additionally, the data were corrected for the receiver gain. The results of the FFT analysis of the echoes were evaluated for *water*, and 10 % kaolin, Table 2.

3 Validation of PPD Design

The new PPD technique can be adaptive with wide application fields. In this research, the objective of the experiments was to evaluate the performance of PPD technique to gather ultrasound measurements for different percentages of solid concentrations in liquid. Invasive measurement is the most direct method of

Table 2 One set size 10, FFT amplitude of IS for tap-water and 10 % kaolin at 2 MHz

Pulse no.	Tape-water, FFT amplitude at 2 MHz	10 % kaolin, FFT amplitude at 2 MHz
1	3.51	3.262
3	2.089	1.558
5	1.243	0.758
7	0.863	0.505
10	0.425	0.208

achieving an accurate measurement. Hence, this method was used for all the experiments carried out for this paper. Using the assumption that 1 ml of tap water equals 1 g in weight, kaolin was mixed with tap water in concentrations of 10, 20, 30 and 40 % to create the different sets of slurries. These concentrations were measured by the weight of kaolin and tap water.

The vessel, shown schematically in Fig. 4, consists of a Plexiglas container and two transducers with a centre frequency of 2 MHz affixed to the outside of the container on opposite sides, using a pitch-catch method. The vessel walls have a thickness w of 3.2 mm and the inside walls are separated by a distance D of 4 cm. The temperature, measured by temperature sensor (LM35), is recorded. For simplicity, the temperatures for all the experiments were kept at room temperature (20 °C), between ± 1 °C [10]. When the vessel is filled with water, less than 10 % of the ultrasound is reflected at the Plexiglas-water interface and the rest is transmitted into the water. At the opposite wall, 91 % is reflected at the water-Plexiglas interface. With the pitch-catch mode operating, the receive transducer records only the first individual signal, as it corresponds to travelling through the slurry. The “signal of interest” (IS) occurs at the shortest time and the path length for the Nth IS (for example the interest received signal of the fifth pulse of set) in the pitch-catch mode is $(2w + D + v \times T)$, where v is the speed of sound through material and T is the total time delay of Nth pulse; and calculated using (1):

$$T = (N - 1)(t_s + t) \quad (1)$$

where the t_s is the pulse spacing time and t is the bipolar pulse width. To obtain attenuation measurements as a function of frequency, the fast Fourier transform (FFT) of the peak of data were corrected for the receiver gain. The results of the FFT analysis of the “peak of interest” are shown for water, 10 %, and 40 % kaolin, Fig. 5. The effects of attenuation for water are clearly observed as the peak amplitude shifts to a smaller frequency as the IS number increases, as shown in Fig. 6.

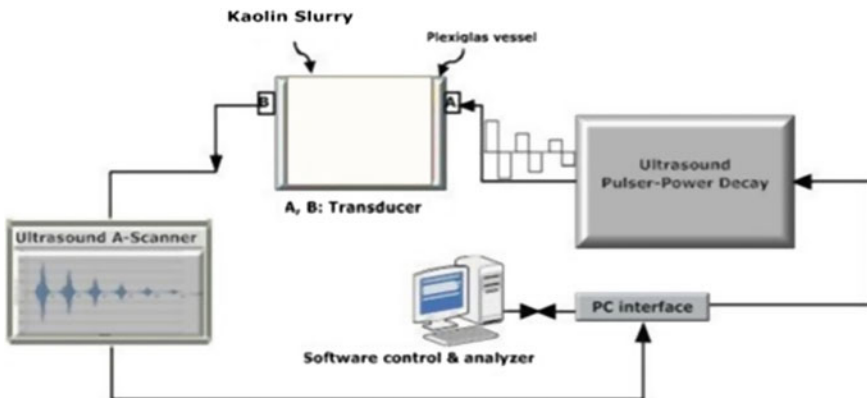


Fig. 4 Sketch experimental setup

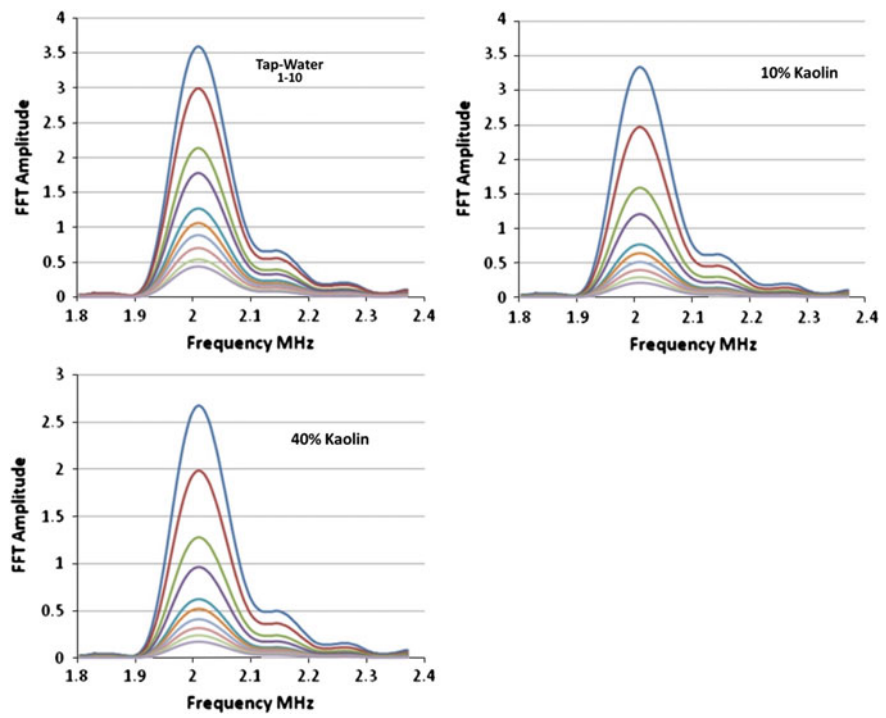
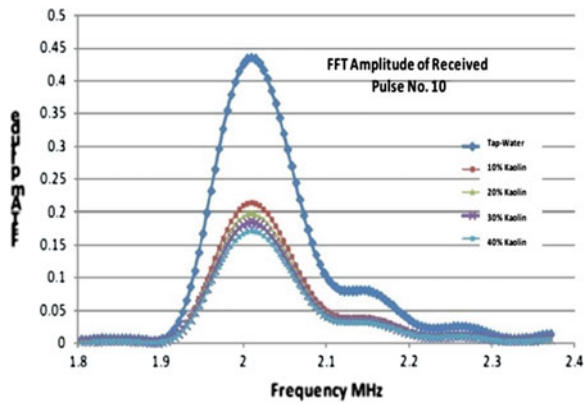


Fig. 5 FFT amplitude for one transmitted set size 10, propagated through water (as reference data), and test 10 % kaolin, and 40 % kaolin

Fig. 6 FFT amplitude comparison of 10th interest signal received for reference data (water) and test data 10, 20, 30, and 40 %



4 Discussion

Data similar to were obtained for the kaolin slurries of 10, 20, 30, and 40 % kaolin by weight. The attenuation due to the 10 % kaolin slurry can be observed in Table 1 by the increasing separation between the FFT amplitudes for water and the slurry. The data for kaolin slurry 10 % wt concentration were compared with that for water by evaluating the natural logarithm of the ratio (FFT amplitude slurry/FFT amplitude water) for a specified frequency (2 MHz). The data exhibit straight lines and the interest signal (IS) for each set pulse was obtained for water and for the slurry. The data were obtained in ten steps, using appropriate decay factor for the transmitted pulses and the receiver gains. Overlapping interest signals (IS) were used to determine the effect of changing the pulser voltage travelling through slurry, and the data were normalized to a pulser voltage of ± 75 V. Additionally, the slopes (S) were obtained. The lower pulses-power such as 40, 35, 30 V greatly affected by low concentrations of slurry than the higher pulse power for one transmitted set as shown in Fig. 3b. For example, the received pulse number 10 has 0.425 V while the amplitude of same pulse travelled through 10 % kaolin was 0.208 V (Table 1); therefore, to detect low concentrations with a long range of measurement, the set size should be increased (i.e., decreasing the decay factor leads to incremental increases in the amount of pulsed-power decay). Decreasing the decay factor leads to an increase in the number of pulses in one transmitted set and causes more generated lower-power pulses; this increase is limited by the high voltage supply. Due to the attenuation of the ultrasound waves propagated through slurry, the relationship between kaolin concentration and the attenuation was determined. The attenuation (α) was calculated using (3) [2].

$$\alpha = \frac{-20}{D} \log_{10} e^S \quad (2)$$

For a given frequency, a plot of the attenuation (α) versus the kaolin weight percentage (w) displays a straight line, Fig. 7, as expected. For example, when the

Fig. 7 Attenuation measurement (dB/cm) versus kaolin concentration wt%

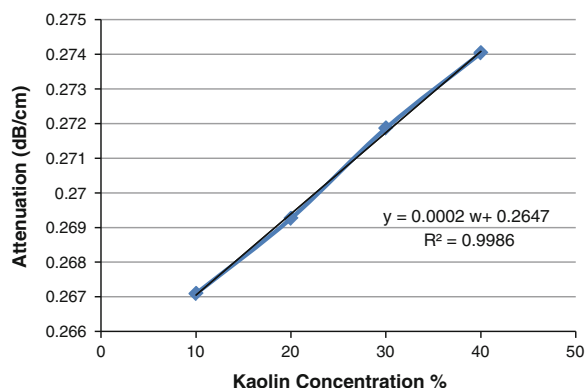


Table 3 The detection comprise between commercial and proposed technique for 10 % kaolin

Set pulse no.	Amplitude of commercial pulser (V)	Amplitude of PPD pulser (V)	Tap water FFT amplitude of		10 % kaolin FFT amplitude of	
			Commercial pulses	PPD pulses	Commercial pulses	PPD pulses
1	75	75	3.51	3.51	3.496	3.494
5	75	55	3.51	1.243	3.496	1.05
10	75	30	3.51	0.425	3.496	0.281

values of the attenuation at 2 MHz are extracted from Fig. 7, such a plot shows the following linear relationship:

$$\alpha = 0.0002w + 0.2647 \quad (3)$$

where the slope has units of dB/cm wt%. Table 3 illustrates the differences between a commercial pulser and the PPD technique.

5 Conclusion

In this research, we developed a self-contained prototype of an ultrasonic decay power pulser system for the detection and measurement of the particle concentration in a fluid. The new technique demonstrates the system's ability to distinguish the presence of particles in a fluid at a concentration lower than 10 %, which is below the limit of detection of the current method with the same device settings.

The PPD ultrasound system that we built generates a multi-pulsed powered decay train for transducer excitation instead of a single negative pulse, which is used in the current pulser. The PPD also delivers variable transmitted ultrasound energy that will increase the level of the received signal interest for SNR improvement, and the accuracy of the concentration measurement is increased. The number of pulses per transmitted set is determined by the user based on the concentration detection level that is required. Although the pulser can generate pulse amplitude up to 75 V, we found that a pulse set with a set size of 10 and initial amplitude of 75 V was sufficient to distinguish the presence of particles in a fluid at a concentration of less than 10 %. The detection of concentration can be improved by increasing the set size, the pulse amplitude that was optimised in this study and the gain of the interest signal amplifier and by optimising the filter's receiver parameters. The interest signal (IS) was digitised and analysed by an ultrasound A-Scan device. The experimental results for the measurement of kaolin concentration based on the PPD technique are shown in Table 3. It can be clearly observed that the lower-order amplitude pulses are more attenuated than the first-order pulses (higher power pulses). *For example*, the echo of the tenth pulse is 0.425 and 0.281 V through water and water 10 % kaolin, respectively, while the echo for first pulse

is 3.51 and 3.494 V. The new technique exhibited high performance as an instrument for measuring the percentage of sedimentation [6, 7], and it is applicable with all container materials. The amount of ultrasound energy received by the transducer can be increased using a transducer with a larger area or by adjusting the position of the transducer and the angle so that the maximum IS amplitude is sensed. The comparison results reveal that the PPD technique is more efficient than the classical technique for the pulser, especially for the detection of very low concentrations (5 %) of particles in a fluid.

References

1. Basso C (2008) Switch-mode power supplies spice simulations and practical designs, 1st edn. McGraw Hill, New York, pp 175–212
2. Chen Q, Freear S, Cowell DMJ (2007) Measurement of solid in liquid content using ultrasound attenuation. In: 5th world congress on industrial process tomography, Bergen, Norway
3. Greenwood MS, Adamson JD, Bamberger JA (2006) Long-path measurements of ultrasonic attenuation and velocity for very dilute slurries and liquids and detection of contaminants. *Ultrasonic* 44:e461–e466
4. Kainka B, Berndt HJ (2001) PC interfaces under Windows. Elektor Electronics (publication), Dorchester, pp 21–34, 248–267
5. Parker D, Lec RM, Pendse HP, Vetelino JF (1990) Ultrasonic sensor for the characterisation of colloidal slurries. In: IEEE ultrasonics symposium, pp 295–2981
6. Salim MS, Abd Malek MF, Heng RBW, Salim NS, Juni KM (2011) A new measurement method of separation percentage for human blood plasma based on ultrasound attenuation. *Int J Phys Sci* 6(30):6891–6898
7. Salim MS, Abd Malek MF, Sabri N, Noaman NM, Juni KM (2013) Novel: time optimization model for centrifugation process: application in human blood-plasma separation. *Meas J*. <http://www.sciencedirect.com/science/article/pii/S0263224113002741>
8. Schmerr LW Jr, Song SJ (2007) Ultrasonic nondestructive evaluation systems: models and measurements. Springer, New York, pp 439–455
9. Secomski W, Nowicki A, Guidi F, Lewin PA (2003) Noninvasive in vivo measurements of hematocrit. *J Ultrasound Med* 22:375–384
10. Shung KK (2006) Diagnostic ultrasound imaging and blood flow measurements. Taylor & Francis Group LLC, Boca Raton, pp 185–210
11. Tang SC, Clement T, Hynynen K (2007) A computer-controlled ultrasound pulser-receiver system for transskull fluid detection using a shear wave transmission technique. *IEEE Trans Ultrason Ferroelectr Freq Control* 54(9):1772–1783

Wireless Sensor Network Wave Propagation in Vegetation

Naseer Sabri, S.A. Aljunid, M.S. Salim, S. Fouad and R. Kamaruddin

Abstract The Wireless Sensor Network (WSN) total path losses of a greenhouse based on the two popular empirical vegetation attenuation models are used to predict the connectivity and the maximum coverage of wireless nodes within the communication path. The foliage imposed effect on the propagating waves is examined, simulated and the total path losses concluded as a function of antenna height and a separation distance of WSN nodes in a field of various densities of vegetation inside a greenhouse. The implemented library of foliage propagation model can be embedded easily with other WSN simulator platforms. The best antennas height based on greenhouse environment and total path loss is shown to be with the 3.5 m and 1 m height for transceivers of main and end nodes, where less total path loss is obtained and perfect connectivity of (100 %) when used with MED vegetation models for all vegetation depths, less than 50 m, while ITU model shows perfect connectivity for same height combination but with less foliage depth of 40 m while it shows 88 % connectivity for higher foliage depth than 40 m.

1 Introduction

Wave propagation analysis based on environmental modeling provides a good initial estimate of the signal characteristics. The ability to have an accurate prediction of radio propagation behavior of a wireless communication system, such as wireless sensor network, is becoming crucial to design scheme of a system where radio wave propagation throughout a foliage medium induces an additional excess

N. Sabri (✉) · S.A. Aljunid
Computer and Communication Engineering School, UniMAP, Kangar, Malaysia
e-mail: nasseersabri@yahoo.com

M.S. Salim · S. Fouad
Laser & Optoelectronics Engineering School, AlNahrain University, Baghdad, Iraq

R. Kamaruddin
Bioprocess Engineering School, UniMAP, Kangar, Malaysia

loss on the propagating components such as direct and reflected waves. Therefore, to provide reliable and enhance network coverage and connectivity of WSN, a precise study and modeling of these effects must be achieved with adequate prediction of the propagation losses [1, 2, 3].

Most of the radio propagation channel models, in a simulation environment, assume an obstacle free propagation channel, and hence, a clear direct path exists between the communication nodes. Thus, in a simulation environment, the performance of wireless communication system is estimated based on those models and the results produced often poorly reflect the real scenario, such as plantation area, in which the presence of vegetation significantly affects the communication between network nodes. As recommended by Meng et al. [4], for short range near ground plantation environment, propagation loss is modeled by an integration of the foliage imposed effect and the effect from the radio wave reflections from the ground and possibly tree canopy. There are different vegetation propagation models reported in published literature [3, 5, 6] and these models are verified by numerous experimental studies in Savage et al. [7], Meng et al. [4], Morataitis et al. [8].

In this chapter, the foliage imposed effect on the propagating waves is examined, simulated and the total path losses are concluded as a function of antenna height and the separation distance of WSN nodes in a field of various densities of vegetation inside a greenhouse. The WSN total path losses of a Greenhouse based on the two popular empirical vegetation attenuation models, namely Weissberger Modified Exponential Decay (MED) and ITU-Recommendation (ITUR) models are modeled, simulated and analyzed to predict the connectivity and the maximum coverage of wireless nodes within the communication path. The implemented library of, foliage propagation model can be embedded easily with other WSN simulators like OMNeT++, NS-2, OPNET simulation platforms [9–11].

2 Precision Agriculture Based WSN

Precision agriculture provides the means for monitoring, evaluating and controlling agricultural practices. It covers a wide range of agricultural trends from daily herd management through horticulture to field crop production. Continuous monitoring of individual crop and its necessities will help growers to potentially recognize a variety of fertilizers, irrigation, ventilation, climate automation and other requirements that provide optimal environment and conditions for crop growth [12]. To determine the behavior of electromagnetic waves, a precise model of propagation must be adopted, however propagation models normally used in wireless communication might not be precisely described the wireless sensor network [13]. WSN nodes are spatially located; usually near the earth's surface thus induce absence of main ray between sender-receiver nodes, although WSN nodes have spatially short distance distribution. Therefore, WSN propagation waves may face obstacle like trees, fence, building and dense foliage.

The newly emerged WSN technology has spread rapidly into various multi-disciplinary applications. Agriculture and farming is one of the industries which have recently redirected their concentration to WSN, looking for this cost effective technology to improve its production and enhance agriculture yield standard [14]. There are several test points equipped with wireless nodes of various sensor types which they are used to collect the local climate, fertilizing and irrigation parameters within different parts of the greenhouse to provide information for successful greenhouse automation system. Cabling of sensor nodes would make the measurement system expensive and vulnerable. Furthermore, the cabled test spots are hard to relocate once they are installed. Thus, WSN is an attractive and cost competent alternative to build the required measurement system. Such as temperature, humidity, light and the carbon dioxide are the most important factors for the quality and productivity of plant growth. Hence continuous monitoring of these environmental variables gives information to the grower to better awareness, how each factor impinges on growth and how to administer maximal crop productiveness [15]. The optimal greenhouse climate adjustment can enable us to improve productivity and achieve remarkable energy savings [7].

3 Physical Aspect Challenges of Applying WSN in Agricultural Fields

Real time radio propagation in a real environment is complex due to the existing of multipath propagation, shadowing and attenuation. In the agriculture field, the wireless communication face challenges in terms of nodes location for wide area mesh coverage and reliable communication link quality above crop canopies. WSN must have the ability to work in a wide range of environments such as bare fields, orchards, vineyards from flat to complex topography and with various weather conditions, all of which affect radio performance [1, 16, 17]. In these environments and situations, the terrain, crop growth, nodes spacing, antenna height and in addition to more common factors will affect the link power budget. For applications inside buildings like greenhouses or warehouses, the radio signal has to go many objects like windows, walls, pallets, machines, which indeed cause a significant degradation in received signal strength. Mostly a 10–20 dB of a received signal level above the sensitivity threshold of the receiver is an accepted value for the link budget [18]. As the crop growth, the density of the leaves increased over time and thus message rate decreases, while when there are less leaves; the message rate increases. Signal propagation above the cross canopy results in attenuation and variance in the received signal strength [19]. Hebel shows that the attenuation and signal strength variance were dependent on line of sight losses and heights less than the Fresnel zone radius [20]. Experimentation in mature corn fields (2.5 m high) with transceivers placed at antenna heights of 1.5 m and 2 m and a distance of 100 m, showed an average 10 dB loss when the transceivers were placed in or across the corn rows [18, 21].

4 Radio Wave Propagation Models in Simulation Tools

The propagation wave between transmit and receive antenna of a radio link is subject to a variety of effects that can alter its amplitude, phase, or frequency. Such propagation effects include; reflection from the ground or large objects, diffraction from edges and corners of terrain or buildings, scattering from foliage or other small objects, attenuation from rain or the atmosphere and Doppler effects from moving antenna.

WSNs simulators use Propagation models to calculate the path losses due to wireless channel impairment. These models are based on real measurements and represent a statistical mean or median of the expected path loss. The predicting of the average received signal power at a given distance from a transceiver node yield to estimate the wireless communication coverage and hence, the calculation of the success rate of packet reception. Despite that, most simulators assume a fix propagation distance for each node where the signal propagates exactly meters or using of free space propagation model (FSPL) or Two-Ray propagation model (T-R). Therefore, radio wave propagation must encounter the impairment of various types of obstacles found between the transceivers and hence more complicated path loss models must presented than these simple models.

Therefore, WSNs simulators need for more efficient and realistic propagation models that encounter the different effects of the environments under vision. The modeling of a common path loss model for all communication systems that take into account all various obstructions, paths, terrain and atmospheric conditions is difficult and not strategic solution. As a result, there are different models found in different types of radio wave propagation for different conditions and phenomena. For agricultural application the foliage effects will represented by the two well known foliage models MED and ITU, and the total path losses will be calculated based on fusion of foliage and FSPL, T-R path loss models.

4.1 Free Space Path Loss (FSPL)

FSPL is a wave propagation model that assumes an ideal propagation conditions along the way between transceivers, it is widely used model in simulation environments. Friis free space equation provides a prediction of the received signal strengths when the transmitter and receiver have a clear, unobstructed line of sight path between them [22, 23]. The FSPL is calculated by;

$$P_{Loss} \text{ dB} = 32.45 + 20 \log_{10}(d_{\text{km}}) + 20 \log_{10}(f_{\text{MHz}}) \quad (1)$$

where f is the frequency in MHz and d is the distance in km between the transceivers. FSPL implies that the received power decreases with distance at a rate of 20 dB/decade. Thus, the received signal power falls off inversely proportional to the square of the distance d between the transmitter and receiver antennas.

4.2 Two-Ray Ground Reflection Model

Ray tracing is a method that uses a geometric approach, and inspects what paths the wireless radio signal follows from transmitter to receiver. Ray-tracing prediction models are good when detailed information about the area is obtainable. While the predicted results may not be applicable to other locations, thus making these models site specific. The fact that for most wireless propagation situations, two paths are found from transmitter to receiver antennas: a direct path and a reflected off the ground path. This model gives a more accurate prediction at a longer distance than the FSPL [22, 23].

Typically in WSNs, wireless sensor nodes are deployed with small elevation from the ground, nearly (0.05 m – 1 m) with low transceiver antenna heights [4]. Therefore, T-R model is a better descriptor for path loss when radio waves propagate near the ground rather than the FSPL model [23]. The T-R path loss model can be calculated by (2).

$$L_{PE}(\text{dB}) = 40 \log_{10}(d) - 20 \log_{10}(h_t) - 20 \log_{10}(h_r) \quad (2)$$

To maximize the connectivity of network nodes due to plane terrain ground reflection the antenna height must clear the first Fresnel zone [22]. The Fresnel zone radius, r in meters, can be calculated by (3).

$$r = 8.657 \sqrt{\frac{d}{f}} \quad (3)$$

where, d is the separation total distance between transceivers in km and f is the frequency in GHz.

The Fresnel zone is the area which includes emitted power. Objects within this zone can cause signal degradation and power defeat from the transmitter to the receiver. The more obstacles within the Fresnel zone, the more wave reflections and phase shifts will be induced, that can lead to a loss in the received power. Moreover, as antenna height gets close to the earth's surface (as with the case of our WSN), the earth's surface goes in the Fresnel zone, making an impediment to communication [24].

5 Wave Propagation Models in Vegetation

Most terrestrial wireless communication systems may require signals to pass through the foliage at some area along its propagation path. Thus, many applications of WSNs, such as military and agricultural applications, may face the challenge of impairing the signal due to the existence of vegetation and crops. The attenuation of radio wave that propagates through vegetation region is considerable, especially at higher frequencies [5]. Many valuable research have been conducted

to model these effects experimentally, most of them are formulated the attenuation as a function of foliage depth and the working frequency. However, the wide range of conditions and types of foliage makes it difficult to develop a generalized prediction procedure.

In the literature, several models to evaluate the excess attenuation due to the presence of foliage in the propagation path can be found. The majority of those models are defined by the expression of (4) [25]. In WSN for agricultural application, a well known models such as Weissberger's Modified Exponential Decay model [6] and ITU Recommendation (ITU-R) [3] is adopted in this work.

$$L_{veg} = A \times f^b \times d^c \quad (4)$$

where L_{veg} is the excess attenuation due to foliage, f is the working frequency and d_f is the depth of a deciduous tree in meters, A, B and C are empirically calculated constants, which are dependent on the type of foliage.

5.1 Weissberger's Model

Weissberger's Modified Exponential Decay model [6, 26, 27] is given by (5). MED model applies when there are dense, dry, leafed trees blocked the propagation path between transceiver antennas within maximum distance of 400 m. Blaunstein [2] indicates that the model covers the frequency range from 230 MHz to 95 GHz.

$$L(\text{dB}) = \begin{cases} 1.33 \times d_f^{0.588} \times f^{0.284} & 14 \leq d_f \leq 400 \\ 0.45 \times d_f \times f^{0.284} & 0 < d_f < 14 \end{cases} \quad (5)$$

where d_f is the depth of foliage along the LOS path in meters f is the frequency in GHz. Attenuation predicted by the vegetation propagation model is in addition to free space and any other non foliage loss.

5.2 Early ITU Vegetation Model

ITU-R model estimates the path loss encountered due to the presence of vegetation along the propagating path. It is an easily applied model that provides results that are fairly consistent with the Weissberger model. The model was proposed for cases where either the transmitter or the receiver is near to a small ($d < 400$ m) grove of trees such that the majority of the signal propagates through the trees. The model is represented by (6) [3].

$$L(\text{dB}) = 0.2f^{0.3}d_f^{0.6} \quad (6)$$

where f the frequency in MHz is d_f is the depth of the foliage along the LOS path in meters.

6 Total Path Loss in Vegetation

The modeling of path propagation loss of a communication channel through vegetation is the integration of foliage loss with the free space loss. The total path loss can be formulated as:

$$P_{\text{Tot-loss}} = P_{\text{FSPL}} + P_{\text{Env-loss}} \quad (7)$$

where $P_{\text{Tot-loss}}$ is the total path loss, P_{FSPL} is the FSL (T-R MODEL) and $P_{\text{Env-loss}}$ is the vegetation loss. The resulting signal power received is given by

$$P_r = P_t - P_{\text{Tot-loss}} \quad (8)$$

While the maximum allowable path loss for WSN can be approximated by

$$\text{Loss}_{\text{thr}} = P_t - R_{\text{sensitivity}} \quad (9)$$

where $R_{\text{sensitivity}}$ is the receiver sensitivity, Loss_{thr} is the maximum path loss in dB. The total path loss cannot exceed Loss_{thr} value without violating the receiver sensitivity. Therefore, for WSN nodes being in the range of communication coverage, the separation distance d must be less than maximum propagation distance d_{max} . The existence of a radio link between nodes is computed by comparing the P_r calculated using propagation model to the Receiver Sensitivity threshold ($R_{\text{sensitivity}}$). While the packet is successfully received by a receiving node if the P_r is greater than $R_{\text{sensitivity}}$ of the receiver and no other packets being sent or received. Contrarily, if P_r is less than $R_{\text{sensitivity}}$, then messages are treated as noise.

A new library (Veg_Path_Loss) of radio wave vegetation loss is built which can be integrated with simulators based C language structure. This library is used to simulate and calculate the total losses based on foliage losses of well known MED or ITU vegetation models with the FSPL and T-R propagation path loss model, also the communication coverage and network connectivity have been included and simulated results can be used to indicate the reliability and performance of the network as well as its provide a tool to investigate the configuration of node placement with respect to a connectivity issue.

7 Simulation Setup of WSN in Agricultural Greenhouse

The simulation assumes a greenhouse of $100\text{ m} \times 100\text{ m}$ area where sensor nodes are deterministically distributed over this area with 10 m separation distance between adjacent nodes. The total number of sensor nodes is assumed to be 100 nodes. This assumption is based on harum manis mango cultivation plan for all season in big greenhouse structure in Perlis, Malaysia. The greenhouse terrain is flat and it mainly consists of sand and soil, with grass covers some parts.

The WSN based agricultural greenhouse application is configured based star topology with the main sink node in the middle. Deployment of wireless sensor nodes in the field could be either random or deterministically based on the application. However, proper planning for deployment of network nodes is extremely important since without proper nodes locations will highly affect the network connectivity and may result in sever network performance degradation. Within greenhouse field, the WSN nodes are deployed based on a pretested of signal strength represented by link quality indicator (LQI) of sensor node within the main node, this test is helpful to guarantee that nodes is at least seen by the main node in free space environment. While the environment is changeable within time due to plantation and growth which affects the communication channel.

To achieve a successful network communication environment, network nodes must be within the communication range. Despite that, there are many scenarios to deploy the network nodes, but the most effective one is the grid based deployment which is one of the most suitable for agricultural application and especially for greenhouse environment. There are different popular pattern of grid mapping of nodes such as square grid, equilateral triangle and normal hexagon. It is seen that the square grid be most effective when a main node located in the center of the grid [28]. Therefore, the square grid mapping is adopted for network node deployment with the main node located in the center of the topology. In this simulation the IEEE 802.15.4 standard is used with operating frequency of 2.4 GHz. The WSN deployed based on non beacon enabled star topology. The wireless channel bit rate is 250 kbps. The transmission power of 0 dBm is chosen for all network nodes and the sensitivity value of -95 dBm based on Omni directional antennas with a typical gain of 1 dBi are used at in the transmitter and receiver.

The integration of free space and T-R path loss models within foliage models is simulated for different vegetation depth and also for different antenna heights. Therefore, the communication network coverage and the connectivity of the nodes within a greenhouse environment based on grid deployment of 10 m apart nodes, within main node in the center are conducted. The antenna height is experience to find the best heights configuration to ensure better link connectivity for various vegetation depths.

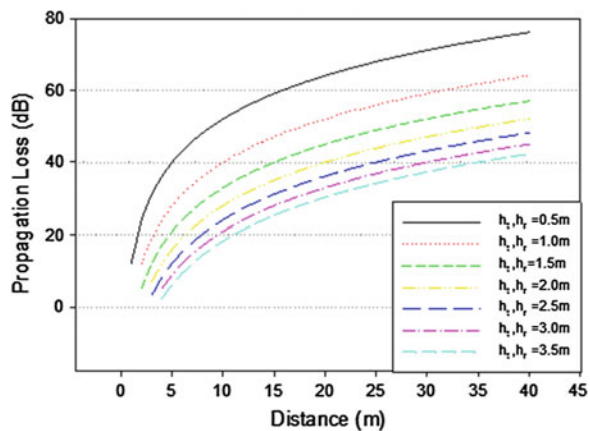
8 Results and Discussion

The impact of antenna height, for the transmitter and the receiver is investigated based on (2). Figure 1 shows the loss as a function of different distances of equal antenna height, as the antenna is far away from the ground, the loss is less, for comparison, antenna height of 0.5 m of a system has a path loss of 64 dB while the same system with an antenna height of 2 m has a loss of 40 dB at 20 m of distance, this reduction in loss must be considered when designing the wireless node's height so maximum height subject to applicable limits which minimize the path loss must be chosen. For application of an agricultural field, the antenna height of the transmitter may be differs in height from the receiver antenna node.

Figure 2 shows that antenna height of a transmitter is above ground by 3.5 m, while the receiver nodes antennas may vary between 0.5, 1, 1.5 or 2 m above ground level. Obviously the attenuation for such system will be higher than a system with same level antenna height (LOS), although this, some application need like this configuration and the path loss prediction computed shows that 1 m receiver antenna height is the most suitable height for agricultural application and for greenhouses specifically, the loss is 57 dB, for distance depth of 50 m among other height losses, also this height adequate the wireless sensor network application inside greenhouse.

For the agricultural applications such as greenhouse in our simulation, the foliage depth range from 1 to 48 m between farthest wireless end node and the based station node with max foliage loss of 16.68 and 21.6 dB for MED and ITU models respectively. The simulation assumes that there are 5 levels of sensor node far away from base station node, these levels is separated by 10 m apart, hence level (i) = 10, 20, 30, 40, 50 m; for $i = 1, 2, 3, 4, 5$. Therefore foliage depth will be in the range of $1 \leq d_f \leq \text{level}(i) - 1$ and the transmitter antenna height is 3.5 m while 1 m

Fig. 1 Path loss versus separation distance for equal antennas height of 0.5, 1, 1.5 and 2 m



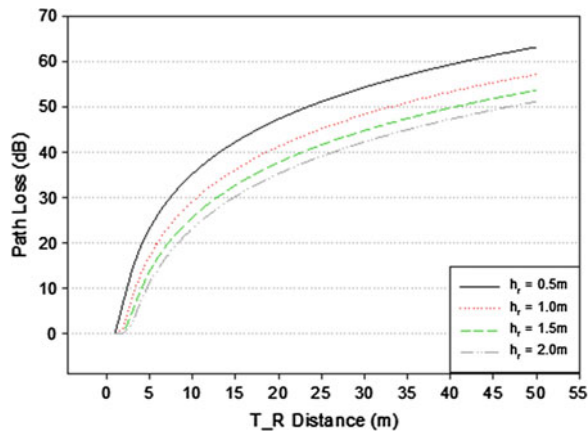


Fig. 2 Path loss versus antenna separation distance for h_t of 3.5 m and h_r of 0.5, 1, 1.5 and 2 m

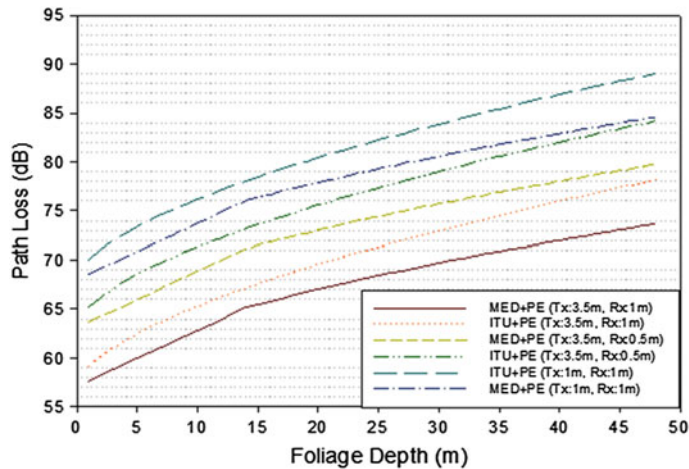


Fig. 3 Antenna height effects on total path loss for MED and ITU vegetation models and T-R model

of height for receiver antenna, these parameters will be applied to MED and ITU-R models.

As a conclusion of the effects of antenna height of the node transceiver on the total path loss are depicted in Fig. 3, the combination of height of 3.5 and 1 m shows less total path loss (73, 78.15 dB for MED and ITU models respectively for max foliage depth) as compared to height of 3.5 and 0.5 m where the total path loss is shown to be 79.78 and 84.18 dB for MED and ITU models respectively. This true since the first Fresnel zone must be clear out of obstacle and maximize the wave strength toward the receiver and the 0.5 m of receiver antenna height is less than the

Fresnel threshold of the first zone. Contrary for equal height of antennas of 1 m, the total path loss is shown to be higher, hence for MED model the loss is found to be 84 and 89 dB for ITU model. This also shows that the best antenna height based on greenhouse environment and total path loss is shown to be with the 3.5 m and 1 m height for transceiver of main and end node.

The main benefits from the simulation is that it easily shows the effects of network nodes height configuration on the reliability and performance of the network, hence it's a good indication of how far the total losses can affect the received signal strength in vegetation environment. This simulation can be used in prior of real time hardware deployment of the WSN nodes with proper calculation of maximum losses that can be damped our network signal propagation and hence degrade the total performance and reliability.

8.1 Communication Coverage Simulation

Usually, based on the maximum coverage range, the wireless sensor network nodes are deployed so that best connectivity is obtained. Hence for a larger communication range of a wireless node, the further it can be deployed. Although that the hardware wireless node has a maximum communication range of 100–120 m within free space line of sight situation, while in many cases of real experiment, communication failed either because proper studies were not been carried out to estimate the communication coverage of the node before deployment in the real environment or the maximum data sheet communication range is achievable only in ideal environment which is little bit different form real time application environment.

Figure 4 shows the communication coverage versus foliage depth for total path loss with various antenna heights. The simulation results shown in figure explain the effects of various vegetation depths between the main node and the end node, the sensing node. The investigation of the network communication coverage and hence the network performance is carried out based on the results shown. Using of only T-R model (PE) alone shows that the coverage distance is uniform at 187 m. Contrarily, the communication coverage decreases with increasing of vegetation depth for both MED and ITU models combined with the T-R model. For example at a height of 3.5, 1 m of antennas the MED+PE model expresses a decreasing of communication coverage from 180 to 134 m for vegetation depth of 1–10 m, respectively. The MED+PE for 3.5, 1 m of antenna height has more coverage range than 3.5, 0.5 m combination height of antennas, also MED+PE shows higher coverage than ITU+PE model and this due to the fact that the ITU model expresses higher loss than MED model (5) and (6).

At antenna height of 0.5 m for both models, the communication range is below 43 m for all vegetation depths. This is true since the height of the antenna has a significant effect on communication range. Overall, the communication range of a network shows a limitation of network nodes placement with respect to main node, the distribution of network nodes must be with consideration of the communication

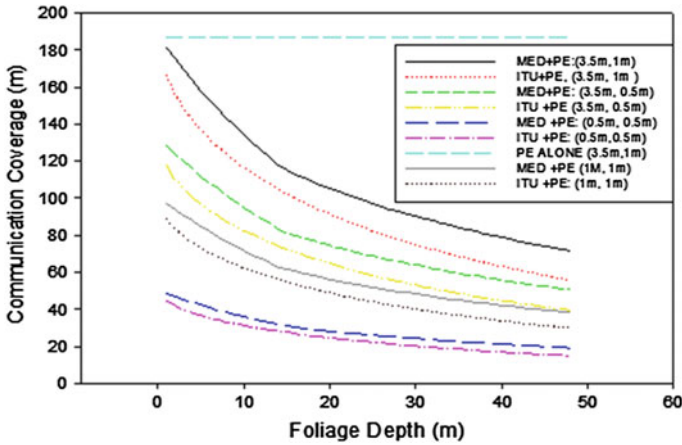


Fig. 4 Communication coverage for various foliage depth and vegetation models, MED and ITU

range of the nodes so that maximum coverage can be achieved and hence network nodes can communicate easily which gives better network performance than a situation where some nodes of the network is out of coverage.

The simulation used, Fig. 4, shows the communication range for the network based on 85 dB sensitivity of the receiver and various depths of vegetation loss, the simulated models shows worthy information on how far a node can be deployed far away from main node and how much vegetation depth can be in the way between them. Hence the real implementation of network nodes in the field has a prior awareness of the maximum dimension that must follow for nodes deployment while simulation results is little deviated from real time implementation due to other effects like grass, iron stand, roof, atmosphere that could contribute in more losses which is neglected during the simulation phase. The antenna heights play an important role in communication coverage and hence will extend the node deployment positions which give more flexibility for the network to cover a wide area with the same capability.

8.2 Network Connectivity Simulation

WSN based on a star topology requires that all wireless end nodes have the ability to direct communication with the main node, hence to achieve maximum connectivity of network nodes, a precise study of nodes positioning should be done prior to implementation. The connectivity simulation of network nodes will identify all the sensor nodes which are within the range of the main node. The connectivity percentage can be defined as:

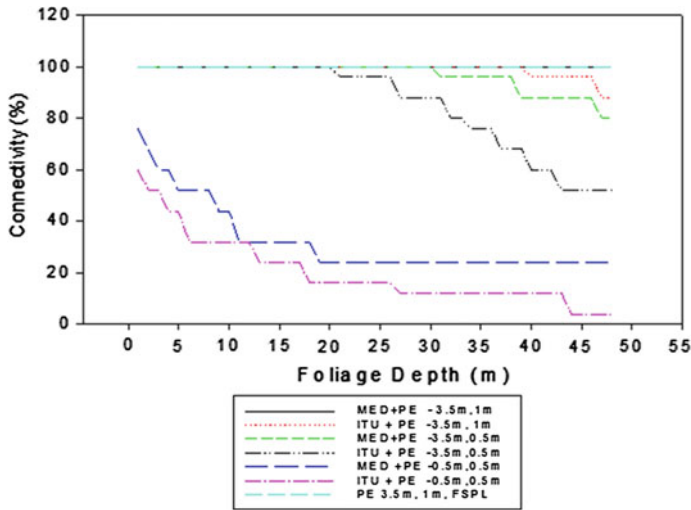


Fig. 5 Network connectivity for different antenna height with propagation models in vegetation

$$Connectivity = \frac{N_{con}}{N_{node}} \times 100\% \quad (10)$$

where N_{con} is the total number of connectivity and N_{node} is total number of nodes in the topology.

Figure 5 shows the simulation result of MED and ITU model with T-R (PE) model at three different antennas height combination and for various vegetation depths. The non vegetation loss of FSPL and PE model shows 100 % of connectivity, an idealistic model of FSPL and antenna height of the T-R model gives 100 % of connectivity.

As the vegetation depth increases, the two vegetation models show less connectivity. The experimental studies in Meng et al. [4] reported that the total path loss for near ground propagation is the accumulation of the foliage and the ground reflection effects. For antennas height of 0.5 m, the MED+PE model shows 76 % of connectivity degraded to 24 % as foliage depth increased from 1 to 20 m for the MED vegetation model while less connectivity of 60 % for the ITU model degraded to 16 % for same foliage depth as shown in Fig. 5.

The MED+PE model at 3.5, 1 m antenna height shows 100 % connectivity, while the ITU+E model shows, at foliage depth of 40 m, decreasing connectivity reached 88 % on the other hand a perfect connectivity is achieved when foliage depth is less than 40 m. Contrary when the antenna height combination for transmit and receive are 3.5 and 0.5 m height the MED+PE model shows connectivity of 100 % when foliage depth is less than 31 m while for more foliage depth greater than 30 m the model shows decrease in node connectivity reached to 80 %. The same behavior of the ITU+PE model is shown for 3.5 and 0.5 m antenna height combination, the

model shows a perfect connectivity before 20 m of foliage depth while decreasing in connectivity shown, reached to 52 %, for higher depth of the foliage.

The ITU+PE model results in Fig. 5 shows lower connectivity as compare to MED+PE model for all antenna height. It's clear shown that when WSN nodes deployed near the ground and its antenna height is low causes higher difficulties in connectivity and hence need either more node power, rearrangement of node positioning or higher nodes' antenna. Contrary when antenna height increases, the percentage of connectivity will be improved.

Based on (3), the analytical calculation proves that greater than 0.9 m of antenna height is required to clear ground reflection. Therefore, based on these results, the antenna height of each end node should be positioned higher above the ground so that optimum connectivity between the nodes is assured. And as shown that the combination of 3.5 and 1 m heights of antennas shows perfect connectivity when used with MED vegetation model for all vegetation depth and ITU model shows perfect connectivity for same height combination but with less foliage depth of 40 m than MED model.

9 Conclusions

The electromagnetic wave propagation in close proximity of the earth's surface can be predicted using propagation models that show the attenuation of traveling signals. Although there are propagation models which are normally used in the conception and analysis of wireless communications networks, these models cannot be directly used in applications that use wireless sensor networks, based on WPAN (Wireless Personal Area Network) technologies, such as IEEE802.15.4. In this chapter, two foliage model, MED and ITU-R, have been simulated, and the height of the antenna for different procedure also have been determined as an effect of loss, basically line of sight communication cannot be achieved in the real environment of agricultural application due to reflection, diffraction and scattering effects.

Inside greenhouses, the distances separate wireless nodes almost are within the range of reliable communication link even the foliage is highly dense. The maximum foliage dense simulated is less by one meter than maximum distance between transmitter and receiver nodes.

In this study, the effect of vegetation on WSN systems' performance is investigated and discussed. The results shown that FSPL and PE models show over optimistic output although that there is vegetation in the transmission path. There is different Radio wave propagation models with different results used to model the radio communication channel. The results show that even for short range, the path loss can be considerable as the radio wave is impairment by vegetation and reflection. Therefore it is crucial to pick the appropriate propagation model based on the environmental effects, since the evaluation of new protocols using inappropriate propagation model may yields of wrong results and conclusions. The communication coverage and network node connectivity are simulated with 100 nodes

spatially equal distributed in $100 \text{ m} \times 100 \text{ m}$ area, the nodes were 10 m apart. The combination of MED+PE propagation model with 3.5 and 1 m antenna height of transmission and receiving nodes respectively shows perfect result with the simulation setup, the communication coverage for MED+PE shows the minimum coverage range of 80 m when foliage depth is 48 m and perfect connectivity. Contrarily ITU+PE model is less than MED+PE model while it still in acceptable results with simulation setup for same heights of antenna.

Results obtained in this study shows that the antenna height of the transceiver for network nodes is one of the most important issues when deploying of WSN system. Hence antenna positioned at ground or near the ground causes short range for WSN and the system will fail when implemented in the real environment. So that to maintain high network node connectivity, a proper node placement strategy must follow that assure of proper height of the antenna with respect to application allowance. Therefore, adequate propagation model should be used to predict antenna heights and path loss to ensure that the system maintains link connectivity.

Finally, the simulations prove its usefulness to use and combine with WSN simulators, hence its assist in the investigation of path loss, antennas height affect, network coverage and network nodes connectivity based on well known foliage vegetation models. The simulators are developed with the use of MATLAB m-files which is based on C like structure programming which make it easy to be combined with simulators like OMNET++, NS2 and OPNET that is based on C structure programming.

References

1. Andrade-Sanchez P, Pierce FJ, Elliot TV (2007) Performance assessment of wireless sensor networks in agricultural settings. In: Proceedings of the 2007 ASABE annual international meeting, Minneapolis, MN, USA
2. Blaunstein N (2000) Radio propagation in cellular networks. Artech House, Norwood, p 172
3. CCIR (1986) Influences of terrain irregularities and vegetation on tropospheric propagation. Technical report, CCIR report
4. Meng YS, Lee YH, Ng BC (2010) Path loss modeling for near-ground VHF radio-wave propagation through forests with tree-canopy reflection effect. Prog Electromagnet Res M 12:131–141
5. I.R.R.P.833-2 (1999) Attenuation in vegetation. Technical report, InternationalTelecom Union, Geneva
6. Weissberger MA (1982) An initial critical summary of models for predicting the attenuation of radio waves by trees. Technical report
7. Savage N, Ndzi D, Seville A, Vilar E, Austin J (2003) Radio wave propagation through vegetation: factors influencing signal attenuation. Radio Sci 38(5):108
8. Moraitis N, Milas VF, Constantinou P (2007) On the empirical model comparison for the land mobile satellite channel. In: VTC spring, pp 1405–1409
9. Chang X (1999) Network simulations with OPNET. In: WSC '99: proceedings of the 31st conference on winter simulation, ACM Press, New York, NY, USA, pp 307–314
10. NS-2 (2012) The network simulator NS-2

11. Varga A (2001) The OMNeT++ discrete event simulation system. In: Proceedings of the European simulation multicongress, SCS—European Publishing House, Prague, Czech Republic, pp 319–324
12. Serodio C, Cunha JB, Morais R, Couto C, Monteiro J (2001) A networked platform for agricultural management systems. *Comput Electron Agric* 31(1):75–90
13. Mestre P, Serodio C, Morais R, Azevedo J, Melo-Pinto P (2010) Vegetation growth detection using wireless sensor networks. In: Proceedings of the world congress on engineering 2010, WCE 2010. Lecture notes in engineering and computer science, vol I, pp 802–807
14. Morais R, Fernandes MA, Matos SGC, Ferreira Serodio P, Reis M (2008) A ZigBee multi-powered wireless acquisition device for remote sensing applications in precision viticulture. *Comput Electron Agric* 62(2):94–106
15. Timmerman GJ, Kamp PGH (2003) Computerised environmental control in greenhouses. *PTC Ede*, pp 15–124
16. Scott T, Wu K, Hoffman D (2006) Radio propagation patterns in wireless sensor networks: new experimental results. In: IWCWC '06: proceeding of the 2006 international conference on communications and mobile 315 computing, ACM Press, New York, NY, USA, pp 857–862
17. Zhou G, He T, Krishnamurthy S, Stankovic JA (2004) Impact of radio irregularity on wireless sensor networks. In: *MobiSys: proceedings of the 2nd international conference on mobile systems, applications, and services*, ACM Press, New York, NY, USA, pp 125–138
18. Tate RF, Hebel MA, Watson DG (2008) WSN link budget analysis for precision agriculture. American society of agricultural and biological engineer, 2008 Providence, Rhode Island 084955. doi:[10.13031/2013.24935](https://doi.org/10.13031/2013.24935)
19. Hebel MA (2006) Meeting wide-area agricultural data acquisition and control challenges through wireless network technology. In: Proceeding of the 4th world congress conference on computers in agriculture and natural resources, Orlando, Florida, USA, 24–26 July 2006
20. Hebel MA, Tate RF, Watson DG (2007) Results of wireless sensor network transceiver testing for agricultural applications. In: Proceedings of the 2007 ASAE annual international meeting, Minneapolis, MN, USA
21. Kirubanand VB, Palaniammal S (2011) Study of performance analysis in wired and wireless network. *Am J Appl Sci* 8:826–832
22. Goldsmith A (2005) *Wireless communications*. Cambridge University Press, New York
23. Rappaport TS (2002) *Wireless communications: principles and practice*, 2nd edn. Prentice Hall PTR, Upper Saddle River
24. John F, Jeffrey Evans J (2010) Predicting ground effects of omnidirectional antennas in wireless sensor networks. *Wirel Sens Netw* 2:879–890
25. Thelen J, Goense D, Langendoen K (2005) Radio wave propagation in potato fields. In: Proceedings of 1st workshop on wireless network measurements (WiNMee), Riva del Garda, Italy
26. Parsons JD (2000) *The mobile radio propagation channel*, 2nd edn. Wiley, Chichester, pp 52–53
27. Seybold JS (2005) *Introduction to RF propagation*. Wiley-IEEE, Hoboken
28. Liu H, Meng Z, Shang Y (2009) Sensor nodes placement for farmland environmental monitoring applications. In: *WiCOM'09: proceedings of the 5th international conference on wireless communications, networking and mobile computing*, IEEE Press, Piscataway, NJ, USA, pp 3193–3196

Fiber Optic Sensors: Short Review and Applications

Naseer Sabri, S.A. Aljunid, M.S. Salim and S. Fouad

Abstract An extensive review of optical fiber sensors and the most beneficial applications is presented in this chapter. Although electrical sensing technologies have been successfully deployed in countless applications, the introduction of optical sensing, engineers and scientists can now perform measurements that were previously impractical or, in some cases, impossible with conventional electrical sensors. At present, many real-world applications already use both approaches to combine both benefits. The inherent advantages of fiber optic sensors such as lightweight, small size, passive, low attenuation, immunity to electromagnetic interference (EMI), wide bandwidth and environmental ruggedness were heavily used to offset their major disadvantages of high cost. Thus Fiber optic sensors (FOSs) have boosted the utility and demand for optical sensors in various military, industry and social fields. FOSs show reliable and rigid sensing tasks over conventional electrical and electronic sensors.

1 Introduction

Recently, FOSs have achieved increased attractiveness and market acceptance. FOSs provides a unique advantages in comparison to traditional sensors and thus FOSs be the successful for certain applications, mainly where conservative sensors are difficult or unfeasible to deploy or cannot offer the same richness of information. Fiber optic sensors offer a wide spectrum of advantages over traditional sensing systems, such as small size and longer lifetime. Immunity to electromagnetic

N. Sabri (✉) · S.A. Aljunid

Computer and Communication Engineering School, UniMAP, Perlis, Malaysia
e-mail: naseersabri@yahoo.com

M.S. Salim · S. Fouad

Laser & Optoelectronics Engineering School, AlNaharain University, Baghdad, Iraq

© Springer Science+Business Media Singapore 2015

F.L. Gaol et al. (eds.), *Recent Trends in Physics of Material Science and Technology*,
Springer Series in Materials Science 204, DOI 10.1007/978-981-287-128-2_19

299

interference, amenability to multiplexing, and high sensitivity make FOs the sensor technology of choice in several fields, including the healthcare and aerospace sectors. This chapter presents an executive review of optical fiber sensors and the most beneficial applications.

The field of fiber optics has undergone tremendous growth and advancement over the last 25 years. The fiber optic communication industry has literally revolutionized the telecommunication industry by providing higher performance, more reliable telecommunication links with ever decreasing bandwidth cost. This revolution is bringing about the benefits of high volume production to component users and a true information superhighway built of glass. Developments fiber optic sensor [1] technology has been a major user of technology associated with the optoelectronic and fiber optic communication industry. Since the past decades, various devices, including fiber optic gyroscopes; sensors of temperature, pressure, and vibration; and chemical probes have been under development of FOS technology. FOSs offer a number of advantages, such as increased sensitivity compared to existing techniques and geometric versatility, which permits into arbitrary shapes [2]. The ability of fiber optic sensors has been enhanced to substitute traditional sensors for acoustics, vibration, electric and magnetic field measurement, acceleration, rotation, temperature, pressure, linear and angular position, strain, humidity, viscosity, chemical measurements and a host of other sensor applications. They can be used in high voltage, high temperature, or corrosive environments due to its dielectric property; In addition, these sensors are compatible with communications systems and have the capacity to carry out remote sensing. Recently, investigation in the field has focused on the development of new materials with non-linear optical properties for important potential applications in photonics. Examples of these materials are the conjugated semiconducting polymers that combine optical properties with the electronic properties of semiconductors. In addition, these conducting polymers have photo luminescent and electroluminescent properties, making them attractive for applications in optoelectronics [3].

Fiber optic sensor technology in turn has often been driven by the development and subsequent mass production of components to support these industries. The inherent advantages of fiber optic sensors which include their ability to be lightweight, of very small size, passive, low attenuation, and low power, immunity to electromagnetic interference (EMI), high sensitivity, wide bandwidth and environmental ruggedness were heavily used to offset their major disadvantages of high cost and unfamiliarity to the end user [4, 5].

2 General Features of Fiber Optical Sensors

The researches and technologies in the field of FOS are aggravated by the anticipation that optical sensors properties have significant advantages compared to conventional sensor types. Recent advances in fiber optics (FOs) and the numerous

advantages of light over electronic systems have boosted the utility and demand for optical sensors in various military, industry and social fields. Environmental and atmospheric monitoring, earth and space sciences, industrial chemical processing and biotechnology, law enforcement, digital imaging, scanning, and printing are exemplars of them. The advances in research and development of FOS devices has extended their applications to various fields of technology, such as medical, chemical, and telecommunications industries. They have been developed to work on a wide variety of physical properties, like temperature, chemical changes, electric and magnetic fields, vibrations, strain, displacement (position), flow, pressure, rotation, radiation, liquid level, light intensity, and color. For performance in harsh environments, FOS proves reliable and rigid sensing devices over conventional electrical and electronic sensors where they have difficulties [1, 6]. Contrary, optical fiber sensors exhibit a number of advantages over other types of sensors; FOSs

- Enable small sensor sizes, do not contaminate their surroundings and are not subject to corrosion.
- Electrical Passiveness: can be used in volatile surroundings.
- Electromagnetic Immunity: perfect for microwave environment, immune to radio frequency interference (RFI) and electromagnetic interference (EMI).
- Resistant to high temperatures and chemically reactive environment: ideal for harsh and hostile environment.
- Compact and Light; perfect match for surface mounting and embedding applications. Require small cable sizes and weights.
- Wide Dynamic Range: ability to monitor a wide range of physical and chemical parameters and thus permit remote sensing.
- Greater Sensitivity and range resolution.
- Electrical Isolation: absolute electrical insulation from lofty electrostatic potential.
- Remote operation over several km lengths without any lead.
- Sensitivity: ideal for deployment in boreholes or measurements in hazardous environment.
- Multiplexing and distribution Capabilities of sensors are sole as they offer measurements at a greater number of points along a single optical cable: ideal for minimizing cable deployment and cable weight, or for monitoring extended structures like pipelines, dams etc.

The advantages of optical sensors are extending to the capacity of optical fiber to send and receive optical signals and over long distances. Nowadays there is a motivation to use of networks of sensors and thus will avoid the conversion process between electronics and photonics separately at each sensing site, hence reducing costs and increasing flexibility.

3 Fiber Optic Sensor Principles

The general structure of an optical fiber sensor system is shown in Figs. 1 and 2. It consists of an optical source (Laser, LED, Laser diode etc.), optical fiber (single or multimode), sensing or modulator element (which transducers the measure and to an optical signal), an optical detector and actuating circuitry (processing electronics, oscilloscope, optical spectrum analyzer etc.). In the FOSs devices, the optical parameters that can be modulated are the amplitude, phase, color (spectral signal), and state of polarization [6, 7]. The optical modulation methods of the sensors involve the following:

Intensity-modulated sensors The variation of the light intensity that is proportional to the perturbing environment can be detected by sensors. The concepts associated with intensity modulation include transmission, reflection, and micro bending. For this, a reflective or transmissive target can be incorporated in the fiber. Other mechanisms that can be used independently or in conjunction with the three primary concepts include absorption, scattering, fluorescence, and polarization.

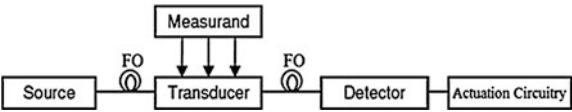


Fig. 1 General structure of an optical fiber sensor

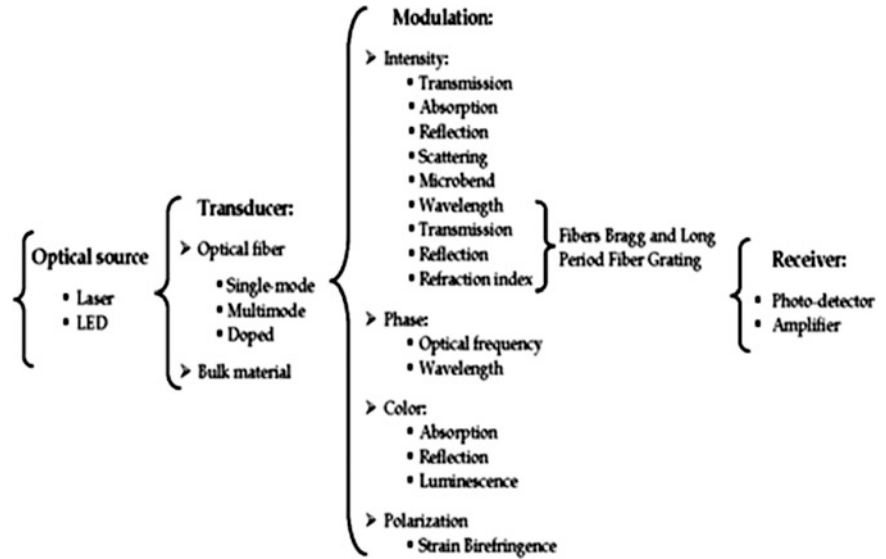


Fig. 2 Basic components of an optical fiber sensor

Intensity modulated sensors normally require more light to function than phase modulated sensors; as a result, they employ large core multimode fibers or bundles of fibers.

Phase-modulated sensors The phase of the light in a sensing fiber is compared to a reference fiber in a device known as an interferometer. These sensors employ a coherent laser light source and two single-mode fibers. The light is split and injected into the reference and sensing fibers. If the light in the sensing fiber is exposed to the perturbing environment, a phase shift occurs between them. The phase shift is detected by the interferometer. Phase modulated sensors are much more accurate than intensity-modulated sensors.

The change in colour is proportional to the changes in the absorption, transmission, reflection, or luminescence of the optical signal, whereas the polarization is related to the strain birefringence.

4 Fiber Optic Sensors Category

The sensing mechanisms of FOS can be sectorized into two main categories;

Direct sensors The modulation of the illuminating light is based on the medium under vision, whereby the composed light is the consequence of back scattering straight from the interrogated medium or medium fluorescence resulted by an optical source (e.g., photometric sensors).

Indirect sensors Employ an intermediary in response to the medium under test property of interest (e.g., temperature, enzyme presence). In addition, indirect sensors are often loosely grouped into two basic classes according to the manner in which the optical fiber is used, sensor can be either an intrinsic one, if the modulation takes place directly in the fiber, or extrinsic, if the modulation is performed by some external transducer [5, 6].

Intrinsic fiber optic sensor (IFOS) The intensity, phase, polarization, wavelength or transit time of light can be modulated. Sensors which modulate light intensity tend to use mainly multimode fibers, while only mono mode cables are used to

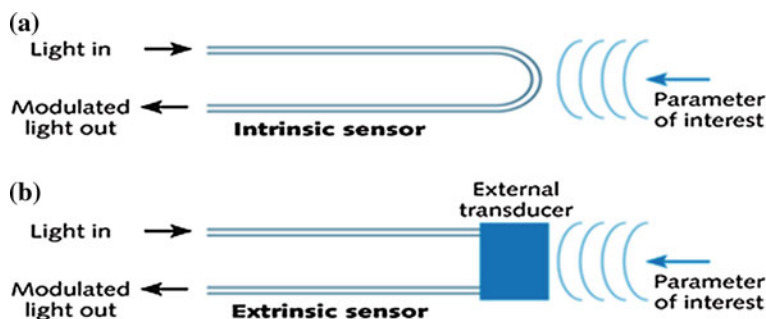


Fig. 3 Types of FOSs. Intrinsic devices (a) and extrinsic devices relay on a transducer (b)

modulate other parameters. An interesting feature of IFOs is that they can provide distributed sensing over distances of up to 1 m. In intrinsic sensors, the variable of interest (physical perturbation) must modify the characteristics of the optical fiber to modify the properties of the light carried by the fiber, Fig. 3a. IFOs directly employ an optical fiber as the sensitive material, sensor head, and also as the medium to transport the optical signal with information of the perturbation environment to be measured. They operate through the direct modulation of the light guided into the optical fiber. The light does not leave the fiber, except at the detection end, the output, of the sensor. These sensors can use interferometric configurations, Fiber Bragg Grating (FBG), Long Period Fiber Grating (LPFG), or special fibers (doped fibers) designed to be sensitive to specific perturbations. Light intensity is the simplest parameter to manipulate in intrinsic sensors because only a simple source and detector are required.

Extrinsic fiber optic sensor (EFOS) The optical fiber is simply used to guide the light to and from a location at which an optical sensor head is located. The sensor head is external to the optical fiber and is usually based on miniature optical components, which are designed to modulate the properties of light in response to changes in the environment with respect to physical perturbations of interest Fig. 3b. Thus, in this configuration, one fiber transmits optical energy to the sensor head. Then this light is appropriately modulated and is coupled back via a second fiber, which guides it to the optical detector.

Extrinsic fiber optic sensors use a fiber optic cable, normally a multimode one, to transmit modulated light from a conventional sensor. A major feature of extrinsic sensors, which makes them so useful in such a large number of applications, is their ability to reach places which are otherwise inaccessible. For example, the insertion of fiber optic cables into the jet engines of aircraft to measure temperature by transmitting radiation into a radiation pyrometer located remotely from the engine. Fiber optic cable can be used in the same way to measure the internal temperature of electrical transformers, where the extreme electromagnetic fields present make other measurement techniques impossible. Extrinsic fiber optic sensors provide excellent protection of measurement signals against noise corruption.

5 Classification of Fibre Optical Sensors

There are many factors that affects on the choice and design of an FOS. Generally it is dedicated by technological and application factors like spatial resolution and localization, sampling volume, accuracy and access to the target site. The collection of these parameters shapes the engineering problem that utters the different system parameters such as the elect of source and detector characterization and arrangement, with appropriate signal analysis based application, for instant, (e.g., background fluorescence subtraction in fiber-based Raman spectroscopy, model fits in fiber-based spatially dependent diffuse reflectance, corrections of fluorescence spectra for tissue attenuation distortions, etc.).

The FOSs may classified based on the spatial distribution of the measurand, such as

- *Point sensors* Where only a single point in space is measured, while it is possible to adopt multiple channels for multiple points exploit multiple points. Like these sensors are Fabry-Perot sensors and single Fibre Bragg Grating (FBG) sensors.
- *Integrated sensors* A physical parameter over a certain spatial section is measured and averaged and thus provides a single value. Such as the using of a deformation sensor to measure the strain over a long base length.
- *Multiplexed sensors* For a single fiber optic cable there is a number of fixed-discrete points where measurand is computed along the cable. Multiplexed FBG's is the ordinary example.
- *Distributed sensor* For a single fiber optic cable, the measurands of the parameter of interest of a certain spatial resolution are computed at any point along the cable. Examples comprise systems based on Rayleigh scattering.

6 Measurands of Optical Sensors

Optical sensing technology can measure almost all of the physical measurands of interest and a very large number of chemical quantities. Table 1 explores the most of optical sensor measurands.

Techniques by which the measurements are made can be generally grouped in three categories depending on.

6.1 Sensing Strategies

The strategies of sensing are based on either measuring the intensity change in one or more light beams or measuring the phase change in the light beams by causing them to interact or interfere with one another. As a result sensors in this class are called either *intensity* or *interferometric sensors*. Techniques used in the case of intensity sensors techniques are based on light scattering, spectral transmission

Table 1 Optical sensor measurands

Temperature	Chemical species
Pressure	Force
Flow	Radiation
Liquid level	pH
Displacement	Humidity
Vibration	Strain
Rotation	Velocity
Magnetic fields	Electric fields
Acceleration	Acoustic fields

changes (i.e., absorption of transmitted light where a simple attenuation is resulted, reflectance changes micro bending or radiative losses, and changes in the modal properties of the fiber. Interferometric sensors have been demonstrated based upon the magneto-optic, the laser-Doppler, or the Sagnac effects, to name a few.

6.2 Degree of Sensing

This class is composed of single point sensor or distributed sensor based on whether sensors operate only at a single point or over a distribution of points. In the case of a point sensor, the transducer may be at the end of a fiber the sole purpose of which is to bring a light beam to and from the transducer. Examples of this sensor type are interferometers bonded to the ends of fibers to measure temperature and pressure. In the case of a distributed sensor, sensing is performed all along the fiber length. Examples of this sensor type are fiber Bragg gratings distributed along a fiber length to measure strain or temperature.

6.3 Role of Optical Fiber

Further distinction is often made in the case of fiber sensors as to whether measurands act externally or internally to the fiber. Where the transducers are external to the fiber and the fiber merely registers and transmits the sensed quantity, the sensors are termed *extrinsic sensors*. Where the sensors are embedded in or are part of the fiber and for this type there is often some modification to the fiber itself, the sensors are termed *internal* or *intrinsic sensors*. Examples of extrinsic sensors are moving gratings to sense strain, fiber-to-fiber couplers to sense displacement, and absorption cells to sense chemistry effects. Examples of intrinsic sensors are those that use micro bending losses in the fiber to sense strain, modified fiber claddings to make spectroscopic measurements, and counter-propagating beams within a fiber coil to measure rotation.

Today, most of the measurands in Table 1 can be sensed by either intensity or interferometric techniques and as either point or distributed effects.

7 Fiber Optical Sensor Types

Fiber optical sensors are sensors that can work in harsh environments where conventional electronic and electrical sensors have difficulties. They can sense various physical properties, such as pressure, position, strain, chemical changes, magnetic and electric fields, flow, vibration, light level, radiation and colour [3, 8].

- Chemical sensors; remote spectroscopy, groundwater and soil contamination.
- Temperature sensors; largest commercially available sensors, range -40 to $1,000$ °C.
- Strain sensors; fiber Bragg gratings (FBG) technology, senses as little as 9 micro strain.
- Biomedical sensors; spectroscopic biomedical sensors, CO_2 , O_2 and pH can be measured simultaneously, flow monitoring by laser dopplerimetry.
- Electrical and magnetic sensors; appealing inherent dielectric nature, less sensitive to electromagnetic interference, small size and safer, they are almost always hybrid.
- Rotation sensor; based on the sagnac effect, two types ring laser gyroscope (RLG) and fiber optic gyroscope (FOG).
- Pressure sensors; based on piezoresistive technique or movable diaphragm, high performance (polarization based sensors); operating pressure ranges from 0 to 70,000 torr.

Displacement and position sensors; one of the first optoelectronic sensors to be developed, simple sensors rely on the change in retro reflectance due to a proximal mirror surface, also referred as liquid level sensors.

8 Fiber Optical Sensing Applications

The distributed and single fiber architecture offer large and lightweight structures benefits of simplify installation and reduce weight. The unique features of noncorrosive and non conductive specificities of the fiber provide highly benefits of industrial and outdoor applications where hazardous gases and voltages might be present. Also, the immunity to EMI removes the need for expensive and often difficult signal conditioning required for measurements near noisy sources such as power transformers. For harsh environmental condition and sensing over long distances, Optical sensing is consider as the perfect solution for applications where conventional electrical sensors such as foil strain gages, thermocouples, and vibrating wires have proven ineffective or difficult to use due to those conditions. In addition, the features and benefits of fiber Bragg grating (FBG) optical sensing include the nonconductive, electrically passive, immune to electromagnetic interference (EMI), induced noise, sensor measurements over very long distances (10 + km), daisy chain multiple sensors on a single fiber. Large and lightweight structures can benefit from the distributed single fiber architecture to simplify installation and reduce weight [6]. The nonconductive and noncorrosive nature of the fiber benefits outdoor and industrial applications where hazardous gases and voltages might be present. Also, the immunity to EMI removes the need for expensive and often difficult signal conditioning required for measurements near noisy sources such as power transformers. FBG optical sensing can benefit many applications in areas such as energy, civil infrastructure, and transportation monitoring.

Chemical field Optical-chemical and biochemical sensing is being researched extensively all over the world, and these sensors are finding an increasing number of applications in industry, environmental monitoring, medicine, biomedicine, and chemical analysis [2]. The main physical phenomena exploited for optical chemical sensing are absorption and fluorescence, although chemical luminescence, Raman scattering, and Plasmon resonance have also been used [9].

Medical field The use of optical fiber technology offers plentiful features that are well suitable for medical applications. Like these features non-toxic and biochemically inert, immune from electromagnetic interference, non-invasively in contact with outer organs such as the skin or surgically. As well, based on their flexibility and thin outer diameter, they can also be positioned into physical cavities (endoscopic approach), inserted interstitially via plainly invasive trocars, (e.g., hollow bore needles), or located intravascularly. Thus, measurands can be carried out in difficult to access parts of the human body with greater concentration sensitivity. Lastly, it is technologically possible to bundle multiple sensors with different measurement capabilities can be fused into a single probe as a packaged instrument, and hence potentially increasing useful information content [10–14]. Optical fiber technology offers a convenient, affordable, safe and effective approach for the delivery and collection of light to and from the tissue region of interest, and has been employed clinically since the 1960s [5].

Energy field Optical sensing offers the ability to address diverse challenges introduced by monitoring power plants and structures that working on generating, distributing and converting power. Whether it is a windmill requiring a lightweight solution or a hydroelectric turbine needing an EMI-resistant system, optical sensing has sole features that make it appropriate for these conventionally hard applications. FBG optical sensing has unique attributes that make it suitable for traditionally difficult applications. Monitoring structures that generate, produce, distribute, and convert electrical power introduces many challenges that can be addressed with FBG optical sensing. Whether it is a windmill requiring a lightweight solution or a hydroelectric turbine needing an EMI-resistant system, optical sensing has unique feature that match perfectly with these traditionally difficult applications. Partial discharge detection uses optical fiber sensors are being tested for use in detecting partial discharges in electrical transformers. Pinpointing such discharges is essential to preventing insulation breakdown and catastrophic failures. For example, monitoring the structural integrity of a wind turbine blade with electrical sensors would often result in noisy measurements because of long copper lead wires. With optical sensing, accurate and noise-free strain measurements on wind turbine blades are possible with little added weight to the structure. Furthermore, the non conductive and distributed nature of optical fibers lends well for many uses in oil and gas applications, including pipeline monitoring and downhole monitoring. Applications of energy FOS devices are wind turbine blade monitoring, pipeline monitoring, power line monitoring, offshore platform monitoring and downhole monitoring.

Agricultural and food Optical sensing can help growers and horticulturists assess crop ripeness, determine chlorophyll and light levels, and evaluate reflectance of petals and leaves to gauge plant health. To assurance the superiority of food

products for the end-user, quality control plays an important role. Optical sensing can be used to classify incoming feed, for online quality control and even for analysing oxygen levels inside closed packaging [15].

Civil field Structural health monitoring systems based on electrical sensors often face significant environmental challenges. An electrical monitoring system would require the installation of countless wires, a lightning grounding system, periodic external calibration, and the potential maintenance of corroded and/or degrading sensors. With an optical sensing solution, these downfalls are all eliminated. The ability to daisy chain multiple sensors on a single fiber greatly reduces the weight and complexity of the system. Furthermore, optical fiber does not corrode or conduct like copper wire, which increases longevity and reduces the risk of damage due to lightning. These attributes, coupled with the fact that optical sensors and NI interrogators do not require calibration, drastically reduce the amount of maintenance required. Applications of civil FOS devices are large building monitoring, bridge and road monitoring, airport landing strip load monitoring and dam monitoring [16].

Transportation field Monitoring systems play a crucial role to guarantee the perfect operation of transportation systems such as ships, railways, airplane, automobile and more. Contrarily, size, harsh environment and weight requirements can expose major challenges to developing an electrical monitoring system optical sensors ease these challenges by providing lightweight distributed sensor measurements that have superior features as aforementioned in addition to the longevity and ease of installation of FBG optical sensors and lack of need for external calibration, these sensing systems can be deployed reliably for decades without needing any maintenance—this is especially beneficial for long-term railway and ship hull monitoring. The ability to have multiple sensors on a single, very thin fiber dramatically reduces the weight of the monitoring system, which is especially important in aerospace applications. Applications of transportation FOS devices are railway monitoring, fuel tank monitoring, airplane wing monitoring, and Ship hull monitoring [16, 17].

9 Conclusions

In conclusion, electrical sensing technologies have been successfully deployed in countless applications. However, with the introduction of optical sensing, engineers and scientists can now perform measurements that were previously impractical or, in some cases, impossible with conventional electrical sensors. At present, many real-world applications already use both approaches to harness the combined benefits of electrical and optical sensing.

Recently, the advances of technology and applications based on fiber optics are progressed very rapidly. As a physical medium, Optical fiber is subjected to various perturbation of at all times. Thus, it experiences size and shape changes and optical (refractive index, mode conversion) changes to a larger or lesser extent depending

upon the nature and the magnitude of the perturbation. It is crucial for communication application to be reliable all times and to minimize various effects on transmitted signals. On the other hand in fiber optic sensing, the response to external influence is deliberately enhanced so that the resulting change in optical radiation can be used as a measure of the external perturbation. Since light is characterized by amplitude (intensity), phase, frequency and polarization, any one or more of these parameters may undergo a change. The fiber acts as a modulator in sensing phase, while in communication, the signal passing through a fiber is already modulated. It also works as a transducer that converts measurands into a corresponding change in the optical radiation.

FOSs offer many advantages such as freedom from EMI, wide bandwidth, compactness, geometric versatility and economy. Indeed of dielectric construction, passive, and high sensitivity when compared to other types of sensors. Specially prepared fibers can withstand high temperature and other harsh environments. In telemetry and remote sensing applications it is possible to use a segment of the fiber as a sensor gauge while a long length of the same or another fiber can convey the sensed information to a remote station. Deployment of distributed and array sensors covering extensive structures and geographical locations is also feasible.

References

1. Kashyap R (1994) Photosensitive optical fibers: devices and applications. *Opt Fiber Tech* 1:17–34
2. Narayanaswamy R, Wolfbeis OS (2004) Optical sensors: industrial, environmental and diagnostic applications. Springer, Berlin
3. Ficocelli P (2008) Fiber optic thermal monitoring systems for SAGD, Petroleum Technology Alliance of Canada, PTAC, technology information session
4. Ghatak A, Thyagarajan K (2000) An introduction to fiber optics. Cambridge University Press, New York. ISBN 0521571200
5. Keiser G (1991) Optical fiber communications, 2nd edn. McGraw-Hill, New York. ISBN 0070336172
6. Yu FTS, Yin S, Ruffin PB (2008) Fiber optic sensors. CRC Press, Taylor & Francis Group, New York. ISBN 9781420053654
7. Eurekalert Distinctive advantages give optical sensors the edge over conventional systems. www.eurekalert.org/pubreleases/2003-08/ti-dag081303.php
8. Sezerman O (2008) Fiber optics distributed strain and temperature sensor (DSTS). www.ozoptics.com. Accessed Sept 2008
9. Korotcenkov G, Cho BK, Narayanaswamy R, Sevilla F (2011) Optical and fiber optic chemical sensors. In: Korotcenkov G (ed) Chemical sensors comprehensive sensor technologies: volume 5 electrochemical and optical sensors, pp. 311–476
10. Welch AJ, van Gemert MJC (2011) Overview of optical and thermal laser-tissue interaction and nomenclature. In: Welch AJ, van Gemert MJC (eds) Optical-thermal response of laser-irradiated tissue, 2nd edn. Berlin, Springer. doi:10.1007/978-90-481-8831-4
11. Udd E (1989) “Overview of fiber optic applications to smart structures,” review of progress in quantitative nondestructive evaluation. Plenum Press, New York
12. Baldini F, Falai A, De Gaudio AR, Landi D, Lueger A, Mencaglia A, Scherr D, Trettnak W (2003) Continuous monitoring of gastric carbon dioxide with optical fibres. *Sens Actuators* 90:132–138

13. Baldini F, Giannetti A (2005) Optical chemical and biochemical sensors: new trends. Proc SPIE 5826:485–499
14. Sensorland (2013) The fibre optical sensors, tissue, 2nd edn. www.sensorland.comaccess. Accessed May 2013
15. oceanoptics.eu (2013) Agriculture and food optimising crop growth and ensuring quality. www.oceanoptics.eu. Accessed May 2013
16. NI (2012) Fiber-optic sensing. <http://www.ni.com/opticalsensing/>
17. Krohn D (2003) Overview of fiber optic sensors. In: KMI's 26th annual Newport conference, October, 2003

Dynamic Effects of Piezoelectric Patch Actuators on Vibrational Response of Non-deterministic Structures: Modelling and Simulations

Azni N. Wahid, Asan G.A. Muthalif and Khairul A.M. Nor

Abstract Many engineering systems such as aircraft and automotive are considered built-up structures, fabricated from components that are classified as deterministic subsystems (DS) and non-deterministic subsystems (Non-DS). The response of Non-DS is sensitive to minor details of material properties, geometry, connections and damping distribution; therefore create problems in vibration control. Hence, the response of Non-DS is estimated using statistical modelling technique such as statistical energy analysis (SEA), in which any external input to the subsystem must be represented in terms of power input. In this research, ensemble average of power delivered by a piezoelectric (PZT) patch actuator to a simply-supported plate when subjected to structural uncertainties is studied using Lagrangian method and obtained by Monte-Carlo simulation. The effects of size and location of the PZT patch actuators on the power delivered to the plate are investigated. It is found that changing the patch location on the structure will not affect the average power supplied by the patch while changing the patch size will change the power magnitude proportionally but with some variations at higher frequency.

1 Introduction

1.1 Deterministic Structures Versus Non-deterministic Structures

Most of engineering systems are subjected to forms of excitation with frequencies that range from low to high which later determine the type of response of the system. If vibrational response characteristics such as displacement, acceleration,

A.N. Wahid · A.G.A. Muthalif (✉) · K.A.M. Nor
Smartstructures, Systems and Control Research Laboratory (S3CRL),
Department of Mechatronics Engineering, International Islamic University Malaysia,
Jalan Gombak, 53100 Kuala Lumpur, Malaysia
e-mail: asan@iium.edu.my

and stress are known precisely as functions of time, the vibration is known as *deterministic vibration* (DS) i.e. can be mathematically described using deterministic method. In addition, systems such as aircraft are usually subjected to excitation at high frequency. The modelling of such systems presents the problem of *high-frequency vibration* [1]. The physical characteristic of a high-frequency vibration is that the precise natural frequencies and shapes of high-order modes are impossible to be calculated hence causing difficulties in vibration analysis and therefore any control effort. The system response subjected to this type of vibration is also very sensitive to uncertainties due to the short wavelength deformation, in which identical structures can produce very different dynamic response properties [2]. Subsystem subjected to this type of vibration falls into *non-deterministic* (Non-DS) subsystem category and therefore statistical modelling method such as Statistical Energy Analysis (SEA) are commonly employed [3, 4].

1.2 Statistical Energy Analysis

Unlike classical vibration analysis which is based on force and displacement, this approach uses energy quantities such as energy, energy density and power. The main aim of the SEA method is to predict energy distribution among subsystems. This is achieved via energy balance equations which involve an expression for power flow [5]. For a single system j , any excitation applied to the subsystem can be considered as power input P_j to the system. Vibrational energy stored in the subsystem due to any input power is E_j . In reality there is also power dissipation, P_{out} , mainly due to internal damping of the structure. This power dissipation (rate of energy dissipation) is given by, $P_{\text{out}} = \omega \eta_j E_j$, where η_j is the internal loss factor of the subsystem. In steady-state, the power input is equal to power loss $P_j = \omega \eta_j E_j$. However, if the subsystem is coupled to another subsystem, k , some power balance holds between both subsystems in which they share vibration energy. The fundamental assumption in SEA is that the power flow between the two subsystems is proportional to the difference in modal energy between them:

$$P_j = \eta_j \omega E_j + \eta_{jk} \omega \left(\frac{E_j}{n_j} - \frac{E_k}{n_k} \right) \quad (1)$$

where ω is the central frequency of a given bandwidth, n_j is the modal density of the subsystem j , E_j is the spatial average energy of subsystem j , $\frac{E_j}{n_j}$ and $\frac{E_k}{n_k}$ are the ensemble average modal energies for subsystem j and k respectively, η_j is the internal loss factor of the subsystem j , and η_{jk} is the coupling loss factor between the subsystem [3, 5].

In SEA, subsystem is not actually of a structural nature as such (in the substructure sense): an SEA subsystem is a group of similar and significant resonant modes (or wavefields) that dominate the response of the structure in a given structural region [3]. A flat plate in general can be represented by three different subsystems: one for the flexural motion and two for the extensional motion (the longitudinal and shear

motion). Hence, a flat plate may be represented by three SEA subsystems with three corresponding energy response variables. In this article, only the flexural motion of the plate is discussed and represented as a SEA subsystem. Hence, only the power injected by the PZT patch actuator to flexural motion of the plate, i.e. bending moment, is considered in the analysis. In calculating the power input from the piezoelectric actuator on the plate, only the induced stress and strain in the piezoelectric materials are considered in the calculation.

Modal density and modal overlap factor are statistically based parameters which are essential to SEA modelling method. Modal density is defined as number of natural frequencies per unit frequency. The concept of modal density is inadequate at low frequency range because its natural frequencies vary with specific condition of the system. In contrast, at higher frequency range (as mode order rises), the natural frequencies are subjected to increasing uncertainties [6] due to its degree of sensitivity to minor details of material properties, geometry, connections and damping distribution of the system [1]. Modal density for a thin flat plate subjected to bending waves is [1]:

$$n(\omega) = \frac{S}{4\pi} \left(\frac{\rho h}{D} \right)^{1/4} \quad (2)$$

where S is structure area (one side), ρ is material density of the plate, h is plate thickness and D is the flexural rigidity of the plate:

$$D = \frac{E}{(1 - \nu^2)} \left(\frac{h^3}{12} \right) \quad (3)$$

ν is Poisson's ratio of the plate.

Modal overlap factor, M is the ratio of the half-power bandwidth to the local average interval between natural frequencies, which is basically the degree of overlap in modal response. M is a dimensionless quantity, proportional to frequency, ω and is given by:

$$M(\omega) = \omega \eta n \quad (4)$$

where η is modal loss factor and n is modal density. At low frequency range, where M is less than unity, individual modal responses are distinctly visible. As M approached unity, the individual modal responses begin to overlap and as M increases beyond unity namely at high frequency range, the modal responses combine to be broader peaks and no distinct resonant peaks are visible in the response. This is illustrated in Fig. 1.

PZT patch actuator has received significant attention for vibration analysis and control due to its unique direct-and-inverse PZT effect. Adding a PZT patch actuator on a structure is equivalent to adding an external moment to the dynamics of the structure [7, 9, 11]. Conveniently, the influence of adding moment on a DS is known; mathematical model describing power supplied by PZT patch actuator is available.

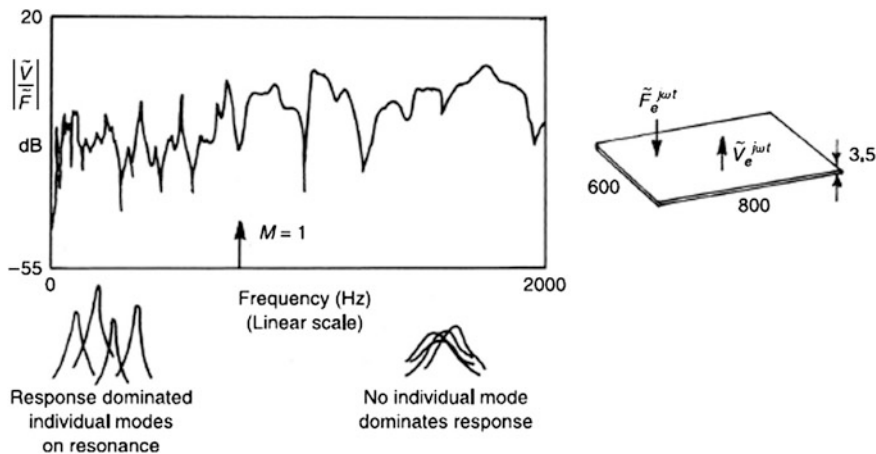


Fig. 1 Frequency response function of a rectangular plate [1]

However, analytical solution to estimate average power delivered by a PZT actuator when attached to a Non-DS is limited and needs to be understood [8]. Such analytical solution is also required when attempting to control vibration of Non-DS as built-up structure using PZT patch actuator.

In this chapter, ensemble average of power given by the PZT patch actuator to a plate when subjected to structural uncertainties is simulated. The effects of size and location of PZT actuator on the power delivered to the plate are investigated and presented. Only these two parameters are on interest considering the practicality in real application. A simply-supported plate with attached PZT patch and distributed point masses is taken as a benchmark model. The distributed point masses are used to introduce uncertainties to the response of the plate i.e. produce the non-deterministic characteristic. Equation of motion for the benchmark model is derived using Lagrangian method and the ensemble average of power delivered to the plate is obtained by Monte-Carlo simulation.

2 Derivation of Equation of Motion

2.1 A Rectangular Plate with Distributed Point Masses Attached with a PZT Patch Actuator

The constitutive equations of a general PZT material are [7]:

$$\sigma_{kl} = c_{ijkl}^E \epsilon_{ij} - e_{kij} E_k \quad (5)$$

$$D_i = e_{ikl} \sigma_{kl} + \zeta_{ik}^T E_k \quad (6)$$

where σ_{kl} and ε_{ij} are components of stress and strain tensors, respectively, c_{ijkl}^E is the elastic stiffness under constant electric field (Hooke's tensor), e_{kij} is the PZT constant, E_k is the electric field, D_i is the electric displacement and ζ_{ik}^T is the dielectric constant under constant strain. For an isotropic, two dimensional element, matrices $\mathbf{c}_{ijkl}^E$ and $\mathbf{e}_{kij}$ can be written as:

$$\mathbf{c}_{ijkl}^E = \begin{bmatrix} \frac{E}{1-\nu^2} & \frac{\nu E}{1-\nu^2} & 0 \\ \frac{\nu E}{1-\nu^2} & \frac{E}{1-\nu^2} & 0 \\ 0 & 0 & \frac{E}{2(1+\nu)} \end{bmatrix}; \quad \mathbf{e}_{kij} = \begin{bmatrix} \frac{E}{1-\nu^2} & \frac{\nu E}{1-\nu^2} & 0 \\ \frac{\nu E}{1-\nu^2} & \frac{E}{1-\nu^2} & 0 \\ 0 & 0 & \frac{E}{2(1+\nu)} \end{bmatrix} \begin{bmatrix} d_{31} \\ d_{32} \\ 0 \end{bmatrix} \quad (7)$$

Therefore, (5) can be rewritten as (subscript 'pzt' refers to piezoelectric material):

$$\begin{bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{xy} \end{bmatrix} = \begin{bmatrix} \frac{E_{pzt}}{1-\nu_{pzt}^2} & \frac{\nu_{pzt} E_{pzt}}{1-\nu_{pzt}^2} & 0 \\ \frac{\nu_{pzt} E_{pzt}}{1-\nu_{pzt}^2} & \frac{E_{pzt}}{1-\nu_{pzt}^2} & 0 \\ 0 & 0 & \frac{E_{pzt}}{2(1+\nu_{pzt})} \end{bmatrix} \begin{bmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{xy} \end{bmatrix} - \begin{bmatrix} \frac{E_{pzt}}{1-\nu_{pzt}^2} & \frac{\nu_{pzt} E_{pzt}}{1-\nu_{pzt}^2} & 0 \\ \frac{\nu_{pzt} E_{pzt}}{1-\nu_{pzt}^2} & \frac{E_{pzt}}{1-\nu_{pzt}^2} & 0 \\ 0 & 0 & \frac{E_{pzt}}{2(1+\nu_{pzt})} \end{bmatrix} \begin{bmatrix} d_{31} \\ d_{32} \\ 0 \end{bmatrix} \frac{V(t)}{t_p} 8$$

where d_{32} and d_{31} are the PZT strain constants. Assuming an isotropic PZT material is used, $d_{32} = d_{31}$. E_{pzt} is the Young's modulus, ν_{pzt} is Poisson's ratio, $V(t)$ is voltage supplied to the material, t_p is the patch thickness; all properties are attributed to the PZT material. Equation (8) will be used to derive the equation of motion of the system consisting of a thin rectangular plate with a PZT patch actuator attached on it.

Consider a rectangular plate with length a , width b and thickness h . The plate has a PZT patch located at distance x_1 along the length and distance y_1 along the width. The patch has $x_2 - x_1$ length, $y_2 - y_1$ width and thickness t_p . A number of point masses are distributed randomly on the plate to create uncertainties (Fig. 2).

The linear strain displacement relationship is represented by the following equation [7]:

$$\epsilon_x = -z \frac{\partial^2 w}{\partial x^2}; \quad \epsilon_y = -z \frac{\partial^2 w}{\partial y^2}; \quad \gamma_{xy} = -2z \frac{\partial^2 w}{\partial x \partial y} \quad (9)$$

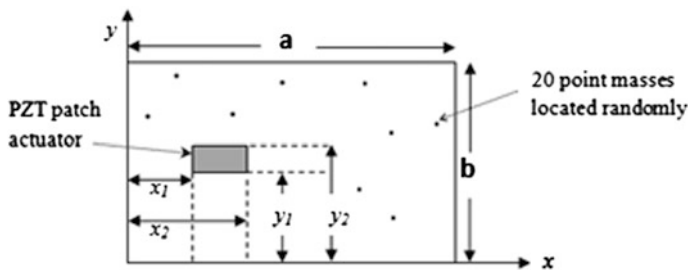


Fig. 2 Schematic diagram of a rectangular plate attached with PZT patch and 20 randomly-located point masses

From (9), ϵ_x , ϵ_y and γ_{xy} are the normal strains and the shear strain of the thin plate respectively. The stresses corresponding to normal strains and the shear strain are represented as:

$$\sigma_x = \frac{E}{1 - \nu^2} (\epsilon_x + \nu \epsilon_y) \quad (10)$$

$$\sigma_y = \frac{E}{1 - \nu^2} (\epsilon_y + \nu \epsilon_x) \quad (11)$$

$$\tau_{xy} = G \gamma_{xy} \quad (12)$$

However, with the existence of PZT patch on the plate, the cross section of the plate will not be uniform therefore stress distribution along the thickness of the plate needs to be modified [9]. The neutral axis of the system is shifted towards the PZT patch and this shift can be calculated by balancing the force and moment across the patch thickness:

$$\int_{-\frac{h}{2}}^{\frac{h}{2}} E(z - z_n) dz + \int_{\frac{h}{2}}^{\frac{h}{2} + t_p} E_{PZT}(z - z_n) dz = 0 \quad (13)$$

which resulted to:

$$z_n = \frac{E_{PZT} t_p (h + t_p)}{2(Eh + E_{PZT} t_p)} \quad (14)$$

where z_n is the shifted neutral axis measured from the plate centre (Fig. 3).

Hence, the linear strain displacement relationships are modified as [10]:

$$\epsilon_x = -(z - z_n) \frac{\partial^2 w}{\partial x^2}; \quad \epsilon_y = -(z - z_n) \frac{\partial^2 w}{\partial y^2}; \quad \gamma_{xy} = -2(z - z_n) \frac{\partial^2 w}{\partial x \partial y} \quad (15)$$

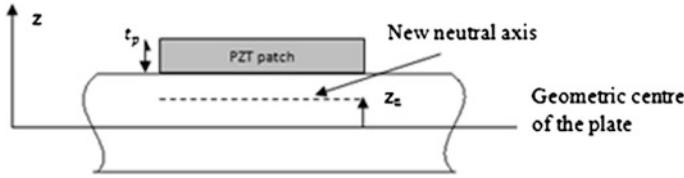


Fig. 3 Cross section along xy-axes to show the shift of neutral axis due to PZT patch attachment

To obtain the equation of motion of the system, Lagrangian method will be employed:

$$\frac{\partial}{\partial t} \left(\frac{\partial L}{\partial \dot{w}_{mn}} \right) - \frac{\partial L}{\partial w_{mn}} = F_{mn} \quad \text{where } L = T - U \quad (16)$$

where T and U are the total kinetic energy and total potential energy of the system, respectively. The total kinetic energy of the system has contributions from the plate, PZT actuator and distributed point masses:

$$\begin{aligned} T_{total} &= T_{plate} + T_{PZT} + T_{dist. point mass} \\ &= \frac{1}{2} \rho h \iint \left(\frac{\partial w}{\partial t} \right)^2 dA + \frac{1}{2} \rho_{pe} t_p \iint \left(\frac{\partial w}{\partial t} \right)^2 S(x, y) dA_p \\ &\quad + \frac{1}{2} \iint \sum_r m_r \delta(x - x_r, y - y_r) \left(\frac{\partial w}{\partial t} \right)^2 dA \end{aligned} \quad (17)$$

where ρ_{pe} is the density of the PZT actuator, A_p is the area of the PZT actuator and $S(x, y) = [H(x - x_1) - H(x - x_2)][H(y - y_1) - H(y - y_2)]$ is a product of Heaviside function which takes care of the presence of PZT patch on the plate. The total potential energy of bending for the system has contributions from the plate and PZT actuator:

$$\begin{aligned} U_{total} &= U_{strain plate} + U_{strain PZT} \\ U_{total} &= \frac{D}{2} \iint \left\{ \left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} \right)^2 - 2(1 - \nu) \left[\left(\frac{\partial^2 w}{\partial x^2} \frac{\partial^2 w}{\partial y^2} \right) - \left(\frac{\partial^2 w}{\partial x \partial y} \right)^2 \right] \right\} dA \\ &\quad + \frac{D_1(x, y)}{2} \iint \left\{ \left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} \right)^2 - 2(1 - \nu) \left[\left(\frac{\partial^2 w}{\partial x^2} \frac{\partial^2 w}{\partial y^2} \right) - \left(\frac{\partial^2 w}{\partial x \partial y} \right)^2 \right] \right\} dA \\ &\quad + \frac{D_2(x, y)}{2} \iint \left\{ \left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} \right)^2 - 2(1 - \nu_{pe}) \left[\left(\frac{\partial^2 w}{\partial x^2} \frac{\partial^2 w}{\partial y^2} \right) - \left(\frac{\partial^2 w}{\partial x \partial y} \right)^2 \right] \right\} dA \\ &\quad + B(t) \iint S(x, y) \left[(d_{31} + \nu_{pe} d_{32}) \frac{\partial^2 w}{\partial x^2} + (d_{32} + \nu_{pe} d_{31}) \frac{\partial^2 w}{\partial y^2} \right] dA \end{aligned} \quad (18)$$

where,

$$D_1(x, y) = d_1 S(x, y) = \frac{E}{(1 - \nu^2)} \left(\frac{h^3}{12} + z_n^2 h \right) S(x, y) \quad (19)$$

$$D_2(x, y) = d_2 S(x, y) = \frac{E_{pe}}{(1 - \nu_{pe}^2)} \left[\frac{t_p^3}{3} + \frac{h^2 t_p}{4} + \frac{h t_p^2}{2} - z_n (h t_p + t_p^2) + z_n^2 t_p \right] S(x, y) \quad (20)$$

$$B(t) = \frac{E_{pe} V(t)}{2(1 - \nu_{pe}^2)} (h + t_p - 2z_n) \quad (21)$$

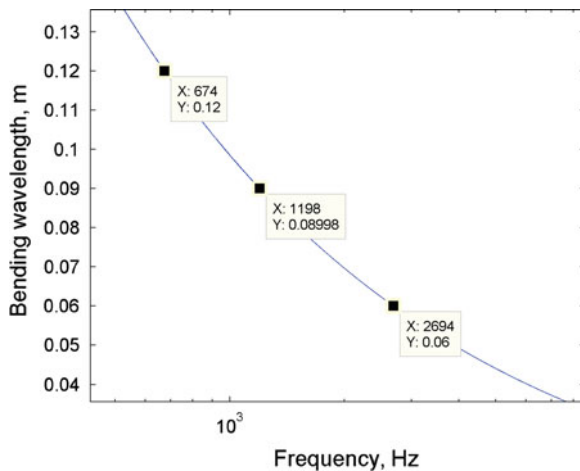
Solving (16) using (17) and (18) and evaluating the integrations lead to the general form of equation of motion:

$$(-\omega^2 [\mathbf{M}] + [\mathbf{K}]) \mathbf{W}_{mn} = \mathbf{F}_{mn} \quad (22)$$

The formula to compute matrices $\mathbf{M}$ and $\mathbf{K}$ is given in Appendix section. Equation (22) will be used to compute Monte-Carlo simulation for vibration response of a non-deterministic rectangular plate attached with a PZT patch actuator. However, the validity of the response using this equation is only up to the frequency where the bending wavelength of the plate is comparable to the PZT patch dimensions at high modal overlap factor, M [11]. Plate bending wavelength is given as [1]:

$$\lambda = \frac{2\pi}{k_B} = \frac{2\pi}{\sqrt{4\omega^2 \rho h / D}} \quad (23)$$

Fig. 4 Plate bending wavelength



From Fig. 4, it can be seen that the PZT patch dimension used in this research ($6 \text{ cm} \times 5 \text{ cm}$), has cut-off frequency at approximately 2 kHz. Therefore, the final EOM obtained in (22) is valid only up to this frequency. The importance of this limitation will be further illustrated in Sect. 4.

3 Power Delivered to a Non-deterministic Thin Rectangular Plate

To obtain the power delivered to the plate, the following equations for energy of the plate due to vibration are first attained [6]:

$$E = \frac{1}{2} \mathbf{W}_{mn}^T \mathbf{K}_C \mathbf{W}_{mn} \quad \text{or} \quad E = \frac{1}{2} \omega^2 \mathbf{W}_{mn}^T \mathbf{M} \mathbf{W}_{mn} \quad (24)$$

which uses complex modal stiffness matrix, $\mathbf{K}_C = \mathbf{K}(1 + j\eta)$ and mass matrix, $\mathbf{M}$ respectively. ω is frequency, η is modal loss factor and $\mathbf{K}$ is stiffness matrix. The extra term $(1 + j\eta)$ is incorporated by assuming a hysteretic damping model, introduced for proportional structural damping effect to the system [7]. Since the system is not coupled to any other subsystem (single system), power input, P_{in} is equal to power dissipated from the system.

$$P_{in} = P_s = \eta \omega E \quad (25)$$

$$P_{in} = \frac{1}{2} \eta \omega \mathbf{W}_{mn}^T \mathbf{K}_C \mathbf{W}_{mn} \quad \text{or} \quad P_{in} = \frac{1}{2} \eta \omega^3 \mathbf{W}_{mn}^T \mathbf{M} \mathbf{W}_{mn} \quad (26)$$

Equation (26) is used to compute the power input to each sample in Monte-Carlo simulation given in Sect. 4.

4 Simulation Studies

A rectangular plate made of steel is used for simulation studies. The properties of the plate and PZT patch actuator used are tabulated in Table 1. The centre of the PZT patch is located at 0.25 of the length and width of the plate. Twenty point masses with total mass 20 % of the plate mass are located randomly on the plate surface. To simulate high-order mode response, high modal overlap factor, M are captured by using high numbers of half-waves in the x direction and y direction; $m = 15$ and $n = 15$ respectively. Modal loss factor of 0.05 is used. Monte-Carlo simulation is employed by taking the response for 50 ensembles (location of point masses is randomized for each simulation), (26) is then used to obtain power delivered by the patch actuator and then the responses are averaged to get the

Table 1 Properties of plate and PZT patch actuator used

Properties	Plate (steel)	PZT actuator (PZT-5A)
Young’s modulus	200×10^9 (Pa)	70×10^9 (Pa)
Density	$7,850$ (kg/m ³)	$7,750$ (kg/m ³)
Length $\times$ width $\times$ thickness	$0.7 \times 0.6 \times 0.001$ (m)	$6 \times 5 \times 0.02$ (cm)
Poisson’s ratio	0.33	0.31
Piezoelectric constant, d_{31}	–	-1.71×10^{-10} (V/m)

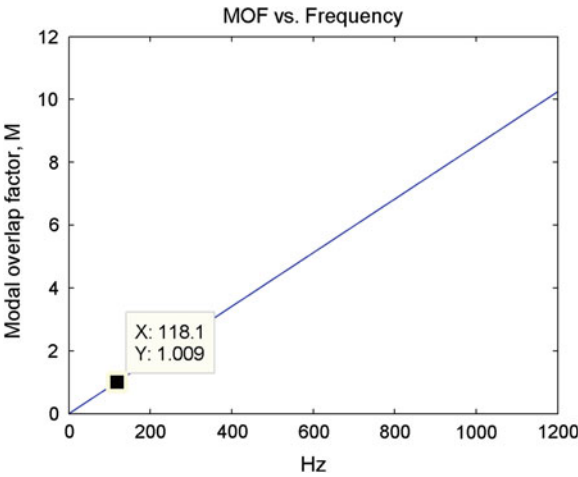
Table 2 Modal overlap factor and its corresponding frequency value

MOF	1	2	3	4	5	6	7	8
Frequency (Hz)	118	237	355	473	592	710	828	947

ensemble average power. Table 2 shows frequency range and its corresponding modal overlap factor for the plate (Fig. 5).

For each ensemble in Fig. 6, the response exhibits small magnitude difference at M less than unity, but significant as frequency increases. Figure 7 shows the closed-up view at deterministic response range i.e. when modal overlap factor, M is less than unity. The ensemble average exhibits distinct narrow peaks which indicate resonant modal response of the system. As M approaches unity, the individual modal responses begin to overlap and as it increases beyond unity, the peaks gradually become broader and indistinct i.e. resonant modal response is not clearly observable. The frequency response of multi-mode system in the high frequency range is very sensitive to small disturbance due to the randomly located point masses in the system. Hence the detail of high frequency response is unpredictable i.e. non-deterministic.

Fig. 5 Plot of modal overlap factor versus frequency



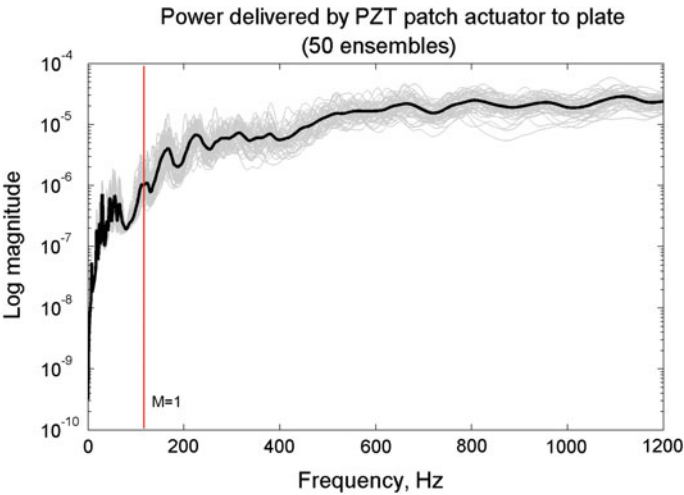
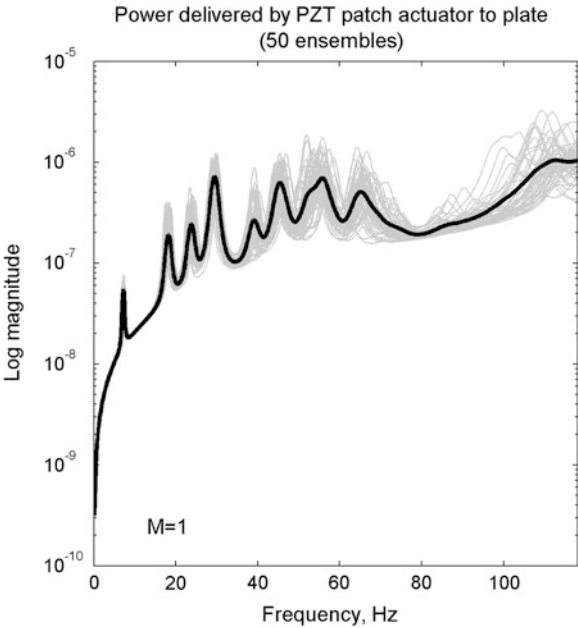


Fig. 6 Ensemble average of power delivered by a PZT actuator to plate (*solid black line*) for 50 ensembles

Fig. 7 Close-up view of deterministic response range ($M < I$) for ensemble average of power delivered by PZT actuator to plate (*solid black line*)



4.1 Influence of PZT Patch Actuator Location and Size on the Average Power Delivered to the Structure

Two case studies are considered in order to further investigate how the average power delivered by PZT patch actuator to the plate is affected when physical parameter of the system is changed.

4.1.1 Case 1: Changing PZT Actuator Location on Plate

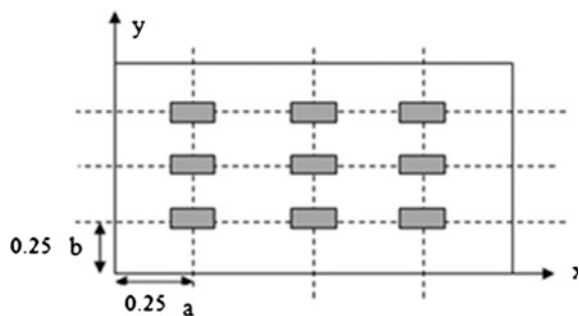
Nine different grid locations on the plate, as shown in Fig. 8, are taken to locate the patch. For each location, ensemble average power delivered by the patch to the plate (50 simulations) is plotted in Figs. 9 and 10. At deterministic response range ($M < 1$) shown in Fig. 10, it can be seen that narrow peaks are exhibited to indicate resonant modal response of the system. When patch is located in the middle of the plate (coordinate (0.5, 0.5), black solid line), it is observed that no modal response exists at even mode numbers (i.e. 2 and 4). This is because mode cannot be excited where force is located. At higher frequency range, it is observed that the peaks become broader and indistinct. In addition, the difference in magnitude between responses at the non-deterministic range is not significant. An important observation can also be made that different PZT actuator location on structure does not affect the ensemble average power delivered.

4.1.2 Case 2: Changing PZT Actuator Size

Five different sizes of PZT actuator are taken; the original size (6 cm $\times$ 5 cm), 0.25, 0.5, 1.5 and twice of the original size are considered separately. Ensemble average power from 50 simulations is plotted in Fig. 11.

From Fig. 11, a general observation can be made that the magnitude of response decreases as the PZT actuator size decreases however the response shape between

Fig. 8 Schematic figure of nine different location of PZT actuator on plate



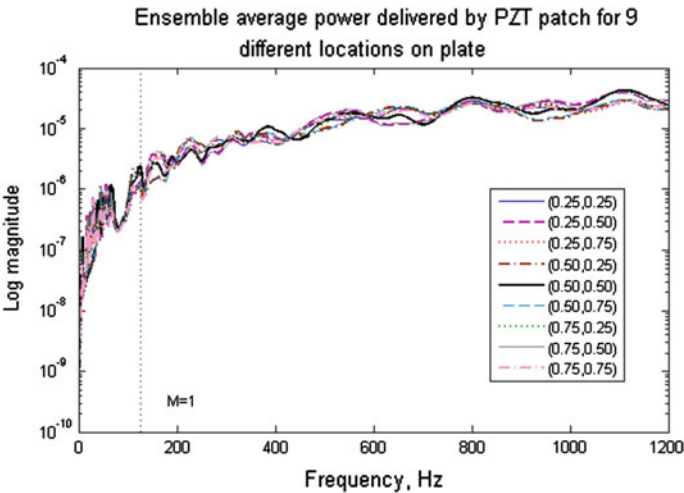


Fig. 9 Ensemble average power delivered (50 ensembles each) by a PZT actuator located at nine different locations on the plate

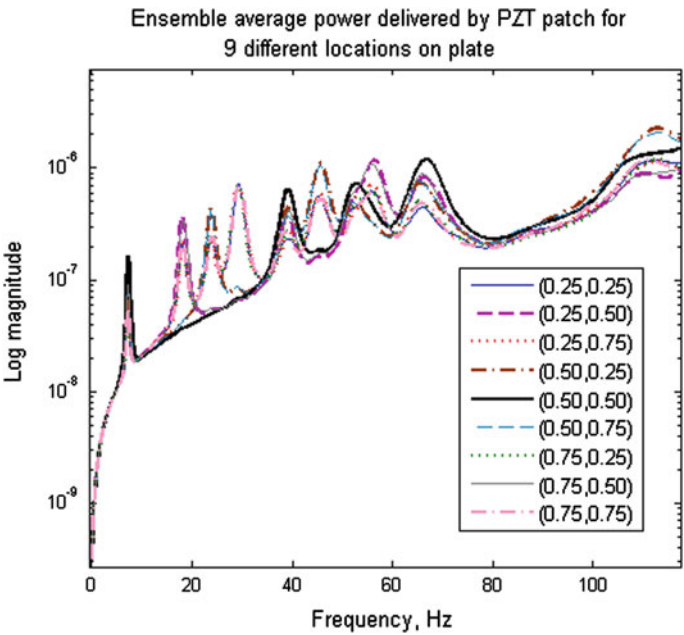


Fig. 10 Close-up view at deterministic response range ($M < 1$) of ensemble average power delivered by a PZT actuator attached at nine different locations on the plate

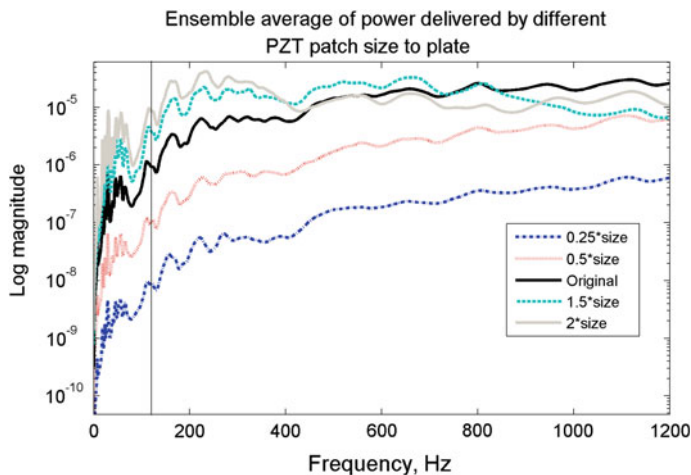


Fig. 11 Ensemble average power delivered (50 ensembles each) by different sizes of PZT patch actuator to the plate

different PZT sizes has no significant difference. There are some variations when the patch size is made too big or too small from the original size which indicates that mass effect becomes important at high frequency. The inertia effect tends to pull down the amplitude of the response in proportion to ω which explains the increasing variance as patch size is made bigger. Another reason is due to the ratio between bending wavelength and the size of the patch. As explained earlier, the response is only valid up to the frequencies where the size of the patch is comparable to the size of bending wavelength. This is illustrated in Fig. 4 in which for 1.5 and 2 times of the original patch size (9 cm $\times$ 7 cm and 12 cm $\times$ 10 cm), the cut-off frequencies are at about 1,000 and 600 Hz respectively. It can be seen at these frequencies, the Lagrangian responses for 1.5 and 2 times patch size drop suddenly (Fig. 11).

5 Conclusions

PZT material is used widely as actuators for excitation or structural vibration control because it is compact and has wide frequency range. Understanding the way vibrational response behaves at high frequency that is by knowing the average power delivered by PZT actuator to a non-deterministic structure is a huge advantage for analysing structural response using SEA method. This research has presented simulation studies on the dynamic effects of a Non-DS when a PZT patch actuator is attached. Parametric studies showed that changing the patch location on the structure will not affect the average power supplied. On the contrary, changing the patch size will change the power magnitude proportionally as larger patch means larger force is applied. However, there are some variations for ensemble

average which becomes more significant as the patch size is made too big or too small. This is due to the inertial effect which pulls down the amplitude of the response in proportion to ω . Also, the response is only valid up to the frequencies where the size of the patch is comparable to the size of bending wavelength.

Appendix

Recall that the kinetic energy and potential energy of the system are:

$$\begin{aligned}
 U &= \frac{D}{2} \iint \left\{ \left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} \right)^2 - 2(1 - \nu) \left[\left(\frac{\partial^2 w}{\partial x^2} \frac{\partial^2 w}{\partial y^2} \right) - \left(\frac{\partial^2 w}{\partial x \partial y} \right)^2 \right] \right\} dA \\
 &+ \frac{1}{2} \iint D_1(x, y) \left\{ \left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} \right)^2 - 2(1 - \nu) \left[\left(\frac{\partial^2 w}{\partial x^2} \frac{\partial^2 w}{\partial y^2} \right) - \left(\frac{\partial^2 w}{\partial x \partial y} \right)^2 \right] \right\} dA \\
 &+ \frac{1}{2} \iint D_2(x, y) \left\{ \left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} \right)^2 - 2(1 - \nu_{pz}) \left[\left(\frac{\partial^2 w}{\partial x^2} \frac{\partial^2 w}{\partial y^2} \right) - \left(\frac{\partial^2 w}{\partial x \partial y} \right)^2 \right] \right\} dA \\
 &+ B(t) \iint S(x, y) \left[(d_{31} + \nu_{pz} d_{32}) \frac{\partial^2 w}{\partial x^2} + (d_{32} + \nu_{pz} d_{31}) \frac{\partial^2 w}{\partial y^2} \right] dA \\
 T &= \frac{1}{2} \rho h \iint \left(\frac{\partial w}{\partial t} \right)^2 dA + \frac{1}{2} \rho_{pz} t_p \iint \left(\frac{\partial w}{\partial t} \right)^2 S(x, y) dA \\
 &+ \frac{1}{2} \iint \sum_r m_r \delta(x - x_r, y - y_r) \left(\frac{\partial w}{\partial t} \right)^2 dA
 \end{aligned}$$

The assumed solution given in double series as in (8), is substituted into these energy equations. The product of multiplication between two double series or its derivatives will have four conditions, i.e. (i) $m = p$ and $n = q$, (ii) $m = p$ and $n \neq q$, (iii) $m \neq p$ and $n = q$ and (iv) $m \neq p$ and $n \neq q$, and there are nine integrals involved which can be solved as follows:

$$\begin{aligned}
 I_1 &= \iint \left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} \right)^2 dA \\
 &= \sum_m \sum_n \sum_p \sum_q W_{mn} W_{pq} \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right] \left[\left(\frac{p\pi}{a} \right)^2 + \left(\frac{q\pi}{b} \right)^2 \right] \\
 &\quad \int_0^a \sin \frac{m\pi x}{a} \sin \frac{p\pi x}{a} dx \int_0^b \sin \frac{n\pi y}{b} \sin \frac{q\pi y}{b} dy \\
 &= \frac{ab}{4} \sum_m \sum_n W_{mn}^2 \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right]^2
 \end{aligned}$$

$$\begin{aligned}
 I_2 &= \iint \frac{\partial^2 w}{\partial x^2} \frac{\partial^2 w}{\partial y^2} dA \\
 &= \sum_m \sum_n \sum_p \sum_q W_{mn} W_{pq} \frac{mnpq\pi^4}{a^2 b^2} \int_0^a \sin \frac{m\pi x}{a} \sin \frac{p\pi x}{a} dx \int_0^b \sin \frac{n\pi y}{b} \sin \frac{q\pi y}{b} dy \\
 &= \frac{ab}{4} \sum_m \sum_n W_{mn}^2 \frac{m^2 n^2 \pi^4}{a^2 b^2}
 \end{aligned}$$

$$\begin{aligned}
 I_3 &= \iint \left(\frac{\partial^2 w}{\partial x \partial y} \right)^2 dA \\
 &= \sum_m \sum_n \sum_p \sum_q W_{mn} W_{pq} \frac{mnpq\pi^4}{a^2 b^2} \int_0^a \cos \frac{m\pi x}{a} \cos \frac{p\pi x}{a} dx \int_0^b \cos \frac{n\pi y}{b} \cos \frac{q\pi y}{b} dy \\
 &= \frac{ab}{4} \sum_m \sum_n W_{mn}^2 \frac{m^2 n^2 \pi^4}{a^2 b^2}
 \end{aligned}$$

$$\begin{aligned}
 I_4 &= \iint S(x, y) \left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} \right)^2 dA \\
 &= \sum_m \sum_n \sum_p \sum_q W_{mn} W_{pq} \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right] \left[\left(\frac{p\pi}{a} \right)^2 + \left(\frac{q\pi}{b} \right)^2 \right] \\
 &\quad \int_{x1}^{x2} \sin \frac{m\pi x}{a} \sin \frac{p\pi x}{a} dx \int_{y1}^{y2} \sin \frac{n\pi y}{b} \sin \frac{q\pi y}{b} dy \\
 &= \sum_m \sum_n W_{mn}^2 \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right]^2 C_{1x} C_{1y} \\
 &\quad + \sum_m \sum_n \sum_{p \neq m} \sum_{q \neq n} W_{mn} W_{pq} \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right] \left[\left(\frac{p\pi}{a} \right)^2 + \left(\frac{q\pi}{b} \right)^2 \right] C_{2x} \cdot C_{2y} \\
 &\quad + \sum_m \sum_n \sum_{p=m} \sum_{q \neq n} W_{mn} W_{pq} \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right] \left[\left(\frac{p\pi}{a} \right)^2 + \left(\frac{q\pi}{b} \right)^2 \right] C_{1x} \cdot C_{2y} \\
 &\quad + \sum_m \sum_n \sum_{p \neq m} \sum_{q=n} W_{mn} W_{pq} \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right] \left[\left(\frac{p\pi}{a} \right)^2 + \left(\frac{q\pi}{b} \right)^2 \right] C_{2x} \cdot C_{1y}
 \end{aligned}$$

$$\begin{aligned}
 I_5 &= \iint S(x, y) \frac{\partial^2 w}{\partial x^2} \frac{\partial^2 w}{\partial y^2} dA \\
 &= \sum_m \sum_n \sum_p \sum_q W_{mn} W_{pq} \frac{mnpq\pi^4}{a^2 b^2} \int_{x1}^{x2} \sin \frac{m\pi x}{a} \sin \frac{p\pi x}{a} dx \int_{y1}^{y2} \sin \frac{n\pi y}{b} \sin \frac{q\pi y}{b} dy \\
 &= \sum_m \sum_n W_{mn}^2 \frac{m^2 n^2 \pi^4}{a^2 b^2} C_{1x} C_{1y} + \sum_m \sum_n \sum_{p \neq m} \sum_{q \neq n} W_{mn} W_{pq} \frac{mnpq\pi^4}{a^2 b^2} C_{2x} C_{2y} \\
 &\quad + \sum_m \sum_n \sum_{p=m} \sum_{q \neq n} W_{mn} W_{pq} \frac{mnpq\pi^4}{a^2 b^2} C_{1x} C_{2y} \\
 &\quad + \sum_m \sum_n \sum_{p \neq m} \sum_{q=n} W_{mn} W_{pq} \frac{mnpq\pi^4}{a^2 b^2} C_{2x} C_{1y}
 \end{aligned}$$

$$\begin{aligned}
I_6 &= \iint S(x, y) \left(\frac{\partial^2 w}{\partial x \partial y} \right)^2 dA \\
&= \sum_m \sum_n \sum_p \sum_q W_{mn} W_{pq} \frac{mnpq\pi^4}{a^2 b^2} \int_{x1}^{x2} \cos \frac{m\pi x}{a} \cos \frac{p\pi x}{a} dx \int_{y1}^{y2} \cos \frac{n\pi y}{b} \cos \frac{q\pi y}{b} dy \\
&= \sum_m \sum_n W_{mn}^2 \frac{m^2 n^2 \pi^4}{a^2 b^2} C_{3x} C_{3y} + \sum_m \sum_n \sum_{p \neq m} \sum_{q \neq n} W_{mn} W_{pq} \frac{mnpq\pi^4}{a^2 b^2} C_{4x} C_{4y} \\
&\quad + \sum_m \sum_n \sum_{p=m} \sum_{q \neq n} W_{mn} W_{pq} \frac{mnpq\pi^4}{a^2 b^2} C_{3x} C_{4y} \\
&\quad + \sum_m \sum_n \sum_{p \neq m} \sum_{q=n} W_{mn} W_{pq} \frac{mnpq\pi^4}{a^2 b^2} C_{4x} C_{3y}
\end{aligned}$$

$$\begin{aligned}
I_7 &= \iint S(x, y) \frac{\partial^2 w}{\partial x^2} dA = - \sum_m \sum_n W_{mn} \left(\frac{m\pi}{a} \right)^2 \int_{x1}^{x2} \sin \frac{m\pi x}{a} dx \int_{y1}^{y2} \sin \frac{n\pi y}{b} dy \\
&= - \sum_m \sum_n W_{mn} \left(\frac{m\pi}{a} \right)^2 C_{5x} C_{5y}
\end{aligned}$$

$$\begin{aligned}
I_8 &= \iint S(x, y) \frac{\partial^2 w}{\partial y^2} dA = - \sum_m \sum_n W_{mn} \left(\frac{n\pi}{b} \right)^2 \int_{x1}^{x2} \sin \frac{m\pi x}{a} dx \int_{y1}^{y2} \sin \frac{n\pi y}{b} dy \\
&= - \sum_m \sum_n W_{mn} \left(\frac{n\pi}{b} \right)^2 C_{5x} C_{5y}
\end{aligned}$$

$$\begin{aligned}
I_9 &= \iint \left(\frac{\partial w}{\partial t} \right)^2 dA = \sum_m \sum_n \sum_p \sum_q \dot{W}_{mn} \dot{W}_{pq} \int_{x1}^{x2} \sin \frac{m\pi x}{a} \sin \frac{p\pi x}{a} dx \int_{y1}^{y2} \sin \frac{n\pi y}{b} \sin \frac{q\pi y}{b} dy \\
&= \sum_m \sum_n \dot{W}_{mn}^2 \frac{ab}{4}
\end{aligned}$$

$$\begin{aligned}
I_{10} &= \iint S(x, y) \left(\frac{\partial w}{\partial t} \right)^2 dA \\
&= \sum_m \sum_n \sum_p \sum_q \dot{W}_{mn} \dot{W}_{pq} \int_{x1}^{x2} \sin \frac{m\pi x}{a} \sin \frac{p\pi x}{a} dx \int_{y1}^{y2} \sin \frac{n\pi y}{b} \sin \frac{q\pi y}{b} dy \\
&= \sum_m \sum_n \dot{W}_{mn}^2 \left\{ \frac{\rho hab}{8} + \frac{\rho_{piezo} t_p}{2} C_{1x} C_{1y} \right\} + \sum_m \sum_n \sum_{p \neq m} \sum_{q \neq n} \dot{W}_{mn} \dot{W}_{pq} \frac{\rho_{piezo} t_p}{2} C_{2x} C_{2y} \\
&\quad + \sum_m \sum_n \sum_{p=m} \sum_{q \neq n} \dot{W}_{mn} \dot{W}_{pq} \frac{\rho_{piezo} t_p}{2} C_{1x} C_{2y} \\
&\quad + \sum_m \sum_n \sum_{p \neq m} \sum_{q=n} \dot{W}_{mn} \dot{W}_{pq} \frac{\rho_{piezo} t_p}{2} C_{2x} C_{1y}
\end{aligned}$$

$$\begin{aligned}
 I_{11} &= \iint \delta(x - x_r, y - y_r) \left(\frac{\partial w}{\partial t} \right)^2 dA \\
 &= \sum_m \sum_n \sum_p \sum_q \dot{W}_{mn} \dot{W}_{pq} \sin \frac{m\pi x_r}{a} \sin \frac{p\pi x_r}{a} \sin \frac{n\pi y_r}{b} \sin \frac{q\pi y_r}{b}
 \end{aligned}$$

where the coefficients C are

- $C_{1x} = \int_{x_1}^{x_2} \sin^2 \frac{m\pi x}{a} dx = \frac{1}{2} \left[(x_2 - x_1) - \frac{a}{2m\pi} (\sin \frac{2m\pi x_2}{a} - \sin \frac{2m\pi x_1}{a}) \right]$
- $C_{1y} = \int_{y_1}^{y_2} \sin^2 \frac{n\pi y}{b} dy = \frac{1}{2} \left[(y_2 - y_1) - \frac{b}{2n\pi} (\sin \frac{2n\pi y_2}{b} - \sin \frac{2n\pi y_1}{b}) \right]$
-
- $C_{2x} = \int_{x_1}^{x_2} \sin \frac{m\pi x}{a} \sin \frac{p\pi x}{a} dx = \frac{a}{2(m-p)\pi} \left[\sin \frac{(m-p)\pi x_2}{a} - \sin \frac{(m-p)\pi x_1}{a} \right]$
 $\quad - \frac{a}{2(m+p)\pi} \left[\sin \frac{(m+p)\pi x_2}{a} - \sin \frac{(m+p)\pi x_1}{a} \right]$
- $C_{2y} = \int_{y_1}^{y_2} \sin \frac{n\pi y}{b} \sin \frac{q\pi y}{b} dy = \frac{b}{2(n-q)\pi} \left[\sin \frac{(n-q)\pi y_2}{b} - \sin \frac{(n-q)\pi y_1}{b} \right]$
 $\quad - \frac{b}{2(n+q)\pi} \left[\sin \frac{(n+q)\pi y_2}{b} - \sin \frac{(n+q)\pi y_1}{b} \right]$
-
- $C_{3x} = \int_{x_1}^{x_2} \cos^2 \frac{m\pi x}{a} dx = \frac{1}{2} \left[(x_2 - x_1) + \frac{a}{2m\pi} (\sin \frac{2m\pi x_2}{a} - \sin \frac{2m\pi x_1}{a}) \right]$
- $C_{3y} = \int_{y_1}^{y_2} \cos^2 \frac{n\pi y}{b} dy = \frac{1}{2} \left[(y_2 - y_1) + \frac{b}{2n\pi} (\sin \frac{2n\pi y_2}{b} - \sin \frac{2n\pi y_1}{b}) \right]$
-
- $C_{4x} = \int_{x_1}^{x_2} \cos \frac{m\pi x}{a} \cos \frac{p\pi x}{a} dx = \frac{a}{2(m-p)\pi} \left[\sin \frac{(m-p)\pi x_2}{a} - \sin \frac{(m-p)\pi x_1}{a} \right]$
 $\quad + \frac{a}{2(m+p)\pi} \left[\sin \frac{(m+p)\pi x_2}{a} - \sin \frac{(m+p)\pi x_1}{a} \right]$
- $C_{4y} = \int_{y_1}^{y_2} \cos \frac{n\pi y}{b} \cos \frac{q\pi y}{b} dy = \frac{b}{2(n-q)\pi} \left[\sin \frac{(n-q)\pi y_2}{b} - \sin \frac{(n-q)\pi y_1}{b} \right]$
 $\quad + \frac{b}{2(n+q)\pi} \left[\sin \frac{(n+q)\pi y_2}{b} - \sin \frac{(n+q)\pi y_1}{b} \right]$
-
- $C_{5x} = \int_{x_1}^{x_2} \sin^2 \frac{m\pi x}{a} dx = \frac{a}{m\pi} (\cos \frac{m\pi x_1}{a} - \cos \frac{m\pi x_2}{a})$
- $C_{5y} = \int_{y_1}^{y_2} \sin^2 \frac{n\pi y}{b} dy = \frac{b}{n\pi} (\cos \frac{n\pi y_1}{b} - \cos \frac{n\pi y_2}{b})$

Thus, the potential and kinetic energies of the system are:

$$\begin{aligned}
 U_{total} = & \sum_m \sum_n W_{mn}^2 \left\{ \frac{abD}{8} \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right]^2 \right. \\
 & + \frac{D_1}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right]^2 C_{1x}C_{1y} - 2(1-\nu) \left(\frac{mn\pi^2}{ab} \right)^2 (C_{1x}C_{1y} - C_{3x}C_{3y}) \right\} \\
 & + \frac{D_2}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right]^2 C_{1x}C_{1y} \right. \\
 & \left. - 2(1-\nu_{pzt}) \left(\frac{mn\pi^2}{ab} \right)^2 (C_{1x}C_{1y} - C_{3x}C_{3y}) \right\} \Bigg\} \\
 & + \sum_m \sum_n \sum_{p \neq m} \sum_{q \neq n} W_{mn} W_{pq} \left\{ \frac{D_1}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right] \left[\left(\frac{p\pi}{a} \right)^2 + \left(\frac{q\pi}{b} \right)^2 \right] C_{2x}C_{2y} \right. \right. \\
 & - 2(1-\nu) \left(\frac{mnpq\pi^4}{a^2b^2} \right) (C_{2x}C_{2y} - C_{4x}C_{4y}) \Bigg\} \\
 & + \frac{D_2}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right] \left[\left(\frac{p\pi}{a} \right)^2 + \left(\frac{q\pi}{b} \right)^2 \right] C_{2x}C_{2y} \right. \\
 & \left. - 2(1-\nu_{pzt}) \left(\frac{mnpq\pi^4}{a^2b^2} \right) (C_{2x}C_{2y} - C_{4x}C_{4y}) \right\} \Bigg\} \\
 & + \sum_m \sum_n \sum_{p=m} \sum_{q \neq n} W_{mn} W_{pq} \left\{ \frac{D_1}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right] \left[\left(\frac{p\pi}{a} \right)^2 + \left(\frac{q\pi}{b} \right)^2 \right] C_{1x}C_{2y} \right. \right. \\
 & - 2(1-\nu) \left(\frac{mnpq\pi^4}{a^2b^2} \right) (C_{1x}C_{2y} - C_{3x}C_{4y}) \Bigg\} \\
 & + \frac{D_2}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right] \left[\left(\frac{p\pi}{a} \right)^2 + \left(\frac{q\pi}{b} \right)^2 \right] C_{1x}C_{2y} \right. \\
 & \left. - 2(1-\nu_{pzt}) \left(\frac{mnpq\pi^4}{a^2b^2} \right) (C_{1x}C_{2y} - C_{3x}C_{4y}) \right\} \Bigg\} \\
 & + \sum_m \sum_n \sum_{p \neq m} \sum_{q=n} W_{mn} W_{pq} \left\{ \frac{D_1}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right] \left[\left(\frac{p\pi}{a} \right)^2 + \left(\frac{q\pi}{b} \right)^2 \right] C_{2x}C_{1y} \right. \right. \\
 & - 2(1-\nu) \left(\frac{mnpq\pi^4}{a^2b^2} \right) (C_{2x}C_{1y} - C_{4x}C_{3y}) \Bigg\} \\
 & + \frac{D_2}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right] \left[\left(\frac{p\pi}{a} \right)^2 + \left(\frac{q\pi}{b} \right)^2 \right] C_{2x}C_{1y} \right. \\
 & \left. - 2(1-\nu_{pzt}) \left(\frac{mnpq\pi^4}{a^2b^2} \right) (C_{2x}C_{1y} - C_{4x}C_{3y}) \right\} \Bigg\} \\
 & + \sum_m \sum_n W_{mn} B(t) \left[(d_{31} + \nu_{pzt}d_{32}) \left(\frac{m\pi}{a} \right)^2 + (d_{32} + \nu_{pzt}d_{31}) \left(\frac{n\pi}{b} \right)^2 \right] C_{5x}C_{5y}
 \end{aligned}$$

$$\begin{aligned}
T_{total} = & \sum_m \sum_n \dot{W}_{mn}^2 \left\{ \frac{\rho hab}{8} + \frac{\rho_{piezo} t_p}{2} C_{1x} C_{1y} \right\} \\
& + \sum_m \sum_n \sum_{p \neq m} \sum_{q \neq n} \dot{W}_{mn} \dot{W}_{pq} \frac{\rho_{piezo} t_p}{2} C_{2x} C_{2y} \\
& + \sum_m \sum_n \sum_{p=m} \sum_{q \neq n} \dot{W}_{mn} \dot{W}_{pq} \frac{\rho_{piezo} t_p}{2} C_{1x} C_{2y} \\
& + \sum_m \sum_n \sum_{p \neq m} \sum_{q=n} \dot{W}_{mn} \dot{W}_{pq} \frac{\rho_{piezo} t_p}{2} C_{2x} C_{1y} \\
& + \sum_m \sum_n \sum_p \sum_q \dot{W}_{mn} \dot{W}_{pq} \sin \frac{m\pi x_r}{a} \sin \frac{p\pi x_r}{a} \sin \frac{n\pi y_r}{b} \sin \frac{q\pi y_r}{b}
\end{aligned}$$

Lagrange's method is used to derive the EOM of the system. Since there is no external force acting on the system, (that is, force is generated internally by the piezoelectric patch), the Lagrange's equation becomes

$$\frac{\partial}{\partial t} \left(\frac{\partial L}{\partial \dot{W}_{mn}} \right) - \frac{\partial L}{\partial W_{mn}} = 0 \quad \text{where} \quad L = T - U$$

the final EOM is as follows:

$$\begin{aligned}
& -\omega^2 \left\{ W_{mn} \left\{ \frac{\rho hab}{4} + \rho_{pz} t_p C_{1x} C_{1y} \right\} \right. \\
& + \frac{\rho_{pz} t_p}{2} \left\{ \sum_{p \neq m} \sum_{q \neq n} W_{pq} C_{2x} C_{2y} + \sum_{p=m} \sum_{q \neq n} W_{pq} C_{1x} C_{2y} + \sum_{p \neq m} \sum_{q=n} W_{pq} C_{2x} C_{1y} \right\} \\
& + \sum_r m_r \sum_p \sum_q W_{pq} \sin \frac{m\pi x_r}{a} \sin \frac{p\pi x_r}{a} \sin \frac{n\pi y_r}{b} \sin \frac{q\pi y_r}{b} \left. \right\} \\
& + 2W_{mn} \left\{ \frac{abD}{8} \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right]^2 \right. \\
& + \frac{D_1}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right]^2 C_{1x} C_{1y} - 2(1-\nu) \left(\frac{mn\pi^2}{ab} \right)^2 (C_{1x} C_{1y} - C_{3x} C_{3y}) \right\} \\
& + \frac{D_2}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right]^2 C_{1x} C_{1y} \right. \\
& \left. \left. - 2(1-\nu_{pz}) \left(\frac{mn\pi^2}{ab} \right)^2 (C_{1x} C_{1y} - C_{3x} C_{3y}) \right\} \right\}
\end{aligned}$$

$$\begin{aligned}
& + \sum_{p \neq m} \sum_{q \neq n} W_{pq} \left\{ \frac{D_1}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right] \left[\left(\frac{p\pi}{a} \right)^2 + \left(\frac{q\pi}{b} \right)^2 \right] C_{2x} C_{2y} \right. \right. \\
& \quad \left. \left. - 2(1 - \nu) \left(\frac{mnpq\pi^4}{a^2 b^2} \right) (C_{2x} C_{2y} - C_{4x} C_{4y}) \right\} \right. \\
& \quad \left. + \frac{D_2}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right] \left[\left(\frac{p\pi}{a} \right)^2 + \left(\frac{q\pi}{b} \right)^2 \right] C_{2x} C_{2y} \right. \right. \\
& \quad \left. \left. - 2(1 - \nu_{pzl}) \left(\frac{mnpq\pi^4}{a^2 b^2} \right) (C_{2x} C_{2y} - C_{4x} C_{4y}) \right\} \right\} \\
& + \sum_{p \neq m} \sum_{q \neq n} W_{pq} \left\{ \frac{D_1}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right] \left[\left(\frac{p\pi}{a} \right)^2 + \left(\frac{q\pi}{b} \right)^2 \right] C_{1x} C_{2y} \right. \right. \\
& \quad \left. \left. - 2(1 - \nu) \left(\frac{mnpq\pi^4}{a^2 b^2} \right) (C_{1x} C_{2y} - C_{3x} C_{4y}) \right\} \right. \\
& \quad \left. + \frac{D_2}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right] \left[\left(\frac{p\pi}{a} \right)^2 + \left(\frac{q\pi}{b} \right)^2 \right] C_{1x} C_{2y} \right. \right. \\
& \quad \left. \left. - 2(1 - \nu_{pzl}) \left(\frac{mnpq\pi^4}{a^2 b^2} \right) (C_{1x} C_{2y} - C_{3x} C_{4y}) \right\} \right\} \\
& + \sum_{p \neq m} \sum_{q \neq n} W_{pq} \left\{ \frac{D_1}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right] \left[\left(\frac{p\pi}{a} \right)^2 + \left(\frac{q\pi}{b} \right)^2 \right] C_{2x} C_{1y} \right. \right. \\
& \quad \left. \left. - 2(1 - \nu) \left(\frac{mnpq\pi^4}{a^2 b^2} \right) (C_{2x} C_{1y} - C_{4x} C_{3y}) \right\} \right. \\
& \quad \left. + \frac{D_2}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right] \left[\left(\frac{p\pi}{a} \right)^2 + \left(\frac{q\pi}{b} \right)^2 \right] C_{2x} C_{1y} \right. \right. \\
& \quad \left. \left. - 2(1 - \nu_{pzl}) \left(\frac{mnpq\pi^4}{a^2 b^2} \right) (C_{2x} C_{1y} - C_{4x} C_{3y}) \right\} \right\} \\
& = -B(t) \left[(d_{31} + \nu_{pzl} d_{32}) \left(\frac{m\pi}{a} \right)^2 + (d_{32} + \nu_{pzl} d_{31}) \left(\frac{n\pi}{b} \right)^2 \right] C_{5x} C_{5y}
\end{aligned}$$

Terms in mass matrix (denoted by M_{mnpq}), stiffness matrix (denoted by K_{mnpq}) and force vector (denoted by F_{mn}) are:

$$\begin{aligned}
M_{mnpq} &= \frac{\rho h a b}{4} + \rho_{pzl} t_p C_{1x} C_{1y} + \frac{\rho_{pzl} t_p}{2} \left\{ \sum_{p \neq m} \sum_{q \neq n} C_{2x} C_{2y} + \sum_{p=m} \sum_{q \neq n} C_{1x} C_{2y} + \sum_{p \neq m} \sum_{q=n} C_{2x} C_{1y} \right\} \\
&+ \sum_r m_r \sum_p \sum_q \sin \frac{m\pi x_r}{a} \sin \frac{p\pi x_r}{a} \sin \frac{n\pi y_r}{b} \sin \frac{q\pi y_r}{b}
\end{aligned}$$

$$\begin{aligned}
K_{mnpq} = & 2 \left\{ \frac{abD}{8} \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right]^2 \right. \\
& + \frac{D_1}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right]^2 C_{1x}C_{1y} - 2(1-\nu) \left(\frac{mn\pi^2}{ab} \right)^2 (C_{1x}C_{1y} - C_{3x}C_{3y}) \right\} \\
& + \frac{D_2}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right]^2 C_{1x}C_{1y} \right. \\
& \left. \left. - 2(1-\nu_{pzt}) \left(\frac{mn\pi^2}{ab} \right)^2 (C_{1x}C_{1y} - C_{3x}C_{3y}) \right\} \right\} \\
& + \sum_{p \neq m} \sum_{q \neq n} \left\{ \frac{D_1}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right] \left[\left(\frac{p\pi}{a} \right)^2 + \left(\frac{q\pi}{b} \right)^2 \right] C_{2x}C_{2y} \right. \right. \\
& \left. \left. - 2(1-\nu) \left(\frac{mnpq\pi^4}{a^2b^2} \right) (C_{2x}C_{2y} - C_{4x}C_{4y}) \right\} \right. \\
& + \frac{D_2}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right] \left[\left(\frac{p\pi}{a} \right)^2 + \left(\frac{q\pi}{b} \right)^2 \right] C_{2x}C_{2y} \right. \\
& \left. \left. - 2(1-\nu_{pzt}) \left(\frac{mnpq\pi^4}{a^2b^2} \right) (C_{2x}C_{2y} - C_{4x}C_{4y}) \right\} \right\} \\
& + \sum_{p \neq m} \sum_{q \neq n} \left\{ \frac{D_1}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right] \left[\left(\frac{p\pi}{a} \right)^2 + \left(\frac{q\pi}{b} \right)^2 \right] C_{1x}C_{2y} \right. \right. \\
& \left. \left. - 2(1-\nu) \left(\frac{mnpq\pi^4}{a^2b^2} \right) (C_{1x}C_{2y} - C_{3x}C_{4y}) \right\} \right. \\
& + \frac{D_2}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right] \left[\left(\frac{p\pi}{a} \right)^2 + \left(\frac{q\pi}{b} \right)^2 \right] C_{1x}C_{2y} \right. \\
& \left. \left. - 2(1-\nu_{pzt}) \left(\frac{mnpq\pi^4}{a^2b^2} \right) (C_{1x}C_{2y} - C_{3x}C_{4y}) \right\} \right\} \\
& + \sum_{p \neq m} \sum_{q=n} \left\{ \frac{D_1}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right] \left[\left(\frac{p\pi}{a} \right)^2 + \left(\frac{q\pi}{b} \right)^2 \right] C_{2x}C_{1y} \right. \right. \\
& \left. \left. - 2(1-\nu) \left(\frac{mnpq\pi^4}{a^2b^2} \right) (C_{2x}C_{1y} - C_{4x}C_{3y}) \right\} \right. \\
& + \frac{D_2}{2} \left\{ \left[\left(\frac{m\pi}{a} \right)^2 + \left(\frac{n\pi}{b} \right)^2 \right] \left[\left(\frac{p\pi}{a} \right)^2 + \left(\frac{q\pi}{b} \right)^2 \right] C_{2x}C_{1y} \right. \\
& \left. \left. - 2(1-\nu_{pzt}) \left(\frac{mnpq\pi^4}{a^2b^2} \right) (C_{2x}C_{1y} - C_{4x}C_{3y}) \right\} \right\} \\
F_{mn} = & -B(t) \left[(d_{31} + \nu_{pzt}d_{32}) \left(\frac{m\pi}{a} \right)^2 + (d_{32} + \nu_{pzt}d_{31}) \left(\frac{n\pi}{b} \right)^2 \right] C_{5x}C_{5y}
\end{aligned}$$

where

$$[\mathbf{M}] = \begin{bmatrix} M_{1111} & M_{1112} & \dots & M_{11pq} \\ M_{1211} & M_{1212} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ M_{mn11} & \dots & \dots & M_{mnpq} \end{bmatrix} \quad [\mathbf{K}] = \begin{bmatrix} K_{1111} & K_{1112} & \dots & K_{11pq} \\ K_{1211} & K_{1212} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ K_{mn11} & \dots & \dots & K_{mnpq} \end{bmatrix}$$

$$[\mathbf{F}] = \begin{bmatrix} F_{11} \\ F_{12} \\ \dots \\ F_{mn} \end{bmatrix}$$

References

1. Fahy F, Gardonio P (2007) Sound and structural vibration—radiation, transmission and response, 2nd edn. Academic Press, Amsterdam
2. Muthalif AGA, Langley RS (2012) Active control of high-frequency vibration: optimization using the hybrid modelling method. *J Sound Vib* 331:969–2983
3. Lyon RH (1975) Statistical energy analysis of dynamical systems: theory and applications. MIT Press, Cambridge
4. Woodhouse J (1981) An approach to the theoretical background of statistical energy analysis applied to structural vibration. *J Acoust Soc Am* 69(6):1695–1709
5. Langley RS, Bercin AN (1994) Wave intensity analysis of high frequency vibrations. *Phil Trans R Soc A: Math Phys Eng Sci* 1681(346):489–499
6. Gardonio P, Brennan MJ (2004) Mobility and impedance methods in structural dynamics advanced applications in acoustics, noise and vibration: chapter 9. Spon Press, London
7. Preumont A (2011) Vibration control of active structures: an introduction, 3rd edn. Springer, Berlin
8. Muthalif AGA (2010) Investigating the average power supplied by a piezoelectric patch actuator when attached to random structures (Application to Vibration Control). CLA presentation of high impact research proposal Kuala Lumpur, Malaysia
9. Jalili N (2010) Piezoelectric based vibration control. Springer, New York
10. Weaver W, Timoshenko SP, Young DH (1990) Vibration problems in engineering, 5th edn. Wiley-Interscience, New York
11. Aoki Y, Gardonio P, Elliott SJ (2006) Strain transducers for active control-lumped parameter model. ISVR technical memorandum No 970ISVR University of Southampton, UK

Solubilities Prediction of Ginger Bioactive Compounds in Liquid Phase of Water by the COSMO-RS Method

Syaripah Za'imah Syed Jaapar, Noor Azian Morad and Yoshio Iwai

Abstract The solubilities in water of four main ginger bioactive compounds, 6-gingerol, 6-shogaol, 8-gingerol and 10-gingerol, were predicted using a conductor-like screening model for real solvent (COSMO-RS) calculations. This study was conducted since no experimental data are available for ginger bioactive compounds solubilities in liquid phase of water. The σ -profiles of these selected molecules were calculated using Gaussian software and the solubilities were calculated using the COSMO-RS method. The solubilities of these ginger bioactive compounds were calculated at 50–200 °C. In order to validate the accuracy of the COSMO-RS method, the solubilities of five hydrocarbon molecules were calculated using the COSMO-RS method and compared with the experimental data in the literature. The selected hydrocarbon molecules were 3-pentanone, 1-hexanol, benzene, 3-methylphenol and 2-hydroxy-5-methylbenzaldehyde. The calculated results of the hydrocarbon molecules are in good agreement with the data in the literature. These results confirm that the solubilities of ginger bioactive compounds can be predicted using the COSMO-RS method. The solubilities of the ginger bioactive compounds are lower than 1.0×10^{-4} at temperatures lower than 130 °C. At 130–200 °C, the solubilities increase dramatically with the highest being 6-shogaol, which is 3.7×10^{-4} mole fraction, and the lowest is 10-gingerol, which is 0.39×10^{-4} mole fraction at 200 °C.

S.Z.S. Jaapar (✉) · N.A. Morad

Malaysia-Japan International Institute of Technology (MJIT), Universiti Teknologi Malaysia (UTM), Jalan Semarak, 54100 Kuala Lumpur, Malaysia
e-mail: szsj_im@yahoo.com

N.A. Morad

e-mail: azian@ic.utm.my

S.Z.S. Jaapar · N.A. Morad

Center of Lipids and Engineering Research (CLEAR), Universiti Teknologi Malaysia (UTM), Jalan Semarak, 54100 Kuala Lumpur, Malaysia

Y. Iwai

Department of Chemical Engineering, Faculty of Engineering, Kyushu University, 744 Motooka, Nishi-Ku, Fukuoka 819-0395, Japan
e-mail: iwai@chem-eng.kyushu-u.ac.jp

1 Introduction

Recently, various alternative and green solvents, such as sub- and supercritical water or fluids have been developed in order to substitute common solvents that have a high level of toxicity. Hot compressed water or subcritical water is an alternative green solvent that has been developed and applied in the extraction process of natural and medicinal plants. This shows that the properties of water have high potential to be discovered, for example, the dielectric constant and viscosity of water. Water is the most abundant, safe, cheap and environmentally friendly pure solvent. Therefore, water seems to be one of crucial importance for further development of the chemistry of subcritical and supercritical fluids [1].

The current study, which was done by Sarip [2] for ginger using subcritical water extraction, indicates that the extracted bioactives by water depend on the temperature of the process. For example, at 135 °C, the bioactive extracted was just 6-gingerol. Whereas, when the extraction was carried out at 165 and 200 °C, more bioactives, such as 6-shogaol and 10-gingerol, were extracted. These temperature conditions are a function of the dielectric constant, which reduces as the temperature elevates [3].

There are four main components of ginger bioactive compounds, which are 6-gingerol, 6-shogaol, 8-gingerol and 10-gingerol. The identification technique of ginger bioactive compounds are well established, so these can be the reference to validate the extraction process using water as the solvent at subcritical conditions or at lower dielectric constant properties. However, physical properties of the bioactives for engineering application such as solubility and diffusivity in solvents are not published up to this date.

The equilibrium and transport properties, such as solubility and diffusion coefficient, are very useful for process design [4]. The solubility of a substance is a fundamental property in order to establish the feasibility of an extraction [5]. The water solubility or the aqueous solubility of a compound is referred to the *equilibrium concentration* or the *saturation concentration* in the aqueous phase [6]. A binary solubility diagram, a plot of T versus x_i shows the liquid–liquid equilibrium (LLE) of a binary system at constant pressure or the pressure effects are negligible [7]. One example of LLE is when the two phases of a binary mixture (liquid–liquid) are at thermodynamic equilibrium. The equilibrium criteria for LLE in a system of i and ii species at certain temperature (T) and pressure (P) in the liquid phases I and II, is written as (1):

$$f_i^I = f_i^{II} \quad (1)$$

When the activity coefficient is used, (1) become

$$x_i^I \gamma_i^I f_i^{\text{ref}} = x_i^{II} \gamma_i^{II} f_i^{\text{ref}} \quad (2)$$

x_i^I , x_i^{II} are the mole fraction of liquid phases I and II for species i, respectively, f_i^{ref} is the fugacity at the reference state. However, at this phase, small effects of

temperature on the activity coefficient, γ can have large influence on the liquid–liquid phase diagram [8]. The activity coefficient, γ is a basic physical-property value needed for the design of separation equipment [9]. The activity coefficient is derived from G^E/RT , which is the expression for the thermodynamic basis for calculation or correlation of LLE [7].

This activity coefficient was used in various thermodynamics model such as UNIQUAC (Universal Quasi-Chemical), UNIFAC (UNIQUAC Functional Group Activity Coefficient) and NRTL (Non-Random Two Liquid Model) to predict the LLE for separation process [10]. Each model has the advantages and disadvantages and has been explained detail by Assael et al. [16]. Most of these models have satisfactory correlation of activity coefficients for binary systems.

In this study, NRTL model was applied to fit the activity coefficients calculated by the COSMO-RS method for the LLE prediction. The equation for NRTL model for the excess Gibbs energy is as (3) below [8]:

$$\frac{G^E}{RT} = x_1 x_2 \left[\frac{\tau_{21} G_{21}}{x_1 + x_2 G_{21}} + \frac{\tau_{12} G_{12}}{x_2 + x_1 G_{12}} \right] \quad (3)$$

From (3), the activity coefficients are:

$$\ln \gamma_1 = x_2^2 \left[\tau_{21} \left(\frac{G_{21}}{x_1 + x_2 G_{21}} \right)^2 + \frac{\tau_{12} G_{12}}{(x_2 + x_1 G_{12})^2} \right] \quad (4)$$

$$\ln \gamma_2 = x_1^2 \left[\tau_{12} \left(\frac{G_{12}}{x_2 + x_1 G_{12}} \right)^2 + \frac{\tau_{21} G_{21}}{(x_1 + x_2 G_{21})^2} \right] \quad (5)$$

The COSMO-RS method has the capability of predicting the thermo-physical properties of liquids, such as activity coefficients from a statistical approach by quantum calculations without experimental data, based on a unimolecular quantum chemical calculation, which is an alternative to the structure-interpolating group contribution methods (GCMs) [11, 12].

In 1995, Klamt [13] had proposed a totally new approach for the calculation of solvation phenomena. It was based on good description of molecules in water and some other solvents from dielectric continuum models. Since then, COSMO-RS method has been improved by Klamt and his group which becomes a powerful tool in molecular modeling, physical chemistry and in chemical engineering.

Basically there are two steps for the COSMO-RS calculations standard method. The first step is to calculate the quantum chemical COSMO for the molecular species involved by extracting the information about the solutes and solvents. Then, the second step performs the COSMO-RS statistical calculations through the COSMOtherm program [11, 14].

The interaction parameters in NRTL equation were determined from the calculated activity coefficients by using the COSMO-RS method. The liquid–liquid equilibria (LLE) were calculated by NRTL model. Then the solubilities were obtained from LLE. The solubilities of binary mixtures for hydrocarbon + water systems were calculated and compared with the data in the literature. The calculated results are in good agreement with the experimental data. The solubilities of ginger bioactive compounds in water for the binary systems were then predicted.

2 COSMO-RS Calculation Procedures

The predictions of solubilities were carried out using the software package COSMOtherm version 3.2 (COSMOlogic GmbH & Co KG, Leverkusen, Germany) and Gaussian software package version 3.09 (GausView, Gaussian, Inc. USA). The calculations were performed on the Density Functional Theory (DFT) level using Gaussian software for each molecule to obtain the COSMO input files for water, hydrocarbon and ginger bioactive molecules. In these calculations, a charge density σ of a segment on the molecular surface was calculated in a virtual conductor. The ideally screened molecule energy with screening charge distribution on the surface was used as the database for COSMO files. The distribution function $p_i(\sigma)$ of the obtained σ (Fig. 1), which is called σ -profile, for each molecule can give the σ -profile of the mixture $p(\sigma)$ by using $p_i(\sigma)$ and the mole fraction x_i of component i in the mixture [15]. Detailed equations and the method of calculation have been described by Shimoyama et al. [15]. The σ -profiles and σ -potential for each molecule were calculated with the COSMOtherm program. The activity coefficients for the binary systems were predicted using the COSMO-RS calculation.

Fig. 1 Schematic illustration of contacting molecular cavities and contact interaction. *Source* Klamt and Eckert [12]

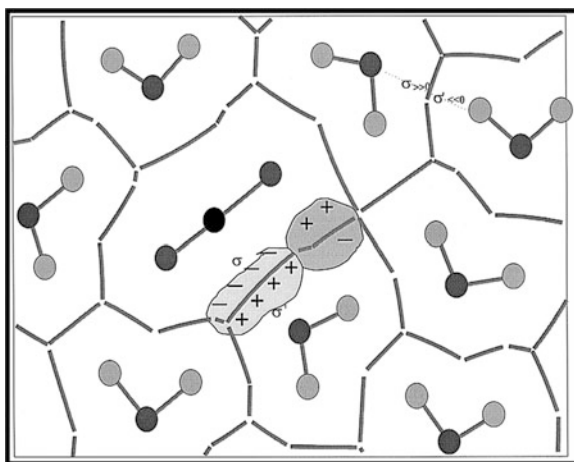
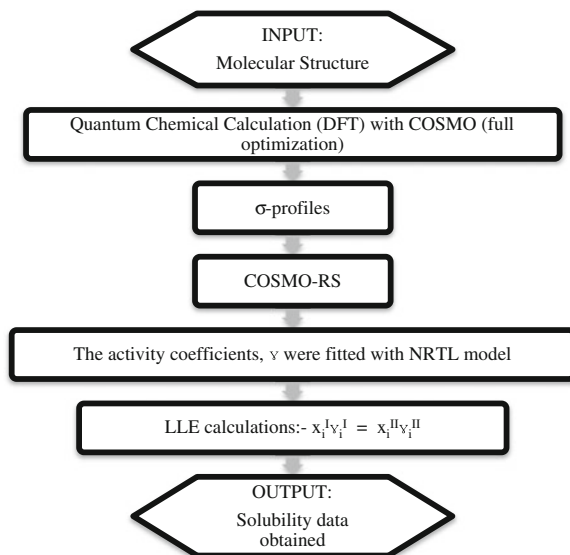


Fig. 2 Flowchart of solubility prediction from COSMO-RS



The activity coefficients were fitted with the NRTL model and the binary parameters in NRTL were obtained. As mentioned earlier, the fugacities of the both phases are the same in LLE. Referring to the (2), the equality of activities between the two liquid phases in LLE is written in (6) as follows:-

$$x_i^I \gamma_i^I = x_i^{II} \gamma_i^{II} \quad (6)$$

where γ is the activity coefficient, and superscripts I and II mean phases I and II, respectively [16].

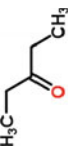
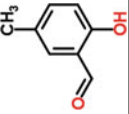
The LLE was calculated using the NRTL model. The solubilities were obtained from the LLE. The COSMO-RS calculations standard method [12] was applied in this research and modified to obtain the solubilities of the molecules from the LLE diagram. Figure 2 shows the flowchart for the calculation of solubilities.

3 Results and Discussion

3.1 Validation of Calculated Results

No experimental data for the solubilities of ginger bioactive compounds solubilities are available to our best knowledge until this date. In order to validate the results, the calculations were done part-by-part for the smaller parts of the ginger bioactive compounds. 3-Pentanone was selected as it has double O bonding, 1-hexanol has $-\text{OH}$ in the formula structure, benzene has ring structure, while 3-methylphenol and 2-hydroxy-5-methylbenzaldehyde have $-\text{OH}$ and $-\text{CH}_3$ connected to the ring structures. Table 1 shows the properties of the selected hydrocarbon molecules.

Table 1 Hydrocarbon molecules selected properties

Hydrocarbon name	3-Pentanone	1-Hexanol	Benzene	3-Methylphenol	2-Hydroxy-5- methylbenzaldehyde
Formula structure					
Formula molecule	$C_5H_{10}O$	$C_6H_{14}O$	C_6H_6	C_7H_8O	$C_8H_8O_2$
Molecular weight	86.13	102.17	78.11	108.14	136.15

These five hydrocarbon molecules were selected because they have a similar formula structure of smaller parts of ginger bioactive compounds and the experimental data are available to prove the reliability of the COSMO-RS method. Figure 3 shows the structures of the minimum (with respect to the COSMO energy) conformers of these selected hydrocarbons. The solubilities of the selected hydrocarbon molecules in five binary systems were then calculated with COSMO-RS and compared with the experimental data, which were 3-pentanone in water, 1-hexanol in water, benzene in water, 3-methylphenol in water, and 2-hydroxy-5-methylbenzaldehyde in water.

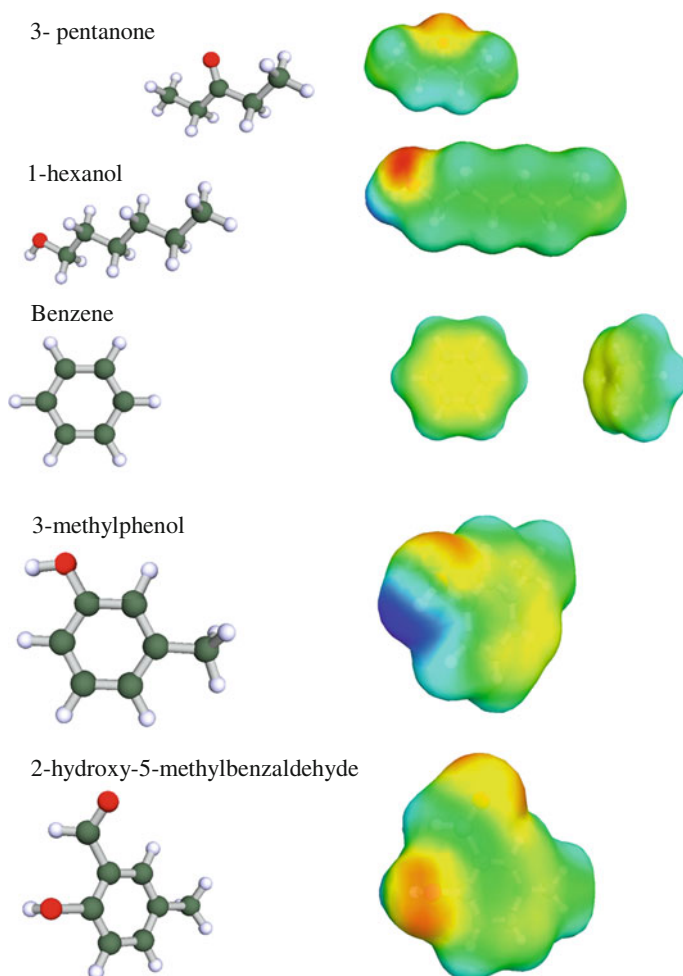


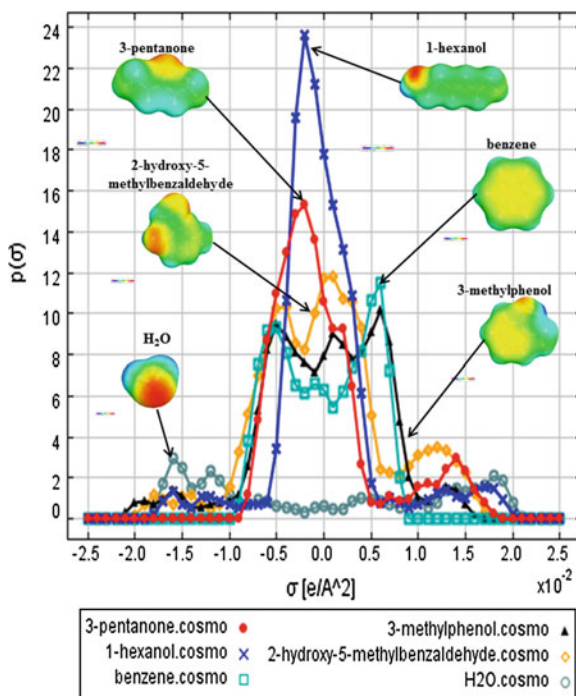
Fig. 3 Screening charge σ and structure of the minimum conformers of 3-pentanone, 1-hexanol, benzene, 3-methylphenol and 2-hydroxy-5-methylbenzaldehyde

3.1.1 σ -Profiles

Figure 4 shows the σ -profiles and COSMO-cavities of six selected compounds. The σ -profile of water is very broad, almost symmetric, with two peaks from both the negative and positive side, which are around -0.016 and $+0.018$ $e/\text{\AA}^2$, resulting from the two polar hydrogen and lone pair of oxygen atoms, respectively. This σ -profile of water was the same as reported by Zhou et al. [17] and almost the same as the results reported by Klamt and Eckert [12].

The interaction energies were determined by the screening charge density σ . The notation of “ σ -profile” is used to describe the distribution $p(\sigma)$, which describes the amount of surface in the ensemble having a screening charge density between σ and $\sigma + d\sigma$. Regarding the definition of the σ -profile, there are two important σ -moments related to the hydrogen bonds that represent the capability of hydrogen bond (HB) donor (−) and acceptor (+) [11, 12]. A molecule is polar enough if the charge density goes beyond ± 0.008 $e/\text{\AA}^2$. The σ -profile of 1-hexanol is narrowed in the middle region due to no significant electrostatic moments and non-polar character with the peak being -0.003 $e/\text{\AA}^2$, which is different to 2-hydroxyl-5-methylbenzal-dehyde, benzene and 3-methylphenyl, which have quite a symmetric σ -profile and are less broad than water. They also show no intensity in both the hydrogen bonding regions, which is beyond ± 0.01 $e/\text{\AA}^2$.

Fig. 4 σ -Profiles of selected hydrocarbons and COSMO-cavities (red positive surface screening charges resulting from negative partial charges within molecule, blue negative surface charges and green neutral charges)



3.1.2 LLE Phase Diagram for Solubilities of Hydrocarbons in Water

The calculated results from the COSMO-RS method are compared to the experimental data for the selected binary systems as shown in Figs. 5, 6, 7, 8 and 9. Figure 5 shows the comparison of the experimental data [18] and the calculated results at temperature 20–120 °C of the solubilities of 3-pentanone in water. The trend of the data and the calculated results are almost the same and really close to each other with average percentage of difference is 9 %. In Fig. 6, there are about 4 % difference between the experimental data [18, 19] and the calculated results for the solubilities of 1-hexanol in water at temperature 20–110 °C. Figure 7 shows the solubilities of benzene [18] in water. The percentage of difference is about 19.27 % at temperature 0–80 °C. While the solubilities of 3-methylphenol in water show about 22.85 % difference between the experimental data [18] and the calculated results as shown in Fig. 8. However, in Fig. 9, the percentage of difference for the solubilities of 2-hydroxy-5-methylbenzaldehyde in water is about 59.11 % which is high. The reasons maybe less of experimental data are available to compare.

The results show that the trends of experimental data and calculated results are almost the same for the binary system. The agreement between the experimental and calculated results in this study is not perfect, but it is satisfactory. In addition, a few papers reported that the calculated and experimental results are in good agreement when using COSMO-RS to predict the thermophysical properties of certain compounds [14, 15, 20, 21].

Fig. 5 Solubilities of 3-pentanone (1) in water (2)

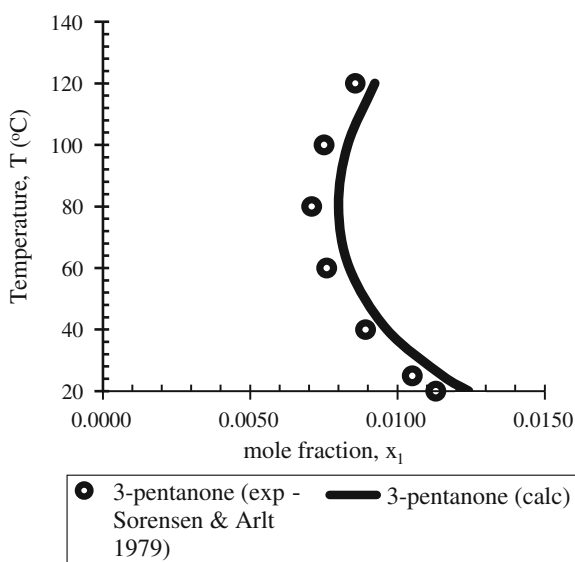
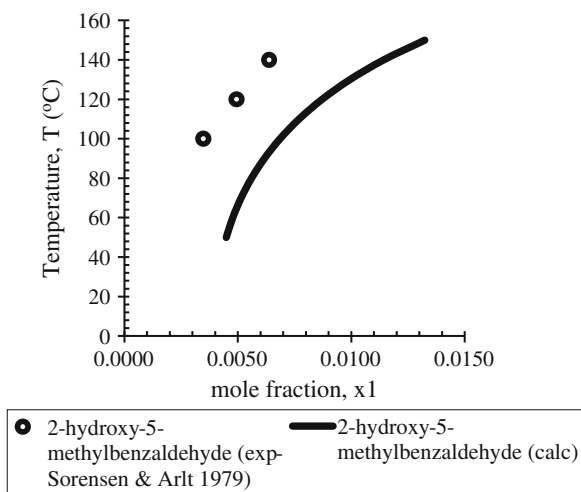


Fig. 9 Solubilities of 2-hydroxy-5-methylbenzaldehyde (1) in water (2)



3.2 Calculated Results for Solubilities of Ginger Bioactive Compounds in Water

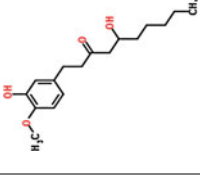
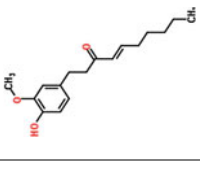
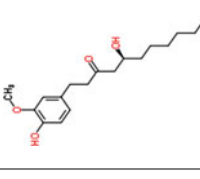
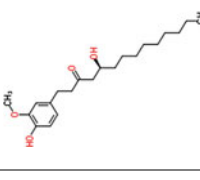
The solubilities of four main components of ginger bioactive compounds were calculated at 50–200 °C where the ginger bioactive compounds are in liquid form at this range of the temperature. Additionally, this range of temperature was selected in order to compare with the Subcritical Water Extraction (SWE) of ginger bioactive compounds, which was done by Sarip [2], and in consideration of the operational cost if higher temperatures were applied as well as the stability of the ginger bioactive itself from the degradation. Table 2 shows the formula structures and molecular weights of four main ginger bioactive compounds.

The solubilities of these ginger bioactive compounds in water were calculated separately in the binary systems which were 6-gingerol in water, 6-shogaol in water, 8-gingerol in water, and 10-gingerol in water using COSMO-RS method. Figure 10 shows the structures of the minimum (with respect to the COSMO energy) conformers of 6-gingerol, 6-shogaol, 8-gingerol and 10-gingerol. The structures for all ginger bioactive compounds are almost the same. The differences are the number of carbon and hydrogen molecules. As for 6-shogaol, the structures has a double bond between carbon molecule and less one ion O^{2-} and two ion H^+ compared to 6-gingerol.

3.2.1 σ -Profiles of Ginger Bioactive Compounds

The σ -profiles and COSMO-cavities of four ginger bioactive compounds and water are shown in Fig. 11. The σ -profile of water is used as the reference for comparison

Table 2 Ginger bioactive compounds properties

Ginger bioactive	6-Gingerol	6-Shogaol	8-Gingerol	10-Gingerol
Formula structure				
Formula molecule	$C_{17}H_{26}O_4$	$C_{17}H_{24}O_3$	$C_{19}H_{30}O_4$	$C_{21}H_{34}O_4$
Molecular weight	294.39	276.37	322.44	350.49

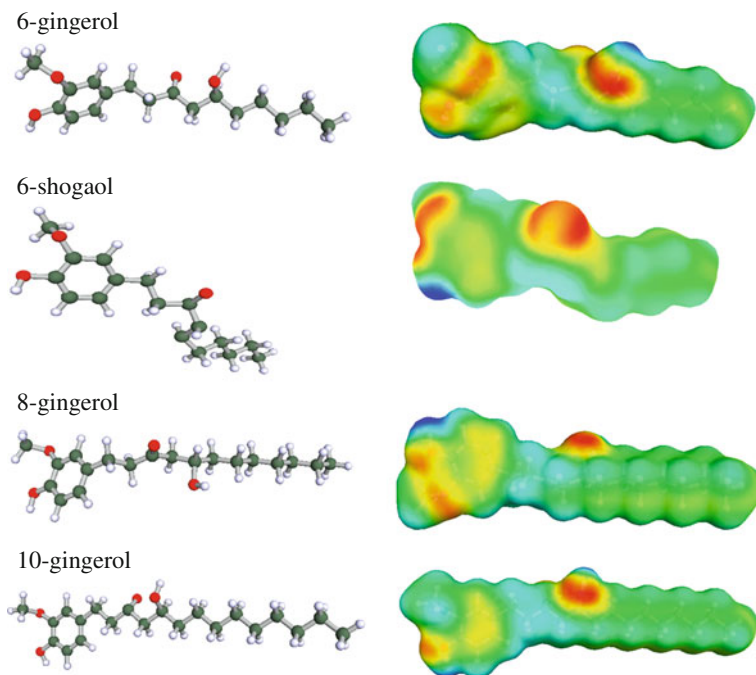


Fig. 10 Screening charge σ and structure of the minimum conformers of 6-gingerol, 6-shogaol, 8-gingerol and 10-gingerol

with the previous results. The σ -profile for each compound of ginger bioactive compound is almost the same, which is narrow in the middle region. 10-Gingerol has the highest peak among the compounds. The σ -profile of 10-gingerol is almost the same as 1-hexanol, as shown in Fig. 4. This σ -profile shows the molecular polarization or interactions [17]. They show that there is no electrostatic moment or non-polar character in the main parts of the ginger bioactive compounds. However, all the ginger bioactive compounds have a peak beyond $+0.01 \text{ e}/\text{\AA}^2$, which represent the presence of the hydrogen bond. All the ginger bioactive compounds show a higher screening charge density, which is in the region of HB acceptor region rather than the HB donor region.

3.2.2 LLE Phase Diagram for Solubilities Prediction of Ginger Bioactive Compounds in Water

The solubility for each ginger bioactive compound was calculated separately in the binary systems (6-gingerol in water, 6-shogaol in water, 8-gingerol in water and 10-gingerol in water). The calculated results of the solubilities for ginger bioactive compounds are plotted in the LLE phase diagram for comparison, as shown in

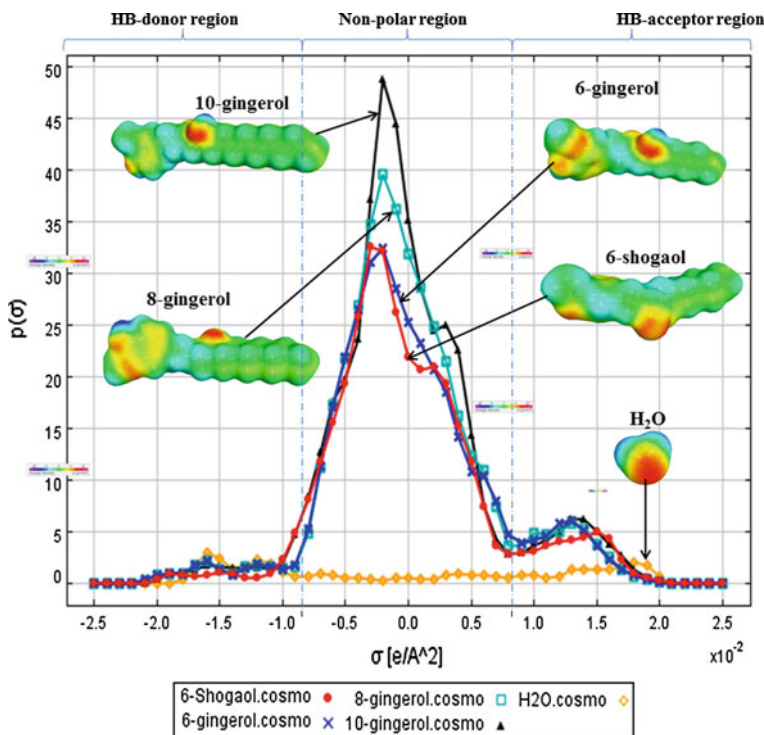
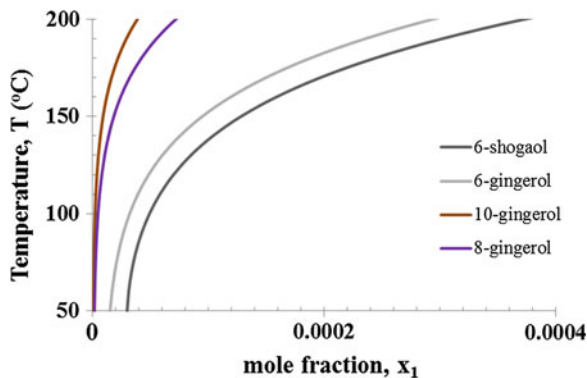


Fig. 11 σ -Profiles of ginger bioactive compounds and COSMO-cavities (*red* positive surface screening charges resulting from negative partial charges within molecule, *blue* negative surface charges and *green* neutral charges)

Fig. 12. These solubilities of ginger bioactive compounds were calculated from 50 to 200 °C. The solubilities of 6-gingerol and 6-shogaol are higher than those of 10-gingerol and 8-gingerol. This is because the molecular structures or molecular weights of 10-gingerol and 8-gingerol are bigger than those of 6-gingerol and 6-shogaol. The trends of solubilities for all ginger bioactive compounds are almost

Fig. 12 Solubilities for ginger bioactive compounds (1) (6-shogaol, 6-gingerol, 8-gingerol and 10-gingerol) in water (2)



the same. The solubilities of these ginger bioactive compounds are very low being less than 1.0×10^{-4} mole fractions when the temperature is below 130 °C. As the temperature increases from 130 to 200 °C, the solubilities of these ginger bioactive compounds increase dramatically. The highest solubility obtained is from 6-shogaol, followed by 6-gingerol, 8-gingerol and 10-gingerol, which are 3.7×10^{-4} , 2.9×10^{-4} , 0.7×10^{-4} and 0.1×10^{-4} mole fraction, respectively, at 200 °C. Ginger bioactive compounds are also sensitive to heat and will degrade as the temperature increases [2].

4 Conclusion

The solubilities of 6-gingerol, 6-shogaol, 8-gingerol and 10-gingerol were calculated with COSMO-RS in order to establish the feasibility of water as the solvent at high temperature with lower dielectric constant. The σ -profiles of 6-gingerol, 6-shogaol, 8-gingerol and 10-gingerol show that the compounds have the slight possibility of polar characteristic with the existence of hydroxide group in the structure. The molecules interaction and polarization are also shown in σ -profile. The COSMO-RS method can be used to calculate the LLE phase diagram. The solubilities of these ginger bioactive compounds increase dramatically after the temperature increase from 130 to 200 °C. This shows that at high temperature, the dielectric constant of water decreased and mimic the properties of common non-polar solvent in engineering application especially extraction. The solubilities of ginger bioactive compounds in hot compressed water can be predicted with COSMO-RS method without extensive experimental work.

Acknowledgments Our sincere gratitude goes to Mr. Aokawa in Department of Chemical Engineering, Kyushu University, Japan. He helped a lot in the COSMO-RS and Gaussian simulation. The authors gratefully acknowledge the financial support from Malaysia-Japan International Institute of Technology (MJIIT), UTM Kuala Lumpur, Malaysia for the Fellowship Program with Kyushu University, Fukuoka, Japan.

References

1. Galkin AA, Lunin VV (2005) Subcritical and supercritical water: a universal medium for chemical reactions. *Russ Chem Rev* 74(1):21–35
2. Sarip MSM (2012) Subcritical water extraction of 6-gingerol and 6-shogaol from *Zingiber officinale*. Universiti Teknologi Malaysia, Malaysia
3. Adam G, Carr RM, Neil RF (2010) Solubility and micronization of griseofulvin in subcritical water. *Ind Eng Chem Res* 49:3403–3410
4. Higashi H, Iwai Y, Arai Y (2001) Solubilities and diffusion coefficients of high boiling compounds in supercritical carbon dioxide. *Chem Eng Sci* 56(10):3027–3044
5. Carr AG, Mammucari R, Foster NR (2011) A review of subcritical water as a solvent and its utilisation for the processing of hydrophobic organic compounds. *Chem Eng J* 172(1):1–17

6. Schwarzenbach RP, Gschwend PM, Imboden DM (2005) Activity coefficient and solubility in water, in environmental organic chemistry. Wiley, New York, pp 133–180
7. Smith JM, Ness HCV, Abbott MM (1996) Introduction to chemical engineering thermodynamics, 5th edn. Chemical engineering series. The McGraw-Hill Companies, New Delhi
8. Prausnitz JM, Lichtenthaler RN, Azevedo EGD (1986) Molecular thermodynamics of fluid-phase equilibria, 2nd edn. Prentice Hall, New Jersey
9. Iwai Y, Yamamoto Y (2013) Concentration dependent surface area parameter model for calculation of activity coefficients. *Fluid Phase Equilib* 337:165–173
10. Banerjee T, Sahoo RJ, Rath SS, Kumar R, Khanna A (2007) Multicomponent liquid–liquid equilibria prediction for aromatic extraction systems using COSMO-RS. *Ind Eng Chem Res* 46(4):1292–1304
11. Klamt A (2003) Prediction of the mutual solubilities of hydrocarbons and water with COSMO-RS. *Fluid Phase Equilib* 206(1–2):223–235
12. Klamt A, Eckert F (2000) COSMO-RS: a novel and efficient method for the a priori prediction of thermophysical data of liquids. *Fluid Phase Equilib* 172(1):43–72
13. Klamt A (1995) Conductor-like screening model for real solvents: a new approach to the quantitative calculation of solvation phenomena. *J Phys Chem* 99(7):2224–2235
14. Freire MG, Santos LMNBF, Marrucho IM, Coutinho JAP (2007) Evaluation of COSMO-RS for the prediction of LLE and VLE of alcohols & ionic liquids. *Fluid Phase Equilib* 255(2):167–178
15. Shimoyama Y, Iwai Y, Takada S, Arai Y, Tsuji T, Hiaki T (2006) Prediction of phase equilibria for mixtures containing water, hydrocarbons and alcohols at high temperatures and pressures by cubic equation of state with GE type mixing rule based on COSMO-RS. *Fluid Phase Equilib* 243(1–2):183–192
16. Assael MJ, Martin TJP, Tsolakis TF (1996) An introduction to their prediction: thermophysical properties of fluids. Imperial College Press, London
17. Zhou T, Chen L, Ye Y, Chen L, Qi Z, Freund H, Sundmacher K (2012) An overview of mutual solubility of ionic liquids and water predicted by COSMO-RS. *Ind Eng Chem Res* 51(17):6256–6264
18. Sorensen JM, Arlt W (1979) Liquid–liquid equilibrium data collection: binary systems, vol V. Chemistry data series. Frankfurt, Dechema (Deutsche Gesellschaft für Chemisches Apparatewesen)
19. Goral M, Wisniewska-Gocłowska B, Maczynski A (2006) Recommended liquid–liquid equilibrium data. Part 4. 1-alkan-ol-water systems. *J Phys Chem Ref Data* 35:141–1391
20. Ferreira AR, Freire MG, Ribeiro JC, Lopes FM, Crespo JG, Coutinho JAP (2011) An overview of the liquid–liquid equilibria of (ionic liquid + hydrocarbon) binary systems and their modeling by the conductor-like screening model for real solvents. *Ind Eng Chem Res* 50(9):5279–5294
21. Freire MG, Ventura SPM, Santos LMNBF, Marrucho IM, Coutinho JAP (2008) Evaluation of COSMO-RS for the prediction of LLE and VLE of water and ionic liquids binary systems. *Fluid Phase Equilib* 268(1–2):74–84