



WAVES AND OPTICS

Harish Parthasarathy



Waves and Optics



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Waves and Optics

Harish Parthasarathy

Professor

Electronics & Communication Engineering
Netaji Subhas Institute of Technology (NSIT)
New Delhi, Delhi-110078



CRC Press

Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business



Manakin
PRESS

First published 2021
by CRC Press
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

and by CRC Press
6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742

© 2021 Manakin Press Pvt Ltd.

CRC Press is an imprint of Informa UK Limited

The right of Harish Parthasarathy to be identified as the author of this work has been asserted by him in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

For permission to photocopy or use material electronically from this work, access www.copyright.com or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. For works that are not available on CCC please contact mpkbooksp permissions@tandf.co.uk

Trademark notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Print edition not for sale in South Asia (India, Sri Lanka, Nepal, Bangladesh, Pakistan or Bhutan).

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloguing-in-Publication Data
A catalog record has been requested

ISBN: 978-0-367-75499-0 (hbk)
ISBN: 978-1-003-16273-5 (ebk)



Table of Contents

Chapter 1: The wave equation with examples from mechanics, optics, electromagnetism and quantum mechanics	1-64
1.1 The Definition of a propagating wave in one, two and three dimensions	1
1.2 Standing waves in one, two and three dimensions	7
1.3 The polarization of a wave	11
1.4 The wave equation in one, two and three dimensions	14
1.5 The polarization of a wave revisited	17
1.6 Basics of fluid dynamics	19
1.7 Waves in a fluid—derivation from first principles	22
1.8 Longitudinal sound/pressure waves in a tube	27
1.9 The difference between transverse and longitudinal waves in terms of wave polarization	28
1.10 Maxwell's equations and the wave equation for the electric and magnetic fields in free space	30
1.11 Solution to Maxwell's equations in terms of retarded potentials satisfying the wave equation with source	32
1.12 The principle of superposition	34
1.13 Diffraction and interference of waves	37
1.14 Green's function for wave equation with sources—Fraunhofer and Fresnel's diffraction	39
1.15 The basic Eikonal equation of geometric optics	42
1.16 Describing the trajectory of light in a medium having spatially varying refractive index	46
1.17 Propagation of light in anisotropic, inhomogeneous and time varying medium	47
1.18 The Schrodinger wave equation in quantum mechanics	48
1.19 The effect of noise on the Schrodinger wave equation—Open systems, ie, coupling of the system to the bath environment	52
1.20 Wave equation with random non-uniform refractive index	53
1.21 The relationship between the wave equation and the Helmholtz equation for waves of given frequency	56
1.22 Waves in a confined region	57
1.23 Schrodinger's wave equation for mixed states in the position kernel domain	62

Chapter 2: Waves in general relativity, quantum gravity, plasma physics and quantum stochastics	65-130
2.1 Gravitational waves	65
2.2 Quantum gravity, the canonical ADM formalism-Schrodinger's equation for the wave function of the space-time metric.	75
2.3 Plasma waves	86
2.4 Evans-Hudson diffusion as a quantum mechanical generalization of the wave equation with noise	93
2.5 Waves in an expanding universe–Newtonian theory of small fluctuations and general relativistic theory of small fluctuations	95
2.6 EMwaves in a curved space-time geometry with inhomogeneous permittivitypermeability tensor	101
2.7 Quantum Optics. Here, the photon field is a quantum electromagnetic field expressible as a superposition of annihilation and creation operators of the photon field with the coefficients of the linear combination being positions of time and space	103
2.8 Quantum optics, notion of a generalized measurement, state collapse after quantum measurement, recovery of states passed through a noisy quantum system, the Knill-Laflamme theorem Stinspring's representation of noisy quantum systems, Information, relative entropy, mutual information and Renyi entropy of quantum systems. Transmission of information over quantum system. The relevance of all this to the wave mechanics of Schrodinger	112
2.9 Controlling the quantum em field produced by electrons and positrons by using a classical em field-An application of Dirac's relativistic wave equation	116
2.10 Calculating the path of a light ray in a static gravitational field	118
2.11 A study of thermal emission by blackholes via Hawking radiation, quantum mechanics of fields in the vicinity of a blackhole and the interactio of electrons,positrons, photons and gravitons with an external noisy bath with application to the design of very large size quantum gates	120

Chapter 3: Analysis of waves in engineering and optical systems, in biological systems, classical and in quantum blackhole physics	131-260
3.1 Wave digital filter design	131
3.2 Large deviation principle in wave-motion	134
3.3 Some more problems in Schrodinger-wave mechanics and Heisenbergmatrix mechanics with relevance to quantum information theory	137
3.4 Questions in optimization techniques	145
3.5 Quantum antennas via the Schrodinger wave equation	148
3.6 Linear algebra for quantum information theory	151
3.7 Transmission lines and waveguides–Questions	153
3.8 Some more matrix inequalities related to quantum information theory	157
3.9 Fresnel and Fraunhoffer diffraction	159
3.10 Surface tension and wave propagation	161
3.11 Klein-Gordon equation in the Schwarzschild space-time with a radialtime independent electromagnetic field and its application to computing the Hawking temperature at which massless/massive particles are emitted from a blackhole	162
3.12 Quantum Belavkin filtering versus classical Kushner-Kallianpur filtering–A comparison	165
3.13 Remark on quantum Belavkin filtering for estimating the state of a quantum vibrating string	179
3.14 Elementary problems in robotics based on damped simple harmonic motion	182
3.15 Approximate solution to the Dirac equation in curved space-time	194
3.16 Some applications of quantum gate design using physical systems	197
3.17 Convergence of perturbation series for nonlinear differential equations	200
3.18 Poiseulle's law and generalized Poiseulle's law for flow through a pipe	204
3.19 Measurement of refractive index	206
3.20 Modes of a vibrating string with applications to particle physics	209

3.21	Hidden Markov Models for estimating the amplitude, frequency and phase of a sinusoidal signal making transitions	212
3.22	The energy-momentum tensor of the Dirac field in a background curved space-time metric	217
3.23	Remark on Noether's theorem on conserved currents	220
3.24	Energy-momentum tensor using the tetrad formalism	222
3.25	Analysis of gravitational waves produced by a finite system of point particles–A perturbation theoretic approach	226
3.26	Heat equation and its solution in R^n , relationship between heat and wave equations, nonlinear heat equations arising as the scaling limit of the simple exclusion process	228
3.27	Study of wave motion of the boundary of single cellular micro-organisms by giving them external stimulus and observing the wave like motion of their boundary walls as well as wave-like fluctuations of the velocity field of the cytoplasmic fluid within them	233
3.28	Snell's laws of reflection and refraction on surfaces separating two uniform media	239
3.29	Spinor form of some equations of mathematical physics: Roger Penrose's theory	244
3.30	Prisms, mirrors and lenses, the general theory	253
3.31	A brief summary of the book	259

Chapter 4:	Probability Theory and Statistics required for random wave motion analysis	261-272
4.1	Summary of contents	261
4.2	Probability spaces and measure theoretic theorems on probability spaces	262
4.3	Basic facts about quantum probability	263
4.4	Some basic classical and quantum stochastic processes	264
4.5	Some applications of classical probability to engineering systems	265
4.6	Quantum stochastic differential equations	267
4.7	Some practical applications of quantum probability	268
4.8	Casting the HP equation in functional derivative form	271

Chapter 5: An introduction to probability and random processes in circuit and field theory from a pedagogical viewpoint	273-356
5.1 Circuit theory concepts from field theory concepts	273
5.2 Graph theoretic analysis of large linear circuits based on KCL and KVL	276
5.3 Two port network theory	276
5.4 Diode and capacitance circuit models	277
5.5 Classical device physics	277
5.6 Device physics using quantum mechanics and quantum electrodynamics	278
5.7 Band theory of a semiconductor by solving Schrodinger's equation	279
5.8 Quantum electrodynamics and quantum field theory	279
5.9 Analyzing random Gaussian and non-Gaussian noise in circuits using higher order correlations and spectra	280
5.10 Noise in nonlinear transistor circuits	282
5.11 Digital electronics	282
5.12 Techniques for analyzing transmission lines	283
5.13 Brownian motion, Poisson processes and stochastic differential equations in circuit theory	284
5.14 Classical and quantum random processes in circuit theory	285
5.15 Simulation of nonlinear ode's and pde's in circuit theory and electromagnetics	285
5.16 Derivation of medium properties from basic physical principles involving motion of individual electrons and magnetic moments in external fields	286
5.17 Partial differential equation methods for analyzing waveguides	287
5.18 Curvilinear coordinate systems and variational methods in engineering electromagnetics	288
5.19 Perturbation theoretic methods in solving electromagnetics problems	290
5.20 Numerical methods in antenna, waveguide and cavity resonator theory	293
5.21 Large deviation theory applied to engineering systems	294
5.22 Robotics based on nonlinear differential equations	295
5.23 Quantization of robot motion	296

5.24	Filtering and control of engineering systems	298
5.25	Quantum many body systems applied to Fermi operator fields and superconductivity	302
5.26	Lie group theory in image processing	307
5.27	Lie group based robotics	308
5.28	Levy process models for jerk noise in robotic systems	310
5.29	Digital systems, classical and quantum gates, design of counters using shift registers and flip-flops	312
5.30	HMM and some of its applications	314
5.31	Quantum Image Processing	314
5.32	Introduce the design aspects of some gadgets through mini-projects some examples	317
5.33	A simple way to introduce quantum electrodynamics	318
5.34	How to teach the theory of non-Abelian gauge theories as non-commutative generalizations of electromagnetism	323
5.35	How to introduce astronomy and cosmology to undergraduates	323
5.36	Quantum image processing revisited	326
5.37	The EKF for arbitrary Markov processes with Gaussian measurementnoise	331
5.38	Quantum scattering theory applied to quantum gate design	333
5.39	Superconductivity for two species and interpretation of the gap function	335
5.40	Introductory quantum information theory	337
5.41	Quantum image processing	340
5.42	Lie group-Lie algebra approach to robot dynamics with two $3 - D$ links, each described by three Euler angles and an overall translational vector $a(t) \in R^3$	347
5.43	Linear algebra and operator theory	350

Chapter 6:	Applications of Lie groups and Lie algebras, filtering, field quantization, Numerical methods for quantum mechanical problems	357-509
6.1	Quantum random walk	357
6.2	Lie group-Lie algebra theoretic coordinate free formulation of the equations of motion of a robot with N 3-D links with the orientation of each link described by an arbitrary element of $SO(3)$ and taking in addition into account a translation of the base pivot of the first link	359

6.3	Numerical methods for computing transition probabilities for photons, gravitons, Klein-Gordon Bosons, Dirac Fermions and non-Abelian matter and gauge particles from inside the critical radius to outside of a Schwarzschild blackhole with quantum gate design applications	360
6.4	Numerical methods for designing quantum gates based on quantum scattering theory for a Schrodinger projectile interacting with a potential	361
6.5	Quantization of a robot in the Lie-group domain when the robot has N 3-D links	362
6.6	Lie group formulation of the single 3-D robot link in the presence of gravitation and external torque	364
6.7	Quantum antennas based on non-Abelian matter and gauge fields	367
6.8	The electroweak theory	371
6.9	Wavelet based system parameter estimation	373
6.10	Applying the EKF to nonlinear circuits involving diodes and transistors	374
6.11	An introduction to classical and quantum error detecting and correcting codes	376
6.12	Orthogonal Latin squares and coding theory	377
6.13	Cyclic codes	378
6.14	Yang-Mills field quantization methods	379
6.15	The Ginzburg-Landau model for superconductivity	383
6.16	Teaching the basics of classical mechanics to school students and first year undergraduates	386
6.17	Teaching Linear algebra and functional analysis to post-graduate students of signal processing	396
6.18	Variants of the Kalman filter	398
6.19	The Cq-coding theorem: Proof based on Quantum Renyi entropy and Shannon's random coding argument	401
6.20	Manual for the Digital Signal Processing and Statistical Signal Processing Laboratory	402
6.21	MATLAB problems on root space decomposition of a Lie algebra	420
6.22	Cartan's criterion for semisimplicity of a Lie algebra	426
6.23	Problems in linear algebra	427
6.24	Problems in non-linear filtering theory	429

6.25	Spectral theorem for bounded self-adjoint operators in a Hilbert space–basic steps	435
6.26	Motion of rigid bodies in electromagnetic fields	439
6.27	Large deviation theory with engineering applications	442
6.28	Lectures in linear algebra for signal processing applications	442
6.29	On the improvement of the signal quality in telephone lines	443
6.30	Quantum Coulomb scattering	445
6.31	Tutorial problems in electromagnetic field theory	450
6.32	Lecture plan for electromagnetic field theory EC-C09	455
6.33	A digression into infinite dimensional vector spaces	457
6.34	Continuation of finite dimensional vector spaces	458
6.35	The general relativistic Maxwell equations in a resonator	459
6.36	Mackey’s theory on the construction of the basic observables in the quantum theory from projective unitary representations of the Galilean group	463
6.37	Hamiltonian density of the electromagnetic field in curved space-time in terms of position and momentum fields	464
6.38	Coulomb scattering	467
6.39	Electromagnetic waves in the Schwarzschild metric	469
6.40	Tutorial problems in electromagnetic field theory	470
6.41	The gravitational n-body problem in general relativity: an approximate treatment	472
6.42	Lecture plan for "Linear Algebra in Signal Processing" -SP-C01	475
6.43	Continuation of finite dimensional vector spaces	476
6.44	Cartan’s equations of structure	477
6.45	Proof of the Riesz representation theorem	479
6.46	Quantum image processing	480
6.47	Inclusion of the Goldstone boson field in the gauge field after symmetry breaking	481
6.48	Some problems in linear algebra	482
6.49	Gravitational N-body problem in general relativity	484
6.50	Multipole radiation fields in the Maxwell theory	488
6.51	How Dirac brackets are used to take care of constraints in Lagrangian and Hamiltonian mechanics	490
6.52	Deep learning of speech models	490
6.53	Some problems on linearization (for the course Linear algebra in signal processing)	491

6.54	Research project proposal for simulating quantum gates of large sizes using quantized ridge waveguide electromagnetic field interacting with quantum dots and also for estimating medium properties on the Angstrom scale	493
6.55	Design of a differentiator using series connection of short circuited transmission line elements	497
6.56	Design of quantum gates by interaction of a quantum em field with gravity	499
6.57	Scattering theory in the interaction picture for time dependent interations	500
6.58	Chapterwise report on Jaspal Khinda's Ph.D thesis	501
6.59	Training a DNN with stochastic inputs with analysis of the robustness against input process and weight matrix fluctuations	504
6.60	Quantum Boltzmann equation	505
6.61	List of Ph.D scholars supervised by Harish Parthasarathy with a brief summary of their theses	506
6.62	The problem of determining the surface current density induced on an antenna surface placed in a nonlinear inhomogeneous and anisotropic medium taking gravitational effects into account	508



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Chapter 1

The wave equation with examples from mechanics, optics, electromagnetism and quantum mechanics

1.1 The Definition of a propagating wave in one, two and three dimensions

1.1.1 Summary

A wave propagating in one dimension is a physical quantity (either a vector or a scalar or even a tensor) having amplitude and phase varying with time and the spatial dimension in such a way that the quantity at the spatial point x at time t equals the same at the origin at time $t - x/c$ where c is the velocity of wave propagation. This is so if the wave propagates along the positive x direction. On the other hand, if it propagates along the negative x direction, then the wave amplitude at (t, x) equals that at $(t + x/c, 0)$. Such wave signals satisfies a one dimensional wave equation and conversely, if a signal defined as a function of spatial location and time satisfies the one dimensional wave equation, then its general solution is a superposition of forward and backward propagating signals of the type mentioned. More generally, a three dimensional wave propagating along the direction of the unit vector $\hat{n} = (n_1\hat{x} + n_2\hat{y} + n_3\hat{z})$ satisfies the property that the wave amplitude at (t, r) equals the wave amplitude at $(t - \hat{n}.r/c, 0)$. In other words, the time taken for the wave to propagate from the origin to a point r in 3-D space equals the projection of the vector joining the origin with that point along the direction \hat{n} divided by the velocity of the wave. Such a wave signal automatically satisfies the three dimensional wave equation and conversely, every solution of the three dimensional wave equation is expressible

as a weighted superposition of such wave signals over different directions \hat{n} . If the wave signal at the origin is a sinusoid of a definite frequency with a phase, ie of the form $A.\cos(\omega t + \phi)$, and if it propagates along the direction \hat{n} , then its amplitude at (t, r) is given by $A.\cos(\omega(t - \hat{n}.r/c) + \phi)$ which means that its phase given by

$$\psi(t, r) = \omega(t - \hat{n}.r/c) + \phi$$

is a constant on the surface

$$\psi(t, r) = \psi_0 \quad \dots \quad (1)$$

At time $t + \tau$, this surface moves to the surface

$$\psi(t + \tau, r) = \psi_0$$

which is the same as

$$\psi(t, r) = \psi_0 - \omega\tau \quad \dots \quad (2)$$

If r_1 falls on the surface (1) and r_2 on the surface (2), then obviously, we obtain on taking the difference of these two equations,

$$\hat{n}.(r_2 - r_1) - c\tau = 0$$

This equation shows that the surfaces of constant phase, ie the wave-fronts travel along the direction \hat{n} with a velocity of c . Now suppose, we consider a non-linear wave, ie, a wave whose phase is not a linear function of (t, r) . Let $\psi(t, r)$ denote is phase. Then the constant phase surfaces or wave fronts are the surfaces described at time t by

$$\psi(t, r) = \psi_0$$

At time $t + dt$, the points r falling on the same constant phase surface satisfy

$$\psi(t + dt, r) = \psi_0$$

If r falls on the first surface and $r + dr$ on the second, then we obtain by taking the difference

$$\psi(t + dt, r + dr) - \psi(t, r) = 0$$

or equivalently,

$$\psi_{,t}(t, r)dt + dr.\nabla_r\psi(t, r) = 0$$

so that the instantaneous wave velocity along the direction \hat{n} at time t is given by

$$c(t, r, \hat{n}) = -(\hat{n}.\nabla_r\psi(t, r))^{-1}\psi_{,t}(t, r)$$

This is called the phase velocity of the wave as opposed to the group velocity which is defined as follows: Let $\omega(k)$ denote the frequency of the wave as a function of the wave vector k . Then the constant phase surfaces are

$$\omega(k)t - k.r = \psi_0$$

which according to the above, gives the phase velocity along the direction \hat{n} as

$$c_p(r, \hat{n}) = (k \cdot n)^{-1} \omega(k)$$

The group velocity on the other hand is computed as follows: Consider two waves having wave vectors k and $k + dk$. The superposition of these two waves is the signal field

$$\begin{aligned} & \cos(\omega(k)t - k \cdot r) + \cos(\omega(k + dk)t - (k + dk) \cdot r) \\ &= -2 \sin((\omega(k) + \omega(k + dk))t/2 - (k + dk/2) \cdot r) \sin((\omega(k + dk) - \omega(k))t/2 - dk \cdot r/2) \\ &= -2 \sin(dk \cdot \nabla_k \omega(k)t/2 - dk \cdot r/2) \cdot \sin(\omega(k)t - k \cdot r) \end{aligned}$$

provided that we neglect quadratic and higher order powers in dk . This superposition has a slowly varying envelope whose space-time dependence is $\sin(dk \cdot \nabla_k \omega(k)t/2 - dk \cdot r/2)$. The velocity of this envelope wave is evidently the vector $\nabla_k \omega(k)$ and this is called the group velocity. It is the velocity of the envelope of a superposition of several waves having wave vector lying within a small differential interval.

If $f(t)$ is the signal at $x = 0$, then since it takes time x/c for the signal to propagate from the origin to x , the signal at x is

$$u_f(t, x) = f(t - x/c)$$

This describes a forward propagating wave. Likewise a backward propagating wave is described by

$$u_b(t, x) = g(t + x/c)$$

Here, c is the velocity of the wave. Superposing the two waves, we get the general form of a wave propagating with a velocity of c :

$$u(t, x) = u_f(t, x) + u_b(t, x) = f(t - x/c) + g(t + x/c)$$

An easy calculation shows that if f, g are twice differentiable functions, then u satisfies the one dimensional wave equation:

$$u_{,tt}(t, x) - c^2 u_{,xx}(t, x) = 0$$

It is not hard to show that the above form of u is the general solution to the wave equation. It should be noted that the solution to the one dimensional wave equation is determined uniquely by two functions f, g which are in turn uniquely determined by the initial conditions $u(0, x) = f(-x/c) + f(x/c)$, $u_{,t}(0, x) = (-1/c)f'(-x/c) + (1/c)f'(x/c)$ upto an arbitrary constant.

1.1.2 Exercises

1. If a signal $A \sin(\omega t)$ propagates radially outwards from the origin, then determine the wave amplitude at (t, r) . Specifically show that

$$\psi(t, r) = A \sin(\omega(t - |r|/c)/|r|)$$

satisfies the three dimensional wave equation.

2. Repeat the previous problem by assuming a signal $f(t)$ at the origin. Show that

$$\psi(t, r) = f(t - |r|/c)/|r|$$

satisfies the three dimensional wave equation.

3. Solve the one dimensional wave equation

$$u_{,tt}(t, x) - c^2 u_{,xx}(t, x) = 0$$

with boundary conditions

$$u(t, 0) = f(t), u(t, L) = g(t)$$

You may use separation of variables. If $f(t), g(t)$ are random processes, then calculate the wave correlations $\mathbb{E}(u(t_1, x_1).u(t_2, x_2))$ in terms of the auto and cross correlations of $f(\cdot)$ and $g(\cdot)$.

4. Solve the two dimensional equation for a vibrating membrane of circular shape with its circular boundary fixed and when the membrane is subject to forcing, like a Tabla or a Mridangam. When the force applied per unit area is a random function $f(t, \rho, \phi)$, then calculate the statistical correlations in the amplitude distribution of the membrane $\langle u(t, \rho, \phi).u(t', \rho', \phi') \rangle$

5. Write down the Lagrangian density of the above vibrating circular membrane subject to classical random forces, then derive the Hamiltonian density and quantize the motion by introducing canonical position and momentum coordinates. Note that the Hamiltonian of the oscillator can be represented as a countable sum of independent harmonic oscillators each having a random forcing.

1.1.3 Advanced Exercises

[1] Define the frequency in cycles per second, the wavelength, the phase velocity, and the group velocity for the components of the wave

$$\psi(t, r) = \int A(k) \exp(-i(\omega(k)t - k.r)) d^3k$$

Show that ψ satisfies the wave equation

$$i\partial_t \psi(t, r) = \omega(-i\nabla) \psi(t, r)$$

and

$$\partial_t^2 \psi(t, r) + \omega(-i\nabla)^2 \psi(t, r) = 0$$

[2] Using the Lorentz transformation, determine the frequency and wave-vector of a wave in a reference frame K' moving w.r.t. the frame K with a uniform velocity of v along the x axis when the wave field in the frame K is given by

$$\psi(t, r) = A \exp(-i(\omega t - k \cdot r))$$

Hence deduce the Doppler effect, namely that if a source of waves moves towards the observer, the observer will record a higher frequency than that emitted by the source and if the source moves away from the observer, the observer will record a lower frequency.

[3] Solve the 3-D wave equation assuming radial symmetry and show that the general causal solution is of the form $s(t - |r|/c)/|r|$.

[4] If we consider the Maxwell equations in free space, show that a solution corresponding to the equations

$$\operatorname{div} E = 0, \operatorname{div} H = 0, \operatorname{curl} E = -j\omega\mu H, \operatorname{curl} H = j\omega\epsilon E$$

which implies

$$\operatorname{div} E = 0, (\nabla^2 + k^2)E = 0, \operatorname{div} H = 0, (\nabla^2 + k^2)H = 0, \operatorname{curl} E = -j\omega\mu H, \operatorname{curl} H = j\omega\epsilon E$$

which in turn implies

$$\operatorname{div} E = 0, \operatorname{div} H = 0, (\nabla^2 + k^2)(r \cdot E) = 0, (\nabla^2 + k^2)(r \cdot H) = 0,$$

can be taken as

$$E(\omega, r) = \sum_{l,m} [c(l, m)f_l(r)\mathbf{L}Y_{lm}(\hat{r}) + (d(l, m)/j\omega\epsilon)\operatorname{curl}(g_l(r)\mathbf{L}Y_{lm}(\hat{r}))],$$

$$H(\omega, r) = \sum_{l,m} [d(l, m)g_l(r)\mathbf{L}Y_{lm}(\hat{r}) - (c(l, m)/j\omega\mu)\operatorname{curl}(f_l(r)\mathbf{L}Y_{lm}(\hat{r}))],$$

where f_l, g_l satisfy the ode

$$f''(r) + (2/r)f'(r) - (l(l+1)/r^2)f(r) + k^2f(r) = 0$$

By making an appropriate transformation of the the independent variable r , show that the solution to this ode can be expressed in terms of the Bessel function $J_{l+1/2}(r)$. Show that

$$\hat{r} \cdot \mathbf{L} = 0$$

where

$$\mathbf{L} = -i\mathbf{r} \times \nabla$$

is the usual angular momentum vector operator in quantum mechanics. What is the physical interpretation of the above solution in terms of wave components polarized perpendicular and parallel to the radial direction of propagation. Also deduce orthogonality relations for the vector valued functions $\mathbf{LY}_{lm}(\hat{r})$, $\text{curl}(f_l(r)\mathbf{LY}_{lm}(\hat{r}))$ on S^2 . Hence explain how the coefficients $c(l, m)$, $d(l, m)$ can be obtained from electric and magnetic field measurements.

Hint: See J.D.Jackson, "Classical Electrodynamics", Wiley.

1.1.4 Points to remember

[1] A forward propagating wave along the x direction is defined by a function of $t - x/c$ where c is the wave velocity while a backward propagating wave is defined by a function of $t + x/c$. Both of these satisfy the wave equation

$$u_{,tt}(t, x) - c^2 u_{,xx}(t, x) = 0$$

and in fact the general solution of this one dimensional wave equation is given by the sum of two such components.

[2] A wave propagating along the direction \hat{n} has its amplitude given by

$$\psi(t, r) = f(t - \hat{n} \cdot r/c)$$

This satisfies the 3-D wave equation

$$\psi_{,t} - c^2 \nabla^2 \psi = 0$$

If $f(t) = A \cos(\omega t + \phi)$ in this expression, then the constant phase surfaces at time t are $t - \hat{n} \cdot r/c = \text{constt.}$ and these surfaces propagate with a velocity c along the direction \hat{n} . The general solution to the 3-D wave equation is a superposition of such solutions along different directions \hat{n} , ie,

$$\psi(t, r) = \int f(\hat{n}, t - \hat{n} \cdot r/c) d\Omega(\hat{n})$$

1.2 Standing waves in one, two and three dimensions

1.2.1 Summary

For sinusoidal waves, ie, monochromatic waves, ie, waves of a definite frequency, an appropriate linear combination of a forward travelling and a backward travelling wave field can result in a field which has nodes at definite spatial points which do not vary with time. Equivalently, for such a superposition, the amplitude as a function of time factorizes into the product of a function of time only and a function of space only. Thus, the amplitude maxima and minima occur at fixed spatial points and at any spatial point, the signal is a sinusoid whose amplitude depends only on the spatial location. Such a superposition takes place when a forward travelling wave meets a barrier and gets reflected as for example in a vibrating string with fixed end points with reflections occurring at the end points. Another example, is a transmission line where a forward propagating voltage or current signal meets a load and gets reflected resulting in a standing wave pattern.

In the previous section, we saw that the general solution to the one dimensional wave equation is given by

$$u(t, x) = f(t - x/c) + g(t + x/c)$$

If we put the requirement that at a set $x = n\Delta, n \in \mathbb{Z}$ of equispaced points on the x axis, the velocity of the wave be zero, then we get

$$-f'(t - n\Delta/c) + g'(t + n\Delta/c) = 0$$

or writing $T = \Delta/c$, this condition becomes

$$f'(t - nT) = g'(t + nT), n \in \mathbb{Z}, t \in \mathbb{R}$$

This equation can be satisfied, for example by requiring that

$$f(t - nT) = g(t + nT), n \in \mathbb{Z}$$

or equivalently,

$$g(t) = f(t + 2nT), n \in \mathbb{Z}, t \in \mathbb{R}$$

and this equation can be satisfied by taking

$$f(t) = \sin(k\pi t/T), g(t) = f(t)$$

for some $k \in \mathbb{Z}$. Corresponding to this choice of f, g , the solution to the wave equation reads

$$\begin{aligned} u(t, x) &= \sin(k\pi(t/T - x/cT)) + \sin(k\pi(t/T + x/cT)) \\ &= 2\sin(k\pi t/T)\cos(k\pi x/cT) \end{aligned}$$

For each integer k we get a solution and a superposition of such solutions for different k 's gives us in a certain sense, a general standing wave

1.2.2 Summary

The polarization of a wave is defined only for vector valued wave fields. It can vary from point to point and from time to time. At each point (t, r) in space-time, the wave field $\psi(t, r)$ has a direction and that direction determines the polarization of the wave. If at the point r in space, the wave field amplitude vector oscillates along a given direction, we say that it is linearly polarized. If however, this vector traces out a circle counterclockwise with time, we say that it is left circularly polarized and if it traces out a circle in the clockwise direction, we say that it is right circularly polarized. If the polarization of the field at a point is along a definite direction for all time, ie, it is linearly polarized, then a small wire antenna at that point will collect maximum signal power only if it is oriented along that direction. If the wire antenna is oriented along some other direction, then it will collect a power proportional to $\cos^2(\theta)$ where θ is the direction between the polarization direction of the wave field and the direction of the wire antenna. If the direction of the wave field at a given point in space traces out an ellipse, then it is said to be elliptically polarized at that point. An elliptically polarized wave can be expressed as a vector superposition of two linearly polarized waves at a given point in space and conversely, a linearly polarized wave can be expressed as a vector superposition of two circularly polarized waves. These fact are seen from the following example:

$$A_x \cdot \cos(\omega t) \hat{x} + A_y \cdot \sin(\omega t) \hat{y}$$

is elliptically polarized and is clearly the sum of two signals linearly polarized along the x and y directions. More generally,

$$A_x \cdot \cos(\omega t + \phi_x) \hat{x} + A_y \cdot \cos(\omega t + \phi_y) \hat{y}$$

is elliptically polarized and is the sum of two linearly polarized signals. Conversely,

$$A \cdot \cos(\omega t + \phi) (n_1 \hat{x} + n_2 \hat{y})$$

is linearly polarized along the direction $n_1 \hat{x} + n_2 \hat{y}$ and can be expressed as

$$\begin{aligned} & ((A/2) \cdot \cos(\omega t + \phi) \hat{n} + (A/2) \cdot \sin(\omega t + \phi) \hat{m}) \\ & + ((A/2) \cdot \cos(\omega t + \phi) \hat{n} - (A/2) \cdot \sin(\omega t + \phi) \hat{m}) \end{aligned}$$

where

$$\hat{n} = n_1 \hat{x} + n_2 \hat{y}, \hat{m} = -n_2 \hat{x} + n_1 \hat{y}$$

demonstrating the fact that a linearly polarized wave is a superposition of two circularly polarized waves at a given point in space. A wave field is said to be transversely polarized if at every point in space, its vectorial direction of oscillation is perpendicular to its direction of propagation and it is said to be longitudinally polarized if its vectorial direction of oscillation is parallel to its direction of oscillation. Examples of transversely polarized waves are electromagnetic waves in free space. This is a consequence of the relation $\operatorname{div} \psi(t, r) = 0$ appearing as Maxwell equations for the electric and magnetic fields. Indeed taking

$\psi(t, r) = \psi(t - \hat{n} \cdot r/c, 0)$ and using $\operatorname{div}\psi = 0$ gives us immediately $\hat{n} \cdot \psi(t, r) = 0$, namely transverse polarization. On the other hand, an equation of the form $\nabla \times \psi(t, r) = 0$ gives us $\hat{n} \times \psi(t, r) = 0$, namely longitudinal polarization. Pressure/sound waves in a tube are longitudinally polarized waves as also are waves in a long spring oscillating along its direction. We can have a wave that is neither transversely polarized nor longitudinally polarized as for example the magnetic field waves in a waveguide supporting a TE mode and the electric field waves in a waveguide supporting a TM mode.

1.2.3 Exercises

[1] Show that the three dimensional wave

$$\psi(t, r) = \sin(\omega t - k \cdot r) + \sin(\omega t + k \cdot r)$$

is a standing wave. Determine the planes where the signal amplitude vanishes. Repeat for

$$\psi(t, r) = \sin(\omega t - k \cdot r) + \sin(\omega t + k \cdot r + \phi)$$

[2] Determine the general form of a standing wave in two dimensions using plane polar coordinates.

[3] Determine the general form of a standing wave in three dimensions using spherical polar coordinates

hints: Use separation of variables in the 2-D wave equation

$$\frac{\partial^2}{\partial t^2} - c^2 \left(\frac{1}{\rho} (\partial/\partial\rho) \rho \partial/\partial\rho + \frac{1}{\rho^2} \partial^2/\partial\phi^2 \right) \psi(t, \rho, \phi) = 0$$

Now assume

$$\psi(t, \rho, \phi) = T(t)R(\rho)\Phi(\phi)$$

and determine the functions $R(\rho)$, $\Phi(\phi)$ and their zeros.

[4] Determine the general form of a standing wave in three dimensions using cylindrical coordinates by solving the 3-D wave equation

$$\frac{\partial^2}{\partial t^2} - c^2 \left(\frac{1}{\rho} (\partial/\partial\rho) \rho \partial/\partial\rho + \frac{1}{\rho^2} \partial^2/\partial\phi^2 + \partial^2/\partial z^2 \right) \psi(t, \rho, \phi, z) = 0$$

using separation of variables. Note that the functions of ρ will be Bessel functions and their zeros will be the nodes of the wave.

[5] Determine the general form of a standing wave in three dimensions using the spherical polar coordinates. Use separation of variables.

[6] If we are given a wave of the form $\psi(t, r) = A \sin(|k|ct - k \cdot r + \phi)$ where A, k, ϕ are independent random variables, then determine the correlation $\mathbb{E}(\psi(t_1, r_1)\psi(t_2, r_2))$ in terms of the pdf's of A and k . Assume that ϕ is uniformly distributed over $[0, 2\pi]$.

1.2.4 Advanced Exercises

[1] Show that if non-causality is allowed then the general solution to the 3-D wave equation having radial symmetry is of the form

$$\psi(t, r) = s_1(t - |r|/c)/|r| + s_2(t + |r|/c)/|r|$$

By means of an example, explain when this defines a standing wave.

[2] Consider the 1-D wave

$$\psi(t, x) = A_1 \cos(k(ct - x) + \phi_1) + A_2 \cos(k(ct + x) + \phi_2)$$

Under what conditions on $A_k, \phi_k, k = 1, 2$ does this represent a standing wave. If $A_1 = A_2 = A$ and $\phi_k, k = 1, 2$ are jointly distributed random variables, then calculate the space-time correlations of $\psi(t, x)$.

[3] Given a wave field $\psi(t, x)$ in one dimension, we wish to approximate it in the least squares sense by a finite superposition of standing waves. Derive an algorithm for this. Assume a fixed frequency. What is the nsr involved, ie, the ratio of the mean square approximation error over a given space-time interval to the mean square value of ψ over the same range ?

1.2.5 Points to remember

[1] A standing wave by definition is a wave (ie, it satisfies the wave equation) that has its nodes at fixed spatial locations.

[2] Another equivalent way of defining a standing wave is that it can be separated into a product of a function of time only and a function of space only.

[3] A standing wave can be represented as a sum of a forward and a backward propagating wave.

[4] Conversely any solution to the one dimensional wave equation can be represented as a superposition of standing waves.

[5] For a standing wave, the maximum and minimum amplitudes of the wave occur at all spatial points at the same time. For example, if the standing wave is $\cos(\omega t)\cos(kx)$, then at all points x , the maximum and minimum amplitudes occur at the times $2n\pi/\omega, (2n + 1)\pi/\omega$ where n assumes integer values.

[6] More generally, in 3-D, for the standing wave $\cos(\omega t)\cos(k.r)$, at every spatial point r the maximum and minimum amplitudes occur at times $2n\pi/\omega, (2n + 1)\pi/\omega$ where n assumes integer values.

1.3 The polarization of a wave

If $\psi(t, r)$ is a wave taking values in \mathbb{R}^3 , its polarization at the space-time point (t, r) is specified by the direction of $\psi(t, r)$. If we place a small wire antenna of infinitesimal length dl directed along the direction \hat{n} , then the signal collected by this antenna is

$$s(t, r, dl\hat{n}) = dl(\hat{n}, \psi(t, r))$$

If we place such infinitesimal antenna sensors at r_1, \dots, r_k of lengths dl_1, \dots, sl_k respectively and oriented along the directions $\hat{n}_1, \dots, \hat{n}_k$, then the weighted output of the received signals can be expressed as

$$x(t) = \sum_{m=1}^k w_m dl_m(\hat{n}_m, \psi(t, r_m))$$

If the waves are all of the same frequency ω , then we can write

$$\psi(t, r) = \text{Re}(\tilde{\psi}(r)\exp(j\omega t))$$

and the output can be expressed as

$$x(t) = \text{Re}(\tilde{x}\exp(j\omega t))$$

where

$$\tilde{x}(t) = \sum_{m=1}^k w_m dl_m(\hat{n}_m, \tilde{\psi}(r_m))$$

In addition, if this vector wave field is corrupted by a noise field $W(t, r) = \text{Re}(\tilde{W}(r)\exp(j\omega t))$, then the output phasor is

$$\tilde{x} = \sum_{m=1}^k w_m dl_m(\hat{n}_m, \tilde{\psi}(r_m)) + \sum_{m=1}^k w_m dl_m(\hat{n}_m, \tilde{W}(r_m))$$

The signal to noise ratio is

$$SNR = \frac{|\sum_{m=1}^k w_m dl_m(\hat{n}_m, \tilde{\psi}(r_m))|^2}{\mathbb{E}|\sum_{m=1}^k w_m dl_m(\hat{n}_m, \tilde{W}(r_m))|^2}$$

and the signal processing problem is to choose the weights $\{w_m\}$ and the antenna polarization directions $\{\hat{n}_m\}$ so that SNR is a maximum.

1.3.1 Exercises

- [1] Assume that the electric field in space is $E(t, r)$ with correlation $\langle E_a(t, r)E_b(t', r') = R_{ab}(t, r|t', r')$. Infinitesimal antenna sensors $dl_k \hat{n}_k$ are placed at $r_k, k = 1, 2, \dots, N$. Finally, a weighted linear combination of the signals at these sensors is formed

with a weight w_k at the sensor at r_k . Let $X(t)$ denote the output signal. Calculate its autocorrelation $\langle X(t) \cdot X(t') \rangle$ and maximize this correlation w.r.t. the weights $\{w_k\}$ given the constraint $\sum_k w_k^2 = W$.

[2] Repeat the previous problem taking the weight constraint as $\sum_{k,j} q_{kj} w_k w_j = W$ where $Q = ((q_{kj}))$ is a real positive definite matrix.

[3] Let $E(t, r)$ denote the electric field in space and let $\Gamma : s \rightarrow R(s), s \in [0, 1]$ be a wire curve in space. The total electric field signal collected by this antenna, ie, potential difference across this wire is the line integral $V(t) = - \int_0^1 (E(t, R(s)), R'(s)) ds$. Determine the statistical correlations and higher order moments of this potential in terms of those of the electric field.

[4] Explain what the polarization diagram of the damped electric field defined by

$$E(t) = [A_x \cos(\omega t) \hat{x} + A_y \cos(\omega t + \phi) \hat{y}] \exp(-\gamma t)$$

will look like. Assume $\gamma > 0$.

Hint: An ellipse which spirals inwards towards the origin.

[5] Express a circularly polarized wave as a superposition of two plane polarized waves and conversely express a plane polarized wave as a linear combination of a left and a right circularly polarized wave.

[6] Express down a wave field at a given spatial point, ie, as a 2-vector valued function of time whose plane of polarization rotates slowly. Does it satisfy the wave equation ? If not determine a second order pde satisfied by such a wave field.

Hint: Consider

$$\psi(t, z) = A_x(t) \cos(\omega t - kz) \hat{x} + A_y(t) \sin(\omega t - kz) \hat{y}$$

where $A_x(t)$ and $A_y(t)$ are slowly time varying functions. Then,

$$\begin{aligned} \psi_{,t} &= (A'_x(t)/A_x(t))\psi_x(t, z)\hat{x} + (A'_y(t)/A_y(t))\psi_y(t, z)\psi_y(t, z)\hat{y} \\ &\quad - (A_x(t)\omega/A_y(t))\psi_y(t, z)\hat{x} + (A_y(t)\omega/A_x(t))\psi_x(t, z)\hat{x} \end{aligned}$$

Likewise calculate $\psi_{,tt}(t, z)$ and relate these two partial derivatives to $\psi_{,z}$ and $\psi_{,zz}$.

1.3.2 Points to remember

[1] Sometimes, a vector field satisfies a wave equation in one, two or three dimensions as in the case of electromagnetism in free space. Then, at each space-time point, the direction of this wave vector field as a function of time determines its polarization.

[2] If at a given point in space, the tip of the wave vector oscillates along one direction and its negative only, the wave field at that spatial point is said to be plane polarized. If the tip moves around a circle, it is said to be circularly

polarized while if it traces an ellipse, it is said to be elliptically polarized. Other kinds of polarization that do not belong to these classes also exist.

[3] Specifically, if at a position $r \in \mathbb{R}^3$, the wave propagating along the direction of k has its amplitude as a function of time t given by

$$\psi(t, r) = A_1 \mathbf{e}_1 \cos(\omega t - k.r + \phi_1) + A_2 \mathbf{e}_2 \cos(\omega t - k.r + \phi_2)$$

then ψ satisfies the wave equation and if further $\operatorname{div}\psi = 0$, then $k.e_1 = k.e_2 = 0$ and hence the wave vector is of the transverse polarization kind, ie, its vector tip moves in a plane perpendicular to the direction k of propagation. We may assume $e_1.e_2 = 0$ and then for general $A_k, \phi_k, k = 1, 2$, the wave is elliptically polarized. Special cases of this include linear polarization when $\phi_1 = \phi_2$ and circular polarization when $A_1 = A_2, \phi_2 = \phi_1 \pm \pi/2$.

[4] If a wave polarized along the direction \hat{n} passes through a polarizer or a linear wire antenna oriented along the direction \hat{m} , then the output of the polarizer/antenna will have its amplitude multiplied by a factor of $\cos(\theta) = \hat{m}.\hat{n}$ and the intensity multiplied by $\cos^2(\theta)$.

[5] If a vector wave field at a given fixed point r in space has the form

$$\psi(t) = \sum_{k=1}^p A_k \mathbf{e}_k \cdot f_k(t)$$

then the output of a linear wire antenna at r oriented along \hat{n} has its amplitude given by

$$x(t) = (\psi(t), \hat{n}) = \sum_{k=1}^p A_k (e_k, \hat{n}) f_k(t)$$

and the average output power of this antenna will be

$$P_x = \sum_{k,j=1}^p A_k A_j (e_k, \hat{n}) (e_j, \hat{n}) R_{kj}$$

where

$$R_{kj} = \lim_{T \rightarrow \infty} T^{-1} \int_0^T f_k(t) f_j(t) dt$$

1.4 The wave equation in one, two and three dimensions

- [a] One dimensional waves in a vibrating string.
- [b] Two dimensional waves in a vibrating membrane.
- [c] Three dimensional waves in optics.

1.4.1 Summary

Prototype examples of waves in one, two and three dimensions, are respectively the vibrating string, the vibrating membrane and electromagnetic waves. These equations follow by applying fundamental physical principles like Newton's laws of motion, Maxwell's theory of electromagnetism, the equation of state of a gas etc. to specific situations and then making approximations. For example, when one applies Newton's second law of motion to each infinitesimal portion of a vibrating string taking the tension along the string into account and then makes the small slope approximation, one ends up with the one dimensional wave equation. When one considers Faraday's law of electromagnetic induction along with Gauss' law, Ampere's law with Maxwell's displacement current correction and the no magnetic monopole condition all in the absence of external charge and current sources, transforms these integral laws to differential form using Gauss' divergence theorem and Stokes theorem, one end up with the Maxwell equations. When one then applies differential operators to both sides of these equations and eliminates certain variables, we end up with the three dimensional wave equations for the electric and magnetic fields in free space. Likewise, the basic fluid dynamical equations, which are consequences of mass conservation and Newton's second law of motion to continuous media, on making small amplitude approximations lead to the one and two dimensional wave equations for waves in the ocean. Finally, just as in the case of the vibrating string, when one writes down Newton's equations of motion for each infinitesimal area element of a membrane fixed at its boundary taking the surface tension into account, after making small slope approximations, one ends up with the two dimensional wave equation. For solving the wave equations, one typically uses separation of variables either in time and cartesian coordinates, or in time and cylindrical coordinates or in time and polar coordinates, depending on the boundary conditions and the symmetry of the problem. These lead to specific kinds of special functions appearing in the solution like spatial sinusoids, spherical harmonics, Bessel's functions etc. When sources are present in the wave equation like charges and currents in the Maxwell theory, force along the length of a vibrating string, force across the area of a vibrating membrane etc, then we get a wave equation with source and these can be solved by determining the Green's function for the wave operator corresponding to the specific boundary condition. Other kinds of generalizations of the wave equation occur in mathematical physics for example in Einstein's theory of gravitation, ie, general relativity when one linearizes the field equations around a background

metric and uses a specific coordinate system. This leads to gravitational waves and gravitational radiation. Likewise in elasticity, the potential energy of the elastic membrane depends on a number of constants like the Young's modulus, Poisson ratio etc. and these may vary from point to point. One sets up the Lagrangian density for such a membrane or medium and ends up with generalized wave equations for the elastic waves.

[a] One dimensional wave. Example: A vibrating string. Let the string be tied at its two ends $x = 0$ and $x = L$. Let $u(t, x)$ be its displacement along the y axis at time t and x -coordinate x . Let $T(t, x)$ denote the tension in the string. Then since an infinitesimal element of the string does not get displaced along the x direction, it must follow that

$$\frac{\partial}{\partial x}(T(t, x)\cos(\theta(t, x))) = 0$$

or equivalently,

$$T(t, x)\cos(\theta(t, x)) = T_0(t)$$

The y component of the equation of motion taking gravity into account is

$$\sigma\sqrt{1 + u_{,x}^2}u_{,tt} = \partial/\partial x(T(t, x)\sin(\theta(t, x))) - \sigma\sqrt{1 + u_{,x}^2}g$$

Here

$$\tan(\theta(t, x)) = u_{,x}(t, x)$$

is the slope of the string. Here σ is the mass per unit length of the string. Eliminating $T(t, x)$ between the above two equations gives us

$$\sigma\sqrt{1 + u_{,x}^2}u_{,tt} - T_0(t)u_{,xx} + \sigma g\sqrt{1 + u_{,x}^2} = 0$$

This is the final form of the exact equation of the vibrating string.

Simplifications: First assume $|u_{,x}| \ll 1$, ie, the slope of the displaced string is small. Second, assume $T_0(t) = T_0$ is a constant. Third, neglect gravity. The result is the one dimensional wave equation for the vibrating string:

$$\sigma u_{,tt} - T_0 u_{,xx} = 0$$

The velocity of these waves is read off immediately as

$$c = \sqrt{T_0/\sigma}$$

1.4.2 Exercises

- [1] If a one dimensional string is fastened at two ends $x = 0$ and $x = L$, both ends of which execute motion along the $y = u$ direction in accordance with the displacements

$$u(t, 0) = f_1(t), u(t, L) = f_2(t)$$

then solve the wave equation

$$u_{,tt}(t, x) - c^2 u_{,xx}(t, x) = 0$$

with these boundary conditions and obtain the space-time correlations $\mathbb{E}(u(t_1, x_1).u(t_2, x_2))$ in terms of $\mathbb{E}(f_1(t)f_1(t'))$, $\mathbb{E}(f_2(t), f_2(t'))$ and $\mathbb{E}(f_1(t)f_2(t'))$.

- [2] If a two dimensional circular membrane whose boundary undergoes random motion in accordance with the equation $u(t, R, \phi) = f(t, \phi)$ where $f(t, \phi)$ is a random force with correlations $\mathbb{E}(f(t, \phi).f(t', \phi'))$ and simultaneously a random force per unit area of the membrane given by $g(t, r, \phi)$ with known correlations is applied, then calculate the space-time correlations of the vibrating membrane. Note that the displacement of the membrane $u(t, r, \phi)$ satisfies the two dimensional wave equation with forcing

$$u_{,tt} - c^2(u_{,rr} + u_{,r}/r + u_{,\phi\phi}/r^2) = g(t, r, \phi), t \geq 0, 0 \leq r \leq R, 0 \leq \phi < 2\pi$$

with the boundary condition

$$u(t, R, \phi) = f(t, \phi), t \geq 0, 0 \leq \phi < 2\pi$$

- [3] Using multivariate Fourier transform theory, solve the wave equation in p time dimensions and q spatial dimensions:

$$\sum_{k=1}^p \frac{\partial^2 u(t, x)}{\partial t_k^2} - c^2 \sum_{k=1}^q \frac{\partial^2 u(t, x)}{\partial x_k^2} = 0$$

Obtain the general solution.

- [4] Show that the wave equation of the previous problem can be derived from the action principle

$$\delta S[u] = 0$$

where

$$S[u] = (1/2) \int_D [(\sum_{k=1}^p \partial u / \partial t_k)^2 - (\sum_{k=1}^q (\partial u / \partial x_k)^2)] d^p t d^q x$$

after imposing appropriate boundary conditions for u on ∂D .

- [5] The solution to the linearized Einstein field equation in vacuum leads to the metric perturbations $h_{\mu\nu}(t, r)$ satisfying the wave equation

$$\partial_\mu \partial^\mu h_{\alpha\beta}(x) = 0$$

provided that the coordinate condition

$$h_{\mu,\nu}^\nu - (1/2)h_{,\mu} = 0, h = h_\alpha^\alpha$$

Express the solution for $h_{\mu\nu}$ as a superposition of plane waves with the coordinate condition leading to certain relationships between the Fourier amplitudes $e_{\mu\nu}(K)$:

$$h_{\mu\nu}(x) = \int [e_{\mu\nu}(K)\exp(i|K|t - iK.r) + \bar{e}_{\mu\nu}(K)\exp(-i|K|t + iK.r)]d^3K$$

Determine the state of polarization of such a wave for a given vector K , like say, K is along the z -axis. Show that there are effectively only five linearly independent coefficients and that these coefficients transform under a rotation by angle θ around the z axis (ie K axis) as rotations by angles zero, θ and 2θ , thereby justifying that gravitons are spin two Bosons. Compare this with the state of polarization of photon, by noting that in the Coulomb gauge, the electric scalar potential is a pure matter field and the magnetic vector potential owing to the Coulomb gauge condition $\operatorname{div}A = 0$ has just two degrees of freedom implying that the photon is a spin one particle but its spin along the axis of propagation can never assume the value zero, ie, it assumes only the values ± 1 . Deduce this result by applying a rotation around the z -axis by an angle θ assuming that K is along the z axis and deduce that (A_x, A_y) gets rotated by an angle θ . It follows from this result that the quantum mechanical state of a photon can be only either a left circularly polarized state, or a right circularly polarized state or a linear combination of these two states. (Reference: The Feynman Lectures on Physics, Vol.III, Narosa Publishers)

1.5 The polarization of a wave revisited

1.5.1 Linear, circular and elliptic polarizations of plane electromagnetic waves

Assume that we have a plane electromagnetic wave of frequency ω travelling along the z axis in free space. In terms of phasors, the amplitude of this wave has the form

$$\mathbf{E}(t, z) = \operatorname{Re}(\tilde{\mathbf{E}}_x \hat{x} + \tilde{\mathbf{E}}_y \hat{y}) \exp(j(\omega t - kz))$$

where \tilde{E}_x and \tilde{E}_y are complex numbers and $k = \omega/c$. We write

$$\tilde{E}_x = A_x \exp(j\phi_x), \tilde{E}_y = A_y \exp(j\phi_y)$$

where

$$A_x = |\tilde{E}_x|, A_y = |\tilde{E}_y|, \phi_x = \operatorname{Arg}(\tilde{E}_x), \phi_y = \operatorname{Arg}(\tilde{E}_y)$$

and then we can write

$$\mathbf{E}(t, z) = A_x \cos(\omega t - kz + \phi_x) \hat{x} + A_y \cos(\omega t - kz + \phi_y) \hat{y}$$

Note that the free space Maxwell equation $\operatorname{div} \mathbf{E} = 0$ is satisfied only because we have assumed the electric field to have only transverse components relative to the direction \hat{z} of propagation. At any point in space having z -coordinate z , we have assuming $\phi_x = \phi_y$ that

$$E_y(t, z)/E_x(t, z) = A_y/A_x$$

for all times t . Thus, this wave is linearly polarized, ie, its vector oscillates along a fixed direction. This is also true if $\phi_y = \phi_x + \pi$ in which case, we get

$$E_y(t, z)/E_x(t, z) = -A_y/A_x$$

for all times t . Now if $\phi_y = \phi_x \pm \pi/2$, it easily follows that

$$E_x(t, z)^2/A_x^2 + E_y(t, z)^2/A_y^2 = 1$$

implying that the E vector traces out an ellipse with semimajor and semiminor axes A_x, A_y or A_y, A_x . Such a wave is said to be elliptically polarized if $A_y \neq A_x$ and circularly polarized if $A_y = A_x$. Finally, for general ϕ_x, ϕ_y, A_x, A_y , we have elliptic polarization but with semimajor and semiminor axis being oriented not along the x and y axis but rather some other two perpendicular directions in the xy plane. To see this, we eliminate $\psi = \omega t - kz$ between the expressions for $E_x(t, z)$ and $E_y(t, z)$. We get

$$\begin{aligned} E_y &= A_y \cdot \cos(\psi + \phi_y) = A_y \cdot \cos(\cos^{-1}(E_x/A_x) + \phi_y - \phi_x)) \\ &= A_y((E_x/A_x)\cos(\phi) - \sqrt{1 - E_x^2/A_x^2}\sin(\phi)) \end{aligned}$$

where $\phi = \phi_y - \phi_x$. This equation gives

$$(E_y - \alpha A_y E_x/A_x)^2 = \beta^2 A_y^2 - \beta^2 A_y^2 E_x^2/A_x^2$$

where

$$\alpha = \cos(\phi), \beta = \sin(\phi)$$

Equivalently,

$$E_y^2 + (A_y/A_x)^2 E_x^2 - 2\alpha A_y E_x E_y/A_x = \beta^2 A_y^2$$

1.5.2 Exercises

- [1] Apply a rotation in the $E_x - E_y$ plane to this equation to bring it to the standard form of an ellipse with semi-major and semi-minor axes oriented along the x and y axes.

[2] If a wave of light polarized along the direction \hat{n} is incident on a polarizer that is oriented along the direction \hat{m} , then its amplitude gets multiplied by $\hat{n} \cdot \hat{m}$ and the out-coming wave is polarized along the direction \hat{m} (This is deduced from the state collapse postulate in quantum mechanics). The intensity therefore gets multiplied by $(\hat{n} \cdot \hat{m})^2$. Now explain how by designing a polarizer that allows the components of light polarized along only the \hat{x} and \hat{y} directions to enter can be used to generate circularly polarized light. Secondly, suppose we have a sequence of polarizers arranged parallelly to each other such that the k^{th} polarizer is oriented along the unit vector \hat{m}_k , $k = 1, 2, \dots, p$. Then if a light wave having amplitude $A_x \cos(\omega t - kz)\hat{x} + A_y \sin(\omega t - kz)\hat{y}$ is incident upon the first polarizer, then what will be the wave amplitude after it comes out of the last polarizer. Assume that the distance between two successive polarizers is d and that the planes of all the polarizers are parallel to the xy plane.

1.6 Basics of fluid dynamics

1.6.1 Exercises

[1] Derive the Navier-Stokes equation using momentum and mass conservation equations taking viscosity into account. The momentum flux tensor is given by

$$\Pi_{ij} = \rho v_i v_j + p \delta_{ij} - \eta(v_{i,j} + v_{j,i}) - \chi(\operatorname{div} v) \delta_{ij}$$

and the momentum density is

$$P_i = \rho v_i$$

The momentum conservation reads

$$P_{i,t} + \Pi_{ij,j} = 0$$

Justify this equation by taking integrals over the volume of the fluid enclosed by a closed surface. The mass density is ρ and the mass flux is ρv_i . Therefore the mass conservation equation reads

$$\rho_{,t} + (\rho v_i)_{,i} = 0$$

[2] Derive the Bernoulli equation

$$p/\rho + v^2/2 + \phi_{,t} + \Phi = \text{constt.}$$

within an incompressible (ie, its density is a constant) and irrotational fluid where ϕ is the velocity potential and Φ is the potential of external forces like gravity. Note that since $\operatorname{div} v = 0$ and $v = \nabla \phi$, it follows that $\nabla^2 \phi = 0$ and hence the viscous contribution to the Bernoulli equation is zero.

[3] Derive the Bernoulli equation:

$$p/\rho + v^2/2 + \Phi = \text{constant}$$

along a streamline if the fluid is only incompressible but a steady state has been achieved. If steady state has not been reached, then along a streamline, the above equation generalizes to

$$p/\rho + v^2/2 + \Phi = constt.$$

along a streamline if viscous effects are not taken into account. Hint:Take the dot product of the Navier-Stokes equation with \hat{v} where \hat{v} is a unit vector along the velocity field at a given space-time point. Note that since we are not assuming irrotationality, there is no velocity potential.

[4] Derive the basic MHD equations and explain how you would solve them by the finite difference method. The basic MHD equations for a conducting fluid are

$$\rho(v_{,t} + (v, \nabla)v) = -\nabla p + \eta\nabla^2v + \sigma(E + v \times B) \times B - \nabla\Phi,$$

$$\rho_{,t} + div(\rho v) = 0,$$

$$divE = 0, divB = 0, curlE = -B_{,t}, curlB = \sigma(E + v \times B) + \epsilon E_{,t},$$

$$\nabla^2\Phi = 4\pi G\rho$$

Study small fluctuations and the dispersion relation in such a MHD fluid by linearizing using

$$v(t, r) = V_0 + \delta v(t, r), E(t, r) = E_0 + \delta E(t, r), B(t, r) = B_0 + \delta B(t, r),$$

$$\Phi(t, r) = \delta\Phi(t, r), \rho(t, r) = \rho_0 + \delta\rho(t, r)$$

[3] Derive the basic Navier-Stokes and equation of continuity equation using Lagrange multiplier fields. Derive consequently the costate equations for the Lagrange multiplier fields and explain how using the Dirac brackets for problems with constraints, the fluid flow problem can be quantized.

[4] Show that for an incompressible fluid, we can represent the velocity field as

$$v = \nabla \times \psi,$$

where

$$div\psi = 0$$

Hence, by taking the curl of the Navier-Stokes equation, eliminate the pressure field and reduce it to an equation involving only ψ . Specialize to 2-D flows, where justifying that the vector field ψ can be taken as $\psi_0(t, x, y)\hat{z}$ and derive a nonlinear pde for ψ_0 in the presence of an external force field per unit volume.

1.6.2 Points to remember

[1] If $v(t, r)$ is the 3-D fluid velocity field, then the acceleration of a given fluid particle at the time-space point (t, r) is given by

$$\begin{aligned}\frac{d}{dt}v(t, r) &= \partial v(t, r)/\partial t + (dx/dt)\frac{\partial v}{\partial x} + (dy/dt)\frac{\partial v}{\partial y} + (dz/dt)\frac{\partial v}{\partial z} \\ &= v_{,t} + v_x v_{,x} + v_y v_{,y} + v_z v_{,z} = \\ &\quad v_{,t} + (v, \nabla)v\end{aligned}$$

This acceleration along the trajectory of the fluid particle is called the material derivative of the velocity and is denoted by Dv/Dt in contrast with $v_{,t}$ which represents just the acceleration of particles at a fixed spatial point. If $v_{,t} = 0$ as happens in steady state, we may still have $Dv/Dt \neq 0$. Dv/Dt can be measured by injecting a drop of ink into the fluid and tracing out its trajectory.

[2] The Navier-Stokes equation can either be derived using Newton's second law of motion by including as forces the pressure gradient force per unit volume and the viscous forces per unit volume along with other external forces like the gravitational force and if the fluid is conducting, the electromagnetic force. It can also be derived by using the conservation of mass and momentum. For this, we define the mass flux and the momentum flux taking into account the stress tensor per unit time due to pressure and viscous forces the latter arising from velocity gradients between two neighbouring surfaces.

[3] The Navier-Stokes equation, the mass conservation equation and the equation of state which relates the pressure to the density give us five equations for the five functions v_x, v_y, v_z, ρ, p .

[4] Any fluid dynamical problem can formally be approximately solved using the finite element method which involves discretizing space and time into pixels and replacing space-time partial derivatives by partial differences. This is a vast subject called Computational Fluid Dynamics (CFD).

[5] To solve a fluid dynamical problem we need the boundary conditions which state that for a viscous fluid both the tangential and normal components of the velocity field vanish on a rigid boundary while for a non-viscous fluid (ie, Eulerian fluid), only the normal component of the velocity vanishes on the boundary.

[6] Fluid dynamical problems with boundary conditions having different kinds of symmetries like spherical, cylindrical, rectangular symmetry can be solved exactly in many cases by choosing our coordinate system appropriately in accordance with the symmetry and expressing the fluid equations in that coordinate system, like spherical polar, cylindrical or Cartesian systems.

[7] 2-D fluids which are incompressible and irrotational can be solved by solving Laplace's equation for the stream function or its dual, the potential function. The Stream function and the potential function satisfy the Cauchy-Riemann equations for the real and imaginary part of an analytic function of a complex variable and hence many 2-D fluid problems with boundary conditions can be solved exactly using functions of a complex variable. The constant velocity potential surfaces are orthogonal to the constant stream function surfaces and hence the complex potential should be such that its imaginary part, namely some surface of constant stream function value determines a boundary which is tangential to the velocity flow.

[8] By introducing the notion of a velocity potential for an irrotational fluid velocity field, it is possible under the assumption of incompressibility to obtain the Bernoulli equation which states that something like the energy density of the fluid is a constant at all space-time points even if viscous effects are taken into account. The various contributions to this energy density come from the kinetic energy per unit mass, the pressure, the velocity potential and the potential of the externally applied conservative force field like gravity.

[9] If viscous effects are not present, then even if irrotationality does not hold good, under the assumption of incompressibility, a quantity like the energy density described above but without the velocity potential is conserved only along a streamline, ie, the directional derivative of this energy along the direction of the velocity vector vanishes everywhere.

1.7 Waves in a fluid—derivation from first principles

1.7.1 General theory

Assume that the sea is one dimensional, ie, flow takes place along the x axis and the height of the sea level at the point (x, y) at time t is $\xi(t, x)$. Assume that the breadth of the sea along the y -axis is w . Let $v_x(t, x, z)$ be the x -component of the velocity of the fluid at (x, y, z) and $v_z(t, x, z)$ the z -component of the velocity of the fluid at (x, y, z) both at time t . The equation of conservation of mass implies

$$-\int_0^{\xi(t,x)} v_{x,x}(t, x, z) w dz = \xi_{,t}(t, x) w$$

or equivalently,

$$-\xi_{,t}(t, x) = \int_0^{\xi(t,x)} v_{x,x}(t, x, z) dz \quad \dots \quad (1)$$

Further, the Navier-Stokes equation gives

$$v_x v_{x,x} + v_z v_{x,z} + v_{x,t} = -p_{,x}/\rho + \nu(v_{x,xx} + v_{x,zz}) \quad \dots \quad (2)$$

$$v_x v_{z,x} + v_z v_{z,z} + v_{z,t} = -p_{,z}/\rho + \nu(v_{z,xx} + v_{z,zz}) - g \quad \dots \quad (3)$$

The mass conservation equation can be expressed as

$$v_{x,x} + v_{z,z} = 0 \quad \dots \quad (4)$$

These equations are four in number for the four functions $v_x(t, x, z)$, $v_z(t, x, z)$, $p(t, x, z)$, $\xi(t, x, z)$. We now make some simplifying approximations: A1:

$$\xi(t, x) = h + \delta\xi(t, x)$$

where $\delta\xi$ is small, ie, the fluctuation in the sea level above or below the constant level h is small. Further, assume that v_x is independent of z . Then, we have, without any approximation,

$$-(h + \delta\xi(t, x))v_{x,x}(t, x) = \delta\xi_{,t}(t, x)$$

Also, (4) then implies that

$$v_z(t, x, z) = -v_{x,x}(t, x)z + f(t, x)$$

where f some function. Then, (2) and (3) give

$$v_{x,t} + v_x v_{x,x} = -p_{,x}/\rho + \nu v_{x,xx} \quad \dots \quad (5)$$

so $p_{,x}(t, x, z)$ is independent of z . Thus, we can write

$$p(t, x, z) = q_1(t, x) + q_2(t, x)z$$

Substituting these expressions into,(5) results in

$$v_{x,t} + v_x v_{x,x} = -(q_{1,x} + q_{2,x}z)/\rho + \nu v_{x,xx}$$

Thus, $q_{2,x}z$ is independent of z and hence $q_{2,x} = 0$. This means that $q_2(t, x) = q_2(t)$. and we get

$$v_{x,t} + v_x v_{x,x} = -q_{1,x}/\rho + \nu v_{x,xx} \quad \dots \quad (6)$$

Likewise, (3) gives

$$f_{,t} - v_x t x z + v_x(f_{,x} - v_{x,x}z) + (f - v_{x,x}z)v_{x,x} = -q_2/\rho + \nu(f_{,xx} - v_{x,xxx}z) - g \quad \dots \quad (7)$$

It follows that

$$v_{x,tx} + v_x v_{x,xx} + v_{x,x}^2 = \nu v_{x,xxx} - a_{,xx}(t, x) \quad \dots \quad (8)$$

and

$$f_{,t} + v_x f_{,x} + f v_{x,x} = -q_2/\rho + \nu f_{,xx} - g - a_{,xx}(t, x) \quad \dots \quad (9)$$

Comparing (6) and (8), we get that

$$q_{1,xx}/\rho = a_{,xx}(t, x)\rho$$

and hence,

$$q_1(t, x) = a(t, x)\rho + b_1(t)x + b_2(t)$$

So our final set of equations is

$$-(h + \delta\xi(t, x))v_{x,x}(t, x) = \delta\xi_{,t}(t, x) \quad \text{--- (a),}$$

$$f_{,t}(t, x) + v_x(t, x)f_{,x}(t, x) + f(t, x)v_{x,x}(t, x) = -q_2(t)/rho + \nu f_{,xx}(t, x) - g - a_{,xx}(t, x) \quad \text{--- (b),}$$

$$v_{x,t}(t, x) + v_x(t, x)v_{x,x}(t, x) = -(a_{,x}(t, x) + b_1(t)) + \nu v_{x,xx} \quad \text{--- (c)}$$

Also when $z = \xi(t, x) = h + \delta\xi(t, x)$, the pressure must be the constant atmospheric pressure p_0 , ie,

$$q_1(t, x) + q_2(t)\xi(t, x) = p_0$$

or equivalently,

$$a(t, x) + b_1(t) + b_2(t)(h + \delta\xi(t, x)) = p_0 \quad \text{--- (d)}$$

(a)-(d) is our final set of four equations for the functions $\delta\xi(t, x), v_x(t, x), f(t, x), a(t, x), b_1(t), b_2(t)$.

We now make some approximations to arrive at the one dimensional wave equation for $\xi(t, x)$. First assuming $v_z = 0$, it follows from (3) that $p(t, x, z) = -\rho gz + A(t, x)$ and since $p(t, x, h + \delta\xi(t, z)) = p_0$, we get

$$p(t, x, z) = p(t, x, z) = p_0 + \rho g(h + \delta\xi(t, x) - z)$$

Still no approximation has been made. We then observe that again without any approximation,

$$v_{x,x}(t, x)(h + \delta\xi(t, x)) + \xi_{,t}(t, x) = 0,$$

$$v_{x,t} + v_x(t, x)v_{x,x}(t, x) = -g\delta\xi_{,x}(t, x) + \nu v_{x,xx}(t, x)$$

However the equation of continuity becomes since $v_z = 0$,

$$v_{x,x} = 0$$

which cannot be satisfied. So we say that $v_{x,x}$ is very small since it is exactly equal to $-v_{z,z}$ which is very small. Also neglecting the nonlinear terms in the above equations, we get the approximate equations

$$hv_{x,x}(t, x) + \delta\xi_{,t}(t, x) = 0,$$

$$v_{x,t} + \rho g\delta\xi_{,x}(t, x) + \nu v_{x,xx}(t, x) = 0$$

If $\nu = 0$, ie, the viscosity is neglected, then these two equations result in the wave equation after eliminating v_x :

$$\delta\xi_{,tt}(t, x) - gh\delta\xi_{,xx}(t, x) = 0$$

This is the same wave equation as above except with a viscous correction term. We leave it as an exercise to determine the dispersion of such a viscous wave, ie, relationship between wavelength and frequency.

1.7.2 Exercises

[1] Prove that if waves in two dimensions in the sea are considered using the following method, then we get after approximation the two dimensional wave equation: The height of the sea level is $\xi(t, x, y)$. The z component of the velocity within the liquid volume is neglected and mass conservation the for the volume over the area $[x, x + dx] \times [y, y + dy]$ yields

$$\xi_{,t}(t, x, y) + (\xi(t, x, y)v_x(t, x, y))_{,x} + (\xi(t, x, y)v_y(t, x, y))_{,y} = 0$$

Writing

$$\xi(t, x, y) = h + \delta\xi(t, x, y)$$

the above equation approximates to

$$\delta\xi_{,t} + h(v_{x,x} + v_{y,y}) = 0 \quad \dots \quad (1)$$

Further, the equation of motion along the x and y directions gives after neglecting the nonlinear part,

$$v_{x,t} = -p_{,x}/\rho$$

where

$$p(t, x, y, z) = \rho g(\xi(t, x, y) - z)$$

In this way we get

$$v_{x,t} + g\delta\xi_{,x} = 0 \quad \dots \quad (2)$$

and likewise,

$$v_{y,t} + g\delta\xi_{,y} = 0 \quad \dots \quad (3)$$

From (1),(2) and (3), derive the two dimensional wave equation for $\delta\xi(t, x, y)$ and show that the velocity of wave propagation is again \sqrt{gh} . Now, improve upon this theory by taking both the nonlinear terms and the viscous terms to get

$$\begin{aligned} & \xi_{,t}(t, x, y) + (\xi(t, x, y)v_x(t, x, y))_{,x} + (\xi(t, x, y)v_y(t, x, y))_{,y} = 0, \\ & v_{x,t} + v_x v_{x,t} + v_y v_{y,t} + g\xi_{,x} + \nu(v_{x,xx} + v_{x,yy}), \\ & v_{y,t} + v_x v_{y,x} + v_y v_{y,y} + g\xi_{,y} + \nu(v_{y,xx} + v_{y,yy}) \end{aligned}$$

These are three equations for the three functions $\xi(t, x, y), v_x(t, x, y), v_y(t, x, y)$. Use perturbation theory to obtain approximate nonlinear wave equations for $\xi(t, x, y)$. For this, you must consider $v_x, v_y, \delta\xi$ to be of the first order of smallness where $\xi = h + \delta\xi$ and introduce a perturbation parameter δ in the quadratic terms to show that these terms are small. After that you must expand the functions $\delta\xi, v_x, v_y$ in powers of δ .

[2] Consider a 2-D incompressible fluid with random velocity field $v_x(x, y), v_y(x, y)$ independent of time. Along with the random pressure field $p(x, y)$, these satisfy the Navier-Stokes equation and the equation of continuity:

$$v_x v_{x,x} + v_y v_{x,y} = \nu(v_{x,xx} + v_{x,yy}) - p_{,x}/\rho\omega,$$

$$v_x v_{y,x} + v_y v_{y,y} = \nu(v_{y,xx} + v_{y,yy}) - p_{,y}/\rho,$$

$$v_{x,x} + v_{y,y} = 0$$

Assuming that the statistical moments of the velocity and pressure field satisfy the hypothesis of homogeneity and isotropy and are zero mean fields, show that the second and third moments can be expressed as (in accordance with homogeneous and isotropic tensors)

$$\langle v_i(r)v_j(r') \rangle = B_{ij}(r' - r) = A(|r - r'|)\delta_{ij} + B(|r - r'|)n_i n_j$$

where $i, j = 1, 2$ with $i = 1$ meaning the x component and $i = 2$ meaning the y component.

$$\begin{aligned} (n_i, i = 1, 2) &= n_1 \hat{x} + n_2 \hat{y} = (r' - r)/|r' - r| \\ \langle v_i(r)v_j(r')v_k(r'') \rangle &= C_{ijk}(r' - r, r'' - r) \\ &= C_1(|r' - r|, |r'' - r|, |r'' - r'|)(\delta_{ij}n_k + \delta_{jk}n_i + \delta_{ki}n_j) \\ &\quad + C_2(|r' - r|, |r'' - r|, |r'' - r'|)n_i n_j n_k \end{aligned}$$

where C_1, C_2 are totally symmetric functions of their arguments and all pressure-velocity correlations are zero. Now derive differential equations for the functions A, B, C_1, C_2 by substituting into the Navier-Stokes equation and the equation of continuity.

1.7.3 Points to remember

[1] Waves motion of the top surface displacement in the sea can be derived by using mass conservation and the Navier-Stokes equation. Certain simplifying assumptions need to be made here to arrive at the wave equation. These assumptions are that the displacement of the sea surface from the equilibrium level $\xi(t, x, y) - h$ is small and the horizontal components of the velocity are also small and finally, the formula for the pressure in a static liquid based on the weight of the water above a unit surface is valid.

[2] If the above assumptions are not made, we still would get a set of three consistent partial differential equations for $v_x(t, x, y), v_y(t, x, y), \xi(t, x, y)$ but we would require a computer to solve this set.

[3] The statistical theory of turbulence was created by A.N.Kolmogorov in which he assumed using isotropicity and homogeneity certain tensor forms for the velocity field and pressure field statistical moments and derived differential equations for these higher order correlations. These equations are valid in the steady state. For a more thorough treatment of this and other fluid dynamical problems, the reader is referred to: Landau and Lifshitz, "Fluid dynamics".

[4] If approximations are not made the nonlinear system of pde's described above can be approximately solved using the solution for the linearized wave equation and perturbation theory by treating the nonlinear terms as small using a perturbation parameter δ tagged to them.

1.8 Longitudinal sound/pressure waves in a tube

1.8.1 General theory

Consider a tube of cylindrical cross section and length L inside which there is only air. Assume that the length of the tube is along the x axis. We consider pressure waves within this tube in the form of longitudinal oscillations of the air column. At time 0, the density of the air column within the length $[x, x + dx]$ is $\rho(0, x)$. The mass of the air within this column at time 0 is therefore $\rho(0, x)dx$. At time t , this infinitesimal air column gets displaced to the interval $[x + u(t, x), x + dx + u(t, x + dx)]$ and hence by mass conservation, the density of air within this displaced column is

$$\rho(t, x + u) = \rho(0, x)dx/(dx + u(t, x + dx)) = \frac{\rho(0, x)}{1 + u_{,x}(t, x)}$$

Assuming an equation of state $p = F(\rho)$, the pressure within this displaced column is

$$p(t, x + u) = F(\rho(0, x)(1 + u_{,x}(t, x))^{-1})$$

The difference between the pressures at $x + u$ and at $x + dx + u + u_{,x}dx = x + dx + u + du$ is

$$\begin{aligned} p(t, x+u) - p(t, x+dx+u+du) &= p(t, x+u) - p(t, x+dx+u+u_{,x}dx) \\ &= p(t, x+u(t, x)) - p(t, x+dx+u(t, x+dx)) = \\ &F(\rho(0, x)(1 + u_{,x})^{-1}) - F(\rho(0, x + dx)(1 + u_{,x} + u_{,xx}dx)^{-1}) \\ &= -dx \frac{d}{dx} F(\rho(0, x)(1 + u_{,x})^{-1}) \end{aligned}$$

This equation leads us immediately to the Newton's second law of motion:

$$\rho(0, x)dx.u_{,tt}(t, x) = -dx \frac{d}{dx} F(\rho(0, x)(1 + u_{,x})^{-1})$$

or equivalently,

$$\rho(0, x)u_{,tt}(t, x) + \frac{d}{dx} F\left(\frac{\rho(0, x)}{1 + u_{,x}(t, x)}\right) = 0$$

This is the exact equation of motion for the longitudinal sound/pressure wave $u(t, x)$.

Simplifying approximations: Assume that $\rho(0, x) = \rho_0$, ie, the initial density is a constant. Further, assume that $|u_{,x}| \ll 1$, ie, the oscillations of the air column are small. Then, we have approximately,

$$F\left(\frac{\rho(0, x)}{1 + u_{,x}}\right) = F(\rho_0(1 - u_{,x})) = F(\rho_0) - F'(\rho_0)\rho_0 u_{,x}$$

and we get with such an approximation, the one dimensional wave equation for u :

$$u_{,tt}(t, x) - F'(\rho_0)u_{,xx}(t, x) = 0$$

The velocity of the sound wave is read off immediately from this equation as

$$c = \sqrt{F'(\rho_0)} = \sqrt{dp(\rho_0)/d\rho}$$

1.8.2 Exercises

[1] Solve approximately the exact nonlinear sound propagation wave equation using perturbation theory by attaching a small perturbation parameter to the nonlinearity.

[2] If the gas within a pipe is charged with charge density proportional to the mass density $\rho(t, x)$ and if an external electric field $E(t, x)$ is applied across the x axis, then what will be the modification to the wave equation for the gas?

[3] If the cross section of the pipe varies with time in accord with the radius of the pipe at (t, x) being $R(t, x)$, then what will the equations of motion of the gas through the pipe look like assuming that the length of the pipe is parallel to the z direction and that the non-vanishing components of the gas velocity are $v_z(t, z, r), v_r(t, z, r)$ where $r = \sqrt{x^2 + y^2}$? Assume that the density of the gas within the pipe also has the same form $\rho(t, z, r)$.

1.8.3 Points to remember

[1] By applying Newton's second law of motion and the equation of mass conservation along with the equation of state of a gas inside a tube, we can derive after approximations, the one dimensional wave equation for the pressure or equivalently the gas displacement inside the pipe. The force on an infinitesimal element of the gas is obtained as a pressure gradient.

[2] If the gas is charged and an external electric field is present, then we can modify the wave equation by taking into account apart from the forces induced by pressure gradients the force of the electric field on the charges.

1.9 The difference between transverse and longitudinal waves in terms of wave polarization

When a wave $\psi(t, r) \in \mathbb{R}^3$ is vector valued, and it propagates along the direction \hat{n} , we can express it as

$$\psi(t, r) = \psi(t - \hat{n} \cdot r/c, 0)$$

It is readily verified that this satisfies the three dimensional wave equation. We say that this wave field is transversely polarized if

$$\hat{n} \cdot \psi(t, r) = 0$$

and that it is longitudinally polarized if

$$\hat{n} \times \psi(t, r) = 0$$

If both of these are not satisfied, then it is partially longitudinal and partially transverse polarized. For example, the TEM waves in empty space are transversely polarized as follows from the Maxwell equations $\nabla \cdot E(t, r) = 0, \nabla \cdot H(t, r) = 0$. Soundwaves in a tube are longitudinally polarized while the electromagnetic field within a waveguide is partially longitudinal and partially transverse polarized. For example, the TE modes have the electric field transversely polarized while the magnetic field partially longitudinal and partially transverse while the TM modes have the magnetic field transversely polarized while the electric field partially longitudinal and partially transversely polarized.

1.9.1 Exercises

[1] Give examples of electromagnetic waves that are partially transverse and partially longitudinally polarized.

[2] If $\phi(t, r) = \phi_0(t).exp(-ik.r)$ in phasor form is the electric scalar potential and if the magnetic vector potential can be neglected, then show that the electric field is longitudinally polarized.

[3] If a wave has wave vector k and can be expressed as a sum of a longitudinally polarized component and a parallelly polarized component as

$$\psi(t, r) = F(\omega t - k.r)\hat{k} + G(\omega t - k.r)(\hat{n} - (k, \hat{n})k/k^2)$$

where \hat{n} is an arbitrary constant unit vector, then calculate $\text{curl}\psi(t, r)$ and $\text{div}\psi(t, r)$. Specifically, show that the former is expressible completely in terms of G only while the latter is expressible in terms of F only.

1.9.2 Points to remember

[1] Transverse waves as in electromagnetics usually follow from the wave equation by a condition like the divergence of the wave vector field vanishes.

[2] Sound waves in a pipe are examples of longitudinal waves, as also are waves in spring.

[3] There may exist waves which have both longitudinal and transverse components as for example, the electromagnetic waves in a waveguide or a cavity resonator.

1.10 Maxwell's equations and the wave equation for the electric and magnetic fields in free space

In free space, the Maxwell equations imply after taking the curl of the two curl equations and making use of the vanishing of the divergence of the electric and magnetic fields, that the electric and magnetic fields satisfy the three dimensional wave equation whose general solution can be expressed as a superposition of plane waves of different frequency and wave vectors but each wave traveling at the speed of light in the medium. The divergence equations in the set of Maxwell equations imply further that in this superposition of plane waves, if we consider a definite frequency and wave vector component, the corresponding amplitudes of the electric and magnetic field vectors will not only be orthogonal to the wave vector, ie, to the direction of propagation of that plane wave but will also be mutually orthogonal. Thus, the unit vectors along the direction of propagation and those along the direction of the electric and magnetic fields will form an orthonormal right handed system of vectors. These waves are therefore called transverse electromagnetic waves, ie, TEM waves. In the presence of charges and currents, the divergence equations appearing in the set of Maxwell equations will have sources and hence if we apply the same manipulations to the Maxwell curl equations and substitute for the divergences of the electric and magnetic fields, we would once again obtain three dimensional wave equations for them but now with vector valued sources. Further, the solution of this system of wave equations with sources will result in the four dimensional convolution of the Green's function

$$G(t, r|t', r') = -\delta(t - t' - |r - r'|/c)/4\pi|r - r'|$$

for the wave operator in free space with the source fields derived from the charge and current densities. At a definite frequency ω , the resulting electric and magnetic field vectors in the far field zone, ie, distant from the charged matter fields which generate them can be expressed in the form of spherical waves, ie, their dependence on t, r is of the form $Re(F(\hat{r})\exp(j(\omega(t - |r|/c))/r)$ which means that these electromagnetic waves propagate in the far field zone as spherical modulated by an amplitude that depends on the direction alone. The Maxwell divergence equations imply that these electric and magnetic field vectors in the far field zone are approximately perpendicular to each other and also the radial direction of propagation. These orthogonality relations hold good provided that we neglect $O(1/r^2)$ terms. Further, in the far field zone, we find that the Poynting vector which gives the power flow per unit area starts with a $1/r^2$ term and this term contributes on integration over the surface of a sphere of radius r , a definite amount of power that is independent of $|r|$. This is because, the infinitesimal surface area term $r^2 d\Omega(\hat{r})$ has an r^2 which cancels out with the $1/r^2$ factor appearing in the Poynting vector. The $O(1/r^n), n \geq 3$ terms do not contribute to any power flow from the source as $r \rightarrow \infty$ for the

same reason, ie, for $n \geq 3$

$$r^2 d\Omega(\hat{r}) \times 1/r^n = r^{2-n} d\Omega(\hat{r}) \rightarrow 0, r \rightarrow \infty$$

In the near field zone, higher powers of $1/r$ than two appearing in the Poynting vector contribute to power efflux and power influx, ie, a constant exchange between the two. These effects are therefore inductive, just as how power gets constantly interchanged between an inductor and a capacitor in an LC circuit. The near field zone is therefore also called the inductive zone.

1.10.1 Exercises

[1] Derive the wave equation with sources for the electric and magnetic field in the presence of electric and magnetic charge and current densities. Start with the Maxwell equations:

$$\operatorname{div} E = \rho/\epsilon, \operatorname{div} H = \rho_m/\mu, \operatorname{curl} E = -\mu H_{,t} - M,$$

$$\operatorname{curl} H = J + \epsilon E_{,t}$$

Show that these equations remain invariant under the transformations

$$\epsilon \rightarrow \mu, \mu \rightarrow \epsilon,$$

$$\rho \rightarrow \rho_m, \rho_m \rightarrow -\rho,$$

$$E \rightarrow H, H \rightarrow -E, J \rightarrow M, M \rightarrow -E$$

and take the curl of the two curl equations making use of the two divergence equations. By applying the principle of superposition, ie, first set $\rho_m = 0, M = 0$, and then set $\rho = 0, J = 0$, show that the particular solution to Maxwell's equations due to the sources can be expressed as

$$E = -\nabla V - A_{,t} - \nabla \times F/\epsilon,$$

$$H = \nabla \times A/\mu - \nabla V_m - F_{,t}$$

where we may impose the Lorentz gauge conditions

$$\operatorname{div} A = -\epsilon \mu V_{,t}, \operatorname{div} F = -\epsilon \mu V_{m,t}$$

without affecting the em fields. Note that the correspondence is complete once we introduce

$$A \rightarrow F, F \rightarrow -A, V \rightarrow V_m, V_m \rightarrow -V$$

and then

$$\nabla^2 V - \epsilon \mu V_{,tt} = -\rho/\epsilon,$$

$$\nabla^2 A - \epsilon \mu A_{,tt} = -\mu J,$$

$$\nabla^2 V_m - \epsilon \mu V_{m,tt} = -\rho_m/\mu,$$

$$\nabla^2 F - \epsilon\mu F_{,tt} = -\epsilon M$$

[2] Continuing from the previous problem, write down in the frequency domain, the far field expressions for the electromagnetic field and hence the Poynting vector in terms of the source J, M near the origin. Deduce in terms of J, M the total power radiated by these sources. This is a fundamental result in antenna theory.

Hint: Consult the book: Constantine Balanis, "Antenna theory", Wiley.

1.10.2 Points to remember

[1] Starting with the Maxwell equations taking into account magnetic charges and current densities and applying the Lorentz gauge conditions on the electric vector, scalar and magnetic vector and scalar potentials, (which do not affect the electric and magnetic fields), we can derive the wave equation with sources for these potentials:

$$(\nabla^2 - \mu\epsilon\partial_t^2)\psi(t, r) = -s(t, r)$$

where the response and source pair (ψ, s) is any one of

$$(A, \mu J), (\Phi, \rho/\epsilon), (F, \epsilon M), (\Phi_m, \rho_m/\epsilon)$$

[2] Alternately, by manipulating the Maxwell equations directly by taking curl, we can derive 3-D wave equations for the electric and magnetic fields with sources appearing as the time derivatives of the electric and magnetic current densities and the spatial gradients of the electric and magnetic charge densities.

[3] Wave equations with sources have both a causal retarded potential solution and a non-causal advanced potential. For physical reasons, we choose the retarded potential solutions which lead to radiation phenomena by sources, ie, $1/r^2$ dependence of the Poynting power flux vector in the far field zone.

1.11 Solution to Maxwell's equations in terms of retarded potentials satisfying the wave equation with source

This has been discussed in the previous chapter.

1.11.1 Exercises

[1] Write down the energy of the electromagnetic field in the Coulomb gauge, ie, $\text{div} A = 0$ and deduce that the scalar potential V is a matter field, ie, it can be expressed as a Coulomb potential corresponding to the charge density ρ without any retardation and that the field energy after removing away the purely matter component can be expressed as

$$U = (1/2) \int (\epsilon_0 |A_{,t}|^2 + \mu_0^{-1} |\text{curl} A|^2) d^3r$$

By passing over to the spatial frequency domain show that this energy along with the equations of motion for the field can be regarded as a continuous superposition of harmonic oscillators in such a way that corresponding to each wave vector k , there are two oscillators polarized perpendicular to k .

[2] Derive the Maxwell-Dirac equations for a system of electrons and positrons interacting with an electromagnetic field:

$$(\partial_t^2 - c^2 \nabla^2) A^\mu = \mu_0 J^\mu,$$

$$J^\mu = -e\psi^* \gamma^0 \gamma^\mu \psi,$$

$$[\gamma^\mu (i\partial_\mu + eA_\mu) - m]\psi = 0$$

Solve this system approximately using perturbation theory.

1.11.2 Points to remember

[1] The retarded potential solution to the 3-D wave equation with source is obtained by Fourier transforming the equation w.r.t. the spatial variables, solving the resulting ode of second order using the forced harmonic oscillator solution, or equivalently via Laplace transforms, and then inverting the spatial Fourier transform to arrive at the retarded potential solution.

[2] If however, em waves are present within a closed boundary, then by applying boundary conditions of the fields on the boundary, we will not in general get the standard retarded potential solution that is valid if the boundary is the surface of an infinite sphere. Rather, our solution will be dictated by the Green's function for the Helmholtz equation with the appropriate boundary conditions. For this and more, see the book by J.D.Jackson, "Classical electrodynamics", Wiley.

1.12 The principle of superposition

In the Young's double slit experiment, if the first slit alone is open and electrons/photons pass through only the first slit, then on the screen, the maximum intensity is observed in front of the first slit with a decaying pattern away from it, while if the second slit alone is open, then maximum intensity is observed in front of slit two. Denote these two intensity patterns by $I_1(r)$ and $I_2(r)$ respectively. We expect therefore that if both the slits are open, we would then observe the intensity distribution $I_1(r) + I_2(r)$, ie, intensity maxima in front of both the slits with decaying intensity away from the slits. But this is not what is observed. In fact, one observes an intensity maximum at the point midway between the two slits and then local intensity maxima and minima, ie fringes away from this midpoint. In other words, we observe an interference pattern which can only be explained by the principle of superposition of waves: $\psi_1(r)$ is the complex amplitude pattern due to the first slit. This means that the wave amplitude is $Re(\psi(r)\exp(j\omega t))$ where the frequency ω is that of the monochromatic wave emitted by the laser or electron gun located behind the slits. Likewise, the wave amplitude due to the second slit is $Re(\psi_2(r)\exp(j\omega t))$. We now add these two amplitude patterns, then square them and take a time average to obtain the intensity distribution for the case when both the slits are open. This gives us the intensity distribution

$$\begin{aligned} I(r) &= \langle (Re(\psi_1(r)\exp(j\omega t)) + Re(\psi_2(r)\exp(j\omega t)))^2 \rangle \\ &= (1/2)|\psi_1(r) + \psi_2(r)|^2 = I_1(r) + I_2(r) + I_{12}(r) \end{aligned}$$

where

$$I_1(r) = |\psi_1(r)|^2/2, I_2(r) = |\psi_2(r)|^2/2, I_{12}(r) = Re(\bar{\psi}_1(r)\psi_2(r))$$

$I_1(r)$ is the intensity pattern on the screen if only the first slit is open, $I_2(r)$ is the intensity pattern on the screen if only the second slit is open and $I_{12}(r)$ is the interference term. Thus, the wave nature of light is justified by the presence of the interference term, likewise De-Broglie's wave-matter duality, supports the fact that electrons can also behave like waves with a wavelength $\lambda = h/p$ if p is the electron momentum and the proof of this wave property of the electron is supported by the interference pattern on the screen when we use an electron gun behind the two slits.

The principle of superposition in quantum mechanics states that if $\psi_k(r), k = 1, 2$ are two wave functions, then the superposition $c_1\psi_1(r) + c_2\psi_2(r)$ where c_1, c_2 are complex constants is also a wave function provided that $\int |\psi_k(r)|^2 d^3r = 1, k = 1, 2$ and c_1, c_2 are scaled so that $\int |c_1\psi_1(r) + c_2\psi_2(r)|^2 d^3r = 1$. In that case $|\psi_k(r)|^2$ is the probability density of the particle's position when its state is $\psi_k, k = 1, 2$ and $|c_1\psi_1(r) + c_2\psi_2(r)|^2$ is its probability density in the superposed state. This latter probability density can be expressed as $|c_1|^2|\psi_1(r)|^2 + |c_2|^2|\psi_2(r)|^2 + 2Re(\bar{c}_1c_2\bar{\psi}_1(r)\psi_2(r))$. Thus, in the superposed state, we do not get a weighted linear combination of the two probability densities but rather, an interference term is also present. This is the fundamental difference between classical and quantum probability.

1.12.1 Summary

When a wave source like a laser or an electron gun emits waves which hit a screen having slits in it, then these slits act as secondary sources which generate waves that interfere on another screen placed in front of the screen having slits. Young's double slit experiment showed an interference pattern on the screen which meant that one has to superpose the wave amplitudes from different slits and then take the intensity, ie, modulus square rather than simply add up the respective intensities. It also confirmed the wave-particle duality of matter postulated by De-Broglie. The diffraction pattern depends on the distance between the slits in Young's experiment and hence can be used to determine the structure of crystals as was done by William Bragg. Specifically, if the distance between two atoms/molecules in a crystal is d and there are N such molecules linearly arranged, then for em waves getting diffracted from these molecules at an angle θ relative to the line joining these molecules, the path difference on the screen corresponding to successive molecules will be $d \cos(\theta)$ and hence the phase difference will be $kd \cos(\theta)$. The total complex amplitude on the screen obtained by superposing all these waves will therefore be proportional to

$$A(\theta) = \sum_{m=0}^{N-1} \exp(jmd \cos(\theta)) = \frac{\exp(jNkd \cos(\theta)) - 1}{\exp(jkd \cos(\theta)) - 1}$$

and hence, the intensity distribution on the screen will be

$$|A(\theta)|^2 = \left[\frac{\sin((N-1)kd \cos(\theta)/2)}{\sin(kd \cos(\theta)/2)} \right]^2$$

This intensity distribution shows a global maximum at $\theta = \pi/2$ and zeros at $\theta = \cos^{-1}(2m\pi/(N-1)kd)$, $m = 1, 2, \dots, N-2$. Here, $k = 2\pi/\lambda$ where λ is the wavelength of the em radiation that is incident upon the crystal. Thus, the intermolecular distance d can be determined from the intensity distribution of the diffracted pattern. Here, the molecules in the crystal act as secondary sources.

The diffraction theory can be developed in full generality as shown in J.D.Jackson's book "Classical Electrodynamics", Wiley. Here, one considers a surface on which the em fields are specified and one solves the Helmholtz/3-D wave equation for the em fields non on the surface in terms of the surface fields by using the Green's function for the Helmholtz equation with appropriate boundary condition. The surface on which the em fields are specified acts as a secondary source here. This theory is a vector form of scalar diffraction theory. However, determining the Green's function for the Helmholtz operator corresponding to boundary conditions on an arbitrary surface is very hard in general and closed forms are not available. Thus, Jackson uses the approximate Green's function, namely, the Green's function $-\exp(jk|r-r'|)/4\pi|r-r'|$ for the Helmholtz operator in free space, ie with vanishing boundary conditions on a sphere of infinite radius to obtain formulas for the diffraction pattern. Further, given a surface on which the em fields are specified, the pattern calculated using this approximate Green's function can be simplified in the far field zone from the surface

where quadratic terms in the coordinates of the surface will not appear. In other words, one gets only linear terms implying that the diffraction amplitude pattern in the far field zone will be proportional to the spatial Fourier transform of the amplitude distribution on the surface. This is known as Fraunhofer diffraction. On the other hand, in the near field zone, the Green's function will involve quadratic terms of the surface coordinates apart from linear terms in the exponential appearing in the approximate Green's function. The resulting near field pattern is known as Fresnel diffraction. We explain all this with full mathematical rigor in what follows.

1.12.2 Exercises

[1] Let $\{|e_k\rangle: 1 \leq k \leq N\}$ be an onb for a Hilbert space \mathcal{H}_1 and $\{|f_m\rangle: 1 \leq m \leq M\}$ an onb for another Hilbert space \mathcal{H}_2 . Thus, $\{|e_k \otimes f_m\rangle: 1 \leq k \leq N, 1 \leq m \leq M\}$ is an onb for $\mathcal{H}_1 \otimes \mathcal{H}_2$. Take any pure state $|\psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$. Express it as the superposition

$$|\psi\rangle = \sum_{k,m} c(k, m) |e_k \otimes f_m\rangle$$

Assuming that this is the joint state of two quantum systems, calculate the mixed state $\rho_k, k = 1, 2$ of the two systems, namely,

$$\rho_1 = Tr_2(|\psi\rangle\langle\psi|), \rho_2 = Tr_1(|\psi\rangle\langle\psi|)$$

Show that ρ_1 and ρ_2 have the same Von-Neumann entropies. Hint: Perform the SVD of the matrix $(c(k, m))$ and express $|\psi\rangle$ as

$$|\psi\rangle = \sum_k \lambda_k |e_k \otimes g_k\rangle$$

where $\{|g_k\rangle\}$ is an on set in \mathcal{H}_2 . Now suppose we take two pure states $|\psi\rangle$ and $|\phi\rangle$ in $\mathcal{H}_1 \otimes \mathcal{H}_2$. What are the mixed states induced in \mathcal{H}_1 and \mathcal{H}_2 by the superposition $c_1|\psi\rangle + c_2|\phi\rangle$ in terms of those induced by $|\psi\rangle$ and $|\phi\rangle$ and an interference term? When does the interference term vanish? Generalize this to the case of p pure states $|\psi_k\rangle, k = 1, 2, \dots, p$ in $\mathcal{H}_1 \otimes \mathcal{H}_2$ by expressing $Tr_m(|\psi\rangle\langle\psi|), m = 1, 2$ in terms of $Tr_m(|\psi_k\rangle\langle\psi_k|), k = 1, 2, \dots, p$ and interference terms $Tr_m(|\psi_k\rangle\langle\psi_r|), m \neq r$ where $|\psi\rangle = \sum_{k=1}^p c(k)|\psi_k\rangle$.

Remark: This exercise shows that the principle of superposition in quantum mechanics holds only for pure states, not for mixed states and it tells us how to calculate the interference terms for mixed states.

1.12.3 Points to remember

[1] The principle of superposition is at the heart of all wave phenomena including light, electromagnetic waves at arbitrary frequency and quantum mechanics. It

states that when waves interfere, we must linearly superpose them and to obtain the resulting intensity, we must take the modulus square of this superposition.

[2] In quantum mechanics, the solution to Schrodinger's equation is a superposition of its energy eigenstates with the coefficients varying harmonically with time. The probability amplitude of getting the system in a certain state, is obtained by taking the inner product of this state with the superposition and the corresponding probability is obtained by forming the modulus square of this probability amplitude.

[3] The principle of superposition is the basis of Feynman's path integral approach to quantum mechanics. It states that along a given path there is a probability amplitude to go from point 1 to point 2 in a given time. This amplitude is given by a pure phase. The total amplitude to go from the first to the second point in the given time t written as $K_t(r_2|r_1)$ is obtained by summing the complex probability amplitudes over all possible paths and this total amplitude after normalization satisfies the Schrodinger wave equation. The path integral method can also be applied to derive all the major results of quantum field theory like the Feynman diagrams for the scattering matrix involving interactions of electrons, positrons and photons.

1.13 Diffraction and interference of waves

1.13.1 General theory

Feynman showed that if $S(r(t), t_1 \leq t \leq t_2)$ is the action functional for a non-relativistic quantum particle to go from $r(t_1) = r_1$ upto $r(t_2) = r_2$ between the times t_1 and t_2 along the path $r(.)$, then after an appropriate normalization (by an infinite constant, the total quantum mechanical amplitude to go from (t_1, r_1) upto (t_2, r_2) can be calculated by a path integral of the form

$$K(t_2, r_2|t_1, r_1) = C \int \exp(iS(r(t), t_1 \leq t \leq t_2)/\hbar) \Pi_{t_1 < t < t_2} dr(t)$$

This means that $|K(t_2, r_2|t_1, r_1)|^2$ is the probability density at r_2 of the particle at time t_2 given that at time t_1 it is located at r_1 . In other words, $K(t_2, r_2|t_1, r_1)$ satisfies Schrodinger's wave equation w.r.t. the variables (t_2, r_2) with the initial condition $K(t_1, r_2|t_1, r_1) = \delta^3(r_2 - r_1)$. Thus, by superposing pure phase terms, we can compute quantum transition probabilities after taking the modulus square. The modulus square tells us that there are interference terms between any two paths and we simply cannot add up the probabilities along each path to get the total transition probability. We have to add complex amplitudes along each path and then take the modulus square in order to get the total probability. The role played by the intensity distribution on the screen in Young's double slit experiment is played by probability in quantum mechanics.

The electric field in free space satisfies the wave/Helmholtz equation

$$(\nabla^2 + k^2)E = 0, \operatorname{div} E = 0$$

Likewise, the magnetic field satisfies

$$(\nabla^2 + k^2)H = 0, \operatorname{div} H = 0$$

These equations are subject to boundary conditions that on the boundary of the surface, the tangential component of the electric field and the normal component of the magnetic field vanish in the case when the boundary is perfect electric conductor. More generally, if the surface electric and magnetic fields are prescribed, then we wish to solve for the electric and magnetic fields inside the region in terms of their surface values. Applying Green's theorem to the volume within the surface gives us assuming that the surface electric field is E_s and surface magnetic field is H_s ,

$$\begin{aligned} & \int_V (E_m(r) \nabla^2 G_{ms}(r, r') - G_{ms}(r, r') \nabla^2 E_m(r)) d^3 r \\ &= \int_{\partial V} (E_m(r) \partial_n G_{ms}(r, r') - G_{ms}(r, r') \partial_n E_m(r)) dS(r) \end{aligned}$$

where $G(r, r')$ is the 3×3 matrix valued Green's function of Helmholtz equation for the boundary, ie,

$$(\nabla^2 + k^2)G_{ms}(r, r') = \delta_{ms}\delta^3(r - r'), \partial_{x'_s} G_{ms}(r, r') = 0$$

and

$$G_{ms}(r, r') = 0, r \in \partial V$$

where summation over the repeated index $m, s = 1, 2, 3$ is implied. This gives us

$$\begin{aligned} & \int_V [E_m(r)(-k^2 G_{ms}(r, r') + \delta_{ms}\delta^3(r - r') - G_{ms}(r, r')(-k^2 E_m(r))] d^3 r = \\ & \int_{\partial V} (E_m(r) \partial_n G_{ms}(r, r') - G_{ms}(r, r') \partial_n E_m(r)) dS(r) \end{aligned}$$

or equivalently, after making the appropriate cancellations,

$$E_s(r') = - \int_{\partial V} E_m(r) \partial_n G_{ms}(r, r') dS(r)$$

This is the fundamental equation of diffraction: The electric field in the interior of V is determined by its values on the bounding surface. We also note the consistency of this formula:

$$\operatorname{div} E(r') = E_{s,s}(r') = - \int_{\partial V} E_m(r) \partial_n \partial_{x'_s} G_{ms}(r, r') dS(r) = 0$$

since by construction of the vector valued Green's function,

$$\partial_{x'_s} G_{ms}(r, r') = 0$$

Note that ∂_n denotes the normal component of the derivative at the surface point w.r.t the variable r which falls on the surface.

1.13.2 Points to remember

[1] The general theory of diffraction involves considering a surface on which an em wave field is incident and then calculating the equivalent surface electric and magnetic current densities on this surface and finally, from these surface current densities, evaluate the response fields produced by them using Green's functions for the given boundary.

[2] Diffraction theory can be used to determine the intermolecular spacing in a crystal by shining an em wavefield on the crystal and observing the diffraction pattern and recalling that the path difference between two diffracted rays coming from two molecules having a separation of d is $d \cos(\theta)$ where θ is the angle between the diffracted rays and the intermolecular displacement vector. When the path difference is an integer multiple of the wave length then the corresponding phase difference is an integer multiple of 2π and the interference is constructive producing an intensity maximum. On the other hand, when the path difference is a half integer multiple of the wavelength, the corresponding phase difference is an odd integer multiple of π producing an intensive minimum (destructive interference).

[3] If an electron/photon gun is placed behind a screen S_1 having two slits separated by a distance of d and a screen S_2 is placed in front of this two slit screed parallelly at some distance from it, then if the first slit is blocked, an intensity maximum on S_2 is observed at a point directly in front of the second slit while if the second slit is blocked, an intensity maximum is observed in front of the first slit. If both the slits are now opened we do not observe intensity maxima in front of both the slits but rather midway between the two slits with a sinusoidally decaying intensity pattern. This effect explain the interference of waves, intensities do not add up only amplitudes add up.

1.14 Green's function for wave equation with sources—Fraunhofer and Fresnel's diffraction

This has already been discussed above but an elementary explanation of these effects would proceed along the following lines. Assume that an em field falls on an aperture D in the xy plane. The surface electric and magnetic current densities on this plane can be expressed as

$$J_s(\omega, x, y) = \hat{z} \times (H_x(\omega, x, y, 0)\hat{x} + H_y(\omega, x, y, 0)\hat{y})$$

and

$$M_s(\omega, x, y) = -\hat{z}(E_x(\omega, x, y, 0)\hat{x} + H_y(\omega, x, y, 0)\hat{y})$$

respectively. The magnetic current density M and the magnetic charge density ρ_m is an artifice introduced into the Maxwell equations to make it symmetric between the electric and magnetic fields. Thus, we can write the magnetic vector potential and the electric vector potential respectively as

$$A(\omega, x, y, z) = (\mu/4\pi) \int_D J_s(\omega, x', y') G(\omega, x, y, z|x', y', 0) dx' dy',$$

$$F(\omega, x, y, z) = (\epsilon/4\pi) \int_D M_s(\omega, x', y') G(\omega, x, y, z|x', y', 0) dx' dy'$$

Applying the Lorentz gauge conditions, gives us the electric scalar potential Φ and the magnetic scalar potential Φ_m as

$$-j\omega\epsilon\mu\Phi = \operatorname{div} A, -j\omega\mu\epsilon\Phi_m = \operatorname{div} F$$

and hence the electric and magnetic fields radiated out from the aperture by diffraction are

$$E(\omega, x, y, z) = -\nabla\Phi - j\omega A + \nabla \times F/\epsilon,$$

$$H(\omega, x, y, z) = -\nabla\Phi_m - j\omega F + \nabla \times A/\mu$$

The Fraunhofer far field and Fresnel near field diffraction patterns can now be computed by substituting for the Greens function G the usual one in free space $\exp(-jk|r - r'|)/|r - r'|$ and making appropriate approximations.

1.14.1 Exercises

[1] Calculate the far field and the near field wave patterns for the anisotropic Helmholtz equation

$$(\nabla^T A \nabla + k^2)\psi(r) = 0$$

where A is a 3×3 real positive definite matrix.

[2] Let $O(r) = A_O \exp(j\phi_O(r))$ denote the object wave field, $R(r) = A_R \exp(j\phi_R(r))$ denote the reference source wave. These two waves interfere and the corresponding intensity pattern

$$I(r) = |A_O \exp(j\phi_O(r)) + A_R \exp(j\phi_R(r))|^2$$

is recorded on a photographic plate so that where the intensity is higher, there is more blackening and where it is less, the blackening is lesser. Show that this is equivalent to recording the pattern $\cos(\phi_O(r) - \phi_R(r))$. Now shine the reference source on the photographic plate after this recording and note that the outgoing wave field has the amplitude pattern

$$\exp(j\phi_R(r)) \cos(\phi_O(r) - \phi_R(r))$$

Show that the phase factor $\exp(j(2\phi_R(r) - \phi_O(r)))$ can be filtered out using a lowpass spatial filter and what we are left with is simply something proportional to $\exp(j\phi_O(r))$, namely the original object amplitude field pattern. In other words, by making the object field interfere with a reference field we are able to record the object field's phase information and hence recover the original 3-D object. This is precisely the science of holography for which Gabor got the Nobel prize. Holography is a method for recording a 3-D object without losing depth information.

[3] Using perturbation theory for differential equations, calculate approximately the far field and near field amplitude and power radiation patterns for the inhomogeneous Helmholtz equation

$$(\nabla^2 + k^2(1 + \delta\chi(r)))\psi(r) = J(r)$$

If $\chi(r)$ is a random field, calculate the correlations in the near and far field amplitude patterns. Now repeat this for the case

$$(\nabla^T(I + \delta K(r))\nabla + k^2(1 + \delta\chi(r)))\psi(r) = J(r)$$

where $K(r)$ is a 3×3 matrix valued function of position.

1.14.2 Points to remember

[1] The Fraunhofer and Fresnel theories of diffraction can be described in a much more abstract setup namely when we have fields satisfying a general d dimensional system of Maxwell's equations. This system involves formulating the Maxwell equations in tensor form as

$$F_{\mu\nu} = A_{\nu,\mu}(x - A_{\mu,\nu}(x), \mu, \nu = 1, 2, \dots, d)$$

Introduce a constant metric $\eta^{\mu\nu}$ and set up the Maxwell equations as

$$F_{,\nu}^{\mu\nu}(x) = -\mu_0 J^\mu(x), F^{\mu\nu} = \eta^{\mu\rho}\eta^{\nu\sigma}F_{\rho\sigma}$$

and then impose the gauge condition

$$A_{,\mu}^\mu = 0$$

to arrive at the generalized wave equation for $A^\mu = \eta^{\mu\nu}A_\nu$:

$$\eta^{\alpha\beta}\partial_\alpha\partial_\beta A^\mu = -\mu_0 J^\mu$$

We determine the Green's function $G(x - x')$ for this wave operator satisfying the appropriate boundary conditions and hence solve:

$$A^\mu(x) = -\mu_0 \int G(x - x')J^\mu(x')d^d x'$$

To specialize this result to diffraction theory, we observe that J^μ will be replaced by a surface current density on a surface induced by an incident d -dimensional em field and then the above integral will become a surface integral over this surface, then we will approximate the integral

$$A^\mu(x) = -\mu_0 \int_S G(x - x') J^\mu(x') dS(x')$$

in the near field and far field zones by studying the behaviour of $G(x)$ in these two zones. The near field zone is the Fresnel zone and the far field zone is the Fraunhofer zone. Typically $x^1 = t$ will be the time coordinate and x^2, \dots, x^d will be the $d - 1$ spatial coordinates and we will look at the Fourier transforms of $A^\mu, G(x), J^\mu(x)$ w.r.t the time coordinate and then make a spatial analysis at each frequency.

1.15 The basic Eikonal equation of geometric optics

1.15.1 General theory

The Eikonal equation is an approximation to the wave equation but is useful when we are only interested in the propagation of wave fronts, ie, constant phase surfaces not on the amplitude of the wave. It is a first order equation in the spatial arguments and is of quadratic algebraic degree. Although we have derived it from the standard three dimensional wave equation or equivalently from the standard three dimensional Helmholtz equation that describe propagation at a definite frequency, it is applicable to even more general situations like Helmholtz equations of the form

$$\sum_{\alpha, \beta=1}^n a(\alpha, \beta, x) \frac{\partial^2 \psi(x)}{\partial x^\alpha \partial x^\beta} + k^2(x) \psi(x) = 0$$

which describe wave propagation in inhomogeneous and anisotropic media. The Eikonal equation involves what is called the geometric optics approximation in which the wave number is very large or equivalently, the wave length is much smaller than the dimensions of the medium in which it propagates. One can improve upon this approximation by using higher order perturbation theory but then the problem of investigating the convergence of the perturbation series would remain. In what follows, we present first the Eikonal equation as an approximate nonlinear partial differential equation for the phase of the wave and then present the nonlinear Schrodinger wave equation obtained by assuming that in a given direction the dimensions of the medium are much larger than the corresponding wavelength of the wave. This is some sort of restricted geometric optics approximation. The second order partial derivative terms appear only for the transverse variables x, y while along the transverse directions x, y , first order partial derivatives with Eikonal like quadratic terms appear along with linear terms. Thus one gets a Schrodinger equation with a nonlinearity if the z axis is

interpreted as time. Alternately, we may have a geometric optics approximation in the transverse directions but not in the longitudinal directions. Then, we get another version of the nonlinear Schrodinger equation with quadratic terms in the first order partial derivative along the longitudinal direction and along the transverse direction, we have only linear terms in the second order transverse partial derivatives. It is an interesting exercise to generalize such an idea to $n > 3$ dimensions where geometric optics approximation holds along the first p co-ordinate directions and along the remaining $n - p$ coordinate directions linear terms in the second order partial differentials of the wavefront appear.

The Helmholtz equation for a wave field is obtained by Fourier transforming the wave equation w.r.t the time argument. It is given by

$$(\nabla^2 + k^2)\psi(r) = 0 \quad \dots \quad (1)$$

where $\psi(r) = \psi(x, y, z) = \hat{\psi}(r, k)$ is defined by

$$\int_{\mathbb{R}} \psi(r, t) \exp(-i\omega t) dt, k = \omega/c$$

and $\psi(r, t)$ satisfies

$$(c^2 \nabla^2 - \partial_t^2)\psi(r, t) = 0$$

We now substitute

$$\psi(r) = \exp(iS(r))$$

into (1). Evaluating the partial derivatives,

$$\nabla \psi = i \nabla S \exp(iS),$$

$$\nabla^2 \psi = (-|\nabla S|^2 + i \nabla^2 S) \exp(iS)$$

so our Helmholtz equation transforms to a nonlinear pde for the phase field S .

$$|\nabla S|^2 - k^2 - i \nabla^2 S = 0 \quad \dots \quad (1)$$

Now, if we assume that propagation takes place only along the x direction, then $-S''(x)/S'(x)^2 = d/dx(1/S')$ and S' is the local wave number which is proportional to the reciprocal of the wavelength. Its rate of change with x is small. Equivalently, writing $S(r) = k.r + \delta S(r)$, we find the exact equation

$$|k + \nabla \delta S(r)|^2 - k^2 - i \nabla^2 \delta S(r) = 0$$

or equivalently,

$$|\nabla \delta S(r)|^2 + 2k \cdot \nabla \delta S(r) = i \nabla^2 \delta S(r)$$

so we have approximately upto first order terms,

$$i \nabla^2 \delta S(r) = 2k \cdot \nabla \delta S(r)$$

and hence, approximately,

$$|k + \nabla\delta S(r)|^2 - k^2 - 2k \cdot \nabla\delta S(r) = 0$$

which can be expressed as

$$|\nabla S(r)|^2 = k^2 + 2k \cdot \nabla\delta S(r)$$

and if we neglect the last term being of the first order of smallness compared to the other terms, we get the Eikonal equation

$$|\nabla S(r)|^2 = k^2$$

Note that

$$|k \cdot \nabla\delta S(r)|/k^2 \approx |\nabla\delta S(r)|/|k|$$

which is very small for small values of the wavelength compared with the dimensions of the region where the wave propagates, ie in geometrical optics. There is another approximation to the wave equation known as the nonlinear Schrodinger equation. Suppose we decompose the wave motion into propagation in the xy plane and along the z direction. Then, Helmholtz equation can be expressed as

$$(\nabla_{\perp}^2 + \partial_z^2 + k^2)\psi(x, y, z) = 0$$

where

$$\nabla_{\perp}^2 = \partial_x^2 + \partial_y^2$$

Now writing

$$\psi(x, y, z) = \exp(iS(r)) = \exp(i(S_0(x, y, z) + k_z z))$$

ie,

$$S(r) = S(x, y, z) = S_0(x, y, z) + k_z z = S_0(r) + k_z z$$

and substituting this into (1) gives us

$$|\nabla S_0 + k_z \hat{z}|^2 - k^2 - i\nabla^2 S_0 = 0$$

This equation is exact and expands to

$$|\nabla S_0|^2 + k_z^2 + 2k_z \partial_z S_0 - k^2 = i\nabla^2 S_0$$

In the special case when $k_z = k$ and $|\partial_z^2 S_0|/k|\partial_z S_0|$ is very small and so also are $|\partial_z^2 S_0|/|\partial_z S_0|^2$ the above equation simplifies to the non-linear Schrodinger equation (with z being interpreted as the time variable and x, y as the spatial variables):

$$i\partial_z S_0 = (-1/2k)\nabla_{\perp}^2 S_0 - (i/2k)|\nabla_{\perp} S_0|^2$$

This is the celebrated non-linear Schrodinger equation. The approximation involved here is that the dimensions of the region in a particular direction are much larger than the wavelength.

1.15.2 Points to remember

[1] When we substitute a pure phase factor $\exp(iS(r))$ into the three dimensional Helmholtz equation (which is a time Fourier transformed version of the wave equation, or equivalently the wave equation at a fixed frequency) and make the geometric optics approximation, namely that the dimensions of the region in which the wave propagates are much larger than the wavelength then second order partial derivatives of the phase S can be neglected and we get the Eikonal equation $|\nabla S|^2 = k^2$. where $\omega = kc$ is the frequency. The surfaces of constant phase are $S(r) - \omega t = \text{const}$ from which, we deduce that the local wave velocity, ie, the velocity of the wavefronts is given by $v = dr/dt$, where

$$S(r + dr) - \omega(t + dt) = S(r) - \omega t$$

or equivalently,

$$dr \cdot \nabla S(r) - \omega dt = 0$$

ie,

$$v \cdot \nabla S(r) = \omega = kc$$

Since the velocity of the wave is directed along the normal to the wavefronts, we can deduce that

$$|v| = c, |\nabla S(r)| = k$$

which is precisely the Eikonal equation. Thus, the Eikonal equation has a natural interpretation in terms of the velocity of wavefronts or equivalently, of constant phase surfaces.

[2] The nonlinear Schrodinger equation is derived from the wave equation by assuming that in certain spatial directions, though not all, the geometric optics approximation is valid.

[3] The approximate refractive index of the medium can be expressed as $n(r) = c/|v| = |\nabla S(r)|/k$ in a medium in which the waves travel at lesser than the speed c of light. In that case, the Eikonal equations has to be modified accordingly.

[4] The Eikonal equation can also be arrived at from Schrodinger's wave equation. Specifically, writing the wave function at energy E as $\psi(r) = \exp(iS(r)/\hbar)$, we get on substituting into the Schrodinger equation

$$(-\hbar^2/2m)\nabla^2\psi(r) + V(r)\psi(r) = E\psi(r)$$

that

$$(-\hbar^2/2m)(i\nabla^2S/\hbar - |\nabla S|^2/\hbar^2) = (E - V(r))$$

and neglecting the $O(1/\hbar)$ term in comparison with the $O(1/\hbar^2)$ term gives us the Eikonal equation for matter waves:

$$|\nabla S(r)|^2 = 2m(E - V(r)) = p(r)^2$$

where $p(r)$ is the momentum of a nonrelativistic particle having energy E and moving in the potential $V(r)$:

$$p(r)^2/2m + V(r) = E$$

[5] The Eikonal approximation involving neglecting partial derivatives of second order can be obtained for more general wave equations, for example, for wave equations in curved space-time having the form

$$\sum_{k,m=1}^N a_{km}(x)\partial_k\partial_m\psi(x) + \sum_{k=1}^N b_k(x)\partial_k\psi(x) + k^2(x)\psi(x) = 0$$

1.16 Describing the trajectory of light in a medium having spatially varying refractive index

Fermat's principle of minimum time requires that the trajectory $s \rightarrow r(s)$ of the light ray in a medium having spatially varying refractive index $n(r)$ be such that the variational principle

$$\delta \int_1^2 |dr(s)|n(r(s))/c_0 = 0$$

be satisfied or equivalently,

$$\delta \int_1^2 |r'(s)|n(r(s))ds = \delta \int \sqrt{x'^2(s) + y'^2(s) + z'^2(s)}n(x(s), y(s), z(s))ds = 0$$

be satisfied.

1.16.1 Exercises

[1] Write down the Euler-Lagrange equations for the trajectory for this problem. Note that we keep the end points fixed, ie, we minimize the travel time of the light ray between two fixed spatial points.

[2] Show that if light in a constant refractive index medium is assumed to travel along a straight line, then Fermat's principle of minimum time leads to Snell's law of refraction when light travels from a medium of constant refractive index n_1 to another medium of constant refractive index n_2 assuming that the boundary between the two media is a plane.

1.17 Propagation of light in anisotropic, inhomogeneous and time varying medium

1.17.1 General theory

If the refractive index along a given direction \hat{m} is expressible as $n(\hat{m}, r, t)$, then the variational problem of determining the trajectory of a photon from its trajectory $r(t)$ would involve writing the Euler-Lagrange equations resulting from the variational principle

$$\delta \int_1^2 |r'(t)| n(r'(t)/|r'(t)|, r(t), t) dt = 0$$

It should be noted that the Euler-Lagrange equations should be solved by taking into account the first integral

$$|r'(t)| = c/n(r'(t)/|r'(t)|, r(t), t)$$

Exercise: If in addition the refractive index is a random field, then by assuming the refractive index to be a small perturbation of a constant isotropic refractive index, use perturbation theory to calculate approximately the equation of the trajectory.

1.17.2 Points to remember

- [1] When a light ray travels through a medium having inhomogeneous and anisotropic refractive index, we can formulate the problem of obtaining its trajectory as a variational principle according to Fermat's principle of least time. The resulting Euler-Lagrange equations give us the trajectory of the light ray as a functional of the refractive index field. Thus, if we have a bunch of light sources that emit light rays in different directions, by measuring their trajectories, we can estimate the refractive index field. A simpler method, is to expand the refractive index field in terms of test functions of the spatial position and estimate the coefficients in this expansion by matching the trajectory values at a set of discrete points using a least squares approach.

1.18 The Schrodinger wave equation in quantum mechanics

1.18.1 General theory

The Schrodinger wave equation derivation is based on two fundamental principles introduced by Max Planck and Louis-De-Broglie. The first states that the energy of a photon of frequency ω is $E = h\omega$ where h equals Planck's constant divided by 2π . The second states that one can associate a wave of wavelength $\lambda = 2\pi h/p$ with a particle of momentum p . Thus since the wave vector is $k = (2\pi/\lambda)\hat{n}$ where \hat{n} is the unit vector along the direction of wave propagation, one can write down for a plane wave

$$\psi(t, r) = \exp(-i(\omega t - k \cdot r))$$

the following

$$ih\partial\psi(t, r)/\partial t = h\omega\psi(t, r) = E\psi(t, r),$$

$$-(h^2/2m)\nabla^2\psi(t, r) = (h^2k^2/2m)\psi(t, r) = (p^2/2m)\psi(t, r) = (E - V)\psi(t, r)$$

where $p^2/2m$ is the kinetic energy of the particle associated with this wave, E the total energy and V the potential energy. Thus, combining these two identities gives us the wave equation

$$(-(h^2/2m)\nabla^2 + V(t, r))\psi(t, r) = E\psi(t, r) = ih\partial\psi(t, r)/\partial t$$

This is known as the Schrodinger wave equation and is at the heart of all quantum mechanical calculations like determining the energy spectrum of atoms and molecules, determining the probability distribution of particles as a function of time given its initial state etc. One does not generally have a plane wave satisfying the Schrodinger wave equation although that was what we started with. The Schrodinger wave equation describes matter waves, ie, waves associated with quantum particles. Only when we combine it with Born's hypothesis that the modulus square of the wave function should be interpreted as the probability density of the particle to be in a certain region of space at a given time does it give a good physical interpretation. The Schrodinger wave equation unlike the usual wave equation is second order in the spatial derivatives but only first order in time and there is a potential term too in it. This asymmetry between time and space is a consequence of using the non-relativistic relation between kinetic energy and momentum: $E = p^2/2m$. If we use Einstein's energy-momentum relation $(E - V)^2 - p^2c^2 - m^2c^4 = 0$, we get a wave equation that is second order in both the spatial and time derivatives:

$$(ih\partial_t - V(t, r))^2 + h^2c^2\nabla^2 - m^2c^4\psi(t, r) = 0$$

This is called the Klein-Gordon (KG) equation. Further, if there is a magnetic vector potential $A(t, r)$ apart from the scalar potential V , the Schrodinger and KG equations get modified respectively to

$$(ih\partial_t + eV(t, r))\psi + (h^2/2m)(\nabla + ieA(t, r)/h)^2\psi = 0,$$

and

$$(ih\partial_t + eV(t, r)) + h^2 c^2 (\nabla + ieA(t, r)/h)^2 - m^2 c^4 \psi(t, r) = 0$$

These equations are for a charged particle of charge $-e$. They follow from classical mechanics that in the presence of an em field with scalar potential V and magnetic vector potential A , the kinetic energy is $E + eV$ and the velocity of the particle is given by $mv = p + eA$. These equations follow from the Lagrangian

$$L(t, r, v) = mv^2/2 + e(V - (v, A))v = dr/dt$$

which yields the correct Lorentz equation of motion

$$mdv/dt = -e(E + v \times B), E = -\nabla V - A_{,t}, B = \nabla \times A$$

The momentum is calculated as

$$p = \partial L / \partial v = mv - A$$

This is in the non-relativistic case. In the relativistic case, the Lagrangian is

$$L(t, r, v) = -mc^2 \sqrt{1 - v^2/c^2} + e(V - (v, A))$$

giving for the momentum

$$p = \partial L / \partial v = \frac{mv}{\sqrt{1 - v^2/c^2}} - eA$$

and the special relativistic Lorentz equation of motion

$$\frac{d}{dt} \frac{mv}{\sqrt{1 - v^2/c^2}} = -e(E + v \times B)$$

The fact that the Schrodinger equation is first order in time yields conservation of probability since the semigroup generated by the Schrodinger evolution is unitary:

$$iU'(t) = H(t)U(t)$$

with $H(t)^* = H(t)$ implies $U(t)^*U(t) = I$ provided that $U(0)^*U(0) = I$. The KG equation does not give a unitary solution in general. The situation is just like this:

$$dx(t)/dt = iax(t)$$

implies

$$x(t) = \exp(iat)x(0), |x(t)| = |x(0)|, t \geq 0$$

while

$$d^2x(t)/dt^2 = -a^2x(t)$$

implies

$$x(t) = \exp(iat)\alpha + \exp(-iat)\beta$$

implies $|x(t)|$ is generally not a constant. Every solution of the former is a solution of the latter but not vice-versa. Owing to the asymmetry between space and time, Schrodinger's equation is unsuitable for special relativistic calculations and owing to the existence of non-unitary solutions of the KG equation, the KG equation is unsuitable for doing any quantum mechanical calculation. Dirac found the way out of this difficulty by factoring using 4×4 anticommuting matrices $\alpha_r, r = 1, 2, 3, \beta$ the quadratic form

$$E^2 - c^2(p_1^2 + p_2^2 + p_3^2) - m^2c^4$$

as

$$(E + c(\alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3) + \beta mc^2)(E - c(\alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3)) - \beta mc^2)$$

and taking his wave equation in the absence of em fields as

$$(E - c(\alpha, p) - \beta mc^2)\psi = 0$$

or in the presence of em fields as

$$(E + ev - c(\alpha, p + eA) - \beta mc^2)\psi = 0$$

With $E = ih\partial_t, p = -ih\nabla$, this becomes a linear pde first order in both the space and time variables and yields non only Lorentz covariance under an appropriately chosen representation of the Lorentz group but also unitarity of the generated semigroup of evolutions. The free Dirac wave field for a particle of mass m also consists of plane waves but with the relationship between wave number and frequency being $h\omega = \pm c\sqrt{m^2c^2 + h^2k^2}$ rather than $\omega = kc$ for mass zero particles like the photon. The same is true for the free KG equation. However, the negative energy states that appear while solving the free Dirac equation correspond to a new particle called the positron which is the antiparticle of the electron. The amplitudes of the KG waves are scalar functions of wave number ie spinless particle however, the amplitudes of the Dirac waves are \mathbb{C}^4 valued functions of the wavenumber, ie three momentum and for each three momentum vector there are four such amplitude vectors. Two for the electron and two for the positron, two being attributed to the fact that electrons and positrons have spin 1/2. More precisely, the solutions of the KG equation

$$(\nabla^2 - (1/c^2)\partial_t^2 - \mu^2)\psi(t, r) = 0$$

can be expressed as a superposition of plane waves:

$$\psi(t, r) = \int [\hat{\psi}_1(K)exp(-i(\omega(K)t - K.r)) + \hat{\psi}_2(K)exp(i(\omega(K)t - K.r))]d^3K$$

where

$$\omega(K) = c\sqrt{K^2 + \mu^2}$$

where

$$\mu = mc^2/h$$

m being the mass of the KG particle. For a given wave number, ie, wave vector, there are only two complex scalar amplitudes $\hat{\psi}_m(K), m = 1, 2$ and if in particular, we put the restriction that the KG field is real, then $\hat{\psi}_2(K) = \hat{\psi}_1(K)^*$. When we however, expand the solutions of the free Dirac equation

$$(c(\alpha, -ih\nabla) + \beta mc^2)\psi(x) = ih\partial_t\psi(x)$$

we get as solutions

$$\psi(x) = \sum_{\sigma=1,2} \int [a(K, \sigma)u(K, \sigma)\exp(-i(\omega(K)t-K.r)) + b(K, \sigma)v(K, \sigma)\exp(i(\omega(K)t-K.r))]d^3K$$

where $a(K, \sigma), b(K, \sigma)$ are complex numbers (which become field operators in the second quantization of the Dirac field) and $u(K, \sigma), v(K, \sigma)$ must necessarily satisfy the algebraic eigenvalue equations

$$\hbar\omega(K)u(K, \sigma) = (hc(\alpha, K) + \beta mc^2)u(K, \sigma), \sigma = 1, 2, K \in \mathbb{R}^3$$

$$-\hbar\omega(K)v(K, \sigma) = (-hc(\alpha, K) + \beta mc^2)v(K, \sigma), \sigma = 1, 2, K \in \mathbb{R}^3$$

The second equation is equivalent to

$$-\hbar\omega(K)v(-K, \sigma) = (hc(\alpha, K) + \beta mc^2)v(-K, \sigma), \sigma = 1, 2, K \in \mathbb{R}^3$$

These equations mean that there are electrons which occupy positive energy states and also electrons which occupy negative energy states, the latter meaning that if we pull out an electron from a negative energy state, we are left with a positive energy hole, which corresponds to the antiparticle of the electron, namely the positron.

1.18.2 Points to remember

[1] Given a relation between the energy of a particle and its momentum, we can construct a wave equation in the quantum theory by replacing the energy by the operator $ih\partial_t$ and the momentum vector operator by $-ih\nabla$. If there is in addition a charge on the particle and a scalar potential V and a magnetic vector potential A , then as in classical mechanics and special relativity, we replace the kinetic energy operator by $ih\partial_t + eV$ and the momentum operator by $-ih\nabla + eA$. Special cases of this give us Schrodinger's equation for non-relativistic mechanics and the Klein-Gordon equation for relativistic quantum mechanics. The above operator corresponds for E and p are respectively obtained by considering a plane wave and applying Planck's quantum hypothesis relating the energy to frequency and De-Broglie's hypothesis relating the momentum to wavelength of a particle. The Schrodinger equation reduces the problem of calculating the energy levels of an atom to obtaining the eigenvalues of a partial differential operator and further Born's hypothesis gives us a nice physical interpretation of the wave function of a particle in terms of probability and current.

[2] The Schrodinger equation is asymmetric in time and space and hence cannot be used as a relativistic wave equation. The Klein-Gordon equation on the other hand is symmetric in time and space but it being second order in time, cannot be used as a valid quantum mechanical wave equation since equations that are second order in time will generally admit in addition to unitary evolution solutions, non-unitary solutions which are unacceptable as they violate the law of total probability conservation. Hence, Dirac factored the energy-momentum relation of special relativity into the product of two linear factors having 4×4 matrix coefficients and thereby obtained the correct description of relativistic quantum mechanics which apart from being Lorentz invariant, implies the KG equation and also tells us the important fact that spin is a relativistic effect that is naturally built into the quantum theory rather than having to artificially introduce a spin-magnetic field interaction as Pauli did. Further, Dirac's equation after incorporating the electromagnetic interaction by replacing the four momentum $p_\mu = -\partial_\mu$ by $p_\mu + eA_\mu$ and applying charge conjugation, shows that there also exists a solution with p_μ replaced by $p_\mu - eA_\mu$, ie, it predicts the existence of the positron and is the first instance of the proof of the existence of antimatter which is perhaps one of the most fundamental discoveries ever made in science.

1.19 The effect of noise on the Schrodinger wave equation—Open systems, ie, coupling of the system to the bath environment

1.19.1 General theory

The coupling of the system to a bath is described by a unitary evolution on the tensor product state of the system and the bath. The total evolution is driven by a system Hamiltonian along with system operators coupled to noisy processes which are operator valued processes in the bath Hilbert space. On tracing out the total density of the system and bath over the bath variables after joint unitary evolution, we obtain an evolution equation of the system state alone which is known as the GKSL equation. This equation describes the evolution of the mixed state of the system, not of a pure system state. In the language of entropy, it can be interpreted as saying that that bath pumps in entropy into the system causing an initial pure state of the system to become mixed. A pure state has zero entropy while a mixed state has positive entropy. An open problem in this field is to determine the Lindblad system coupling operators that would guarantee monotonic increase of the system entropy.

1.20 Wave equation with random non-uniform refractive index

1.20.1 General theory

Here $\epsilon(\omega, r)$ is the random frequency dependent permittivity of the medium while the permeability μ is a constant. The Maxwell curl equations are

$$\operatorname{curl} E(\omega, r) = -j\omega\mu H(\omega, r),$$

$$\operatorname{curl} H(\omega, r) = j\omega\epsilon(\omega, r)E(\omega, r)$$

To this we append the other two Maxwell equations

$$\operatorname{div}(\epsilon E) = 0, \operatorname{div} H = 0$$

We get

$$\operatorname{div} E = -(\nabla(\log(\epsilon)), E)$$

and hence

$$(\nabla^2 + \omega^2\mu\epsilon)E + \nabla(\nabla(\log\epsilon), E) = 0$$

This is the fundamental wave/Helmholtz equation for propagation of the electric field in a medium having frequency dependent non-uniform permittivity. We now solve this approximately using perturbation theory. Write

$$\epsilon(\omega, r) = \epsilon_0(1 + \delta\chi(\omega, r))$$

where χ is a random function and δ is a small perturbation parameter. Then, we have

$$\log\epsilon = \log(\epsilon_0) + \log(1 + \delta\chi) = \log(\epsilon_0) + \sum_{n \geq 1} (-1)^{n-1}\delta^n\chi^n/n$$

and so

$$\nabla(\log(\epsilon)) = \sum_{n \geq 1} (-1)^{n-1}\delta^n\nabla(\chi^n)/n$$

Writing

$$k^2 = \omega^2\mu\epsilon_0$$

we can thus express the above exact equation as

$$\begin{aligned} & (\nabla^2 + k^2)E(\omega, r) + k^2\delta\chi(\omega, r)/n)E(\omega, r) \\ & + \sum_{n \geq 1} (-1)^{n-1}(\delta^n/n)\nabla(\nabla\chi^n(\omega, r), E(\omega, r)) = 0 \end{aligned}$$

We now expand

$$E(\omega, r) = \sum_{n \geq 0} \delta^n E_n(\omega, r)$$

Substituting this expression into the above equation and equating equal powers of δ gives us

$$(\nabla^2 + k^2)E_0(\omega, r) = 0,$$

and for $n \geq 1$

$$\begin{aligned} & (\nabla^2 + k^2)E_n(\omega, r) + k^2\chi(\omega, r)E_{n-1}(\omega, r) \\ & + \sum_{m=0}^{n-1}((-1)^{n-m-1}/(n-m))\nabla(\nabla\chi^{n-m}(\omega, r), E_m(\omega, r)) = 0 \end{aligned}$$

and letting $G_k(r, r')$ denote the Green's function for the Helmholtz operator $\nabla^2 + k^2$ corresponding to the given boundary conditions, we can recursively compute the iterates $E_n(\omega, r)$, $n = 1, 2, \dots$ as

$$\begin{aligned} E_n(\omega, r) &= - \int G_k(r, r')[k^2\chi(\omega, r')E_{n-1}(\omega, r')] \\ &+ \sum_{m=0}^{n-1}((-1)^{n-m-1}/(n-m))\nabla(\nabla\chi^{n-m}(\omega, r'), E_m(\omega, r'))]d^3r' \end{aligned}$$

Using this expression, knowing the moments $\langle E_{n_1}(\omega, r_1) \otimes E_{n_2}(\omega, r_2) \otimes \dots \otimes E_{n_m}(\omega, r_m) \rangle$ or more generally, with some of the terms in this tensor product replaced by their complex conjugates for $n_1, \dots, n_m \leq n-1, m = 1, 2, \dots$, we can calculate the same for $n_1, \dots, n_m \leq n, m = 1, 2, \dots$. This provides a recursive algorithm for calculating the statistical properties of the scattered electric field by the frequency dependent inhomogeneous random medium.

We now note that upto $O(\delta)$, the above wave equation approximates to

$$(\nabla^2 + k^2)E(\omega, r) + \delta k^2\chi(\omega, r)E(\omega, r) + \delta\nabla(\nabla\chi(\omega, r), E(\omega, r)) = 0$$

which has the approximate solution

$$E(\omega, r) = E_0(\omega, r) + \delta.E_1(\omega, r)$$

where

$$(\nabla^2 + k^2)E_0(\omega, r) = 0$$

and

$$(\nabla^2 + k^2)E_1(\omega, r) = -[k^2\chi(\omega, r)E_0(\omega, r) + \nabla(\nabla\chi(\omega, r), E_0(\omega, r))]$$

The general solution for E_0 keeping in mind $\text{div}E_0 = 0$ (which is obtained by equating to zero the $O(1)$ term in $\text{div}(\epsilon E) = 0$) is given by

$$E_0(\omega, r) = \int F(\omega, \hat{n})\exp(-i\omega\hat{n}.r)/c d\Omega(\hat{n})$$

where

$$k = \omega/c, c = 1/\sqrt{\mu\epsilon_0}, \hat{n} \cdot F(\omega, \hat{n}) = 0$$

Assume that the correlations of the incident electric field E_0 in the frequency domain are of the form

$$\langle F(\omega, \hat{n}) \cdot F(\omega', \hat{n}')^* \rangle = 2\pi\delta(\omega - \omega')\delta(\hat{n} - \hat{n}')S(\omega, \hat{n})$$

Then, we can calculate the incident electric field correlations in the time domain as

$$\begin{aligned} E_0(t, r) &= \int E_0(\omega, r) \exp(i\omega t) d\omega / (2\pi) \\ &= \int F(\omega, \hat{n}) \exp(i\omega(t - \hat{n} \cdot r/c)) d\omega d\Omega(\hat{n}) / (2\pi) \\ &= \int f(t - \hat{n} \cdot r/c, \hat{n}) d\Omega(\hat{n}) \end{aligned}$$

where

$$f(t, \hat{n}) = \int F(\omega, \hat{n}) \exp(i\omega t) d\omega / (2\pi)$$

Then,

$$\begin{aligned} \langle f(t, \hat{n}) \cdot f(t', \hat{n}')^T \rangle &= (2\pi)^{-2} \int \langle F(\omega, \hat{n}) \cdot F(\omega', \hat{n}')^* \rangle \exp(i(\omega t - \omega' t')) d\omega d\omega' \\ &= (2\pi)^{-1} \delta(\hat{n} - \hat{n}') \int S(\omega, \hat{n}) \exp(i\omega(t - t')) d\omega \end{aligned}$$

and hence

$$\langle E_0(t, r) \cdot E_0(t', r')^T \rangle = (2\pi)^{-1} \int S(\omega, \hat{n}) \exp(i\omega(t - t' - \hat{n} \cdot (r - r')/c)) d\omega d\Omega(\hat{n})$$

which shows that the electric field $E_0(t, r)$ is a wide sense stationary random field in space-time. We note that

$$\begin{aligned} &\langle E_0(\omega, r) \cdot E_0(\omega', r')^* \rangle = \\ &2\pi\delta(\omega - \omega') \int S(\omega, \hat{n}) \exp(-i\omega \hat{n} \cdot (r - r')/c) d\Omega(\hat{n}) \end{aligned}$$

We now evaluate the correlations in the scattered field assuming that $\chi(\omega, r)$ is a random field statistically uncorrelated with the incident electric field $E_0(\omega, r)$. We get by solving the equation for the scattered field E_1 ,

$$E_1(\omega, r) = - \int G_k(r, r')[k^2 \chi(\omega, r') E_0(\omega, r') + \nabla(\nabla \chi(\omega, r'), E_0(\omega, r'))] d^3 r'$$

and hence, assuming a scalar Green's function G_k ,

$$\begin{aligned} &\langle E_{1a}(\omega, r) \cdot E_{1b}(\omega', r')^* \rangle = \\ &\int G_k(r, r_1) G_{k'}(r', r'_1)^* \langle (k^2 \chi(\omega, r_1) E_{0a}(\omega, r_1) + (\chi_{,s}(\omega, r_1) E_{0s}(\omega, r_1)),_a) \\ &\times (k'^2 \chi(\omega', r'_1)^* E_{0b}(\omega', r'_1)^* + (\chi_{,m}(\omega', r'_1)^* E_{0m}(\omega', r'_1)^*,_b) d^3 r_1 d^3 r'_1 \end{aligned}$$

1.21 The relationship between the wave equation and the Helmholtz equation for waves of given frequency

1.21.1 General theory

The most general form of a linear wave equation arising out of a time invariant system is a pde of the from

$$P(r, \nabla_r \partial_t) \psi(t, r) = 0 \quad (1)$$

where $P(r, p, E)$ is a polynomial in the variables $p \in \mathbb{R}^3$ and $E \in \mathbb{R}$ with coefficients depending on $r = (x, y, z)$ but not on t . We assume a certain ordering of the non-commuting operators r, ∇_r . In fact using the standard Heisenberg commutation relations, we can always arrange that the terms in r appear to the left of the terms involving ∇_r . If we define

$$\hat{\psi}(\omega, r) = \int_{-\infty}^{\infty} \psi(t, r) \exp(-j\omega t) dt$$

then the above pde reduces to

$$P(r, \nabla_r, j\omega) \hat{\psi}(\omega, r) = 0 \quad (2)$$

(1) is a generalized wave equation and (2) is the generalized Helmholtz equation corresponding to it. Special cases of such equations arise in the Maxwell theory in situations when the relationships between the fields $E(t, r), D(t, r)$ and $H(t, r), B(t, r)$ have the form

$$D(t, r) = \int_{-\infty}^t \epsilon(t-s, r) E(s, r) ds, \quad B(t, r) = \int_{-\infty}^t \mu(t-s, r) H(s, r) ds$$

where ϵ, μ are 3×3 matrix valued functions. Such relations can be easily deduced from physical models of an electron moving in an external field subject to a spring and damping force acting about the equilibrium position of the electron. More generally, if nonlinear damping or nonlinear spring binding forces are accounted for, then one would obtain the following generalizations of the above equations:

$$D(t, r) = \sum_{n \geq 1} \int_{s_1, \dots, s_n \geq 0} \epsilon_n(s_1, s_2, \dots, s_n, r) (E(t-s_1, r) \otimes E(t-s_2, r) \otimes \dots \otimes E(t-s_n, r)) ds_1 \dots ds_n,$$

$$B(t, r) = \sum_{n \geq 1} \int_{s_1, \dots, s_n \geq 0} \epsilon_n(s_1, s_2, \dots, s_n, r) (H(t-s_1, r) \otimes H(t-s_2, r) \otimes \dots \otimes H(t-s_n, r)) ds_1 \dots ds_n,$$

1.21.2 Exercises

- [1] Transform these integral relations in the time domain to integral relations in the frequency domain and set up the Maxwell equations in such a field-dependent (ie, nonlinear) inhomogeneous and anisotropic medium.

1.22 Waves in a confined region

1.22.1 General theory

When we look at em waves within a waveguide or a cavity resonator, ie confine the em waves to stay within a boundary, then the TEM condition of free space breaks down owing to the necessity of imposing boundary conditions on the em fields at the walls of the container. For example if we look at the electric field inside a waveguide whose length is oriented along the z -direction, then we must impose the condition that E_z vanishes on the boundary since the boundary is assumed to be a perfect conductor. But since there can exist surface currents in a perfect conductor, the same cannot be said about H_z . The origin of surface currents on the walls can be traced out to Ampere's law with Maxwell's displacement current correction term. Indeed, the line integral of the transverse magnetic field around the boundary of a cross section of the guide that is parallel to the xy plane is non-zero and hence there must be a net conduction plus displacement current correction term. Since there is no conduction current in the interior of the guide, there must be a net displacement current along the z direction and hence E_z cannot vanish in general. This means that the electric field can have a longitudinal component.

1.22.2 Exercises

[1] Assume that the axis of a waveguide with perfectly conducting walls is parallel to the z axis and that the cross section of the guide is parallel to the xy plane. Assume that in the frequency domain, the electric and magnetic fields can be expressed as superpositions of phasors of the form

$$E(x, y).exp(-\gamma z), H(x, y).exp(-\gamma z)$$

Let

$$E_{\perp} = E_x \hat{x} + E_y \hat{y}, H_{\perp} = H_x \hat{x} + H_y \hat{y}$$

By simplifying the Maxwell equations

$$\text{curl } E = -j\omega\mu H, \text{curl } H = j\omega\epsilon E$$

show that

$$E_{\perp} = (-\gamma/h^2)(\nabla_{\perp} E_z) - (j\omega\mu/h^2)\nabla_{\perp} H_z \times \hat{z},$$

$$H_{\perp} = (-\gamma/h^2)(\nabla_{\perp} H_z) + (j\omega\epsilon/h^2)\nabla_{\perp} E_z \times \hat{z}$$

where

$$h^2 = \omega^2\epsilon\mu + \gamma^2$$

Hint: Use the identities

$$\begin{aligned} \text{curl } E &= (\nabla_{\perp} - \gamma\hat{z}) \times (E_{\perp} + E_z\hat{z}) = \\ &= \nabla_{\perp} \times E_{\perp} + \nabla_{\perp} E_z \times \hat{z} - \gamma\hat{z} \times E_{\perp} \end{aligned}$$

and note that the first term has only a z component while the latter two components have only x and y components. By using the fact that the tangential and normal components of E and H vanish on the boundary ∂D of the guide, show that E_z, H_z satisfy respectively the Dirichlet and Neumann boundary conditions for the Helmholtz equation:

$$(\nabla_{\perp}^2 + h^2)(E_z, H_z) = 0,$$

$$E_z|_{\partial D} = 0, \partial H_z / \partial \hat{n}|_{\partial D} = 0$$

where \hat{n} is the unit outward normal to the boundary (It falls in the xy plane). Show that for rectangular waveguides, the modal eigenvalues h^2 are the same for both the TE and TM modes while this is not so for cylindrical waveguides. Specifically, show that for cylindrical guides, the modal eigenvalues for the TE mode are determined by the roots of the Bessel function while those for the TM mode are determined by the roots of the derivative of the Bessel function.

[2] With reference to the previous problem, assume that in the xy plane, there are orthogonal curvilinear coordinates (q_1, q_2) such that $q_1 = c_0$ is the waveguide boundary. Using Lame's coefficients and the formulae for gradient and the Laplacian in orthogonal curvilinear coordinates, determine the Helmholtz equations satisfied by E_z, H_z and the associated boundary conditions in terms of the q_1, q_2 system. Give a variational method for obtaining these Helmholtz equations with the boundary condition and hence formulate a finite element algorithm for computing the modal eigenvalues and eigenfunctions for the TE and TM modes.

[3] With reference to the previous problem, assume that the permittivity and permeability of a guide at a given frequency are functions of x and y or equivalently of q_1 and q_2 and determine a pair of coupled generalized Helmholtz equations for E_z, H_z . Show that they do not generally decouple into TE and TM solutions as they did for the case when the permittivity and permeability were constants. Derive a perturbation theoretic algorithm for computing the shift in the modal eigenvalues and eigenfunctions when the permittivity and permeability are in homogeneous but small perturbations of constants, ie,

$$\epsilon(x, y) = \epsilon_0(1 + \delta\chi_e(x, y)), \mu(x, y) = \mu_0(1 + \delta\chi_m(x, y))$$

or equivalently,

$$\epsilon(q_1, q_2) = \epsilon_0(1 + \delta\chi_e(q_1, q_2)), \mu(x, y) = \mu_0(1 + \delta\chi_m(q_1, q_2))$$

[4] With reference to the previous problems, assume that the permittivity and permeability are small anisotropic and inhomogeneous perturbations of constant, ie,

$$\epsilon(q_1, q_2, z) = \epsilon_0(1 + \delta\chi_e(q_1, q_2, z)),$$

$$\mu(q_1, q_2,) = \mu_0(1 + \delta.\chi_m(q_1, q_2, z))$$

where χ_e, χ_m are 3×3 matrices having the expansion

$$\chi_e(q_1, q_2, z) = \sum_{n \in \mathbb{Z}} \chi_{e,n}(q_1, q_2) \exp(-jn\beta z),$$

$$\chi_m(q_1, q_2, z) = \sum_{n \in \mathbb{Z}} \chi_{m,n}(q_1, q_2) \exp(-jn\beta z)$$

Assume that the electric and magnetic fields have the expansions

$$E(x, y, z) = \sum_{n \in \mathbb{Z}} E_n(x, y) \exp(-(\gamma + n\beta)z),$$

$$H(x, y, z) = \sum_{n \in \mathbb{Z}} H_n(x, y) \exp(-(\gamma + n\beta)z),$$

Substitute these into the Maxwell curl equations and apply perturbation theory with δ as the perturbation parameter to determine the shift in the modal eigenvalues and eigenfunction caused by the small inhomogeneity and anisotropy. Here $\beta = 2\pi/d$ where d is the length of the guide. Also apply this theory to the cavity resonator problem when the the planes $z = 0, d$ are closed by perfectly conducting surfaces so that the E_\perp, H_z vanish on these surfaces. Calculate the perturbation to the resonator oscillation frequency caused by this inhomogeneity and anisotropy. First attempt the case when the permittivity and permeability are isotropic scalar functions and then attempt the anisotropic case.

Remarks: For the waveguide problem, we do not have to perturb the frequency while for the cavity resonator problem, since the square of the frequency appears as an eigenvalue, we have to perturb the frequency also to obtain its change under the introduction of an inhomogeneity and anisotropy. The Maxwell curl equations for the waveguide problem give

$$\nabla_\perp E_{nz} \times \hat{z} - (\gamma + n\beta) \hat{z} \times E_{n\perp} = -j\omega\mu_0 H_{n\perp} - j\omega\delta\mu_0 \sum_k (\chi_{m,n-k} H_k)_{perp},$$

$$\nabla_\perp H_{nz} \times \hat{z} - (\gamma + n\beta) \hat{z} \times H_{n\perp} = j\omega\epsilon_0 E_{n\perp} + j\omega\delta\epsilon_0 \sum_k (\chi_{e,n-k} E_k)_{perp},$$

$$\hat{z} \cdot \nabla_\perp \times E_{n\perp} = -j\omega\mu_0 H_{nz} - j\omega\mu_0 \delta \sum_k (\chi_{m,n-k} H_k)_z$$

$$\hat{z} \cdot \nabla_\perp \times H_{n\perp} = j\omega\epsilon_0 E_{nz} + j\omega\epsilon_0 \delta \sum_k (\chi_{e,n-k} E_k)_z$$

Writing

$$E_n = E_n^{(0)} + \delta.E_n^{(1)} + O(\delta^2), H_n = H_n^{(0)} + \delta.H_n^{(1)} + O(\delta^2)$$

we get on equating terms of $O(\delta^r), r = 0, 1$ successively,

$$\nabla_\perp E_{nz}^{(0)} \times \hat{z} - (\gamma + n\beta) \hat{z} \times E_{n\perp}^{(0)} = -j\omega\mu_0 H_{n\perp}^{(0)}$$

$$\nabla_{\perp} H_{nz}^{(0)} \times \hat{z} - (\gamma + n\beta) \hat{z} \times H_{n\perp}^{(0)} = j\omega\epsilon_0 E_{n\perp}^{(0)}, n \in \mathbb{Z}$$

and

$$\begin{aligned} \nabla_{\perp} E_{nz}^{(1)} \times \hat{z} - (\gamma + n\beta) \hat{z} \times E_{n\perp}^{(1)} &= -j\omega\mu_0 H_{n\perp}^{(1)} - j\omega\mu_0 \sum_k (\chi_{m,n-k} H_k^{(0)})_{perp} \\ \nabla_{\perp} H_{nz}^{(1)} \times \hat{z} - (\gamma + n\beta) \hat{z} \times H_{n\perp}^{(1)} &= j\omega\epsilon_0 E_{n\perp}^{(1)} + j\omega\epsilon_0 \sum_k (\chi_{e,n-k} E_k^{(0)})_{perp} \\ \hat{z} \cdot \nabla_{\perp} \times E_{n\perp}^{(0)} &= -j\omega\mu_0 H_{nz}^{(0)}, \\ \hat{z} \cdot \nabla_{\perp} \times H_{n\perp}^{(0)} &= j\omega\epsilon_0 E_{nz}^{(0)}, \\ \hat{z} \cdot \nabla_{\perp} \times E_{n\perp}^{(1)} &= -j\omega\mu_0 H_{nz}^{(1)} - j\omega\mu_0 \sum_k (\chi_{m,n-k} H_k^{(0)})_z \\ \hat{z} \cdot \nabla_{\perp} \times H_{n\perp}^{(1)} &= j\omega\epsilon_0 E_{nz}^{(1)} + j\omega\epsilon_0 \sum_k (\chi_{e,n-k} E_k^{(0)})_z \end{aligned}$$

For the cavity resonator problem, the relevant equations are

$$\begin{aligned} \nabla_{\perp} E_{nz} \times \hat{z} - (\gamma + n\beta) \hat{z} \times E_{n\perp} &= -j(\omega + \delta\omega_1) \mu_0 H_{n\perp} - j\omega\delta\mu_0 \sum_k (\chi_{m,n-k} H_k)_{perp} \\ \nabla_{\perp} H_{nz} \times \hat{z} - (\gamma + n\beta) \hat{z} \times H_{n\perp} &= j(\omega + \delta\omega_1) \epsilon_0 E_{n\perp} + j\omega\delta\epsilon_0 \sum_k (\chi_{e,n-k} E_k)_{perp}, \\ \hat{z} \cdot \nabla_{\perp} \times E_{n\perp} &= -j(\omega + \delta\omega_1) \mu_0 H_{nz} - j\omega\mu_0 \delta \sum_k (\chi_{m,n-k} H_k)_z \\ \hat{z} \cdot \nabla_{\perp} \times H_{n\perp} &= j(\omega + \delta\omega_1) \epsilon_0 E_{nz} + j\omega\epsilon_0 \delta \sum_k (\chi_{e,n-k} E_k)_z \end{aligned}$$

where ω is the unperturbed frequency and the perturbed frequency is $\omega + \delta\omega_1 + O(\delta^2)$. In these equations, we must expand

$$E_n = E_n^{(0)} + \delta E_n^{(1)} + O(\delta^2),$$

$$H_n = H_n^{(0)} + \delta H_n^{(1)} + O(\delta^2)$$

and calculate ω_1 using the electric and magnetic field unperturbed eigenfunctions. The boundary conditions are that H_z must vanish when $z = 0, d$ and E_x, E_y must also vanish when $z = 0, d$. In view of the Maxwell equation

$$(\epsilon E)_x, x + (\epsilon E)_y, y + (\epsilon E)_z, z = 0$$

the boundary conditions on E_x, E_y are equivalent to requiring that

$$(\epsilon_{xz} E_z), x + (\epsilon_{yz} E_z), y + \epsilon_{zx} E_{x,z} + \epsilon_{zy} E_{y,z} + (\epsilon_{zz} E_z), z = 0$$

where $z = 0, d$. Further, in the isotropic case, this is equivalent to $E_{z,z} = 0, z = 0, d$. Thus, in the isotropic inhomogeneous case, H_z will depend on z like $\sin(\pi p z/d)$ and E_z like $\cos(\pi p z/d)$ where p is an integer. Thus for the resonator problem, we should use the expansions

$$H_z = \sum_n H_{nz}(q_1, q_2) \sin(n\pi z/d), E_z = \sum_n E_{nz}(q_1, q_2) \cos(n\pi z/d)$$

Note that this means that

$$H_z = H_z^{(0)} + \delta \cdot H_z^{(1)}, E_z = E_z^{(0)} + \delta \cdot E_z^{(1)}$$

where

$$H_z^{(0)} = \sum_n H_{nz}^{(0)}(q_1, q_2) \sin(n\pi z/d), E_z^{(0)} = \sum_n E_{nz}^{(0)}(q_1, q_2) \cos(n\pi z/d)$$

$$H_z^{(1)} = \sum_n H_{nz}^{(1)}(q_1, q_2) \sin(n\pi z/d), E_z^{(1)} = \sum_n E_{nz}^{(1)}(q_1, q_2) \cos(n\pi z/d)$$

We leave it as an exercise to pursue this line and use tools similar to time independent perturbation theory in quantum mechanics to calculate the perturbation in the eigenfrequencies caused by the inhomogeneity of the medium.

1.22.3 Points to remember

[1] To study the propagation of waves in a waveguide whose cross section is parallel to the xy plane and whose length is parallel to the z axis, it turns out useful to assume the z-dependence of the fields as $\exp(-\gamma z)$ where γ depends on the frequency of operation. We then separate out the electric and magnetic field components into a transverse part and a longitudinal part and substitute these into the two transverse components of the Maxwell curl equations in the frequency domain, thus obtaining the transverse components of the electromagnetic field in terms of the transverse gradients of their longitudinal components. The longitudinal components of the Maxwell curl equations then gives us decoupled Helmholtz equations for the longitudinal components of the em fields. Owing to this decoupling, it is possible to express any solution as the sum of a TE solution and a TM solution.

[2] In the case when the guide's permittivity and permeability are inhomogeneous or anisotropic or both, the decoupling into TE and TM modes generally cannot be done. The TE mode is characterized by the vanishing of the longitudinal component of the electric field on the boundary surface while the TM mode is characterized by the vanishing of the normal component of the magnetic field which leads to in the isotropic case, the vanishing of normal derivative of the longitudinal magnetic field component on the surface. Here, we assume that the boundary wall of the guide is a perfect electric conductor (PEC).

[3] The different boundary conditions associated with the TE and TM modes in general, gives us different values of the modes, ie, of the cutoff frequencies. In the special case of a rectangular waveguide however, these modes coincide. In the case of a cylindrical guide, they are respectively defined in terms of the zeroes of the Bessel functions and the zeroes of the Besse function derivatives.

1.23 Schrodinger's wave equation for mixed states in the position kernel domain

1.23.1 Points to remember

[1] The starting point for an intuitive derivation of the Schrodinger equation is based on Planck's quantum hypothesis that energy $E = \hbar\omega$ and De-Broglie's matter wave duality relation $\lambda = h/p$ or equivalently, $p = \hbar k$ where ω, k are the frequency and wave number of the wave and p is the momentum of the particle associated with a matter wave. By using the expression for a plane wave $\psi(t, r) = \exp(-i(\omega t - k \cdot r))$, and the classical mechanics relation $E = p^2/2m + V(r)$, we then easily derive

$$((-h^2/2m)\nabla^2 + V(r))\psi(t, r) = i\hbar\partial_t\psi(t, r)$$

In other words $i\hbar\partial_t$ is the energy operator and $-i\hbar\nabla$ is the momentum operator.

[2] The Schrodinger equation for a mixed state in the position kernel domain is obtained by using the fact that the equilibrium mixed state kernel in the position domain can be expanded as

$$\rho(r, r') = \sum_k p_k \psi_k(r) \bar{\psi}_k(r')$$

where E_k is the energy of the Hamiltonian operator corresponding to the stationary state $\psi_k(r)$. It is easy to see that if $H(r, r')$ is the position kernel representation of the Hamiltonian operator, then ρ satisfies

$$H\rho - \rho H = 0$$

where

$$AB(r, r') = \int A(r, r'') d^3r'' B(r'', r')$$

for two operators A, B having position kernels

$$A(r, r') = \langle r | A | r' \rangle, B(r, r') = \langle r | B | r' \rangle$$

If we take $p_k = \exp(-\beta E_k)/Z(\beta)$, then ρ becomes the Gibbs density. If X is any observable, its average in the state ρ is a combination of classical and quantum averages:

$$\begin{aligned} \langle X \rangle_\rho &= \text{Tr}(\rho X) = \int \rho(r, r') X(r', r) d^3r d^3r' \\ &= \sum_k p_k \int \bar{\psi}_k(r') X(r', r) \psi_k(r) d^3r d^3r' \\ &= \sum_k p_k \langle \psi_k | X | \psi_k \rangle \end{aligned}$$

p_k is the classical probability that the system is in the state ψ_k and $\langle \psi_k | X | \psi_k \rangle$ is the quantum average of X given that the system is in the pure state ψ_k . We

can have other kinds of mixed states which vary with time in accordance with Schrodinger's equation. For example, suppose $\phi_k(t, r), k = 1, 2, \dots$ satisfy the time varying Schrodinger equation

$$ih\partial_t\phi_k(t, r) = H\phi_k(t, r)$$

and are also mutually orthogonal. Then we define

$$\rho_t(r, r') = \sum_k p_k \phi_k(t, r) \bar{\phi}_k(t, r')$$

It is easily seen that

$$ih\partial_t\rho_t = H\rho_t - \rho_t H = [H, \rho_t]$$

The probability density of the position of a particle in this state at time t is given by

$$\rho_t(r, r) = \sum_k p_k |\phi_k(t, r)|^2$$

$|\phi_k(t, r)|^2$ is the probability density of the position of the particle at time t , given that it is in the state ϕ_k while p_k is the classical probability that the particle is in the state ϕ_k . We have the obvious relation

$$\langle \chi | \rho_t | \chi \rangle = \sum_k p_k |\langle \chi | \phi_k \rangle|^2 \geq 0$$

for any state χ and further,

$$Tr(\rho_t) = \sum_k \int |\phi_k(t, r)|^2 d^3r = \sum_k p_k = 1$$

An obvious method for constructing such states is to start with the complete orthonormal eigenbasis $\psi_k(r), k = 1, 2, \dots$ of the Hamiltonian operator H . Apply a unitary operator U to each of these states giving

$$\phi_k(r) = U\psi_k(r) = \sum_m u(m, k)\psi_m(r)$$

Allowing these states to evolve under the Hamiltonian H resulting in

$$\phi_k(t, r) = \exp(-itH)\phi_k(r) = \sum_m u(m, k)\exp(-iE_m t)\psi_m(r)$$

and then using the above construction of $\rho_t(r, r')$. The idea of using mixed states is basically to study the effects of noise on a quantum system. Noise transforms pure states into mixed ones and hence zero entropy states into positive entropy states. This is clarified by the Hudson-Parthasarathy theory of quantum noise, where a unitary evolution governs the joint dynamics of system and bath and hence pure states on the system \otimes bath space are transformed again into pure states and after tracing out this evolved pure state over the bath variables, the system ends up in a mixed state. Thus, the bath pumps entropy into the system.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Chapter 2

Waves in general relativity, quantum gravity, plasma physics and quantum stochastics

2.1 Gravitational waves

2.1.1 Summary

Here, we study the linearized Einstein field equations which represent a good approximation in the weak gravitational field limit. The Ricci tensor then after an appropriate transformation of coordinates transforms into the standard D'Alembertian wave operator acting on the metric perturbations and on the rhs of the Einstein field equations, we have the energy-momentum tensor of the matter field. Thus, in the weak field limit, the metric perturbations are expressed using the standard retarded potential formulas that we study in electromagnetics with the source field being the matter energy-momentum tensor rather than a current density field. We can then make a far field approximation which leads to formulae for the metric perturbations of the form

$$h_{\mu\nu}(t, r) = (\exp(j\omega r/c)/r) \int F_{\mu\nu}(\omega, \hat{r}) \exp(j\omega \hat{r}.r'/c) \exp(-j\omega t) d\omega d^3r'$$

as one obtains in antenna theory. Using this expression for the metric perturbation, we can calculate upto quadratic orders in the metric perturbation, the energy-momentum tensor of the radiated gravitational field at large distances from the matter source. It should be noted that the energy-momentum pseudo-tensor of the gravitational field is obtained as proportional to the non-linear component of the Einstein tensor $R^{\mu\nu} - R.g^{\mu\nu}/2$ since when added to

the energy-momentum tensor of the matter field, its ordinary four divergence vanishes owing to the same identity holding good for the linear component of the Einstein tensor. This vanishing of the ordinary divergence implies a conservation law and hence one interprets the nonlinear component of the Einstein tensor as being the energy-momentum pseudo-tensor of the gravitational field. The $(0, r)$, $r = 1, 2, 3$ components of this energy-momentum pseudo-tensor yield the energy flux of the gravitational field and using its far field formula, we can evaluate by considering only the quadratic terms in the metric perturbations, the total gravitational energy per unit time radiated by the matter distribution. This is just like how one computes by integrating the Poynting vector over the surface a large sphere the total electromagnetic energy radiated per unit time by the charges and currents within the sphere. It should further be noted that if we solve the linearized Einstein field equations in free space, we obtain superposition of plane waves. If we consider one such plane wave propagating along a given direction, say the \hat{z} -direction using harmonic coordinates, then we obtain coordinate conditions for the complex amplitudes of these plane waves. These coordinate conditions show that there are essentially only five linearly independent amplitudes as opposed to the ten metric tensor components. Further, these five linearly independent amplitudes can be grouped in such a way that under a rotation by angle θ around the direction of propagation, these amplitudes get phase changed by $1, \exp(\pm i\theta)$ and $\exp(\pm 2i\theta)$. Thus, we infer that if an elementary particle called the graviton were associated with such a wave, then this particle must have spin two just as photons associated with an electromagnetic wave have spin one. We can also obtain using second order perturbation theory, frequency coupling in the gravitational waves. For this, we must expand the Ricci tensor upto linear and quadratic terms in the metric perturbations and their space-time partial derivatives and set up the field equations for a given energy-momentum tensor of matter. We then first neglect the quadratic terms and solve for the metric in the way as mentioned above, ie by the use of retarded potentials and then make far field approximations. We then calculate the quadratic contribution to the Einstein tensor by substituting this first order solution and again solve a linear wave equation for the metric perturbation with source taken as this quadratic contribution. We add the two solutions. If the energy momentum tensor was Fourier transformed w.r.t. time, the quadratic contribution to the Ricci tensor would contain products of the energy-momentum tensor Fourier transform at different frequency pairs and hence the solution to the metric upto second order will also contain such contributions.

We can also consider the measurement of a gravitational wave using photons as was done recently in the Lego apparatus at Louisiana. This would involve shining a laser on a screen and observing a deflection of the laser point on the screen when a gravitational wave interacted with it. To make calculations about this we would need to solve Maxwell's equations in a gravitational field when the gravitational field is a wave. Treating the gravitational wave as a small perturbation, the Maxwell equations in curved space-time ($F^{\mu\nu} \sqrt{-g}_{,\nu} = 0$ with $F^{\mu\nu} = g^{\mu\alpha} g^{\nu\beta} F_{\alpha\beta}, F_{\alpha\beta} = A_{\beta,\alpha} - A_{\alpha,\beta}$) can be solved approximately using first

order perturbation theory w.r.t. the metric perturbations of flat space-time or metric perturbations of a curved space-time background and the result would be that the em four potential gets corrected by a term that is bilinear in the incident em field and the metric perturbations. Then by matching this approximate solution to that observed on the screen one can estimate the strength and other properties like frequencies in the gravitational wave.

2.1.2 Discussion

Gravitational waves predicted by Einstein's general relativity: Corrections to planar waves by second degree nonlinear terms in the Ricci tensor.

Einstein's field equation in vacuum is

$$R_{\mu\nu} = 0$$

We write

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}(x)$$

where $\eta_{\mu\nu}$ is the Minkowski metric of flat space-time and $h_{\mu\nu}(x)$ is a small perturbation of this. We expand $R_{\mu\nu}$ upto second degree in the $h'_{\mu\nu}$ s. Writing

$$g = \eta + h$$

we have

$$g^{-1} = (I + \eta h)^{-1}\eta = \eta - \eta h\eta + \eta h\eta h\eta + O(h^2)$$

In terms of components, we have upto quadratic orders in h ,

$$g^{\mu\nu} = \eta_{\mu\nu} - h^{\mu\nu} + h^{\mu\alpha}h_\alpha^\nu$$

$$\begin{aligned} \Gamma_{\mu\alpha}^\alpha &= g^{\alpha\beta}\Gamma_{\beta\mu\alpha} = \\ &= (1/2)g^{\alpha\beta}g_{\alpha\beta,\mu} \\ &= (1/2)(\eta_{\alpha\beta} - h^{\alpha\beta})h_{\alpha\beta,\mu} \\ &= (1/2)(h_{,\mu} - h^{\alpha\beta}h_{\alpha\beta,\mu}) \\ \Gamma_{\mu\alpha,\nu}^\alpha &= \\ &= (1/2)(h_{,\mu\nu} - (h^{\alpha\beta}h_{\alpha\beta,\mu}),\nu) \\ \Gamma_{\mu\nu}^\alpha &= (1/2)(\eta_{\alpha\beta} - h^{\alpha\beta})(h_{\beta\mu,\nu} + h_{\beta\nu,\mu} - h_{\mu\nu,\beta}) \\ &= (1/2)(h_{\mu,\nu}^\alpha + h_{\nu,\mu}^\alpha - h_{\mu\nu}^\alpha) \\ &\quad - (1/2)h^{\alpha\beta}(h_{\beta\mu,\nu} + h_{\beta\nu,\mu} - h_{\mu\nu,\beta}) \end{aligned}$$

So

$$\begin{aligned} \Gamma_{\mu\nu,\alpha}^\alpha &= \\ &= (1/2)(h_{\mu,\alpha\nu}^\alpha + h_{\nu,\alpha\mu}^\alpha - \square h_{\mu\nu}) \end{aligned}$$

$$\begin{aligned}
& -(1/2)(h^{\alpha\beta}(h_{\beta\mu,\nu} + h_{\beta\nu,\mu} - h_{\mu\nu,\beta})),_{\alpha} \\
& \Gamma_{\mu\nu}^{\alpha}\Gamma_{\alpha\beta}^{\beta} = \\
& (\eta_{\alpha\rho}\eta_{\beta\sigma}/4)(h_{\rho\mu,\nu} + h_{\rho\nu,\mu} - h_{\mu\nu,\rho}) \\
& \times(h_{\sigma\alpha,\beta} + h_{\sigma\beta,\alpha} - h_{\alpha\beta,\sigma}) \\
& = (1/4)(h_{\mu,\nu}^{\alpha} + h^{\alpha\nu,\mu} - h_{\mu\nu}^{\rho}) \\
& \times(h_{\alpha,\beta}^{\beta} + h_{,\alpha}^{\beta} - h_{\alpha,\beta}^{\beta})
\end{aligned}$$

and finally,

$$\begin{aligned}
& \Gamma_{\mu\beta}^{\alpha}\Gamma_{\nu\alpha}^{\beta} = \\
& (\eta_{\alpha\rho}\eta_{\beta\sigma}/4)(h_{\rho\mu,\beta} + h_{\rho\beta,\mu} - h_{\mu\beta,\rho}) \\
& \times(h_{\sigma\nu,\alpha} + h_{\sigma\alpha,\nu} - h_{\nu\alpha,\sigma}) \\
& = (1/4)(h_{\mu,\beta}^{\alpha} + h_{\beta,\mu}^{\alpha} - h_{\mu\beta}^{\alpha}) \\
& \times(h_{\nu,\alpha}^{\beta} + h_{\alpha,\nu}^{\beta} - h_{\nu\alpha}^{\beta})
\end{aligned}$$

Thus, upto second order perturbations, our Einstein field equations in vaccum become

$$\begin{aligned}
& (1/2)(h_{,\mu\nu} - (h^{\alpha\beta}h_{\alpha\beta,\mu}),_{\nu} \\
& -(1/2)(h_{\mu,\alpha\nu}^{\alpha} + h_{\nu,\alpha\mu}^{\alpha} - \square h_{\mu\nu}) \\
& +(1/2)(h^{\alpha\beta}(h_{\beta\mu,\nu} + h_{\beta\nu,\mu} - h_{\mu\nu,\beta})),_{\alpha} \\
& -(1/4)(h_{\mu,\nu}^{\alpha} + h^{\alpha\nu,\mu} - h_{\mu\nu}^{\rho})h_{,\alpha} \\
& +(1/4)(h_{\mu,\beta}^{\alpha} + h_{\beta,\mu}^{\alpha} - h_{\mu\beta}^{\alpha}) \\
& \times(h_{\nu,\alpha}^{\beta} + h_{\alpha,\nu}^{\beta} - h_{\nu\alpha}^{\beta}) = 0
\end{aligned}$$

or equivalently, separating out the linear and quadratic components, we have

$$\begin{aligned}
& (h_{,\mu\nu} - h_{\mu,\alpha\nu}^{\alpha} - h_{\nu,\alpha\mu}^{\alpha} + \square h_{\mu\nu}) \\
& -(h^{\alpha\beta}h_{\alpha\beta,\mu}),_{\nu} + (h^{\alpha\beta}(h_{\beta\mu,\nu} + h_{\beta\nu,\mu} - h_{\mu\nu,\beta})),_{\alpha} \\
& -(1/2)(h_{\mu,\nu}^{\alpha} + h^{\alpha\nu,\mu} - h_{\mu\nu}^{\rho})h_{,\alpha} \\
& +(1/2)(h_{\mu,\beta}^{\alpha} + h_{\beta,\mu}^{\alpha} - h_{\mu\beta}^{\alpha}) \\
& \times(h_{\nu,\alpha}^{\beta} + h_{\alpha,\nu}^{\beta} - h_{\nu\alpha}^{\beta}) = 0
\end{aligned}$$

On adopting the standard harmonic coordinate condition

$$h_{\nu,\mu}^{\mu} - (1/2)h_{,\nu} = 0$$

this wave equation further simplifies to

$$\square h_{\mu\nu}$$

$$\begin{aligned}
& -(h^{\alpha\beta}h_{\alpha\beta,\mu}),\nu + (h^{\alpha\beta}(h_{\beta\mu,\nu} + h_{\beta\nu,\mu} - h_{\mu\nu,\beta})),\alpha \\
& -(1/2)(h_{\mu,\nu}^\alpha + h^{\alpha\nu,\mu} - h_{\mu\nu}^\rho)h_{,\alpha} \\
& +(1/2)(h_{\mu,\beta}^\alpha + h_{\beta,\mu}^\alpha - h_{\mu\beta}^\alpha)(h_{\nu,\alpha}^\beta + h_{\alpha,\nu}^\beta - h_{\nu\alpha}^\beta) = 0
\end{aligned}$$

Now consider the case when matter is present described by an energy-momentum tensor $T_{\mu\nu}(x)$ where $x = (t, r)$. The Einstein field equations

$$R_{\mu\nu} - (1/2)Rg_{\mu\nu} = -8\pi GT^{\mu\nu}$$

can be expressed in the alternative form as

$$R_{\mu\nu} = -8\pi G(T_{\mu\nu} - (1/2)Tg_{\mu\nu})$$

and its linearized version is

$$\begin{aligned}
& (1/2)(h_{,\mu\nu} - h_{\mu,\alpha\nu}^\alpha - h_{\nu,\alpha\mu}^\alpha + \square h_{\mu\nu}) \\
& = -8\pi GS_{\mu\nu}
\end{aligned}$$

where

$$S_{\mu\nu} = T_{\mu\nu} - (1/2)T\eta_{\mu\nu}$$

with

$$T_{\mu\nu} = (p + \rho)v_\mu v_\nu - p\eta_{\mu\nu}$$

and

$$T = p - 3\rho$$

Here,

$$v_\mu = \eta_{\mu\nu}v^\nu$$

with v^μ being the four velocity of the matter field. Alternately, we define

$$h'_{\mu\nu} = h_{\mu\nu} - (1/2)h\eta_{\mu\nu}$$

and find that the linearized form of R is given by

$$\begin{aligned}
R & \approx \eta_{\mu\nu}R_{\mu\nu} \approx \\
& = (1/2)\eta_{\mu\nu}(h_{,\mu\nu} - h_{\mu,\alpha\nu}^\alpha - h_{\nu,\alpha\mu}^\alpha + \square h_{\mu\nu}) \\
& = \square h - h_{,\alpha\beta}^{\alpha\beta}
\end{aligned}$$

so that the linearized form of the Einstein tensor

$$G_{\mu\nu} = R_{\mu\nu} - (1/2)Rg_{\mu\nu}$$

is

$$\begin{aligned}
& (1/2)(h_{,\mu\nu} - h_{\mu,\alpha\nu}^\alpha - h_{\nu,\alpha\mu}^\alpha + \square h_{\mu\nu}) \\
& -(1/2)(\square h - h_{,\alpha\beta}^{\alpha\beta})\eta_{\mu\nu}
\end{aligned}$$

If we now put the coordinate condition

$$h_{\nu,\mu}^\mu - (1/2)h_{,\mu} = 0$$

this simplifies to

$$G_{\mu\nu} \approx \square(1/2)(h_{\mu\nu} - (1/2)h\eta_{\mu\nu})$$

and the linearized Einstein field equations can be expressed in the form

$$\square h'_{\mu\nu} = -16\pi G T_{\mu\nu}$$

where

$$h'_{\mu\nu} = h_{\mu\nu} - (1/2)h\eta_{\mu\nu}$$

the solution to which is

$$h'_{\mu\nu}(t, r) = -4\pi G \int T_{\mu\nu}(t - |r - r'|, r') d^3 r' / |r - r'|$$

Another equivalent form of the linearized field equations is

$$\square h_{\mu\nu} = -16\pi G S_{\mu\nu}$$

the solution to which is

$$h_{\mu\nu}(t, r) = -4\pi G \int S_{\mu\nu}(t - |r - r'|, r') d^3 r' / |r - r'|$$

The far field approximate solution to this equation is

$$h_{\mu\nu}(t, r) = (-4\pi G/r) \int S_{\mu\nu}(t - r + \hat{r}.r', r') d^3 r'$$

or equivalently in the frequency domain

$$\begin{aligned} \hat{h}_{\mu\nu}(\omega, r) &= \int_{\mathbb{R}} h_{\mu\nu}(t, r) \exp(-i\omega t) dt = \\ &-4\pi G (\exp(-ikr)/r) \int \hat{S}_{\mu\nu}(\omega, r') \exp(ik\hat{r}.r') d^3 r' \\ &= -4\pi G (\exp(-ikr)/r) \tilde{S}_{\mu\nu}(\omega, k\hat{r}) \end{aligned}$$

where

$$\tilde{S}_{\mu\nu}(\omega, \mathbf{k}) = \int S_{\mu\nu}(t, r) \exp(-i(\omega t - \mathbf{k}.r)) dt d^3 r$$

Equivalently,

$$\hat{h}'_{\mu\nu}(\omega, r) = -4\pi G (\exp(-ikr)/r) \tilde{T}_{\mu\nu}(\omega, k\hat{r})$$

In all these expressions, $k = \omega/c = \omega$ on choosing our units so that $c = 1$. Now, we are in a position to calculate the total energy radiated by our matter system

in the form of gravitational waves. First, we note that the exact Einstein field equations are

$$G_{\mu\nu}^{(1)} + G_{\mu\nu}^{(2)} = -8\pi GT_{\mu\nu}$$

where $G_{\mu\nu}^{(1)}$ is the linear component of $G_{\mu\nu}$ and is given by

$$\begin{aligned} G_{\mu\nu}^{(1)} &= \\ (1/2)(h_{,\mu\nu} - h_{\mu,\alpha\nu}^\alpha - h_{\nu,\alpha\mu}^\alpha + \square h_{\mu\nu}) \\ &\quad - (1/2)(\square h - h_{,\alpha\beta}^{\alpha\beta})\eta_{\mu\nu} \\ &= (1/2)(\square(h_{\mu\nu} - h\eta_{\mu\nu}) + (1/2)(h_{,\mu\nu} - h_{\mu,\alpha\nu}^\alpha - h_{\nu,\alpha\mu}^\alpha + h_{,\alpha\beta}^{\alpha\beta}\eta_{\mu\nu}) \end{aligned}$$

It is easily seen that the ordinary four divergence of $G_{\mu\nu}^{(1)}$ vanishes:

$$\begin{aligned} 2G_{\mu\nu}^{(1),\nu} &= \eta_{\nu\rho}G_{\mu\nu,\rho}^{(1)} = \\ \square(h_{\mu\nu}^\nu - h_{,\mu}^\nu) &+ \\ (h_{,\mu\nu}^\nu - h_{\mu,\alpha\nu}^{\alpha,\nu} - h_{\nu,\alpha\mu}^{\alpha,\nu} + h_{,\alpha\beta\mu}^{\alpha\beta}) \\ &= 0 \end{aligned}$$

even if we do not impose any coordinate condition on $h_{\mu\nu}$. $G_{\mu\nu}^{(2)}$ is defined by $G_{\mu\nu} - G_{\mu\nu}^{(1)}$ and it therefore follows from the Einstein field equations that the ordinary four divergence of $-8\pi GT_{\mu\nu} - G_{\mu\nu}^{(2)}$ must necessarily vanish. Equivalently,

$$(T_{\mu\nu} + (8\pi G)^{-1}G_{\mu\nu}^{(2)})^{\cdot\nu} = 0$$

which means that $(8\pi G)^{-1}G_{\mu\nu}^{(2)}$ is to be interpreted as the energy-momentum pseudo-tensor of the gravitational field. Such an interpretation ensures that the total energy and momentum of the matter field and the gravitational field is conserved. We write

$$\tau_{\mu\nu} = (8\pi G)G_{\mu\nu}^{(2)}$$

and approximate $G_{\mu\nu}^{(2)}$ by retaining only the terms quadratic in the $h_{\mu\nu}$ and its partial derivatives. $h_{\mu\nu}$ will be calculated as above using the linearized Einstein field equations in the presence of matter and we can thus estimate $\tau^{\mu 0}$ in terms of the energy-momentum tensor of the matter field. This would give us the energy momentum flux of the gravitational field in the frequency domain and that would complete the story of gravitational radiation. According to the above calculation, we have on neglecting cubic and higher order terms $h_{\mu\nu}$ and its partial derivatives,

$$\begin{aligned} R_{\mu\nu}^{(2)} &= \\ (-1/2)(h^{\alpha\beta}h_{\alpha\beta,\mu}),_\nu \\ &\quad + (1/2)(h^{\alpha\beta}(h_{\beta\mu,\nu} + h_{\beta\nu,\mu} - h_{\mu\nu,\beta})),_\alpha \\ &\quad - (1/4)(h_{\mu,\nu}^\alpha + h^{\alpha\nu,\mu} - h_{\mu\nu}^\rho)h_{,\alpha} \end{aligned}$$

$$\begin{aligned} & + (1/4)(h_{\mu,\beta}^\alpha + h_{\beta,\mu}^\alpha - h_{\mu\beta}^{\alpha}) \\ & \times (h_{\nu,\alpha}^\beta + h_{\alpha,\nu}^\beta - h_{\nu\alpha}^{\beta}) \end{aligned}$$

Thus,

$$\begin{aligned} R^{(2)\mu\nu} &= (g^{\mu\alpha} g^{\nu\beta} (R_{\mu\nu}^{(1)} + R_{\mu\nu}^{(2)}))^{(2)} \\ &= [(\eta_{\mu\alpha} - h^{\mu\alpha})(\eta_{\nu\beta} - h^{\nu\beta})(R_{\mu\nu}^{(1)} + R_{\mu\nu}^{(2)})]^{(2)} \\ &= -(\eta_{\nu\beta} h^{\mu\alpha} + \eta_{\mu\alpha} h^{\nu\beta} R_{\mu\nu}^{(1)} + \eta_{\mu\alpha} \eta_{\nu\beta} R_{\mu\nu}^{(2)}) \end{aligned}$$

Further, upto quadratic orders,

$$R = R^{(1)} + R^{(2)} = (\eta_{\mu\nu} - h^{\mu\nu})(R_{\mu\nu}^{(1)} + R_{\mu\nu}^{(2)})$$

so that

$$\begin{aligned} R^{(2)} &= \eta_{\mu\nu} R_{\mu\nu}^{(2)} - h^{\mu\nu} R_{\mu\nu}^{(1)}, \\ G_{\mu\nu} &= G_{\mu\nu}^{(1)} + G_{\mu\nu}^{(2)}, \\ G^{\mu\nu} &= (\eta_{\mu\alpha} - h^{\mu\alpha})(\eta_{\nu\beta} - h^{\nu\beta})(G_{\alpha\beta}^{(1)} + G_{\alpha\beta}^{(2)}) \end{aligned}$$

so that

$$G^{(2)\mu\nu} = \eta_{\mu\alpha} \eta_{\nu\beta} G_{\alpha\beta}^{(2)} - (\eta_{\nu\beta} h^{\mu\alpha}) + \eta_{\mu\alpha} h^{\nu\beta} G_{\alpha\beta}^{(1)}$$

where

$$\begin{aligned} G_{\mu\nu}^{(1)} &= R_{\mu\nu}^{(1)} - (1/2)R^{(1)}\eta_{\mu\nu} \\ G_{\mu\nu}^{(2)} &= R_{\mu\nu}^{(2)} - (1/2)R^{(2)}\eta_{\mu\nu} + (1/2)R^{(1)}h^{\mu\nu} \\ &= R_{\mu\nu}^{(2)} - (1/2)\eta_{\mu\nu}(\eta_{\alpha\beta} R_{\alpha\beta}^{(2)} - h^{\alpha\beta} R_{\alpha\beta}^{(1)}) \\ &\quad + (1/2)R^{(1)}h^{\mu\nu} \end{aligned}$$

Assuming the harmonic coordinate condition, we have seen that the far field approximation to the metric perturbations $h_{\mu\nu}(x)$ is given by

$$h_{\mu\nu}(t, r) = (-4\pi G/r) \int T_{\mu\nu}(t - r + \hat{r}.r', r') d^3 r'$$

and if we assume that the energy-momentum tensor of matter varies harmonically with time with a frequency ω , so that we can write

$$T_{\mu\nu}(t, r) = \text{Re}(\hat{T}_{\mu\nu}(\omega, r) \exp(j\omega t))$$

then in phasor notation,

$$h_{\mu\nu}(t, r) = -(4\pi G/r) \text{Re}(\exp(j\omega(t - r)) \tilde{T}_{\mu\nu}(\omega, \omega \hat{r}))$$

where

$$\tilde{T}_{\mu\nu}(\omega, k) = \int \hat{T}_{\mu\nu}(\omega, r) \exp(jk.r) d^3 r$$

To calculate the energy-momentum pseudo-tensor of the gravitational field, $\tau^{\mu\nu}$ upto quadratic orders in $h_{\mu\nu}$, we note that it can be expressed in the form

$$\begin{aligned}\tau^{\mu\nu}(t, r) &= C_1(\mu\nu\rho\sigma\alpha\beta\gamma\delta)h_{\rho\sigma,\alpha}h_{\beta\gamma,\delta} \\ &+ C_2(\mu\nu\rho\sigma\alpha\beta\gamma\delta)h_{\rho\sigma}h_{\alpha\beta,\gamma\delta}\end{aligned}$$

We calculate the time averaged value of this using the above phasor formula for $h_{\mu\nu}$. To compute gravitational radiation, only terms upto $(1/r^2)$ need to be retained. Writing

$$F_{\mu\nu}(\omega, \hat{r}) = -4\pi G \tilde{T}_{\mu\nu}(\omega, \omega\hat{r})$$

we have

$$h_{\mu\nu}(t, r) = \text{Re}(F_{\mu\nu}(\omega, \hat{r})\exp(j\omega(t - r)))/r$$

and hence we can write

$$h_{\mu\nu,0} = \text{Re}(j\omega F_{\mu\nu}(\omega, \hat{r})\exp(j\omega(t - r)))/r$$

$$h_{\mu\nu,m} = -e_m \text{Re}(j\omega F_{\mu\nu}(\omega, \hat{r})\exp(j\omega(t - r)))/r + O(1/r^2)$$

where

$$e_m = \hat{r}_m = x_m/r$$

is the m^{th} Cartesian component of the radial direction \hat{r} with $m = 1, 2, 3$.

2.1.3 Points to remember

[1] Gravity is not a force; it is a curvature of the space-time manifold.

[2] The weak principle of equivalence states that gravitational and inertial masses are proportional. Therefore, a gravitational field can be locally cancelled out by passing over to a locally inertial frame, ie, a freely falling elevator over a small interval of time.

[3] The weak principle of equivalence implies that laws of motion of particles can be formulated so that they are valid in all frames.

[4] The strong principle of equivalence states that all the laws of physics can be expressed in tensor notation, ie, they can be formulated in a way that they hold good in all reference frames. These laws include the laws of motion on classical and quantum levels, classical and quantum field theories including Maxwell's electromagnetic field theory, Yang-Mills non-Abelian gauge theories, their quantum versions and Einstein's theory of gravitation, both at the classical and quantum levels.

[5] The strong principle of equivalence in conjunction with Newton's inverse square law as a weak gravitational field approximation leads to the Einstein-Maxwell-Dirac equations for the gravitational field driven by the energy-momentum tensor of the Dirac matter field and the Maxwell radiation field.

[6] The weak field limit of the Einstein field equations along with a coordinate condition leads to the existence of gravitational waves in which the perturbed

metric tensor satisfies the usual wave equation. When matter and radiation sources are present in addition, then the weak field limit of the Einstein field equations leads to gravitational waves driven by the energy-momentum tensor of matter and radiation. The solution for the metric perturbation therefore appears as retarded potentials with sources coming from matter and electromagnetic radiation. It should be noted that the energy-momentum tensor of the Dirac matter field must be calculated by incorporating the spinor connection of the gravitational field and accordingly modifying the Lagrangian density of the Dirac field interacting with the Maxwell and gravitational fields. Likewise, if we are to include non-Abelian gauge and matter fields, then the corresponding Yang-Mills Lagrangian density must be modified by taking into account the spinor connection of the gravitational field apart from the non-Abelian gauge connection fields. After this incorporation, the energy-momentum tensor of the Yang-Mills matter and gauge fields must be computed and added to the right side of the Einstein field equations.

[7] The gravitational field has a pseudo energy-momentum tensor which when added to the energy momentum tensor of the matter, radiation and other non-Abelian fields, is conserved, ie, its ordinary four divergence is zero. The energy-momentum pseudo-tensor of the gravitational field is computed approximately from the second degree terms in the Einstein-Ricci tensor in the metric perturbations. By substituting the retarded potential solution for the metric perturbations with energy-momentum sources coming from Abelian and non-Abelian matter fields, radiation and non-Abelian gauge fields, into the expression for the energy-momentum pseudo-tensor for the gravitational fields, and integrating the corresponding energy flux over a large sphere of radius R , we can determine the total power radiated out in the form of gravitational waves just as we do in conventional antenna theory by integrating the Poynting vector over a large sphere with the electric and magnetic fields computed in the far field approximation from the current sources starting from the retarded potential formula.

[8] Particles moving in a gravitational field satisfy the geodesic equation, ie, the equation of shortest proper length on the curved space-time manifold. This space-time curve appears to the observer as accelerated motion.

[9] We can have a background gravitational field whose small perturbations travel as gravitational waves satisfying the wave equations with space-time dependent coefficients.

[10] Photons and material particles get affected by a gravitational wave. The interaction between photons and a gravitational field is described by the Lagrangian density of the em field in curved background space-time. Any Lagrangian density must be a scalar field, ie, invariant under all diffeomorphisms of the space-time manifold.

[11] The equations of motion of an em field in a curved background space-time that result from the Lagrangian formalism are the usual Maxwell equations but with ordinary partial derivatives replaced by covariant derivatives in order to get tensor equations for it is only tensor equations that are diffeomorphism invariant, ie, tensor laws hold in all reference frames.

2.2 Quantum gravity, the canonical ADM formalism-Schrodinger's equation for the wave function of the space-time metric

2.2.1 Summary

In the ADM formulation, one writes down the action for general relativity, ie, for the metric tensor in such a way that the spatial and temporal contributions of the metric are clearly separated out. In other words, one first decomposes the metric of space-time into a purely spatial component and a component that is orthogonal to the spatial component. This is achieved by first choosing a reference coordinate system x^μ in contrast to the given coordinate system X^μ . We then assume that the three dimensional surface $x^0 = t = \text{const}$ denoted by Σ_t is imbedded inside the four dimensional continuum (X^μ) of space-time. The vectors tangential to Σ_t are then precisely $(\partial X^\mu / \partial x^a = X_{,a}^\mu : \mu = 0, 1, 2, 3)$ where $a=1,2,3$. The vector $T^\mu = (X_{,t}^\mu)$ is not tangential to Σ_t but in general it is not even normal to this surface. So we subtract off a component from T^μ that is tangential to Σ_t and denote the resulting normal component by $N.n^\mu$ where N is a normalizing factor. The metric of space time in the X^μ system can then be decomposed as

$$g^{\mu\nu} = q^{\mu\nu} + n^\mu n^\nu$$

where

$$q^{\mu\nu} = q^{ab} X_{,a}^\mu X_{,b}^\nu$$

is purely spatial, ie it is constructed as a combination of outer product of vectors that are tangential to Σ_t . It is easily shown that the inverse of the 3×3 matrix $((q^{ab}))$ is $((\tilde{g}_{ab}))$ where $\tilde{g}_{\mu\nu}$ is the metric in the x^μ system. The next step is to define a spatial covariant derivative. This is done by choosing a spatial vector u_μ , ie, $u_\mu n^\mu = 0$ and constructing $\nabla_\mu u_\nu$ and then projecting it onto the spatial surface by considering

$$D_\mu u_\nu = q_\mu^{\mu'} q_\nu^{\nu'} \nabla_{\mu'} u_{\nu'}$$

We then iterate this process to define the double spatial covariant derivative of a spatial vector as

$$D_\rho D_\mu u_\nu = q_\rho^{\rho'} q_\mu^{\mu'} q_\nu^{\nu'} \nabla_{\rho'} D_{\mu'} u_{\nu'}$$

ie we repeatedly take the ordinary covariant derivative and project it onto the surface Σ_t . Thus, one defines the spatial Riemann curvature tensor as

$$[D_\mu, D_\nu] u_\rho = \tilde{R}_{\mu\nu\rho}^\alpha u_\alpha$$

where u_α is spatial. We then define the spatial tensor

$$K_{\mu\nu} = q_\mu^{\mu'} q_\nu^{\nu'} \nabla_{\mu'} n_{\nu'}$$

and prove that with neglect of a covariant divergence, the actual 4-D Riemann tensor $R_{\nu\rho\sigma}^\mu$ splits into the sum of a spatial component and a component that is a quadratic form in $K_{\mu\nu}$. Correspondingly, the true curvature scalar when multiplied by the invariant four volume element $\sqrt{-g(X)}d^4X$ and integrated over the whole of \mathbb{R}^4 splits into the sum of two components. The first component is purely spatial, ie, it is a function of q_{ab} and its spatial derivatives $q_{ab,c}, N, N^a$ only and the second contains $q_{ab,t}$. Here, N^a is defined by the decomposition

$$T^\mu = N^a X_{,a}^\mu + N n^\mu$$

into a purely spatial part, ie, tangential to Σ_t and a part Nn^μ that is normal to Σ_t . Further, $\sqrt{-g(X)}$ is expressible entire in a very elementary way in terms of N, N^a, q_{ab} . The time derivatives of N and N^a do not appear in the action integral. Thus this is a problem of constraints and rather than Poisson brackets, one must use the Dirac bracket. Further, since the component in $R\sqrt{-g}$ involving $q_{ab,t}$ is a quadratic form in $K_{\mu\nu}$ or equivalently in $K_{ab} = X_{,a}^\mu X_{,b}^\nu K_{\mu\nu}$, it follows that the momentum fields P^{ab} that are canonically conjugate to the position fields and are given by $\delta S/\delta q_{ab,t}$ are easily computed as linear forms in K_{ab} . This enables one to write down the Hamiltonian constraint operator in terms of P^{ab}, q_{ab}, N, N^a easily. It is from this discussion obvious that the Hamiltonian density of the gravitational field will be quadratic in P^{ab} and hence can be expressed as

$$\mathcal{H} = G_{abcd}(q)P^{ab}P^{cd} - \tilde{R}(q)\sqrt{q}$$

where $G_{abcd}(q)$ and $\tilde{R}(q)$ are functions of only q_{ab} and its spatial partial derivatives $q_{ab,c}$. This led Wheeler and De-Witt to write down the "Schrodinger equation of the universe" in the form

$$\left(\int (G_{abcd}\partial^2/\partial q_{ab}\partial q_{cd} + \tilde{R}(q)\sqrt{q})d^4X \right) \psi(q) = 0$$

This equation was subsequently analyzed by Hartle and Hawking in a paper called "The wave function of the universe" and solved for the De-Sitter metric justifying why the current radius of the universe measured by analyzing the red-shift of galaxies is the most probable.

2.2.2 Discussion

Quantum gravity–Schrodinger's wave equation for the metric of space-time

[a] Hamiltonian formulation of general relativity.

$$g^{\mu\nu} = q^{\mu\nu} + n^\mu n^\nu$$

$$q^{\mu\nu} n_\nu = 0$$

$$X_{,t}^\mu = T^\mu = N^\mu + N n^\mu$$

$$N^\mu n_\mu = 0$$

$$\begin{aligned} N^\mu &= N^a X_{,a}^\mu \\ q^{\mu\nu} &= q_{ab} X_{,a}^\mu X_{,b}^\nu \end{aligned}$$

Orthogonality of $X_{,a}^\mu$ and n^μ implies

$$g_{\mu\nu}(T^\mu - N^a X_{,a}^\mu)X_{,b}^\nu = 0 \quad (1)$$

So N^a satisfies

$$\tilde{g}_{0b} = g_{\mu\nu} T^\mu X_{,b}^\nu = q_{ab} N^a$$

where

$$\begin{aligned} q_{ab} &= g_{\mu\nu} X_{,a}^\mu X_{,b}^\nu = \tilde{g}_{ab} \\ g^{\mu\nu} &= \tilde{g}^{ab} X_{,a}^\mu X_{,b}^\nu + \tilde{g}^{0a}(X_{,0}^\mu X_{,a}^\nu + X_{,0}^\nu X_{,a}^\mu) + \tilde{g}^{00} X_{,0}^\mu X_{,0}^\nu \\ &= q^{ab} X_{,a}^\mu X_{,b}^\nu + n^\mu n^\nu \end{aligned}$$

where

$$q^{ab} = \tilde{g}^{ab}$$

and we make use of the orthogonality relations (1).

Spatial covariant derivative: Let u_μ be a spatial vector, ie, $X_{,a}^\mu u_\mu = 0$. Then $q^{\mu\nu} u_\nu = 0$ and we define

$$D_\mu u_\nu = q_\mu^{\mu'} q_\nu^{\nu'} \nabla_{\mu'} u_{\nu'}$$

Writing

$$u_a = u_\mu X_{,a}^\mu$$

we define

$$D_a u_b = X_{,a}^\mu X_{,b}^\nu D_\mu u_\nu$$

Then,

$$\begin{aligned} D_a u_b &= X_{,a}^\mu X_{,b}^\nu \nabla_\mu u_\nu = \\ &= X_{,a}^\mu X_{,b}^\nu (u_{\nu,\mu} - \Gamma_{\mu\nu}^\alpha u_\alpha) \\ &= X_{,b}^\nu (u_{\nu,a} - X_{,a}^\mu \Gamma_{\mu\nu}^\alpha u_\alpha) \\ &= (u_{b,a} - u_\nu X_{,ab}^\nu - X_{,a}^\mu X_{,b}^\nu \Gamma_{\mu\nu}^\alpha u_\alpha) \\ &= u_{b,a} - u_\nu X_{,ab}^\nu - \Gamma_{\alpha\mu\nu} X_{,a}^\mu X_{,b}^\nu X_{,c}^\alpha u^c \\ &= u_{b,a} - (q_{\mu\nu} X_{,c}^\mu X_{,ab}^\nu + \Gamma_{\alpha\mu\nu} X_{,a}^\mu X_{,b}^\nu X_{,c}^\alpha) u^c \\ &= u_{b,a} - \gamma_{cab} u^c \end{aligned}$$

where the spatial connection coefficients γ_{ab}^c are defined by

$$\gamma_{cab} = (q_{\mu\nu} X_{,c}^\mu X_{,ab}^\nu + \Gamma_{\alpha\mu\nu} X_{,a}^\mu X_{,b}^\nu X_{,c}^\alpha)$$

Now,

$$\begin{aligned} [D_\mu, D_\nu] u_\alpha &= D_\mu D_\nu u_\alpha - D_\nu D_\mu u_\alpha \\ D_\mu D_\nu u_\alpha &= q_\mu^{\mu'} q_\nu^{\nu'} q_\alpha^{\alpha'} \nabla_{\mu'} q_{\nu'}^{\nu''} q_{\alpha'}^{\alpha''} \nabla_{\nu''} u_{\alpha''} \end{aligned}$$

$$\begin{aligned}
& \nabla_{\mu'} q_{\nu'}^{\nu''} q_{\alpha'}^{\alpha''} \nabla_{\nu''} u_{\alpha''} \\
= & \nabla_{\mu'} (\delta_{\nu'}^{\nu''} - n^{\nu''} n_{\nu'}) (\delta_{\alpha'}^{\alpha''} - n^{\alpha''} n_{\alpha'}) \nabla_{\nu''} u_{\alpha''} \\
= & \delta_{\nu'}^{\nu''} \delta_{\alpha'}^{\alpha''} \nabla_{\mu'} \nabla_{\nu''} u_{\alpha''} \\
& - \nabla_{\mu'} \delta_{\nu'}^{\nu''} n^{\alpha''} n_{\alpha'} \nabla_{\nu''} u_{\alpha''} \\
& - \nabla_{\mu'} \delta_{\alpha'}^{\alpha''} n^{\nu''} n_{\nu'} \nabla_{\nu''} u_{\alpha''} \\
& + \nabla_{\mu'} n^{\nu''} n_{\nu'} n^{\alpha''} n_{\alpha'} \nabla_{\nu''} u_{\alpha''} \\
= & \nabla_{\mu'} \nabla_{\nu'} u_{\alpha'} \\
& - \nabla_{\mu'} n^{\alpha''} n_{\alpha'} \nabla_{\nu'} u_{\alpha''} \\
& - \nabla_{\mu'} n^{\nu''} n_{\nu'} \nabla_{\nu''} u_{\alpha'} \\
& + \nabla_{\mu'} n^{\nu''} n_{\nu'} n^{\alpha''} n_{\alpha'} \nabla_{\nu''} u_{\alpha''}
\end{aligned}$$

So using the orthogonality of n_μ with q_ν^μ the derivation property of the covariant derivative, we get

$$\begin{aligned}
D_\mu D_\nu u_\alpha = & \\
q_\mu^{\mu'} q_\nu^{\nu'} q_\alpha^{\alpha'} \nabla_{\mu'} \nabla_{\nu'} u_{\alpha'} & \\
- q_\mu^{\mu'} q_\nu^{\nu'} q_\alpha^{\alpha'} n^{\alpha''} (\nabla_{\mu'} n_{\alpha'}) \nabla_{\nu'} u_{\alpha''} & \\
- q_\mu^{\mu'} q_\nu^{\nu'} q_\alpha^{\alpha'} n^{\nu''} (\nabla_{\mu'} n_{\nu'}) (\nabla_{\nu''} u_{\alpha'}) &
\end{aligned}$$

Defining

$$K_{\mu\nu} = q_\mu^{\mu'} q_\nu^{\nu'} \nabla_{\mu'} n_{\nu'}$$

we have that $K_{\mu\nu} = K_{\nu\mu}$ by using the property that n_μ is the normal to a $3-D$ surface and hence of the form

$$n_\mu = F.G_{,\mu}$$

where F, G are scalar field. From the symmetry of the connection,

$$\begin{aligned}
\nabla_\mu n_\nu &= n_{\nu,\mu} - n_{\mu,\nu} \\
= F_{,\mu} G_{,\nu} - F_{,\nu} G_{,\mu} &= (\log F)_{,\mu} n_\nu - (\log F)_{,\nu} n_\mu
\end{aligned}$$

which implies that

$$q_\mu^{\mu'} q_\nu^{\nu'} \nabla_\mu n_\nu = 0$$

since

$$q_\mu^\mu n_\mu = 0$$

This proves the required symmetry of $K_{\mu\nu}$. In the course of the above proof, we have shown that

$$\begin{aligned}
[D_\mu, D_\nu] u_\alpha = & \\
q_\mu^{\mu'} q_\nu^{\nu'} q_\alpha^{\alpha'} [\nabla_{\mu'}, \nabla_{\nu'}] u_{\alpha'} &
\end{aligned}$$

$$\begin{aligned} & -q_\nu^{\nu'} n^{\alpha'} K_{\mu\alpha} \nabla_{\nu'} u_{\alpha'} \\ & + q_\mu^{\mu'} n^{\alpha'} K_{\nu\alpha} \nabla_{\mu'} u_{\alpha'} \end{aligned}$$

and since $u_\alpha n^\alpha = 0$ because u_α is assumed to be a spatial vector, we get

$$\begin{aligned} [D_\mu, D_\nu] u_\alpha &= \\ q_\mu^{\mu'} q_\nu^{\nu'} q_\alpha^{\alpha'} R_{\mu'\nu'\alpha'}^{\beta'} u_{\beta'} \\ + q_\nu^{\nu'} K_{\mu\alpha} (\nabla_{\nu'} n^{\alpha'}) u_{\alpha'} \\ - q_\mu^{\mu'} K_{\nu\alpha} (\nabla_{\mu'} n^{\alpha'}) u_{\alpha'} \\ = q_\mu^{\mu'} q_\nu^{\nu'} q_\alpha^{\alpha'} R_{\mu'\nu'\alpha'}^{\beta'} u_{\beta'} \\ + (K_{\mu\alpha} K_\nu^\beta - K_{\nu\alpha} K_\mu^\beta) u_\beta \end{aligned}$$

In deriving this formula, we have made use of the following identity for spatial vectors:

$$q_\mu^{\mu'} u_{\mu'} = u_\mu$$

We now note that

$$D_\nu n_\alpha = q_\nu^{\nu'} q_\alpha^{\alpha'} \nabla_{\nu'} n_{\alpha'} = K_{\nu\alpha}$$

Also,

$$\begin{aligned} R_{\mu\nu\alpha}^\beta n_\beta &= [\nabla_\mu, \nabla_\nu] n_\alpha \\ R_{\mu\nu} &= R_{\mu\nu\beta}^\beta = R_{\mu\nu\rho}^\beta \delta_\beta^\rho \\ &= R_{\mu\nu\rho}^\beta (q_\beta^\rho + n^\rho n_\beta) = \\ R_{\mu\nu\rho}^\beta q_\beta^\rho &+ R_{\mu\nu\rho}^\beta n^\rho n_\beta = \\ q_\beta^\rho R_{\mu\nu\rho}^\beta &+ n^\rho (n_{\mu:\nu:\rho} - n_{\mu:\rho:\nu}) \end{aligned}$$

Thus,

$$\begin{aligned} R &= g^{\mu\nu} R_{\mu\nu} = (q^{\mu\nu} + n^\mu n^\nu) (q_\beta^\rho R_{\mu\nu\rho}^\beta + n^\rho (n_{\mu:\nu:\rho} - n_{\mu:\rho:\nu})) \\ &= q^{\mu\nu} q_\beta^\rho R_{\mu\nu\rho}^\beta + \\ q^{\mu\nu} n^\rho (n_{\mu:\nu:\rho} &- n_{\mu:\rho:\nu}) + q_\beta^\rho R_{\mu\nu\rho}^\beta n^\mu n^\nu \\ + n^\mu n^\nu n^\rho (n_{\mu:\nu:\rho} &- n_{\mu:\rho:\nu}) \\ &= q^{\mu\nu} q_\beta^\rho R_{\mu\nu\rho}^\beta + \\ g^{\mu\nu} n^\rho (n_{\mu:\nu:\rho} &- n_{\mu:\rho:\nu}) + q_\beta^\rho R_{\mu\nu\rho}^\beta n^\mu n^\nu \\ = q^{\mu\nu} q_\beta^\rho R_{\mu\nu\rho}^\beta + \\ g^{\mu\nu} n^\rho (n_{\mu:\nu:\rho} &- n_{\mu:\rho:\nu}) + q_\beta^\rho R_{\mu\nu\rho}^\beta n^\mu n^\nu \\ = q^{\mu\nu} q_\beta^\rho R_{\mu\nu\rho}^\beta + \end{aligned}$$

$$\begin{aligned}
& n^\rho(n_{:\nu:\rho}^\nu - n_{:\rho:\nu}^\nu) + q_\beta^\rho R_{\mu\nu\rho}^\beta n^\mu n^\nu \\
&= n^\rho(n_{:\nu:\rho}^\nu - n_{:\rho:\nu}^\nu) - q^{\rho\beta} R_{\beta\nu\rho}^\mu n_\mu n^\nu \\
&= n^\rho(n_{:\nu:\rho}^\nu - n_{:\rho:\nu}^\nu) - q^{\rho\beta} n^\nu (n_{\beta:\nu:\rho} - n_{\beta:\rho:\nu}) \\
&= n^\rho(n_{:\nu:\rho}^\nu - n_{:\rho:\nu}^\nu) - (g^{\rho\beta} - n^\rho n^\beta) n^\nu (n_{\beta:\nu:\rho} - n_{\beta:\rho:\nu})
\end{aligned}$$

After neglecting a total covariant divergence, this quantity is the same as

$$\begin{aligned}
& -n_{:\rho}^\rho n_{:\nu}^\nu + n_{:\nu}^\rho n_{:\rho}^\nu \\
& -n^\nu(n_{:\nu:\rho}^\rho - n_{:\rho:\nu}^\rho) + n^\rho n^\beta n^\nu (n_{\beta:\nu:\rho} - n_{\beta:\rho:\nu}) \\
& = -n_{:\rho}^\rho n_{:\nu}^\nu + n_{:\nu}^\rho n_{:\rho}^\nu \\
& -n^\nu(n_{:\nu:\rho}^\rho - n_{:\rho:\nu}^\rho)
\end{aligned}$$

Again, after neglecting a total covariant divergence, this is the same as

$$\begin{aligned}
& -n_{:\rho}^\rho n_{:\nu}^\nu + n_{:\nu}^\rho n_{:\rho}^\nu \\
& + n_{:\rho}^\nu n_{:\nu}^\rho - n_{:\nu}^\nu n_{:\rho}^\rho \\
& = 2(n_{:\rho}^\nu n_{:\nu}^\rho - (n_{:\nu}^\nu)^2)
\end{aligned}$$

Now,

$$\begin{aligned}
K_{\rho'}^{\nu'} n_{:\nu'}^{\rho'} &= \\
q_\nu^{\nu'} q_{\rho'}^\rho n_{:\rho'}^\nu n_{:\nu'}^{\rho'} &= \\
n_{:\rho'}^{\nu'} n_{:\nu'}^{\rho'}
\end{aligned}$$

since $q_\nu^{\nu'} = \delta_\nu^{\nu'} + n^{\nu'} n_\nu$ and $n_\nu n_{:\rho}^\nu = 0$. Further,

$$\begin{aligned}
K_\rho^\nu n_{:\nu}^\rho &= K_{rho}^\nu (q_{\rho'}^\rho + n_{\rho'}^\rho) (q_\nu^{\nu'} + n_\nu^{\nu'}) n_{:\nu'}^{\rho'} \\
&= K_\rho^\nu K_\nu^\rho
\end{aligned}$$

since

$$K_\rho^\nu n^\rho = 0, K_\rho^\nu n_\nu = 0$$

Thus, we get

$$n_{:\rho}^\nu n_{:\nu}^\rho = K_\rho^\nu K_\nu^\rho$$

Likewise,

$$\begin{aligned}
K = K_\nu^\nu &= q_\rho^\nu n_{:\nu}^\rho = (\delta_\rho^\nu - n_\rho^\nu n_\rho) n_{:\nu}^\rho \\
&= n_{:\nu}^\nu
\end{aligned}$$

since

$$n_\rho n_{:\nu}^\rho = 0$$

Putting all this together, we get the result that R differs from

$$\tilde{R} + 2(K_\nu^\mu K_\mu^\nu - K^2)$$

by a total covariant divergence where \tilde{R} is the spatial part of the curvature, ie,

$$\tilde{R} = R_{\mu\nu\rho}^\beta q_\beta^\rho q^{\mu\nu}$$

We next note that for a spatial vector u_μ , ie, $q_\nu^\mu u_\mu = 0$, we can write

$$[D_\mu, D_\nu]u_\alpha = \hat{R}_{\mu\nu\alpha}^\beta u_\beta$$

We wish to show that

$$\hat{R} = q^{\mu\nu} q_\beta^\alpha \hat{R}_{\mu\nu\alpha}^\beta$$

depends only on $q_{ab}, q_{ab,c}, q_{ab,cd}$, ie on the spatial metric and its spatial partial derivatives. We note that we can write

$$\hat{R} = q^{ab} q^{cd} X_{,a}^\mu X_{,b}^\nu X_{,c}^\alpha X_{,d}^\beta \hat{R}_{\beta\mu\nu\alpha}$$

So it is enough to show that

$$\hat{R}_{abcd} = X_{,a}^\mu X_{,b}^\nu X_{,c}^\alpha X_{,d}^\beta \hat{R}_{\beta\mu\nu\alpha}$$

depends only on q_{ab} and its spatial partial derivatives, ie, w.r.t x^c . But that is obvious. Note that

$$q^{\mu\nu} = q^{ab} X_{,a}^\mu X_{,b}^\nu$$

$$((q^{ab})) = ((q_{ab}))^{-1}$$

where

$$q_{ab} = \tilde{g}_{ab}$$

We can write

$$D_\mu u_\alpha = q_\mu^{\mu'} q_\alpha^{\alpha'} \nabla_{\mu'} u_{\alpha'}$$

Also,

$$\begin{aligned} & q^{\mu'\mu} q^{\alpha'\alpha} \nabla_{\mu'} u_{\alpha'} = \\ & q^{ab} q^{cd} X_{,a}^{\mu'} X_{,b}^{\nu'} X_{,c}^{\alpha'} X_{,d}^{\beta'} \nabla_{\mu'} u_{\alpha'} \\ & = q^{ab} q^{cd} X_{,a}^{\mu'} X_{,b}^{\nu'} X_{,c}^{\alpha'} X_{,d}^{\beta'} (u_{\alpha',\mu'} - \Gamma_{\mu'\alpha'}^\sigma u_\sigma) \end{aligned}$$

Note further that we can write

$$\begin{aligned} u^\sigma &= X_{,a}^\sigma u^a, \\ & \Gamma_{\sigma\mu\nu} X_{,c}^\sigma X_{,a}^\mu X_{,b}^\nu \\ &= (1/2)(g_{\sigma\mu,\nu} + g_{\sigma\nu,\mu} - g_{\mu\nu,\sigma}) X_{,c}^\sigma X_{,a}^\mu X_{,b}^\nu \end{aligned}$$

and

$$g_{\sigma\mu,\nu} X_{,c}^\sigma X_{,a}^\mu X_{,b}^\nu =$$

$$(q_{\sigma\mu,nu} + (n_\sigma n_\mu)_{,\nu}) X_{,c}^\sigma X_{,a}^\mu X_{,b}^\nu \\ = q_{\sigma\mu,b} X_{,c}^\sigma X_{,a}^\mu$$

since

$$n_\mu X_{,a}^\mu = 0$$

Now,

$$q_{\sigma\mu} X_{,c}^\sigma X_{,a}^\mu = g_{\sigma\mu} X_{,c}^\sigma X_{,a}^\mu = \tilde{g}_{ca} = q_{ca}$$

since

$$g_{\sigma\mu} = q_{\sigma\mu} + n_\sigma n_\mu$$

and

$$n_\sigma X_{,c}^\sigma = 0$$

This gives

$$q_{ca,b} = q_{\sigma\mu,b} X_{,c}^\sigma X_{,a}^\mu + q_{\sigma\mu} (X_{,cb}^\sigma X_{,a}^\mu + X_{,c}^\sigma X_{,ab}^\mu)$$

By permutation of these spatial indices, we get

$$q_{cb,a} = q_{\sigma\mu,a} X_{,c}^\sigma X_{,b}^\mu + q_{\sigma\mu} (X_{,ca}^\sigma X_{,b}^\mu + X_{,c}^\sigma X_{,ba}^\mu)$$

$$q_{ab,c} = q_{\sigma\mu,c} X_{,b}^\sigma X_{,a}^\mu + q_{\sigma\mu} (X_{,bc}^\sigma X_{,a}^\mu + X_{,b}^\sigma X_{,ac}^\mu)$$

Adding the first two and subtracting the third equation gives us

$$q_{ca,b} + q_{cb,a} - q_{ab,c} = q_{\sigma\mu,b} X_{,c}^\sigma X_{,a}^\mu + q_{\sigma\mu,a} X_{,c}^\sigma X_{,b}^\mu - q_{\sigma\mu,c} X_{,c}^\sigma X_{,a}^\mu \\ + 2q_{\sigma\mu} X_{,c}^\sigma X_{,ab}^\mu$$

Thus, we get

$$\begin{aligned} D^\mu u^\alpha &= \\ q^{\mu' \mu} q^{\alpha' \alpha} \nabla_{\mu'} u_{\alpha'} &= \\ q^{ab} q^{cd} X_{,a}^{\mu'} X_{,b}^\mu X_{,c}^{\alpha'} X_{,d}^\alpha \nabla_{\mu'} u_{\alpha'} &= \\ = q^{ab} q^{cd} X_{,a}^{\mu'} X_{,b}^\mu X_{,c}^{\alpha'} X_{,d}^\alpha (u_{\alpha',\mu'} - \Gamma_{\sigma\mu'\alpha'} u^\sigma) & \\ = u_{\alpha',a} q^{ab} q^{cd} X_{,b}^\mu X_{,c}^{\alpha'} X_{,d}^\alpha & \\ - \Gamma_{\sigma\mu'\alpha'} q^{ab} q^{cd} X_{,a}^{\mu'} X_{,b}^\mu X_{,c}^{\alpha'} X_{,d}^\alpha u^\sigma & \\ = u_{\alpha',a} q^{ab} q^{cd} X_{,b}^\mu X_{,c}^{\alpha'} X_{,d}^\alpha & \\ - (1/2)[(q_{ea,c} + q_{ec,a} - q_{ac,e}) + q_{\sigma\rho} X_{,e}^\sigma X_{,ac}^\rho] u^e q^{ab} q^{cd} X_{,b}^\mu X_{,d}^\alpha & \end{aligned}$$

Now, we note that

$$q_{\sigma\mu} X_{,e}^\sigma u^e = q_{\sigma\mu} u^\sigma = u_\mu$$

and further,

$$u_{\alpha',a} X_{,c}^{\alpha'} = (u_{\alpha'} X_{,c}^{\alpha'})_{,a} - u_{\alpha'} X_{,ac}^{\alpha'}$$

$$= u_{c,a} - u_\rho X_{,ac}^\rho$$

and then after making the appropriate cancellations, we get

$$D^\mu u^\alpha =$$

$$q^{ab} q^{cd} X_{,b}^\mu X_{,d}^\alpha [u_{c,a} - \Gamma_{eac}^{(q)} u^e]$$

where

$$\Gamma_{eac}^{(q)} = (1/2)(q_{ea,c} + q_{ec,a} - q_{ac,e})$$

This result tells us that the spatial covariant derivative of a spatial vector has the same form of a usual covariant derivative but with the connection replaced by the connection of the spatial metric q_{ab} .

Some additional remarks: We write

$$T^\mu = X_{,0}^\mu = N^\mu + N n^\mu, N^\mu = N^a X_{,a}^\mu$$

so that to obtain orthogonality between $X_{,b}^\mu$, $b = 1, 2, 3$ and n^μ , we must choose N^a so that

$$g_{\mu\nu}(T^\mu - N^a X_{,a}^\mu)X_{,b}^\nu = 0$$

or equivalently,

$$\tilde{g}_{0b} - \tilde{g}_{ba}N^a = 0$$

Writing

$$q_{ab} = \tilde{g}_{ab}$$

and

$$q^{ab} = ((q_{ab}))^{-1}$$

we get

$$\begin{aligned} N^a &= q^{ab}\tilde{g}_{b0} \\ g^{\mu\nu} &= \tilde{g}_{\alpha\beta}X_{,\alpha}^\mu X_{,\beta}^\nu = \\ \tilde{g}^{00}X_{,0}^\mu X_{,0}^\nu &+ \tilde{g}^{0a}(X_{,0}^\mu X_{,a}^\nu + X_{,0}^\nu X_{,a}^\mu) + \\ \tilde{g}^{ab}X_{,a}^\mu X_{,b}^\nu &= \\ = \tilde{g}^{ab}X_{,a}^\mu X_{,b}^\nu &+ \tilde{g}^{00}(N^\mu + N n^\mu)(N^\nu + N n^\nu) \\ + \tilde{g}^{0a}(X_{,a}^\nu(N^\mu + N n^\mu) &+ X_{,a}^\mu(N^\nu + N n^\nu)) \\ = (\tilde{g}^{ab}X_{,a}^\mu X_{,b}^\nu &+ \tilde{g}^{0a}(N^\mu X_{,a}^\nu + N^\nu X_{,a}^\mu) + \tilde{g}^{00}N^\mu N^\nu) \\ + (\tilde{g}^{00}N^2 n^\mu n^\nu) & \\ + (N\tilde{g}^{00}(N^\mu n^\nu + N^\nu n^\mu) &+ N\tilde{g}^{0a}(n^\mu X_{,a}^\nu + n^\nu X_{,a}^\mu)) \end{aligned}$$

The first bracket is the purely spatial part, the second bracket is the part that is orthogonal to the spatial part and the third bracket is the cross-term part. We shall show that the cross term part vanishes and then evaluate explicitly the

spatial part and the part orthogonal to the spatial part. The cross term part vanishes because

$$\begin{aligned} & \tilde{g}^{00}N^\mu + \tilde{g}^{0a}X_{,a}^\mu \\ &= X_{,a}^\mu(\tilde{g}^{00}N^a + \tilde{g}^{0a}) \end{aligned}$$

and

$$\begin{aligned} & \tilde{g}_{ba}(\tilde{g}^{00}N^a + \tilde{g}^{0a}) \\ &= \tilde{g}^{00}\tilde{g}_{ba}N^a + (\delta_b^0 - \tilde{g}_{b0}\tilde{g}^{00}) \\ &= \tilde{g}^{00}(\tilde{g}_{ba}N^a - g_{b0}) = 0 \end{aligned}$$

implying that

$$N^\mu + \tilde{g}^{0a}X_{,a}^\mu = 0$$

This proves that the cross-term part vanishes. For the term normal to the spatial part, we find that assuming the normalization $n_\mu n^\mu = 1$,

$$\begin{aligned} N^2 &= g_{\mu\nu}(T^\mu - N^\mu)(T^\nu - N^\nu) \\ &= g_{\mu\nu}(T^\mu - N^\mu)T^\nu \end{aligned}$$

since $T^\mu - N^\mu$ is proportional to n^μ which is orthogonal to N^μ . Thus,

$$\begin{aligned} N^2 &= g_{\mu\nu}T^\mu T^\nu - g_{\mu\nu}N^a X_{,a}^\mu T^\nu \\ &= \tilde{g}_{00} - N^a \tilde{g}_{0a} \\ &= \tilde{g}_{00} - \tilde{g}_{0a} q^{ab} \tilde{g}_{0b} \\ &= 1/\tilde{g}^{00} \end{aligned}$$

where we use the fact that $((q^{ab}))$ is the inverse of $((q_{ab})) = ((\tilde{g}_{ab}))$. This proves that $\tilde{g}^{00}N^2 = 1$ and hence the term in the decomposition of $g^{\mu\nu}$ that is orthogonal to the spatial part is simply $n^\mu n^\nu$. Finally, we evaluate the spatial part and prove that it equals $q^{ab}X_{,a}^\mu X_{,b}^\nu$. The spatial part is given by

$$\begin{aligned} &= (\tilde{g}^{ab}X_{,a}^\mu X_{,b}^\nu + \tilde{g}^{0a}(N^\mu X_{,a}^\nu + N^\nu X_{,a}^\mu) + \tilde{g}^{00}N^\mu N^\nu) \\ &= X_{,a}^\mu X_{,b}^\nu(\tilde{g}^{ab} + \tilde{g}^{0a}N^b + \tilde{g}^{0b}N^a + \tilde{g}^{00}N^a N^b) \end{aligned}$$

So, we must prove that

$$(\tilde{g}^{ab} + \tilde{g}^{0a}N^b + \tilde{g}^{0b}N^a + \tilde{g}^{00}N^a N^b) = q^{ab} \quad (1)$$

Note that the LHS of this equation is given by

$$\tilde{g}^{ab} + \tilde{g}^{0a}q^{bc}\tilde{g}_{0c} + \tilde{g}^{0b}q^{ac}\tilde{g}_{0c} + \tilde{g}^{00}q^{ac}q^{bd}\tilde{g}_{0c}\tilde{g}_{0d}$$

We get on multiplying the LHS of (1) by q_{ca} , the result

$$q_{ca}(\tilde{g}^{ab} + \tilde{g}^{0a}N^b + \tilde{g}^{0b}N^a + \tilde{g}^{00}N^a N^b)$$

$$= q_{ca}\tilde{g}^{ab} + q_{ca}\tilde{g}^{0a}N^b + \tilde{g}^{0b}\tilde{g}_{c0} + \tilde{g}^{00}\tilde{g}_{c0}N^b \dots \quad (2)$$

where we made use of the identity

$$q_{ca}N^a = \tilde{g}_{c0}$$

We now observe that

$$q_{ca}\tilde{g}^{ab} = \tilde{g}_{ca}\tilde{g}^{ab} = \delta_c^b - \tilde{g}_{c0}\tilde{g}^{0b}$$

Thus, (2) equals

$$\begin{aligned} \delta_c^b + N^b(\tilde{g}^{00}\tilde{g}_{c0} + q_{ca}\tilde{g}^{0a}) \\ = \delta_c^b + N^b\delta_c^0 = \delta_c^b \end{aligned}$$

which completes the proof.

2.2.3 Points to remember

[1] In the quantum theory of gravity, we require to define in terms of the metric tensor appropriate position and momentum fields and then by applying the Legendre transformation, write down the Hamiltonian of the gravitational field in terms of these fields.

[2] Canonical quantum gravity can be derived by first choosing a reference frame x^μ and embedding the constant $x^0 = t$ 3-D surface Σ_t inside our given four dimensional space-time described by coordinates X^μ . We then consider the decomposition of $X_{,t}^\mu$ into a part that is tangential to Σ_t and a part normal to Σ_t . Such an orthogonal decomposition yields a corresponding orthogonal decomposition of the metric tensor in the X^μ frame into a part that is tangential to Σ_t and a part normal to Σ_t . Mixing terms cancel out. The vectors tangential to Σ_t are called spatial vectors and vectors normal to Σ_t are called normal vectors. Corresponding to this decomposition, we defined the spatial covariant derivatives of spatial vectors and hence the spatial components of the curvature tensor. Using this scheme, we decompose the invariant four dimensional integral of the curvature scalar R into the sum of a purely spatial part, ie involving only the spatial components of the metric tensor q_{ab} , $a, b = 1, 2, 3$ and a purely normal part. During this process, the spatial components of $X_{,t}^\mu$ denoted by N^a and the normalizing factor N of the unit normal that appears in this decomposition also appear. The purely spatial part of the curvature scalar contains only spatial derivatives of q_{ab} while the purely normal part contains time derivatives of q_{ab} which appear quadratically in the normal component of the curvature scalar and this facilitates computation of the momentum fields corresponding to q_{ab} and hence the Hamiltonian. The Hamiltonian equations for the corresponding wave functions of the metric therefore assume a Schrodinger like form first discovered by Wheeler and De-Witt, known as the wave-function of the universe. The other position fields N, N^a do not appear with time derivatives and hence the canonical momenta corresponding to them are zero. This compels us to use Dirac's theory of constraints leading to the Dirac bracket replacing the Poisson bracket. Hartle and Hawking have successfully used this equation in a restricted sense to calculate the probability of the De-Sitter universe having the current value and have shown this probability to be large thus justifying why our present universe is structured in this way.

2.3 Plasma waves

2.3.1 Summary

Consider a plasma made up of p species of charged particles where the k^{th} species particle has a charge e_k , mass m_k and a Boltzmann distribution function $f_k(t, r, v)$ in phase space. Its dynamics is described by the Boltzmann equation with a collision term that causes each species of particle to interact with particles of its own species and also with particles of the other species apart from its interaction with the electromagnetic field. Let $\sigma_{kj}(v_1, v, \hat{n})$ denote the scattering cross section for two particles of species k and j respectively. This means that if a particle of species k traveling with velocity v_1 interacts with a particle of species j traveling with velocity v with the minimum relative distance between the two particles being s and after the interaction, the relative velocity between the two particles $v'_1 - v'$ after the scattering process is directed along the direction \hat{n} relative to the relative line of initial motion $v_1 - v$, then the number of particles scattered within a unit solid angle around \hat{n} divided by the incident relative flux $|v_1 - v|$ can be calculated as

$$\sigma_{kj}(v_1, v, \hat{n}) = sdsd\phi/d\Omega(\hat{n})$$

or equivalently, assuming azimuthal symmetry,

$$\sigma_{kj}(v_1, v, \theta) = sds/\sin(\theta)d\theta$$

Then it is clear that the rate at which the number of particles of species k per unit phase volume in $r - v$ space is given by

$$(\partial f_k(t, r, v)/\partial t)_{coll} = \sum_j \int [-\sigma_{kj}(v_1, v, \hat{n})f_k(t, r, v)f_j(t, r, v_1) + \sigma_{jk}(v'_1(v_1, v, \hat{n}), v'(v_1, v, \hat{n}), \hat{n})f_j(t, r, v'_1)f_k(t, r, v')] d^3v_1 d\Omega(\hat{n})$$

where $v'_1(v_1, v, \hat{n})$, $v'(v_1, v, \hat{n})$ denote the initial velocities of particles of species j and k respectively which after scattering will cause the first particle to have a velocity v_1 and the second to have a velocity v so that the final relative velocities $v_1 - v$ is directed along the unit vector $-\hat{n}$ relative to the initial relative velocities $v'_1 - v'$. Conservation of the number of particles of the different species (provided that no chemical reactions between the particles takes place) implies that

$$\partial f_k(t, r, v)/\partial t + (v, \nabla_r) f_k(t, r, v) + (e_k/m_k)(E(t, r) + v \times B(t, r), \nabla_v) f_k(t, r, v) = (\partial f_k(t, r, v)/\partial t)_{coll}$$

We linearize the collision term around the Maxwell velocity distribution $f_{k0}(v)$ for each species and then denoting the change in the distribution function relative to the equilibrium Maxwell distribution by $\delta f_k(t, r, v)$ and assuming the electric and magnetic fields generated within the plasma to be small, we obtain after linearization the system of Vlasov equations:

$$\delta f_{k,t} + (v, \nabla_r) \delta f_k + (e_k/m)(E(t, r) + v \times B(t, r), \nabla_v) f_k(t, r, v) = -\delta f_k(t, r, v)/\tau(v),$$

$$\begin{aligned} \operatorname{div} E &= \sum_k e_k \int \delta f_k(t, r, v) d^3 v, \operatorname{div} B = 0, \\ \operatorname{curl} E &= -B_{,t}, \operatorname{curl} B = \mu J + \mu \epsilon E_{,t}, J = \sum_k e_k \int v \delta f_k(t, r, v) d^3 v \end{aligned}$$

We can substitute into these equations plane wave oscillatory solutions

$$\begin{aligned} E(t, r) &= \operatorname{Re}(E_0 \cdot \exp(j(\omega t - k \cdot r))), B(t, r) = \operatorname{Re}(B_0 \cdot \exp(j(\omega t - k \cdot r))), \\ \delta f_k(t, r, v) &= \operatorname{Re}(\delta f_{k0} \cdot \exp(j(\omega t - k \cdot r))) \end{aligned}$$

and by setting the coefficient determinant to zero derive the dispersion relation for plasma oscillations, ie, the relationship between the frequency ω and wave vector k .

2.3.2 Discussion

[a] The Boltzmann particle distribution function $f(t, r, v)$ is assumed to be a small perturbation of the Maxwell velocity distribution function $f_0(v) = (m/2\pi kT)^{3/2} \exp(-mv^2/2kT)$. The Maxwell distribution function satisfies Boltzmann's stationary equation in the absence of collisions and external electric field forces although a magnetic field may be imposed:

$$\begin{aligned} \partial f_0 / \partial t &= 0, (v, \nabla_r) f_0 = 0, \\ (v \times B, \nabla_v) f_0 &= 0 \end{aligned}$$

If however we impose an electrostatic field described by a potential $\Phi(r)$ and the plasma particle have a charge q , then f_0 is replaced by the Gibbs distribution function

$$f_0(r, v) = Z(\beta)^{-1} \exp(-\beta(mv^2/2 + q\Phi(r)))$$

where

$$\beta = 1/kT, Z(\beta) = \int \exp(-\beta(mv^2/2 + q\Phi(r))) d^3 v d^3 r$$

and then $f_0(r, v)$ again satisfies the stationary Boltzmann equation with electric field $E(r) = -\nabla\Phi(r)$, arbitrary magnetic field and zero collision forces:

$$f_{0,t} = 0, (v, \nabla_r) f_0(r, v) + q/m(-\nabla\Phi(r) + v \times B(t, r), \nabla_v) f_0(r, v) = 0$$

If we assume that the electric field is a weak perturbation of the electrostatic field $-\nabla\Phi$ and the magnetic field is also weak, ie

$$E(t, r) = -\nabla\Phi_0(r) + \delta E(t, r), B(t, r) = \delta B(t, r)$$

then the particle distribution function will be a weak perturbation of the equilibrium distribution $f_0(r, v)$ and we can write

$$f(t, r, v) = f_0(r, v) + \delta f(t, r, v)$$

and we obtain using first order perturbation theory with the relaxation time approximation for the collision term:

$$\delta f_{,t} + (v, \nabla_r) \delta f + (q/m)(\delta E + v \times \delta B, \nabla_v) f_0 = -\delta f / \tau(v)$$

which simplifies to

$$\delta f_{,t} + (v, \nabla_r) \delta f - (q/kT)(\delta E, v) f_0 + \delta f / \tau(v) = 0$$

Taking Fourier transforms w.r.t space and time, ie, defining

$$\hat{\delta}f(\omega, k, v) = \int \delta f(t, r, v) \exp(-i(\omega t - k \cdot r)) dt d^3r$$

we get

$$(i(\omega - (k, v)) + 1/\tau(v)) \hat{\delta}f(\omega, k, v) = (q/kT) \int_{\mathbb{R} \times \mathbb{R}^3} (v, \delta E(t, r)) f_0(r, v) \exp(-i(\omega t - k \cdot r)) dt d^3r$$

or equivalently,

$$\begin{aligned} \hat{\delta}f(\omega, k, v) &= (q/kT)(i(\omega - (k, v)) + 1/\tau(v))^{-1} \int (v, \delta E(t, r)) f_0(r, v) \exp(-i(\omega t - k \cdot r)) dt d^3r \\ &= (q/kT)(i(\omega - (k, v)) + 1/\tau(v))^{-1} \int (v, \delta \hat{E}(\omega, r)) f_0(r, v) \exp(-i(\omega t - k \cdot r)) dt d^3r \end{aligned}$$

Defining the kernel

$$K(\omega, r, v) = (q/kT)(2\pi)^{-3}(i(\omega - (k, v)) + 1/\tau(v))^{-1} \exp(i(k, r)) d^3k$$

we find that the above formula can be expressed after inverse Fourier transforming w.r.t the spatial arguments alone,

$$\hat{\delta}f(\omega, r, v) = \int K(\omega, r - r', v) f_0(r', v) (v, \delta \hat{E}(\omega, r')) d^3v d^3r'$$

The current density produced by this fluctuation in the particle distribution function is

$$\begin{aligned} J(t, r) &= q \int v \delta \hat{f}(\omega, v, r) d^3v = \\ &= q \int K(\omega, r - r', v) f_0(r', v) v (v, \delta \hat{E}(\omega, r')) d^3v d^3r' \\ &= \int \sigma(\omega, r, r') \delta \hat{E}(\omega, r') d^3r' \end{aligned}$$

where, the conductivity matrix kernel $\sigma(\omega, r, r') \in \mathbb{C}^{3 \times 3}$ is given by

$$\sigma(\omega, r, r') = q \int K(\omega, r - r', v) f_0(r', v) v v^T d^3v \in \mathbb{C}^{3 \times 3}$$

It should be noted at this stage that the electric and magnetic field fluctuations are determined by coupling the above formula for the conductivity to the Maxwell equations, ie,

$$\begin{aligned} \operatorname{curl} \delta \hat{E}(\omega, r) &= -j\omega \delta \hat{B}(\omega, r), \\ \operatorname{curl} \delta \hat{B}(\omega, r) &= \mu J(\omega, r) + j\omega \epsilon \delta \hat{E}(\omega, r) \\ &= \mu \int \sigma(\omega, r, r') \delta \hat{E}(\omega, r') d^3 r' + j\omega \mu \epsilon \delta \hat{E}(\omega, r) \end{aligned}$$

By combining, these two equations with the equations

$$\operatorname{div} \delta \hat{B} = 0, \operatorname{div} J(\omega, r) = -j\omega \rho(\omega, r), \operatorname{div} \delta \hat{E}(\omega, r) = \rho(\omega, r)/\epsilon$$

we get the generalized wave equation

$$\nabla(\operatorname{div} \delta \hat{E}(\omega, r)) - \nabla^2 \delta \hat{E}(\omega, r) = -j\omega \mu \int \sigma(\omega, r, r') \delta \hat{E}(\omega, r') d^3 r' + \omega^2 \mu \epsilon \delta \hat{E}(\omega, r)$$

or equivalently,

$$\begin{aligned} (\nabla^2 + \omega^2 \mu \epsilon) \delta \hat{E}(\omega, r) - (j/\omega \epsilon) \nabla(\operatorname{div}(\sigma \cdot \delta \hat{E})(\omega, r)) \\ - j\omega \mu (\sigma \cdot \delta \hat{E})(\omega, r) = 0 \end{aligned}$$

If $\|\sigma\| \ll \omega \epsilon$, this equation approximates to

$$(\nabla^2 + \omega^2 \mu \epsilon) \delta \hat{E}(\omega, r) - j\omega \mu (\sigma \cdot \delta \hat{E})(\omega, r) = 0$$

[b] Other approaches.

$$E(t, r) = E_0(r) + \delta E_1(t, r) + \delta^2 E_2(t, r) + \dots, E_0(r) = -\nabla \Phi(r)$$

$$B(t, r) = B_0(r) + \delta B_1(t, r) + \delta^2 B_2(t, r) + \dots$$

where E_k, B_k are of the k^{th} order of smallness. Likewise, for the distribution function,

$$f(t, r, v) = f_0(r, v) + \delta f_1(t, r, v) + \delta^2 f_2(t, r, v) + \dots$$

where

$$f_0(r, v) = C(\beta) \exp(-\beta(mv^2/2 + q\Phi(r)))$$

Substituting these into the Boltzmann equation

$$f_{,t} + (v, \nabla) f + (q/m)(E + v \times B, \nabla_v) f = \delta C(f)(t, r, v)$$

where

$$C(f) = \int \sigma(v_1, v, \hat{n}|v'_1, v')(f(t, r, v'_1)f(t, r, v') - f(t, r, v_1)f(t, r, v)) d^3 v_1 d\Omega(\hat{n})$$

where $v'_1 v'$ are easily computed as functions of v_1, v and the scattering direction \hat{n} by applying the energy and momentum conservation equations for binary elastic particle collisions:

$$v_1 + v = v'_1 + v', v_1^2 + v^2 = v'^2 + v'^2$$

The Boltzmann equation is to be combined with the Maxwell equations

$$\operatorname{div} E = \rho/\epsilon, \operatorname{div} H = 0, \operatorname{curl} E = -\mu H_{,t}, \operatorname{curl} H = \epsilon E_{,t} + \mu J$$

where

$$\rho(t, r) = \int f(t, r, v) d^3 v, J(t, r) = \int v f(t, r, v) d^3 v$$

The zeroth order equations are

$$(v, \nabla) f_0(r, v) + (q/m)(E_0(r) + v \times B_0(r)) f_0(r, v) = 0$$

which is easily verified to be identically satisfied. Further,

$$\operatorname{div} E_0 = -\nabla^2 \Phi(r) = \rho_0/\epsilon = (q/\epsilon) \int f_0(r, v) d^3 v = (qD(\beta)/\epsilon) \exp(-\beta\Phi(r))$$

Here,

$$D(\beta) = N / \int \exp(-\beta\Phi(r)) d^3 r$$

with N being the total number of charged particles. This equation must be solved for $\Phi(r)$. The first order equations are

$$\begin{aligned} & f_{1,t}(t, r, v) + (v, \nabla_r) f_1(t, r, v) + (q/m)(E_0(r) + v \times B_0(r)) f_1(t, r, v) + (q/m) E_1(t, r) f_0(r, v) \\ &= \int \sigma(v_1, v, \hat{n} | v'_1 v') (f_0(r, v'_1) f_0(r, v') - f_0(r, v_1) f_0(r, v)) d^3 v_1 d\Omega(\hat{n}) \quad (1) \end{aligned}$$

This is an equation for two unknown functions $f_1(t, r, v), E_1(t, r)$. To get another equation, we use the Maxwell equations

$$\operatorname{div} E_1 = \int f_1(t, r, v) d^3 v / \epsilon, \operatorname{curl} E_1 = -B_{1,t},$$

$$\operatorname{curl} B_1 = q\mu \int v f_1(t, r, v) d^3 v + \mu\epsilon E_{1,t}$$

These equations imply

$$\nabla(\operatorname{div} E_1) - \nabla^2 E_1 = -q\mu \int v f_{1,t} d^3 v - \mu\epsilon E_{1,tt}$$

or equivalently,

$$\nabla^2 E_1 - \mu\epsilon E_{1,tt} = q\mu \int v f_{1,t} d^3 v + (q/\epsilon) \nabla \left(\int f_1(t, r, v) d^3 v \right) \quad (2)$$

Note that $\rho_1 = \int q f_1 d^3 v$ and $\text{div} E_1 = \rho_1 / \epsilon$. In general, the k^{th} order equations are

$$\begin{aligned} & f_{k,t}(t, r, v) + (v, \nabla_r) f_k(t, r, v) + (q/m)(E_0(r) + v \times B_0(r), \nabla_v) f_k(t, r, v) \\ & + \sum_{m=1}^k (E_m(t, r) + v \times B_m(t, r), \nabla_v) f_{k-m}(t, r, v) \\ & = \sum_{m=0}^{k-1} \int \sigma(v_1, v, \hat{n}|v'_1, v') (f_m(t, r, v'_1) f_{k-m-1}(t, r, v) - f_m(t, r, v_1) f_{k-m-1}(t, r, v)) d^3 v_1 d\Omega(\hat{n}) \quad ---(3) \end{aligned}$$

This equation is to be solved for f_k, E_k, B_k in conjunction with the k^{th} order terms in the Maxwell equations, ie,

$$\text{div} E_k(t, r) = (q/\epsilon) \int f_k(t, r, v) d^3 v$$

$$\text{curl} E_k = -B_{k,t},$$

$$\text{div} B_k = 0,$$

$$\text{curl} B_k = \mu \int v f_k(t, r, v) d^3 v + \mu \epsilon E_{k,t}$$

These latter equations give on taking the curl and using the remaining Maxwell equations,

$$\nabla^2 E_k(t, r) - \mu \epsilon E_{k,tt}(t, r) = \mu \int v f_{k,t}(t, r, v) d^3 v + (q/\epsilon) \nabla \int f_k(t, r, v) d^3 v \quad ---(4)$$

and

$$\nabla^2 B_k(t, r) - \mu \epsilon B_{k,tt}(t, r) = \mu \nabla \times \int v f_k(t, r, v) d^3 v \quad ---(5)$$

2.3.3 Points to remember

[1] A system of charged particles of different species forms a plasma. The motion of these charged particles in motion generates electromagnetic fields apart from those externally applied. The equations governing the macroscopic dynamics of such a plasma and the electromagnetic fields is the Boltzmann kinetic transport equation for the particle distribution function along with the Maxwell equations for the em fields. Each charged species of the plasma has a distribution function in phase space at a given time where by phase space, we mean position-velocity space, ie, \mathbb{R}^6 . The Boltzmann kinetic transport equation essentially states that the number of particles of each species is conserved, which means that the material derivative of the each species distribution function over a small volume around a given point in phase space equals the rate at which that species of particles moves into this phase volume due to collisions with other particles

minus the rate at which particles from within this phase volume goes out due to collisions. Binary collision theory enables us to calculate the scattering cross sections for such problems thereby enabling us to evaluate the collision term in the transport equation.

[2] The Maxwell equations in a plasma are driven by the average charge and current densities at different space-time points obtained from the different charged species using their distribution function. The average is carried out by integrating the distribution function multiplied by unity or the velocity over the velocity space.

[3] The Maxwell velocity distribution satisfies the time independent Boltzmann equation in the absence of external forces and the collision term.

[4] The Gibbs distribution $\exp(-E(r, v)/kT)$ where the total energy $E(r, v) = mv^2/2 + \Phi(r)$ is the sum of the kinetic and time independent potential energy satisfies the time independent Boltzmann equation with force $-\nabla\Phi(r)$ per particle. This is a generalization of the previous point.

[5] When the stationary equilibrium distribution function is perturbed slightly by time varying electromagnetic perturbations with the relaxation time approximation taken for the collision term, then we can derive a dispersion relation for the oscillations of the perturbations in the distribution function and perturbations in the electromagnetic field by assuming these perturbations to travel like plane waves along any given direction. This dispersion relation gets modified in the presence of a strong constant magnetic field along a given direction. The dispersion relation between frequency and the wave vector is obtained by equating to zero the determinant of the system of linear algebraic equations obtained from the linearized Boltzmann equation and the Maxwell equations taking into account the charge and current density oscillations induced by the perturbation in the Boltzmann distribution function.

[6] If a small space-time varying electric field is applied to a Boltzmann particle in equilibrium, then its distribution function gets perturbed slightly and we can calculate the current density induced by this perturbed distribution function as a linear functional of the perturbing electric field. This calculation gives us the approximate conductivity tensor kernel in the frequency domain.

2.4 Evans-Hudson diffusion as a quantum mechanical generalization of the wave equation with noise

2.4.1 Discussion

We start our discussion with Evans-Hudson diffusion in quantum stochastics as generalizations of the heat and wave equations.

The wave equation in one dimension is

$$u_{,tt}(t, x) - c^2 u_{,xx}(t, x) = 0$$

We can write this as

$$u_{,t} = v, v_{,t} = c^2 u_{,xx}$$

or equivalently, combining this into a single vector pde, gives us

$$\frac{d}{dt} \begin{pmatrix} u \\ v \end{pmatrix} = L \begin{pmatrix} u \\ v \end{pmatrix}$$

where L is the matrix operator

$$L = \begin{pmatrix} 0 & 1 \\ c^2 \partial_x^2 & 0 \end{pmatrix}$$

Writing ξ for $[u, v]^T$, we can express this as

$$d\xi_t = L\xi_t dt$$

When stochastic terms are included, this becomes

$$d\xi_t = L\xi_t dt + \sigma dB_t$$

where B_t is two-vector valued standard Brownian motion field defined on \mathbb{R} . This is like a wave equation driven by a random source. Thus, if f is a smooth function defined on $L^2(\mathbb{R}) \otimes \mathbb{R}^2$, we have

$$\begin{aligned} df(\xi_t) &= f'(\xi_t)d\xi_t + (\sigma^2/2)f''(\xi_t)dt \\ &= (f'(\xi_t)L\xi_t + (\sigma^2/2)f''(\xi_t))dt + \sigma f'(\xi_t)dB_t \end{aligned}$$

This equation should be interpreted in the variational derivative sense, ie, in terms of Frechet/Gateaux derivatives. For example, suppose we let

$$f(\xi_t) = \int S(\xi_t(x))dx$$

then $f'(\xi_t)L\xi_t$ should be interpreted as

$$\int S'(\xi_t(x))L\xi_t(x)dx,$$

$f''(\xi_t)$ should be interpreted as

$$\int S''(\xi_t(x))dx$$

and $f'(\xi_t)dB_t$ should be interpreted as

$$\int f'(\xi_t(x))dB_t(x)dx$$

Here, we use the Ito formula in the form

$$(dB_t(x))^2 = dt$$

We can more generally, consider functionals of the form

$$f(\xi_t) = \int S(\xi_t(x_1), \dots, \xi_t(x_n))dx_1\dots dx_n$$

and then arrive at

$$\begin{aligned} df(\xi_t) &= \sum_k \int S_{,k}(\xi_t(x_1), \dots, \xi_t(x_n))d\xi_t(x_k)dx_1\dots dx_n \\ &\quad + (1/2) \sum_{k,m} \int S_{,km}(\xi_t(x_1), \dots, \xi_t(x_n))dx_t(x_k)d\xi_t(x_m)dx_1\dots dx_n \\ &= \left(\sum_k \int S_{,k}(\xi_t(x_1), \dots, \xi_t(x_n))L\xi_t(x_k)dx_1\dots dx_n \right) dt \\ &\quad + (\sigma/2) \left(\sum_{k,m} \int Tr(S_{,km}(\xi_t(x_1), \dots, \xi_t(x_n))R(t, x_k, x_m))dx_1\dots dx_n \right) dt \\ &\quad + \sigma \sum_k \int S_{,k}(\xi_t(x_1), \dots, \xi_t(x_n))dB_t(x_k)dx_1\dots dx_n \end{aligned}$$

where we assume the Ito formula

$$dB_t(x)dB_t(y)^T = R(t, x, y)dt$$

We define the linear operator j_t acting on the space of such functionals f defined on $L^2(\mathbb{R}) \otimes \mathbb{C}^2$ as

$$j_t(f) = f(\xi_t) = \int S(\xi_t(x_1), \dots, \xi_t(x_n))dx_1\dots dx_n$$

Then

$$j_t(cf + g) = cj_t(f) + j_t(g), j_t(fg) = j_t(f)j_t(g)$$

and the above Ito formula can be expressed in the form of a commutative Evans-Hudson flow as:

$$dj_t(f) = j_t(\theta_0(f))dt + \int j_t(\theta_1(x, f))dB_t(x)dx$$

where for $f : \mathcal{L}^2(\mathbb{R}) \otimes \mathbb{C}^2$ of the form

$$f(\xi) = \int S(\xi(x_1), \dots, \xi(x_n))dx_1 \dots dx_n$$

we define the linear operators $\theta_0, \theta(x, .)$ acting on the space of such functionals by

$$\begin{aligned} \theta_0(f) &= \sum_k \int S_{,k}(\xi(x_1), \dots, \xi(x_n))L\xi(x_k)dx_1 \dots dx_n \\ &+ (\sigma/2)(\sum_{k,m} \int Tr(S_{,km}(\xi_t(x_1), \dots, \xi_t(x_n))R(t, x_k, x_m))dx_1 \dots dx_n) \\ \theta_1(x, f) &= \sigma \sum_k \int S_{,k}(\xi(x_1), \dots, \xi(x_n))\delta(x - x_k)dx_1 \dots dx_n \\ &= \sigma \sum_k \int S_{,k}(\xi(x_1), \dots, \xi(x_{k-1}), \xi(x), \xi(x_{k+1}), \dots, \xi(x_n))dx_1 \dots dx_{k-1}dx_{k+1} \dots dx_n \end{aligned}$$

2.5 Waves in an expanding universe–Newtonian theory of small fluctuations and general relativistic theory of small fluctuations

2.5.1 Summary

In the general relativistic theory of small fluctuations, we start with the homogeneous and isotropic metric of Robertson and Walker (RW), ie,

$$d\tau^2 = dt^2 - \frac{S^2(t)}{1 - kr^2}(dr)^2 - S^2(t)r^2(d\theta^2 + \sin^2(\theta)d\phi^2)$$

and set up the Einstein field equations for the energy-momentum tensor corresponding to comoving velocity fields, ie,

$$V^0 = 1, V^r = 0, r = 1, 2, 3$$

and show that this velocity field satisfies the geodesic equation for the RW metric. We then set up the Einstein field equations taking as our energy-momentum tensor of matter

$$T^{\mu\nu} = (\rho(t) + p(t))V^\mu V^\nu - p(t)g^{\mu\nu}$$

and show that the RW metric is consistent with these field equations provided that the functions $S(t), \rho(t), p(t)$ satisfy a set of two coupled ordinary differential equations in t .

We then consider a Minkowskian metric $\eta_{\mu\nu}$ and assume that the Baryon number $n(x)$ satisfies the number conservation equation $(nv^\mu)_{;\mu} = 0$. Let $\sigma(x)$ denote the entropy per Baryon. Then by the first law of thermodynamics for each Baryon,

$$T(x)d\sigma(x) = d(\rho(x)/n(x)) + p(x)d(1/n(x))$$

where T is the temperature field. By combining this equation with the energy-momentum conservation equation

$$(T^{\mu\nu} + \Delta T^{\mu\nu})_{,\nu} = 0$$

where

$$T^{\mu\nu} = (\rho + p)v^\mu v^\nu - pg^{\mu\nu}$$

and $\Delta T^{\mu\nu}$ is the energy-momentum contribution due to viscous and thermal effects, we obtain an expression for the rate of entropy increase per unit volume in a comoving reference frame at a point in terms of $\Delta T^{\mu\nu}$ and the velocity field v^μ . The condition for entropy increase in accordance with the second law of thermodynamics then constrains the tensor $\Delta T^{\mu\nu}$ to have a specific form namely as

$$\Delta T^{\mu\nu} = H_\alpha^\mu H_\beta^\nu \tilde{\Delta} T^{\alpha\beta}$$

where

$$H_\nu^\mu = \delta_\nu^\mu - v^\mu v_\nu$$

and

$$-\chi_1(T)(v^{\mu,\nu} + v^{\nu,\mu} - (2/3)v^\rho_\nu \eta^{\mu\nu}) + \chi_2(T)(T^{\mu\nu} + T^{\nu\mu})$$

In the general theory of relativity, partial derivatives w.r.t. the space-time coordinates are replaced by covariant derivatives and the metric $\eta^{\mu\nu}$ of flat space-time is replaced by the metric $g^{\mu\nu}(x)$ of curved space-time.

The general relativistic theory of small fluctuations involves writing

$$v^\mu(x) = V^\mu + \delta v^\mu(x),$$

$$g_{\mu\nu}(x) = g_{\mu\nu}^{(0)} + \delta g_{\mu\nu}(x)$$

where $g_{\mu\nu}^{(0)}(x)$ is the RW metric and $\delta g_{\mu\nu}$ is the metric perturbation. Further, $\rho(x) = \rho_0(t) + \delta\rho(x)$ and $p(x) = p_0(t) + \delta p(x)$ where $\rho_0(t), p_0(t)$ are the density and pressures corresponding to the RW metric. We solve the Maxwell equations for electromagnetic wave propagation in the RW metric:

$$(F^{\mu\nu} \sqrt{-g^{(0)}})_{,\nu} = 0$$

where

$$F_{\mu\nu} = A_{\nu,\mu} - A_{\mu,\nu}, F^{\mu\nu} = g^{(0)\mu\alpha} g^{(0)\nu\beta} F_{\alpha\beta}$$

Once we obtain these unperturbed em waves, we form the electromagnetic energy-momentum tensor:

$$S^{\mu\nu} = (-1/4)F^{\alpha\beta} F_{\alpha\beta} g^{(0)\mu\nu} + F^{\mu\alpha} F^{\nu\beta} g_{\alpha\beta}^{(0)}$$

The linearized Einstein field equations in the presence of the CMBR (Cosmic microwave background radiation) are then given by

$$\delta R_{\mu\nu} = -8\pi G(\delta T^{\mu\nu} - (1/2)\delta T.g^{(0)\mu\nu} - (1/2)T\delta g^{\mu\nu}) - 8\pi GS^{\mu\nu}$$

This equation determines the dynamics of $\delta g_{\mu\nu}(x)$, $\delta v^r(x)$, $r = 1, 2, 3$, $\delta\rho(x)$ and $\delta p(x) = p'(\rho_0(t))\delta\rho(x)$. In the absence of the electromagnetic fields in the CMBR $F^{\mu\nu}$, we get homogeneous linear equations for the above perturbations which have oscillatory solutions while if the CMBR is taken into account, the former solutions form the general solution for the homogeneous part which has to be added to a particular solution of the inhomogeneous part to get the general solution for metric, velocity and density perturbations. It should be noted here, that since the CMBR contributions to the current evolution of the universe is very weak, we have treated its corresponding term $-8\pi GS^{\mu\nu}$ as being a first order perturbation. Once we have solved for the metric perturbations, we can then determine the next order correction to the CMBR by solving Maxwell's equations in the perturbed metric $g_{\mu\nu}^{(0)} + \delta g_{\mu\nu}$. The next order correction to the metric perturbations can then be determined by going back to the perturbed Einstein field equation and substituting in place of $S^{\mu\nu}$ its corrected version formed from the corrected version of $F^{\mu\nu}$.

2.5.2 Discussion

(a) The Newtonian theory of small fluctuations.

$$v(t, r) = H(t)r + \delta v(t, r)$$

where $H(t)$ is Hubble's constant. Likewise, we expand the density as

$$\rho(t, r) = \rho_0 + \delta\rho(t, r)$$

The pressure becomes

$$\begin{aligned} p(t, r) &= p(\rho_0) + p'(\rho_0)\delta\rho(t, r) \\ &= p_0 + K\delta\rho(t, r) \end{aligned}$$

on using the equation of state. The gravitational potential is also expanded as

$$\Phi = \Phi_0 + \delta\Phi(t, r)$$

The Navier-Stokes, Equation of continuity and Newton's inverse square law after linearization give

$$\begin{aligned} (H(t)r, \nabla)\delta v(t, r) + (\delta v(t, r), \nabla)H(t)r + \delta v_{,t}(t, r) &= \\ -K\nabla\delta\rho(t, r)/\rho_0 - \nabla\Delta\Phi(t, r) + \nu\nabla^2\delta v(t, r) &\dots\dots\dots(1) \\ \delta\rho_{,t}(t, r) + \text{div}(\rho_0\delta v(t, r) + \delta\rho(t, r)H(t)r) &= 0 \end{aligned}$$

or equivalently,

$$\delta\rho_{,t}(t, r) + \rho_0 \operatorname{div}(\delta v(t, r)) + H(t)(r, \nabla)\delta\rho(t, r) + 3H(t)\delta\rho(t, r) = 0 \quad (2)$$

$$\nabla^2\delta\Phi(t, r) = 4\pi G\delta\rho(t, r) \quad (3)$$

We look for a solution of the form

$$\delta v(t, r) = \delta v_0(t) \exp(ik.r/S(t)),$$

$$\delta\rho(t, r) = \delta\rho_0(t) \exp(ik.r/S(t)),$$

$$\delta\Phi(t, r) = \delta\Phi_0(t) \exp(ik.r/S(t))$$

where $S(t)$ is some function of time to be determined. This form of the solution corresponds to the cosmological principle that as the universe expands with the scale factor $S(t)$, the wavelength will also correspondingly increase proportionally to this scale factor. Substituting these expressions into (1) gives us

$$\begin{aligned} & iH(t)\delta v_0(t)(r, k)/S(t) + H(t)\delta v_0(t) + \delta v'_0(t) - \delta v_0(t)i(k, r)S'(t)/S^2(t) \\ &= -K(\delta\rho_0(t)/\rho_0)ik/S(t) - \delta\Phi_0(t)ik/S(t) + \nu\delta v_0(t)(-k^2/S^2(t)) \end{aligned}$$

It follows that we must necessarily have

$$H(t) = S'(t)/S(t),$$

$$H(t)\delta v_0(t) + \delta v'_0(t) + K\delta\rho_0(t)k/S(t) + \delta\Phi_0(t)ik/S(t) + \nu k^2\delta v_0(t)/S^2(t) = 0 \quad (4)$$

This is a vector ordinary differential equation that relates the 3-vectors $\delta v_0(t)$, k , $\delta v'_0(t)$ and the scalar $\delta\Phi_0(t)$. Likewise, (2) gives

(b) The general relativistic theory of small fluctuations.

(c) The Wheeler-De Witt equation: The Schrodinger wave equation of the universe. The Einstein-Hilbert action for the gravitational field has the form (Thomas Thiemann:Modern Canonical Quantum General Relativity”, Cambridge University Press) given by:

This leads us to the Wheeler-De-Witt equation for the Hamiltonian density

$$H = G^{abcd}P_{ab}P_{cd} + R\sqrt{-q}$$

where $q = ((q_{ab}))$ is the spatial part of the metric and R is the spatial part of the curvature scalar. Here, P_{ab} is obtained as the canonical momentum corresponding to the position field q_{ab} , ie, if S is the action, then

$$P_{ab} = \frac{\delta S}{\delta q_{ab,t}}$$

G^{abcd} are function of the position field q_{ab} and its spatial partial derivatives $q_{ab,c}$. Therefore the Schrodinger equation for the universe can be obtained by setting

$$P_{ab} = -i\delta/\delta q_{ab}$$

and it results in

$$(-G^{abcd}\frac{\delta^2}{\delta q_{ab}\delta q_{cd}} + R\sqrt{-g})\psi(q) = 0$$

This is the stationary state Schrodinger equation of the universe. There are other ways of determining the wave function of the universe, for example, by using Feynman's path integral. Let $S(g)$ denote the action and let Σ be a three dimensional surface that separates two four dimensional regions M_+ and M_- . For example, Σ can be the event horizon $r = 2GM/c^2$ of the Schwarzschild blackhole so that M_+ is the region outside the blackhole $r > 2GM/c^2$ and M_- is the region inside the blackhole $r < 2GM/c^2$. Then, the wave function ψ on the surface Σ may be obtained by the action integral $\psi(q_{ab}) = \int_{M_+} \exp(iI(g))dg$ where the boundary condition during the path integral evaluation is that the metric on Σ is fixed at q_{ab} . Hawking has shown the equivalence of these two definitions.

2.5.3 Points to remember

[1] In the Newtonian theory of small fluctuations, the unperturbed velocity field is assumed to follow Hubble's law $V(t, r) = H(t)r$ and one writes down the Navier-Stokes and mass conservation equation for this velocity field assuming that the unperturbed density and pressure fields are also functions of t only. When small perturbations depending on both position and time are introduced into the Navier-Stokes dynamics, the equation of mass conservation and the equation of state taking into account the Poisson equation for the inhomogeneous gravitational potential perturbation we get after linearization of these equations a set of linear pde's for the velocity, density and potential perturbations using which dispersion relations for oscillations of these perturbed quantities can be derived assuming that the wavelength of these plane waves scales proportionally to $S(t)$, the radius of the universe where $S'(t)/S(t) = H(t)$ arises as a natural consequence of the these equations. These dispersion relations give us information about how galaxies are formed and how they evolve with time once we understand that galaxies in our universe are small inhomogeneous perturbations of matter in an otherwise homogeneous and isotropic universe.

[2] In the general relativistic theory of small fluctuations, one starts with the Robertson-Walker metric for a homogeneous and isotropic universe and derives using the Einstein field equations with the assumption that the matter density and pressure are functions of coordinate time only, a set of two differential equations relating the radius $S(t)$ of the universe, the unperturbed density $\rho(t)$ and the pressure $p(t)$. The third equation is the equation of state which relates $p(t)$ to $\rho(t)$. Solving these equations with different assumptions yields the well known Friedmann models for our expanding universe. So much as far as there are no homogeneities. The unperturbed velocity field satisfies in this model the comoving property, ie, the choice of the Robertson-Walker coordinate system is such that particles which have fixed spatial coordinates at all times are the geodesics. This is just like putting a small blob on the surface of a balloon having a grid of latitudes and longitudes and noting that as the balloon expands, the latitude and longitude of the blob remain the same. It should be noted that the differential equations for $S(t), \rho(t), p(t)$ derived from general relativity can also be derived from basic Newtonian mechanics if we assume that our universe is an expanding ball having a fixed amount of matter within it distributed homogeneously. The Einstein field equations then have the interpretation that the net rate of change of energy of matter within a sphere of radius r is the rate at which pressure forces do work on this sphere from outside and the second Einstein field equation is obtained by writing down the conservation of the total kinetic and gravitational potential energy of a point mass placed on the surface of the expanding universe. Only the identification of the constants like the energy of the point mass in terms of curvature cannot be achieved by the Newtonian theory. Now we perturb the Robertson-Walker metric slightly by an inhomogeneous metric perturbation $\delta g_{\mu\nu}(t, r)$ with a coordinate gauge condition $\delta g_{0\mu}(t, r) = 0$ and also perturb the comoving velocity by $\delta v^k(t, r)$ and the density by $\delta\rho(t, r)$ and then taking into account the contribution to the energy-momentum tensor of matter coming from viscous and thermal effects, write down the perturbed/linearized Einstein field equations around the homogeneous and isotropic Robertson-Walker solution and derive partial differential equations for these perturbed quantities and from there the dispersion relations assuming that the wavelength of the perturbation waves scales proportionally to the wavelength. It should be noted that the viscous and thermal contributions to the energy-momentum tensor of the matter field are derived by assuming Baryon number conservation in conjunction with the first and second laws of thermodynamics. A complete account of this can be found in Steven Weinberg, "Gravitation and Cosmology; Principles and Applications of the General Theory of Relativity", Wiley.

2.6 EM waves in a curved space-time geometry with inhomogeneous permittivity-permeability tensor

2.6.1 Discussion

A^μ is the contravariant four potential. $A_\mu = g_{\mu\nu}A^\nu$ is the covariant four potential. The covariant EM field tensor is

$$F_{\mu\nu} = A_{n\nu,\mu} - A_{\mu,\nu}$$

In special relativity, $F_{\mu\nu}$ consist of E and B components. The D and H components are obtained respectively by operating on E and B by the permittivity tensor and inverse permeability tensor. In general relativity, we represent this tensor by $C^{\mu\nu\alpha\beta}(x)$ so that the tensor $\tilde{F}^{\mu\nu}$ which consists of the D and H components is given by

$$\tilde{F}^{\mu\nu} = C^{\mu\nu\alpha\beta}(x)F_{\alpha\beta}(x)$$

and the Maxwell equations

$$\operatorname{div}D = \rho, \operatorname{curl}H = J + \partial D/\partial t$$

of special relativity become in general relativity, the tensor equation

$$(\tilde{F}^{\mu\nu}\sqrt{-g}),_\nu = -J^\mu\sqrt{-g}$$

or equivalently,

$$(C^{\mu\nu\alpha\beta}(x)F_{\alpha\beta}(x)\sqrt{-g}),_\nu = -J^\mu\sqrt{-g}$$

We wish to derive the generalized wave equation from this for the EM four potential by imposing a gauge condition. The gauge condition imposed is

$$(A_\nu g^{\mu\nu}\sqrt{-g}),_\mu = 0$$

Writing in perturbation form,

$$C^{\mu\nu\alpha\beta}(x) = \eta^{\mu\alpha}\eta^{\nu\beta} + \delta C^{\mu\nu\alpha\beta}(x)$$

and

$$\sqrt{-g} = 1 + h/2, g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$$

we get approximately, ie, upto linear orders in $h_{\mu\nu}$ and in $\delta C^{\mu\nu\alpha\beta}(x)$,

$$[\eta^{\mu\alpha}\eta^{\nu\beta} + \delta C^{\mu\nu\alpha\beta}(x) + \eta^{\mu\alpha}\eta^{\nu\beta}h(x)/2]F_{\alpha\beta}(x)],_\nu = 0$$

in the absence of external currents and charges. Likewise, the gauge condition written in perturbative form is

$$[(\eta^{\mu\nu} - h^{\mu\nu} + \eta^{\mu\nu}h/2)A_\nu],_\mu = 0$$

Writing in perturbative form

$$A_\mu = A_\mu^{(0)} + A_\mu^{(1)}$$

this gauge condition gives on equating order zero and order one terms respectively,

$$\eta_{\mu\nu} A_{\nu,\mu}^{(0)} = 0,$$

$$\eta_{\mu\nu} A_{\nu,\mu}^{(1)} + ((h\eta_{\mu\nu}/2 - h^{\mu\nu}) A_\nu^{(0)})_{,\mu} = 0$$

For ease of notation, we write

$$A^{\mu(0)} = \eta_{\mu\nu} A_\nu^{(0)}, \quad A^{\mu(1)} = \eta_{\mu\nu} A_\nu^{(1)}$$

and then our gauge conditions read

$$A_{,\mu}^{\mu(0)} = 0,$$

$$A_{,\mu}^{\mu(1)} = (P_\nu^\mu A^{(0)\nu})_{,\mu}$$

where

$$P_\nu^\mu(x) = h_\nu^\mu(x) - h(x)\delta_\nu^\mu/2$$

where raising and lowering of indices are performed w.r.t the Minkowski metric $\eta_{\mu\nu}$. Our wave equation

$$[\eta^{\mu\alpha}\eta^{\nu\beta} + \delta C^{\mu\nu\alpha\beta}(x) + \eta^{\mu\alpha}\eta^{\nu\beta}h(x)/2]F_{\alpha\beta}(x)]_{,\nu} = 0$$

becomes on equating zeroth and first order terms,

$$\partial^\alpha \partial_\alpha A_\mu^{(0)} = 0,$$

$$\begin{aligned} & \eta_{\mu\alpha}\eta_{\nu\beta}(A_{\beta,\alpha\nu}^{(1)} - A_{\alpha,\beta\nu}^{(1)}) \\ & + (\delta C^{\mu\nu\alpha\beta} + \eta_{\mu\alpha}\eta_{\nu\beta}h/2)F_{\alpha\beta}^{(0)}_{,\nu} = 0 \end{aligned}$$

or equivalently, on using the gauge condition, with

$$\begin{aligned} & \square = \partial_\alpha \partial^\alpha = \eta_{\alpha\beta} \partial_\alpha \partial_\beta \\ & - \square A^{(1)\mu} + \eta_{\mu\alpha}(P_\sigma^\rho A^{(0)\sigma})_{,\rho\alpha} \\ & + ((\delta C^{\mu\nu\alpha\beta} + \eta_{\mu\alpha}\eta_{\nu\beta}h/2)F_{\alpha\beta}^{(0)})_{,\nu} = 0 \end{aligned}$$

This is the generalized first order perturbative approximation to the wave equation for the electromagnetic four potential in the presence of a gravitational field and a medium having non-uniform, anisotropic permittivity and permeability.

2.7 Quantum Optics. Here, the photon field is a quantum electromagnetic field expressible as a superposition of annihilation and creation operators of the photon field with the coefficients of the linear combination being positions of time and space

2.7.1 Discussion

$$A_\mu(t, r) = \sum_{k=1}^N a(k) f_{k\mu}(t, r) + a(k)^* \bar{f}_{k\mu}(t, r)$$

The Boson commutation rules are followed:

$$[a(k), a(m)^*] = \delta_{km}$$

This gives us

$$[A_\mu(t, r), A_\nu(t', r')] = 0, [A_\mu(t, r)^*, A_\nu(t', r')^*] = 0,$$

$$[A_\mu(t, r), A_\nu(t', r')^*] = \sum_k [f_{k\mu}(t, r) \bar{f}_{k\nu}(t', r') - \bar{f}_{k\mu}(t, r) f_{k\nu}(t', r')]$$

When this optical field (Electromagnetic four potential) is incident upon an atom described by canonical positions $Q = (Q_1, Q_2, Q_3)$ and momenta $P = (P_1, P_2, P_3)$ satisfying the commutation relations $[Q_i, Q_j] = 0, [P_i, P_j] = 0, [Q_i, P_j] = i\delta_{ij}$, then the interaction energy between the atom and the quantum EM field is given upto linear orders in the $a(k), a(k)^*$ by

$$H_I(t) = (ie/2m)((P, A) + (A, P)) - eA_0$$

where $A = (A_r)_{r=1}^3$. This interaction Hamiltonian can be expressed using $P = -i\nabla_r, Q = r$ as

$$\begin{aligned} H_I(t) &= (ie/2m)(-2i(A, \nabla) - idivA) - eA_0 \\ &= (e/m)(A(t, r), \nabla) + e(divA/2m - A_0) \end{aligned}$$

This interaction Hamiltonian is a special case of the following: The Hilbert space of the atom is \mathcal{H}_A and the Hilbert space of the EM field is \mathcal{H}_F . The atomic Hamiltonian is $H_A \otimes I_F$ while the field Hamiltonian is $I_A \otimes H_F$. The interaction Hamiltonian is of the form

$$H_I(t) = \sum_k (X_k(t, r) \otimes a(k) + X_k(t, r)^* \otimes a(k)^*)$$

for example in the above example taken from Schrodinger's wave mechanics, we have

$$\operatorname{div} A = \sum_k (f_{km,m}(t, r) \otimes a(k) + \bar{f}_{km,m}(t, r) \otimes a(k)^*)$$

and

$$(A, \nabla) = \sum_k (f_{km}(t, r) \partial_m \otimes a(k) + \bar{f}_{km}(t, r) \partial_m \otimes a(k)^*)$$

$$A_0 = \sum_k (f_{k0}(t, r) \otimes a(k) + \bar{f}_{k0}(t, r) \otimes a(k)^*)$$

It is therefore easily seen that the operators $f_{km}(t, r) \partial_m = f_{km}(t, Q) i P_m$ and $f_{k0}(t, Q)$ play the role of the atomic observables $X_k(t, r)$. These observables in the Q -representation become infinite dimensional matrices. (See for example the book by P.A.M. Dirac, "Principles of quantum mechanics", Oxford. The effect of the interaction between light and matter is to cause the matter to make transitions between its stationary state. If we adopt the interaction representation, meaning thereby, replace $X_k(t, r)$ by $\exp(itH_A)X_k(t, Q)\exp(-itH_A)$ and $a(k)$ by $\exp(-itH_F).a(k).\exp(-itH_F) = a(k)\exp(-i\omega_k t)$ where we assume that

$$H_F = \sum_k \omega_k a(k)^* a(k)$$

is the field Hamiltonian which can be derived using

$$H_F = (1/2) \int ((\partial A / \partial t)^2 + (\nabla \times A)^2) d^3 r$$

and we absorb the factor $\exp(-i\omega_k t)$ into $X_k(t, Q)$ and likewise for their adjoints, we can still represent our interaction Hamiltonian in the interaction picture in the form

$$H_I(t) = \sum_k (X_k(t, r) \otimes a(k) + X_k(t, r)^* \otimes a(k)^*)$$

where the atomic operators $X_k(t, r) = X_k(t, Q)$ are redefined as

$$\exp(-i\omega_k t).\exp(itH_A)X_k(t, Q).\exp(-itH_A)$$

Using first order perturbation theory, the approximate probability for the atom to make a transition from a stationary state $|n\rangle$ to a stationary state $|m\rangle$ in time T with the optical field remaining in the coherent state

$$|\phi(u)\rangle = \exp(-\|u\|^2/2)|e(u)\rangle$$

is given by

$$\begin{aligned} P_T(|n\rangle \rightarrow |m\rangle) &= |\int_0^T [\langle m|X_k(t, Q)|n\rangle \langle \phi(u)|a(k)|\phi(u)\rangle \\ &\quad + \langle m|X_k(t, Q)^*|n\rangle \langle \phi(u)|a(k)^*|\phi(u)\rangle] dt|^2 \\ &= |\int_0^T [\langle m|X_k(t, Q)|n\rangle u_k + \langle m|X_k(t, Q)|n\rangle \bar{u}_k] dt|^2 \end{aligned}$$

Suppose we take into account quantum noise in the sense of Hudson and Parthasarathy which means that some part of the quantum electromagnetic field is described by time varying operators, ie, the creation and annihilation process operators. Then we would have to replace u_k by $u_k(t)$ during the integration process.

The Poisson distribution for photons in quantum optics: Let $a(k), a(k)^*, k = 1, 2, \dots, N$ be as above. The energy of the photon field is given by

$$H_F = \sum_{k=1}^N \omega_k a(k)^* a(k)$$

The density operator corresponding to this Hamiltonian is

$$\rho = \exp(-\beta H_F) / \text{Tr}(\exp(-\beta H_F))$$

Let $|n_1, \dots, n_N\rangle, n_k = 0, 1, 2, \dots, k = 1, 2, \dots, N$ denote the standard orthonormal basis for the photon oscillators. Then,

$$H_F |n_1, \dots, n_k\rangle = \left(\sum_{k=1}^N \omega_k n_k \right) |n_1, \dots, n_N\rangle$$

So

$$\exp(-\beta H_F) |n_1, \dots, n_N\rangle = \exp\left(-\beta \sum_{k=1}^N \omega_k n_k\right) |n_1, \dots, n_N\rangle$$

and therefore the partition function for the photon field is given by

$$\begin{aligned} Z(\beta) &= \text{Tr}(\exp(-\beta H_F)) = \sum_{n_1, \dots, n_N \geq 0} \exp\left(-\beta \sum_{k=1}^N \omega_k n_k\right) \\ &= \prod_{k=1}^N (1 - \exp(-\beta \omega_k))^{-1} \end{aligned}$$

Now consider the coherent state of the photon field:

$$|u\rangle = |u_1, \dots, u_N\rangle = \exp(-\|u\|^2/2) \sum_{n_1, \dots, n_N \geq 0} u_1^{n_1} \dots u_N^{n_N} |n_1, \dots, n_N\rangle / \sqrt{n_1! \dots n_N!}$$

Then, we find for the probability of this state occurring,

$$P(u) = \langle \phi(u) | \rho | \phi(u) \rangle = \langle \phi(u) | \exp(-\beta H_F) | \phi(u) \rangle / Z(\beta)$$

Now,

$$\begin{aligned} \langle \phi(u) | \exp(-\beta H_F) | \phi(u) \rangle &= \exp(-|u|^2) \sum_{n \geq 0, m \geq 0} \langle n | \exp(-\beta H_F) | m \rangle u^m \bar{u}^n / \sqrt{m! n!} \\ &= \exp(-|u|^2) \sum_{n \geq 0} \exp(-\beta m \cdot \omega) |u|^{2m} / m! = \exp\left(\sum_k |u_k|^2 (\exp(-\beta \omega_k) - 1)\right) \end{aligned}$$

Thus,

$$P(u) = \frac{\exp(-\sum_k |u_k|^2(1 - \exp(-\beta\omega_k)))}{\prod_k (1 - \exp(-\beta\omega_k))}$$

We can also find out the probability of the state of photons being $|n_1, \dots, n_N\rangle$ given that the field is in a coherent state $|\phi(u)\rangle$. This probability distribution on the number of photons in a coherent state is

$$\begin{aligned} P(n_1, \dots, n_N|u) &= |\langle n_1, \dots, n_N | \phi(u) \rangle|^2 = \\ &\exp(-|u|^2)(\prod_{k=1}^N |u_k|^{2n_k} / n_k!) \end{aligned}$$

which is a product of N independent Poisson distributions having means $|u_k|^2, k = 1, 2, \dots, N$. In the Hudson-Parthasarathy theory of quantum stochastic processes, more can be said. Specifically, suppose we have a conservation operator $\lambda(H)$ where H is a Hermitian matrix. Then the moment generating function of this Boson Fock-space observable in the coherent state $|\phi(u)\rangle$ is given by

$$\begin{aligned} \langle \phi(u) | \exp(-\beta\lambda(H)) | \phi(u) \rangle &= \exp(-|u|^2) \langle e(u) | \exp(-\beta\lambda(H)) | e(u) \rangle \\ &= \exp(-|u|^2) \langle e(u) | e(\exp(-\beta H) u) \rangle = \exp(-|u|^2) \exp(\langle u | \exp(-\beta H) | u \rangle) \end{aligned}$$

If H has the spectral representation

$$H = \int_{\mathbb{R}} \lambda E(d\lambda)$$

then it follows that this moment generating function evaluates to

$$\begin{aligned} &\exp(-|u|^2) \exp\left(\int_{\mathbb{R}} \exp(-\beta\lambda) d \langle u | E(\lambda) | u \rangle\right) \\ &= \exp\left(-\int_{\mathbb{R}} (1 - \exp(-\beta\lambda)) d \langle u | E(\lambda) | u \rangle\right) \end{aligned}$$

which is the moment generating function of a Mixture of independent Poisson distributions. For example, suppose H is a finite Hermitian matrix of size $d \times d$ with spectral representation

$$H = \sum_{k=1}^d |e_k \rangle \langle e_k| \lambda_k$$

Then, the above moment generating function of $\lambda(H)$ is given by

$$\exp\left(-\sum_{k=1}^d (1 - \exp(-\beta\lambda_k)) |\langle e_k | u \rangle|^2\right)$$

which is the moment generating function of the classical random variable

$$X = \sum_{k=1}^d \lambda_k N_k$$

where N_1, \dots, N_d are independent classical Poisson random variables with means $| < e_k | u > |^2, k = 1, 2, \dots, d$ respectively.

Exercises:

[1] Let $z \in \mathbb{C}^n$ and define

$$|e(z)\rangle = 1 \oplus \bigoplus_{n \geq 1} z^{\otimes n} / \sqrt{n!}$$

Note that $|e(z)\rangle$ is the exponential vector in the Boson Fock space $\Gamma_s(\mathbb{C}^n)$ which is isomorphic as a Hilbert space to $L^2(\mathbb{R}^n)$. Show that

$$\langle e(z) | e(u) \rangle = \exp(\langle z | u \rangle), z, u \in \mathbb{C}^n$$

Deduce the Glauber-Sudarshan representation

$$\pi^{-n} \int |e(z)\rangle d^n z d^n \bar{z} \exp(-\|z\|^2) \langle e(z)| = I$$

where if $z = x + jy$, by $d^n z d^n \bar{z}$, we mean $d^n x d^n y$ and I is the identity operator in $\Gamma_s(\mathbb{C}^n)$. To prove this you can multiply the lhs of the above equation by $\langle e(w) |$ to the left and by $|e(v)\rangle$ to the right and show by elementary multivariate Gaussian integration that

$$\begin{aligned} & \pi^{-n} \int \langle e(w) | e(z) \rangle \langle e(z) | e(v) \rangle \exp(-\|z\|^2) d^n z d^n \bar{z} \\ &= \pi^{-n} \int \exp(\langle w | z \rangle) \exp(\langle z | v \rangle) \exp(-\langle z | z \rangle) d^n z d^n \bar{z} = \exp(\langle w | v \rangle) = \langle e(w) | e(v) \rangle \end{aligned}$$

[2] Let $X_k(t), k = 1, 2, \dots, p$ be $N \times N$ matrices depending on time and $(a_k, a_k^*), k = 1, 2, \dots, p$ define independent Harmonic oscillator annihilation-creation operator pairs, ie

$$[a_k, a_j^*] = \delta_{kj}, [a_k, a_j] = 0, [a_k^*, a_j^*] = 0$$

Solve the Schrodinger equation

$$i\psi_{,t}(t, \mathbf{r}_1, \dots, \mathbf{r}_p) = H(t)\psi(t, \mathbf{r}_1, \dots, \mathbf{r}_p)$$

where

$$H(t) = H_A \otimes I + I_N \otimes H_F + H_I(t)$$

where H_A is an $N \times N$ Hermitian matrix representing the Hamiltonian of the atom, $H_F = \sum_{k=1}^p \omega(k) a_k^* a_k$ acts in $L^2(\mathbb{R}^p)$ representing the total field Hamiltonian. $H_I(t)$ represents the interaction Hamiltonian between the atom and the field and has the form

$$H_I(t) = \sum_{k=1}^N (X_k(t) \otimes a_k + X_k(t)^* \otimes a_k^*)$$

Now write down the Schrodinger equation in the interaction representation for mixed states:

$$i\rho'(t) = [\tilde{H}_I(t), \rho(t)]$$

where $\tilde{H}_I(t)$ is the interaction picture Hamiltonian defined by

$$\tilde{H}_I(t) = \exp(it(H_A \otimes I + I \otimes H_F)).H_I(t).\exp(it(H_A \otimes I + I \otimes H_F))$$

$$= \sum_{k=1}^p (\tilde{X}_k(t) \otimes a_k + \tilde{X}_k(t)^* \otimes a_k^*)$$

where

$$\tilde{X}_k(t) = \exp(itH_A)X_k(t).\exp(-itH_A).\exp(-i\omega(k)t)$$

Now to solve this Schrodinger equation, we represent the interaction picture density as

$$\rho(t) = \pi^{-n} \int F(t, z, \bar{z}) \otimes |e(z)\rangle \langle e(z)| \exp(-\|z\|^2) d^n z d^n \bar{z}$$

where $F(t, z, \bar{z})$ is an $N \times N$ matrix. From the Schrodinger equation, derive a pde for F using the formulae

$$a_k \rho(t) = \pi^{-n} \int z_k F(t, z, \bar{z}) \otimes |e(z)\rangle \langle e(z)| \exp(-\|z\|^2) d^n z d^n \bar{z},$$

$$\begin{aligned} a_k^* \rho(t) &= \pi^{-n} \int F(t, z, \bar{z}) \otimes (\partial |e(z)\rangle / \partial z_k) \langle e(z)| \exp(-\|z\|^2) d^n z d^n \bar{z} \\ &= -\pi^{-n} \int (\partial(F(t, z, \bar{z}) \exp(-\|z\|^2)) / \partial z_k) \otimes |e(z)\rangle \langle e(z)| d^n z d^n \bar{z} \end{aligned}$$

on using integration by parts, and further, taking the adjoint of the above equations and using the fact that $F(t, z, \bar{z})$ is Hermitian matrix,

$$\rho(t) a_k^* = \pi^{-n} \int \bar{z}_k F(t, z, \bar{z}) \otimes |e(z)\rangle \langle e(z)| \exp(-\|z\|^2) d^n z d^n \bar{z},$$

$$\rho(t) a_k = -\pi^{-n} \int (\partial(F(t, z, \bar{z}) \exp(-\|z\|^2)) / \partial \bar{z}_k) \otimes |e(z)\rangle \langle e(z)| d^n z d^n \bar{z}$$

Derive from the Schrodinger equation in the interaction representation for mixed states, the following equivalent pde satisfied by F :

$$iF_{,t}(t, z, \bar{z}) =$$

$$\begin{aligned} &\sum_{k=1}^p [z_k \tilde{X}_k(t) F(t, z, \bar{z}) + \tilde{X}_k(t)^* (-\partial F(t, z, \bar{z}) / \partial z_k + \bar{z}_k F(t, z, \bar{z})) \\ &- \bar{z}_k F(t, z, \bar{z}) \tilde{X}_k(t)^* + (\partial F(t, z, \bar{z}) / \partial \bar{z}_k) - z_k F(t, z, \bar{z})] \tilde{X}_k(t) \end{aligned}$$

Reference: Mandel and Wolf, "Optical Coherence and Quantum Optics", Cambridge University Press.

Points to remember

[1] In quantum optics, we are concerned with the interaction of light with matter on an atomic scale. Thus, the Hilbert space for describing this interaction is the tensor product of the atomic Hilbert space and the em field Hilbert space. The total Hamiltonian is the sum of an atomic Hamiltonian acting in the atomic Hilbert space, the field Hamiltonian acting in the field Hilbert space, or more precisely a Boson Fock space and an interaction Hamiltonian acting in the tensor product of these two spaces.

[2] The quantum em field is described by a sequence of quantum harmonic oscillator creation and annihilation operators. These operators can either create or annihilate a photon with any specified momentum and helicity. The vector potential $A(t, r)$ of the quantum em field may thus be represented as a linear Hermitian combination of creation and annihilation operators of independent harmonic oscillators whose linear combination coefficients are vector valued functions of time and position. The position variables appearing in these linear combination coefficients should be regarded as atomic observables.

[3] The atomic Hamiltonian is the sum of the electron kinetic and potential energy of binding with the nucleus. It can therefore be expressed as a function of the position and momentum observables with the momentum observable being $-ih$ times the gradient operator w.r.t. the position observable.

[4] The interaction Hamiltonian between the quantum em field and the atom is obtained by replacing the momentum $P = -ih\nabla$ in the atomic Hamiltonian with $P + eA(t, r)$ and then adding to this Hamiltonian a term $-e\Phi(t, r)$ where Φ is the scalar electric potential obtained by applying the Lorentz gauge condition $\Phi_{,t}/c^2 + \operatorname{div} A = 0$. After expanding this, we find that the interaction Hamiltonian is expressible as a linear combination of the creation and annihilation operators tensored with atomic observables. This interaction Hamiltonian provides a general scheme for describing the interaction Hamiltonian of a quantum em field with any kind of atomic system like spin j systems. It is always a linear superposition of atomic operators tensored with the field creation operators and the adjoints of the corresponding atomic operators tensored with the field annihilation operators.

[5] Once the interaction Hamiltonian has been identified, it is an elementary matter to formulate the Schrodinger equation in the interaction picture. The idea is to use the Glauber-Sudarshan method of resolving the identity operator in Boson Fock space with an integral of the outer product of bra and ket coherent vectors of the bath w.r.t. their complex arguments. Thus, the density operator in the interaction picture can also be expanded as an integral of an atomic observable valued function $F(t, u, \bar{u})$ of time and the complex vector appearing in the Glauber-Sudarshan representation tensored with the outer product of bra and ket coherent vectors, substitute this expression into the interaction picture Schrodinger equation, use the elementary properties of the action of creation and annihilation operators on coherent vectors and thus transform the Schrodinger

equation into a partial differential equation for the function $F(t, u, \bar{u})$ that is first order in time and in the variables z, \bar{z} . Since for spin j atomic systems, F is a $2j+1 \times 2j+1$ matrix valued function, the complexity of the problem is greatly reduced from an infinite dimensional one for operators in the tensor product of the system Hilbert space \mathbb{C}^{2j+1} and the Boson Fock space to a finite dimensional one for operators in the finite dimensional system Hilbert space \mathbb{C}^{2j+1} .

[6] Conventional quantum communication can be described as the transmission of a mixed quantum state through a TPCP quantum channel, ie, trace preserving completely positive linear map. Such a communication can equivalently be described as picking one out of a set of p unitary operators with some probability, applying this operator in the adjoint representation to the input state. The result is convex linear combination of noiseless unitary channels acting on the input state. Recovery of the input state from the measured output state means the design of another TPCP map, ie, recovery operators that when applied to the received output state outputs the input state and the recovery operators should not depend on the input state, only on the noisy quantum channel operators. The fundamental result in this direction that ensures not only recovery of the input state under certain specified conditions but also gives an algorithm to construct the recovery operators from the channel operators is the Knill-Laflamme theorem which states the existence of recovery operators and how to construct them whenever the range of the input state is contained in a subspace \mathcal{C} called the code space and further the channel noise operators should be such that for any pair N_1, N_2 of them, $PN_2^*N_1P$ should be proportional to P where P is the orthogonal projection onto the code space. Intuitively this means that the noise operators should not displace the code space.

[7] The concept of measurement in quantum mechanics is not very easy to comprehend since it has no classical analogue. The Heisenberg uncertainty principle states that if two observables do not commute, then there exists at least one state in which the product of the variances of the two observables is positive. In particular, we have the major result that if q, p form a canonical position-momentum pair, so that $[q, p] = i\hbar$, then in any state ρ , we have

$$\text{Tr}((\rho(q - \langle q \rangle)^2(p - \langle p \rangle)^2) \geq \hbar^2/4$$

where for any observable X , $\langle X \rangle = \text{Tr}(\rho X)$. This principle means that if two observables do not commute, they do not have a joint probability distribution at least in one state. This also means that if we measure one of such a pair of observables and note its outcome, then the state of the system collapses to the state in which that observable has the outcome measured. In that collapsed state, the second observable will have infinite variance so that the product of the two variances in this collapsed state ($0 \times \infty$) is positivity. This principle of collapse can be formulated first in the following simplified form: Let ρ be the state of a system and X an observable having spectral decomposition $X = \sum_k |e_k\rangle x(k) \langle e_k|$ where $x(k), k = 1, 2, \dots$ are distinct real numbers. If we measure X and note its outcome as $x(k)$ (which will occur with a probability

of $\text{Tr}(\rho|e_k\rangle\langle e_k|) = \langle e_k|\rho|e_k\rangle$, then after this measurement, the state of the system collapses to the pure state $|e_k\rangle$ or equivalently $|e_k\rangle\langle e_k|$. If however some of the eigenvalues of X are equal, we can write

$$X = \sum_k c(k)P_k, \quad \sum_k P_k = I$$

where $c(k), k = 1, 2, \dots$ are the distinct eigenvalues of X and

$$P_k = \sum_{m:x(m)=c(k)} |e_m\rangle\langle e_m|$$

is the orthogonal onto the eigenspace of X with eigenvalue $c(k)$. This subspace can also be expressed as $\mathcal{N}(X - c(k))$. Then, if X is measured and $c(k)$ is the noted outcome, the state of the system collapses to

$$\rho_k = \frac{P_k \rho P_k}{\text{Tr}(\rho P_k)}$$

If however, X is measured but the outcome is not noted, then the state of the system collapses to

$$\sigma = \sum_k p(k)\rho_k = \sum_k P_k \rho P_k$$

where

$$p(k) = \text{Tr}(\rho P_k)$$

is the probability of getting the outcome $c(k)$.

The notion of measurement involved here is the PVM, ie projection valued measurement $\{P_k\}$, ie, we can forget about the observable X and consider a PVM to consist of just operators $\{P_k\}$ such that $P_k^* = P_k = P_k^2$ and $\sum_k P_k = I$. The most general kind of measurement that includes PVM as a special case is a POVM, ie, a positive operator valued measurement which consists of a family $\{M_k\}$ of operators with $M_k \geq 0, \sum_k M_k = I$. In that case if this POVM is measured and the k^{th} outcome associated with M_k is measured, then the system collapses to the state

$$\rho_k = \frac{\sqrt{M_k} \rho \sqrt{M_k}}{\text{Tr}(\rho M_k)}$$

If however, the measurement $\{M_k\}$ is made but the outcome is not noted, then the system collapses to the state

$$\sigma = \sum_k \sqrt{M_k} \rho \cdot \sqrt{M_k}$$

Note that this is a valid state since

$$\sigma \geq 0, \text{Tr}(\sigma) = \sum_k \text{Tr}(\rho M_k) = \text{Tr}(\rho \cdot \sum_k M_k) = \text{Tr}(\rho) = 1$$

2.8 Quantum optics, notion of a generalized measurement, state collapse after quantum measurement, recovery of states passed through a noisy quantum system, the Knill-Laflamme theorem Stinespring's representation of noisy quantum systems, Information, relative entropy, mutual information and Renyi entropy of quantum systems. Transmission of information over quantum system. The relevance of all this to the wave mechanics of Schrodinger

2.8.1 Discussion

Let ρ_{AB} be a mixed state of a quantum system in the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. A' 's state is then

$$\rho_A = Tr_B(\rho_{AB})$$

while B' 's state is

$$\rho_B = Tr_A(\rho_{AB})$$

Example: ρ_{AB} or for that matter any linear operator in $\mathcal{H}_A \otimes \mathcal{H}_B$ can be represented as

$$\rho_{AB} = \sum_{k=1}^d X_{Ak} \otimes X_{Bk}$$

where X_{Ak} acts in \mathcal{H}_A while X_{Bk} acts in \mathcal{H}_B . A' 's state is then

$$\rho_A = Tr_B(\rho_{AB}) = \sum_k Tr(X_{Bk}) X_{Ak}$$

and B' 's state is

$$\rho_B = Tr_A(\rho_{AB}) = \sum_k Tr(X_{Ak}) X_{Bk}$$

A unitary evolution of a state is given by

$$\rho \rightarrow U\rho U^*$$

where

$$UU^* = U^*U = I$$

Such an evolution preserves the trace, positivity and the Von-Neumann entropy:

$$H(\rho) = -Tr(\rho \log(\rho)) = H(U\rho U^*)$$

Such an evolution also transforms a pure state to a pure state:

$$|\psi><\psi| \rightarrow U|\psi><\psi|U^* = (U|\psi>)(U|\psi>)^*$$

Since the entropy of a pure state is zero, it follows that a unitary evolution transforms a zero entropy state to a zero entropy state. We now note that if ρ_{AB} is a mixed state in $\mathcal{H}_A \otimes \mathcal{H}_B$, then it has a spectral representation

$$\rho_{AB} = \sum_k \lambda_k |e_k><e_k|$$

where $\lambda_k \geq 0$, $\sum_k \lambda_k = 1$ and $\{|e_k>\}$ is an ONB for $\mathcal{H}_A \otimes \mathcal{H}_B$. The state

$$|x> = \sum_k \sqrt{\lambda_k} |e_k \otimes f_k>$$

is a pure state in $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_R$ where \mathcal{H}_R is another Hilbert space with ONB $\{f_k\}$. We have

$$Tr_R(|x><x|) = \sum_k \lambda_k |e_k><e_k| = \rho_{AB}$$

Thus $|x>$ is a purification of ρ_{AB} with reference system R . Further in the state $|x>$ for \mathcal{H} , the state of the system R is given by

$$\rho_R = Tr_{AB}(|x><x|) = \sum_k \lambda_k |f_k><f_k|$$

It is therefore easy to see that

$$H(\rho_{AB}) = - \sum_k \lambda_k \log(\lambda_k) = H(\rho_R)$$

More generally, if we have Hilbert spaces \mathcal{H}_k , $k = 1, 2, \dots, N$ and we choose a set $I = \{i_1, \dots, i_k\} \subset \{1, 2, \dots, N\}$ with complementary set $J = \{j_1, \dots, j_{N-k}\}$, then if $|x>$ is a pure state in $\otimes_{k=1}^N \mathcal{H}_k$, we define the states

$$\rho_I = Tr_{\mathcal{H}_J}(|x><x|), \rho_J = Tr_{\mathcal{H}_I}(|x><x|)$$

where

$$\mathcal{H}_I = \otimes_{m=1}^k \mathcal{H}_{i_m}, \mathcal{H}_J = \otimes_{m=1}^{N-k} \mathcal{H}_{j_m}$$

then it is easy to see that

$$H(\rho_I) = H(\rho_J)$$

Now consider a random Schrodinger evolution, ie, with probability $p(k)$, we apply the unitary U_k for $k = 1, 2, \dots, n$ to a quantum state where $\sum_k p(k) = 1$. Then the output state is

$$T(\rho) = \sum_k p(k) U_k \rho U_k^*$$

Obviously, T is a trace preserving completely positive (TPCP) map. T transforms a pure state to a pure state if $p(k) = 1$ for some k and $p(j) = 0$ for $j \neq k$. In general, T transforms a pure state to a mixed state and hence generates entropy. Now let E_1, \dots, E_n be operators such that

$$\sum_{k=1}^n E_k^* E_k = I$$

and define

$$T(\rho) = \sum_k E_k \rho E_k^*$$

Again T is a TPCP map and thus describes noisy quantum evolution. It is easy to see that the above random Schrodinger map is a special case of this by letting $E_k = \sqrt{p(k)} U_k$. Note that $E_k : \mathcal{H} \rightarrow \mathcal{H}'$ where \mathcal{H} is the input Hilbert space and \mathcal{H}' is the output Hilbert space. We define the operator

$$E = [E_1, E_2, \dots, E_n] : \mathcal{H} \otimes \mathbb{C}^n \rightarrow \mathcal{H}'$$

Then, we can write

$$T(\rho) = E(I_n \otimes \rho)E^*$$

We now define

$$F = \begin{pmatrix} E_1 \\ E_2 \\ \dots \\ E_n \end{pmatrix}$$

F maps \mathcal{H} into $\mathcal{H}' \otimes \mathbb{C}^n$ and we have the result that

$$F^* F = \sum_k E_k^* E_k = I$$

where I is the identity in \mathcal{H} . Further, we have

$$Tr_1(F \rho F^*) = \sum_k E_k \rho E_k^* = T(\rho)$$

Thus, F is an isometry mapping \mathcal{H} into $\mathcal{H}' \otimes \mathbb{C}^n$. F can be extended to a unitary matrix in $\mathcal{H}' \otimes \mathbb{C}^n$ by adding columns to it. We denote this unitary matrix by

$$U = [F|G]$$

Then, we have since

$$F = U \begin{pmatrix} I \\ 0 \end{pmatrix} = US$$

say, where

$$S = \begin{pmatrix} I \\ 0 \end{pmatrix}$$

that

$$T(\rho) = Tr_1(US\rho S^*U^*)$$

We note that

$$\begin{aligned} S\rho S^* &= \begin{pmatrix} \rho & 0 \\ 0 & 0 \end{pmatrix} \\ &= |e_0>< e_0| \otimes \rho \end{aligned}$$

where

$$|e_0> = \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix}$$

so that

$$|e_0>< e_0| = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

This gives us finally Stinespring's representation:

$$T(\rho) = Tr_1(U(|e_0>< e_0| \otimes \rho)U^*)$$

2.8.2 Points to remember

[1] Given a joint mixed state of two systems A and B , we can obtain the marginal state of A by partial tracing out the original state over B and likewise, the marginal state of B by partial tracing over A .

[2] A quantum operation involving noise acting on a state can be modeled by a TPCP map which in turn is equivalent to applying one of a set of unitary gates to a mixed state with some probability.

[3] Stinespring's representation of a quantum operation on a mixed state states that any quantum noisy operation on a mixed state can be obtained by adjoining another Hilbert space with the original one via the tensor product map, choosing a pure state in the second Hilbert space also called the bath, then tensoring the original state with this pure state and finally applying a unitary operator in the adjoint representation on this tensor product state and finally tracing out over the second Hilbert space, ie, over the bath space. This process has a nice physical interpretation: First couple the system with the noisy bath, evolve jointly the system and bath according to unitary dynamics so that the total probability of a particle being either in the system or in the bath is unity, and then average over the bath variables. When the partial tracing over the

bath is performed, unitarity is broken and a pure state on joint system and bath space becomes a mixed state on system space. This means that quantum noise is introduced into the system from the bath and when we look at the system alone, we do not get unitary dynamics. Although we have described this situation in the static case, the Hudson-Parthasarathy quantum Ito theory explains how this can be achieved dynamically. In other words, the GKSL equation for the mixed state is TPCP but not unitary and can be dilated into a unitary evolution by tensor producting the original system with a bath Hilbert space. The physical interpretation of tracing out over bath variables resulting in the initial pure state of the system becoming mixed after evolution represents pumping in of entropy from the bath into the system. In summary, unitary evolution implies noiseless evolution implies conservation of entropy while non-unitary evolution implies noisy evolution implies change in entropy.

2.9 Controlling the quantum em field produced by electrons and positrons by using a classical em field-An application of Dirac's relativistic wave equation

2.9.1 Discussion

Statement of the problem: The second quantized Dirac field $\psi(x)$ in the presence of a classical em field $A_\mu^{(c)}(x)$ satisfies the Dirac equation

$$[\gamma^\mu(i\partial_\mu + eA_\mu^{(c)}(x)) - m]\psi(x)$$

We assume this classical em field to be small and expand ψ as a first order perturbation series:

$$\psi(x) = \psi^{(0)}(x) + \psi^{(1)}(x)$$

Then using first order perturbation theory, we get

$$[\gamma^\mu(i\partial_\mu) - m]\psi^{(0)}(x) = 0$$

$$[\gamma^\mu i\partial_\mu - m]\psi^{(1)}(x) = -e\gamma^\mu A_\mu^{(c)}(x)\psi^{(0)}(x)$$

So $\psi^{(0)}$ can be expanded in the usual way in terms of electron-positron creation and annihilation operators,

$$\psi^{(0)}(x) = \int [u(P, \sigma)a(P, \sigma)\exp(-ip.x) + v(P, \sigma)b(P, \sigma)^*\exp(ip.x)]d^3P$$

where

$$p^0 = E(P) = \sqrt{m^2 + P^2}$$

The momentum space wave functions $u(P, \sigma), v(P, \sigma)$ satisfy the free Dirac equation:

$$(\gamma.p - m)u(P, \sigma) = 0, (\gamma.p + m)v(P, \sigma) = 0, \sigma = 1, 2$$

These are obtained by substituting $\psi^{(0)}(x)$ into Dirac's equation. We note that with $\alpha^r = \gamma^0\gamma^r, r = 1, 2, 3$, we have

$$((\alpha, P) + \beta m)u(P, \sigma) = E(P)u(P, \sigma),$$

$$((\alpha, P) + \beta m)v(-P, \sigma) = -E(P)v(-P, \sigma)$$

confirming that a negative energy electron when removed is equivalent to a positive energy positron. The solution for $\psi^{(1)}$ is given by

$$\psi^{(1)}(x) = \int K(x - x')(-e\gamma^\mu A_\mu^{(c)}(x'))\psi^{(0)}(x')d^4x'$$

where

$$K(x) = \int (\gamma.p - m)^{-1} \exp(ip.x) d^4p / (2\pi)^4$$

2.9.2 Points to remember

- [1] We write down Dirac's relativistic wave equation in an electromagnetic field as is done in all the standard textbooks on quantum mechanics, like for example in Dirac's own book "The principles of quantum mechanics". We first recall from basic qed that the unperturbed second quantized Dirac field, ie, in the absence of any electromagnetic field is a free field which can be expanded in terms of electron annihilation operator fields and positron creation fields in momentum space. After the classical em perturbation is switched on, we can calculate the perturbation in the second quantized Dirac field caused by the perturbation. This computation will involve solving a linear pde first order with constant coefficients with source defined by a bilinear coupling between the classical em field and the creation and annihilation operators of the free Dirac field. The solution to this system yields the correction to the free Dirac field expressed in terms of the electron propagator kernel and is bilinear in the creation and annihilation operators of the free Dirac field and the external classical em four potential field. We can thus compute the change in the Dirac four current density caused by this perturbation in the second quantized Dirac field expressed in the form of a quadratic function of creation and annihilation operators of the free Dirac field and the externally applied em four potential. By applying the retarded potential theory, we can thus calculate the quantum electromagnetic field produced by this second quantized current density perturbation and hence evaluate its statistical moments in space-time for a given state of the Dirac field, ie, a state that specifies the number of electrons and positrons and their four momenta and spins. Now we can use this expression to control the externally applied em potential so that the quantum statistical moments of the quantum em field radiated out by the second quantized Dirac currents match a given pattern.

2.10 Calculating the path of a light ray in a static gravitational field

The metric of space-time is $g_{\mu\nu}(r)$ where $r = (x, y, z)$. This metric is time independent. The photon travels along a null geodesic:

$$0 = d\tau^2 = g_{\mu\nu}dx^\mu dx^\nu$$

or equivalently, parametrizing the trajectory by t , we have

$$0 = g_{00}(r(t)) + 2g_{0m}(r(t))dx^m/dt + g_{mk}(r(t))(dx^m(t)/dt)(dx^k(t)/dt) \quad (1)$$

The other equations are obtained from the variational principle

$$\delta \int (g_{\mu\nu}(dx^\mu/d\lambda)(dx^\nu/d\lambda))^{1/2} d\lambda = 0$$

which give

$$g_{0\nu}dx^\nu/d\tau = K \quad (2)$$

$$(d/d\lambda)(g_{m\nu}dx^\nu/d\lambda) = g_{\mu\nu,m}(dx^\mu/d\lambda)(dx^\nu/d\lambda), m = 1, 2, 3$$

where K is an infinite constant. The second gives the standard geodesic equations in the form

$$d^2x^r/d\tau^2 + \Gamma_{00}^r(dt/d\tau)^2 + 2\Gamma_{0m}^r(dt/d\tau)(dx^m/d\tau) + \Gamma_{ms}^r(dx^m/d\tau)(dx^s/d\tau) = 0 \quad (3)$$

where $r = 1, 2, 3$. The null geodesic condition (1) is a first integral of the motion. We note that (3) can be expressed as

$$(d^2x^r/d\tau^2)(dt/d\tau)^{-2} + \Gamma_{00}^r + 2\Gamma_{0m}^r(dx^m/dt) + \Gamma_{ms}^r(dx^m/dt)(dx^s/dt) = 0 \quad (4)$$

Now,

$$dx^r/d\tau = (dx^r/dt)(dt/d\tau),$$

$$d^2x^r/d\tau^2 = (d^2x^r/dt^2)(dt/d\tau)^2 + (dx^r/dt)(d^2t/d\tau^2)$$

so (4) is the same as

$$\begin{aligned} d^2x^r/dt^2 + \Gamma_{00}^r + 2\Gamma_{0m}^r(dx^m/dt) + \Gamma_{ms}^r(dx^m/dt)(dx^s/dt) \\ + (dx^r/dt)(d^2t/d\tau^2)(dt/d\tau)^{-2} = 0 \end{aligned} \quad (5)$$

Now, by (2)

$$g_{00}dt/d\tau + g_{0r}dx^r/d\tau = K$$

which gives on differentiating w.r.t τ ,

$$g_{00,m}(dt/d\tau)(dx^m/d\tau) + g_{00}d^2t/d\tau^2 + g_{0r,m}(dx^r/d\tau)(dx^m/d\tau) + g_{0r}d^2x^r/d\tau^2 = 0$$

and hence multiplying this equation by $(dt/d\tau)^{-2}$ gives us

$$g_{00}(d^2t/d\tau^2)(dt/d\tau)^{-2} + g_{00,m}dx^m/dt + g_{0r,m}(dx^r/dt)(dx^m/dt)$$

$$+g_{0r}(d^2x^r/d\tau^2)(dt/d\tau)^{-2}=0 \quad (6)$$

Now from the fundamental geodesic equation, we have

$$(dt/d\tau)^{-2}[d^2x^r/d\tau^2+\Gamma_{00}^r(dt/d\tau)^2+2\Gamma_{0m}^r(dt/d\tau)(dx^m/d\tau)+\Gamma_{ms}^r(dx^m/d\tau)(dx^s/d\tau)]=0$$

or equivalently,

$$\begin{aligned} & (d^2x^r/d\tau^2)(dt/d\tau)^{-2}= \\ & -[\Gamma_{00}^r+2\Gamma_{0m}^r dx^m/dt+\Gamma_{ms}^r(dx^m/dt)(dx^s/dt)] \end{aligned} \quad (7)$$

Substituting (7) into (6), we solve for $(d^2t/d\tau^2)(dt/d\tau)^{-2}$ and then substitute this expression into (5) to obtain a second order differential equation for the null geodesic trajectory $x^r(t)$, $r = 1, 2, 3$. We leave this as an exercise.

Remark: An alternate way to get at $(d^2t/d\tau^2)(dt/d\tau)^{-2}$ directly is to us the fourth component of the geodesic equation

$$d^2t/d\tau^2+\Gamma_{00}^0(dt/d\tau)^2+\Gamma_{m0}^0(dx^m/d\tau)(dt/d\tau)+\Gamma_{ms}^0(dx^m/d\tau)(dx^s/d\tau)=0$$

which gives on multiplying by $(dt/d\tau)^{-2}$, the equation

$$\begin{aligned} & (d^2t/d\tau^2)(dt/d\tau)^{-2}= \\ & -[\Gamma_{00}^0+\Gamma_{m0}^0 dx^m/dt+\Gamma_{ms}^0(dx^m/dt)(dx^s/dt)] \end{aligned} \quad (8)$$

resulting thereby on combining with (5) in the geodesic equation

$$\begin{aligned} & d^2x^r/dt^2+\Gamma_{00}^r+2\Gamma_{0m}^r(dx^m/dt)+\Gamma_{ms}^r(dx^m/dt)(dx^s/dt) \\ & -(dx^r/dt)[\Gamma_{00}^0+\Gamma_{m0}^0 dx^m/dt+\Gamma_{ms}^0(dx^m/dt)(dx^s/dt)]=0 \end{aligned} \quad (9)$$

This equation is valid for any geodesic, not necessarily null. To get the null condition, we must in addition impose on its first integral the condition that $d\tau = 0$, ie,

$$g_{00}+2g_{0r}dx^r/dt+g_{rs}(dx^r/dt)(dx^s/dt)=0 \quad (10)$$

(9) and (10) are the two fundamental equations for null geodesics valid even if the metric is time dependent. However, in the time independent case, suppose that in addition, the metric does not depend on one of the spatial coordinates, say x^3 . Then, from the variational principle, we have in addition to (2),

$$g_{3\mu}dx^\mu/d\tau=K' \quad (11)$$

where K' is another infinite constant. We have the result that the ratio $K'/K=\beta$ is finite. Then, we get by forming this ratio,

$$\frac{g_{3\mu}dx^\mu/dt}{g_{0\mu}dx^\mu/dt}=\beta$$

or equivalently,

$$(g_{30}-\beta g_{00})+(g_{31}-\beta g_{01})dx^1/dt+(g_{32}-\beta g_{02})dx^2/dt+(g_{33}-\beta g_{03})dx^3/dt=0 \quad (12)$$

This equation along with the null condition (1) furnishes us with two equations for the trajectory. If in addition, we assume that the geodesic is in the plane $x^3 = \text{const}$, then these two equations give us two first order differential equations for the planar trajectory $x^1(t), x^2(t)$ from which we can obtain a first order differential equation for the trajectory ie x^1 as a function of x^2 . We leave the details to the reader.

2.11 A study of thermal emission by blackholes via Hawking radiation, quantum mechanics of fields in the vicinity of a blackhole and the interaction of electrons,positrons, photons and gravitons with an external noisy bath with application to the design of very large size quantum gates

2.11.1 Discussion

Consider first a Klein-Gordon field $\psi(x)$ in the vicinity of the Schwarzschild blackhole. It satisfies the equation

$$(g^{\mu\nu} \sqrt{-g} \psi_{,\mu}),_\nu + \mu^2 \psi \sqrt{-g} = 0$$

which becomes

$$\begin{aligned} & \alpha(r)^{-1} r^2 \sin(\theta) \psi_{,00} - (\alpha(r) r^2 \sin(\theta) \psi_{,r}),_r \\ & - (\sin(\theta) \psi_{,\theta}),_\theta - \sin(\theta)^{-1} \psi_{,\phi\phi} + \mu^2 r^2 \sin(\theta) \psi = 0 \end{aligned}$$

Suppose we are interested in purely radial dynamics. Then ψ is a function of (t, r) only and we get

$$\alpha(r)^{-1} r^2 \psi_{,00}(t, r) - (\alpha(r) r^2 \psi_{,r}(t, r)),_r + \mu^2 r^2 \psi(t, r) = 0$$

where $\alpha(r) = 1 - 2m/r$, $m = GM$, $c = 1$. In the frequency domain, this becomes

$$(\alpha(r)^{-1} r^2 \omega^2 - \mu^2 r^2) \psi(\omega, r) + (\alpha(r) r^2 \psi_{,r}(\omega, r)),_r = 0$$

We denote two linearly independent solutions of this equation by $f_{\pm}(\omega, r)$. More generally, if we admit angular dependence also, then we would obtain in the frequency domain,

$$\begin{aligned} & (\alpha(r)^{-1} r^2 \omega^2 - \mu^2 r^2) \psi(\omega, r, \theta, \phi) + (\alpha(r) r^2 \psi_{,r}(\omega, r, \theta, \phi)),_r \\ & + (\sin(\theta)^{-1} \psi_{,\theta}(\omega, r, \theta, \phi)),_\theta + \sin(\theta)^{-2} \psi_{,\phi\phi}(\omega, r, \theta, \phi) \\ & = 0 \end{aligned}$$

Adopting the separation of variables

$$\psi(\omega, r, \theta, \phi) = F_{lm}(\omega, r)Y_{lm}(\theta, \phi)$$

we find that the angular part gives a term proportional to Y_{lm} (these are the spherical harmonics) and the radial equation becomes

$$\begin{aligned} & (\alpha(r)^{-1}r^2\omega^2 - \mu^2r^2)F_{lm}(\omega, r) + (\alpha(r)r^2F_{lm,r}(\omega, r)),_r \\ & -l(l+1)F_{lm}(\omega, r) = 0 \end{aligned}$$

This second order ode for $F_{lm}(\omega, r)$ in the r -variable has two linearly independent solutions, one representing the radially outgoing waves and the other representing the radially incoming waves. We denote these two linearly independent solutions by $F_{lm\pm}(\omega, r)$. The general solution to the original Klein-Gordon equation in space-time near the Schwarzschild blackhole can thus be represented as

$$\psi(t, r, \theta, \phi) = \sum_{l,m} \int_{\mathbb{R}} [C_1(l, m, \omega)F_{lm+}(\omega, r) + C_2(l, m, \omega)F_{lm-}(\omega, r)]Y_{lm}(\theta, \phi)\exp(i\omega t)d\omega$$

In quantum blackhole physics, the coefficients $C_s(l, m, \omega)$, $s = 1, 2$, $l = 0, 1, 2, \dots$, $m = -l, -l+1, \dots, l-1, l$ become operators in a second quantized Hilbert space satisfying certain commutation relations which can be determined only using the Lagrangian-Hamiltonian formalism.

The Hamiltonian of the scalar field is derived from the Lagrangian density

$$L = (1/2)g^{\mu\nu}\sqrt{-g}\psi_{,\mu}\psi_{,\nu} - \mu^2\psi^2\sqrt{-g}/2$$

Note that the KG field ψ is real. The canonical momentum density is

$$p(x) = \partial L / \partial \psi_{,0} = g^{0\nu}\sqrt{-g}\psi_{,\nu}$$

and hence

$$\psi_{,0} = (p/g^{00}\sqrt{-g}) - (g^{0m}\psi_{,m}/g^{00}\sqrt{-g})\psi_{,m}$$

and hence the Hamiltonian density of the field is

$$\begin{aligned} H &= p\psi_{,0} - L = \\ & g^{0\nu}\sqrt{-g}\psi_{,\nu}\psi_{,0} - (1/2)g^{\mu\nu}\sqrt{-g}\psi_{,\mu}\psi_{,\nu} + \mu^2\psi^2\sqrt{-g}/2 \\ &= (1/2)g^{00}\sqrt{-g}(\psi_{,0})^2 - (1/2)g^{rs}\sqrt{-g}\psi_{,r}\psi_{,s} + \mu^2\psi^2\sqrt{-g}/2 \\ &= (1/2)g^{00}\sqrt{-g}[(p/g^{00}) - (g^{0m}\psi_{,m}/g^{00})\psi_{,m}]^2 \\ &\quad - (1/2)g^{\mu\nu}\sqrt{-g}\psi_{,\mu}\psi_{,\nu} + \mu^2\psi^2\sqrt{-g}/2 \\ &= H(p(x), \psi(x), \nabla\psi(x)) \end{aligned}$$

We can then compute the partition function of the KG field using the Euclidean path integral. For this, we assume a time independent metric, like the Schwarzschild/Kerr/Reissner-Nordström/Kerr-Newman metric. This guarantees that the Hamiltonian of the KG field does not depend explicitly on time. The partition function of the KG field is then given by (we set $t = 0$, $\beta = 1/kT$)

Here, $\beta = 1/kT$ where T is the temperature. From $Z(\beta)$, we derive the free energy of the KG field:

$$F = -kT \log(Z) = -\log(Z(\beta))/\beta$$

Now, the entropy S is given by

$$F = U - TS$$

or equivalently,

$$\begin{aligned} S &= k(U - F)/kT = k\beta(U - F) = k\beta(-\log(Z(\beta)), \beta + \log(Z(\beta))/\beta) \\ &= (-k/\beta^2)[(\log Z)/\beta], \beta \end{aligned}$$

Thus, the entropy of the KG field near a blackhole can be in practice, computed as a function of the temperature and the blackhole parameters like mass and angular momentum. In particular, for the Schwarzschild blackhole, the entropy becomes a function of the temperature and the mass of the blackhole M .

Rindler space: We transform the Schwarzschild coordinates t, r to (t', r') so that the metric appears in the form

$$d\tau^2 = F(t', r')(dt'^2 - dr'^2) - f(t', r')(d\theta^2 + \sin^2(\theta)d\phi^2)$$

This can be done as follows: Write $\alpha(r) = 1 - 2m/r$. Then,

$$\begin{aligned} \alpha(r)dt^2 - \alpha(r)^{-1}dr^2 &= \\ \alpha(r)(dt^2 - \alpha(r)^{-2}dr^2) \end{aligned}$$

We set

$$dr' = \alpha(r)^{-1}dr = dr/(1 - 2m/r) = rdr/(r - 2m)$$

which gives on integration,

$$r' = r + 2m \log(|r - 2m|)$$

We do not transform t , ie, $t' = t$ and our transformed metric assumes the form

$$d\tau^2 = \alpha(r)(dt^2 - dr'^2) - r^2(d\theta^2 + \sin^2(\theta)d\phi^2)$$

Near the event horizon $r = 2m$ and $r' \rightarrow -\infty$, we have approximately,

$$r' = 2m \log(|r - 2m|)$$

or equivalently,

$$r = 2m \pm \exp(r'/2m)$$

$$\begin{aligned} \alpha(r) &= 1 - 2m/r \approx 1 - 2m/(2m \pm \exp(r'/2m)) = 1 - 1/(1 \pm (2m)^{-1} \cdot \exp(r'/2m)) \\ &\approx \pm(2m)^{-1} \cdot \exp(r'/2m) \end{aligned}$$

Our approximate metric becomes

$$d\tau^2 = \pm(2m)^{-1} \exp(r'/2m)(dt^2 - dr'^2) - 4m^2 \exp(r'/2m)(d\theta^2 + \sin^2(\theta)d\phi^2)$$

The radial null geodesics near the event horizon are therefore approximately given by

$$r' = \pm t$$

In fact, this is exactly true without any approximation, ie, the radial null geodesics of the Schwarzschild metric are

$$r' = \pm t$$

or equivalently,

$$r + 2m \log(|r - 2m|) = \pm t$$

or equivalently,

$$r \pm t + 2m \log(|r - 2m|) = 0$$

We take a harmonic photonic wave of frequency ω which means that its complex amplitude as a function of time t is given by $A \exp(i\omega t)$. This wave travels from within the event horizon $r = 2m - \epsilon$ to outside the event horizon $r = 2m + \epsilon$. The variation of the complex amplitude of this photon as a wave with (t, r) should therefore be

$$A \exp(i\omega(t - r - 2m \log(2m - r)))$$

The phase of this wavefront is $\omega(t - r - 2m \log(2m - r))$ and it is a constant along the surface $t - r - 2m \log(r - 2m)$ in agreement with the above formula for the photon trajectory. When this photon wave is within the event horizon, its amplitude is

$$|A \exp(i\omega(t - r - 2m \log(2m - (2m - \epsilon))))|^2 = |A|^2$$

and when it is outside the event horizon, its amplitude is

$$\begin{aligned} |A \exp(i\omega(t - r - 2m \log(2m - (2m + \epsilon))))|^2 &= |A|^2 |\exp(i\omega \cdot 2m \log(-\epsilon))|^2 \\ &= |A|^2 \exp(i\omega \cdot 2m(i\pi + \log(\epsilon)))^2 = |A|^2 \exp(-4m\pi\omega) \end{aligned}$$

This means that when the photon comes out of the event horizon of the black-hole, its energy has decreased by the factor $\exp(-4m\pi\omega)$. Such a decrease in its energy should correspond to its emission of thermal radiation at a temperature T where $\exp(-h\omega/2\pi kT) = \exp(-4m\pi\omega)$. The temperature of the radiation should therefore be given by

$$h/2\pi kT = 4m\pi$$

or equivalently,

$$T = h/8\pi^2 mk = hc^2/8\pi^2 GMk$$

where h is Planck's constant, c is the speed of light, G is Newton's gravitational constant, M is the mass of the blackhole and k is Boltzmann's constant. This is called the Hawking temperature.

Another way to derive Hawking radiation: Consider the radial wave equation for a scalar field $\psi(t, r)$ in the Schwarzschild metric. It is given by

$$\alpha(r)^{-1}r^2\psi_{,tt} - (\alpha(r)r^2\psi_{,r})_{,r} = 0$$

Note that we are assuming a massless particle. For massive particles, this radial wave equation should be replaced by

$$\alpha(r)^{-1}r^2\psi_{,tt} - (\alpha(r)r^2\psi_{,r})_{,r} + \mu^2\psi(t, r) = 0$$

Taking the Fourier transform w.r.t time, the former radial wave equation becomes the radial Helmholtz equation

$$(\alpha(r)r^2\psi_{,r}(\omega, r))_{,r} + \omega^2\alpha(r)^{-1}r^2\psi(\omega, r) = 0$$

It can be rearranged as

$$\alpha(r)r^2\psi_{,rr}(\omega, r) + (\alpha(r)r^2)_{,r}\psi_{,r}(\omega, r) + \omega^2\alpha(r)^{-1}r^2\psi(\omega, r) = 0$$

This equation describes the radial propagation of a wave of frequency ω in the vicinity of the Schwarzschild blackhole. We look for a solution of the form

$$\psi(\omega, r) = \exp(iS(\omega, r))$$

If we write

$$S(\omega, r) = S_R(\omega, r) + iS_I(\omega, r)$$

where S_R, S_I are real, then the intensity of the wave at r is given by

$$I(\omega, r) = |\exp(iS(\omega, r))|^2 = \exp(-2S_I(\omega, r))$$

and by letting $r \rightarrow 2m$, we would be able to calculate the temperature T of Hawking radiation, by setting

$$\exp(-2S_I(\omega, 2m)) = \exp(-h\omega/2\pi kT)$$

This corresponds to the fact that the probability of a photon of frequency ω being emitted is equals $a \exp(-h\omega/2\pi kT)$ according to the Gibbs-Planck quantum hypothesis. Now, we obtain an equation for S . We have

$$\psi_{,r} = iS' \cdot \exp(iS), \psi_{,rr} = (iS'' - S'^2)\exp(iS)$$

and substituting this into the above radial Helmholtz equation gives

$$\alpha(r)r^2(iS'' - S'^2) + (\alpha(r)r^2)'iS' + \omega^2\alpha(r)^{-1}r^2 = 0 \quad (1)$$

Writing

$$S = S_R + iS_I$$

we get

$$\alpha(r)r^2(-S_I'' + iS_R'' - (S_R'^2 - S_I'^2 + 2iS_R'S_I'))$$

$$+(\alpha(r)r^2)'(-S'_I + iS'_R) + \omega^2\alpha(r)^{-1}r^2 = 0$$

Noting that

$$\begin{aligned} (\alpha(r)r^2)' / (\alpha(r)r^2) &= (\log(\alpha(r)r^2))' \\ (\log(r^2 - 2mr))' &= (2r - 2m) / (r(r - 2m)) \end{aligned}$$

this equation can be rearranged as

$$\begin{aligned} &(-S''_I + iS''_R - (S'^2_R - S'^2_I + 2iS'_R S'_I)) \\ &+ (2(r - m)/r(r - 2m))(-S'_I + iS'_R) + \omega^2\alpha(r)^{-2} = 0 \end{aligned}$$

Equating the real and imaginary parts of this equation gives us

$$-S''_I - S'^2_R + S'^2_I - (2(r - m)/r(r - 2m))S'_I + \omega^2\alpha(r)^{-2} = 0, \dots \quad (2)$$

$$S''_R - 2S'_R S'_I + (2(r - m)/r(r - 2m))S'_R = 0 \dots \quad (3)$$

(1) approximately gives (Noting that S is actually $2\pi S/h$ and hence linear terms in S and its partial derivatives are of order $1/h$ while square terms are of order $1/h^2$ which are therefore much larger than the former in magnitude)

$$-\alpha(r)r^2S'^2 + \omega^2\alpha(r)^{-1}r^2 = 0$$

or equivalently,

$$S' = \pm\omega\alpha(r)^{-1} = \pm\omega r/(r - 2m)$$

This integrates to give

$$S(\omega, r) = \pm\omega(r + 2m.\log(r - 2m))$$

and hence

$$S(\omega, r) = \pm\omega(r + 2m(i\pi + \log(2m - r))), r < 2m,$$

$$S(\omega, r) = \pm\omega(r + 2m.\log(r - 2m)), r > 2m$$

so that the ratio of $|\exp(iS(\omega, r))|^2$ for $r > 2m$ and $r < 2m$ is given approximately by

$$\exp(-4\pi m\omega) = \exp(-4\pi GM\omega/c^2)$$

and equating this to $\exp(-\hbar\omega/2\pi kT)$, we get the desired value of the Hawking temperature T .

The occurrence of Hawking radiation causes the evolution of the wave function ψ of the Klein-Gordon field to be non-unitary. This means that Hawking radiation can be used to design non-unitary quantum gates. One way to do this would be to start with a set of N pairs of desired input and output wave functions $(\psi_{1k}(r), \psi_{2k}(r)), k = 1, 2, \dots, N$ and then to construct using solutions to the KG equation an evolution operator T for which $T\psi_{1k} \approx \psi_{2k}, k = 1, 2, \dots, N$. The correct approach however would be to use instead of the KG equation, the

Dirac equation in the Schwarzschild metric for the determination of the evolution operator T . Our tetrad for the Schwarzschild metric is $V_{a\mu}$ where

$$V_{0\mu}dx^\mu = f(r)dt, V_{1\mu}dx^\mu = f(r)^{-1}dr,$$

$$V_{2\mu}dx^\mu = rd\theta, V_{3\mu}dx^\mu = r.\sin(\theta)d\phi$$

where

$$f(r) = \sqrt{\alpha(r)} = \sqrt{1 - 2m/r}$$

We thus find,

$$V_{00} = f(r), V_{01} = V_{02} = V_{03} = 0,$$

$$V_{10} = 0, V_{11} = f(r)^{-1}, V_{12} = V_{13} = 0,$$

$$V_{20} = V_{21} = 0, V_{22} = r, V_{23} = 0,$$

$$V_{30} = V_{31} = V_{32} = 0, V_{33} = r.\sin(\theta)$$

To compute the spinor connection, we need the covariant derivatives $V_{a\mu:\nu}$. We have

$$V_{a\mu:\nu} = V_{a\mu,\nu} - \Gamma_{\mu\nu}^\rho V_{a\rho}$$

Thus, we need to tabulate 64 covariant derivatives

$$V_{a\mu:\nu}, a, \mu, \nu = 0, 1, 2, 3$$

For $(a, \mu, \nu) = (0, 0, 0)$, we find that

$$\begin{aligned} V_{00:0} &= V_{00,0} - \Gamma_{00}^\rho V_{0\rho} = -\Gamma_{00}^1 V_{01} \\ &= (1/2)g^{11}g_{00,1}V_{01} = 0 \end{aligned}$$

For $(0, 0, 1)$, we find that

$$\begin{aligned} V_{00:1} &= V_{00,1} - \Gamma_{01}^\rho V_{0\rho} \\ &= f'(r) - \Gamma_{01}^0 V_{00} = f'(r) - g^{00}\Gamma_{001}V_{00} \\ &= f'(r) - (1/2)g^{00}g_{00,1}f(r) = f'(r) - (1/2)f(r)\alpha'(r)/\alpha(r) \end{aligned}$$

For $(0, 0, 2)$, we have

$$\begin{aligned} V_{00:2} &= V_{00,2} - \Gamma_{02}^\rho V_{0\rho} = 0 \\ V_{00:3} &= V_{00,3} - \Gamma_{03}^\rho V_{0\rho} = 0 \\ V_{01:0} &= V_{01,0} - \Gamma_{10}^\rho V_{0\rho} \\ &= -\Gamma_{10}^0 V_{00} = (-1/2)g^{00}g_{00,1}f(r) = (-1/2)\alpha'(r)f(r)/\alpha(r) \\ V_{01:1} &= V_{01,1} - \Gamma_{11}^\rho V_{0\rho} = -\Gamma_{11}^1 V_{01} = 0 \\ V_{01:2} &= V_{01,2} - \Gamma_{12}^\rho V_{0\rho} = -\Gamma_{12}^2 V_{02} = 0 \\ V_{01:3} &= V_{01,3} - \Gamma_{13}^\rho V_{0\rho} = -\Gamma_{13}^1 V_{01} = 0 \\ V_{02:0} &= -\Gamma_{20}^\rho V_{0\rho} = -\Gamma_{20}^0 V_{02} = 0 \end{aligned}$$

$$\begin{aligned}
V_{02:1} &= -\Gamma_{21}^\rho V_{0\rho} = -\Gamma_{21}^2 V_{02} = 0 \\
V_{02:2} &= -\Gamma_{22}^\rho V_{0\rho} = -\Gamma_{22}^0 V_{00} = 0 \\
V_{02:3} &= -\Gamma_{23}^\rho V_{0\rho} = -\Gamma_{23}^3 V_{03} = 0 \\
V_{03:0} &= -\Gamma_{30}^\rho V_{0\rho} = 0 \\
V_{03:1} &= -\Gamma_{31}^\rho V_{0\rho} = -\Gamma_{31}^3 V_{03} = 0 \\
V_{03:2} &= -\Gamma_{32}^\rho V_{0\rho} = -\Gamma_{32}^0 V_{00} = 0 \\
V_{03:3} &= -\Gamma_{33}^\rho V_{0\rho} = -\Gamma_{33}^0 V_{00} = 0- \\
V_{10:0} &= V_{10,0} - \Gamma_{00}^\rho V_{1\rho} = -\Gamma_{00}^1 V_{11} = \\
(1/2)g^{11}g_{00,1}f(r)^{-1} &= (-1/2)\alpha(r)\alpha'(r)/f(r) \\
V_{10:1} &= V_{10,1} - \Gamma_{01}^\rho V_{1\rho} = \\
-\Gamma_{01}^0 V_{10} - \Gamma_{01}^1 V_{11} &= -\Gamma_{01}^1 V_{11} \\
&= (-1/2)g^{11}g_{11,0}/f(r) = 0 \\
V_{10:2} &= V_{10,2} - \Gamma_{02}^\rho V_{1\rho} = 0 \\
V_{10:3} &= V_{10,3} - \Gamma_{03}^\rho V_{1\rho} = 0 \\
V_{11:0} &= V_{11,0} - \Gamma_{10}^\rho V_{1\rho} \\
&= -\Gamma_{10}^1 V_{11} = 0 \\
V_{11:1} &= V_{11,1} - \Gamma_{11}^\rho V_{1\rho} \\
&= (1/f(r))' - \Gamma_{11}^1 V_{11} = (1/f(r))' - (1/2)g^{11}g_{11,1}/f(r) \\
&= (1/f)' - (1/2)\alpha'/\alpha f \\
V_{11:2} &= -\Gamma_{12}^\rho V_{1\rho} = -\Gamma_{12}^2 V_{12} = 0 \\
V_{11:3} &= -\Gamma_{13}^\rho V_{1\rho} = -\Gamma_{13}^1 V_{11} = 0 \\
V_{12:0} &= -\Gamma_{20}^\rho V_{1\rho} = -\Gamma_{20}^1 V_{11} = 0 \\
V_{12:1} &= -\Gamma_{21}^\rho V_{1\rho} = -\Gamma_{21}^1 V_{11} = 0 \\
V_{12:2} &= -\Gamma_{22}^\rho V_{1\rho} = -\Gamma_{22}^1 V_{11} = \\
(1/2)g^{11}g_{22,1}/f &= r\alpha(r)/f(r) \\
V_{12:3} &= -\Gamma_{23}^\rho V_{1\rho} = -\Gamma_{23}^1 V_{11} = 0 \\
V_{13:0} &= -\Gamma_{30}^\rho V_{1\rho} = -\Gamma_{30}^1 V_{11} = 0 \\
V_{13:1} &= -\Gamma_{31}^\rho V_{1\rho} = -\Gamma_{31}^1 V_{11} = 0 \\
V_{13:2} &= -\Gamma_{32}^\rho V_{1\rho} = -\Gamma_{32}^1 V_{11} = 0 \\
V_{13:3} &= -\Gamma_{33}^\rho V_{11} = (1/2)g^{11}g_{33,1}V_{11} = r.\alpha(r).sin^2(\theta)f(r) \\
V_{20:0} &= -\Gamma_{00}^\rho V_{2\rho} = -\Gamma_{00}^2 V_{22} = 0 \\
V_{20:1} &=
\end{aligned}$$

2.11.2 Points to remember

[1] In classical blackhole physics a particle takes infinite coordinate time but finite proper time to arrive at the critical Schwarzschild radius $r_c = 2m = 2GM/c^2$ either from outside the event horizon or from within it. In particular, particles from within the blackhole will never arrive outside its event horizon for an external observer to see. Thus, no emission of particles can take place from within the blackhole for an external observer.

[2] In quantum blackhole physics as propounded first by Stephen Hawking, there is a probability for thermal emission of photons, electrons, positrons, gravitons or other elementary particles satisfying the non-Abelian gauge equations of Yang and Mills. This emission occurs at a temperature called the Hawking temperature.

[3] The basic method of computing the Hawking temperature is to start with the Klein-Gordon equation in the Schwarzschild curved space-time background and extract out only its radial part. Partial derivatives w.r.t time are replaced by multiplication with $j\omega$ as is done in arriving at the Helmholtz equation from the wave equation. We substitute $\exp(iS(\omega, r)/\hbar)$ for its solution and evaluate using perturbation theory the approximate imaginary part of S , say $S_I(\omega, r)$. Then in accordance with Feynman's path integral approach to quantum mechanics,

$$\begin{aligned} & |\exp(i(S(\omega, r_c + \epsilon) - S(\omega, r_c - \epsilon)))|^2 \\ &= \exp(-2(S_I(\omega, r_c + \epsilon) - S_I(\omega, r_c - \epsilon))) \end{aligned}$$

is the probability of a photon to get emitted from just within the critical radius to just outside the critical radius at frequency ω . But according to the Gibbs-Boltzmann principle in statistical mechanics, this probability must be equal to $\exp(-\hbar\omega/kT)$. We thus obtain the value of the Hawking temperature T at which emission occurs. If the mass of the KG particle is set equal to zero, then the Hawking temperature comes out to be frequency independent.

[4] Blackholes can emit all kinds of particles not only photon. It may even gravitons, electrons, positrons of the elementary particles appearing in Yang-Mills non-Abelian gauge theories. To determine for example the probability of emission of an electron/positron, we solve Dirac's equation in the Schwarzschild background and calculate the probability density of the wave function outside the critical radius after time t_0 given that at time $t = 0$ the wave function was nonzero only within the critical radius.

[5] To calculate the probability of emission of gravitons, we look at small perturbations of the Schwarzschild metric and write down Lagrangian density from which the wave equations for these perturbations is derived. We then write down the corresponding Hamiltonian density and the Hamiltonian in which the position fields will be the metric perturbations and the momentum fields will be obtained by differentiating the Lagrangian density w.r.t the time derivative of the position fields. The resulting Schrodinger equation will then give us the wave function of the position fields outside the critical radius given that at time $t = 0$ the wave function of the metric perturbations was concentrated

within the critical radius. The theory of classical metric perturbations has been studied intensively by S.Chandrasekhar in his book "The mathematical theory of blackholes", Oxford University Press.

[6] A Schwarzschild blackhole also has an entropy given by $S = A/4G$ where $A = 4\pi r_c^2$ is the event horizon surface area. This formula can be deduced using the thermodynamic relation

$$\frac{\partial S}{\partial U} = 1/T$$

where $T = T(M)$ is Hawking's temperature and $U = Mc^2$ is the energy of the blackhole.

[7] A blackhole produces entropy. This can be seen from the following argument. Let at time $t = 0$ the state of the gravitons and other particles within the critical radius be a pure state $|\psi_I\rangle$ and the state of the gravitons and other particles outside the critical radius at $t = 0$ be $|\psi_O\rangle$. Then the total state of all the particles within and outside at time $t = 0$ is the tensor product $|\psi_I\rangle|\psi_O\rangle$. This is so because particles initially inside the critical radius cannot propagate outside and vice-versa. Now this state of particles evolves under the Hamiltonians of the gravitational field plus the Hamiltonians of all the other particles and their interactions. So after time t the total state of the particles is $|\psi(t)\rangle = U(t)(|\psi_I\rangle|\psi_O\rangle)$. The state of the particles outside the blackhole at time $t > 0$ is therefore the mixed state

$$\rho_O(t) = Tr_I(|\psi(t)\rangle\langle\psi(t)|)$$

A pure state has zero entropy and a mixed state has finite entropy. Thus, entropy for an observer outside the critical radius has been created by the blackhole.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Chapter 3

Analysis of waves in engineering and optical systems, in biological systems, classical and in quantum blackhole physics

3.1 Wave digital filter design

(Mridul's Phd work)

3.1.1 General theory

We can design digital filters based on cascading of wave digital filter blocks as follows.

If V, I are respectively the voltage and current along a transmission line and Z_0 is the characteristic impedance, then the forward and backward wave components at any point along the line are

$$a = (V + Z_0 I)/2, b = (V - Z_0 I)/2$$

For a loss line, these complex phasors at two points on the line separated by a distance of d are related by

$$a(2) = \exp(-j\beta d)a(1), b(2) = \exp(j\beta d)b(1)$$

where $\beta = \omega/u$. In the time domain, this translates to

$$a(2, t) = a(1, t - d/u), b(2, t) = b(1, t + d/u)$$

so that d/u is the time delay. Assuming that the points sampled along the line are uniformly spaced by a distance of Δ , these relations become where n means $n\Delta$

$$a_{k+1}(n) = a_k(n-1), b_{k+1}(n) = b_k(n+1)$$

A cascade of such 2×2 blocks with some matrix multipliers introduced in between can result in a digital filter having almost any arbitrary rational transfer function. This means that a typical block will have the following matrix transfer function

$$H(z) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} z^{-1} & 0 \\ 0 & z \end{pmatrix}$$

and a general 2×2 FIR matrix transfer function will be obtained by multiplying several such blocks with different values of a, b, c, d .

We describe here in some detail, how to match the filter coefficients of an FIR filter of fixed length p to the impulse response of an IIR filter when the coefficients of the FIR filter depend on a set of q parameters which can be varied over a given domain.

Let

$$H(z) = \sum_{n=0}^{\infty} h[n]z^{-n}$$

be the transfer function of an IIR filter and let our FIR filter to be designed have the transfer function

$$A(z) = \sum_{n=0}^N a_k(\theta)z^{-k}$$

where

$$\theta = [\theta_1, \dots, \theta_q]^T$$

It is clear from Parseval's theorem that the error energy between these two transfer functions is given by

$$E(\theta) = (1/2\pi) \int_{-\pi}^{\pi} |H(e^{j\omega}) - A(e^{j\omega})|^2 d\omega =$$

$$\sum_{k=0}^N |a_k(\theta) - h[k]|^2 + \sum_{k>N} |h[k]|^2$$

We neglect the second sum since it does not depend on the adjustable parameter vector θ and so effectively, our error energy may be redefined as

$$E(\theta) = \sum_{k=0}^N |a_k(\theta) - h[k]|^2$$

This is usually minimized using the gradient search algorithm:

$$\theta[k+1] = \theta[k] - \mu \cdot \nabla E(\theta[k]), k = 0, 1, 2, \dots$$

or equivalently, in component form,

$$\theta_m[k+1] = \theta_m[k] - \mu \cdot \frac{\partial E(\theta[k])}{\partial \theta_m}, m = 1, 2, \dots, p$$

In other words, if the energy increases with increase in θ_m , then $\partial E / \partial \theta_m$ is positive and hence at the next iteration we decrease θ_m by making the decrease proportional to the positive quantity $\partial E / \partial \theta_m$ and likewise, if E decreases with increase in θ_m , then $\partial E / \partial \theta_m$ is negative and hence at the next iteration, we increase θ_m by making this increase proportional to the positive quantity $-\partial E / \partial \theta_m$. There are however some drawbacks with this method. These are

- [a] We are allowed to vary θ only over a given subset $D \subset \mathbb{R}^p$,
- [b] The energy function $E(\theta)$ may have many local minima and if our initial choice of the parameter θ is near a local minimum, then the gradient search algorithm will cause the iteration to converge to that local minimum unless the adaptation step size μ is not large enough. To avoid this, we may choose several initial values of θ spread out over the entire domain D of the allowed choice set and then run the gradient search algorithm. Different choices would converge to different local minima and one particular choice would converge to a global minimum. Another way is the following. This works well in the special case when the parameter vector has just one real component, ie, $\theta \in \mathbb{R}$. In this case, if θ is allowed to vary only over an interval $[\alpha, \beta]$ of the real line, then we simply plot its graph and see where the global minimum is attained. Specifically, we divide this interval into N parts:

$$[\alpha + r\Delta, \alpha + (r+1)\Delta], r = 0, 1, \dots, N-1, \Delta = (\beta - \alpha)/N$$

and plot the points

$$(\alpha + r\Delta, E(\alpha + r\Delta)), r = 0, 1, \dots, N-1$$

From the graph, we can read the value of θ at which the global minimum is attained.

Another approach is the following: When the interval $[\alpha, \beta]$ for θ to vary is given, we run the gradient algorithm by choosing $\theta[0]$ in $[\alpha, \beta]$ and if at some iteration, $k > 0$, we get $\theta[k] < \alpha$, then we change its value to $\theta[k] = \alpha + \epsilon$ while if $\theta[k] > \beta$, then we change its value to $\beta - \epsilon$ where ϵ is a small number, ie, much smaller than $\beta - \alpha$. A rough sketch of such a computer programme would be

for $n = 1 : M$,

$$\theta[n+1] = \theta[n] - \mu E'(\theta[n])$$

If $\theta[n+1] < \alpha$, then define $\theta[n+1] = \alpha + \epsilon$ else if $\theta[n+1] > \beta$, then define $\theta[n+1] = \beta - \epsilon$ end;

Further, to avoid landing at a local minimum, we start with a large value of the adaptation constant μ and gradually as the iteration proceeds we reduce its value so that final values of the iterates do not hop around the global minimum.

3.1.2 Points to remember

[1] In general, the filter design problem is to first choose a parametric function $H(z|\theta)$ for a filter with θ being an unknown parameter vector, for example this may be an FIR filter of length p with θ as the vector of filter coefficients. Given a desired transfer function $H_d(z)$, we may define the distance between H and H_d in many ways, like

$$d(H(.|\theta), H_d(.)) = \max_{z \in D} |H(z|\theta) - H_d(z)|$$

or

$$d(H(.|\theta), H_d(.))^2 = \int_D |H_d(z|\theta) - H_d(z)|^2 dx dy$$

where D is a subset of \mathbb{R}^2 and $z = x + jy$, or even with Γ as a contour like the unit circle, as

$$d(H(.|\theta), H_d(.))^2 = \left| \left(\int_{\Gamma} |H_d(z|\theta) - H_d(z)|^2 z^{-1} dz \right) \right|$$

Each choice of the distance measure yields a different approximant to the given filter.

[2] In the theory of wave digital filters, we take a lossless transmission line of length d and relate using basic transmission line theory, the voltage and current phasors $(V_1(\omega, 0), I_1(\omega, 0))$ at one end to the same $(V_2(\omega, d), I_2(\omega, d))$ at the other end via a 2×2 complex matrix. The corresponding incoming and outgoing wave phasor amplitudes $a(1) = (V(\omega, 0) + R_0 I(\omega, 0))/2, b(1) = (V(\omega, 0) - R_0 I(\omega, 0))/2$ at one end are related in a very simple way to the corresponding amplitudes $a(2) = (V(\omega, d) + R_0 I(\omega, d))/2, b(2) = (V(\omega, d) - R_0 I(\omega, d))/2$ at the other end through matrix elements like $\exp(\pm j\omega d/u)$. Thus, if time is measured in units of d/u , this Tx line block can be represented in terms of single delay difference equations and by cascading a sequence of blocks, we can hope to approximate any given FIR or even IIR filters. Filters designed in this way are called wave digital filters owing to the analogy of forward and backward wave amplitudes and the corresponding reflection coefficients appearing in transmission line theory.

3.2 Large deviation principle in wave-motion

3.2.1 Discussion

The wave equation in one dimension driven by a white-Gaussian noise field has the form

$$u_{,t}(t, x) - c^2 u_{,xx}(t, x) = \sqrt{\epsilon} w(t, x), 0 \leq x \leq L$$

with boundary conditions

$$u(t, 0) = u(t, L) = 0, t \geq 0$$

Here w is a zero mean Gaussian random field with covariance

$$\mathbb{E}(w(t, x)w(t', x')) = \delta(t - t')\delta(x - x')$$

We can develop the solution in a half sine-wave Fourier series:

$$u(t, x) = \sum_{n \geq 1} u_n(t) \phi_n(x)$$

where

$$\phi_n(x) = (2/L)^{1/2} \sin(n\pi x/L)$$

We also note that the random noise field $w(t, x)$ can be expanded as

$$w(t, x) = \sum_{n \geq 1} w_n(t) \phi_n(x)$$

Then, we require that

$$\mathbb{E}(w_n(t) w_m(t')) = \delta_{n,m} \delta(t - t')$$

so that using

$$\delta(x - x') = \sum_{n \geq 1} \phi_n(x) \phi_n(x')$$

we get the correct expression for the covariance of w . Thus, $\{w_n(\cdot)\}$ are independent standard white Gaussian noise processes and can be represented as

$$w_n(t) = B'_n(t)$$

where $\{B_n(\cdot)\}$ are independent standard Brownian motion processes. Substituting this into the above wave equation gives us

$$u''_n(t) + \omega_n^2 u_n(t) = \sqrt{\epsilon} w_n(t), n = 1, 2, \dots$$

where

$$\omega_n = n\pi c/L$$

Equivalently, in terms of Ito stochastic differential equations,

$$du_n(t) = v_n(t) dt, dv_n(t) = -\omega_n^2 u_n(t) dt + \sqrt{\epsilon} dB_n(t), n \geq 1$$

Using Schilder's theorem, we obtain the rate function of the processes $u_n(t), 0 \leq t \leq T, n = 1, 2, \dots$, the expression

$$I(u_n(t), n \geq 1, 0 \leq t \leq T) =$$

$$(1/2) \sum_{n \geq 1} \int_0^T (u''_n(t) + \omega_n^2 u_n(t))^2 dt$$

Rate function for one dimensional stochastic filtering problems. The state process $x(t) \in \mathbb{R}$ satisfies the sde

$$dx(t) = f_t(x(t)) dt + \sqrt{\delta} g_t(x(t)) dB(t)$$

and the measurement process is

$$dz(t) = x(t)dt + \sqrt{\epsilon}\sigma dv(t)$$

where B, v are independent Brownian motion processes. If $\epsilon = 0$, the process x can be recovered exactly from the measurement z since there is no measurement noise. If $\delta = 0$, then the process x is non-random and can be uniquely determined from its initial value $x(0)$ by solving the differential equation

$$x'(t) = f_t(x(t)), t \geq 0$$

The EKF for this system is

$$d\hat{x}(t) = f_t(\hat{x}(t))dt + \sigma^{-2}\epsilon^{-1}P(t)(dz(t) - \hat{x}(t)dt),$$

$$P'(t) = 2f_t(\hat{x}(t))P(t) + \delta g_t(\hat{x}(t))^2 - \sigma^{-2}\epsilon^{-1}P(t)^2$$

We see that if $\epsilon \rightarrow 0$, then the second Riccati equation gives $P(t) \rightarrow 0$ which means that the state estimate is perfect. This agrees with what we said earlier, namely, that zero measurement noise implies perfect state recovery. In general, we define

$$e(t) = x(t) - \hat{x}(t)$$

and obtain approximately,

$$\begin{aligned} de(t) &= f'_t(\hat{x}(t))e(t)dt + \sqrt{\delta}g_t(\hat{x}(t))dB(t) - \sigma^{-2}\epsilon^{-1}P(t)(e(t)dt + \sqrt{\epsilon}\sigma dv(t)) \\ &= (f'(\hat{x}(t)) - \sigma^{-2}\epsilon^{-1}P(t))e(t)dt + \sqrt{\delta}g_t(\hat{x}(t))dB(t) - \sigma^{-1}\epsilon^{-1/2}P(t)dv(t) \end{aligned}$$

Since $P(t)$ scales as ϵ , we define

$$Q(t) = P(t)/\epsilon$$

Then the above equation along with the Riccati equation can be expressed as

$$\begin{aligned} de(t) &= (f'(\hat{x}(t)) - \sigma^{-2}Q(t))e(t)dt + \sqrt{\delta}g_t(\hat{x}(t))dB(t) \\ &\quad - \sqrt{\epsilon}\sigma^{-1}Q(t)dv(t), \\ Q'(t) &= 2f_t(\hat{x}(t))Q(t) + \epsilon\delta g_t(\hat{x}(t))^2 - \sigma^{-2}Q(t)^2 \end{aligned}$$

This equation implies that conditioned on the state observer $\hat{x}(t), 0 \leq t \leq T$, the rate function for $e(t), 0 \leq t \leq T$ is given by

$$I(e|\hat{x}) =$$

3.3 Some more problems in Schrodinger-wave mechanics and Heisenberg-matrix mechanics with relevance to quantum information theory

3.3.1 Discussion

[a] Initially, the system is in a pure state $|f\rangle$. A random unitary operator $U(\theta)$ where θ has the probability distribution $P(d\theta)$ acts on this state, then the output state is a mixed state:

$$\rho = \int U(\theta)|f\rangle\langle f|U(\theta)^*P(d\theta)$$

The entropy of the input state is zero while the entropy of the output state is non-zero. Such a situation occurs for example when the em field within a waveguide depends on random parameters like amplitudes and phases and this field is incident on an atom or a quantum harmonic oscillator that is initially in a pure state. The computations can be carried out in the interaction picture causing the state after interacting with the random em guide field to be mixed. Thus, the waveguide field described as a classical random em field pumps in entropy into the atomic system.

[b] A more accurate description of the above phenomena is as follows: The system is initially in the pure state $|f\rangle$ while the photon bath is in another pure state $|g\rangle$ like for example, a coherent state. Then the initial state of the system and bath is the pure state $|f \otimes g\rangle = |f\rangle|g\rangle$. The atomic/system observables are defined in the Hilbert space \langle while the quantum em field observables are defined in the Hilbert space \mathcal{H} , like for example, the Boson Fock space. The Hamiltonian of the system and bath is of the form

$$H(t) = H_S \otimes I_B + I_S \otimes H_B + V_{SB}(t)$$

where I_B is the identity operator in the bath space \mathcal{H} , I_S is the identity operator in the system Hilbert space \mathfrak{h} , H_S is the system Hamiltonian acting in \mathfrak{h} , H_B is the bath/em field Hamiltonian acting in \mathcal{H} while $V_{SB}(t)$ is the interaction Hamiltonian acting in the tensor product space $\mathfrak{h} \otimes \mathcal{H}$. In the interaction picture, the state of the system after time T is

$$\rho_S(T) = Tr_B(U(T)|f \otimes g\rangle\langle f \otimes g|U(T)^*)$$

where

$$U(T) = T\{\exp(-i \int_0^T \tilde{V}_{SB}(t)dt)\},$$

with

$$\tilde{V}_{SB}(t) = (\exp(itH_S) \otimes \exp(itH_B))V_{SB}(t)(\exp(-itH_S) \otimes \exp(-itH_B))$$

Writing

$$U(T) = \sum_k U_k \otimes W_k$$

where U_k is a linear operator acting in \mathfrak{h} while W_k is a linear operator acting in \mathcal{H} , the unitarity condition on $U(T)$ becomes

$$\sum_{k,m} U_m^* U_k \otimes W_m^* W_k = I = \sum_{k,m} U_m U_k^* \otimes W_m W_k^*$$

and we find that

$$\rho_S(T) = \sum_{k,m} \langle g | W_m^* W_k | g \rangle (U_k | f \rangle \langle f | U_m^*)$$

Once again we find that the initial system state is pure and hence has zero entropy while the final system state is mixed and therefore has non-zero entropy, implying therefore that the bath has pumped entropy into the system. It is possible to describe this latter quantum process using the former classical probabilistic method, ie, by deriving $\rho_S(T)$ as the action of a classically random unitary operator acting on the initial system state alone. This is achieved as follows: The matrix ($\langle g | W_m^* W_k | g \rangle$) is positive definite and hence we can write its spectral decomposition as

$$\langle g | W_m^* W_k | g \rangle = \sum_l p(l) e(l,k) \bar{e}(l,m)$$

where $p(l) \geq 0$ and

$$\begin{aligned} \sum_k e(l,k) \bar{e}(l',k) &= \delta(l,l'), \\ \sum_l e(l,k) \bar{e}(l,m) &= \delta(k,m) \end{aligned}$$

Thus,

$$\begin{aligned} \rho_S(T) &= \sum_{l,k,m} p(l) e(l,k) \bar{e}(l,m) (U_k | f \rangle \langle f | U_m^*) \\ &= \sum_l p(l) (V_l | f \rangle \langle f | V_l^*) \end{aligned}$$

where

$$V_l = \sum_k e(l,k) U_k$$

It is easy to verify that by absorbing a positive factor from $p(l)$ into V_l , we can make V_l unitary operators and simultaneously guarantee $\sum_l p(l) = 1$. We leave this as an exercise.

Remark: In Cq coding theory, we have a source emitting an alphabet x with probability $p(x)$ and we encode this alphabet into a quantum state $\rho(x)$ in a Hilbert space and transmit this state over a noiseless channel. The output state

is then $\sum_x p(x)\rho(x)$ and its entropy is $H(\sum_x p(x)\rho(x))$. The entropy of the output state given the input signal is $\sum_x p(x)H(\rho(x))$. Thus, the information transmitted over the channel is

$$I(p, \rho) = H\left(\sum_x p(x)\rho(x)\right) - \sum_x p(x)H(\rho(x))$$

In case the transmitted states $\rho(x)$ are all pure, we have $H(\rho(x)) = 0$ and the information transmitted becomes

$$I(p, \rho) = H\left(\sum_x p(x)\rho(x)\right)$$

The generalization to continuous random sources is elementary:

$$I(p, \rho) = H\left(\int p(\theta)\rho(\theta)d\theta\right) - \int p(\theta)H(\rho(\theta))d\theta$$

where $p(\theta)$ is the probability density of the emitted source signal θ .

3.3.2 Quantum hypothesis testing

Let ρ, σ be two mixed states in a Hilbert space \mathcal{H} .

We wish to design a POVM test T for deciding whether ρ or σ is the true state. It is known from basic operator theory (See M.Hayashi, "An introduction to quantum information") that the optimal test that minimizes the probability of miss given an upper bound on the false alarm probability (the quantum Neyman-Pearson test) is of the form

$$T = \{\rho - c.\sigma > 0\}$$

where $\{A > 0\}$ for a Hermitian operator A , denotes the orthogonal projection onto the space spanned by the eigenvectors of A having positive eigenvalues. Then fixing the false alarm probability as ϵ means that c must be selected so that the false alarm probability is

$$P_F = Tr(\sigma\{\rho - c\sigma > 0\}) = \epsilon$$

and the probability of miss is then

$$P_M = Tr(\rho\{\rho - c\sigma \leq 0\})$$

We have

$$\begin{aligned} P_M &= Tr((\rho - c.\sigma)\{\rho - c.\sigma \leq 0\}) + c.Tr(\sigma.\{\rho - c.\sigma \leq 0\}) \\ &\leq c.(1 - P_F) = c.(1 - \epsilon) \\ &\quad \sqrt{\rho}\{\rho - c.\sigma \leq 0\}\sqrt{\rho} \end{aligned}$$

$$\begin{aligned}
&= \sqrt{\rho} \{ \rho^s \leq c^s \sigma^s \} \sqrt{\rho} \\
&\sqrt{\rho} \{ 1 \leq c^s \rho^{-s/2} \sigma^s \rho^{-s/2} \} \sqrt{\rho} \\
&\leq c^s \rho^{(1-s)/2} \sigma^s \rho^{(1-s)/2}, s > 0
\end{aligned}$$

This gives on taking trace,

$$P_M \leq c^s \text{Tr}(\rho^{1-s} \sigma^s) = \exp(s(\log(c) - D_s(\rho|\sigma)))$$

where

$$D_s(\rho|\sigma) = -s^{-1} \cdot \log(\text{Tr}(\rho^{1-s} \sigma^s))$$

Note that application of L'Hospital rule gives

$$\lim_{s \rightarrow 0} D_s(\rho|\sigma) = \text{Tr}(\rho(\log(\rho) - \log(\sigma))) = D(\rho|\sigma)$$

namely, the quantum relative entropy between σ and ρ . So, we get the result that if $\log(c) = D_s(\rho|\sigma) - \delta$, then

$$P_M \leq \exp(-s\delta)$$

On the other hand, we have for $s > 0$,

$$\begin{aligned}
\epsilon = P_F &= \text{Tr}(\sigma \cdot (\rho - c \cdot \sigma > 0)) = \text{Tr}(\sigma \{ \rho^s > c^s \sigma^s \}) \\
&\leq c^{-s} \text{Tr}(\sqrt{\sigma} \cdot \sigma^{-s/2} \rho^s \sigma^{-s/2} \sqrt{\sigma}) \\
&= c^{-s} \text{Tr}(\sigma^{1-s} \rho^s) = c^{-s} \exp(-sD_s(\sigma|\rho)) \\
&= \exp(-s(D_s(\sigma|\rho) + \log(c)))
\end{aligned}$$

Replacing s by $1 - s$ in this inequality with $s < 1$, we get

$$\begin{aligned}
P_F &\leq \exp(-(1-s)((s/(1-s))D_s(\rho|\sigma) + \log(c))) \\
&= \exp(-s \cdot D_s(\rho|\sigma) + (s-1)\log(c)), s < 1
\end{aligned}$$

It follows that assuming $s \in (0, 1)$, by selecting c so that

$$\log(c) = D_s(\rho|\sigma) - \delta$$

we get

$$P_M \leq \exp(-s\delta), P_F \leq \exp((1-s)\delta - D_s(\sigma|\rho))$$

Now consider the problem of testing the hypothesis between the tensor product state $\rho^{\otimes n}$ and $\sigma^{\otimes n}$ where $n = 1, 2, \dots$. Following the same logic, we get by considering the test $\{\rho^{\otimes n} - c^n \cdot \sigma^{\otimes n} > 0\}$, the following results

$$\begin{aligned}
P_M &= P_M(n) \leq \\
c^{ns} (\text{Tr}(\rho^{1-s} \sigma^s))^n &= \exp(s(\log(c) - D_s(\rho|\sigma)))^n \\
&= \exp(ns(\log(c) - D_s(\rho|\sigma)))
\end{aligned}$$

and

$$P_F = P_F(n) \leq \exp(n(-sD_s(\rho|\sigma) + (s-1)\log(c)))$$

It follows that by choosing

$$\log(c) = D_s(\rho|\sigma) - \delta$$

we get

$$P_M(n) \leq \exp(-ns\delta), P_F(n) \leq \exp(n((1-s)\delta - D_s(\rho|\sigma)))$$

Letting

$$0 < \delta < D_s(\rho|\sigma)/(1-s)$$

we get the result that

$$P_M(n), P_F(n) \rightarrow 0, n \rightarrow \infty$$

These inequalities are valid for all $s \in (0, 1)$. Moreover, from the above discussion, we have for all $s > 0$ and all $c > 0$ the following:

$$P_M(n) \leq \exp(ns(\log(c) - D_s(\rho|\sigma))),$$

$$P_F(n) \leq \exp(-s(D_s(\sigma|\rho) + \log(c)))$$

Thus,

$$\limsup_{n \rightarrow \infty} n^{-1} \cdot \log(P_M(n)) \leq s(\log(c) - D_s(\sigma|\rho)),$$

$$\limsup_{n \rightarrow \infty} n^{-1} \cdot \log(P_F(n)) \leq -s(D_s(\sigma|\rho) + \log(c))$$

for all $s > 0$. For $s \rightarrow 0+$, this gives

$$\limsup_{n \rightarrow \infty} n^{-1} \cdot \log(P_M(n)) \leq -s \log(c) - s D(\sigma|\rho) + o(s)$$

$$\limsup_{n \rightarrow \infty} n^{-1} \cdot \log(P_F(n)) \leq -s D(\rho|\sigma) - s \log(c) + o(s),$$

3.3.3 Some useful matrix inequalities in quantum information theory

[1]

$$S(\rho|\sigma) = \text{Tr}(\rho \cdot (\log(\rho) - \log(\sigma)))$$

is convex as a function of a pair of density matrices (ρ, σ) , ie, if $0 \leq \alpha \leq 1$, then

$$S(\alpha\rho_1 + (1-\alpha)\rho_2 | \alpha\sigma_1 + (1-\alpha)\sigma_2) \leq$$

$$\alpha S(\rho_1|\sigma_1) + (1-\alpha)S(\rho_2|\sigma_2)$$

This is deduced from Lieb's inequality. Now, if P is an orthogonal projection, then for an operator X , its pinching by P is defined to be the operator

$$PXP + (1-P)X(1-P)$$

More generally, if P_1, \dots, P_n are orthogonal projections such that $P_i P_j = 0, i \neq j$ and $\sum_{j=1}^n P_j = I$, then the pinching of X by $\{P_j : 1 \leq j \leq n\}$ is defined to be the operator

$$\mathcal{C}_P(X) = \sum_{j=1}^n P_j X P_j$$

We claim that if ρ, σ are density operators and $\{P_j\}$ a pinching as above, then

$$S(\mathcal{C}_P(\rho)|\mathcal{C}_P(\sigma)) \leq S(\rho||\sigma)$$

This inequality states that performing a quantum operation/measurement on two states cannot increase the relative entropy between them. This result is also known as "monotonicity of the quantum relative entropy".

Let P be an orthogonal projection. By an appropriate unitary transformation of our basis, we can assume that

$$P = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$$

Accordingly define the unitary operator

$$U = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}$$

Also if X is an operator of the same size, we can partition it in the same way as

$$X = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}$$

We then find that

$$\begin{aligned} PXP + (1 - P)X(1 - P) &= \begin{pmatrix} X_{11} & 0 \\ 0 & X_{22} \end{pmatrix} \\ &= \frac{1}{2}(UXU^* + X) \end{aligned}$$

Hence, if f is an operator convex function, then we have the result that for any orthogonal projection P and operator X , we have

$$\begin{aligned} f(PXP + (1 - P)X(1 - P)) &= f\left(\frac{X + UXU^*}{2}\right) \leq (1/2)(f(UXU^*) + f(X)) = \\ &= (1/2)(Uf(X)U^* + f(X)) = Pf(X)P + (1 - P)f(X)(1 - P) \end{aligned}$$

Taking trace gives

$$\text{Tr}(f(\mathcal{C}_P(X))) \leq \text{Tr}(f(X))$$

where

$$\mathcal{C}_P(X) = PXP + (1 - P)X(1 - P)$$

More generally, let \mathcal{H} be a finite dimensional Hilbert space and $f : \mathcal{B}(\mathcal{H})^k \rightarrow \mathbb{R}$ be an operator convex function, ie, for $X_1, \dots, X_k, Y_1, \dots, Y_k \in \mathcal{B}(\mathcal{H})$

$$\begin{aligned} f(\alpha X_m + (1-\alpha)Y_m, m = 1, 2, \dots, k) &\leq \alpha f(X_m, m = 1, 2, \dots, k) + (1-\alpha)f(Y_m, m \\ &= 1, 2, \dots, k) \end{aligned}$$

then by the same argument as above, we have

$$\begin{aligned} f(PX_m P + (1-P)X_m(1-P), m = 1, 2, \dots, k) &= f((UX_m U^* + X_m)/2, m = 1, 2, \dots, k) \\ &\leq (f(UX_m U^*, m = 1, 2, \dots, k) + f(X_m, m = 1, 2, \dots, k))/2 \\ &= (f(X_m, m = 1, 2, \dots, k) + Uf(X_m, m = 1, 2, \dots, k)U^*)/2 \end{aligned}$$

which gives on taking trace,

$$Tr(f(PX_m P + (1-P)X_m(1-P), m = 1, 2, \dots, k)) \leq Tr(f(X_m, m = 1, 2, \dots, k))$$

We further have the following obvious result: Let $f : \mathcal{B}(\mathcal{H})^k \rightarrow \mathbb{R}$ be unitarily invariant and convex. Then for any pinching $\{P_k\}$, we have

$$f(\mathcal{C}_P(X_1), \dots, \mathcal{C}_P(X_k)) \leq f(X_1, \dots, X_k)$$

In all these discussions, we assume $X_j, Y_j, j = 1, 2, \dots, k$ to be positive definite matrices. Here,

$$\mathcal{C}_P(X) = \sum_k P_k X P_k$$

This follows from the fact that any pinching is a product of 2×2 pinchings and if P is an orthogonal projection operator and U the associated unitary matrix, then

$$f(PXP + (1-P)X(1-P)) = f((X + UXU^*)/2) \leq (f(X) + f(UXU^*))/2 = f(X)$$

since f is convex and unitarily invariant.

Consider now the function

$$S(x, y) = x \cdot (\log(x) - \log(y))$$

defined on $\mathbb{R}_+ \times \mathbb{R}_+$. This function is convex since its Hessian is ≥ 0 :

$$\begin{aligned} H_S(x, y) &= \begin{pmatrix} S_{xx} & S_{xy} \\ S_{xy} & S_{yy} \end{pmatrix} \\ &= \begin{pmatrix} 1/x & -1/y \\ -1/y & x/y^2 \end{pmatrix} \end{aligned}$$

We have

$$Tr(H_S(x, y)) = 1/x + x/y^2 > 0,$$

$$\det H_S(x, y) = 0$$

proving that $H_S(x, y) \geq 0$. Lieb proved that for any matrix X and $t \in [0, 1]$ the function

$$f_t(T, S) = \text{Tr}(XT^tX^*S^{1-t})$$

is concave on the space of pairs (T, S) of positive definite matrices. It follows that

$$F(T, S) = \lim_{t \rightarrow 0+} f_t(T, S)/t = \text{Tr}(X \log(T) X^* S) - \text{Tr}(X X^* S \log(S))$$

is also concave. Taking $X = I$ gives the result that

$$S(S, T) = \text{Tr}(S(\log(S) - \log(T)))$$

is convex on the space (S, T) of pairs of positive definite matrices. We also note that $S(S, T)$ is unitarily invariant, ie, if U is a unitary matrix, then

$$S(USU^*, UTU^*) = S(S, T)$$

Hence, the above pinching results are applicable and establish the fact that

$$S(\mathcal{C}_P(S), \mathcal{S}_P(T)) \leq S(S, T)$$

which in the context of density matrices and quantum information theory implies that mutual information between two states cannot increase after performing a generalized measurement or a quantum operation.

3.3.4 Points to remember

[1] Entropy can be pumped into a quantum system by coupling it to a noisy bath. For example, suppose that the em field within a waveguide consists of TE and TM waves with random amplitudes and phases. Then suppose we place an atom or a quantum harmonic oscillator within the guide so that it interacts with the random classical em field. The interaction Hamiltonian can then be a random operator built out of the position and momentum observables of the atom, the randomness arising due to the random amplitudes and phases of the guide's em field. The Schrodinger unitary evolution operator in the interaction picture thus becomes a random operator and the action of this unitary operator on an initially pure atomic state will retain purity but this purity will be destroyed once we average the resulting state w.r.t. the probability measure of the random amplitudes and phases of the em fields. An approximate computation of the resulting mixed state of the atom can be made and its Von-Neumann entropy can be computed using approximate formulas for the logarithm of a perturbed matrix and also approximate perturbation theoretic formulas for the eigenvalues of a perturbed matrix.

[2] A more accurate way to model the interaction of the guide's em field with the atom is to regard the em field as a quantum em field built out of creation and annihilation operators with the em field bath being in a coherent

state. The interaction Hamiltonian of this quantum em field with the atom can be expressed as a bilinear function of these field operators and the atomic observables. We can then compute the unitary evolution of the tensor product of the initial pure atomic state with the bath coherent state after time T . This state will again be pure but when we calculate its partial trace over the bath state, the result is a mixed state for the atomic system and its entropy can be approximately calculated using matrix perturbation theory. It should be noted that in the previous model, we assume that the em field is a classical random field while in the latter case, it is a quantum field. The role played by averaging over classical randomness in the previous case is played by partial tracing over the bath state in this case. This case gives a more accurate description of how nature works.

[3] In the problem of discriminating between two quantum states ρ, σ , we have to construct a decision operator $0 \leq T \leq I$ so that $Tr(\rho T)$ is the probability of making a correct decision given that ρ is the true state and $Tr(\sigma T)$ is the probability of making an erroneous decision given that σ is the true state. Given a fixed false alarm error probability $Tr(\sigma T)$, we have to choose T so that $Tr(\rho T)$ is a maximum and this problem can easily be solved by considering the spectral resolution of $\rho - c\sigma$ where c is a real number. We can also talk of a large deviation principle here, ie try to design a test T_n for each positive integer n so that for a given error probability of the second kind $Tr(\sigma^{\otimes n} T_n)$, $Tr(\rho^{\otimes n} T_n)$ is a maximum or equivalently the error probability of the first kind $Tr(\rho^{\otimes n} (I - T_n))$ is a minimum. Denoting these error probabilities by ϵ_n and μ_n , it can be shown as in the Large deviation principle for classical hypothesis testing that we can choose a sequence of tests T_n such that $\limsup_{n \rightarrow \infty} n^{-1} \log(\mu_n) < 0$ and $\limsup_{n \rightarrow \infty} n^{-1} \log(\epsilon_n) < -D(\sigma|\rho) + \delta$ for arbitrary $\delta > 0$ where $D(\sigma|\rho)$ is the quantum relative entropy between σ and ρ . In particular, we can choose a sequence of tests such that both the error probabilities converge to zero as $n \rightarrow \infty$.

[4] Some other useful quantum information inequalities are the pinching inequality for operator convex functions and the monotonicity of the quantum relative entropy.

3.4 Questions in optimization techniques

[1](a) If A is an $n \times m$ real matrix not having full column rank, then explain how you would calculate the orthogonal projection operator onto $\mathcal{R}(A)$.

(b) If K is a closed convex subset of an infinite dimensional Hilbert space \mathcal{H} , then explain using Apollonius' theorem, how you would calculate the orthogonal projection onto K . In the course of your discussion, define a Cauchy sequence in a normed linear space and the condition for this space to be complete, ie, a Hilbert space.

[2] Write down using variational calculus the algorithm for computing the input $u(t) \in \mathbb{R}^d$ so that the cost functional

$$S[x, u, \mu] = \int_0^T \mathcal{L}(x(t), u(t), t) dt - \mu^T \psi(x(T))$$

is a minimum. Here, the state vector $x(t) \in \mathbb{R}^n$ satisfies the equation

$$x'(t) = f(t, x(t), u(t)), t \geq 0, x(0) = x_0$$

where

$$f : \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^d \rightarrow \mathbb{R}^n$$

is a smooth map. Describe two applications of this problem, one to minimum fuel problems and another to minimum time problems with end-time constraints.

[3] Describe the stochastic Bellman-Hamilton-Jacobi optimal control equations for minimizing the expected cost

$$\mathbb{E}\left[\sum_{n=0}^N L(x[n], u[n], n) | x[0]\right]$$

when $x[n]$ satisfies the following discrete time stochastic difference equation

$$x[n+1] = f(x[n], u[n+1], n) + w[n+1], n \geq 0$$

with $x[n] \in \mathbb{R}^N, u[n] \in \mathbb{R}^d$ and we impose the constraint that $u[n]$ be a non-random function of $x[n]$ only, ie, $u[n] = K(x[n])$ where $K : \mathbb{R}^N \rightarrow \mathbb{R}^d$ is a non-random map. Here, $w[n], n = 1, 2, \dots$ is assumed to be an iid sequence in \mathbb{R}^d with pdf $p(w)$.

hint: Consider the minimization problem

$$V(x[n], n) = \min_{u[k], n < k \leq N} \mathbb{E}\left[\left(\sum_{k=n}^N L(x[k], u[k], k)\right) | x[n]\right]$$

[4] [a] Derive the Klein-Gordon-Higgs equation

$$(\partial_t^2 - c^2 \nabla^2) \phi(t, r) + \mu^2 \phi(t, r) + \epsilon V(\phi(t, r)) = 0$$

from a variational principle and obtain its general solution upto $O(\epsilon)$ using perturbation theory.

[b] Derive Maxwell's equations from the variational principle

$$\delta\left[\frac{1}{2} \int (|E(t, r)|^2 - |H(t, r)|^2) dt d^3r - \int (\rho(t, r) \Phi(t, r) - (J(t, r), A(t, r))) dt d^3r\right] = 0$$

where

$$E = -\nabla \Phi - \partial_t A, H = \nabla \times A$$

and the variation is taken w.r.t the fields $\Phi(t, r), A(t, r)$.

[5] Consider the HP noisy Schrodinger equation

$$dU(t) = (-(iH + LL^*/2)dt + LdA(t) - L^*dA(t)^*)U(t)$$

where H, L, L^* are operators in the system Hilbert space \mathfrak{h} while $A(t), A(t)^*$ are the usual annihilation and creation processes in the Boson Fock space $\Gamma_s(L^2(\mathbb{R}_+))$. State the quantum Ito formula and using this prove the unitarity of $U(t)$ for all $t \geq 0$ provided that $U(0)$ is unitary in $\mathfrak{h} \otimes \Gamma_s(L^2(\mathbb{R}_+))$. Deduce the optimal Belavkin filter equation for the HP state

$$j_t(X) = U(t)(X \otimes I)U(t)^*, X \in \mathcal{B}(\mathfrak{h}), X^* = X$$

in terms of the non-demolition measurement algebra

$$\eta_t^0 = \sigma\{Y_s^o : s \leq t\}, Y_t^o = U(t)^*(I \otimes Y_t^i)U(t),$$

$$Y_t^i = cA(t) + \bar{c}A(t)^*, c \in \mathbb{C}$$

Note that the filter is of the form

$$\pi_t(X) = \mathbb{E}(j_t(X)|\eta_\theta)$$

where the expectations are calculated in the state $|f \otimes \phi(u)\rangle$, with $|f\rangle \in \mathfrak{h}$ and $|\phi(u)\rangle$ is the bath coherent state. To derive the filter assume it has the form

$$d\pi_t(X) = F_t(X)dt + G_t(X)dY_t^o$$

where $F_t(X), G_t(X)$ are in the Abelian algebra η_t^o . Prove in the course of your discussion that

$$U(T)^*(I \otimes Y_t^i)U(T) = Y_t^o, T \geq t$$

and hence deduce the non-demolition property of the measurement, ie,

$$[Y_t^o, j_s(X)] = 0, s \geq t, [Y_s^o, Y_t^o] = 0 \forall s, t \geq 0$$

Finally, explain how by performing infinitesimal control unitaries of the form $U^c(t, t+dt) = \exp(iZ \otimes dY_t^i)$ with $Z \in \mathcal{B}(\mathfrak{h}), Z^* = Z$ on the Belavkin filtered state, you can get rid of the Lindblad noise terms involving L, L^* in the Belavkin state evolution equation. Use the reference probability method in your calculations.

[6] Consider the Banach space $X = C[0, 1]$ of all continuous functions on $[0, 1]$ with norm given by

$$\|f\| = \max_{0 \leq t \leq 1} |f(t)|$$

Calculate the Frechet derivative in this space of the function

$$F : X \rightarrow \mathbb{R}$$

defined by

$$F(x) = \int_0^1 x^2(t) dt$$

In the course of your derivation define the Frechet and Gateaux derivatives of a function on a Banach space.

[7] [a] Calculate the left and right invariant Haar measures on the group of 3×3 lower triangular matrices with ones on the diagonal. [b] Give one application of the Haar measure to the image processing problem of estimating the group element that transforms one image field into another. [c] Define the Galilean group of transformations on \mathbb{R}^2 and explain how you would estimate an element from this group, namely rotation angle, translation and velocity vector from a pair of images, the second being a transform of the first. What if scaling is also taken into account.

3.5 Quantum antennas via the Schrodinger wave equation

3.5.1 Discussion

The Schrodinger wave equation in the second quantized picture is an operator field $\psi(t, r)$ that annihilates an electron located at r at time t . The Lagrangian density from which the Schrodinger equation is derived is

$$L = i\psi^*(\partial_t\psi) - (1/2m)(\nabla\psi^*, \nabla\psi) - V\psi^*\psi$$

Exercise: Verify that the variational principle

$$\delta \int L d^3r dt$$

where the variation is carried out w.r.t. ψ and ψ^* leads to the Schrodinger equation. We now compute the Hamiltonian density corresponding to this Lagrangian density:

$$\pi = \delta L / \delta \psi_{,t} = i\psi^*$$

Thus, application of the Legendre transform gives the Hamiltonian density as

$$\mathcal{H} = \pi\psi - L = (1/2m)(\nabla\psi^*, \nabla\psi) + V\psi^*\psi$$

The total energy/Hamiltonian of the Schrodinger field is then

$$H = \int \mathcal{H} d^3r = (1/2m) \int (\nabla\psi^*, \nabla\psi) d^3r + \int V\psi^*\psi d^3r$$

which gives after integration by parts,

$$H = (-1/2m) \int \psi^* \nabla^2 \psi d^3r + \int V \psi^* \psi d^3r$$

The canonical commutation relations are

$$[\psi(t, r), \pi(t, r')] = i\delta^3(r - r')$$

which results in the annihilation-creation operator commutation relations

$$[\psi(t, r), \psi^*(t, r')] = \delta^3(r - r')$$

Note that in Schrodinger's theory, positrons do not exist. The electron field is present within the cavity resonator B with boundary surface ∂B . We solve the stationary Schrodinger equation

$$(-1/2m) \nabla^2 \psi_n(r) + V \psi_n(r) = E_n \psi_n(r), n \geq 1$$

for $r \in B$ with boundary conditions $\psi(r) = 0$ for $r \in \partial B$ resulting in the eigenfunction expansion

$$\psi(t, r) = \sum_{n \geq 1} c(n) \psi_n(r) \exp(-iE_n t)$$

For the second quantized Hamiltonian, we get on using the orthonormality of $\psi_n(\cdot)$, $n = 1, 2, \dots$ in B that

$$H = \sum_n E(n) c(n)^* c(n)$$

The commutation relations

$$[\psi(t, r), \psi^*(t, r')] = \delta^3(r - r')$$

imply

$$\sum_{n,m} [c(n), c(m)^*] \psi_n(r) \psi_m^*(r') \exp(i(E_m - E_n)t) = \delta^3(r - r')$$

Combining this with the identity

$$\sum_n \psi_n(r) \psi_n(r')^* = \delta^3(r - r')$$

this results in the commutation relations

$$[c(n), c(m)^*] = \delta(n, m)$$

These are Bosonic commutation relations which do not admit the Pauli exclusion principle. So rather than commutation relations, we must postulate anticommutation relations

$$[\psi(t, r), \psi(t, r')^*]_+ = \delta^3(r - r')$$

which result in

$$[c(n), c(m)]_+ = \delta(n, m)$$

Now we compute the charge and current density operators as

$$\rho(t, r) = -e\psi(t, r)^*\psi(t, r), J(t, r) = (ie/2m)(\psi^*(t, r)\nabla\psi(t, r) - \psi(t, r)\nabla\psi^*(t, r))$$

and we find that these quantities can be expressed in terms of the operators $c(n), c(n)^*$ as

$$\rho(t, r) = -e \sum c(n)^* c(m) \psi_n(r)^* \psi_m(r) \exp(i(E_n - E_m)t),$$

$$J(t, r) = (ie/2m) \sum_{n,m} c(n)^* c(m) (\psi_n(r)^* \nabla \psi_m(r) - \psi_m(r) \nabla \psi_n(r)^*) \exp(i(E_n - E_m)t)$$

Note that $c(n)$ annihilates an electron in the state $\psi_n(\cdot)$ while $c(n)^*$ creates an electron in the state $\psi_n(\cdot)$.

Problem: Assume that the initial state of the system is $|f\rangle = |n_1, \dots, n_k\rangle$, ie, one electron each in the states $\psi_{n_j}, j = 1, 2, \dots, k$ with the other states empty. Calculate the moments

$$\langle f | \rho(t_1, r_1) \dots \rho(t_s, r_s) | f \rangle, \langle f | J_{l_1}(t_1, r_1) \dots J_{l_s}(t_s, r_s) | f \rangle$$

using the obvious identities

$$c(m) |n_1, \dots, n_k\rangle = \sum_{j=1}^k \delta(m, n_j) |n_1, \dots, \hat{n}_j, \dots, n_k\rangle,$$

$$c(m)^* |n_1, \dots, n_k\rangle = |m, n_1, \dots, n_k\rangle$$

if $m \neq n_j, j = 1, 2, \dots, k$ and zero otherwise. Using these identities, calculate the s^{th} moments of the electromagnetic four potential

$$\langle f | A^{\mu_1}(t_1, r_1) \dots A^{\mu_s}(t_s, r_s) | f \rangle$$

by making use of the retarded potential formula

$$A^\mu(t, r) = (1/4\pi) \int J^\mu(t - |r - r'|/c, r') d^3r' / |r - r'|$$

3.6 Linear algebra for quantum information theory

3.6.1 Discussion

Let A be an $n \times n$ Hermitian matrix. Let e_1, \dots, e_n be a complete orthonormal set of eigenvectors of A with eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$:

$$A|e_j\rangle = \lambda_j |e_j\rangle, j = 1, 2, \dots, n, \langle e_k | e_j \rangle = \delta(k, j)$$

Let \mathcal{M} be any k dimensional subspace of \mathbb{C}^n . Then \mathcal{M} has a non-zero intersection with $V_k = \text{span}\{e_k, e_{k+1}, \dots, e_n\}$ since the latter is $n - k + 1$ dimensional and

$$\dim(\mathcal{M} \cap V_k) = \dim \mathcal{M} + \dim V_k - \dim(\mathcal{M} + V_k) \geq k + n - k + 1 - n = 1$$

It follows that

$$\min_{x \in \mathcal{M}, \|x\|=1} \langle x, Ax \rangle \leq \lambda_k$$

On the other hand, choosing $\mathcal{M} = \text{span}\{e_1, \dots, e_k\}$, we get

$$\min_{x \in \mathcal{M}, \|x\|=1} \langle x, Ax \rangle = \lambda_k$$

We have thus proved the fundamental inequality,

$$\max_{\dim \mathcal{M}=k} \min_{x \in \mathcal{M}, \|x\|=1} \langle x, Ax \rangle = \lambda_k$$

Alternately, suppose \mathcal{M} is any $n - k + 1$ dimensional subspace of \mathbb{C}^n . Then, \mathcal{M} has a non-zero intersection with $\text{span}\{e_1, \dots, e_k\}$ and hence

$$\max_{x \in \mathcal{M}, \|x\|=1} \langle x, Ax \rangle \geq \lambda_k$$

On the other hand choosing $\mathcal{M} = \text{span}\{e_k, e_{k+1}, \dots, e_n\}$, we get

$$\max_{x \in \mathcal{M}, \|x\|=1} \langle x, Ax \rangle = \lambda_k$$

Thus, we get

$$\min_{\dim \mathcal{M}=n-k+1} \max_{x \in \mathcal{M}, \|x\|=1} \langle x, Ax \rangle = \lambda_k$$

These variational formulas can be used to approximately determine the energy levels of a quantum system given its Hamiltonian after appropriate truncation.

Now let A be as above and consider the 2^n -dimensional vector space $\wedge V$ where $V = \mathbb{C}^n$. This means that

$$\wedge V = \mathbb{C} \oplus \bigoplus_{k=1}^n \wedge^k V$$

Consider the operators

$$\Lambda_r(A) = \bigoplus_{k=0}^r I_n^{\wedge k-1} \wedge A \wedge I_n^{\wedge n-k}, r = 0, 1, \dots, n$$

$\Lambda_r(A)$ acts in $\wedge^r V$ and an orthonormal basis for $\wedge^r V$ is given by

$$\mathcal{B}_r = \{c(n, r)e_{i_1} \wedge \dots \wedge e_{i_r} : 1 \leq i_1 < i_2 < \dots < i_r \leq n\}$$

where

$$|c(n, r)|^2 \binom{n}{r} = 1$$

Here, we define

$$u_1 \wedge \dots \wedge u_r = \sum_{\sigma \in S_r} \text{sgn}(\sigma) u_{\sigma 1} \otimes \dots \otimes u_{\sigma r}$$

It is easy to see that \mathcal{B}_r is a complete onb of eigenvectors of $\Lambda_r(A)$ with eigenvalues

$$\lambda_{i_1} + \dots + \lambda_{i_r}$$

for the eigenvector

$$c(n, r)e_{i_1} \wedge \dots \wedge e_{i_r}$$

It follows that the maximum eigenvalue of $\Lambda_r(A)$ equals $\lambda_1 + \dots + \lambda_r$ and hence by the above variational principle applied to $\Lambda_r(A)$, we get

$$\max_{x \in \wedge^r V, \|x\|=1} \langle x, \Lambda_r(A)x \rangle = \lambda_1 + \dots + \lambda_r$$

Note that we also obviously have the above equality if the maximum is taken over all x of the form $x = f_1 \wedge f_2 \wedge \dots \wedge f_r$ where $f_j \in V$, $\langle f_j | f_l \rangle = \delta(j, l)$. This results in the following variational principle:

$$\max_{f_1, \dots, f_r \in V, \langle f_j, f_l \rangle = \delta(j, l), 1 \leq j, l \leq r} \sum_{j=1}^r \langle f_j, Af_j \rangle = \sum_{j=1}^r \lambda_j$$

This is called Ky-Fan's maximum principle.

Some other inequalities related to quantum information theory:

[a] Matrix Holder inequality : If A, B are positive definite matrices and $p, q > 1, 1/p + 1/q = 1$, then

$$\text{Tr}(AB) \leq (\text{Tr}(A^p))^{1/p} \cdot (\text{Tr}(B^q))^{1/q}$$

To prove this, we use Hadamard's three line theorem: If $f(z)$ is non-negative, analytic and bounded in the strip $0 \leq \text{Re } z \leq 1$, then for $z = x + iy$ with $x \in [0, 1]$, we have

$$f(z) \leq M_0^{1-x} M_1^x$$

where

$$M_0 = \sup_{y \in \mathbb{R}} f(iy), M_1 = \sup_{y \in \mathbb{R}} f(1 + iy)$$

Now put $p = 1/z$

$$f(z) = \frac{|Tr(A^z B^{1-z})|}{|Tr(A)^z \cdot Tr(B)^{1-z}|}$$

Then, $f(z)$ satisfies the above requirements. We have

$$f(iy) = 1, |f(1 + iy)| \leq Tr(A)/Tr(A) = 1$$

So, we get by applying the theorem,

$$f(x) \leq 1, x \in (0, 1)$$

which results in the desired inequality.

Remark:

3.7 Transmission lines and waveguides—Questions

[1] Derive formulas for the distributed parameters R, L, G, C of a coaxial transmission line with infinitesimal parameters ϵ, μ, σ when the radii of the inner and outer core are respectively a and b .

[2] Assume that a cylindrical cavity resonator is made of a perfect conductor of length d and radius R . The region $0 \leq z \leq d_1$ is filled with water while the region $d_1 \leq z \leq d$ is filled with air. Assume water has permittivity and permeability (ϵ_1, μ_1) while air has the corresponding parameters as (ϵ_2, μ_2) . Denote the cylindrical components of the electric field and magnetic field components in the two regions respectively by $(E_\rho^{(k)}, E_\phi^{(k)}, E_z^{(k)}), (H_\rho^{(k)}, H_\phi^{(k)}, H_z^{(k)})$, $k=1,2$ respectively. Assume that in the two regions, the dependences on z of the H_z components of the TE modes are of the form $\sin(\alpha z)$ and $\sin(\beta(d - z))$. This would guarantee that H_z vanishes at $z = 0$ and $z = d$. Express $E_\rho(\rho, \phi, z), E_\phi(\rho, \phi, z), H_\rho(\rho, \phi, z), H_\phi(\rho, \phi, z)$ in the two regions in terms of H_z in the two regions. Hence, set up the appropriate Helmholtz equations for H_z in both the regions and obtain the general solutions for the TE mode by applying the boundary conditions that E_ρ, E_ϕ vanish at $z = 0, d$ and are continuous at $z = d_1$, H_ρ, H_ϕ are continuous at $z = d_1$ (no surface current can exist at a dielectric interface), and finally, μH_z is continuous at $z = d_1$. Repeat for the TM modes.

[3] Calculate the far field radiation pattern produced by a rectangular cavity resonator with all walls being perfect electric conductors when the oscillation frequency is

$$\omega[mnp] = (m^2/a^2 + n^2/b^2 + p^2/d^2)^{1/2} \pi \sqrt{\epsilon \mu}$$

Before presenting your final expression for the far field Poynting vector, present the expressions for (a) the electric and magnetic field components within the resonator as a function of (x, y, z, t) , and (b) the induced surface current density on the surface walls of the resonator.

[4] A lossy transmission line has distributed parameters R, L, G, C . The load is $Z_L(\omega)$. Write down the expressions for (1) the propagation constant $\gamma(\omega)$, (2) The VSWR after defining it, assume a lossless line for this case, (3) The distance between two successive voltage maxima and between two voltage minima, assume a lossless line for this case (4) the distance of the first voltage maximum and first voltage minimum from the load, assume a lossless line for this case. Explain how you would calculate the wavelength of the propagating wave in the lossless case as well as the reflection coefficient (both magnitude and phase) in terms of the VSWR and the quantities defined in (3) and (4). Finally, if the frequency of operation changes from ω to ω' , compute the new VSWR in terms of the original VSWR at a distance of d from the load assuming a lossless line and purely resistive load.

[5] (a) Write down a series expansion for the inverse Fourier transform of $\exp(\gamma(\omega)z)$ where $\gamma(\omega) = \sqrt{(R + j\omega L)(G + j\omega C)}$ by writing it as

$$\gamma(\omega) = j\omega\sqrt{LC}\sqrt{(1 - jR/\omega L)(1 - jG/\omega C)}$$

and assuming that $R/L, G/C \ll |\omega|$. Use generalized functions for expressing your answer. (b) Repeat (a) without making any assumption. For this, partition the frequency line into three parts assuming $R/L > G/C$ and doing a Taylor-Laurent expansion for $\gamma(\omega)$ in each of the three regions. Use this formula to obtain a series expansion for the voltage as a function of z and time t for a lossy line when the input source voltage is $v_s(t)$ and source resistance is R_s .

hint: Use the fact that $(j\omega)^r F(\omega)$ has inverse Fourier transform $d^r f(t)/dt^r$ while $(j\omega)^{-r} F(\omega)$ has inverse Fourier transform equal to the r -fold integral of $f(t)$ from zero to t . You will also have to use the formula for the inverse Fourier transform of $u(\omega - \omega_0)$ where u is the unit step function. This is obtained by applying duality to the formula

$$\mathcal{F}u(t) = 1/j\omega + \pi\delta(\omega)$$

[6] A transmission line of length d and characteristic impedance $Z_0(\omega)$ is terminated by a load of $Z_L(\omega)$. Two stubs of length L_1, L_2 respectively terminated by impedances of $Z_{L_1}(\omega)$ and $Z_{L_2}(\omega)$ are located at distances of d_1 and d_2 respectively from the load. Find (a) The VSWR at all points along the line, (b) the impedance seen at the input $z = d$ of the line. (c) Two relations between L_1, L_2, d_1, d_2 for matching to a line of characteristic impedance $Z'_0(\omega)$. Assume the frequency of operation ω to be fixed.

[7] If a rectangular cavity resonator supports the modes $(n_1, m_1, p_1), (n_2, m_2, p_2)$ and (n_3, m_3, p_3) with respective oscillation frequencies of $\omega_1, \omega_2, \omega_3$, then determine the dimensions a, b, d of the resonator.

[8] Solve the transmission line equations approximately upto $O(\delta)$ for inhomogeneous distributed parameters, ie, the line equations taking into account line loading are

$$V_{,z}(\omega, z) + (R_0 + j\omega L_0)I(\omega, z) + \delta(R_1(z) + j\omega C_1(z))I(\omega, z) = J(\omega, z),$$

$$I_{,z}(\omega, z) + (G_0 + j\omega C_0)V(\omega, z) + \delta(G_1(z) + j\omega C_1(z))V(\omega, z) = E(\omega, z)$$

Assume source conditions

$$V(\omega, 0) = V_s(\omega), V(\omega, d) = Z_L(\omega)I(\omega, d)$$

3.7.1 Points to remember

[1] The Tx line equations are first order coupled pde's in the voltage and current along the line as a function of (t, z) . These are derived by applying the KCL and KVL to an infinitesimal section of the line with each infinitesimal section consisting of a series resistance and inductance and a parallel conductance and capacitance. Elimination of the current variable leads to the voltage variable satisfying a damped wave equation and vice versa. If the resistance and conductance per unit length are neglected ie the line is lossless, then these damped wave equations become the usual one dimensional wave equation.

[2] By taking Fourier transforms w.r.t. the time variable in the Tx line equations, we get a second order ordinary linear ordinary differential equation in the spatial variable for the voltage and current which can be solved immediately to obtain the line voltage and current in the form of a forward wave and a backward wave. These forward wave will be damped with spatial variation from the source while the backward wave will be amplified. For a lossless line no damping or amplification will be there, ie, the propagation constant will be purely imaginary.

[2] The reflection coefficient at any point on the line is the ratio of the amplitude of the backward wave to that of the forward wave. At the load end, the reflection coefficient admits a simple expression in terms of the load impedance and the characteristic impedance of the line. As we move away from the load, this reflection coefficient varies as $\exp(2\gamma z)$ where γ is the propagation constant. At any point on the line, the relation between the input line impedance and the reflection coefficient is the same as it is at the load end. This enables us to compute the input impedance at any point on the line.

[3] The VSWR is the ratio of the maximum voltage amplitude to the minimum voltage amplitude along the line. For a lossless line, it has a simple expression in terms of the magnitude of the reflection coefficient at the load end. The phase of the reflection coefficient at the load end for a lossless line and the propagation constant determines the location of the voltage maxima and minima w.r.t. the load. If these positions are known, we can then calculate the propagation constant and the phase of the reflection coefficient. Therefore, for a lossless line, knowing the VSWR and the location of the first voltage maximum/minimum as well as the distance between successive voltage maxima or minima, we can determine the reflection coefficient at the load end and the propagation constant and hence the reflection coefficient anywhere along the line and hence the input impedance at any point on the line.

[4] If a line is terminated by a load equal to its characteristic impedance, then there is no reflected wave, ie, all the power on the line is transmitted to the load. Thus, it is of interest to attach open and short circuited stubs at different points on a loaded line to make its input impedance match that of another given line. This problem can be formulated analytically and solved graphically using the Smith chart.

[5] If we express the relationship between the normalized input impedance and the reflection coefficient in terms of its real part and imaginary part, then the constant resistance curves and constant reactance curves in the reflection coefficient plane are circles. This is the fundamental property on which the construction of the Smith chart is based. As we move away from the load, the curve traced by a point on the Smith chart is a circular arc moving clockwise provided that the line is lossless, while if it is lossy, the curve is an inward going spiral.

[6] Nonuniform lines are analyzed using perturbation theory for differential equations or better still using Fourier series expansions of the distributed parameters and the line voltage and current in terms of the spatial variable and then followed by perturbation theory.

3.8 Some more matrix inequalities related to quantum information theory

3.8.1 Discussion

Consider a square matrix A and consider the block matrix

$$X = \begin{pmatrix} I & A^* \\ A & I \end{pmatrix}$$

Suppose $X \geq 0$. Then, A is a contraction, ie,

$$A^*A \leq I$$

To prove this, we choose an arbitrary matrix C of the same size as A and consider the inequality

$$[x^*, x^*C^*]X \begin{pmatrix} x \\ Cx \end{pmatrix} \geq 0$$

This expands to give

$$\|x\|^2 + \|Cx\|^2 + (\langle x, A^*Cx \rangle + \langle x, C^*Ax \rangle) \geq 0$$

This gives

$$I + C^*C + A^*C + C^*A \geq 0$$

Replacing C by $-C$ in this inequality gives

$$I + C^*C \geq A^*C + C^*A$$

This is true for all C . Now choose

$$C = A(A^*A)^{-1/2}$$

assuming that A is non-singular. Then,

$$C^*C = I$$

and we get

$$I \geq (A^*A)^{1/2}$$

which proves the claim.

3.8.2 Points to remember

[1] Quantum information theory is about the transmission of classical bits and quantum states over noisy channels and recovering the classical bits/quantum states from the received state with as small error probability as possible. In quantum information theory, we study the maximum rate at which classical/quantum information can be transmitted over the channel after appropriate encoding of the source strings or quantum state with zero limiting error probability of decoding. Classical information of an information source is given by the Shannon entropy while quantum information of a quantum state is given by the Von-Neumann entropy. Classical information theory is a special case of quantum information theory for commuting states and observables. The maximum rate at which information can be transmitted reliably over a quantum channel is the maximum mutual information of the channel over all input source probability distributions while the maximum rate at which classical information can be transmitted reliably over a quantum channel after encoding each classical source alphabet into a quantum state is given by the maximum of the classical-quantum mutual information over all source probability distributions. This is a theorem due to Winter and Holevo. The maximum rate at which quantum information can be transmitted over a quantum noisy channel is as yet an unsolved problem although many conjectures about this have been made.

[2] Matrix inequalities in quantum information theory like the matrix Holder's inequality enable us to get bounds on the error probability. These inequalities are usually derived starting from Renyi's entropy and that for the Von-Neumann entropy as a limiting case of the Renyi entropy.

[3] Another place in quantum information theory where inequalities are required is in entangled assisted quantum communication. When two persons share an entangled state, then by transmitting classical bits to each other, they can transmit a quantum state as for example in quantum teleportation. This transmission can be achieved at infinite speed in contradiction with the special theory of relativity that sets the speed of light as the limit at which information/energy can be transmitted. When complete entanglement is not possible, then errors are involved in the decoding process and the estimation of these error probabilities involves making use of quantum information theoretic inequalities involving things like the Fidelity between two states.

[4] Other kinds of quantum information inequalities are important like the concavity of quantum entropy, joint convexity of quantum relative entropy and the monotonicity of the quantum relative entropy under quantum operations. These inequalities and many more are derived from the fundamental Lieb inequality for operators. These can be found in the book by Rajendra Bhatia, "Matrix Analysis", Springer. Another important set of inequalities in quantum information theory is based on Rayleigh's variational method for calculating the eigenvalues of a Hermitian matrix or the singular values of any matrix in descending or ascending order.

3.9 Fresnel and Fraunhoffer diffraction

3.9.1 Discussion

A wave source from the plane area $(x, y) \in D$ having amplitude $f(x, y)$ per unit area produces a radiation field pattern

$$\psi(X, Y, Z) = \int_D f(x, y) \exp(-jk\sqrt{(X-x)^2 + (Y-y)^2 + Z^2}) dx dy / \sqrt{(X-x)^2 + (Y-y)^2 + Z^2}$$

where $k = \omega/c$. When $|Z|$ is very large as compared to $|X|, |Y|, |x|, |y|$ as happens in diffraction theory using the Young double-slit experiment, we can make a binomial approximation

$$\begin{aligned} \sqrt{(X-x)^2 + (Y-y)^2 + Z^2} &\approx Z(1 + ((X-x)^2 + (Y-y)^2)/2Z^2) \\ &= Z + ((X-x)^2 + (Y-y)^2)/2Z \end{aligned}$$

and we get the approximation

$$\psi(X, Y, Z) \approx |Z|^{-1} \exp(jkZ) \int_D f(x, y) \exp(((X-x)^2 + (Y-y)^2)/2Z) dx dy$$

This is the Fresnel diffraction formula. If further, we have that $|X|, |Y| \gg |x|, |y|$, then we have in addition the approximation

$$(X-x)^2 + (Y-y)^2 \approx X^2 + Y^2 - 2(Xx + Yy)$$

and the above diffraction formula approximates to

$$\psi(X, Y, Z) \approx |Z|^{-1} \exp(jk(Z + (X^2 + Y^2)/2Z)) \int_D f(x, y) \exp(-jk(Xx + Yy)) dx dy$$

This formula is known as the Fraunhoffer diffraction formula and tells us that the far field amplitude radiation pattern is the $2 - D$ spatial Fourier transform of the amplitude field over the source area D .

We wish to generalize this formula for source surfaces of arbitrary shape. Let D be an open connected set in \mathbb{R}^3 with boundary surface ∂D . The wave field $\psi(r), t = (x, y, z)$ satisfies Helmholtz equation

$$(\nabla^2 + k^2)\psi(r) = 0, r \in D$$

and on the boundary ∂D at each point, we specify either ψ or $\partial\psi/\partial\hat{n}$. Let $G_k(r, r')$ denote the Green's function corresponding to this boundary condition, ie,

$$(\nabla^2 + k^2)G_k(r, r') = \delta^3(r - r'), r, r' \in D$$

with $G_k(r, r') = 0$ if ψ is specified at $r \in \partial D$ and $\partial G_k(r, r')/\partial\hat{n} = 0$ if $\partial\psi/\partial\hat{n}$ is specified at $r \in \partial D$. Then, Green's theorem gives

$$\int_D (G_k(r, r') \nabla^2 \psi(r) - \psi(r) \nabla^2 G_k(r, r')) d^3r$$

$$= \int_{\partial D} (G_k(r, r') \partial \psi(r) / \partial \hat{n} - \psi(r) \partial G_k(r, r') / \partial \hat{n}) dS(r)$$

This gives us

$$\psi(r') = \int_{\partial D} (-G_k(r, r') \partial \psi(r) / \partial \hat{n} + \psi(r) \partial G_k(r, r') / \partial \hat{n}) dS(r), r' \in D \quad (1)$$

Now, we make the free space approximation for the Green's function:

$$G_k(r, r') \approx \frac{\exp(-jk|r - r'|)}{4\pi|r - r'|}$$

and calculate $\psi(r')$, $r' \in D$ in terms of its values or its normal derivative on ∂D . This will yield the generalized diffraction formula on which we can impose the Fresnel and Fraunhofer approximations. We have

$$\begin{aligned} \partial G_k(r, r') / \partial \hat{n} &= \\ \exp(-jk|r - r'|)(-jk(\hat{n}, r - r')/|r - r'| - (\hat{n}, r - r')/|r - r'|^3) \end{aligned}$$

and substituting this into (1) yields the desired diffraction formula. This situation corresponds to an incident wave falling on the surface ∂D which then becomes a source and generates a diffracted wave field within D . We can equivalently consider the exterior of ∂D , ie, $D^c = \mathbb{R}^3 - D$ as the region where the diffraction pattern is to be computed. The bounding surfaces for this region are ∂D with its normal being directed inward and the surface of an infinite radius sphere where all the fields vanish. In this case, we can consider the far field and near field approximations and obtain respectively the Fraunhofer and Fresnel diffraction patterns. In the far field zone, we have $|r| \gg |r'|$ and in the near field zone, $|r| \approx |r'|$.

Exercise: Compute the Fraunhofer and Fresnel diffraction pattern approximations stated in the above paragraph.

3.9.2 Points to remember

[1] When a wave field falls on a surface, it reradiates to produce a diffraction pattern in accordance to the spatial Fourier version of retarded potential theory. The near field pattern is called the Fresnel diffraction pattern and is obtained by retaining quadratic terms in the exponential of the Green's function while the far field pattern is obtained by neglecting the quadratic terms and taking into account only the linear terms in the exponential of the Green's function. This means that the far field amplitude pattern is the spatial Fourier transform of the amplitude on the screen.

[2] In the case of em waves falling on a surface, we can determine the Fresnel and Fraunhofer patterns by applying Green's theorem to the surface or equivalently, by calculating the equivalent surface electric and magnetic current densities on the surface induced by the incident em wave and then applying the retarded potential formula to compute the reradiated (ie, diffracted) field pattern in terms of the induced surface current densities.

3.10 Surface tension and wave propagation

3.10.1 Discussion

Assume that we have a water bubble of radius R in the form of a hemisphere with the flat surface being located on the water surface. The temperature of the air within the bubble is T . The volume of the hemispherical bubble is

$$V = 2\pi R^3/3$$

Let σ denote the surface tension along the circular boundary of the flat surface part of the bubble. If p is the pressure within the bubble, then we have

$$pV = nKT$$

where K is Boltzmann's constant and n is the number of air molecules within the bubble. The total pressure force that pulls the spherical surface part of the bubble away along the z axis is given by

$$\begin{aligned} & \int_0^{\pi/2} p 2\pi R^2 \sin(\theta) \cos(\theta) d\theta \\ &= \pi R^2 p \end{aligned}$$

This pressure force must be balanced by the surface tension force, ie,

$$\pi p R^2 = \tau \cdot 2\pi R$$

or equivalently,

$$\tau = pR/2$$

Note that the surface tension acts on the circular boundary along the negative z direction. Now we consider small fluctuations in the radius of the bubble. If $R(t)$ denotes the bubble radius at time t , then the pressure force along the z direction is

$$p(t)\pi R^2, p(t) = nkT/V(t) = nkT/(4\pi R^3(t)/3) = 3nkT/4\pi R^3(t)$$

and if $\tau(t)$ is the surface tension, we have the following relations: The total internal energy of the air molecules within the bubble is given by $U_i(t) = mCT(t)$ where C is the specific heat per unit mass of air. The total gravitational potential energy of the air within the bubble is

$$U_g(t) = \int_0^{R(t)} \rho(t) g z \pi (R^2(t) - z^2) dz, \rho(t) = 3m/4\pi R^3(t)$$

Since no heat is added from outside to the bubble, the thermodynamic relation

$$0 = dQ = dU + pdV, U = U_i + U_g$$

gives us one relation between $R(t)$ and $T(t)$. The other relation is $\tau = p(t)R(t)/2$ assuming that the surface tension τ does not vary with time.

Exercise: Derive using the above formulas, a differential equation satisfied by $R(t)$ and using small oscillation theory, solve approximately for $\delta R(t) = R(t) - R(0)$ by the method of linearization.

3.10.2 Points to remember

[1] The surface tension on the circumference of the circular base of a water bubble balances the air pressure within the bubble so as to prevent it from breaking. At a given air temperature, using the ideal gas law, we can compute the pressure as a function of the bubble radius and then if we assume the bubble radius to be oscillating, we can compute its equation of motion using energy conservation taking into account the gravitational potential energy of the air within the bubble. The total kinetic energy of the air molecules within the bubble can be computed using Maxwell's equipartition formula giving the relation between the kinetic energy of a molecule and the temperature. The rate of change of the kinetic plus potential energy of the bubble must be equal to the net rate at which pressure forces from outside and the surface tension on the bottom circumference do work on the bubble and this leads to bubble oscillations.

3.11 Klein-Gordon equation in the Schwarzschild space-time with a radial time independent electromagnetic field and its application to computing the Hawking temperature at which massless/massive particles are emitted from a blackhole

3.11.1 Discussion

The KG equation for a complex scalar field $\psi(t, r, \theta, \phi) = \psi(t, \mathbf{r}) = \psi(x)$ in the presence of an external em field described by the static four potential $A_\mu(\mathbf{r})$ is given by

$$(\partial_\mu + ieA_\mu)g^{\mu\nu}\sqrt{-g}(\partial_\nu + ieA_\nu)\psi(x) + \mu^2\sqrt{-g}\psi(x) = 0$$

This can be derived from the variational principle

$$\delta S[\psi, \psi^*] = 0,$$

where

$$S[\psi, \psi^*] = \int L[\psi(x), \psi^*(x), \partial_\mu \psi(x), \partial_\mu \psi^*(x)] d^4x$$

and

$$L = (1/2)g^{\mu\nu}((\partial_\mu - ieA_\mu)\psi)^*((\partial_\nu + ieA_\nu)\psi)\sqrt{-g} - \mu^2\psi^*\psi\sqrt{-g}/2$$

We now look at the Schwarzschild metric which is diagonal in the time and spherical polar coordinates:

$$g_{00} = \alpha(r), g_{11} = -\alpha(r)^{-1}, g_{22} = -r^2, g_{33} = -r^2 \sin^2(\theta),$$

so that

$$g^{00} = \alpha(r)^{-1}, g^{11} = -\alpha(r), g^{22} = -1/r^2, g^{33} = -1/(r^2 \sin^2(\theta))$$

The KG equation becomes for the Schwarzschild metric in the frequency domain, ie $\partial_t \rightarrow i\omega$,

$$\begin{aligned} & -g^{00}\sqrt{-g}(\omega + eA_0)^2\psi + (\partial_r + ieA_r)g^{11}\sqrt{-g}(\psi_{,r} + ieA_1\psi) \\ & + (\partial_\theta + ieA_2)g^{22}\sqrt{-g}(\psi_{,\theta} + ieA_2\psi) \\ & (\partial_\phi + ieA_3)g^{33}\sqrt{-g}(\psi_{,\phi} + ieA_3\psi) = \mu^2\psi\sqrt{-g} \end{aligned}$$

This becomes

$$\begin{aligned} & \alpha(r)^{-1}(\omega + eA_0)^2\psi + \frac{1}{r^2}(\partial_r + ieA_1)\alpha(r)r^2(\psi_{,r} + ieA_1\psi) \\ & + \frac{1}{r^2 \sin \theta}(\partial_\theta + ieA_2)\sin(\theta)(\psi_{,\theta} + ieA_2\psi) \\ & + \frac{1}{r^2 \sin^2(\theta)}(\partial_\phi + ieA_3)^2\psi = \mu^2\psi \end{aligned}$$

If we set $\alpha(r) = 1$, ie $m = 0$, then the usual Klein-Gordon equation of special relativity is obtained. We substitute into this equation

$$\psi(\omega, r, \theta, \phi) = \exp(iS(\omega, r, \theta, \phi))$$

and obtain using

$$\begin{aligned} \psi_{,r} &= iS_{,r}\psi, \psi_{,\theta} = iS_{,\theta}\psi, \psi_{,\phi} = iS_{,\phi}\psi, \\ \psi_{,rr} &= (iS_{,r} - S_{,r}^2)\psi, \psi_{,\theta\theta} = (iS_{,\theta\theta} - S_{,\theta}^2)\psi, \\ \psi_{,\phi\phi} &= (iS_{,\phi\phi} - S_{,\phi}^2)\psi \end{aligned}$$

then we get

$$\begin{aligned} & (\partial_r + ieA_1)\alpha(r)r^2(\psi_{,r} + ieA_1\psi) = \\ & (\partial_r + ieA_1)(r^2 - 2mr)(\psi_{,r} + ieA_1\psi) = \\ & 2(r - m)(\psi_{,r} + ieA_1\psi) + (r^2 - 2mr)(\psi_{,rr} + ieA_{1,r}\psi + 2ieA_1\psi_{,r} - e^2 A_1^2\psi) \\ & = (r^2 - 2mr)\psi_{,rr} + (2(r - m) + (r^2 - 2mr)2ieA_1)\psi_{,r} + (2(r - m)ieA_1 \\ & \quad + (r^2 - 2mr)ieA_{1,r} - e^2(r^2 - 2mr)A_1^2)\psi \end{aligned}$$

and likewise,

$$\begin{aligned} & (\partial_\theta + ieA_2)\sin(\theta)(\psi_{,\theta} + ieA_2\psi) \\ &= \sin(\theta)\psi_{,\theta\theta} + (2ieA_2\sin(\theta) + \cos(\theta))\psi_{,\theta} \\ &+ (ie(\sin(\theta)A_2)_{,\theta} - e^2A_2^2\sin(\theta))\psi \end{aligned}$$

and finally,

$$\begin{aligned} & (\partial_\phi + ieA_3)^2\psi = \\ & \psi_{,\phi\phi} + 2ieA_3\psi_{,\phi} + (ieA_{3,\phi} - e^2A_3^2)\psi \end{aligned}$$

The resulting problem is one in which there are two perturbation parameters m, e . We can formulate this problem in the following form:

$$(L_0 + mL_1 + eL_2 + emL_3 + e^2mL_4 + em^2L_5 + e^2m^2L_6)\psi = 0$$

where some of the linear partial differential operators $L_k, k = 0, 1, \dots, 6$ depend on the frequency ω and also on the particle mass μ . To solve this pde perturbatively, we need to make some assumption regarding the relative order of magnitudes of m and e . Let us say that $m = O(e^s)$, ie, we put $m = ke^s$ where k is an integer or a fraction. Then, the above pde acquires the form

$$(L_0 + ke^sL_1 + eL_2 + ke^{s+1}L_3 + ke^{s+2}L_4 + k^2e^{2s+1}L_5 + k^2e^{2s+2}L_6)\psi = 0$$

3.11.2 Points to remember

- [1] The Klein-Gordon equation for a particle in the gravitational field defined by a metric and also in the presence of an electromagnetic field can be expressed using the Laplace-Beltrami operator, ie ordinary divergence of the scalar field taking into account the em potential as a connection followed by a covariant divergence with a connection term coming from the electromagnetic potential. This equation described the dynamics of a scalar field in the presence of gravity and electromagnetic forces. When solved using perturbation theory, it can be used to compute the probability density of scalar particles outside the event horizon of a Schwarzschild blackhole after time t and if we work in the frequency domain, assuming a static em field, it can be used to compute the probability density of particles just outside the event horizon at a given frequency which on equating to $\exp(-\hbar\omega/kT)$ can be used to compute the Hawking temperature at a given frequency in terms of the background em field.

3.12 Quantum Belavkin filtering versus classical Kushner-Kallianpur filtering—A comparison

3.12.1 Summary

Quantum systems are described by the Schrodinger wave equation or by duality between states and observables, by the Heisenberg matrix mechanics equations. In the presence of noise, additional terms involving coupling between the bath processes and the system observables contribute to the evolution and the Hamiltonian in the process acquires a non-Hermitian term which tends to dissipate the energy produced by the noise fluctuations. Such a balance between fluctuation and dissipation ensures preservation of the unitarity of the joint evolution of system and bath. This is basically the heart of the Hudson-Parthasarathy noisy Schrodinger equation. Even without introducing a Boson Fock space for the bath it is possible to just add classically random terms like Brownian motion differentials modulated by Hermitian system potential operators and simultaneously introducing an Ito correction term to the Hamiltonian that is not Hermitian and adequately describes the fluctuation -dissipation theorem but adding classical noise does not capture all the possible degrees of freedom required by the GKSL equation for density evolution. Consequently various kinds of processes like the damped harmonic oscillator cannot be described by classical noise followed by stochastic averaging. This is because the potentials that modulate classical noise have to be Hermitian while the Lindblad operators can even be non-Hermitian and yet preserving the TPCP (Trace preserving completely positive map) property. In fact, the Hudson-Parthasarathy theory precisely dilates the most general GKSL equation even with non-Hermitian system operators to a differential equation that preserves the unitarity of the evolution of the system \otimes bath space. The dilation is achieved by tensoring the system space with the bath space modeled as a Boson Fock space rather than a simple $L^2(P)$ space where P is a classical probability distribution. Filtering theory can be developed for the HP noisy evolution of a system observable by making non-demolition measurements. Such a filter called the Belavkin filter after V.P.Belavkin is a non-commutative generalization of the Kushner-Kallianpur filter and differs from the latter by the presence of extra factor of order \hbar and \hbar^2 owing to the non-commutativity of the observables. In the special case when all the system observables and the random Belavkin density matrix are diagonal operators in the system Hilbert space, the Belavkin filter reduces to the classical Kushner-Kallianpur filter. One can develop the Belavkin filter even for quantum fields. For example, consider a field satisfying the three dimensional wave equation with or without potential terms, ie a wave equation of the form

$$\partial_t^2 \phi(t, r) - c^2 \nabla^2 \phi(t, r) + V(\phi(t, r)) = 0$$

We choose an orthonormal basis for $L^2(\mathbb{R}^3)$ and expand the quantum field as a linear combination of these basis functions. The coefficients will then become an infinite sequence of operator valued functions of time which satisfy the Heisenberg matrix equations which can be derived by expressing the Lagrangian of

the classical field in terms of these coefficients, applying the Legendre transformation to get the Hamiltonian of the classical field in terms of these time dependent coefficients and then introducing commutation/anticommutation relations between these coefficients now regarded as operators. These commutation/anticommutation relations can be derived from those for the original field regarded as operators and using the CCR or CAR for the position and momentum fields. The position field is the quantum field $\phi(x)$ itself and the momentum field is $\delta\mathcal{L}/\delta\partial_0\phi(x)$. Having derived the commutation/anticommutation relations for this sequence of operator coefficients varying with time, we express the Hamiltonian of the field by integrating out the Hamiltonian density over all of space in terms of these operator coefficients. To this Hamiltonian, we add quantum noise terms in the sense of Hudson and Parthasarathy and then set up the noise HP-Schrodinger equation. In other words, the system is given by a quantum field operator or equivalently by an infinite sequence of operators dependent on time alone and the bath is another type of quantum field whose underlying Hilbert space is the Boson Fock space. This is required otherwise we cannot describe random time evolution of bath noise processes. These bath quantum fields are vectors in the Boson Fock space expressed as functionals of vectors in a Hilbert space like $L^2(\mathbb{R}_+) \otimes \mathbb{C}^d$ where the $L^2(\mathbb{R}_+)$ part is required to describe random time evolution. The crux of the Hudson-Parthasarathy is to replace \mathbb{C}^d valued vectors by \mathbb{C}^d valued functions of time and thereby obtain quantum stochastic processes whose time samples have different kinds of statistics in different states of the bath. This setup provides a dynamics of quantum wave motion in a potential in the presence of quantum noise perturbations to the dynamics.

3.12.2 Discussion

The Belavkin quantum filter gives an extra $O(h)$ correction factor to the classical nonlinear filter arising from the non-commutativity of the system and measurement observables. Here, h is Planck's constant. We illustrate this by a special example in which the classical state space is \mathbb{R} and the corresponding quantum system state space is $\mathfrak{h} = L^2(\mathbb{R})$.

The classical state model is a diffusion process $x(t)$ satisfying a 1-D sde

$$dx(t) = f(x(t))t + g(x(t))dB(t)$$

where B is a Brownian motion process. The classical measurement model is

$$dz(t) = \chi(x(t))dt + \sigma_v dV(t)$$

where V is another Brownian motion independent of B . It is well known that if $\phi(x)$ is a classical system observable and we define

$$\pi_t(\phi) = \mathbb{E}(\phi(x(t))|\eta_t^o]$$

where

$$\eta_t^o = \sigma(z(s) : s \leq t)$$

so that $\pi_t(\phi)$ is the minimum mean square estimate of $\phi(x(t))$ based on noisy measurements collected upto time t , then $\pi_t(\phi)$ satisfies the Kushner-Kallianpur stochastic pde

$$d\pi_t(\phi) = \pi_t(\theta_0(\phi))dt + (\pi_t(\phi.\chi) - \pi_t(\phi).\pi_t(\chi))(dz(t) - \pi_t(\chi)dt)$$

where

$$\theta_0 = f(x)d/dx + (g(x)^2/2)d^2/dx^2$$

is the generator of the Markov process $x(t)$. For a proof of this see for example: A.Jazwinski,"Stochastic processes and filtering theory".

We note that the above classical sde can be expressed as a special commutative case of an Evans-Hudson diffusion by defining

$$j_t(\phi) = \phi(x(t))$$

so that j_t becomes a homomorphism from the algebra of bounded twice differentiable functions with bounded first and second order derivatives on \mathbb{R} into the algebra of bounded random variables on a probability space:

$$j_t(c\phi + \psi) = c\phi(x(t)) + \psi(x(t)) = cj_t(\phi) + j_t(\psi),$$

$$j_t(\phi.\psi) = (\phi.\psi)(x(t)) = \phi(x(t)).\psi(x(t)) = j_t(\phi).j_t(\psi)$$

By Ito's formula,

$$d\phi(x(t)) = \theta_0(\phi)(x(t))dt + \theta_1(\phi)(x(t))dB(t)$$

or equivalently,

$$dj_t(\phi) = j_t(\theta_0(\phi))dt + j_t(\theta_1(\phi))dB(t)$$

where

$$\theta_1 = g(x)d/dx$$

In the quantum context, the state process $x(t)$ is replaced by $j_t(X)$ where X is a bounded self-adjoint operator in a system Hilbert space \mathfrak{h} and

$$j_t(X) = U(t)^*XU(t) = U(t)^*(X \otimes I)U(t)$$

where $U(t)$ is a unitary evolution on the Hilbert space

$$\mathfrak{h} \otimes \Gamma_s(L^2(\mathbb{R}_+))$$

and satisfies the HP noisy Schrodinger equation

$$dU(t) = (-iH + LL^*/2)dt - LdA(t) + L^*dA(t)^*)U(t)$$

An easy calculation using the quantum Ito's formula

$$dA(t)dA(t)^* = dt, (dA(t))^2 = (dA(t)^*)^2 = dA(t)^*dA(t) = dt dA(t) = dt dA(t)^* = 0$$

shows that

$$d(U(t)^*U(t)) = dU^*.U + U^*dU + dU^*dU = 0$$

so that $U(t)$ is unitary since $U(0) = I$ is unitary. j_t is now again a homomorphism from the Banach algebra of bounded operators in \mathfrak{h} into the Banach algebra of bounded operators in $\mathfrak{h} \otimes \Gamma_s(L^2(\mathbb{R}_+))$. Further another application of quantum Ito formula to the equation

$$dj_t(X) = dU(t)^*XU(t) + U(t)^*XdU(t) + dU(t)^*XdU(t)$$

gives

$$dj_t(X) = j_t(\theta_0(X))dt + j_t(\theta_1(X))dA(t) + j_t(\theta_2(X))dA(t)^*$$

where now

$$\theta_0(X) = i[H, X] - (1/2)(L[L^*, X] + [X, L]L^*)$$

$$= i[H, X] - (1/2)(LL^*X + XLL^* - 2LXL^*),$$

$$\theta_1(X) = [L, X], \theta_2(X) = [X, L^*]$$

We now describe the Belavkin non-demolition measurement. The input measurement is

$$Y^i(t) = A(t) + A(t)^*$$

and the output measurement is

$$Y^o(t) = U(t)^*Y^i(t)U(t) = U(t)^*(I \otimes Y^i(t))U(t)$$

It is well known that

$$Y^o(t) = U(T)^*Y^i(t)U(T) \forall T \geq t$$

using which the non-demolition property of $Y^o(\cdot)$ follows, namely if the output measurement algebra is defined to be

$$\eta^o(t) = \sigma\{Y^o(s) : s \leq t\}$$

where now σ denotes the Von-Neumann algebra generated by its arguments, then we have

$$[\eta^o(t), j_s(X)] = 0, \forall s \geq t$$

and

$$[\eta^o(t), \eta^o(s)] = 0 \forall t, s \geq 0$$

Another application of the quantum Ito formula shows that

$$dY^o(t) = dY^i(t) + j_t(L + L^*)dt \quad (1)$$

The Belavkin filter for

$$\pi_t(X) = \mathbb{E}[j_t(X)|\eta^o(t)]$$

where the expectations are carried out with the bath in the vacuum coherent state $|e(0)\rangle$ is given by (J.Gough and Koestler, "Quantum Filtering in Coherent States")

$$d\pi_t(X) = \pi_t(\theta_0(X))dt + (\pi_t(LX+XL^*) - \pi_t(L+L^*)\pi_t(X))(dY^o(t) - \pi_t(L+L^*)dt) --- (2)$$

In the special commutative case, we choose L to be multiplication operator by the bounded real function $\chi(x)/2$ in $L^2(\mathbb{R})$ and X to be multiplication by the bounded real function $\phi(x)$ in $L^2(\mathbb{R})$ where we assume that $\mathfrak{h} = L^2(\mathbb{R})$. In that case (1) becomes

$$dY^o(t) = dB(t) + \chi(x(t))dt$$

when we take $j_t(\psi) = \psi(x(t))$ for $x(t)$ being a classical diffusion as above. Hence, the classical measurement model is recovered in this special commutative case. More generally, if we choose $L = \chi(x)/2 + i\mu(x, p)$ where μ is any Hermitian operator as a function of x, p , we get $L + L^* = \chi(x)$ but

$$LX + XL^* = \chi(x)\phi(x) + i[\mu(x, p), \phi(x)]$$

The second term is of $O(h)$ and can be viewed as a quantum correction term in the classical Kushner-Kallianpur filter. For example, $\mu(x, p) = \nu(x)p + p\nu(x)$ gives us

$$\begin{aligned} [\mu(x, p), \phi(x)] &= \nu(x)[p, \phi(x)] + [p, \phi(x)]\nu(x) \\ &= -2ih\nu(x)\phi'(x) \end{aligned}$$

where h is taken as Planck's constant divided by 2π .

Further, assuming that

$$H = p^2/2m + V(q), q = x, p = -ihd/dx$$

and

$$L = (q + ip)/2, L^* = (q - ip)/2$$

we get with $X = \phi(x)$,

$$\begin{aligned} [H, \phi] &= [p^2/2m, \phi] = (2m)^{-1}([p, \phi]p + p[p, \phi]) = (h/2m)(-i\phi'(x)p - ip\phi'(x)) \\ &= (-h^2/2m)(\phi'd/dx + (d/dx)\phi') = (-h^2/2m)(\phi''(x) + 2\phi'(x)d/dx) \end{aligned}$$

This formula should be interpreted as $(2\pi/h)[H, \phi]$ to get agreement with physics, ie, we should take

$$[H, \phi] = (-h/2m)(\phi''(x) + 2\phi'(x)d/dx)$$

Choosing as a special case $\phi(x) = x = q$, we get

$$[H, \phi] = [H, x] = (-h/m)d/dx = -ip/m$$

Further, we have for the above choice of L ,

$$L + L^* = q = x$$

Further, as before, using the notation h for $h/2\pi$, we have

$$\begin{aligned}\theta_1(q) &= [L, q] = [q + ip, q]/2 = h/2, \\ \theta_2(q) &= [q, q - ip]/2 = h/2 \\ \theta_0(q) &= i[H, q] - (1/2)(L[L^*, q] + [q, L]L^*) \\ &= p/m - (1/2)(-h(q + ip)/4 - h(q - ip)/4) = p/m + hq/4\end{aligned}$$

so that the Evans-Hudson equation for q becomes

$$dj_t(q) = j_t(p/m + hq/4)dt + h(dA(t) + dA(t)^*)/2 -- (3)$$

Remark: More generally,

$$\begin{aligned}\theta_1(\phi) &= [q + ip, \phi(q)]/2 = h\phi'(q)/2 \\ \theta_2(\phi) &= [\phi(q), q - ip]/2 = [ip, \phi(q)]/2 = h\phi'(q)/2 \\ \theta_0(\phi) &= (i/h)[p^2/2m, \phi(q)] - (1/8)((q + ip)[q - ip, \phi(q)] \\ &\quad + [\phi(q), q + ip](q - ip)) \\ &= (1/2m)(p\phi'(q) + \phi'(q)p) - (h/8)(-(q + ip)\phi'(q) - \phi'(q)(q - ip)) \\ &= (1/2m)(p\phi'(q) + \phi''(q)p) + (h/8)(2q\phi'(q) - h\phi''(q)) \\ &= (2m)^{-1}(p\phi'(q) + \phi'(q)p) + hq\phi'(q)/4 - (h^2/8)\phi''(q)\end{aligned}$$

We further have

$$\begin{aligned}\theta_1(p) &= [L, p] = [q + ip, p]/2 = ih/2, \\ \theta_2(p) &= [p, L^*] = [p, q - ip]/2 = -ih/2 \\ [L^*, p] &= [q - ip, p]/2 = ih/2, \\ [p, L] &= [p, q + ip]/2 = -ih/2, \\ [H, p] &= [V(q), p] = ihV'(q)\end{aligned}$$

so we get

$$\begin{aligned}\theta_0(p) &= -V'(q) - (1/8)((q + ip)[q - ip, p] + [p, q + ip](q - ip)) \\ &= -V'(q) - (1/8)(ih(q + ip) - ih(q - ip)) \\ &= -V'(q) + hp/4\end{aligned}$$

and so, along with (3), we have the Evans-Hudson diffusion

$$dj_t(p) = j_t(-V'(q) + hp/4)dt + (ih/2)d(A(t) - A(t)^*) -- (5)$$

The Belavkin filter for q becomes

$$d\pi_t(q) = \pi_t(\theta_0(q))dt + (\pi_t(Lq + qL^*) - \pi_t(q)^2)(dY^o(t) - \pi_t(q)dt)$$

and for p , it becomes

$$d\pi_t(p) = \pi_t(\theta_0(p))dt + (\pi_t(Lp + pL^*) - \pi_t(q)\pi_t(p))(dY^o(t) - \pi_t(p)dt)$$

we note that the measurement model becomes in this special case,

$$dY^o = dB(t) + j_t(L + L^*)dt = j_t(q)dt + dB(t), B = A + A^*$$

ie we are making noisy measurements of the position only. The Belavkin filter for $\phi(q)$ is

$$d\pi_t(\phi) = \pi_t(\theta_0(\phi))dt + (\pi_t(L\phi(q) + \phi(q)L) - \pi_t(\phi)\pi_t(q))(dY^o(t) - \pi_t(q)dt)$$

with

$$\begin{aligned} L\phi(q) + \phi(q)L &= [(q + ip)\phi(q) + \phi(q)(q - ip)]/2 \\ &= q\phi(q) + i[p, \phi(q)]/2 = q\phi(q) - h\phi'(q) \end{aligned}$$

Making all the substitutions, we find that

$$\begin{aligned} d\pi_t(\phi(q)) &= \pi_t((2m)^{-1}(p\phi'(q) + \phi'(q)p) + hq\phi'(q)/4 - (h^2/8)\phi''(q))dt \\ &\quad + (\pi_t(q\phi(q)) - \pi_t(\phi(q))\pi_t(q) - h\pi_t(\phi'(q)))(dY^o(t) - \pi_t(q)dt) \end{aligned}$$

It should be noted that classically, we have $q'(t) = p(t)/m$ and so

$$d\phi(q(t))/dt = \phi'(q)p/m$$

and the term

$$(2m)^{-1}(p\phi'(q) + \phi'(q)p)$$

appearing in the above Belavkin equation should be regarded as a quantum generalization of this formula with the other two terms respectively of $O(h)$ and $O(h^2)$ are the Lindblad noise correction terms. To get closer to the classical analogy, as at the beginning, we now take

$$L = ihg(q)p = g(q)d/dq = gD$$

so that

$$L^* = -(d/dq)g(q) = -Dg$$

and

$$L + L^* = -g'(q) = \chi(q)$$

say. Further, assume that the Hamiltonian is given by $H = -(if(q)D + iDf(q))/2 = (f(q)p + pf(q))/2, D = d/dq$. Then, the Evans-Hudson diffusion becomes for $X = \phi(q)$ is obtained as follows:

$$\theta_1(\phi) = [L, \phi(q)] = ([g(q)d/dq, \phi(q)] = g(q)\phi'(q)$$

$$\begin{aligned}
\theta_2(\phi) &= [\phi(q), L^*] = [\phi(q), -(d/dq)g(q)] = g(q)\phi'(q) \\
\theta_0(\phi) &= (i/h)[H, X] - (1/2)(L[L^*, X] + [X, L]L^*) \\
&= (1/2)(gD[Dg, \phi] + [\phi, gD]Dg) + (i/2h)[f(q)p + pf(q), \phi(q)] \\
&= (1/2)(gD\phi'g - g\phi'Dg) + f(q)\phi'(q) = (g/2)[D, \phi']g + f\phi' = f\phi' + g^2\phi''/2
\end{aligned}$$

which coincides with the classical generator of the above-mentioned diffusion process. The Evans-Hudson diffusion is then similar to the classical case discussed above:

$$\begin{aligned}
dj_t(\phi(q)) &= j_t(\theta_0(\phi))dt + j_t(\theta_1(\phi))dA + j_t(\theta_2(\phi))dA^* \\
&= j_t((gD + g^2D^2/2)\phi)dt + j_t(gD\phi)dB, B = A + A^*
\end{aligned}$$

The measurement process is

$$dY^o(t) = j_t(L + L^*)dt + dB(t) = \chi(q(t))dt + dB(t), \chi(q) = -g'(q)$$

Everything is commutative here. We next look at the Belavkin filter for this case:

$$\begin{aligned}
LX + XL^* &= gD\phi - \phi Dg = g\phi' + g\phi D - \phi g' - \phi g D \\
&= g\phi' - \phi g' = \chi\phi + g\phi'
\end{aligned}$$

The second term $g\phi'$ is the quantum correction term to the classical term. Specifically, the classical filter obtained as

$$d\pi_t(\phi) = \pi_t(\theta_0(\phi))dt + (\pi_t(\chi\phi) - \pi_t(\chi)\pi(\phi))(dY^o(t) - \pi_t(\chi)dt)$$

gets replaced in the quantum scenario by

$$d\pi_t(\phi) = \pi_t(\theta_0(\phi))dt + (\pi_t(\chi\phi) - \pi_t(\chi) + \pi_t(g\phi'))(dY^o(t) - \pi_t(\chi)dt)$$

Of course, there is a flaw in our argument since in the classical case, the process noise and measurement noise B, V are assumed to be uncorrelated, not so in the quantum set up where in this specific example, they are the same.

Remark: If we use $X(L + L^*) = \phi\chi$ in place of $LX + XL^*$, we would recover the classical result. Hence the non-commutativity of L and X produces an $O(h)$ correction term in the Belavkin filter.

More examples: Consider the classical Hamiltonian system in classical white noise with damping:

$$dq(t) = p(t)dt, dp(t) = -V'(q(t))dt - \gamma p(t)dt + \sigma dB(t)$$

The state space here is \mathbb{R}^2 , and for $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ a bounded, twice differentiable function, we define

$$j_t(\phi) = \phi(q(t), p(t))$$

Then, j_t is a homomorphism from the Banach algebra of twice continuously defined bounded functions with bounded first and second order partial derivatives

into the Banach algebra of bounded random variables in a given probability space. We have

$$\begin{aligned} dj_t(\phi) &= D_1\phi(q(t), p(t))dq(t) + D_2\phi(q(t), p(t))dp(t) + (1/2)D_2^2\phi(q(t), p(t))(dp(t))^2 \\ &= j_t((pD_1 - (V'(q) + \gamma p)D_2 + \sigma^2 D_2^2/2)\phi)dt + j_t(\sigma D_2\phi)dB(t) \end{aligned}$$

So defining the operators

$$\theta_0 = pD_1 - (V'(q) + \gamma p)D_2 + \sigma^2 D_2^2/2,$$

$$\theta_1 = \sigma D_2$$

we get the classical Evans-Hudson flow:

$$dj_t(\phi) = j_t(\theta_0(\phi))dt + j_t(\theta_1(\phi))dB(t)$$

When we take a measurement of the form

$$dz(t) = \chi(q(t), p(t))dt + \sigma dV(t)$$

where B, V are independent Brownian motion processes, the Kushner-Kallianpur filtering equation for the resulting system is

$$d\pi_t(\phi) = \pi_t(\theta_0(\phi))dt + (\pi_t(\chi\phi) - \pi_t(\chi)\pi_t(\phi))(dz(t) - \pi_t(\chi)dt)$$

In the special cases when $\phi(q, p) = \phi(q)$ and $\phi(q, p) = p$, the classical Evans-Hudson flows are

$$dj_t(\phi(q)) = j_t(p\phi'(q))dt,$$

$$dj_t(p) = j_t(-(V'(q) + \gamma p))dt + \sigma dB(t)$$

which are precisely a rewriting of the original sde of the system. The Kushner-Kallianpur equations in these cases with further $\chi = \chi(q)$ reduce to

$$d\pi_t(\phi(q)) = \pi_t(p\phi'(q))dt + (\pi_t(\chi(q)\phi(q)) - \pi_t(\chi(q))\pi_t(\phi(q))(dz(t) - \pi_t(\chi(q))dt),$$

$$d\pi_t(p) = \pi_t(-(V'(q) + \gamma p))dt + (\pi_t(p\chi(q)) - \pi_t(\chi(q))\pi_t(p))(dz(t) - \pi_t(p)dt)$$

Now we look at the quantum version of this flow using the Hudson-Parthasarathy equation: Let

$$H = p^2/2 + V(q), L = \chi(q, p)/2 + i\mu(q, p), [q, p] = ih, \mu(q, p)^* = \mu(q, p), \chi(q, p)^* = \chi(q, p)$$

Our HP equation is taken as

$$dU(t) = (-(iH + LL^*/2)dt - LdA + L^*dA^*)U(t)$$

The measurement process is

$$Y^o(t) = U(t)^*Y^i(t)U(t), Y^i(t) = A(t) + A(t)^*$$

and the bath is in the vacuum state. The system Hilbert space is now

$$\mathfrak{h} = L^2(\mathbb{R}^2)$$

We take $X = \phi(q, p)$ now a differential operator \mathfrak{h} with $p = -ihd/dq$. We can write our measurement differential as

$$dY^o(t) = dY^i(t) + j_t(L + L^*)dt = j_t(\chi)dt + dB(t)$$

which is similar to the above classical measurement model. The quantum Evans-Hudson flow is now as before

$$dj_t(X) = j_t(\theta_0(X))dt + j_t(\theta_1(X))dA + j_t(\theta_2(X))dA^*$$

where

$$\begin{aligned}\theta_0(X) &= (i/h)[H, X] - (1/2)(LL^*X + XLL^* - 2LXL^*) \\ &= (i/h)[H, X] - (1/2)(L[L^*, X] + [X, L]L^*) \\ \theta_1(X) &= [L, X], \theta_2(X) = [X, L^*]\end{aligned}$$

We consider a special case of the above choice of L to analyze quantum effects:

$$L = \chi(q)/2 + iap, L^* = \chi(q)/2 - iap, L + L^* = \chi(q), a \in \mathbb{R}$$

Assume further that $X = \phi = \phi(q)$. Then

$$[L^*, X] = [\chi(q)/2 - iap, \phi(q)] = -ah\phi'(q),$$

$$[X, L] = -ah\phi'(q),$$

$$(i/h)[H, X] = [p^2/2, \phi(q)] = (1/2)(p\phi'(q) + \phi'(q)p)$$

$$\begin{aligned}\theta_0(\phi) &= (1/2)(p\phi'(q) + \phi'(q)p) - (1/2)((\chi(q)/2 + iap)(-ah\phi'(q)) - ah\phi'(q)(\chi(q)/2 - iap)) \\ &= (1/2)(p\phi'(q) + \phi'(q)p) + (1/2)(ah\chi(q)\phi'(q) + ia^2h(p\phi'(q) - \phi'(q)p)) \\ &= (1/2)(p\phi'(q) + \phi'(q)p) + (1/2)(ah\chi(q)\phi'(q) + a^2h^2\phi''(q))\end{aligned}$$

The quantum Evans-Hudson flow for the operator $\phi(q)$ now reads

$$\begin{aligned}dj_t(\phi) &= j_t((1/2)(\{p, \phi'(q)\} + (1/2)(ah\chi(q)\phi'(q) + a^2h^2\phi''(q)))dt \\ &\quad + j_t(ah\phi'(q))dB(t), B = A + A^*\end{aligned}$$

Comparing this equation with the corresponding classical version above, we see the appearance of extra $O(h)$ and $O(h^2)$ factors. In particular, choosing $\phi(q) = q$ gives us

$$dj_t(q) = j_t(p + (1/2)a\chi(q))dt + adB(t)$$

We now compute the quantum Evans-Hudson diffusion for p . We have

$$[L, p] = -ih\chi'(q)/2 = [L^*, p]$$

Thus,

$$\theta_1(p) = -ih\chi'(q)/2, \theta_2(p) = ih\chi'(q)/2$$

and

$$\begin{aligned} \theta_0(p) &= (i/h)[V(q), p] - (1/2)(L[L^*, p] + [p, L]L^*) \\ &= -V'(q) - (1/2)((\chi(q)/2 + iap)(ih\chi'(q)/2) + (-ih\chi'(q)/2)(\chi(q)/2 - iap)) \\ &= -V'(q) + (ah/4)\{p, \chi'(q)\} \end{aligned}$$

Thus, the quantum Evans-Hudson flow for p is given by

$$dj_t(p) = j_t(-V'(q) + (ah/4)\{p, \chi'(q)\})dt - j_t(\chi'(q))ih(dA(t) - dA(t)^*)$$

Now we set up the Belavkin filter equations for $\phi(q)$ and p . The constant part c_0 of $\chi'(q)$ gives the damping term $c_0 ahp/4$ in the first term on the right and we can choose c_0 so that it equals $-\gamma p$. In that way we get the classical version plus some quantum corrections involving Planck's constant. We now set up the corresponding Belavkin quantum filter equations for the observables $\phi(q)$ and p . We have

$$\begin{aligned} L\phi(q) + \phi(q)L^* &= (\chi(q)/2 + iap)\phi(q) + \phi(q)(\chi(q)/2 - iap) \\ &= \chi(q)\phi(q) + ia[p, \phi(q)] = \chi(q)\phi(q) + ha\phi'(q) \end{aligned}$$

So the Belavkin filter is

$$\begin{aligned} d\pi_t(\phi(q)) &= \pi_t((1/2)(\{p, \phi'(q)\} + (1/2)(ah\chi(q)\phi'(q) + a^2h^2\phi''(q)))dt \\ &\quad + (\pi_t(\chi(q)\phi(q) + ha\phi'(q)) - \pi_t(\chi(q))\pi_t(\phi(q))(dz(t) - \pi_t(\phi(q))dt) \end{aligned}$$

We see the presence of the extra quantum factor $ha\phi'(q)$ in this filter proportional to Planck's constant that is not present in the classical filter. Finally,

$$Lp + pL^* = (\chi(q)p + p\chi(q))/2$$

$$d\pi_t(p) = \pi_t(-V'(q) + (ah/4)\{p, \chi'(q)\})dt + (\pi_t(\chi(q)p + p\chi(q))/2) - \pi_t(\chi(q))\pi_t(p)(dY^o(t) - \pi_t(p)dt)$$

Again the difference with the classical filter for p is seen in the presence of factors involving h and in the non-commutativity of $\chi(q), p$.

Conclusions: In the Belavkin filter $O(h)$ and $O(h^2)$ terms appear in contrast with the classical filter owing to the non-commutativity especially in the term $LX + XL^* \neq (L + L^*)X$ and also in other Lindblad factors. To illustrate this, we have taken the examples of quantization of a one dimensional classical diffusion and then a Hamiltonian system with two dimensional phase space comprising of a kinetic energy plus potential energy as the system Hamiltonian along with diffusion and dissipation terms used classically in Kubo's fluctuation-dissipation theory.

3.12.3 Exercises

- [1] Describe how you can formulate the incorporation of classical white noise in the Schrodinger evolution equation of a quantum system and maintain unitary evolution. Consider the noise to be mixtures of Brownian and Poisson differentials.

Hint:

$$dU(t) = [-(iH + P)dt - i \sum_k V_k dB_k(t) + \sum_k S_k dN_k(t)]U(t)$$

where

$$H^* = H, V_k^* = V_k, S_k^* S_k + S_k + S_k^* = 0,$$

$$P = (1/2) \sum_k V_k^2$$

- [2] Describe the three kinds of basic quantum noise processes in the HP sense, namely, how one constructs using the Weyl representation in Boson Fock space, the creation, annihilation and conservation processes.

- [3] Describe the Quantum Ito table with proof.

- [4] Prove that coherent vectors span a dense linear manifold in the Boson Fock space.

- [5] Derive the GKSL equation from the HP equation when the bath is in a coherent state by using the method of unitary state evolution followed by tracing out over the bath degrees of freedom.

- [6] Interpret the meaning of coherent states and the creation and annihilation operator fields constructed using the Weyl representation in terms of the algebra of an infinite sequence of independent quantum harmonic oscillators.

- [7] Derive the classical Kushner-Kallianpur filter using both the methods, one, direct application of Bayes rule and Ito's formula and two, using the reference probability approach mentioned in the paper by Gough and Kostler, "Quantum filtering in coherent states".

- [8] Define the notion of non-demolition measurements in the sense of Belavkin and why owing to the fact that they form an Abelian algebra commuting with future values of the HP observable evolutions there is no Heisenber uncertainty involved which leads us to the existence of conditional expectations needed in Belavkin's quantum filter.

- [9] Give an example showing that two observables do not commute and their quantum joint characteristic function in some state does not satisfy positive definitivity which is why they do not have a joint probability distribution function. Hint: Consider the Pauli spin matrices.

- [10] Derive Belavkin's filter when the input measurement is a mixture of quantum Brownian motions and quantum Poisson processes. Show that if $dY^o(t)$ is the output measurement differential, then Belavkin's filter can be expressed in the form

$$d\pi_t(X) = F_t(X)dt + \sum_{k \geq 1} G_{kt}(X)(dY^o(t))^k$$

and using Gough and Kostler's reference probability approach to the process

$$dC(t) = C(t) \sum_{k \geq 1} f_k(t) (dY^o(t))^k$$

derive an infinite sequence of linear equations for obtaining $F_t(X), G_{kt}(X), k \geq 1$.

[11] Derive an algorithm for simulating the HP and Belavkin filter equations by taking matrix elements w.r.t truncated orthonormal sets constructed from an onb for the system Hilbert space tensored with a finite o.n. set in the bath space obtained by applying the Gram-Schmidt orthonormalization process to a set of coherent/exponential vectors.

[12] Describe Lec-Bouten's method of quantum control for partially removing Lindblad noise terms in the Belavkin filter. The algorithm is based on choosing a system observable Z and applying the control unitary $U^c(t, t+dt) = \exp(iZdY^o(t))$ in the adjoint representation to the controlled state $\rho_c(t)$ at time t after passing it through the Belavkin filter during time $t, t+dt$ to obtain $\rho_c(t+dt)$. If Z is chosen appropriately, then one Lindblad term can be removed. After obtaining $\rho_c(t+dt)$, we take the next differential measurement and based on this apply the Belavkin filter followed by Lec-Bouten control to obtain $\rho_c(t+2dt)$ etc. In this way the complete quantum control algorithm is obtained for removing a part of the Lindblad noise. Hint: Read the Ph.D thesis of Lec-Bouten. Note that the Belavkin filtered state at time $t+dt$ is

$$\rho_B(t+dt) = \rho_c(t) + F_t(\rho_c(t))dt + G_t(\rho_c(t))(dY^o(t) - Tr(\rho_c(t)(M+M^*))dt) --- (1)$$

where

$$F_t(\rho_c(t)) = \theta^*(\rho_c(t))$$

where θ^* consists of the Hamiltonian and the Lindblad terms. Everything is commutative here. After applying the above infinitesimal control, the state becomes

$$\rho_c(t+dt) = U_c(t, t+dt)\rho_B(t+dt)U_c(t, t+dt)^*$$

This transformation will touch only the first term $\rho_c(t)$ in (1) and Z can be chosen so that this transformed term contains a term that cancels out one of the Lindblad terms in $\theta^*(\rho_c(t))$. For doing this, we have to view $\pi_t(Z)$ as a classical random process $Tr(\rho_c(t)Z)$ where $\rho_c(t)$ is also regarded as classical random process with values in the space of system matrices. Then,

$$U_c(t, t+dt)\rho_c(t)U_c(t, t+dt)^* =$$

$$(1+iZdY^o(t))\rho_c(t)(1-idY^o(t)Z^*) = \rho_c(t) + Z\rho_c(t)Z^*dt + i(Z\rho_c(t)dY^o(t) - dY^o(t)\rho_c(t)Z^*)$$

This will cancel one of the Lindblad terms involving Z . The other terms involving Z will be canceled from the next term in the Belavkin filter, ie, from

$$(1+iZdY^o(t))G_t(\rho_c(t))(dY^o(t) - \pi_t(M+M^*)dt)(1-idY^o(t)Z^*)$$

The term which matters in this expansion is

$$i(ZG_t(\rho_c(t)) - G_t(\rho_c(t))Z^*)$$

and this will cancel out the other Lindblad term involving Z . Replace Z by iZ . Note then that $U^c(t, t + dt)$ is then no longer unitary. So try $U^c(t, t + dt) = \exp(ZdY^o(t) - dY^o(t)Z^*)$ which is unitary.

[13] With regard to the previous problem, compute Z for optimal state tracking, ie, so that given $\rho_c(t)$, we have that $\rho_c(t + dt)$ is as close as possible in the Hilbert-Schmidt norm to a given state ρ_g . Solve this optimization problem using the gradient search method.

[14] Using the formula for the differential of an exponential in Lie group theory (See V.S.Varadarajan, "Lie Groups, Lie Algebras and their Representations", Springer), calculate the rate at which the Von-Neumann entropy of the GKSL state changes with time and derive conditions that the Lindblad operators should satisfy for this entropy rate to be positive in accordance with the second law of thermodynamics. Also derive formulas for the expected value of the rate of Belavkin filtered state entropy with the expectation being taken when the bath is in a coherent state. For this computation, you may use the fact that in the coherent state, the innovation process $W(t) = Y^o(t) - \int_0^t \pi_s(M + M^*)ds$ is a classical Brownian motion relative to the σ -field generated by the output measurements. Thus classical probability can be used to evaluate the required expectation. You may also have to use perturbation theory to approximately solve the Belavkin filter equation. The small perturbation parameter must be attached to the terms in the Belavkin filter that are nonlinear in the state.

3.13 Remark on quantum Belavkin filtering for estimating the state of a quantum vibrating string

3.13.1 Discussion

The wave equation for a one dimensional string attached to zero at its endpoints is given by

$$u_{,tt}(t, x) - u_{,xx}(t, x) = f(t, x), 0 \leq x \leq 1, u(t, 0) = u(t, 1) = 0$$

This equation can be derived from the Lagrangian density

$$\begin{aligned} \mathcal{L}(t, x, u, u_{,t}, u_{,x}) = \\ (1/2)(u_{,t}^2 - u_{,x}^2) + f(t, x)u(t, x) \end{aligned}$$

To quantize this motion, we first express the Lagrangian $L = \int_0^1 \mathcal{L} dx$ in terms of the spatial Fourier series coefficients of u :

$$u(t, x) = \sum_{n \geq 1} u_n(t) \sqrt{2} \cdot \sin(n\pi x)$$

$$f(t, x) = \sum_{n \geq 1} f_n(t) \sqrt{2} \cdot \sin(n\pi x)$$

Substituting this into the equation of motion gives us

$$u_n''(t) + (n\pi)^2 u_n(t) = f_n(t), n \geq 1$$

The corresponding Lagrangian is

$$\begin{aligned} L(t, \{u_n(t), u'_n(t)\}) = \\ (1/2) \sum_{n \geq 1} (u_n'^2(t) - n^2 \pi^2 u_n^2(t)) + f_n(t)u_n(t) \end{aligned}$$

We next compute the Hamiltonian of the string:

$$p_n(t) = \frac{\partial L}{\partial u'_n} = u'_n$$

so

$$\begin{aligned} H(t, \{u_n, p_n\}) = \sum_n p_n u'_n - L = \\ \sum_n [(1/2)(p_n^2 + n^2 \pi^2 u_n^2) - f_n(t)u_n] \end{aligned}$$

We write q_n in place of u_n so that

$$H(t) = H(t, \{q_n, p_n\}) = \sum_n [(1/2)(p_n^2 + n^2 \pi^2 q_n^2) - f_n(t)q_n]$$

To quantize this system, we postulate the canonical commutation relations:

$$[q_n, p_m] = i\delta_{n,m}$$

The noisy HP Schrodinger equation is

$$dU(t) = (-(iH(t) + LL^*/2)dt - LdA + L^*dA^*)U(t)$$

and the flow is

$$j_t(X) = U(t)(X \otimes I)U(t)^*$$

giving the Evans-Hudson flow equations

$$dj_t(X) = j_t(\theta_0(X))dt + j_t(\theta_1(X))dA + j_t(\theta_2(X))dA^*$$

where

$$\begin{aligned} \theta_0(X) &= i[H(t), X] - (1/2)(LL^*X + XLL^* - 2LXL^*), \\ &= i[H(t), X] - (1/2)(L[L^*, X] + [X, L]L^*) \\ \theta_1(X) &= [L, X], \theta_2(X) = [X, L^*] \end{aligned}$$

In particular,

$$\begin{aligned} dj_t(q_n) &= dq_n(t) = \\ j_t(\theta_0(q_n))dt + j_t(\theta_1(q_n))dA + j_t(\theta_2(q_n))dA^*, \\ dj_t(p_n) &= dp_n(t) = j_t(\theta_0(p_n))dt + j_t(\theta_1(p_n))dA + j_t(\theta_2(p_n))dA^*, \end{aligned}$$

We take

$$L = \chi(q) + i \sum_n a_n p_n = \chi(q) + ia^T p$$

where

$$q = (q_n), p = (p_n), a_n \in \mathbb{C}$$

and then we have

$$[L, q_n] = a_n, [L^*, q_n] = -\bar{a}_n, [L, p_n] = i\chi_{,n}(q), [L^*, p_n] = i\chi_{,n}(q)$$

where we assume that

$$\chi(q)^* = \chi(q),$$

and use the notation

$$\chi_{,n}(q) = \frac{\partial \chi(q)}{\partial q_n}$$

We have

$$\begin{aligned} \theta_0(q_n) &= i[H(t), q_n] - (1/2)(L[L^*, q_n] + [q_n, L]L^*) \\ &= p_n - (1/2)(-\bar{a}_n(\chi(q) + ia^T p) - a_n(\chi(q) - ia^* p)) \\ &= p_n + (Re(a_n)\chi(q) - Im(a_n a^*)p) \\ \theta_1(q_n) &= a_n, \theta_2(q_n) = \bar{a}_n, \end{aligned}$$

$$\begin{aligned}
\theta_0(p_n) &= i[H(t), p_n] - (1/2)(L[L^*, p_n] + [p_n, L]L^*) \\
&= i[n^2\pi^2q_n^2/2 - f_n(t)q_n, p_n] - (1/2)((\chi(q) + ia^T p)(i\chi_{,n}(q)) - i\chi_{,n}(q)(\chi(q) - ia^T p)) \\
&\quad = -n^2\pi^2q_n + f_n(t) + (1/2)\{a^T p, \chi_{,n}(q)\} \\
\theta_1(p_n) &= [\chi(q), p_n] = i\chi_{,n}(q), \\
\theta_2(p_n) &= -i\chi_{,n}(q)
\end{aligned}$$

Thus, the noisy Heisenberg equations (Evans-Hudson flow) for $\{q_n, p_n\}$ are given by

$$\begin{aligned}
dq_t(q_n) &= (j_t(p_n) + (Re(a_n)j_t(\chi(q)) - \sum_m Im(a_n\bar{a}_m)j_t(p_m))dt \\
&\quad + (a_n dA(t) + \bar{a}_n dA(t)^*), \\
dj_t(p_n) &= [-n^2\pi^2j_t(q_n) + f_n(t) + (1/2)j_t(\{a^T p, \chi_{,n}(q)\})]dt \\
&\quad + i j_t(\chi_{,n}(q))(dA(t) - dA(t)^*)
\end{aligned}$$

Equivalently, writing

$$X(t) = j_t(X) = U(t)^*(X \otimes I)U(t)$$

for any system observable X , we can express the above Heisenberg equations as

$$\begin{aligned}
dq_n(t) &= [p_n(t) + Re(a_n)\chi(q(t)) - \sum_m Im(a_n\bar{a}_m)p_m(t)]dt \\
&\quad + (a_n dA(t) + \bar{a}_n dA(t)^*), \\
dp_n(t) &= [-n^2\pi^2q_n(t) + f_n(t) + (1/2)\{a^T p(t), \chi_{,n}(q(t))\}]dt \\
&\quad + i\chi_{,n}(q(t))(dA(t) - dA(t)^*)
\end{aligned}$$

The measurement process is

$$Y^o(t) = U(t)(I \otimes Y^i(t))U(t)^*, Y^i(t) = A(t) + A(t)^*$$

and the Belavkin filter is

$$d\pi_t(q_n) = \pi_t($$

Left as an exercise to complete.

3.13.2 Points to remember

- [1] By quantizing a vibrating string into an infinite sequence of harmonic oscillators, we can add Lindblad noise terms to describe damping. Alternately, we can dilate the corresponding GKSL equation into the HP noisy Schrodinger equation by taking into account creation and annihilation processes of the bath and then try to estimate the state of the system based on Belavkin's theory of non-demolition measurements.

3.14 Elementary problems in robotics based on damped simple harmonic motion

3.14.1 Summary

The dynamical equations of a single robot are derived from the Euler-Lagrange equations choosing as our Lagrangian

$$L(q, q', t) = (1/2)q'^T M(q)q' - V(q) + \tau(t)^T q + d(t)^T q$$

where $q(t) \in \mathbb{R}^d$ is the angular position vector for the d links, $M(q)$ is the mass moment of inertia matrix of size $d \times d$ expressed as a function of the link angular coordinates. $V(q)$ is the potential in which the robot moves, like for example, the gravitational field. $\tau(t)$ is the torque vector applied to the different links, this can come either from the motors present at the links or from an external human operator applying a force $f(t)$ at the end-effector, ie, a the tip of the last link. $f(t)$ is transformed into a torque $\tau_f(t, q)$ using D'Alembert's principle of virtual work: If $r(q)$ denotes the position of the end-effector, then

$$(f(t), dr) = (\tau_f(t), dq)$$

so that

$$\tau_f(t) = \tau_f(t, q) = (f(t), dr/dq) = (dr(q)/dq)^T f(t)$$

where $dr(q)/dq$ is the Jacobian matrix of the transformation $q \rightarrow r(q)$. $d(t)$ represents a small disturbance to the torque. The Euler-Lagrange equations for the above system are given by

$$(M(q)q')' + V'(q) - (1/2)q'^T M'(q)(q' \otimes I_d) - \tau(t) - \tau_f(t, q) - d(t) = 0$$

These equations can be cast in the form

$$M(q)q'' + N(q, q') = \tau(t) + \tau_f(t, q) + d(t)$$

Suppose now that we are given a solution $q_0(t)$ to the above equations of motion in the absence of disturbance:

$$M(q_0(t))q_0''(t) + N(q_0(t), q_0'(t)) - \tau(t) - \tau_f(t, q_0(t)) = 0$$

and we seek a solution to the problem when the small disturbance is also present. We assume that the resulting solution $q(t)$ deviates from $q_0(t)$ by a small amount $\delta q(t)$:

$$q(t) = q_0(t) + \delta q(t)$$

Linearization gives us

$$M(q_0(t))\delta q''(t) + M'(q_0(t))(I \otimes q_0''(t))\delta q(t) + N_{,1}(q_0(t), q_0'(t))\delta q(t) + N_{,2}(q_0(t), q_0'(t))\delta q'(t) - d(t) = 0$$

This equation for $\delta q(t)$ is of the form of a simple harmonic oscillator with damping and a disturbance force $d(t)$ but with the mass matrix, the damping coefficient matrix and the spring constant matrix all being functions of time. In

order to analyze the effects of small disturbances on the dynamics of robot, we may therefore start with simple harmonic models. For example, we could try to design the pd coefficients K_p, K_d so that a desired trajectory $x_d(t), 0 \leq t \leq T$ is tracked by minimizing

$$\mathbb{E} \int_0^T (x(t) - x_d(t))^2 dt$$

when $x(t)$ satisfies the SHM equation with tracking error feedback:

$$x''(t) + \gamma x'(t) + K_0 x(t) = K_p(x_d(t) - x(t)) + K_d(x'_d(t) - x'(t)) + \sigma B'(t)$$

where $B(t)$ is a standard Brownian motion process or equivalently, in stochastic differential form,

$$dx(t) = v(t)dt,$$

$$dv(t) = [-(\gamma v(t) + K_0 x(t)) + K_p(x_d(t) - x(t)) + K_d(x'_d(t) - v(t))]dt + \sigma dB(t)$$

The computation of the above average tracking error energy in terms of K_p and K_d is not very easy and even after it has been done, the error energy will be a highly nonlinear function of K_p, K_d which is difficult to optimize. The other route is to assume that these pd coefficients are approximately known and assume that the true pd coefficients are small perturbations of the guess coefficients and then using perturbation theory for stochastic differential equations, we compute the average error energy upto quadratic terms in the pd coefficient perturbation which is then easy to optimize by just solving a system of two coupled linear algebraic equations. This system can also be quantized but before discussing how this is done, we shall mention the important problem of teleoperation and relate it to the damped simple harmonic motion of two oscillators driven by mutual feedback thereby justifying its appearance in a book like this on wave motion. Assume that two robots, one called the master robot and the other the slave robot act in teleoperation, ie, they are far separated from each other, like for example, they are located in two different countries. Taking into account teleoperation (TOP) delay, the differential equations for the two robots are

$$M_m(q_m)q''_m + N_m(q_m, q'_m) = \tau_m(t) + \sum_{k=1}^p C_m[k](q_s(t-kT) - q_m(t-kT)) + W_m(t),$$

$$M_s(q_s)q''_s + N_s(q_s, q'_s) = \tau_s(t) + \sum_{k=1}^p C_s[k](q_m(t-kT) - q_s(t-kT)) + W_s(t)$$

where W_m, W_s are respectively master and slave noise processes. This equation can be linearized as before to obtain linear stochastic differential equations with delay for the master and slave deviations $\delta q_m(t)$ and $\delta q_s(t)$ which are like damped coupled simple harmonic equations with TOP delay terms and noise. If we assume that the feedback coefficients are small perturbations of known values, ie,

$$C_m[k] = C_{m0}[k] + \delta C_m[k], C_s[k] = C_{s0}[k] + \delta C_s[k]$$

then we can use perturbation theory to solve for the perturbations $\delta q_m(t), \delta q_s(t)$ in terms of the perturbed feedback coefficients. The resulting tracking energy

$$\mathbb{E} \int_0^T \| \delta q_m(t) - \delta q_s(t) \|^2 dt$$

will then be quadratic functions of $\{\delta C_m[k], \delta C_s[k]\}$ and this quadratic function is easily minimized assuming quadratic constraints on them, ie, we cannot choose these coefficient perturbations too large for that would require spending more energy. Finally, we turn to the quantum aspect of pd control. In the theory of quantum stochastic dynamics, a fundamental role is played by the Evans-Hudson flow which has the form

$$d\dot{j}_t(X) = j_t(\theta_b^a(X)) d\Lambda_a^b(t)$$

where $\langle \cdot \rangle$ is the system Hilbert space, the Boson Fock space $\Gamma_s(L^2(\mathbb{R}_+) \otimes \mathbb{C}^d)$ is the noise Hilbert space on which the noise operators $\Lambda_b^a(t), a, b = 0, 1, 2, \dots$ satisfy the Hudson-Parthasarathy quantum Ito table:

$$\Lambda_0^0(t) = t, d\Lambda_b^a(t)d\Lambda_d^c(t) = \epsilon_d^a d\Lambda_b^c(t)$$

These processes contain quantum Brownian motions and quantum Poisson processes. The algebra of linear operators in the system Hilbert space \mathfrak{h} is denoted by $\mathcal{A}(\mathfrak{h})$ and the algebra of linear operators in $\Gamma_s(L^2[0, t] \otimes \mathbb{C}^d)$ is denoted by \mathcal{B}_t . $j_t : \mathcal{A}(\mathfrak{h}) \rightarrow \mathcal{B}_t$ is a homomorphism provided that the structure maps (linear) $\theta_b^a : \mathcal{A}(\mathfrak{h}) \rightarrow \mathcal{A}(\mathfrak{h})$ satisfy certain identities known as the structure equations. These are derived by applying the quantum Ito table to

$$j_t(XY) = j_t(X)j_t(Y)$$

and making use of the Evans-Hudson flow equations. The main idea of quantum control is to allow the structure maps θ_b^a to depend on some unknown parameter vector α and select α so that in a given state ρ in the *system* \otimes *noise* space, the tracking error energy

$$Tr(\rho \left(\int_0^T j_t(X) - X_d(t) \right)^2 dt)$$

is minimized. Here, $X_d(t)$ is a desired operator valued process in the *system* \otimes *noise* Hilbert space. $j_t(X)$ is computed approximately in terms of X and $\{\Lambda_b^a(t)\}$ by applying perturbation theory to the Evans-Hudson flow:

$$dj_t(X) = j_t(\theta_0^0(X))dt + \epsilon \sum_{a \geq 1 \text{ or } b \geq 1} j_t(\theta_b^a(X)) d\Lambda_a^b(t)$$

assuming that the solution to be expandable in a power series in ϵ :

$$j_t = j_t^{(0)} + \epsilon j_t^{(21)} + \epsilon^2 j_t^{(2)} + \dots + \epsilon^n j_t^{(n)} + \dots$$

substituting an equating equal powers of ϵ so that $j_t^{(k)}, k = 0, 1, 2, \dots$ can be recursively calculated:

$$\begin{aligned} dj_t^{(0)}(X) &= j_t(\theta_0^0(X))dt, \\ dj_t^{(n)}(X) &= j_t^{(n)}(\theta_0^0(X)) + \sum_{a \geq 1 \text{ or } b \geq 1} j_t^{(n-1)}(\theta_b^a(X))d\Lambda_a^b(t) \end{aligned}$$

A more direct way would be to iterate

$$j_t^{(n)}(X) = X + \int_0^t j_s^{(n-1)}(\theta_b^a(X))d\Lambda_a^b(s)ds, n \geq 1,$$

$$j_t^{(0)}(X) = X$$

as was originally done by Hudson and Parthasarathy to prove the existence and uniqueness of solutions to quantum stochastic differential equations.

3.14.2 Discussion

[1] Consider a damped harmonic oscillator with differential equation

$$x''(t) + \gamma x'(t) + \omega_0^2 x(t) = K_p(x_d(t) - x(t)) + K_d(x'_d(t) - x'(t)) + \sigma B'(t)$$

where x_d is the desired trajectory to be tracked and $B(\cdot)$ is standard Brownian motion. Express this equation in stochastic differential format, ie, as

$$\begin{aligned} d \begin{pmatrix} x(t) \\ v(t) \end{pmatrix} &= \\ \begin{pmatrix} 0 & 1 \\ -\omega_0^2 - K_p & -\gamma - K_d \end{pmatrix} \begin{pmatrix} x(t) \\ v(t) \end{pmatrix} dt \\ &+ \begin{pmatrix} 0 \\ K_p x_d(t) + K_d x'_d(t) \end{pmatrix} dt + \sigma \begin{pmatrix} 0 \\ 1 \end{pmatrix} dB(t) \end{aligned}$$

Design an algorithm for choosing K_p, K_d in such a way so that

$$\mathcal{E}(K_p, K_d) = \mathbb{E} \int_0^T ((x(t) - x_d(t))^2 + \alpha(v(t) - x'_d(t))^2) dt$$

is a minimum. This will involve a highly nonlinear optimization. To simplify matters, assume

$$K_p = K_{p0} + \delta K_p, K_d = K_{d0} + \delta K_d$$

where K_{p0}, K_{d0} are known and expand $\mathcal{E}(K_{p0} + \delta K_p, K_{d0} + \delta K_d)$ upto second degree in $\delta K_p, \delta K_d$ and hence derive optimal linear equations for calculating $\delta K_p, \delta K_d$ which minimize this approximate error energy.

Remark: Writing

$$A = A(K_p, K_d) =$$

$$\begin{pmatrix} 0 & 1 \\ -\omega_0^2 - K_p & -\gamma - K_d \end{pmatrix}$$

and

$$f(t) = f(t, K_p, K_d) = K_p x_d(t) + K_d x'_d(t)$$

and denoting the state transition matrix

$$\Phi(t) = \Phi(t, K_p, K_d) = \exp(tA)$$

show that

$$\begin{aligned} x(t) &= \int_0^t \Phi_{12}(t-s|\theta)(f(s|\theta)ds + \sigma dB(s)) \\ v(t) &= \int_0^t \Phi_{22}(t-s|\theta)(f(s|\theta)ds + \sigma dB(s)) \end{aligned}$$

where

$$\theta = [K_p, K_d]^T$$

Deduce that

$$\mathbb{E}(x(t)) = \int_0^t \Phi_{12}(t-s|\theta)f(s|\theta)ds,$$

$$\mathbb{E}(v(t)) = \int_0^t \Phi_{22}(t-s|\theta)f(s|\theta)ds,$$

$$\begin{aligned} Var(x(t)) &= \mathbb{E}(x(t)^2) - (\mathbb{E}(x(t)))^2 = \\ &\quad \sigma^2 \int_0^t \Phi_{12}(s|\theta)^2 ds, \end{aligned}$$

$$Var(v(t)) = \sigma^2 \int_0^t \Phi_{22}(s|\theta)^2 ds$$

and hence obtain a formula for $\mathcal{E}(K_p, K_d)$.

[2] Master and slave robots are simple harmonic oscillators. The aim is to make the slave follow the master using position and velocity error feedback with the latter driven by a hand operator force $f_m(t)$ and the slave experiences an environmental force $f_s(t)$ whose effect is to be communicated to the master via position and velocity error feedback. Thus, the master and slave robots move according to teleoperation. As a simplified model, we do not assume any teleoperation delay. The system of equations for the master and slave positions $x_m(t), x_s(t)$ are therefore given by

$$x''_m(t) + \gamma_m x'_m(t) + \alpha_m x_m(t) = K_{mp}(x_s(t) - x_m(t)) + K_{md}(x'_s(t) - x'_m(t)) + f_m(t) + \sigma_m B'_m(t),$$

$$x''_s(t) + \gamma_s x'_s(t) + \alpha_s x_s(t) = K_{sp}(x_m(t) - x_s(t)) + K_{sd}(x'_m(t) - x'_s(t)) + f_s(t) + \sigma_s B'_s(t),$$

where $B_m(\cdot), B_s(\cdot)$ are independent Brownian motion processes. Formulate these equations as a linear stochastic differential equation for the state vector

$$\xi(t) = [x_m(t), x'_m(t), x_s(t), x'_s(t)]^T$$

driven by the bivariate standard Brownian motion process $[B_m(t), B_s(t)]^T$, and formulate the energy function to be minimized w.r.t the parameter vector

$$\theta = [K_{mp}, K_{md}, K_{sp}, K_{sd}]^T$$

The energy function is defined as

$$\mathcal{E}(\theta) = \mathbb{E}\left[\int_0^T [x_m(t) - x_s(t), x'_m(t) - x'_s(t)] Q \begin{pmatrix} x_m(t) - x_s(t) \\ x'_m(t) - x'_s(t) \end{pmatrix} dt\right]$$

where Q is a 2×2 real positive definite matrix.

[3] Performance of parameter estimation algorithms using perturbation theory. Consider the problem of estimating the parameter vector θ in the sde

$$x'(t) = f(t, x(t), \theta) + G(t, x(t), \theta)W(t)$$

where $x(t) \in \mathbb{R}^n$ and $W(t)$ is n -dimensional white Gaussian noise. $f(t, x(t), \theta)$ is an $n \times 1$ vector, $G(x(t), \theta)$ is an $n \times n$ matrix and $\theta \in \mathbb{R}^p$. The Maximum likelihood estimate of θ based on the data $\{x(t) : 0 \leq t \leq T\}$ is given by

$$\begin{aligned} \hat{\theta} &= \operatorname{argmin}_{\theta} \int_0^T \|G(t, x(t), \theta)^{-1}(x'(t) - f(t, x(t), \theta))\|^2 dt \\ &= \operatorname{argmin}_{\theta} \int_0^T (x'(t) - f(t, x(t), \theta))^T (G(t, x(t), \theta)G(t, x(t), \theta)^T)(x'(t) - f(t, x(t), \theta)) dt \end{aligned}$$

This optimization may be too complex. An approximation is obtained by starting with a guess parameter value θ_0 and writing $\theta = \theta_0 + \delta\theta$ so that the above equation approximates to

$$\delta x'(t) = F_0(t)\delta x(t) + F_1(t)\delta\theta + G_0(t)W(t)$$

where

$$\begin{aligned} x(t) &= x_0(t) + \delta x(t), x'_0(t) = f(t, x_0(t), \theta_0), F_0(t) = \frac{\partial f(t, x_0(t), \theta_0)}{\partial x}, \\ F_1(t) &= \frac{\partial f(t, x_0(t), \theta_0)}{\partial \theta}, \\ G_0(t) &= G(t, x_0(t), \theta_0) \end{aligned}$$

The maximum likelihood estimate of $\delta\theta$ is then

$$\begin{aligned} \hat{\delta\theta} &= \operatorname{argmin}_{\delta\theta} \int_0^T \|G_0(t)^{-1}(\delta x'(t) - F_0(t)\delta x(t) - F_1(t)\delta\theta)\|^2 dt \\ &= \left(\int_0^T F_1(t)^T (G_0(t)G_0(t)^T)^{-1} F_1(t) dt \right)^{-1} \left(\int_0^T F_1(t)^T (G_0(t)G_0(t)^T)^{-1} (\delta x'(t) - F_0(t)\delta x(t)) dt \right) \end{aligned}$$

Problem: Substitute the solution

$$\delta x(t) = \int_0^t \Phi(t,s)(F_1(s)\delta\theta + G_0(s)dB(s)) \dots \quad (a)$$

where $\Phi(t,s)$ is the state transition matrix corresponding to $F_0(t)$ and hence calculate the covariance matrix of $\delta\theta$. A better approximation is to take our linearized model as

$$d\delta x(t) = (F_0(t)\delta x(t) + F_1(t)\delta\theta)dt + G_0(t)dB(t) + G_1(t)(\delta x(t) \otimes dB(t))$$

[4] Approximate performance analysis for parameter estimates in a damped harmonic oscillator with forcing. The parameter vector is

$$\theta = [\gamma, \alpha]^T$$

and the equation of motion is

$$x''(t) + \gamma x'(t) + \alpha x(t) = f(t) + \sigma B'(t), t \geq 0$$

The negative log-likelihood function for estimating this parameter vector is

$$L(x(t), t \in [0, T] | \theta) =$$

$$\int_0^T (x''(t) + \gamma x'(t) + \alpha x(t) - f(t))^2 dt$$

Setting the derivative of L w.r.t. θ to zero gives its ML estimate as

$$\hat{\theta}(T) = -P(T)^{-1}q(T)$$

where

$$P(T) = \begin{pmatrix} \int_0^T x'^2(t) dt & \int_0^T x'(t)x(t) dt \\ \int_0^T x'(t)x(t) dt & \int_0^T x^2(t) dt \end{pmatrix}$$

$$q(T) = \left[\int_0^T (x''(t) - f(t))x'(t) dt, \int_0^T (x''(t) - f(t))x(t) dt \right]$$

We wish to determine approximately

$$\mathbb{E}[\hat{\theta}(T)], Cov(\hat{\theta}(T))$$

Let $\Phi(t,s)$ denote the state transition matrix corresponding to

$$A = \begin{pmatrix} 0 & 1 \\ -\alpha & -\gamma \end{pmatrix}$$

Thus,

$$\Phi(t,s) = \exp((t-s)A)$$

Then,

$$x(t) = \int_0^t \Phi_{12}(t, s)(\sigma dB(s) + f(s)ds), x'(t) = \int_0^t \Phi_{22}(t, s)(\sigma dB(s) + f(s)ds) \quad \dots \quad (1)$$

We next evaluate the mean and covariance of $\hat{\theta}(T)$ upto $O(\sigma^2)$. For this we note that

$$x''(t) = f(t) - \sigma B'(t) - \gamma x'(t) - \alpha x(t) \quad \dots \quad (2)$$

where $x(t), x'(t)$ are given by (1). Then,

$$\begin{aligned} & \int_0^T (x''(t) - f(t))x'(t)dt = \\ & - \int_0^T x'(t)(\sigma B'(t) + \gamma x'(t) + \alpha x(t))dt \\ & = -\sigma \int_0^T x'(t)dB(t) + \gamma x(T) + \alpha \int_0^T x(t)dt \\ & = -\sigma \int_{0 < s < t < T} \Phi_{22}(t-s)(\sigma dB(s) + f(s)ds)dB(t) - \gamma \int_0^T \Phi_{12}(T-s)(\sigma dB(s) + f(s)ds) \\ & \quad - \alpha \int_{0 < s < t < T} \Phi_{12}(t-s)(\sigma dB(s) + f(s)ds)dt \end{aligned}$$

[4] Quantum versions of controller design based on Evans-Hudson flows.
The HP equations are specified as

$$dU(t) = -(iH(\alpha) + L(\alpha)L(\alpha)^*/2)dt - L(\alpha)dA(t) + L(\alpha)^*dA(t)^*)U(t)$$

where

$$L(\alpha) = \sum_{k=1}^p \alpha_k L_k,$$

$$H(\alpha) = H_0 + \sum_{k=1}^p \alpha_k V_k$$

where H_0, V_k are Hermitian operators in the system Hilbert space \mathfrak{h} while $\alpha_k \in \mathbb{R}$ are control parameters. The Evans-Hudson flow corresponding to this situation is described as follows. Let X be an observable in $\mathcal{B}(\mathfrak{h})$. We define

$$j_t(X) = U(t)^* X U(t) = j_t(X|\alpha), U(t) = U(t|\alpha)$$

Then, by quantum Ito's formula, $U(t|\alpha)$ is unitary for all $\alpha \in \mathbb{R}^p$ and we have

$$dj_t(X) = j_t(\theta_0(X|\alpha))dt + j_t(\theta_1(X|\alpha))dA(t) + j_t(\theta_2(X|\alpha))dA(t)^*$$

We take non-demolition measurements

$$Y^o(t) = U(t)^* Y^i(t) U(t), Y^i(t) = A(t) + A(t)^*$$

and assuming that the bath is in the vacuum state, we obtain the Belavkin filter

$$d\pi_t(X) = \pi_t(\theta_0(X))dt + (\pi_t(LX + XL^*) - \pi_t(L + L^*)\pi_t(X))(dY^o(t) - \pi_t(L + L^*)dt)$$

Here,

$$\begin{aligned}\theta_0(X) &= i[H, X] - (1/2)(LL^*X + XLL^* - 2LXL^*) \\ &= i[H, X] - (1/2)(L[L^*, X] + [X, L]L^*) \\ \theta_1(X) &= [L, X], \theta_2(X) = [X, L^*]\end{aligned}$$

The Belavkin filter can be implemented provided that we know α since L, L^*, H are functions of the parameter α . The point is that we do not know α and we wish at each stage to update it so that the state estimate $\pi_t(X)$ tracks a desired trajectory $X_d(t)$ that is measurable w.r.t the measurement algebra $\eta^o(t) = \sigma(Y^o(s) : s \leq t)$ just as $\pi_t(X)$ is. We therefore update α according to the gradient algorithm

$$d\alpha(t) = -\mu(t)dt \frac{\partial}{\partial \alpha} (\mathbb{E}(\pi_t(X|\alpha(t)) - X_d(t))^2)$$

the expectation being taken in the vacuum coherent state of the bath. Thus, we get to know $\alpha(t+dt) = \alpha(t) + d\alpha(t)$, using which we update $H(\alpha(t)), L(\alpha(t))$ to $H(\alpha(t+dt))$ and $L(\alpha(t+dt))$ respectively. The above expectation is carried out in the state $|f \otimes e(0)\rangle$ where $|f\rangle \in \mathfrak{h}$ and $|e(0)\rangle$ is the bath vacuum state, ie, vacuum exponential state. These values of the updated operators are substituted into the Belavkin filter equation to get the updated state estimate $\pi_{t+dt}(X)$. The process continues. Another approach would be to simply solve the Belavkin filter approximately using perturbation theory. One way to do this is to observe that all the quantities in the Belavkin filter equation commute and hence it can be regarded as a classical sde. We write E_{km} for a matrix with a 1 at the (k, m) and $(m, k)^{th}$ positions and zero at all the other positions. Then, the Belavkin filter becomes a coupled nonlinear sde for the processes $\pi_t(E_{km})$. We write

$$\theta_0(E_{km}) = \sum_{k',m'} \theta_0(k', m' | k, m, \alpha) E_{k'm'}$$

$$LE_{km} + E_{km}L^* = \sum_{k',m'} L(k', m' | k, m, \alpha) E_{k'm'}$$

$$L + L^* = \sum_{k',m'} M(k', m' | k, m, \alpha) E_{k'm'}$$

Then, the Belavkin filter can be expressed as

$$d\pi_t(E_{km}) = \sum_{k',m'} \theta_0(k', m' | k, m, \alpha) \pi_t(E_{k'm'}) dt$$

$$+ \left[\sum_{k',m'} L(k',m'|k,m,\alpha) \pi_t(E_{k'm'}) - \left(\sum_{k',m'} M(k',m'|k,m,\alpha) \pi_t(E_{k'm'}) \right) \pi_t(E_{km}) \right] dW(t)$$

where

$$\begin{aligned} W(t) &= Y^o(t) - \int_0^t \pi_s(L + L^*) ds \\ &= Y^o(t) - \sum_{k',m'} M(k',m'|k,m,\alpha) \int_0^t \pi_s(E_{k'm'}) ds \end{aligned}$$

[5] Derivation of the Lindblad equation from Hamiltonian dynamics after averaging over random parameters. The Schrodinger unitary evolution is

$$iU'(t) = (H_0 + \epsilon V(t))U(t)$$

Writing

$$U(t) = \exp(-itH_0)W(t),$$

we get

$$iW'(t) = \tilde{V}(t)W(t), \quad \tilde{V}(t) = \exp(itH_0)V(t).\exp(-itH_0)$$

For $\Delta \rightarrow 0$, we have,

$$\begin{aligned} W(\Delta) &= I + W'(0)\Delta + W''(0)\Delta^2/2 + O(\Delta^3) \\ &= I - i\Delta\tilde{V}(0) + (\Delta^2/2)(-\tilde{V}(0)^2 - i\tilde{V}'(0)) + O(\Delta^3) \end{aligned}$$

Now suppose $V(t) = V(t|\theta)$ depends on a random parameter vector θ . Then, we write

$$\tilde{V}(t) = \tilde{V}(t|\theta)$$

and hence

$$W(\Delta) = I - i\Delta\tilde{V}(0|\theta) - (\Delta^2/2)(\tilde{V}(0|\theta)^2 + i\tilde{V}'(0|\theta)) + O(\Delta^3)$$

We write

$$L(\theta) = \tilde{V}(0|\theta), \quad M(\theta) = \tilde{V}(0|\theta)^2 + i\tilde{V}'(0|\theta)$$

so that

$$W(\Delta) = W(\Delta|\theta) = I - i\Delta L(\theta) - (\Delta^2/2)M(\theta)$$

Note that $L(\theta)$ is Hermitian while $M(\theta)$ is not necessarily Hermitian. The state at time Δ in the interaction picture is given by

$$\rho(\Delta) = \int W(\Delta|\theta)\rho(0).W(\Delta|\theta)^*P(d\theta)$$

We have

$$\begin{aligned} \rho(\Delta|\theta) &= W(\Delta|\theta)\rho(0).W(\Delta|\theta) \\ &= (I - i\Delta L - (\Delta^2/2)M).\rho(0).(I + i\Delta L - (\Delta^2/2)M^*) \end{aligned}$$

$$= \rho(0) - i\Delta[L, \rho(0)] - (\Delta^2/2)(-2L\rho(0)L + M\rho(0) + \rho(0)M^*)$$

with neglect of $O(\Delta^3)$ terms. We note that

$$M = L^2 + iK$$

where

$$K = \tilde{V}'(0|\theta), L = \tilde{V}(0|\theta)$$

Hence, we can equivalently express the above relation as

$$\rho(\Delta|\theta) = \rho(0) - i\Delta[L, \rho(0)] - (\Delta^2/2)(i[K, \rho(0)] + L^2\rho(0) + \rho(0)L^2 - 2L\rho(0)L)$$

which explicitly shows the trace preserving property, ie,

$$\text{Tr}(\rho(\Delta|\theta)) = \text{Tr}(\rho(0))$$

3.14.3 Points to remember

[1] Linearization of the dynamical equations of a robot gives a linear second order linear differential equation for the fluctuations in the angular positions of the link. This second order differential equation in general has time varying coefficients but as a starting point, assuming that the unperturbed trajectory does not vary too fast, we can model the coefficients of this linear second order ode or sde as being constants.

[2] The problem of filtering and control of such a second order linear sde driven by white noise can be easily achieved using the Kalman filter. The Kalman filter estimates the state perturbation from continuous noisy measurements taken on the position and velocity. Once the position and velocity perturbations have been estimated, we can control the SHMD (Simple harmonic oscillator with damping) robot by feeding back a force proportional to the difference between fluctuations of the desired trajectory and the Kalman filter observer output.

[3] Two robots in teleoperation can be approximated by two SHMD's driven by white Gaussian noise. Taking teleoperation delays into account, we can feedback into the two robots integer multiples of the TOP delay of the master-slave position and velocities into both the robots thereby ensuring that the slave robot tracks the master and vice-versa. The advantage of this is that if for example, the slave robot is performing a surgery, then the effect of the environment in which it is moving can be felt by the master robot via the above mentioned TOP feedback and accordingly the master moves. Further the effect of a human operator force on the master robot can be felt by the slave robot via TOP feedback and the slave performs the surgery accordingly.

[4] Optimal control of a robot modeled as an SHMD with non-random and random forces by adjusting the pd feedback coefficients can be achieved using

a block processing algorithm. The idea is to solve the linear sde in terms of the driving forces and the pd coefficients and then calculate the mean square deviation of its position and velocity from the desired ones, integrated over time and then optimize this error energy w.r.t. the pd coefficients. Generally such an optimization is hard to do owing to the highly nonlinear dependence of the error energy on the pd coefficients. A simpler approximate way to carry out this optimization may therefore be devised in the following way: First assume some guess values of the pd coefficients and then assume that the true pd coefficients are small perturbations of these guess values and Taylor expand the error energy upto quadratic terms in the pd coefficient perturbation. This can be achieved using perturbation theory for linear sde's. After this, we set the partial derivatives of the error energy w.r.t the pd coefficient perturbations to zero thereby obtaining linear algebraic equations for the pd coefficient perturbations which are easily solved.

[5] The observer based control scheme can also be generalized to the quantum context as was first done by V.P.Belavkin. This involves first modeling a quantum SHMD with noise by a quantum stochastic differential equation (qsde) with Hamiltonian taken as that of an SHM and damping and quantum noise induced by Lindblad operators of the system coupled to a Boson Fock space bath. If the qsde is used to obtain the system state evolution alone by tracing out over the bath variables with the bath in a coherent state, then we obtain the GKSL equation which in the dual Heisenberg form gives the dynamics of a quantum SHMD as can be verified by calculating the Heisenberg evolution of the position and momentum of the oscillator. In other words, the unitary dilation of the GKSL equation by introducing a bath and quantum noise processes in the sense of Hudson and Parthasarathy can be regarded as the equivalent of the classical Langevin sde for a SHMD driven by noise. We can then take non-demolition measurements from the bath passed through the HP system and use Belavkin's quantum filter to obtain good estimates of the Heisenberg noisy observables or by duality the state of system. The Belavkin state filter is a stochastic commutative Schrodinger equation because all the variables occurring in it depend on the non-demolition measurements which form an Abelian algebra. After filtering the HP state via the Belavkin filter, we can further reduce noise or create quantum trajectory tracking by applying infinitesimal unitary control depending on appropriate system observables and the measured output noise differential or on giving an error feedback into the Evans-Hudson evolution for noisy observables (ie HP noisy Heisenberg observables). We can make the structure maps in the Evans-Hudson flow depend on the feedback coefficient parameters in such a way that unitarity of the associated HP evolution is not destroyed and simultaneously, the error energy between the Evans-Hudson observable and a given quantum observable in a given state of the system and bath after integration over time is a minimum. This method is the exact quantum analogue of the classical block processing based pd coefficient estimation for trajectory tracking.

3.15 Approximate solution to the Dirac equation in curved space-time

3.15.1 Discussion

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + h_{\mu\nu}(x) = \eta_{ab}V_\mu^a(x)V_\nu^b(x)$$

Write the tetrad as

$$V_\mu^a(x) = \delta_\mu^a + \xi_\mu^a(x)$$

Then,

$$\begin{aligned} h_{\mu\nu}(x) &= \eta_{ab}(\delta_\mu^a \xi_\nu^b + \delta_\nu^b \xi_\mu^a(x) + \xi_\mu^a(x) \xi_\nu^b(x)) \\ &= \xi_{\mu\nu}(x) + \xi_{\nu\mu}(x) + \xi_{a\mu}(x) \xi_\nu^a(x) \end{aligned}$$

or with neglect of quadratic and higher order terms in ξ_μ^a , we get

$$h_{\mu\nu} = \xi_{\mu\nu} + \xi_{\nu\mu}$$

We may therefore assume without loss of generality that

$$\xi_{\mu\nu} = \xi_{\nu\mu} = h_{\mu\nu}/2$$

Note that raising and lowering of indices are carried out w.r.t the Minkowski metric $\eta_{\mu\nu}$. We compute the spinor connection of the gravitational field upto linear orders in ξ_μ^a :

$$\Gamma_\mu = (1/2)J^{ab}V_{a\nu}V_{b;\mu}^\nu, J^{ab} = (1/4)[\gamma^a, \gamma^b]$$

Now,

$$\begin{aligned} V_{a\nu}V_{b;\mu}^\nu &= (\eta_{a\nu} + \xi_{a\nu})(V_{b,\mu}^\nu + \Gamma_{\rho\mu}^\nu V_b^\rho) \\ &= (\eta_{a\nu} + \xi_{a\nu})(\xi_{b,\mu}^\nu + \Gamma_{\rho\mu}^\nu(\delta_b^\rho + \xi_b^\rho)) \\ &= \xi_{ab,\mu} + \eta_{a\nu}\Gamma_{b\mu}^\nu = \xi_{ab,\mu} + \Gamma_{ab\mu} \\ &= \xi_{ab,\mu} + (\xi_{ab,\mu} + \xi_{a\mu,b} - \xi_{b\mu,a}) \end{aligned}$$

and since

$$J^{ab} = -J^{ba}$$

we get upto linear orders in ξ_μ^a the result that

$$\begin{aligned} \Gamma_\mu &= (1/2)J^{ab}(\xi_{a\mu,b} - \xi_{b\mu,a}) \\ &= J^{ab}\xi_{a\mu,b} \end{aligned}$$

We now improve this approximation by also considering second order contributions from ξ_μ^a to Γ_μ . We have

$$h_{\mu\nu} = \xi_{\mu\nu}(x) + \xi_{\nu\mu}(x) + \xi_{a\mu}(x) \xi_\nu^a(x)$$

Then,

$$\begin{aligned}
 V_{a\nu}V_{b:\mu}^\nu &= \\
 \eta_{a\nu} + \xi_{a\nu} &(\xi_{b,\mu}^\nu + \Gamma_{\rho\mu}^\nu(\delta_b^\rho + \xi_b^\rho)) \\
 &= \xi_{ab,\mu} + \eta_{a\nu}\Gamma_{\rho\mu}^\nu\delta_b^\rho + \\
 &\quad \eta_{a\nu}\xi_b^\rho\Gamma_{\rho\mu}^\nu + \\
 &\quad \xi_{a\nu}\xi_{b,\mu}^\nu + \\
 &\quad \xi_{a\nu}\delta_b^\rho\Gamma_{\rho\mu}^\nu \\
 &= \xi_{ab,\mu} + \eta_{a\nu}\Gamma_{b\mu}^\nu + \\
 &\quad \eta_{a\nu}\xi_b^\rho\Gamma_{\rho\mu}^\nu + \\
 &\quad \xi_{a\nu}\xi_{b,\mu}^\nu + \\
 &\quad \xi_{a\nu}\Gamma_{b\mu}^\nu
 \end{aligned}$$

We now note that upto second order terms in ξ_μ^a , we have

$$\begin{aligned}
 \xi_{a\nu}\Gamma_{b\mu}^\nu &= \\
 \eta_{\nu\alpha}\xi_{a\nu}\Gamma_{\alpha b\mu} &= \\
 \xi_a^\nu\Gamma_{\nu b\mu} &= (\xi_a^\nu)(\xi_{\nu b,\mu} + \xi_{\nu\mu,b} - \xi_{b\mu,\nu})
 \end{aligned}$$

where we are assuming without loss of generality that

$$\xi_{\mu\nu} = \xi_{\nu\mu}$$

Next,

$$\begin{aligned}
 \eta_{a\nu}\xi_b^\rho\Gamma_{\rho\mu}^\nu &= \\
 \xi_b^\rho\Gamma_{a\rho\mu} &= \\
 \xi_b^\rho(\xi_{a\rho,\mu} + \xi_{a\mu,\rho} - \xi_{\mu\rho,a}) &=
 \end{aligned}$$

Further,

$$\begin{aligned}
 \eta_{a\nu}\Gamma_{b\mu}^\nu &= \Gamma_{ab\mu} = \\
 (\xi_{ab,\mu} + \xi_{a\mu,b} - \xi_{b\mu,a}) &+ \\
 (1/2)[(\xi_{ca}\xi_b^c)_{,\mu} + (\xi_{ca}\xi_\mu^c)_{,b} - (\xi_{c\mu}\xi_b^c)_{,a}] &=
 \end{aligned}$$

Combining all these equations, we get upto second order in ξ_μ^a , Dirac's equation in a gravitational field is

$$[i\gamma^a V_a^\mu(x)(\partial_\mu + \Gamma_\mu(x)) - m]\psi(x) = 0$$

which gives upto linear orders in ξ :

$$[i\gamma^c(\delta_c^\mu + \xi_c^\mu(x))(\partial_\mu + J^{ab}\xi_{a\mu,b}) - m]\psi(x) = 0$$

neglecting the terms quadratic in ξ in this equation gives

$$[(i\gamma^\mu \partial_\mu - m) + (i\gamma^c \xi_c^\mu \partial_\mu + i\gamma^c J^{ab} \xi_{ac,b}(x))] \psi(x) = 0$$

Separating out the partial derivatives into the time and spatial components gives us

$$\begin{aligned} & [(\gamma^0 + \gamma^c \xi_c^0) i \partial_0 + (\gamma^r + \gamma^c \xi_c^r) i \partial_r \\ & + i\gamma^c J^{ab} \xi_{ac,b} - m] \psi(x) = 0 \end{aligned}$$

Now, we note that

$$\begin{aligned} (\gamma^0 + \gamma^c \xi_c^0)^{-1} &= \gamma^0 - \gamma^0 \gamma^c \gamma^0 \xi_c^0 \\ &= \gamma^0 - \gamma^0 \xi_0^0 + \gamma^r \xi_r^0 \end{aligned}$$

with neglect of quadratic and higher powers of ξ . Thus, again with neglect of quadratic and higher powers of ξ , we have ($r, s = 1, 2, 3$)

$$\begin{aligned} & (\gamma^0 + \gamma^c \xi_c^0)^{-1} (\gamma^r + \gamma^c \xi_c^r) = \\ & \gamma^0 \gamma^r + \gamma^0 \gamma^c \xi_c^r - \gamma^0 \gamma^r \xi_0^0 + \gamma^s \gamma^r \xi_s^0 \end{aligned}$$

and the approximate Dirac equation obtained is thus given by

$$\begin{aligned} & [i \partial_0 + (\alpha^r (1 - \xi_0^0) + \alpha^\mu \xi_\mu^r + \gamma^s \gamma^r \xi_s^0) i \partial_r \\ & + i \alpha^\mu J^{ab} \xi_{a\mu,b} - m \gamma^0 (1 - \xi_0^0 + \alpha^r \xi_r^0)] \psi(x) = 0 \end{aligned}$$

Writing

$$p_\mu = i \partial_\mu, p^\mu = \eta^{\mu\nu} p_\nu,$$

this equation can be rearranged as

$$(p^0 - \alpha^r p^r - m - \delta H) \psi = 0$$

where δH is the first order perturbation to the Dirac Hamiltonian by the gravitational field. It is given by ($P_r = -p_r = p^r, r = 1, 2, 3, (\alpha, P) = \alpha_r P_r$)

$$\begin{aligned} \delta H &= -\xi_0^0 (\alpha, P) + \alpha^\mu \xi_\mu^r P_r + \gamma^s \gamma^r \xi_s^0 P_r \\ &- i \alpha^\mu J^{ab} \xi_{a\mu,b} - m \gamma^0 (\xi_0^0 - \alpha^r \xi_r^0) \end{aligned}$$

This matrix valued differential operator is not Hermitian. We must extract out its Hermitian part and then rewrite the perturbed Dirac Hamiltonian for computing the transition probabilities for the electron between two states.

3.15.2 Points to remember

[1] To obtain a covariant version of Dirac's relativistic equation in a background curved space-time metric, we must use the tetrad formalism to generate Dirac matrix fields and also introduce a spinor connection for the gravitational field both of which ensure that the modified Dirac equation is invariant under local Lorentz transformations, not only under global Lorentz transformations.

[2] Once Dirac's equation in curved space-time has been formulated, the question of solving it approximately using perturbation theory must be developed. To do so, we observe that the spinor connection of the gravitational field is expressible easily in terms of the tetrad and hence if the gravitational field is weak, we can upto first order of smallness, express the metric perturbations linearly in terms of the tetrad fields which are assumed to be small perturbations of the standard Minkowski tetrad. In this way, after extracting out the Hermitian part of the generalized Dirac Hamiltonian in curved space-time, we can apply the standard methods of perturbation theory for pde's to derive formulas for the perturbation in the wave function caused by the gravitational field. If the gravitational field is time independent, then we can used this formalism with time independent perturbation theory to compute the first order shift in the energy levels of an atom caused by a static gravitational field. On the other hand, if the gravitational field is time dependent, then we can compute for example, the transition probability per unit time for an electron between two stationary states of the Dirac Hamiltonian in say the Coulomb potential under the influence of gravitational radiation coming from a time varying gravitational wave source like a pulsar or the collision between two blackholes.

3.16 Some applications of quantum gate design using physical systems

3.16.1 Discussion

[1] The classical FFT of signals used in spectral analysis of signals as well as in computing the system output for a given input is an algorithm of $O(N \log N)$ complexity. On the other hand, the quantum Fourier transform using Hadamard and other control unitary gates is an $O(N)$ complexity algorithm.

[2] Quantum communication using techniques like teleportation and superdense coding are infinitely faster than classical communication using bit strings encoded into electromagnetic signals and transmitted over a channel. In teleportation, two members share a maximally entangled state of d dits and by transmitting $\log_2(2d)$ classical bits, it is possible to transmit $\log_2(d)$ -qubits of quantum information. Transmitting a classical bit involves only transmitting a one or a zero while transmitting a qubit involves transmitting three real numbers and using entangled states, this can be carried out at infinite speed. The basis of this is the EPR paradox which is actually not a paradox. Two members share an entangled state consisting of two particles of spin 1/2. Owing to entanglement, if the first member measures spin +1/2 of the particle, then the second member will measure spin -1/2 for his particle and vice-versa. Thus, even when the members are far separated, transmission of information at infinite speed is possible.

[3] While estimating quantum observables like the spin of an electron or its position and momentum from noisy measurements, classical filtering breaks down owing to the presence of Heisenberg uncertainty principles manifested in the form of non-commutativity of observables. So $O(h)$ corrections are involved in the results of classical filtering theory. Measuring quantum observables becomes important primarily owing to the fact that in the future all computers will get replaced by quantum computers which rather than registers to store bits will involve quantum registers to store qubits, ie, the states of atomic particles and when we measure the signals in these quantum registers, the collapse postulate will take place thereby disturbing the state of the registers. So for accurately measuring signals in registers that are changing with time, we would need to use non-demolition measurements and quantum filtering ideas based on such measurements.

[4] An important problem in physics is to determine the energy levels of a quantum system like an atom, a molecule or a crystal consisting of an array of atoms or molecules. The quantum system will be described by a Hamiltonian operator H and its action on a given initial state after time t will be determined by the Schrodinger unitary operator $U(t) = \exp(-itH)$. If $E_n, n = 1, 2, \dots$ are the energy levels of the system, then the eigenvalues of $U(t)$ will be $\exp(-itE_n), n = 1, 2, \dots$. So the quantum computation problem of estimating the phases of a unitary matrix using the quantum Fourier transform (see Nielsen and Chuang, "Quantum computation and quantum information, Cambridge University Press") can be applied to this situation to estimate the energy levels of the atomic system. If instead, we have a semiconductor device with energy bands then the band structure can be determined by a similar method. Quantum gates designed using physical systems can be used in the phase estimation problem especially in the quantum Fourier transform part.

[5] In quantum communication, we wish to transmit a state ρ over a noisy quantum channel and recover this state at the receiving end. For example, we can excite an atom at the transmitter end with a message electromagnetic field, thereby altering its state and transmit this state over a channel whose recovery at the receiving end will enable us to determine accurately what the message electromagnetic field was. This is done using the Knill-Laflamme theorem (See Nielsen and Chuang, "Quantum computation and quantum information", Cambridge University Press). The noisy channel produces a received state given by

$$\mathfrak{k}_1(\rho) = \sum_{m=1}^p E_m \rho E_m^*$$

and at the receiving end, we need to determine recovery operators $\{R_k\}$ that depend only on the channel noise operators $\{E_m\}$, so that the output state of the receiver

$$\mathfrak{k}_2 \circ \mathfrak{k}_1(\rho) = \sum_m R_m \mathfrak{k}_1(\rho) R_m^* = \rho$$

This perfect recovery is possible only if the channel noise operators $\{E_m\}$ and the range space of the transmitted state ρ satisfy the conditions of the conditions of the Knill-Laflamme theorem. In this case, the design of the recovery operators $\{R_m\}$ can be reduced to a quantum gate design problem by using Stinespring/Choi-Kraus representation of a quantum operation in terms of a reference state, a unitary gate and a measurement involving partial trace. The design of the unitary gate can be achieved using physical systems. Further, the simulation of the quantum channel $\mathfrak{k}_1 = \{E_m\}$ can also be accomplished using unitary gate design in the Choi-Kraus representation:

$$\|_1(\rho) = Tr_2(U(\rho \otimes \rho_0)U^*)$$

(Reference: M.Hayashi, "Quantum Information", Springer, 2017).

[6] In Cq communication, we wish to transmit a string of N statistical independent symbols from an alphabet A with each symbol x having a probability $p(x)$ associated with it. We associate a quantum state $\rho(x)$ with each symbol x in the alphabet. Thus, in order to transmit the string $(x_1x_2\dots x_N)$ we transmit the quantum state $\rho_N(x_1, \dots, x_N) = \rho(x_1) \otimes \dots \otimes \rho(x_N)$. Assuming the channel to be noiseless, at the receiver end, we wish to decode the string (x_1, x_2, \dots, x_N) from the received state ρ_N with very small error probability. The Cq coding theorem proved by Winter and Holevo (See the book K.R.Parthasarathy, "Coding theorems of classical and quantum information theory", Hindustan book agency) states that, for all sufficiently large N , we can select M_N distinct source alphabet strings u_1, \dots, u_N of length N so that after these strings are encoded into quantum states as above, then detection operators D_1, \dots, D_M can be constructed at the receiver end, so that the correct detection probabilities $Tr(\rho_N(u_k)D_k)$, $k = 1, 2, \dots, M$ are arbitrarily close to unity (ie, converge to unity as $N \rightarrow \infty$) provided that $\lim_{N \rightarrow \infty} \log(M_N)/N \leq C$ where C is the Cq channel capacity defined by

$$C = C(\rho) = \max_{\{p(x)\}} [S(\sum_x p(x)\rho(x)) - \sum_x p(x)S(\rho(x))]$$

The design of the detection operators D_k , $k = 1, 2, \dots, M$ can be achieved by reducing POVM measurements to PVM measurements and then implementing the PVM measurements using the Choi-Kraus representation based on design of a unitary matrix and a partial trace operation in which the unitary matrix can be designed using quantum physical systems.

[7] Quantum control: When we design a quantum computer with quantum registers to store states of quantum systems, we need to measure the state without disturbing it as mentioned above. Therefore, we make non-demolition measurements on the evolving register states and use these measurements to construct the Belavkin filter [John Gough and Kostler, "Quantum filtering in coherent states") which gives us a good estimate of the state of the register. However, the resulting state estimate evolves according to the Belavkin filter equations which contains apart from the Hamiltonian component, Lindblad noise as well as measurement noise. By using Lec-Bouten's algorithm for quantum control based on applying infinitesimal unitary operators to the Belavkin evolved state (Lec Bouten, Filtering and control in quantum optics, Phd thesis), we can partially get rid of the Lindblad noise thereby resulting in almost the exact Hamiltonian dynamics of the quantum system whose state is stored in the quantum register. For however implementing this unitary control algorithm based on system observables and non-demolition measurements, we need to design the corresponding unitary gate at each control interval and this can be done in a fast way using unitary gates designed using real quantum physical systems.

Reference: Kumar Gautam and Harish Parthasarathy, "Project proposal for quantum gate design using physical systems with applications to computing, communication and control", Technical report, NSIT, 2018.

3.17 Convergence of perturbation series for nonlinear differential equations

3.17.1 Summary

When we take damped simple harmonic oscillations perturbed by a small nonlinearity, the equation of motion is given by

$$x''(t) + \gamma x'(t) + \omega^2 x(t) = \epsilon f(x(t))$$

This equation can be expressed in state variable form as a first order nonlinear differential equation for the vector valued process $\xi(t) = [x(t), x'(t)]^T$:

$$\xi'(t) = A\xi(t) + \epsilon g(\xi(t))$$

We attempt to solve this equation by perturbation theory:

$$\xi(t) = \xi_0(t) + \sum_{k \geq 1} \epsilon^k x_{ik}(t)$$

Substituting gives us after Taylor expanding the nonlinearity g around ξ_0 and equating equal powers of ϵ :

$$\begin{aligned} \xi'_0(t) &= A\xi_0(t), \\ \xi'_n(t) - A\xi_n(t) &= \sum_{k_1+2k_2+\dots+(n-1)k_{n-1}=n-1} \frac{g^{(k_1+\dots+k_{n-1}}(\xi_0(t))}{k_1! \dots k_{n-1}!} (\xi_1(t)^{\otimes k_1} \otimes \dots \otimes \xi_{n-1}(t)^{\otimes (k_{n-1})}) \end{aligned}$$

This equation is easily solved for $\xi_n(t)$ in terms of ξ_1, \dots, ξ_{n-1} . It is easy to see that if we assume an inequality of the form

$$\|\xi_k(t)\| \leq \alpha[k] \exp(\beta[k]t), \quad 0 \leq k \leq n-1, \quad t \in [0, T]$$

then $\beta[0]$ is simply the maximum of the real parts of the eigenvalues of A and by applying the triangle inequality to the solution for $\xi_n(t)$, we obtain an inequality of the form

$$\|\xi_n(t)\| \leq \alpha[n] \exp(\beta[n]t)$$

where

$$\beta[n] = \max(\beta[1] + \dots + \beta[n-1], \beta[0])$$

and

$$\alpha[n] = 2\max_{t \in [0, T]} \sum_{k_1+2k_2+\dots+(n-1)k_{n-1}=n-1} \frac{\|g^{(k_1+\dots+k_{n-1})}(\xi_0(t))\|}{k_1! \dots k_{n-1}!} \alpha[1]^{k_1} \dots \alpha[n-1]^{k_{n-1}} / (k_1\beta[1] + \dots + k_{n-1}\beta[n-1] - \beta[0])$$

If the denominator does not vanish (small denominator principle), then $\alpha[n]$ is a finite positive number. The perturbation series will then converge over $t \in [0, T]$ provided that ϵ is small enough so that

$$\sum_{n \geq 1} \alpha[n] \epsilon^n \exp(\beta[n]T) < \infty$$

In particular, if the limit

$$L = \limsup_{n \rightarrow \infty} \alpha[n]^{1/n} \exp(\beta[n]T/n) < \infty$$

then the series will converge for

$$\epsilon < 1/L$$

The convergence of perturbation series based on small denominators has been explored extensively in problems of celestial mechanics starting with Laplace and Poincare and culminating in the pioneering works of Kolmogorov, Arnold, Moser, Birkhoff and Carl Ludwig Siegel. Perturbation theory for differential equations is commonly used in quantum mechanics where its desired to find out the change in the energy levels of an atom and its eigenstates when its Hamiltonian gets perturbed by a small potential (time independent perturbation theory) and also in calculating the transition probability per unit of an electron in an atom making a transition between two stationary states when perturbed by an external electromagnetic field (time dependent perturbation theory). It is also used extensively in quantum field theory to calculate the scattering matrix for electrons, positrons and photons interacting with each other. The crucial idea here is that the perturbing Hamiltonian is expressible as the product of second quantized operator fields integrated out over space and to calculate the scattering matrix elements one must use the Dyson series for expanding the time ordered version of $\exp(-i \int_{-\infty}^{\infty} H(t) dt)$ where $H(t)$ is the sum of the unperturbed Hamiltonians of the electron-positron fields, the photon field and their interactions. Various terms in the perturbation series for the scattering matrix elements are usually calculated using the Feynman diagrams but the convergence of this

series is very hard to prove or disprove. When divergence is shown, then one adopts the Renormalization technique developed by Freeman Dyson involving scaling of the fields, charges and masses. But explaining all this would require another book and we just refer the reader to the masterpiece "Steven Weinberg, "The quantum theory of fields", Vol. I, Cambridge University Press.

3.17.2 Discussion

Consider first the scalar equation

$$x'(t) = ax(t) + \epsilon f(x(t))$$

where $f(x)$ has a Taylor series converging over the whole of \mathbb{R} , ie,

$$f(x) = \sum_{k \geq 0} f^{(k)}(0)x^k/k!$$

with

$$\sum_{k \geq 0} |f^{(k)}(0)||x|^k/k! < \infty, x \in \mathbb{R}$$

For example $f(x)$ can be a polynomial or more generally, a finite linear combination of exponentials multiplied by polynomials. Expanding the solution $x(t)$ as a perturbation series in powers of ϵ , ie,

$$x(t) = x(t, \epsilon) = x_0(t) + \sum_{n \geq 1} \epsilon^n x_n(t)$$

and substituting this series into the above differential equation gives after equating terms of equal powers of ϵ ,

$$x'_0(t) = ax_0(t),$$

and for $n \geq 1$,

$$x'_n(t) - ax_n(t) = \sum_{\substack{k_1+2k_2+\dots+(n-1)k_{n-1}=n-1, k_j \geq 0, j=1,2,\dots,n-1}} \frac{f^{(k_1+\dots+k_{n-1})(0)} x_1(t)^{k_1} \dots x_{n-1}(t)^{k_{n-1}}}{k_1! \dots k_{n-1}!}$$

This equation has the solution assuming that $x_n(0) = 0, n \geq 1$,

$$x_n(t) =$$

$$\int_0^t \exp(a(t-\tau)) d\tau \left[\sum_{\substack{k_1+2k_2+\dots+(n-1)k_{n-1}=n-1, k_j \geq 0, j=1,2,\dots,n-1}} \frac{f^{(k_1+\dots+k_{n-1})(0)} x_1(\tau)^{k_1} \dots x_{n-1}(\tau)^{k_{n-1}}}{k_1! \dots k_{n-1}!} \right] \dots \quad (1)$$

We assume that

$$|x_k(t)| \leq b[k].\exp(\beta[k]t), t \in \mathbb{R}, k = 0, 1, \dots, n - 1$$

and determine $\beta[n]$ in terms of $\beta[k], k \leq n - 1$ as follows. Substitute these exponential bounds into (1) to get

$$\begin{aligned} |x_n(t)| &\leq \exp(at) \sum_{k_1+2k_2+\dots+(n-1)k_{n-1}=n-1} \frac{|f^{(k_1+k_2+\dots+k_{n-1})}|}{k_1! \dots k_{n-1}!} \\ &\times b[1]^{k_1} \dots b[n-1]^{k_{n-1}} \int_0^t \exp((k_1\beta[1] + \dots + k_{n-1}\beta[n-1] - a)\tau) d\tau \\ &= \exp(at) \sum_{k_1+2k_2+\dots+(n-1)k_{n-1}=n-1} \frac{|f^{(k_1+k_2+\dots+k_{n-1})}|}{k_1! \dots k_{n-1}!} \\ &\times b[1]^{k_1} \dots b[n-1]^{k_{n-1}} (k_1\beta[1] + \dots + k_{n-1}\beta[n-1] - a)^{-1} \\ &\times [\exp((k_1\beta[1] + \dots + k_{n-1}\beta[n-1] - a)t) - 1] \end{aligned}$$

3.17.3 Points to remember

[1] Perturbation theory for solving nonlinear ode's and pde's or even linear ode's and pde's is based on attaching a small perturbation parameter to the nonlinear component or to the more difficult to analyze component of the differential equation (like for example, those involving non-constant coefficients) and then to expand the solution as a power series in the is parameter and equate different powers of this parameter yielding thereby a sequence of elementary differential equations for each approximant in the perturbation series in terms of the previous approximants, which are easily solved. This technique has been widely used in quantum mechanics by Schrodinger, Born, Pauli, Dirac and Jordan to obtain the shift in the energy levels and eigenstates of an atom when perturbed by a static potential or the approximate transition probability per unit time between two states of an atom when perturbed by time varying em radiation. Although for most problems we can write down such a perturbation series, it remains an open question that in general, how large can the pertrubation parameter be for the convergence of this series in some norm on an appropriate Banach space of functions. In this section, we have partially developed solution to this problem for the case of a linear time invariant system described by a linear vector differential equation perturbed by a small non-linearity.

3.18 Poiseulle's law and generalized Poiseulle's law for flow through a pipe

3.18.1 Discussion

Consider a pipe of cross section D having length L inside which an incompressible fluid of viscosity η flows. Assume that the length of the pipe is parallel to the z axis so that the cross-section D of the pipe is parallel to the xy plane. Neglect the effect of gravity. Let $v_z(x, y) \hat{z}$ denote the velocity field in the pipe. We may more generally assume that the boundary curve ∂D of the pipe's cross section is the curve $q_1 = c$ where c is a constant and (q_1, q_2) is a system of orthogonal coordinates in the xy plane. For example, $q_1 = r, q_2 = \phi$ where $r = \sqrt{x^2 + y^2}, \phi = \tan^{-1}(y/x)$. The two dimensional Laplacian can be expressed in the $q_1 - q_2$ system in terms of Lame's coefficients

$$H_k = \sqrt{(\partial x / \partial q_k)^2 + (\partial y / \partial q_k)^2}, k = 1, 2$$

Note that using the orthogonality of the $q_1 - q_2$ system, it is easily shown that

$$H_k = |\nabla q_k|^{-1}, k = 1, 2$$

The 2-D Laplacian is given by

$$\partial^2 / \partial x^2 + \partial^2 / \partial y^2 = (H_1 H_2)^{-1} ((\partial / \partial q_1)(H_2 / H_1) \partial / \partial q_1 + (\partial / \partial q_2)(H_1 / H_2) \partial / \partial q_2)$$

So the Navier-Stokes equation for steady flow in the pipe reduces to

$$\eta \nabla^2 v_z(q_1, q_2) - p_{,z} = 0$$

It follows that $p_{,z}$ must be independent of z . Thus, we write

$$-p_{,z} = f(q_1, q_2)$$

so that we if the pressure gradient $f(q_1, q_2)$ is known, we can solve for v_z the equation

$$(\psi(q_1, q_2) v_{z,q_1})_{,q_1} + (\psi(q_1, q_2)^{-1} v_{z,q_2})_{,q_2} + \chi(q_1, q_2) f(q_1, q_2) / \eta = 0$$

where

$$\psi(q_1, q_2) = H_2 / H_1, \chi(q_1, q_2) = H_1 H_2$$

This equation is to be solved with the boundary condition

$$v_z(c, q_2) = 0$$

ie the velocity vanishes on the boundary of the pipe. The volume flow per unit time through the pipe is then given in terms of the pressure gradient by

$$Q = \int_D v_z(q_1, q_2) H_1 H_2(q_1, q_2) dq_1 dq_2$$

and the mass flow per unit time is ρQ . We now make the following observation. Since $v_x = v_y = 0$ or equivalently, $v_{q_1} = v_{q_2} = 0$, it follows from the xy components of the Navier-Stokes equation that $p_{,x} = p_{,y} = 0$, ie, $p_{,z}$ cannot depend on x, y or z , ie the pressure gradient must be a constant. So, we write $-p_{,z} = f$ where $f = (p_A - p_B)/L$ where p_A and p_B are the pressures at the two ends. Thus, the above z-component of the Navier-Stokes equation simplifies to

$$(\psi(q_1, q_2)v_{z,q_1})_{,q_1} + (\psi(q_1, q_2)^{-1}v_{z,q_2})_{,q_2} + f\chi(q_1, q_2)/\eta = 0$$

We consider now the special case of a cylindrical pipe of radius R . Then, $q_1 = r, q_2 = \phi, c = R$. Thus, $H_1 = 1, H_2 = r$ and the the above z-component of the Navier-Stokes equation becomes

$$r^{-1}(rv_{z,r}(r, \phi))_{,r} + r^{-2}v_{z,\phi\phi}(r, \phi) + f/\eta = 0$$

$$v_z(R, \phi) = 0, 0 \leq \phi < 2\pi$$

It is not clear how to get the general solution to this equation. However, if we assume v_z to independent of ϕ , ie, $v = v_z(r)\hat{z}$, then this equation simplifies to

$$(rv'_z(r))' = -rf/\eta, v_z(R) = 0$$

and we get on integrating twice,

$$rv'_z(r) = -r^2f/2\eta + c_1, v_z(r) = -fr^2/4\eta + c_1\log(r) + c_2$$

Since $v_z(0)$ must be finite, we must have $c_1 = 0$ and further since $v_z(R) = 0$, we must have

$$c_2 = fR^2/4\eta$$

so that the solution is

$$v_z(r) = f(R^2 - r^2)/4\eta$$

The volume flow per unit time is then

$$\begin{aligned} Q &= \int_0^R 2\pi rv_z(r)dr = (\pi f/2\eta) \int_0^R r(R^2 - r^2)dr \\ &= (\pi f/2\eta)(R^4/4) = \pi f R^4 / 8\eta = \pi(p_A - p_B)R^4 / 8\eta L \end{aligned}$$

This is called Poiseulle's law.

3.18.2 Points to remember

[1] Within a pipe with a constant pressure gradient p' between the two pipe ends, there is effectively only the longitudinal component of fluid velocity which depends on the transverse coordinates. The Navier Stokes equation for such motion in steady state therefore becomes a linear pde for this velocity component of the form

$$\nabla_\perp^2 v_z(x, y) = -p'/\rho\eta$$

where the rhs is a constant. Solving this in plane polar coordinates in place of (x, y) for a pipe with a circular cross section with the boundary condition that v_z vanishes when $x^2 + y^2 = R^2$ where R is the pipe radius gives us Poiseulle's law.

[2] The same equation as above can be applied to a pipe of arbitrary cross section by expressing the transverse Laplacian ∇_{\perp}^2 in terms of the orthogonal curvilinear coordinates (q_1, q_2) with the assumption that $q_1 = \text{constt}$ defines the shape of the boundary.

[3] Time varying versions of Poiseulle's law can be derived by solving

$$v_{z,t}(t, x, y) = \nu \nabla_{\perp}^2 v_z(t, x, y) - p'(t, x, y)/\rho$$

with the appropriate boundary conditions.

3.19 Measurement of refractive index

3.19.1 Summary

There are many methods of estimating the refractive index of a medium even in the case when it varies from point to point. One of the earliest methods concerns measurement of the refractive index of a transparent fluid like water. We fill a bucket with the fluid and place a coin at its bottom. We choose a line of sight and determine the position of the coin as seen from above, ie, from air. Then we remove the coin and again measure the coin's position as seen from the same point. The angular difference between the two measurements then gives the refractive index of the fluid relative to air. Thus, let n_F denote the refractive index of the fluid and n_A that of air. $n = n_F/n_A$ is the refractive index of the fluid relative to air. When the fluid is present, a ray of light emanating from the coin and making an angle α with the unit normal to the fluid surface plane will get deflected in air so that this deflected/refracted ray makes an angle β with the same normal where $\sin(\alpha)/\sin(\beta) = n$. In the absence of the fluid, the ray will make the same angle α with the normal to the plane. Thus, knowing α, β , n may be determined.

In the case of an inhomogeneous medium, let $\epsilon(\omega, r), \mu(\omega, r)$ denote the permittivity and permeability at the point $r = (x, y, z)$ in the medium when the frequency of light is ω . The refractive index of the medium at this frequency and at the point r is then

$$n(\omega, r) = c_0 \sqrt{\epsilon \mu(\omega, r)}$$

where c_0 is the velocity of light in vacuum. We can measure it by taking measurements of the electromagnetic field at a discrete set of points and using interpolation. The basic idea here is to start with the Maxwell equations in frequency domain:

$$\operatorname{curl} E(\omega, r) = -j\omega \mu(\omega, r) H(\omega, r),$$

$$\operatorname{curl} H(\omega, r) = j\omega \epsilon(\omega, r) E(\omega, r),$$

$$\operatorname{div}(\epsilon E(\omega, r)) = 0, \operatorname{div}(\mu H(\omega, r)) = 0$$

From these, we derive

$$\operatorname{div} E = (-\nabla \log \epsilon, E), \operatorname{div} H = -(\nabla \log \mu, H)$$

and hence

$$\begin{aligned} -\nabla^2 E - \nabla(\nabla \log \epsilon, E) &= -j\omega(\nabla \mu \times H + \mu \nabla \times H) \\ &= -j\omega(\nabla \mu \times H + j\omega \mu \epsilon E) \end{aligned}$$

and

$$-\nabla^2 H - \nabla(\nabla \log \mu, H) = j\omega(\nabla \epsilon \times E - j\omega \epsilon \mu H)$$

which can be rearranged as

$$(\nabla^2 + \omega^2 \mu \epsilon)E + \nabla(\nabla \log \epsilon, E) - j\omega \nabla \mu \times H = 0 \quad \dots \quad (1)$$

and

$$(\nabla^2 + \omega^2 \mu \epsilon)H + \nabla(\nabla \log \mu, H) + j\omega \nabla \epsilon \times E = 0 \quad \dots \quad (2)$$

(1) and (2) constitute our fundamental wave propagation equations for an electromagnetic wave in a medium whose permittivity and permeability depend on both frequency and position. The idea is to expand the functions $\epsilon(\omega, r)$ and $\mu(\omega, r)$ as a linear combination of basis functions $\psi_n(r)$, $n = 1, 2, \dots, N$:

$$\begin{aligned} \epsilon(\omega, r) &= \epsilon_0(1 + \delta \sum_{n=1}^N c_n \psi_n(r)), \\ \mu(\omega, r) &= \mu_0(1 + \delta \sum_{n=1}^N d_n \psi_n(r)) \end{aligned}$$

substitute these into the above equations after approximating these equations upto $O(\delta)$ taking discrete measurements of the electric and magnetic fields at at least $2N$ spatial points and using the least squares method to estimate $\{c_n, d_n\}$. We can also use higher order perturbation theory by expanding $\log(\epsilon) = \log(\epsilon_0) + \log(1 + \delta \cdot \sum_{n=1}^N c_n \psi_n(r))$ and $\log \mu = \log(\mu_0) + \log(1 + \delta \cdot \sum_{n=1}^N d_n \psi_n(r))$ upto $O(\delta^K)$ and use the em field measurement data combined with the gradient search algorithm to estimate c_n, d_n

3.19.2 Discussion

If a light ray propagates along a path $s \rightarrow r(s)$ in a 3D-medium having refractive index $n(r) = n(x, y, z)$, then its velocity at the point r is $c_0/n(r)$ where c_0 is the velocity of light in vacuum. To compute the path of the light ray, we apply Fermat's principle of least time, by minimizing the total time taken for the light ray to propagate from a fixed point to another fixed point in the medium. This time taken is

$$T_{AB} = \int_{s_A}^{s_B} |r'(s)|n(r(s))ds/c_0, r(s_A) = r_A, r(s_B) = r_B$$

$$d/ds(\partial L/\partial r') = \partial L/\partial r$$

where

$$L(r, r') = |r'| n(r) = n(x, y, z) \sqrt{x'^2 + y'^2 + z'^2}$$

Thus, we get

$$(n(r)r''/|r'|)' = |r'|\nabla n(r)$$

Once we track the path of the light ray, we can estimate the refractive index of the medium by using the method of moments, namely, expand

$$n(r) = \sum_{m=1}^p a(m)\psi_m(r)$$

where $\psi_m(r)$ are basis functions and substitute this expression into the above trajectory differential equation to get

$$\sum_m a(m)(r''(s)/|r'(s)|)\psi_m(r(s))' = |r'(s)| \sum_m a(m)\nabla\psi_m(r(s))$$

From this equation, the coefficients $a(m)$'s are to be estimated. This may be done by defining the parameter s appropriately, and by measuring a large finite set of trajectory points $r(s_k), k = 1, 2, \dots, N$, we get linear equations for $a(m)$ which can be estimated by taking $a(0) = 1$ and using the standard least squares algorithm to estimate $a(m), 1 \leq m \leq p$.

3.19.3 Points to remember

[1] There are three methods to treat the measurement of refractive index.

[a] Pour the liquid into a bucket and place coin at the bottom of the bucket. Observe from above the direction of the coin's image one directly without the liquid and then with the liquid. From the angle between these two line of sights, by applying Snell's law of refraction, the refractive index can be calculated

[b] If the refractive index is a function of the spatial location and the direction at that location, then we derive an expression for the time taken by a light ray to travel from one point to another in this medium. This computation involves only noting that the speed of light at that spatial point along the tangential direction to the light ray curve is simply the speed of light in vacuum divided by the local refractive index. In this way, the total time taken for the ray can be expressed as an action integral where the corresponding Lagrangian is a function of $r(s)$ and $r'(s)$. By applying Fermat's principle of minimum time and variational calculus, we derive a differential equation for the light trajectory $s \rightarrow r(s)$ in term of the refractive index field. Taking measurements of the trajectory, we can in principle solve the inverse problem of estimating the refractive index field from the trajectory measurements for different starting points and different initial directions.

[c] In this method, we do not talk about light trajectories. Rather, we treat light as an electromagnetic wave field and assume in accordance with basic physical principles that the permittivity and permeability of the medium are functions of frequency and spatial position. We set up the Maxwell field equations in such a medium and by applying perturbation theory in which the permittivity and permeability are assumed to be small perturbations of constants, approximately evaluate the scattered em fields for given incident fields. These scattered fields will be linear functionals of the permittivity and permeability deviations and hence by measuring the scattered em field at a discrete set of spatial points, we can estimate the permittivity and permeability functions of frequency and position. In this method, we use the method of moments to expand the non-uniform permittivity and permeability as linear combinations of test functions with the coefficients being frequency dependent. These coefficients are estimated by a least squares method from scattered em field measurements at a fixed frequency. Thus, the refractive index as a function of frequency and spatial position may be estimated.

3.20 Modes of a vibrating string with applications to particle physics

3.20.1 Discussion

A closed curve Γ at time t is parametrized as $(t, s) \rightarrow X^\mu(t, s)$ where μ varies over D indices, ie, the dimension of space-time is D . s varies over $[0, 1]$ with $X^\mu(t, 0) = X^\mu(1, 1)$. The string is therefore closed. As this string moves in space, it covers a two dimensional manifold. In order to get analogy with the vibrating string, the action functional of this string is taken as

$$S(X(t, .)) = \int_{\mathbb{R}} \int_0^1 (1/2) g_{\mu\nu}(X(t, s))(X_{,t}^\mu(t, s)X_{,t}^\nu(t, s) - (1/2)X_{,s}^\mu(t, s)X_{,s}^\nu(t, s)) dt ds$$

where $g_{\mu\nu}(X)$ is the metric of the D -dimensional space. If there is an interaction of this string with a generalized electromagnetic field, then this interaction action is defined by

$$S_{int}(X) = -e \int B_{\mu\nu}(X(t, s)) dX^\mu(t, s) \wedge dX^\nu(t, s) dt ds$$

where $B_{\mu\nu}(X)$ is a skew symmetric tensor and is interpreted as a 2-tensor electromagnetic potential. This is analogous to the usual interaction $-e \int A_\mu(t, r(t)) dx^\mu(t)$ of a charged particle with an electromagnetic field described by a four potential A_μ . A charged particle traces out a world line while a string traces out a world sheet. So these two concepts are analogous. It should be noted that

$$dX^\mu(t, s) \wedge dX^\nu(t, s) =$$

$$(X_{,t}^\mu X_{,s}^\nu - X_{,s}^\mu X_{,t}^\nu)(t, s) dt ds$$

By adding the string action with its interaction action with the generalized em field, we can write down the Euler-Lagrange pde for the string $X^\mu(t, s)$.

Exercise: Set up the Euler-Lagrange equations for the above string interacting with a generalized em field.

The field equations for the generalized em field: We are assuming that $B_{\mu\nu}(X) = -B_{\nu\mu}(X)$. These are the generalized em potentials and then we define the generalized em field tensor as

$$F_{\mu\nu\sigma} = B_{\mu\nu,\sigma} + B_{\nu\sigma,\mu} + B_{\sigma\mu,\nu}$$

It is easily verified from the antisymmetry and the assumed tensorial property of $B_{\mu\nu}$ under diffeomorphisms that $F_{\mu\nu\sigma}(X)$ is a tensor. We raise and lower indices using the metric tensor inverse ($g^{\mu\nu}(X)$) and the metric tensor ($(g_{\mu\nu}(X))$ respectively. The quantity $F_{\mu\nu\sigma}F^{\mu\nu\sigma}$ is a scalar field and therefore, the action

$$S[B] = \int F_{\mu\nu\sigma}F^{\mu\nu\sigma}\sqrt{-g}d^dX$$

is an invariant. This may be taken as our action functional for the generalized em field. The total action for a set of M strings interacting with this generalized em field is then taken as

$$\begin{aligned} S(X_k, B_{\mu\nu}) = & \\ & \sum_{k=1}^M \int_{\mathbb{R}} \int_0^1 (1/2)g_{\mu\nu}(X_k(t, s))(X_{k,t}^\mu(t, s)X_{k,t}^\nu(t, s) - (1/2)X_{k,s}^\mu(t, s)X_{k,s}^\nu(t, s))dt ds \\ & + K_1 \int F_{\mu\nu\sigma}F^{\mu\nu\sigma}\sqrt{-g}d^dX \\ & + K_3 \sum_{k=1}^M e_k \int B_{\mu\nu}(X_k(t, s))(X_{k,t}^\mu X_{k,s}^\nu - X_{k,s}^\mu X_{k,t}^\nu)(t, s)dt ds \end{aligned}$$

Exercise: Write down the equations of motion of these M strings interacting with the generalized em field and simultaneously write down the field equations for the generalized em fields interacting with these M charged strings.

Generalization to n -dimensional strings. An n -dimensional string in D -dimensional space-time traces out an $n + 1$ -dimensional world sheet given by the mapping

$$(t, s_1, \dots, s_n) \rightarrow X^\mu(t, s_1, \dots, s_n)$$

Let $g_{\mu\nu}(X)$ denote the metric of the D -dimensional space-time. The action functional for the string is taken as

$$S_S[X] = \int g_{\mu\nu}(X(t, s_1, \dots, s_n))(X_{,t}^\mu X_{,t}^\nu - \sum_{k,j=1}^n a_{kj}(t, s_1, \dots, s_n)X_{,s_k}^\mu X_{,s_j}^\nu)dt ds_1 \dots ds_n$$

where $((a_{kj}(t, s_1, \dots, s_n)))$ is a $D \times D$ positive definite matrix valued function of (t, s_1, \dots, s_n) . For example, we can take $a_{kj} = \delta_{kj}$. Likewise, generalized em potential is described by the $(n+1)$ -indexed tensor

$$B_{\mu_1 \dots \mu_{n+1}}(X)$$

and it is assumed to satisfy the skew symmetric property:

$$B_{\mu_1, \dots, \mu_{n+1}}(X) = \text{sgn}(\sigma) B_{\mu_1, \dots, \mu_n}(X), \sigma \in S_{n+1}$$

The skew-symmetric generalized em field tensor is then defined by

$$F_{\mu_1 \dots \mu_{n+2}}(X) = \sum_{(\mu_1, \dots, \mu_{n+2})} B_{(\mu_1 \dots \mu_{n+1}, \mu_{n+2})}$$

where the sum is a cyclic sum, ie, over all the $n + 1$ cyclic permutations of the indices μ_1, \dots, μ_{n+2} . It is easy to prove using the tensorial and skew-symmetric property of B that F is also a tensor under diffeomorphisms of \mathbb{R}^D .

The action functional for the interaction between the generalized em field and the string is taken as the invariant

$$\begin{aligned} S_{SEM}(B) &= \int B_{\mu_1 \dots \mu_{n+1}}(X) dX^{\mu_1} \wedge \dots \wedge dX^{\mu_{n+1}} \\ &= \int B_{\mu_1 \dots \mu_{n+1}}(X) \sum_{\sigma \in S_{n+1}} \text{sgn}(\sigma) X_{,t}^{\mu_{\sigma 1}} X_{,s_1}^{\mu_{\sigma 2}} \dots X_{,s_n}^{\mu_{\sigma(n+1)}} dt ds_1 \dots ds_n \end{aligned}$$

3.20.2 Points to remember

[1] A moving string in d spatial and one time dimension is a closed curve at each time $t \geq 0$ defined by a map $s \rightarrow X(t, s)$ from $[0, 1]$ into \mathbb{R}^d . We can write down the Lagrangian for the string as the difference between its kinetic and potential energies defined as quadratic forms in the string velocity and position just as is done in non-relativistic Newtonian mechanics for vibrating string but with the difference that the metric is that of a curved space so that it is applicable to general relativity. The Euler-Lagrange equations then determine its classical motion, ie, a world sheet instead of a world line for particles.

[2] The interaction between charged string and a string em field, is the two dimensional generalization of the familiar action for interaction of an em field with a charged particle $-q \int A_\mu dx^\mu$. The generalization is $\int B_{\mu\nu} dX^\mu \wedge dX^\nu$. Here, $B_{\mu\nu}$ is the string em potential that is skew-symmetric and the antisymmetric field tensor $F_{\mu\nu} = A_{\nu,\mu} - A_{\mu,\nu}$ that represents the em field for charged particles is replaced for strings by

$$F_{\mu\nu\rho} = B_{\mu\nu,\rho} + B_{\nu\rho,\mu} + B_{\rho\mu,\nu}$$

The Lagrangian density of the string em field is the scalar $F_{\mu\nu\rho} F^{\mu\nu\rho}$. In this way a consistent theory of strings moving in a gravitational field and interacting with generalized em fields can be developed.

[3] String theory can be generalized to multimensional strings not just one dimensional by considering them as maps $(s_1, \dots, s_n) \in [0, 1]^n \rightarrow X(t, s_1, \dots, s_n) \in \mathbb{R}^d$.

[4] The energy-momentum tensor of a string can be computed from its Lagrangian and incorporated into Einstein's general theory of relativity as also can the energy-momentum tensor of the generalized em field.

[5] Quantization of the string can be achieved by first constructing the canonical momentum fields as $p_\mu(t, s) = \partial L / \partial X^\mu(t, s)$ and then applying the Legendre transformation to construct the Hamiltonian density and then Fourier series analyzing the string Lagrangian and Hamiltonian w.r.t. the spatial variable s . Quantization can be carried out in the Lagrangian formalism using Feynman's path integral or in the Hamiltonian formalism using the Schrodinger-Heisenberg operator wave-function formalism.

3.21 Hidden Markov Models for estimating the amplitude, frequency and phase of a sinusoidal signal making transitions

3.21.1 Summary

We consider here a wave field whose amplitude, frequency and phase can make sudden transitions at discrete times with given transition probabilities. The transition probability matrix here can be assumed to be known or at least a rough initial estimate of these transition probabilities are known. Over non-overlapping time slots, the amplitude, frequency and phase of the signal remain constant and when we make measurements of such a signal, measurement noise invariably enters into the system. Thus, we have a Hidden Markov Model (HMM) for the entire process. The problem is that we take measurements of the signal over all the time slots and from these observations we wish to make an estimate of the amplitude, frequency and phase over each time slot using the maximum a posteriori (MAP) method. For this, we evaluate the probability of making the given observation. The evaluation of this joint probability distribution involves knowledge of the transition probabilities from one slot to another of the amplitude, frequency and phase and also knowledge of the statistics of the measurement noise over each time slot. We may also seek to re-estimate the transition probabilities from the given measurements by assuming ergodicity of the measured process. Thus, if $q[n], n = 1, 2, \dots, N$ is the sequence of amplitude, frequency and phase triplets over N slots with transition probabilities

$$Pr(q[n+1] = b | q[n] = a) = A(b|a)$$

where a, b take values in a finite set, and the measured process over the n^{th} time slot is represented as

$$f(t) = f(t, q[n]) = f_0(t, q[n]) + w(t), 0 \leq t \leq T$$

then we estimate the state sequence $\{q[n]\}$ by maximizing

$$(\prod_{n=0}^{N-1} A(q[n+1]|q[n]).\prod_{t \in [0,T]} p_w(f(t, q[n]) - f_0(t, q[n])))$$

where p_w is the probability density of the noise sample $w(t)$. The above MAP maximization can also be carried out on a real time basis just as one does in non-linear filtering theory. This recursive estimation is known in the literature as the Viterbi algorithm. Re-estimation of the transition probabilities based on the ergodic hypothesis can be achieved using

$$\hat{A}[b|q] = \frac{\sum_{n=0}^{N-1} Pr(q[n+1] = b, q[n] = a | f(t), 0 \leq t \leq NT)}{\sum_{n=0}^{N-1} Pr(q[n] = a | f(t), 0 \leq t \leq NT)}$$

where these conditional probabilities are computed using Bayes rule and the assumed values of the transition probabilities $A[b|a]$.

3.21.2 Discussion

At time n , the amplitude, frequency and phase of the signal are respectively $(A[n], \omega[n], \phi[n])$. We assume that this vector takes values in a finite set $E = \{\xi[1], \dots, \xi[M]\}$ and the transition probability for this is

$$p((A[n+1], \omega[n+1], \phi[n+1]) = \xi[k] | (A[n], \omega[n], \phi[n]) = \xi[j]) = \\ a(j, k), j, k = 1, 2, \dots, M, a(j, k) \geq 0, \sum_{k=1}^M a(j, k) = 1$$

Given that the triplet of amplitude frequency and phase of the signal is $(A[n], \omega[n], \phi[n])$, the recorded signal has noise in it and is given by

$$O_n[k] = A[n].\cos(\omega[n]k + \phi[n]) + w_n[k], k = 1, 2, \dots, N$$

where $w_n[k]$ are iid random variables with pdf f_w . Thus, the observation sequence over the n^{th} time slot is given by

$$O_n = ((O_n[k]))_{k=1}^N \in \mathbb{R}^N$$

We take measurements of this signal over K time slots and from these measurements wish to estimate the amplitude, frequency and phases over each of these slots and also re-estimate the transition probabilities using the classic Viterbi algorithm. The MAP of the states $\{(A[n], \omega[n], \phi[n]) : n = 1, 2, \dots, K\}$ given the measurements $(O_n, n = 1, 2, \dots, K)$ is given by

$$\{\hat{\xi}[n] : 1 \leq n \leq K\} = Argmax_{\xi[n] \in E, n=1,2,\dots,K} \Pi_{n=1}^K \Pi_{k=1}^N f_w(O_n[k] \\ - A[n].\cos(\omega[n]k + \phi[n])). \Pi_{n=1}^{K-1} a(\xi[n], \xi[n+1])$$

where

$$(A[n], \omega[n], \phi[n]) = \xi[n], 1 \leq n \leq K$$

Simulational examples:

(a) Signal generation algorithm.

Assume that $w_n[k]$ is $N(0, \sigma^2)$. We first generate the observation sequence $O_n(k), k = 1, 2, \dots, N$ in each time slot $n = 1, 2, \dots, K$ as follows.

for $n = 1 : K$

for $j = 1 : M$

```

if
(A[n],ω[n],φ[n])' = ξ[:,j]
u=rand;
for m = 1 : M
if
    a(j,1) + ... + a(j,m-1) ≤ u < a(1,1) + ... + a(j,m)
    (A[n+1],ω[n+1],φ[n+1]) = (ξ[1,m],ξ[2,m],ξ[3,m])'
for k = 1 : N
    w(n+1,k) = σ² * randn;
    O(n+1,k) = A[n+1].cos(ω[n+1]k + φ[n+1]) + w(n+1,k)
end;

(b) Amplitude, frequency, phase sequence estimation algorithm.
for n₁ = 1 : M
for n₂ = 1 : M ...
for nₖ = 1 : M
for r = 1 : K
    (A[r],ω[r],φ[r]) = ξ[:,nᵣ]
end;
Evaluate

```

$C[n_1, \dots, n_K] = \prod_{r=1}^K \prod_{k=1}^N f_w(O[r, k] - A[r].\cos(\omega[r]k + \phi[r])).\prod_{j=1}^{K-1} a(n_j, n_{j+1})$

end;end;...end;

Choose that sequence (n_1, \dots, n_K) for which $C(n_1, \dots, n_K)$ is a maximum.

(Reference: Rabiner and Schaffer, "A tutorial on Hidden Markov Models", Proc.IEEE)

Hidden-Markov models for estimating image field parameters following a Markovian evolution law

References: [1] Sunil Joshi, H.Parthasarathy and T.K.Rawat, Technical report, NSIT, 2018.

[2] Rajneesh Vasisht, H.Parthasarathy and T.K.Rawat, Technical report, NSIT, 2018.

3.21.3 Exercises

[1] Consider one dimensional wave motion over the successive time intervals $[nT, (n+1)T], n = 0, 1, 2, \dots$ with respective wave velocities $c[n], n = 0, 1, 2, \dots$ so that over the time interval $[nT, (n+1)T]$, the wave amplitude $u(t, x)$ satisfies

$$u_{,tt}(t, x) - c[n]^2 u_{,xx}(t, x) = 0, 0 \leq x \leq L, u(t, 0) = u(t, L) = 0$$

where w is a white Gaussian noise field with zero mean and covariance

$$\mathbb{E}(w(t, x)w(t', x')) = \sigma^2 \delta(t - t')\delta(x - x')$$

[2] Compare the results of non-linear filtering theory with those of estimating the states of a Markov chain from observations whose statistics at any discrete time instant depend only on the state of the Markov chain at that time instant.

[3] A particle follows the geodesic equation on a Riemannian manifold with white noise forcing. By discretizing the equations of motion show that the position and velocity pair forms a Markov chain and explain how by taking noisy measurements of the radiation field emitted by the charged particle, one can estimate the particle's trajectory using MAP theory.

[4] the magnetic vector potential generated by a random current density field $J(t, r)$ satisfies

$$(\nabla^2 - c^{-2}\partial_t^2)A(t, r) = -\mu J(t, r)$$

Assuming that

$$J(t, r) = \sum_k B'_k(t)\psi_k(r)$$

where $B_k(t)$'s are independent Brownian motions and also expanding $A(t, r)$ in terms of these test functions $\psi_k(r)$ as

$$A(t, r) = \sum_k a_k(t)\psi_k(r)$$

derive a system of coupled stochastic differential equations satisfied by $(a_k(t), b_k(t) = a'_k(t))_{k \geq 1}$. Discretize this system thereby obtaining a Markov chain for $(a_k(n\delta), b_k(n\delta)) : k \geq 1, n = 0, 1, 2, \dots$. Assume now that noisy measurements on the electric field

$$E(t, r) = -A_{,t}(t, r)$$

are taken at discrete times. Cast the problem of estimating the states of the Markov chain, ie, the magnetic vector potential in terms of these noisy measurements and hence deduce how the current density at different times is to be estimated. You may have to discretize space into pixels in order to solve this problem.

[4] The wave function of a quantum mechanical particle evolves according to Schrodinger's equation with a random potential:

$$d|\psi(t)\rangle = [-(iH + V^2/2)dt - VdB(t)]|\psi(t)\rangle$$

Noisy measurement of the average value of an observable X are taken at discrete times t_1, t_2, \dots :

$$O[k] = \langle \psi(k\delta)|X|\psi(k\delta)\rangle + V[k], k = 1, 2, \dots$$

where the measurement noise samples $V[k]$ are white Gaussian noise samples independent of $B(\cdot)$. Using Hidden Markov theory formulate a method for estimating the wave function at the discrete times $k\delta, k = 0, 1, 2, \dots$

[5] Repeat the previous problem when the Schrodinger equation is replaced by the Dirac wave equation for a relativistic particle with noise:

$$d|\psi(t)\rangle = [-i[(\alpha, -idt\nabla + \sum_k L_k(r)dB_k(t)) + dt\beta m] - \sum_k (\alpha, L_k(r))^2 dt/2]|\psi(t)\rangle$$

Here,

$$L_k(r) = (L_{k1}(r), L_{k2}(r), L_{k3}(r))$$

are scalar multiplication operators and the measurement is

$$dZ(t) = \langle \psi(t) | X | \psi(t) \rangle dt + dV(t)$$

First apply the EKF to estimate the wave function in real time and then discretize both the Dirac equation and the measurement model and apply HMM theory to estimate the wave function at discrete times.

3.21.4 Points to remember

[1] Many physical processes in nature are based on HMM. The underlying system state undergoes transitions according to a Markov law and when the system is in a specific state q , a signal whose statistics depends on q only is measured. The goal is to estimate the state sequence from the measurement sequence. This problem with real-time estimates appears in the classical Kushner-Kallianpur filtering theory.

[2] The state sequence is estimated based on the observation sequence using the MAP method by applying Bayes' rule. The other important problem is the re-estimation of the state probabilities and the one-step transition probabilities of the underlying Markov chain based on the measured observation sequence. This is needed because in general, we do not have exact values of the initial probability of the Markov chain state and also of the one step transition probabilities. These can be estimated by starting with guess values and then assuming conditional ergodicity, ie we improve on our initial guess of the state probability by computing the time average of its conditional probability given the observation sequence divided by the joint probability of the observation sequence and likewise we improve on our initial guess of the one step Markov transition probability by computing the time average of the conditional joint probability of states at two successive times given the observation sequence and divide this by our estimate of the marginal state probability described above based on time averages. It is an open problem to prove that by repeating this iterative process, under what conditions on the initial guess for the marginal state probabilities and the one step transition probability will the sequence of these estimated probabilities converge.

3.22 The energy-momentum tensor of the Dirac field in a background curved space-time metric

3.22.1 Discussion

The general relativistic Dirac equation is

$$[\gamma^a V_a^\mu(x)(i\partial_\mu + eA_\mu(x) + i\Gamma_\mu(x)) - m]\psi(x) = 0$$

where $\Gamma_\mu(x)$, the spinor connection of the gravitational field is given by

$$\Gamma_\mu = (1/2)V_{a\nu}V_{b;\mu}^\nu J^{ab}, J^{ab} = (1/4)[\gamma^a, \gamma^b]$$

and $V_{a\mu}$ is the tetrad of the metric:

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + h_{\mu\nu}(x) = \eta_{ab}V_\mu^a(x)V_\nu^b(x)$$

We write

$$V_\mu^a(x) = \delta_\mu^a + \xi_\mu^a(x)$$

Then we get

$$\begin{aligned} h_{\mu\nu}(x) &= \eta_{ab}(\delta_\mu^a\delta_\nu^b + \delta_\mu^a\xi_\nu^b + \delta_\nu^b\xi_\mu^a \\ &\quad + \xi_\mu^a\xi_\nu^b) \end{aligned}$$

which implies that we exactly have

$$h_{\mu\nu} = \eta_{b\mu}\xi_\nu^b + \eta_{a\nu}\xi_\mu^a + \xi_{a\mu}\xi_{b\nu}$$

Note that the tetrad indices a, b, c etc. are raised and lowered using the Minkowski metric η_{ab} while the tensor indices $\mu, \nu, \rho, \sigma, \alpha, \beta$ etc. are raised and lowered using the metric tensor $g_{\mu\nu}(x)$ and its inverse $g^{\mu\nu}(x)$. The Lagrangian density for the Dirac field in the background gravitational and em field may be taken as

$$\begin{aligned} \mathcal{L}_D(\psi, \bar{\psi}, \psi_{,\mu}, \bar{\psi}_{,\mu}) &= \\ V_a^\mu(x)Re[\psi(x)^*\gamma^0\gamma^a(i\partial_\mu + i\Gamma_\mu + eA_\mu) - m\gamma^0]\psi(x)]\sqrt{-g(x)} \\ &= V_a^\mu(x)Re[\psi(x)^*\alpha^a(i\partial_\mu + i\Gamma_\mu + eA_\mu) - m\beta]\psi(x)]\sqrt{-g(x)} \end{aligned}$$

where

$$\alpha^a, a = 0, 1, 2, 3$$

are all 4×4 Hermitian matrices with

$$\alpha^0 = I$$

Here a bar denotes complex conjugate while a star denotes conjugate transpose. The usual definition of the energy-momentum tensor gives the following energy-momentum tensor of the Dirac field:

$$T_\mu^\nu = \mathcal{L}\delta_\mu^\nu - \psi_{,\mu}^T \frac{\partial \mathcal{L}}{\partial \psi_{,\nu}} - \psi_{,\mu}^* \frac{\partial \mathcal{L}}{\partial \bar{\psi}_{,\nu}}$$

We get as a column vector,

$$\partial \mathcal{L} / \partial \psi_{,\nu} = (i/2) V_a^\nu \alpha^{aT} \bar{\psi}(x)$$

$$\partial \mathcal{L} / \partial \bar{\psi}_{,\nu} = (-i/2) V_a^\nu \alpha_a \psi(x)$$

Thus,

$$T_\mu^\nu = \mathcal{L} \delta_\mu^\nu + \text{Im}(\psi_{,\mu}^*(x) \alpha^a \psi(x)) V_a^\nu(x)$$

Here, we are assuming that the tetrad $V_a^\mu(x)$ is a real field unlike the Penrose tetrad.

Using the same idea, we now compute the energy-momentum tensor of the electromagnetic field in a background curved space-time. The Lagrangian density is

$$\mathcal{L} = (-1/4) F_{\mu\nu} F^{\mu\nu}$$

where

$$F_{\mu\nu} = A_{\nu,\mu} - A_{\mu,\nu}$$

and

$$F^{\mu\nu} = g^{\mu\alpha} g^{\nu\beta} F_{\alpha\beta}$$

The energy-momentum tensor of the em field is then

$$T_\mu^\nu = \mathcal{L} \delta_\mu^\nu - (\partial \mathcal{L} / \partial A_{\rho,\nu}) A_{\rho,\mu}$$

Now,

$$\partial \mathcal{L} / \partial A_{\rho,\nu} = F^{\rho\nu}$$

and so

$$T_\mu^\nu = (-1/4) F_{\alpha\beta} F^{\alpha\beta} \delta_\mu^\nu - F^{\rho\nu} A_{\rho,\mu}$$

Now, in free space, ie, in the absence of charge and current densities in the concerned region, Maxwell's equations gives us

$$(F^{\rho\nu} \sqrt{-g})_{,\rho} = 0$$

and hence after neglecting a perfect four divergence following multiplication by $\sqrt{-g}$, T_μ^ν is the same as

$$\tilde{T}_\mu^\nu = (-1/4) F_{\alpha\beta} F^{\alpha\beta} \delta_\mu^\nu - F^{\rho\nu} (A_{\rho,\mu} - A_{\mu,\rho})$$

$$= T_\mu^\nu = (-1/4) F_{\alpha\beta} F^{\alpha\beta} \delta_\mu^\nu + F^{\rho\nu} F_{\rho\mu}$$

To take into account the presence of charges and currents in the concerned region of space-time, we add an interaction term $-J^\mu A_\mu$ to the Lagrangian density of the em field giving

$$\mathcal{L} = (-1/4) F_{\mu\nu} F^{\mu\nu} - J^\mu A_\mu$$

Then, the energy-momentum tensor of the em field is computed as before:

$$\partial \mathcal{L} / \partial A_{\rho,\nu} = F^{\rho\nu}$$

So

$$\begin{aligned} T_\mu^\nu &= \mathcal{L}\delta_\mu^\nu - \frac{\partial\mathcal{L}}{\partial A_{\rho,\nu}}A_{\rho,\mu} \\ &= ((-1/4)F_{\alpha\beta}F^{\alpha\beta} - J^\rho A_\rho)\delta_\mu^\nu - F^{\rho\nu}(A_{\rho,\mu} - A_{\mu,\rho}) \\ &\quad - F^{\rho\nu}A_{\mu,\rho} \end{aligned}$$

Now, after neglecting a four divergence, we get the following identity

$$\begin{aligned} -F^{\rho\nu}A_{\mu,\rho}\sqrt{-g} &= (F^{\nu\rho}A_\mu\sqrt{-g})_{,\rho} - A_\mu(F^{\nu\rho}\sqrt{-g})_{,\rho} \\ &= A_\mu J^\nu\sqrt{-g} \end{aligned}$$

So if we assume that our currents are fixed and known, then in this background current density field, the energy-momentum of the em field is

$$\begin{aligned} T_\mu^\nu &= (-1/4)F_{\alpha\beta}F^{\alpha\beta}\delta_\mu^\nu + F^{\nu\rho}F_{\mu\rho} \\ &\quad - J^\rho A_\rho\delta_\mu^\nu + J^\nu A_\mu \end{aligned}$$

This equation can be combined with the expression for the Dirac current in an electromagnetic field and a background gravitational field to obtain the total energy-momentum tensor of the electromagnetic field and the matter field consisting of electrons and positrons in a background gravitational field.

3.22.2 Exercises

- [1] Write down the energy-momentum tensor of the Dirac field in curved space-time using the Newman-Penrose tetrad formalism and specialize to the Kerr metric for rotating blackholes.

Hint: See the book S.Chandrasekhar, "The mathematical theory of black-holes", Oxford University Press.

- [2] Write down the energy-momentum tensor of the Dirac field in curved space-time interacting with a Yang-Mills non-Abelian gauge field. Take into account also the gauge field Lagrangian density.

3.22.3 Points to remember

- [1] There are two ways to derive the energy-momentum of a field given its Lagrangian density in a curved space-time background. The first is to take the variation of the corresponding invariant action w.r.t the metric field and the coefficient of $\sqrt{-g}\delta g_{\mu\nu}$ gives the energy-momentum tensor of the field in view of the Einstein field equations and the fact that the variation of the Einstein-Hilbert action w.r.t. the metric yields the Einstein tensor as the coefficient of $\sqrt{-g}\delta g_{\mu\nu}$. Therefore, if the Einstein-Hilbert action for gravitation is added to the matter action and the variation w.r.t. $\delta g_{\mu\nu}$ of the result is set to zero,

we get the Einstein field equations provided that we interpret the coefficient of $\sqrt{-g}\delta g_{\mu\nu}$ in the variation of the matter action as the energy-momentum tensor of the matter field. The same holds for the electromagnetic field and also for non-Abelian matter and gauge fields. However, it should be noted that the spinor connection of the gravitational field must be taken into account along with the electromagnetic $U(1)$ gauge field and/or the non-Abelian gauge field while writing down the matter action like the Dirac action or the Yang-Mills non-Abelian matter action interacting with the gauge field and the gravitational field. Further, in a curved space-time, the constant Dirac γ matrices must be replaced by their tetrad modulations in order to guarantee that the anticommutator of the tetrad modulated Dirac matrices have the metric tensor of curved space-time as their anticommutator and not the Minkowski metric of flat space-time. Equivalently, the tetrad modulations of the Dirac matrices along with the introduction of the spinor connection of the gravitational field guarantees not only global Lorentz invariance of the Dirac equation but also local Lorentz invariance, ie, invariance under Lorentz transformations that depend on the space-time coordinates.

[2] The second method of deriving the energy-momentum tensor is to differentiate the Lagrangian density w.r.t. the space-time coordinates assuming it not to depend explicitly on the space-time coordinates and then use the Euler-Lagrange equations of motion for the field to obtain the vanishing four divergence of a tensor defined in terms of the field Lagrangian and its partial derivatives w.r.t. the four gradients of the field. Alternately, this definition can be regarded as the manifestation of Noether's theorem which in this case states that if the Lagrangian density is invariant under space-time translations then the energy-momentum tensor is a conserved Noether current. However using this method does not generally lead to a symmetric energy-momentum tensor and additional terms must be added to make it symmetric without destroying the vanishing of the four divergence.

3.23 Remark on Noether's theorem on conserved currents

3.23.1 Discussion

Noether's theorem in the context of particles and fields with application to obtaining first integrals of the motion of particles and fields enabling a closer approach to integrating the equations of motion of a field. If $\mathcal{L}(x, \phi_k(x), \phi_{k,\mu}(x))$ is a Lagrangian density that is invariant under infinitesimal transformations of the field such as

$$\phi_k(x) \rightarrow \phi_k(x) + \epsilon \cdot \mathcal{F}_k(x, \phi(x))$$

where $\epsilon \rightarrow 0$, then we infer that

$$0 = (\partial \mathcal{L} / \partial \phi_k) \mathcal{F}_k(x, \phi) + (\partial \mathcal{L} / \partial \phi_{k,\nu}) \mathcal{F}_k(x, \phi)_{,\nu}$$

which implies on using the field equations

$$\partial_\mu(\partial\mathcal{L}/\partial\phi_{k,\mu}) = \partial\mathcal{L}/\partial\phi_k$$

that

$$\partial_\mu J^\mu = 0$$

where

$$J^\mu = (\partial\mathcal{L}/\partial\phi_{k,\mu})\mathcal{F}_k(x, \phi)$$

Specific cases of this conservation law include: (a) Conservation of energy and momentum when the Lagrangian density is invariant under space-time translations, ie,

$$\phi_k(x) \rightarrow \phi_k(x + \epsilon) = \phi_k(x) + \epsilon^\mu \phi_{k,\mu}$$

(b) Conservation of the angular momentum when the Lagrangian density is invariant under space-time rotations ie under Lorentz transformations:

$$\begin{aligned} \phi_k(x^\mu) &\rightarrow \phi_k(x^\mu + \epsilon_\nu^\mu x^\nu) \\ &= \phi_k(x) + \epsilon \omega_\nu^\mu \phi_{k,\mu}(x)x^\nu \end{aligned}$$

where $\omega_{\mu\nu} = -\omega_{\nu\mu}$ are constants, ie they determine an infinitesimal Lorentz transformation of space-time.

3.23.2 Exercises

[1] What are the conserved currents in the Yang-Mills non-Abelian gauge theory when explicit dependence on space and time coordinates is not there.

[2] If the Lagrangian density of a field is invariant under space-time rotations, then write down the conserved currents.

[3] If the Lagrangian $L(q, q')$ of a single particle where $q' = dq/dt$ is invariant under a one parameter group $\{g_t : t \in \mathbb{R}\}$ of diffeomorphisms of \mathbb{R}^3 , then write down a first integral of the motion in terms of the generator $A(q) = dg_t(q)/dt|_{t=0}$ of the group. Give examples illustrating this result from classical mechanics, namely cases corresponding to energy, momentum and angular momentum conservation.

3.23.3 Points to remember

[1] If the Lagrangian of a system of particles is invariant under a Lie group of transformations, then it is also invariant under infinitesimal transformations from the Lie group and using this combined with the Euler-Lagrange equations of motion, we can derive conserved quantities, ie, first integrals of the motion. If the dimension of the Lie group under which the Lagrangian is invariant is N , then we get N independent first integrals of the motion. In particular, if the Lagrangian does not contain time explicitly, then it is invariant under time translations which leads to energy conservation.

[2] More generally, if the Lagrangian density of a system of fields is invariant under a Lie group of field transformations, then it is also invariant under infinitesimal transformations of the Lie group which when combined with the Euler-Lagrange equations for the fields, leads to the conservation of Noether currents expressible in terms of the fields and the Lie group generators acting on the fields. A special case is when the Lagrangian density does not explicitly depend on the space-time coordinates. In this case, the Noether currents are the energy-momentum tensor components and we get conservation of the energy-momentum tensor.

3.24 Energy-momentum tensor using the tetrad formalism

3.24.1 Summary

A tetrad system defines a locally inertial frame for a metric. It is useful in many problems where the Einstein field equations or the Einstein-Maxwell equations in the standard coordinate system are very hard to solve and even simplify. A typical example is the Kerr metric corresponding to a rotating blackhole. The tetrad formalism replaces the space-time coordinates by one forms and one obtains the Riemann-Christoffel curvature tensor relative to such a tetrad basis in the form of scalars which are easily computed using Cartan's equations of structure. The basic reference for such computations is S.Chandrasehar, "The mathematical theory of blackholes", Oxford University Press, 1985. Further, using the tetrad formalism, the Dirac equation in a curved space-time is also easily formulated owing to the fact that the Dirac γ matrices must be replaced by their tetrad modulated versions in order to obtain the anticommutation relations

$$\gamma^\mu(x)\gamma^\nu(x) + \gamma^\nu(x)\gamma^\mu(x) = g^{\mu\nu}(x)$$

and further in deriving the form of the spinor connection for the gravitational field which is easily expressed in terms of a tetrad basis. The connection must ensure that under local Lorentz transformations, the Dirac equation remains invariant and such a connection is most easily constructed using tetrads. The best reference for this problem is contained in the book Steven Weinberg, "Gravitation and Cosmology", Principles and applications of the general theory of relativity", Wiley. In fact, once one constructs the spinor connection for the gravitational field, it is a straightforward matter to write down the total Lagrangian for the gravitational field, the matter field consisting of electrons and positrons interacting with the gravitational field and the electromagnetic field and the electromagnetic field interacting with the gravitational field. The Dirac current density interacting with the electromagnetic field automatically is contained in the Dirac electron-positron Lagrangian. The total Lagrangian density for these fields is of the form

$$\begin{aligned} L(V_\mu^a, V_{\mu,\nu}^a, A_\mu, \psi, \psi^*, \psi_{,\mu}, \psi_{,\mu}^*) &= K_1.g^{\mu\nu}\sqrt{-g}(\Gamma_{\mu\nu}^\alpha\Gamma_{\alpha\beta}^\beta - \Gamma_{\mu\beta}^\alpha\Gamma_{\nu\alpha}^\beta) \\ &+ K_2 Re(\psi^*\gamma^0[V_a^\mu\gamma^a(i\partial_\mu + eA_\mu + (i/2)J^{ab}V_b^\nu V_{b\nu;\mu}) - m]\psi)\sqrt{-g} \\ &\quad K_3 F_{\mu\nu}F^{\mu\nu}\sqrt{-g} \end{aligned}$$

where V_μ^a is the tetrad defined by

$$g_{\mu\nu} = \eta_{ab} V_\mu^a V_\nu^b$$

with η_{ab} denoting the Minkowski metric. Taking the variation of the corresponding action w.r.t $g_{\mu\nu}$ gives the Einstein field equations in the presence of electron-positron matter and em radiation. The coefficient of $\sqrt{-g}\delta g_{\mu\nu}$ appearing when the Dirac part is varied gives the energy-momentum tensor of the matter field while the coefficient of the same quantity appearing when the Maxwell electromagnetic part is varied gives the energy-momentum tensor of the radiation field. Another way to specify the energy-momentum tensor of the matter and radiation fields is via the Noether theory which states that if \mathcal{L} is a Lagrangian density of any set of fields $\phi_a, a = 1, 2, \dots, N$, and if further \mathcal{L} does not explicitly depend on the space-time coordinates, then this Lagrangian density is invariant under space-time translations and hence under the Euler-Lagrange equations, the tensor

$$T_\mu^\nu = \mathcal{L}\delta_\mu^\nu - (\partial\mathcal{L}/\partial\phi_{a,\nu})\phi_{a,\mu}$$

has a vanishing four divergence:

$$T_{\mu,\nu}^\nu = 0$$

which implies that this tensor should be regarded as the energy-momentum tensor of the set of fields $\phi_a, a = 1, 2, \dots, N$. However, such a procedure may not in general yield a symmetric energy-momentum tensor and hence we have to add terms to symmetrize it in such a way that the four divergence of the added terms also vanishes under the field equations of motion. Such a situation appears in the Maxwell theory. Yet another method of specifying the energy momentum tensor of the matter field, is to carry out the variation of the action w.r.t. the tetrad $V_\mu^a(x)$. The coefficient of $\sqrt{-g}\delta V_\mu^a$ (Note that $\sqrt{-g} = \det((V_\mu^a))$) then yields the energy momentum tensor after contracting one of its indices with the tetrad. The energy momentum tensor in such a situation thus has a scalar index and four vector index. This is clearly explained in the discussion.

3.24.2 Discussion

$$g_{\mu\nu} = \eta_{ab} V_\mu^a V_\nu^b$$

Let \mathcal{L} be the scalar Lagrangian density of a matter/radiation field. The corresponding action $S = \int \mathcal{L}\sqrt{-g}d^4x$ is a diffeomorphism invariant. From the Einstein-Hilbert action principle, it is known that $T^{\mu\nu} = \delta S/\sqrt{-g}\delta g_{\mu\nu}$ is the energy-momentum tensor of the field apart from a constant multiplier. Now,

$$\sqrt{-g} = V = \det((V_\mu^a))$$

Also,

$$\delta g_{\mu\nu} = \eta_{ab}(V_\mu^a \delta V_\nu^b + V_\nu^b \delta V_\mu^a)$$

So, we can write

$$\begin{aligned}\delta_g S &= T^{\mu\nu} \sqrt{-g} \delta g_{\mu\nu} \\ &= T^{\mu\nu} V \eta_{ab} (V_\mu^a \delta V_\nu^b + V_\nu^b \delta V_\mu^a)\end{aligned}$$

and assuming $T^{\mu\nu}$ to be symmetric, this equals

$$\begin{aligned}T^{\mu\nu} V (V_\mu^a \delta V_{a\nu} + V_\nu^a \delta V_{a\mu}) \\ = 2T^{\mu\nu} V \cdot V_\mu^a \delta V_{a\nu}\end{aligned}$$

So we get,

$$\delta_g S / \delta V_{a\nu} = 2T^{\mu\nu} \cdot V \cdot V_\mu^a$$

or equivalently,

$$(1/2) V_\nu^b \delta_g S / V \cdot \delta V_{a\nu} = T^{\mu\nu} V_\mu^a V_\nu^b = T^{ab}$$

this is the energy-momentum scalar, ie, energy momentum tensor relative to the tetrad. All its ten components are scalars under diffeomorphism. We can use this formula to compute the energy-momentum tensor of the Dirac field relative to the tetrad. The Lagrangian density of the Dirac field in curved space-time is

$$\mathcal{L} = Re[V_a^\mu(x) \psi^*(x) \alpha^a (i\partial_\mu + i\Gamma_\mu) \psi(x) - m\psi^*(x) \beta \psi(x)] \sqrt{-g}$$

where

$$\Gamma_\mu = (1/2) J^{ab} V_a^\nu V_{b\nu;\mu}, J^{ab} = (1/4)[\gamma^a, \gamma^b]$$

We have

$$g_{\mu\nu} = \eta^{ab} V_{a\mu} V_{b\nu}$$

so that

$$\begin{aligned}\delta g_{\mu\nu} &= \eta^{ab} (V_{a\mu} \delta V_{b\nu} + V_{b\nu} \delta V_{a\mu}) \\ &= V_\mu^a \delta V_{a\nu} + V_\nu^a \delta V_{a\mu} \\ V_a^\mu V_\mu^b &= \delta_a^b\end{aligned}$$

3.24.3 Exercises

[1] Solve using perturbation theory, the Yang-Mills matter and gauge field equations when the Lagrangian density is given by

$$\begin{aligned}\mathcal{L} &= \psi(x)^* [(\gamma^0 \gamma^\mu \otimes I_N)(iI_{4N} \partial_\mu + eA_\mu^\alpha(x) I_4 \otimes \tau_\alpha) - m\gamma^0 \otimes I_N] \psi(x) \\ &\quad + K_0 \cdot F_{\mu\nu}^\alpha F^{\mu\nu\alpha}\end{aligned}$$

where

$$\begin{aligned}\nabla_\mu &= \partial_\mu + ieA_\mu^\alpha \tau_\alpha, \\ ieF_{\mu\nu} &= ieF_{\mu\nu}^\alpha \tau_\alpha = [\nabla_\mu, \nabla_\nu]\end{aligned}$$

where τ_α are Hermitian generators of the gauge group $U(N)$ with structure constants $C(\alpha\beta\rho)$ defined by

$$[\tau_\alpha, \tau_\beta] = iC(\alpha\beta\rho)\tau_\rho$$

[2] Write down the energy-momentum tensor of the Yang-Mills $U(N)$ Gauge fields and matter fields without taking curvature of space-time into account.

[3] Write down the energy-momentum tensor of the Yang-Mills $U(N)$ Gauge fields and matter fields by taking the background curvature into account.

3.24.4 Points to remember

[1] The metric of space-time can be expressed in tetrad form which transforms the coordinate system into a locally inertial frame. The action integral for the gravitational field and matter fields can be expressed using the tetrad basis in place of the original metric. This involves introducing the gravitational connection for the concerned matter field, for example the gravitational spinor connection for the Dirac field. The connection is defined in terms of the tetrad and replaces the usual four gradient by the covariant gradient taking the gravitational connection into account. This guarantees that under local Lorentz transformations, the Dirac equation in curved space-time will remain invariant. The variation of the action functional w.r.t the metric tensor is then replaced by its variation w.r.t the tetrad basis and we can derive the tetrad components of the energy-momentum tensor from that. The matter field can be either the Dirac field or more generally, a coupling of the Yang-Mills and Dirac matter fields coupled to the electromagnetic $U(1)$ gauge field and the non-Abelian gauge fields. The energy momentum tensor of all the matter and gauge fields can be determined by summing up the Lagrangian density of the matter field taking into account the gravitational connection apart from the electromagnetic four potential $U(1)$ connection and the non-Abelian Yang-Mills connection with the Lagrangian densities of the gauge fields. The variation of this matter-gauge Lagrangian density w.r.t the metric gives the energy-momentum tensor of the matter and gauge fields. Equivalently, the variation of the same w.r.t. the tetrad basis gives the energy-momentum tetrad components. Einstein's field equations in tetrad form are sometimes more convenient to handle than in four vector form. See for example the masterpiece : S.Chandrasekhar, "The mathematical theory of blackholes", Oxford University Press.

3.25 Analysis of gravitational waves produced by a finite system of point particles—A perturbation theoretic approach

3.25.1 Summary

Here, the energy-momentum tensor has the form

$$T^{\mu\nu} = c^2 \sum_k (-g(x_k))^{-1/2} m_k \delta^3(x - x_k) (dx_k^\mu/d\tau) (dx_k^\nu/dt)$$

It is easy to see that

$$\int T^{\mu\nu} \sqrt{-g} d^4x = c^2 \sum_k m_k \int (dx_k^\mu/d\tau) dx_k^\nu$$

is a tensor and since $\sqrt{-g}d^4x$ is an invariant, it follows that $T^{\mu\nu}$ is a tensor. By considering the special relativistic approximation, it follows that $T^{\mu\nu}$ is the energy-momentum tensor for the finite system of point particles having rest masses $\{m_k\}$. More generally, if S is any scalar field, then

$$\int T^{\mu\nu} S \sqrt{-g} d^4x = c^2 \sum_k m_k \int S(x_k) (dx_k^\mu/d\tau) dx_k^\nu$$

is an invariant from which the tensorial property of $T^{\mu\nu}$ is inferred. In analogy with the Schwarzschild metric, in order to determine the gravitational field at large distances from these point particles, we assume the metric to have the form

$$d\tau^2 = (1 + 2\phi/c^2 + h_{00}/c^4) dt^2 - (1/c^2)((1 - 2\phi/c^2) - h_{ij}/c^4) dx^i dx^j - 2h_{0j} dx^0 dx^j/c^3$$

so that

$$\begin{aligned} g_{00} &= 1 + 2\phi/c^2 + h_{00}/c^4, \\ g_{ij} &= -((1 - 2\phi/c^2)\delta_{ij}/c^4 + h_{ij}/c^6), \\ g_{0j} &= -h_{0j}/c^3 \end{aligned}$$

and substitute these expressions into the Ricci tensor $R_{\mu\nu}$, as well as into $T_{\mu\nu}$ retaining terms only upto $O(1/c^4)$ and applying appropriate coordinate conditions, set up the Einstein field equations upto $O(1/c^4)$:

$$R_{\mu\nu} = (-8\pi G/c^4)(T_{\mu\nu} - Tg_{\mu\nu}/2)$$

and hence determine the corrections to gravitational waves upto $O(1/c^4)$. Moreover, the approximate solutions for the metric tensor in terms of the positions and four velocities of the system of particles can be used in the geodesic equations for the particle motion to obtain the effect of mutual gravity on the particle motion. Then using this corrected four velocities, we can get corrections to the energy-momentum tensor of matter to obtain corrections to the metric and the process can be repeated. This is in fact the general relativistic version of the Newtonian gravitational many body problem.

3.25.2 Exercises

[1] Solve the Einstein field equations $R_{\mu\nu} = 0$ in vacuum upto second degree terms showing explicitly the frequency coupling in the linearized plane gravitational waves.

[2] Calculate the total energy per unit time radiated in the far field zone due to gravitational radiation coming from matter near the origin whose energy-momentum tensor has the form

$$T_{\mu\nu}(t, r) = \text{Re}(\tilde{T}_{\mu\nu}(r)\exp(j\omega t))$$

For this, you must first assume the metric to have the form

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}(t, r)$$

and evaluate the Einstein tensor

$$G^{\mu\nu} = R^{\mu\nu} - Rg^{\mu\nu}/2$$

upto quadratic orders in the metric perturbations $h_{\mu\nu}$ and its partial derivatives. Then you must observe that the linear component in $G^{\mu\nu}$ has a vanishing ordinary four divergence and hence if $G^{(2)\mu\nu}$ component, we get by neglecting the cubic and higher components that

$$[T^{\mu\nu} + (8\pi G)^{-1}G^{(2)\mu\nu}],_{\nu} = 0$$

and hence $(8\pi G)^{-1}G^{(2)\mu\nu}$ must be interpreted as the energy-momentum pseudo-tensor of the gravitational field. Thus, the rate of flow of gravitational energy due to radiation per unit area must be given by $(8\pi G)^{-1}G^{(2)r0}$, $r = 1, 2, 3$ evaluate the surface integral of this quantity over a sphere of large radius and hence determine the rate at which the matter emits gravitational energy (Reference: Steven Weinberg, "Gravitation and Cosmology; Principles and applications of the general theory of relativity", Wiley). You must calculate $G^{(2)\mu\nu}$ using $h_{\mu\nu}$ obtained as the retarded potential solution to the linearized Einstein field equations driven by the matter field.

[3] Write down the approximate general relativistic Lagrangian for a system of N point particles of rest masses m_1, \dots, m_N moving under mutual gravitational interaction. Your answer must be expressed as a sum of the Newtonian Lagrangian with corrections from special and general relativity (Reference: Landau and Lifshitz, "The classical theory of fields").

3.25.3 Points to remember

[1] The gravitational many problem is a well known problem in celestial mechanics that admits closed form solutions only for two masses and certain isolated solutions like Lagrange's triangle for three bodies. For more bodies, we do not know about any closed form solution except for the total energy and total angular momentum integral.

[2] In general relativity, we can also talk about the gravitational many body problem. We first write down the energy-momentum tensor field for N point masses as a function of their four velocities and their positions. We have then to solve the Einstein field equations given this energy-momentum tensor. But the four velocities of the N particles are governed by geodesics defined by the solution metric and hence we are not able to obtain any closed form solution. The only way out of this difficulty is to obtain approximate solutions by using perturbation theory. As an initial approximation, we assume a Schwarzschild like metric with a gravitational potential $\phi(r)/c^2$ appearing in the metric plus terms of order $1/c^3$ and $1/c^4$. We substitute these expressions for the metric tensor components into the energy-momentum tensor and write down the $O(1/c^2)$, $O(1/c^3)$ and $O(1/c^4)$ perturbation theoretic terms in the Einstein field equations. We then accordingly solve for the metric in terms of the positions and four velocities of the particles and substitute this metric into the geodesic equations to get corrections to the motion. This approximation is based on the fact that at large distances from the system of N particles, the matter distribution will appear spherical and so a Schwarzschild-like solution can be taken as the starting metric.

3.26 Heat equation and its solution in R^n , relationship between heat and wave equations, nonlinear heat equations arising as the scaling limit of the simple exclusion process

3.26.1 Summary

It is well known that the Maxwell equations in a conducting medium lead to damped wave equations, ie, wave equations of the form

$$\nabla^2\psi(t, r) - (1/c^2)\partial_t^2\psi(t, r) - \gamma\partial_t\psi(t, r) = 0$$

Where γ is proportional to the conductivity. In the limit of very high conductivity, this equation approximates to the heat equation

$$\partial_t\psi(t, r) = \gamma^{-1}\nabla^2\psi(t, r)$$

whose solution was first written by Einstein as the convolution of a Gaussian density with a variance proportional to time with the initial value of the wave, ie, at time $t = 0$. This equation was written when Einstein worked out the probability law of a pollen particle in a warm liquid executing Brownian motion owing to random hits by the surrounding molecules. Subsequently the same heat equation was discovered for the temperature distribution in a metal obeying linear heat flow, ie, the heat flux is proportional to the negative temperature gradient. Today in mathematical physics, especially in statistical mechanics, there has emerged a field called interacting particle systems and more specifically, the simple exclusion process in which particles are located at the sites of a lattice and make transitions from occupied to vacant sites with the transition rate and

location controlled by Poisson processes, ie, a particle waits for an exponential time and then chooses a site according to a probability distribution provided that the site is empty and then makes a transition. Mathematicians are interested in what happens when the lattice becomes a continuous medium like a d -dimensional torus owing to the spacing of the sites getting smaller and smaller and simultaneously time is also scaled according to number dependent on the lattice spacing. Eminent mathematicians like Varadhan, Yau, Olla, Kipnis, Guo, Papanicolaou, Rezhakanlou and Quastel have contributed to this problem. In simple terms, Varadhan has explained how such scaling limits lead to a heat equation for the limiting density of the particles with diffusion coefficient depending nonlinearly on the density, ie, generalizations of the Burgers equation. It is therefore natural to ask whether second order time derivative terms can also arise in such problems when external forces act on the particles. Such equations obviously cannot occur for Markov processes like the simple exclusion process but if we have a bivariate process that is Markov, then we may as in the case of the simple harmonic oscillator be able to derive differential equations that are second order in time and second order in space with density dependent coefficients. Keeping this objective in mind, we have presented a section on scaling limits highlighting some of the major results of Varadhan et.al.

3.26.2 Discussion

Nonlinear waves in the scaling limit of the simple exclusion process. We have a lattice $X = \mathbb{Z}_N^d$ and if a point x in this lattice is occupied at time t , we put $\eta_t(x) = 1$ and $\eta_t(x) = 0$ otherwise. For each pair $(x, y) \in X \times X$, we have an independent Poisson process $N_t(x, y)$ with a rate of $\lambda p(x, y)$ and transitions of particles from occupied sites to vacant sites take place according to these Poisson clocks with a probability distribution of $p(x, y)$ for a particle to jump from x to y given that it jumps. Thus, $\sum_{y \in X} p(x, y) = 1$. We then obtain the following sde for the exclusion process:

$$d\eta_t(x) = \sum_{y \neq x} [\eta_t(y)(1 - \eta_t(x))dN_t(y, x) - \eta_t(x)(1 - \eta_t(y))dN_t(x, y)]$$

Scaling limits for such processes have been derived in the literature (See "Collected papers of S.R.S.Varadhan", Vol.4, Particle systems and their large deviations, Hindustan book agency). A different kind of model would be the following. Let $v_t(x)$ denote any physical quantity associated with a particle at the point x at time t . Then $v_t(x)$ will increase in the time interval $[t, t + dt]$ if a transition takes place from some $y \neq x$ to x while it will decrease if a transition takes place from x to some $y \neq x$ in this infinitesimal time interval. For example, we could write

$$dv_t(x) = \sum_{y \neq x} [\eta_t(y)(1 - \eta_t(x))dN_t(y, x) - \eta_t(x)(1 - \eta_t(y))dN_t(x, y)] \quad \dots \quad (1)$$

Again, the site x will have an increase in the the number of particles from zero to one in the time interval $[t, t + dt]$ proportional to the physical quantity $v_t(x)$. For example, $v_t(x)$ may be the particle current at x . Thus, we have

$$d\eta_t(x) = v_t(x)dt \quad (2)$$

Remark: The particle current from our physical intuition from site y to site x should be proportional to $\eta_t(y) - \eta_t(x)$, so ideally speaking, we should replace (2) by

$$d\eta_t(x) = \sum_{y \neq x} a(x, y)v_t(y)dt \quad (3)$$

where $a(x, y), y, x \in X$ are real numbers. We now look at the mean and variance propagation equations in (1) and (2). First, suppose we approximate (1) by taking expectations on both sides and replacing $\mathbb{E}f(\eta_t)$ by $f(\mathbb{E}\eta_t)$ for any nonlinear function $f : X \rightarrow \mathbb{R}$. Let

$$\mathbb{E}\eta_t(x) = \rho_t(x), \eta_t(x) - \rho_t(x) = \delta\eta_t(x),$$

$$\mathbb{E}(v_t(x)) = V_t(x), v_t(x) - V_t(x) = \delta v_t(x)$$

Then the above equations become approximately,

$$\begin{aligned} dV_t(x)/dt &= \lambda \sum_{y \neq x} \rho_t(y)(1 - \rho_t(x))p(y, x) - \rho_t(x)(1 - \rho_t(y))p(x, y)] \\ d\rho_t(x)/dt &= \sum_{y \neq x} a(x, y)V_t(y) \end{aligned}$$

Consider a special case of these equations when

$$p(y, x) = p(y - x) = p(x - y)$$

We then get

$$\begin{aligned} dV_t(x)/dt &= \lambda \sum_{y \neq x} p(y - x)(\rho_t(y) - \rho_t(x)) \\ &= \lambda \sum_z (p(z)(\rho_t(x + z) - \rho_t(x))) \\ d\rho_t(x)/dt &= \sum_z a(z)V_t(x + z) \end{aligned}$$

where we further assume that

$$a(y, x) = a(y - x)$$

Now replacing x by $\theta = x/N \in [0, 1]^d$ and Taylor expanding these equations followed by truncation, we get that the above equations approximate to

$$dV_t(\theta)/dt = \lambda \sum_z p(z)((z/N, \nabla \rho_t(\theta)) + (1/2)(z/N, \nabla \nabla^T \rho_t(\theta), z/N))$$

which in the limit of $N \rightarrow \infty$ converges assuming that

$$\sum_z p(z)z/N \rightarrow b, \sum_z zz^T p(z)/N^2 \rightarrow C$$

to

$$dV_t(\theta)/dt = (b, \nabla) \rho_t(\theta) + (1/2)Tr(C \nabla \nabla^T \rho_t(\theta)),$$

$$d\rho_t(\theta)/dt = a_0 V_t(\theta) + (a_1, \nabla) V_t(\theta)$$

where

$$a_0 = \sum_z a(z), a_1 = \sum_z z.a(z)/N$$

Here, we are making the approximation

$$V_t(x+z) \approx V_t(\theta) + (z/N, \nabla)V_t(\theta)$$

Thus, we get after eliminating $V_t(\theta)$ that

$$\begin{aligned} d^2\rho_t(\theta)/dt^2 &= a_0 dV_t(\theta)/dt + (a_1, \nabla)dV_t(\theta)/dt \\ &= a_0[(b, \nabla)\rho_t(\theta) + (1/2)Tr(C\nabla\nabla^T\rho_t(\theta))] \\ &\quad + (a_1, \nabla)[(b, \nabla)\rho_t(\theta) + (1/2)Tr(C\nabla\nabla^T\rho_t(\theta))] \end{aligned}$$

Retaining only upto second order partial derivatives w.r.t θ , we get

$$\begin{aligned} d^2\rho_t(\theta)/dt^2 &= a_0(b, \nabla)\rho_t(\theta) + (a_0/2)Tr(C\nabla\nabla^T\rho_t(\theta)) \\ &\quad + (a_1, \nabla)(b, \nabla)\rho_t(\theta) \end{aligned}$$

This is a generalized damped wave equation for $\rho_t(\theta)$ and is easily solved using d -dimensional Fourier series on the torus.

3.26.3 Exercises

[1] Carry out the Taylor expansion upto infinite degree and write down the resulting pde for this system as $N \rightarrow \infty$ after making appropriate assumptions on the behaviour of the moments of $p(.)$. For this, you must replace $a(z)$ by $a_N(z)$ and assume appropriate behaviour of $a_N(z)$ as $N \rightarrow \infty$.

Fluctuations: We rewrite (1) as

$$\begin{aligned} dv_t(x) &= \sum_{y \neq x} [\eta_t(y)(1 - \eta_t(x))dM_t(y, x) - \eta_t(x)(1 - \eta_t(y))dM_t(x, y)] \\ &\quad + \lambda \sum_z [p(z)(\eta_t(x+z) - \eta_t(x))dt] \end{aligned}$$

where $M_t(x, y)$ is the Martingale

$$M_t(x, y) = N_t(x, y) - \lambda p(x, y)t$$

Regarding the Martingale M_t as zero mean noise that is of small amplitude, we get approximately

$$\begin{aligned} d\delta v_t(x) &= \sum_z [p(z)(\delta\eta_t(x+z) - \delta\eta_t(x))]dt \\ &\quad + \sum_z [\rho_t(x+z)(1 - \rho_t(x))dM_t(x+z, x) - \rho_t(x)(1 - \rho_t(x+z))dM_t(x, x+z)] \end{aligned}$$

which when combined with

$$d\delta\eta_t(x) = \sum_z a(z)\delta v_t(x+z)$$

can be easily used to calculate the evolution of the moments

$$\mathbb{E}[\delta\eta_t(x)\delta\eta_t(y)], \mathbb{E}[\delta v_t(x)\delta v_t(y)], \mathbb{E}[\delta\eta_t(x)\delta v_t(y)]$$

using Ito's formula for Poisson processes.

[2] Formulate the approximate mean and covariance propagation equations for the stochastic differential equation

$$dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dB(t) + \int_{u \in A} f(t, X(t), u)(dN(t, du) - \lambda(t, u)dudt)$$

where the noise terms are small. Here, $N(t, du)$ is a Poisson random field with rate

$$\mathbb{E}[N(t, du)] = \Lambda(t, u)du, \Lambda(t, u) = \int_0^t \lambda(s, u)ds$$

3.26.4 Points to remember

[1] The simple exclusion process is a Markov process model for modeling the motion of particles in a finite lattice. A lattice site may be occupied by a particle or not occupied. If one site is occupied and another site is unoccupied, then there is a small probability of the particle from the occupied site to jump to the unoccupied site. The clock that controls such a jump is a Poisson process whose rate depends on the positions of the two sites. The entire state of the system at any time t is specified completely if we know at that time which sites are occupied and which are unoccupied. The Markov process generator for this process can be easily written down.

[2] Scaling limits talk about how the empirical density of the lattice scaled to fall within the torus behaves as the size of the lattice becomes infinite and simultaneously time is also scaled accordingly. It has been shown in the literature that the empirical density then converges to a deterministic density that satisfies nonlinear versions of the heat equation and the rate at which this empirical density converges to the nonrandom density which is a function of time and position on the multidimensional torus can be evaluated using a large deviation principle, ie, the rate function for the sequence of empirical densities as the lattice becomes finer and finer can be obtained. Methods based on ideas like relative entropy between two probability distributions and its computation using Girsanov's theorem for evaluating the Radon-Nikodym derivative between two probability measures that are required for the relative entropy computation.

[3] By a slight modification of the simple exclusion process with scaling limits discussed above, we can introduce positions and velocities of interacting particles and can derive instead of nonlinear versions of heat equation, nonlinear versions of the wave equation in the scaling limit.

3.27 Study of wave motion of the boundary of single cellular micro-organisms by giving them external stimulus and observing the wave like motion of their boundary walls as well as wave-like fluctuations of the velocity field of the cytoplasmic fluid within them

3.27.1 Summary

The study of small velocity and density perturbations in the velocity field of the cytoplasm inside a living cell when a small perturbing force is applied to its boundary is a difficult problem in wave motion. It is well known that the velocity and density perturbations can be studied using linearization of the Navier-Stokes and mass conservation equation. However, when the boundary perturbations are also considered, then we have to make a time varying diffeomorphism of the coordinate system so that in the transformed coordinates, the boundary will appear to be non-changing with time. Such a time varying diffeomorphism is also one of the unknown functions like the perturbed velocity and density field but only after applying the boundary condition that the velocity and density vanish on the boundary can a complete set of linearized equations be developed. This idea is originally due to T.Kato, who proposed variational principles based on "form theory" for formulating partial differential equations. The action integral is taken over a complicated boundary but by making an appropriate diffeomorphic change of the coordinate system, this action can be reduced to one involving a simpler boundary.

3.27.2 Discussion

When singular cellular organisms like Amoeba and Paramecium move by changing the shape of their cell boundary, we wish to classify from the velocity field of the fluid within them, the nature of the microorganism. Storing the entire velocity field of the cytoplasm fluid within each cell requires an enormous amount of memory. So we adopt a DNN approach for such classification. Specifically, we apply an external force along the boundary of the living cell, causing the boundary shape to change to $\partial B(t)$ at time t . Here, $B(t)$ is the two dimensional region within the cell at time t and $\partial B(t)$ is the corresponding boundary. The velocity field of the cytoplasm at time t is $v(t, x, y) = v_x(t, x, y)\hat{x} + v_y(t, x, y)\hat{y}$ for $(x, y) \in B(t)$ with the boundary condition that $v(t, x, y) = 0$ for $(x, y) \in \partial B(t)$. Setting up the Navier-Stokes equations and the equation of continuity for the cytoplasmic fluid, ie,

$$\operatorname{div}(\rho(t, x, y)v(t, x, y)) + \partial_t \rho(t, x, y) = 0, (x, y) \in B(t),$$

$$\rho(t, x, y)(\partial_t v(t, x, y) + (v(t, x, y), \nabla)v(t, x, y)) = -\nabla p(t, x, y) + \eta \nabla^2 v(t, x, y)$$

along with the equation of state of the fluid

$$p(t, x, y) = F(\rho(t, x, y))$$

we get in all four equations for the four functions v_x, v_y, ρ, p . These equations are to be solved subject to the above stated time varying boundary conditions. We now apply as input at the input layer of our DNN, the boundary data $\partial B(t)$ and at the output layer, we take measurements and match it to the space-time samples of the velocity field of the cytoplasm. In between the input and the output layers, there are p layers. If the signals at the nodes of the k^{th} layer are $x_{k1}(t), \dots, x_{k,L}(t)$, the signals at the nodes of the $(k+1)^{th}$ layer are

$$x_{k+1,m}(t) = \sigma\left(\sum_{j=1}^L w_{kmj} x_{kj}(t)\right), 1 \leq m \leq L, k = 1, 2, \dots, N-1$$

where σ is the usual nonlinear sigmoidal function used in neural networks. The output of the final ie N^{th} layer $x_{N,m}(t), m = 1, 2, \dots, L$ is matched to the measured velocity fields as a function of time at $L/2$ spatial points, ie, to $(v_x(t, x_i, y_i), v_y(t, x_i, y_i) : 1 \leq i \leq L/2\}$ by adapting the weights with time until a steady state is reached. These trained weights are then used as the classifying parameters for the given microorganism cell.

Experimental setup: (1) A good powerful microscope for viewing the cell motion.

(2) Micro-probes for piercing the microorganism without destroying it with sensors at its ends for recording the cytoplasmic fluid velocity field at that specific spatial location within the micro-organism.

(3) A sensitive recorder of the vibrations of the probe that would record the fluid velocity from its vibrations at the other end.

(4) A method for introducing charge into the micro-organism that would respond to an applied electromagnetic field and move in accordance with not only the boundary forces but the electromagnetic force on the charged fluid within it.

(5) An alternate to using probes for taking measurements of the velocity field: We can use X-rays to photograph the fluid inside the cell at different times and hence estimate the velocity field. Further, X-rays can also be used to record the density field fluctuations since where the density is more, the X-ray photograph will display darker spots and vice-versa.

This problem raises the important issue of solving partial differential equations in space-time within a region whose boundary varies with time. For example, consider the 3-D wave equation with forcing

$$u_{,tt}(t, r) - c^2 \nabla^2 u(t, r) - f(t, r) = 0, r \in B(t)$$

with the boundary condition

$$u(t, r) = 0, r \in \partial B(t)$$

This equation can be derived from the action principle $\delta S = 0$ where

$$S[u] = \int dt \int_{B(t)} ((1/2)(\partial_t u)^2 - (1/2)(\nabla u)^2 + fu) d^3 r$$

We can make a transformation

$$r = R(t, \xi), x \in \mathbb{R}^3$$

so that when $r \in B(t)$, it follows that $\xi \in B(0)$. This would then reduce the problem to solving a pde with vanishing boundary conditions on a constant boundary. The above action can be expressed as follows in terms of (t, ξ) :

$$\begin{aligned} du(t, R(t, \xi))/dt &= u_{,t}(t, R) + (\nabla u(t, R), R_{,t}), \\ \nabla_\xi u(t, R) &= R_{,\xi}^T \nabla_r u(t, R) \end{aligned}$$

These equations can be solved for $\nabla_r u$ and $u_{,t}$ in terms of $v_{,t}(t, \xi)$ and $\nabla_\xi v(t, \xi)$ where $v(t, \xi) = u(t, R(t, \xi))$ and substituted into the above action integral to get an action of the form

$$S[v] = \int dt \int_{B(0)} L(t, \xi, v(t, \xi), v_{,t}(t, \xi), \nabla_\xi v(t, \xi)) d^3 \xi$$

Note that the change of measure

$$d^3 r = \det(R_{,\xi}(t, \xi)) d^3 \xi$$

must be used here. The corresponding Euler-Lagrange equations for v are easily written down and the boundary conditions on v are non-time varying, ie, $v(t, \xi) = 0, \xi \in \partial B(0)$. This problem has been addressed in a masterly way in the book

T.Kato, "Perturbation theory for linear operators", Springer.

Finite element method for solving for the fluid velocity and density within a fluctuating boundary.

Consider the action principle

$$\begin{aligned} S[\lambda_x, \lambda_y, \mu, v_x, v_y, \rho] = & \int \lambda_x(t, x, y)(\rho(t, x, y)(v_x v_{x,x} + v_y v_{x,y} + v_{x,t} + p'(\rho(t, x, y))\rho_{,x} \\ & - \eta(v_{x,xx} + v_{x,yy}) - f_x(t, x, y))) dx dy dt \\ & + \int \lambda_y(t, x, y)(\rho(t, x, y)(v_x v_{y,x} + v_y v_{y,y} + v_{y,t} + p'(\rho)\rho_{,y} - \eta(v_{y,xx} + v_{y,yy}) - f_y(t, x, y))) dx dy dt \\ & + \int \mu(t, x, y)((\rho v_x)_{,x} + (\rho v_y)_{,y} - \rho_{,t}) dx dy dt \end{aligned}$$

The Euler-Lagrange equations derived from

$$\delta S = 0$$

give (a) The Navier-Stokes equations when the variation is carried out w.r.t λ_x, λ_y , (b) The equation of continuity when the variation is carried out w.r.t μ , (c) when the variation is carried out w.r.t v_x , we get that $\delta_{v_x} S = 0$ gives us

$$\begin{aligned} & \lambda_x \rho v_{x,x} - (\lambda_x \rho v_x)_{,x} - (\lambda_x \rho v_y)_{,y} - (\lambda_x \rho)_{,t} \\ & - \eta \lambda_{x,xx} - \eta \lambda_{x,yy} - \rho \mu_{,x} \\ & + \lambda_y \rho v_{y,x} = 0 \end{aligned}$$

and likewise for $\delta_{v_y} S = 0$ gives us the same equation after interchanging λ_x with λ_y and v_x with v_y . The equation $\delta_\rho S = 0$ gives us

$$\begin{aligned} & \lambda_x (v_x v_{x,x} + v_y v_{x,y} + v_{x,t}) - (\lambda_x p'(\rho))_{,x} + \lambda_x p''(\rho) \rho_{,x} \\ & + \lambda_y (v_x v_{y,x} + v_y v_{y,y} + v_{y,t}) - (\lambda_y p'(\rho))_{,y} + \lambda_y p''(\rho) \rho_{,y} \\ & + \mu_{,t} - \mu_{,x} v_x - \mu_{,y} v_y = 0 \end{aligned}$$

Solving these equations is hard but we can devise a finite element method involving dividing the fluid region and time interval into pixels and assigning a $v_x, v_y, \rho, \lambda_x, \lambda_y, \mu$ value to each pixel and replacing the above integral by a finite sum followed by setting the partial derivatives in this finite sum w.r.t. all these variables at each of these pixels to zero. At the pixels through which the cell boundary passes, we assign zero velocity and zero density. If the boundary varies with time in a way in which we do not know, then we assume that a change of the coordinate system of the form $x = R_x(t, u, v), y = R_y(t, u, v)$ results in the boundary $u = \text{constt}$. We then perform such a change of the variables using these unknown functions R_x and R_y and set the variational derivative of S w.r.t. these functions also to zero to give us the equations for the time varying boundary.

Remark: Consider first the simple problem of writing down the variational equations $\delta S = 0$ when the action integral has the form

$$S[\phi] = \int_{t_1}^{t_2} dt \int_{\partial B(t)} L(\phi(t, r), \phi_{,t}(t, r), \nabla \phi(t, r)) d^2 r$$

where $\partial B(t)$ is a time varying closed curve in \mathbb{R}^2 . The Dirichlet conditions $\phi(t, r) = 0, r \in \partial B(t)$ are assumed. We make the coordinate transformation $r = R(t, u, v)$ so that $r \in \partial B(t)$ iff $(u, v) \in \Gamma$ where Γ is the square $u = 0, 1$ or $v = 0, 1$. Writing

$$\psi(t, u, v) = \phi(t, R(t, u, v))$$

we find that the boundary condition on ψ becomes

$$\psi(t, u, v) = 0, u = 0, 1, v = 0, 1$$

The action S can now be expressed as

$$\begin{aligned} S[\phi] = S_1[\psi] = \int_{t_1 < t < t_2, 0 < u < 1, 0 < v < 1} & L_1(\psi(t, u, v), \psi_{,t}(t, u, v), \psi_{,u}(t, u, v), \psi_{,v}(t, u, v), \\ & R_{,t}(t, u, v), R_{,u}(t, u, v), R_{,v}(t, u, v)) du dv \end{aligned}$$

To obtain L_1 we first note that

$$d^2r = \det(R_{,\xi}(t, \xi)) d^2\xi, \xi = (u, v)$$

and next

$$\psi_{,\xi} = \phi_{,r}(t, R) R_{,\xi}$$

so that

$$\phi_{,r}(t, R) = \psi_{,\xi}(R_{,\xi})^{-1}$$

in row vector format, or equivalently, in column vector format, we get by taking the transpose of the above equation that

$$\nabla_r \phi = (R_{,\xi}(t, \xi))^{-T} \nabla_\xi \psi(t, \xi)$$

Further,

$$\psi_{,t}(t, \xi) = \phi_{,t}(t, R) + \phi_{,r}(t, R) R_{,t}$$

or equivalently,

$$\begin{aligned} \phi_{,t} &= \psi_{,t} - (\nabla_r \phi)^T R_{,t} \\ &= \psi_{,t}(t, \xi) - (\nabla_\xi \psi(t, \xi))^T R_{,\xi}(t, \xi)^{-1} R_{,t}(t, \xi) \end{aligned}$$

Making these substitutions in L gives us L_1 in the above form, ie,

$$\begin{aligned} L_1(\psi(t, \xi), \psi_{,t}(t, \xi), \nabla_\xi \psi(t, \xi), R_{,t}(t, \xi), R_{,\xi}(t, xi)) \\ = L(\phi(t, R), \phi_{,t}(t, R), \nabla_r \phi(t, R)) \det(R_{,\xi}(t, \xi)) \end{aligned}$$

with

$$R = R(t, \xi)$$

Owing to the elementary boundary conditions after this transformation, the finite element method for minimizing L_1 can easily be formulated by dividing the u-v space, ie, $[0, 1]^2$ into discrete pixels and imposing the boundary condition that ψ vanishes on the boundary pixels.

Reference: Vijayant Agrawal and Harish Parthasarathy, "Estimating the shape of structures within a fluid with fluctuating boundary from measurements of the velocity and density fields of the fluid within the boundary, Technical Report, NSIT, 2018.

3.27.3 Exercises

[1] Write down the fluid dynamic equations in two dimensions using plane polar coordinates.

[2] Write down the fluid dynamic equations for a conducting fluid explaining how they get coupled to the Maxwell equations via the force per unit volume $J \times B, J = \sigma(E + v \times B)$. Describe the dispersion relations in the presence of a

given constant magnetic field component $B_0\hat{z}$ apart from the small electric and magnetic field fluctuations. Use linearization of the fluid dynamic equations.

[3] Describe the fluid dynamic equations within a cell whose boundary changes with time in a given prescribed manner, ie, you may assume a coordinate transformation $(x, y) = R(t, u, v)$ where R is a given function and the boundary is $u = 0, 1, v = 0, 1$.

3.27.4 Points to remember

[1] The cytoplasm within a cell can be modeled using basic fluid dynamics in a region with a time varying boundary. When we apply a force on the boundary, we change the dynamics of the density and velocity field of the cytoplasm and by recording the input force and the boundary shape, we can in principle solve for the velocity field as a function of these two. We can then design a DNN (Deep Neural Network) that takes as input the force applied on the boundary and trains the weights of the DNN so that the output matches with the computed velocity field. In this way, for cells having different kinds of diseases, we can train the weights and store them. Now, given a new cell, we apply a similar force on its boundary and record the boundary shape using which, we compute the velocity field of the cytoplasm and compute the weights from these inputs and outputs. We match the evolved weights with the set of trained weights and classify the disease according to which weight pattern the evolved one is closest to.

[2] By introducing Lagrange multiplier fields, the equations of motion of a fluid can be cast as a variational problem to which one may apply the finite element method for solving it. Further, for a fluid dynamical problem with a time varying boundary, one can transform the coordinate system so that the transformation equations involve time also and such that the boundary conditions for the transformed problem assume a simple time invariant form.

[3] Using the Lagrangian of the fluid as described above by the introduction of Lagrange multipliers, we can by applying the Legendre transform, derive the Hamiltonian of the fluid and hence formulate the Schrodinger equation for the quantum fluid.

3.28 Snell's laws of reflection and refraction on surfaces separating two uniform media

3.28.1 Summary

The general theory of aberrations can also be explained from this. Suppose we have a surface $z = \psi(x, y)$ that separates two uniform media. On the side $z < \psi(x, y)$, the medium parameters are (ϵ_1, μ_1) while on the side $z > \psi(x, y)$, the parameters are (ϵ_2, μ_2) . The incident electric wave field phasor is in medium 1, is of a definite frequency and is expressed as a superposition of plane TEM wave:

$$E_i(r) = \int E_{i0}(\hat{n}) \exp(-j\omega \hat{n} \cdot r / c_1) d\Omega(\hat{n})$$

where

$$c_1 = c_0 / \sqrt{\epsilon_1 \mu_1}, \hat{n} \cdot E_{i0}(\hat{n}) = 0$$

The reflected electric wave field is of the same form

$$E_r(r) = \int E_{r0}(\hat{n}) \exp(-j\omega \hat{n} \cdot r / c_1) d\Omega(\hat{n})$$

where

$$\hat{n} \cdot E_{r0}(\hat{n}) = 0$$

and finally, the transmitted field is of the form

$$E_t(r) = \int E_{t0}(\hat{n}) \exp(-j\omega \hat{n} \cdot r) d\Omega(\hat{n})$$

where

$$\hat{n} \cdot E_{t0}(\hat{n}) = 0$$

These solutions follow from the Maxwell equations in free space. The corresponding magnetic fields in the incident, reflected and transmitted waves are obtained from the Maxwell curl equations

$$\operatorname{curl} E = -j\omega \mu H$$

so

$$H_i(r) = \eta_1^{-1} \int \hat{n} \times E_{i0}(\hat{n}) \exp(-j\omega \hat{n} \cdot r / c_1) d\Omega(\hat{n}),$$

$$H_r(r) = \eta_1^{-1} \int \hat{n} \times E_{r0}(\hat{n}) \exp(-j\omega \hat{n} \cdot r / c_1) d\Omega(\hat{n}),$$

$$H_t(r) = \eta_2^{-1} \int \hat{n} \times E_{t0}(\hat{n}) \exp(-j\omega \hat{n} \cdot r / c_2) d\Omega(\hat{n}),$$

where

$$\eta_k = \sqrt{\mu_k / \epsilon_k}, k = 1, 2$$

is the impedance of the medium $k, k = 1, 2$. The boundary conditions are that the tangential components of the electric and magnetic field are continuous at the interface $z = \psi(x, y)$. This means that

$$\begin{aligned} & E_i(x, y, \psi(x, y)) + E_r(x, y, \psi(x, y)) - \\ & (E_i(x, y, \psi(x, y)), (\psi_{,x}(x, y)\hat{x} + \psi_{,y}(x, y)\hat{y} - \hat{z}))(\psi_{,x}(x, y)\hat{x} + \psi_{,y}(x, y)\hat{y} - \hat{z}) / (\psi_{,x}(x, y)^2 + \psi_{,y}(x, y)^2 + 1) \\ & - (E_r(x, y, \psi(x, y)), (\psi_{,x}(x, y)\hat{x} + \psi_{,y}(x, y)\hat{y} - \hat{z}))(\psi_{,x}(x, y)\hat{x} + \psi_{,y}(x, y)\hat{y} - \hat{z}) / (\psi_{,x}(x, y)^2 + \psi_{,y}(x, y)^2 + 1) \\ & = E_t(x, y, \psi(x, y)) - \\ & (E_t(x, y, \psi(x, y)), (\psi_{,x}(x, y)\hat{x} + \psi_{,y}(x, y)\hat{y} - \hat{z}))(\psi_{,x}(x, y)\hat{x} + \psi_{,y}(x, y)\hat{y} - \hat{z}) / (\psi_{,x}(x, y)^2 + \psi_{,y}(x, y)^2 + 1) \end{aligned}$$

and likewise with E_i, E_r, E_t replaced by the above expressions for H_i, H_r, H_t . From these equations which hold good for all x, y we must solve for $E_{r0}(\hat{n})$ and $E_{t0}(\hat{n})$ in terms of the function $E_{i0}(\hat{n})$ where \hat{n} varies over the entire unit sphere. In general, this is a very complex problem and so for obtaining concrete results, we discuss below the special case when the surface is a plane, namely the xy plane $z = 0$.

3.28.2 Discussion

Prove Snell's laws of reflection and refraction for plane waves incident on a plane boundary that separates two media having permittivity and permeability pairs $(\epsilon_k, \mu_k), k = 1, 2$. Calculate the ratio of the reflected power and transmitted power fluxes to this incident power flux for the case in which the incident wave electric field vector makes an angle $\pi/2 - \alpha_i$ w.r.t the plane containing the incident, reflected and refracted ray.

hint: Assume $z = 0$ to be the boundary separating the two media, so that for $z < 0$, the medium has parameters (ϵ_1, μ_1) and for $z > 0$, the parameters (ϵ_2, μ_2) . The wave-number in medium 1 is $\beta_1 = \omega\sqrt{\epsilon_1\mu_1}$ and that in medium 2 is $\beta_2 = \omega\sqrt{\epsilon_2\mu_2}$. The incident electric field phasor for perpendicular polarization ($\alpha = 0$) is

$$E_i(r) = E_{i0} \exp(-j\beta_1 \hat{n}_i \cdot r) \hat{x}$$

where

$$\hat{n}_i = \hat{y} \sin(\theta_i) + \hat{z} \cos(\theta_i)$$

with θ_i being the angle of incidence. On the other hand, for parallel polarization ($\alpha = \pi/2$), the magnetic field is polarized perpendicular to the plane of incidence:

$$H_i = H_{i0} \exp(-j\beta_1 \hat{n}_i \cdot r) \hat{x}$$

The corresponding electric field is calculated from Maxwell's equations as

$$\nabla \times H_i = j\omega\epsilon_1 E_i$$

so that

$$E_i = (-\beta_1/\omega\epsilon_1) \hat{n}_i \times \hat{x} H_{i0} \exp(-j\beta_1 \hat{n}_i \cdot r)$$

Note that

$$\beta_1/\omega\epsilon_1 = \eta_1 = \sqrt{\mu_1/\epsilon_1}$$

For an arbitrary polarization, (ie mixture of parallel and perpendicular polarization), the electric field is a linear combination of these two cases:

$$E_i = E_{i0}(\cos(\alpha_i)\hat{x} + \sin(\alpha_i)\hat{n}_i \times \hat{x})\exp(-j\beta_1\hat{n}_i \cdot \hat{r})$$

Note that

$$\hat{n}_i \cdot \hat{r} = y \cdot \sin(\theta_i) + z \cdot \cos(\theta_i)$$

and hence,

$$\hat{n}_i \times \hat{x} = -\hat{z} \cdot \sin(\theta_i) + \hat{y} \cdot \cos(\theta_i)$$

so that incident electric field phasor is

$$E_i(r) = E_{i0}(\cos(\alpha_i)\hat{x} + \sin(\alpha_i)\cos(\theta_i)\hat{y} - \sin(\alpha_i)\sin(\theta_i)\hat{z})\exp(-j\beta_1\hat{n}_i \cdot \hat{r})$$

Likewise, the reflected electric field has the form

$$E_r(r) = E_{r0}(\cos(\alpha_r)\hat{x} + \sin(\alpha_r)\cos(\theta_i)\hat{y} + \sin(\alpha_r)\sin(\theta_r)\hat{z})\exp(-j\beta_1\hat{n}_r \cdot \hat{r})$$

where

$$\hat{n}_r = \hat{y} \cdot \sin(\theta_r) - \hat{z} \cdot \cos(\theta_r)$$

so that

$$\hat{n}_r \cdot \hat{r} = y \cdot \sin(\theta_r) - z \cdot \cos(\theta_r)$$

and finally, the transmitted/refracted electric field phasor is

$$E_t(r) = E_{t0}(\cos(\alpha_t)\hat{x} + \sin(\alpha_t)\cos(\theta_t)\hat{y} - \sin(\alpha_t)\sin(\theta_t)\hat{z})\exp(-j\beta_2\hat{n}_t \cdot \hat{r})$$

where

$$\hat{n}_t = \hat{y} \cdot \sin(\theta_t) + \hat{z} \cdot \cos(\theta_t)$$

so that

$$\hat{n}_t \cdot \hat{r} = y \cdot \sin(\theta_t) + z \cdot \cos(\theta_t)$$

The magnetic phasor fields corresponding to the electric field in the incident, reflected and transmitted waves are respectively found using the Maxwell equation

$$\nabla \times E = -j\omega\mu H$$

This gives

$$-j\beta_1\hat{n}_i \times E_i = -j\omega\mu_1 H_i,$$

or

$$H_i = (\beta_1/\omega\mu_1)\hat{n}_i \times E_i = \hat{n}_i \times E_i / \eta_1$$

and likewise,

$$H_r = \hat{n}_r \times E_r / \eta_1, H_t = \hat{n}_t \times E_t / \eta_2$$

Now

$$\hat{n}_i \times E_i = \hat{n}_i \times (\cos(\alpha_i)\hat{x} + \sin(\alpha_i)\hat{n}_i \times \hat{x})E_{i0}\exp(-j\beta_1\hat{n}_i \cdot \hat{r})$$

$$= (\cos(\alpha_i)\hat{n}_i \times \hat{x} - \sin(\alpha_i)\hat{x})E_{i0}.\exp(-j\beta_1\hat{n}_i.r)$$

where we use

$$\hat{n}_i.\hat{x} = 0$$

In this way, we find for the magnetic field component in each ray,

$$H_i = \eta_1^{-1}(\cos(\alpha_i)\hat{n}_i \times \hat{x} - \sin(\alpha_i)\hat{x})E_{i0}.\exp(-j\beta_1\hat{n}_i.r)$$

$$H_r = \eta_1^{-1}(\cos(\alpha_r)\hat{n}_r \times \hat{x} - \sin(\alpha_r)\hat{x})E_{r0}.\exp(-j\beta_1\hat{n}_r.r)$$

$$H_t = \eta_2^{-1}(\cos(\alpha_t)\hat{n}_t \times \hat{x} - \sin(\alpha_t)\hat{x})E_{t0}.\exp(-j\beta_2\hat{n}_t.r)$$

Equating the tangential components (ie, x and y components) of the electric field at the boundary $z = 0$ gives us

$$\begin{aligned} & E_{i0}(\cos(\alpha_i)\hat{x} + \sin(\alpha_i)\cos(\theta_i)\hat{y}).\exp(-j\beta_1 y.\sin(\theta_i)) \\ & + E_{r0}(\cos(\alpha_r)\hat{x} + \sin(\alpha_r)\cos(\theta_r)\hat{y}).\exp(-j\beta_1 y.\sin(\theta_r)) \\ & = E_{t0}(\cos(\alpha_t)\hat{x} + \sin(\alpha_t)\cos(\theta_t)\hat{y}).\exp(-j\beta_2 y.\sin(\theta_t)) \end{aligned}$$

Since this equation holds for all $y \in \mathbb{R}$, we must have

3.28.3 Exercises

[1] Show that in the individual cases of parallel and perpendicular polarization involving reflection and refraction of an em plane wave at the interface of two dielectrics, the power flux is conserved, ie, calculate first the Poynting vectors for the incident, reflected and transmitted em fields:

$$S_i = |E_i|^2 \hat{n}_i / 2\eta_1, S_r = |E_r|^2 \hat{n}_r / 2\eta_1, S_t = |E_t|^2 \hat{n}_t / 2\eta_2$$

and prove the sum of the reflected and transmitted power fluxes at the interface equals power flux of the incident wave at the interface, ie,

$$S_i.\hat{z} = S_r.(-\hat{z}) + S_t.\hat{z}$$

[2] Describe the dispersion relation for em waves in a conducting medium with parameters (ϵ, μ, σ) .

[3] Describe the transmitted and reflected em fields at the interface of two conductors with parameters $(\epsilon_k, \mu_k, \sigma_k), k = 1, 2$ corresponding to normal incidence. Calculate the amount of power dissipated per unit area of the interface in the second medium. For this you must calculate $\int_0^\infty \sigma_2 |E_t|^2 dz / 2$. Use the results of the previous problem to describe damped plane waves in the two media. Note that in this problem, the frequency ω is fixed. In fact you must first show using time domain analysis that the frequency remains unchanged during the process of reflection and transmission.

[4] Calculate the reflectivity $R = |E_r|^2 / |E_i|^2$ and transmittivity $T = |E_t|^2 / |E_i|^2$ for the case of normal incidence at the interface of two dielectrics taking their conductivities into account. Hence explain why metals are good reflectors of light as compared with insulators.

3.28.4 Points to remember

[1] The solution to Maxwell's equations in free space show that the electric and magnetic fields can be expressed as a superposition of plane waves all travelling with the velocity of light in that medium corresponding to the given frequency. Therefore, it is of interest to consider an electromagnetic field in which the electric field and magnetic fields are single plane waves, ie, have just the same wave vector. If such a wave is incident from one dielectric medium onto the interface between this medium and another medium, then there will also be a reflected and transmitted em wave. The total em field on the first side, ie, the side of incidence will be the sum of the em fields in the incident and reflected waves while the em field on the second, ie, transmitted side will have just one component. Maxwell's equations show that in any plane wave, the electric field vector, the magnetic field vector and the wave vector, ie, the wave vector magnitude multiplied by the direction of propagation are mutually perpendicular. Further, from the boundary conditions of Maxwell's equations, it is easily seen that the wave vectors of the incident, reflected and transmitted waves are coplanar (this plane is called the plane of incidence) and that the frequency of the wave remains the same in the incident, reflected and transmitted em wave fields. It is also clear from the boundary conditions that if the electric field in the incident wave is polarized perpendicular to this plane of incidence, then the same will be true for the reflected and transmitted electric fields. This situation is called "perpendicular polarization". Likewise, if the electric field in the incident wave is polarized parallel to the plane of incidence, then the same will be true for the reflected and transmitted fields. This situation is called "parallel polarization". More generally, we can form a linear combination of these two extreme cases of polarization and conclude from the principle of superposition (applicable because Maxwell's equation is are linear) that if in the incident wave field, the electric field vector makes an angle α with the plane of incidence then the same will be true in the reflected and transmitted waves. It should be noted that since the electric and magnetic field vectors in a plane wave are mutually perpendicular, it follows that if the electric field is parallelly polarized, then the magnetic field will be perpendicularly polarized and vice versa. The boundary conditions state that at a dielectric interface, there cannot be any surface current density and hence not only are the tangential and normal components of the electric field E and magnetic field $B = \mu H$ continuous at the boundary, but so also is the tangential component of H continuous. The E field on the first side is the sum of the E fields in the incident and reflected wave as is correspondingly true for the H field and the E field on the second side is just the transmitted field as is also true for the H and B fields. The only linearly independent boundary conditions that one gets for this situation of reflection and transmission through the interface between two dielectrics are the continuity of the tangential components of E and H . Continuity of the normal component of B does not give us any new equation. From these two boundary conditions, we can derive formulas for the reflection and transmission coefficients R, T which are respectively the ratios of the reflected E field to incident E field amplitude and transmitted E

field to incident E field amplitude. Obviously, $1 + R = T$ follows for the case of perpendicular polarization. These coefficients are expressible in terms of the angle of incidence and the values of ϵ, μ on both the sides. We can also prove using these boundary conditions the Snell's laws of reflection and refraction. More generally, we can consider a plane em wave incident on a dielectric interface of arbitrary shape and apply the boundary conditions. Although concrete results are not obtained, we can in principle obtain an algorithm for computing the reflected wave field and the transmitted wave fields which will not generally be single plane waves. These will be superpositions of plane waves of the same frequency but with wave vectors pointed along different directions.

Finally for a planar dielectric interface, we can by using the formulas for R and T derive expressions for the incident flux of power P_i through a unit area of the interface, the reflected flux of power P_r through a unit area of the interface and the transmitted flux of power P_t through a unit area of the interface. These are computed by taking the dot product of the unit normal (with a reversed sign in the case of the reflected wave) to the interface with the average Poynting vector in the respective incident, reflected and transmitted waves. After making these computations, we can deduce the law of energy conservation: $P_i = P_r + P_t$.

3.29 Spinor form of some equations of mathematical physics:Roger Penrose's theory

3.29.1 Summary

Penrose has shown (Ref:Roger Penrose and Wolfgang Rindler, Spinor and twistor approach to physics, two volume set) that almost all the major classical and quantum field theories known today are expressible more naturally in spinor form rather than in vector tensor form. The idea to first choose a Penrose null tetrad and construct generalized Pauli spin matrices which depend on the space and time coordinates. These four matrices give us the required transformation law from four vector fields to spinor fields and vice-versa. Using this idea, the first major success was achieved by S.Chandrasekhar when he was able to solve the Dirac equation in the Kerr metric. Already, Penrose had earlier proposed the formulation of of Dirac's equation for the electron in an arbitrary curved space-time background. But first explicit solutions to these equations for specific space-times that occur in reality like the Schwarzhild and Kerr metrics were given by Chandrasekhar. Dirac's equation in curved space-time is important because by solving in near the event horizon of a blackhole, one can derive expressions for the generalized Hawking temperature at which electrons, positrons and photons will be emitted by the blackhole. Just as one makes a transition form the curved space-time coordinates to a locally inertial frame using a tetrad basis which results in associating four scalar fields with each four vector field, one can also make a transition from spinor fields with two components in curved space-time to two scalar fields in the same curved space-time using a dyad basis. It is by using a dyad basis in conjunction with the generalized Pauli spin matrices that one is able to define 2×2 spinor matrix fields and then 2×2 spinor scalar fields corresponding to a four vector field. The advantage of the spinor formalism is that many times when there are highly nonlinear tensor field equations, by using spinors these reduce to elementary spinor field equations.

3.29.2 Discussion

$l^\mu, n^\mu, m^\mu, \bar{m}^\mu$ form the Newman Penrose tetrad of four vector fields satisfying $l.l = n.n = m.m = \bar{m}.\bar{m} = l.m = m.n = \bar{m}.n = 0$ and $l.n = 1, m.\bar{m} = -1$ where by $a.b$ we mean $g_{\mu\nu}a^\mu b^\nu$. We form the generalized Pauli spin matrices

$$\sigma^\mu = \begin{pmatrix} l^\mu & m^\mu \\ \bar{m}^\mu & n^\mu \end{pmatrix}, \mu = 0, 1, 2, 3$$

Its components, we denote by $\sigma_{ab}^\mu, a, b = 0, 1$. These are Hermitian matrices since l, n are real. The fundamental spinor corresponding to the metric $g_{\mu\nu}$ is the skew symmetric 2×2 matrix $\epsilon_{AB'}$ and its inverse is denoted by $\epsilon^{AB'}$. Just as the tetrad l, n, m, \bar{m} takes us from the noninertial frame with metric $g_{\mu\nu}$ to a locally inertial frame with metric η_{ab} having constant coefficients, so that

$$g^{\mu\nu} = l^\mu n^\nu + l^\nu n^\mu - m^\mu \bar{m}^\nu - \bar{m}^\mu m^\nu$$

we can introduce a dyad $\zeta_a^A, A, a = 0, 1$ which takes us from the non-inertial spinor metric ϵ^{AB} to the locally inertial spinor metric ϵ_{ab} by the formula

$$\epsilon_{ab} = \epsilon_{AB} \zeta_a^A \zeta_B^B$$

or equivalently,

$$\epsilon^{AB} = \epsilon^{ab} \zeta_a^A \zeta_b^B$$

where

$$, ((\epsilon_{ab})) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

and $((\epsilon^{ab}))$ is its inverse, namely its negative. We have the correspondence that

$$\sigma_{\#b} = \sigma_{AB} \zeta_a^A \zeta_B^B$$

where $((\sigma_{AB}^\mu)), \mu = 0, 1, 2, 3$ are the usual Pauli spin matrices that one encounters in special relativity. We can write

$$l_\mu = \sigma_{\#b} o_a o_b, m_\mu = \sigma_{\#b} o_a k_b,$$

$$\bar{m}^\mu = \sigma_{ab}^\mu k^a o^b, n^\mu = \sigma_{ab}^\mu k^a k^b$$

where

$$(o^a) = [1, 0]^T, (k^a) = [0, 1]^T$$

Alternately, we define

$$o_a = \zeta_a^A o_A, k_a = \zeta_a^A k_A$$

and

$$o^a = \epsilon^{ab} o_b, k^a = \epsilon^{ab} k_b$$

and hence write in terms of the flat space time Pauli matrices

$$l^\mu = \sigma_{AB}^\mu o^A o^B, m^\mu = \sigma_{AB}^\mu o^A k^B,$$

$$\bar{m}^\mu = \sigma_{AB}^\mu k^A o^B, n^\mu = \sigma_{AB}^\mu k^A k^B$$

The identities

$$o^A o_A = \epsilon^{AB} o_A o_B = 0, k^A k_A = \epsilon^{AB} k_A k_B = 0$$

give respectively

$$l.l = 0 = n.n$$

The identities

$$o^A k_A = \epsilon^{AB} o_A k_B = 1, k^A o_A = \epsilon^{AB} k_A o_B = -1$$

imply respectively

$$l.n = 1, m.\bar{m} = -1$$

We note that

$$o^A = \zeta_a^A o^a, k^A = \zeta_a^A k^a$$

and also

$$\epsilon^{ab} \zeta_a^A \zeta_b^B = \epsilon^{AB}$$

This equation is the dyad analogue of expressing the metric tensor in terms of a tetrad.

Covariant derivative of a dyad: Let ξ^A be a spinor field. In locally inertial coordinates, its components are

$$\zeta_A^a \xi^A = \xi^a$$

where

$$\zeta_A^a \zeta_b^A = \delta_b^a, \zeta_A^a \zeta_a^B = \delta_A^B$$

Now consider a four vector field A_μ and its covariant derivative tensor $A\mu : \nu$. From physical intuition, we must have

$$\begin{aligned} \sigma_{AB}^\mu \sigma_{CD}^\nu A_{\mu:\nu} &= \sigma_{AB}^\nu (\sigma_{CD}^\mu A_\mu)_{:\nu} \\ &= \sigma_{AB}^\nu (\sigma_{CD:\nu}^\mu A_\mu + \sigma_{CD}^\mu A_{\mu:\nu}) \\ &= \sigma_{AB}^\mu \sigma_{CD:\nu}^\mu A_\mu + \sigma_{AB}^\nu \sigma_{CD}^\mu A_{\mu:\nu} \end{aligned}$$

Thus, we must necessarily have the relation

$$\sigma_{AB:\nu}^\mu = 0$$

Further, we have

$$g^{\mu\nu} = \sigma_{AB}^\mu \sigma_{CD}^\nu \epsilon^{AC} \epsilon^{BD}$$

or equivalently,

$$g_{\mu\nu} = \sigma_\mu^{AB} \sigma_\nu^{CD} \epsilon_{AC} \epsilon_{BD}$$

Taking the covariant derivative on both sides gives

$$0 = g_{\mu\nu:\rho} = \sigma_\mu^{AB} \sigma_\nu^{CD} (\epsilon_{AC:\rho} \epsilon_{BD} + \epsilon_{AC} \epsilon_{BD:\rho})$$

Note that

$$\begin{aligned}\sigma_{AB}^{\mu} \sigma_{\nu}^{AB} &= \delta_{\nu}^{\mu}, \\ \sigma_{AB}^{\mu} \sigma_{\mu}^{CD} &= \delta_A^C \delta_B^D\end{aligned}$$

by definition. It follows that

$$\epsilon_{AC:\rho} \epsilon_{BD} + \epsilon_{AC} \epsilon_{BD:\rho} = 0$$

From this, it follows that

$$\epsilon_{AC:\mu} = 0$$

Now we wish to define the covariant derivative of a spinor. From intuitive considerations, if ξ^A, η^A are two spinor fields, then

$$\begin{aligned}\sigma_{CD}^{\mu} (\xi^A \eta^B)_{:\mu} &= (\xi^A \eta^B)_{:CD} \\ &= \xi_{:CD}^A \eta^B + \xi^A \eta_{:CD}^B\end{aligned}$$

Also, since

$$\xi^a = \xi^A \zeta_A^a$$

is a scalar field because ζ_A^a is a dyad, it follows that

$$\xi^A = \xi^a \zeta_a^A$$

and hence

$$\xi_{:\mu}^A = \xi_{,\mu}^a \zeta_a^A + \xi^a \zeta_{a:\mu}^A$$

Then,

$$\sigma_{BA}^{\mu} \xi_{:\mu}^A = \sigma_{BA}^{\mu} \zeta_a^A \xi_{,\mu}^a + \xi^a \sigma_{BA}^{\mu} \zeta_{a:\mu}^A$$

Hence,

$$\begin{aligned}\sigma_{BA}^{\mu} \zeta_b^B \xi_{:\mu}^A &= \sigma_{BA}^{\mu} \zeta_b^B \zeta_a^A \xi_{,\mu}^a \\ &\quad + \xi^a \sigma_{BA}^{\mu} \zeta_b^B \zeta_{a:\mu}^A \\ &= \sigma_{ba}^{\mu} \xi_{,\mu}^a + \xi^a \sigma_{BA}^{\mu} \zeta_b^B \zeta_{a:\mu}^A\end{aligned}$$

We now use the fact that

$$\sigma_{ba}^{\mu}$$

is one of the four vectors $l^\mu, m^\mu, \bar{m}^\mu, n^\mu$ for each $a, b = 0, 1$. Therefore,

$$\sigma_{ba:\nu}^{\mu} = \sigma_{ba,\nu}^{\mu} + \Gamma_{\rho\nu}^{\mu} \sigma_{ba}^{\rho}$$

This implies that

$$\begin{aligned}\sigma_{BA}^{\mu} \zeta_b^B \zeta_{a:\mu}^A &= \\ \sigma_{ba:\mu}^{\mu} - \sigma_{BA}^{\mu} \zeta_b^B \zeta_a^A &= \end{aligned}$$

(a) Dirac's relativistic wave equation in spinor form.

$$\begin{aligned}\sigma_{AB}^\mu P_{:\mu}^A &= \\ \sigma_{ab}^\mu \zeta_A^a \zeta_B^b P_{:\mu}^A &= \\ \sigma_{ab}^\mu zeta_B^b [(\zeta_A^a P^A)_{:\mu} - P^A \zeta_{A:\mu}^a] &= \\ = \sigma_{ab}^\mu \zeta_B^b (P_{,\mu}^a - P^A \zeta_{A:\mu}^a) &\end{aligned}$$

Now,

$$\begin{aligned}\sigma_{ab}^\mu \zeta_B^b \zeta_{A:\mu}^a &= \\ (\sigma_{ab}^\mu \zeta_B^b \zeta_A^a)_{:\mu} - \sigma_{ab}^\mu \zeta_{B:\mu}^b \zeta_A^a - \sigma_{ab:\mu}^\mu \zeta_B^b \zeta_A^a &= \\ = -\sigma_{ab}^\mu \zeta_{B:\mu}^b \zeta_A^a - \sigma_{ab:\mu}^\mu \zeta_B^b \zeta_A^a &\end{aligned}$$

since

$$\sigma_{AB}^\mu = \sigma_{ab}^\mu \zeta_B^b \zeta_A^a$$

and

$$\sigma_{AB:\nu}^\mu = 0$$

Thus,

$$\sigma_{ab}^\mu (\zeta_{A:\mu}^a \zeta_B^b + \zeta_A^a \zeta_{B:\mu}^b) + \sigma_{ab:\mu}^\mu \zeta_A^a \zeta_B^b = 0$$

In fact, we have the more general result by the same logic that

$$0 = \sigma_{AB:\nu}^\mu = \sigma_{ab}^\mu (\zeta_{A:\nu}^a \zeta_B^b + \zeta_A^a \zeta_{B:\nu}^b) + \sigma_{ab:\nu}^\mu \zeta_A^a \zeta_B^b$$

From this, we easily deduce that

$$\sigma_{ab}^\mu \zeta_{A:\nu}^a + \sigma_{ab}^\mu \zeta_A^a \zeta_{B:\nu}^b + \sigma_{ab:\nu}^\mu \zeta_A^a \zeta_B^b$$

The Dirac equation is expressed in the form

$$\sigma_{ab}^\mu \zeta_A^a P_{:\mu}^A = iQ_a = imQ_A \zeta_A^a,$$

$$\sigma_{ba}^\mu \zeta_A^a Q_{:\mu}^A = im\bar{P}_A \zeta_A^a$$

In flat space-time, the Dirac equation is usually expressed in the form

$$i\partial_t \psi = (\alpha, -i\nabla) \psi + \beta m \psi$$

or equivalently, writing

$$\psi = [\phi^T, \chi^T]^T$$

where ϕ, χ are spinors, as

$$i\partial_t \phi = (\sigma, -i\nabla) \chi + m\phi,$$

$$i\partial_t \chi = (\sigma, -i\nabla) \phi - m\chi$$

Writing

$$P = \phi + \chi, Q = \phi - \chi,$$

the Dirac equation can equivalently be expressed as

$$i\partial_t P = (\sigma, -i\nabla)P + mQ,$$

$$i\partial_t Q = -(\sigma, -i\nabla)Q + mP$$

These are equivalent to

$$\partial_t P = -(\sigma, \nabla)P - imQ,$$

$$\partial_t Q = (\sigma, \nabla)Q - imP$$

The first equation can be expressed in terms of spinor components as

$$\partial_t P^A + (\sigma^r)_{AB} \partial_r P^B = -imQ_A$$

and noting that

$$Q^A = \epsilon^{AB} Q_B$$

the second equation is the same as

$$\partial_t Q_A - (\sigma^r)_{AB} \partial_r Q_B = -imP^A$$

or equivalently, replacing A by C and premultiplying by $\epsilon^{AC} = -\epsilon_{AC}$ gives us

$$\partial_t Q^A + \epsilon_{AC} (\sigma^r)_{CB} \partial_r Q_B = imP_A$$

or equivalently,

$$\partial_t Q^A + \epsilon_{AC} (\sigma^r)_{CB} \epsilon_{BD} Q^D = imP_A$$

Now,

$$\epsilon \sigma^1 \epsilon = i\sigma^2 \sigma^1 i\sigma^2 = -\sigma^2 \sigma^1 \sigma^2 = -i\sigma^2 \sigma^3 = \sigma^1,$$

$$\epsilon \sigma^2 \epsilon = -(\sigma_2)^3 = -\sigma^2,$$

$$\epsilon \sigma^3 \epsilon = -\sigma^2 \sigma^3 \sigma^2 = -i\sigma^1 \sigma^2 = \sigma^3$$

Hence, noting that σ^1, σ^3 have all entries real while σ^2 has all entries purely imaginary, we can express the second equation as

$$\partial_t Q^A + (\bar{\sigma}^r)_{AB} \partial_r Q^B = imP_A$$

Renaming \bar{Q}_A as Q_A and taking the conjugate of this equation, it follows that Dirac's equation in flat space-time can be expressed as

$$\partial_t P^A + (\sigma^r)_{AB} \partial_r P^B = -im\bar{Q}_A,$$

$$\partial_t Q^A + (\sigma_r)_{AB} \partial_r Q^B = -im\bar{P}_A$$

or equivalently in four vector notation as

$$\sigma_{AB}^\mu \partial_\mu P^B = -im\bar{Q}_A,$$

$$\sigma_{AB}^\mu \partial_\mu Q^B = -im\bar{P}_A$$

We now look at the curved space-time situation:

$$\begin{aligned} \sigma_{ab}^\mu \zeta_A^a P_{;\mu}^A &= \\ \sigma_{ab}^\mu \zeta_A^a (P^b \zeta_b^A)_{;\mu} &= \\ \sigma_{ab}^\mu (\zeta_A^a \zeta_{b;\mu}^A P^b + \zeta_b^A P_{,\mu}^b) &= \\ \zeta_b^A \sigma_{ab}^\mu P_{,\mu}^b + \sigma_{ab}^\mu \zeta_A^a \zeta_{b;\mu}^A P^b & \end{aligned}$$

3.29.3 Exercises

[1] Write down the

- (b) Yang-Mills equations in spinor form.
- (c) Maxwell equations in spinor form.
- (d) The Einstein field equations in spinor form.

Hint: Let the covariant derivative of the Yang-Mills particle be given by

$$\nabla_\mu = \partial_\mu + ieA_\mu^\alpha(x)\tau_\alpha$$

where τ_α are the Hermitian generators of the gauge group $U(N)$. Replace this covariant derivative by the spinor covariant derivative

$$\nabla_{AB} = \sigma_{AB}^\mu \nabla_\mu$$

Calculate the spinorial curvature

$$F_{ABCD} = [\nabla_{AB}, \nabla_{CD}] = \sigma_{AB}^\mu \sigma_{CD}^\nu [\nabla_\mu, \nabla_\nu]$$

and express the Lagrangian density of the gauge field $Tr(F_{\mu\nu}F^{\mu\nu})$ in terms of F_{ABCD} and

$$F^{ABCD} = \epsilon^{AA'} \epsilon^{BB'} \epsilon^{CC'} \epsilon^{DD'} F_{A'B'C'D'}$$

Now after writing down the Yang-Mills gauge field Lagrangian density in spinorial form, we write down the Yang-Mills matter field Lagrangian density in spinorial form by starting with the Lagrangian density

$$(1/2)(\nabla_\mu \psi)^*(\nabla^\mu \psi) - (\mu^2/2)\psi^*\psi$$

and writing

$$\nabla_\mu = \sigma_\mu^{AB} \nabla_{AB},$$

and taking

$$\psi = [(\phi^A)^T, (\chi^A)^T]^T$$

The field equations then follow from elementary variational calculus.

Maxwell's equations in spinorial form are obtained by replacing the electromagnetic four potential A_μ by

$$A_{BC} = \sigma_{BC}^\mu A_\mu$$

and the field tensor by

$$\begin{aligned} F_{ABCD} &= \sigma_{AB}^\mu \sigma_{CD}^\nu (A_{\nu,\mu} - A_{\mu,\nu}) \\ &= \sigma_{AB}^\mu \sigma_{CD}^\nu F_{\mu\nu} \end{aligned}$$

The Maxwell equations

$$F_{,\nu}^{\mu\nu} = J^\mu$$

are then expressed as

$$F_{,CD}^{ABCD} = J^{AB}$$

where

$$J^{AB} = \sigma_\mu^{AB} J^\mu,$$

This is okay provided that we assume a flat space-time background. If we assume a curved space-time background, we have to use the dyad formalism combined with the Newman-Penrose tetrad ie, generalized Pauli spin matrices. We then get

$$\sigma_{cd}^\mu \zeta_c^c \zeta_d^d F_{:\mu}^{ABCD} = 0$$

or equivalently,

$$\sigma_{cd}^\mu \zeta_c^c \zeta_d^d (F^{abpq} \zeta_a^A \zeta_b^B \zeta_p^C \zeta_q^D)_{:\mu} = 0$$

We then use the formulas for the spin coefficients expressed in terms of dyads combined with the fact that F^{abpq} is a scalar and hence its covariant derivative is an ordinary derivative. Finally, for expressing the Einstein field equations in spinorial form, we start with the metric tensor $g_{\mu\nu}$ expressed in spinor form using

$$g_{\mu\nu} \sigma_{AB}^\mu \sigma_{CD}^\nu = g_{ABCD}$$

or equivalently,

$$g_{\mu\nu} = \sigma_\mu^{AB} \sigma_\nu^{CD} g_{ABCD}$$

and then express the curvature tensor in spinor form in terms of g_{ABCD} . Note that the covariant derivative in gtr when acting on covariant vectors can be expressed in spinorial form as

$$\begin{aligned} \nabla_\mu A_\rho &= \partial_\mu A_\rho - \Gamma_{\mu\rho}^\nu A_\nu \\ &= \partial_\mu (A_{bc} \sigma_\rho^{bc}) - \Gamma_{\mu\rho}^\nu A_{bc} \sigma_\nu^{bc} \\ &= A_{bc} \sigma_{\rho,\mu}^{bc} + \sigma_\rho^{bc} A_{bc,\mu} \\ &\quad - \Gamma_{\mu\rho}^{bc} A_{bc} \end{aligned}$$

Finally,

$$\nabla_{pq} A_{rs} = \sigma_{pq}^\mu \sigma_{rs}^\rho \nabla_\mu A_\rho$$

Simplify this expression and transform this locally inertial/scalar spinor form of the covariant derivative into noninertial spinor form using dyads:

$$\nabla_{PQ} A_{RS} = \zeta_P^p \zeta_Q^q \zeta_R^r \zeta_S^s \nabla_{pq} A_{rs}$$

Continue this further to obtain the Riemann-Christoffel curvature tensor in spinor form remembering that

$$((\sigma_{ab}^\mu)) = \begin{pmatrix} l^\mu & m^\mu \\ \bar{m}^\mu & n^\mu \end{pmatrix}$$

and

$$\sigma_{ab}^\mu \zeta_A^a \zeta_B^b = \sigma_{AB}^\mu$$

where $((\zeta_A^a))$ is the inverse of $((\zeta_a^A))$.

3.29.4 Points to remember

[1] All the tensor equations in physics can be represented in spinor form. A prototype equation of this kind is Dirac's relativistic equation for the electron. Although this equation does not appear to be in tensor form, it is a tensor equation relative to a representation of the Lorentz group, ie, the Dirac representation. The algebra of this representation is the Clifford algebra generated by the Dirac $4 \times 4 \gamma$ -matrices.

[2] To transform tensors into spinors in curved space-time, we must construct general relativistic versions of the Pauli spin matrices. This is achieved using the Newman-Penrose tetrad. Further, just as the tetrad formalism in general relativity transforms four vectors into four scalars, we can define dyads that transform spinors in general relativity into scalars. This enables us to define the covariant derivative of spinors and hence formulate the Dirac equation in terms of covariant derivatives of two general relativistic spinors. The four generalized Pauli spin matrices are constructed from the Newman-Penrose tetrad and are used to transform four-vector fields into spinor fields and vice-versa. If this transformation is to be consistent, we are forced to conclude that the covariant derivative for spinors must be defined in such a way that the covariant derivative of the generalized Pauli spin matrix is zero.

[3] The advantage of using the spinor formulation of the laws of physics in place of the vector and tensor formulation is that since quadratic forms in the spinors after generalized Pauli transformation yield four vectors, namely the Newman-Penrose tetrad, it follows that sometimes nonlinear tensor equations in mathematical physics admit an elementary linear set of equations for the associated spinor field.

3.30 Prisms, mirrors and lenses, the general theory

A convex lens is the interior of the intersection of two solid spheres made of glass. We can use Snell's laws of reflection, compute the trajectory of a ray on the other side of the lens given an incident ray on the former side. More generally, consider a lens that is the intersection of the regions $\psi_1(x, y, z) > 0$ and $\psi_2(x, y, z) < 0$ with this intersection being filled with a material of refractive index n_2 , like say glass. The medium in the regions $I : \psi_1(x, y, z) < 0$ and $II : \psi_2(x, y, z) > 0$ is air of refractive index n_1 . Consider a ray of light defined by the equation $s \rightarrow \mathbf{R}_1(s) = (X_1(s), Y_1(s), Z_1(s))$ coming from the region I and hitting the surface $\psi_1(x, y, z) = 0$. The intersection of this ray with this first surface of the lens is obtained by solving the following equation for s :

$$\psi_1(X_1(s), Y_1(s), Z_1(s)) = 0$$

Let s_0 denote its solution. The unit outward normal to this surface at s_0 is given by

$$\hat{m}_1(s_0) = \nabla\psi_1(R_1(s_0))/|\nabla\psi_1(R_1(s_0))|,$$

Then, the angle of incidence is $\theta_i(s_0)$ given by

$$\cos(\theta_i(s_0)) = -(R'_1(s_0), \hat{m}_1(s_0))$$

Let $\theta_r(s_0)$ denote the corresponding angle of refraction. Then, by Snell's law,

$$\sin(\theta_i(s_0))/\sin(\theta_r(s_0)) = n_2/n_1$$

or equivalently,

$$\theta_r(s_0) = \sin^{-1}(n_1 \cdot \sin(\theta_i(s_0))/n_2)$$

The refracted ray falls in the plane containing $\hat{m}_1(s_0)$ and $R'_1(s_0)$. Thus a vector along the direction of the refracted ray is of the form

$$\hat{m}_2(s_0) = \alpha\hat{m}_1(s_0) + \beta R'_1(s_0)$$

Any point on the refracted ray is therefore of the form

$$\mathbf{R}_2(s) = R_1(s_0) + (s - s_0)\hat{m}_2(s_0)$$

Since this ray makes an angle $\theta_r(s_0)$ with the vector $-\hat{m}_1(s_0)$ at $R_1(s_0)$, we must have

$$\cos(\theta_r(s_0)) = -(\hat{m}_2(s_0), \hat{m}_1(s_0))$$

This gives us one equation for α, β . The other equation is obtained from the condition

$$1 = |\alpha\hat{m}_1(s_0) + \beta R'_1(s_0)|$$

Thus, $R_2(s)$ for arbitrary s can be computed easily. Now, the intersection point of this ray $R_2(s)$ with the second surface of the lens is obtained by solving

$$\psi_2(R_2(s)) = 0$$

and we denote by s'_0 the corresponding solution. Let θ'_i, θ'_r denote the corresponding angles of incidence and refraction when the ray travels from within the glass to the outside air. The unit vector along the outward normal to the second surface ψ_2 and s'_0 is given by

$$\hat{m}'_1(s'_0) = \nabla\psi_2(R_2(s'_0)) / |\nabla\psi_2(R_2(s'_0))|$$

and the angle of incidence at this point of the ray R_2 is given by

$$\cos(\theta'_i) = (\hat{m}'_1(s'_0), m_2(s_0))$$

and finally, θ'_r is given by

$$\sin(\theta'_i) / \sin(\theta'_r) = n_1 / n_2$$

Further the transmitted ray $R_3(s)$ has its unit tangent vector in the plane containing $\hat{m}'_1(s'_0)$ and $\hat{m}_2(s_0)$. Thus, we can write

$$R_3(s) = R_2(s'_0) + (s - s'_0)(\alpha' \hat{m}_2(s_0) + \beta' \hat{m}'_1(s'_0))$$

where one equation for α', β' is

$$|\alpha' \hat{m}_2(s_0) + \beta' \hat{m}'_1(s'_0)| = 1$$

and the other equation is

$$\cos(\theta'_r) = (\alpha' \hat{m}_2(s_0) + \beta' \hat{m}'_1(s'_0), \hat{m}'_1(s'_0))$$

Thus, we have completely determined the transmitted ray. We can then take an object, place it before such a lens and draw the transmitted ray for each incident ray coming from the object. From the set of transmitted rays, the image field can be calculated after applying an appropriate phase change in accordance with the length of the rays in the air medium and length of the rays in the glass medium.

All these calculations are based on the Fresnel theory since they take place at small distances from the object source. Consider for example a spherical lens of focal length f . The radius of curvature of the lens $R = 2f$. A point on the surface of the lens has coordinates $(X, Y, \sqrt{R^2 - X^2 - Y^2})$. consider now an plane object field on one side of the lens having amplitude distribution $A(X, Y)$ on the plane (X', Y', Z_0) where Z_0 is fixed. When a ray of light from this point on the object plane hits the lens, its complex amplitude on the lens surface becomes

$$A(X', Y') \exp(-jk\sqrt{(X - X')^2 + (Y - Y')^2 + (R - Z_0)^2})$$

approximately where we have approximated the z-coordinate $\sqrt{R^2 - X^2 - Y^2}$ on the lens surface by R since $|X|, |Y| \ll R$ provided that the focal length of the lens is sufficiently large. Thus, the amplitude on a screen located on the other side of the lens that coincides with the plane Z_1 is given by

$$\tilde{A}(x, y) = \int A(X', Y') \exp(-jk\sqrt{(X - X')^2 + (Y - Y')^2 + (R - Z_0)^2}).$$

$$\times \exp(-jk\sqrt{(x-X)^2 + (y-Y)^2 + (Z_1 - R)^2}) dX' dY' dXdY$$

Making the approximations

$$|X - X'|, |Y - Y'| \ll |R - Z_0|, |x - X|, |y - Y| \ll |Z_1 - R|$$

gives us the approximate Fresnel formula for the complex amplitude on the screen after omitting a constant phase factor:

$$\tilde{A}(x, y) = \int A(X', Y') \exp(-jk((X - X')^2 + (Y - Y')^2)/2|R - Z_0|).$$

$$\times \exp(-jk((x - X)^2 + (y - Y)^2)/2|Z_1 - R|) dXdY dX' dY'$$

Writing $|R - Z_0| = R_0, |Z_1 - R| = R_1$, we get

$$\tilde{A}(x, y) = \int A(X', Y') \exp(-jk(X^2 + Y^2)/2R_0) \exp(-jk((x + X - X')^2 + (y + Y - Y')^2)/2R_1) dXdY dX' dY'$$

$$= \int A(X', Y') \exp(-jk(X^2(1/2R_0 + 1/2R_1) + Y^2(1/2R_0 + 1/2R_1))) \exp(-jk((x - X')^2 + (y - Y')^2)/2R_1)$$

$$\times \exp(-jk(x(X - X') + y(Y - Y'))/2R_1) dx dY dX' dY'$$

We perform first the integration over X, Y , the points on the lens's surface and then neglect this unimportant phase term. The result is

$$\tilde{A}(x, y) = \int A(X', Y') \exp(-jk((x - X')^2 + (y - Y')^2)/2R_1) dX' dY'$$

and if we assume that the object span is small, then $X'^2 + Y'^2 \ll R_1^2$ and we get after approximately the following expression for the amplitude on the screen:

$$\tilde{A}(x, y) = \exp(-jk(x^2 + y^2)/2R_1) \int A(X', Y') \exp(jk(xX' + yY')/R_1) dX' dY'$$

This shows that apart from a non-constant phase factor, the image produced by a lens on a screen is the spatial Fourier transform of the object amplitude distribution.

3.30.1 Exercises

- [1] Consider a paraboloid as a surface of revolution:

$$x^2 + y^2 = 4az$$

Assume that a light ray parallel to the z axis coming from $z = +\infty$ hits the inner parabolic surface at $(x_0, y_0, z_0 = (x_0^2 + y_0^2)/4a)$. Calculate the equation of the reflected ray and show that it passes through the focus of the parabola at

$(0, 0, a)$. Show also that the equation of a parabola $y^2 = 4ax$ in the xy plane can be obtained as the locus of points the ratio of whose perpendicular distance from a vertical line to that from a given point is a constant.

[2] A convex lens is constructed as the region of intersection of the two spheres $x^2 + y^2 + z^2 \leq a^2$ and $(x - b)^2 + y^2 + z^2 \leq c^2$. A ray of light hits the second convex surface at the point $P : (x_0, y_0, z_0) = \sqrt{c^2 - (x_0 - b)^2 - y_0^2}$ from the air side and this ray has direction cosines (n_x, n_y, n_z) . It then gets refracted into the lens glass. Assume that this refracted ray hits the inner first surface at Q . Determine the coordinates of Q and hence the equation of the refracted ray.

[3] Assuming that a ray parallel to the axis of a lens gets refracted on the other side so that it passes through the focus at $(f, 0)$. Also assume that a ray passing through the centre of the lens passes straight through on the other side without getting bent. Now, take an object like a stick of length d placed vertically at $(-u, 0)$. Consider the rays coming from this object from the points $(-u, d)$ one passing through the centre of the lens and the other going parallel to the lens and getting refracted to pass through the focus at $(f, 0)$. By considering the point of intersection of these two rays, show that the image of the stick will be formed at $x = v$ where $1/u + 1/v = 1/f$. Also determine the magnification of the image.

[4] Derive the formula $1/u - 1/v = 1/f$ for a convex mirror where the negative sign before $1/v$ stands for the image being virtual and also derive the formula $1/u + 1/v = 1/f$ for a concave mirror.

[5] Draw the ray diagram of a telescope consisting of two convex lenses the objective lens being of focal length f_1 and the eyepiece of focal length f_2 where $f_2 \ll f_1$. Assuming that these two lenses are separated by a distance d determine a formula for the magnification.

[6] Explain why the focal length of a concave or convex lens or mirror must be half its radius of curvature.

[7] Why does the frequency of light not get altered when it passes from a medium of one refractive index to a medium of another refractive index while the wavelength gets altered. Explain this using Maxwell's equations. Explain in this context Newton's experiment of the splitting of white light into several colours when it passes through a prism. Use the fact that the refractive index of a medium is frequency dependent. Explain from basic physics of an electron bound to the nucleus moving in an externally applied electric field by calculating the resulting dipole moment of the electron-nucleus pair per unit volume.

[8] By considering the Hamiltonian of an electron bound to the nucleus by a radial potential and getting perturbed by an external electromagnetic field, explain using second order perturbation theory with the electronic charge as the perturbation parameter, how the wave function of the electron gets altered after it interacts with the em field. Assuming an initial Gibbs state for the atom in terms of the temperature and the stationary states with their corresponding energy levels, determine the mixed state after time t under the interaction. Explain how one can using this formula, calculate the average electric and magnetic dipole moments of the electron after time t in terms of the electric and magnetic vector potentials. Thereby develop a theory which explains the inho-

mogeneity, anisotropy and field dependence of the permittivity and permeability of a medium comprising of several atoms each with one free electron.

[9] Gravitational lensing: Using the Schwarzschild metric for a spherical mass distribution with mass M and radius R , calculate the equation of the trajectory of a light ray coming from a point object behind the spherical mass and gets bent by the gravitational field of the mass. Deduce that by extrapolating the bent rays coming from both the sides of the mass that two images of the point object will be formed for an observer looking from the opposite side and evaluate the angle between the two images as seen by the observer.

hint: Assuming the spherical mass to have its centre located at the origin and the rays of light propagating in the xy plane, ie, $\theta = \pi/2$, the Schwarzschild metric is given by

$$d\tau^2 = \alpha(r)dt^2 - \alpha(r)^{-1}dr^2/c^2 - r^2d\phi^2/c^2, \alpha(r) = 1 - 2GM/rc^2$$

Light rays travel along null geodesics $d\tau^2 = 0$ and hence

$$\alpha(r) - \alpha(r)^{-1}r'^2/c^2 - r^2\phi'^2/c^2 = 0 \quad \dots \quad (1)$$

where

$$\xi' = d\xi/dt$$

Further, the Euler Lagrange geodesic equations obtained from the action principle $\delta \int (d\tau/d\lambda)d\lambda = 0$ give us

$$\alpha(r)dt/d\tau = K, r^2d\phi/d\tau = L$$

where K, L are infinite constants but $L/K = \beta$ is a finite constant. Thus,

$$r^2\phi' = \beta\alpha(r) \quad \dots \quad (2)$$

So finally, we get from (1) and (2),

$$r^4/\beta^2\alpha(r) - \alpha(r)^{-1}(dr/d\phi)^2/c^2 - r^2/c^2 = 0$$

Replace $r = 1/u$ and derive the equation

$$1/\beta^2\alpha(1/u) - \alpha(1/u)^{-1}(du/d\phi)^2/c^2 - u^2/c^2 = 0$$

Rearrange this equation to get

$$(du/d\phi)^2 + u^2 - 1/\beta^2 - 2GMu^3/c^2 = 0$$

Now recognize the last term as the general relativistic correction term and apply perturbation theory to get the approximate trajectory of the light ray. Note that on differentiating this equation w.r.t ϕ , we get

$$u''(\phi) + u(\phi) = 3GMu^2/c^2$$

If the general relativistic correction term on the rhs is not present, then we have a straight line solution in the form $u(\phi) = A\cos(\phi)$. Apply perturbation theory to this unperturbed solution and hence calculate the angle between the two images of the point.

3.30.2 Points to remember

[1] The properties of a lens can all be derived from Snell's law of refraction, ie, by applying this principle repeatedly, the path of a light ray falling on one surface of a lens/prism gets refracted and again gets refracted at the other lens surface and by calculating the intersection of two emerging rays coming from some point on the object, the position of the image of that point can be determined. It should be noted that the refractive index of glass is frequency dependent and hence rays of different frequencies get bent by different angles according to Snell's law. This is the basis of Newton's fundamental experiment showing that white light is a mixture em waves of all frequencies. Further, for a convex lens or a concave lens in which the curvature of the surface is a constant the relation $1/u \pm 1/v = 1/f$ holds where u is the distance of the object from the lens, v is the distance of the image from the lens and f , the focal length is one half the lens' radius of curvature. The plus or minus sign is chosen according as the image is real or virtual. The same result is valid for convex and concave mirrors. It can be shown by applying Snell's law of refraction that a ray parallel to the axis of a convex lens bends on the other side so as to pass through its focus, ie, a point on the axis at a distance f from its centre.

[2] By using Fresnel's diffraction theory, it is not hard to show that if light from a planar object placed parallel to convex lens falls on the lens, then the amplitude pattern of the image formed on a flat screen placed on the other side of the lens parallel to it is the two dimensional spatial Fourier transform of the object amplitude.

[3] If the objective lens of a telescope has a large focal length f_o and the eyepiece of the telescope has a small focal length f_e , then the telescope magnifies far field objects. The magnification is given by the ratio f_o/f_e as can be readily shown by drawing the ray diagram, applying the equation $1/u + 1/v = 1/f$ and using properties of similar triangles.

[4] It is instructive to draw the ray diagram for convex and concave lenses and mirrors, and determine conditions under which real and virtual images are formed by these and also the associated magnification for different values of u and a given f . For example, we may explain using $1/u + 1/v = 1/f$ for a convex lens that if $u < f$, then $v < 0$, ie, the image is virtual while if $u > f$, then $v > 0$ and the image is real. Likewise, for a convex mirror, if $u > f$, we have $v < 0$ using $1/u - 1/v = 1/f$. This is why the image of traffic seen in a convex mirror in a car is virtual and diminished in size.

[5] According to Einstein's general theory of relativity, gravity bends light and the deflection of a light ray passing in the vicinity of a spherical gravitating object can be calculated by using the Schwarzschild metric of a spherical distribution of matter. Thus, if a star is located behind a spherical mass, its light ray will bend around the matter sphere, causing an observer looking from the front to see two images of the star separated by an angle of $4GM/Rc^2$ where M, R are respectively the mass and radius of the matter sphere. This calculation can be performed by writing down the geodesic equation of a particle in the $\theta = \pi/2$ plane of the Schwarzschild metric and imposing the null condition

to derive the trajectory equation $\phi \rightarrow r(\phi)$ of the light ray. The result is the standard Newtonian result for ordinary massive particles with the presence of a small general relativistic correction term.

3.31 A brief summary of the book

Where do wave equations arise, how to approximate highly nonlinear problems in mechanics, fluid dynamics, electromagnetism, plasma physics, general relativity and quantum mechanics by elementary linear wave equations and to develop techniques for solving them like separation of variables, Fourier series and transforms, how to simulate wave equations using computer software by discretizing the spatial and temporal domains.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Chapter 4

Probability Theory and Statistics required for random wave motion analysis

4.1 Summary of Contents

1. Probability spaces.
2. Random variables
3. Lebesgue integral w.r.t. a measure and w.r.t. a probability measure
4. Cornerstone theorems of Lebesgue's theory of integration: Monotone convergence, Fatou's lemma and dominated convergence.
5. Moments of a random variable, mean, variance, skewness and higher moments.
6. The direct product of a finite number of measures and Fubini's theorem on change in the order of integration.
7. Karl Pearson's correlation coefficient.
8. Absolute continuity of measures, the Radon-Nikodym theorem and probability distributions and densities.
9. Conditional expectation and conditional probability.
10. Quantum probability space: Events, observables and states in a quantum probability space.
11. Joint probability distribution of classical random variables.
12. Heisenberg's uncertainty principle for non-commuting observables and the impossibility of defining joint probabilities for non-commuting observables in the quantum theory.
13. Positive definitivity of the joint characteristic function of classical random variables.
14. Bochner's theorem in classical probability.
15. Non-positive definitivity of joint characteristic functions of non-commuting observables in the quantum theory.
16. The inequalities of John Bell and the consequent impossibility of constructing a hidden variable theory for quantum mechanics.

17. Classical stochastic processes: Brownian motion, Poisson process and more generally Levy processes.
18. Ito's formula for Brownian motion and Poisson processes.
19. Quantum Brownian motion and quantum Poisson processes: Noncommutative generalizations of classical BM and classical PP.
20. Realizing classical stochastic processes using families of operators in Boson Fock space in a coherent state.
21. Bernoulli random variables in classical and quantum probability.
22. Stochastic differential equations driven by Brownian motion: Existence and Uniqueness.
23. Stochastic differential equations driven by Levy processes.
24. The generalized Ito-Doleans Dade-Meyer differential rule for discontinuous Martingales.
25. Girsanov's formula for the change of drift in an sde.
26. Quantum stochastic differential equations in the sense of Hudson and Parthasarathy; Existence and Uniqueness of solutions.
27. Some practical applications of quantum probability.
 - [a] Computation of scattering cross sections.
 - [b] The wave function of our universe: Wheeler-Dewitt-Hartle-Hawking's theory. The Hartle-Hawking theory of the probability distribution of the radius of our universe with a scalar field coupling.
 - [c] Computation of tunneling probabilities of a quantum particle through a potential barrier with applications to tunneling diode.
 - [d] Probability of induced transitions in a laser; The interaction of Light with atoms—The Glauber-Sudarshan theory.
 - [e] Energy bands in a semiconductor—the theory based on Bloch wave functions; Specialization to the Kronig-Penney model.
 - [f] Modeling quantum systems coupled to a noisy bath—The Hudson-Parthasarathy equation and its solution using Maasen's kernel approach and using the functional form of the Glauber-Sudarshan representation.

4.2 Probability spaces and measure theoretic theorems on probability spaces

A triplet (Ω, \mathcal{F}, P) where Ω is called the sample space of the experiment, \mathcal{F} is the σ -field of events and P is a probability measure.

2. Random variables
3. Lebesgue integral w.r.t. a measure and w.r.t. a probability measure
4. Cornerstone theorems of Lebesgue's theory of integration: Monotone convergence, Fatou's lemma and dominated convergence.
5. Moments of a random variable, mean, variance, skewness and higher moments.
6. The direct product of a finite number of measures and Fubini's theorem on change in the order of integration.
7. Karl Pearson's correlation coefficient.
8. Absolute continuity of measures, the Radon-Nikodym theorem and probability distributions and densities.
9. Conditional expectation and conditional probability.

4.3 Basic facts about quantum probability

10. Quantum probability space: Events, observables and states in a quantum probability space.
11. Joint probability distribution of classical random variables.
12. Heisenberg's uncertainty principle for non-commuting observables and the impossibility of defining joint probabilities for non-commuting observables in the quantum theory.
13. Positive definitivity of the joint characteristic function of classical random variables.
14. Bochner's theorem in classical probability.
15. Non-positive definitivity of joint characteristic functions of non-commuting observables in the quantum theory.
16. The inequalities of John Bell and the consequent impossibility of constructing a hidden variable theory for quantum mechanics.

Given three random variables X_1, X_2, X_3 , each assuming values ± 1 only, we observe that

$$X_3(X_1 - X_2) \leq 1 - X_1 X_2$$

Taking expectations gives

$$\mathbb{E}(X_1 X_3) - \mathbb{E}(X_2 X_3) \leq 1 - \mathbb{E}(X_1 X_2)$$

Interchanging X_1 and X_2 gives

$$\mathbb{E}(X_2 X_3) - \mathbb{E}(X_1 X_3) \leq 1 - \mathbb{E}(X_1 X_2)$$

Thus,

$$|\mathbb{E}(X_1 X_3) - \mathbb{E}(X_2 X_3)| \leq 1 - \mathbb{E}(X_1 X_2)$$

This is called Bell's inequality for Bernoulli random variables. It is violated in certain cases for quantum Bernoulli random variables. For example, consider $\sigma_k, k = 1, 2, 3$ ie, the Pauli spin matrices. These matrices have eigenvalues ± 1 but do not commute. Choose unit directions $\hat{n}_1, \hat{n}_2, \hat{n}_3$ and consider the observables

$$X_k = (\hat{n}_k, \sigma) = n_{k1}\sigma_1 + n_{k2}\sigma_2 + n_{k3}\sigma_3, k = 1, 2, 3$$

Then,

$$X_k X_m = (\hat{n}_k, \hat{n}_m) + i(\sigma, \hat{n}_k \times \hat{n}_m)$$

Now, we can choose the state

$$\rho = (1/2)I_2$$

Then,

$$Tr(\rho \sigma_k) = 0, k = 1, 2, 3$$

and we have

$$Tr(\rho X_k X_m) = (\hat{n}_k, \hat{n}_m), k, m = 1, 2, 3$$

Let θ_{km} be the angle between \hat{n}_k and \hat{n}_m . Then to verify Bell's inequality, for the quantum observables X_1, X_2, X_3 , we require that

$$|\cos(\theta_{13}) - \cos(\theta_{23})| \leq 1 - \cos(\theta_{12})$$

for all points on the unit sphere. Now, suppose we choose \hat{n}_3 as the north pole and $|\theta_1 - \theta_2| < \epsilon$. Then by an appropriate choice of ϵ , Bell's inequality will be violated.

4.4 Some basic classical and quantum stochastic processes

17. Classical stochastic processes: Brownian motion, Poisson process and more generally Levy processes.

18. Ito's formula for Brownian motion and Poisson processes.

19. Quantum Brownian motion and quantum Poisson processes: Noncommutative generalizations of classical BM and classical PP.

20. Realizing classical stochastic processes using families of operators in Boson Fock space in a coherent state.

21. Bernoulli random variables in classical and quantum probability.

22. Stochastic differential equations driven by Brownian motion: Existence and Uniqueness.

23. Stochastic differential equations driven by Levy processes.

24. The generalized Ito-Doleans Dade-Meyer differential rule for discontinuous Martingales.

Kolmogorov's Martingale inequality, Doob's Martingale inequality, convergence theory of Martingales based on the down-crossing theorem, construction of the stochastic integral w.r.t L^2 Martingales. Exponential Martingales and the Martingale version of the change of measure theorem.

25. Girsanov's formula for the change of drift in an sde.

Consider two diffusion processes $x_1(t), x_2(t)$ with same diffusion coefficient $\sigma(x)$ but different drift coefficients $b_1(x), b_2(x)$ respectively. Thus, they satisfy the sde's

$$dx_k(t) = b_k(x(t))dt + \sigma(x(t))dB(t), k = 1, 2$$

Let $P_k, k = 1, 2$ denote the probability measures induced by them on the spaces $C[0, T]$ starting at a fixed point $x(0)$. Then, we have from intuitive considerations, after defining $a(x) = \sigma(x)\sigma(x)^T, b(x) = b_2(x) - b_1(x)$,

$$\begin{aligned} R(x) &= \frac{dP_2}{dP_1}(x) = \exp\left((-1/2) \int_0^T [(dx(t) - b_2(x(t))dt)^T (a(x(t))dt)^{-1} (dx(t) - b_2(x(t))dt) \right. \\ &\quad \left. - (dx(t) - b_1(x(t))dt)^T (a(x(t))dt)^{-1} (dx(t) - b_1(x(t))dt)]\right) \\ &= \exp\left(\int_0^T [dx(t)^T a(x(t))^{-1} b(x(t)) - (1/2)(b_2(x(t))^T a(x(t))^{-1} b_2(x(t)) - b_1(x(t))^T a(x(t))^{-1} b_1(x(t))dt]\right) \end{aligned}$$

It should be noted that if x is a diffusion process with drift-diffusion coefficient pair $(b_1(x), \sigma(x))$ w.r.t the probability measure P_1 , then x is a diffusoin process with drift-diffusion coefficient pair $(b_2(x), \sigma(x))$ w.r.t the probability measure $R.P_1 = P_2$. In particular, suppose $\sigma(x) = I$ and $b_1(x) = 0$. Then, P_1 is the Wiener measure, $b_2(x) = b(x)$ and writing $B(t) = x(t)$, we have that $B(t)$ is a Brownian motion process w.r.t P_1 while

$$R(B) = \exp\left(\int_0^T b(B(t))^T dB(t) - (1/2) \int_0^T \|b(B(t))\|^2 dt\right)$$

with $B(t)$ being a Brownian motion with drift $b(.)$ w.r.t $P_2 = R.P_1$, ie, $B(t) - \int_0^t b(B(s))ds$ is a Brownian motion process w.r.t $R.P_1$. This is Girsanov's formula for diffusion process.

4.5 Some applications of classical probability to engineering systems

26. Disturbance observer models and disturbance rejection methods.

Abstract: Here, we describe the general kind of disturbance observer used in nonlinear systems and more specifically in robotics. The idea in the design of the disturbance observer is that its governing differential equations should be based on only the instantaneous angular position and velocity of the robot or equivalently only on the current state of the nonlinear system governed by the state variable dynamics. It should not involve the acceleration of the robot or equivalently, the derivative of the state of the nonlinear system. Without taking into account random noise in the state dynamics, we evaluate the rate of change of the Lyapunov energy of the disturbance estimation error and derive conditions under which this error converges to zero as time goes to ∞ under the condition that in this asymptotic limit, the disturbance converges to a constant dc level. After taking random noise into account, we evaluate the performance in the sense of mean value of the rate of change of the Lyapunov energy and study conditions under which this value is asymptotically bounded.

Consider the robot system

$$M(q)q'' + N(q, q') = \tau(t) + d(t) + w(t)$$

where q is the angular position vector of the links, τ is the external torque, d is the disturbance and w is WGN. Consider the equations

$$\hat{d}(t) = z(t) + p(q'(t)),$$

$$z'(t) = L(q(t), q'(t))(\tau(t) - N(q(t), q'(t)) - \hat{d}(t))$$

Then, we have

$$\hat{d}'(t) = L(N - \tau - \hat{d}) + p'(q')q''$$

$$= L(N - \tau - \hat{d}) + p'(q')M(q)^{-1}(\tau - N + d + w)$$

$$L = p'(q')M(q)^{-1}$$

we get

$$\hat{d}'(t) = L(d - \hat{d} + w)$$

Assume that

$$w_1 = d - \hat{d}$$

is WGN. Then writing $\epsilon = w_1 + w$, it follows that ϵ is also WGN. So our system of equations for the state and disturbance observer after disturbance rejection becomes

$$\begin{aligned} M(q)q'' + N(q, q') &= \tau + d - \hat{d} + w \\ &= \tau + \epsilon, \\ \hat{d}' &= L(q, q')\epsilon \end{aligned}$$

and the EKF can be applied to this system. More generally, consider a system of the form

Other models for disturbance rejection in other kinds of systems.

$$x'(t) = f(t, x(t)) + d(t) + w(t)$$

where $x(t), d(t), w(t) \in \mathbb{R}^n$. The explicit dependence of f on time arises because of a known input. d is the disturbance and w is WGN. We design the disturbance observer as follows:

$$\hat{d}(t) = z(t) + p(x(t))$$

where

$$z'(t) = -L(t, x(t))(f(t, x(t)) + \hat{d}(t))$$

Then, we get

$$\hat{d}'(t) = p'(x(t))((f(t, x(t)) + d(t) + w(t)) - L(t, x(t))(f(t, x(t)) + \hat{d}(t)))$$

Assume that

$$L(t, x(t)) = p'(x(t))$$

Then,

$$\hat{d}'(t) = p'(x(t))(d(t) - \hat{d}(t) + w(t))$$

which is appropriate for a disturbance observer provided that $p'(x(t))$ is positive definite.

Lyapunov energy based disturbance observer: Let $w(t) = 0$ and define

$$V(t) = (1/2)(d(t) - \hat{d}(t))^T J(x(t))(d(t) - \hat{d}(t))$$

Then,

$$V'(t) = -(1/2)(d(t) - \hat{d}(t))^T [K(x(t))^T J(x(t)) + J(x(t))K(x(t)) - \sum_i J_{,i}(x(t))x'_i(t)](d(t) - \hat{d}(t))$$

where

$$K(x) = p'(x)$$

and we neglect $d'(t)$, ie, at large times, the disturbance is nearly a dc disturbance. Let $\max_{i,t} x'_i(t) = v$ and $\min_{i,t} x'_i(t) = u$. Then, if we can ensure that $J(x)$ is positive definite and simultaneously

$$K(x)^T J(x) + J(x) K(x) - \sum_i v_i J_{,i}(x)$$

is positive definite where each v_i is either v or u , then our disturbance observer will converge to the true disturbance. More precisely, for the rate of change of the Lyapunov energy to be negative, it is sufficient that

$$K(x)^T J(x) + J(x) K(x) - v_0 (\sum_i \|J_{,i}(x)\|) I \geq 0$$

where $\|\cdot\|$ is the spectral norm of matrices and $v_0 = \max_{i,t} |x_i(t)|$. Taking random noise into account:

$$\begin{aligned} V'(t) = & -(1/2)(d(t) - \hat{d}(t) + w(t))^T [K(x(t))^T J(x(t)) + J(x(t)) K(x(t))] \\ & - \sum_i J_{,i}(x(t)) x'_i(t) (d(t) - \hat{d}(t) + w(t)) \end{aligned}$$

and hence, if we assume that $d - \hat{d} = w_1$ is WGN and define $\epsilon = w_+ w_1$.

Conclusions: We have based on the Lyapunov energy method designed a general kind of disturbance observer based on the instantaneous state of a nonlinear dynamical system and have applied to the robot problem. The disturbance observer guarantees that in the absence of noise, and the condition that the disturbance converges to a constant dc value, the disturbance error will converge to zero. In the presence of random noise, we need to make an analysis of the ensemble averaged Lyapunov energy rate of change. That will be the subject of a future paper.

4.6 Quantum stochastic differential equations

27. Quantum stochastic differential equations in the sense of Hudson and Parthasarathy; Existence and Uniqueness of solutions.

$$j_t(X) = X + \int_0^t j_s(\theta_b^a(X)) d\Lambda_a^b(s) ds$$

To solve this, we construct the iterative scheme:

$$j_t^{(n+1)}(X) = X + \int_0^t j_s^{(n)}(\theta_b^a(X)) d\Lambda_a^b(s), n \geq 0$$

Then

$$\begin{aligned} & \langle f \otimes \phi(u) | j_t^{(n+1)}(X) - j_t^{(n)}(X) | f \otimes \phi(u) \rangle \\ &= \int_0^t \langle f \otimes \phi(u) | (j_s^{(n)}(\theta_b^a(X)) - j_s^{(n-1)}(\theta_b^a(X))) | f \otimes \phi(u) \rangle u^b(s) \bar{u}_a(s) ds \end{aligned}$$

from which the desired convergence results are derived. Specifically, we define

$$\Delta_{n+1}(t) = \langle f \otimes \phi(u) | j_t^{(n+1)}(X) - j_t^{(n)}(X) | f \otimes \phi(u) \rangle$$

Then, we get

$$\Delta_{n+1}(t) \leq \int_0^t \Delta_n(s) \| u(s) \|^2 ds$$

which by application of Gronwall's inequality yields

$$\Delta_{n+1}(t) \leq C \left(\int_0^t \| u(s) \|^2 ds \right)^n / n!$$

from which existence of a solution to the qsde is easily established.

4.7 Some practical applications of quantum probability

4.7.1 Computation of scattering cross sections

A, B are two Hamiltonians, $B=A+V$. The wave operators are

$$\begin{aligned} \Omega_+ &= \lim_{t \rightarrow \infty} \exp(itB) \cdot \exp(-itA), \\ \Omega_- &= \lim_{t \rightarrow -\infty} \exp(itB) \cdot \exp(-itA) \end{aligned}$$

The scattering matrix is

$$S = \Omega_+^* \Omega_-$$

Let $|\phi_i\rangle$ be the input free state at energy E . Then the output free state at the same energy E is $|\phi_o\rangle = S(E)|\phi_i\rangle$. The probability of scattering within a cone \mathcal{C} is given by

$$\| \chi_C S(E) \phi_i \|^2$$

Further, the probability current density after scattering is given by

$$J = (ie/2m)(\phi_o \nabla \phi_o^* - \phi_o^* \nabla \phi_o)$$

and we can compute using this, the probability that per unit time particles scattered into a solid angle Ω_0 at infinite radius as

$$\lim_{r \rightarrow \infty} \int_{\Omega_0} J(r) \cdot \hat{r} \cdot r^2 d\Omega(\hat{r})$$

from which the scattering cross section is easily evaluated by taking the incident state ϕ_i as corresponding to a single particle moving per unit area per unit time at a perpendicular distance between $[b, b + db]$ from the scattering centre. In this analysis, we take $A = -h^2 \nabla^2 / 2m$ and $V = V(r)$ to be the potential of the scattering centre. The standard reference for such problems is Werner Amrein, "Hilbert space methods in quantum mechanics", CRC press.

4.7.2 The wave function of our universe

Wheeler-Dewitt-Hartle-Hawking's theory. The Hartle-Hawking theory of the probability distribution of the radius of our universe with a scalar field coupling.

4.7.3 Computation of tunneling probabilities of a quantum particle through a potential barrier with applications to tunneling diode

Computations based on quasi-classical quantum mechanics. This analysis is also applicable to some quantum field theoretic problems like the quantized KG field perturbed by a Higgs potential. Such a problem in a finite region, say a finite box can be regarded using standard Fourier series analysis as a infinite sequence of 3-D harmonic oscillators with mutual interaction generated by the Higgs potential.

4.7.4 Probability of induced transitions in a laser

The interaction of Light with atoms—The Glauber-Sudarshan theory.

4.7.5 Energy bands in a semiconductor

—the theory based on Bloch wave functions; Specialization to the Kronig-Penney model.

4.7.6 Modeling quantum systems coupled to a noisy bath

—The Hudson-Parthasarathy equation and its solution using Maasen's kernel approach and using the functional form of the Glauber-Sudarshan representation.

$$dU(t) = (-(iH + P)dt + L_1 dA + L_2 dA^* + S d\Lambda)U(t) = dW(t).U(t)$$

The joint state of system and bath is

$$\rho(t) = U(t)\rho(0)U(t)^*$$

which satisfies

$$\begin{aligned} d\rho(t) &= dU(t)\rho(0)U(t)^* + U(t)\rho(0)dU(t)^* + dU(t)\rho(0)dU(t)^* \\ &= dW(t)\rho(t) + \rho(t)dW(t)^* + L_1\rho(t).L_1^*dt + Sd\Lambda\rho(t)S^* + S\rho(t)dA^*L_1^* \\ &\quad + L_1dA\rho(t)S^* \end{aligned}$$

Note that

$$\begin{aligned} dW.\rho + \rho.dW^* &= \\ -i[H, \rho]dt - \{P, \rho\}dt + L_1dA\rho + L_2dA^*\rho + \rho.dA^*L_1^* + \rho.dA.L_2^* + (Sd\Lambda\rho + d\Lambda\rho S^*) \end{aligned}$$

To solve this Schrodinger equation for the joint state ρ of system and bath, we assume

$$\rho(t) = \int F(t, u, \bar{u}) \otimes |e(u)\rangle \langle e(u)| du.d\bar{u}$$

where

$$du = \Pi_{t \geq 0} du(t), d\bar{u} = \Pi_{t \geq 0} d\bar{u}(t)$$

are path measures. Note that $u \in L^2(\mathbb{R}_+)$ which is isomorphic to $l^2(\mathbb{Z}_+)$. We observe that $\rho(t)$ by definition commutes with $dA(t)$, $dA(t)^*$ and $d\Lambda(t)$ provided that we assume that the initial bath state commutes with these operators and this along with quantum Ito's formula has been used in the derivation of the above relations. If we do not make such an assumption, then we would get

$$\begin{aligned} d\rho(t) &= dW.\rho + \rho.dW^* - i[H, \rho]dt - \{P, \rho\}dt + L_1dA.\rho.dA^*L_1^* + L_1dA.\rho.dA.L_2^* \\ &\quad + L_1dA.\rho.d\Lambda.S^* + L_2dA^*.\rho.dA^*.L_1^* + L_2dA^*.\rho.dA.L_2^* + L_2dA^*.\rho.d\Lambda.S^* \\ &\quad + Sd\Lambda.\rho.dA^*L_1^* + Sd\Lambda.\rho.dA.L_2^* + Sd\Lambda.\rho.d\Lambda.S^* \end{aligned}$$

Now,

$$dA\rho.dA^* = \int F(t, u, \bar{u})dA(t)|e(u)\rangle \langle e(u)|dA(t)^*dud\bar{u} = 0(dt^2)$$

and hence this term can be neglected. Note that

$$dA(t)|e(u)\rangle \langle e(u)|dA(t)^* = |u(t)|^2 dt^2 |e(u)\rangle \langle e(u)|$$

Further,

$$dA(t)|e(u)\rangle \langle e(u)| = u(t)dt.|e(u)\rangle \langle e(u)|$$

$$|e(u)\rangle \langle e(u)|dA(t)^* = \bar{u}(t)dt|e(u)\rangle \langle e(u)|$$

$$\begin{aligned} dA(t)^*|e(u)\rangle \langle e(u)| &= \frac{d}{d\epsilon}|e(u + \epsilon\chi_{[t, t+dt]})\rangle \langle e(u)||_{\epsilon=0} \\ &= \chi_{[t, t+dt]}\frac{\delta}{\delta u(t)}|e(u)\rangle \langle e(u)| \end{aligned}$$

$$dA(t)^*|e(u)\rangle \langle e(u)|dA(t) = \chi_{[t, t+dt]}\frac{\delta^2}{\delta u(t)\delta\bar{u}(t)}|e(u)\rangle \langle e(u)|$$

$$dA(t)|e(u)\rangle \langle e(u)|dA(t) = u(t)dt.\chi_{[t, t+dt]}\frac{\delta}{\delta\bar{u}(t)}(|e(u)\rangle \langle e(u)|)$$

which can be neglected.

4.8 Casting the HP equation in functional derivative form

$$dU(t) = (-(iH + P)dt + L_1 dA + L_2 dA^* + S d\Lambda)U(t)$$

Let

$$U(t) = \int F(t, u, \bar{u}) \otimes |e(u) > < e(u)| du d\bar{u}$$

So,

$$\begin{aligned} L_1 dA(t)U(t) &= \int (L_1 F(t, u, \bar{u})) \otimes dA(t) |e(u) > < e(u)| du d\bar{u} \\ &= dt \int u(t) L_1 F(t, u, \bar{u}) \otimes |e(u) > < e(u)| du d\bar{u} \\ L_2 dA(t)^* &= \int L_2 F(t, u, \bar{u}) \otimes dA(t)^* |e(u) > < e(u)| du d\bar{u} \\ &= dt \int L_2 F(t, u, \bar{u}) \otimes \frac{\delta}{\delta u(t)} |e(u) > < e(u)| du d\bar{u} \\ &= -dt \int L_2 \frac{\delta F(t, u, \bar{u})}{\delta u(t)} |e(u) > < e(u)| du d\bar{u} \end{aligned}$$

Also

$$\begin{aligned} S d\Lambda U(t) &= \int SF(t, u, \bar{u}) \otimes d\Lambda |e(u) > < e(u)| du d\bar{u} \\ &\quad \int SF(t, u, \bar{u}) \otimes dt^{-1} dA^* dA |e(u) > < e(u)| du d\bar{u} \\ &= \int u(t) SF(t, u, \bar{u}) \otimes dA^* |e(u) > < e(u)| du d\bar{u} \\ &= dt \int u(t) SF(t, u, \bar{u}) \otimes \frac{\delta}{\delta u(t)} |e(u) > < e(u)| du d\bar{u} \\ &= -dt \int \frac{\delta}{\delta u(t)} (u(t) F(t, u, \bar{u})) \otimes |e(u) > < e(u)| du d\bar{u} \end{aligned}$$

Thus, the HP Schrodinger equation translates to the following functional differential equation:

$$\begin{aligned} \frac{\partial F(t, u, \bar{u})}{\partial t} &= (-iH + P)F(t, u, \bar{u}) \\ &\quad + u(t)L_1 F(t, u, \bar{u}) - L_2 \frac{\delta F(t, u, \bar{u})}{\delta u(t)} - S \frac{\delta(u(t)F(t, u, \bar{u}))}{\delta u(t)} \end{aligned}$$



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Chapter 5

An introduction to probability and random processes in circuit theory from a pedagogical viewpoint

5.1 Circuit theory concepts from field theory concepts

In order to learn the fundamentals of electrical and electronics engineering, the student just when he enters into college should first brush up his fundamentals of electromagnetic field theory, the reason being that electric and magnetic fields form the fundamental building blocks of this branch of engineering. The voltage difference between two terminals on a bread -board on which a circuit has been built is simply the line integral of the electric field along any path between the two points. The current flowing through a wire having a small cylindrical cross section is the surface integral of the current density over a cross section of the wire and this current density J can be calculated in principle once the electric and magnetic fields are known using the Maxwell equation

$$\text{curl}H = J + \epsilon \frac{\partial E}{\partial t} \quad \dots \quad (1)$$

Kirchhoff's current law (KCL) states that the sum of currents emanating from a given node in a circuit is zero. This is provided that the node cannot accumulate charges. In the language of field theory, the KCL is expressed by the restricted

charge conservation equation

$$\operatorname{div} J = 0 \quad \dots \quad (2)$$

which when integrated over a closed surface containing the node yields after applying the Gauss' divergence theorem $\int_S J \cdot ndS = 0$ where S is a closed surface enclosing the node and this is precisely the KCL. If the node is big, it can accumulate charges and then the generalization of the KCL would read

$$\operatorname{div} J + \frac{\partial \rho}{\partial t} = 0 \quad \dots \quad (3)$$

which in integral form is

$$\int J \cdot ndS = -\frac{dQ(t)}{dt} \quad \dots \quad (4)$$

where $Q(t)$ is the charge contained in the node at time t . This charge conservation equation can in turn be derived from the Maxwell equations:

$$\operatorname{curl} H = J + \epsilon \frac{\partial E}{\partial t}, \operatorname{div} E = \rho/\epsilon \quad \dots \quad (5)$$

On taking the divergence of the first and using the second, we obtain (3). Likewise, the Kirchhoff voltage law which states that the algebraic sum of voltage around a closed loop of a circuit is zero can be derived by starting with Faraday's law of electromagnetic induction

$$\operatorname{curl} E = -\mu \frac{\partial H}{\partial t}$$

which gives on integrating over the flat surface S whose boundary Γ is the given circuit and applying Stokes' theorem,

$$\int_{\Gamma} E \cdot dr = -\frac{d\Phi}{dt}$$

where Φ is the magnetic flux through S and its contribution comes only from the inductors in the circuit. Thus, $V_R + V_C + V_i = \int_{\Gamma} E \cdot dr$ represents the algebraic sum of voltage drops in the circuit Γ coming from resistors, capacitors and voltage sources while $V_L = -\frac{d\Phi}{dt}$ is the algebraic sum of the emf/voltage drops across all the inductors in the circuit. This equation therefore reads

$$V_R + V_C + V_i + V_L = 0$$

which is the KVL. The question which naturally arises in ones mind is that why do we have to introduce such field theoretic concepts which are harder to grasp first rather than directly talk about resistors, capacitors, inductors, voltage and current sources and the relationship between current and voltage for these elements along with the KCL and KVL. The reason being that the circuits

assembled by us on Bread-boards are lumped parameter circuits, ie they consist of discrete distributions of circuit element while on the other hand, when we go by train we observe massive transmission lines carrying current attached to tall pylons and also we observe large antenna dishes for our television and internet basis being fed in by transmission lines. These objects, ie, transmission lines, waveguides and antennas are examples of distributed parameter networks, ie, we have a continuous distribution of resistances, capacitances and inductances which may be functions of the spatial location along the network. These networks are important in power transmission from the generator to our houses, receiving image signals on our television screen and in transmitting messages encoded as electromagnetic fields from our antennas into space whose signals we collect using receiver antennas. They also form a major part of electrical and electronics engineering and can be modeled and understood only using the Maxwell field theory. Thus a thorough understanding of the Maxwell equations enables us to grasp both lumped parameter circuits and distributed parameter circuits in one stroke. This is the precise reason for introducing field theoretic concepts right at the beginning of our electronics engineering curriculum.

Properties of resistances, capacitances and inductances also follow from field theoretic analysis naturally. For example, the field theoretic version of Ohm's law $J = \sigma E$ and the definition of voltage as the line integral of the electric field, the definition of current as the surface integral of the current density and the definition of resistance as the line integral $\int_1^2 dr/\sigma A(r)$ where dr is the line element, $A(r)$ is the cross sectional area at r and σ is the conductivity which may also vary with r , r being the distance parameter along the wire, all are natural consequence of classical field theory and lead to the familiar Ohm's law $V = IR$. Likewise, the $Q = CV$ relation between charge on the plates and voltage between the plates is a consequence of the one dimensional version of Gauss' law $d(\epsilon E_z)/dz = 0$ in between the plates and the boundary condition derived from Gauss' law that ϵE_z on the surface of the plates gives us the surface charge density. The field theoretic viewpoint thus enables us to solve more complex problems like the determination of the resistance or capacitance between two surfaces of arbitrary shape when the medium inside has non-uniform conductivity and non-uniform resistivity. Coming over now to inductance, the starting point for determining the inductance of a coil is Ampere's law in integral form $\int_{\Gamma} H \cdot dr = I$ which in differential form reads $\text{curl } H = J$. This states that the line integral of the magnetic field H around a closed loop equals the net current flowing through the loop. This equation therefore yields the magnetic field within the loop as a linear function of I . The magnetic flux density $B = \mu H$ where μ may even be field dependent or even have memory as in Hysteresis effects is then computed and its flux through the loop surface is calculated yielding the magnetic flux Φ as a function of the current which when combined with Faraday's law of induction (once again a field equation) in integral form yields the EMF of an inductor $E = -d\Phi/dt$ as being proportional to the current if the permeability is a constant or it may even be a complicated nonlinear functional of the current if the medium has Hysteresis.

All this discussion shows us that to clearly formulate the circuit equations KCL, KVL and obtain the current voltage relations for resistances, capacitances and inductances, we require a thorough grounding in electromagnetic field theory.

5.2 Graph theoretic analysis of large linear circuits based on KCL and KVL

Once the field theoretic concept has been understood, we can start teaching the fundamentals of electronics starting with the elementary KCL,KVL analysis of lumped parameter networks, the Laplace domain representation of impedances constructed using resistances, capacitances and inductances, using the Laplace domain technique to replace the differential equation formulation of KCL,KVL equations by linear algebraic s -domain equations, solving these algebraic equations to obtain s -domain expressions for currents and voltages in the circuit and then inverse Laplace transforming these expressions to obtain time domain transient expressions for the currents and voltage. Then we introduce sinusoidal steady state analysis by replacing the s variable with the imaginary $j\omega$ variable with ω being the operating frequency corresponding to the voltage/current ac source. When sources of different frequencies are present, we teach them how to block all sources except one of them, solve for the voltages and currents in the circuit using the phasor technique ie the $j\omega$ method, transform phasors into time domain expressions and then superpose all the resulting time domain expressions, each one being obtained by blocking all the sources except one of them. We then introduce graph theory for programming a computer to solve large sized linear circuits. The notion of incidence matrix, reduced incidence matrix and cutset matrix for formulating the KCL is first taught. Then, the f-loop and mesh matrices are introduced for formulating the KVL. After that a fundamental relation connecting the f-loop matrix and f-cutset matrix $Q_f^T B_f = 0$ is proved and used along with the elemental relations(obtained by writing down the element current-voltage relations in the s -domain) in matrix form, noting that controlled sources yield non-diagonal entries. We then assemble these equations to obtain a complete set of matrix equations for the chord currents and another complete set of matrix equations for the branch voltages in the f-cutset. Graph theory enables us to develop a systematic algorithm for programming a computer to solve any given linear network however large. We simply have to enter the cutset and loop matrices, or equivalently the graph topology and the elemental relation matrices and the computer does the rest.

5.3 Two port network theory

The next stage is to introduce the students to two port networks both linear and nonlinear taking as examples two port networks derived from linear circuit problems like T and Π blocks and two port networks obtained while modeling BJT and C-MOS transistor elements. A lot of emphasis must be given on two port

theory including linearizing nonlinear two port systems because it is this branch of electronics which will enable them to understand modern electronic gadgets based on transistors like oscillators, amplifiers. The notion of linearization is related to ac response of an amplifier/oscillator around a quiescent operating point. After linearizing, how to analyze the resulting ac/small-signal equivalent circuits using the KCL and KVL plays a crucial role. The Ebers-Moll two port model of a BJT transistor giving the collector and emitter currents in terms of exponential combinations of the base-collector and base-emitter voltages should be introduced and the associated linearized model and the H and G parameter small-signal equivalent circuits derived using these linearized equations. Analyzing amplifiers in the different frequency bands starting from these linearized models should be taught. When two or more transistors are present in the circuit as in the differential amplifier, then how to do a small signal ac analysis using the linearized H or G parameter models must be dealt with in detail.

5.4 Diode and capacitance circuit models

The next kind of related problems concern diode circuit models which may be taught either before or during the course of discussing transistor circuits. clipping and clamping circuits involving diode resistor and diode-resistor-capacitor combinations may be introduced at this stage. Here, we assume the diode to be ideal, ie, constant forward resistance and infinite reverse resistance. Other piecewise linear models may also be introduced but at it should be emphasized in the class that these models are all approximations to the exponential current-voltage model of a pn-junction diode. Half wave, full wave and bridge rectifiers with analysis based on the ideality of the diode can be explained with their response to sinusoidal signals stressing on Fourier series analysis of the rectified voltage and computation of the ripple factor. When capacitances are introduced into the circuit to reduce the ripple, then how to compute by elementary transient RC circuit behaviour, the ripple factor can also be taught at this stage. Diode circuits used in communication system theory like the envelope detector for AM-demodulation with elementary qualitative analysis of their working can be brought in here.

5.5 Classical device physics

It is my opinion that device physics should be brought in at this stage and not earlier because device physics is at a higher technical level of complexity. The inquisitive student will definitely ask how one arrives at the exponential formulas for the currents in diodes and transistors. Some erroneous models exist in the literature on the mathematical models for two port systems especially the MOS transistor which has three terminals: a gate, a source and a drain. The channel between the source and the drain carries the current and the width of

this channel is controlled by the gate source voltage. The current through the source-drain channel depends on the width of this channel and this yields the drain current as $I_D = \alpha(V_{GS} - V_T)^2$. We need another equation to complete the two port description of the MOS transistor. Normally, this equation is set to be $I_G = 0$, ie, zero gate current. Then, $I_D = I_S$. However, this is not exact since in between the gate and source plates, there is a resistance and also a capacitance and likewise between the gate and drain plates. Thus, a more accurate second equation would be to take

$$I_G = V_{GS}/R_{GS} + V_{GD}/R_{GD} + C_{GS}dV_{GS}/dt + C_{GD}dV_{GD}/dt$$

In this way after linearizing, we get an accurate model for the MOS. Such innovative models as a possible improvement over textbook models may be suggested in the classroom with the hope that the students will be able to apply electromagnetic theory to compute the parameters $\alpha V_T, R_{GS}, R_{GD}, C_{GS}, C_{GD}$.

To understand device physics, some amount of statistical mechanics especially the Gibbs distribution should be taught. Then how to apply Poisson's electrostatic equation in one dimension with boundary conditions to compute the widths and potential barriers of the space-charge layers for a doped p-n junction diode should be introduced. The diffusion equation giving the diffusion current of minority carriers in a doped semiconductor based on concentration gradients should be derived in the classroom. The drift current in terms of the applied electric field and notions such as mobility, relaxation time etc. can be taught and how one models the total current in a pn junction as the sum of a drift component and a diffusion component must be emphasized. The portion on classical device physics may be concluded by deriving the two diode/Ebers-Moll model for a BJT transistor with improvements in the model by introducing base-emitter and base-collector junction capacitances may be briefly dealt with.

5.6 Device physics using quantum mechanics and quantum electrodynamics

Device physics at a more advanced level can be introduced in the form of mini-projects. This would involve notions like the photo-electric effect and its application to the design of photoelectric cell that would produce current when light falls on it, the design of tunneling diode invented by Esaki that is based on the quantum mechanical tunneling effect etc. The photo electric effect cannot be successfully explained without introducing quantum electrodynamics involving the description of a quantum electromagnetic field interacting with matter in the form of the electron-positron second quantized Dirac field. The perturbation to the Dirac field caused by the quantum electromagnetic wave is calculated using time dependent perturbation theory and is expressed in terms of the creation and annihilation operator fields of the electrons, positrons and photons and the average value of this current is calculated in a given state of the matter and field. The Dirac equation with electromagnetic interactions should here also involve the classical Coulomb interaction between the electron and the nucleus.

Quantum tunneling theory can be introduced by starting with the stationary Schrodinger equation for a particle moving in one dimension in a rectangular potential barrier. Solution to the Schrodinger equation in the three regions following by application of the boundary conditions yields the transmission and reflection coefficients in terms of the particle's energy and the height and width of the potential barrier.

5.7 Band theory of a semiconductor by solving Schrodinger's equation

Another important problem in device physics involving the use of quantum mechanics is the computation of the band structure in a semiconductor. This involves treating such a crystal as a lattice of ions generating a three dimensional periodic potential with possibly three different periods along the three linearly independent directions of the lattice. The Bloch wave functions are introduced by applying periodic boundary conditions and the resulting Schrodinger equation is analyzed using Fourier series expansions for the periodic potential and for the Bloch wave function. By solving this in some special cases like the Kronig-Pinney model where the potential is a δ -function train, we can in principle determine where the energy levels are more clustered, ie, we determine the band structure. This enables us in particular to calculate the valence and conduction bands in a semiconductor which are important in deriving the Gibbs exponential formulas for the current in a semiconducting pn junction.

5.8 Quantum electrodynamics and quantum field theory

Quantum mechanics and quantum field theory should perhaps not be taught at the classroom level; these concepts should be introduced as projects supplemented with practical design methods, like for example, removing the top cap of a transistor to obtain a photo-cell and measure the light intensity and frequency falling on such a photo-cell, measuring the current produced by the photo cell and tallying it with theoretical computations based on quantum electrodynamics as developed by Feynman, Schwinger, Tomonaga and Dyson.

While introducing concepts in quantum field theory, we should emphasize on Fermion field quantization, for that will enable us to develop the theory of superconductivity. The idea is to start with the second quantized Schrodinger Hamiltonian obtained by integrating $\psi^*(r)H\psi(r)$ w.r.t d^3r over space where H is the Schrodinger Hamiltonian and $\psi(r)$ is Fermionic field satisfying the canonical anticommutation rule $\{\psi(r)^*, \psi(r')\} = \delta^3(r - r')$. To this second quantized Hamiltonian, we add the pairing Hamiltonian which consists of fourth degree terms in the Fermionic field, two of which appear without an adjoint and two appear with an adjoint. Introduction of this term enables us to describe the

Cooper phenomena in superconductors involving the pairing of Fermions to produce Bosons which travel inside the metal with almost zero resistance. The average current density in a superconductor must be computed by starting with the Schrodinger expression for probability current as a bilinear from in $\psi(r), \psi(r)^*$ and taking its average w.r.t. the Gibbs state for the unperturbed Hamiltonian. This is much more simplified when we introduce the temperature Green's function $G(t, r|t', r') = Tr(\exp(-\beta H)T\{\psi(t, r)\psi(t', r')^*\})/Tr(\exp(-\beta H))$ and derive a differential equation for this using the Heisenberg dynamics for the field operators $\psi(t, r)$ that get simplified by using the anticommutation rules for the Fermionic field operators.

5.9 Analyzing random Gaussian and non-Gaussian noise in circuits using higher order correlations and spectra

The next stage while teaching electronics to undergraduate is to introduce classical probability and random processes as a mathematical tool for analyzing fluctuations in circuits. For example, thermal noise caused by the incessant random motion of electrons at finite temperature produces a noisy component of current in a circuit whose power spectral density depends on the transfer function of the circuit. The thermal noise current itself is nearly white, ie, it has a constant spectral density of $kT/2$. However, quantum effects are considered, we get a more accurate model as $kT/2$ plus a Planckian spectrum. This thermal noise having a spectral density of $S_n(\omega)$ generated in a given resistance produces a noisy current having spectral density $|H(\omega)|^2 S_n(\omega)$ at an element in the circuit that sees a transfer function of $H(\omega)$ relative to the resistance where the thermal noise is generated. The student of electronics should be taught how to connect the end terminals of a resistor in a heat bath to the CRO and see the spiky noisy voltage waveform. The other thing which they should be encouraged to pursue is the use of the spectrum analyzer to analyze how much power is contained in each frequency band of a voltage waveform including noise.

Random noise is also present in the input voltage source of a circuit. If the circuit is linear and time invariant, then the output power spectral density of noise is easily computed using the transfer function method. If the system is time varying, then its response may be described by a linear integral equation of the form

$$y(t) = \int h(t, \tau)x(\tau)d\tau$$

Thus, if $x(\cdot)$ is a random process, we can compute the statistical moments of the output as

$$\langle y(t_1) \dots y(t_n) \rangle = \int (\prod_{k=1}^n h(t_k, \tau_k)) \langle x(\tau_1) \dots x(\tau_n) \rangle d\tau_1 \dots d\tau_n$$

In the special case of time invariant linear systems, $h(t, \tau)$ is a function of $t - \tau$

only, and we get by assuming the input to be a stationary process, the following relationship between the multivariate Fourier transforms of the input and output higher moments,

$$P_{y,n}(\omega_1, \dots, \omega_{n-1}) = H(\omega_1) \dots H(\omega_{n-1}) \bar{H}(\omega_1 + \dots + \omega_{n-1}) P_{x,n}(\omega_1, \dots, \omega_{n-1})$$

For $n = 2$ this reduces to the usual formula for the psd and hence cannot be used to estimate the phase of the transfer function. For $n \geq 3$, the output polyspectrum is sensitive to the phase of the system and can be used to estimate it after the magnitude of the transfer function has been estimated from the psd. The same can be carried out for discrete time LTI systems with the only difference that the inverse Fourier transform is carried out over $[-\pi, \pi)$ and the signals are defined over \mathbb{Z} rather than the continuous time case where both the signal time domain and the signal Fourier transform domains are \mathbb{R} .

Random process theory can also be introduced from the correlation viewpoint as directly solutions of dynamical equations. For example, if the LTI system has a transfer function

$$H(s) = \frac{\sum_{k=0}^{q-1} b[q+1-k]s^k}{s^p + \sum_{k=0}^{p-1} a[p+1-k]s^k}$$

then the output $y(t)$ is related to the input by the differential equation

$$y^{(p)}(t) + \sum_{k=0}^{p-1} a[p+1-k]y^{(k)}(t) = \sum_{k=0}^{q-1} b[q+1-k]x^{(k)}(t)$$

From this equation we obtain the differential equation for the propagation of the output and input-output correlations:

$$\begin{aligned} \frac{\partial^p}{\partial t^p} < y(t)x(t') > + \sum_{k=0}^{p-1} a[p+1-k] \frac{\partial^k}{\partial t^k} < y(t)x(t') > \\ &= \sum_{k=0}^{q-1} b[q+1-k] \frac{\partial^k}{\partial t^k} < x(t)x(t') >, t \neq t' \end{aligned}$$

and

$$\begin{aligned} \frac{\partial^p}{\partial t^p} < y(t)y(t') > + \sum_{k=0}^{p-1} a[p+1-k] \frac{\partial^k}{\partial t^k} < y(t)y(t') > \\ &= \sum_{k=0}^{q-1} b[q+1-k] \frac{\partial^k}{\partial t^k} < x(t)y(t') >, t > t' \end{aligned}$$

and these equations can be solved to obtain the input-output correlations and the output correlations from initial conditions on $\frac{\partial^k}{\partial t^k} < y(t)x(t') > |_{t=t'}, k = 0, 1, \dots, p-1$ and $\frac{\partial^k}{\partial t^k} < y(t)y(t') > |_{t=t'}, k = 0, 1, \dots, p-1$. These equations are valid even when the coefficients $a[k] = a[k, t], b[k, t]$ depend on time. Numerical methods for programming in MATLAB to solve the above equations after discretization may be introduced at this stage.

5.10 Noise in nonlinear transistor circuits

Modeling and estimation of nonlinear transistor and diode circuits can also be introduced at this stage. This would enable us to introduce the student to the stochastic calculus of Ito by modeling nonlinear circuits with thermal noise and noisy inputs using stochastic differential equations. The Fokker-Planck-Kolmogorov forward and backward equations for the transition probability density of the state of the circuit, like a subset of the nodal voltages and currents can be solved for by developing numerical methods for solving the forward Kolmogorov forward equation using the method of moments. For that matter, modeling the propagation of noise in any non-linear circuit built out of BJT and C-MOS transistors using Brownian motion and stochastic calculus is lacking in our current syllabus for M.Tech and doctoral students. We can also include spiky noise in the form of Compound Poisson processes and more generally, model the nonlinear circuit using stochastic differential equations driven by Levy processes, ie processes with independent increments whose characteristic functional is determined by the Levy-Khintchine formula. After modeling, comes the process of estimating the circuit parameters and the circuit state based on noisy measurements at certain nodes and certain links. Developing this theme further, we get an extended state vector comprising the state vector and the unknown parameter vector which can be estimated on a real time basis using the Kushner-Kallianpur nonlinear filter. This would also motivate the student to study quantum stochastic differential equations and quantum real time filtering which are important in quantum circuits at the atomic level. Approximating the Kushner-Kallianpur infinite dimensional filter leads to the extended Kalman filter which is easily implementable since it involves only evolution of the conditional mean of the extended state and the corresponding error covariance matrix. The student of electronics and signal processing should definitely know how to implement the EKF for a nonlinear circuit.

5.11 Digital electronics

Now we come to how digital electronics is to be taught. First, the design of basic digital gates like NOT, AND, NAND, OR, EXOR, should be designed using transistor elements where the high voltage is treated as a one and a low voltage as a zero. The design and operation of basic flip-flops like the D , T , JK , RS flip-flops and how by linking these to a shift register, we can use a clock cause the register to count cyclically a binary sequence. The design of clocks for sequential digital circuits using the 555 timer or using pulse width modulation (PWM). This involves feeding one input terminal of an opamp with a ramp voltage and the other with a dc voltage whose amplitude can be varied arbitrarily. Thus, when the ramp input crosses the other dc input level, the opamp triggers, from a 1 state to a zero state or vice-versa. The output is therefore precisely a PWM signal whose width is controlled by the slope of the ramp and the other dc input. Thus, clocks of varying duty cycle can be designed using this method. The

notion of a truth table and a Karnaugh map for simplifying Boolean expressions must be introduced while designing digital circuits like the half adder, full adder, binary to decimal conversion etc. It must also be introduced while designing counters with a shift register connected to flip flops described by their excitation tables. The inputs to the flip-flops are Boolean functions of their outputs designed according to their excitation table. While teaching digital electronics, the DFT and FFT may be introduced for the computer aided spectral analysis of signals. The FFT is implemented using an array of butterfly blocks, each taking two inputs and produces two outputs. If the length of the sequence whose FFT is to be computed is N , then there are $\log_2 N$ layers of butterflies and consequently the algorithmic complexity is $O(N \log_2(N))$ as compared to $O(N^2)$ required by the DFT. In this context, the student may be taught how to discretize a linear circuit in time, so that it becomes a system of difference equations in discrete time and then perform an FFT of these equations using a computer programme and hence calculate the output spectrum in the discrete frequency domain.

5.12 Techniques for analyzing transmission lines

After teaching analog and digital electronics along with device physics (which requires the bare minimum of electromagnetic field theory), the student should be taught in full detail electromagnetic field theory along with its applications to antennas, transmission lines and waveguides. Transmission line theory does not require much of emft except at one point where one draws a comparison between transmission lines and waveguides by taking to infinite parallel conducting plates both parallel to say the xz plane and separated by a distance of d metres and writes down the Maxwell curl equations relating $E_y(t, z), H_x(t, z)$. The basic aim of teaching transmission line theory from a mathematical standpoint should be emphasized, namely, the notion of replacing ordinary differential equations in time for lumped parameter circuit with the notion of a partial differential equation in length and time for distributed parameter circuits. Solving the Tx line equations in the case of uniform distributed parameters and the consequent introduction of notions like propagation constant, characteristic impedance, VSWR, voltage and current maxima and minima, reflection coefficient along the line, the computation of input impedance and matching (to give zero reflection coefficient, so that all the power is transmitted to the load), using the smith chart for stub matching, derivation of the smith chart constant resistance and constant reactance circles must be introduced. Techniques of analyzing non-uniform Tx lines using the Fourier series in the spatial variable, the notion of an infinite sequence of propagation constants appearing as eigenvalues of infinite sized matrices and its physical interpretation in terms of a cascade of infinitesimal Tx line blocks may be done. Solving the nonuniform line approximately in the frequency domain using perturbation theory for ode's may also be mentioned. Taking hysteresis non-linearities with memory into account in the inductance part of the line equations and capacitive non-linearities

with memory into account in the capacitance part of the line equations can be introduced in the form of projects. This will give the student an idea of solving non-linear integro-differential equations approximately using perturbation theory and the consequent estimation of these nonlinearities by analyzing the spectra and polyspectra of the line current and voltage in the frequency domain. One point missed out in our syllabus on Tx lines is line loading, ie, if we have a constant external voltage per unit length and an external line current appearing due to stray environmental effects, then how to solve the line equations, specifically, how to solve the pde's

$$\frac{\partial i(t, z)}{\partial z} + C(z) \frac{\partial v(t, z)}{\partial t} + G(z)v(t, z) = j(t, z),$$

$$\frac{\partial v(t, z)}{\partial z} + L(z) \frac{\partial i(t, z)}{\partial t} + R(z)i(t, z) = e(t, z),$$

which translate in the frequency domain to

$$\frac{\partial I(\omega, z)}{\partial z} + (G(z) + j\omega C(z))V(\omega, z) = J(\omega, z),$$

$$\frac{\partial V(\omega, z)}{\partial z} + (R(z) + j\omega L(z))I(\omega, z) = E(\omega, z),$$

with source and load conditions

$$V_s(\omega) - Z_s(\omega)I(\omega, 0) = V(\omega, 0),$$

$$V(\omega, d) - I(\omega, d)Z_L(\omega) = 0$$

then how to solve for $V(\omega, z)$, $I(\omega, z)$ and evaluate their statistical moments taking into account possible randomness in the line loading functions $j(t, z)$, $e(t, z)$.

5.13 Brownian motion, Poisson processes and stochastic differential equations in circuit theory

While teaching random process theory to electronics and communication students or for that matter to any engineering student, two of the most important processes that must be covered in full detail are the Brownian motion and the Poisson process and Ito's formula for them, how stochastic differential equations driven by these processes are formulated, existence and uniqueness of the solutions to sde's using Lipshitz conditions on their coefficients are proved etc. Typically, non-Gaussian processes are derived by passing Gaussian processes or more precisely white Gaussian noise through nonlinearities and this fact is conveyed best through Ito's theory of sde's. At the same time, the Stroock-Varadhan martingale formulation of diffusion processes can be introduced. This

involves defining a diffusion process in a different way, ie, rather than as a solution to Ito's sde with drift and diffusion coefficients satisfying certain Lipshitz condition, we define it as a process for which a certain class of functionals of this process satisfies the Martingale property. This formulation includes the Ito definition as a special case and it covers even more general processes for which Lipshitz conditions on the coefficients do not hold. Moreover, unlike the Ito theory which is based on sample paths, the Stroock-Varadhan theory focusses on the existence and uniqueness of probability measures on the space of continuous sample paths starting at a given point. Even more, unlike the Ito theory, the Stroock-Varadhan theory does not even require the generator of the process to be defined in terms of drift and diffusion coefficients which are functions of the current time and current state, it only requires the drift and diffusion coefficients to be progressively measurable functions of the sample paths of the process.

Solving sde's using perturbation theory for sde's gives the engineering student a readily available method for plotting solutions.

5.14 Classical and quantum random processes in circuit theory

While introducing the notion of random signals in electronic circuits, the clear distinction between classical random processes and quantum random processes must be stressed owing to the fact that at the atomic scale, simultaneous measurability of noise at two different times is not possible in every state of the bath if the corresponding operators do not commute as follows from Heisenberg's uncertainty principle. This means that generally, we cannot talk about the joint probability distribution function of the time samples of the random process although one can formally, as in the Hudson-Parthasarathy theory talk about quantum stochastic differential equations and the corresponding Evans-Hudson flow that describes noisy Heisenberg dynamics. However, in some special commutative cases, one can derive classical Brownian motion and classical Poisson processes from their generalized quantum versions.

5.15 Simulation of nonlinear ode's and pde's in circuit theory and electromagnetics

The electronics student must learn how to derive the Runge-Kutta algorithm and use it to simulate linear and non-linear circuits described by ordinary differential equations. He must also learn to simulate pde's like the Tx line equations in the time-space domain by discretizing all the independent variables in the pde. At a given time, all the spatial samples of the signal field/fields should be stacked into a single vector using something like Kronecker tensor products and Lexicographic ordering to obtain an ordinary differential equation for the

resulting vector w.r.t. the time variable. Using this tool, he should learn to simulate the Maxwell curl equations which are first order in time. He may also learn how to use this technique to simulate the wave equation with source for the magnetic vector and electric scalar potential by converting this second order in time and space equation into a pair of two first order in time but second order in space differential equations. This should be tried on the Navier-Stokes equations of fluid dynamics and also on Maxwell's equations with field and position dependent permittivities and permeabilities and hysteresis nonlinearities:

$$D(\omega, r) = \epsilon_0 E(\omega, r) + \epsilon_0 \sum_{n \geq 1} \delta^m \int \chi_n(\omega, \omega_1, \dots, \omega_n) \bigotimes_{m=1}^n E(\omega_m, r) d\omega_1 \dots d\omega_n$$

$$B(\omega, r) = \mu_0 H(\omega, r) + \mu_0 \sum_{m \geq 1} \delta^m \int \chi_{Bn}(\omega, \omega_1, \dots, \omega_n) \bigotimes_{m=1}^n H(\omega_m, r) d\omega_1 \dots d\omega_n$$

More generally, it should be noted that D will also depend on H and likewise, B will also depend on E .

5.16 Derivation of medium properties from basic physical principles involving motion of individual electrons and magnetic moments in external fields

The derivation of these two equations from first principles using classical mechanics and then a generalization of this using quantum mechanics should be introduced. For the first part, we write down the classical equation of motion of displacement of an electron from its equilibrium position w.r.t the nucleus when an electromagnetic field is applied as

$$M(r)\xi'(t) + \Gamma(r)\xi'(t) + K_0(r)\xi(t) + \delta K_1(r, \xi(t)) = -e(E(t, r+\xi(t)) + \xi'(t) \times B(t, r+\xi(t)))$$

This equation is solved using perturbation theory $K_0\xi$ represents the linear/harmonic part of the binding force while $K_1(r, \xi)$ represents the anharmonic part of the binding force. These are obtained by expanding the binding potential around the equilibrium position. The dipole moment per unit volume, ie, the polarization is computed as

$$P(t, r) = -N(r)e\xi(t)$$

where N is the number density of the atoms. Then, D is expressed as

$$D = \epsilon_0 E + P$$

Likewise, to describe the Hysteresis effect we modify the law of precession of magnetic moments in an electromagnetic field as

$$m'(t) = \gamma_1 m(t) \times B(t) + \gamma_2 p(t) \times E(t)$$

where $m(t)$ is the magnetic dipole moment of the atom and $p(t)$ is its electric dipole moment, ie, $p(t) = \xi(t) = -e\xi(t)$. It should be noted that the magnetization is the magnetic moment per unit volume $M(t) = N(r)m(t)$ and H is computed using

$$B = \mu_0 H + M$$

It should be noted that $m(t) \times B(t)$ is the torque on a magnetic dipole while $p(t) \times E(t)$ is the torque on an electric dipole. Further, the angular momentum L of the atomic dipole is related to its magnetic dipole moment m by

$$m = -eL/2m_e$$

This expression does not take spin into account. Spin is a purely quantum mechanical effect and if that is accounted for, then we get the modified equation

$$m = -e(L + 2s)/2m_e$$

where s is the spin angular momentum. This expression can be obtained using Dirac's relativistic wave equation by showing that the quantity $J = L + 2s$ is conserved in a radial potential and hence should be interpreted as the total angular momentum.

5.17 Partial differential equation methods for analyzing waveguides

Having covered the portion on Tx lines, the student should be taught how to apply Maxwell's electromagnetic field theory to waveguides and antennas. These two topics form the core part on applications of partial differential equations. The wave guide equations are obtained by separating the gradient and curl operations into a longitudinal component and a transverse components in accordance with the assumption that the longitudinal dependence on the fields is $\exp(-\gamma z)$ and then substituting these expressions into the transverse parts of the Maxwell curl equations. Doing so at a given frequency enables us to obtain concrete expressions for the transverse component of the em fields in terms of the transverse gradients of the longitudinal field components. The last of the Maxwell curl equations then gives us Helmholtz equations for the longitudinal field components. Even when the permittivity and permeability have spatial dependence on the transverse variables alone along with frequency, one can in the same way derive the same expressions for the transverse fields in terms of the longitudinal fields. However, the longitudinal components of the Maxwell curl equations will not give us decoupled Helmholtz equations for the longitudinal E and H components as was possible in the constant permittivity and permeability case. We instead get two coupled generalized Helmholtz equations for these components with coefficients depending on frequency and the transverse spatial components.

In the homogeneous permittivity and permeability case, the two decoupled 2-D Helmholtz equations for E_z, H_z are solved by applying the appropriate boundary conditions, ie, the tangential component E_z vanishes on the boundary and the normal derivative of H_z vanishes on the boundary (the latter ensures that the normal component of H will vanish on the boundary). These boundary conditions yield at a given frequency, possible discrete values of the propagation constant γ and the resulting solution is a superposition of these modes. These solutions can be separated into two classes ie, the TE and TM modes. In the TE mode, there is no longitudinal E and hence all components of the em field are derived from H_z alone while in the TM mode, there is no longitudinal H and hence all components of the em field are derived from E_z alone.

5.18 Curvilinear coordinate systems and variational methods in engineering electromagnetics

In the case of rectangular guides, it is an accidental coincidence that the propagation constants for both the TE and TM modes are the same while this is not so in the case of cylindrical guides and guides of arbitrary cross section. The TM modes are derived solutions to the Helmholtz equation for E_z with the Dirichlet boundary condition, ie E_z vanishes on the boundary while the TE modes are derived with the Neumann boundary condition, ie, the normal derivative of H_z vanish on the boundary. These two different kinds of boundary conditions for E_z, H_z satisfying the 2-D Helmholtz equation are responsible for the two different kinds of the propagation constants and hence cutoff frequency. For a given mode, ie, a given eigenvalue of the 2-D Helmholtz equation with the appropriate boundary condition, we get a cutoff frequency for that mode, ie, that frequency at which the propagation constant becomes purely imaginary. In the case of cylindrical guides, the cutoff frequency for the TM mode is expressed in terms of the zeros of the Bessel functions while for the TE mode it is expressed in terms of the zeros of the derivative of the Bessel functions. More generally, for a waveguide with arbitrary smooth cross section, we may select an orthogonal coordinate system in the xy plane, namely $(q_1(x, y), q_2(x, y))$ so that $q_1 = c_0$ is the boundary curve of the guide. Then, the 2-D Laplacian and hence the 2-D Helmholtz equation can be expressed in the $q_1 - q_2$ system using Lame's coefficients and although the Helmholtz equation becomes more complex than that using the xy coordinate system, the boundary conditions are simple, namely, for the TM mode, we apply the Dirichlet condition: $E_z(q_1, q_2) = 0$ where $q_1 = c_0$ and for the TE mode, we apply the Neumann condition: $\frac{\partial H_z(q_1, q_2)}{\partial q_1} = 0$ when $q_1 = c_0$. It is possible to formulate the Helmholtz and generalized Helmholtz equation with Dirichlet or Neumann boundary conditions using a variational method. For example, the generalized Helmholtz equation

$$\operatorname{div}(\epsilon_1(x, y)\nabla E_z(x, y)) + h^2\epsilon_2(x, y)E_z(x, y) = 0, (x, y) \in D$$

with E_z prescribed on the boundary ∂D can be formulated using variational calculus as

$$\delta S = 0$$

with $\delta E_z = 0$ on the boundary ∂D , where

$$S = (1/2) \int_D (\nabla_{\perp} E_z(x, y)^*, \epsilon(x, y) \nabla_{\perp} E_z(x, y)) dx dy$$

$$- (h^2/2) \int_D \epsilon_2(x, y) E_z(x, y)^* E_z(x, y) dx dy$$

Minimization of S w.r.t E_z, E_z^* with the Dirichlet boundary condition implying that $\delta E_z, \delta E_z^* = 0$ on ∂D yields the desired equation. This minimization problem may be approximated using the finite element method (FEM) by partitioning D into triangles, and assigning values of E_z at the vertices of the inner triangles, assigning $E_z = 0$ on the boundary vertices and for each triangle, approximating E_z within the triangle by a linear function $ax + by + c$ with a, b, c determined by the vertex field values and the vertex coordinates and then calculating the integral S by substituting this approximation in terms of the inner vertex field values and minimizing this quadratic function of the inner vertex field values to get a set of homogeneous linear equations with h^2 appearing as generalized eigenvalues of the resulting matrices. These modal eigenvalues are approximately obtained by setting the determinant of the matrix pencil to zero and once the values of h^2 are approximately known, the propagation constant values γ may be determined using $\omega^2 \mu_0 \epsilon_0 + \gamma^2 = h^2$ which yields the cutoff frequencies as $\omega_c = h/\sqrt{\epsilon_0 \mu_0}$. For each mode h , we get a cutoff frequency. As the partition into triangles becomes finer and finer, we get a better and better approximation for the modes. It should be mentioned that when we make a transition from the true continuous problem to the approximate discrete problem, the quantity to be minimized is

$$(1/2) E^* (A - h^2 B) E$$

where E is the vector of inner vertex potentials and setting the gradient of this to zero gives us the approximate modal equations

$$(A - h^2 B) E = 0$$

which gives us the modal equations

$$\det(A - h^2 B) = 0$$

Plotting of the field lines in a given cross section using some software may also be taught. For example, if we wish to plot the transverse electric field lines, we would have to solve the first order differential equation

$$\frac{dx}{E_x(x, y)} = \frac{dy}{E_y(x, y)}$$

using finite differences.

5.19 Perturbation theoretic methods in solving electromagnetics problems

Having discussed wave-guide theory from the standpoint of solving the Helmholtz equations with appropriate boundary conditions, the course should aim at developing perturbation theoretic methods for solving the coupled (for E_z and H_z) generalized Helmholtz equations when the permittivity and permeability depend on frequency and the transverse coordinates. This can be achieved by writing

$$\begin{aligned} E_z &= E_+ z^{(0)} + \delta.E_z^{(1)} + O(\delta^2), \\ H_z &= H_z^{(0)} + \delta.H_z^{(1)} + O(\delta^2), \\ \gamma &= \gamma^{(0)} + \delta.\gamma^{(1)} + O(\delta^2) \end{aligned}$$

where the permittivity and permeability are expressed as small perturbations of the homogeneous cases:

$$\epsilon(\omega, x, y) = \epsilon_0(1 + \delta\chi_e(\omega, x, y)),$$

$$\mu(\omega, x, y) = \mu_0(1 + \delta.chi_B(\omega, x, y))$$

By making these substitutions into the generalized Helmholtz equations and equating coefficients of δ^0 and δ^1 respectively, we derive the unperturbed and first order perturbed equations for the fields and propagation constants respectively. The unperturbed equations are same as those for the homogeneous case while the first order perturbed equations give us not only the perturbation in the fields caused by the inhomogeneity but also perturbations in the propagation constant γ . This problem is similar to the problem of perturbation theory for eigenvalues. We should note that the perturbation to the propagation constant can equivalently be viewed as a perturbation to the modal eigenvalue h^2 :

$$\begin{aligned} h^2 &= \omega^2 \epsilon_0 \mu_0 + \gamma^2 = \omega^2 \mu_0 \epsilon_0 + (\gamma^{(0)})^2 + \delta.2.\gamma^{(0)}\gamma^{(1)} + O(\delta^2) \\ &= \lambda^{(0)} + \delta.\lambda^{(1)} + O(\delta^2) \end{aligned}$$

where

$$\begin{aligned} \lambda^{(0)} &= \omega^2 \mu_0 \epsilon_0 + (\gamma^{(0)})^2, \\ \lambda^{(1)} &= 2\gamma^{(0)}\gamma^{(1)} \end{aligned}$$

When the unperturbed modes are degenerate as happens for the rectangular wave-guide case, we must use the theory of secular determinants to obtain the modal splitting. Both $E_z^{(0)}$ and $E_z^{(1)}$ must satisfy the same boundary conditions, ie, must vanish on the boundary. Likewise, $H_z^{(0)}$ and $H_z^{(1)}$ must also satisfy the same boundary conditions, ie, their normal derivative must vanish on the boundary. The unperturbed fields $E_z^{(0)}$ and $H_z^{(0)}$ determine an orthonormal basis for $L^2(D) \otimes \mathbb{C}^2$ which can be used to calculate the perturbed modes as is done in time independent perturbation theory in quantum mechanics. The same circle of ideas is used in the cavity resonator (DRA, ie, dielectric resonator

antenna) case. When the medium is homogeneous, we know the solution in terms of the eigenvalues of the 2-D Helmholtz equation with side boundary conditions and top and bottom surface boundary conditions, ie, H_z depends on z as $\sin(\pi pz/d)$ where p is an integer while E_z depends on z as $\cos(\pi pz/d)$. This is because E_x, E_y vanish on the top and bottom surfaces ($z = 0, d$) and hence

$$\partial E_z / \partial z = -(\partial E_x / \partial x + \partial E_y / \partial y) = 0$$

while H_z being the normal component of the magnetic field on the top and bottom surfaces, must necessarily vanish since these surfaces are perfect conductors. Further, wherever γ appears in the waveguide problem, we replace it with $-\partial/\partial z$. It follows that for the TE modes in a homogeneous resonator, the possible frequencies of oscillation are given by

$$\omega_H(n, p)^2 \mu_0 \epsilon_0 - (\pi p/d)^2 = h_H(n)^2$$

where $h_H[n]^2$ are the modal eigenvalues for the 2-D Helmholtz problem

$$(\nabla_{\perp}^2 + h_H^2) H_z(x, y) = 0, (x, y) \in D$$

with the Neumann boundary condition

$$\partial H_z(x, y) / \partial \hat{n} = 0, (x, y) \in \partial D$$

and likewise, for the TM modes in a homogeneous resonator, the possible frequencies of oscillation are given by

$$\omega_E(n, p)^2 \mu_0 \epsilon_0 - (\pi p/d)^2 = h_E(n)^2$$

where $h_E(n)^2$ are the modal eigenvalues for the 2-D Helmholtz problem

$$(\nabla_{\perp}^2 + h_E^2) E_z(x, y) = 0, (x, y) \in D$$

with the Dirichlet boundary condition

$$E_z(x, y) = 0, (x, y) \in \partial D$$

Again by applying perturbation theory, we can derive the shift in the oscillation frequencies produced by small inhomogeneities in the permittivity and permeability. The idea, as in the case of non-uniform transmission lines is to expand the electric and magnetic fields as a Fourier series in z with a period of d multiplied by a propagation constant factor $\exp(-\gamma z)$ and also expand the perturbations in the permittivity and permeability as Fourier series in z with coefficients depending on ω and the transverse variables x, y and then apply perturbation theory to derive the shift in the oscillation frequencies. Whilst applying perturbation theory, it should be borne in mind that the terms in the Fourier series expansion of E, H, ϵ, μ involving $\exp(j2\pi nz/d), n \neq 0$ should be multiplied by the small perturbation parameter since these terms would be absent if the permittivity and permeability did not depend on z .

The next topic to be covered in detail is antenna theory. How to compute the far field approximation to the electromagnetic field pattern and hence the Poynting vector by making the binomial approximation for $|r - r'|$ with $|r| \gg |r'|$. The important points to be covered here are as follows:

[1] Derivation of the far field electric and magnetic fields upto $O(1/r)$ in terms of the spatial Fourier transform of the current density.

[2] Computation of the far field Poynting vector, ie, the $O(1/r^2)$ term and integrate the radial component of this w.r.t $r^2 d\Omega(\hat{r})$ to obtain the total power radiated by the current field.

[3] Derive formulas for the radiation resistance for antennas of different kinds like the infinitesimal dipole, straight wire circular loop etc. using the results of [2].

[4] Derive formulas for the surface current density induced on the antenna surface when an incident em wave field falls on it. Assume some surface current density, calculate the electric field generated by this surface current density and set the tangential component of this electric field plus the incident electric field on the surface to zero to obtain the well known Pocklington integral equations for the surface current density.

[5] Apply the theory developed in [4] to compute the current induced on the driven elements in a Yagi array by solving the integral equations using the method of moments. The method of moments can also be used to solve the Pocklington integral equations for the surface current density on an antenna sheet by expanding this density as a linear combination of test functions defined as functions of the surface coordinates and calculating the coefficients from the integral equations by forming the inner products with the test functions.

[6] Teach the basic notions of antenna arrays especially the pattern multiplication theorem and how it can be applied to the design of the Chebyshev array which has a constant sidelobe level. In this context the notions of directivity of an antenna should be introduced and also the radar equation which states how much power can a receiver antenna aperture collect from a transmitter antenna. The directivity is the maximum of

$$D(\theta, \phi) = \frac{U(\theta, \phi)}{(4\pi)^{-1} \int U(\theta', \phi') \sin(\theta') d\theta' d\phi'}$$

where $U(\theta, \phi)$ is the power emitted by the antenna per unit solid angle in the far field zone.

[7] Teach the students about horn antennas, ie, the E -plane sectoral horn and the H -plane sectoral horn. Basically, the electromagnetic field feeding the horn antenna is achieved via a rectangular waveguide. When it meets the feeding mouth of the antenna, we calculate the distance from that point to a point on the radiating surface, usually a spherical surface, of the horn antenna and hence the phase difference between each point on the feeding mouth and each point on the radiating surface. Taking this phase difference into account and the fields on the feeding mouth surface, we can calculate the em field on the radiating surface of the horn and hence the induced surface electric and surface

magnetic current density on the horn surface as $M_S = \hat{n} \times E$ and $J_S = \hat{n} \times H$ from which using the standard retarded potential theory, the far field em field pattern is obtained.

[8] Teach about DRA's, ie, cavity resonators as antennas. Compute as mentioned earlier during the discussion on waveguides and cavity resonators, the electric and magnetic fields on the resonator surfaces at each of the discrete set of oscillation frequencies and hence obtain the induced surface electric and magnetic surface current densities on the resonator surfaces from which by applying retarded potentials the far field radiation field patterns can be obtained. It should be noted that DRA's operate at only a discrete set of frequencies unlike conventional antennas because the boundary conditions on the fields on the closed surface of the antenna causes the eigenfrequencies to assume only a discrete set of values.

5.20 Numerical methods in antenna, waveguide and cavity resonator theory

[9] Teach about numerical methods for calculating the input impedance of an antenna as well as the mutual impedance between two antennas. Basic emphasis should be on the method of moments. Given an antenna surface, we divide its area into N little pixels such that at one of the pixels a source current I_0 is given. Then we calculate the surface current density on the antenna surface by assuming a certain density, calculating the electric field produced by this density and equating its tangential components on all the pixels except the feed pixel to zero giving thereby a set of $N - 1$ equations for the surface current density at all the pixels except the feed pixel where the surface current density is known in terms of the feed current I_0 and the pixel area.

[9] Teach about the finite element method for casting DRA problems having inhomogeneous permittivity and permeability. The basic technique is based on the following idea. Assume a volume D bounded by a closed surface ∂D with a wave field $\psi(r)$ inside satisfying the generalized Helmholtz equation

$$\operatorname{div}(A(r)\nabla\psi(r)) - \lambda b(r)\psi(r) = 0$$

for $r \in D$ and $\psi(r) = 0$ for $r \in \partial D$. The possible frequencies of oscillation of this wave field correspond to the discrete set of eigenvalues λ . This equation can be derived from the variational principle

$$\delta S(\psi) = 0$$

with $\delta\psi = 0$ on ∂D , where

$$S = (1/2) \int (\nabla\psi(r), A(r)\nabla\psi(r)) d^3r + (1/2)\lambda \int b(r)\psi(r)^2 d^3r$$

assuming that $A(r)$ is a symmetric 3×3 matrix. This minimization and the consequent determination of the eigenvalues λ can be cast in discrete form using

the finite element method based on 3-dimensional pixels by assigning zero values to ψ on the boundary pixels.

5.21 Large deviation theory applied to engineering systems

While teaching the applications of probability theory to electronic circuits, definitely some attention should be given to the more recent development of Large deviation theory. In this theory, the measured random variable or random process depends on small amplitude noise and when the noise amplitude goes to zero, the measured signal becomes deterministic. Then one can ask, that if we assign a small perturbation parameter ϵ to the noise amplitude in the signal, how fast does the probability $P_\epsilon(B)$ of the noisy signal x_ϵ to enter a forbidden region B for the associated deterministic signal go to zero when the perturbation parameter goes to zero. This probability in large deviation theory is expressed in the form $\exp(-\inf_{x \in B} I(x)/\epsilon)$, or more precisely,

$$\lim_{\epsilon \rightarrow 0} \epsilon \cdot \log(P_\epsilon(B)) = -\inf_{x \in B} I(x)$$

where $I(x)$ is the rate function determined as the Legendre transform of the limiting moment generating function of x_ϵ after some normalizations, ie,

$$I(x) = \sup_{\lambda} (\langle \lambda, x \rangle - \bar{\Lambda}(\lambda))$$

where

$$\bar{\Lambda}(\lambda) = \lim_{\epsilon \rightarrow 0} \epsilon \cdot \Lambda_\epsilon(\lambda/\epsilon)$$

with

$$\Lambda_\epsilon(\lambda) = \log \mathbb{E}[\exp(\langle \lambda, x_\epsilon \rangle)]$$

This fundamental theorem in probability theory generalizes the large deviations theory for independent random variables due to Cramer and also Sanov's theorem for the large deviations of empirical measures of independent random variables and can be used to compute asymptotic statistical behaviour for empirical measures for Markov and more generally a large class of stationary processes. It is called the Gartner-Ellis theorem and its proof is not very easy. A good reference for reading about this principle is Amir Dembo and Ofer Zeitouni, "Large deviations, techniques and applications", Springer. This theory can be used to compute asymptotic probabilities for voltages and currents in linear and nonlinear circuits to cross certain amplitude barriers assuming that the thermal noise in the resistors is of very small amplitudes. This branch of probability has not been used anywhere in the engineering curriculum in India although it is a fundamental tool in applied probability. The modern theory of large deviations in arbitrary topological vector spaces is due to the famous Indian Mathematician S.R.S.Varadhan for which he received the Abel prize which is awarded for "Lifetime achievements in Mathematics" and is the mathematical equivalent of the Nobel prize.

5.22 Robotics based on nonlinear differential equations

An introduction to classical and quantum robotics. Courses on this topic can be taught to electronics students in the final year, for such courses will enable the student to see how electronics can be used to control mechanical objects. This subject is perhaps the most important existing link between electronics and mechanical engineering. There are many ways to introduce this subject. One way is to start directly with a d -link robot, each of whose links is a one dimensional rod having lengths $l_k, k = 1, 2, \dots, d$ and making angles $q_k(t), k = 1, 2, \dots, d$ with the base. The kinetic energy can be obtained by d one dimensional integrations in the form

$$K(t) = (1/2)q'^T(t)M(q(t))q'(t)$$

where

$$q(t) = [q_1(t), \dots, q_d(t)]^T \in [0, 2\pi]^d$$

and $M(q)$ is a positive definite $d \times d$ matrix whose elements are affine linear functions of the trigonometric functions $\cos(q_k), \sin(q_k)$. Let $V(q)$ denote the gravitational potential energy of the robot and $\psi(q, q')$ the frictional torque at the link joints where motors are attached providing an external torque vector $\tau(t)$. The Euler-Lagrange equations of motion are

$$\frac{d}{dt} \frac{\partial K}{\partial q'} - \frac{\partial K}{\partial q} - \frac{\partial V}{\partial q} - \psi(q, q') - \tau(t) = 0$$

This results in

$$\frac{d}{dt}(M(q)q') - (1/2)(I_d \otimes q'^T)M'(q)q' - V_{,q} - \psi(q, q') - \tau = 0$$

where

$$M'(q) = \begin{pmatrix} \frac{\partial M}{\partial q_1} \\ \vdots \\ \frac{\partial M}{\partial q_d} \end{pmatrix} \in \mathbb{R}^{d^2 \times d}$$

Define

$$N(q, q') = M'(q)^T(q' \otimes q') - (1/2)(I_d \otimes q'^T)M'(q)q' - V_{,q} - \psi(q, q')$$

This component of the torque is built out of the centrifugal terms, the Coriolis terms, the gravitational terms and the frictional terms. Then, the robot dynamical equation can be expressed as the following system of d nonlinearly coupled second order differential equations for $q(t)$.

$$M(q)q'' + N(q, q') = \tau(t) \quad \text{--- (1)}$$

5.23 Quantization of robot motion

If we wish to quantize this motion, we would introduce the Hamiltonian (obtained by Legendre transforming the conservative part of the dynamics)

$$H(t) = H(q, p, t) = (1/2)p^T M(q)^{-1}p + V(q) - \tau(t)^T q$$

and incorporating the dissipative terms using Lindblad operators. The final equation for the evolution of the density operator $\rho(t)$ in the Hilbert space $L^2((0, 2\pi)^d)$ is

$$i\rho'(t) = [H(t), \rho(t)] - (1/2)\left(\sum_{k=1}^p L_k^* L_k \rho + \rho L_k^* L_k - 2L_k \rho L_k^*\right)$$

The dual Heisenberg-GKSL equation for observables in this Hilbert space is

$$\begin{aligned} X'(t) &= i[H(t), X(t)] + (i/2)\sum_{k=1}^p (L_k^* L_k X + X L_k^* L_k - 2L_k^* X L_k) \\ &= i[H, X] + (i/2)\sum_k (L_k^*[L_k, X] + [X, L_k^*]L_k) \end{aligned}$$

By choosing the Lindblad operators L_k as appropriate functions of q, p where q, p satisfy the canonical commutation relations

$$[q_k, p_m] = i\delta_{km}$$

we can derive (1) from the equations

$$\begin{aligned} q' &= i[H, q] + (i/2)\sum_k (L_k^*[L_k, q] + [q, L_k^*]L_k), \\ p' &= i[H, p] + (i/2)\sum_k (L_k^*[L_k, q] + [q, L_k^*]L_k) \end{aligned}$$

For example, taking

$$L_k = F_k(q, p)$$

we get

$$\begin{aligned} [L_k, q_r] &= [F_k(q, p), q_r] = -iF_{k,d+r}(q, p) \\ [L_k, p_r] &= [F_k(q, p), p_r] = iF_{k,r}(q, p) \end{aligned}$$

This gives

$$\begin{aligned} &i\sum_k L_k^*[L_k, q_r] + [q_r, L_k^*]L_k \\ &= \sum_k [F_k(q, p)^* F_{k,d+r}(q, p) - F_{k,d+r}(q, p)^* F_k(q, p)] \end{aligned}$$

and likewise,

$$\begin{aligned} i \sum_k L_k^*[L_k, p_r] + [p_r, L_k^*]L_k \\ = - \sum_k [F_k(q, p)^* F_{k,r}(q, p) - F_{k,r}(q, p)^* F_k(q, p)] \end{aligned}$$

We wish to choose the functions F_k so that the GKSL observable equations for q, p result in the robot equations taking friction into account. Note that

$$\begin{aligned} i[H, q_r] &= (i/2)[p^T M(q)^{-1} p, q_r] = (i/2)[p_m(M(q)^{-1}))_{ms} p_s, q_r] \\ &= (1/2)((p^T M(q)^{-1})^T + M(q)^{-1} p) \end{aligned}$$

so that

$$i[H, q] = M(q)^{-1} p + (1/2)[p^T, M(q)^{-1}] = M(q)^{-1} p - (i/2)div_q M(q)^{-1}$$

Also,

$$\begin{aligned} i[H, p_r] &= i[(1/2)p^T M(q)^{-1} p + V(q) - \tau(t)^T q, p_r] \\ &= -V_{,r}(q) + \tau_r(t) + (i/2)p^T[M(q)^{-1}, p_r]p \\ &= -V_{,r}(q) + \tau_r(t) - (1/2)p^T(M(q)^{-1}),_r p \end{aligned}$$

Our observable GKSL equations for q, p are therefore

$$\begin{aligned} q' &= i[H, q] + (i/2) \sum_k (L_k^*[L_k, q] + [q, L_k^*]L_k) \\ &= M(q)^{-1} p - (i/2)div_q M(q)^{-1} + (1/2) \sum_k [F_k(q, p)^* F_{k,d+r}(q, p) - F_{k,d+r}(q, p)^* F_k(q, p)] \\ p' &= i[H, p] + (i/2) \sum_k (L_k^*[L_k, p] + [p, L_k^*]L_k) \\ &= (-1/2)\nabla_q(p^T M(q)^{-1} p) - \nabla_q V(q) - (1/2) \sum_k [F_k(q, p)^* F_{k,r}(q, p) - F_{k,r}(q, p)^* F_k(q, p)] + \tau(t) \end{aligned}$$

We wish these GKSL equations to yield our classical Hamiltonian equations, ie, the terms involving the operators F_k should correspond to frictional effects.

The other approach to robot quantization is via Feynman's path integral method which requires only the Lagrangian. But the path integral is hard to evaluate owing to the highly nonlinear dependence of $M(q)$ on q and further if frictional forces are present, this method will not work.

In the context of quantization of the robot dynamics, it is of interest to develop a time independent perturbation method for obtaining the stationary states and the discrete energy levels of a quantum robot in the presence of only a gravitational field and no frictional forces. This means that we must determine the eigenfunctions and eigenvalues of the Hamiltonian

$$H = (1/2)p^T M(q)^{-1} p + V(q)$$

when $[q_r, p_s] = i\delta_{rs}$. To do so we write

$$M(q) = M_0 + M_1(q)$$

where M_0 is a constant $d \times d$ symmetric diagonal matrix while $M_1(q)$ is regarded as a perturbation consisting of the trigonometric function terms in the $q_i - q_j$ variables. We can correspondingly write

$$\begin{aligned} M(q)^{-1} &= M_0^{-1} + \sum_{n \geq 1} (M_0^{-1} M_1(q))^n M_0^{-1} \\ &= M_0^{-1} + \delta J(q) \end{aligned}$$

and hence,

$$H = H_0 + \delta H, H_0 = (1/2)p^T M_0^{-1} p + V(q), \delta H = p^T \delta J(q) p / 2$$

If the eigenfunctions and eigenvalues of H_0 are known, then by standard time independent quantum mechanical perturbation theory, their perturbed versions can be computed to any order, for example, the first order shift in the energy level $E_n^{(0)}$ of the unperturbed Hamiltonian would be $\langle n | \delta J(q) | n \rangle$ where $H_0 | n \rangle = E_n^{(0)} | n \rangle$.

5.24 Filtering and control of engineering systems

Kalman and Extended Kalman filter methods may be introduced into the robotics courses both from the classical and from the quantum viewpoint. In the classical case, this may be done by first developing the Kushner-Kallianpur filtering equations for general Markov processes with WGN measurement errors. Both by directly applying Bayes rules and by using the reference probability approach, ie, the orthogonality principle (John Gough and Kostler, "Quantum filtering in coherent states") this derivation can be achieved. The KK filter is infinite dimensional in the sense that in order to compute the j^{th} conditional moments, we require the r^{th} moment for $r > j$ and hence we end up with an infinite sequence differential equations for the conditional moments. The EKF considerably simplifies this computation by making a second order Taylor approximation in the computation of the conditional moments thereby requiring only differential equations for the first two conditional moments in order to obtain the state estimate, ie, the first conditional moment.

The problem with the EKF is that it is in some sense a Gaussian approximation for non-Gaussian processes, ie, we still retain the Gaussian assumption after it passes through nonlinear systems. This problem is overcome by the UKF (The Unscented Kalman filter) and we propose to discuss this here.

The UKF: Let ξ_1, ξ_2, \dots be iid $N(0, I)$ random vectors. If P is a positive definite matrix, then $\sqrt{P}\xi_k, k = 1, 2, \dots$ are iid $N(0, P)$ random vectors. Now the state equations are

$$X(t+1) = F(t, X(t)) + W(t+1)$$

and the measurement equation is

$$Z(t) = h(t, X(t)) + V(t)$$

where W, V are discrete time iid normal random vectors. Let

$$Y(t) = \{Z(s) : s \leq t\}$$

ie the aggregate of all measurements collected upto time t . In the EKF, we write

$$\hat{X}(t+1|t) = \mathbb{E}(X(t+1)|Y(t)) = \mathbb{E}(F(t, X(t))|Y(t)) \approx F(t, \mathbb{E}(X(t)|Y(t))) = F(t, \hat{X}(t|t))$$

which is not a good approximation when F is highly nonlinear. We cannot take the expectation operator inside a highly nonlinear function without making a big error. In the UKF, we replace this by

$$\hat{X}(t+1|t) = K^{-1} \sum_{a=1}^K F(t, \hat{X}(t|t) + \sqrt{P(t|t)}\xi_a)$$

This is justified because $e(t|t) = X(t) - \hat{X}(t|t)$ is zero conditional mean with conditional covariance $P(t|t)$ and so has $\sqrt{P(t|t)}\xi_a$. The UKF computation of $P(t+1|t)$ likewise proceeds as

$$\begin{aligned} P(t+1|t) &= cov(X(t+1) - \hat{X}(t+1|t)|Y(t)) = cov(F(t, X(t)) + W(t+1) - \hat{X}(t+1|t)|Y(t)) \\ &= cov(F(t, X(t)) - \hat{X}(t+1|t)) = cov(F(t, \hat{X}(t|t) + e(t|t)) - \hat{X}(t+1|t)) \\ &= K^{-1} \sum_{a=1}^K (F(t, \hat{X}(t|t) + \sqrt{P(t|t)}\xi_a) - \hat{X}(t+1|t))''^T \end{aligned}$$

where $('')$ the same expression as the bracket before it. A more convenient way to write this is

$$\begin{aligned} e(t+1|t) &= X(t+1) - \hat{X}(t+1|t), \\ e^a(t+1|t) &= F(t, \hat{X}(t|t) + \sqrt{P(t|t)}\xi_a) - \hat{X}(t+1|t), \\ P(t+1|t) &= K^{-1} \cdot \sum_{a=1}^K e^a(t+1|t) \cdot (e^a(t+1|t))^T \end{aligned}$$

In the UKF to compute $\hat{X}(t+1|t+1)$ and $P(t+1|t+1)$ approximately, we use the well known formula that if U, V are jointly Gaussian random vectors, then

$$\mathbb{E}(U|V) = \mathbb{E}(U) + cov(U, V) \cdot cov(V)^{-1} (V - \mathbb{E}(V))$$

and

$$\text{cov}(U|V) = \text{cov}(U) - \text{cov}(U, V).\text{cov}(V).\text{cov}(V, U)$$

In our context, we shall use it in the following form:

$$\hat{X}(t+1|t+1) = \mathbb{E}(X(t+1)|Y(t), Z(t+1)) \approx$$

$$\hat{X}(t+1|t) - \text{cov}(e(t+1|t), e_Z(t+1|t)).\text{cov}(Z(t+1)|Y(t))^{-1}(Z(t+1) - \hat{Z}(t+1|t))$$

where we define

$$\begin{aligned} e_Z(t+1|t) &= h(t+1, X(t+1)) + V(t+1) - K^{-1} \sum_{a=1}^K h(t+1, \hat{X}(t+1|t) + e^a(t+1|t)) \\ &= h(t+1, \hat{X}(t+1|t) + V(t+1) + e(t+1|t)) - K^{-1} \sum_{a=1}^K h(t+1, \hat{X}(t+1|t) + e^a(t+1|t)) \end{aligned}$$

and hence

$$\begin{aligned} \text{cov}(e(t+1|t), e_Z(t+1|t)) &\approx \\ K^{-1} \sum_{b=1}^K e^b(t+1|t). (e_Z^b(t+1|t))^T & \end{aligned}$$

where

$$e_Z^b(t+1|t) = h(t+1, \hat{X}(t+1|t) + e^b(t+1|t)) - K^{-1} \sum_{a=1}^K h(t+1, \hat{X}(t+1|t) + e^a(t+1|t))$$

Also, in the UKF approximation,

$$\text{cov}(Z(t+1)|Y(t)) = \text{cov}(e_Z(t+1)|Y(t)) = P_V + K^{-1} \sum_{b=1}^K e_Z^b(t+1|t). (e_Z^b(t+1|t))^T$$

and

$$\begin{aligned} \hat{Z}(t+1|t) &= \mathbb{E}(h(t+1, X(t+1))|Y(t)) = \mathbb{E}(h(t+1, \hat{X}(t+1|t) + e(t+1|t))|Y(t)) \\ &= K^{-1} \cdot \sum_{a=1}^K h(t+1, \hat{X}(t+1|t) + e^a(t+1|t)) \end{aligned}$$

Finally, in the UKF approximation,

$$\begin{aligned} P(t+1|t+1) &= \text{cov}(X(t+1) - \hat{X}(t+1|t+1)|Y(t), Z(t+1)) \\ &= \text{cov}(X(t+1) - \hat{X}(t+1|t)|Y(t), Z(t+1)) \\ &= \text{cov}(e(t+1|t)|Y(t), Z(t+1)) = \\ \text{cov}(e(t+1|t)|Y(t)) - \text{cov}(e(t+1|t), Z(t+1)|Y(t)).\text{cov}(Z(t+1)|Y(t))^{-1}.\text{cov}(e(t+1|t), Z(t+1)|Y(t))^T \end{aligned}$$

where

$$\text{cov}(e(t+1)|Y(t)) = P(t+1|t),$$

$$\text{cov}(e(t+1|t), Z(t+1)|Y(t)) = \text{cov}(e(t+1|t), e_Z(t+1|t)|Y(t))$$

$$= K^{-1} \cdot \sum_{a=1}^K e^a(t+1|t) \cdot (e_Z^a(t+1|t))^T,$$

$$\text{cov}(Z(t+1)|Y(t)) = \text{cov}(e_Z(t+1)|Y(t)) = K^{-1} \cdot \sum_{a=1}^K e_Z^a(t+1) \cdot (e_Z^a(t+1))^T + P_V$$

This completes the description of the UKF

Simultaneous filtering and control of a robot to track a given trajectory must be taught both at the classical level and at the quantum level. The reason for this is that we cannot obtain accurate noiseless measurements of the robot state at any given time using optical encoders. We can only obtain noisy versions of some functions of the state at each time, but in order to control the robot so that it tracks a given trajectory, we must give a an error feedback torque to it between the desired state trajectory and the current state trajectory. Since the current state trajectory measurements are noisy, we must first use a filter or an observer that will produce real time estimates of the robot trajectory based on noisy measurements and then our error to be fed back as a torque to the robot dynamics will be between the desired state and the filtered/observed state. The EKF is just one kind of observer. In general, an observer to the robot state dynamics

$$dX(t) = F(t, X(t))dt + G(t, X(t))dB(t)$$

will be constructed as

$$d\hat{X}(t) = F(t, \hat{X}(t))dt + L(t, \hat{X}(t))(dZ(t) - h(t, \hat{X}(t))dt) \quad \dots \quad (1)$$

where the measurement differential is

$$dZ(t) = h(t, X(t))dt + dV(t) \quad \dots \quad (2)$$

Heuristically, this process of error feedback can be interpreted as follows. Suppose the observer gain $L(t, \hat{X}(t))$ is a positive definite matrix. The estimated output at time t is $h(t, \hat{X}(t))dt$ and if this is much smaller than the true output $dZ(t)$, then $dZ(t) - h(t, \hat{X}(t))dt > 0$ and the positivity of L ensures that we give positive error feedback in order to cause an increase in $\hat{X}(t)$ so that the estimated output increases toward the true output and vice versa. Herein we are assuming that h is some sort of increasing function of the state $X(t)$. This crude justification is valid in one dimension but also works in multiple dimensions by diagonalizing the matrix L having positive eigenvalues. Having thus obtained the state estimate, we may given an error feedback for tracking the desired state trajectory $X_d(t)$ by modifying the state dynamics to

$$dX(t) = F(t, X(t))dt + G(t, X(t))dB(t) + K(t)(X_d(t) - \hat{X}(t))dt \quad \dots \quad (3)$$

where $K(t)$ is the feedback gain matrix. The set of equations (1), (2) and (3) can be linearized around $\hat{X}(t)$ obtaining thereby linear sde's for the two kinds of

errors, one the trajectory tracking error $e(t) = X_d(t) - X(t)$ and two the state observer estimation error $f(t) = X(t) - \hat{X}(t)$. We note that $X_d(t) - \hat{X}(t) = e(t) + f(t)$. It is a simple matter to perform this and obtain approximate mean and covariance propagation equations for these in the form of coupled linear sde's.

Lyapunov energy theory can also be used to construct the feedback controller and also study the asymptotics of the rate of change of its average over all statistical ensembles.

5.25 Quantum many body systems applied to Fermi operator fields and superconductivity

The next important topic to be taught or initiated as projects is the computation of the current in a superconductor by an external classical or an external quantum electromagnetic field. We adopt the BCS theory of superconductivity in which the electron operator field satisfies the canonical anticommutation relations in the second quantized picture and the Hamiltonian of this field consists of a Schrodinger part with a Hartree-Fock interaction part and also a BCS part involving contribution to the potential by Cooper pairs. When two fermions combine and form a Cooper pair which is a Boson, these Bosons travel through the Fermi liquid with almost zero resistance yielding to the generation of supercurrents. The easiest way to compute this supercurrent is through the field Hamiltonian approach followed by Schrodinger's representation of the current density in terms of the wave field operators, their adjoints and the magnetic vector potential. The average of this current density in the Gibbs state can be expressed as the Gibbs average of quadratic forms in the wave operator fields which constitute the Green's function and the dynamics of the Green's function can be expressed easily by making use of the Heisenberg dynamics of the wave-field operators.

Specifics of the BCS model: We start with the wave-field Fermionic operators

$$\psi_a(r), \psi_a(r)^*, r \in \mathbb{R}^3$$

satisfying the anticommutation relations

$$\{\psi_a(r), \psi_b(r')^*\} = \delta_{ab}\delta^3(r - r'), \{\psi_a(r), \psi_b(r')\} = 0, \{\psi_a(r)^*, \psi_b(r')^*\} = 0$$

After time t , under the influence of the self-consistent Hartree-Fock Hamiltonian taking into account the Cooper pairing, the evolution is still unitary with the unitary evolution operator depending on the wave fields. Hence, the same anticommutation relations are valid at a fixed time t for $\psi_a(t, r), \psi_a(t, r'), a = 1, 2, \dots, p$. The Hartree-Fock Hamiltonian with Cooper pairing is derived from the Hamiltonian

$$K = (-\hbar^2/2m) \int \psi_a(r)^* (\nabla + ieA(t, r)/\hbar) \psi_a(r) d^3r$$

$$+ \int (V_0(r) + V(t, r))\psi_a(r)^*\psi_a(r) + \int g_2(r)\psi_a(r)^*\psi_a(r)\psi_b(r)^*\psi_b(r)d^3r$$

It should be now noted that the Hartree-Fock-BCS Hamiltonian is derived from this by replacing the product in the last term by

$$\begin{aligned} \psi_a(r)^*\psi_a(r)\psi_b(r)^*\psi_b(r) &\approx \\ <\psi_a(r)^*\psi_b(r)> \psi_a(r)\psi_b(r)^* - <\psi_a(r)^*\psi_b(r)^*> \psi_a(r)\psi_b(r) \\ - <\psi_a(r)\psi_b(r)> \psi_a(r)^*\psi_b(r)^* \end{aligned}$$

Keeping this in mind, we define the gap functions

$$\Delta_{ab}(r) = <\psi_a(r)\psi_b(r)>$$

and another auxiliary number density and correlation functions

$$N_{ab}(r) = <\psi_a(r)^*\psi_b(r)>$$

We thus postulate the following HF-BCS Hamiltonian

$$K(t) = K_0(t) + V_{HF} + V_{BCS}$$

where

$$\begin{aligned} K_0(t) &= (-h^2/2m) \int \psi_a(r)^*(\nabla + ieA(t, r)/h)^2\psi_a(r)d^3r \\ &+ \int (V_0(r) + V(t, r))\psi_a(r)^*\psi_a(r), \\ V_{HF} &= \int g_1(r)N_{ab}(r)\psi_a(r)\psi_b(r)^*d^3r, \end{aligned}$$

$$V_{BCS} = \int (g_2(r)\Delta_{ab}(r)\psi_a(r)^*\psi_b(r)^* + \bar{g}_2(r)\bar{\Delta}_{ba}(r)\psi_a(r)\psi_b(r))d^3r$$

where we note that

$$\bar{\Delta}_{ba}(r) = <\psi_a(r)^*\psi_b(r)^*>$$

In these formulae, averages are taken w.r.t the Gibbs state

$$\rho(\beta) = \exp(-\beta K_{00})/Z(\beta), Z(\beta) = \text{Tr}(\exp(-\beta K_{00}))$$

where

$$K_{00} = (-h^2/2m) \int \psi_a(r)^*\nabla^2\psi_a(r)d^3r + \int V_0(r)\psi_a(r)^*\psi_a(r)d^3r$$

The Heisenberg evolution equations for the Fermionic fields are

$$-ih\psi_{a,t}(t, r) = [K(t), \psi_a(t, r)]$$

Now, using the Fermionic field anticommutation relations, we have

$$[K(t), \psi_a(r)] = T_1 + T_2 + T_3 + T_4$$

where

$$\begin{aligned} T_1 &= (-h^2/2m) \int [\psi_b(r')^*(\nabla' + ieA(t, r')/h)^2 \psi_b(r'), \psi_a(r)] d^3r' \\ &= (h^2/2m)(\nabla + ieA/h)^2 \psi_a(r) \end{aligned}$$

where we used

$$\begin{aligned} \psi_b(r')\psi_a(r) &= -\psi_a(r)\psi_b(r'), \\ \psi_b(r')^*\psi_a(r) &= \delta_{ab}\delta^3(r - r') - \psi_a(r)\psi_b(r') \end{aligned}$$

Note that $A(t, r)$ and $V(t, r)$ are the quantum em potentials expressible in terms of creation and annihilation operators of the photon field. Specifically, in the Lorentz gauge, they can be expressed in the form

$$A(t, r) = \sum_k a_k f_k(t, r) + a_k^* \bar{f}_k(t, r),$$

$$V(t, r) = -c^2 \int_0^t \operatorname{div} A(s, r) ds = \sum_k a_k g_k(t, r) + a_k^* \bar{g}_k(t, r)$$

where

$$g_k(t, r) = -c^2 \int_0^t \operatorname{div} f_k(s, r) ds$$

In the literature while computing the supercurrent density as

$$J(t, r) = <(ie/2m)(\psi(t, r)^*(\nabla + ieA(t, r)/h)\psi(t, r) - ((\nabla - ieA(t, r)/h)\psi(t, r)^*)\psi(t, r))>$$

we assume that $A(t, r)$ is a classical field and that the average is taken w.r.t $\rho(\beta)$. However, we are assuming that that em field is a quantum field and hence we may compute the average current density w.r.t the Fermion-Boson state $\rho(\beta, u) = \rho(\beta) \otimes |\phi(u)\rangle \langle \phi(u)|$ where $|\phi(u)\rangle$ is the coherent state for the photon field:

$$|\phi(u)\rangle = \exp(-\|u\|^2/2) \sum_{\mathbf{n} \geq 0} u_1^{n_1} \dots u_k^{n_k} \dots |\mathbf{n}\rangle / \sqrt{n_1! n_2! \dots}$$

where

$$\mathbf{n} = (n_1, n_2, \dots),$$

$$|\mathbf{n}\rangle = |n_1, n_2, \dots\rangle = a_1^{*n_1} a_2^{*n_2} \dots |0\rangle / \sqrt{n_1! n_2! \dots}$$

We next observe that

$$\begin{aligned} T_2 &= \left[\int (V_0(r') + V(t, r')) [\psi_b(r')^* \psi_b(r'), \psi_a(r)] d^3r' \right. \\ &\quad \left. - (V_0(r) + V(t, r)) \psi_a(r) \right] \\ T_3 &= \left[\int g_1(r') N_{bc}(r') [\psi_b(r') \psi_c(r')^*, \psi_a(r)] d^3r' \right. \end{aligned}$$

$$= g_1(r)N_{ba}(r)\psi_b(r)$$

(summation over the repeated index b is understood). Finally,

$$\begin{aligned} T_4 &= \int (g_2(r')\Delta_{bc}(r')[\psi_b(r')^*\psi_c(r')^*, \psi_a(r)] + \bar{g}_2(r')\bar{\Delta}_{cb}(r')[\psi_b(r')\psi_c(r'), \psi_a(r)])d^3r' \\ &= \int g_2(r')\Delta_{bc}(r')(\psi_b(r')^*(\delta_{ca}\delta^3(r - r') - \psi_a(r')\psi_c(r'))d^3r' \\ &= g_2(r)\Delta_{ba}(r)\psi_b(r)^* - g_2(r)\Delta_{ac}(r)\psi_c(r)^* \\ &= 2g_2(r)\Delta_{ba}(r)\psi_b(r)^* \end{aligned}$$

Combining all these expressions finally results in the following Schrodinger equation for the Fermionic field:

$$\begin{aligned} &i\hbar\psi_{a,t}(t, r) = \\ &(-\hbar^2/2m)(\nabla + ieA(t, r)/\hbar)^2\psi_a(t, r) + (V_0(r) + V(t, r))\psi_a(t, r) \\ &- g_1(r)N_{ba}(r)\psi_b(t, r) - 2g_2(r)\Delta_{ba}(r)\psi_b(t, r)^* \end{aligned}$$

This equation is nonlinear since the functions $\Delta_{ab}(r)$ and $N_{ab}(r)$ depend on $\psi_c(r), \psi_c^*(r), c = a, b$ which are respectively the values of $\psi_c(t, r)$ and $\psi_c(t, r)^*$ at $t = 0$. Now, we define two kinds of Green's functions:

$$G_{ab}(t, r|t', r') = \langle T\{\psi_a(t, r)\psi_b(t', r')^*\} \rangle,$$

$$F_{ab}(t, r|t', r') = \langle T\{\psi_a(t, r)\psi_b(t', r')\} \rangle$$

where T denotes the Fermionic time ordering operator:

$$T\{X(t)Y(t')\} = \theta(t - t')X(t)Y(t') - \theta(t' - t)Y(t')X(t)$$

We first observe that

$$\Delta_{ab}(r) = -\lim_{t' \rightarrow t+, r' \rightarrow r} F_{ba}(t', r'|t, r) \quad (1)$$

$$N_{ab}(r) = -\lim_{t' \rightarrow t+, r' \rightarrow r} G_{ba}(t', r'|t, r) \quad (2)$$

We next derive a differential equation for these two Green's functions. Before that we note that the average current density in the tensor product state ρ for the Fermionic and photonic field can be expressed in terms of the Green's function G_{ab} :

$$\begin{aligned} J(t, r) &= \langle (ieh/2m)(\psi_a(t, r)^*(\nabla + ieA(t, r)/\hbar)\psi_a(t, r) - ((\nabla - ieA(t, r)/\hbar)\psi_a(t, r)^*)\psi_a(t, r)) \rangle \\ &= (-ieh/2m)\lim_{t' \rightarrow t+, r' \rightarrow r} \nabla_r \langle T\{\psi_a(t, r)\psi_a(t', r')^*\} \rangle \\ &\quad + (ieh/2m)\lim_{t' \rightarrow t+, r' \rightarrow r} \nabla_r \langle T\{\psi_a(t', r')\psi_a(t, r)^*\} \rangle \\ &\quad + (e^2/m) \langle A(t, r) \rangle \lim_{t' \rightarrow t+, r' \rightarrow r} [G_{aa}(t', r'|t, r) \\ &= \lim_{t' \rightarrow t+, r' \rightarrow r} [(-ieh/2m)\nabla_r[G_{aa}(t, r|t', r') - G_{aa}(t', r'|t, r)] \end{aligned}$$

$$+(e^2/m) < A(t, r) > G_{aa}(t', r' | t, r)] -- (3)$$

The differential equations for the Green's functions are obtained as follows:

$$\partial_t G_{ab}(t, r | t', r') = \partial_t [< \psi_a(t, r) \psi_b^*(t', r') > \theta(t-t') - < \psi_b^*(t', r') \psi_a(t, r) > \theta(t'-t)]$$

$$= \delta(t-t') < \{ \psi_a(t, r), \psi_b^*(t, r') \} > + [< (\partial_t \psi_a(t, r)) . \psi_b^*(t', r') > \theta(t-t')$$

$$- < \psi_b^*(t', r') \partial_t \psi_a(t, r) > \theta(t'-t)]$$

$$= \delta_{ab} \delta(t-t') \delta^3(r-r') + < T\{(\partial_t \psi_a(t, r)) . \psi_b^*(t', r')\} >$$

Thus, we get using the above Schrodinger equation for ψ_a after correlating with ψ_b^* and following it up with the time ordering operation, and likewise, correlating with ψ_b followed by time ordering,

$$ih \partial_t G_{ab}(t, r | t', r') = ih \delta(t-t') \delta^3(r-r')$$

$$-(h^2/2m)(\nabla + ieA(t, r)/h)^2 G_{ab}(t, r | t', r') + (V_0(r) + V(t, r)) G_{ab}(t, r | t', r')$$

$$-g_1(r) N_{ca}(r) G_{cb}(t, r | t', r') - 2g_2(r) \Delta_{ca}(r) \bar{F}_{bc}(t', r' | t, r) -- (4)$$

$$ih \partial_t F_{ab}(t, r | t', r') =$$

$$-(h^2/2m)(\nabla + ieA(t, r)/h)^2 F_{ab}(t, r | t', r') + (V_0(r) + V(t, r)) F_{ab}(t, r | t', r')$$

$$-g_1(r) N_{ca}(r) F_{cb}(t, r | t', r') - 2g_2(r) \Delta_{ca}(r) \bar{G}_{bc}(t', r' | t, r) -- (5)$$

Equations (1)-(5) constitute the fundamental equations for calculating the current in a superconductor in the presence of an electromagnetic field, quantum or classical.

Reference:Fetter and Walecka, "Quantum theory of many particle systems", Dover.

We shall discuss some approximate methods for solving for the Green's functions G_{ab}, F_{ab} in the presence of an em field when the solutions are known in its absence.

In matrix notation, we can express the above differential equations for the Green's functions as

$$\partial_t G(t, t') = ih \delta(t-t') I + L_0 G(t, t') + e A_t . L_1 G(t, t') + (e^2/2m) A_t^2 G(t, t')$$

$$-g_1 N^T . G(t, t') - 2g_2 \Delta^T \bar{F}^T(t', t)$$

and likewise,

5.26 Lie group theory in image processing

The next topic which must also be introduced at the undergraduate level in the form of small projects is the applications of Lie-group and Lie-algebra theory to image processing and robotics. In image processing problems, the typical way in which group theory arises is as follows: We are given a manifold \mathcal{M} like $\mathbb{R}^2, \mathbb{R}^2, S^2$ etc. on which a Lie group G acts. For example, G may be $E(2)$, the Euclidean group $\mathbb{R}^2 \otimes_s SO(2)$ of rotations and translations of the plane, or $E(3)$, the Euclidean group $\mathbb{R}^3 \otimes_s SO(3)$ of translations and rotations of three dimensional space, or the Galilean group $(\mathbb{R}^3 \times \mathbb{R}^3) \otimes_s (\mathbb{R} \times SO(3))$ of spatial translations, uniform spatial motions, time delay/advancement, rotations acting on the space-time manifold $\mathbb{R} \times \mathbb{R}^3$, or more generally, we can also include the scaling group \mathbb{R}_+ in the Galilean group. Now given an image field $f : \mathcal{M} \rightarrow \mathbb{C}$, we may apply a G -transformation element $g \in G$ on this image field and then observe the transformed image f_1 after corrupted by noise:

$$f_1(x) = f(g^{-1}x) + w(x), x \in \mathcal{M}$$

One problem would be to estimate g from measurements of the original and transformed and blurred image f_1 . For example, if $w(x)$ is a white Gaussian noise field, the maximum likelihood estimator of g would be something like

$$\hat{g} = \operatorname{argmin}_g \int_{\mathcal{M}} |f_1(x) - f(g^{-1}x)|^2 h(x) d\mu(x) \quad \text{--- (a)}$$

where μ is a G -invariant measure on \mathcal{M} and $1/2h(x)$ is the spectral strength of $w(x)$, ie,

$$\mathbb{E}(w(x)w(y)) = \delta(x - y)/2h(x)$$

One way to construct the invariant measure μ on \mathcal{M} would be to take a point $x_0 \in \mathcal{M}$ and consider the map $\tau : G \rightarrow \mathcal{M}$ by $\tau(g) = gx_0$. Then, if μ_G is a left invariant Haar measure on G , it follows that $\mu = \mu_G \circ \tau^{-1}$ is a G -invariant measure on \mathcal{M} . This is because for any $g \in G$,

$$\begin{aligned} \int_{\mathcal{M}} f(gx) d\mu(x) &= \int_{\mathcal{M}} f(gx) d\mu_G \circ \tau^{-1}(x) \\ &= \int_G f(g\tau(h)) d\mu_G(h) = \int_G f(ghx_0) d\mu_G(h) = \int_G f(hx_0) d\mu_G(h) \\ &= \int_{\mathcal{M}} f(x) d\mu_G \circ \tau^{-1}(x) = \int_{\mathcal{M}} f(x) d\mu(x) \end{aligned}$$

The problem of solving the optimization problem in (a) is very hard and involves a computationally intensive search over G . The theory of group representations greatly simplifies this optimization problem and this can be used as a starting point in courses for introducing the theory of group representations. For example, if π is a representation of G , we get in the absence of noise,

$$\hat{f}_1(\pi) = \int_G f_1(hx_0) \pi(h) dh = \int_G f(g^{-1}hx_0) \pi(h) dh$$

$$= \int f(hx_0)\pi(gh)dh = \pi(g)\hat{f}(\pi)$$

and hence by comparing $\hat{f}_1(\pi)$ and $\hat{f}(\pi)$ for different representations π , we get to know $\pi(g)$ for different π using which reconstruction of g can be achieved.

5.27 Lie group based robotics

In the context of robotics, suppose we have d 3-D links with the $j + 1^{th}$ link attached to the j^{th} link at a fixed point p_j at time $t = 0$, we find that a point ξ in the k^{th} link at time $t = 0$ moves after time t to the point

$$\xi(t) = R_1(t)p_1 + R_2(t)R_1(t)p_2 + \dots + R_{k-1}(t)\dots R_1(t)p_{k-1} + R_k(t)R_{k-1}(t)\dots R_1(t)(\xi - p_{k-1})$$

where $R_1(t)$ is the rotation suffered by the first link about the origin and for $k = 2, 3, \dots$, $R_k(t)$ is the rotation suffered by the k^{th} link relative to the $k - 1^{th}$ link around its point of attachment $R_{k-1}(t)\dots R_1(t)p_{k-1}$. We define

$$S_k(t) = R_k(t)R_{k-1}(t)\dots R_1(t) \in SO(3), k = 1, 2, \dots$$

Thus,

$$\xi(t) = \sum_{j=1}^{k-1} S_j(t)p_j + S_k(t)(\xi - p_{k-1})$$

and this gives the kinetic energy of the k^{th} link as

$$\begin{aligned} T_k(t) &= (\rho/2) \int_{B_k} \| \xi'(t) \|^2 d^3\xi \\ &= (\rho/2) \int_{B_k} \| \sum_{j=1}^{k-1} S'_j(t)p_j + S'_k(t)(\xi - p_{k-1}) \|^2 d^3\xi \end{aligned}$$

The total kinetic energy of the d-link robot is therefore of the form

$$T(t) = \sum_{k=1}^d T_k(t) = (1/2) \sum_{k,m=1}^d Tr(S'_k(t)J_{km}S'_m(t))$$

where J_{km} are constant 3×3 matrices satisfying $J_{km}^T = J_{mk}$. The gravitational potential energy of the robot is given by

$$V(t) = g \sum_{k=1}^d M_k(e_3^T S_k(t)l_k + \sum_{j=1}^{k-1} e_3^T S_j(t)p_j)$$

where l_k is the position vector of the centre of gravity of the k^{th} link relative to its point p_{k-1} of attachment to the $k - 1^{th}$ link at time $t = 0$. This shows

that $V(t)$ is a linear function of $\{S_k(t) : 1 \leq k \leq d\}$ while $T(t)$ is a quadratic function of $\{S'_k(t) : 1 \leq k \leq d\}$. The Lagrangian of the robot defined as

$$L(t, S_k(t), S'_k(t), 1 \leq k \leq d) = T(t) - V(t)$$

is therefore linear in the $S_k(t)$'s and quadratic in the $S'_k(t)$'s. The external torques act at the joints $p_0 = 0, p_1, \dots, p_{d-1}$ via motors. Let $\tau_k(t)$ denote the machine torque pseudo-vector at the k^{th} joint. Then, its contribution to the Lagrangian must be described appropriately. The k^{th} link suffers a rotation $R_k(t) = S_k(t)S_{k-1}(t)^{-1}$ relative to the $k-1^{th}$ link. In time t to $t+dt$, the angle pseudo-vector of rotation suffered by the k^{th} link relative to the $k-1^{th}$ link is $d\theta\hat{n}$ where $dR_k(t)\xi = d\theta\hat{n} \times R_k(t)\xi$ for any vector ξ . Thus, if $X = (X_1, X_2, X_3)$ denote the generators of $SO(3)$, then

$$dR_k(t)\xi = d\theta(t)(\hat{n}.X)R_k(t)\xi$$

or

$$d\theta(\hat{n}.X) = dR_k(t).R_k(t)^{-1}$$

This equation can be solved for $d\theta\hat{n}$ by using the identity $Tr(X_iX_j) = -2\delta_{ij}$ and therefore, $Tr((\hat{n}.X)X_k) = -2n_k$. The energy spent by the motor in performing this infinitesimal rotation is $\tau_k(t)^T\hat{n}d\theta$. We have from the above,

$$d\theta n_k = d\theta(-1/2)Tr((\hat{n}.X)X_k) = (-1/2)Tr(dR_k(t).R_k(t)^{-1}X_k)$$

and the contribution to the action functional (ie, time integral of the Lagrangian) by the motors is given by

$$\begin{aligned} S_T &= \sum_{k=1}^d \int_0^t \tau_k(t)^T \hat{n}(t) d\theta(t) = \\ &= (-1/2) \sum_{k=1}^d \sum_{j=1}^3 \int_0^t \tau_{kj}(t) Tr(R'_k(t)R_k(t)^{-1}X_j) dt \end{aligned}$$

We define the antisymmetric torque tensor by

$$\mathcal{T}_k(t) = \sum_{j=1}^3 \tau_{kj}(t) X_j$$

Then, we can write

$$S_T = (-1/2) \sum_{k=1}^d \int_0^t Tr(R'_k(t)R_k(t)^{-1}\mathcal{T}_k(t)) dt = \int_0^t L_T dt$$

where L_T is the motor torque contribution to the Lagrangian and is given by

$$L_T = (-1/2) \sum_{k=1}^d Tr(R'_k(t)R_k(t)^{-1}\mathcal{T}_k(t))$$

5.28 Levy process models for jerk noise in robotic systems

The robot differential equations have the general form

$$dX(t) = F(t, X(t), \theta)dt + \sum_{k=1}^p G_k(t, X(t), \theta)dN_k(t)$$

where $N_k(\cdot)$, $k = 1, 2, \dots, p$ are independent Poisson processes with rates λ_k , $k = 1, 2, \dots, p$ respectively. Our measurement model is

$$dZ(t) = h(t, X(t))dt + \sigma_V dV(t)$$

where $V(\cdot)$ is standard vector valued Brownian motion. Let $Z_t = \{Z(s) : s \leq t\}$. We wish to construct the EKF for estimating $X(t), \theta$ dynamically. Thus,

$$d\theta(t) = 0$$

and our extended state vector is

$$\xi(t) = [X(t)^T, \theta(t)^T]^T$$

This is a Markov process and its generator is given by

$$\mathbb{E}[d\psi(X(t), \theta(t))|X(t) = x, \theta(t) = \theta]$$

$$= dt[\psi_{,x}(t, \theta)^T F(t, x, \theta) + \sum_k \lambda_k (\psi(x + G_k(t, x, \theta), \theta) - \psi(x, \theta))] = dt.K_t\psi(x, \theta)$$

K is the generator of the process $\xi(t)$ and its kernel is given by

$$K_t(x, \theta|x', \theta') = F(t, x, \theta^T)(\nabla\delta(x-x'))\delta(\theta-\theta') + \sum_k \lambda_k (\delta(x-x'+G_k(t, x, \theta)) - \delta(x-x'))\delta(\theta-\theta')$$

The Kushner-Kallianpur filter for $\phi(x, \theta)$ with

$$\pi_t(\phi) = \mathbb{E}[\phi(X(t), \theta(t))|Z_t]$$

is given by

$$d\pi_t(\phi) = \pi_t(K_t\phi)dt + \sigma_V^{-2}(\pi_t(h_t\phi) - \pi_t(h_t)\pi_t(\phi))(dZ(t) - \pi_t(h_t)dt)$$

We then make the EKF approximations. First observe that

$$\begin{aligned} K_t\xi &= K_t[x^T, \theta^T]^T = [F(t, x, \theta)^T, 0^T]^T + \sum_k \lambda_k [G_k(t, x, \theta)^T, 0^T]^T \\ &= [F(t, x, \theta)^T + \sum_k \lambda_k G_k(t, x, \theta)^T, 0^T]^T \end{aligned}$$

So,

$$\pi_t(K_t \xi) \approx [F(t, \hat{X}(t), \hat{\theta}(t))^T + \sum_k \lambda_k G_k(t, \hat{X}(t), \hat{\theta}(t))^T, 0^T]^T$$

Also,

$$\begin{aligned} \pi_t(\xi \cdot h_t) - \pi_t(\xi) \pi_t(h_t) &\approx \\ \pi_t(\delta \xi(t) \cdot h'_t(\hat{\xi}(t)) \delta \xi(t)) \\ &= P_\xi(t) h'_t(\hat{\xi}(t))^T \end{aligned}$$

This gives the approximate EKF equation for the conditional mean:

$$\begin{aligned} d\hat{\xi}(t) &= [F(t, \hat{\xi}(t))^T + \sum_k \lambda_k G_k(t, \hat{\xi}(t))^T, 0^T]^T dt \\ &+ \sigma_V^{-2} P_\xi(t) h'_t(\hat{\xi}(t))^T (dZ(t) - h_t(\hat{\xi}(t))) dt \end{aligned}$$

For the error covariance

$$P_\xi(t) = Cov(\delta \xi(t) | Z_t), \delta \xi(t) = \xi(t) - \hat{\xi}(t)$$

we start with

$$\begin{aligned} d(\delta \xi(t) \delta \xi(t)^T) &= (d\delta \xi(t)) \cdot \delta \xi(t)^T + \delta \xi(t) \cdot d\delta \xi(t)^T \\ &+ d\delta \xi(t) \cdot d\delta \xi(t)^T \end{aligned}$$

Now,

$$d\delta \xi(t) = d\xi(t) - d\hat{\xi}(t) =$$

$$\begin{aligned} &\tilde{F}(t, \xi(t)) dt + \tilde{G}_k(t, \xi(t)) dN_k(t) - \tilde{F}(t, \hat{\xi}(t)) dt - \sum_k \lambda_k \tilde{G}_k(t, \hat{\xi}(t)) dt - \sigma_V^{-2} P_\xi(t) H_t^T (dZ - h_t(\hat{\xi}(t))) dt \\ &\approx (\tilde{F}(t, \xi(t)) - \tilde{F}(t, \hat{\xi}(t))) dt + \sum_k \tilde{G}_k(t, \hat{\xi}(t)) dM_k(t) \\ &- \sigma_V^{-2} P_\xi(t) H_t^T (H_t \delta \xi(t) + \sigma_V dV(t)) \\ &\approx \tilde{F}'(t, \hat{\xi}(t)) \delta \xi(t) dt + \sum_k \tilde{G}_k(t, \hat{\xi}(t)) dM_k(t) \\ &- \sigma_V^{-2} P_\xi(t) H_t^T (H_t \delta \xi(t) + \sigma_V dV(t)) \end{aligned}$$

where

$$M_k(t) = N_k(t) - \lambda_k t$$

is a discontinuous Martingale. It follows (using $dV \cdot dV^T = dt I, dM_k \cdot dM_j = \delta_{kj} dN_k, dV \cdot dN_k = 0$) that

$$dP_\xi(t)/dt \approx$$

$$\begin{aligned} &\tilde{F}'(t, \hat{\xi}(t)) P_\xi(t) + P_\xi(t) F'(t, \hat{\xi}(t))^T - \sigma_V^{-2} P_\xi(t) H_t^T H_t P_\xi(t) \\ &+ \sum_k \lambda_k \tilde{G}_k(t, \hat{\xi}(t)) G_k(t, \hat{\xi}(t))^T \end{aligned}$$

This forms the EKF for Poisson noise driven state with white Gaussian measurement noise. Note the notation used:

$$H_t = h'_t(\hat{\xi}(t)), \tilde{F}(t, \xi) = [F(t, \xi)^T, 0]^T, \tilde{G}_k(t, \xi) = [G_k(t, \xi)^T, 0]^T$$

5.29 Digital systems, classical and quantum gates, design of counters using shift registers and flip-flops

For a long time, a lot of digital systems based on implementing Boolean functions and counters has been taught to electronics students in undergraduate courses. Now the time has come to explain that digital gates are generally irreversible and moreover, only a limited amount of classical information can be transmitted over a classical channel ie a channel having a i/o transition probability matrix $((p(k, m)))_{1 \leq k, m \leq N}$. The Boolean function that represents the output of a JK flip-flop giving $Q_{n+1} = f(Q_n, J_n, K_n) = \bar{J}_n \bar{K}_n Q_n + \bar{J}_n \bar{K}_n Q_n + J_n \bar{K}_n$ or that of a D or T flip flop in the form $Q_{n+1} = f(Q_m, D_n) = D_n, Q_{n+1} = f(Q_n, T_n) = Q_n \bar{T}_n + \bar{Q}_n T_n$ should be perhaps introduced via its truth table. The excitation table for the JK flip-flop should then be introduced giving possible values of (J_n, K_n) as Boolean functions of Q_n, Q_{n+1} and then how such an excitation table may be used to design a counter in a shift register. How the figure eight consisting of seven linear bulbs can be used to represent all numbers from zero to nine and which of these lights should be switched on to get a decimal number in this range from a four digit binary number using the theory of Boolean functions should be introduced. After introducing all this, the microprocessor should be taught explaining how to write programs using its standard commands like *MOV*, *MVI*, *PUSH*, *JUMP* for some simple operations like solving simple difference equations, adding numbers from one to N , calculating the DFT of a sequence etc. The functions of the Data bus and address bus should be taught by simple commands like given a binary data loaded on the data bus through an input port and given a specified register location defined by a binary data string on the address bus, how to transfer the data from the data bus to the register should be explained. Simple binary operations like the half-adder, full-adder, multiplication with floating points should be introduced via Boolean functions and how these functions are built into the microprocessor should be taught.

Having taught all these, the irreversibility of gates such as not and, nand, or, exor etc. which have just one binary output and two binary inputs should be explained. In this context, the notion of a 2×2 unitary quantum gate acting on single qubit state and its reversibility should be explained. The notion of an r qubit gate as a unitary $2^r \times 2^r$ matrix acting on an r -qubit quantum gate may be touched upon. Specific examples like the NOT gate which is a two qubit quantum gate, one qubit being a controlling qubit and the other qubit

being the controlled qubit may be dealt with. The SWAP gate, phase gate, Hadamard gates, the Fredkin gates and how these gates may be used to build the quantum Fourier transform gate of size $2^r \times 2^r$ acting on r qubits using just $O(2^r)$ operations in contrast to $O(r \cdot 2^r)$ operations required by a classical FFT should be explained. How to use entanglement as a resource for transmitting a single qubit quantum state using just transmission of two classical bits should be explained (teleportation). More generally how to use entanglement to transmit r -qubits by transmitting just $2r$ classical bits should be explained. The notion of a qubit consisting of a 2×1 complex vector having unit norm expressible as a superposition of a $|0\rangle$ state and a $|1\rangle$ state may be explained. How a qubit carries far more information than a classical bit may be touched upon. The interpretation of an r -qubit quantum state in terms of quantum probabilities must be explained thoroughly. In this context, the notion of mixed quantum states and how the state collapses after a measurement is made in the quantum theory unlike the classical theory should be explained.

Some applications of quantum computation like phase finding, order finding, factorizing a number etc. are very important. These ideas revolve around applying a control unitary operator to a state, taking a measurement following this application and obtaining the required result with a high probability on making a measurement of some component of the state.

Other applications revolve around quantum communication which is essentially of two kinds, one Cq channel communication in which a classical source alphabet with some probability distribution is encoded into a quantum state. The Cq-channel consists of the transformation $x \rightarrow \rho(x)$ of each source alphabet into a quantum state on the same Hilbert space and then to recover the sent alphabet with high probability by making a measurement on the received state. Just as in Shannon's classical information theory, the error probabilities will be large, however, if we consider a string $u = (x_1 x_2 \dots x_n)$ of source alphabets, and encode this string into the tensor product state $\rho(u) = \bigotimes_{k=1}^n \rho(x_k)$ then if n is sufficiently large, we can form M_n such strings with M_n greater than $2^{n(\max_p I(p, \rho) - \delta)}$ such that the error probability is arbitrarily small and δ is also arbitrarily small. Here

$$I(p, \rho) = S\left(\sum_x p(x)\rho(x)\right) - \sum_x p(x)S(\rho(x))$$

is the mutual information between the source signal and the received state, $\{p(x)\}$ is the source probability distribution and S is the Von-Neumann entropy function of a state. Moreover, the converse of this coding theorem also holds, ie, if we choose any set of M_n string of source alphabets such that $\log(M_n)/n > \max_p I(p, \rho) + \delta$ for some $\delta > 0$, then no matter what decision operators we choose, the error probability will converge to unity as $n \rightarrow \infty$. There are many ways to prove this fundamental result based on actual construction of the decision POVM operators and random coding arguments. These latest developments which generalize Shannon's coding theorem are due to Holevo and Winter. The notion of detection of the source string with small error probability means that for any $\epsilon, \delta > 0$ and all sufficiently large n , we can

construct $M = M_n$ sequences u_{n1}, \dots, u_{nM_n} of source alphabets, each of length n , positive operators $D_{n1}, \dots, D_{nM}, M = M_n$ such that $\sum_{k=1}^M D_{nk} \leq I$ and $\text{Tr}(\rho(u_{nk})D_{nk}) > 1 - \epsilon, k = 1, 2, \dots, M_n$ such that $\log(M_n)/n > C - \delta$ where $C = \max_p I(p, \rho)$ and further C is the largest such number ie if $\log(M_n)/n > C + \delta$ for all sufficiently large n , then $\min_k \text{Tr}(\rho(u_{nk})D_{nk}) \rightarrow 0$. C is called the Cq capacity of the channel. My feeling is that it is more advantageous to prove the quantum noiseless Schumacher compression theorem and the quantum Cq noisy Shannon coding theorem and then show the classical theorems of Shannon are special cases of this.

5.30 HMM and some of its applications

[1] Application to an image processing problem. We wish to compress a dynamically changing image field. In the continuous spatial domain, the image intensity field $\phi(t, x, y), (x, y) \in D, t \in [0, NT]$ is modeled by the pde

$$\sum_{a,b=1}^p D_{ab}(t) \partial^2 \phi(t, x, y) / \partial x_a \partial x_b + w(t, x, y) = \partial \phi(t, x, y) / \partial t$$

ie a diffusion equation with slowly time varying diffusion matrix $D(t) = ((D_{ab}(t)))$. We assume that the diffusion matrix $D(t)$ is slowly varying in the sense that it is a constant over the time slot $[mT, (m+1)T]$ and then makes a transition to another matrix in the next time slot $[(m+1)T, (m+2)T]$. The aim is firstly to estimate the diffusion matrix values over each slot and then to reestimate their transition probabilities from the measurements of ϕ over the whole time duration $[0, NT]$. Assume that the diffusion matrix assumes values in a finite set $\mathcal{E} = \{D_1, \dots, D_K\}$ and that the Markov transition probabilities are $\pi(k, m) = P(D(nT + 0) = D_m | D(nT - 0) = D_k)$. Assuming then that the image field has been discretized in space and time, we replace the above equation by its discretized version

$$\phi(t+1, x, y) = \phi(t, z, y) + \delta w(t, x, y) + \delta_0 \sum_{a,b} D_{ab}[n] \Delta_a \Delta_b \phi(t, x, y), nT \leq t < (n+1)T$$

and the probability of making these observations can be expressed as

5.31 Quantum Image Processing

After introducing concepts such as diffusion equation based edge smoothing and image enhancement and histogram equalization of classical images (ie, constructing a probability density for image pixel grey-scale levels by maximizing the entropy of the pdf over each image patch subject to moment constraints) the course should aim at dealing with problems of modeling the electromagnetic

field coming from the image as a quantum em field described in terms of creation and annihilation operators. When such a quantum em field hits an object that is moving or rotating, some of it gets scattered and some of it passes through the object. If we model the dynamically varying object parameters as $\theta(t), t \geq 0$, and the quantum em field as

$$A(t, r) = \sum_k (a_k f_k(t, r) + a_k^* \bar{f}_k(t, r)), [a_k, a_j^*] = \delta_{kj}$$

then after passing through the moving object, this received em field has the form

$$A_1(t, r) = \sum_k a_k \chi_k(\theta(s) : s \leq t) + a_k^* \bar{\chi}_k(\theta(s) : s \leq t)$$

and after this field interacts with an atomic receiver or a lattice of atomic receivers, each of which has a free valence electron, the interaction Hamiltonian can be expressed as

$$V(t) = \sum_k [a_k \otimes \xi_k(\theta(s) : s \leq t) + a_k^* \otimes \xi_k^*(\theta(s) : s \leq t)]$$

where ξ_k are atomic observables depending on the parameter history in some specific way. Now let $|p\rangle$ be the initial state of the lattice of atomic electrons and let the coherent state $|\phi(u)\rangle$ be the initial state of the quantum em field. Thus,

$$a_k |\phi(u)\rangle = u_k |\phi(u)\rangle,$$

$$a_k^* |\phi(u)\rangle = (\partial/\partial u_k - \bar{u}_k) |\phi(u)\rangle$$

The initial state of the atomic lattice and the field is the tensor product $|\phi(u)p\rangle$ and the final state is $|\phi(u)q\rangle$. The first order transition probability from the former to the latter in time T is obtained using quantum mechanical time dependent perturbation theory as

$$P_T(|p\rangle \rightarrow |q\rangle |u\rangle) =$$

$$|\int_0^T \langle \phi(u)q | V(t) | \phi(u)p \rangle \exp(i(E(q) - E(p))t) dt|^2$$

Now,

$$\langle \phi(u)q | V(t) | \phi(u)p \rangle =$$

$$\begin{aligned} & \sum_k (\langle \phi(u) | a_k | \phi(u) \rangle \langle q | \xi_k | p \rangle + \langle \phi(u) | a_k^* | \phi(u) \rangle \langle q | \xi_k^* | p \rangle) \\ &= 2\text{Re}[\sum_k u_k \langle q | \xi_k(\theta(s) : s \leq t) | p \rangle] \end{aligned}$$

which means that by measuring the transition probabilities $P_T, T \geq 0$ between two atomic states with the quantum image field remaining in the same coherent state, we get information about the object motion parameters $\theta(t), t \geq 0$.

If for example, the moving object is a fan, then $\theta(t)$ satisfies a second order differential equation with noise present and we can use the EKF to estimate it dynamically based on noisy measurements of the transition probabilities. More generally, if the object is also of quantum size, then its parameter $\theta(t)$ will be an observable in another Hilbert space which may satisfy a quantum stochastic differential equation like the Hudson-Parthasarathy equation and then we must use the Belavkin quantum filter to estimate this observable dynamically from our measurements.

Other problems in quantum image processing:

[a] How to teach the basic concepts in quantum image processing based on static filters. The quantum image field is a state of the electromagnetic field that can be expressed either as a pure state

$$|I\rangle = \sum_{k,m} c(k,m) |k, m\rangle$$

or more generally as a mixed state of the form

$$\rho_I = \sum_{k,m,j,l} |k, m\rangle p(k, m, j, l) \langle j, l|$$

In these expressions, $|k, m\rangle$ denotes the state of the k^{th} image pixel when its grey scale level is specified by the index m . Thus, if $|I\rangle$ is the quantum image and a measurement in the standard basis $\{|k, m\rangle \langle k, m| : k, m\}$ is made, then the probability that the k^{th} pixel will be selected and that this pixel is has the grey scale level m is given by $|c(k, m)|^2$. On the other hand, if the image field is in the mixed state ρ_I , then if the above measurement is made, the probability that the k^{th} pixel is selected and that it has the grey scale level m is given by

$$\langle k, m | \rho_I | k, m \rangle = p(k, m, k, m)$$

By processing a quantum image, we mean that we apply a quantum operation \mathbf{t} to it, ie, \mathbf{t} is a TPCP map and hence admits the Stinespring-Choi-Kraus representation,

$$(\rho) = \sum_r E_r \rho E_r^*, \sum_r E_r^* E_r = I$$

Thus, the output image state after processing and adding quantum noise is given by

$$\rho_o = \sum_r E_r \rho_I E_r^*$$

The problem is to recover ρ_I from ρ_o by the use of recovery operators. This problem is solved completely by the Knill-Laflamme theorem which states that

[b] If we are to apply Belavkin's theory of filtering based on non-demolition measurements, we must first couple the image field to a bath described by Boson-Fock space and introduce creation, annihilation and conservation quantum noise

processes into the joint unitary evolution of the image field and bath. After unitarily evolving for time t , we partially trace out over the bath state to obtain the output image state. This state is of the form (1) to which the Knill-Laflamme theorem can be applied:

$$\begin{aligned}\rho_o &= Tr_2(U(T)(\rho_I \otimes |\phi(u)\rangle\langle\phi(u)|)U(T)^*) \\ &= \sum_r E_r \rho_I E_r^*\end{aligned}$$

where the system operators E_r are defined in terms of the system operators that drive the Hudson-Parthasarathy unitary evolution $U(t)$, ie, in terms of H, L_1, L_2, S, P where

$$dU(t) = (-(iH + P)dt + L_1 \otimes dA(t) - L_2 \otimes dA(t)^* + S(t) \otimes d\Lambda(t))U(t)$$

Exercise, by using time dependent perturbation theory with the perturbation parameter δ attached to L_1, L_2 and δ^2 attached to P , give approximate formulae for the operators E_k in terms of L_1, L_2, P, H . Assume that the bath is in the coherent state $|\phi(u)\rangle$.

5.32 Introduce the design aspects of some gadgets through mini-projects some examples

[a] Consider a pendulum rod with a magnet attached to its end. Below the rod, when it is in the vertical hanging state is present an electromagnetic coil powered by an ac source of frequency ω . The pendulum moves under the joint action of the torque induced by gravity and the torque produced by the electromagnetic coil acting on the magnetic dipole of the suspended magnet. This torque can be computed as $\mu \times B(t, \theta)$ where μ is the magnetic moment of the permanent magnet attached to the pendulum rod and $B(t, \theta)$ is the magnetic field produced by the electromagnetic coil on at the site of the pendulum magnet when the pendulum makes an angle θ with the vertical. The equation of motion of the pendulum taking into account both the the torques and a frictional retarding torque proportional to the angular velocity of the pendulum can thus be expressed as

$$ml^2\theta''(t) = -\gamma\theta'(t) - mgl\sin(\theta(t)) + \mu B(t, \theta(t))$$

Since the current in the electromagnet is ac with a frequency of say ω rad/sec, we can write

$$B(t, \theta) = B_0(\theta)\sin(\omega t)$$

and we have in perturbation theoretic form, the equation of motion

$$ml^2\theta''(t) + \gamma\theta'(t) + (mgl - B'_0(0))\theta(t) = \delta(-mgl(\sin(\theta(t)) - \theta(t)) +$$

Remark: A working model of this gadget was shown to me by my colleague Professor D.V.Gadre.

5.33 A simple way to introduce quantum electrodynamics

There are broadly speaking two ways to introduce quantum electrodynamics to undergraduate and post-graduate students, one the operator theoretic formalism of Schwinger, Tomonaga and Dyson and secondly the path integral formalism of Richard Feynman. The second is more intuitive and enables us to directly obtain transition probabilities for scattering of electrons, positrons and photons and also things like vacuum polarization in which a photon splits into an electron and positron (pair creation) and then again combines (pair annihilation) to once again obtain a photon. The representation of pair production and annihilation is depicted using Feynman diagrams as a circle with two photon lines attached to it, one an incoming photon line and two an outgoing photon line. Let D denote the bare photon propagator. Then the correction to the photon propagator caused by vacuum polarization yields the corrected photon propagator in the form

$$D_c = (D^{-1} + \Sigma)^{-1} \approx D - D\Sigma.D$$

The exact representation of this equation is

$$D_c^{-1} = D^{-1} + \Sigma$$

or equivalently, after premultiplying this equation by D_c and postmultiplying by D , we get

$$D = D_c + D_c\Sigma D$$

or after rearranging,

$$D_c = D - D_c\Sigma D$$

This equation is exact. Another equivalent exact representation of this corrected propagator equation is

$$D_c = D - D\Sigma D_c$$

Likewise, the path integral approach enables us to calculate the self energy of the electron which states that an electron accelerates thereby emitting electromagnetic radiation and then the electron interacts with this self induced radiation and moves. Here, Σ is calculated using integrals of products of the electron propagator. The self-energy of the electron is represented by a Feynman diagram as before but with the photon and electron propagators interchanged. This gives the radiation corrected electron propagator as

$$S_{ce} = S_e - S_e D_p S_{ce}$$

The reason why the Feynman path integral approach can be used to derive these corrected propagators is that the Feynman path integrals for electrons and photons can readily expressed as the sum of a sequence of Gaussian integrals with imaginary covariance matrices and then we can use formulae for the higher order moments of a Gaussian random vector to obtain these propagator corrections.

For example, consider the action functional for photons, electrons and positrons interacting with each other:

$$S[A, \psi] = (-1/16\pi) \int F_{\mu\nu} F^{\mu\nu} d^4x + \int \bar{\psi} \gamma^\mu (i\partial_\mu + eA_\mu) - m \psi d^4x$$

where

$$\bar{\psi} = \psi^* \gamma^0, F_{\mu\nu} = A_{\nu,\mu} - A_{\mu,\nu}$$

This action integral can be expressed as the sum of a quadratic functional of A_μ , a quadratic functional of ψ, ψ^* and an interaction term that is cubic in $A_\mu, \psi, \bar{\psi}$. The trick to calculating the corrected electron or photon propagator is to expand the exponential of the cubic term and then perform the path integration. These expanded terms will appear as moments of a Gaussian distribution functional. For example, the corrected electron propagator is

$$\begin{aligned} S_{eclm}(x, y) &= \int \exp(i \int (L_{EM} + L_D + L_{int}) d^4z) \psi_l(x) \bar{\psi}_m(y) \Pi_{x \in \mathbb{R}^4} d\psi_l(z) d\bar{\psi}_l(z) dA_\mu(z) \\ &= \int \exp(i \int (L_{EM}(z) + L_D(z)) d^4z (1 + \sum_{n \geq 1} ((i \int L_{int}(z) d^4z)^n / n!)) \psi_l(x) \bar{\psi}_m(y) \Pi_{x \in \mathbb{R}^4, k, \mu} d\psi_k(z) d\bar{\psi}_k(z) dA_\mu(z) \\ &= \int \exp(i \int L_D d^4z) \psi_l(x) \bar{\psi}_m(y) d\psi . d\bar{\psi} + \\ &+ \int \exp(i \int (L_{EM}(z) + L_D(z)) d^4z) \sum_{n \geq 1} ((i \int L_{int}(z) d^4z)^n / n!) \psi_l(x) \bar{\psi}_m(y) d\psi . d\bar{\psi} . dA \end{aligned}$$

The first term on the rhs is the free electron propagator, ie, without any radiation corrections and the second term is the radiation introduced corrections to the electron propagator. Here,

$$\begin{aligned} \int L_{EM}(z) d^4z &= (-1/8\pi) \int (A_{\nu,\mu}(z) A^{\nu,\mu}(z) - A_{\nu,\mu}(z) A^{\mu,\nu}(z)) d^4z \\ &= \int A_\nu(z) K_{p\nu\mu}(z, z') A_\mu(z') d^4z d^4z' \end{aligned}$$

where

$$\begin{aligned} K_{p\nu\alpha}(z, z') &= (-1/8\pi) \eta^{\nu\alpha} \eta^{\mu\beta} \partial_\mu \partial_\beta \delta^4(z - z') \\ &+ (1/8\pi) \eta^{\mu\alpha} \eta^{\nu\beta} \partial_\mu \partial_\beta \delta^4(z - z') \\ &= (1/8\pi) [\partial^\alpha \partial^\nu - \eta^{\nu\alpha} \square] \delta^4(z - z') \end{aligned}$$

where

$$\square = \partial^\alpha \partial_\alpha$$

Also,

$$\int L_D(z) d^4z = \int \bar{\psi}(z) (\gamma^\mu i\partial_\mu - m) \psi(z) d^4z$$

$$= \int \bar{\psi}(z) K_e(z, z') \psi(z') d^4z d^4z'$$

where

$$K_e(z, z') = (\gamma^\mu i\partial_\mu - m)\delta^4(z - z')$$

In the four-momentum domain, we can express these kernels as

$$K_{p\mu\nu}(z, z') = (2\pi)^{-4} \int \hat{K}_{\mu\nu}(p) \exp(-ip.(z - z')) d^4p$$

where

$$\hat{K}_{\mu\nu}(p) = \int K_{p\mu\nu}(z) \exp(ip.z) d^4z = \eta^{\mu\nu} p^2 - p^\mu p^\nu$$

where

$$p^2 = p^\mu p_\mu = \eta^{\mu\nu} p_\mu p_\nu$$

The 4×4 matrix with entries $\hat{K}_{\mu\nu}(p)$ is not invertible since

$$\hat{K}_{\mu\nu}(p)p_\nu = 0$$

To calculate the bare photon propagator, using the standard formula for the covariance of Gaussian random vector, we require it to be invertible. Therefore we replace $\hat{K}_{\mu\nu}(p)$ by $\eta^{\mu\nu} p^2 - \alpha p^\mu p^\nu$ where $\alpha \neq 1$ and after making all the calculations using the inverse of this as the bare photon propagator, we let $\alpha \rightarrow 1$. Consider the equation

$$(p^2 \eta^{\mu\nu} - \alpha p^\mu p^\nu) a_\nu = b^\mu$$

This becomes,

$$p^2 a^\mu - \alpha p^\mu (p.a) = b^\mu$$

and hence,

$$a^\mu = (b^\mu + \alpha(p.a)p^\mu)/p^2$$

whence,

$$p.a = b.p/p^2 + \alpha(p.a)$$

so that

$$p.a = b.p/p^2(1 - \alpha)$$

from which we get

$$a^\mu = b^\mu/p^2 + \alpha(b.p)p^\mu/(p^2)^2(1 - \alpha) = Q_\nu^\mu b^\nu$$

or equivalently,

$$a_\mu = b_\mu/p^2 + \alpha(b.p)p_\mu/(p^2)^2(1 - \alpha) = Q_{\mu\nu} b^\nu$$

where

$$Q_{\mu\nu} = \eta_{\mu\nu}/p^2 + \alpha p_\mu p_\nu/(p^2)^2(1 - \alpha) \quad (1)$$

This should be taken as the photon propagator. It should at this point be noted that during calculations involving the photon propagator, the interaction energy involved is $\int J^\mu A_\mu d^3x$. The corresponding transition probability amplitudes or scattering matrix elements caused by this interaction will therefore according the Dyson series involve terms like

$$\begin{aligned} & < T\{J^{\mu_1}(x_1)A_{\mu_1}(x_1)\dots J^{\mu_k}(x_k)A_{\mu_k}(x_k)\} > \\ & = < T\{J^{\mu_1}(x_1)\dots J^{\mu_k}(x_k)\} > < T\{A_{\mu_1}(x_1)\dots A_{\mu_k}(x_k)\} > \end{aligned}$$

and the second term by the Gaussian moment theory is a product of photon propagators. Here $J^\mu(x) = -e\bar{\psi}(x)\gamma^\mu\psi(x)$ is the Dirac current. This Dirac current is conserved, ie, $\partial_\mu J^\mu = 0$. Thus in the momentum domain we must have $p_\mu J^\mu = 0$ which means that

$$Q_{\mu\nu}J^\nu = \eta_{\mu\nu}p^2J^\nu/p^2$$

Thus for all practical purposes involving interaction Hamiltonians between the current and the electromagnetic field we can drop out the second term in (1) and hence the effective photon propagator is

$$Q_{\mu\nu} = \frac{\eta_{\mu\nu}}{p^2}$$

Another way to see this is a consequence of the Lorentz gauge condition $\partial^\mu A_\mu = 0$. This implies that $\partial^\mu Q_{\mu\nu} = 0$ or in the momentum domain $p^\mu Q_{\mu\nu} = 0$ which is why the second term in (1) is not of any physical consequence during the calculations of the scattering matrix. In the calculation of the corrected photon propagator using the above mentioned scheme, we'll get the correction term as

$$\int \exp(i \int (L_{EM}(z) + L_D(z))) (\sum_{n \geq 1} (i \int L_{int}(z) d^4z)^n / n!) ((A_\mu(x)A_\nu(y)d\psi.d\bar{\psi}.dA$$

The $n = 1$ term for an external photon line corresponding to a classical em field A_μ^c is

$$\int \exp(i \int (L_{EM}(z) + L_D(z))) \int \bar{\psi}(z)\gamma^\rho\psi(z)A_\rho^c(z)d^4z A_\mu(x)A_\nu(y)$$

This evaluates to using the Gaussian formula for the second order moments,

$$D_{\mu\nu}(x, y) \int A_\rho^c(z) Tr(S_e(z, z)) d^4z$$

which does not make much sense. Here,

$$S_e(z, z') = \int \exp(i \int L_D(z) d^4z) \psi(z)\bar{\psi}(z')d\psi.d\bar{\psi}$$

is the electron propagator and it equals using the Gaussian formula for the second moments,

$$S_e(z, z') = \int \hat{S}_e(p) \exp(-ip.(z - z')) d^4 p$$

where

$$\hat{S}_e(p) = (\gamma.p - m)^{-1} = \frac{\gamma.p + m}{p^2 - m^2} >$$

However if we consider the $n = 2$ term assuming that the em fields are quantum fields, then we get as the correction term to the photon propagator:

$$D_2(x, y) = \int \exp(i \int (L_{EM}(z) + L_D(z))) \int \bar{\psi}(z) \gamma^\rho \psi(z) \bar{\psi}(z') \gamma^\alpha \psi(z') \\ \times A_\rho(z) A_\alpha(z') d^4 z d^4 z' A_\mu(x) A_\nu(y) d\psi.d\bar{\psi}.dA$$

This is an eighth Gaussian moment computation. It is easily seen from this expression that the "Gaussian random variables" $\{\psi(z), \bar{\psi}(z) : z \in \mathbb{R}^4\}$ are "statistically independent" of the "Gaussian random variables" $\{A_\mu(z) : z \in \mathbb{R}^4, \mu = 0, 1, 2, 3\}$. We shall be denoting the bare photon propagator $Q_{\mu\nu}$ by $D_{\mu\nu}$ and the bare electron propagator by S_e . The various terms that appear after performing the above Gaussian moment integral can be expressed as

$$(\gamma^\rho)_{ab} (\gamma^\alpha)_{cd} \int [< T\{\bar{\psi}_a(z) \psi_b(z) \bar{\psi}_c(z') \psi_d(z')\} > \\ \times < T\{A_\rho(z) A_\alpha(z') A_\mu(x) A_\nu(y)\} > d^4 z d^4 z' \\ = \gamma^\rho)_{ab} (\gamma^\alpha)_{cd} \int [< T\{\bar{\psi}_a(z) \psi_b(z)\} > < T\{\bar{\psi}_c(z') \psi_d(z')\} > + \\ < T\{\bar{\psi}_a(z) \bar{\psi}_c(z')\} > < T\{\psi_b(z) \psi_d(z')\} > + < T\{\bar{\psi}_a(z) \psi_d(z')\} > < T\{\psi_b(z) \bar{\psi}_c(z')\} > \\ \times [< T\{A_\rho(z) A_\alpha(z')\} > < T\{A_\mu(x) A_\nu(y)\} > + \\ < T\{A_\rho(z) A_\mu(x)\} > < T\{A_\alpha(z') A_\nu(y)\} > + < T\{A_\rho(z) A_\nu(y)\} > < T\{A_\alpha(z') A_\mu(x)\} >] d^4 z d^4 z'$$

The only terms which are of significance in the above expansion for determining corrections to the photon propagator are

$$(\gamma^\rho)_{ab} (\gamma^\alpha)_{cd} \int [< T\{\bar{\psi}_a(z) \psi_d(z')\} > < T\{\psi_b(z) \bar{\psi}_c(z')\} >] \\ \times [< T\{A_\rho(z) A_\mu(x)\} > < T\{A_\alpha(z') A_\nu(y)\} > + < T\{A_\rho(z) A_\nu(y)\} > < T\{A_\alpha(z') A_\mu(x)\} >] d^4 z d^4 z'$$

These evaluate to

$$(\gamma^\rho)_{ab} (\gamma^\alpha)_{cd} \int S_{da}(z' - z) S_{bc}(z - z') [D_{\rho\mu}(z - x) D_{\alpha\nu}(z' - y) \\ + D_{\rho\nu}(z - y) D_{\alpha\mu}(z' - x)] d^4 z d^4 z' \\ = \int Tr(\gamma^\rho S(z - z') \gamma^\alpha S(z' - z)) [D_{\rho\mu}(z - x) D_{\alpha\nu}(z' - y) \\ + D_{\rho\nu}(z - y) D_{\alpha\mu}(z' - x)] d^4 z d^4 z'$$

5.34 How to teach the theory of non-Abelian gauge theories as non-commutative generalizations of electromagnetism

First we start with describing electromagnetism as a $U(1)$ gauge theory of fields coupled to the Schrodinger electron field and Dirac electron-positron field. The Schrodinger equation can be expressed as

$$((-h^2/2m)(\nabla + ieA/h)^2 - eV)\psi = ih\partial_t\psi$$

Now suppose we make a $U(1)$ transformation on ψ :

$$\psi \rightarrow \exp(i\phi/h)\psi = \tilde{\psi}$$

Note that ψ and ϕ are both functions of the space-time coordinates $x = (t, r)$. Then we have

$$\begin{aligned}\psi &= \exp(-i\phi/h)\tilde{\psi} \\ \nabla\psi &= \exp(-i\phi/h)(-ie\nabla\phi/h + \nabla)\tilde{\psi} \\ (\nabla + ieA/h)\psi &= \exp(-i\phi/h)(\nabla + ieA/h - ie\nabla\phi/h)\tilde{\psi}\end{aligned}$$

Hence,

$$(\nabla + ieA/h)^2\psi = \exp(-i\phi)(\nabla + ieA/h - i\nabla\phi/h)^2\tilde{\psi}$$

Also,

$$ih\partial_t\psi = \exp(-i\phi/h)(ih\partial_t + e\partial_t\phi)\tilde{\psi}$$

Thus, the above Schrodinger equation can be expressed as

$$(ih\partial_t + e(V + \partial_t\phi))\tilde{\psi} = (-h^2/2m)(\nabla + (ie/h)(A - \nabla\phi))^2\tilde{\psi}$$

Thus, the $U(1)$ -transformed wave function $\tilde{\psi}$ satisfies the same Schrodinger equation but with V replaced by $V + \partial_t\phi$ and A replaced $A - \nabla\phi$. This transformation of the potentials is the standard gauge transformation that leaves the electromagnetic field invariant. Note that the em fields are defined in terms of the potentials V, A by

$$E = -\nabla V - \partial_t A, B = \nabla \times A$$

More generally, if the wave function ψ takes values in \mathbb{C}^N .

5.35 How to introduce astronomy and cosmology to undergraduates

Astronomy must be taught first of all by teaching the students to construct an inverting telescope using two convex lenses of focii f_1 and f_2 with $f_1 \gg f_2$. These two lenses are rolled inside thick paper tubes and slided into one another.

Looking through the eyepiece lens f_1 with the objective lens f_2 pointed at a distant object, by sliding the two tubes inside each other until the distant scene appears focussed and magnified, ie, brought near, the telescope can be used to see clearly the craters of the moon, the rings of Saturn and the four moons of Jupiter. The magnification of a distant object can be shown to equal f_1/f_2 by drawing the ray diagram and using properties of similar triangles. The student must be asked to derive this formula on his own. The telescope should never be pointed at the sun and viewed through the eyepiece. This would instantly burn the retina and total blindness would result. The sunspots must be projected via the telescope using a plain white sheet on the eyepiece side with the objective lens pointed at the sun and the tubes slided one inside the other until focus is obtained. The next step to be taught in astronomy is the notion of redshift of receding galaxies. Powerful radio telescopes are required to detect this red-shift thereby establishing the expansion of the universe and the fact that the galaxies are comoving, ie, the velocity of a star relative to an observer at a given time depends only on its comoving radial coordinate. This fact is expressed using Hubble's law $v = H(t)r$. This should be the starting point for introducing the Robertson-Walker metric for a homogeneous and isotropic universe in terms of comoving coordinates, ie, the geodesics correspond to constant spatial coordinates. Comoving coordinates can best be illustrated by marking coloured points on a balloon with latitudes and longitudes and blowing the balloon so that it expands radially outwards with the marked points always having the same latitudes and longitudes. The scale factor of the universe $S(t)$ enters into the Robertson-Walker metric and Hubble's constant is then $H(t) = S'(t)/S(t)$ since the radial distance of a point at time t having comoving radial coordinate r is $r(t) = S(t)r$ and hence its radial velocity is $v(t, r) = S'(t)r = (S'(t)/S(t))r(t) = H(t)r(t)$, ie, the radial velocity of a star is proportional to its radial distance from the observer located at the origin.

After this basic introduction, the notion of maximally symmetric spaces should be introduced by proving the theorem that if the space is an N dimensional Riemannian manifold, then it can have at-most $N(N + 1)/2$ linearly independent Killing vectors which may be chosen as $N(N - 1)/2$ rotations and N quasi-translations. The characterization of Killing vectors as those vector fields that define infinitesimal coordinate transformations leaving the metric invariant should be taught and the differential equations satisfied by them. The radial trajectories of light rays in Robertson-Walker metric should be introduced by illustrating how the redshift of receding galaxies can be explained from this. If L is the luminosity of an object having comoving radial coordinate r_1 and it emits a light pulse during the time interval $[t_1, t_1 + dt_1]$, then this pulse arrives at r_0 during the time interval $[t_0, t_0 + dt_0]$ is such a way that

$$dt/S(t) = dr$$

so that

$$r_1 - r_0 = \int_{t_1}^{t_0} dt/S(t) = \int_{t_1+dt_1}^{t_0+dt_0} dt/S(t)$$

yielding the red-shift formula,

$$dt_0/S(t_0) = dt_1/S(t_1)$$

or in terms of frequencies of the pulse,

$$\nu_0/\nu_1 = dt_1/dt_0 = S(t_1)/S(t_0)$$

For an expanding universe, $S(t_1) < S(t_0)$ and hence,

$$\nu_1 > \nu_0$$

implying the redshift. This formula is valid for zero spatial curvature spaces. For other curvatures k taking values $-1, 0, 1$, the redshift formula is the same:

$$\int_{r_0}^{r_1} dr/\sqrt{1 - kr^2} = \int_{t_1}^{t_0} dt/S(t) = \int_{t_1+dt_1}^{t_0+dt_0} dt/S(t)$$

The implications of this redshift formula are several, starting from calculations of the light energy collected by the aperture of a telescope at comoving radial coordinate r_0 due to a source of luminosity L located at the comoving radial coordinate r_1 . The energy of a photon of frequency ν is $h\nu$. Hence, if L is the luminosity of the source, ie the energy emitted by the source per unit time, then $(Ldt_1/h\nu_1)$ is the number of photons emitted by the source per unit time and the number of photons collected by a telescope of aperture area A_r at the origin is thus

$$(Ldt_1/h\nu_1)A_r/4\pi(S(t_0)r_1)^2$$

These photons arrive with a frequency of ν_0 over the time interval $[t_0, t_0 + dt_0]$ and hence the total power collected by the receiver telescope is

$$P_r = [(Ldt_1/h\nu_1)A_r/4\pi(S(t_0)r_1)^2]h\nu_0/dt_0$$

Now using the formula,

$$\nu_0/\nu_1 = S(t_1)/S(t_0) = dt_1/dt_0$$

we can write this as

$$P_r = LS(t_1)^2A_r/(4\pi S(t_0)^4r_1^2)$$

and the students may be taught how to use this formula for calculating the luminosity of the source.

Prerequisites for cosmology: While illustrating the fundamental principles of cosmology from an observational viewpoint as well as from the theoretical viewpoint, it is instructive to give the students a thorough grounding in tensor calculus and Riemannian geometry while simultaneously illustrating the basic principles of general relativity like the geodesic equations as equations describing the motion of a particle in a gravitational field, the notion of covariant

derivative and its applications to formulating the Maxwell equations and Dirac equation in coordinate independent formats. The weak and strong principles of equivalence should be stated and explained by means of examples. The weak principle of equivalence states that the laws of motion do not depend on the observer. This implies as regards the motion of material particles in a constant gravitational field that the gravitational field can be cancelled out by performing our experiments in a freely falling elevator. More generally, a non-constant gravitational field can be locally cancelled out by using a freely falling elevator within an infinitesimal region of space-time. The weak principle of equivalence leads us immediately to the geodesic equations of motion of material particles in a gravitational field while simultaneously interpreting the gravitational field as a distortion in the geometry of flat space-time. The strong principle of equivalence goes even further as it leads to the prediction that gravity affects all the physical phenomena like electromagnetism, fluid flow, the laws of quantum mechanics and quantum field theory. It states that by passing over to a locally freely falling frame, one can locally cancel out the effect of gravity on any physical phenomenon. Thus, it naturally leads to the fact that ordinary partial derivatives in the field equations for any theory must be replaced by covariant derivatives. Use of covariant derivatives guarantees that the field equations are all tensor equations and hence valid in all frames, ie, for all observers when they are valid in one frame. Thus the strong principle of equivalence can alternately be stated as "All the laws of physics are valid in all frames". The experiments of Eotvos for testing the principle of equivalence should be included by interpreting the weak principle of equivalence in the form "gravitational mass an inertial mass of a body are proportional".

5.36 Quantum image processing revisited

The state of the quantum image at time $t = 0$ is given by

$$|\psi_I(t)\rangle = \sum_{k=0}^{N-1} \sum_{m=0}^{K-1} c(k, m|\theta) |k, m\rangle$$

where θ is the unknown parameter vector to be estimated and $|k, m\rangle$ denotes the state of the image when the k^{th} pixel is measured and the grey scale level of that pixel is m . In this model, there is interference between the grey scale level of different pixels. For example, if pixels and grey scale levels are measured, ie, if the measurement system is $\mathbf{M} = \{|k, m\rangle\langle k, m| : k, m\}$, then the probability of getting $|k, m\rangle$ is $|c(k, m|\theta)|^2$. On the other hand, if we choose a measurement system in which one of the measurements is $|\chi\rangle = (|k_1, m_1\rangle + |k_2, m_2\rangle)/\sqrt{2}$ with outcome ω , then the probability of getting this superposed state is

$$|\langle \chi | \psi_I(t) \rangle|^2 = |c(k_1, m_1|\theta) + c(k_2, m_2|\theta)|^2/2$$

in contrast to the value $(|c(k_1, m_1|\theta)|^2 + |c(k_2, m_2|\theta)|^2)/2$ that we would get for the total intensity if the image field was classical. We can now allow this image

field to evolve under a Hamiltonian dynamics with Ornstein-Uhlenbeck noise according to Schrodinger's equation and by measuring the transition probabilities to different states at different times, attempt to estimate the parameter vector θ of the initial image. The noisy dynamics can be expressed in the form

$$id|\psi_I(t)\rangle/dt = (H_0 + w(t)V)|\psi_I(t)\rangle, t \geq 0,$$

$$dw(t) = -\gamma w(t)dt + \sigma dB(t)$$

with the initial condition

$$|\psi_I(0)\rangle = \sum_{k,m} c(k, m|\theta)|k, m\rangle$$

Upto quadratic orders in the noise w the solution is

$$\begin{aligned} |\psi_I(t)\rangle &= U_0(t)|\psi_I(0)\rangle - i \int_0^t U_0(t-s)VU_0(s)|\psi_I(0)\rangle w(s)ds \\ &\quad - \int_{0 < s_2 < s_1 < t} U_0(t-s_1)VU_0(s_1-s_2)VU_0(s_2)|\psi_I(0)\rangle w(s_2)w(s_1)ds_2 ds_1 \end{aligned}$$

The transition probability at time t to a state $|\phi\rangle$ is then given upto quadratic orders in the noise moments (the noise is assumed to have zero mean)

$$\begin{aligned} \mathbb{E}[|\langle \phi | \psi_I(t) \rangle|^2] &= \\ |\langle \phi | U_0(t) | \psi_I(0) \rangle|^2 &+ \int_{0 < s_1, s_2 < t} \langle \phi | U_0(t-s_1) VU_0(s) | \psi_I(0) \rangle \langle \phi | U_0(t-s_2) VU_0(s_2) | \psi_I(0) \rangle^* R_w(s_1, s_2) ds_1 ds_2 \\ - 2\text{Re}[\langle \phi | U_0(t) | \psi_I(0) \rangle \int_{0 < s_2 < s_1 < t} &\langle \phi | U_0(t-s_1) VU_0(s_1-s_2) VU_0(s_2) | \psi_I(0) \rangle^* R_w(s_1, s_2) ds_1 ds_2] \end{aligned}$$

Substituting in this expression $|\psi_I(0)\rangle = \sum_{k,m} c(k, m|\theta)|k, m\rangle$, we get the transition probability to the state $|\phi\rangle$ as a function of θ and hence by measuring the transition probability to different states at different times, we can estimate the parameter θ .

Now, given the original image $|\psi_I\rangle = |\psi_I(0)\rangle$, suppose we superpose to it a water-mark image $|\psi_W\rangle = \sum_{k,m} d(k, m)|k, m\rangle$. The resulting image will be

$$|\psi\rangle = K(|\psi_I\rangle + |\psi_W\rangle) = K(\sum_{k,m} (c(k, m) + d(k, m))|k, m\rangle)$$

We wish to extract out the original image from this water-marked image or equivalently, the water-marked image from this image. For that we must assume some sort of selectivity, ie, for certain values of k, m $c(k, m)$ dominates over $d(k, m)$ and for the other values of k, m , $c(k, m)$ is small. Then we can use selective filters, ie linear transformations on the underlying Hilbert space to extract out the original image. Other methods, involve, first applying a linear transformation T like a quantum cosine transform [1] to the original image,

then superposing it with the water-mark image and then applying the inverse of T to the resulting watermarked image. The resulting image will be

$$T^{-1}(|T|\psi_I\rangle + |\psi_w\rangle) = |\psi_I\rangle + T^{-1}|\psi_w\rangle$$

In this output image, it may happen that the indices k, m where ψ_I has larger coefficients heavily dominate the corresponding coefficients of $T^{-1}|\psi_w\rangle$ and this may enable us to separate out the two images. Let $S = T^{-1}$. Then, to extract out the original image as accurately as possible, we should minimize $\|S|\psi_w\rangle\|$ subject to the constraint that $\||\psi_I\rangle + S|\psi_w\rangle\| = 1$. Equivalently, we minimize

$$\langle \psi_w | S^* S | \psi_w \rangle - \lambda (\| \psi_I \|^2 + \langle \psi_w | S^* S | \psi_w \rangle + \langle \psi_I | S \psi_w \rangle + \langle \psi_w | S^* | \psi_I \rangle - 1)$$

w.r.t S . More generally, we may design the filter S for a whole ensemble of watermark images by minimizing

$$E(S, S^*, \lambda_n, n = 1, 2, \dots) =$$

$$\sum_n \langle \psi_{wn} | S^* S | \psi_{wn} \rangle - \sum_n \lambda_n (\| \psi_I \|^2 + \langle \psi_{wn} | S^* S | \psi_{wn} \rangle + \langle \psi_I | S \psi_{wn} \rangle + \langle \psi_{wn} | S^* | \psi_I \rangle - 1)$$

Setting the variational derivative of E w.r.t S^* to zero gives us a set of linear equations for S :

$$S \sum_n (1 - \lambda_n) |\psi_{wn}\rangle \langle \psi_{wn}| = \sum_n \lambda_n |\psi_I\rangle \langle \psi_{wn}|$$

The Lagrange multipliers are then determined by substituting

$$S = [\sum_n (1 - \lambda_n) |\psi_{wn}\rangle \langle \psi_{wn}|]^{-1} [\sum_n \lambda_n |\psi_I\rangle \langle \psi_{wn}|]$$

into the constraint equations

$$\| \psi_I \|^2 + \langle \psi_{wn} | S^* S | \psi_{wn} \rangle + \langle \psi_I | S \psi_{wn} \rangle + \langle \psi_{wn} | S^* | \psi_I \rangle = 1$$

Filters of this kind can be realized using non-random and random unitary evolutions as well as by introducing quantum noise into the system and then forming partial traces to obtain Lindblad type evolutions. However in this latter formalism, the quantum image must not be expressed as a pure state, rather as a mixed state.

A more accurate method for describing this superposition process is to put the restriction that all the pure states at each stage of processing have unit norm. The initial image state is

$$|\psi_I\rangle = \sum_{k,m} c(k, m) |k, m\rangle$$

and the assumption that $\| |\psi_I\rangle \| = 1$ means that

$$\sum_{k,m} |c(k, m)|^2 = 1$$

The watermark image is

$$|\psi_w\rangle = \sum_{k,m} d(k, m) |k, m\rangle$$

and for this to have unit norm, we must have

$$\sum_{k,m} |d(k, m)|^2 = 1$$

Note that

$$\langle k, m | k', m' \rangle = \delta(k, k') \delta(m, m')$$

After applying the processing operator T (a linear operator) to the original image state and performing the normalization, the state of the resulting image becomes

$$|\psi_1\rangle = T|\psi_I\rangle / \| T|\psi_I\rangle \|$$

Then we superpose this state with the water-mark state and normalize the resulting state to get the state

$$|\psi_2\rangle = (|\psi_1\rangle + |\psi_w\rangle) / \| |\psi_1\rangle + |\psi_w\rangle \|$$

The inverse of the operator T is then applied to this state and then normalized to get

$$|\psi_3\rangle = (T^{-1}|\psi_2\rangle) / \| T^{-1}|\psi_2\rangle \|$$

The problem is then to determine T so that

$$\| |\psi_3\rangle - |\psi_I\rangle \|$$

is a minimum. More generally, we can consider processing using a single linear operator T of a sequence of pairs of original and watermark states $|\psi_{I,k}\rangle, |\psi_{w,k}\rangle, k = 1, 2, \dots, N$. Then, T would have to be designed to minimize

$$f(T) = \sum_{k=1}^N \| |\psi_{3,k}\rangle - |\psi_{I,k}\rangle \| ^2$$

where

$$|\psi_{3,k}\rangle = (T^{-1}|\psi_{2,k}\rangle) / \| T^{-1}|\psi_{2,k}\rangle \|,$$

$$|\psi_{2,k}\rangle = (|\psi_{1,k}\rangle + |\psi_{w,k}\rangle) / \| |\psi_{1,k}\rangle + |\psi_{w,k}\rangle \|,$$

$$|\psi_{1,k}\rangle = T|\psi_{I,k}\rangle / \| T|\psi_{I,k}\rangle \|$$

This is a complex non-linear optimization problem and can be looked upon as training the operator T using sample quantum images and sample water-mark

images so that after superposing each pair and applying the processing operator T , we get the best possible reconstruction. The algorithm based on gradient search for minimizing $f(T)$ is to choose a small number ϵ and an adaptation constant μ and assuming that the state vectors are of size $p \times 1$ so that the operator T will be of size $p \times p$, and then let E_{nm} denote the $p \times p$ size matrix with a one at the $(n, m)^{th}$ entry and a zero at the other entries, and then implement the recursion

$$T_{nm}[k+1] = T_{nm}[k] - (\mu/\epsilon)(f(T + \epsilon E_{nm}) - f(T))$$

Note that in the limit $\epsilon \rightarrow 0$, $(f(T + \epsilon E_{nm}) - f(T))/\epsilon$ converges to $\frac{\partial f(T)}{\partial T_{nm}}$ so that in this limit the above recursion becomes the gradient search algorithm. The MATLAB programme for this is as follows:

```

for r = 1 : L
    sum = 0;
    for k = 1 : N
        psi1[:, k] = T[:, (r - 1) * L + 1 : r * L] * psiI[:, k] / norm(T[:, (r - 1) * L + 1 :
r * L] * psiI[:, k]);
        psi2[:, k] = (psi1[:, k] + psiw[:, k]) / norm(psi1[:, k] + psiw[:, k]);
        psi3[:, k] = (T[:, (r - 1) * L + 1 : r * L]-1 * psi2[:, k]) / norm(T[:, (r - 1) * L + 1 :
r * L]-1 * psi2[:, k]);
        sum = sum + norm(psi3[:, k] - psiI[:, k])2;
    end;

```

Now repeat the same steps with $T[:, (r - 1) * L + 1 : r * L]$ replaced by $T[:, (r - 1) * L + 1 : r * L] + \epsilon E_{nm}$ for each $n, m = 1, 2, \dots, p$ and implement the above stated recursion.

Noise analysis (approximate) ψ_N is a noise vector. The noise corrupted quantum background image is $\psi_1 = (\psi_I + \psi_N) / \| \psi_I + \psi_N \|$. This is processed by the operator T and normalized to obtain

$$\psi_2 = T\psi_1 / \| T\psi_1 \|$$

After superposing this with the watermark image ψ_w , the image state is

$$\psi_3 = (\psi_2 + \psi_w) / \| \psi_2 + \psi_w \|$$

and after applying the inverse operation, followed by state normalization, the resulting image is

$$\psi_4 = T^{-1}\psi_3 / \| T^{-1}\psi_3 \|$$

The estimation error energy is then

$$\mathbb{E}(\| \psi_4 - \psi_I \|^2)$$

and this must be calculated approximately by expanding the state estimation error only upto linear orders in ψ_N .

5.37 The EKF for arbitrary Markov processes with Gaussian measurement noise

$X(t)$ is a Markov process having generator K_t with kernel $K_t(x, y)$. The measurement process is

$$dz(t) = h(X(t))dt + \sigma_v dV(t)$$

where $V(\cdot)$ is standard vector valued Brownian motion. The Kushner-Kallianpur filter for the observable $\phi(X(t))$ is

$$d\pi_t(\phi) = \pi_t(K_t\phi))dt + \sigma_v^{-2}(\pi_t(\phi.h) - \pi_t(\phi)\pi_t(h))^T(dz(t) - \pi_t(h)dt)$$

where

$$\pi_t(\phi) = \mathbb{E}(\phi(X(t))|\eta_t), \eta_t = \{z(s) : s \leq t\}$$

We write

$$\hat{X}(t) = \pi_t(X), P(t) = cov(\delta X(t)|\eta_t), \delta X(t) = X(t) - \hat{X}(t)$$

We have approximately, (with $\delta X = X - \hat{X}(t)$),

$$\begin{aligned} \pi_t(X.h^T) &= \pi_t(Xh(X)) = \pi_t((\hat{X}(t) + \delta X).(h(\hat{X}(t)) + h'(\hat{X}(t))\delta X)^T) \\ &= \hat{X}(t).h(\hat{X}(t)) + \pi_t(\delta X.\delta X^T)h'(\hat{X}(t))^T \\ &= \hat{X}(t).h(\hat{X}(t)) + P(t)h'(\hat{X}(t))^T \end{aligned}$$

and hence,

$$\pi_t(X.h^T) - \pi_t(X)\pi_t(h^T) \approx P(t)H_t^T$$

where

$$H_t = h'(\hat{X}(t))$$

Also,

$$\pi_t(K_t X) = \int \pi_t(K_t(x, y))ydy \approx \int K_t(\hat{X}(t), y)ydy$$

We write

$$F_t(X) = \int K_t(X, y)ydy$$

and hence we have approximately,

$$d\hat{X}(t) = F_t(\hat{X}(t))dt + \sigma_v^{-2}P(t)H_t^T(dz(t) - h(\hat{X}(t))dt)$$

where

$$H_t = h'(\hat{X}(t))$$

We define

$$b_t(x) = \int K_t(x, y)(y - x)dy$$

and

$$a_t(x) = \int K_t(x, y)(y - x)(y - x)^Tdy$$

We note that

$$b_t(x)dt = \mathbb{E}(dX(t)|X(t) = x)$$

and

$$a_t(x)dt = \mathbb{E}(dX(t).dX(t)^T|X(t) = x)$$

since $dX(t) = X(t + dt) - X(t)$ and

$$\mathbb{E}(X(t + dt)|X(t) = x) = x + dtK_t x = x + \int K_t(x, y)ydy$$

Note that

$$\int K_t(x, y)dy = 0$$

and hence, we may equivalently write

$$b_t(x) = \int K_t(x, y)ydy = F_t(x)$$

Now,

$$P(t) = \mathbb{E}(\delta X(t).\delta X(t)^T|\eta_t)$$

$$d(\delta X(t).\delta X(t)^T) = d\delta X(t).\delta X(t)^T + \delta X(t).d\delta X(t)^T + d\delta X(t).d\delta X(t)^T$$

and

$$d\delta X(t).\delta X(t)^T = dX(t).\delta X(t)^T - d\hat{X}(t).\delta X(t)^T$$

Now,

$$\mathbb{E}(dX(t).\delta X(t)^T|\eta_t, X(t)) = dt.F_t(X(t)).\delta X(t)^T$$

and hence

$$\begin{aligned} \mathbb{E}(dX(t).\delta X(t)^T|\eta_t) &= dt.\mathbb{E}(F_t(X(t)).\delta X(t)^T|\eta_t) \\ &\approx dt.F'_t(\hat{X}(t))P(t) \end{aligned}$$

since

$$\mathbb{E}(\delta \hat{X}(t)|\eta_t) = 0$$

and $\hat{X}(t)$ is η_t -measurable. Also, we have approximately,

$$\begin{aligned} \mathbb{E}(d\hat{X}(t).\delta X(t)^T|\eta_t) &= \\ E[(F_t(\hat{X}(t))dt + \sigma_v^{-2}P(t)H_t^T(dz(t) - h(\hat{X}(t))dt))\delta X(t)^T|\eta_t] \\ &= \mathbb{E}[(F_t(\hat{X}(t))dt + \sigma_v^{-2}P(t)H_t^T(h(X(t))dt + \sigma_v dV(t) - h(\hat{X}(t))dt))\delta X(t)^T|\eta_t] \\ &= dt\sigma_v^{-2}P(t)H_t^T\mathbb{E}[h'(\hat{X}(t))\delta X(t)\delta X(t)^T|\eta_t] \\ &= dt.\sigma_v^{-2}P(t)H_t^TH_tP(t) \end{aligned}$$

Taking the transpose of this equation gives the same result. Finally,

$$\begin{aligned} d\delta X(t).d\delta X(t)^T &= (dX(t) - d\hat{X}(t)).(dX(t) - d\hat{X}(t))^T \\ &= dX(t).dX(t)^T + d\hat{X}(t).d\hat{X}(t)^T = \end{aligned}$$

$$dX(t).dX(t)^T + \sigma_v^{-4} P(t) H_t^T (\sigma_v^2 dt) H_t P(t) = dX(t).dX(t)^T + \sigma_v^{-2} P(t) H_t^T H_t P(t)$$

and

$$\mathbb{E}(dX(t).dX(t)^T | \eta_t) \approx a(\hat{X}(t))dt$$

Combining all this, we get the estimation error covariance evolution

$$P'(t) = F'_t(\hat{X}(t))P(t) + P(t)F'_t(\hat{X}(t))^T + a(\hat{X}(t)) - \sigma_v^{-2} P(t) H_t^T H_t P(t)$$

These equations form the EKF for the general case when the state process is an arbitrary Markov process with known infinitesimal generator.

5.38 Quantum scattering theory applied to quantum gate design

This theory illustrates the most important feature of quantum mechanics in infinite dimensional Hilbert spaces that are not present in the finite dimensional setting.

Mourre's conjugate operator method: Let T be a bounded Hermitian operator in a Hilbert space with spectral measure $E(.)$ so we can write

$$T = \int_{-M}^M x dE(x)$$

for some $0 < M \infty$. Let A be another Hermitian operator in the same Hilbert space so that $B = [T, iA]$ is bounded and there exists an $a > 0$ such that for some interval J in \mathbb{R} , we have

$$E(J)BE(J) \geq a.E(J)$$

Suppose $\delta > 0$ and $J = [b, d]$. Define $F = E[b, d] = E(J)$. Then, for all $\lambda \in J = [b + \delta, d - \delta]$, we have by the spectral theorem,

$$\| (T - \lambda - i\mu)^{-1} F^\perp \| \leq (\delta^2 + \mu^2)^{-1/2} \leq 1/\delta$$

where

$$F^\perp = I - F = E((-\infty, b - \delta)) + E((d + \delta, \infty))$$

It follows that for any vector g in the Hilbert space, and for all $\lambda \in J$,

$$\| F^\perp g \| \leq \| T - \lambda - i\mu \| / \delta \leq K \| g \| / \delta$$

where

$$K^2 = \max((M - b)^2 + \mu^2, (M + b)^2 + \mu^2, (M - d)^2 + \mu^2, (M + d)^2 + \mu^2)$$

Note that

$$\| (T - \lambda - i\mu)g \| = \| (T - \lambda - i\mu + i\epsilon B - i\epsilon B)g \|$$

$$\leq \|N_\epsilon g\| + \epsilon \|Bg\|$$

Thus,

$$\begin{aligned}\|F^\perp g\| &= \| (T - \lambda - i\mu)^{-1} F^\perp (N_\epsilon - i\epsilon B) g \| \\ &\leq \delta^{-1} [\|N_\epsilon g\| + \epsilon \|Bg\|]\end{aligned}$$

Note that we have used the identity

$$T - \lambda - i\mu = N_\epsilon - i\epsilon B$$

Here, we have defined

$$N_\epsilon = T - \lambda - i\mu + i\epsilon B, G_\epsilon = N_\epsilon^{-1}$$

We have further from the assumed Mourre inequality,

$$\|g\|^2 = \|Fg\|^2 + \|F^\perp g\|^2,$$

$$a < Fg, Fg > = a < g, Fg > \leq < g, FBg >$$

$$\begin{aligned}&= < g, Bg > - < g, BF^\perp g > - < g, F^\perp Bg > + < g, F^\perp BF^\perp g > \\ &\leq < g, Bg > + 2 \|Bg\| \cdot \|F^\perp g\| + \|B\| \|F^\perp g\|^2\end{aligned}$$

Then,

$$i\epsilon < g, Bg > = < g, N_\epsilon g > - < g, (T - \lambda - i\mu)g >$$

Also,

$$\begin{aligned}2i \operatorname{Im} < g, N_\epsilon g > &= < g, (N_\epsilon - N_\epsilon^*)g > \\ &= 2i < g, (\mu + \epsilon B)g >\end{aligned}$$

so that

$$< g, Bg > = (\operatorname{Im}(< g, N_\epsilon g >) - \mu < g, g >)/\epsilon$$

We can using the earlier inequality conclude that for a fixed, δ and for all sufficiently small ϵ and for sufficiently small μ dependent upon ϵ , we have

$$< g, g > \leq (K_1 + K_2/\epsilon) \|g\| \cdot \|N_\epsilon g\|$$

where K_1, K_2 depend on δ . From this, we obtain the fundamental inequality

$$\|G_\epsilon g\| \leq (K_1 + K_2/\epsilon) \|g\|$$

Exercise:

5.39 Superconductivity for two species and interpretation of the gap function

$$\begin{aligned}
 H = & (-1/2m) \int [\psi_1(x)^*((\nabla + ieA(x)/m)^2 + \mu(x))\psi_1(x) + \psi_2(x)^*(\nabla + ieA(x)/m)^2\psi_2(x)]d^3x \\
 & \int [V_1(x) < \psi_1(x)^*\psi_1(x) > \psi_2(x)^*\psi_2(x) + V_2(x) < \psi_2(x)^*\psi_2(x) > \psi_1(x)^*\psi_1(x)]d^3x \\
 & + \int [V_3(x) < \psi_1(x)\psi_2(x) > \psi_1(x)^*\psi_2(x)^* + \bar{V}_3(x) < \psi_2(x)^*\psi_1(x)^* > \psi_2(x)\psi_1(x)]d^3x \\
 & = H_1 + H_2 + H_3
 \end{aligned}$$

The terms involving V_1, V_2 are the Hartree-Fock terms while the terms involving $V_3(x)$ and $\bar{V}_3(x)$ are the Cooper pair terms. The field operator anticommutation rules are

$$\{\psi_a(x), \psi_b(x')^*\} = \delta_{ab}\delta^3(x - x'), a, b = 1, 2$$

assuming $t = t'$. All the other anticommutators are zero. Define the Green's functions

$$G_{ab}(x|x') = < T\{\psi_a(x)\psi_b(x')^*\} >, a, b = 1, 2,$$

$$F_{ab}(x|x') = < T\{\psi_a(x)\psi_b(x')\} >, a, b = 1, 2$$

Obtain the differential equations for these Green's functions using the self consistency condition

$$F_{12}(x|x) = \Delta(x) = < \psi_1(x)\psi_2(x) >$$

and

$$G_{11}(x|x) = - < \psi_1(x)^*\psi_1(x) >$$

$$G_{22}(x|x) = - < \psi_2(x)^*\psi_2(x) >$$

We observe that using the anticommutation rules,

$$G_{ab}(x|x')^* = -G_{ba}(x'|x),$$

$$F_{ab}(x|x')^* = - < T(\psi_b(x')^*\psi_a(x)) >$$

Note also that

$$G_{ab}(x|x') = - < T(\psi_b(x')^*\psi_a(x)) >,$$

$$F_{ab}(x|x') = -F_{ba}(x'|x)$$

Make use of the field operator anticommutation relations and the Heisenberg equations of motion

$$i\psi_{a,t}(x) = [H, \psi_a(x)]$$

Note that this equation also implies

$$i\psi_{a,t}^*(x) = [H, \psi_a^*(x)]$$

These Heisenberg equations yield the Schrodinger equations for the wave operators. Solve using perturbation theory for the Green's function by treating the magnetic vector potential terms as a perturbation. Note that

$$[H_1, \psi_1(x)] = (1/2m)((\nabla + ieA)^2 + \mu(x)\psi_1(x),$$

$$[H_1, \psi_2(x)] = (1/2m)(\nabla + ieA)^2\psi_2(x),$$

$$[H_2, \psi_1(x)] = -V_2(x) < \psi_2(x)^*\psi_2(x) > \psi_1(x)$$

$$[H_2, \psi_2(x)] = -V_1(x) < \psi_1(x)^*\psi_1(x) > \psi_2(x)$$

$$[H_3, \psi_1(x)] = -V_3(x) < \psi_1(x)\psi_2(x) > \psi_2(x)^*$$

$$[H_3, \psi_2(x)] = V_3(x) < \psi_1(x)\psi_2(x) > \psi_1(x)^*$$

The Heisenberg equations of motion now become

$$\psi_{1,t}(x) = (i/2m)((\nabla + ieA)^2 + \mu(x))\psi_1(x)$$

$$-iV_2(x)N_2(x)\psi_1(x) - iV_3(x)\Delta(x)\psi_2(x)^*$$

$$\psi_{2,t}(x) = (i/2m)((\nabla + ieA)^2 + \mu(x))\psi_2(x) - iV_1(x)N_1(x)\psi_2(x) + iV_3(x)\Delta(x)\psi_1(x)^*$$

For a superconductor, the Cooper-pair gap function contribution $\Delta(x)$ dominates as compared to the Hartree-Fock contribution $N_1(x)$, $N_2(x)$ and so these equations approximate to

$$\psi_{1,t}(x) = (i/2m)((\nabla + ieA)^2 + \mu(x))\psi_1(x) - iV_3(x)\Delta(x)\psi_2(x)^*$$

$$\psi_{2,t}(x) = (i/2m)((\nabla + ieA)^2 + \mu(x))\psi_2(x) + iV_3(x)\Delta(x)\psi_1(x)^*$$

These constitute the fundamental field operator equations for a superconductor. From these equations, the Green's functions are easily seen to satisfy

$$iG_{11,t}(x|x') = i\delta^4(x - x') - (1/2m)((\nabla + ieA(x))^2 + \mu(x))G_{11}(x|x')$$

$$+V_3(x)\Delta(x) < T(\psi_2(x)^*\psi_1(x')^*) >$$

$$iF_{12,t}(x|x') = (i/2m)((\nabla + ieA)^2 + \mu)F_{12}(x|x') + V_3\Delta < T(\psi_2(x)^*\psi_2(x')) >$$

$$iG_{22,t}(x|x') = i\delta^4(x - x') - (1/2m)((\nabla + ieA)^2 + \mu)G_{22}(x|x')$$

$$-V_3\Delta < T(\psi_1(x)^*\psi_2(x')^*) >$$

$$iF_{21,t}(x|x') = -(1/2m)((\nabla + ieA)^2 + \mu)F_{21}(x|x') - V_3(x)\Delta(x) < T(\psi_1(x)^*\psi_1(x')) >$$

Now,

$$< T(\psi_2(x)^*\psi_1(x')^*) > = - < T(\psi_1(x')\psi_2(x)) >^* = -F_{12}(x'|x)^* = F_{21}(x|x')^*$$

$$< T(\psi_2(x)^*\psi_2(x')) > = - < T(\psi_2(x')\psi_2(x)^*) > = -G_{22}(x'|x)$$

$$< T(\psi_1(x)^*\psi_2(x')^*) > = - < T(\psi_2(x')\psi_1(x)) >^* = -F_{21}(x'|x)^* = F_{12}(x|x')^*$$

$$< T(\psi_1(x)^*\psi_1(x')) > = - < T(\psi_1(x')\psi_1(x)^*) > = -G_{11}(x'|x)$$

So we finally get

$$\begin{aligned} iG_{11,t}(x|x') &= i\delta^4(x - x') - (1/2m)((\nabla + ieA(x))^2 + \mu(x))G_{11}(x|x') \\ &\quad + V_3(x)\Delta(x)F_{21}(x|x')^* \end{aligned}$$

$$\begin{aligned} iG_{22,t}(x|x') &= i\delta^4(x - x') - (1/2m)((\nabla + ieA(x))^2 + \mu(x))G_{22}(x|x') + V_3(x)\Delta(x)F_{21}(x'|x)^* \\ iF_{21,t}(x|x') &= -(1/2m)((\nabla + ieA(x))^2 + \mu(x))F_{21}(x|x') + V_3(x)\Delta(x)G_{11}(x'|x) \end{aligned}$$

Now consider the special case when there is no magnetic field and $\mu(x), V_3(x)$ are constants μ, V_3 . Then we may assume that the gap function $\Delta(x)$ is also a constant Δ and we shall also be assuming that G_{11}, G_{22}, F_{21} depend on only $x - x'$. We shall therefore write $G_{11}(x - x'), G_{22}(x - x'), F_{21}(x - x')$ in place of $G_{11}(x|x'), G_{22}(x|cx'), F_{21}(x|x')$ respectively. Taking the Four dimensional Fourier transform of the above equations then gives us

$$\begin{aligned} (\omega - k^2/2m + \mu)G_{11}(\omega, k) - V_3\Delta F_{21}(-\omega, -k)^* &= i, \\ (\omega - k^2/2m + \mu)G_{22}(\omega, k) - V_3\Delta F_{21}(\omega, k)^* &= i, \\ (\omega - k^2/2m + \mu)F_{21}(\omega, k) - V_3\Delta G_{11}(-\omega, -k) &= 0 \end{aligned}$$

We can replace the third equation by

$$(-\omega + k^2/2m + \mu)F_{21}(-\omega, -k) - V_3\Delta G_{11}(\omega, k) = 0$$

We also recall that

$$G_{11}(x - x')^* = -G_{11}(x' - x)$$

and hence

$$G_{11}(\omega, k)^* = -G_{11}(\omega, k)$$

Thus, the third equation is the same as

$$(\omega - k^2/2m + \mu)F_{21}(-\omega, -k)^* + V_3\Delta G_{11}(\omega, k) = 0$$

5.40 Introductory quantum information theory

The following problems must be addressed in this course:

[1] Notion of a pure and mixed state and observables in finite dimensional Hilbert spaces.

[2] Notion of expectation of an observable in a state and the duality between the Schrodinger wave mechanics, Heisenberg's matrix mechanics and Dirac's interaction picture mechanics for quantum mechanics in finite dimensional Hilbert spaces.

[3] Lieb's concavity inequality and its application to proving concavity of the Von-Neumann entropy $S(\rho) = -Tr(\rho.log(\rho))$ of a state.

[4] The information theoretic distance between two quantum states ρ and σ :

$$D(\rho|\sigma) = \text{Tr}(\rho(\log(\rho) - \log(\sigma)))$$

[5] Proof of the joint convexity of the information distance.

[6] The notion of measurement in the quantum theory: Discussion of both PVM and POVM and the proof that any POVM can be extended to a PVM on a larger Hilbert space:

$$k_M(\rho) = \sum_k (\sqrt{M_k} \rho \sqrt{M_k}) = \text{Tr}_2(U(\rho \otimes |e_0\rangle\langle e_0|))U^*$$

Other versions are

$$\text{Tr}_2 \sum_k E_k (\rho \otimes \rho_0) E_k = \sum_k \sqrt{M_k} \cdot \rho \cdot \sqrt{M_k}$$

for an appropriate choice of ρ_0 depending on $\{M_k\}$ only.

$$\sqrt{M_k} \cdot \rho \cdot \sqrt{M_k} = \text{Tr}_2(E_k(\rho \otimes \rho_0)E_k)$$

[7] Measurement as a quantum operation, ie, as a noisy Schrodinger evolution.

[8] Proof of the decrease in the information between two states after a quantum operation as a consequence of its joint convexity:

$$D(k_M(\rho)|k_M(\sigma)) \leq D(\rho|\sigma)$$

This is a special case of a fundamental result on operator convex functions. If $f(X)$ is operator convex, and $Z_k, k = 1, 2, \dots, p$ are operators such that $\sum_k Z_k^* Z_k = I$, then for Hermitian matrices $X_k, k = 1, 2, \dots, p$ we have

$$f\left(\sum_k Z_k X_k Z_k^*\right) \leq \sum_k Z_k f(X_k) Z_k^*$$

[7] Notion of how a measurement on one state causes a change in the other state linked to the previous via a tensor product: Let the state be $\rho_A \otimes \rho_B$ defined in the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. Let $\{M_k^B\}$ be a POVM on system B . Then, after the measurement, the overall state collapses to

$$\sum_k \rho_A \otimes \sqrt{M_k^B} \rho_B \sqrt{M_k^B}$$

By partial tracing, it follows that after the measurement, the state of A remains unchanged while that of B collapses to

$$\sum_k \sqrt{M_k^B} \rho_B \cdot \sqrt{M_k^B}$$

More generally, if the joint state of A and B is ρ_{AB} with marginals ρ_A and ρ_B , then after the measurement, on B , the state of A collapses to

$$\sum_k Tr_B((I_A \otimes \sqrt{M_k^B})\rho_{AB}(I_A \otimes \sqrt{M_k^B}))$$

while that of B becomes

$$\sum_k \sqrt{M_k^B} \rho_B \sqrt{M_k^B}$$

[9] Notion of Renyi relative entropy, quantum relative entropy as a limiting case of Renyi entropy, applications of Renyi entropy to determining the asymptotic bound on the error probability (quantum large deviations) in the binary hypothesis testing problem, ie, quantum Stein's inequality. Convexity properties of Renyi relative entropy.

[10] Examples of noisy quantum channels, ie, TPCP maps on the state and their realization using noisy Schrodinger evolution. One important quantum channel is obtained as a transformation from states in $\mathcal{H}_A \otimes \mathcal{H}_B$ into states in \mathcal{H}_A by making measurements on B :

$$T(\rho) = \sum_k Tr_B(I_A \otimes \sqrt{M_k^B})\rho.(I_A \otimes \sqrt{M_k^B})$$

Another example is obtained by taking states ρ_1, \dots, ρ_K in \mathcal{H}_B , a state ρ_A in \mathcal{H}_A , a POVM $\{M_k^A\}$ in \mathcal{H}_A and defining a state in \mathcal{H}_B by

$$T(\rho_A) = \sum_{k=1}^K Tr(\rho_A M_k^A) \rho_k$$

[11a] Entropy typical projections and Bernoulli typical projections.

[11b] The Feinstein-Khintchine fundamental lemma in classical information theory and its quantum generalization as the greedy algorithm for constructing codewords and the corresponding detection operators.

[11c] Schumacher quantum noiseless compression as a generalization of Shannon's noiseless coding theorem.

[13] Mutual information for Cq channels as a special case of quantum relative entropy: Let $\bar{\rho} = \sum_x p(x)\rho(x)$

$$D\left(\sum_x p(x)|x\rangle\langle x| \otimes \rho(x) \middle| \sum_x p(x)|x\rangle\langle x| \otimes \bar{\rho}\right)$$

5.41 Quantum image processing

$|\psi_I(0) >$ is the given/background quantum image field and $|\psi_w(0) >$ is the quantum watermark image field. Both are of unit length. We wish to design a unitary operator U so that

$$|\psi_f > = U^{-1} \frac{(U|\psi_I(0) > + |\psi_w(0) >)}{\| U|\psi_I(0) > + |\psi_w(0) > \|}$$

is as close as possible to $|\psi_I(0) >$. Physically, this means that we first process the background image with the filter U , superpose the watermark image onto it and then after appropriate normalization of this superposed image, we apply the inverse filter U^{-1} to obtain a quantum image (normalized) that is as close as possible to the original quantum image $|\psi_I(0) >$ in the Euclidean norm. Further, we shall be interested in realizing the unitary operator U using a Hamiltonian system which follows the Schrodinger evolution, with the Hamiltonian dependent upon a parameter vector θ and shall also assume that noise is present during the Schrodinger evolution process. The aim shall then be to estimate the parameter θ by minimizing the mean square distance between $|\psi_f >$ and $|\psi_I(0) >$ w.r.t θ . We note that U becomes a random unitary operator once there is noise present in the Schrodinger evolution. U shall be obtained as $U(T)$, the Schrodinger evolution operator after T seconds. The unitary evolution equation assuming classical white Gaussian noise in the dynamics is described by an Ito stochastic differential equation with an Ito correction term to guarantee unitarity of the evolution. The dynamics is

$$dU(t) = [-(iH(\theta) + V^2(\theta)/2)dt - iV(\theta)dB(t)]U(t)$$

and its approximate solution upto second order in V is given by the truncated Dyson series:

$$U(t) = U_0(t)W(t), U_0(t) = \exp(-itH_0),$$

$$dW(t) = [-\tilde{V}(t)^2 dt/2 - i\tilde{V}(t)dB(t)]W(t), \tilde{V}(t) = U_0(t)^* V U_0(t)$$

Thus upto quadratic orders in V , we have

$$W(T) = I - \int_0^T \tilde{V}(t)^2 dt/2 - i \int_0^T \tilde{V}(t)dB(t) - \int_{0 < t_2 < t_1 < T} \tilde{V}(t_1)\tilde{V}(t_2)dB(t_2)dB(t_1)$$

We have

$$\begin{aligned} \mathcal{F}(\theta) &= \| |\psi_f > - |\psi_I(0) > \| ^2 = \\ &\| U^{-1} \frac{(U|\psi_I(0) > + |\psi_w(0) >)}{\| U|\psi_I(0) > + |\psi_w(0) > \|} - |\psi_I(0) > \| ^2 \\ &= \| \frac{|\psi_I(0) > + U^* |\psi_w(0) >}{(2(1 + Re(< \psi_w(0) | U^* |\psi_I(0) >))^{1/2}} - |\psi_I(0) > \| ^2 \\ &= \| (1 - a)|\psi_I(0) > + U^* |\psi_w(0) > \| ^2 / a^2 \end{aligned}$$

where

$$a^2 = 2(1 + \operatorname{Re}(\langle \psi_w(0) | U | \psi_I(0) \rangle))$$

Thus,

$$\begin{aligned}\mathcal{F}(\theta) &= [(1-a)^2 + 1 + 2(1-a)\operatorname{Re}(\langle \psi_w(0) | U | \psi_I(0) \rangle)]/a^2 \\ &= (2+a^2-2a+(1-a)(a^2-2))/a^2 = (2+a^2-2a+a^2-2-a^3+2a)/a^2 \\ &= (2a^2-a^3)/a^2 = 2-a\end{aligned}$$

So minimizing $\mathbb{E}[\mathcal{F}(\theta)]$ is equivalent to maximizing

$$\begin{aligned}\mathbb{E}(a) &= \mathbb{E}(1 + \operatorname{Re}(\langle \psi_w(0) | U | \psi_I(0) \rangle))^{1/2} \\ &\approx 1 + (1/2)\mathbb{E}[\operatorname{Re}(\langle \psi_w(0) | U | \psi_I(0) \rangle)] - (1/8)\mathbb{E}[(\operatorname{Re}(\langle \psi_w(0) | U | \psi_I(0) \rangle))^2]\end{aligned}$$

We have

$$U = U(T) \approx U_0(T) + \sum_k \theta_k X_k + \sum_{k,m} \theta_k \theta_m Y_{km}$$

where we assume that H_0 is independent of θ and

$$V = V(\theta) = \sum_k \theta_k V_k,$$

so that

$$\tilde{V}(t) = \sum_k \theta_k \tilde{V}_k(t)$$

Then,

$$W(T) = I - \int_0^T \tilde{V}(t)^2 dt / 2 - i \int_0^T \tilde{V}(t) dB(t) - \int_{0 < t_2 < t_1 < T} \tilde{V}(t_1) \tilde{V}(t_2) dB(t_2) dB(t_1)$$

and so

$$X_k = -i \int_0^T \tilde{V}_k(t) dB(t),$$

$$Y_{km} = -(1/2) \int_0^T \tilde{V}_k(t) \tilde{V}_m(t) dt - \int_{0 < t_2 < t_1 < T} \tilde{V}_k(t_1) \tilde{V}_m(t_2) dB(t_2) dB(t_1)$$

$$\operatorname{Re}(\langle \psi_w(0) | U | \psi_I(0) \rangle) =$$

$$\begin{aligned}&\operatorname{Re}(\langle \psi_w(0) | U_0(T) | \psi_I(0) \rangle) + \sum_k \theta_k \operatorname{Re}(\langle \psi_w(0) | X_k | \psi_I(0) \rangle) \\ &+ \sum_{k,m} \theta_k \theta_m \operatorname{Re}(\langle \psi_w(0) | Y_{km} | \psi_I(0) \rangle)\end{aligned}$$

Also

$$\mathbb{E}(X_k) = 0, \mathbb{E}(Y_{km}) = (-1/2) \int_0^T \tilde{V}_k(t) \tilde{V}_m(t) dt$$

$$\begin{aligned}
& \mathbb{E}[Re(<\psi_w(0)|U|\psi_I(0)>)] = \\
& Re(<\psi_w(0)|U_0(T)|\psi_I(0)>) + \sum_{k,m} \theta_k \theta_m Re(<\psi_w(0)|\mathbb{E}(Y_{km})|\psi_I(0)>) \\
& \mathbb{E}[(Re(<\psi_w(0)|U|\psi_I(0)>))^2] = \\
& (Re(<\psi_w(0)|U_0(T)|\psi_I(0)>)^2 + 2Re(<\psi_w(0)|U_0(T)|\psi_I(0)>).\sum_{k,m} \theta_k \theta_m Re(<\psi_w(0)|\mathbb{E}Y_{km}|\psi_I(0)>)) \\
& + \sum_{k,m} \theta_k \theta_m \mathbb{E}[Re(<\psi_w(0)|X_k|\psi_I(0)>).Re(<\psi_w(0)|X_m|\psi_I(0)>)]
\end{aligned}$$

We note that

$$\begin{aligned}
& \mathbb{E}[Re(<\psi_w(0)|X_k|\psi_I(0)>).Re(<\psi_w(0)|X_m|\psi_I(0)>)] \\
& = \int_0^T Im(<\psi_w(0)|\tilde{V}_k(t)|\psi_I(0)>).Im(<\psi_w(0)|\tilde{V}_m(t)|\psi_I(0)>)dt
\end{aligned}$$

The objective function to be maximized upto second order terms in the parameters is thus

$$\theta^T B \theta = \sum_{k,m} b(k,m) \theta_k \theta_m$$

where

$$\begin{aligned}
b(k,m) &= (1/2)Re(<\psi_w(0)|\mathbb{E}(Y_{km})|\psi_I(0)>) - (1/8)\mathbb{E}[Re(<\psi_w(0)|X_k|\psi_I(0)>).Re(<\psi_w(0)|X_m|\psi_I(0)>)] \\
&= (-1/4) \int_0^T <\psi_w(0)|\tilde{V}_k(t)\tilde{V}_m(t)|\psi_I(0)> dt \\
&- (1/8) \int_0^T Im(<\psi_w(0)|\tilde{V}_k(t)|\psi_I(0)>).Im(<\psi_w(0)|\tilde{V}_m(t)|\psi_I(0)>)dt
\end{aligned}$$

This maximization must be carried out subject to an energy constraint on the noise. In other words, we must distribute the parameters θ amongst the noise sources $V_k dB(t)$ in such a way that the total noise energy is fixed at E_N , ie,

$$\int_0^T \mathbb{E}Tr(Z(t)(\sum_k \theta_k \tilde{V}_k(t))^2) dt = E_N$$

or equivalently,

$$\sum_{k,m} a(k,m) \theta_k \theta_m = E_N$$

where

$$a(k,m) = \int_0^T Tr(Z(t)\tilde{V}_k(t)\tilde{V}_m(t))dt$$

This constrained maximization problem is therfore equivalent to maxmizing

$$Q(\theta, \lambda) = \theta^T B \theta - \lambda(\theta^T A \theta - E_N)$$

and the solution is that we choose that generalized eigenvector θ_0 of the generalized eigenvalue problem

$$(B - \lambda A)\theta = 0$$

corresponding to the maximum generalized eigenvalue λ_0 , ie, the maximum root of the equation

$$\det(B - \lambda A) = 0$$

and with the normalization condition

$$\theta_0^T A \theta_0 = E_N$$

The case of quantum noise. Designing optimal processors for quantum watermarking in the presence of quantum noise described using the HP quantum stochastic calculus. The correct model for noise corrupting a quantum system evolving according to a Hamiltonian is described by the Hudson-Parthasarathy quantum stochastic calculus. This formalism provides an even more general framework for modeling noise in quantum systems than is provided by the above classical Brownian motion and Ito calculus formalism. The reason behind this is that the master equation for the mixed system state when it is coupled to a bath known as the GKSL equation in full generality can be derived only from the Hudson-Parthasarathy noisy Schrodinger unitary evolution equation on system \otimes noise space by partial tracing over the bath space. From this generalized GKSL equation, we can derive by selecting appropriate Lindblad operators the Heisenberg dynamical equations for dissipative systems which is not possible using the classical Brownian motion model. For this reason, in the following section, we assume that the bath is in a coherent state, couple the system quantum background image state and quantum watermark state to the bath via the tensor product. We then assume that the Lindblad noise operators appearing in the Hudson-Parthasarathy noisy Schrodinger equation are linear combinations of some operators with the coefficients in this linear combination being unknown real parameters which are estimated by first approximately solving the Hudson-Parthasarathy equation upto second order terms in the parameters for the unitary evolution and then choosing these parameters such that after preprocessing the background image coupled to the bath, superposing with the watermark image coupled to the bath followed by normalization and then post-processing the resulting superposed bath coupled image state by the inverse of the unitary evolution, the resulting state has a minimum norm square error w.r.t. the original background image field coupled to the bath. For evaluating these norms, we make use of the way in which the creation and annihilation processes in the Hudson-Parthasarathy quantum stochastic calculus act on the coherent states of the bath. For full details, we refer the reader to [K.R.Parthasarathy, "An introduction to quantum stochastic calculus", Birkhauser, 1992].

The evolution equation is

$$dU(t) = (-(iH + P)dt + LdA(t) - L^*dA(t)^*)U(t)$$

where by quantum Ito's formula,

$$dA \cdot dA^* = dt$$

and to ensure unitarity of the evolution, we have by quantum Ito's formula,

$$P = LL^*/2$$

The solution using Dyson series upto quadratic orders in L, L^* is given by

$$U(t) = U_0(t)W(t)$$

where

$$U_0(t) = \exp(-itH_0)$$

and

$$W(t) = I + \int_0^T (\tilde{L}(t)dA(t) - \tilde{L}^*(t)dA(t)^*)$$

$$\begin{aligned} &+ \int_{0 < s < t < T} (\tilde{L}(t)dA(t) - \tilde{L}^*(t)dA(t)^*).(\tilde{L}(s)dA(s) - \tilde{L}^*(s)dA(s)^*) \\ &\quad - \int_0^T \tilde{P}(t)dt \end{aligned}$$

where

$$\begin{aligned} \tilde{L}(t) &= U_0(t)^*LU_0(t), \tilde{L}^*(t) = U_0(t)^*L^*U_0(t), \\ \tilde{P}(t) &= \tilde{L}(t)\tilde{L}(t)^*/2 \end{aligned}$$

and we can write

$$W = W(T) = I + W_1 + W_2$$

where W_1 consists of first order terms in the noise operators:

$$W_1 = \int_0^T (\tilde{L}(t)dA(t) - \tilde{L}(t)^*dA(t)^*)$$

and W_2 consists of quadratic terms in the noise operators:

$$W_2 =$$

$$+ \int_{0 < s < t < T} (\tilde{L}(t)dA(t) - \tilde{L}^*(t)dA(t)^*).(\tilde{L}(s)dA(s) - \tilde{L}^*(s)dA(s)^*) - \int_0^T \tilde{P}(t)dt$$

We write

$$L = \theta_1 L_1 + \theta_2 L_2, L^* = \theta_1 L_1^* + \theta_2 L_2^*$$

and design the real parameters $\theta_k, k = 1, 2$ so that the gate W in the interaction picture at time T is optimal ie, such that when we write

$$U = U(T) = U_0(T)W$$

then

$$\mathcal{E}(\theta_1, \theta_2) =$$

$$\left\| U^* \left(\frac{U|\psi_I \otimes \phi(u)\rangle + |\psi_w \otimes \phi(u)\rangle}{\|U|\psi_I \otimes \phi(u)\rangle + |\psi_w \otimes \phi(u)\rangle\|} \right) - |\psi_I \otimes \phi(u)\rangle \right\|^2$$

is a minimum. Equivalently, since $U^*U = I$, we have

$$\mathcal{E}(\theta_1, \theta_2) =$$

$$\left\| \left(\frac{|\psi_I \otimes \phi(u)\rangle + U^*|\psi_w \otimes \phi(u)\rangle}{\|U|\psi_I \otimes \phi(u)\rangle + |\psi_w \otimes \phi(u)\rangle\|} \right) - |\psi_I \otimes \phi(u)\rangle \right\|^2$$

We have

$$\begin{aligned} & \| |\psi_I \otimes \phi(u)\rangle + U^*|\psi_w \otimes \phi(u)\rangle \|^2 = \\ & 2 + 2\operatorname{Re}(\langle \psi_w \otimes \phi(u) | U |\psi_I \otimes \phi(u)\rangle) = a^2 \end{aligned}$$

say. We have approximately,

$$\begin{aligned} & \langle \psi_w \otimes \phi(u) | U |\psi_I \otimes \phi(u)\rangle = \\ & \langle \psi_w \otimes \phi(u) | (I + W_1 + W_2) |\psi_I \otimes \phi(u)\rangle \\ & = 1 + \langle \psi_w \otimes \phi(u) | W_1 |\psi_I \otimes \phi(u)\rangle + \langle \psi_w \otimes \phi(u) | W_2 |\psi_I \otimes \phi(u)\rangle \end{aligned}$$

and

$$\begin{aligned} & \langle \psi_w \otimes \phi(u) | W_1 |\psi_I \otimes \phi(u)\rangle \\ & = \langle \psi_w \otimes \phi(u) | \int_0^T \tilde{L}(t) dA(t) - \tilde{L}(t)^* dA(t)^* |\psi_I \otimes \phi(u)\rangle \\ & = \int_0^T [\langle \psi_w | \tilde{L}(t) |\psi_I \rangle u(t) - \langle \psi_w | \tilde{L}(t)^* |\psi_I \rangle \bar{u}(t)] dt \\ & = \theta_1 q_1 + \theta_2 q_2 \end{aligned}$$

where

$$q_k = \int_0^T [\langle \psi_w | \tilde{L}_k(t) |\psi_I \rangle - \langle \psi_w | \tilde{L}_k^*(t) |\psi_I \rangle] dt, k = 1, 2$$

Also,

$$\begin{aligned} & \langle \psi_w \otimes \phi(u) | W_2 |\psi_I \otimes \phi(u)\rangle = \\ & + \int_{0 < s < t < T} \langle \psi_w \otimes \phi(u) | (\tilde{L}(t) dA(t) - \tilde{L}^*(t) dA(t)^*) \cdot (\tilde{L}(s) dA(s) - \\ & \quad \tilde{L}^*(s) dA(s)^*) |\psi_I \otimes \phi(u)\rangle - \int_0^T \langle \psi_w | \tilde{P}(t) |\psi_I \rangle dt \\ & = \int_{0 < s < t < T} [\langle \psi_w | \tilde{L}(t) \tilde{L}(s) |\psi_I \rangle u(t) u(s) - \langle \psi_w | \tilde{L}(t) \tilde{L}(s)^* |\psi_I \rangle \bar{u}(s) u(t) - \\ & \quad \langle \psi_w | \tilde{L}(t)^* \tilde{L}(s) |\psi_I \rangle \bar{u}(t) u(s) + \langle \psi_w | \tilde{L}(t)^* \tilde{L}(s)^* |\psi_I \rangle \bar{u}(t) \bar{u}(s)] dt ds \end{aligned}$$

$$-(1/2) \int_0^T <\psi_w|\tilde{L}(t)\tilde{L}(t)^*|\psi_I> dt$$

$$= \sum_{k,m=1}^2 q_{km}\theta_k\theta_m$$

where the matrix elements q_{km} are easily evaluated in terms of the system operators L_1, L_2 . Specifically,

$$\begin{aligned} q_{km} = & \\ = & \int_{0 < s < t < T} [<\psi_w|\tilde{L}_k(t)\tilde{L}_m(s)|\psi_I> u(t)u(s) - <\psi_w|\tilde{L}_k(t)\tilde{L}_m(s)^*|\psi_I> \bar{u}(s)u(t) - \\ & <\psi_W|\tilde{L}_k(t)^*\tilde{L}_m(s)|\psi_I> \bar{u}(t)u(s) + <\psi_w|\tilde{L}_k(t)^*\tilde{L}_m(s)^*|\psi_I> \bar{u}(t)\bar{u}(s)] dt ds \\ & -(1/2) \int_0^T <\psi_w|\tilde{L}_k(t)\tilde{L}_m(t)^*|\psi_I> dt, k, m = 1, 2 \end{aligned}$$

We write

$$q_{km}^R = Re(q_{km}), q_k^R = Re(q_k)$$

and then get

$$a^2 = 2 + 2.q_k^R\theta_k + 2.q_{km}^R\theta_k\theta_m$$

and

$$\mathcal{E}(\theta_1, \theta_2) =$$

$$\begin{aligned} & \| (1-a)|\psi_I \otimes \phi(u)> + U^*|\psi_w \otimes \phi(u)> \|^2 / a^2 \\ & = [(1-a)^2 + 1 + 2(1-a)Re(<\psi_w \otimes \phi(u)|U|\psi_I \otimes \phi(u)>)]/a^2 \\ & = [(1-a)^2 + 1 + (1-a)(a^2 - 2)]/a^2 = 2 - a \end{aligned}$$

The minimization of this quantity w.r.t θ_1, θ_2 is equivalent to maximizing a . If we wish to estimate θ_1, θ_2 by minimizing the sum of recovery error energies based on p pairs ($|\psi_{I,k}>, |\psi_{w,k}>$), $k = 1, 2, \dots, p$, then we will compute $a_k(\theta_1, \theta_2)$ for each pair in the above manner, ie, by replacing $|\psi_I>$ and $|\psi_w>$ respectively with $|\psi_{I,k}>$ and $|\psi_{w,k}>$ and maximize

$$a(\theta_1, \theta_2) = \sum_{k=1}^p a_k(\theta_1, \theta_2)$$

5.42 Lie group-Lie algebra approach to robot dynamics with two 3 – D links, each described by three Euler angles and an overall translational vector $a(t) \in R^3$

Suppose

$$R(t) = R_z(\phi(t))R_x(\theta(t))R_z(\psi(t)) = \exp(\phi(t)X_3).\exp(\theta(t)X_1).\exp(\psi(t)X_3)$$

where

$$X_k^T = -X_k, [X_k, X_m] = \epsilon(kmr)X_r$$

We find that

$$R'(t) = \xi(t)R(t)$$

where

$$\begin{aligned} \xi(t) &= [\phi'(t)X_3 + \theta'(t)\exp(\phi(t)ad(X_3))(X_1) + \psi'(t)\exp(\phi(t)ad(X_3)).\exp(\theta(t)ad(X_1))(X_3)] \\ &= \phi'X_3 + \theta'(X_1.\cos(\phi) + X_2.\sin(\phi)) + \psi'.\exp(\phi.ad(X_3))(X_3.\cos(\theta) - X_2.\sin(\theta)) \\ &= \phi'X_3 + \theta'(X_1.\cos(\phi) + X_2.\sin(\phi)) + \psi'(X_3.\cos(\theta) - \sin(\theta).(X_2.\cos(\phi) - X_1.\sin(\phi))) \\ &\quad = \alpha(t)X_1 + \beta(t)X_2 + \gamma(t)X_3 \end{aligned}$$

where

$$\alpha(t) = \theta'.\cos(\phi) + \psi'.\sin(\theta).\sin(\phi),$$

$$\beta(t) = \theta'.\sin(\phi) - \psi'.\sin(\theta).\cos(\phi),$$

$$\gamma(t) = \phi' + \psi'.\cos(\theta)$$

Kinetic energy of first link is given by

$$\begin{aligned} K_1(t) &= (\rho/2) \int_{B_1} \| a'(t) + R_1(t)\xi \|^2 d^3\xi \\ &= (M/2) \| a'(t) \|^2 + (1/2)Tr(R'_1(t)J_1R'_1(t)^T) \\ &\quad + Ma'(t)^T R_1(t) d_1 \end{aligned}$$

where

$$\begin{aligned} J_1 &= (\rho/2) \int_{B_1} \xi \cdot \xi^T d^3\xi \\ M d_1 &= \rho \int_{B_1} \xi \cdot d^3\xi \end{aligned}$$

Note that $R_1(t)$ is the same as $R(t)$ but with ϕ_1, θ_1, ψ_1 replacing ϕ, θ, ψ and $R'_1(t) = \xi_1(t)R_1(t)$ with $x_1(t)$ the same as $\xi(t)$ with the same replacements. Note that

$$Tr(R'_1 J_1 R'^T_1) = -Tr(\xi_1^2 R_1 J_1 R_1^T)$$

$$\begin{aligned}
&= -[\alpha^2 Tr(X_1^2 R_1 J_1 R_1^T) + \beta^2 Tr(X_2^2 R_1 J_1 R_1^T) + \gamma^2 Tr(X_3^2 R_1 J_1 R_1^T) \\
&\quad + \alpha\beta Tr((X_1 X_2 + X_2 X_1) R_1 J_1 R_1^T) + \beta\gamma Tr((X_2 X_3 + X_3 X_2) R_1 J_1 R_1^T) \\
&\quad + \gamma\alpha Tr((X_3 X_1 + X_1 X_3) R_1 J_1 R_1^T)]
\end{aligned}$$

Note that

$$\begin{aligned}
R_1 &= R_1(t) = R_z(\phi_1(t)) R_x(\theta_1(t)) R_z(\psi_1(t)) \\
&= \exp(\phi_1(t) X_3) \cdot \exp(\theta_1(t) X_1) \cdot \exp(\psi_1(t) X_3) \\
R_z(\phi_1) &= \begin{pmatrix} \cos(\phi_1) & -\sin(\phi_1) & 0 \\ \sin(\phi_1) & \cos(\phi_1) & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
R_x(\theta_1) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_1) & -\sin(\theta_1) \\ 0 & \sin(\theta_1) & \cos(\theta_1) \end{pmatrix}
\end{aligned}$$

We write

$$C_{kk} = C_{kk}(\psi, \theta, \phi) == -Tr(X_k^2 R_1 J_1 R_1^T), k = 1, 2, 3,$$

$$C_{km} = C_{km}(\psi, \theta, \phi) = -(1/2)Tr((X_k X_m + X_m X_k) R_1 J_1 R_1^T), k, m = 1, 2, 3$$

Then, we can write the rotational part of the kinetic energy of the first link as

$$\begin{aligned}
&(1/2)Tr(R'_1(t) J_1 R'_1(t)^T) = \\
&(1/2)[C_{11}\alpha_1^2 + C_{22}\beta_1^2 + C_{33}\gamma^2 + 2C_{12}\alpha_1\beta_1 + 2C_{23}\beta_1\gamma_1 + 2C_{31}\gamma_1\alpha_1)] \\
&= (1/2)[C_{11}(\theta'_1 \cdot \cos(\phi_1) + \psi'_1 \cdot \sin(\theta_1) \cdot \sin(\phi_1))^2 + C_{22}(\theta'_1 \cdot \sin(\phi_1) \\
&\quad - \psi'_1 \cdot \sin(\theta_1) \cdot \cos(\phi_1))^2 + C_{33}(\phi'_1 + \psi'_1 \cdot \cos(\theta_1))^2 \\
&\quad - 2C_{12}(\theta'_1 \cdot \cos(\phi_1) + \psi'_1 \cdot \sin(\theta_1) \cdot \sin(\phi_1))(\theta'_1 \cdot \sin(\phi_1) - \psi'_1 \cdot \sin(\theta_1) \cdot \cos(\phi_1)), \\
&\quad + 2C_{23}(\theta'_1 \cdot \sin(\phi_1) - \psi'_1 \cdot \sin(\theta_1) \cdot \cos(\phi_1))(\phi'_1 + \psi'_1 \cdot \cos(\theta_1))] \\
&\quad + 2C_{31}(\theta'_1 \cdot \cos(\phi_1) + \psi'_1 \cdot \sin(\theta_1) \cdot \sin(\phi_1))(\phi'_1 + \psi'_1 \cdot \cos(\theta_1))]
\end{aligned}$$

This can be arranged in the form

$$\begin{aligned}
(1/2)Tr(R'_1 J_1 R_1) &= (1/2)[M_{11}(\phi_1, \theta_1, \psi_1)\phi'^2_1 + M_{22}(\phi_1, \theta_1, \psi_1)\theta'^2_1 \\
&\quad + M_{33}(\phi_1, \theta_1, \psi_1)\psi'^2_1 + 2M_{12}(\phi_1, \theta_1, \psi_1)\phi'_1\theta'_1 \\
&\quad + 2M_{23}(\phi_1, \theta_1, \psi_1)\theta'_1\psi'_1 + M_{31}(\phi_1, \theta_1, \psi_1)\theta'_1\psi'_1] \\
&= (1/2)[\phi'_1, \theta'_1, \psi'_1] M(\phi_1, \theta_1, \psi_1) \begin{pmatrix} \phi'_1 \\ \theta'_1 \\ \psi'_1 \end{pmatrix}
\end{aligned}$$

The total kinetic energy of the first link is therefore

$$K_1(t) = (1/2)[\phi'_1, \theta'_1, \psi'_1] M(\phi_1, \theta_1, \psi_1) \begin{pmatrix} \phi'_1 \\ \theta'_1 \\ \psi'_1 \end{pmatrix} + (M_1/2) \| a'(t) \|^2 + M a'(t)^T R_1(t) d_1$$

The potential energy of the first link is

$$V_1(t) = M_1 g (R_1(t) d_1)_3 = M_1 g e_3^T R_1(t) d_1 = V_1(\phi_1(t), \theta_1(t), \psi_1(t))$$

The Kinetic energy of the second link is

$$K_2(t) = (\rho/2) \int_{B_2} \| d/dt(a(t) + R_1(t)p + R_2(t)(\xi - p)) \|^2 d^3\xi$$

where $R_2 R_1$ has been denoted by R_2 . It should be noted that if we use $R_2 R_1$ in place of R_2 , then physically this means that first we apply R_1 the combination of both the links around the pivot of the first link and then we apply R_2 to the second link relative to the point p at which the second link is attached to the first link. We find that

$$K_2(t) = (M_2/2) \| a'(t) + R'_1(t)p \|^2 + (1/2) Tr(R'_2(t) J_2 R'_2(t)^T) + M_2 (a'(t) + R'_1(t)p)^T R'_2(t) d_2$$

where

$$Md_2 = \rho \int_{B_2} (\xi - p) d^3\xi,$$

$$J_2 = (\rho/2) \int (\xi - p)(\xi - p)^T d^3\xi$$

$R_2(t)$ is the same as $R_1(t)$ but with ϕ_2, θ_2, ψ_2 replacing ϕ_1, θ_1, ψ_1 and we can write

$$R'_2(t) = \xi_2(t) R_2(t)$$

where $\xi_2(t)$ is the same as $\xi_1(t)$ but with ϕ_2, θ_2, ψ_2 replacing ϕ_1, θ_1, ψ_1 . Likewise, $M_{ab}(\phi_1, \theta_1, \psi_1)$ is replaced by $\tilde{M}_{ab}(\phi_2, \theta_2, \psi_2)$ where \tilde{M}_{ab} is calculated in the same way as M_{ab} but with J_1 replaced by J_2 and ϕ_1, θ_1, ψ_1 replaced by ϕ_2, θ_2, ψ_2 . Thus, the kinetic energy of the second link can be expressed as

$$\begin{aligned} K_2(t) &= (M_2/2) \| a'(t) + \xi_1(t) R_1(t)p \|^2 \\ &\quad + M_2 (a'(t) + \xi_1(t) R_1(t)p)^T \xi_2(t) R_2(t) d_2 \\ &\quad + (1/2)[\phi'_2, \theta'_2, \psi'_2] \tilde{M}(\phi_2, \theta_2, \psi_2) \begin{pmatrix} \phi'_2 \\ \theta'_2 \\ \psi'_2 \end{pmatrix} \end{aligned}$$

5.43 Linear algebra and operator theory

This course should be taught not by directly proving theorems but by taking examples from quantum computation and quantum mechanics in finite dimensional Hilbert spaces. The basic theorems that must be covered in such a course are:

- [1] Rank-Nullity theorem in finite dimensional Hilbert spaces.
- [2] Nullity, deficiency and index of a linear operator in infinite dimensional Hilbert spaces.
- [3] Computing the exponential of a linear operator in finite and infinite dimensional Banach spaces with applications to solving state variable problems, ie, linear differential equations in Banach and Hilbert spaces.
- [4] Various kinds of norms in finite and infinite dimensional vector spaces: The L^2 -norm, the sup-norm, the L^p norms for $p \geq 1$.
- [5] The Cauchy-Schwarz and Holder's inequality in finite and infinite dimensional Hilbert and Banach spaces.
- [6] The graph of a mapping between two Banach spaces.
- [7] The closure of a graph and a core of an operator: if $D_0 \subset D(T)$ is a linear manifold such that $\{(x, Tx) : x \in D_0\}$ is dense in $Gr(T) = \{(x, Tx) : x \in D(T)\}$, then D_0 is called a core for T .
- [8] convergence and weak convergence in Banach spaces.
- [9] The Baire category theorem and its application to proving the uniform boundedness principle: If X, Y are Banach spaces and $T_n : X \rightarrow Y, n = 1, 2, \dots$ is a sequence of bounded linear operators such that for each $x \in X$, the sequence $\{\|T_n x\|\}_n$ is bounded, then $\{\|T_n\|\}_n$ is a bounded sequence.
- [10] Eigenvalues and eigenvectors of a linear operator in finite dimensional vector spaces, diagonalable and non-diagonalable operators, characteristic and minimal polynomials.
- [11] Direct sum decompositions and orthogonal direct sum decompositions of a vector space.
- [12] The primary decomposition theorem and the Jordan canonical form of a matrix.
- [13] Bounded and Compact operators in infinite dimensional Banach spaces.
- [14] Relative boundedness and relative compactness of operators with applications to quantum scattering theory.
- [15] Solving linear differential equations in Hilbert spaces and its application to the Schrodinger and Heisenberg quantum mechanics.
- [16] Solving linear differential equations in Banach spaces using perturbation theory and its application to the interaction picture and Dyson series in quantum mechanics for calculating transition probabilities.
- [17] Applications of operator theory in infinite dimensional Hilbert spaces to quantum field theory especially in defining the annihilation and creation operator fields, construction of the Fock spaces for handling an infinite number of distinguishable and indistinguishable particles.
- [18] Using symmetrization and anti-symmetrization of the tensor product of a finite number of vectors to construct examples of Bosonic and Fermionic

states for a finite number of indistinguishable particles.

[19] The spectral theorem for finite dimensional normal operators, infinite dimensional compact Hermitian operators and infinite dimensional bounded and unbounded Hermitian operators in a Hilbert space. Proofs of all these theorems should be included in the course.

[20] Decomposition theorems for operators: The spectral decomposition, the QR decomposition based on Gram-Schmidt orthonormalization process, the polar and singular value decompositions.

[21] linear differential equations with random operators with applications to quantum mechanics.

[22] Explicit formulas for the scattering matrix for Schrodinger operators.

[24] Explicit formula for the scattering matrix for Dirac operators.

[25] Proof of the essential self-adjointness of the Schrodinger and Dirac operators, ie, the self-adjointness of the closure of the Schrodinger operators and Dirac operators in the space of infinitely differentiable functions with compact support.

[26] Examples of how to define the domain of unbounded operators like the position, momentum, kinetic energy and energy operators.

[27] When is a symmetric operator essentially self-adjoint.

[28] Proof that the adjoint of an operator is closed and hence any symmetric operator is closable.

[29] Existence of the wave operators.

If A, B are two Hermitian operators with non-trivial continuous spectrum, then we wish to define the domains D_{\pm} on which the limits

$$\Omega_{\pm} = \lim_{t \rightarrow \pm\infty} \exp(itB) \cdot \exp(-itA)$$

exist. $\Omega_+ f$ exists if $\int_0^\infty \| (B - A) \cdot \exp(-itA) f \| dt < \infty$. In particular, if we are able to find a potential $V = B - A$ such that $\| V \cdot \exp(-it) f \|$ decays as t^a as $t \rightarrow \infty$ where $a < -1$, then $\Omega_+ f$ exists. Some interesting results exist in this direction like notions of relative smoothness. Suppose we formally consider Parseval's relation in the form

$$\begin{aligned} \int_{\mathbb{R}} \exp(i\lambda t) \cdot \exp(-\mu|t|) \cdot V \cdot \exp(-itA) dt &= V[-(i\lambda - \mu - iA)^{-1} + (i\lambda + \mu - iA)^{-1}] \\ &= V[i(A - \lambda + i\mu)^{-1} - i(A - \lambda - i\mu)^{-1}] = -iV[R(\lambda + i\mu) - R(\lambda - i\mu)] \end{aligned}$$

where $R(\xi)$ denotes the resolvent of A . Then, Parseval's theorem gives

$$\begin{aligned} &\int_{\mathbb{R}} \| \exp(-\mu|t|) \cdot V \cdot \exp(-itA) f \|^2 dt \\ &= (2\pi)^{-1} \int_{\mathbb{R}} \| V \cdot (R(\lambda + i\mu) - R(\lambda - i\mu)) f \|^2 d\lambda \end{aligned}$$

A particular case of this equation is obtained by replacing f with $E(J)f$ where J is a finite interval and $E(\cdot)$ is the spectral measure of A . We then get

$$\int_{\mathbb{R}} \exp(-2|\mu|t) \| V \cdot \exp(-itA) E(J) f \|^2 dt$$

$$\begin{aligned}
&= (2\pi)^{-1} \int_{\mathbb{R}} \| V.(R(\lambda + i\mu) - R(\lambda - i\mu))E(J)f \|_{}^2 d\lambda \\
&= (2\pi)^{-1} \int_{\mathbb{R}} \| V.E(J)(R(\lambda + i\mu) - R(\lambda - i\mu))f \|_{}^2 d\lambda
\end{aligned}$$

since $E(J)$ commutes with A and hence also with the resolvent of A , $R(z) = (A - z)^{-1}$. We now observe that if $\lambda \notin J$, then assuming that J is a closed interval, we have

$$\lim_{\mu \rightarrow 0} E(J)(R(\lambda + i\mu) - R(\lambda - i\mu)) = 0$$

Thus, we get the useful identity,

$$\begin{aligned}
&\lim_{\mu \rightarrow 0+} (2\pi)^{-1} \int_J \| V.E(J).(R(\lambda + i\mu) - R(\lambda - i\mu))f \|_{}^2 d\lambda \\
&= \int_{\mathbb{R}} \| V.E(J).exp(-itA).f \|_{}^2 dt
\end{aligned}$$

We now observe that

$$\begin{aligned}
E(J)(R(\lambda + i\mu) - R(\lambda - i\mu)) &= \int_J ((x - \lambda + i\mu)^{-1} - (x - \lambda - i\mu)^{-1})E(dx) \\
&= \int_J \frac{2i\mu}{(x - \lambda)^2 + \mu^2} E(dx)
\end{aligned}$$

and hence if J is a subset of the absolutely continuous spectrum of A , then we have for any vectors f, g and $\lambda \in J$,

$$\langle g, E(J)(R(\lambda + i\mu) - R(\lambda - i\mu))f \rangle =$$

$$\int_J \frac{2i\mu}{(x - \lambda)^2 + \mu^2} (d \langle g, E(x)f \rangle / dx) dx$$

which converges as $\mu \rightarrow 0$ to

$$\int_J \delta(x - \lambda) (d \langle g, E(x)f \rangle / dx) dx = 2\pi i d \langle g, E(\lambda)f \rangle / d\lambda$$

Now we relate this discussion to the existence of wave operators. Let $W(t) = exp(iBt).exp(-iAt)$ where A, B are two Hermitian operators. Then, for $0 \leq t_1 < t_2$, we have

$$W(t_2) - W(t_1) = i \int_{t_1}^{t_2} exp(iBt)(B - A).exp(-iAt) dt$$

Suppose

$$B - A = C_1^* C_2$$

then we have

$$\begin{aligned} | \langle g, (W(t_2) - W(t_1))f \rangle |^2 &= \left| \int_{t_1}^{t_2} \langle C_1 \exp(-iBt)g, C_2 \exp(-iAt)f \rangle dt \right|^2 \\ &\leq \left(\int_{t_1}^{t_2} \|C_1 \exp(-iBt)g\|^2 dt \right) \cdot \left(\int_{t_1}^{t_2} \|C_2 \exp(-iAt)f\|^2 dt \right) \end{aligned}$$

It follows that if $\|C_1 \exp(-iBt)g\|^2$ and $\|C_2 \exp(-iAt)f\|^2$ are integrable on $[0, \infty)$, then

$$\lim_{t_1, t_2 \rightarrow \infty} |\langle g, W(t_2)f \rangle - \langle g, W(t_1)f \rangle| = 0$$

and hence $\lim_{t \rightarrow \infty} \langle g, W(t)f \rangle$ exists. This fact suggests the strong link between the existence of wave operators and the notion of relative smoothness of an operator. An operator V is A -relatively smooth on J is

$$\lim_{\mu \rightarrow 0+} \int_J \|V.E(J)(R(\lambda + i\mu) - R(\lambda - i\mu))\|^2 d\lambda < \infty$$

We have already noted that this inequality implies that

$$\int_{\mathbb{R}} \|V.E(J)\exp(-itA)f\|^2 dt < \infty$$

for any f and hence we deduce by the above argument that if $V = C_1^* C_2 = B - A$ is such that C_1 is B -smooth on J and C_2 is A -smooth on J , then

$$\lim_{t \rightarrow \infty} \langle E(J_2)g, W(t)E_0(J_1)f \rangle$$

exists where E is the spectral measure of B and E_0 is the spectral measure of A and f belongs to the absolutely continuous spectrum of A while g belongs to the absolutely continuous spectrum of B . Equivalently, we may assume that f, g are arbitrary but J_1 is an interval of absolute continuity of $E_0(\cdot)$ while J_2 is an interval of absolute continuity of $E(\cdot)$. This proves in some sense the asymptotic completeness of the wave operator $W_+ = \lim_{t \rightarrow \infty} W(t)$, ie, W_+ is defined on the entire space of absolutely continuous vectors for A and W_+ maps this space $\mathcal{H}_{ac}(A)$ into $\mathcal{H}_{ac}(B)$, the space of absolutely continuous vectors for B .

[30] Design of quantum gates using perturbed Hamiltonian operators combined with Schrodinger dynamics.

[31] Basic classical and Cq information theory. Proof of the Cq Shannon coding theorem.

[32] Teleportation via shared entanglement.

[33] Generation of almost completely entangled states from mixed states.

[35] Generation of mixed states from entangled states.

[36] Mourre's conjugate operator method in proving asymptotic completeness of the wave operators in scattering theory. Relatively smooth operators.

[37] Trace class and Hilbert-Schmidt operators. Generalization of the svd to trace class operators in infinite dimensional Hilbert spaces.

[38] Hilbert-Schmidt operators are compact, finite rank operators are compact operators, every compact operator in a Banach space is the operator norm limit of a sequence of finite rank operators.

[39] Conditions on a symmetric operator to be essentially self-adjoint and self-adjoint. A is symmetric if $A \subset A^*$. A symmetric operator A is essentially self-adjoint iff $R(A \pm i)^{\perp} = 0$ ie if $R(A \pm i)$ are dense in \mathcal{H} . A symmetric operator A is self-adjoint iff $R(A \pm i) = \mathcal{H}$

[40] Application of relative boundedness theory to proving self-adjointness of an operator under a small perturbation.

[41] Problems involving the existence of wave operators for time dependent scattering potentials.

[42] Notion of the essential spectrum of a self-adjoint operator in an infinite dimensional Hilbert space and its characterization.

[43] Design of large sized quantum gates using quantum electrodynamics.

Example: An electron interacts with the em field of the nucleus gets scattered and then releases a photon and again gets scattered. The em field of the nucleus in the four momentum domain is defined by $\gamma^\mu A_\mu(Q)$ while the initial four momentum of the electron is p . After scattering by the nucleus, its momentum becomes $p + Q$ and then it releases a photon of momentum q and gets scattered again to a momentum of p' . By four momentum conservation, we have $p + Q = p' + q$. The scattering amplitude after taking into account all the momentum components of the nuclear field is proportional to

$$\begin{aligned} & \int \bar{u}(p') \gamma^\nu q_\nu S(p+Q) \gamma^\mu A_\mu(Q) u(p) d^4 Q \\ &= u(p')^* \left[\int \gamma^\nu S(p+Q) \gamma^\mu u(p) (p+Q - p')_\nu A_\mu(Q) d^4 Q \right] u(p) \end{aligned}$$

where

$$S(p) = (\gamma \cdot p - m + i\epsilon)^{-1}$$

is the electron propagator. After an appropriate scaling, it follows that the generator of the scattering matrix corresponding to this problem is proportional to

$$G(p', p) = \int \gamma^\nu S(p+Q) \gamma^\mu (p+Q - p')_\nu A_\mu(Q) d^4 Q$$

This generator is to be matched to a given generator $G_0(p', p)$ by tuning the nuclear field $A_\mu(Q)$ so that

$$\int \| G(p', p) - G_0(p', p) \|^2 W(p', p) d^4 p d^4 p'$$

is a minimum subject to energy constraints on the nuclear potential like the total energy in the field integrated over time is fixed. In the four momentum

domain, this constraint reads

$$\int \left[\sum_{r=1}^3 (jQ_r A_0(Q) - jQ^0 A_r(Q))^2 + (1/2) \sum_{r,s=1}^3 (jQ_r A_s(Q) - jQ_s A_r(Q))^2 \right] d^4 Q = E$$

This quadratic constraint can be expressed in the form

$$\int K^{\mu\nu}(Q) A_\mu(Q) A_\nu(Q)^* d^4 Q = E$$

Optimization of the nuclear field for the design of a quantum gate is thus a least squares problem and hence in a course on introductory linear algebra and functional analysis, we must introduce various kinds least squares problems and their solutions using generalized inverses of matrices. All kinds of generalized inverses like the least squares generalized inverse, the minimum norm generalized inverse and the Moore-Penrose least squares minimum norm generalized inverse of a matrix can be obtained in a very elementary manner from the svd of a matrix.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Chapter 6

Applications of Lie groups and Lie algebras, filtering, field quantization, Numerical methods for quantum mechanical problems

6.1 quantum random walk

Here, we introduce a new notion of a quantum random walk on a graphene lattice. The state of the system at time t is of the form

$$|\psi(t)\rangle = |i, j\rangle \otimes [a(0), a(1), a(2), a(3)]^T$$

The state of the coin at time t is $[a(0), a(1), a(2), a(3)]^T$. If $a(0)$ is the outcome of the coin, then the state of the random walker at time $t+1$ is $|i+1, j\rangle$ while if it is $a(1)$, then the state of the walker at time $t+1$ is $|i-1, j\rangle$ and if it is $a(2)$, then the walker state at time $t+1$ is $|i, j+1\rangle$ and finally, if it is $a(3)$, then the state is $|i, j-1\rangle$. Here, the quantum random walk takes place on the 2-D lattice. and the above transition dynamics can be described as

$$\begin{aligned} |\psi(t+1)\rangle &= U(t+1, t)|\psi(t)\rangle = |i+1, j\rangle \otimes [a(0), 0, 0, 0]^T + \\ &|i-1, j\rangle \otimes [0, a(1), 0, 0]^T + |i, j+1\rangle \otimes [0, 0, a(2), 0]^T + |i, j-1\rangle \otimes [0, 0, 0, a(3)]^T \end{aligned}$$

When interpreted classically, the walker moves one step up with probability $|a(0)|^2$, one step down with probability $|a(1)|^2$, one step to the right with probability $|a(2)|^2$ and one step down with probability $|a(3)|^2$. However, unlike the

classical case this is not the end of the story. We can superpose different states and get interference states. moreover, in the above unitary dynamics, the coin's state does not change. We can also incorporate coin state changes into this dynamics. However, it would be good to see some examples in which (a) the coin state also changes and the probability of the random walker being in a state $|i, j\rangle$ after time t is computed given the initial state is $|i_0, j_0\rangle$, (b) Some mention of how one can extend the quantum walk to continuous time quantum processes and how noise can be incorporated into the dynamics. Some ideas of how Hamiltonians can be used to describe this quantum dynamics can be mentioned. Also computation of the probability after time t of the walker and the coin being in some superposition state can be explained by means of some examples. This would enable us to actually see the difference between the classical random walk and the quantum random walk. In the classical random walk, we talk only about the probability of the walker being at the site (i, j) after time t is given classical text's like William Feller's. Here, we talk about the probability amplitude. It may be mentioned that noise can be incorporated into the walker dynamics in the form

$$|\psi(t+1)\rangle = (U_0(t+1, t) + \delta U(t+1, t))|\psi(t)\rangle$$

where $\delta U(t+1, t)$ is a random operator so that the unitarity condition holds:

$$(U_0(t+1, t) + \delta U(t+1, t))^*(U_0(t+1, t) + \delta U(t+1, t)) = I$$

Some work in the literature describe a nice extension of classical single qubit teleportation theory to two qubit teleportation. In single qubit teleportation theory, Alice and Bob share a two qubit entangled state. Alice then adjoins a single qubit state $|\psi_A\rangle$ to this entangled state thereby acquiring possession of two qubits as compared to Bob who has one qubit entangled to one of the two qubits of Alice's state. Alice performs a measurement on her two qubits and after she notes her outcome described by two classical bits, she transmits this two classical bit outcome to Bob who then applies an appropriate one qubit gate based on his two received classical bits to recover $|\psi_A\rangle$. Depending on Alice's measurement outcome, the overall state of Alice and Bob collapse to a three qubit state and hence Bob is able to recover the state $|\psi_A\rangle$ which Alice wishes to transmit by applying an appropriate gate to his qubit depending upon Alice's collapsed outcome.

Some works in the literature offer a simple extension of single qubit teleportation to two qubit teleportation. The method involves a third person Charlie so that Alice, Bob and Charlie share a six qubit entangled state with each of them having two qubits. Alice adjoins to this entangled state, another two qubit state $|\psi_A\rangle$ via the tensor product that she wishes to transmit to Bob. For that purpose, she makes a measurement on four of her qubits and informs the result to Bob and Charlie via classical communication. Charlie then applies an appropriate gate to two of his qubits based on Alice's classical communication

and reports his result to Bob via classical communication. Based on Alice's and Charlie's classical communication, Bob applies a gate to his two qubits to recover the two qubit $|\psi_A\rangle$ that Alice wishes to transmit. A fundamental new idea is involved here, namely that involving communication from a third party. The following possible generalization may also be considered: If there are N persons sharing an entangled state $|\psi_E\rangle$ of nk qubits, with each person holding k qubits, and if the first person wishes to transmit a k qubit state $|\psi_1\rangle$ to the second person, then he can adjoin his state to this entangled state so that the overall state of the N persons is $|\psi_1\rangle|\psi_E\rangle$. The first person then applies a POVM measurement $\{M_m\}$ to this $2k$ qubit system and notes the outcome m . The state of the overall system then collapses to $\sqrt{M_m} \otimes I_{(n-1)k} |\psi_1 \otimes \psi_E\rangle$ apart from a normalization factor. When the first person communicates his outcome m to all the remaining $k-1$ users, the these $k-1$ users can each apply k -qubit gates to their state in a sequence depending on the outcome m causing the state of the overall system to collapse and after receiving their reports via classical measurements, the second person can apply a gate to decode $|\psi_1\rangle$. Is it possible to develop a general theory along these lines ?

6.2 Lie group-Lie algebra theoretic coordinate free formulation of the equations of motion of a robot with N 3-D links with the orientation of each link described by an arbitrary element of $SO(3)$ and taking in addition into account a translation of the base pivot of the first link

Exercise: Write down the Lagrangian of a system of connected 3-D links taking into account a rotation matrix valued function of time for each link by expressing the kinetic and potential energies of this system respectively as quadratic forms in the time derivatives of the rotation matrices and as linear functions of the rotation matrices. Set up the Euler-Lagrange equations for this system in the Lie algebra domain by expressing each rotation matrix as the exponential of an $SO(3)$ Lie algebra element and using the standard formula for the differential of the exponential map.

Reference: V.S.Varadarajan, "Lie groups, Lie algebras and their representations", Springer, 1984.

6.3 Numerical methods for computing transition probabilities for photons, gravitons, Klein-Gordon Bosons, Dirac Fermions and non-Abelian matter and gauge particles from inside the critical radius to outside of a Schwarzschild blackhole with quantum gate design applications

Exercise: Set up the Lagrangian for each of the above theories in a background metric and derive by applying the Legendre transformation, the corresponding Hamiltonian for the associated particle in the background metric of the blackhole and derive the wave function evolution from these equations. In the case of gravitons, set up the Lagrangian and hence the Hamiltonian for small metric perturbations around the background classical metric and hence derive the wave function evolution for the metric perturbations. Show that if initially, the wave function is concentrated over a region in the interior of the blackhole, then finally, ie, after evolution for a finite time duration, the wave function will be non-zero even at points to the exterior of the blackhole. Such a tunneling phenomenon of particles across the Schwarzschild radius is not possible classically.

References:

- [1] Hawking, "Particle emission by blackholes"
- [2] Hartle and Hawking, "The wave function of the universe"
- [3] S.Chandrasekhar, "The mathematical theory of blackholes."
- [4] Steven Weinberg, "Gravitation and Cosmology: Principles and applications of the general theory of relativity", Wiley.
- [5] Steven Weinberg, "The quantum theory of fields, Vol.II, Cambridge University Press.

The proposal: The wave equation for a KG particle in the metric $g_{\mu\nu}$ of a Schwarzschild blackhole and in the presence of an external em field is given by

$$[g^{\mu\nu}(x)(i\nabla_\mu + eA_\mu(x))(i\partial_\nu + eA_\nu(x)) - m^2]\phi(x) = 0$$

We wish to design a unitary transformation U_d in the Hilbert space of states that will map a set of initial states $\phi_k(r), k = 1, 2, \dots, N$ into a corresponding set of output states $\psi_k(r), k = 1, 2, \dots, N$. For this, we solve the above KG equation perturbatively w.r.t. the em potentials and choose the potentials so that a least squares match after a fixed time T occurs w.r.t. the final states starting at the corresponding initial states. The problem is that the KG equation being second order in time gives forward as well as backward travelling waves which is not compatible with the philosophy of time evolution in quantum mechanics. We therefore look at the Dirac and non-Abelian gauge field and matter field

equations in quantum mechanics using the gravitational connection. Let Γ_μ be the spinor connection of the gravitational field [4] and let $A_\mu^\alpha(x)\tau_\alpha$ where $\tau_\alpha, \alpha = 1, 2, \dots, p$ are the Hermitian Lie algebra generators of the non-Abelian unitary gauge group. Then taking into account the gauge fields, the aim is to match pairs of input-output states, each having $4N$ components and satisfying the Yang-Mills equation [5]

$$[(\gamma^a V_a^\mu(x) \otimes I_N)(i\partial_\mu + i\Gamma_\mu(x) \otimes I_N) + (I_4 \otimes \tau_\alpha)A_\mu^\alpha(x) - mI_{4N}]\psi(x) = 0$$

where $V_a^\mu(x)$ is the tetrad of the gravitational field. The gauge fields $A_\mu^\alpha(x)$ satisfy the gauge field equations but we shall assume that they can be controlled to get an optimal match between the pairs of input-output states. When we assume the presence of an external control current source for the Yang-Mills gauge field generation, then we must optimize the evolution w.r.t. these current sources $J_\mu^\alpha(x)$ and not w.r.t the gauge fields. The above matter field as well as the gauge field equations are derived from the Lagrangian density

$$\begin{aligned} Re[\psi(x)^*(\gamma^0 \otimes I_N)[(\gamma^a V_a^\mu(x) \otimes I_N)(i\partial_\mu + i\Gamma_\mu(x) \otimes I_N) + (I_4 \otimes \tau_\alpha)A_\mu^\alpha(x) - mI_{4N}]\psi(x)] \\ + K_1.Tr(F_{\mu\nu}F^{\mu\nu}(x)) + K_2.J_\alpha^\mu(x)A_\mu^\alpha(x) \end{aligned}$$

where

$$F_{\mu\nu} = [\nabla_\mu, \nabla_\nu]$$

$$\partial_\mu + \Gamma_\mu(x) \otimes I_N + iA_\mu^\alpha(x)I_4 \otimes \tau_\alpha$$

These equations are nonlinear and must be solved perturbatively to obtain approximate expressions for the evolved matter states in terms of the control non-Abelian current density and then the optimization must be performed.

6.4 Numerical methods for designing quantum gates based on quantum scattering theory for a Schrodinger projectile interacting with a potential

6.4.1 General theory

It is a well known result in quantum scattering theory (W.O.Amrein, "Hilbert space methods in quantum mechanics") that if H_0 is the unperturbed Hamiltonian of the projectile and $H = H_0 + V$ is its perturbed form by a scattering centre, then the scattering matrix at energy λ can be expressed as

$$\begin{aligned} S(\lambda) = E_0(d\lambda)\Omega_+^*\Omega_-E_0(d\lambda)/d\lambda &= I + 2\pi i.E_0(d\lambda)(V - V(H - \lambda)^{-1}V).E_0(d\lambda)/d\lambda \\ &= I + 2\pi iR(\lambda) \end{aligned}$$

where $E_0(\cdot)$ is the spectral measure of the unperturbed Hamiltonian H_0 . This formula enables us to express the scattering matrix generator $R(\lambda)$ entirely in terms of the scattering potential V or more precisely its spatial Fourier transform and that this representation enables us to express the scattering generator kernel as

$$R_\lambda(\omega'|\omega)$$

where ω, ω' are unit vectors in \mathbb{R}^3 corresponding to the initial and final directions of the momentum of the projectile. Thus, we can design a quantum gate by choosing the potential V so that this kernel matches a given Hermitian kernel $R_g(\omega|\omega')$ corresponding to a unitary gate whose rows and columns are indexed by points on the unit sphere.

Reference: W.O.Amrein, "Hilbert space methods in quantum mechanics".

6.5 Quantization of a robot in the Lie-group domain when the robot has N 3-D links

6.5.1 General theory

The total kinetic energy of the system can be expressed in the form

$$K(t) = \sum_{n=1}^N (\rho/2) \int_{B_k} \| a'(t) + R'_1(t)p_1 + R'_2(t)p_2 + \dots + R'_{k-1}(t)p_{k-1} + R'_k(t)(\xi - p_k) \|^2 d^3\xi$$

where $a(t) \in \mathbb{R}^3$ is the translation vector of the pivot of the first link and

$$R_k(t) = S_k(t)S_{k-1}(t)\dots S_1(t)$$

with $S_1 = R_1$ and S_k being the rotation applied to the k^{th} link relative to the $(k-1)^{th}$ link. This can be expressed as

$$\begin{aligned} K(t) &= \sum_{k=1}^N [(M_k/2) \| a'(t) + R'_1(t)p_1 + R'_2(t)p_2 + \dots + R'_{k-1}(t)p_{k-1} \|^2 \\ &\quad + (1/2) \text{Tr}(R'_k(t)J_k R'_k(t)^T) + M_k(a'(t) + R'_1(t)p_1 + \dots + R'_{k-1}(t)p_{k-1}, R'_k(t)d_k)] \\ &= K_0(a(t), a'(t), R_k(t), R'_k(t), k = 1, 2, \dots, N) \end{aligned}$$

The potential energy of this link system can be expressed as

$$\begin{aligned} V(t) &= V_0(a(t), R_k(t), k = 1, 2, \dots, N) = \\ &= \sum_{k=1}^N M_k(e_3, a(t) + R_1(t)p_1 + R_2(t)p_2 + \dots + R_{k-1}(t)p_{k-1} + R_k(t)d_k) \end{aligned}$$

The Lagrangian of the system taking into account the constraints $R_k(t) \in SO(3)$, $k = 1, 2, \dots, N$ is given by

$$L(a(t), a'(t), R_k(t), R'_k(t), \Lambda_k(t), k = 1, 2, \dots, N) = \\ K_0 - V_0 - \sum_{k=1}^N \text{Tr}(\Lambda_k(t)(R_k(t)^T R_k(t) - I))$$

The Euler-Lagrange equations give

$$(d/dt)\delta L/\delta a'(t) = \delta L/\delta a, \\ (d/dt)\delta L/\delta R'_k(t) - \delta L/\delta R_k(t) = 0, \\ \delta L/\delta \Lambda_k(t) = 0$$

These give

$$(d/dt)\delta L_0/\delta a'(t) = \delta L_0/\delta a, \\ (d/dt)\delta L_0/\delta R'_k(t) - \delta L_0/\delta R_k(t) = (\Lambda_k + \Lambda_k^T)R_k^T \\ R_k^T R_k = I$$

for $k = 1, 2, \dots, N$, where

$$L_0 = K_0 - V_0$$

Eliminating the Lagrange multipliers, we get as our equations of motion,

$$(d/dt)\delta L_0/\delta a'(t) = \delta L_0/\delta a, \\ X_k^T = X_k$$

where

$$X_k = [(d/dt)\delta L_0/\delta R'_k(t) - \delta L_0/\delta R_k(t)]R_k$$

The last two equations can be expressed in the form

$$P_k R_k = R_k^T P_k^T$$

where

$$P_k = (d/dt)\delta L_0/\delta R'_k(t) - \delta L_0/\delta R_k(t)$$

Thus,

$$P_k R_k R_k^T = R_k^T P_k^T R_k^T$$

If this equation ensures that $R_k \in SO(3)$, then we must have

$$P_k = R_k^T P_k^T R_k^T$$

or equivalently,

$$R_k P_k R_k = P_k^T$$

This identity can be verified by direct calculation. Now, we attempt the quantization problem. We write for the canonical momentum corresponding to R_k :

$$\pi_k = \delta L/\delta R'_k$$

6.6 Lie group formulation of the single 3-D robot link in the presence of gravitation and external torque

6.6.1 General theory

Suppose the pivot of the top is fixed. After time t , the top suffers a rotation $R(t) \in SO(3)$. Its kinetic energy can be expressed as

$$K(t) = (1/2)Tr(R'(t)JR'(t)^T)$$

Let $\Omega(t)$ be the angular velocity tensor in the fixed frame. Thus,

$$R'(t) = \Omega(t)R(t)$$

and

$$R'' = \Omega'R + \Omega R' = (\Omega' + \Omega^2)R = ZR$$

where

$$Z = \Omega' + \Omega^2$$

Note that

$$\Omega^T = -\Omega$$

since

$$R^T R = I$$

gives on differentiation

$$R'^T R + R^T R' = 0$$

or since $R' = \Omega R$,

$$R^T \Omega^T R + R^T \Omega R = 0$$

so that

$$\Omega^T + \Omega = 0$$

Now the potential energy can be expressed as

$$V(t) = mge_e^T R(t)d = Tr(R(t)A), A = mgde_3^T$$

Finally, suppose τ_1, τ_2, τ_3 are the torques around the three coordinate axes and correspondingly, $\omega_1, \omega_2, \omega_3$ are the three angular velocities. The contribution to the torque Lagrangian is

$$\sum_k \tilde{\tau}_k(t)\omega_k(t)$$

where

$$\tilde{\tau}_k(t) = \int_0^t \tau_k(s)ds$$

If for example, we write in the scalar case, $\omega = \phi'(t)$, then the above Torque Lagrangian becomes

$$L_T = -\tilde{\tau}(t)\phi'(t)$$

and when we add this to the Lagrangian of the free system, we get the correct equations of motion:

$$L = L_0 + L_T, d/dt(\partial L/\partial \dot{\phi}) = \partial L/\partial \phi$$

gives

$$d/dt(\partial L_0/\partial \dot{\phi}) - \tilde{\tau}' - \partial L_0/\partial \phi = 0$$

or equivalently,

$$d/dt(\partial L_0/\partial \dot{\phi}) - \partial L_0/\partial \phi = \tau$$

Note that we could also use $\tau\phi$ for the torque Lagrangian, but the former is preferable as it leads to an immediate generalization to the general case of 3-D rotations. In the general case, we define the torque tensor as the antisymmetric matrix

$$T(t) = \tau_1 X_1 + \tau_2 X_2 + \tau_3 X_3$$

where X_1, X_2, X_3 are the standard Lie algebra generators of $SO(3)$. Likewise,

$$\Omega = \omega_1 X_1 + \omega_2 X_2 + \omega_3 X_3$$

Then, we write

$$\tilde{T}(t) = \int_0^t T(s)ds = \sum_k \tilde{\tau}_k \omega_k$$

The torque Lagrangian is then

$$-\sum_k (\tilde{\tau}_k \omega_k) = (1/2)Tr(\tilde{T}(t)\Omega(t)) = (1/2)Tr(\tilde{T}(t)R'(t)R(t)^T)$$

The total Lagragian of the 3-D link taking gravitation and torque into account and also the constraint $R^T R = I$ via the Lagrange multiplier $\Lambda(t)$ is then

$$L(R, R', \Lambda) = (1/2)Tr(R' J R'^T) - Tr(RA) + (1/2)Tr(\tilde{T} R' R^T) - Tr(\Lambda(R^T R - I))$$

Setting the variational derivative of L w.r.t. $R(t)$ to zero then gives

$$JR''^T - A - (1/2)(R^T \tilde{T})' - (1/2)R'^T \tilde{T} - (\Lambda + \Lambda^T)R^T = 0$$

This equation can be expressed after using the constraint $R^T = R^{-1}$ to

$$JR''^T R - AR - (1/2)R^T TR = \Lambda + \Lambda^T$$

It follows that the lhs is symmetric, ie,

$$JR''^T R - AR - (1/2)R^T TR = R^T R'' J - R^T A^T + (1/2)R^T TR -- (1)$$

Note that

$$T^T = -T$$

Still (1) is not in an implementable form since it requires the constraint $R^T R = I$ to be imposed as an additional equation. We can however obtain an implementable equation even without having to use Euler angles as follows. Using $R'' = ZR$, we write (1) as

$$JR^T Z^T R - AR - R^T TR + R^T A^T - R^T ZRJ = 0 \quad \dots \quad (2)$$

Note that

$$Z = \Omega' + \Omega^2$$

We define

$$Y = R^T \Omega R = R^{-1} \Omega R = R^{-1} R'$$

Then,

$$\begin{aligned} Y' &= -R^{-1} R' R^{-1} \Omega R + R^{-1} \Omega' R + R^{-1} \Omega R' \\ &= -R^{-1} \Omega^2 R + R^{-1} \Omega' R + R^{-1} \Omega^2 R = R^{-1} \Omega' R \end{aligned}$$

Also,

$$Y^2 = R^{-1} \Omega^2 R$$

Thus,

$$Y' + Y^2 = R^{-1} (\Omega' + \Omega^2) R = R^{-1} ZR = R^T ZR$$

Also

$$Y^T = -Y, Y^{2T} = Y^2, Y'^T = -Y'$$

Hence, we can write (2) in the form

$$J(-Y' + Y^2) - (Y' + Y^2)J + R^T A^T - AR - R^T TR = 0 \quad \dots \quad (3)$$

Our final set of Lie theoretic equations for the 3-D link is therefore

$$-[J, Y']_+ + [J, Y^2] + R^T A^T - AR - R^T TR = 0 \quad \dots \quad (4a)$$

$$R' = RY \quad \dots \quad (4b)$$

$$Y^T = -Y \quad \dots \quad (4c)$$

6.7 Quantum antennas based on non-Abelian matter and gauge fields

6.7.1 General theory

The matter action has the form

$$L_M(\psi, D_\mu \psi) = \psi^* \gamma^0 [i\gamma^\mu D_\mu - m] \psi$$

where D_μ is the gauge covariant derivative

$$D_\mu = \partial_\mu - ieA_\mu = \partial_\mu - ieA_\mu^a \tau_a$$

with $\{\tau_a\}$ being the Hermitian generators of the gauge group which is a subgroup of $U(N)$. Thus, $\exp(it\tau_a)$ is an $N \times N$ unitary matrix for all indices a and $t \in \mathbb{R}$. The curvature of this covariant derivative gives the gauge field tensor:

$$-ieF_{\mu\nu} = [D_\mu, D_\nu] = -ie(A_{\nu,\mu} - A_{\mu,\nu}) - e^2[A_\mu, A_\nu]$$

or equivalently,

$$\begin{aligned} F_{\mu\nu} &= A_{\nu,\mu} - A_{\mu,\nu} - ieA_\mu^a A_\nu^b [\tau_a, \tau_b] \\ &= A_{\nu,\mu} - A_{\mu,\nu} + eA_\mu^a A_\nu^b C(abc)\tau_c \\ &= F_{\mu\nu}^c \tau_c \end{aligned}$$

where

$$F_{\mu\nu}^c = A_{\nu,\mu}^c - A_{\mu,\nu}^c + eC(abc)A_\mu^a A_\nu^b$$

summation over the repeated gauge Lie algebra index c being implied. The gauge field Lagrangian density can be taken as $g_{ab}F_{\mu\nu}^a F^{\mu\nu b}$ where $g_{ab} = \text{Tr}(\tau_a \tau_b)$. To see that this is invariant under the gauge group action, we need only note that by definition if $g(x)$ is a space-time dependent element of the gauge group, then under this action on the matter wave function $\psi(x)$, we require the gauge field A_μ to transform to A'_μ in such a way so that

$$D'_\mu g\psi = gD_\mu \psi, D'_\mu = \partial_\mu - ieA'_\mu$$

Equivalently,

$$D'_\mu = \partial_\mu - ieA'_\mu = gD_\mu g^{-1} = g(\partial_\mu - ieA_\mu)g^{-1}$$

as differential operators, ie,

$$\begin{aligned} -ieA'_\mu &= g(\partial_\mu g^{-1}) - ieg.A_\mu.g^{-1} \\ &= -(\partial_\mu g)g^{-1} - ieg.A_\mu.g^{-1} \end{aligned}$$

or equivalently,

$$A'_\mu = g.A_\mu.g^{-1} - (i/e)(\partial_\mu g)g^{-1} \quad (1)$$

This equation is the non-Abelian generalization of the $U(1)$ gauge transformation of the electromagnetic potential and the corresponding phase transformation of the wave function in Schrodinger's and Dirac's wave equation in the presence of an electromagnetic field. The $U(1)$ phase transformation of the wave function is $\psi(x) \rightarrow \exp(i\phi(x))\psi(x)$ where $\phi(x) \in \mathbb{R}$ so that $\exp(i\phi(x)) \in U(1)$. Correspondingly if $A_\mu(x) \in \mathbb{R}$ is the electromagnetic four potential, then its transformation law as per the required transformation of the covariant derivative is given by

$$\exp(i\phi(x))(\partial_\mu + ieA_\mu(x))\exp(-i\phi(x)) = \partial_\mu + ieA'_\mu(x)$$

which results in

$$ieA'_\mu = -i\phi_{,\mu} + ieA_\mu$$

or equivalently,

$$A'_\mu = A_\mu - (1/e)\phi_{,\mu} \quad \dots \quad (2)$$

This is precisely the Lorentz gauge transformation of the em potential which leaves the em field invariant. (1) is the non-Abelian ie $U(N)$ generalization of the Abelian $U(1)$ situation (2).

If the matter action is built out of the scalar quantities $\psi^*\psi, \psi^*D_\mu\psi, (D^\mu\psi)^*D_\mu\psi$, then it will automatically be gauge invariant since for $g = g(x) \in U(N)$,

$$\psi'^*\psi' = (g\psi)^*(g\psi) = \psi^*g^*g\psi = \psi^*\psi,$$

$$\psi'^*D'_\mu\psi' = (g\psi)^*gD_\mu\psi = \psi^*g^*gD_\mu\psi = \psi^*D_\mu\psi,$$

$$(D'^\mu\psi')^*(D_\mu\psi') = (gD^\mu\psi)^*(gD_\mu\psi) = (D^\mu\psi)^*(D_\mu\psi)$$

More generally, if $L_G(\psi_0, \psi_1, \dots, \psi_K, \psi_0^*, \psi_1^*, \dots, \psi_M^*)$ is invariant under the gauge group ie, $L_G(g\psi_1, \dots, g\psi_K, \psi_1^*g^*, \dots, \psi_M^*g^*) = L_G(\psi_1, \dots, \psi_K, \psi_1^*, \dots, \psi_M^*)$, then the following matter Lagrangian density

$$L_G(\psi, D_{\mu_1}D_{\mu_2}, \dots, D_{\mu_k}\psi, (\mu_1, \dots, \mu_k) \in I, \psi^*, (D^{\nu_1} \dots D^{\nu_m}\psi)^*, (\nu_1, \dots, \nu_m) \in J)$$

will be invariant under local gauge group transformations. To ensure in addition Lorentz invariance, the matter Lagrangian density must in general be a function of $\psi^*\psi$,

$$(D^{\mu_1}D^{\mu_2} \dots D^{\mu_k}\psi)^*(D_{\mu_1} \dots D_{\mu_k}\psi), k = 1, 2, \dots$$

the Einstein summation convention being implied. On the other hand, the gauge field tensor $F_{\mu\nu}$ has the obvious transformation law

$$-ieF'_{\mu\nu} = [D'_\mu, D'_\nu] = [gD_\mu g^{-1}, gD_\nu g^{-1}] = g[D_\mu, D_\nu]g^{-1} = -ieg.F_{\mu\nu}.g^{-1},$$

ie,

$$F'_{\mu\nu} = g.F_{\mu\nu}.g^{-1}$$

for any local $g = g(x)$ in the gauge group. Thus, if the gauge Lagrangian density is any function of $Tr(F_{\mu\nu}F^{\mu\nu})$, it will be both Lorentz and gauge invariant. Now,

$$g_{ab}F^{\mu\nu a}F^b_{\mu\nu} = Tr(F^{\mu\nu a}F^b_{\mu\nu}\tau_a\tau_b)$$

$$= Tr(F_{\mu\nu}F^{\mu\nu})$$

which obviously gauge invariant, since

$$Tr(F'_{\mu\nu}F^{\mu\nu'}) = Tr(gF_{\mu\nu}g^{-1}.g.F^{\mu\nu}g^{-1}) = Tr(F_{\mu\nu}.F^{\mu\nu})$$

Now we come to the quantum antenna part. First we talk only about electrons, positrons and photons. Let $a(p, \sigma)$, $b(p, \sigma)$, $c(k, s)$ denote respectively the annihilation operators of the electron, positron and photon of respectively momentum p, p, k and spin σ, σ, s . Let $J_{\mu,ext}(x)$ denote the external classical current density field and $J_{\mu,ext}^\alpha(x)$ the non-Abelian current density for the non-Abelian gauge field. The total Lagrangian density for the matter and gauge fields is given by

$$\begin{aligned} L = & K_1 F_{\mu\nu a} F^{\mu\nu a} + K_2 \psi^*(\gamma^0 \otimes I_N)(\gamma^\mu \otimes (I_N(i\partial_\mu + eA_\mu) + eB_\mu^a \tau_a) - mI_{4N})\psi \\ & + K_2 F_{\mu\nu} F^{\mu\nu} + K_3 Tr(G_{\mu\nu} G^{\mu\nu}) \end{aligned}$$

where

$$\begin{aligned} -ieG_{\mu\nu} &= [\partial_\mu - ieB_\mu, \partial_\nu - ieB_\nu] \\ &= -ie(B_{\nu,\mu} - B_{\mu,\nu}) - e^2[B_\mu, B_\nu] \end{aligned}$$

The total current density of the Maxwell-Dirac field is given by

$$J_e^\mu(x) = -e\psi^*(x)(\gamma^0 \gamma^\mu \otimes I_N)\psi(x) \quad (a)$$

while the total current density of the non-Abelian gauge-Dirac field is given by

$$J_a^\mu = -e\psi^*(x)(\gamma^0 \gamma^\mu \otimes \tau_a)\psi(x) \quad (b)$$

The interaction Lagrangians between the former current density and the Maxwell field and between the latter current density and the non-Abelian gauge fields are respectively given by

$$-J_e^\mu A_\mu, -J_a^\mu B_\mu^a$$

If in addition, there are external Maxwell and non-Abelian current classical densities J_{ext}^μ and $J_{ext,a}^\mu$, then the total Lagrangian density of the matter and gauge fields is given by

$$\begin{aligned} L = & K_1 F_{\mu\nu a} F^{\mu\nu a} + K_2 \psi^*(\gamma^0 \otimes I_N)(\gamma^\mu \otimes (I_N(i\partial_\mu + eA_\mu) + eB_\mu^a \tau_a) - mI_{4N})\psi \\ & + K_2 F_{\mu\nu} F^{\mu\nu} + K_3 Tr(G_{\mu\nu} G^{\mu\nu}) \\ & + J_{ext}^\mu A_\mu + J_{ext,a}^\mu B_\mu^a \end{aligned}$$

The field equations derived from the action principle are:

$$F_{,\nu}^{\mu\nu} = -\mu_0(J_e^\mu + J_{ext}^\mu) \quad (c),$$

$$(\partial_\nu + eC(cba)A_\nu^b)G^{\mu\nu a} = J_{ext,c}^\mu + J_c^\mu - \dots \quad (d)$$

which can be better expressed in covariant notation as

$$[\nabla_\nu, G^{\mu\nu}] = J_{ext}^\mu + J^\mu$$

where

$$J_{ext}^\mu = J_{ext,a}^\mu \tau^a, J^\mu = J_a^\mu \tau^a$$

where the Roman indices a, b, c are lowered using the metric ((g_{ab})) and raised using the inverse of this metric ((g^{ab})) and the Greek indices μ, ν are lowered and raised using the Minkowski metric ($(\eta_{\mu\nu})$) and its inverse which is itself. The idea is to first solve for the free Dirac field in the absence of the gauge potentials A_μ, B_μ^a and express this free Dirac wave field in terms of electron annihilation operators $a(p, \sigma)$ and positron creation operators $b(p, \sigma)^*$:

$$\psi(x) = \int (u(p, \sigma)a(p, \sigma)\exp(-ip.x) + v(p, \sigma)b(p, \sigma)^*\exp(ip.x))d^3p$$

Then we use this wave function to calculate the internal currents $J_{e\mu}, J_\mu^a$ in terms of the electron and positron creation and annihilation operators based on the formulas (a) and (b). Then, we solve perturbatively the Maxwell and non-Abelian gauge field equations (c) and (d) expressing thereby the gauge fields A_μ and B_μ^a in terms of the external current sources and the electron-positron operators. Apart from this, we have a free photon field expressed in terms of the photon creation and annihilation operators $c(k, s)^*, c(k, s)$. This satisfies (c) with the rhs set to zero and can be added to the above particular solution. The free photon field is given by a superposition of plane waves:

$$A_\mu^{free} = \int (c(k, s)e_\mu(k, s)\exp(-ik.x) + c(k, s)^*\bar{e}_\mu(k, s)\exp(ik.x))d^3k$$

The resulting gauge fields are therefore expressible as functions of the operator fields in momentum space

$$\{a(p, \sigma), b(p, \sigma), c(k, s), a(p, \sigma)^*, b(p, \sigma)^*, c(k, s)^*\}$$

The free Gibbs state for the photons, electrons and positrons is given by

$$\rho = \exp(-\beta H)/Z(\beta)$$

where

$$Z(\beta) = \text{Tr}(\exp(-\beta H))$$

and

$$H_{ph} + H_{ep},$$

$$H_{ph} = \int |K|c(k, s)^*c(k, s)d^3K, H_{ep} = \int E(P)(a(p, \sigma)^*a(p, \sigma) + b(p, \sigma)b(p, \sigma)^*)d^3P$$

where

$$k = (|K|, K), p = (E(P), P), E(P) = \sqrt{P^2 + m^2}$$

Note the commutation and anticommutation relations

$$\begin{aligned}[c(k, s), c(k', s')^*] &= \delta^3(k - k')\delta_{s, s'}, [c(k, s), c(k', s')] = 0 \\ [a(p, \sigma), a(p', \sigma')^*]_+ &= \delta^3(p - p')\delta_{\sigma, \sigma'}, \\ [b(p, \sigma), b(p', \sigma')^*]_+ &= \delta^3(p - p')\delta_{\sigma, \sigma'},\end{aligned}$$

and all the other anticommutators amongst a, b, a^*, b^* vanish. Further,

$$\begin{aligned}[a(p, \sigma), c(k, s)] &= 0, [a(p, \sigma)^*, c(k, s)] = 0 \\ [b(p, \sigma), c(k, s)] &= 0, [b(p, \sigma)^*, c(k, s)] = 0\end{aligned}$$

Having solved for the gauge fields in term of the external sources and the Dirac currents, we can in practice determine the external current sources so that

$$C_1 \int Tr(\rho(A_\mu(x) - A_\mu^d(x))^2 w^\mu(x) d^4x + C_2 \int Tr(\rho(B_\mu^a(x) - B_\mu^{a,d}(x))^2 q_a^\mu(x) d^4x$$

is a minimum where $w^\mu(x)$ and $q_a^\mu(x)$ are positive weighting functions and A_μ^d, B_μ^a are the desired potentials, which may be classical. In this way, we can control the external current sources so that the electromagnetic field and the Yang-Mills non-Abelian gauge fields are as close as possible to given values of these when the state of the electrons, positrons and photons is prescribed. It should be noted that we do not have unlike the electromagnetic field, a free Yang-Mills gauge field because the free Yang-Mills gauge field equations are non-linear unlike the free em field equations. Even if we were to find a particular solution to the free Yang-Mills gauge field, we cannot superpose it onto the solution with current sources to get another solution because of the nonlinearity of the Yang-Mills field equations. We can at the most hope to perturbatively solve the Yang-Mills gauge field with current sources treating the nonlinear terms in the field equations as small perturbations. The zeroth order term in the Yang-Mills field equations is the linear wave equation with sources like the em field and we can introduce creation and annihilation operators as coefficients in the solution to the zeroth order field equations. The higher order perturbation terms in the Yang-Mills gauge field equations will then be expressed recursively in terms of the lower order terms and thus we can in such a situation, solve for the Yang-Mills potentials in terms of the creation and annihilation operators of the linear component of the field equations and the creation and annihilation operators of the Dirac field of electrons and positrons that enter into the source currents. In this way, we would get a complete quantum field operator theoretic picture of the non-Abelian gauge fields.

6.8 The electroweak theory

The gauge boson Lagrangian density is

$$(1/2)(A_{\nu, \mu} - A_{\mu, \nu} - \alpha A_\mu \times A_\nu)^2 + (1/2)(B_{\nu, \mu} - B_{\mu, \nu})^2$$

where $A_\mu = (A_\mu^1, A_\mu^{2,3}) \in \mathbb{R}^3$ and $B_\mu \in \mathbb{R}$. The matter field Lagrangian density after taking into account interactions with the gauge Bosons A_μ, B_μ is given by

$$\bar{l}_e(\gamma^\mu(i\partial_\mu + e\tau.A_\mu + eyB_\mu)l_e$$

Here, $l_e = [e^T, \nu_e^T]^T$ is the electronic Lepton field. e is the electron field while ν_e is the electron type neutrino field. Apart from the sum of these two Lagrangians which respect the $SU(2) \otimes U(1)$ gauge symmetry, we have a scalar doublet $\phi = [\phi_1, \phi_2]^T$ which interact with the Leptonic field and the electronic field causing symmetry breaking. This interaction Lagrangian density is given by

$$\bar{l}_{eL}G\phi.e_R$$

where G is a constant matrix,

$$l_{eL} = (1 + \gamma_5/2)l_e, e_R = (1 - \gamma_5)e/2$$

Clearly this interaction term is quadratic in the electronic field e and hence this symmetry breaking term gives mass to the electrons. Finally, there is the Lagrangian density of the scalar field ϕ which respects the $SU(2) \times U(1)$ symmetry. It is given by

$$(1/2)|(i\partial_\mu + e\tau.A_\mu + eyB_\mu)\phi|^2$$

Note that y is a constant 2×2 Hermitian matrix while $\tau = (\tau_a, a = 1, 2, 3)$ is a triplet of 2×2 Hermitian matrices. In the ground state of ϕ say $\phi_0 = [0, v^T]^T$ three of the gauge bosons acquire mass while one of them, the photon does not acquire mass. This can be seen by considering ther term

$$\begin{aligned} |(\tau.A_\mu + yB_\mu)\phi_0|^2 &= \phi_0^*\tau_a\tau_b\phi_0 A_\mu^a A_\mu^b + \phi_0^*y^2\phi_0 B_\mu B^\mu \\ &\quad + 2Re(\phi_0^*y\tau_a\phi_0)B_\mu A_\mu^a \end{aligned}$$

We may choose the matrix y such that $Re(\phi_0^*y\tau_a\phi_0) = 0, a = 1, 2, 3, \phi_0^*y^2\phi_0 = 0$ and then it will follow that the gauge field B_μ has no mass and hence can be interpreted as the electromagnetic four potential. Even if this is not possible, we can choose a set of four independent linear combinations of $A_\mu^a, a = 1, 2, 3, B_\mu$ such that in one of the linear combinations, the square term coefficient vanishes and hence this combination corresponds to a massless field, namely the photon field. Finally, we can show that the symmetry breaking term gives mass to the electrons. Indeed, in the ground state ϕ_0 of ϕ , this term becomes

$$\bar{l}_{eL}G\phi_0.e_R = v[e^*, \nu_e^*]\gamma^0(1 + \gamma_5)G \begin{pmatrix} 0 \\ (1 - \gamma_5)e/2 \end{pmatrix}$$

which contains a mass term e^*e giving the electron it mass. In summary, we have an electroweak Lagrangian density that respects the $SU(2) \times U(1)$ symmetry for matter fields comprising electrons, neutrinos and scalar particles along with gauge boson fields. When symmetry breaking occurs with the scalar field in the ground state, then it gives mass to the electrons and not to the neutrinos and in addition, it gives mass to three of the gauge bosons and does not give mass to one of the gauge bosons, namely the photon.

6.9 Wavelet based system parameter estimation

Consider a mechanical system with damping and random forces:

$$x''(t) + \gamma x'(t) + U'(x(t), \theta) = \sigma dB(t)/dt$$

θ is the parameter vector to be estimated and $B(\cdot)$ is Brownian motion. Suppose we assume

$$U(x, \theta) = U_0(x) + \sum_{k=1}^r U_k(x)\theta_k$$

Take noisy measurements on $x(t)$, of the form

$$dz(t) = h(x(t))dt + dV(t)$$

We may assume $x(t) \in \mathbb{R}^n$, $\gamma \in \mathbb{R}^{n \times n}$, $U : \mathbb{R}^n \times \mathbb{R}^p \rightarrow R$, $\sigma \in \mathbb{R}^{n \times d}$, $B(t) \in \mathbb{R}^d$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^q$ and $B(t), V(t)$ are respectively independent \mathbb{R}^d and \mathbb{R}^q valued standard Brownian motion processes. We can use the EKF to estimate the state $x(t)$ and the parameter vector θ even if the latter is slowly time varying with known noisy dynamics. However, the wavelet based method involves directly measuring $x(t)$, calculating its wavelet coefficients as well as those of $x'(t)$ and $x''(t)$:

$$a[n] = \int x(t)\phi_n(t)dt, b[n] = \int x'(t)\phi_n(t)dt, c[n] = \int x''(t)\phi_n(t)dt$$

and that of $U'_k(x(t))$, $k = 0, 1, \dots, d$:

$$e_k[n] = \int U'_k(x(t))\phi_n(t)dt$$

so that the above equation becomes

$$c[n] + \gamma.b[n] + e_0[n] + \sum_{k=1}^r \theta_k e_k[n] = \sigma w[n]$$

or equivalently,

$$c[n] + \gamma.b[n] + e_0(n) + E[n]\theta = \sigma w[n], n = 1, 2, \dots$$

where

$$E[n] = [e_1[n], \dots, e_r[n]]$$

where

$$w[n] = \int B'(t)\phi_n(t)dt = \int \phi_n(t)dB(t)$$

Note that the $w[n], n = 1, 2, \dots$ are iid standard normal vectors since the ϕ'_n 's form an orthonormal set. Thus, the mle of θ based on wavelet coefficient measurements assuming σ is a scalar is given by

$$\hat{\theta}_{ML} = \operatorname{argmin}_{\theta} \sum_n \| c[n] + \gamma.b[n] + e_0[n] + E[n]\theta \| ^2$$

$$= -\left(\sum_n E[n]^T E[n]\right)^{-1} \left(\sum_n E[n]^T (c[n] + \gamma.b[n] + e_0[n])\right)$$

and we find that

$$\begin{aligned}\hat{\theta}_{ML} - \theta &= \\ -\sigma \left(\sum_n E[n]^T E[n]\right)^{-1} \left(\sum_n E[n]^T w[n]\right)\end{aligned}$$

The parameter estimation error covariance is thus (note that this is an unbiased estimator)

$$\begin{aligned}Cov(\hat{\theta}_{ML} - \theta) &= Cov(\hat{\theta}_{ML}) = \\ \sigma^2 \left(\sum_n E[n]^T E[n]\right)^{-1}\end{aligned}$$

6.10 Applying the EKF to nonlinear circuits involving diodes and transistors

If in the Ebers-Moll model of a transistor, the current is expanded upto quadratic orders in the base-emitter and base-collector voltages, and the input voltage is an OU process, then by appropriately defining the state vector in terms of nodal voltages, we can express the circuit dynamics (obtained by applying the KCL and KVL) in the form

$$dX(t) = (A_1 X(t) + A_2(X(t) \otimes X(t)) + A_3 u(t))dt + A_4 dB(t)$$

with the measurement being taken on one or more of the state variables:

$$dZ(t) = H X(t)dt + \sigma_V dV(t)$$

In the discrete domain, ie, after time discretization, these stochastic differerential equations become stochastic difference equations of the form

$$X[n+1] = C_1 X[n] + C_2(X[n] \otimes X[n]) + C_3 u[n] + C_4 W[n+1]$$

$$Z[n] = H X[n] + \sigma_V V[n]$$

Consider the general discrete time case:

$$X[n+1] = f_n(X[n]) + g_n(X[n])W[n+1],$$

$$Z[n] = h_n(X[n]) + V[n+1]$$

where $W[.]$ and $P_V^{-1/2}.V[.]$ are iid standard vector normal processes. Let $F_n = f'_n(\hat{X}[n|n])$, $H_{n+1} = h_{n+1}(\hat{X}[n+1|n])$, $G_n = g_n(\hat{X}[n|n])$ and

$$e[n|n] = X[n] - \hat{X}[n|n], e[n+1|n] = X[n+1] - \hat{X}[n+1|n],$$

$P[n|n] = Cov(e[n|n]|Z_n) = cov(X[n]|Z_n)$, $P(n+1|n) = cov(e[n+1|n]|Z_n) = cov(X[n+1]|Z_n)$
where

$$Z_n = \{Z[k] : k \leq n\}$$

We have approximately,

$$\begin{aligned} \hat{X}[n+1|n] &= \mathbb{E}(f_n(X[n])|Z_n) \approx f_n(\hat{X}[n|n]) + (1/2)f_n''(\hat{X}[n|n])Vec(P[n|n]) \\ e[n+1|n] &\approx X[n+1] - f_n(\hat{X}[n|n]) - (1/2)f_n''(\hat{X}[n|n])Vec(P[n|n]) \\ &= f_n(X[n]) - f_n(\hat{X}[n|n]) + G_n W[n+1] + G'_n W[n+1] - (1/2)f_n''(\hat{X}[n|n])Vec(P[n|n]) \\ &\approx f'_n(\hat{X}[n|n])e[n|n] + G_n W[n+1] + g'_n(\hat{X}[n|n])(e[n|n] \otimes W[n+1]) \\ &\quad + (1/2)f_n''(\hat{X}[n|n])(e[n|n] \otimes e[n|n]) - (1/2)f_n''(\hat{X}[n|n])Vec(P[n|n]) \end{aligned}$$

which gives on taking covariance

$$\begin{aligned} P[n+1|n] &= F_n P[n|n] F_n^T + G_n G_n^T + G'_n (P[n|n] \otimes I) G_n'^T \\ &\quad + (1/2) F'_n Cov(e[n|n] \otimes e[n|n]|Z_n) F_n'^T \end{aligned}$$

Also,

$$\begin{aligned} \hat{X}[n+1|n+1] &= \hat{X}[n+1|n] + K[n+1](Z[n+1] - h_{n+1}(\hat{X}[n+1|n])) \\ &\quad - (1/2)h''_{n+1}(\hat{X}[n+1|n])Vec(P[n+1|n]) \\ &= \hat{X}[n+1|n] + K[n+1](Z[n+1] - H_{n+1} - (1/2)H''_{n+1}Vec(P[n+1|n])) \end{aligned}$$

The Kalman gain is determined by minimizing

$$\begin{aligned} Tr(Cov(e[n+1|n+1]|Z_n)) &= \\ Tr(Cov(e[n+1|n] - K[n+1](h_{n+1}(X[n+1]) - H_{n+1} - H''_n Vec(P[n+1|n]))|Z_n)) &= \\ Tr(Cov(e[n+1|n] - K[n+1](H'_{n+1}e[n+1|n] + V[n+1] + (1/2)H''_{n+1}(e[n+1|n] \otimes e[n+1|n]) \\ &\quad - Vec(P[n+1|n]))|Z_n)) \\ \approx Tr[(I - K[n+1]H'_{n+1})P[n+1|n](I - K[n+1]H'_{n+1})^T] &+ Tr(K[n+1]P_V K[n+1]^T) + \\ + Tr(K[n+1]H''_{n+1}Cov(e[n+1|n] \otimes e[n+1|n]|Z_n)H''_{n+1}^T K[n+1]^T) & \end{aligned}$$

provided that we neglect cubic moments of $e[n+1|n]$ conditioned on Z_n . Finally,

$$\begin{aligned} P[n+1|n+1] &= Cov(e[n+1|n+1]) \\ &= Cov(X[n+1] - \hat{X}[n+1|n] - K[n+1](Z[n+1] - H_{n+1} - (1/2)H''_{n+1}Vec(P[n+1|n])) \\ &= Cov(e[n+1|n] - K[n+1](H'_{n+1}e[n+1|n] + V[n+1] + (1/2)H''_{n+1}(e[n+1|n] \otimes e[n+1|n]) \\ &\quad - Vec(P[n+1|n]))) \\ &= (I - K[n+1]H'_{n+1})P[n+1|n].(I - K[n+1]H'_{n+1})^T + K[n+1]P_V K[n+1]^T + \\ &\quad (1/2)H''_{n+1}.Cov(e[n+1|n] \otimes e[n+1|n]).H''_{n+1}^T \end{aligned}$$

6.11 An introduction to classical and quantum error detecting and correcting codes

In the classical coding theory context, we have a finite Abelian group A with $\mu(A) = N$. Let C be a subgroup of A^n with $\mu(C) = N$. Suppose that for any two distinct $c_1, c_2 \in C$, we have $w(c_1^{-1}c_2) \geq d$ where $w(g)$ is the number of non-zero entries in the sequence $g \in A^n$. This is equivalent to saying that $w(c) \geq d$ for all $c \in C - \{e\}$. Suppose further that for some $c \in C$, $w(c) = d$. Then, we say that C is a group code with minimum distance d . Consider now the coset group $X = A^n/C$. We can choose a finite set $\{g_1 = e, g_2, \dots, g_r\} \subset A^n$ such that $g_k C \cap g_m C = \phi$ for all $k \neq m$ and further, $A^n = \bigcup_{k=1}^r g_k C$. Let $t = [(d-1)/2]$ and let $c_1, c_2 \in C$ be arbitrary distinct elements. Suppose $s \in A^n$ is such that $w(s) \leq t$. Then let $c = c_2^{-1}c_1$. We have that $w(c_2^{-1}c_1 s) = w(cs) \geq d - t = d - [(d-1)/2] \geq d - (d-1)/2 = (d+1)/2 > t$. This means that the "Hamming distance" between $c_1 s$ and c_2 is greater than t . It follows that C is a t -error correcting code. It should also be noted that C is a $d-1$ -error detecting code. This means that if $c \in C$ and $s \in A^n$ with $w(s) \leq d-1$, then cs cannot be in C unless $w(s) = 0$, ie, $s = e^n$ in which case $cs = c$ implying thereby $d-1$ error correction. On the other hand, if $s_1, s_2 \in A^n$ with $w(s_k) \leq t, k = 1, 2$, then for any two distinct $c_1, c_2 \in C$, we have $w((c_2 s_2)^{-1}(c_1 s_1)) > 0$ implying thereby t -error correction.

In the quantum context, suppose \mathcal{H} is a Hilbert space. Then, let \mathcal{N}_t denote the subspace of $\mathcal{B}(\mathcal{H}^{\otimes n})$ spanned by all elements of the form $\bigotimes_{k=1}^n X_k$ where $X_k \in \mathcal{B}(\mathcal{H})$ and atmost t elements from the sequence X_1, X_2, \dots, X_n differ from the identity operator. Then, in the context of the Knill-Laflamme theorem, if C is a subspace of $\mathcal{H}^{\otimes n}$ that corrects N_t , then we say that C is a t -error correcting quantum code. This means that if $N_1, N_2 \in \mathcal{N}_t$, we have that

$$PN_2^*N_1P = \lambda(N_2^*N_1)P$$

where $\lambda(N_2^*N_1)$ is a complex scalar and P is the orthogonal projection onto C .

Some examples of error correcting quantum codes: Let G be a finite group acting on a set X and let P be a spectral measure on a σ -field \mathcal{F} of X -subsets assuming values in $\mathcal{P}(\mathcal{H})$, the lattice of orthogonal projections in \mathcal{H} . Assume that U is a unitary representation of G and that (G, U, P) is an imprimitivity system in the sense of Wigner and Mackey, ie,

$$U(g)P(E)U(g)^* = P(g.E), g \in G, E \in \mathcal{F}$$

We choose $E \in \mathcal{F}$ and a subset H of G and wish to determine when the quantum code $P(E)$ can detect and correct the linear manifold of noise operators spaned by $\{U(g) : g \in H\}$. For error detection, we must have

$$P(E)U(g)P(E) = \lambda(g)P(E), g \in H$$

This is equivalent to requiring that

$$P(E)U(g)P(E)U(g)^* = \lambda(g)P(E)U(g)^*$$

or equivalently,

$$P(E \cap gE) = \lambda(g)P(E)U(g)^*, g \in H$$

It should be noted that necessary condition for the above error detection is given by

$$P(E)U(g)^*P(E).P(E)U(g)P(E) = |\lambda(g)|^2 P(E)$$

or equivalently,

$$P(E \cap g^{-1}E) = |\lambda(g)|^2 P(E), g \in H$$

This implies that for each $g \in H$, either $E \cap g^{-1}E = \phi$ or else $g^{-1}E = E$. This result can be interpreted classically that the code E detects the noise H only if each $g \in H$ either takes the code E outside it or else leaves the code E invariant. Note that if $E \cap g^{-1}E = \phi$, then $\lambda(g) = 0$ while if $g^{-1}E = E$, then $|\lambda(g)| = 1$.

6.12 Orthogonal Latin squares and coding theory

Let $\lambda(x, y)$ be a square array with $1 \leq x, y \leq N$. We say that it is a Latin square if firstly, each row and each column of this array is a permutation of $1, 2, \dots, N$ and $\lambda(x, y) = \lambda(x, y')$ implies $y = y'$ and likewise, $\lambda(x, y) = \lambda(x', y)$ implies $x = x'$. In other words, any two rows of the array λ are permutations of each other with no two entries in their same column being equal and likewise, any two columns of the array λ are permutations of each other with no two entries in their same row being equal. Two Latin squares λ and μ are said to be orthogonal if $(\lambda(x, y), \mu(x, y)) = (\lambda(x', y'), \mu(x', y'))$ implies that $x = x'$ and $y = y'$.

An example of orthogonal Latin squares: Let A be a finite Abelian group and $\tau : A \rightarrow A$ a group isomorphism. Define $\lambda(x, y) = \tau(x)y, x, y \in A$. Then λ is a Latin square defined as a map $\lambda : A \times A \rightarrow A$. Indeed, $\tau(x)y = \tau(x)y'$ implies $y = y'$ and $\tau(x)y = \tau(x')y$ implies $\tau(x) = \tau(x')$ which implies $x = x'$. Now let $\tau_k, k = 1, 2$ be two isomorphisms of A such that $\tau_2^{-1}o\tau_1$ has no fixed point except e . Then $\lambda_k(x, y) = \tau_k(x)y, j = 1, 2$ define two mutually orthogonal Latin squares. Indeed, $(\tau_1(x)y, \tau_2(x)y) = (\tau_1(x')y', \tau_2(x')y')$ implies $\tau_1(x)y = \tau_1(x')y', \tau_2(x)y = \tau_2(x')y'$ which implies $\tau_k(x'^{-1}x) = y'y^{-1}, k = 1, 2$. This in turn implies $\tau_1(x'^{-1}x) = \tau_2(x'^{-1}x)$ which implies $x'^{-1}x$ is a fixed point of $\tau_2^{-1}\tau_1$ and hence by hypothesis, $x' = x$ and therefore also $y' = y$, thus establishing the orthogonality of the Latin squares $\lambda_k, k = 1, 2$. The converse is also easily seen to be true by just reversing the argument: If τ_1, τ_2 are two isomorphisms of A such that the corresponding Latin squares constructed above orthogonal, then e is the only fixed point of the isomorphism $\tau_2^{-1}o\tau_1$.

Now let $\lambda_j, j = 1, 2, \dots, n$ be mutually orthogonal Latin squares for the same set E . Then, consider the $n + 2$ -length sequences

$$c(x, y) = (x, y, \lambda_1(x, y), \dots, \lambda_n(x, y),)x, y \in E$$

The total number of these sequences is $\mu(E)^2$. Let us compute the minimum Hamming distance of the code C whose elements are these sequences. Suppose $x = x', y \neq y'$. Then, $\lambda_k(x, y) \neq \lambda_k(x, y') = \lambda_k(x', y')$, $k = 1, 2, \dots, n$, by the definition of a Latin square. Likewise, if $y = y'$ but $x \neq x'$. In both of these cases, we therefore see that the distance between the codewords $c(x, y)$ and $c(x', y')$ is $n + 1$. On the other hand, if $x \neq x', y \neq y'$, then by definition of orthogonal Latin squares, it follows that by considering the ordered pair (λ_k, λ_j) that if $\lambda_k(x, y) = \lambda_k(x', y')$ for some k , then $\lambda_j(x, y) \neq \lambda_j(x', y')$ for all $j \neq k$. Therefore, in this case, the Hamming distance between the codewords $c(x, y)$ and $c(x', y')$ is $2 + n - 1 = n + 1$. This proves that the minimum Hamming distance of C is $n + 1$.

6.13 Cyclic codes

Let p be a prime number and n a positive integer. It is known that if $q = p^n$, then there exists a field $\mathbb{F} = \mathbb{F}_q$ with q elements. Let $\mathbb{F}_q[x]$ denote the ring of polynomials in the variable x with coefficients from \mathbb{F}_q . It is known that there exists at least one primitive element $\alpha \in \mathbb{F}_q$, ie, $\mathbb{F}_q = \{\alpha^k, k = 0, 1, \dots, q-1\}$. For any $\beta \in \mathbb{F}_q$, denote by M_β the minimal polynomial of β , ie, M_β is the unique monic polynomial with coefficients from \mathbb{F}_q of least degree for which $M_\beta(\beta) = 0$. Clearly, M_β is irreducible, since if $M_\beta = PQ$ where $P, Q \in \mathbb{F}_q[x]$ are such that $\deg P, \deg Q > 1$, then either P or Q is a constant polynomial, for otherwise P or Q will be a polynomial of degree smaller than M_β that annihilates β which is a contradiction to the definition of M_β . Now suppose f is any irreducible monic polynomial in \mathbb{F}_q . Let $I_f = f(x)\mathbb{F}_q[x]$, ie, the I_f is the ideal in $\mathbb{F}_q[x]$ generated by f . Note that f is indeed the generator of I_f , for if g is the generator of I_f , then from the division algorithm, g divides f which is impossible unless g is either a constant or f itself since f is assumed to be irreducible. Now consider the ring $F_f = \mathbb{F}_q[x]/I_f$. We claim that F_f is a field with q^d elements. Indeed, F_f is an Abelian group under addition. Further, F_f is closed under multiplication all by definition of addition and multiplication operations in quotient ring. It remains to prove that each non-zero element in F_f has a multiplicative inverse. For assume that $\deg f = d$. Then, since f is irreducible, a basis for F_f is $[1], [x], \dots, [x^{d-1}]$ where $[h] = h + I_f$ for any $h \in \mathbb{F}_q[x]$. Hence, any $X \in F_f$ is of the form $[a_0 + a_1x + \dots + a_{d-1}x^{d-1}]$ with $a_0, \dots, a_{d-1} \in \mathbb{F}_q$. Hence, the cardinality of F_f is q^d . Suppose $[a], [b] \in F_f, 0 < \deg a, \deg b < d$ are such that $[ab] = [0]$. Then, f divides ab and since f is irreducible, it follows that f divides either a or b which is impossible since $\deg a, \deg b < d$. Thus, there are no zero divisors in F_f , ie, F_f is a division ring. Further, if $a \in \mathbb{F}_q[x], \deg a < d$ is non zero, then $[a], [xa], \dots, [x^{d-1}a]$ form a basis for F_f since if some linear combination of these is zero, say $\sum_{k=0}^{d-1} c_k[x^k a] = 0$, then $a(x) \sum_{k=0}^{d-1} c_k x^k$ is divisible by f which is impossible since f is irreducible, unless all the c'_k s are zero. Thus the elements $[pa]$ as p runs over all polynomials in $\mathbb{F}_q[x]$ having degree smaller than d are all distinct and hence $[p][a] = [pa] = [1]$ for some such

p and further such a $[p]$ is unique, for if $[q][a] = [1]$ for another such q , then $[p - q][a] = [(p - q)a]0$ which implies that $(p - q)a$ is divisible by f which is impossible since $\deg p, \deg q, \deg a < d$ and f is irreducible. This proves that F_f is a field.

Remark: When p is a prime, we know that the field \mathbb{F}_p exists. There is a theorem that when p is a prime and n any positive integer, then there exists an irreducible polynomial f of degree $n - 1$ over \mathbb{F}_p . Thus, the above result implies that F_f is a field with $q = p^n$ elements. In other words, for any prime p and positive integer n , there exists a field with $q = p^n$ elements.

6.14 Yang-Mills field quantization methods

The general theory of quantizing gauge invariant actions.

Let $I[\phi]$ be the action functional of matter and gauge fields. Let G denote the gauge group and assume that $\exp(I[\phi])D\phi$ is G -invariant, ie, for all $\Lambda \in G$, we have

$$\exp(I[\phi_\Lambda])D\phi_\Lambda = \exp(I[\phi])D\phi$$

A special case of this is when the action $I[\phi]$ and the path measure $D\phi$ are individually G -invariant. Consider the path integral

$$S = \int G(\phi)B(f[\phi])F(f[\phi])D\phi, G(\phi) = \exp(iI[\phi])$$

Here, f is a gauge fixing functional. Replacing ϕ by ϕ_Λ for some $\Lambda \in G$, we get

$$\begin{aligned} S &= \int G(\phi_\Lambda)B(f[\phi_\Lambda])F(f[\phi_\Lambda])D\phi_\Lambda \\ &= \int G(\phi)B(f[\phi_\Lambda])F(f[\phi_\Lambda])D\phi \end{aligned}$$

Let $\rho(\Lambda)d\Lambda$ denote the Haar measure on G . Then, it is a well known theorem in measure theory that

$$1/\rho(\Lambda) = \det(d\Lambda o \lambda/d\lambda)|_{\lambda=e}$$

where e is the identity transformation of the gauge group. Now we define

$$F(f[\phi]) = \det(Df[\phi_\lambda]/D\lambda)|_{\lambda=e}$$

Then, we have

$$\begin{aligned} F(f[\phi_\Lambda]) &= \det(Df[\phi_\Lambda]/D\Lambda)\det(D\Lambda o \lambda/D\lambda)|_{\lambda=e} \\ &= \det(Df[\phi_\Lambda]/D\Lambda)\rho(\Lambda)^{-1} \end{aligned}$$

and hence,

$$\int B(f[\phi_\Lambda])F(f[\phi_\Lambda])\rho(\Lambda)D\Lambda$$

$$\begin{aligned}
&= \int B(f[\phi_\Lambda]) \det(Df[\phi_\Lambda]/D\Lambda) D\Lambda \\
&= \int B[f] Df
\end{aligned}$$

provided that we assume that the map $\Lambda \rightarrow f[\phi_\Lambda]$ is surjective, ie it maps G onto the space of all gauge fixing functionals f for any fixed set of matter and gauge fields ϕ . Clearly, this implies that

$$S = (D/C) \int G[\phi] D\phi$$

where

$$D = \int B[f] Df, C = \int_G \rho(\Lambda) D\Lambda$$

This argument shows that the path integral for calculating propagators and scattering matrix amplitudes is for all practical purposes, independent of the gauge fixing functional B provided that we take into account the correct Jacobian determinant corresponding to this gauge fixing functional.

As an application of this circle of ideas, consider the problem of quantizing the Yang-Mills non-Abelian matter and gauge fields. The covariant derivative is

$$\nabla_\mu = \partial_\mu + ieA_\mu^\alpha \tau_\alpha$$

The covariant field tensor is

$$\begin{aligned}
ieF_{\mu\nu} &= [\nabla_\mu, \nabla_\nu] = ie(A_{\nu,\mu} - A_{\mu,\nu}) - e^2[A_\mu, A_\nu] \\
&= ieF_{\mu\nu}^\alpha \tau_\alpha
\end{aligned}$$

where

$$F_{\mu\nu}^\alpha = (A_{\nu,\mu}^\alpha - A_{\mu,\nu}^\alpha) - eC(\beta\gamma\alpha)A_\mu^\beta A_\nu^\gamma$$

where the τ_α are Hermitian generators satisfying the commutation relations

$$[\tau_\alpha, \tau_\beta] = iC(\alpha\beta\gamma)\tau_\gamma$$

The Lagrangian density of the gauge field is

$$L = F_{\mu\nu}^\alpha F^{\mu\nu\alpha}$$

provided we assume that

$$Tr(\tau_\alpha \tau_\beta) = K.\delta_{\alpha\beta}$$

which we may always do by the Gram-Schmidt process of orthonormalization. We expand this Lagrangian density as

$$\begin{aligned}
L &= 2\eta_{\nu\rho}\eta_{\mu\sigma}[A_{\nu,\mu}^\alpha A_{\rho,\sigma}^\alpha - A_{\nu,\mu}^\alpha A_{\sigma,\rho}^\alpha] \\
&\quad - 2eC(\beta\gamma\alpha)\eta_{\nu\rho}\eta_{\mu\sigma}A_\mu^\beta A_\nu^\gamma(A_{\rho,\sigma}^\alpha - A_{\sigma,\rho}^\alpha)
\end{aligned}$$

$$+e^2 C(\beta\gamma\alpha)C(\beta'\gamma'\alpha')\eta_{\mu\rho}\eta_{\nu\sigma}A_\mu^\beta A_\nu^\gamma A_\rho^{\beta'} A_\sigma^{\gamma'}$$

This Lagrangian density contains second, third and fourth degree terms unlike the case of emft where only second degree terms appear. Therefore, computing the propagator in closed form is impossible but we can obtain the first few extra terms using perturbation theory treating the cubic and fourth degree terms as perturbations and using the theory of Gaussian integrals. We take our gauge fixing functional as

$$f[A] = a^2 \int \eta_{\mu\nu} A_\mu^\alpha A_\nu^\alpha d^4x$$

This is Lorentz invariant. It appears in the path integral as

$$B(f[A]) = \exp(-f[A]/2)$$

We can express this using standard Gaussian path integrals as

$$B(f[A]) = \int \exp(ia \int A_\mu^\alpha g_\mu^\alpha d^4x) \cdot \exp(-(1/2) \int g_\mu^\alpha g^{\mu\alpha} d^4x) Dg$$

This means that we are adding extra terms to the Yang-Mills Lagrangian density given by

$$L_1[A, g] = a \cdot A_\mu^\alpha g_\mu^\alpha + (i/2) g_\mu^\alpha g^{\mu\alpha}$$

Note that the g 's are auxiliary fields. Likewise, we need to represent the Jacobian functional $F(f[A])$ using path integrals over other auxiliary fields. We recall that

$$F(f[A]) = \det(Df[A_\lambda]/D\lambda)|_{\lambda=0}$$

where λ is an infinitesimal gauge transformation of gauge fields. We take as our infinitesimal local gauge transformation

$$g(x) = I + i\epsilon^a(x)\tau_a$$

and then the gauge transformation of the gauge field is given upto linear orders in ϵ^a as

$$\begin{aligned} A'_\mu &= g A_\mu g^{-1} - (i/e)(g \partial_\mu g^{-1}) \\ &= A_\mu + i\epsilon^b [\tau_b, A_\mu] - (1/e)\epsilon^b_{,\mu} \tau_b \end{aligned}$$

or equivalently, in terms of the structure constants of the gauge group,

$$A_\mu^{a'} = A_\mu^a - C(bca)\epsilon^b A_\mu^c - (1/e)\epsilon^a_{,\mu}$$

This gives the variational derivative

$$\delta A_\mu^a(x)/\delta\epsilon^b(y) = -[C(bca)A_\mu^c(x) - (1/e)\delta_{ab}\partial_\mu]\delta^4(x-y)$$

If we had instead chosen the gauge fixing functional as

$$f^a = f^a[A] = \eta_{\mu\nu} A_{\mu,\nu}^a = A_{;\mu}^{a\mu}$$

and

$$B[f] = B[f[A]] = \exp(-(xi/2) \int f^a f^a d^4x) = \exp(-(\xi/2)f[A])$$

then, we would choose an auxiliary field g^a and represent $B[f]$ as the Gaussian Path integral

$$B[f] = \int \exp\left(\int (if^a g^a - g^a g^a / 2\xi) d^4x\right) Dg$$

This means that the Lagrangian density of the Yang-mills gauge field acquires an extra term

$$f^a g^a + ig^a g^a / 2\xi$$

Now to represent the determinant term in the path integral, we use a set $\omega^a, \bar{\omega}^a$ of Fermionic fields and represent

$$F[f[A]] = \det(Df[A_\lambda]/D\lambda)|_{\lambda=0}$$

Writing ϵ in place of λ , we have

$$A_\mu^{a'} = A_\mu^a - C(bca)\epsilon^b A_\mu^c - (1/e)\epsilon_{,\mu}^a$$

Thus,

$$f^{a'} = A_{,\mu}^{a'} = f^a - C(bca)(\epsilon^b A^{\mu c})_{,\mu} - (1/e)\square\epsilon^a$$

and hence

$$Df[A_\epsilon]/D\epsilon$$

is represented by the kernel

$$\begin{aligned} K_{ad}(x, y) &= [C(bca)A^{\mu c}(x)\delta_{bd}\partial_\mu - (1/e)\delta_{ad}\square]\delta^4(x - y) \\ &= [A^{\mu c}(x)C(dca)\partial_\mu - (1/e)\delta_{ad}\square]\delta^4(x - y) \end{aligned}$$

From the theory of Berezin's Fermionic path integrals, we have

$$\det(K) = \int \exp(-(1/2) \int (\omega^{a*}(x)K_{ad}(x, y)\omega^d(y)d^4x d^4y) D\omega. D\omega^*$$

Further,

$$\begin{aligned} &\int \omega^{a*}(x)K_{ad}(x, y)\omega^d(y)d^4x d^4y = \\ &\int \omega^{a*}(x)\omega^d(y)[A^{\mu c}(x)C(dca)\partial_\mu - (1/e)\delta_{ad}\square]\delta^4(x - y)d^4x d^4y \\ &= \int C(dca)A^{\mu c}(x)\omega^{a*}(x)\omega_{,\mu}^d(x)d^4x \\ &\quad - (1/e) \int (\square\omega^{a*}(x))\omega^a(x)d^4x \\ &= \int C(dca)A_{,\mu}^{\mu c}(x)\omega^{a*}(x)\omega_{,\mu}^d(x)d^4x \end{aligned}$$

$$+(1/e) \int \eta_{\mu\nu} \omega_{,\mu}^{a*}(x) \omega_{,\nu}^a(x) d^4x$$

Thus, the total action for the Yang-Mills gauge field taking auxiliary and ghost fields into account is given by

$$\begin{aligned} S_{tot}(A, g, g^* \omega, \omega^*) = \\ \int [F_{\mu\nu}^\alpha F^{\mu\nu\alpha} + Re(f^{a*} g^a) - Im(g^{a*} g^a)/2\xi + C(dca) A_{,\mu}^{\mu c} Re(\omega^{a*} \omega_{,\mu}^d) \\ +(1/e) \eta_{\mu\nu} \omega_{,\mu}^{a*} \omega_{,\nu}^a] d^4x \end{aligned}$$

where

$$f^a = A_{,\mu}^{\mu a}$$

6.15 The Ginzburg-Landau model for superconductivity

$A_\mu(x)$ is the external em potential. $\psi(x)$ is the wave function of the electron. The Lagrangian density of the em field is $K_1 F_{\mu\nu} F^{\mu\nu}$ where $F_{\mu\nu} = A_{\nu,\mu} - A_{\mu,\nu}$. The Lagrangian of the electron has terms $|(\nabla + ieA/h)\psi|^2 + a|\psi|^2 - b|\psi|^4$. Under a $U(1)$ gauge transformation, $\psi(x) \rightarrow \exp(ie\phi(x)/h)\psi(x) = \psi(x)'$ and $A(x) \rightarrow A(x)'$ so that $|(\nabla + ieA'/h)\psi'|^2 = |(\nabla + ieA/h)\psi|^2$. This implies that

$$ieA'/h + ie\nabla\phi/h = ieA/h$$

or equivalently,

$$A' = A - e\nabla\phi/h$$

Under this gauge transformation, the total Lagrangian density of the em field and the electron field is invariant, ie, we have an unbroken $U(1)$ symmetry. Note that the total Lagrangian density of the em field and the electron field is

$$L = K_1 F_{\mu\nu} F^{\mu\nu} - |(\nabla + ieA/h)\psi|^2 - a|\psi|^2 + b|\psi|^4$$

We now write

$$\psi(x) = \rho(x). \exp(-ie\phi(x)/h)$$

where $\rho(x) \geq 0$ and $\phi(x) \in \mathbb{R}$. This decomposition can be viewed as saying that ρ is the unbroken part while the exponential in ϕ is the broken Goldstone part. This can be understood by noting that

$$(\nabla + ieA/h)\psi = \exp(-i\phi/h)[(\nabla\rho + ieA\rho/h) - ie\nabla\phi/h]$$

so that

$$|\nabla + ieA/h)\psi|^2 = |\nabla\rho|^2 + (e^2\rho^2/h^2)(A - \nabla\phi)^2$$

Thus, the Lagrangian of the em field plus electron field is then in the absence of electric fields,

$$L = (-1/2)(\nabla \times A)^2 - |\nabla \rho|^2 - (e^2 \rho^2/h^2)(A - \nabla \phi)^2 - a\rho^2 + b\rho^4$$

This expression shows that the Goldstone part of the electron field ϕ does not contain any mass term corresponding to the general theory that broken symmetries are associated with the production of massless particles. We also note that the unbroken part (involving only $\rho(x)$ and its derivatives) of the Lagrangian is invariant under $U(1)$. The symmetry is broken only if ρ is non-zero and the mass term associated with the broken symmetry can be obtained by passing over into a minimum energy state when $\rho(x) = <\rho>$ is a constant and A is a pure gauge $\nabla\phi$. Writing $\rho(x) = <\rho> + \delta\rho(x)$, we find that the coefficient of $-(\delta\rho(x))^2/2$ in the Lagrangian density is

$$2a - 12b <\rho>^2$$

and this determines the mass of the unbroken symmetry particles. Now we come to the electromagnetic part. The Lagrangian component $(A - \nabla\phi)^2$ attains its minimum when A is a pure gauge: $A = \nabla\phi$. In this case, we find by integrating once around a loop Γ that bounds a surface S within the superconductor that the magnetic flux through S is given by

$$\Phi_B = \int_S B \cdot dS = \int_{\Gamma} A \cdot dr = \int_{\Gamma} \nabla\phi \cdot dr = \Delta\phi$$

where $\Delta\phi$ is the change in phase of the Goldstone part of the electron field when we go once around the closed loop Γ . But since the wave function must be single valued, we find that

$$ie\Delta\phi/h = 2n\pi$$

or equivalently,

$$\Delta\phi = 2n\pi h/e, n \in \mathbb{Z}$$

In other words, the magnetic flux is quantized:

$$\Phi_B = 2n\pi h/e, n \in \mathbb{Z}$$

This flux is quantized only when the superconductor is in its equilibrium, ie, minimum energy state. The Euler-Lagrange field equation for A corresponding to the Lagrangian density

$$L = (-1/2)(\nabla \times A)^2 - |\nabla \rho|^2 - (e^2 \rho^2/h^2)(A - \nabla \phi)^2 - a\rho^2 + b\rho^4$$

is

$$\nabla \times B = -(e^2 \rho^2/h^2)(A - \nabla \phi), B = \nabla \times A = \nabla \times (A - \nabla \phi)$$

This clearly displays the broken symmetry: If ρ is non zero, we take its ground state value $<\rho>$ and then the Maxwell equation for B contains the vector

potential A which is not gauge invariant. The Euler-Lagrange equation for the unbroken electron field component ρ is

$$\nabla^2 \rho - (e^2 \rho / h^2)(A - \nabla\phi)^2 - a\rho + 4b\rho^3 = 0$$

If $\rho = 0$, these field equations simplify to the gauge invariant Maxwell equation

$$\nabla \times B = 0$$

When $\rho \neq 0$, we consider the ground state solution for ρ , namely the constant $\langle \rho \rangle$. Replacing ρ by this ground-state value, the field equations become after assuming $\operatorname{div}(A - \nabla\phi) = 0$,

$$\nabla^2(A - \nabla\phi) = (e^2 \langle \rho \rangle^2 / h^2)(A - \nabla\phi)$$

We note that the equation for ρ is trivially satisfied when ρ is taken as the constant $\langle \rho \rangle$ provided that we take $A = \nabla\phi$ and

$$a = 4b \langle \rho \rangle^3$$

The above differential equation for $A - \nabla\phi$ shows that the penetration depth of the magnetic field (ie the depth upto which $A - \nabla\phi$ is non-zero is given by

$$(e^2 \langle \rho \rangle^2 / h^2)^{-1/2} = he / \langle \rho \rangle$$

We have thus successfully used the Ginzburg-Landau theory to successfully explain the Meissner effect, ie the near expulsion of a magnetic field from a superconductor and also the phenomenon of flux quantization, ie when the superconductor is in its ground state, $A - \nabla\phi = 0$ which implies quantization of the magnetic flux in integer multiples of $2\pi\hbar/e$. Actually e should be replaced by $2e$ in view of the formation of Cooper pairs more successfully explained by the BCS microscopic theory of superconductivity. The next aspect is to use the Ginzburg-Landau theory of superconductivity to explain the Josephson current as an ac current between two superconductors maintained at constant potential difference. This phenomenon enables us to estimate the Planck's constant accurately in terms of the electronic charge by measuring the frequency of the Josephson current. The Hamiltonian density H in analogy with the above formalism but containing apart from the magnetic vector potential, also the scalar electric potential and taking time variations into account contains a term

$$L(A_\mu - \partial_\mu\phi)$$

which attains its minimum when $A_\mu = \partial_\mu\phi = \phi_{,\mu}$, ie, when the em four potential is a pure gauge. it follows that

$$-\frac{\partial L}{\partial A_\mu} = J_\mu = \partial L / \partial \phi_{,\mu}$$

is the four current. In particular, if H is the Hamiltonian, then the charge s

$$Q = \frac{\partial H}{\partial V} = -\partial H / \partial \phi_{,0}$$

where V is the electric scalar potential. In other words, the charge Q and potential form a dual canonical pair. Suppose we had two superconductors maintained at a constant potential difference of ΔV . Then, the above equation gives for the phase difference $\Delta\phi$ of the Cooper pair field ϕ between the two superconductors by duality,

$$\Delta\phi_{,0} = \partial H / \partial Q = \Delta V$$

so that

$$\Delta\phi(t) = t\Delta V$$

$\Delta\phi(t)$ is the phase difference in the spatial domain, ie, it can be regarded as being proportional to the spatial gradient of the Cooper phase between the two superconductors. The current between the superconductors is given by

$$\begin{aligned} I(t) &= S \cdot \partial L(A - \Delta\phi(t)/d) / \partial A|_{A=0} = S \cdot L'(-\Delta\phi(t)/d) \\ &= -S L'(-t\Delta V/d) \end{aligned}$$

where S is the cross sectional area of the region between the two superconductors and d is the separation between the superconductors. Now, L and hence L' is a periodic function of its argument with a period of $T = 2\pi h/e$. (Instead of regarding $L(A - \nabla\phi)$ as a function of $A - \nabla\phi$, we can regard it as a function of $A - \Delta\phi/d$ and when $\Delta\phi$ changes by $2\pi h/e$, L should not change. Thus the asserted claim. It follows that the current $I(t)$ is periodic in t with a period of

$$\tau = Td/\Delta V = 2\pi h/e\Delta V$$

This is the famous Josephson current formula. By measuring the frequency

$$\omega = 2\pi/\tau = e\Delta V/h$$

we can get to determine the ratio e/h .

6.16 Teaching the basics of classical mechanics to school students and first year undergraduates

The basic topics that must be introduced in chronological order are

[1a] The notion of displacement, velocity and acceleration: $x(t)$ is the distance covered along the x axis. The total displacement at time t is $|x(t) - x(0)|$ while the total distance covered by the particle upto time t is $\int_0^t |dx(t)| = \int_0^t |x'(t)|dt$. In three dimensions, the displacement at time t is the vector

$$\mathbf{r}(t) = x(t)\hat{x} + y(t)\hat{y} + z(t)\hat{z}$$

The displacement magnitude at time t assuming $\mathbf{r}(0) = 0$ is

$$|\mathbf{r}(t)| = \sqrt{x^2(t) + y^2(t) + z^2(t)}$$

while the total distance covered by the particle upto time t is

$$\int_0^t |\mathbf{dr}(t)| = \int_0^t |\mathbf{r}'(t)| dt$$

Exercise: Show using the triangle inequality for integrals that the total distance covered is always \geq the total displacement.

[1b] Description of instantaneous velocity and acceleration via Newtonian differential calculus.

If $\mathbf{r}(t)$ is the displacement at time t , then the instantaneous velocity at time t is

$$\mathbf{v}(t) = \mathbf{r}'(t) = x'(t)\hat{x} + y'(t)\hat{y} + z'(t)\hat{z}$$

The instantaneous speed at time t is

$$|\mathbf{v}(t)| = \sqrt{v_x(t)^2 + v_y(t)^2 + v_z(t)^2}, v_x(t) = x'(t), v_y(t) = y'(t), v_z(t) = z'(t)$$

[2] Newton's laws of motion.

[3] The solution to Newton's second law of motion for uniformly accelerated particles using Newtonian integral calculus. Generalization to non-uniformly accelerated particles.

[4] The motion of projectiles and their graphs.

[5] Vector addition.

[6] Free body diagrams and their application to the study of motion of bodies on an inclined plane, and motion of bodies attached by strings over pulleys and with slide along inclined planes.

[7] Taking static and dynamic friction into account while discussing sliding of bodies over inclined planes and on flat surfaces.

[8] The one dimensional simple harmonic oscillator.

[9] The simple pendulum: Approximate solution by the process of linearization.

[10] Power series expansion of the solution to the simple pendulum using elliptic integrals.

[11] The motion of a system of particles with external and internal forces.

[12] Conservation laws for momentum, energy and angular momentum.

[13] The three dimensional equations of motion in rectangular, cylindrical and spherical polar coordinates.

[14] The two dimensional equations of motion in rectangular and plane polar coordinates.

[15] The two body central force problem. Energy and angular momentum integrals.

[16] Kepler's laws of motion for motion under the inverse square law of gravitation. Derivation from the energy and angular momentum integrals.

[17] The three body problem. Solutions of Lagrange (Rotating equilateral triangles).

[18] The Lagrangian formulation of mechanics. Derivation of the Euler-Lagrange equations from variational calculus.

[19] The Hamiltonian formulation of mechanics using Legendre's transformation. Derivation of the Hamilton equations of motion from Lagrange's equations of motion.

[20] The existence of periodic orbits for Hamiltonian systems.

[21] The stability problem in Hamiltonian mechanics.

[22] The restricted three body problem: Motion of a particle in a periodic gravitational field generated by the rotation of two bodies around their centre of mass.

[23] Stability of an orbit.

In the dynamical system

$$x'(t) = f(x(t))$$

If $x_0(t), t \geq 0$ is a trajectory satisfying the above differential equation, then we say that this trajectory is stable if there exists a $\delta > 0$ and $x(0)$ such that $\|x(0) - x_0(0)\| < \delta$, we have that $\sup_t |x(t) - x_0(t)| < \infty$, ie, a small perturbation of the initial conditions implies that the trajectory for all times remains within a bounded region around the unperturbed trajectory. We say that the trajectory $x_0(t)$, is asymptotically stable if given any $\epsilon > 0$ there exists a $\delta > 0$ and a $0 < T < \infty$ such that $|x(0) - x_0(0)| < \delta$ implies that for all $t > T$, we have $|x(t) - x_0(t)| < \epsilon$. This condition is equivalent to requiring that for $|x(0) - x_0(0)| < \delta$, we have that $\lim_{t \rightarrow \infty} |x(t) - x_0(t)| = 0$.

[24] Mechanics of particles subject to random forces. Calculation of the statistical moments of the trajectories using perturbation theory for differential equations.

[25] Hamiltonian systems in noise described by Ito's stochastic differential equations taking damping and white noise fluctuating forces into account.

$$q_{,t} = H_{,p}(q, p), p_{,t} = -H_{,q}(q, p) - \zeta p + W(t)$$

$$dW(t) = -\gamma W(t)dt + \sigma dB(t)$$

The state of the system is $\xi(t) = [q(t)^T, p(t)^T, W(t)^T]^T$. For its pdf $f(t, \xi)$, we have the Fokker Planck equation

$$f_{,t} = -\operatorname{div}_q(H_{,p}f) + \operatorname{div}_p(H_{,q}f) + \zeta \cdot \operatorname{div}_p(pf) - W^T \nabla_p f + \gamma \operatorname{div}_W(Wf) + (\sigma^2/2) \nabla_W^2 f$$

The equilibrium density $f_e(q, p, W)$ satisfies

$$\{H, f\} + \zeta \cdot (nf + p^T \nabla_p f) - W^T \nabla_p f + \gamma(nf + W^T \nabla_W f) + (\sigma^2/2) \nabla_W^2 f = 0$$

[26] The normal form for Hamiltonian systems.

$$H = H(qp)$$

$$q' = H_{,p} = qH'(qp), p' = -H_{,q} = -pH'(qp)$$

Thus,

$$q'p + qp' = 0$$

Writing $\omega = qp$ gives $\omega' = 0$. So $\omega(t) = \text{const}t = \omega(0)$ and

$$q' = qH'(\omega(0)), p' = -pH'(\omega(0))$$

This gives us the complete solution:

$$q(t) = q(0)\exp(t.H'(\omega(0))), p(t) = p(0).\exp(-tH'(\omega(0))), \omega(0) = q(0).p(0)$$

[27] The motion of rigid bodies. Notion of moment of inertia.

[28] Rolling friction and the motion of a cylinder and a sphere on an inclined plane.

[29] The Lagrangian of a spinning top: Solution using Euler angles.

[30] Quasi-periodic motion of a Hamiltonian system with several degrees of freedom.

[31] Hamilton-Jacobi theory, the action-angle variables and its application to quasi-periodic motion of a mechanical system.

[32] The general theory of dynamical systems in continuous and discrete time. The Poincare recurrence theorem, Lyapunov exponents, invariant measure, ergodicity of a dynamical system, The Kolmogorov-Sinai entropy of a dynamical system. Spectral measure of a phase volume preserving dynamical system.

[33] Lyapunov's stability theory of a dynamical system.

[34] Applications of Lie group theory to mechanics: Noether's theorem on conserved currents associated with the invariance of a Lagrangian/Hamiltonian under a Lie group of transformations.

[35] Motion of a particle in a rotating frame: The Coriolis and centrifugal forces. Application to determining the angular velocity of the earth from the deflection of a falling body and from the deflection Foucault pendulum away from its oscillating plane at any given latitude.

[36] Application of classical mechanics to robot dynamics with several 3-D links and translation of the base pivot point.

[37] The Poisson bracket formulation of classical mechanics.

[38] Transition from classical mechanics to quantum mechanics using the Poisson-Lie bracket correspondence.

[39] Transition from classical mechanics to quantum mechanics using the Lagrangian Feynman path integral formalism.

[40] The classical theory of fields as a limiting case of classical mechanics when the number of particles becomes infinitely large. This correspondence is based on expanding the fields using orthonormal bases for the underlying Hilbert space and treating each coefficient as corresponding to a particle.

[41] The quantum theory of fields as a limiting case of the quantum mechanics of a finite system of particles based on Feynman's path integral approach and also based on the Hamiltonian operator theoretic formalism.

[42] The general relativistic n -body problem.

Let $g_{\mu\nu}(x)$ denote the metric of space time and let $x_i(\tau_i), i = 1, 2, \dots, N$ denote the world lines of these particles. Here, τ_i is the proper time for the i^{th} particle:

$$d\tau_i = \sqrt{g_{\mu\nu}(x_i)dx_i^\mu dx_i^\nu}$$

or equivalently,

$$\begin{aligned} d\tau_i(t)/dt &= \sqrt{g_{\mu\nu}(x_i(t))dx_i^\mu(t)dx_i^\nu(t)} \\ &= \sqrt{g_{00}(x_i) + g_{0m}(x_i)x_i^{m'}(t) + g_{km}(x_i(t))x_i^{k'}(t)x_i^{m'}(t)} \end{aligned}$$

where t is the coordinate time, ie, the time as measured by a clock located at an infinite distance from the n -bodies where the space-time manifold is flat Minkowskian. The energy-momentum tensor of the matter field of the n point particles is given by

$$T^{\mu\nu}(x) = (-g(x))^{-1/2} \sum_{i=1}^n m_i \delta^3(x - x_i) (dx_i^\mu/dt)(dx_i^\nu/dt)(dt/d\tau_i)$$

We note that

$$\int T^{\mu\nu}(x) \sqrt{-g(x)} d^4x = \sum_i m_i (dx_i^\mu/d\tau_i) dx_i^\nu$$

which is a tensor quantity since m_i is a scalar and dx_i^ν and $dx_i^\mu/d\tau_i$ are 4-vectors. The Einstein field equations are to be solved for the metric as well as for the world lines $x_i(t), i = 1, 2, \dots, n$ of the n particles. In order to do so, we make some approximations. We take as our metric a small perturbation of the Schwarzschild metric at large distances from the n particles where this system of particles will appear as a spherical body producing a gravitational potential ϕ .

$$\begin{aligned} d\tau^2 &= (1 + 2\phi/c^2 + h_{00}/c^4)dt^2 - [(1 - 2\phi/c^2)\delta_{ij}/c^2]dx^i dx^j \\ &\quad + 2(h_{0i}/c^3)dtdx^i = \\ &\quad g_{00}dt^2 + 2g_{0i}dtdx^i + g_{ij}dx^i dx^j \end{aligned}$$

We then compute using this metric the components of the Ricci tensor $R_{\mu\nu}$, ie, R_{00}, R_{0i}, R_{ij} upto $O(1/c^4)$ and substitute these into the Einstein field equations

$$R_{00} = K(T_{00} - Tg_{00}/2), R_{0i} = K(T_{0i} - Tg_{0i}/2),$$

$$R_{ij} = K(T_{ij} - Tg_{ij}/2)$$

where

$$K = -8\pi G/c^3$$

Exercise: Check that the units match on both sides of the Einstein field equations.

hint: R_{00} has the units of $1/t^2$. R_{ij} has the units of $1/c^2 t^2 = 1/d^2$ while R_{0i} has the units of $1/d \cdot t = 1/ct^2 = c/d^2$ where d means distance. On the other hand, T^{00} has the units of m/t^3 . T^{ij} has the units of $md^2/t^5 = mc^2/t^3$. T^{0i} has the units of $md/t^4 = mc/t^3$. It follows that T_{00} has the units of m/t^3 , T_{ij} has the units of $m/d^2 t = m/c^2 t^3$. Finally, T_{0i} has the units of m/ct^3 . Thus, KT_{00} has the units of $Gm/c^3 t^3 = d/ct^3 = 1/t^2$ which coincides with the units of R_{00} . KT_{0i} has the units of $Gm/c^4 t^3 = d/c^2 t^3 = 1/ct^2$ which coincides with the units of R_{0i} . Finally, KT_{ij} has the units of $Gm/c^5 t^3 = d/c^3 t^3 = 1/c^2 t^2$ which coincides with that of R_{ij} . This proves as regards units are concerned, the consistency of the Einstein field equations

$$R_{\mu\nu} = K(T_{\mu\nu} - Tg_{\mu\nu}/2)$$

[43] The dynamics of particles in electromagnetic fields.

In special relativistic mechanics, the equations of motion of a particle in an electromagnetic field using the four vector notation are

$$m \frac{d^2 x^\mu}{d\tau^2} = q F^{\mu\nu} \frac{dx_\nu}{d\tau}$$

where m is the rest mass of the particle (an invariant) and q its charge (an invariant). This can be expressed in three vector notation as

$$m \frac{d}{dt} (\gamma \cdot d\mathbf{r}/dt) = q\gamma (E + (d\mathbf{r}/dt) \times \mathbf{B})$$

where

$$\gamma = (1 - v^2/c^2)^{-1/2}, \mathbf{v} = d\mathbf{r}/dt$$

If relativistic effects are neglected, then we get the approximation

$$d^2\mathbf{r}/dt^2 = (q/m)(\mathbf{E} + (d\mathbf{r}/dt) \times \mathbf{B})$$

One important problem here is to study the motion of charged particles in periodically time varying em fields, ie,

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

This problem is a generalization of Hill's problem which involves the study of orbits in a periodic potential. Here, owing to the presence of the magnetic field, we have in addition velocity dependent potentials which are not covered by Hill's problem. To get approximate solutions, we define $\omega = 2\pi/T$ and expand the electric and magnetic fields as Fourier series in t :

$$\mathbf{E}(t, \mathbf{r}) = \sum_n \mathbf{E}_n(\mathbf{r}) \cdot \exp(in\omega t),$$

$$\mathbf{B}(t, \mathbf{r}) = \sum_n \mathbf{B}_n(\mathbf{r}).\exp(in\omega t)$$

We also expand the solution as a Fourier series in t assuming it to be periodic with period T . Assume

$$\mathbf{E}_n(\mathbf{r}) = \sum_{k \geq 0} \mathbf{e}(n, k) \mathbf{r}^{\otimes k},$$

$$\mathbf{B}_n(\mathbf{r}) = \sum_{k \geq 0} \mathbf{b}(n, k) \mathbf{r}^{\otimes k}$$

and writing

$$\mathbf{r}(t) = \sum_n \mathbf{c}(n) \exp(in\omega t)$$

we get

$$\begin{aligned} \mathbf{E}_n(\mathbf{r}(t)) &= \sum_{n, k, m_1, \dots, m_k} \mathbf{e}(n, k) \mathbf{c}(m_1) \otimes \dots \otimes \mathbf{c}(m_k) \exp(i(m_1 + \dots + m_k)\omega t) \\ &= \sum_m \mathbf{e}_n(m) \exp(jm\omega t) \end{aligned}$$

where

$$\mathbf{e}_n(m) = \sum_{k, m_1, \dots, m_k : m_1 + \dots + m_k = m} \mathbf{e}(n, k) (\mathbf{c}(m_1) \otimes \dots \otimes \mathbf{c}(m_k))$$

[44] The dynamics of electromagnetic fields in the presence of charged matter coupled to the dynamics of charged matter in electromagnetic fields.

[45] The dynamics of quantum electromagnetic fields and charged electrons and positrons in quantum field theory. Description of the interaction via the scattering matrix. Computing the scattering matrix using the operator theoretic formalism and using the path integral formalism of Feynman. Derivation of the Feynman diagrammatic rules using both the approaches.

[46] Rotating stroboscopic fan according to quantum mechanical laws with quantum noise coming from a bath. The image quantum em field passes through this rotating quantum fan window and at the output interacts with an atom causing transitions. From measurement of the transition probabilities, the parameters of the fan dynamics are to be estimated.

Classical fan dynamics derived from motor dynamics:

$$J\theta''(t) + \gamma\theta'(t) + \alpha\theta(t) = IBL\sin(\theta(t)) + w(t)$$

The Hamiltonian of the fan in the absence of noise and damping is

$$H(\theta, p) = p^2/2J + \alpha\theta^2/2 + IBL\cos(\theta)$$

This gives the correct Hamilton equations of motion:

$$\theta' = H_{,p}, p' = -H_{,\theta}$$

To quantize this motion and take bath noise into account, we must formulate the Schrodinger evolution in the form of a unitary dilation in system \otimes bath space using the Hudson- Parthasarathy quantum stochastic calculus. The evolution equation is

$$dU(t) = (-(iH + P)dt + LdA(t) - L^*dA(t)^*)U(t)$$

where

$$P = LL^*/2$$

The terms LdA and L^*dA^* account for noise while the term $-Pdt$ accounts for damping. The background quantum em field is

$$A(t, x, y) = \sum_k (f_k(t, x, y)a_k + \bar{f}_k(t, x, y)a_k^*)$$

where

$$[a_k, a_m^*] = \delta_{km}$$

and the a_k, a_k^* are independent of system and bath processes that describe the quantum fan. The fan window function is

$$K(\phi - \theta(t)) = j_t(K(\phi - \theta))$$

where

$$j_t(\theta) = U(t)^*\theta.U(t) = \theta(t)$$

which satisfies the Hudson-Parthasarathy-noisy Heisenberg dynamics

$$\begin{aligned} dj_t(\theta) &= j_t(i[H, \theta] - (1/2)(P\theta + \theta P - 2L\theta L^*)dt \\ &\quad + j_t(\theta L - L^*\theta)dA(t) + j_t(L^*\theta - \theta L)dA(t)^*) \\ dj_t(p) &= j_t(i[H, p] - (1/2)(Pp + pP - 2LpL^*)dt \\ &\quad + j_t(pL - L^*p)dA(t) + j_t(L^*p - pL)dA(t)^*) \end{aligned}$$

These are the quantum mechanical non-commutative generalizations of the classical rotating fan equation in the presence of damping and noise. We have already introduced here three Hilbert spaces, one, the Hilbert space \mathcal{H}_1 of the fan system, two the bath Hilbert space \mathcal{H}_2 in which the fan is rotating and three the Hilbert space \mathcal{H}_3 on which the electromagnetic field operators $a_k, a_k^*, k = 1, 2, \dots$ are defined. There is a fourth Hilbert space, namely that of the atomic system receiver that interacts with the quantum em field after it gets quantum windowed by the rotating fan. The total Hilbert space is the tensor product of these four Hilbert spaces. The windowed quantum em field is given by

$$X(t, x, y) = A(t, x, y)K(\phi - \theta(t)) = A(t, x, y)K(\tan^{-1}(y/x) - \theta(t))$$

We assume that A and hence X is directed along the z direction. When this quantum magnetic vector potential interacts with an atom, the interaction Hamiltonian is given by

$$H_I(t) = (e/2m_0)(P_z X + X P_z) = (-ie/m_0)X(t, x, y)\partial/\partial z$$

We assume that the atom is a $3 - D$ oscillator and the probability amplitude of this atomic oscillator making a transition from the stationary state $|n_1, n_2, n_3\rangle$ to the state $|m_1, m_2, m_3\rangle$ in time T with the bath remaining in the coherent state $|\phi(u)\rangle = \exp(-\|u\|^2/2)|e(u)\rangle$ for some $u \in L^2(\mathbb{R}_+)$ and the quantum image field behind the fan window being in the coherent state $|\phi(v)\rangle$ where $v = ((v_k)) \in l^2(\mathbb{Z}_+)$. We expand the solution to $U(t)$ upto quadratic orders in the noise operators L, L^* (This may be done more easily by attaching a perturbation tag ϵ to the noise operators L and L^* and hence the perturbation tag ϵ^2 to P). Thus writing

$$U(t) = U_0(t) + U_1(t) + U_2(t) + \dots$$

where $U_k(t)$ is of the k^{th} order in the noise operators, we obtain using standard second order perturbation theory,

$$dU_0(t) = -iHU_0(t)dt,$$

$$dU_1(t) = -iHU_1(t)dt + LU_0(t)dA(t) - L^*U_0(t)dA(t)^*$$

$$dU_2(t) = -iHU_2(t)dt - PU_0(t)dt + LU_1(t)dA(t) - L^*U_1(t)dA(t)^*$$

Solving these with the obvious initial conditions

$$U_0(0) = I, U_k(0) = 0, k \geq 1$$

gives us

$$U_0(t) = \exp(-itH),$$

$$U_1(t) = \int_0^t (U_0(t-s)(LU_0(s)dA(s) - L^*U_0(s)dA(s)^*),$$

$$U_2(t) = \int_0^t U_0(t-s)(-PU_0(s) + LU_1(s)dA(s) - L^*U_1(s)dA(s)^*)$$

$$= - \int_0^t U_0(t-s)PU_0(s)ds + \int_{0 < s_2 < s_1 < t} U_0(t-s_1)LU_0(s_1-s_2)LU_0(s_2)dA(s_2)dA(s_1)$$

$$- \int_{0 < s_2 < s_1 < t} U_0(t-s_1)LU_0(s_1-s_2)L^*U_0(s_2)dA(s_2)^*dA(s_1)$$

$$- \int_{0 < s_2 < s_1 < t} U_0(t-s_1)L^*U_0(s_1-s_2)LU_0(s_2)dA(s_2)dA(s_1)^*$$

$$+ \int_{0 < s_2 < s_1 < t} U_0(t-s_1)L^*U_0(s_1-s_2)L^*U_0(s_1)dA(s_2)^*dA(s_1)^*$$

The required transition probability in time $[0, T]$ is given approximately by where we denote

$$|n_1, n_2, n_3, \phi(u), \phi(v) \rangle = |n, u, v \rangle, |m_1, m_2, m_3, \phi(u), \phi(v) \rangle = |m, u, v \rangle$$

$$\begin{aligned} P_T(|n_1, n_2, n_3, \phi(u), \phi(v) \rangle \rightarrow |m_1, m_2, m_3, \phi(u), \phi(v) \rangle) &= \\ P_T(|n, u, v \rangle \rightarrow |m, u, v \rangle) &= \end{aligned}$$

$$-i \int_0^T \langle m, u, v | \tilde{H}_I(t) | n, u, v \rangle dt - \int_{0 < t_1 < t_2 < T} \langle m, u, v | \tilde{H}_I(t_1) \tilde{H}_I(t_2) | n, u, v \rangle dt_1 dt_2$$

where $\tilde{H}_I(t)$ is the interaction Hamiltonian between the output quantum em field and the atomic oscillator in the interaction picture. It is given by

$$\tilde{H}_I(t) = (e/m_0) \tilde{X}(t, x, y) \tilde{P}_z(t)$$

where

$$\tilde{X}(t, x, y) = \tilde{A}(t, x, y) K(\phi - \tilde{\theta}(t))$$

$$\tilde{P}_z(t) = ((P_z - iM\omega_0 Z)/2) \exp(-i\Omega_0 t) + (P_z + iM\omega_0 Z) \exp(i\Omega_0 t)/2$$

with $P_z = -i\partial/\partial Z$. Ω_0 is the natural frequency of the atomic oscillator. It should be noted that the unperturbed Hamiltonian of the system is the sum of (a) The quantum image field Hamiltonian $\sum_k \omega_k a_k^* a_k$, (b) the Hamiltonian of the fan motor described above and (c) the Hamiltonian of the atomic oscillator

$$(P_x^2 + P_y^2 + P_z^2)/2M + M\Omega_0^2(X^2 + Y^2 + Z^2)/2$$

Also

$$\tilde{\theta}(t) = W(t)^* \theta W(t)$$

where $W(t)$ is the same as $U(t)$ but with $U_0(t)$ removed, ie, upto second order terms,

$$\begin{aligned} W(t) &= U_0(t)^* U(t) = \\ I &+ \int_0^t (U_0(-s)(LU_0(s)dA(s) - L^*U_0(s)dA(s)^*)) \\ - \int_0^t U_0(-s)PU_0(s)ds &+ \int_{0 < s_2 < s_1 < t} U_0(-s_1)LU_0(s_1 - s_2)LU_0(s_2)dA(s_2)dA(s_1) \\ - \int_{0 < s_2 < s_1 < t} U_0(-s_1)LU_0(s_1 - s_2)L^*U_0(s_2)dA(s_2)^*dA(s_1) \\ - \int_{0 < s_2 < s_1 < t} U_0(-s_1)L^*U_0(s_1 - s_2)LU_0(s_2)dA(s_2)dA(s_1)^* \\ + \int_{0 < s_2 < s_1 < t} U_0(-s_1)L^*U_0(s_1 - s_2)L^*U_0(s_1)dA(s_2)^*dA(s_1)^* \end{aligned}$$

Finally,

$$\tilde{A}(t, x, y) = \sum_k [f_k(t, x, y)a_k \cdot \exp(-i\omega_k t) + \bar{f}_k(t, x, y)a_k^* \cdot \exp(i\omega_k t)]$$

We shall denote

$$g_k(t, x, y) = f_k(t, x, y)\exp(-i\omega_k t)$$

so that

$$\tilde{A}(t, x, y) = \sum_k g_k(t, x, y)a_k + \bar{g}_k(t, x, y)a_k^*$$

6.17 Teaching Linear algebra and functional analysis to post-graduate students of signal processing

[1] 2×2 matrices, computing their matrix exponential, solving linear vector differential equations with time varying coefficients using Dyson series using an application in quantum mechanics. Example of NH_3 molecule in an electric field. When N is above the plane of the triangle formed by the three H 's, it has an energy E_1 while when it is below it has an energy E_2 owing to the interaction of the electric dipole moment of the molecule with the electric field. If in the absence of the electric field, the energies in the two states are the same, say E_0 and if μ is the electric dipole moment along the z axis that is perpendicular to the plane of the molecule and \mathcal{E} is the electric field along the z direction, then $E_1 = E_0 - \mu\mathcal{E}$, $E_2 = E_0 + \mu\mathcal{E}$ and if $-iA$ is the transition probability amplitude per unit time from either the up state to the down state or vice-versa, then the Hamiltonian of the molecule in the electric field is

$$H(t) = \begin{pmatrix} E_1(t) & A \\ A & E_2(t) \end{pmatrix} =$$

$$\begin{pmatrix} E_0 - \mu\mathcal{E}(t) & A \\ A & E_0 + \mu\mathcal{E}(t) \end{pmatrix}$$

Now if the electric field is time independent, we can calculate the eigenvalues of this Hamiltonian as a function of the electric field and if it is time dependent, we can calculate the approximate transition probabilities from up state to down state and vice versa using time dependent perturbation theory in two state quantum mechanics and from these transition probability measurements, we can determine the electric dipole moment of the molecule. This is known as the Ammonia maser and has been excellently described in "The Feynman lectures on physics, vol. III".

[2] Illustrating the basic concepts in linear algebra using 2×2 matrices including computing the Jordan canonical form and functions of matrices like

the matrix exponential using the Jordan canonical form. Another interesting application in the theory of 2×2 Hermitian matrices is describing the evolution of the wave function for a two state molecule having a spin magnetic dipole moment interacting with a magnetic field. The magnetic field is assumed to be of the form

$$\mathbf{B}(t) = B_0 \hat{z} + \delta \cdot \mathbf{B}_1(t)$$

where B_0 is a strong constant magnetic field along the z axis and $\delta \cdot \mathbf{B}_1(t)$ is a small time varying perturbation to this magnetic field, it may for example have a specific frequency in which case the problem becomes one in NMR (nuclear magnetic resonance) spectroscopy. The interaction Hamiltonian between the spin of the molecule and the magnetic field is given by

$$H(t) = ge(\sigma, \mathbf{B}(t))/2m = H_0 + \delta \cdot H_1(t)$$

where

$$\begin{aligned} H_0 &= ge\sigma_3 B_0/2m = (geB_0/2m) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ &= \omega_0 \sigma_3 \end{aligned}$$

where

$$\omega_0 = geB_0/2m$$

is the Larmor frequency of precession and

$$H_1(t) = ge(\sigma, \mathbf{B}_1(t))/2m = (ge/2m) \begin{pmatrix} B_{1z}(t) & B_{1x}(t) - iB_{1y}(t) \\ B_{1x}(t) + iB_{1y}(t) & -B_z(t) \end{pmatrix}$$

Exercise: When $\mathbf{B}_1(t) = \mathbf{B}_{10}\cos(\omega t)$, then determine the approximate resonant solution to Schrodinger's equation when ω is close to ω_0 . This solution precisely displays the phenomenon of nuclear magnetic resonance.

[3] Give examples of non-diagonable matrices using similarity transformations of Jordan blocks. Using Jordan blocks, give examples of cases when the minimal polynomial does not equal the characteristic polynomial. Show using the same example that a matrix is diagonable iff the geometric multiplicity of each eigenvalue of the matrix equals its algebraic multiplicity. In general, we have that geometric multiplicity \leq algebraic multiplicity.

[4] Using the Jordan decomposition of a matrix T in the form

$$T = \sum_{k=1}^r (\lambda_k P_k + N_k)$$

where the P'_k 's are projections, N'_k 's are nilpotent operators and all the P'_k 's and N'_k 's commute with each other with $P_k P_j = P_k N_j = 0$, $k \neq j$, where specifically, we have

$$N_k = (T - \lambda_k)P_k$$

show that if $N_k^{m_k} = 0, N_k^{m_k-1} \neq 0$, then

$$\cdot(T) = \sum_{k=1}^r [f(\lambda_k)P_k + f'(\lambda_k)N_k/1! + f''(\lambda_k)N_k^2/2! + \dots + f^{(m_k-1)}(\lambda_k)N_k^{m_k-1}/(m_k-1)!]$$

whenever f is $m_k - 1$ times differentiable at λ_k for each k .

[5] Study of root systems in Lie algebra theory and Cartan's classification of the simple Lie algebras. Dynkin diagrams for root systems.

6.18 Variants of the Kalman filter

Maximum Co-entropy Kalman filter.

State model:

$$x[k+1] = F_k(x[k]) + W[k+1],$$

Measurement model:

$$z[k] = h_k(x[k]) + V[k]$$

$$z[k+1] - h_{k+1}(x[k+1]) = V[k+1],$$

$$x[k+1] - \hat{x}[k+1|k] = e[k+1|k]$$

and $V[k+1]$ is independent of $e[k+1|k]$.

$$Cov(e[k+1|k]) = P[k+1|k]$$

$\hat{x}[k+1|k+1]$ is computed by minimizing the energy of the error process

$$[V[k+1]^T, e[k+1|k]^T]^T$$

or more precisely, of its normalized version

$$[(P_V^{-1/2}V[k+1])^T, (P[k+1|k]^{-1/2}e[k+1|k])^T]^T$$

This random vector has zero mean and identity covariance. So if $\mathcal{E}[\xi]$ is an energy function, we minimize

$$\psi(x[k+1]) = \mathcal{E}[[P_V^{-1/2}(z[k+1] - h_{k+1}(x[k+1]))^T, (P[k+1|k]^{-1/2}(x[k+1] - \hat{x}[k+1|k]))^T]^T]$$

w.r.t. $x[k+1]$ to obtain $\hat{x}[k+1|k+1]$. This minimization is accomplished iteratively. In the maximum co-entropy method, we choose $\mathcal{E}(\xi)$ to be a truncated version of the exponential in the standard normal density, ie,

$$\mathcal{E}_N(\xi) = \sum_{k=1}^N \|\xi\|^{2k} / 2^k k!$$

Note that

$$\hat{x}[k+1|k] = F_k(\hat{x}[k|k]) + (1/2)F_k''(\hat{x}[k|k])Vec(P[k|k])$$

To complete the algorithm, we require $P[k+1|k+1]$. For this, we take $N = 1$ as in the EKF. We note that for $N = 1$,

$$\begin{aligned} \psi(x[k+1]) &= (1/2)(z[k+1] - h_{k+1}(x[k+1]))^T P_V^{-1}(z[k+1] - h_{k+1}(x[k+1])) + \\ &(1/2)(x[k+1] - \hat{x}[k+1|k])^T P[k+1|k]^{-1}(x[k+1] - \hat{x}[k+1|k]) \end{aligned}$$

Making the approximation

$$h_{k+1}(x[k+1]) = h_{k+1}(\hat{x}[k+1|k]) + h'_{k+1}(\hat{x}[k+1|k])(x[k+1] - \hat{x}[k+1|k])$$

this minimization is accomplished by solving a system of linear equations. If we are to be more accurate, ie, use the co-entropy method, then this linearized approximation must be substituted into the expression for $\psi(x[k+1])$ and then minimized by an iterative algorithm like the gradient search algorithm.

Application to the robot state estimation problem in the presence of disturbance. The robot dynamical equations are

$$M(q)q' + N(q, q') = \tau(t) + d(t) + w(t)$$

The disturbance observer equation is

$$z' = L(q, q')(-\tau + N - \hat{d}), \hat{d} = p(q') + z$$

Thus, we get

$$\begin{aligned} \hat{d}' &= p'(q')q'' + z' \\ p'(q')M^{-1}(\tau + d + w - N) - L(\tau - N + \hat{d}) \end{aligned}$$

Assuming

$$L(q, q') = p'(q')M(q)^{-1}$$

we get

$$\hat{d}' = L(q, q')(d - \hat{d} + w)$$

We assume that $d - \hat{d} = \tilde{w}$ is WGN. Then, $= \sigma B' = w + \tilde{w}$ is WGN and our basic equations for the robot dynamics and disturbance observer can be expressed in the form after subtracting out the disturbance estimate,

$$M(q)q'' + N(q, q') = \tau + \sigma B', \hat{d}' = \sigma L(q, q')B'$$

We define our state vector as

$$\xi = [q^T, q'^T, \hat{d}^T]^T$$

and then we can express our state equations in stochastic differential form as

$$dq = q'dt,$$

$$dq' = (-M(q)^{-1}N(q, q') + \tau)dt + \sigma M(q)^{-1}dB(t),$$

$$d\hat{d} = \sigma L(q, q')dB(t)$$

We wish to design the maximum co-entropy Kalman filter for estimating $\xi(t)$ based on the angular position measurements:

$$dz = qdt + dV$$

To do so, we first discretize the above coupled sde and then implement the maximum co-entropy filter in the form [1]. The discretized state evolution equations are

$$q[n+1] = q[n] + \Delta.\omega[n],$$

$$\omega[n+1] = \omega[n] + \Delta.(F_n(q[n], \omega[n]) + \sigma G(q[n])W[n+1]/\sqrt{\Delta},$$

$$\hat{d}[n+1] = \hat{d}[n]) + L(q[n], \omega[n])W[n+1]/\sqrt{\Delta}$$

where $W[n]$ is a sequence of iid standard normal random vectors and

$$F_n(q[n], \omega[n]) = (-M(q[n])^{-1}N(q[n], \omega[n]) + \tau[n])$$

$$\tau[n] = \tau(n\Delta),$$

$$G(q[n]) = M(q[n])^{-1}$$

and the measurement model is

$$y[n] = q[n]\Delta + \sigma.V[n].\sqrt{\Delta}$$

where $y[n] = z[n] - z[n-1]$ and $V[n]$ is a sequence of iid standard iid normal random vectors. The state equations can be expressed in the form

$$\xi[n+1] = K_n(\xi[n]) + J[\xi[n])W[n+1]$$

where

$$\xi = [q^T, \omega^T, \hat{d}^T]^T$$

and

$$K_n(\xi) =$$

6.19 The Cq-coding theorem:Proof based on Quantum Renyi entropy and Shannon's random coding argument

Let $x \in A$ be encoded into the state $W(x)$. Let $p(\cdot)$ be a probability distribution on the finite alphabet A . Let $\phi(1), \dots, \phi(N)$ be sequences in A^n . Let

$$\pi_k = \{W(\phi(k)) > 2N.W_p^{\otimes n}\}$$

where

$$W_p = \sum_{x \in A} p(x)W_x,$$

$$W(\phi(k)) = \otimes_{m=1}^n W(\phi_m(k))$$

where

$$\phi(k) = (\phi_1(k), \phi_2(k), \dots, \phi_n(k)) \in A^n$$

Let

$$Y_k = (\sum_{j=1}^N \pi_j)^{-1/2} \cdot Y_k \cdot (\sum_{j=1}^N \pi_j)^{-1/2}, 1 \leq k \leq N$$

Then

$$0 \leq Y_k \leq I, \sum_{k=1}^N Y_k = I$$

The error probability in using the code ϕ with each codeword $\phi(k)$ transmitted with the uniform probability of $1/N$ (this is plausible by the typical sequence argument) and the detection operators $\{Y_k\}$ is given by

$$\begin{aligned} P(\epsilon|\phi) &= N^{-1} \sum_{k=1}^N \text{Tr}(W(\phi(k))(1 - Y_k)) \\ &\leq N^{-1} \sum_{k=1}^N \text{Tr}(W(\phi(k))(2(1 - \pi_k) + 4 \cdot \sum_{j:j \neq k} \pi_j)) \end{aligned}$$

Now if we take the average of this error probability with respect to the probability distribution $p^{\otimes n}$ on A^n , ie, the codewords $\phi(k), k = 1, 2, \dots, N$ are independent random vectors, with probabilities $p^{\otimes n}(\phi(k)) = \prod_{m=1}^N p(\phi_m(k))$, then we get the following expression for the upper bound average error probability:

$$\begin{aligned} \mathbb{E}[P(\epsilon|\phi)] &\leq \\ 2N^{-1} \sum_{k=1}^N \mathbb{E}[\text{Tr}(W(\phi(k)).\{W(\phi(k)) \leq 2N.W_p^{\otimes n}\})] \\ + 4N^{-1} \cdot \sum_{1 \leq k \neq j \leq N} \mathbb{E}[\text{Tr}(W(\phi(k)).\{W(\phi(j)) \geq 2N.W_p^{\otimes n}\})] \end{aligned}$$

Now, for $s > 0$, we have

$$\begin{aligned} & \mathbb{E}[Tr(W(\phi(k))\{W(\phi(k)) \leq 2N.W_p^{\otimes n}\})] \\ & \leq \mathbb{E}[Tr(W(\phi(k))^{1-s}(2N)^s(W_p^{\otimes n})^s)] \\ & = (2N)^s \left(\sum_{x \in A} p(x) Tr(W(x)^{1-s}.W_p^s) \right)^n \\ & = 2^s \cdot \exp(ns.(R + s^{-1} \log[\sum_{x \in A} p(x) Tr(W(x)^{1-s} W_p^s)])) \end{aligned}$$

By using L'Hopital's rule, we find that

$$\begin{aligned} & \lim_{s \rightarrow 0+} s^{-1} \log[\sum_{x \in A} p(x) Tr(W(x)^{1-s} W_p^s)] \\ & = Tr(W_p \log(W_p)) - \sum_x p(x) W(x) \log(W(x)) \\ & = -I(p, W) \end{aligned}$$

Thus, we get the result that if $R < I(p, W)$, then

$$\mathbb{E}P(\epsilon|\phi) \rightarrow 0, n \rightarrow \infty$$

6.20 Manual for the Digital Signal Processing and Statistical Signal Processing Laboratories

6.20.1 List of Experiments based on the above theory

Before attempting these experiments, prepare a write-up containing the following:

- 1.Aim
- 2.List of equipment required.
- 3.Procedure.
- 4.Observations.
- 5.Results and Conclusions.

6.20.2 Linear Algebra and Probability theory in Signal Processing

Aim: To perform experiments related to filter design, signal and parameter estimation algorithms in linear systems having noise, like the Wiener filter, standard linear model, the Kalman filter, estimating the spatial frequencies in the DOA problem using eigensubspace based high resolution algorithms, order and time recursive estimation of parameters in AR models, estimating the numerator and denominator polynomials of rational IIR systems by giving an impulse input and measuring the impulse response, making a performance analysis of the estimation algorithms, describe the statistical hypothesis testing problem to discriminate between two quantum states in finite dimensional Hilbert spaces, study properties of white noise generated in heated circuits, determine experimentally the spectrum of noise, design numerical techniques for computing the capacity of an m-ary channel, verify via MATLAB simulations, Shannon's second coding theorem, namely that if we transmit information at a rate lesser than the capacity, then we can design a code that will be able to recover the transmitted symbol emitted by the source with an arbitrarily small probability of error,

[1] Design an FIR filter with p taps based on the frequency sampling method for a given transfer function $H_d(\omega)$, $|\omega| \leq \pi$ and plot the error between the output of the designed filter and the given filter for a given input sequence by taking $H_d(\omega)$ to be an IIR filter having a rational transfer function.

[2] Simulate the binary hypothesis testing problem so that under hypothesis H_1 , the measured signal is $s_1[n] + w[n]$, $1 \leq n \leq N$ and under hypothesis H_0 , the measured signal is $s_0[n] + w[n]$, $1 \leq n \leq N$. Take for example, s_1 and s_0 as ASK,PSK or FSK signals. Here, $w[n]$ is iid $N(0, \sigma^2)$. Calculate the log-likelihood ratio and empirically by doing several simulations, estimate the error probabilities and compare with the theoretical values.

[3] Simulate two quantum states ρ, σ , ie, positive definite $N \times N$ complex matrices of unit trace and calculate the optimal decision operator $0 \leq T \leq I_N$ for deciding between these two hypotheses. For a fixed value of the error probability $P_F = Tr(\sigma T) = c$, determine T such that the probability of detection $P_D = Tr(\rho T)$ is a maximum. Implement this algorithm.

hint: Using Lagrange multipliers, we have to maximize

$$\begin{aligned} f(T, \lambda) &= Tr(\rho T) - \lambda.(Tr(\sigma T) - c) \\ &= Tr((\rho - \lambda\sigma)T) + \lambda c \end{aligned}$$

For fixed λ , this is a maximum w.r.t T if T is chosen as the orthogonal projection

$$E(\lambda) = \{\rho - \lambda\sigma > 0\}$$

Finally, λ is determined by the constraint condition

$$Tr(\sigma.E(\lambda)) = c$$

- [4] Consider the least squares problem of estimating the parameter vector $\theta \in \mathbb{R}^p$ in the linear model

$$X = A\theta + W$$

where A is an $m \times n$ matrix of rank $r < \min(m, n)$. Determine using the SVD of A , the minimum norm least squares generalized inverse A^+ of A and show that the corresponding estimate of θ is

$$\hat{\theta} = A^+ X$$

Verify by Monte-Carlo simulations that this is indeed the minimum norm least squares estimator of θ .

[5] Connect a heated up resistor to a CRO and display plots of the random waveform. Determine its spectral density from empirical measurements of the noisy waveform. Compare this spectral density with that obtained by connecting the resistor to a spectral analyzer. Repeat when the resistor is connected to a capacitor and/or inductor in a single closed loop. Measure the spectrum of the voltage across the resistor, capacitor and inductor and verify with theoretical calculations by assuming that the noise voltage appearing at the resistor terminals can be replaced by a white noise source of spectral density $kT/2$ or more precisely by a noise source with Planckian spectrum.

[6] Write a program to calculate the capacity $C(q)$ of an M -ary channel with transition probabilities $((q(i, j))_{1 \leq i, j \leq M}$, ie, $q(i, j) \geq 0, \sum_j q(i, j) = 1$. By means of a search over all codewords of length n where n is a large integer and non-overlapping decision regions, show that we can select $N(n)$ different codewords with $n^{-1} \log(N(n)) < C$ and $Pr(\text{error}) < \epsilon$ where $\epsilon > 0$ is given. Assume any probability distribution of the source on $\{1, 2, \dots, N\}$. Conversely, show again by an exhaustive search that there is an $\epsilon > 0$ such that no matter how large n is, we cannot obtain non-overlapping decision regions that will result in an error probability $< \epsilon$ with N codewords when $n^{-1} \log(N) > C$.

6.20.3 Linear Algebra in Signal Processing

- [1] Write a MATLAB program to simulate the state variable equation

$$X[n + 1] = AX[n] + GW[n + 1]$$

where $X[n]$ is a vector in \mathbb{R}^p , A is a $p \times p$ matrix, G is a $p \times q$ matrix and $W[n]$ is a sequence of iid standard Gaussian vectors in \mathbb{R}^q . For simulation purposes, assume that $p = 4, q = 2$. Write a program to compute the time averaged autocorrelation of $X[n]$, ie,

$$R_{XX} = N^{-1} \cdot \sum_{n=1}^N X[n]X[n]^T$$

Compare with the theoretically obtained autocorrelation of $X[n]$.

[2] Write a program to calculate the eigenvalues of an $N \times N$ matrix. Your program must be based on calculating the coefficients of the characteristic polynomial of the matrix and then using the MATLAB command "roots".

[3] Write a MATLAB program to calculate the coefficients of a monic polynomial given its roots. Compare your result with the MATLAB command "poly".

[4] Write a program to Gram-Schmidt orthonormalize a given set of p linearly independent vectors of size $N \times 1$ where $p \leq N$. Take for example, $N = 20, p = 10$ and check your result by computing the inner product between the Gram-Schmidt orthonormalized vectors.

[5] Generate an $N \times p$ real matrix X and write a program to calculate the orthogonal projection P onto $\mathcal{R}(X)$, ie,

$$P = X(X^T X)^{-1} X$$

Check by taking a few examples that if $\xi \in \mathbb{R}^p$, and $\eta \in \mathbb{R}^N$, then

$$\| \eta - P\eta \| \leq \| \eta - A\xi \|$$

ie, $P\eta$ is the best approximation to η in $\mathcal{R}(A)$.

[6] In the previous problem, if δX is a small perturbation of X , then calculate the corresponding perturbation in P , both using theoretical linearization and directly using MATLAB.

[7] Write a MATLAB program to generate N samples of the impulse response $h[n]$ of an IIR filter

$$H(z) = \frac{\sum_{k=0}^q b[k]z^{-k}}{1 + \sum_{k=1}^p a[k]z^{-k}}$$

Now using the relation

$$h[n] + \sum_{k=1}^p a[k]h[n-k] = b[n]$$

estimate $a[k], b[l]$ from the sequence $h[n]$ using Prony's method and Shank's method. Now add WG noise to the measured impulse response $h[n]$ and do the same estimation. Calculate the mean squared error energy nsrs as a function of the variance of the noise. Make a theoretical computation of the same and compare the results.

hint: For $n > q$, we have

$$h[n] + \sum_{k=1}^p a[k]h[n-k] = 0$$

Apply the least squares algorithm to this expression for estimating $\{a[k]\}$ from $\{h[n]\}$.

[8] Generate the signal and measurement processes $s[n], x[n]$ using

$$s[n] = as[n - 1] + w[n], x[n] = s[n] + v[n]$$

where $|a| < 1$ and $w[n]$ is an iid $N(0, \sigma_w^2)$ sequence and $v[n]$ is an iid $N(0, \sigma_v^2)$ sequence independent of the $w[n]$ sequence. Calculate the transfer function of the causal Wiener filter for estimating $s[n]$ based on $x[k], k \leq n$. Now, pass $x[n]$ through this filter to obtain the estimate $\hat{s}[n]$. Calculate the SNR

$$SNR = \frac{\sum_n s[n]^2}{\sum_n (s[n] - \hat{s}[n])^2}$$

and compare with the theoretically obtained value

$$SNR(th) = \frac{\mathbb{E}(s[n]^2)}{\mathbb{E}(s[n] - \hat{s}[n])^2}$$

Assume all processes to be stationary.

[9] Write a program to implement the Kalman filter for the state-variable system

$$X[n+1] = AX[n] + GW[n+1]$$

with the measurement model

$$Z[n] = HX[n] + V[n]$$

where $X[n] \in \mathbb{R}^p$, $A \in \mathbb{R}^{p \times p}$, $G \in \mathbb{R}^{p \times p}$, $Z[n] \in \mathbb{R}^q$, $H \in \mathbb{R}^{q \times p}$ and $W[n]$ is iid $N(0, I_p)$ while $V[n]$ is iid $N(0, Q)$ independent of the sequence $W[.]$ where Q is a positive definite matrix of size $q \times q$. Calculate the time-averaged SNR defined by

$$SNR = \frac{\sum_n X[n]^T X[n]}{\sum_n e[n]^T e[n]}$$

where

$$e[n] = e[n|n] = X[n] - \hat{X}[n|n]$$

with

$$\hat{X}[n|n] = \mathbb{E}(X[n]|Z[k], k \leq n)$$

and compare with the theoretically obtained value

$$SNR(th) = \frac{\sum_n \mathbb{E}(X[n]^T X[n])}{\sum_n \text{Tr}(P[n|n])}$$

where

$$P[n|n] = \text{Cov}(e[n|n]) = \mathbb{E}(e[n|n] \cdot e[n|n]^T)$$

Straight forward calculations show that

$$\hat{X}[n+1|n] = A\hat{X}[n|n],$$

$$P[n+1|n] = Cov(X[n+1] - \hat{X}[n+1|n]|Z_n) = AP[n|n]A^T + GG^T,$$

$$\hat{X}[n+1|n+1] = \hat{X}[n+1|n] + K[n+1](Z[n+1] - H\hat{X}[n+1|n])$$

where $K[n+1]$ is chosen to minimize

$$\mathbb{E} \| X[n+1] - \hat{X}[n+1|n+1] \|^2$$

$$= Tr((I - K[n+1]H)P[n+1|n](I - K[n+1]H)^T + K[n+1]QK[n+1]^T)$$

and finally,

$$P[n+1|n+1] = Cov(X[n+1] - \hat{X}[n+1|n+1]) =$$

$$Cov(X[n+1] - \hat{X}[n+1|n] - K[n+1](H(X[n+1] - \hat{X}[n+1|n]) + V[n+1]))$$

$$= (I - K[n+1]H)P[n+1|n](I - K[n+1]H)^T + K[n+1]QK[n+1]^T$$

[10] Generate an AR time series as

$$x[n] = - \sum_{k=1}^p a[k]x[n-k] + w[n]$$

where $w[n]$ is iid $N(0, \sigma^2)$ and $a[k]'s$ are such that the polynomial

$$A(z) = 1 + \sum_{k=1}^p a[k]z^{-k}$$

has all its roots within the unit circle (Generate such a polynomial using the "poly" command in MATLAB).

Now implement the Levinson-Durbin algorithm to estimate the AR coefficients $a[k]$ using time averaged estimates of the autocorrelation matrix. Calculate the reflection coefficients and implement the system as a Lattice filter. Calculate empirically the SNR involved in simulating the system using the Lattice filters of different orders upto order p . Demonstrate via such simulations that the SNR keeps increasing as the order increases from 0 to p .

hint: The Levinson-Durbin algorithm is derived as follows: Let $a_p[k], k = 1, 2, \dots, p$ denote the p^{th} order predictor of the stationary process $x[n]$ having autocorrelation $R[k] = \mathbb{E}(x[n]x[n-k])$. Thus, the p^{th} -order prediction of $x[n]$ is given by

$$\hat{x}[n] = - \sum_{k=1}^p a_p[k]x[n-k]$$

The prediction coefficients are determined by minimizing

$$\mathbb{E}e[n]^2, e[n] = A_p(z)x[n] = x[n] + \sum_{k=1}^p a_p[k]x[n-k]$$

so that

$$A_p(z) = 1 + \sum_{k=1}^p a_p[k]z^{-k}$$

The optimal normal equations are then

$$(\partial/\partial a_p[k])\mathbb{E}e[n]^2 = 0$$

and these result in the orthogonality principle

$$\mathbb{E}(e[n]x[n-k]) = 0, 1 \leq k \leq p$$

with the minimum prediction error energy

$$E[p] = \mathbb{E}e[n]^2 = \mathbb{E}(e[n]x[n])$$

These formulae can be expanded to give

$$\sum_{m=1}^p R[k-m]a_p[m] = -R[k], 1 \leq k \leq p,$$

$$E[p] = R[0] + \sum_{k=1}^p a_p[k]R[k]$$

or equivalently,

$$\mathbf{R}_p \mathbf{a}_p = -\mathbf{r}[1:p] \quad \dots \quad (1)$$

$$E[p] = R[0] + \mathbf{a}_p^T \mathbf{r}[1:p] \quad \dots \quad (2)$$

Now define the $p \times p$ inversion matrix \mathbf{J}_p that takes every vector $[\xi_1, \dots, \xi_p]^T$ to its reversed version $[\xi_p, \dots, \xi_1]^T$. Then we have

$$\mathbf{J}_p[x[n-1], \dots, x[n-p]]^T = [x[n-p], x[n-p+1], \dots, x[n-1]]^T$$

and hence by wide-sense stationarity,

$$\mathbf{J}_p \mathbf{R}_p \mathbf{J}_p^T = \mathbf{R}_p \quad \dots \quad (3)$$

Define

$$\mathbf{b}_p = \mathbf{J}_p \mathbf{a}_p \quad \dots \quad (4)$$

Then since $\mathbf{J}_p^2 = \mathbf{I}_p$, we get from (1),

$$\mathbf{J}_p \mathbf{R}_p \mathbf{J}_p \mathbf{b}_p = -\mathbf{r}[p:1]$$

or in view of (3)

$$\mathbf{R}_p \mathbf{b}_p = -\mathbf{r}[p : 1]$$

Now the $p + 1^{th}$ order predictor is given by

$$\mathbf{R}_{p+1} \mathbf{a}_{p+1} = -\mathbf{r}[1 : p + 1] \quad \dots \quad (5)$$

$$E[p + 1] = R[0] + \mathbf{a}_{p+1}^T \mathbf{r}[1 : p + 1] \quad \dots \quad (6)$$

(5) can be expressed as

$$\mathbf{R}_p \mathbf{a}_{p+1}[1 : p] + \mathbf{r}[p : 1] a_{p+1}[p + 1] = -\mathbf{r}[1 : p] \quad \dots \quad (7),$$

$$\mathbf{r}[p : 1]^T \mathbf{a}_{p+1}[1 : p] + R[0] a_{p+1}[p + 1] = -R[p + 1] \quad \dots \quad (8)$$

Assume

$$\mathbf{a}_{p+1}[1 : p] = \mathbf{a}_p - K \mathbf{b}_p$$

Then, we get on substitution,

$$\mathbf{R}_p(\mathbf{a}_p[1 : p] - K \mathbf{b}_p) + \mathbf{r}[p : 1] a_{p+1}[p + 1] = -\mathbf{r}[1 : p],$$

which becomes on using the equations of the p^{th} order predictor,

$$-\mathbf{r}[1 : p] + K \mathbf{r}[p : 1] + \mathbf{r}[p : 1] a_{p+1}[p + 1] + \mathbf{r}[1 : p] = 0$$

or equivalently,

$$K = -a_{p+1}[p + 1] \quad \dots \quad (9)$$

Further, we get on using (8),

$$R[p + 1] + \mathbf{r}[p : 1]^T (\mathbf{a}_p - K \mathbf{b}_p) - R[0] K = 0$$

Now using

$$E[p] = R[0] + \mathbf{a}_p^T \mathbf{r}[1 : p] = R[0] + \mathbf{b}_p^T \mathbf{r}[p : 1]$$

we get

$$R[p + 1] + \mathbf{a}_p^T \mathbf{r}[p : 1] - K(E[p] - R[0]) - KR[0] = 0$$

or equivalently,

$$R[p + 1] + \mathbf{a}_p^T \mathbf{r}[p : 1] - KE[p] = 0$$

so that

$$\mathbf{a}_{p+1}[p + 1] = -K = -\frac{R[p + 1] + \mathbf{a}_p^T \mathbf{r}[p : 1]}{E[p]}$$

We finally require an update formula for $E[p + 1]$:

$$\begin{aligned} E[p + 1] &= R[0] + \mathbf{a}_{p+1}^T \mathbf{r}[1 : p + 1] \\ &= R[0] + \mathbf{a}_{p+1}[1 : p]^T \mathbf{r}[1 : p] - KR[p + 1] \\ &= R[0] + (\mathbf{a}_p - K \mathbf{b}_p)^T \mathbf{r}[1 : p] - KR[p + 1] \end{aligned}$$

$$\begin{aligned}
&= E[p] - K \mathbf{a}_p^T \mathbf{r}[p : 1] - KR[p + 1] \\
&= E[p] - K(R[p + 1] + \mathbf{r}[p : 1]^T \mathbf{a}_p) = (1 - K^2)E[p]
\end{aligned}$$

It is usual to denote K by $K[p + 1]$.

Remark: Solving a system of p linear equations in p variables by the Gauss elimination method requires $O(p^3)$ multiplications while in this case, we solve for the p^{th} order predictor starting from the zeroth order predictor using $O(p^2)$ multiplications, ie, $O(p)$ multiplications to go from p to $p + 1$. This is because the matrix \mathbf{R}_p has the centro-symmetric property (3).

[11] Implement the RLS and RLS lattice algorithms for time recursive estimation and simultaneous time and order recursive estimation respectively for the system given in the previous problem.

Specifically, generate a time series $x[n]$ according to the ARMA difference equation

$$x[n] + \sum_{k=1}^p a[k]x[n - k] = \sum_{k=0}^p b[k]w[n - k]$$

where $w[n]$ is WGN of unit variance. The $b[k]$'s can be selected arbitrarily and the $a[k]$'s are selected using the poly command in MATLAB so that

$$A(z) = 1 + \sum_{k=1}^p a[k]z^{-k}$$

has all its roots inside the unit circle. Now define the data vector

$$\mathbf{x}_n = [x[n], x[n - 1], \dots, x[0]]^T \in \mathbb{R}^{n+1 \times 1}$$

$$z^{-r} \mathbf{x}_n = [x[n - r], x[n - r - 1], \dots, x[0], 0, \dots, 0]^T \in \mathbb{R}^{n+1 \times 1}$$

which follows by taking $x[n] = 0$ for $n < 0$. Generate the data matrix

$$\mathbf{X}_{n,m} = [z^{-1} \mathbf{x}_n, \dots, z^{-m} \mathbf{x}_n] \in \mathbb{R}^{n+1 \times m}$$

Construct the optimal p^{th} order predictor $= \mathbf{a}_{n,p} \in \mathbb{R}^{p \times 1}$ so that

$$\| \mathbf{x}_n + \mathbf{X}_{n,p} \mathbf{a} \| ^2$$

is a minimum. The answer is

$$\mathbf{a}_{n,p} = -(\mathbf{X}_{n,p}^T \mathbf{X}_{n,p})^{-1} \mathbf{X}_{n,p}^T \mathbf{x}_n$$

Compute the corresponding p^{th} order forward prediction error vector as

$$\mathbf{e}_f(n|p) = \mathbf{x}_n + \mathbf{X}_{n,p} \mathbf{a}(n|p)$$

Also compute the p^{th} order backward prediction filter $\mathbf{b} = \mathbf{b}_{n,p}$ by minimizing

$$\| z^{-p} \mathbf{x}_n + \mathbf{X}_{n+1,p} \mathbf{b} \| ^2$$

Show that it is given by

$$\mathbf{b}_{n,p} = (\mathbf{X}_{n+1,p}^T \mathbf{X}_{n+1,p})^{-1} \mathbf{X}_{n+1,p} z^{-p} \mathbf{x}_n$$

Compute the corresponding p^{th} order backward prediction error vector defined by

$$\mathbf{e}_b(n|p) = z^{-p} \mathbf{x}_n + \mathbf{X}_{n+1,p} \mathbf{b}_{n,p}$$

Let $\mathbf{P}_{n,p}$ denote the orthogonal projection onto $\mathcal{R}(\mathbf{X}_{n,p})$. Show that

$$\mathbf{e}_f(n|p) = \mathbf{P}_{n,p}^\perp \mathbf{x}_n = (\mathbf{I} - \mathbf{P}_{n,p}) \mathbf{x}_n$$

$$\mathbf{e}_f(n-1|p) = \mathbf{P}_{n,p}^\perp z^{-p-1} \mathbf{x}_n$$

Now write a program to compute these two errors recursively with increasing p . Also write a program to compute these two error vectors with increasing n .

Remark: This amounts to developing recursive formulas for computing the projection operator updates for $\mathbf{P}_{n,p}$ recursively with increasing n and with increasing p . The time update formula for these projection operators is achieved by noting what it becomes when one extra row is added to $\mathbf{X}_{n,p}$ at the top while the order update formula is achieved by noting what it becomes when one extra column is added to $\mathbf{X}_{n,p}$ at the end.

Now consider another process $y[n]$ generated by passing $x[n]$ through an FIR filter and then adding noise to it. Estimate the p^{th} order FIR filter $\mathbf{h} = \mathbf{h}_{n,p}$ so that

$$\| \mathbf{y}_n - \mathbf{X}_{n+1,p} \mathbf{h} \|^2$$

is a minimum. In view of the Gram-Schmidt orthonormalization process, this filter $\mathbf{h}_{n,p}$ can also be realized by minimizing

$$\| \mathbf{y}_n - (g[0] \mathbf{e}_b(n|p-1) + g[1] \mathbf{e}_b(n|p-2) + \dots + g[p-1] \mathbf{e}_b(n|0)) \|^2$$

This minimization is easy to carry out since the vectors $\mathbf{e}_b(n|k)$, $k = 0, 1, \dots, p-1$ are mutually orthogonal. Carry out this minimization and hence implement the FIR filter $\mathbf{h}_{n,p}$ as a Lattice filter.

[12] Write a program to calculate in the inverse of a matrix when we add one extra row and one extra column using a recursive procedure, ie, the inverse of the appended matrix should be expressed in terms of the inverse of the original matrix.

hint: For block structured matrices, if

$$X = \begin{pmatrix} P & Q \\ R & S \end{pmatrix}$$

and

$$X^{-1} = \begin{pmatrix} \tilde{P} & \tilde{Q} \\ \tilde{R} & \tilde{S} \end{pmatrix}$$

then

$$\begin{aligned} P\tilde{P} + Q\tilde{R} &= I, P\tilde{Q} + Q\tilde{S} = 0, \\ R\tilde{P} + S\tilde{R} &= 0, \\ R\tilde{Q} + S\tilde{S} &= I \end{aligned}$$

[13] Simulate a two dimensional array signal vector in the form

$$X[n, m] = \sum_{k=1}^p A[k] \exp(j(\omega(1, k)n + \omega(2, k)m)) + W[n, m]$$

where W is spatial WGN and $A[k]$ are independent normal random variables with

$$\mathbb{E}(A[k]A[m]) = P[k]\delta[k - m]$$

Implement the 2-D MUSIC and 2-D ESPRIT algorithms for estimating the 2-D frequency pairs $(\omega(1, k), \omega(2, k))$, $k = 1, 2, \dots, p$ from the data $X[n, m]$, $1 \leq n, m \leq N$.

hint: Use the Kronecker tensor product and refer to the papers on 2-D MUSIC and 2-D ESPRIT by H.Parthasarathy, S.Prasad and S.D.Joshi published in Elsevier, Signal Processing (1994) and IEEE transactions on Signal processing (1995).

Express the array signal data in the form

$$\mathbf{X} = \mathbf{E}(\theta)\mathbf{a} + \mathbf{W}$$

and its shifted versions as

$$\mathbf{X}_1 = \mathbf{E}(\theta)\Phi_1\mathbf{a} + \mathbf{W}_1,$$

$$\mathbf{X}_2 = \mathbf{E}(\theta)\Phi_2\mathbf{a} + \mathbf{W}_2$$

Evaluate the array signal correlations as

$$\mathbf{R}_{XX} = \mathbf{E}(\theta)\mathbf{P}\mathbf{E}(\theta)^* + \sigma^2\mathbf{I},$$

$$\mathbf{R}_{XX_1} = \mathbf{E}(\theta)\mathbf{P}\Phi_1^*\mathbf{E}(\theta) + \sigma^2\mathbf{Z}_1,$$

$$\mathbf{R}_{XX_2} = \mathbf{E}(\theta)\mathbf{P}\Phi_2^*\mathbf{E}(\theta) + \sigma^2\mathbf{Z}_2$$

Where

$$\theta = \{(\omega(1, k), \omega(2, k)) : 1 \leq k \leq p\}$$

and

$$\mathbf{E}(\theta) = \text{Col}[\mathbf{e}(\omega(1, k)) \otimes \mathbf{e}(\omega(2, k)) : 1 \leq k \leq p]$$

where

$$\mathbf{e}(\omega) = ((\exp(jn\omega)))_{n=0}^{N-1} \in \mathbb{C}^N$$

$$\Phi_1 = \text{diag}[\exp(j\omega(1, k)) : 1 \leq k \leq p],$$

$$\Phi_2 = \text{diag}[\exp(j\omega(2, k)) : 1 \leq k \leq p]$$

Now calculate the signal and noise eigenvectors of \mathbf{R}_{XX} . If $\mathbf{v}_{p+1}, \dots, \mathbf{v}_N$ are the set of orthonormal eigenvectors, then they are orthogonal to the signal subspace $\mathcal{R}(\mathbf{E}(\theta))$ and hence θ can be determined by minimizing

$$F(\omega_1, \omega_2) = \sum_{k=p+1}^n |\mathbf{v}_n^*(\mathbf{e}(\omega_1) \otimes \mathbf{e}(\omega_2))|^2$$

or equivalently by locating the peaks of the MUSIC pseudo-spectrum estimator

$$P(\omega_1, \omega_2) = \frac{1}{F(\omega_1, \omega_2)}$$

This is the $2 - D$ MUSIC algorithm. In the $2 - D$ ESPRIT algorithm, we compute the rank reducing numbers of

$$\begin{aligned} & \mathbf{E}(\theta)\mathbf{P}(\mathbf{I} - \gamma\Phi_k^*)\mathbf{E}(\theta)^* \\ &= \mathbf{R}_{XX} - \sigma^2\mathbf{I} - \mathbf{R}_{XX_k} + \sigma^2\mathbf{Z}_k, k = 1, 2 \end{aligned}$$

and then pair them. Equivalently, we determine the rank reducing pairs (γ_1, γ_2) that fall on the two torus of the matrix pencil

$$\begin{aligned} & \mathbf{E}(\theta)\mathbf{P}(2\mathbf{I} - \gamma_1\Phi_1^* - \gamma_2\Phi_2^*)\mathbf{E}(\theta)^* \\ &= 2(\mathbf{R}_{XX} - \sigma^2\mathbf{I}) - \gamma_1(\mathbf{R}_{XX_1} - \sigma^2\mathbf{Z}_1) \\ & \quad - \gamma_2(\mathbf{R}_{XX_2} - \sigma^2\mathbf{Z}_2) \end{aligned}$$

Note that the noise eigenvalue σ^2 is determined as the minimum eigenvalue of \mathbf{R}_{XX} .

6.20.4 MATLAB problems in waves and optics

[1] Write a MATLAB program to simulate the evolution of the electromagnetic field in space-time given the current density and the charge density. Do the simulation of the fields within a cube of side length L . Do it directly by a space-time discretization of the Maxwell equations and then in the frequency domain using the 3-D FFT for the spatial variables.

[2] Simulate the Maxwell em field for a given charge and current density within a cube by 3-D Fourier series analysis by using the wave equations satisfied by the magnetic vector potential and electric scalar potential in the Lorentz gauge. Discretize in time the spatial Fourier series for the potentials.

[3] Write a MATLAB program to simulate the evolution of the wave function of the Dirac wave field in a given electromagnetic field within a cube by using

3-D spatial Fourier series. Repeat when in addition the Schwarzschild is present as the curved space-time background.

[4] Take an n -dimensional vector space like say $V = \mathbb{C}^n$. Consider the Fermion Fock space $\Lambda V = \mathbb{C} \oplus \bigoplus_{k=1}^n \Lambda^k V$ and define the annihilation and creation operator fields on this space by

$$a(u)(u_1 \wedge u_2 \wedge \dots \wedge u_k) = u \wedge u_1 \wedge u_2 \wedge \dots \wedge u_k$$

and

$$a(u)^*(u_1 \wedge \dots \wedge u_k) = \sum_{m=1}^k (-1)^{m-1} \langle u | u_m \rangle u_1 \wedge \dots \wedge u_{m-1} \wedge u_{m+1} \dots \wedge u_k$$

Show that these two operators are adjoint to each other and verify their anticommutation relations using MATLAB experiments.

[5] Numerically evaluate the transition probability from one stationary state to another for a quantum harmonic oscillator in the presence of quantum noise in the form of creation and annihilation processes. Assume that the bath remains in the coherent state $|\phi(u)\rangle$. hint: The HP equation is

$$dU(t) = (-(iH + P)dt + LdA(t) - L^*dA(t)^*)U(t)$$

writing

$$U(t) = U_0(t)W(t), U_0(t) = \exp(-itH), P(t) = U_0(t)^*PU_0(t), L(t) = U_0(t)^*LU(t)$$

we get

$$dW(t) = [-P(t)dt + L(t)dA(t) - L(t)^*dA(t)^*]W(t)$$

6.20.5 MATLAB exercises on classical and quantum information theory

[1] Take a finite alphabet, say $A = \{1, 2, \dots, N\}$. Generate a probability distribution $p(x), x \in A$ on A . Choose a small $\delta > 0$ and generate all the Bernoulli typical sequences $T(n, p, \delta)$ for a given positive integer n . Note that a sequence $u \in T(n, p, \delta)$ iff

$$|N(x|u) - np(x)| \leq \delta \sqrt{np(x)(1 - p(x))}, \forall x \in A$$

where x is the number of times x appears in the sequence u . Verify by simulation the following inequalities: [a]

$$np(x) - \delta \sqrt{np(x)(1 - p(x))} \leq N(x|u) \leq np(x) + \delta \sqrt{np(x)(1 - p(x))}, \forall x \in A$$

[b]

$$2^{-nH(p) - K\delta\sqrt{n}} \leq p^{\otimes n}(u) \leq 2^{-nH(p) + K\delta\sqrt{n}}$$

where

$$H(p) = - \sum_{x \in A} p(x) \log(p(x))$$

and

$$K = \sum_{x \in A} \sqrt{p(x)(1-p(x))} \leq a/2$$

Let $\mu(E)$ denote the number of elements in the set E . Then, [c]

$$\mu(T(n, p, \delta)) \leq 2^{nH(p) + K\delta\sqrt{n}}$$

[d]

$$p^{\otimes n}(T(n, p, \delta)^c) \leq a/\delta^2$$

[d]

$$p^{\otimes n}(T(n, p, \delta)) \geq 1 - a/\delta^2$$

[2] Quantum Schumacher compression.

Simulate a density matrix ρ in \mathbb{C}^p , ie, a positive definite unit trace matrix. Let its spectral decomposition be given by

$$\rho = \sum_{j=1}^p |j\rangle p(j) \langle j|$$

Construct the following spectral decomposition of the matrix $\rho^{\otimes n}$:

$$\rho^{\otimes n} = \sum_{j_1, \dots, j_n=1}^p |j_1, \dots, j_n\rangle p(j_1) \dots p(j_n) \langle j_1, \dots, j_n|$$

We write

$$\mathbf{j} = (j_1, \dots, j_n)$$

and

$$p^{\otimes n}(\mathbf{j}) = p(j_1) \dots p(j_n)$$

Then, we can write

$$\rho^{\otimes n} = \sum_{\mathbf{j}} |\mathbf{j}\rangle p^{\otimes n}(\mathbf{j}) \langle \mathbf{j}|$$

Define for $\delta > 0$, the δ -typical sequence space of length n :

$$T(n, p, \delta) = \{ \mathbf{j} \mid |N(x|\mathbf{j}) - np(x)| \leq \delta \sqrt{np(x)(1-p(x))} \forall x = 1, 2, \dots, p \}$$

Then construct the δ -typical projection

$$E(n, \rho, \delta) = \sum_{\mathbf{j} \in T(n, p, \delta)} |\mathbf{j}\rangle \langle \mathbf{j}|$$

Empirically verify the following identities

$$\begin{aligned} p^{\otimes n}(T(n, p, \delta)) &\geq 1 - p/\delta^2 \\ 2^{nH(\rho) - K\delta\sqrt{n}} &\leq \mu(T(n, p, \delta)) \leq 2^{nH(\rho) + K\delta\sqrt{n}}, \\ 2^{nH(\rho) - K\delta\sqrt{n}} E(n, \rho, \delta) &\leq \rho^{\otimes n} E(n, \rho, \delta) \leq 2^{nH(\rho) + K\delta\sqrt{n}} E(n, \rho, \delta) \end{aligned}$$

as operator inequalities,

$$Tr(E(n, \rho, \delta)) = \mu(T(n, p, \delta)),$$

$$2^{nH(\rho) - K\delta\sqrt{n}} \leq \mu(T(n, p, \delta)) \leq 2^{nH(\rho) + K\delta\sqrt{n}},$$

If F is any orthogonal projection with $Tr(\rho^{\otimes n} F) > 1 - \epsilon$, then

$$\begin{aligned} 1 - \epsilon < Tr(\rho^{\otimes n} F) &\leq Tr(\rho^{\otimes n} E(n, \rho, \delta) F) + Tr(\rho^{\otimes n} (1 - E(n, \rho, \delta))) \\ &\leq 2^{-nH(\rho) + K\delta\sqrt{n}} Tr(F) + 2^{-nH(\rho) + K\delta\sqrt{n}} \end{aligned}$$

Thus,

$$Tr(F) \geq (1 - \epsilon) \cdot 2^{nH(\rho)} - 2^{K\delta\sqrt{n}}$$

so that writing F_n in place of F , we get

$$\liminf_{n \rightarrow \infty} n^{-1} \cdot \log(Tr(F_n)) \geq H(\rho)$$

This shows that we cannot achieve a compression rate better than $H(\rho)$ and any compression rate smaller than $H(\rho)$ is achievable if we assume that the compression should involve arbitrary small error of state recovery probability.

[3] Verification of the converse to Shannon's noisy coding theorem for classical channels.

$\nu_x(y), x \in A, y \in B$ are the channel transition probabilities. For $u \in A^n$ and $v \in B^n$, we define

$$\nu_u(v) = \prod_{k=1}^n \nu_{u(k)}(v(k))$$

ie the channel is a DMC (Discrete memoryless channel). Suppose that there exist codewords, $u_1, \dots, u_M \in A^n$ with $P_{u_k} = P$ for all k and disjoint sets $D_1, \dots, D_K \in B^n$ so that $\nu_{u_k}(D_k) > 1 - \epsilon \forall k$. Let For any $u \in A^n$, let $T(n, \nu_u, \delta)$ be the set of all sequences $v \in B^n$ with the property that for all $x \in A, y \in B$, we have

$$|N(y|v_x) - \nu_x(y)N(x|u)| < \delta \sqrt{N(x|u)\nu_x(y)(1 - \nu_x(y))}$$

This condition is equivalent to requiring that

$$v_x \in T(N(x|u), \nu_x, \delta) \forall x \in A$$

Then we have for $v \in T(n, \nu_u, \delta)$, the inequalities

$$N(x|u)\nu_x(y) - \delta \sqrt{N(x|u)\nu_x(y)(1 - \nu_x(y))} \leq N(y|v_x) \leq N(x|u)\nu_x(y) + \delta \sqrt{N(x|u)\nu_x(y)(1 - \nu_x(y))}$$

$\forall x \in A$. Note that for such v ,

$$\Pi_{y \in B} \nu_x(y)^{N(y|v_x)} = \nu_x^{\otimes N(x|u)}(v_x)$$

and

$$\Pi_{x \in A} \nu_x^{\otimes N(x|u)}(v_x) = \nu_u(v)$$

Thus, it follows from the above equation that for $v \in T(n, \nu_u, \delta)$, we have

$$2^{-\sum_{x \in A} N(x|u)H(\nu_x) - K\delta\sqrt{n}} \leq \nu_u(v) \leq 2^{-\sum_{x \in A} N(x|u)H(\nu_x) + K\delta\sqrt{n}}$$

where

$$H(\nu_x) = - \sum_{y \in A} \nu_x(y) \cdot \log(\nu_x(y))$$

Writing

$$H(Y|X) = \sum_{x \in A} P(x) H(\nu_x)$$

we get

$$2^{-nH(Y|X) - K\delta\sqrt{n}} \leq \nu_u(v) \leq 2^{-nH(Y|X) + K\delta\sqrt{n}}, v \in T(n, \nu_u, \delta), P_u = P$$

Now we note that

$$\begin{aligned} \nu_u(T(n, \nu_u, \delta)) &= \Pi_{x \in A} \nu_x^{\otimes N(x|u)}(T(N(x|u), \nu_x, \delta)) \\ &\geq (1 - 1/\delta^2)^a \geq (1 - a/\delta^2) \end{aligned}$$

by Chebyshev's inequality. Thus by the union bound,

$$\begin{aligned} \nu_{u_k}(D_k \cap T(n, \nu_{u_k}, \delta)) \\ \geq 1 - \epsilon - a/\delta^2 \end{aligned}$$

Let

$$q(y) = \sum_{x \in A} P(x) \nu_x(y), y \in B$$

It is not hard to prove that if $v \in T(n, \nu_u, \delta)$ with $P_u = P$, then

$$v \in T(n, q, \delta\sqrt{a})$$

(we shall prove this at the end). Thus,

$$T(n, \nu_{u_k}, \delta) \subset T(n, q, \delta\sqrt{a}), k = 1, 2, \dots, M$$

We thus have

$$\begin{aligned} 1 - \epsilon - a/\delta^2 &\leq \nu_{u_k}(D_k \cap T(n, \nu_{u_k}, \delta)) \\ &\leq 2^{-nH(Y|X) + K\delta\sqrt{n}} \mu(D_k \cap T(n, \nu_{u_k}, \delta)) \\ &\leq 2^{-nH(Y|X) + K\delta\sqrt{n}} \mu(D_k \cap T(n, q, \delta\sqrt{a})) \end{aligned}$$

and hence summing over k using the disjointness of the D_k gives us

$$\begin{aligned} M(1 - \epsilon - a/\delta^2) &\leq 2^{-nH(Y|X)+K\delta\sqrt{n}} \mu(T(n, q, \delta\sqrt{a})) \\ &\leq 2^{-nH(Y|X)+nH(Y)+K_1\delta\sqrt{n}} \end{aligned}$$

which proves the converse of Shannon's noisy coding theorem.

[4] Verification of the direct part of Shannon's coding theorem. Let

$$q(y) = \sum_{x \in A} P(x) \nu_x(y), y \in B$$

Define

$$V_\delta = \{(x, y) \in A \times B : |\log(\nu_x(y)/q(y)) - I(X, Y)| < \delta\}$$

and

$$V_\delta^x = \{y : (x, y) \in V_\delta\}$$

We note that if the input source r.v. X has the distribution $P()$ so that the output Y has the distribution q when the channel transition probability distribution is $\nu_x(y)$, then

$$\mathbb{E} \log(\nu_X(Y)/q(Y)) = \sum_{x,y} P(x) \nu_x(y) \log(\nu_x(y)/q(y)) = H(Y) - H(Y|X) = I(X, Y)$$

By Chebyshev's inequality, we have

$$\omega(V_\delta) > 1 - \alpha/\delta^2$$

where

$$\alpha = \text{Var}(\log(\nu_X(Y)/q(Y)))$$

where

$$\omega(x, y) = P(x) \nu_x(y), (x, y) \in A \times B$$

Now let $\epsilon > 0$ and choose distinct $x_1, \dots, x_M \in A$ and $V_1, \dots, V_M \subset B$ such that M is maximal subject to the constraints the V'_j 's are disjoint, $\nu_{x_j}(V'_j) > 1 - \epsilon$, $j = 1, 2, \dots, M$. We wish to then derive a lower bound on M . To this end, choose $x_1 \in A$ such that $\nu_{x_1}(V_{\delta}^{x_1}) > 1 - \epsilon$. This is possible since

$$\begin{aligned} 1 - \epsilon &< 1 - \alpha/\delta^2 \leq \omega(V_\delta) = \sum_{x \in A} P(x) \nu_x(V_\delta^x) \\ &\leq \max_{x \in A} \nu_x(V_\delta^x) \end{aligned}$$

and we are assuming that

$$\epsilon > \alpha/\delta^2$$

Define

$$V_1 = V_\delta^{x_1}$$

Now if $\nu_x(V_\delta^x \cap V_1^c) \leq 1 - \epsilon$ for all $x \in A$, we stop. Otherwise, we choose $x_2 \in A$ such that $\nu_{x_2}(V_\delta^{x_2} \cap V_1^c) > 1 - \epsilon$ and define $V_2 = V_\delta^{x_2} \cap V_1^c$. If $\nu_x(V_\delta^x \cap V_1^c \cap V_2^c) \leq 1 - \epsilon$ for all $x \in A$, we stop else choose $x_3 \in A$ so that $\nu_{x_3}(V_\delta^{x_3} \cap V_1^c \cap V_2^c) > 1 - \epsilon$. This process must obviously terminate at some point since our alphabets are finite. In other words, we have selected a finite positive integer M , distinct elements $x_1, \dots, x_M \in A$ and disjoint sets $V_1, \dots, V_M \subset B$ such that $\nu_{x_j}(V_j) > 1 - \epsilon, j = 1, 2, \dots, M$. It is therefore clear that for all $x \in A$, we have

$$\nu_x(V_\delta^x \cap (\bigcup_{k=1}^M V_k)^c) \leq 1 - \epsilon$$

By Chebyshev' inequality, we also have

$$\begin{aligned} 1 - \alpha/\delta^2 &= \omega(V_\delta) = \sum_{x \in A} P(x) \nu_x(V_\delta^x) \\ &= \sum_{x \in A} P(x) \nu_x(V_\delta^x \cap (\bigcup_{k=1}^M V_k)^c) \\ &\quad + \sum_{x \in A} P(x) \nu_x(V_\delta^c \cap \bigcup_{k=1}^M V_k) \\ &\leq 1 - \epsilon + q\left(\bigcup_{k=1}^M q(V_k)\right) \\ &= 1 - \epsilon + \sum_{k=1}^M q(V_k) \quad \text{--- (1)} \end{aligned}$$

because the V'_k 's are pairwise disjoint. Now, $y \in V_x^\delta$ implies

$$I(X, Y) - \delta \leq \log(\nu_x(y)/q(y)) \leq I(X, Y) + \delta$$

implies,

$$2^{I(X, Y) - \delta} q(y) \leq \nu_x(y) \leq 2^{I(X, Y) + \delta} q(y)$$

implies

$$2^{I(X, Y) - \delta} q(V_x^\delta) \leq \nu_x(V_\delta^x) \leq 2^{I(X, Y) + \delta} q(V_x^\delta)$$

In particular since $V_k \subset V_\delta^{x_k}$, it follows from this inequality that

$$q(V_k) \leq 2^{-I(X, Y) + \delta}$$

since

$$\nu_x(V_\delta^x) \leq 1$$

Thus, we get from (1),

$$1 - \alpha/\delta^2 \leq 1 - \epsilon + M \cdot 2^{-I(X, Y) + \delta}$$

and hence the required lower bound on M is derived:

$$M \geq (\epsilon - \alpha/\delta^2) \cdot 2^{I(X, Y) - \delta}$$

6.20.6 MATLAB exercises involving convergence analysis of the LMS algorithm

[1] Generate a sequence $\mathbf{X}(n), n = 1, 2, \dots$ of iid $N(\mathbf{0}, \mathbf{R})$ random vectors. Generate a desired signal $d[n] = \mathbf{h}^T \mathbf{X}[n] + w[n]$ where $w[n], n = 1, 2, \dots$ of iid $N(0, \sigma_w^2)$ random variables. Now, estimate \mathbf{h} using the LMS algorithm as

$$\begin{aligned}\mathbf{h}[n+1] &= \mathbf{h}[n] - \mu \nabla_{\mathbf{h}} (d[n] - \mathbf{h}[n]^T \mathbf{X}[n])^2 \\ &= (\mathbf{I} - 2\mu \mathbf{X}[n] \mathbf{X}[n]^T) \mathbf{h}[n] + \mu d[n] \mathbf{X}[n]\end{aligned}$$

Compute the asymptotic mean and covariance of the filter weight vector $\mathbf{h}[n]$ and compare with theoretical expressions.

[2] Use the LMS algorithm to estimate the parameters of a second order nonlinear Volterra system defined by

$$y[n] = \mathbf{h}^T \mathbf{X}[n] + \mathbf{g}^T (\mathbf{X}[n] \otimes \mathbf{X}[n]) + w[n]$$

where

$$\mathbf{X}[n] = [x[n], x[n-1], \dots, x[n-p]]^T$$

with $x[n]$ being WGN of variance σ_x^2 and $w[n]$ also being WGN of variance σ_w^2 . The LMS algorithm is

$$\begin{aligned}\mathbf{h}[n+1] &= \mathbf{h}[n] - \mu \nabla_{\mathbf{h}} (y[n] - \mathbf{h}[n]^T \mathbf{X}[n] - \mathbf{g}[n]^T (\mathbf{X}[n] \otimes \mathbf{X}[n]))^2 \\ &= \mathbf{h}[n] + 2\mu e[n] \mathbf{X}[n], \\ \mathbf{g}[n+1] &= \mathbf{g}[n] - \mu \nabla_{\mathbf{g}} (y[n] - \mathbf{h}^T \mathbf{X}[n] - \mathbf{g}^T (\mathbf{X}[n] \otimes \mathbf{X}[n]))^2 \\ &= \mathbf{g}[n] + 2\mu e[n] \mathbf{X}[n] \otimes \mathbf{X}[n]\end{aligned}$$

with

$$e[n] = y[n] - \mathbf{h}[n]^T \mathbf{X}[n] - \mathbf{g}[n]^T (\mathbf{X}[n] \otimes \mathbf{X}[n])$$

Make an approximate statistical performance analysis of this algorithm.

6.21 MATLAB problems on root space decomposition of a Lie algebra

[1] Consider the Lie algebra $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{C})$. This consists of all $n \times n$ complex matrices having trace zero. Let E_{ij} denote the $n \times n$ matrix with a zero at all positions except at the $(i, j)^{th}$ position where it is a one:

$$(E_{ij})_{p,q} = \delta_{ip} \delta_{jq}$$

Consider the vector subspace \mathfrak{h} of $\mathfrak{sl}(n, \mathbb{C})$ consisting of all diagonal matrices having zero trace. For $i \neq j$, consider the vector subspace \mathfrak{g}_{ij} spanned by the single element E_{ij} . Show that

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{i \neq j} \mathfrak{g}_{ij}$$

and that this is precisely a root space decomposition of \mathfrak{g} with \mathfrak{h} as the Cartan subalgebra and $\mathfrak{g}_{ij}, i \neq j$ as the one dimensional root subspaces. For $i < j$, let α_{ij} denote the element of \mathfrak{h}^* defined by the property

$$\alpha_{ij}(H) = H_{ii} - H_{jj}, H \in \mathfrak{h}$$

Show that the set $\pm\alpha_{ij}, i < j$ exhaust all the roots of \mathfrak{g} relative to the Cartan algebra \mathfrak{h} . Show that we can take the set

$$\Delta_+ = \{\alpha_{ij}, i < j\}$$

as the positive roots and

$$\Delta_- = -\Delta_+$$

as the negative roots. Show that for this definition of positive roots, the set

$$S = \{\alpha_{i,j+1}, 1 \leq i \leq n-1\}$$

forms a set of simple roots. Draw the Dynkin diagram for this set of simple roots.

[2] Consider the Lie-algebra $\mathfrak{g} = \mathfrak{so}(2n, \mathbb{C})$ consisting of all $2n \times 2n$ complex skew-symmetric matrices having trace zero. Show that this indeed defines a Lie algebra. Let the $2n \times 2n$ matrices E_{ij} be defined as earlier. Define

$$F_{ij} = E_{ij} - E_{ji}, 1 \leq i < j \leq 2n$$

Show that $F_{ij} \in \mathfrak{g}$ and that for $i < j, k < m$, we have

$$\begin{aligned} [F_{ij}, F_{km}] &= [E_{ij}, E_{km}] + [E_{ji}, E_{km}] - [E_{ij}, E_{mk}] - [E_{ji}, E_{mk}] \\ &= \delta_{jk}E_{im} - \delta_{im}E_{kj} + \delta_{ik}E_{jm} - \delta_{jm}E_{ki} \\ &\quad - \delta_{jm}E_{ik} + \delta_{ik}E_{mj} - \delta_{ik}E_{jm} + \delta_{jm}E_{ki} \\ &= \delta_{jk}E_{im} - \delta_{im}E_{kj} - \delta_{jm}E_{ik} + \delta_{ik}E_{mj} \end{aligned}$$

Interpret this result as a generalization of the Lie algebra commutation relations for 3×3 rotation matrices. Write down the root space decomposition for \mathfrak{g} and draw its Dynkin diagram.

Remarks: The matrices $F_{i,i+1}, i = 1, 2, \dots, n-1$ commute with each other and the above formula implies for $k < m$,

$$\begin{aligned} [F_{i,i+1}, F_{km}] &= \delta(i+1, k)E_{im} - \delta(i, m)E_{k,i+1} \\ &\quad - \delta(i+1, m)E_{ik} + \delta(i, k)E_{m,i+1} \end{aligned}$$

If $k > i+1$, this equals zero. If $m < i$, then again we get zero. If $k = i+1$, we get

$$E_{im}$$

while if $m = i$, we get $-E_{k,i+1}$.

[3] Let \mathfrak{g} be a complex semisimple Lie algebra with a Cartan subalgebra \mathfrak{h} . Let α be a root and let X_α, Y_α be root vectors with eigenvalues $\alpha(H)$ and $-\alpha(H)$ for $H \in \mathfrak{h}$, ie,

$$[H, X_\alpha] = \alpha(H)X_\alpha, [H, Y_\alpha] = -\alpha(H)Y_\alpha$$

For any $H \in \mathfrak{h}$, by the Jacobi identity,

$$[H, [X_\alpha, Y_\alpha]] + [X_\alpha, [Y_\alpha, H]] + [Y_\alpha, [H, X_\alpha]] = 0$$

This gives

$$[H, [X_\alpha, Y_\alpha]] + \alpha(H)[X_\alpha, Y_\alpha] + \alpha(H)[Y_\alpha, X_\alpha] = 0$$

which simplifies to

$$[H, [X_\alpha, Y_\alpha]] = 0$$

Since \mathfrak{h} is maximal Abelian, it follows that

$$[X_\alpha, Y_\alpha] \in \mathfrak{h}$$

We write

$$[X_\alpha, Y_\alpha] = c(\alpha)H_\alpha$$

For any $H \in \mathfrak{h}$, we have

$$\begin{aligned} B(H, [X_\alpha, Y_\alpha]) &= B([H, X_\alpha], Y_\alpha) \\ &= \alpha(H)B(X_\alpha, Y_\alpha) \end{aligned}$$

or equivalently,

$$c(\alpha)B(H, H_\alpha) = \alpha(H).B(X_\alpha, Y_\alpha)$$

We choose $c(\alpha) = B(X_\alpha, Y_\alpha)$ (this may be done by scaling H_α appropriately) and hence $H_\alpha \in \mathfrak{h}$ satisfies

$$B(H, H_\alpha) = \alpha(H), H \in \mathfrak{h}$$

Now consider the Lie-algebra $\mathfrak{sl}_\alpha(2, \mathbb{C})$ generated by $\{H_\alpha, X_\alpha, Y_\alpha\}$. By the above discussion, they satisfy the commutation relations

$$[H_\alpha, X_\alpha] = \alpha(H_\alpha)X_\alpha,$$

$$[H_\alpha, Y_\alpha] = -\alpha(H_\alpha)Y_\alpha,$$

$$[X_\alpha, Y_\alpha] = B(X_\alpha, Y_\alpha)H_\alpha$$

We may scale X_α and Y_α so that

$$B(X_\alpha, Y_\alpha) = 1$$

and then defining

$$H_\alpha = [X_\alpha, Y_\alpha]$$

we get

$$c(\alpha) = 1$$

so that

$$[H_\alpha, X_\alpha] = \alpha(H_\alpha)X_\alpha,$$

$$[H_\alpha, Y_\alpha] = -\alpha(H_\alpha)Y_\alpha,$$

$$[X_\alpha, Y_\alpha] = H_\alpha$$

Then,

$$\alpha(H_\alpha) = B(H_\alpha, H_\alpha) = B(H_\alpha, [X_\alpha, Y_\alpha]) =$$

$$B([H_\alpha, X_\alpha], Y_\alpha) = \alpha(H_\alpha)B(X_\alpha, Y_\alpha) = \alpha(H_\alpha)$$

which verifies the consistency of our commutation relations. Note that to obtain the standard commutation relations for $\mathfrak{sl}(2, \mathbb{C})$, namely

$$[\hat{H}_\alpha, X_\alpha] = 2X_\alpha,$$

$$[\hat{H}_\alpha, Y_\alpha] = -2Y_\alpha,$$

$$[X_\alpha, Y_\alpha] = \hat{H}_\alpha$$

we must scale H_α , ie, we must replace H_α by $\hat{H}_\alpha = 2H_\alpha/\alpha(H_\alpha)$ in which case, we would have

$$B(\hat{H}_\alpha, H) = 2\alpha(H)/\alpha(H_\alpha)$$

and in particular,

$$B(\hat{H}_\alpha, H_\alpha) = 2$$

provided that after scaling H_α , we scale X_α and Y_α appropriately (this scaling would not affect the first two commutation relations which are eigen relations).

Now consider the above commutation relations for the generators of $\mathfrak{sl}_\alpha(2, \mathbb{C})$. Let β be another positive root. If $X_{\beta+r\alpha}$ is a weight vector for the adjoint representation of $\mathfrak{sl}_\alpha(2, \mathbb{C})$ where r is an integer, then its weight is $\beta + r\alpha$, ie,

$$[H_\alpha, X_{\beta+r\alpha}] = (\beta + r\alpha)(H_\alpha)X_{\beta+r\alpha}$$

$$= (\beta(H_\alpha) + r\alpha(H_\alpha))X_{\beta+r\alpha}$$

If then $[X_\alpha, X_{\beta+r\alpha}]$ is non-zero, then it is a weight vector $X_{\beta+(r+1)\alpha}$ of $\mathfrak{sl}_\alpha(2, \mathbb{C})$ with weight $\beta + (r+1)\alpha$. likewise, $[Y_\alpha, X_{\beta+r\alpha}] = X_{\beta+(r-1)\alpha}$ if non-zero, is a weight vector of $\mathfrak{sl}_\alpha(2, \mathbb{C})$ with weight $\beta + (r-1)\alpha$. It follows from the finite dimensionality of \mathfrak{g} the there exist non-negative integers p, q such that $[Y_\alpha, X_{\beta-q\alpha}] = 0$, $[X_\alpha, Y_{\beta+r\alpha}] = 0$ but $X_{\beta+r\alpha}, -q \leq r \leq p$ are all non-zero. Consider then the $p+q+1$ dimensional vector space

$$V_{\beta,\alpha} = \text{span}\{X_{\beta+r\alpha} : -q \leq r \leq p\}$$

This vector space is invariant under $ad(X_\alpha), ad(Y_\alpha)$ and hence also under

$$ad(H_\alpha) = ad[X_\alpha, Y_\alpha] = [ad(X_\alpha), ad(Y_\alpha)]$$

Since $\text{Tr}([A, B]) = 0$ for finite matrices, it follows that

$$\text{Tr}(ad(H_\alpha)|_{V_{\beta,\alpha}}) = 0$$

and hence

$$\sum_{r=-q}^p (\beta + r\alpha)(H_\alpha) = 0$$

or

$$\beta(H_\alpha)(p+q+1) + \alpha(H_\alpha) \sum_{r=-q}^p r = 0$$

or

$$(p+q+1)\beta(H_\alpha) + (p-q)(p+q+1)\alpha(H_\alpha)/2 = 0$$

or

$$2\beta(H_\alpha)/\alpha(H_\alpha) = q - p$$

In other words, for any two roots α, β of the semisimple Lie algebra \mathfrak{g} , the quantity

$$2\beta(H_\alpha)/\alpha(H_\alpha) \in \mathbb{Z}$$

These integers are called the Cartan integers when α, β are simple and positive. We note that

$$\alpha(H_\beta) = B(H_\alpha, H_\beta)$$

and hence if $\{\alpha_1, \dots, \alpha_l\}$ are the simple positive roots of \mathfrak{g} , the Cartan integers

$$a_{ij} = 2 < \alpha_i, \alpha_j > / < \alpha_j, \alpha_j >$$

are all integers. The simple positive roots are linearly independent and any positive root is a non-negative integer linear combination of the simple positive roots while any negative root is a non-positive integer linear combination of the simple positive roots. Moreover owing to the non-singularity of the invariant bilinear form B on \mathfrak{h} , it follows that the map $(\alpha, \beta) \rightarrow B(H_\alpha, H_\beta) = < \alpha, \beta >$ defines an inner product on the vector space spanned by the simple roots, ie on the space \mathfrak{h}^* . Thus, we get using the Schwarz inequality,

$$0 \leq a_{\alpha,\beta} a_{\beta,\alpha} = \frac{4 < \alpha, \beta >^2}{< \alpha, \alpha > < \beta, \beta >} < 4$$

whenever α, β are distinct roots. This inequality provides the key route to Cartan's classification of the simple Lie algebras.

[1] Group representation theory applied to determination of the camera orientation.

Initially, let the camera screen be parallel to the xy plane so that its unit normal is \hat{z} . Let the camera lens be focussed so that it can take pictures on its screen of a scene $f(x, y, z) = f(\mathbf{r})$ at a distance d from it. Let a rotation

$$R = R(\phi, \theta, \psi) = R_z(\phi)R_x(\theta)R_z(\psi)$$

be applied to the camera screen, so that after this has been done, the plane of the screen is spanned by the orthonormal vectors $R\hat{x}$ and $R\hat{y}$ and the normal to the screen is $R\hat{z}$. The image field recorded by the camera screen with ξ_x denoting the coordinate distance along the $R\hat{x}$ direction and ξ_y the coordinate distance along the $R\hat{y}$ direction is given by

$$g(\xi_x, \xi_y) = f(d.R\hat{z} + \xi_x.R\hat{x} + \xi_y.R\hat{y}) = f(R\chi) = g(\chi)$$

where we write

$$\chi = [\xi_x, \xi_y, d]^T$$

and then for any $S \in SO(3)$, we have

$$g(S\chi) = f(RS\chi)$$

and then if π is an irreducible representation of $SO(3)$, we get

$$\begin{aligned} \int g(S\chi)\pi(S)^*d\mu(S) &= \int_{SO(3)} f(RS\chi)\pi(S)^*d\mu(S) \\ &= \int f(S\chi)\pi(R^{-1}S)^*d\mu(S) = (\int f(S\chi)\pi(S)^*d\mu(S))\pi(R) \end{aligned}$$

If we apply this formula to several scaled versions $s\chi, s > 0$ of the position vector χ of the image field relative to the camera centre, we then get

$$f_s(\chi) = f(s\chi), g_s(\chi) = g(s\chi) = f(s.R\chi)$$

so that

$$\begin{aligned} \int_{SO(3)} g_s(S\chi)\pi(S)^*d\mu(S) &= \int f(s.RS\chi)\pi(S)^*d\mu(S) = \int_{SO(3)} f(s.S\chi)\pi(R^{-1}S)^*d\mu(S) \\ &= (\int f(s.S\chi)\pi(S)^*d\mu(S))\pi(R) \end{aligned}$$

and hence by calculating the matrices

$$\int_{SO(3)} g_s(S\chi)\pi(S)^*d\mu(S) \text{ and } (\int f(s.S\chi)\pi(S)^*d\mu(S))$$

for different values of s , we can estimate $\pi(R)$ by solving linear equations. In particular, taking the standard 3-D representation $\pi(R) = R$ of $SO(3)$, we can obtain R easily by solving a system of linear equations. So the whole problem of estimating the camera rotation amounts to determining the Haar measure μ on $SO(3)$. This can be done using a standard theorem in the theory of differential forms: Let $\omega_1(g), \dots, \omega_n(g)$ be a set of n linearly independent left invariant one forms on G . then, a left invariant Haar measure on G is given by

$$d\mu = \omega_1 \wedge \dots \wedge \omega_n$$

Thus, if we choose a coordinates $g = (g_1, \dots, g_n)$ for G , and write

$$\omega_k(g) = \omega_{km}(g)dg_m$$

then

$$d\mu(g_1, \dots, g_n) = \det[(\omega_{km}(g))].dg_1 \dots dg_n$$

Equivalently, by duality, it follows that if $X_1(g), \dots, X_n(g)$ are linearly independent left invariant vector fields on G , then a left invariant Haar measure on G is given by

$$d\mu(g_1, \dots, g_n) = \det[(X_{km}(g))]^{-1}dg_1 \dots dg_n$$

where

$$X_k(g) = X_{km}(g)\partial/\partial g_m$$

We now calculate a set of three linearly independent left invariant vector fields on $SO(3)$ as a linear combination of

$$\partial/\partial\phi, \partial/\partial\theta, \partial/\partial\psi$$

where ϕ, θ, ψ are the three Euler angles. Using the above theorem, this would then enable us to express the Haar measure on $SO(3)$ in the form

$$M(\phi, \theta, \psi).d\phi.d\theta.d\psi$$

6.22 Cartan's criterion for semisimplicity of a Lie algebra

Let \mathfrak{g} be a Lie algebra and $B(\cdot, \cdot)$ its bilinear symmetric invariant form, ie,

$$B(X, Y) = \text{Tr}(ad(X).ad(Y)), X, Y \in \mathfrak{g}$$

We say that \mathfrak{g} is semisimple iff B is a non-degenerate, ie, $B(X, Y) = 0 \forall Y \in \mathfrak{g}$ implies $X = 0$. We shall now show that semisimplicity of \mathfrak{g} implies that if \mathfrak{a} is any ideal in \mathfrak{g} , then \mathfrak{a}^\perp is an ideal in \mathfrak{g} and that we have the direct sum decomposition

$$\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$$

That \mathfrak{a}^\perp is also an ideal in \mathfrak{g} does not require \mathfrak{g} to be semisimple. It simply follows from the identity

$$B([X, Y], Z) = B(X, [Y, Z]), X, Y, Z \in \mathfrak{g}$$

Indeed, if $X \in \mathfrak{a}^\perp$ and $Z \in \mathfrak{a}, Y \in \mathfrak{g}$, then

$$B([X, Y], Z) = B(X, [Y, Z]) = 0$$

since \mathfrak{a} implies that $[Y, Z] \in \mathfrak{a}$. Since $Z \in \mathfrak{a}$ is arbitrary $[X, Y] \in \mathfrak{a}^\perp$ proving that \mathfrak{a}^\perp is an ideal in \mathfrak{g} . It is easy to see that

$$\dim \mathfrak{a} + \dim \mathfrak{a}^\perp = \dim \mathfrak{g}$$

So it remains to prove that

$$\mathfrak{a} \cap \mathfrak{a}^\perp = \{0\}$$

First we observe that $\mathfrak{a} \cap \mathfrak{a}^\perp$ is an Abelian ideal in \mathfrak{g} . That it is an ideal follows immediately from the obvious fact that the intersection of ideals is an ideal. That it is Abelian follows from the fact that if $X, Y \in \mathfrak{a} \cap \mathfrak{a}^\perp$ and $\mathfrak{Z} \in \mathfrak{g}$ is arbitrary, then

$$B([X, Y], Z) = B(X, [Y, Z]) = 0$$

since $X \in \mathfrak{a}$ and $[Y, Z] \in \mathfrak{a}^\perp$ because \mathfrak{a}^\perp is an ideal and $Y \in \mathfrak{a}^\perp$. Thus, $[X, Y]$ is orthogonal to \mathfrak{g} and since $B(., .)$ is assumed to be non-degenerate, it follows that $[X, Y] = 0$ proving the Abelian property of $\mathfrak{a} \cap \mathfrak{a}^\perp$.

Now let $X \in \mathfrak{a} \cap \mathfrak{a}^\perp$. Then let $Z \in \mathfrak{g}$. Let \mathfrak{b} be any complementary subspace of $\mathfrak{a} \cap \mathfrak{a}^\perp$. Then, $ad(Z)ad(X)$ maps $\mathfrak{a} \cap \mathfrak{a}^\perp$ into $\{0\}$ since $[X, \mathfrak{a} \cap \mathfrak{a}^\perp] = 0$ by the Abelian property of $\mathfrak{a} \cap \mathfrak{a}^\perp$. Further, $ad(Z).ad(X)$ maps \mathfrak{b} also into $\mathfrak{a} \cap \mathfrak{a}^\perp$ in view of the fact that $\mathfrak{a} \cap \mathfrak{a}^\perp$ is an ideal in \mathfrak{g} . It follows that

$$B(Z, X) = Tr(ad(Z).ad(X)) = 0$$

and hence by the non-degeneracy of B on \mathfrak{g} , it follows that $X = 0$, ie, $\mathfrak{a} \cap \mathfrak{a}^\perp = \{0\}$ thereby establishing the required claim.

6.23 Problems in linear algebra

[1] If \mathcal{F} is a commuting family of linear operators on a finite dimensional complex inner product space, that show that this family is simultaneously triangulable w.r.t an ONB.

[2] If \mathcal{F} is a commuting family of diagonalable linear operators on a finite dimensional complex vector space, then show that \mathcal{F} is simultaneously diagonalable.

[3] If \mathcal{F} is a commuting family of normal operators in an inner product space, then show that this family is simultaneously diagonalable w.r.t. an ONB.

[4] Let W be a proper T -invariant subspace of a complex finite dimensional vector space V and let $\alpha \notin W$. Let $s(\alpha, T, W)$ denote the T -conductor of α into W , ie, the set of all polynomials f for which $f(T)\alpha \in W$. Show that $s(\alpha, T, W)$ is an ideal in $\mathbb{C}[t]$ and that if p_0 denotes its monic generator, then $\deg p_0 > 0$ and therefore if c is a root of p_0 , then $p(t) = (t - c)q(t)$ where $q \in \mathbb{C}[t]$. Deduce that $\beta q(T)\alpha \notin W$ and that $(T - c)\beta \in W$. Use this argument combined with induction to show that every operator on V is triangulable.

[5] Discuss the notion of Cartan algebra for a complex semisimple Lie algebra and the associated root space decomposition. Specifically, let \mathfrak{g} be a any Lie algebra and for $X, Y \in \mathfrak{g}$, define

$$B(X, Y) = Tr(ad(X).ad(Y))$$

We shall say that \mathfrak{g} is a semisimple Lie algebra if B is non-degenerate. If this is so then show that if \mathfrak{a} is any ideal in \mathfrak{g} and \mathfrak{a}^\perp consists of all elements $Y \in \mathfrak{g}$ for which $B(X, Y) = 0 \forall X \in \mathfrak{a}$, then

$$\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$$

ie, every invariant subspace \mathfrak{a} for the family of operators $ad(\mathfrak{g})$ acting on \mathfrak{g} has a unique complementary invariant subspace that is orthogonal to \mathfrak{a} . Conversely, show that if every invariant subspace for $ad(\mathfrak{g})$ has a complementary invariant subspace (not necessary orthogonal), then \mathfrak{g} is semisimple. Now, let for any complex number c and $X \in \mathfrak{g}$, the subspace $\mathfrak{g}(X, c)$ be defined as the set of all $Y \in \mathfrak{g}$ for which $(ad(X) - c)^m(Y) = 0$ for some positive integer m . Show that the primary decomposition implies the existence of a finite set of distinct complex numbers c_1, \dots, c_r and positive integers m_1, \dots, m_r which are minimal such that

$$\mathfrak{g} = \bigoplus_{m=1}^r \mathfrak{g}(X, c_m)$$

Note that if $\mathfrak{g}(X, 0)$ is nonzero, then $ad(X)$ is nilpotent and vice-versa. Let $H_0 \in \mathfrak{g}$ be such that

$$\dim \mathfrak{g}(H_0, 0) = \min \dim_{X \in \mathfrak{g}} \mathfrak{g}(X, 0)$$

then show that if \mathfrak{g} is any Lie algebra, not necessarily semisimple, then $\mathfrak{h} = \mathfrak{g}(H_0, 0)$ is a nilpotent subalgebra of \mathfrak{g} . If further, \mathfrak{g} is assumed to be semisimple, then show that $\mathfrak{h} = \mathfrak{g}(H_0, 0)$ is a maximal Abelian subalgebra of \mathfrak{g} , and hence we have by the simultaneous diagonability theorem in Linear algebra, the root space decomposition:

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Delta} \mathfrak{g}_\alpha$$

where Δ is a finite subset of \mathfrak{h}^* such that

$$[H, X] = \alpha(H)X, \forall H \in \mathfrak{h}, X \in \mathfrak{g}_\alpha, \alpha \in \Delta$$

[6] Discuss the root-space decomposition of the complex classical Lie algebras

$$\mathfrak{sl}(n, \mathbb{C}), \mathfrak{so}(2n, \mathbb{C}), \mathfrak{so}(2n+1, \mathbb{C}), \mathfrak{sp}(2n, \mathbb{C})$$

and draw their Dynkin diagrams.

Remarks:

[a] $\mathfrak{sl}(n, \mathbb{C})$ consists of all $n \times n$ complex matrices having zero trace. The corresponding Lie group $SL(n, \mathbb{C})$ is the set of all $n \times n$ complex matrices having unit determinant.

[b] $\mathfrak{so}(n, \mathbb{C})$ is the set of all $n \times n$ complex skew-symmetric matrices having zero trace. The corresponding Lie group $SO(n, \mathbb{C})$ is the set of all $n \times n$ complex

orthogonal matrices having unit determinant. By orthogonality of X , we mean that $X^T = X^{-1}$.

[c] $\mathfrak{sp}(2n, \mathbb{C})$ is the set of all $2n \times 2n$ complex matrices X such that

$$X^T J + JX = 0$$

where

$$J = \begin{pmatrix} \mathbf{0}_n & \mathbf{I}_n \\ -\mathbf{I}_{n*} & \mathbf{0}_n \end{pmatrix}$$

The corresponding Lie group $Sp(n, \mathbb{C})$ is the set of all $n \times n$ complex matrices \mathbf{g} such that

$$\mathbf{g}^T J \mathbf{g} = \mathbf{I}_{2n}$$

This group appears in Hamiltonian mechanics.

6.24 Problems in non-linear filtering theory

[1] The magnetic vector potential and the electric scalar potential satisfy the 3-D wave equations with source:

$$(\partial_t^2 - c^2 \nabla^2) A(t, r) = \mu c^2 J(t, r),$$

$$(\partial_t^2 - c^2 \nabla^2) \Phi(t, r) = \mu c^4 \rho(t, r)$$

We assume that the sources J, ρ contain components that are white Gaussian w.r.t the time variable. Then we can express the above pde's in state variable form

$$\begin{aligned} \frac{d}{dt} \begin{pmatrix} A(t, r) \\ V(t, r) \\ \Phi(t, r) \\ \Psi(t, r) \end{pmatrix} &= \begin{pmatrix} V(t, r) \\ c^2 \nabla^2 A(t, r) \\ \psi(t, r) \\ c^2 \nabla^2 \Phi(t, r) \end{pmatrix} + \begin{pmatrix} 0 \\ \mu c^2 J(t, r) \\ 0 \\ \mu c^4 \rho(t, r) \end{pmatrix} \\ &= \begin{pmatrix} 0 & I & 0 & 0 \\ c^2 \nabla^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & I \\ 0 & 0 & c^2 \nabla^2 & 0 \end{pmatrix} \begin{pmatrix} A(t, r) \\ V(t, r) \\ \Phi(t, r) \\ \Psi(t, r) \end{pmatrix} \\ &\quad + S(t, r) + W(t, r) \end{aligned}$$

where $S(t, r)$ is the noiseless part of the source term and $W(t, r)$ is the noisy part. We now discretize space into pixels along the x, y, z directions and assume that x, y, z after this spatial discretization take values $1, 2, \dots, N$. We denote by e_x the $N \times 1$ vector with a one at the x^{th} position and zeros at all the other positions. Then, we define the vectors

$$e(xyz) = e_x \otimes e_y \otimes e_z \in \mathbb{R}^{N^3 \times 1},$$

$$\begin{aligned}
A_t &= \sum_{x,y,z=1}^N A(t, x, y, z) \otimes e(xyz) \in \mathbb{R}^{3N^3 \times 1}, \\
V_t &= \sum_{x,y,z=1}^N V(t, x, y, z) \otimes e(xyz) \in \mathbb{R}^{3N^3 \times 1}, \\
\Phi_t &= \sum_{x,y,z=1}^N \Phi(t, x, y, z) e(xyz) \in \mathbb{R}^{N^3 \times 1}, \\
\psi_t &= \sum_{x,y,z=1}^N \psi(t, x, y, z) e(xyz) \in \mathbb{R}^{N^3 \times 1}
\end{aligned}$$

The Laplacian operator ∇^2 when acting on a vector field like A after discretization becomes a $3N^3 \times 3N^3$ matrix D_3 while when acting on a scalar field like Φ becomes an $N^3 \times N^3$ matrix D . Specifically, these matrices are

$$D_3 = \Delta^{-2} \sum_{x,y,z=1}^N I_3 \otimes e(xyz) [(e_{x+1} + e_{x-1} - 2e_x) \otimes e_y \otimes e_z] +$$

$$(e_x \otimes (e_{y+1} + e_{y-1} - 2e_y) \otimes e_z) + (e_x \otimes e_y \otimes (e_{z+1} + e_{z-1} - 2e_z))^T$$

It is easily verified using properties of the Kronecker tensor product that

$$\begin{aligned}
D_3 A_t &= \Delta^{-2} \sum_{xyz} [A(t, x+1, y, z) + A(t, x-1, y, z) + A(t, x, y+1, z) + A(t, x, y-1, z) + A(t, x, y, z+1) \\
&\quad + A(t, x, y, z-1) - 6A(t, x, y, z)] \otimes e(xyz) \\
&= \sum_{xyz} \tilde{\nabla}^2 A(t, x, y, z) \otimes e(xyz)
\end{aligned}$$

where $\tilde{\nabla}^2$ is the discretized Laplacian acting on functions defined on \mathbb{Z}^3 . Likewise,

$$\begin{aligned}
D &= \Delta^{-2} \sum_{x,y,z=1}^N e(xyz) [(e_{x+1} + e_{x-1} - 2e_x) \otimes e_y \otimes e_z] + \\
&\quad (e_x \otimes (e_{y+1} + e_{y-1} - 2e_y) \otimes e_z) + (e_x \otimes e_y \otimes (e_{z+1} + e_{z-1} - 2e_z))^T
\end{aligned}$$

The gradient operator ∇ while acting on a scalar field after discretization becomes a $3N^3 \times N^3$ matrix defined by

$$G = \Delta^{-1} \sum_{x,y,z=1}^{N^3} \left(\begin{array}{c} ((e_{x+1} - e_x) \otimes e_y \otimes e_z)^T \\ (e_x \otimes (e_{y+1} - e_y) \otimes e_z)^T \\ (e_x \otimes e_y \otimes (e_{z+1} - e_z))^T \end{array} \right) \otimes e(xyz)$$

It is easily verified that

$$G\Phi_t = \sum_{xyz} \tilde{\nabla}\Phi(t, x, y, z) \otimes e(xyz)$$

where $\tilde{\nabla}$ is the discretized version of the gradient operator acting on scalar functions defined on \mathbb{Z}^3 :

$$\tilde{\nabla}\Phi(t, x, y, z) = \Delta^{-1} \begin{pmatrix} \Phi(t, x+1, y, z) - \Phi(t, x, y, z) \\ \Phi(t, x, y+1, z) - \Phi(t, x, y, z) \\ \Phi(t, x, y, z+1) - \Phi(t, x, y, z) \end{pmatrix}$$

The discretized version of the electric field is

$$E_t = \sum_{xyz} E(t, x, y, z) \otimes e(xyz) \in \mathbb{R}^{3N^3 \times 1}$$

and that of the magnetic field is

$$H_t = \sum_{xyz} H(t, x, y, z) \otimes e(xyz) \in \mathbb{R}^{3N^3 \times 1}$$

where $E(t, x, y, z)$ is the discretized version of $-\nabla\Phi - V$ and $H(t, x, y, z)$ is the discretized version of $\nabla \times A/\mu$. Thus,

$$E_t = -G\Phi_t - V_t,$$

and to determine H_t similarly as

$$H_t = CA_t$$

where C is the $3N^3 \times 3N^3$ discretized version of the curl operator. Specifically, let G_x, G_y, G_z denote discretized versions of the partial derivatives $\partial/\partial x, \partial/\partial y, \partial/\partial z$ respectively acting on scalar fields. Thus,

$$G_x = \Delta^{-1} \sum_{xyz} e(xyz)((e_{x+1} - e_x) \otimes e_y \otimes e_z)^T,$$

$$G_y = \Delta^{-1} \sum_{xyz} e(xyz)(e_x \otimes (e_{y+1} - e_y) \otimes e_z)^T,$$

$$G_z = \Delta^{-1} \sum_{xyz} e(xyz)(e_x \otimes e_y \otimes (e_{z+1} - e_z))^T$$

Also define the matrices P_x, P_y, P_z which extract out the x, y, z components respectively of a vector field, ie,

$$P_x A_t = A_{xt}, P_y A_t = A_{yt}, P_z A_t = A_{zt}$$

so that

$$P_x = \sum_{xyz} e(xyz)([1, 0, 0] \otimes e(xyz)^T) \in \mathbb{R}^{N^3 \times 3N^3}$$

$$P_y = \sum_{xyz} e(xyz)([0, 1, 0] \otimes e(xyz)^T) \in \mathbb{R}^{N^3 \times 3N^3}$$

$$P_z = \sum_{xyz} e(xyz)([0, 0, 1] \otimes e(xyz)^T) \in \mathbb{R}^{N^3 \times 3N^3}$$

Then,

$$CA_t = \sum_{xyz} (\tilde{\nabla} \times A(t, x, y, z)) \otimes e(xyz)$$

$$= [f_1 \otimes (G_y P_z - G_z P_y) + f_2 \otimes (G_z P_x - G_x P_z) + f_3 \otimes (G_x P_y - G_y P_x)] A_t$$

where

$$f_1 = [1, 0, 0]^T, f_2 = [0, 1, 0]^T, f_3 = [0, 0, 1]^T$$

Thus, we can write

$$C = [f_1 \otimes (G_y P_z - G_z P_y) + f_2 \otimes (G_z P_x - G_x P_z) + f_3 \otimes (G_x P_y - G_y P_x)] \in \mathbb{R}^{3N^3 \times 3N^3}$$

The state and measurement model with the measurement model being based on em field measurements at discrete set of $M << N^3$ pixels can now be formulated in the form

$$\begin{aligned} & \frac{d}{dt} \begin{pmatrix} A_t \\ V_t \\ \Phi_t \\ \psi_t \end{pmatrix} \\ &= \begin{pmatrix} 0 & I & 0 & 0 \\ c^2 D_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & I \\ 0 & 0 & c^2 D & 0 \end{pmatrix} \begin{pmatrix} A_t \\ V_t \\ \Phi_t \\ \psi_t \end{pmatrix} \\ & \quad + S_t + W_t \\ Z_t^E &= L^E E_t + \epsilon_{1t}, \\ Z_t^H &= L^H H_t + \epsilon_{2t}, \end{aligned}$$

where

$$E_t = -G\Phi_t - V_t,$$

$$H_t = \mu^{-1} CA_t$$

Accordingly, the EKF can be formulated. Here L^E and L^H have respectively M_1 and M_2 rows where $M_1 + M_2 = M$. This measurement model can be expressed as

$$\begin{aligned} Z_t &= \begin{pmatrix} Z_t^E \\ Z_t^H \end{pmatrix} = \\ & \begin{pmatrix} 0 & -L^E & -L^E G & 0 \\ \mu^{-1} L^H C & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} A_t \\ V_t \\ \Phi_t \\ \psi_t \end{pmatrix} + \epsilon_t \end{aligned}$$

State equations in the presence of inhomogeneous and anisotropic permittivity and permeability.

$$\begin{aligned} \operatorname{div} D &= \rho, \operatorname{div}(B) = 0, \\ \operatorname{curl} E &= -B_{,t}, \operatorname{curl} H = J + D_{,t} \\ D_a(t, r) &= \epsilon_0 E_a(t, r) + \delta \epsilon_0 \int \chi_{e1ab}(t_1, r) E_b(t - t_1, r) dt_1 \\ &+ \delta^2 \epsilon_0 \int \chi_{e2abc}(t_1, t_2, r) E_b(t_1, r) E_c(t_2, r) dt_1 dt_2 + \dots \\ B_a(t, r) &= \mu_0 H_a(t, r) + \delta \mu_0 \int \chi_{m1ab}(t_1, r) H_b(t - t_1, r) dt_1 \\ &+ \delta^2 \epsilon_0 \int \chi_{m2abc}(t_1, t_2, r) H_b(t_1, r) H_c(t_2, r) dt_1 dt_2 + \dots \end{aligned}$$

Such a model may for example be derived from a nonlinear differential equation of the form

$$\partial D_a(t, r) / \partial t = F_{ea}(t, r, E(t, r), D(t, r))$$

or more generally, by increasing the order of the differential equation as in

$$\partial^n D_a(t, r) / \partial t^n = F_{ea}(t, r, E(t, r), \partial^m D(t, r) / \partial t^m, m = 0, 1, \dots, n-1 | \theta)$$

or sometimes

$$D_a(t, r) = \epsilon_0 E_a(t, r) + P_a(t, r)$$

where

$$\partial^n P_a(t, r) / \partial t^n = F_{ea}(t, r, E(t, r), \partial^m P(t, r) / \partial t^m, m = 0, 1, \dots, n-1 | \theta)$$

and likewise for the magnetic field,

$$H(t, r) = B(t, r) / \mu_0 - M(t, r)$$

where $M(t, r)$ is the magnetization/magnetic moment per unit volume and it satisfies a differential equation of the form

$$\partial^n M_a(t, r) / \partial t^n = F_{ma}(t, r, H(t, r), \partial^m M(t, r) / \partial t^m, m = 0, 1, \dots, n-1 | \theta)$$

Example from polarization and hysteresis nonlinearities. An electron bound to an atom via a binding potential $U(r, \xi)$ at the point r with ξ denoting the displacement of the electron from the nucleus satisfies in an external electric field $E(t, r)$ the differential equation

$$M_0(r) \xi''(t) + \Gamma(r) \xi'(t) + \partial U(r, \xi) / \partial \xi = -eE(t, r)$$

If there are $N(r)$ atoms per unit volume, then the polarization vector field is given by

$$P(t, r) = -eN(r)\xi(t)$$

Thus, P satisfies the differential equation

$$M(r)P_{,tt}(t, r) + \Gamma(r)P_{,t}(t, r) + U_{,\xi}(r, -P(t, r)/eN(r)) = -eE(t, r)$$

and likewise the precession of a magnetic dipole in a magnetic field results in an angular momentum-torque relation

$$M_{,t}(t, r) = \gamma(r)M(t, r) \times B(t, r)$$

which can be solved to express $M(t, r)$ as a Volterra series in $B(t, r)$, or equivalently,

$$M_{,t}(t, r) = \gamma(r)\mu_0 M(t, r) \times H(t, r)$$

which gives M as a Volterra series in H . Now consider the Maxwell equations. We wish to derive non-linear wave equations for the magnetic vector and electric potentials and then apply nonlinear filtering theory to these non-linear state equations when noisy measurements on the electric and magnetic fields are taken at discrete spatial points. The equation

$$\operatorname{div} B = 0$$

implies

$$B = \nabla \times A$$

The equation

$$\nabla \times E + B_{,t} = 0$$

then implies

$$\nabla \times (E + A_{,t}) = 0$$

which implies

$$E = -\nabla\Phi - A_{,t}$$

The equation

$$\operatorname{div} D = \rho$$

means that

$$\epsilon_0 \operatorname{div} E + \operatorname{div} P = \rho$$

which implies

$$-\nabla^2\Phi - \operatorname{div} A_{,t} + \epsilon_0^{-1} \operatorname{div} P = \rho/\epsilon_0$$

and hence after applying the Lorentz gauge condition

$$\operatorname{div} A = -\epsilon_0 \mu_0 \Phi_{,t}$$

we get

$$\nabla^2\Phi - c^{-2}\Phi_{,tt} = -\rho/\epsilon_0 - \operatorname{div} P/\epsilon_0$$

As seen above, P satisfies a non-linear differential equation of the form

$$P_{,tt}(t, r) = F_e(t, r, P_{,t}(t, r), P(t, r), E(t, r)) = F_e(t, r, P_{,t}, -\nabla\Phi - A_{,t})$$

Likewise, for the magnetic field, we have using the equation

$$\nabla \times B = c^{-2} E_{,t} + \mu_0 (J + \nabla \times M + P_{,t})$$

and hence

$$\nabla(\nabla \cdot A) - \nabla^2 A = c^{-2} E_{,t} + \mu_0 (J + \nabla \times M + P_{,t})$$

or after applying the Lorentz gauge condition,

$$\nabla^2 A - c^{-2} A_{,tt} = -\mu_0 (J + \nabla \times M + P_{,t})$$

where the magnetization M satisfies the differential equation

$$M_{,t}(t, r) = \gamma(r) \mu_0 M(t, r) \times H(t, r)$$

By introducing the additional state variables

$$V = A_{,t}, \psi = \Phi_{,t}, \chi = P_{,t}$$

we get a set of nonlinear partial differential equations for the state vector $\xi = [A^T, V^T, \Phi, \psi, P^T, \chi^T, M^T]^T$ which is first order in time and hence we can discretize this system with respect to the spatial arguments to obtain a nonlinear state equation for ξ_t .

6.25 Spectral theorem for bounded self-adjoint operators in a Hilbert space—basic steps

Assume $T^* = T$, $\|T\| \leq 1$. We construct $|T|$ as the positive square root of T^2 using an iterative process stated at the end of this paragraph. Now define for and self-adjoint operator X , its positive and negative parts as

$$X_+ = (|X| + X)/2, X_- = (|X| - X)/2$$

Then,

$$X_+, X_- \geq 0, X = X_+ - X_-, |X| = X_+ X_-$$

For $\lambda \in \mathbb{R}$ define $E(\lambda)$ to be the orthogonal projection onto $\mathcal{N}((T - \lambda)_+)$. Then for $\lambda \leq \mu$, we have

$$E(\lambda) \leq E(\mu)$$

i.e., $\{E(\lambda), \lambda \in \mathbb{R}\}$ is an increasing family of orthogonal projections. In fact, we have for $\lambda \leq \mu$,

$$(T - \lambda)_+ \geq T - \lambda \geq T - \mu$$

and since

$$|T - \lambda| \geq T - \lambda$$

Therefore, we have

$$(T - \lambda)_+ \geq (1 - E(\mu))(T - \lambda)_+(1 - E(\mu)) \geq (1 - E(\mu))(T - \mu)(1 - E(\mu)) = (T - \mu)_+$$

Remark:

$$\mathcal{R}(1 - E(\mu)) = \mathcal{N}(T - \mu)_+^\perp$$

But

$$\mathcal{R}((T - \mu)_-) \subset \mathcal{N}(T - \mu)_+$$

since

$$(T - \mu)_+ \cdot (T - \mu)_- = |T - \mu|^2 - (T - \mu)^2 = 0$$

Thus,

$$\mathcal{N}((T - \mu)_+^\perp) \subset \mathcal{R}((T - \mu)_-^\perp)$$

Thus,

$$\mathcal{R}(1 - E(\mu)) \subset \mathcal{R}((T - \mu)_-^\perp)$$

Now, $\mathcal{R}((T - \mu)_-)$ and hence $\mathcal{R}((T - \mu)_-^\perp)$ are T -invariant by the self-adjointness of T and the commutativity of T and $(T - \mu)_-$. Hence, it is also $(T - \mu)_-$ invariant. Thus,

$$\mathcal{R}((T - \mu)_-(1 - E(\mu))) \subset \mathcal{R}((T - \mu)_-) \cap \mathcal{R}((T - \mu)_-) = \{0\}$$

Note: As our iteration method will show, the square root of a bounded positive operator is constructed as the limit of a sequence of polynomials in the operator and hence this square root commutes with the original operator). Thus,

$$(T - \mu)_-(1 - E(\mu)) = 0$$

Thus,

$$(1 - E(\mu))(T - \mu)(1 - E(\mu)) = (1 - E(\mu))(T - \mu)_+(1 - E(\mu)) = (T - \mu)_+$$

since

$$(T - \mu)_+ E(\mu) = 0, E(\mu)(T - \mu)_+ = 0$$

Thus, we have finally proved that if $\mu \geq \lambda$, then

$$(T - \lambda)_+ \geq (T - \mu)_+$$

and therefore since both sides here are positive operators, it follows that

$$\mathcal{N}((T - \lambda)_+) \subset \mathcal{N}(T - \mu)_+$$

or equivalently,

$$E(\lambda) \leq E(\mu)$$

Assume now that $\|T\| = M < \infty$. We can then show that for $\lambda > M$, $(T - \lambda)_+ = 0$ and for $\lambda < -M$, $(T - \lambda)_- = 0$. It follows that for $\lambda > M$, $E(\lambda) = I$ and for $\lambda < -M$, $\mathcal{R}((T - \lambda)_-) = \mathcal{H}$ and hence $\mathcal{R}((T - \lambda)_-)^\perp = \{0\}$. But,

$$\mathcal{R}((T - \lambda)_-) \subset \mathcal{R}((T - \lambda)_+)$$

and hence

$$\mathcal{R}((T - \lambda)_+^\perp) \subset \mathcal{R}((T - \lambda)_-^\perp) = 0$$

ie,

$$\mathcal{R}(T - \lambda)_+^\perp = 0, \lambda < -M$$

But,

$$\mathcal{R}((T - \lambda)_+^\perp) = \mathcal{N}((T - \lambda)_+)$$

by the self-adjointness of the concerned operator, and hence, we get

$$\mathcal{N}((T - \lambda)_+) = \{0\}, \lambda < -M$$

or equivalently,

$$E(\lambda) = 0, \lambda < -M$$

Thus, $E(\lambda)$ increases from 0 to I as λ increases from $-M$ to M . We next wish to show that $\lambda < \mu$ implies

$$\lambda.(E(\mu) - E(\lambda)) \leq T(E(\mu) - E(\lambda)) \leq \mu.(E(\mu) - E(\lambda))$$

In fact, $x \in \mathcal{R}(E(\mu) - E(\lambda))$ implies $x \in \mathcal{R}(E(\lambda))^\perp$ implies $x \in \mathcal{N}((T - \lambda)_+^\perp)$. Now, $y \in \mathcal{R}((T - \lambda)_+)$ implies $y \in \mathcal{N}((T - \lambda)_+^\perp)$ and therefore,

$$\mathcal{N}((T - \lambda)_+) \subset \mathcal{R}((T - \lambda)_+^\perp)$$

Also,

$$\mathcal{N}((T - \lambda)_+^\perp) \subset \mathcal{R}((T - \lambda)_+)$$

because, $x \in \mathcal{R}((T - \lambda)_+^\perp)$ implies $\langle x, (T - \lambda)_+y \rangle = 0 \forall y$ implies $\langle (T - \lambda)_+x, y \rangle = 0, \forall y$ implies $x \in \mathcal{N}((T - \lambda)_+^\perp)$ and hence taking the orthogonal complement on both sides gives the result. Thus,

$$\begin{aligned} (T - \lambda)(E(\mu) - E(\lambda)) &= (T - \lambda)((1 - E(\lambda))(E(\mu) - E(\lambda))) \\ &= (T - \lambda)_+(1 - E(\lambda))(E(\mu) - E(\lambda)) \\ &\quad (T - \lambda)_-(1 - E(\lambda))(E(\mu) - E(\lambda)) \end{aligned}$$

But by the above observation,

$$\begin{aligned} (T - \lambda)_-(1 - E(\lambda))x &\subset (T - \lambda)_-\mathcal{N}((T - \lambda)_+^\perp) \\ &\subset (T - \lambda)_-(\mathcal{R}((T - \lambda)_+)) = \{0\} \end{aligned}$$

since

$$(T - \lambda)_-(T - \lambda)_+ = 0$$

This proves that

$$\begin{aligned} (T - \lambda)(E(\mu) - E(\lambda)) &= (T - \lambda)((1 - E(\lambda))(E(\mu) - E(\lambda))) \\ &= (T - \lambda)_+(1 - E(\lambda))(E(\mu) - E(\lambda)) \geq 0 \end{aligned}$$

Likewise,

$$(T - \mu)(E(\mu) - E(\lambda)) = (T - \mu)(E(\mu)(E(\mu) - E(\lambda)))$$

and

$$(T - \mu)_+ E(\mu) = 0$$

so

$$(T - \mu)(E(\mu) - E(\lambda)) = -(T - \mu)_- E(\mu)(E(\mu) - E(\lambda)) \leq 0$$

thereby proving the claim. Now, we prove the strong right continuity of the increasing family of orthogonal projections $\{\{E(\lambda) : \lambda \in \mathbb{R}\}$. Since this is an increasing family of orthogonal projections, we have for $\mu > \lambda$ and any vector x ,

$$\| (E(\mu) - E(\lambda))x \|^2 = \| E(\mu)x \|^2 - \| E(\lambda)x \|^2 \geq 0$$

Thus, we have that the limit

$$\lim_{\mu \downarrow \lambda, \mu > \lambda} \| E(\mu)x \|$$

exists. Hence $\{E(\mu)x\}_{\mu > \lambda}$ forms a Cauchy sequence as $\mu \downarrow \lambda$, and hence the limit

$$E(\lambda+)x = \lim_{\mu \downarrow \lambda} E(\mu)x$$

exists. We write $F(\lambda)x = E(\lambda+)x - E(\lambda)x$ and we have to show that this vanishes (Note that x is arbitrary). We have for $\mu > \lambda$,

$$(A - \lambda)(E(\mu) - E(\lambda))x = (A - \lambda) + (E(\mu) - E(\lambda))x$$

Hence taking $\lim \mu \downarrow \lambda$ and using the boundedness of $(A - \lambda)$ and $(A - \lambda)_+$ that

$$(A - \lambda)F(\lambda)x = (A - \lambda)_+ F(\lambda)x$$

However,

$$(E(\mu) - E(\lambda))x \in \mathcal{R}(E(\mu)) = \mathcal{N}((T - \mu)_+)$$

and hence

$$(A - \mu)(E(\mu) - E(\lambda))x = 0, \mu > \lambda$$

Taking the limit $\mu \downarrow \lambda$ gives

$$(A - \lambda)F(\lambda)x = 0$$

Thus,

$$(A - \lambda)_+ F(\lambda)x = 0$$

and hence

$$F(\lambda)x \in \mathcal{N}((A - \lambda)_+) \subset \mathcal{N}((A - \mu)_+), \mu > \lambda$$

Thus,

$$(E(\mu) - E(\lambda))F(\lambda)x = 0, \mu > \lambda$$

and hence letting $\mu \downarrow \lambda$ results in

$$F(\lambda)^2x = 0$$

But obviously, $F(\lambda)$ is an orthogonal projection and hence $F(\lambda)^2 = F(\lambda)$. This proves that

$$F(\lambda)x = 0$$

and since this holds for all x , it follows that

$$F(\lambda) = 0$$

ie,

$$E(\lambda+) = E(\lambda)$$

when interpreted in the strong convergence sense.

6.26 Motion of rigid bodies in electromagnetic fields

This section of the book is inspired by some practical gadgets constructed in the laboratory of one of my colleagues Professor Dhananjay Gadre.

[a] The motion of one magnet in the magnetic field generated by another magnet. A magnet placed at the origin of the coordinates, from a microscopic viewpoint of view consists of small current loops all oriented along the same direction (a Ferromagnet/permanent magnet). The magnetic field generated by this in space without taking relativity into account is

$$B(t, r) = (\mu/4\pi)\mathbf{m}(t) \times \mathbf{r}/r^3$$

where we take into account the possibility that the magnetic dipole moment can change with time, for example, if the magnet rotates around an axis or moves with some velocity.

Note: If relativity is taken into account in the form of retarded potentials, then a loop Γ carrying a time varying current $I(t)$ produces a magnetic vector potential field

$$\mathbf{A}(t, \mathbf{r}) = (\mu/4\pi) \int_{\Gamma} I(t - |\mathbf{r} - \mathbf{r}'|/c, \mathbf{r}') d\mathbf{r}' / |\mathbf{r} - \mathbf{r}'|$$

Upto $O(1/c^2)$, this equals

$$A(t, r) = (\mu/4\pi)(I(t) \int_{\Gamma} d\mathbf{r}' / |\mathbf{r} - \mathbf{r}'| + (I''(t)/2c^2) \int_{\Gamma} d\mathbf{r}' |\mathbf{r} - \mathbf{r}'|)$$

In the far field approximation, this evaluates to

$$(\mu I(t)/4\pi r^3) \int_{\Gamma} d\mathbf{r}' \mathbf{r}' \cdot \mathbf{r} - (\mu I''(t)/8\pi c^2 r) \int_{\Gamma} d\mathbf{r}' \mathbf{r}' \cdot \mathbf{r}'$$

where we have made use of the fact that Γ is closed for this implies $\int_{\Gamma} d\mathbf{r}' = 0$. Now,

$$\int_{\Gamma} \mathbf{r} \cdot \mathbf{r}' d\mathbf{r}' = \mathbf{a} \times \mathbf{r}$$

where

$$\mathbf{a} = (1/2) \int_{\Gamma} \mathbf{r}' \times d\mathbf{r}'$$

Indeed, with \mathbf{a} as defined above, we have

$$\begin{aligned} \mathbf{a} \times \mathbf{r} &= (1/2) \int_{\Gamma} (\mathbf{r} \cdot \mathbf{r}' d\mathbf{r}' - \mathbf{r}' \cdot \mathbf{r} d\mathbf{r}') \\ &= \int_{\Gamma} \mathbf{r} \cdot \mathbf{r}' d\mathbf{r}' \end{aligned}$$

since

$$\begin{aligned} \int_{\Gamma} \mathbf{r}' \cdot \mathbf{r} d\mathbf{r}' &= \int_{\Gamma} (d(\mathbf{r}' \cdot \mathbf{r} \mathbf{r}') - \mathbf{r}' \cdot \mathbf{r} d\mathbf{r}') \\ &= - \int_{\Gamma} \mathbf{r}' \cdot \mathbf{r} d\mathbf{r}' \end{aligned}$$

Thus, the magnetic vector potential in the far field zone (which is valid for small current loops) is given upto $O(1/c^2)$ by

$$\mathbf{A}(t, \mathbf{r}) = (\mu/4\pi)(\mathbf{m}(t) \times \mathbf{r}/r^3 - \mathbf{m}''(t) \times \mathbf{r}/2c^2 r)$$

where

$$\mathbf{m}(t) = I(t)\mathbf{a}$$

is the magnetic dipole moment of the loop.

[3] The force of interaction between two magnets. Assume that the first magnet is an infinitesimal magnetic dipole \mathbf{m}_1 placed at \mathbf{r}_1 and that the second magnet is again an infinitesimal magnetic dipole \mathbf{m}_2 placed at \mathbf{r}_2 . Let $\mathbf{B}_k(r)$ denote the magnetic field produced by the k^{th} magnet, $k = 1, 2$. The total energy in the magnetic field generated by the two magnets is

$$U = (2\mu)^{-1} \int (\mathbf{B}_1(r) + \mathbf{B}_2(r))^2 d^3r$$

and hence the interaction energy between the two magnets is

$$U_I = \mu^{-1} \int (\mathbf{B}_1(r), \mathbf{B}_2(r)) d^3r$$

This energy is a function of $\mathbf{r}_1, \mathbf{r}_2$, the positions of the two magnets. Thus we write

$$U_I = U_I(\mathbf{r}_1, \mathbf{r}_2)$$

It follows that the force exerted by the first magnet on the second magnet is given by

$$\mathbf{F}_{12} = -\nabla_{r_2} U_I(\mathbf{r}_1, \mathbf{r}_2)$$

Now there are other alternate ways to express this force. For example, let $\mathbf{A}_1(\mathbf{r})$ denote the magnetic vector potential generated by the first magnet. Then,

$$p\mathbf{A}_1(\mathbf{r}) = (\mu/4\pi|\mathbf{r} - \mathbf{r}_1|^3)\mathbf{m}_1 \times (\mathbf{r} - \mathbf{r}_1)$$

We have that

$$\begin{aligned} U_I &= \mu^{-1} \int (\mathbf{B}_2(\mathbf{r}), \nabla \times \mathbf{A}_1(\mathbf{r})) d^3r \\ &= \mu^{-1} \int (\nabla \times \mathbf{B}_2(\mathbf{r}), \mathbf{A}_1(\mathbf{r})) d^3r \end{aligned}$$

(on integrating by parts). Using the Maxwell equation

$$\operatorname{curl} \mathbf{B}_2 = \mu \mathbf{J}_2(\mathbf{r})$$

we get

$$U_I = \int (\mathbf{J}_2(\mathbf{r}), \mathbf{A}_1(\mathbf{r})) d^3r$$

where $\mathbf{J}_2(\mathbf{r})$ is the current density within the second magnet caused by the electrons spinning around the nucleus in an oriented way in order to form a magnetic dipole.

Let us now provide a quantum mechanical derivation of the force between the two magnets. Suppose that the second magnet is regarded as an atomic system with energy operator $E(r - r_2, -ih\nabla)$. After it interacts with the magnetic field generated by the first magnet, its energy operator becomes $E(r - r_2, -ih\nabla + eA_1(r))$ where we assume that $-e$ is the equivalent charge of the second magnet comprising the sum of all the electronic charges in it. Then we solve for the wave function of the second magnet using the Schrodinger equation:

$$E(r - r_2, -ih\nabla + eA_1(r))\psi_n(r|r_2) = E_n(r_2), n = 1, 2, \dots$$

We next compute the force on the second magnet caused by the first magnet when the latter is in the state ψ_n . This force is given by

$$-\nabla_{r_2} E_n(r_2) = -E'_n(r_2)$$

Assuming finally that the second magnet is in the Gibbs state it follows that the average force experienced by the second magnet is given by

$$\mathbf{F}_{12}(\mathbf{r}_2) = \sum_n \exp(-\beta E_n(r_2)) E'_n(r_2) / \sum_n \exp(-\beta E_n(r_2))$$

One can proceed even further by computing the quantum magnetic field produced by the first magnet. Let $a_k, a_k^*, k = 1, 2, \dots$ denote the annihilation and creation operators of the photon field produced by the first magnet. Let ψ denote the second quantized Dirac wave function of the electron-positron field within the first magnet. The 4-current density operator field generated by the first magnet can then be expressed as

$$J^\mu(t, r) = -e\psi(t, r)^* \gamma^0 \gamma^\mu \psi(t, r)$$

and hence by the retarded potential theory, the total quantum magnetic field produced by the first magnet is given by

$$\mathbf{A}(t, r) = (e/4\pi) \int_{V_1} J(\mathbf{r}_1) \times (\mathbf{r} - \mathbf{r}_1) d^3 r_1 / |\mathbf{r} - \mathbf{r}_1|^3$$

In addition, we may add to this a contribution from the free photon field

$$\mathbf{A}_{ph}(t, r) = \int (2|K|)^{-1/2} [c(K, s) \exp(-ik.x) + c(K, s)^* \exp(ik.x)] d^3 K$$

Note that the second quantized Dirac field has the plane wave expansion

$$\psi(t, r) = \int [a(p, \sigma) \mathbf{u}(p, \sigma) \exp(-ip.x) + b(p, \sigma)^* \mathbf{v}(p, \sigma) \exp(ip.x)] d^3 P$$

where the integration is carried out on the mass shell

$$\mathbf{P}^2 + m^2 = p^{02} = E(\mathbf{P})^2$$

6.27 Large deviation theory with engineering applications

- [1] Cramer's theorem on large deviations for sample averages of iid random variables.
- [2] Sanov's theorem on large deviations for the empirical distribution of a sequence of independent random variables.
- [3] Derivation of Sanov's theorem from Cramer's theorem.
- [4] The Gartner-Ellis theorem on the rate function for empirical means of non-independent random variables when the limiting logarithmic moment generating function exists.
- [5] Varadhan's integral lemma and Bryc's inverse Varadhan lemma for the limiting logarithmic moment generating function of a function of a family of r.v's having the large deviation property. Applications to the computation of the low temperature limit of the partition function in statistical mechanics.
- [6]

6.28 Lectures in Linear Algebra for Signal Processing Applications

- [1] Spectral theorem for compact Hermitian operators in an infinite dimensional Hilbert space.
- [2] Spectral theorem for bounded Hermitian operators in an infinite dimensional Hilbert space.
- [3] The Cq-Shannon coding theorem.

The basic Classical-Quantum noisy Shannon coding theorem based on proofs given by Andreas Winter

$\rho(x)$ is a state in a Hilbert space \mathcal{H} for each $x \in A$ where A is a finite alphabet. For $u \in A^n$, we write $\rho(u) = \otimes_{x \in A} \rho(x)^{\otimes N(x|u)}$ where the order of the tensor products over x is specified according to a given ordering of A . We also write $P_u(x) = N(x|u)/n$. Here, $N(x|u)$ is the number of times x appears in the sequence u . We also define $\bar{\rho} = \sum_{x \in A} P(x)\rho(x)$ where P is a given probability distribution on A .

6.29 On the improvement of the signal quality in telephone lines

- [1] Adaptive echo cancellation implementation using digital filters.

Summary: Let A and B be speakers communicating over a line. A 's speech is $x[n]$ while B 's speech is $y[n]$. The signal $z[n]$ received by A is B 's speech plus A 's echo:

$$z[n] = y[n] + \sum_{k=0}^p g[k]x[n-k] = y[n] + e[n]$$

This echo is produced by A 's speech signal impinging at points on the line where there is some distortion/non-uniformity. Assuming that A 's speech $x[n]$ and B 's speech $y[n]$ are uncorrelated, it follows that A 's speech will be highly correlated with his own echo $e[n]$ but not with B 's speech $y[n]$. So if A passes his own speech signal $xz[n]$ through a filter/adaptive filter with aim of predicting his received signal $z[n]$, then he will be able to predict only the echo component $e[n]$ and hence by subtracting the filter output $\hat{e}[n]$ from his received signal $z[n]$, he can recover a good estimate of B 's signal:

$$\hat{y}[n] = z[n] - \hat{e}[n] = x[n] + e[n] - \hat{e}[n]$$

Let $h[n]$ be the filter used by A . The filter output is

$$\hat{e}[n] = \sum_{k=0}^q h[k]x[n-k]$$

and A minimizes

$$E[h] = \mathbb{E}(z[n] - \hat{e}[n])^2$$

w.r.t. $\{h[n]\}$. Now observing that the processes $x[.]$ and $y[.]$ are uncorrelated, the error energy is

$$E[g] = \mathbb{E}(y[n]^2) + \mathbb{E}(e[n] - \hat{e}[n])^2$$

and hence minimizing this w.r.t h is equivalent to minimizing

$$F[h] = \mathbb{E}(e[n] - \hat{e}[n])^2$$

w.r.t g . It is a straightforward matter to express the optimum g in terms of the statistical correlations in $x[.]$ and the echo generating system impulse response $h[.]$ and hence calculate the SNR:

$$SNR = \frac{\mathbb{E}(y[n]^2)}{\mathbb{E}(e[n] - \hat{e}[n])^2}$$

Since statistical correlations are not typically available, we use adaptive LMS and RLS algorithms for updating the filter weights. These adaptive filters will also take into account situations when the echo producing deformities shift and change slowly with time causing the corresponding echo generating impulse response $h[n]$ to vary slowly with time. The adaptive LMS algorithm is

$$\begin{aligned} h_{n+1}[r] &= h_n[r] - \mu \frac{\partial}{\partial h_n[r]} (z[n] - \sum_{m=0}^q h_n[m]x[n-m])^2, r = 0, 1, \dots, q \\ &= h_n[r] + 2\mu x[n-r].(z[n] - \sum_{m=0}^q h_n[m]x[n-m]) \end{aligned}$$

or in vectorial form,

$$\mathbf{h}_{n+1} = (\mathbf{I} - 2\mu \mathbf{x}_n \mathbf{x}_n^T) \mathbf{h}_n + 2\mu z[n] \mathbf{x}_n$$

where

$$\mathbf{h}_n = [h_n[0], \dots, h_n[q]]^T, \mathbf{x}_n = [x[n], x[n-1], \dots, x[n-q]]^T$$

More generally, there may be nonlinearities in the line causing the echo. In this case, the echo model is

$$e[n] = G(x[n], x[n-1], \dots, x[n-p] | \theta)$$

where θ is an unknown parameter vector and we use the gradient LMS algorithm to update θ so that

$$\mathbb{E}(z[n] - G(\mathbf{x}_n | \theta))^2$$

is minimized. The LMS update equations read

$$\begin{aligned} \theta[n+1] &= \theta[n] - \mu \nabla_\theta (z[n] - G(\mathbf{x}_n | \theta[n]))^2 \\ &= \theta[n] + 2\mu (\nabla_\theta G(\mathbf{x}_n | \theta[n])) (z[n] - G(\mathbf{x}_n | \theta[n])) \end{aligned}$$

After convergence, the echo can be determined and subtracted off from A's received signal $z[n]$. Some of the open problems in this formulation are to compute the approximate mean and covariance propagation equations for the parameters $\theta[n]$ when $x[.]$ and $y[.]$ are independent stochastic processes with known statistics.

In implementing the RLS algorithm, we would be required to calculate the filter coefficient vector \mathbf{h}_n recursively in time by minimizing

$$\sum_{k=0}^n \lambda^{n-k} (z[k] - \mathbf{h}_n^T \mathbf{x}_k)^2$$

or in the nonlinear case, calculate $\theta[n]$ recursively in time by minimizing

$$\sum_{k=0}^n \lambda^{n-k} (z[k] - G(\mathbf{x}_k | \theta[n]))^2$$

- [2] Adaptive line enhancement in telephone lines based on the difference between the correlation time lags for signal and noise processes.
- [3] Noise and echo cancellation in conference calls, ie, more than two users communicating with each other through a common server/exchange.
- [4] Use of queueing theory ideas in the transmission of information packets over telephone lines.

6.30 Quantum Coulomb scattering

$$H_0 = P^2, H = H_0 + V(Q), V(Q) = \alpha/|Q|$$

The wave operators do not exist since

$$\int \| V(Q)U_t^0 f \| dt = \infty$$

for $f(Q)$ in $C^\infty(\mathbb{R}^n)$. We therefore define

$$X_t(P) = X_t = (\gamma/2|P|)\log(t), t > 0$$

and find that

$$dX_t/dt = \gamma/2|P|t$$

We prove the existence of the modified wave operator

$$\Omega_+ = \lim_{t \rightarrow \infty} U_{-t} U_t^0 \exp(-iX_t)$$

defined on C^∞ . Here,

$$U_t^0 = \exp(-itH_0), U_t = \exp(-itH)$$

We have

$$\begin{aligned} \Omega_+ f &= f + \int_0^\infty d/dt(U_{-t} U_t^0 \exp(-iX_t)) dt \\ &= f + \int_0^\infty U_{-t} (iV(Q) - i\gamma/2|P|t) U_t^0 \cdot \exp(-iX_t) f dt \end{aligned}$$

So to prove the existence of Ω_+ we must show that

$$\int_0^\infty \| (V(Q) - \gamma/2|P|t) U_t^0 \exp(-iX_t) f \| dt < \infty$$

provided that γ is chosen appropriately. We have

$$U_{-t}^0 Q U_t^0 = \exp(itad(P^2))(Q) = Q + 2tP$$

and on the other hand defining

$$Z_t = Z_t(Q) = \exp(-iQ^2/4t)$$

gives

$$Z_t P Z_t^* = \exp(-iad(Q^2)/4t)(P) = P + Q/2t = (Q + 2tP)/2t$$

Thus,

$$U_{-t}^0 (V(Q) - \gamma/2|P|t) U_t^0 = \alpha/|Q + 2tP| - \gamma/2|P|t$$

We therefore choose $\gamma = \alpha$ to get

$$\begin{aligned} & \| V(Q) - \gamma/2|P|t) U_t^0 \cdot \exp(-iX_t) f \| = \\ & \| U_{-t}^0 (\alpha/|Q| - \gamma/2|P|t) U_t^0 \cdot \exp(-iX_t) f \| \\ & = \alpha \| (1/|Q + 2tP| - 1/2|P|t) f_t \| \end{aligned}$$

where

$$f_t = \exp(-iX_t) f$$

Intuitively, it is clear that the function $t \rightarrow \| (|Q + 2Pt|^{-1} - |2Pt|^{-1}) f_t \|$ should be integrable over $[1, \infty)$ because of the cancellation of the $O(1/t)$ term. Specifically, we can argue out intuitively that with $D = (Q, P) + (P, Q) = 2(Q, P) - 2in$

$$\begin{aligned} & |Q + 2Pt|^{-1} - |2Pt|^{-1} = (Q^2 + 4P^2 t^2 + 2tD)^{-1/2} - |2Pt|^{-1} \\ & = [(4P^2 t^2)^{-1} [1 + Q^2 (4P^2)^{-1} t^{-2} + D(2P^2)^{-1} t^{-1}]]^{1/2} - |2Pt|^{-1} \end{aligned}$$

should be integrable in t for large t . However, we need a more rigorous proof. It is not clear from the above expression that the $O(1/t)$ term cancels out. However, we can also try the expansion

$$\begin{aligned} & (Q^2 + 4P^2 t^2 + 2tD)^{-1/2} - |2Pt|^{-1} \\ & = [(4P^2 t^2 + 2Dt)^{-1} (1 + Q^2 (4t^2 P^2 + 2tD)^{-1})]^{1/2} - |2tP|^{-1} \end{aligned}$$

For large t , this behaves as

$$(4P^2 t^2 + 2Dt)^{-1/2} - |2tP|^{-1}$$

If we assume that P^2 and D commute, then we can obtain using a binomial expansion the above expression as being of order $1/t^2$ for large t which guarantees integrability. But these two operators actually do not commute. We therefore adopt a different approach to prove the existence of the modified wave operators.

$$Z_t P Z_t^* = (Q + 2Pt)/2t$$

So

$$|Q + 2Pt|^{-1} - |2Pt|^{-1} = Z_t|2Pt|^{-1}Z_t^* - |2Pt|^{-1} = (2t)^{-1}(Z_t|P|^{-1}Z_t^* - |P|^{-1})$$

Now, by differentiation followed by integration,

$$Z_t|P|^{-1}Z_t^* - |P|^{-1} = \int_0^t Z_s[iQ^2/4s^2, |P|^{-1}]Z_s^* ds$$

Now,

$$[Q^2, |P|^{-1}] = [Q_j, |P|^{-1}]Q_j + Q_j[Q_j, |P|^{-1}]$$

with summation over the repeated index j being implied. Now,

$$[Q_j, |P|^{-1}] = -iP_j/|P|^3$$

So,

$$[Q^2, |P|^{-1}] = -i[(Q, P/|P|^3) + (P/|P|^3, Q)]$$

and we get

$$Z_t|P|^{-1}Z_t^* - |P|^{-1} = (1/4) \int_0^t s^{-2} Z_s[(Q, P/|P|^3) + (P/|P|^3, Q)]Z_s^* ds$$

Other methods: We write

$$\nu(Q) = \phi(Q)\alpha/|Q|, \mu(Q) = \psi(Q)\alpha/|Q|$$

where $\phi(x) = 1$ for $|x| > 1$, $\phi(x) = 0$, $|x| < 1/2$ and ϕ is in $C^\infty(\mathbb{R}^n)$. Likewise, $\psi(x) = 1 - \phi(x)$ satisfies $\psi(x) = 0$ for $|x| > 1$, $\psi(x) = 1$, $|x| < 1/2$ and ψ is of course C^∞ . It is clear that since ν is C^∞ , it follows that

$$|\nu(x)| \leq K_1(1 + |x|^2)^{-1/2}$$

Indeed, we have that this inequality holds trivially for $|x| < 1/2$ and for $|x| > 1$, we have

$$(1 + |x|^2)^{-1/2} = |x|^{-1}(1 + |x|^{-2})^{-1/2} \geq 2/|x|$$

In the range $1/2 < |x| < 1$, we assume that $\phi(x) \in [0, 1]$. We also obviously have

$$|\nabla \nu(x)| \leq K_2(1 + |x|^2)^{-1},$$

$$|\nabla^2 \nu(x)| \leq K_3(1 + |x|^2)^{-3/2}$$

Now,

$$d/dt(U_{-t}U_t^0 \exp(-iX_t)) = iU_{-t}(V(Q) - dX_t/dt)U_t^{(0)}\exp(-iX_t)$$

Note that $X_t = X_t(P)$ commutes with $U_t^0 = \exp(-itP^2)$.

$$dX_t/dt = \alpha/|2Pt| = V(2Pt)$$

So,

$$\begin{aligned}
& \| (d/dt(U_{-t}U_t^0 \exp(-iX_t))f \| \leq \\
& \| (V(Q) - V(2Pt))U_t^0 \exp(-iX_t)f \| \\
= & \| (\nu(Q) - \nu(2Pt) + \mu(Q) - \mu(2Pt))U_t^0 \exp(-iX_t)f \| \\
\leq & \| (\nu(Q) - \nu(2Pt))U_t^0 \exp(-iX_t)f \| \\
+ & \| \mu(Q)U_t^0 \exp(-iX_t)f \| + \| \mu(2Pt)U_t^0 \exp(-iX_t)f \|
\end{aligned}$$

We note that $\mu(2Pt), U_t^0, \exp(-iX_t)$ all commute. Thus,

$$\begin{aligned}
& \| \mu(2Pt)U_t^0 \exp(-iX_t)f \| = \| \mu(2Pt)f \| \\
= & \left(\int |\mu(2tk)|^2 |\tilde{f}(k)|^2 d^n k \right)^{1/2} \rightarrow 0, t \rightarrow \infty
\end{aligned}$$

Note that we are assuming that there is a $\delta > 0$ such that $\tilde{f}(k) = 0$ for $|k| < \delta$. Now more can be said. We have

$$\begin{aligned}
& \left(\int |\mu(2tk)|^2 |\tilde{f}(k)|^2 d^n k \right)^{1/2} = (2t)^{-n/2} \left[\int |\mu(k)|^2 |\tilde{f}(k/2t)|^2 d^n k \right]^{1/2} \\
= & (2t)^{-n/2} \left[\int_{|k| \leq 1} |\mu(k)|^2 |\tilde{f}(k/2t)|^2 d^n k \right]^{1/2} \\
\leq & c(f) t^{-n/2}
\end{aligned}$$

which is integrable w.r.t. t over $[\delta, \infty)$ for any $\delta > 1$ provided that $n \geq 3$. Now,

$$\begin{aligned}
& \| \mu(Q)U_t^0 \exp(-iX_t)f \| = \\
= & \| U_{-t}^0 \mu(Q)U_t^0 \exp(-iX_t)f \| = \\
& \| \mu(Q + 2Pt) \exp(-iX_t).f \| \\
= & \| Z_t \mu(2Pt) Z_t^* \exp(-iX_t)f \|
\end{aligned}$$

where

$$Z_t = \exp(-iQ^2/4t)$$

Now,

$$\begin{aligned}
& \| Z_t \mu(2Pt) Z_t^* \exp(-iX_t)f \|^2 \\
= & \| \mu(2Pt) Z_t^* \exp(-iX_t).f \|^2 = \int |\mu(2kt) \tilde{f}_t(k)|^2 d^n k
\end{aligned}$$

where $\tilde{f}_t(k)$ is the Fourier transform of $\exp(-ix^2/4t) \exp(-iX_t(-i\nabla))f(x)$. Further, this equals

$$(2t)^{-n} \int_{|k| \leq 1} |\mu(k) \tilde{f}_t(k/2t)|^2 d^n k$$

Thus,

$$\| \mu(Q)U_t^0 \exp(-iX_t)f \| \leq d(f) t^{-n/2}$$

which is again integrable over $[\delta, \infty)$ provided that $n \geq 3$.

Remark:

$$\| \mu(2Pt)Z_t^* \exp(-iX_t)f \| \leq \| \mu(2Pt)Z_t^* < Q >^{-2n} \| \cdot \| < Q >^{2n} \exp(-iX_t)f \|$$

where

$$< Q >^2 = 1 + Q^2, < Q >^{-2n} = (1 + Q^2)^{-n}$$

Now, for $q \geq 2$

$$\| \mu(2Pt) < Q >^{-2n} \| \leq \| \mu(2tx) \|_q \cdot \| |(1+x^2)^{-n}| \|_q$$

and

$$\| \mu(2tx) \|_q = \left(\int |\mu(2tx)|^q d^n x \right)^{1/q} = (2t)^{-n/q} \| \mu(x) \|_q$$

Further,

$$\int (1+x^2)^{-nq} d^n x = c_n \int_0^\infty (1+r^2)^{-nq} r^{n-1} dr$$

is finite if $n - 1 - 2nq < -1$ or equivalently, if $q > 1/2$. Further, $t^{-n/q}$ is integrable over $[\delta, \infty)$ if $n/q > 1$, ie, if $n > q$.

Remark: Let ϕ, ψ be defined on \mathbb{R}^n . Let $2 \leq q < \infty$. Then, we claim that

$$\| \phi(P)\psi(Q) \| \leq \| \phi \|_q \cdot \| \psi \|_q$$

To see this, we define

$$u = 1 - 2/q, q \geq 2, 1/p = 1 - 1/q$$

Then, $0 \leq u \leq 1$. For any function f on \mathbb{R}^n define its Fourier transform to be Tf . Then, we have the obvious identities

$$\| Tf \|_2 = \| f \|_2,$$

$$\| Tf \|_\infty \leq C \| f \|_1$$

Define

$$\chi(u) = (\| Tf \|_q / \| f \|_p)$$

Then,

$$\chi(0) = 1, \chi(1) \leq C$$

So by Hadamard's three line theorem,

$$\chi(u) = \| Tf \|_q / \| f \|_p \leq \chi(0)^{1-u} \chi(1)^u = C^u = C^{1-2/q} = K(q), q \geq 2$$

say. $K(q)$ depends only on q . We write this fundamental inequality as

$$\| Tf \|_q \leq K(q) \| f \|_p, q \geq 2, 1/q + 1/p = 1$$

Now, Holder's inequality implies that if

$$1/p + 1/q = 1/r$$

then

$$\| fg \|_r \leq \| f \|_p \cdot \| g \|_q$$

Thus, we get

$$\| \phi(Q)\psi(P)f \|_2 \leq \| \phi \|_q \cdot \| \psi(P)f \|_{q'}$$

where

$$1/q + 1/q' = 1/2$$

Further, as observed above,

$$\begin{aligned} \| \psi(P)f \|_{q'} &\leq C_1 \| T\psi(P)f \|_p \\ &= C_1 \| \psi(Q)Tf \|_p \leq C_1 \| \psi \|_q \cdot \| Tf \|_{p'} \\ &\leq C_1 C_2 \| \psi \|_q \cdot \| f \|_{p''} \end{aligned}$$

where we choose p so that

$$1/p + 1/q' = 1$$

The other numbers are chosen so that

$$1/q + 1/p' = 1/p, 1/p' + 1/p'' = 1$$

This means that

$$1/p'' = 1 - 1/p' = 1 - 1/p + 1/q = 1/q' + 1/q = 1/2$$

This proves that

$$\| \phi(Q)\psi(P)f \|_2 \leq C_1 C_2 \| \phi \|_q \cdot \| \psi \|_q \cdot \| f \|_2$$

and hence

$$\| \phi(Q)\psi(P) \| \leq \| \phi \|_q \cdot \| \psi \|_q, q \geq 2$$

6.31 Tutorial problems in electromagnetic field theory

[1] A point charge Q moves along the trajectory $t \rightarrow \mathbf{R}(t)$. Its velocity is therefore $\mathbf{V}(t) = \mathbf{R}'(t)$. Show that the charge and current densities in space corresponding to this point charge are respectively

$$\rho(t, r) = Q\delta^3(r - \mathbf{R}(t)), \mathbf{J}(t, r) = Q\mathbf{V}(t)\delta^3(r - \mathbf{R}(t))$$

Evaluate the retarded magnetic vector potential

$$A(t, r) = (\mu/4\pi) \int J(t - |r - r'|/c) d^3 r' / |r - r'| =$$

$$\begin{aligned}
&= (\mu/4\pi) \int J(t', r') \delta(t' - t + |r - r'|/c) d^3 r' d^3 t' / |r - r'| \\
&= (Q\mu/4\pi) \int \mathbf{V}(t') \delta^3(r' - R(t')) \delta(t' - t + |r - r'|/c) d^3 r' d^3 t' / |r - r'| \\
&= (Q\mu/4\pi) \int \mathbf{V}(t') \delta(t' - t + |r - R(t')|/c) dt' / |r - R(t')|
\end{aligned}$$

Evaluate this integral using the identity

$$\delta(f(x)) = \sum_k \delta(x - x_k) / |f'(x_k)|$$

where $\{x_k\}$ are the roots of $f(x)$.

[2] Evaluate the dispersion relation for a conducting fluid in an electromagnetic field. The fluid equations coupled to the Maxwell equations are

$$\begin{aligned}
\rho(t, r)((v(t, r), \nabla)v(t, r) + v_{,t}(t, r)) \\
= -\nabla P(\rho(t, r)) + \eta \nabla^2 v(t, r) + \sigma(E(t, r) + v(t, r) \times B(t, r)) \times B(t, r), \\
\operatorname{div}(\rho(t, r)v(t, r)) + \rho_{,t}(t, r) = 0
\end{aligned}$$

$$\operatorname{curl} E(t, r) = -B_{,t}(t, r), \operatorname{curl} B(t, r) = \mu \sigma(E(t, r) + v(t, r) \times B(t, r)) + \mu \epsilon E_{,t}(t, r)$$

We write

$$\begin{aligned}
\mathbf{v}(t, r) &= V_0 + \delta v(t, r), \rho(t, r) = \rho_0 + \delta \rho(t, r), \\
E(t, r) &= E_0 + \delta E(t, r), B(t, r) = B_0 + \delta B(t, r)
\end{aligned}$$

For the unperturbed equations to hold good, we require

$$E_0 = 0, V_0 \times B_0 = 0$$

The perturbed equations then become

$$\begin{aligned}
\rho_0((V_0, \nabla)\delta v(t, r) + \delta v_{,t}(t, r)) \\
= -P'(\rho_0)\nabla\delta\rho(t, r) + \eta\nabla^2\delta v(t, r) + \sigma(\delta E(t, r) + V_0 \times \delta B(t, r) + \\
\delta v(t, r) \times B_0) \times B_0 --- (1)
\end{aligned}$$

$$\rho_0 \operatorname{div} \delta v(t, r) + (V_0, \nabla\delta\rho(t, r)) + \delta \rho_{,t}(t, r) = 0 --- (2)$$

$$\operatorname{curl} \delta E(t, r) = -\delta B_{,t}(t, r) --- (3),$$

$$\operatorname{curl} \delta B(t, r) \mu \sigma(\delta E(t, r) + \delta v(t, r) \times B_0 + V_0 \times \delta B(t, r)) --- (4)$$

Eqns.(1)-(4) are the fundamental linearized equations for the four functions $\delta v(t, r), \delta \rho(t, r), \delta E(t, r), \delta B(t, r)$ using which we can derive dispersion relations assuming that these variables behave as plane waves, ie,

$$\delta v(t, r) = \operatorname{Re}(\delta v_0 \exp(i(\omega t - K.r))),$$

$$\delta \rho(t, r) = \operatorname{Re}(\delta \rho_0 \exp(i(\omega t - K.r))),$$

$$\delta E(t, r) = \operatorname{Re}(\delta E_0 \exp(i(\omega t - K.r))),$$

$$\delta B(t, r) = \operatorname{Re}(\delta B_0 \exp(i(\omega t - K.r)))$$

After making these substitutions into the above linearized pde, we get homogeneous linear algebraic equations for $\delta v_0, \delta \rho_0, \delta E_0, \delta B_0$ and hence setting the determinant of the associated matrix to zero, we get the dispersion relation, ie, relation between ω and \mathbf{K} .

[3] Computing the average electric and magnetic dipole moment of an atom in a Gibbs state after it gets perturbed by a static electromagnetic field.

$$\mathbf{E}(r) = -\nabla\Phi(r), \mathbf{B}(r) = \operatorname{curl} \mathbf{A}(r)$$

We may assume the gauge condition

$$\operatorname{div} \mathbf{A}(r) = 0$$

Unperturbed atomic Hamiltonian

$$H_0 = (-\hbar^2/2m)\nabla^2 + V_0(r)$$

Perturbed atomic Hamiltonian

$$H = (-\hbar^2/2m)(\nabla + ie\mathbf{A}(r)/\hbar)^2 + V_0(r) - e\Phi(r)$$

We may write

$$H = H_0 + eV_1 + e^2V_2$$

where

$$V_1 = -\Phi(r) - (ih/m)(\mathbf{A}(r), \nabla), V_2 = A(r)^2/2m$$

Assume that a complete orthonormal basis for the unperturbed stationary states is $|\psi_{nk}^{(0)}\rangle, k = 1, 2, \dots, d_n, n = 1, 2, \dots$ with the distinct energy levels $E_n^{(0)}, n = 1, 2, \dots$ so that $\{\psi_{nk}^{(0)} : k = 1, 2, \dots, d_n\}$ being an orthonormal basis for $\mathcal{N}(H_0 - E_n^{(0)})$, ie, the unperturbed energy level $E_n^{(0)}$ has a degeneracy of d_n . The first order shifts in the wave functions and the energy levels are given by

$$e|\phi_{nr}^{(1)}\rangle = e \sum_{m,l} |\phi_{ml}^{(0)}\rangle \frac{\langle \phi_{ml}^{(0)} | V_1 | \phi_{nr}^{(0)} \rangle}{(E_n^{(0)} - E_m^{(0)})}$$

where

$$|\phi_{nr}^{(0)}\rangle = \sum_{k=1}^{d_n} c_n(r, k) |\psi_{nk}^{(0)}\rangle$$

and $((c_n(r, k)))_{k=1}^{d_n} = \mathbf{c}_n(r)$ is the eigenvector of the $d_n \times d_n$ "secular matrix" $((\langle \psi_{nk}^{(0)} | V_1 | \psi_{nl}^{(0)} \rangle))_{1 \leq k, l \leq d_n}$ with eigenvalue $E_{nr}^{(1)}$ where $r = 1, 2, \dots, d_n$. These eigenvectors may be assumed to be a complete orthonormal basis for \mathbb{C}^{d_n} . Note that just as $|\psi_{nk}^{(0)}\rangle, k = 1, 2, \dots, d_n$ forms an onb for $\mathcal{N}(H_0 - E_n^{(0)})$, so also $|\phi_{nr}^{(0)}\rangle = \sum_{k=1}^{d_n} c_n(r, k) |\psi_{nk}^{(0)}\rangle, r = 1, 2, \dots, d_n$ also forms an onb for the same subspace. It is these "rotated eigenvectors of the unperturbed Hamiltonian

which get perturbed and causing the degeneracy to be lifted. Specifically, the perturbed eigenfunction of $H_0 + eV_1$ corresponding to the eigenvalues $E_n^{(0)} + eE_{nr}^{(1)}$ is $|\phi_{nr}^{(0)}\rangle + e|\phi_{nr}^{(1)}\rangle$ for each $r = 1, 2, \dots, d_n$ upto $O(e)$. The perturbed Gibbs density operator upto $O(e)$ is then

$$\rho + \delta\rho = \frac{\sum_{n,r}(|\phi_{nr}^{(0)} + e|\phi_{nr}^{(1)}\rangle)exp(-\beta(E_n^{(0)} + eE_{nr}^{(1)}))(\langle\phi_{nr}^{(0)}| + e\langle\phi_{nr}^{(1)}|)}{Z^{(0)}(\beta) + eZ^{(1)}(\beta)}$$

where

$$Z^{(0)}(\beta) = \sum_{n,r} d_n exp(-\beta E_n^{(0)}),$$

$$Z^{(1)}(\beta) = -\beta \sum_{n,r} exp(-\beta E_{nr}^{(0)}) E_{nr}^{(1)}$$

and the numerator evaluates upto $O(e)$ to

$$\sum_{n,r} exp(-\beta E_n^{(0)}) |\phi_{nr}^{(0)}\rangle \langle\phi_{nr}^{(0)}| + e[\sum_{n,r} exp(-\beta E_n^{(0)}) (|\phi_{nr}^{(0)}\rangle \langle\phi_{nr}^{(1)}| + |\phi_{nr}^{(1)}\rangle \langle\phi_{nr}^{(0)}|) - \beta E_{nr}^{(1)} |\phi_{nr}^{(0)}\rangle \langle\phi_{nr}^{(0)}|]$$

Thus,

$$\rho = Z^{(0)}(\beta)^{-1} \sum_{n,r} exp(-\beta E_n^{(0)}) |\phi_{nr}^{(0)}\rangle \langle\phi_{nr}^{(0)}|$$

and

$$\delta\rho = -e\rho.Z^{(1)}(\beta)/Z^{(0)}(\beta) + eZ^{(0)}(\beta)^{-1} [\sum_{n,r} exp(-\beta E_n^{(0)}) (|\phi_{nr}^{(0)}\rangle \langle\phi_{nr}^{(1)}| + |\phi_{nr}^{(1)}\rangle \langle\phi_{nr}^{(0)}|) - \beta E_{nr}^{(1)} |\phi_{nr}^{(0)}\rangle \langle\phi_{nr}^{(0)}|]$$

The average electric dipole moment of the atom upto $O(e^2)$ is thus given by

$$p + \delta p = -e \langle \mathbf{r} \rangle = -e Tr((\rho + \delta\rho)\mathbf{r})$$

where

$$p = -e Tr(\rho\mathbf{r}) = -e Z^{(0)}(\beta)^{-1} \sum_{n,r} d_n exp(-\beta E_n^{(0)}) \langle\phi_{nr}^{(0)}|\mathbf{r}|\phi_{nr}^{(0)}\rangle$$

is the unperturbed dipole moment of the atom which is usually zero in the ground state (owing to spherical symmetry of the binding potential $V(r)$). δp is the perturbation to the dipole moment caused by application of the external static electromagnetic field:

$$\begin{aligned} \delta p &= -e Tr(\delta\rho\mathbf{r}) = \\ &(e^2 Z^{(1)}(\beta)/Z^{(0)}(\beta)).Tr(\rho\mathbf{r}) \\ &- e^2 Z^{(0)}(\beta)^{-1} [\sum_{n,r} exp(-\beta E_n^{(0)}) (2Re(\langle\phi_{nr}^{(1)}|\mathbf{r}|\phi_{nr}^{(0)}\rangle) - \beta E_{nr}^{(1)} \langle\phi_{nr}^{(0)}|\mathbf{r}|\phi_{nr}^{(0)}\rangle)] \end{aligned}$$

[4] Computing the average time varying electric and magnetic dipole moment of an atom perturbed by a time varying electromagnetic field taking the spin of the electron into account. The Hamiltonian is defined in the Hilbert space

$$\mathcal{H} = \mathbb{C}^2 \otimes L^2(\mathbb{R}^3) = L^2(\mathbb{R}^3) \oplus L^2(\mathbb{R}^3)$$

and is given by the Pauli formula

$$H(t) = (-\hbar^2/2m)(\nabla + ieA(t, r)/\hbar)^2 + V_0(r) - e\Phi(t, r) + eh(\sigma, B(t, r))/2m$$

In the special case, when the magnetic field is spatially homogeneous, we can take

$$A(t, r) = B(t) \times r/2$$

and get

$$\begin{aligned} H(t) &= (-\hbar^2/2m)\nabla^2 + V_0(r) - (ie\hbar/2m)(B(t) \times r, \nabla) - e\Phi(t, r) + eh(\sigma, B(t))/2m \\ &= (-\hbar^2/2m)\nabla^2 + V_0(r) - e\Phi(t, r) + e(B(t), \mathbf{L} + h\sigma)/2m + (e^2/8m)(B(t) \times r)^2 \end{aligned}$$

where

$$\mathbf{L} = -i\hbar r \times \nabla$$

is the angular momentum pseudo-vector operator. We note that the electric field in this case is

$$E(t, r) = -\nabla\Phi(t, r) - B'(t) \times r/2$$

We shall assume this latter case and hence write

$$H(t) = I_2 \otimes H_0 + eV_1(t) + e^2V_2(t)$$

where

$$H_0 = (-\hbar^2/2m)\nabla^2 + V_0(r),$$

$$V_1(t) = -I_2 \otimes \Phi(t, r) + (e/2m)(B(t), \mathbf{L} + h\sigma),$$

$$V_2(t) = (1/8m)I_2 \otimes (B(t) \times r)^2$$

Remark on second order degenerate perturbation theory.

$$H_0|\phi_{nr}^{(0)}\rangle = E_n^{(0)}|\phi_{nr}^{(0)}\rangle, \quad r = 1, 2, \dots, d_n$$

$$(H_0 - E_n^{(0)})|\phi_{nr}^{(1)}\rangle = (E_{nr}^{(1)} - V_1)|\phi_{nr}^{(0)}\rangle,$$

$$(H_0 - E_n^{(0)})|\phi_n^{(2)}\rangle = (E_n^{(2)} - V_2)|\phi_{nr}^{(0)}\rangle + (E_{nr}^{(1)} - V_1)|\phi_{nr}^{(1)}\rangle$$

Taking the scalar product on both sides with $\langle \phi_{nm}^{(0)} |$ gives

$$\begin{aligned} 0 &= E_n^{(2)}\delta_{mr} - \langle \phi_{nm}^{(0)} | V_2 | \phi_{nr}^{(0)} \rangle + E_{nr}^{(1)} \langle \phi_{nm}^{(0)} | \phi_{nr}^{(1)} \rangle \\ &\quad - \langle \phi_{nm}^{(0)} | V_1 | \phi_{nr}^{(1)} \rangle \end{aligned}$$

This is an inconsistent equation since for $m \neq r$, we get several different relations between the potentials V_1 and V_2 . It is therefore not clear what the approach to second order degenerate perturbation theory should be.

6.32 Lecture plan for electromagnetic field theory EC-C09

- [1] Basics of vector analysis and vector calculus (2)
 - [a] Scalar and vector fields in space in Cartesian co-ordinates. (1)
 - [b] Transformation of coordinates between cartesian, cylindrical and spherical polar coordinates. (1)
 - [c] Transformation of unit vectors in the three systems. (2)
 - [d] Gauss' and Stokes integral theorems with proofs, Computing the flux of a vector field from a closed surface using Gauss' law and directly, computing the circulation of a vector field around a closed wire and using Stokes' theorem. (3)

- [2] Electrostatics
 - [a] Coulomb's law for the electric field produced by static charges and the principle of superposition for electric fields. (1)
 - [b] Coulomb's law and derivation of Gauss' law of electrostatics from Coulomb's law; Derivation of Coulomb's law from Gauss' law assuming spherical symmetry. Gauss' law in integral and differential form. (2)
 - [c] Boundary conditions for the electrostatic field. The physical meaning of the boundary conditions in terms of surface charge density and potential. (1)
 - [d] The electrostatic potential, derivation from the irrotationality of the static electric field. (1)
 - [e] Work done in transfer of a charge in an electric field and its relation to the electrostatic potential. (1)
 - [f] The energy contained in an electric field in space. (1)
 - [g] Poisson's equation for the electrostatic field in the presence of a static charge distribution, Solving Laplace's equation with boundary conditions of various kinds, spherical surfaces, discs and concentric discs, rectangular boundary. (1)
 - [h] Uniqueness theorem for the Poisson equation with Dirichlet, Neumann and mixed boundary conditions. (1)
 - [i] Green's function for the Poisson equation and solution of Poisson's equation within a boundary with distributed charge density and Dirichlet/Neumann/mixed boundary conditions on the boundary. Expressing the solution to Poisson's equation in terms of the Green's function. (1)
 - [j] Examples of Green's function computation: Rectangular box in two and three dimensions, sphere region between two concentric spheres, sector, disk in the plane, annulus between two concentric circles. Cylinder, space between two concentric cylinders in three dimensions. (2)
 - [k] The electric field of an infinitesimal dipole. (1)
 - [l] Dielectrics and polarization. Charge density and surface charge densities associated with the polarization field. The electric displacement vector D , dielectric constant, susceptibility, Gauss law for the displacement vector in terms of free charges. Capacitance between two surfaces with constant and spatially

varying permittivity. Boundary conditions for tangential components of E and normal component of D at the interface between two surfaces. (2)

[3] Magnetostatics.

[a] Ampere's law and the no magnetic monopole condition in integral and differential form. (1)

[b] The magnetic vector potential (1)

[c] Derivation of the Biot-Savart law (1)

[d] Magnetic circuits. (1)

[e] Calculating the magnetic field for various configurations like straight wire, circular loop, infinite sheet carrying a constant surface current density. (2)

[4] Faraday's law of induction. Computing the induced emf for various problems, applications to electromagnetic brakes, induction heating and eddy current heating. (2)

[5] The equation of continuity, ie, charge conservation for time dependent fields. (1)

[6] The displacement current correction term to Ampere's law as a remedy to the violation of charge conservation. (1)

[7] Electromagnetic waves in free space. (1)

[8] Reflection and refraction of em waves at the boundary between two dielectrics, Snell's laws of reflection and refraction. (2)

[9] The complete solution to the Maxwell equations using retarded potential theory after application of the Lorentz gauge conditions. (2)

[10] Poynting's theorem on the flux of power and energy density in the electromagnetic field. (1)

[11] Other topics in electromagnetics

[a] An introduction to far field radiation patterns produced by Antennas. (2)

[b] An introduction to waveguides and cavity resonators. (2)

[c] Multipole expansion of the static electrostatic potential. (1)

[d] Multipole radiation from a system of time varying charges and currents, Basics of antenna theory. (2)

[e] Radiation by accelerating charges: The Lienard-Weichert potentials. (2)

[f] Quantization of the em field within a cavity resonator. Interaction of the quantum em field with an electron bound to an atom described by Dirac's relativistic wave equation.

[g] Quantum electrodynamics in free space and within a cavity resonator: Interaction between electrons, positrons and photons within a cavity resonator.

[h] Design of very large size quantum gates from the interaction of a quantum em field with a second quantized Dirac field.

Note: This schedule also covers tutorials. 42 lectures in all.

Lecture plan for "Linear Algebra in Signal Processing"

- [1] Fields, vector spaces, linear independence, basis, dimension of a vector space.
- [2] Co-ordinate representations of vectors relative to a basis, Matrix transformation of a vector relative to two bases.
- [3] Linear transformations on a vector space and from one vector space into another.
- [4] Matrix representation of a linear transformation relative to a basis and similarity transformation of the matrix of a linear transformation relative to two bases.
- [5] Range, nullspace, rank and nullity of a linear transformation.
- [6] Subspaces, direct sum of subspaces.
- [8] Invariant subspaces of a linear transformation, the matrix of a linear transformation relative to a basis that is an extension of a basis of an invariant subspace.

6.33 A digression into infinite dimensional vector spaces

- [9] Infinite dimensional vector spaces, examples.
 - [a] Cauchy sequences, Banach and Hilbert spaces
 - [b] Bounded linear operators in Banach and Hilbert spaces.
 - [c] Compact operators in a Banach space.
 - [d] Resolvent and spectrum of a linear operator in a Banach space.
 - [e] Spectra of compact operators in a Banach space.
 - [f] Graph of an operator in a Banach space, closed and closable operators in a Banach space, Closure of a closable operator.
 - [g] Densely defined operators in a Hilbert space, the adjoint of a densely defined operator, uniqueness of the adjoint, symmetric operators, self-adjoint operators.
 - [h] Orthogonal projections in a Hilbert space (existence and uniqueness), Spectral theorem for compact, bounded and unbounded self-adjoint operators in a Hilbert space.
 - [i] Bounded linear functionals on a Hilbert space and the Riesz representation.
 - [j] Functions of a self-adjoint operator in a Hilbert space in terms of spectral integrals
 - [k] Applications of the spectral theorem to quantum scattering theory, existence of wave operators, completeness of the wave operators, scattering matrix and explicit formula for the scattering matrix in the momentum domain as a unitary operator in $L^2(S^2)$ in terms of the free particle Hamiltonian and the scattering potential.

$$X = \int_0^\infty \| V(Q)U_t^0(P)f \| .dt < \infty$$

Here,

$$U_t^0(P) = \exp(-itP^2)$$

$$U_{-t}^0 Q U_t^0 = \exp(itad(P^2))(Q) = Q + 2Pt$$

Let

$$Z_t(Q) = \exp(-iQ^2/4t)$$

Then,

$$Z_t P Z_t^* = \exp(-iad(Q^2)/4t)(P) = P + Q/2t$$

so

$$Z_t V(2Pt) Z_t^* = V(Q + 2Pt) = U_{-t}^0 V(Q) U_t^0$$

and hence,

$$X = \int_0^\infty V(2Pt) Z_t(Q)^* f \| dt$$

6.34 Continuation of finite dimensional vector spaces

- [l] Proof of the primary decomposition theorem.
- [m] Proof of the Jordan decomposition theorem and algorithms for computation of the Jordan factors.
- [n] Eigenvalues, eigenvectors, characteristic and minimal polynomials of a matrix, algebraic and geometric multiplicity of an eigenvalue, criteria for diagonalability of a linear operator spectral theorem for finite dimensional normal matrices.
- [o] Functions of a linear operator using Cauchy's integral theorem for functions of a complex variable.
- [p] Calculation of the canonical projections on the eigenspaces in terms of Cauchy's integral theorem.
- [q] Perturbation theory for eigenvalues and eigenvectors of a matrix–The Rayleigh-Schrodinger theory.
- [r] Decomposition theorems of matrix theory.
- [1] The spectral decomposition
- [2] The polar decomposition
- [3] The singular value decomposition
- [4] The QR decomposition based on Gram-Schmidt orthonormalization process.
- [5] The root space decomposition of a semisimple Lie algebra.
- [s] An introduction to group representation theory.
- [1] Definition of a Lie group and its Lie algebra as the tangent space at the identity.
- [2] The Lie algebra viewed as the class of left invariant vector fields on the Lie group.
- [3] The Lie bracket as commutator of vector fields.
- [4] Vector fields and one forms on a differentiable manifold.

- [5] Transformation of tangent vectors and vector fields on a manifold under a diffeomorphism.
- [6] Transformation of one forms under a diffeomorphism.
- [7] Tensor fields on a manifold and their transformation under a diffeomorphism.
- [8] General theory of skewsymmetric forms on a manifold.
- [9] The differential of a skewsymmetric form.
- [10] The Lie derivative of a vector field, of a one form and then of a general tensor field.
- [11] Basics of Riemannian geometry: Parallel translation, covariant derivative, connection, torsion and curvature of a connection, Cartan's equations of structure.
- [3] [a] Interaction of quantized cavity resonator fields with the Dirac electron and then with the second quantized Dirac field.
- [b] Effect of gravity on the quantized em field in a cavity resonator.

6.35 The general relativistic Maxwell equations in a resonator

$$(F^{\mu\nu}\sqrt{-g}),_{\nu} = 0$$

These equations are general relativistic versions of the usual Maxwell equations

$$\operatorname{curl} E = -\mu H_{,t}, \operatorname{curl} B = \mu \epsilon E_{,t}$$

We define the electric field components as

$$E_r = -F^{0r} = F^{r0}, r = 1, 2, 3$$

and the magnetic field as

$$B_1 = -F^{23}, B_2 = -F^{31}, B^3 = -F^{12}$$

Then the general relativistic Maxwell equations are

$$(E_r \sqrt{-g}),_r = 0,$$

$$\begin{aligned} (B_1 \sqrt{-g}),_3 - (B_3 \sqrt{-g}),_1 &= \mu \epsilon (E_2 \sqrt{-g}),_t, \\ (B_2 \sqrt{-g}),_1 - (B_1 \sqrt{-g}),_2 &= \mu \epsilon (E_3 \sqrt{-g}),_t \\ (B_3 \sqrt{-g}),_2 - (B_2 \sqrt{-g}),_3 &= \mu \epsilon (E_1 \sqrt{-g}),_t \end{aligned}$$

We shall instead define our electric and magnetic fields as

$$E_r = F_{0r} = A_{r,0} - A_{0,r}, B_1 = -F_{23} = -A_{3,2} + A_{2,3},$$

$$B_2 = -F_{31} = -A_{1,3} + A_{3,1}, B_3 = -F_{12} = A_{1,2} - A_{2,1}$$

Then, the Maxwell equations

$$F_{\mu\nu,\rho} + F_{\nu\rho,\mu} + F_{\rho\mu,\nu} = 0, (F^{\mu\nu}\sqrt{-g})_{,\nu} = (g^{\mu\alpha}g^{\nu\beta}\sqrt{-g}F_{\alpha\beta})_{,\nu} = 0$$

The first of these gives

$$\operatorname{curl} E + B_{,t} = 0$$

or equivalently,

$$\operatorname{curl} E + \mu H_{,t} = 0$$

ie this equation is unaffected by gravitation. The second gives

$$(F^{r0}\sqrt{-g})_{,0} + (F^{rs}\sqrt{-g})_{,s} = 0$$

or equivalently,

$$(g^{\mu r}g^{\nu 0}\sqrt{-g}F_{\mu\nu})_{,0} + (g^{\mu r}g^{\nu s}\sqrt{-g})_{,s} = 0$$

For weak gravitational fields described by small metric perturbations as

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}(x)$$

we get upto $O(h\mu\nu)$, the system

$$\operatorname{curl} H + \epsilon E_{,t} = X(t, r)$$

where $X(t, r) = X(x)$ is bilinear in (E, H) and $h_{\mu\nu}$ and their space-time partial derivatives provided that we retain term only upto linear orders in $h_{\mu\nu}$ and its partial derivatives, ie, $X_k(x)$ is of the form

$$C_1(km\mu\nu\rho)(E_m(x)h_{\mu\nu}(x))_{,\rho} + C_2(km\mu\nu\rho)(B_mh_{\mu\nu})_{,\rho}$$

Now assume the metric to be static, ie, time independent and we look for solutions for the electromagnetic field that oscillate with a frequency of ω within the cavity resonator. That would enable us to apply perturbation theory for linear pde's to determine the shift in the resonator frequencies caused by gravitational effects. Writing

$$E = E^{(0)} + E^{(1)}, B = B^{(0)} + B^{(1)}$$

and

$$\omega = \omega_0 + \delta\omega,$$

where

$$\omega^{(0)} = \omega_0(mnp), E^{(0)} = E_{mnp}^{(0)}(\omega_0(mnp), r) = E_{mnp}^{(0)}(r)$$

$$H^{(0)} = H_{mnp}^{(0)}(\omega_0(mnp), r) = H_{mnp}^{(0)}(r)$$

we get the unperturbed equations as

$$\operatorname{curl} E_{mnp}^{(0)}(r) = -j\omega_0(mnp)\mu H_{mnp}^{(0)},$$

$$\operatorname{curl} H_{mnp}^{(0)}(r) = j\omega_0(mnp)\epsilon E_{mnp}^{(0)}(r)$$

The first order perturbed equations are

$$\operatorname{curl} E^{(1)} + j\omega_0 \mu H^{(1)} + j\delta\omega \cdot \mu \cdot H_{mnp}^{(0)} = 0$$

$$\operatorname{curl} H^{(1)} - j\omega_0 \epsilon E^{(1)} - j\delta\omega \epsilon E_{mnp}^{(0)} = X_{mnp}$$

where $X_{mnp}(\omega_0, r)$ is obtained by replacing (E, H) in X by $(E_{mnp}^{(0)}, H_{mnp}^{(0)})$ respectively. We now have

$$\langle E_{mnp}^{(0)}, \operatorname{curl} E^{(1)} \rangle = - \int_C \operatorname{div}(E_{mnp}^{(0)} \times E^{(1)}) d^3r + \langle \operatorname{curl} E_{mnp}^{(0)}, E^{(1)} \rangle$$

All the volume integrals are taken over C , the volume of the resonator. Now, by Gauss' integral theorem,

$$\int_C \operatorname{div}(E_{mnp}^{(0)} \times E^{(1)}) d^3r = \int_{\partial C} (\hat{n} \times E_{mnp}^{(0)}, E^{(1)}) dS = 0$$

since the tangential component of the unperturbed electric field vanishes of the boundary walls of the resonator. Also,

$$\langle \operatorname{curl} E_{mnp}^{(0)}, E^{(1)} \rangle = j\omega_0 \mu \langle H_{mnp}^{(0)}, E^{(1)} \rangle$$

Also,

$$\begin{aligned} \langle E_{mnp}^{(0)}, \operatorname{curl} H^{(1)} \rangle &= - \int_C \operatorname{div}(E_{mnp}^{(0)} \times H^{(1)}) d^3r + \langle \operatorname{curl} E_{mnp}^{(0)}, H^{(1)} \rangle \\ &= j\omega_0 \mu \langle H_{mnp}^{(0)}, H^{(1)} \rangle \end{aligned}$$

Thus, we get two equations

$$j\omega_0 \mu \langle H_{mnp}^{(0)}, E^{(1)} \rangle + j\omega_0 \mu \langle E_{mnp}^{(0)}, H^{(1)} \rangle + j\delta\omega \cdot \mu \cdot \langle E_{mnp}^{(0)}, H_{mnp}^{(0)} \rangle = 0 \quad (1),$$

$$\begin{aligned} j\omega_0 \mu \langle H_{mnp}^{(0)}, H^{(1)} \rangle - j\omega_0 \epsilon \langle E_{mnp}^{(0)}, E^{(1)} \rangle - j\delta\omega \epsilon \langle E_{mnp}^{(0)}, E_{mnp}^{(0)} \rangle \\ = \langle E_{mnp}^{(0)}, X_{mnp} \rangle \end{aligned} \quad (2)$$

Again by taking the inner product with $H_{mnp}^{(0)}$ instead of $E_{mnp}^{(0)}$ we get on using the fact that the tangential component of $E^{(1)}$ vanishes on the boundary,

$$-j\omega_0 \epsilon \langle E_{mnp}^{(0)}, E^{(1)} \rangle + j\omega_0 \mu \langle H_{mnp}^{(0)}, H^{(1)} \rangle$$

$$+ j\delta\omega \mu \langle H_{mnp}^{(0)}, H_{mnp}^{(0)} \rangle = 0 \quad (3),$$

$$-j\omega_0 \epsilon \langle E_{mnp}^{(0)}, H^{(1)} \rangle - j\omega_0 \epsilon \langle H_{mnp}^{(0)}, E^{(1)} \rangle$$

$$- j\delta\omega \epsilon \langle H_{mnp}^{(0)}, E_{mnp}^{(0)} \rangle = \langle H_{mnp}^{(0)}, X_{mnp} \rangle \quad (4)$$

Since $\langle E_{mnp}^{(0)}, H_{mnp}^{(0)} \rangle = 0$, it follows that eqns (1)-(4) can be rearranged as

$$j\omega_0 \mu \langle H_{mnp}^{(0)}, E^{(1)} \rangle + j\omega_0 \mu \langle E_{mnp}^{(0)}, H^{(1)} \rangle = 0 \quad (1')$$

$$\begin{aligned}
& j\omega_0\mu < H_{mnp}^{(0)}, H^{(1)} > - j\omega_0\epsilon < E_{mnp}^{(0)}, E^{(1)} > - j\delta\omega\epsilon < E_{mnp}^{(0)}, E_{mnp}^{(0)} > \\
& = < E_{mnp}^{(0)}, X_{mnp} > \dots (2') \\
& - j\omega_0\epsilon < E_{mnp}^{(0)}, E^{(1)} > + j\omega_0\mu < H_{mnp}^{(0)}, H^{(1)} > \\
& \quad + j\delta\omega\mu < H_{mnp}^{(0)}, H_{mnp}^{(0)} > = 0 \dots (3'), \\
& - j\omega_0\epsilon < E_{mnp}^{(0)}, H^{(1)} > - j\omega_0\epsilon < H_{mnp}^{(0)}, E^{(1)} > \\
& = < H_{mnp}^{(0)}, X_{mnp} > \dots (4')
\end{aligned}$$

Subtracting (2') from (3') gives us

$$\begin{aligned}
& j\delta\omega[\mu < H_{mnp}^{(0)}, H_{mnp}^{(0)} > + \epsilon < E_{mnp}^{(0)}, E_{mnp}^{(0)} >] \\
& = - < E_{mnp}^{(0)}, X_{mnp} >
\end{aligned}$$

which yields the desired formula for the shift in the frequency. Likewise from (1') and (4'), we get a consistency condition

$$< H_{mnp}^{(0)}, X_{mnp} > = 0 \dots (5)$$

(5) gives us a normalization condition, ie a scaling condition for $H_{mnp}^{(0)}$ to be satisfied in terms of $E_{mnp}^{(0)}$. More precisely, we can write X_{mnp} as a linear functional of $E_{3mnp}^{(0)}$ and $H_{3mnp}^{(0)}$ and then we replace $H_{3mnp}^{(0)}$ by $\lambda H_{3mnp}^{(0)}$ so that (5) is satisfied.

Remark: The unperturbed fields $E^{(0)}, H^{(0)}$ are linear superpositions of $Re(E_{mnp}^{(0)}(r)\exp(j\omega_{mnp}t))$ and $Re(H_{mnp}^{(0)}(r)\exp(j\omega_{mnp}t))$ and both of these fields can be expressed in terms of the z-components of these fields. This is a well known fact in waveguide-cavity resonator theory.

Quantization of the resonator fields in the presence of a background gravitational field. The Lagrangian density is

$$\mathcal{L} = K F_{\mu\nu} F^{\mu\nu} \sqrt{-g}$$

We choose our position fields as

$$A^\mu, \mu = 0, 1, 2, 3$$

Then,

$$\begin{aligned}
F_{\mu\nu} &= A_{\nu,\mu} - A_{\mu,\nu} = (g_{\nu\rho} A^\rho)_{,\mu} - (g_{\mu\rho} A^\rho)_{,\nu} \\
&= (g_{\nu\rho,\mu} - g_{\mu\rho,\nu}) A^\rho + g_{\nu\rho} A^\rho_{,\mu} - g_{\mu\rho} A^\rho_{,\nu}
\end{aligned}$$

We can define the canonical momentum fields as

$$\pi_\rho = \frac{\partial \mathcal{L}}{\partial A_{,0}^\rho} = 2K \sqrt{-g} F^{\mu\nu} \frac{\partial F_{\mu\nu}}{\partial A_{,0}^\rho}$$

and then by applying the Legendre transformation, the Hamiltonian density is

$$\mathcal{H} = \pi_\rho A_0^\rho - \mathcal{L}$$

Now,

$$\frac{\partial F_{\mu\nu}}{\partial A_{,0}^\rho} = g_{\nu\rho}\delta_\mu^0 - g_{\mu\rho}\delta_\nu^0$$

and hence,

$$\begin{aligned}\pi_\rho &= 2K\sqrt{-g}F^{\mu\nu}(g_{\nu\rho}\delta_\mu^0 - g_{\mu\rho}\delta_\nu^0) \\ &= 2K\sqrt{-g}(g_{\nu\rho}F^{0\nu} - g_{\mu\rho}F^{\mu 0}) \\ &= 4K\sqrt{-g}g_{\mu\rho}F^{0\mu}\end{aligned}$$

Thus, the Hamiltonian density is

$$\mathcal{H} = 4K\sqrt{-g}g_{\mu\rho}F^{0\mu}A_{,0}^\rho - K\sqrt{-g}F_{\mu\nu}F^{\mu\nu}$$

or

$$\mathcal{H}/K\sqrt{-g} = 4g_{\mu\rho}F^{0\mu}A_{,0}^\rho - F_{\mu\nu}F^{\mu\nu}$$

Alternately, suppose we use as our canonical position variables A_μ , we define the electric and magnetic fields by

$$E_r = A_{r,0} - A_{0,r} = F_{0r}, B_1 = A_{2,3} - A_{3,2} = -F_{23}, B_2 = -F_{31}, B_3 = -F_{12}$$

then the canonical momenta are

$$\begin{aligned}\pi^\rho &= \partial\mathcal{L}/\partial A_{\rho,0} = 2K\sqrt{-g}F^{\mu\nu}\partial F_{\mu\nu}/\partial A_{\rho,0} \\ &= 2K\sqrt{-g}F^{\mu\nu}(\delta_\nu^\rho\delta_\mu^0 - \delta_\mu^\rho\delta_\nu^0)\end{aligned}$$

so the Hamiltonian density now becomes

$$\begin{aligned}\mathcal{H} &= \pi^\rho A_{\rho,0} - \mathcal{L} = \\ &[2KF^{0\nu}A_{\nu,0} - 2KF^{\mu 0}A_{\mu,0} - KF_{\mu\nu}F^{\mu\nu}]\sqrt{-g} \\ &= K\sqrt{-g}[4F^{0\mu}A_{\mu,0} - F_{\mu\nu}F^{\mu\nu}]\end{aligned}$$

6.36 Mackey's theory on the construction of the basic observables in the quantum theory from projective unitary representations of the Galilean group

Galilean group element: $(a, u, s, g), (b, v, t, h)$ where $a, u \in \mathbb{R}^3, (s, g) \in \mathbb{R} \times SO(3)$. Let G be the Galilean group. Then

$$G = V \otimes_s H, H = \mathbb{R} \times SO(3), V = \mathbb{R}^3 \times \mathbb{R}^3$$

Buy (a, u, s, g) , we mean $(a, u)o(s, g)$. Its composition law is

$$((a, u)o(s, g))o((b, v)o(t, h)) = ((s, g)[(b, v)] + (a, u), (s, g)o(t, h))$$

where

$$(s, g)[(b, v)] = (s, g)o(b, v)o(s, g)^{-1} = (s, g)o(b, v)o(-s, g^{-1})$$

$$\begin{aligned} (s, g)o(b, v)o(-s, g^{-1})(t, r) &= (s, g)o(b, v).(t-s, g^{-1}r) \\ &= (s, g)o(t-s, g^{-1}r+v(t-s)+b) \\ &= (t, r+gv(t-s)+gb) = (g(b-vs), gv)(t, r) \end{aligned}$$

Thus, we get

6.37 Hamiltonian density of the electromagnetic field in curved space-time in terms of position and momentum fields

$$\begin{aligned} \mathcal{L} &= K\sqrt{-g}F_{\mu\nu}F^{\mu\nu} \\ \pi^\rho &= \partial\mathcal{L}/\partial A_{\rho,0} = \\ &\quad 4K\sqrt{-g}F^{0\rho} \\ \mathcal{H} &= \pi^\rho A_{\rho,0} - \mathcal{L} = \\ &\quad K\sqrt{-g}(4F^{0\rho}A_{\rho,0} - F_{\mu\nu}F^{\mu\nu}) \\ &= K\sqrt{-g}(4F^{0m}A_{m,0} - F_{\mu\nu}F^{\mu\nu}) \end{aligned}$$

Now,

$$\begin{aligned} \pi^\rho &= 4K\sqrt{-g}g^{0\alpha}g^{\rho\beta}F_{\alpha\beta} \\ &= 4K\sqrt{-g}[g^{00}g^{0m}F_{0m} + g^{0m}g^{00}F_{m0} + g^{0m}g^{\rho k}F_{mk}] \\ &= 4K\sqrt{-g}[(g^{00}g^{0m} - g^{0m}g^{00})F_{0m} + g^{0m}g^{\rho k}F_{mk}] \end{aligned}$$

We need to solve for $A_{\rho,0}$ in terms of $A_{\rho,m}$ and π^m . Note that $\pi^0 = 0$. This is one of the constraints to be incorporated using the Dirac bracket formalism. We write the above equation as

$$\pi^s = 4K\sqrt{-g}[(g^{00}g^{sm} - g^{0m}g^{s0})F_{0m} + g^{0m}g^{sk}F_{mk}]$$

We define the 3×3 matrix

$$\gamma = ((\gamma^{sm})), \gamma^{sm} = g^{00}g^{sm} - g^{0m}g^{s0}$$

and then write the above as

$$\pi^s/4K\sqrt{-g} = \gamma^{sm}F_{0m} + g^{0m}g^{sk}F_{mk}$$

Now writing $((\gamma_{sm}))$ as the inverse of the matrix $((\gamma^{sm}))$, we get

$$F_{0s} = \gamma_{sm}[\pi^m/4K\sqrt{-g} - g^{0r}g^{km}F_{rk}]$$

Now we define the electric and magnetic fields as

$$E_s = F_{0s}, B_s = -\epsilon(skm)F_{km}$$

Then, we can write

$$E_s = \gamma_{sm}[\pi^m/4K\sqrt{-g} + g^{0r}g^{km}\epsilon(rkl)B_l]$$

or equivalently,

$$\pi^s/4K\sqrt{-g} = \gamma^{sm}E_m - g^{0m}g^{sk}\epsilon(mkl)B_l$$

We then have

$$\begin{aligned} \mathcal{H} &= \pi^s A_{s,0} - \mathcal{L} = \\ K\sqrt{-g}[(\gamma^{sm}E_m - g^{0m}g^{sk}\epsilon(mkl)B_l)A_{s,0} - F_{\mu\nu}F^{\mu\nu}] \end{aligned}$$

We note that

$$\begin{aligned} \int \pi^\rho A_{\rho,0} d^3x &= \int \pi^s A_{s,0} d^3x = \\ \int 4K\sqrt{-g}F^{0s}A_{s,0}d^3x &= 4K \int \sqrt{-g}F^{0s}(A_{s,0} - A_{0,s} + A_{0,s})d^3x \\ &= 4K \int F^{0s}F_{0s}\sqrt{-g}d^3x + 4K \int \sqrt{-g}F^{0s}A_{0,s}d^3x \\ &= 4K \int F^{0s}F_{0s}\sqrt{-g}d^3x + 4K \int (F^{0s}A_0\sqrt{-g}),_s d^3x - 4K \int (F^{0s}\sqrt{-g}),_s A_0 d^3x \\ &= 4K \int F^{0s}F_{0s}\sqrt{-g}d^3x \end{aligned}$$

in the absence of charges since in the absence of charges, the Maxwell equations read

$$(F^{0s}\sqrt{-g}),_s = 0$$

and also we have noted that the spatial volume integral of a three divergence is zero by Gauss' theorem. This results in

$$\mathcal{H} = 4KF^{0s}F_{0s}\sqrt{-g} - KF_{\mu\nu}F^{\mu\nu}\sqrt{-g}$$

$$K\sqrt{-g}(4F^{0s}F_{0s} - F_{\mu\nu}F^{\mu\nu}) = K\sqrt{-g}(2F^{0s}F_{0s} - F_{rs}F^{rs})$$

Noting that

$$F_{0s} = E_s, F_{rs} = -\epsilon(rsm)B_m$$

we have

$$\begin{aligned} F^{0s} &= (g^{00}g^{sm} - g^{0m}g^{0s})F_{0m} + g^{0k}g^{sm}F_{km} \\ &= \gamma^{sm}E_m + g^{0k}g^{sm}\epsilon(kml)B_l \end{aligned}$$

Also

$$\begin{aligned} F^{rs} &= (g^{rm}g^{s0} - g^{r0}g^{sm})F_{m0} + g^{rm}g^{sk}F_{mk} \\ &= -\beta^{rms}E_m - g^{rm}g^{sk}\epsilon(mkp)B_p \end{aligned}$$

where

$$\beta^{rms} = g^{rm}g^{s0} - g^{r0}g^{sm}$$

Thus,

$$\begin{aligned} \mathcal{H} &= K\sqrt{-g}(2F^{0s}F_{0s} - F_{rs}F^{rs}) \\ &= 2K\sqrt{-g}(\gamma^{sm}E_m + g^{0k}g^{sm}\epsilon(kml)B_l)E_s \\ &\quad - K\sqrt{-g}\epsilon(rsm)(\beta^{rls}E_l + g^{rl}g^{sk}\epsilon(lkp)B_p)B_m \end{aligned}$$

Exercise: [1] Verify that in the case of flat Minkowskian space-time, this expression reduces to the familiar special relativistic formula

$$\mathcal{H} = (1/2)(E^2 + B^2)$$

provided that we take $K = -1/4$. Note that in flat space-time, $\gamma^{sm} = -\delta_{sm} = g^{sm}$, $g^{00} = 1$, $g^{0s} = 0$.

(2) Denote

$$\mathcal{H}_0 = (1/2)(E^2 + B^2)$$

Then show that if the metric is

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + h_{\mu\nu}(x)$$

i.e. represented as a small perturbation of flat space-time, we can then write

$$\mathcal{H} = \mathcal{H}_0 + C_1(rs, x)E_rE_s + C_2(rs, x)B_rB_s + C_3(rs, x)E_rB_s$$

where $C_j(rs, x)$ are linear combinations of $h_{\mu\nu}(x)$. Now we are in a position to express the Hamiltonian in terms of the canonical position and momentum fields. Note that the B_s are functionals of the position fields $A_r(x)$ while the equation

$$\pi^s/4K\sqrt{-g} = \gamma^{sm}E_m - g^{0m}g^{sk}\epsilon(mkl)B_l$$

gives

$$E_s = \gamma_{sm}(-\pi^s/\sqrt{-g} + g^{0m}g^{sk}\epsilon(mkl)B_l)$$

Exercise: Using the above equations, express the Hamiltonian density in terms of the position and momentum fields $A_s, \pi^s, s = 1, 2, 3$ in a curved space-time upto linear orders in the metric perturbations.

6.38 Coulomb scattering

The wave operator Ω_+ does not exist. In fact, writing

$$H_0 = P^2/2m, V(Q) = Ze^2/|Q|$$

we find that

$$\int_0^\infty \|V(Q)\exp(-itH_0)f\| dt$$

is infinite. In fact, we have

$$\exp(itH_0)Q.\exp(-itH_0) = \exp(it.ad(P^2)/2m)(Q) = Q + Pt/m$$

and on the other hand,

$$\exp(-imQ^2/2t)P.\exp(imQ^2/2t) = \exp(-im.ad(Q^2)/2t)(P)$$

$$= P + mQ/t = (m/t)(Q + Pt/m)$$

Thus,

$$\begin{aligned} \exp(itH_0)V(Q).\exp(-itH_0) &= V(Q + Pt/m) = V((t/m)(P + mQ/t)) \\ &= Z_t V(Pt/m) Z_t^* \end{aligned}$$

where

$$Z_t = Z_t(Q) = \exp(-imQ^2/2t)$$

Thus,

$$\begin{aligned} \|V(Q).\exp(-itH_0)f\| &= \|V(Pt/m).Z_t^*f\| \\ &\leq \|V(Pt/m)\langle Q \rangle^{-n}\| \cdot \|\langle Q \rangle^n f\| \end{aligned}$$

where

$$\langle Q \rangle = (1 + Q^2)^{1/2}$$

Assuming that

$$\int (1 + Q^2)^n |f(Q)|^2 d^3Q < \infty$$

It follows that the wave operator Ω_+ will exist provided that

$$t \rightarrow \|V(Pt/m)\langle Q \rangle^{-n}\|$$

is integrable on $[0, \infty)$. Now for any $q \geq 2$, we have

$$\rightarrow \|V(Pt/m)\langle Q \rangle^{-n}\| \leq \|V(Qt/m)\|_q \cdot \|\langle Q \rangle^{-n}\|_q$$

We have

$$\|\langle Q \rangle^{-n}\|_q = \left(\int (1 + Q^2)^{-nq/2} d^3Q \right)^{1/q}$$

and for this integral to be finite, we require that $2 - nq/2 < -1$ or equivalently, $nq > 6$. On the other hand, for $t \rightarrow \|V(Qt/m)\|_q$ to be integrable over $[0, \infty)$, we can derive the condition:

$$\left(\int |V(Qt/m)|^q d^3Q \right)^{1/q} = (m^3/3t^3)^{1/q} \left(\int |V(Q)|^q d^3Q \right)^{1/q}$$

and this is integrable if $t^{-3/q}$ is integrable over (δ, ∞) . This will happen provided that $3/q > 1$ or equivalently, $q < 3$ and also $V \in L^q(\mathbb{R}^3)$. Thus for the existence of the wave operator Ω_+ acting on a function $f(Q)$, we require three conditions:

Scattering matrix in quantum mechanics

$$\begin{aligned}\Omega_+ &= I + i \int_0^\infty U(-t)VU^0(t)dt \\ \Omega_+^* &= I - i \int_0^\infty U^0(-t)VU(t)dt \\ \Omega_- &= I - i \int_{-\infty}^0 U(-t)VU^0(t)dt \\ &= I - i \int_0^\infty U(t)VU^0(-t)dt \\ \Omega_-^* &= I + i \int_0^\infty U^0(t)VU(-t)dt = I + i \int_{-\infty}^0 U^0(-t)VU(t)dt \\ \Omega_+^* - \Omega_-^* &= - \int_{-\infty}^\infty U^0(-t)VU(t)dt \\ R = S - I &= \Omega_+^* \Omega_- - \Omega_-^* \Omega_- = (\Omega_+^* - \Omega_-^*) \Omega_- \\ &= - \int_{\mathbb{R}} U^0(-t)VU(t)\Omega_- dt = - \int_{\mathbb{R}} U^0(-t)V\Omega_- U^0(t)dt\end{aligned}$$

where we have used

$$U(t)\Omega_- = \Omega_- U^0(t)$$

Thus,

$$\begin{aligned}-R &= \int_{\mathbb{R}} U^0(-t)V(I - i \int_0^\infty U(s)VU^0(-s)ds)U^0(t)dt \\ &= \int_{\mathbb{R}} U^0(-t)VU^0(t)dt - i \int_{\mathbb{R} \times \mathbb{R}_+} U^0(-t)VU(s)VU^0(t-s)dtds \\ &\quad - i \int_{\mathbb{R}^3 \times \mathbb{R}_+} E_0(d\lambda)V \exp(-isH)VE_0(d\mu) \exp(i(\lambda - \mu)t + i\lambda s)dt ds \\ &= 2\pi \int_{\mathbb{R}^2} \delta(\lambda - \mu)E_0(d\lambda)VE_0(d\mu) \\ &\quad - 2\pi \int_{\mathbb{R}^2} \delta(\lambda - \mu)E_0(d\lambda)V(H - \lambda)^{-1}.VE_0(d\mu)\end{aligned}$$

Equivalently,

$$-R/2\pi = \int_{\mathbb{R}} E'_0(\lambda) V E'_0(\lambda) d\lambda - \int_{\mathbb{R}} E'_0(\lambda) V \cdot (H - \lambda)^{-1} \cdot V \cdot E'_0(\lambda) d\lambda$$

This gives us explicitly the form of the scattering matrix at energy λ :

$$S(\lambda) - I = R(\lambda) = -2\pi(E'_0(\lambda)(V - V \cdot (H - \lambda)^{-1}V)E'_0(\lambda))$$

It is clear that the matrix element $\langle f | S(\lambda) | g \rangle$ exists for all λ if and only if f, g belong to the absolutely continuous spectrum of H_0 , for that would imply the existence of the derivatives $E'_0(\lambda)f = dE_0(\lambda)f/d\lambda$ and $E'_0(\lambda)g = dE_0(\lambda)g/d\lambda$.

6.39 Electromagnetic waves in the Schwarzschild metric

$$g_{00} = \alpha(r) = 1 - 2m/r, g_{11}(r) = -\alpha(r)^{-1}, g_{22} = -r^2, g_{33} = -r^2 \sin^2(\theta)$$

The Maxwell equations are

$$(F^{\mu\nu}\sqrt{-g}),_{\nu} = 0$$

$$F^{\mu\nu} = g^{\mu\mu}g^{\nu\nu}F_{\mu\nu}$$

(No summation). So,

$$\sum_{\nu}(g^{\mu\mu}g^{\nu\nu}\sqrt{-g}F_{\mu\nu}),_{\nu} = 0$$

or

$$\sum_{\nu}(g^{\mu\mu}g^{\nu\nu}\sqrt{-g}(A_{\nu,\mu} - A_{\mu,\nu})),_{\nu} = 0$$

or

$$\begin{aligned} & \sum_{\nu}(g^{\mu\mu}g^{\nu\nu}\sqrt{-g}),_{\nu}(A_{\nu,\mu} - A_{\mu,\nu}) \\ & + \sum_{\nu}(g^{\mu\mu}g^{\nu\nu}\sqrt{-g})(A_{\nu,\mu\nu} - A_{\mu,\nu\nu}) = 0 \end{aligned}$$

The gauge condition used is

$$(A^{\mu}\sqrt{-g}),_{\mu} = 0$$

or

$$(g^{\mu\mu}\sqrt{-g}A_{\mu}),_{\mu} = 0$$

or

$$(g^{\nu\nu}\sqrt{-g}),_{\nu}A_{\nu} + g^{\nu\nu}\sqrt{-g}A_{\nu,\nu} = 0$$

Differentiating this equation gives

$$(g^{\nu\nu}\sqrt{-g}),_{\nu\mu}A_{\nu} + (g^{\nu\nu}\sqrt{-g}),_{\nu}A_{\nu,\mu}$$

$$+(g^{\nu\nu}\sqrt{-g})_{,\mu}A_{\nu,\nu} + g^{\nu\nu}\sqrt{-g}A_{\nu,\mu\nu} = 0$$

Thus, our wave equation becomes

$$\begin{aligned} & \sum_{\nu}(g^{\mu\mu}g^{\nu\nu}\sqrt{-g})_{,\nu}(A_{\nu,\mu} - A_{\mu,\nu}) \\ & - \sum_{\nu}(g^{\mu\mu}g^{\nu\nu}\sqrt{-g})A_{\mu,\nu\nu} \\ & - \sum_{\nu}g^{\mu\mu}(g^{\nu\nu}\sqrt{-g})_{,\mu}A_{\nu,\nu} \\ & - \sum_{\nu}(g^{\nu\nu}\sqrt{-g})_{,\nu\mu}g^{\mu\mu}A_{\nu} \\ & - \sum_{\nu}(g^{\nu\nu}\sqrt{-g})_{,\nu}g^{\mu\mu}A_{\nu,\mu} = 0 \end{aligned}$$

which simplifies to

6.40 Tutorial problems in electromagnetic field theory

[1] Describe the motion of n bar magnets moving under their mutual interaction.

hint: Let m_k denote the magnetic moment of the k^{th} magnet and assume that at time t , its centre of mass located at $R_k(t)$. Assume that it has suffered a rotation $S_k(t)$ around its centre of mass at time t relative to its orientation \hat{l}_k at time 0. Then, its kinetic energy at time t is given by

$$T_k(t) = M_k|R'_k(t)|^2/2 + (1/2)\xi'_k(t)^T J_k(\xi_k(t))\xi'_k(t)$$

where $\xi_k(t) = (\phi_k(t), \theta_k(t), \psi_k(t))$ is the Euler angle triplet suffered by the k^{th} magnet. We can write

$$S_k(t) = R_z(\phi_k(t))R_x(\theta_k(t))R_z(\psi_k(t)) = R(\phi_k(t), \theta_k(t), \psi_k(t)) = R(\xi_k(t))$$

and if V_k is the region of space occupied by the k^{th} magnet at time $t = 0$, then the rotational part of the kinetic energy of the k^{th} magnet is

$$\begin{aligned} T_k^R(t) &= (\rho_k/2) \int_{V_k} \| S'_k(t)r \|^2 d^3r \\ &= (1/2)Tr(S'_k(t)\eta_k S'_k(t)^T), \eta_k = \rho_k \int_{V_k} rr^T d^3r \end{aligned}$$

and ρ_k is the density of the k^{th} magnet. Note that

$$S'_k(t) = (\xi'_k(t)^T \nabla_{\xi})R(\xi_k(t)) = \xi'_{km}(t)R_{,m}(\xi_k(t))$$

so that

$$T_k^R(t) = (1/2)\xi'_{km}(t)\xi'_{kp}(t)Tr(R_{,m}(\xi_k(t))\eta_k R_{,p}(\xi_k(t))^T)$$

with summation over the repeated indices m, p being implied. Thus,

$$J_k(\xi_k) = ((Tr(R_{,m}(\xi_k)\eta_k R_{,p}(\xi_k)^T)))_{1 \leq m, p \leq 3}$$

The total kinetic energy of the system of magnets is

$$T(t) = \sum_{k=1}^n M_k |R'_k(t)|^2 / 2 + (1/2) \sum_{k=1}^n Tr(\xi'_k(t)^T J_k(\xi_k(t)) \xi_k(t))$$

The total potential energy of interaction of the system of n magnets is given by

$$V(t) = \sum_{1 \leq k < j \leq n} mu^{-1} \int (B_k(t, r), B_j(t, r)) d^3r$$

where $B_k(t, r)$ is the magnetic field produced by the k^{th} magnet. It is given by (in a non-relativistic approximation)

$$\begin{aligned} B_k(t, r) &= \nabla \times \mu m_k (S_k(t) l_k) \times (r - R_k(t)) / 4\pi |r - R_k(t)|^3 \\ &= \nabla \times A_k(t, r) \end{aligned}$$

where $A_k(t, r)$ is the magnetic vector potential generated by the k^{th} magnet:

$$A_k(t, r) = \mu m_k (S_k(t) l_k) \times (r - R_k(t)) / 4\pi |r - R_k(t)|^3$$

We note that integration by parts or equivalently application of Gauss' integral theorem gives

$$\begin{aligned} \int (B_k, B_j) d^3r &= \int (\text{curl } A_k, B_j) d^3r = \int \text{div}(A_k \times B_j) d^3r - \int (A_k, \text{curl } B_j) d^3r \\ &= - \int (A_k, \text{curl } B_j) d^3r \end{aligned}$$

The interaction potential energy of the system can be expressed as

$$\begin{aligned} V &= (\mu/16\pi^2) \sum_{1 \leq k < j \leq n} m_k m_j (\text{curl}(S_k(t) l_k \times (r - R_k(t))) / |r - R_k(t)|^3), \\ &\quad \text{curl}(S_j(t) l_j (r - R_j(t))) / |r - R_j(t)|^3 d^3r \end{aligned}$$

This is a function of the Euler angles $\xi_k(t)$, $k = 1, 2, \dots, n$ and also of the positions $R_k(t)$, $k = 1, 2, \dots, n$ of the centres of masses of the magnets. Thus, one can in principle set up the Lagrangian $T(t) - V(t)$ of the system as a function of $\{R_k, \xi_k, R'_k, \xi'_k, k = 1, 2, \dots, n\}$ and write down the equations of motion of the system.

6.41 The gravitational n-body problem in general relativity: an approximate treatment

The energy-momentum tensor of the system is given by

$$\begin{aligned} T^{\mu\nu}(x) &= \sum_{i=1}^n m_i \delta^3(x - x_i) (-g(x))^{-1/2} \frac{dx_i^\mu}{d\tau_i} \frac{dx_i^\nu}{d\tau_i} \frac{d\tau_i}{cdt} \\ \int T^{\mu\nu}(x) \sqrt{-g(x)} d^4x &= \sum_i m_i \int \frac{dx_i^\mu}{d\tau_i} \frac{dx_i^\nu}{d\tau_i} d\tau_i \\ &= \sum_i m_i \int \frac{dx_i^\mu}{d\tau_i} dx_i^\nu \end{aligned}$$

which is an invariant. We assume the metric generated by this n -body system to be $g_{\mu\nu}(x)$ and of the form

$$\begin{aligned} d\tau^2 &= (1 + 2\phi(x)/c^2 + h_{00}(x)/c^4)(cdt)^2 - [(1 - 2\phi(x)/c^2)\delta_{km} + h_{km}(x)/c^4]dx^k dx^m \\ &\quad + 2h_{0k}(x)cdt.dx^k/c^3 \end{aligned}$$

where

$$x^0 = ct, x^1 = x, x^2 = y, x^3 = z$$

Thus,

$$\begin{aligned} g_{00} &= 1 + 2\phi/c^2 + h_{00}/c^4, g_{km} = -(1 - 2\phi/c^2)\delta_{km} - h_{km}/c^4, \\ g_{0k} &= h_{0k}/c^3 \end{aligned}$$

We calculate

$$\begin{aligned} -g &= (1 + 2\phi/c^2)(1 - 2\phi/c^2)^3 + O(1/c^4) = 1 - 4\phi/c^2 + O(1/c^4) \\ (-g)^{-1/2} &= 1 + 2\phi/c^2 \end{aligned}$$

Also,

$$\begin{aligned} d\tau/cdt &= (1 + 2\phi/c^2 - (1 - 2\phi/c^2)v^2/c^2)^{1/2} + O(1/c^3) \\ &= 1 + \phi/c^2 + O(1/c^3) \end{aligned}$$

Then,

$$(-g)^{-1/2}.d\tau/dt = (1 + 2\phi/c^2)(1 + \phi/c^2) + O(1/c^3) = 1 + 3\phi/c^2 + O(1/c^3)$$

Taking $K = -8\pi G/c^2$, we find that firstly the metric tensor coefficients $g_{\mu\nu}$ are dimensionless and hence the Ricci tensor $R_{\mu\nu}$ has dimensions of $1/d^2$ where d is length and $K.T_{\mu\nu}$ has the units of $(G/c^2) \times \text{mass density} = G\rho/c^2$. This in turn has the units of GM/d^3c^2 which can be expressed as $(GM/d^2)(1/dc^2)$ which

has units of g/dc^2 , ie, $(d/t^2 d)/(d/t)^2 = 1/d^2$, ie $1/d^2$. Thus, $K T_{\mu\nu}$ and $R_{\mu\nu}$ have the same units and hence, the Einstein field equations are

$$R_{\mu\nu} = K(T_{\mu\nu} - Tg_{\mu\nu}/2)$$

where

$$T = g_{\mu\nu} T^{\mu\nu} = \sum_i m_i (-g(x))^{-1/2} \delta^3(x - x_i) d\tau_i/dt$$

We write

$$S_{\mu\nu} = T_{\mu\nu} - Tg_{\mu\nu}/2$$

Then,

$$S_{\mu\nu} = \sum_i m_i (-g)^{-1/2} \delta^3(x - x_i) (u_{i\mu} u_{i\nu} - (d\tau_i/cdt) g_{\mu\nu}/2)$$

Now, with neglect of $O(1/c^3)$ terms,

$$\begin{aligned} (u_i^0 u_i^0 - (d\tau_i/cdt) g_{00}/2)(-g)^{-1/2} &= (cdt/d\tau_i)^2 - (g_{00}/2)d\tau_i/cdt)(-g)^{-1/2} \\ &= [(1 - \phi/c^2)^2 - (1/2)(1 + 2\phi/c^2)(1 + \phi/c^2)](1 + 2\phi/c^2) \\ &= [1 - 2\phi/c^2 - (1/2)(1 + 3\phi/c^2)](1 + 2\phi/c^2) \\ &= [1/2 - 5\phi/c^2][1 + 2\phi/c^2] = 1/2 - 4\phi/c^2 \end{aligned}$$

Hamilton-Jacobi equation for the motion of a particle in the Schwazrchild metric.

$$\alpha(r) = 1 - 2m_0/r, m_0 = GM_0/c^2$$

HJ equation is

$$g^{\mu\nu} S_{,\mu} S_{,\nu} = m^2$$

Thus,

$$\alpha(r)^{-1} S_{,t}^2 - \alpha(r) S_{,r}^2 - r^{-2} S_{,\theta}^2 - (r \sin(\theta))^{-1} S_{,\phi}^2 - m^2 = 0$$

$S = S(t, r, \theta, \phi)$ is the HJ action function. We separate the variables:

$$S(t, r, \theta, \phi) = -Et + M\phi + S_1(r)$$

S is assumed to be independent of θ since the motion is taking place in the $\theta = \pi/2$ plane. Substituting this gives

$$\alpha(r)^{-1} E^2 - \alpha(r) S'_1(r)^2 - r^{-2} M^2 - m^2 = 0$$

or

$$S'_1(r) = (\alpha(r)^{-2} E^2 - \alpha(r)^{-1} (M^2/r^2 + m^2))^{1/2}$$

We have that

$$\partial S / \partial M = constt.$$

so that if ΔS is the change in S when the particle moves from a point P_1 at a radial distance of r_1 to a point P_2 at a radial distance of r_2 , then the corresponding change in the angle is $\Delta\phi$ given by

$$\Delta\phi = -\partial\Delta S_1/\partial M$$

We therefore find that

$$\Delta\phi = (-\partial/\partial M) \int_{r_1}^{r_2} dr. (\alpha(r)^{-2} E^2 - \alpha(r)^{-1} (M^2/r^2 + m^2))^{1/2}$$

In particular, if the particle is a photon, then we get this change in the angle by putting $m = 0$ as

$$\Delta\phi = (-\partial/\partial M) \int_{r_1}^{r_2} dr. (\alpha(r)^{-2} E^2 - \alpha(r)^{-1} (M^2/r^2))^{1/2}$$

Remark on Hamilton-Jacobi theory. Consider the action integral

$$S(t, q(t)) = \int_0^t L(s, q(s), q'(s)) ds$$

taken along the classical trajectory q between times 0 and t . We get by varying the final time and coordinate point:

$$\begin{aligned} & S_{,t}(t, q(t))\delta t + S_{,q}(t, q(t))(\delta q(t) + q'(t)\delta t) \\ & L(t, q(t), q'(t))\delta t + \int_0^t (L_{,q}\delta q(s) + L_{,q'}\delta q'(s)) ds \\ & = L(t, q(t), q'(t))\delta t + L_{,q'}(s, q(s), q'(s))\delta q(s)|_0^t \\ & = L(t, q(t), q'(t))\delta t + L_{,q'}(t, q(t), q'(t))\delta q(t) \end{aligned}$$

where we make use of the Euler-Lagrange equations satisfied by the classical trajectory q . Thus noting that

$$p(t) = L_{,q'}(t, q(t), q'(t))$$

we can write on equating coefficients of δt and $\delta q(t)$,

$$S_{,t}(t, q) + S_{,q}(t, q)q' = L(t, q, q'), S_{,q}(t, q) = p$$

Equivalently,

$$S_{,t}(t, q) = L - pq' = -H(t, q, p), S_{,q}(t, q) = p$$

so that $S(t, q)$ satisfies the HJ equation

$$S_{,t}(t, q) + H(t, S_{,q}(t, q)) = 0$$

There is another way to derive the HJ equation such that the action function S will generate a canonical transformation in such a manner that the transformed positions and momenta are constants of the motion and the transformed Hamiltonian is zero. This then provides us immediately with the solutions to the equations of motion in terms of the action function. If the two systems (Q, P, K) and (q, p, H) are to correspond to the same equations of motion, then the difference between the two corresponding Lagrangians must be a total time derivative, ie,

$$P.Q' - K = p.q' - H + dF/dt$$

Taking $F = F(t, q, Q)$, we then get

$$p.q' - H = P.Q' - K + F_{,t} + F_{,q}.q' + F_{,Q}.Q'$$

and hence

$$K = H + F_{,t}, p = F_{,q}, P = -F_{,Q}$$

We now choose the generating function F so that $K = 0$. Then, F must necessarily satisfy the pde

$$H(q, F_{,q}(t, q, Q)) + F_{,t}(t, q, Q) = 0$$

and then since $K = 0$, the new canonical variables Q, P are constants of the motion. Thus, the equations of motion relative to the original system of canonical coordinates are given by

$$P + F_{,Q}(t, q(t), Q) = 0, p(t) = F_{,q}(t, q(t), Q)$$

Here, Q, P are $2n$ constants determined by the initial conditions. It is easily seen that F is identified with the action function along the classical path described above.

6.42 Lecture plan for "Linear Algebra in Signal Processing"-SP-C01

- [1] Fields, vector spaces, linear independence, basis, dimension of a vector space. (1)
 - [2] Co-ordinate representations of vectors relative to a basis, Matrix transformation of a vector relative to two bases. (1)
 - [3] Linear transformations on a vector space and from one vector space into another. (1)
 - [4] Matrix representation of a linear transformation relative to a basis and similarity transformation of the matrix of a linear transformation relative to two bases. (1)
 - [5] Range, nullspace, rank and nullity of a linear transformation, the rank-nullity theorem. (1)

[6] Subspaces, direct sum of subspaces. (1)

[8] Invariant subspaces of a linear transformation, the matrix of a linear transformation relative to a basis that is an extension of a basis of an invariant subspace. (1)

Digression into infinite dimensional vector spaces

[9] Infinite dimensional vector spaces, examples. (1)

[a] Cauchy sequences, Banach and Hilbert spaces. (1)

[b] Bounded linear operators in Banach and Hilbert spaces. (2)

[c] Compact operators in a Banach space. (2)

[d] Resolvent and spectrum of a linear operator in a Banach space. (1)

[e] Spectra of compact operators in a Banach space. (1)

[f] Graph of an operator in a Banach space, closed and closable operators in a Banach space, Closure of a closable operator. (1)

[g] Densely defined operators in a Hilbert space, the adjoint of a densely defined operator, uniqueness of the adjoint, symmetric operators, self-adjoint operators. (1)

[h] Orthogonal projections in a Hilbert space (existence and uniqueness), Spectral theorem for compact, bounded and unbounded self-adjoint operators in a Hilbert space. (2)

[i] Bounded linear functionals on a Hilbert space and the Riesz representation. (1)

[j] The spectral theorem for bounded and unbounded Hermitian operators in a Hilbert space, functions of a self-adjoint operator in a Hilbert space in terms of spectral integrals. (1)

[k] Applications of the spectral theorem to quantum scattering theory, existence of wave operators, completeness of the wave operators, scattering matrix and explicit formula for the scattering matrix in the momentum domain as a unitary operator in $L^2(S^2)$ in terms of the free particle Hamiltonian and the scattering potential. (2)

6.43 Continuation of finite dimensional vector spaces

[l] Proof of the primary decomposition theorem. (2)

[m] Proof of the Jordan decomposition theorem and algorithms for computation of the Jordan factors. (3)

[n] Eigenvalues, eigenvectors, characteristic and minimal polynomials of a matrix, algebraic and geometric multiplicity of an eigenvalue, criteria for diagonalability of a linear operator spectral theorem for finite dimensional normal matrices. (2)

[o] Functions of a linear operator using Cauchy's integral theorem for functions of a complex variable. (1)

[p] Calculation of the canonical projections on the eigenspaces in terms of Cauchy's integral theorem. (1)

- [q] Perturbation theory for eigenvalues and eigenvectors of a matrix–The Rayleigh-Schrodinger theory. (2)
 - [r] Decomposition theorems of matrix theory.
 - [1] The spectral decomposition (1)
 - [2] The polar decomposition (1)
 - [3] The singular value decomposition (1)
 - [4] The QR decomposition based on Gram-Schmidt orthonormalization process. (1)
 - [5] The root space decomposition of a semisimple Lie algebra. (2)
 - [s] An introduction to group representation theory. (1)
 - [1] Definition of a Lie group and its Lie algebra as the tangent space at the identity. (1)
 - [2] The Lie algebra viewed as the class of left invariant vector fields on the Lie group. (1)
 - [3] The Lie bracket as commutator of vector fields. (1)
 - [4] Vector fields and one forms on a differentiable manifold. (1)
 - [5] Transformation of tangent vectors and vector fields on a manifold under a diffeomorphism. (1)
 - [6] Transformation of one forms under a diffeomorphism. (1)
 - [7] Tensor fields on a manifold and their transformation under a diffeomorphism. (1)
 - [8] General theory of skewsymmetric forms on a manifold. (1)
 - [9] The differential of a skewsymmetric form. (1)
 - [10] The Lie derivative of a vector field, of a one form and then of a general tensor field. (1)
 - [11] Basics of Riemannian geometry: Parallel translation, covariant derivative, connection, torsion and curvature of a connection, Cartan's equations of structure. (2)
 - [12] Application of group representation theory to image processing problems: Pattern recognition and feature extraction, estimation the group transformation element.(2)
 - [13] Linearization of nonlinear systems. (2)
- Total number of lectures (Tentative): 57.

6.44 Cartan's equations of structure

First we compute the torsion of a connection.

$$\begin{aligned}
 \nabla_X(Y) &= \nabla_{X^a e_a}(Y^b e_b) = X^a(e_a(Y^b)e_b + Y^b \nabla_{e_a} e_b) \\
 &= X^a(e_a(Y^b)e_b + Y^b \omega_{ab}^c e_c) \\
 &= X^a(e_a(Y^b) + \omega_{ac}^b Y^c)e_b
 \end{aligned}$$

Thus,

$$\nabla_X(Y) - \nabla_Y X = X^a e_a(Y^b) - Y^a e_a(X^b) + (\omega_{ac}^b - \omega_{ca}^b) X^a Y^c e_b$$

It follows that

$$\begin{aligned} T^b(X, Y) &= e^b(T(X, Y)) = e^b(\nabla_X Y - \nabla_Y X - [X, Y]) = \\ &= X^a e_a(Y^b) - Y^a e_a(X^b) + (\omega_{ac}^b - \omega_{ca}^b) X^a Y^c - e^b([X, Y]) \\ &= X(e^b(Y)) - Y(e_b(X)) + (\omega_{ac}^b - \omega_{ca}^b) X^a Y^c - e^b([X, Y]) \end{aligned}$$

Now,

$$XY = X^a e_a(Y^b e_b) = X^a Y^b e_a e_b + X^a e_a(Y^b) e_b$$

so that

$$[X, Y] = X^a Y^b [e_a, e_b] + (X^a e_a(Y^b) - Y^a e_a(X^b)) e_b$$

Now, if ω is any one form, we have in local coordinates

$$\omega = \omega_a dx^a$$

so that

$$d\omega(X, Y) = \omega_{a,b} dx^b \wedge dx^a (X, Y) = \omega_{a,b} (X^b Y^a - X^a Y^b)$$

On the other hand,

$$X(\omega(Y)) = X^a \partial_a (\omega_b Y^b) = X^a \omega_{b,a} Y^b + \omega_b X^a Y^b_{,a}$$

so that

$$X(\omega(Y)) - Y(\omega(X)) = \omega_{b,a} (X^a Y^b - X^b Y^a) + \omega([X, Y])$$

and hence,

$$d\omega(X, Y) = X(\omega(Y)) - Y(\omega(X)) - \omega([X, Y])$$

Combining all these identities, we obtain finally Cartan's first equation of structure:

$$\begin{aligned} T^b(X, Y) &= de^b(X, Y) + (\omega_{ac}^b - \omega_{ca}^b) X^a Y^c \\ &= de^b(X, Y) + (\omega_{ac}^b - \omega_{ca}^b) e^a(X) e^c(Y) \end{aligned}$$

or equivalently,

$$\begin{aligned} T^b &= de^b + (\omega_{ac}^b - \omega_{ca}^b) e^a \otimes e^c \\ &= de^b + \omega_{ac}^b (e^a \otimes e^c - e^c \otimes e^a) \\ &= de^b + \omega_{ac}^b e^a \wedge e^c \end{aligned}$$

Equivalelntly, defining the one form

$$\omega_a^b = \omega_{ac}^b e^c$$

we can write

$$T^b = de^b + e^a \wedge \omega_a^b$$

This is Cartan's first equation of structure.

6.45 Proof of the Riesz representation theorem

Let \mathcal{H} be Hilbert space and f a bounded linear function on \mathcal{H} . Then there exists a unique vector $z_f \in \mathcal{H}$ such that $f(x) = \langle z_f, x \rangle \forall x \in \mathcal{H}$.

Proof: Let

$$\mathcal{N}(f) = \{x \in \mathcal{H} : f(x) = 0\}$$

Then since f is a nonzero functional, it is clear that $\mathcal{N}(f) \neq \mathcal{H}$. Further, $\mathcal{N}(f)$ is a closed subspace of \mathcal{H} since f is bounded and hence continuous. Hence, $\mathcal{N}(f)^\perp \neq \{0\}$. So we choose any $0 \neq z \in \mathcal{N}(f)^\perp$. Then consider

$$\xi = x - f(x)z/f(z)$$

Note that $f(z) \neq 0$ for otherwise

$$z \in \mathcal{N}(f)^\perp \cap \mathcal{N}(f) = \{0\}$$

which is a contradiction. This means that $\xi \in \mathcal{H}$ is well defined. Further,

$$f(\xi) = f(x)f(z)/f(z) = 0$$

ie

$$\xi \in \mathcal{N}(f)$$

It follows that

$$\langle z, \xi \rangle = 0$$

and hence

$$0 = \langle z, x - f(x)z/f(z) \rangle = \langle z, x \rangle - f(x) \|z\|^2/f(z)$$

from which, we get

$$f(x) = (f(z)/\|z\|^2) \langle z, x \rangle = \langle z_f, x \rangle, x \in \mathcal{H}$$

where

$$z_f = \bar{f}(z)z/\|z\|^2$$

This proves the existence of the representative vector z_f of the linear functional f . Uniqueness follows from $\langle z_f, x \rangle = \langle u, x \rangle, x \in \mathcal{H}$ implies $\langle z_f - u, x \rangle = 0, x \in \mathcal{H}$ implies $z_f - u = 0$.

Wedge product of skew-multilinear functionals

$$f(x_{\sigma 1}, \dots, x_{\sigma m}) = \text{sgn}(\sigma) f(x_1, \dots, x_m), \sigma \in S_m$$

$$g(x_{\rho 1}, \dots, x_{\rho n}) = \text{sgn}(\rho) g(x_1, \dots, x_n), \rho \in S_n$$

Define

$$f \wedge g(x_1, \dots, x_{m+n}) = C(m, n) \sum_{\tau \in S_{m+n}} f(x_{\tau 1}, \dots, x_{\tau m}).g(x_{\tau(m+1)}, \dots, x_{\tau(m+n)})$$

Problem: Choose $C(m, n)$ so that \wedge becomes an associative product.

6.46 Quantum image processing

Let $((X(n, m))_{1 \leq n, m \leq N})$ be a classical image intensity field with $0 \leq X(n, m) < 1$. Define the corresponding quantum image state vector by

$$|\psi_X\rangle = \bigotimes_{n,m=1}^N (\sqrt{1 - X(n, m)^2} \exp(i\phi(n, m))|0\rangle + X(n, m)\exp(i\chi(n, m))|1\rangle)$$

where the tensor product is taken in lexicographic order. For $x_1, x_2, \dots, x_{N^2} = 0, 1$, we can write

$$|\psi_X\rangle = \sum_{x_1, \dots, x_{N^2}=0,1} c_X(x_1, \dots, x_{N^2}) |x_1 x_2 \dots x_{N^2}\rangle$$

and it is clear that

$$\langle \psi_X | \psi_X \rangle = 1$$

It is clear that $c_X(x_1, \dots, x_{N^2})$ is a multilinear combination of the complex numbers

$$(\beta(n, m)\exp(i\phi(n, m)) = \sqrt{1 - X(n, m)^2}\exp(i\phi(n, m)), \alpha(n, m)$$

$$= X(n, m)\exp(i\chi(n, m))), n, m = 1, 2, \dots, N^2$$

such that the first term in these pairs appears r times and the second term appears $N^2 - r$ times where r depends on the binary sequence $x_1 \dots x_{N^2}$. We process the image $|\psi_X\rangle$ using a quantum unitary gate U and try to remove noise from it by making the processed state $U|\psi_X\rangle$ as close as possible in norm to a given state. We note that

$$\sum_{x_1, \dots, x_{i-1}, 1, x_{i+1}, \dots, x_{N^2}} |c_X(x_1, \dots, x_{N^2})|^2 = X(n, m)^2,$$

$$\sum_{x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_{N^2}} |c_X(x_1, \dots, x_{N^2})|^2 = 1 - X(n, m)^2,$$

Another problem in quantum image processing is as follows. Take another classical image $((X_w(n, m)))$, considered as noise and add it to the signal image $((X(n, m)))$ to get the noisy image $Y(n, m) = X(n, m) + X_w(n, m)$. Then, normalize $Y(n, m)$ so that its pixel values all appear in $[0, 1]$. This can be done by defining $b = \max(Y(n, m))$, $a = \min(Y(n, m))$ and defining

$$Z(n, m) = \frac{Y(n, m) - a}{b - a}$$

Then as before represent $Z(n, m)$ by the quantum state $|\psi_Z\rangle$, do quantum processing on this state to get $U|\psi_Z\rangle$ with U chosen in such a way that $\|U|\psi_Z\rangle - |\psi_X\rangle\|$ is a minimum.

6.47 Inclusion of the Goldstone boson field in the gauge field after symmetry breaking

$\psi(x)$ is the original matter wave function. The group $G \subset U(n)$ acts on such wave functions. We write

$$\psi(x) = \gamma(x)\tilde{\psi}(x)$$

where $\gamma(x)$ is a representative element in G/H with H the unbroken subgroup. $\tilde{\psi}(x)$ transforms according to H . We write for $g \in G$,

$$g\psi(x) = g\gamma(x)\tilde{\psi}(x) = \gamma'(g, x)h(g, x)\tilde{\psi}(x)$$

where $\gamma'(g, x)$ is a representative element of $g\gamma(x)H$ and $h(g, x) \in H$. We abbreviate $\gamma'(g, x)$ to $\gamma'(x)$. Likewise, consider the gauge fields $A_\mu(x) = A_\mu^a(x)T_a$ where T_a are the Lie algebra generators of G . The group G acts on $A(x)$ in the adjoint representation. Just as we can remove the Goldstone contribution $\gamma(x)$ from $\psi(x)$ to obtain $\tilde{\psi}(x) = \gamma(x)^{-1}\psi(x)$, so also we can remove the Goldstone contribution from $A_\mu(x)$ to obtain $\tilde{A}_\mu(x)$:

$$\tilde{A}_\mu(x) = \gamma(x)^{-1}A_\mu(x)\gamma(x)$$

Note that that gauge covariant derivative is

$$\partial_\mu + ieA_\mu(x)$$

Which transforms under local G -transformations as

$$g(x)(\partial_\mu + ieA_\mu(x))g(x)^{-1} = \partial_\mu + ieA'_\mu(x)$$

This yields the gauge field transformation law as

$$ieA'_\mu(x) = ieg(x)A_\mu(x)g(x)^{-1} + g(x)(\partial_\mu g(x)^{-1})$$

or equivalently,

$$A'_\mu(x) = g(x).A_\mu(x).g(x)^{-1} - (i/e)g(x)(\partial_\mu g(x)^{-1})$$

Now after removing the Goldstone modes $\gamma(x)$ from the matter field $\psi(x)$, the remaining matter field $\tilde{\psi}(x)$ transformed according to the broken subgroup H . Likewise, after removing the Goldstone modes from the gauge field $A_\mu(x)$, the remaining gauge field $\tilde{A}_\mu(x)$ plus the gauge field derived from the Goldstone modes $\gamma(x)$ of the matter field $\psi(x)$ should transform according to the adjoint representation of the broken subgroup H . We shall see that this is indeed the case. We have with the abbreviations $\gamma'(x) = \gamma(g, x)$, $h'(x) = h(g, x)$,

$$g(x)\gamma(x) = \gamma'(x)h(x)$$

so that

$$(\partial_\mu g(x))\gamma(x) + g(x)\partial_\mu\gamma(x) = (\partial_\mu\gamma'(x))h(x) + \gamma'(x)\partial_\mu h(x)$$

Thus,

$$\begin{aligned} & \gamma(x)^{-1}g(x)^{-1}(\partial_\mu g(x))\gamma(x) + \gamma(x)^{-1}\partial_\mu\gamma(x) \\ &= h(x)^{-1}(\gamma'(x)^{-1}\partial_\mu\gamma'(x))h(x) + h(x)^{-1}\partial_\mu h(x) \end{aligned}$$

Also, the equation

$$A'_\mu(x) = g(x).A_\mu(x).g(x)^{-1} - (i/e)g(x)(\partial_\mu g(x)^{-1})$$

implies

$$\begin{aligned} A'_\mu(x) &= \gamma'(x)h(x)\gamma(x)^{-1}A_\mu(x)\gamma(x)h(x)^{-1}\gamma'(x)^{-1} \\ &\quad - (i/e)(g\partial_\mu g^{-1}) \\ &= \gamma'(x)h(x)\tilde{A}_\mu(x)h(x)^{-1}\gamma'(x)^{-1} - (i/e)(g\partial_\mu g^{-1}) \end{aligned}$$

so that

$$\begin{aligned} \tilde{A}'_\mu(x) &= \gamma'(x)^{-1}A'_\mu(x)\gamma'(x) = h(x)\tilde{A}_\mu h(x)^{-1} - (i/e)\gamma'(x)^{-1}g(x)(\partial_\mu g(x)^{-1})\gamma'(x) \\ &= \gamma'(x)^{-1}A'_\mu(x)\gamma'(x) = h(x)\tilde{A}_\mu h(x)^{-1} + (i/e)\gamma'(x)^{-1}(\partial_\mu g(x))g(x)^{-1}\gamma'(x) \\ &= h(x)\tilde{A}_\mu h(x)^{-1} + (i/e)h\gamma^{-1}(g(x)^{-1}\partial_\mu g(x))\gamma h^{-1} \end{aligned}$$

6.48 Some problems in linear algebra

[1] Derive the fourth order predictor using the Levinson-Durbin algorithm for the process $x[n]$ which satisfies

$$x[n] = s[n] + w[n], s[n] = as[n-1] + v[n]$$

where $w[.]$ is iid $N(0, \sigma_w^2)$, $v[.]$ is iid $N(0, \sigma_v^2)$ and $w[.]$ and $v[.]$ are statistically independent processes.

[2] Calculate using the Gram-Schmidt orthonormalization process a set of five orthonormal vectors in \mathbb{C}^5 that span the same subspace as that spanned by the vectors

$$\begin{aligned} \mathbf{f}_1 &= [1, 0, 0, 0, 0]^T, \mathbf{f}_2 = [1, 1, 0, 0, 0]^T, \\ \mathbf{f}_3 &= [1, 1, 1, 0, 0]^T, \mathbf{f}_4 = [1, 1, 1, 1, 0]^T, \mathbf{f}_5 = [1, 1, 1, 1, 1]^T \end{aligned}$$

[3] Write down explicitly the law of composition and inversion of the Galilean group defined by

$$G = \{(\mathbf{a}, \mathbf{v}, s, \mathbf{g}) : \mathbf{a}, \mathbf{v} \in \mathbb{R}^3, s \in \mathbb{R}, \mathbf{g} \in SO(3)\}$$

Calculate the space-time $4 - D$ Fourier transform of

$$\psi(t, \mathbf{r}) = f(\tau^{-1}(t, \mathbf{r}))$$

where

$$\tau = (\mathbf{a}, \mathbf{v}, s, \mathbf{g})$$

in terms of the $4 - D$ space-time Fourier transform of ψ defined by

$$\hat{\psi}(\omega, \mathbf{k}) = \int f(t, \mathbf{r}) \exp(j(\omega t - \mathbf{k} \cdot \mathbf{r})) dt d^3 r$$

[4] Calculate $\exp(t\mathbf{A})$ where

$$\mathbf{A} = \mathbf{S} \mathbf{J} \mathbf{S}^{-1}$$

with

$$\mathbf{J} = \text{diag}[\mathbf{J}_{\lambda_1}, \tilde{\mathbf{J}}_{\lambda_1}, \mathbf{J}_{\lambda_2}]$$

where

$$\mathbf{J}_{\lambda_1} = \begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix},$$

$$\mathbf{J}_{\lambda_2} = \begin{pmatrix} \lambda_2 & 1 & 0 \\ 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_2 \end{pmatrix}$$

$$\tilde{\mathbf{J}}_{\lambda_1} = \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{pmatrix}$$

Here, \mathbf{S} is an arbitrary 8×8 non-singular matrix and $\lambda_1 \neq \lambda_2$. Also determine the minimal and characteristic polynomials of \mathbf{A} .

[5] Write short notes on any two of the following:

[a] The spectral theorem for normal operators in a finite dimensional inner product space.

[b] The spectral theorem for bounded Hermitian operators in an infinite dimensional Hilbert space.

[c] Functions of a matrix using its Jordan decomposition.

[d] The rank-nullity theorem.

6.49 Gravitational N-body problem in general relativity

$x^0 = ct, x^1, x^2, x^3$ are cartesian coordinates.

$$\begin{aligned} g_{00} &= 1 + 2\phi/c^2 + h_{00}/c^4, g_{rs} \\ &= -(1 - 2\phi/c^2)\delta_{rs} + h_{rs}/c^4, g_{0r} = h_{0r}/c^3 \end{aligned}$$

All computations are upto $O(1/c^4)$.

$$g = -(1 - 4\phi/c_2) + O(1/c_4)$$

$$\begin{aligned} \sqrt{-g} &= 1 - 2\phi/c^2 + O(1/c^4), (-g)^{-1/2} = 1 + 2\phi/c^2 \\ (v_i^\mu) &= (1, 0, 0, 0) \end{aligned}$$

Energy-momentum tensor of the discrete matter distribution

$$\begin{aligned} T^{\mu\nu} &= \sum_i m_i (-g)^{-1/2} \delta^3(x - x_i) v_i^\mu v_i^\nu d\tau_i / dt \\ &= \sum_i m_i (-g)^{-1/2} \delta^3(x - x_i) u_i^\mu u_i^\nu dt / d\tau_i \\ v_i^\mu &= dx_i^\mu / d\tau, u_i^\mu = dx_i^\mu / dt (u_i^0 = 1) \\ d\tau_i / dt &= (1 + 2\phi/c^2 - u_i^2/c^2)^{1/2} = 1 + \phi/c^2 - u_i^2/2c^2 + O(1/c^4) \\ v_{i0} &= g_{00} v_i^0 = g_{00} = 1 + 2\phi/c^2 + O(1/c^4) \end{aligned}$$

Thus,

$$\begin{aligned} T_{00} &= \sum_i m_i \delta^3(x - x_i) (-g)^{-1/2} (1 + 2\phi/c^2)^2 (1 - \phi/c^2 + u_i^2/2c^2) \\ &= \sum_i m_i \delta^3(x - x_i) (1 + 2\phi/c^2) (1 + 4\phi/c^2) (1 - \phi/c^2 + u_i^2/2c^2) + O(1/c^4) \\ &= \sum_i m_i \delta^3(x - x_i) (1 + 5\phi_i/c^2 + u_i^2/2c^2) + O(1/c^4) \end{aligned}$$

where by ϕ_i , we mean $\phi(x_i)$. Further,

$$\begin{aligned} T_{rs} &= \sum_i m_i \delta^3(x - x_i) (1 + 2\phi/c^2) (1 - 2\phi/c^2)^2 u_i^r u_i^s (1 - \phi/c^2 + u_i^2/2c^2) + O(1/c^4) \\ &= \sum_i m_i \delta^3(x - x_i) (1 + 5\phi_i/c^2 u_i^2/2c^2) u_i^r u_i^s + O(1/c^4) \end{aligned}$$

Note that u_i^2 is the same order of magnitude as ϕ . This can be seen by considering the relationship between the orbital velocity of a body around another body of mass M in classical Newtonian gravity. Further,

$$T_{0r} = - \sum_i m_i \delta^3(x - x_i) (-g)^{-1/2} (1 + 2\phi/c^2) (1 - 2\phi/c^2) u_i^r (1 - \phi/c^2 + u_i^2/2c^2) + O(1/c^4)$$

$$\begin{aligned}
&= - \sum_i m_i \delta^3(x - x_i) (1 + 2\phi/c^2) u_i^r (1 - \phi/c^2 + u_i^2/2c^2) + O(1/c^4) \\
&= - \sum_i m_i \delta^3(x - x_i) (1 + \phi/c^2 + u_i^2/2c^2) u_i^r + O(1/c^4) \\
T = g_{\mu\nu} T^{\mu\nu} &= \sum_i m_i \delta^3(x - x_i) (-g)^{-1/2} d\tau_i/dt \\
&= \sum_i m_i \delta^3(x - x_i) (1 + 3\phi/c^2 - u_i^2/2c^2) + O(1/c^4) \\
S_{00} &= T_{00} - T g_{00}/2 \\
&= \sum_i m_i \delta^3(x - x_i) [(1 + 5\phi/c^2) - (1 + 3\phi/c^2 - u_i^2/2c^2)(1 + 2\phi/c^2)/2] + O(1/c^4) \\
&= \sum_i m_i \delta^3(x - x_i) [1/2 + 5\phi/2c^2 + u_i^2/4c^2] + O(1/c^4)
\end{aligned}$$

Also,

$$\begin{aligned}
S_{0r} &= T_{0r} - T g_{0r}/2 = \\
&- \sum_i m_i \delta^3(x - x_i) (1 + \phi/c^2 + u_i^2/2c^2) u_i^r + O(1/c^3)
\end{aligned}$$

Then, upto $O(1/c^4)$, we have

$$\begin{aligned}
R_{00} &= \Gamma_{0\alpha,0}^\alpha - \Gamma_{00,\alpha}^\alpha - \Gamma_{00}^\alpha \Gamma_{\alpha\beta}^\beta + \Gamma_{0\beta}^\alpha \Gamma_{0\alpha}^\beta \\
&= \Gamma_{0k,0}^k - \Gamma_{00,k}^k - \Gamma_{00}^k [\Gamma_{k0}^0 + \Gamma_{km}^m] + 2\Gamma_{00}^k \Gamma_{k0}^0
\end{aligned}$$

Note that $g_{km,0} = c^{-1} g_{km,t}$ is $O(1/c^3)$ while $g_{00,k}$ and $g_{km,r}$ are $O(1/c^2)$. Simplifying further, we get

$$\begin{aligned}
R_{00} &= \\
&\Gamma_{0k,0}^k - \Gamma_{00,k}^k - \Gamma_{00}^k \Gamma_{km}^m + \Gamma_{00}^k \Gamma_{k0}^0
\end{aligned}$$

Now,

$$\begin{aligned}
\Gamma_{0k}^k &= g^{km} \Gamma_{m0k} = (1/2) g^{km} g_{km,0} = (-1/2) g_{kk,0} = (-1/2c) g_{kk,t} \\
&= (-3/c^3) \phi_{,t}
\end{aligned}$$

So,

$$\begin{aligned}
\Gamma_{0k,0}^k &= (-3/c^4) \phi_{,tt} \\
\Gamma_{00}^k &= g^{km} \Gamma_{m00} = (g^{km}/2)(2g_{m0,0} - g_{00,m}) = \\
&- h_{k0,t}/c^4 + \phi_{,k}/c^2 + \phi\phi_{,k}/c^4 + h_{00,k}/2c^4
\end{aligned}$$

Thus,

$$\begin{aligned}
\Gamma_{0k,k}^k &= -h_{k0,kt}/c^4 + (1/c^2) \nabla^2 \phi + \operatorname{div}(\phi \nabla \phi)/c^4 + \nabla^2 h_{00}/2c^4 \\
\Gamma_{00}^k \Gamma_{km}^m &= -(\phi_{,k}/4c^2) g_{mm,k} = (-3\phi_{,k}/2c^2)(\phi_{,k}/c^2)
\end{aligned}$$

$$= -3(\nabla\phi)^2/2c^4$$

$$\Gamma_{00}^k \Gamma_{k0}^0 = (\phi_{,k}/2c^2) g_{00,k} = \phi_{,k} \phi_{,k}/c^4 = (\nabla\phi)^2/c^4$$

Combining all these, we get upto $O(1/c^4)$,

$$R_{00} =$$

$$(-3/c^4)\phi_{,tt} - [-h_{k0,kt}/c^4 + (1/c^2)\nabla^2\phi + \text{div}(\phi\nabla\phi)/c^4 + \nabla^2 h_{00}/2c^4]$$

$$+ 5(\nabla\phi)^2/2c^4$$

$$= -\nabla^2\phi/c^2 + (1/c^4)[-3\phi_{,tt} + h_{k0,kt} + 3(\nabla\phi)^2/2 - \phi\nabla^2\phi - \nabla^2 h_{00}]$$

So our first Einstein field equation

$$R_{00} = (-8\pi G/c^2)S_{00}$$

reads upto $O(1/c^4)$ as

$$-\nabla^2\phi/c^2 + (1/c^4)[-3\phi_{,tt} + h_{k0,kt} + 3(\nabla\phi)^2/2 - \phi\nabla^2\phi - \nabla^2 h_{00}]$$

$$= (-8\pi G/c^2) \sum_i m_i \delta^3(x - x_i) [1/2 + 5\phi_i/2c^2 + u_i^2/4c^2]$$

Equating $O(1/c^2)$ on both sides gives

$$\nabla^2\phi = 4\pi G \sum_i m_i \delta^3(x - x_i)$$

which is Newton's inverse square law of gravitation. Equating $O(1/c^4)$ terms on both sides gives

$$[-3\phi_{,tt} + h_{k0,kt} + 3(\nabla\phi)^2/2 - \phi\nabla^2\phi - \nabla^2 h_{00}]$$

$$- 20\pi G \sum_i \phi(x) m_i \delta^3(x - x_i) - 2\pi G \sum_i m_i \delta^3(x - x_i) u_i^2$$

and since the Newtonian solution ϕ is known, we can immediately solve for the $O(1/c^2)$ perturbation h_{00} to ϕ using the Coulomb inverse square law kernel. We next compute R_{m0} :

$$R_{m0} = \Gamma_{m\alpha,0}^\alpha - \Gamma_{m0,\alpha}^\alpha - \Gamma_{m0}^\alpha \Gamma_{\alpha\beta}^\beta$$

$$+ \Gamma_{m\beta}^\alpha \Gamma_{0\alpha}^\beta$$

Now, again upto $O(1/c^4)$, we have

$$\Gamma_{m\alpha,0}^\alpha = \Gamma_{m0,0}^0 - \Gamma_{mk,0}^k = \Gamma_{0m0,0} - (g^{kr}\Gamma_{rmk}),_0$$

$$= g_{00,m0}/2 + (1/2)((1 + 2\phi/c^2)g_{kk,m}),_0$$

$$= \phi_{,mt}/c^3 - (3/2c^3)\phi_{,mt} = -\phi_{,mt}/2c^3$$

$$\begin{aligned}
& \Gamma_{m0,\alpha}^{\alpha} = \Gamma_{m0,0}^0 + \Gamma_{m0,k}^k \\
& = \Gamma_{0m0,0} - \Gamma_{km0,k} = (1/2)g_{00,m0} - (1/2)(g_{mm,0m} + g_{k0,mk} - g_{m0,kk}) \\
& = \phi_{,mt}/c^3 + \phi_{,mt}/c^3 - h_{k0,mk}/2c^3 + \nabla^2 h_{m0}/2c^3 \\
& \quad \Gamma_{m0}^{\alpha} \Gamma_{\alpha\beta}^{\beta} = \\
& \quad \Gamma_{m0}^0 \Gamma_{0k}^k = g^{00} \Gamma_{0m0} g^{km} \Gamma_{m0k} \\
& = (1/4)g^{00} g_{00,mg} g^{km} g_{km,0} = -\phi_{,m}/c^3 \phi_{,mt} = -\phi_{,m} \phi_{,mt}/c^3 \\
& \quad \Gamma_{m\beta}^{\alpha} \Gamma_{0\alpha}^{\beta} = \\
& \quad \Gamma_{mk}^0 \Gamma_{00}^k + \Gamma_{m0}^k \Gamma_{0k}^0 \\
& = g^{00} \Gamma_{0mk} g^{kk} \Gamma_{k00} + g^{kk} \Gamma_{km0} g^{00} \Gamma_{00k} \\
& = 0
\end{aligned}$$

upto $O(1/c^4)$. Note that $g_{km,0} = g_{km,t}/c$ is $O(1/c^3)$. Thus, finally, upto $O(1/c^4)$, we have

$$\begin{aligned}
R_{m0} &= -\phi_{,mt}/2c^3 - [2\phi_{,mt}/c^3 - h_{k0,mk}/2c^3 + \nabla^2 h_{m0}/2c^3] \\
&\quad + \phi_{,m} \phi_{,mt}/c^3 \\
&= -5\phi_{,mt}/2c^3 + h_{k0,mk}/2c^3 - \nabla^2 h_{m0}/c^3 + \phi_{,m} \phi_{,mt}/c^3
\end{aligned}$$

Since $S_{m0} = (-1/2)Tg_{m0} = (-1/2)\sum_i m_i \delta^3(x - x_i)(-g(x))^{-1/2}h_{m0}(x)/c^3$ in the approximation in which the particles have zero velocities, so that $(8\pi G/c^2)S_{m0} = O(1/c^5)$, the Einstein field equations give

$$R_{m0} = 0$$

ie,

$$-5\phi_{,mt}/2 + h_{k0,mk}/2 - \nabla^2 h_{m0} + \phi_{,m} \phi_{,mt} = 0$$

This equation must be solved for h_{m0} . Finally, we compute

$$\begin{aligned}
R_{km} &= \Gamma_{k\alpha,m}^{\alpha} - \Gamma_{km,\alpha}^{\alpha} - \Gamma_{km}^{\alpha} \Gamma_{\alpha\beta}^{\beta} \\
&\quad + \Gamma_{k\beta}^{\alpha} \Gamma_{m\alpha}^{\beta}
\end{aligned}$$

We now observe that upto $O(1/c^4)$, the Lagrangian of the system of N bodies does not involve h_{rs} ; it involves only ϕ, h_{00}, h_{0r} . Indeed,

$$L = - \sum_i m_i c^2 (g_{\mu\nu}(x_i) u_i^\mu u_i^\nu)$$

Note that the corresponding action integral is $S = \int L dt$. Thus, We get

$$(g_{\mu\nu} u^\mu u^\nu)^{1/2} = (g_{00} + 2g_{0r} u^r + g_{rs} u^r u^s/c^2)^{1/2}$$

$$\begin{aligned}
&= (1 + 2\phi/c^2 + h_{00}/c^4 + 2h_{0r}u^r/c^4 - (1 - 2\phi/c^2)u^2/c^2)^{1/2} + O(1/c^5) \\
&= 1 + \phi/c^2 - u^2/2c^2 + h_{00}/2c^4 + h_{0r}u^r/c^4 - (1/8c^4)(2\phi - u^2)^2 + 2\phi u^2/c^4 + O(1/c^5)
\end{aligned}$$

Thus, upto $O(1/c^4)$, to compute the total Lagrangian of this system of interacting particles upto $O(1/c^2)$ we require only ϕ, h_{00}, h_{0r} and not h_{rs} . This Lagrangian is given by

$$L = \sum_i [m_i u_i^2/2 - m_i \phi(x_i)/2 - m_i h_{00}(x_i)/2c^2 + (m_i/8c^2)(2\phi(x_i) - u_i^2)^2 - m_i \phi(x_i)u_i^2/c^2]$$

Remark: When we pass to a rotating frame of reference, the infinitesimal spatial coordinate change is

$$(d\mathbf{r} - \omega \times \mathbf{r} dt)^2 = (d\mathbf{r})^2 + (\omega \times \mathbf{r})^2 dt^2 - 2(\omega \times \mathbf{r}, d\mathbf{r})dt$$

where ω can be a function of \mathbf{r} . Therefore the change in the metric under this local transformation is

$$c^2 dt^2 (1 - (\omega \times \mathbf{r})^2/c^2) - (d\mathbf{r})^2 + 2(\omega \times \mathbf{r}, d\mathbf{r})dt$$

This justifies the coefficient of $c dt d\mathbf{r}$ being $O(1/c^3)$ provided that $r\omega$ is of order $1/c^2$ when general relativistic effects are taken into account. In general relativity, the Schwarzschild radius of a spherical blackhole is $2GM/c^2$ and hence at this scale the order of magnitude of the term $r\omega$ must be of the order $2GM\omega/c^2$ which justifies our claim.

6.50 Multipole radiation fields in the Maxwell theory

$$X_{lm}(\hat{r}) = LY_{lm}(\hat{r}), L = -ir \times \nabla$$

$$\int Y_{lm}(\hat{r}) \bar{Y}_{l'm'}(\hat{r}) dS(\hat{r}) = \delta_{ll'} \delta_{mm'}$$

where Y_{lm} are the spherical harmonics. We have

$$\hat{r} \cdot X_{lm} = 0$$

We wish that $f(r)X_{lm}(\hat{r})$ satisfies the Helmholtz equation, ie,

$$(\nabla^2 + k^2)(f(r)X_{lm}(\hat{r})) = 0$$

Since the vector operator L commutes with ∇^2 , we get using

$$\nabla^2 = r^{-1} \frac{\partial^2}{\partial r^2} r - L^2/r^2$$

that

$$r^{-1}(rf(r))'' + (k^2 - l(l+1)/r^2)f(r) = 0$$

This equation has two linearly independent solutions, say $j_l(kr), h_l(kr)$. Note that

$$L.X_{lm} = L^2 Y_{lm} = l(l+1)Y_{lm},$$

$$\hat{r}.X_{lm} = 0$$

$$\int (LY_{lm}, LY_{l'm'}) dS = \int \int (Y_{lm}, L^2 Y_{l'm'}) dS = l(l+1) \delta_{ll'} \delta_{mm'}$$

$$\operatorname{curl} \operatorname{curl} X_{lm} = \operatorname{curl} (\operatorname{curl} (LY_{lm})) = \nabla(\nabla \cdot LY_{lm}) - \nabla^2 LY_{lm}$$

$$i\nabla \times LY_{lm} = \nabla \times (r \times \nabla Y_{lm}) =$$

$$(\nabla^2 Y_{lm})r - 2\nabla Y_{lm} - r(d/dr)\nabla Y_{lm}$$

$$= -l(l+1)Y_{lm}\hat{r}/r - 2\nabla Y_{lm}$$

So,

$$i\hat{r} \cdot \nabla \times LY_{lm} = -l(l+1)Y_{lm}/r - 2dY_{lm}/dr$$

$$= -l(l+1)Y_{lm}/r$$

$$i\operatorname{div} LY_{lm} = \nabla \cdot r \times \nabla Y_{lm} = \nabla \times r \cdot \nabla Y_{lm} - r \cdot \nabla \times \nabla Y_{lm} = 0$$

$$\operatorname{curl}(f(r)LY_{lm}) = f'(r)\hat{r} \times LY_{lm} + f(r)\operatorname{curl} LY_{lm}$$

$$i\hat{r} \times L\psi = \hat{r} \times (r \times \nabla \psi) = \mathbf{r}\psi_{,r} - r\nabla\psi$$

So

$$i\hat{r} \times LY_{lm} = -r\nabla Y_{lm}$$

Combining these equations, we get

$$\operatorname{curl}(f(r)LY_{lm}) = i f'(r)r \nabla Y_{lm} + f(r)(i(l(l+1)Y_{lm}\hat{r}/r + 2i\nabla Y_{lm})$$

$$= (if'(r)r + 2if(r))\nabla Y_{lm}(\hat{r}) + il(l+1)(\hat{r}/r)f(r)Y_{lm}(\hat{r})$$

This shows that

$$\hat{r} \cdot \operatorname{curl}(f(r)LY_{lm}(\hat{r})) = il(l+1)f(r)Y_{lm}(\hat{r})/r$$

Then, if S_r denotes the surface of a sphere of radius r and $dS_r(\hat{r})$ is differential surface area, we have

$$\begin{aligned} & \int_{S_r} ((LY_{l'm'})^*(\hat{r}), \operatorname{curl}(f(r)LY_{lm}(\hat{r}))) dS_r(\hat{r}) \\ &= \int_{S_r} (\bar{Y}_{l'm'})(L \cdot \nabla Y_{lm})(if'(r)r + 2if(r)) dS_r(\hat{r}) = 0 \end{aligned}$$

since $L \cdot \nabla = 0$. On the other hand, we have

$$\int_{S_r} ((f(r)LY_{lm}(\hat{r}))^*, LY_{l'm'}(\hat{r})) dS_r(\hat{r})$$

$$= \int_{S_r} \bar{f}(r) l(l+1) \bar{Y}_{lm}(\hat{r}) Y_{l'm'}(\hat{r}) dS_r(\hat{r}) = l(l+1) \bar{f}(r) \delta_{l,l'} \delta_{m,m'}$$

Problems in differential geometry

[1] Show that when acting on k -forms,

$$di_X + i_X d = L_X$$

where X is any vector field and i_X is the contraction by X while L_X is the Lie derivative by X .

Reference: Ralph Abraham and Jerrold E.Marsden, "Foundations of Mechanics", Chelsea.

6.51 How Dirac brackets are used to take care of constraints in Lagrangian and Hamiltonian mechanics

Exercise: When there are constraints on the canonical position and momenta of a Hamiltonian system, the canonical commutation relations are not compatible with it. Thus, we must replace the Poisson brackets in classical mechanics and the Lie bracket in quantum mechanics by a different kind of bracket— The Dirac bracket between observables. These constraints ensure that the unconstrained positions and momenta satisfy the correct equations of motion when the Dirac bracket is used and the constraints are compatible with the Dirac bracket. The basic reference for Dirac brackets is:

Reference:

[1] Steven Weinberg, "The quantum theory of fields", vol.1, Cambridge University Press.

6.52 Deep learning of speech models

Let $\{x_k(n)\}$ be the speech sequence of speaker A for a given text T_k and let $\{y_k(n)\}$ be the speech sequence of speaker B for the same text. We wish to learn this transformation, ie, design an neural network based on this data that will transform a speech sequence of A $\{x(n)\}$ corresponding to any other text T to the corresponding speech sequence of B corresponding to the same text T . For this we first observe that the speech sequences have large dynamic range and hence each sample would require $\log_2(N)$ bits for quantization where N is 16, 32 etc. while on the other hand, if we use a linear predictive coder (LPC) for the speech data, then the prediction errors will be of small dynamic range requiring

say just one bit for quantization. This is because of the highly correlated nature of the speech signals. Thus, we may represent $x_k(n)$ as

$$x_k(n) = - \sum_{m=1}^p a_k(m)x_k(n-m) + w_k(n)$$

where $w_k(n)$ can be represented by a single bit and the LPC coefficient vector a_k would require $p.\log_2(N)$ bits for effective quantization. This coefficient vector should be selected so that

$$\sum_n (x_k(n) + \sum_{m=1}^p a_k(m)x_k(n-m))^2$$

is a minimum. If the speech data length is $K \approx 100$ samples while the LPC filter length is $p \approx 5$, then without the use of LPC, we would require $K.\log_2(N)$ bits to store the data while after LPC, we would require just $p.\log_2(N) + K$ bits (K bits for the set of N noise samples each sample requiring just a single bit). We have for $N = 16$,

$$K.\log_2(N) = 100 \times 4 = 400$$

bits for data storage without LPC and with LPC

$$p.\log_2(N) + K = 5 \times 4 + 100 = 120$$

bits for storage. For highly correlated speech samples, use of LPC thus drastically reduces the required amount of data storage for effective reconstruction.

6.53 Some problems on linearization (for the course Linear algebra in signal processing)

[1] Consider the Boltzmann equation

$$\begin{aligned} f_{,t}(t, r, v) + (v, \nabla_r) f(t, r, v) + (q/m)(v \times B(t, r), \nabla_v) f(t, r, v) \\ = \frac{f_0(v) - f(t, r, v)}{\tau_v} \quad \text{---(1)} \end{aligned}$$

where

$$f_0(v) = C \exp(-mv^2/2kT)$$

is the Maxwell distribution that solves the equilibrium Boltzmann equation

$$(v, \nabla_r) f_0 + (q/m)(v \times B, \nabla_v) f_0 = 0$$

We assume that

$$B(t, r) = B_0 + \delta B(t, r), \delta B(t, r) = \operatorname{Re}(\delta B_0 \exp(i(\omega t - k.r)))$$

The first order perturbed Boltzmann equation is

$$\begin{aligned} \delta f_{,t}(t, r, v) + (v, \nabla_r) \delta f(t, r, v) - (q/m)(B_0 \times v, \nabla_v) \delta f(t, r, v) \\ - (\delta B(t, r) \times v, \nabla_v) f_0(v) + \delta f(t, r, v)/\tau_v = 0 \quad \dots \quad (2) \end{aligned}$$

Note that we have neglected the second order term

$$(\delta B(t, r) \times v, \nabla_v) \delta f(t, r, v)$$

Write

$$\delta f(t, r, v) = Re(\delta f_0(v).exp(i(\omega t - k.r)))$$

so that the linearized equation (2) reads

$$\begin{aligned} i\omega \delta f_0(v) - i(v, k) \delta f_0(v) - (q/m)(B_0 \times v, \nabla_v) \delta f_0(v) \\ + \delta f(t, r, v)/\tau_v = 0 \quad \dots \quad (3) \end{aligned}$$

where we have used

$$(\delta B(t, r) \times v, \nabla_v) f_0(v) = 0$$

The Maxwell equation

$$\text{curl} \delta B(t, r) = \mu \int v \delta f(t, r, v) d^3 v$$

becomes

$$-ik \times \delta B_0 = \mu \int v \delta f_0(v) d^3 v$$

[2] Consider a system of N electrons moving in the electrostatic field generated by themselves and a nucleus of charge Ze . As a first order approximation to the joint wave function of the electrons (neglecting spin), we assume that the wave function is separable so that if $\psi_k(t, r_k)$ is the wave function of the k^{th} electron then the electric potential seen by it is given by

$$V_k(t, r_k) = \frac{-Ze^2}{|r_k|} + e^2 \sum_{j \neq k} \int \frac{|\psi_j(t, r_j)|^2}{|r_k - r_j|} d^3 r_j$$

Now write down the approximate Schrodinger equation for the k^{th} electrons as

$$i\psi_{k,t}(t, r_k) = (-1/2m)\nabla_k^2 \psi_k(t, r_k) + V_k(t, r_k)\psi_k(t, r_k), k = 1, 2, \dots, N \quad \dots \quad (1)$$

This is a system of N nonlinearly coupled pde's. Obtain the linearized version of these equations by expanding

$$\psi_k(t, r_k) = \psi_{k0}(t, r_k) + \delta\psi_k(t, r_k)$$

where the ψ_{k0} satisfy the decoupled linear Schrodinger equations

$$i\psi_{k0,t}(t, r_k) = (-1/2m)\nabla_k^2 \psi_k(t, r_k) - Ze^2/|r_k|, k = 1, 2, \dots, N$$

and the $\delta\psi_k$ satisfy coupled linearized equations

$$\begin{aligned} i\delta\psi_{k,t}(t, r_k) = & (-Ze^2/|r_k|)\delta\psi_k(t, r_k) + \\ e^2 \sum_{j \neq k} \int & (|r_k - r_j|^{-1} \cdot (|\psi_{j0}(t, r_j)|^2 \delta\psi_k(t, r_k) + \psi_{k0}(t, r_k)(\psi_{j0}(t, r_j)^* \delta\psi_j(t, r_j) + \\ & \psi_{j0}(t, r_k)^* \delta\psi_j(t, r_j))) d^3 r_j \end{aligned}$$

Indicate a numerical scheme by which you can obtain solutions to these linearized equations. Also indicate a second order perturbation theoretic analysis of the nonlinear Schrodinger equation (1).

[3] Repeat problem [2] by taking the spin of the electron and the Pauli exclusion principle into account. Specifically, let σ_k denote the spin index of the k^{th} electron and write $\xi_k = (r_k, \sigma_k)$, $k = 1, 2, \dots, N$. Here, $\sigma_k = 1, 2$ according as the k^{th} electron has spin $+1/2$ or spin $-1/2$ along the z direction. The wave function of the entire system of N electrons must be antisymmetric w.r.t. the variables ξ_1, \dots, ξ_N . So as a first order approximation, we assume it to be of "separable form"

$$\psi(t, \xi_1, \dots, \xi_N) = (N!)^{-1} \sum_{\tau \in S_N} sgn(\tau) \bigotimes_{k=1}^N \psi_{\tau k}(t, \xi_k)$$

Substitute this expression into the Schrodinger equation

$$\begin{aligned} i\psi_{,t}(t, \xi_1, \dots, \xi_N) = & [(-1/2m) \sum_{k=1}^N \nabla_k^2 - \sum_k Ze^2/|r_k| \\ & + e^2 \sum_{k < j} |r_k - r_j|^{-1}] \psi(t, \xi_1, \dots, \xi_N) \end{aligned}$$

From this multiparticle linear equation, we derive single particle coupled nonlinear pde's by forming the inner product with $\psi_k(t, \xi_k)$. We leave this as an exercise to the reader.

6.54 Research project proposal for simulating quantum gates of large sizes using quantized ridge waveguide electromagnetic field interacting with quantum dots and also for estimating medium properties on the Angstrom scale

The waveguide em fields can be expressed in terms of H_z, E_z . We can thus expand

$$\mathbf{E}(x, y, z, t) = \sum_n Re(c[n].exp(j\omega t - \gamma_n(\omega)z)) \psi_n(x, y)$$

$$\mathbf{B}(x, y, z, t) = \sum_n Re(d[n].exp(j\omega t - \beta_n(\omega)z)\chi_n(x, y))$$

where $\psi_n(x, y), \chi_n(x, y) \in \mathbb{R}^3$ are orthogonal functions on the cross section of the guide. If in addition, we put lids on the top and bottom of the guide so that it becomes a resonator, then the frequency takes discrete values and hence we can write the fields as

$$\mathbf{E}(x, y, z, t) = \sum_n Re(c[n].exp(j\omega_1[n]t))\psi_n(x, y, z),$$

$$\mathbf{B}(x, y, z, t) = \sum_n Re(d[n].exp(j\omega_2[n]t).\chi_n(x, y, z))$$

where now $\psi_n, \chi_n, n = 1, 2, \dots$ are mutually orthogonal functions on the resonator volume with values in \mathbb{R}^3 . When we pass over to the quantum theory, then $c[n], d[n]$ become creation and annihilation operators of quantum harmonic oscillators. To see this, we calculate the energy of the electromagnetic field within the resonator V as

$$\begin{aligned} U &= (1/2) \int_V (|E|^2/2 + |B|^2/2) dx dy dz = \\ &(1/4) \sum_n (|c[n]|^2 + |d[n]|^2) \end{aligned}$$

We now look at the interaction of this quantum em field with an atom placed within the resonator. For this, we need to determine the electromagnetic four potential in terms of the electric and magnetic fields. We adopt the Coulomb gauge, so that $\text{div}A = 0$ and $\nabla^2\Phi = 0$ since there is no charge density within the guide. Thus, we can take $\Phi = 0$ and

$$E = -A_{,t}, B = \text{curl}A$$

This gives

$$A(x, y, z, t) = - \sum_n Im(c[n].exp(j\omega_1[n]t)/\omega_1[n])\psi_n(x, y, z),$$

In case that there is charged matter within the resonator, we would express the corresponding four current density as a second quantized Dirac current field. Now consider the case when there is a small inhomogeneous perturbation in the permittivity:

$$\epsilon = \epsilon_0(1 + \delta.\chi(\omega, r))$$

in the frequency domain. The Maxwell equations $\text{div}B = 0, \text{curl}E = -E_{,t}$ do not get changed and so they imply that $B = \text{curl}A, E = -\nabla\Phi - j\omega A$. We assume the Coulomb gauge $\text{div}A = 0$. Then the other Maxwell equations are

$$\text{div}(\epsilon E) = \rho, \text{curl}B = \mu_0\epsilon J + j\omega\epsilon E$$

and we get upto $O(\delta)$, the equation

$$\operatorname{div} E + \delta(\nabla\chi, E) = \rho(1 - \delta\chi)$$

which implies in view of the Coulomb gauge that

$$\nabla^2\Phi + \delta(\nabla\chi, \nabla\Phi + j\omega A) = -(1 - \delta\chi)\rho/\epsilon_0$$

or equivalently,

$$\nabla^2\Phi + \rho/\epsilon_0 = \delta[\chi\rho/\epsilon_0 - (\nabla\chi, \nabla\Phi) - j\omega(\nabla\chi, A)]$$

From the frequency dependence of χ , it is clear that now Φ is now no longer a matter field. We write

$$\Phi = \Phi_0 + \delta.\Phi_1 + O(\delta^2), A = A_0 + \delta.A_1 + O(\delta^2)$$

so that Φ_0 becomes a matter field satisfying Poisson's equation

$$\nabla^2\Phi_0 = -\rho/\epsilon_0$$

while Φ_1 is no longer matter field and it satisfies

$$\nabla^2\Phi_1 = \chi\rho/\epsilon_0 - (\nabla\chi, \nabla\Phi_0) - j\omega(\nabla\chi, A_0)$$

Likewise, the other Maxwell equation

$$\operatorname{curl} B = \mu_0 J + j\omega\mu_0\epsilon_0(1 + \delta\chi)E$$

gives on substituting for B, E in terms of the potentials and applying the Coulomb gauge,

$$\nabla^2 A + \omega^2\mu_0\epsilon_0 A - j\omega\mu_0\epsilon_0\nabla\Phi + \delta.j\omega\mu_0\epsilon_0\chi E + \mu_0 J = 0$$

In the absence of external charges and currents, we have $\rho = 0, J = 0, \Phi_0 = 0$ and these field equations simplify to

$$\nabla^2\Phi_1 + j\omega(\nabla\chi, A_0) = 0,$$

$$\nabla^2 A_0 + \omega^2\mu_0\epsilon_0 A_0 = 0,$$

$$\begin{aligned} \nabla^2 A_1 + \omega^2\mu_0\epsilon_0 A_1 - j\omega\mu_0\epsilon_0\nabla\Phi_1 \\ + \omega^2\mu_0\epsilon_0\chi A_0 = 0 \end{aligned}$$

The unperturbed magnetic vector potential A_0 satisfies the standard wave equation and hence can be expanded as a superposition of creation and annihilation operators in momentum space, while we can easily solve the above equations for the perturbations Φ_1, A_1 in terms of A_0 and hence also express these as linear combinations of the creation and annihilation operators of the free unperturbed field. It remains to calculate the Hamiltonian of the perturbed em field upto

$O(\delta)$ expressing the perturbation as a linear combination of the creation and annihilation operators. The Hamiltonian of the perturbed em field is upto $O(\delta)$ given by

$$H = \int [(\epsilon_0/2)(\delta\nabla\Phi_1 + A_{0,t} + \delta A_{1,t})^2 + (1/2\mu_0)(\nabla \times (A_0 + \delta A_1))^2] d^3r$$

As explained above it can be approximated as

$$H = H_0 + \delta.H_1(t)$$

where H_0 is the Hamiltonian of the unperturbed em field, ie, in the absence of inhomogeneities in the medium and $H_1(t)$ is a perturbation to this Hamiltonian caused by medium inhomogeneities. We can abstract the essential content of this theory and write the fields as

$$\begin{aligned} A_0(t, r) &= \sum_k (\mathbf{f}_k(t, r)a_k + \bar{\mathbf{f}}_k(t, r)a_k^*), \Phi_0(t, r) = 0, \\ A_1(t, r) &= \sum_k (\mathbf{g}_k(t, r)a_k + \bar{\mathbf{g}}_k(t, r)a_k^*), \\ \Phi_1(t, r) &= \sum_k (h_k(t, r)a_k + \bar{h}_k(t, r)a_k^*) \end{aligned}$$

where the function \mathbf{f}_k is a plane wave with some polarization independent of the medium inhomogeneities while the functions \mathbf{g}_k, h_k are medium dependent functions. When this quantum electromagnetic field interacts with an atom, the interaction Hamiltonian is derived from the basic atomic Hamiltonian

$$H_A(t) = (\mathbf{P} + e\mathbf{A}(t, r))^2/2m - Ze^2/|\mathbf{r} - \mathbf{r}_0| - e\Phi_1(t, r)$$

with

$$\mathbf{P} = -ih\nabla_r$$

and the interaction part of this Hamiltonian is given by

$$H_I(t) = (-ieh/m)(\mathbf{A}(t, \mathbf{r}), \nabla_r) - e\Phi_1(t, \mathbf{r})$$

with

$$\mathbf{A}(t, \mathbf{r}) = \mathbf{A}_0(t, \mathbf{r}) + \mathbf{A}_1(t, \mathbf{r})$$

and $\Phi_1(t, \mathbf{r})$ being linear combinations of the field creation and annihilation operators a_k^*, a_k and hence by studying atomic transitions using this interaction Hamiltonian assuming that the quantum electromagnetic field is in a coherent state or in a superposition of coherent states, we can hope to estimate properties of the medium, namely the susceptibility functions on which $\mathbf{g}_k(t, \mathbf{r})$ and $h_k(t, \mathbf{r})$ depend. We can also modify this theory for Dirac particles in which case the interaction Hamiltonian will be

$$e(\alpha, \mathbf{A}(t, \mathbf{r})) - e\Phi_1(t, \mathbf{r})$$

while the unperturbed atomic Hamiltonian is

$$H_{A0} = (\alpha, \mathbf{P}) + \beta m - Ze^2/|\mathbf{r} - \mathbf{r}_0|$$

where $\alpha = (\alpha_1, \alpha_2, \alpha_3)$, β are the four anticommuting Dirac matrices. We can also manipulate the medium susceptibility so as to produce a desired evolution for the atom or atom and field and thereby realize large sized quantum unitary gates.

Reference: Lalit Kumar, Ph.D thesis, NIT, 2018.

6.55 Design of a differentiator using series connection of short circuited transmission line elements

We take N lossless lines with characteristic impedances R_1, \dots, R_N , propagation constants $j\beta_k = j\omega/u_k$, $u_k = \sqrt{L_k C_k}$, $k = 1, 2, \dots, N$ and lengths d_1, \dots, d_N . If $I(\omega)$ is the input current to each of these lines, then the corresponding line voltages at the input will be

$$V_k(\omega) = I(\omega) \cdot jR_k \cdot \tan(\beta_k d_k) = jR_k \tan(\omega d_k / u_k), k = 1, 2, \dots, N$$

If we connect these lines in series with some of the polarities reversed, then the net output voltage of these lines will be

$$V(\omega) = I(\omega) \sum_{k=1}^N jS_k \cdot \tan(\lambda_k \omega)$$

where

$$\lambda_k = d_k / u_k, S_k = \pm R_k$$

Alternately if we apply the same voltage $V_1(\omega)$ to all these lines connected in parallel, then the net input current will be

$$I_1(\omega) = V(\omega) \sum_{k=1}^N -jR_k^{-1} \cot(\lambda_k \omega)$$

We choose the former arrangement and note that since for distinct $\lambda_1, \dots, \lambda_N$, the functions $\omega \rightarrow \tan(\lambda_k \omega)$ are linearly independent, it is possible by choosing N large enough to approximate any function $f(\omega)$ with arbitrarily small error over a given bounded frequency range. To construct a differentiator we therefore try to approximate its impedance $j\omega$ by

$$V(\omega)/I(\omega) = \sum_{k=1}^N jS_k \cdot \tan(\lambda_k \omega)$$

by choosing the weights $\{S_k\}$ appropriately. For this we approximate $\tan(\lambda_k \omega)$ by a polynomial of degree N using a truncated Taylor expansion:

$$\tan(x) = \sum_{k=1}^N c(k)x^k, c(1) = 1$$

and then our approximation problem reduces to equating coefficients of $\omega^m, m = 1, 2, \dots, N$ on both sides of

$$\begin{aligned}\omega &\approx \sum_{k=1}^N S_k \tan(\lambda_k \omega) \\ &\approx \sum_{k=1}^N S_k \sum_{m=1}^N c(m)(\lambda_k \omega)^m\end{aligned}$$

We then get

$$\sum_{k=1}^N S_k \lambda_k^m = \delta[m - 1], m = 1, 2, \dots, N$$

This can be expressed in Van-Der-Monde matrix notation as

$$\begin{pmatrix} \lambda_1 & \lambda_2 & \lambda_3 & \dots & \lambda_N \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 & \dots & \lambda_N^2 \\ \dots & \dots & \dots & \dots & \dots \\ \lambda_1^N & \lambda_2^N & \lambda_3^N & \dots & \lambda_N^N \end{pmatrix} \begin{pmatrix} S_1 \\ S_2 \\ S_3 \\ \vdots \\ S_N \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

It is well known that the above $N \times N$ Van-Der-Monde matrix is invertible and hence we obtain a solution. Another way to approach this problem is via a weighted least square method. Let $W(\omega)$ be a positive function. Then, we choose $\{S_k\}_{k=1}^N$ so that

$$E(\mathbf{S}) = \int_{-\sigma}^{\sigma} W(\omega) (\omega - \sum_{k=1}^N S_k \tan(\lambda_k \omega))^2 d\omega$$

is a minimum. Setting the partial derivatives of E w.r.t. S_k to zero gives us the optimal normal equations:

$$\mathbf{AS} = \mathbf{b}, \mathbf{S} = \mathbf{A}^{-1}\mathbf{b}$$

where

$$\mathbf{A} = ((a(k, m))), a(k, m) = \int_{-\sigma}^{\sigma} W(\omega) \tan(\lambda_k \omega) \cdot \tan(\lambda_m \omega) d\omega,$$

$$\mathbf{b} = ((b(k))), b(k) = \int_{-\sigma}^{\sigma} W(\omega) \omega \cdot \tan(\lambda_k \omega) d\omega$$

6.56 Design of quantum gates by interaction of a quantum em field with gravity

The gravitational field is a controllable background classical field. Next, we assume that the free gravitational field is also quantized and we design a gate built out of the creation and annihilation operator fields of the em and gravitational fields.

The free em field can be expressed as

$$F_{\mu\nu}(x) = \sum_k f((\mu\nu k, x) a_k + \bar{f}(\mu\nu k, x) a_k^*)$$

The free gravitational field assuming it to be weak is

$$\delta g_{\mu\nu}(x) = \sum_k (h(\mu\nu k, x) c_k + \bar{h}(\mu\nu k, x) c_k^*)$$

where a_k, a_k^* are photon annihilation and creation operators while c_k, c_k^* are graviton annihilation and creation operators. The Lagrangian density of the em field plus its interaction Lagrangian density with the gravitational field is given by

$$L = (-1/4) F_{\mu\nu} F^{\mu\nu} \sqrt{-g} = (-1/4) g^{\mu\alpha} g^{\nu\beta} \sqrt{-g} F_{\mu\nu} F_{\alpha\beta}$$

We regard A_μ as the pure electromagnetic covariant four potential without any gravitational field component. Then $A^\mu = g^{\mu\alpha} A_\alpha$ is the contravariant electromagnetic four potential with gravitational components. We could also start our analysis the other way round, namely, by regarding A^μ as the pure contravariant em four potential and $A_\mu = g_{\mu\alpha} A^\alpha$ as the covariant em four potential containing gravitational components. However we choose the former formalism for then

$$F_{\mu\nu} = A_{\nu,\mu} - A_{\mu,\nu}$$

is the pure em field tensor not containing any gravitational components. Then upto first order gravitational perturbations, we have

$$g_{\mu\nu} = \eta_{\mu\nu} + \delta g_{\mu\nu},$$

$$g^{\mu\nu} = \eta_{\mu\nu} - \delta g^{\mu\nu}$$

$$\sqrt{-g} = 1 + \delta g/2$$

where

$$\delta g^{\mu\nu} = \eta_{\mu\alpha}\eta_{\nu\beta}\delta g_{\alpha\beta},$$

$$\delta g = \eta_{\mu\nu}\delta g_{\mu\nu}$$

6.57 Scattering theory in the interaction picture for time dependent interactions

$$H(t) = H_0 + V(t)$$

$$\partial_t U(t, s) = -iH(t)U(t, s), U(s, s) = I$$

$$U_0(t) = \exp(-itH_0)$$

$$\Omega_+(t_2, t_1) = U(t_2, t_1)^{-1}U_0(t_2 - t_1) = U(t_2, t_1)^*U_0(t_2 - t_1), t_2 > t_1,$$

$$\Omega_-(t'_2, t'_1) = U(t'_1, t'_2)^{-1}U_0(t'_1 - t'_2) = U(t'_1, t'_2)^*U_0(t'_1 - t'_2), t'_2 > t'_1$$

Then,

$$\begin{aligned} S(t_2, t'_2 | t_1, t'_1) &= \Omega_+(t_2, t_1)^* \Omega_-(t'_2, t'_1) = \\ &U_0(t_1 - t_2)U(t_2, t_1).U(t'_1, t'_2)^*U_0(t'_1 - t'_2) \end{aligned}$$

In the interaction picture, we define $W(t_2, t_1)$ by

$$U(t_2, t_1) = U_0(t_2)W(t_2, t_1)U_0(-t_1)$$

Then we have

$$\begin{aligned} i\partial U(t_2, t_1) &= H_0U_0(t_2)W(t_2, t_1)U_0(-t_1) + iU_0(t_2)\partial_{t_2}W(t_2, t_1)U_0(-t_1) \\ &= H(t_2)U(t_2, t_1) = (H_0 + V(t_2))U_0(t_2)W(t_2, t_1)U_0(-t_1) \end{aligned}$$

which simplifies to

$$i\partial_{t_2}W(t_2, t_1) = \tilde{V}(t_2)W(t_2, t_1), W(t_1, t_1) = I$$

Likewise,

$$U(t'_1, t'_2) = U_0(t'_1)W(t'_1, t'_2)U_0(-t'_2)$$

and hence

$$S(t_2, t'_2 | t_1, t'_1) = U_0(t_1)W(t_2, t_1)U_0(t'_2 - t_1)W(t'_1, t'_2)^*U_0(-t'_2)$$

where for $t > s$,

$$W(t, s) = I + \sum_{n \geq 1} (-i)^n \int_{s < t_n < \dots < t_1 < t} \tilde{V}(t_1) \dots \tilde{V}(t_n) dt_1 \dots dt_n$$

with

$$\tilde{V}(t) = U_0(-t)V(t)U_0(t)$$

This gives in particular on taking $t'_2 = t_1, t'_1 = -t_2$

$$\begin{aligned} S(t_2| -t_2) &= S(t_2, t_1| t_1, -t_2) = U_0(t_1)W(t_2, t_1)W(-t_2, t_1)^*U_0(-t_1) \\ &= U_0(t_1)W(t_2, -t_2)U_0(-t_1) \end{aligned}$$

In particular, if $t_1 = 0$, then

$$S(t_2| -t_2) = I + \sum_{n \geq 1} (-i)^n \int_{-t_2 < s_n < \dots < s_1 < t_2} \tilde{V}(s_1) \dots \tilde{V}(s_n) ds_1 \dots ds_n$$

In the limit $t_2 \rightarrow \infty$, we get the standard form of the scattering matrix as a Dyson series used in quantum field theory for deriving the Feynman diagrams:

$$\begin{aligned} S(\infty) &= S = I + \sum_{n \geq 1} (-i)^n \int_{-\infty < s_n < \dots < s_1 < \infty} \tilde{V}(s_1) \dots \tilde{V}(s_n) ds_1 \dots ds_n \\ &= T\{\exp(-i \int_{-\infty}^{\infty} \tilde{V}(t) dt)\} \end{aligned}$$

Note: $W(t'_1, t'_2)^* = W(t'_2, t'_1), t'_2 > t'_1$.

6.58 Chapterwise report on Jaspal Khinda's Ph.D thesis

Chapter 1 provides an introduction to existing microstrip patch antennas (MSPA) as a solution to the need for drastically reducing the size of antennas to be useful in the current wireless revolution. Various patch shapes are discussed here for operation at microwave frequencies. The kind of substrates used, feed techniques for MSPA's along with the existing mathematical models in the literature for analyzing the MSPA's design and performance are mentioned. Advantages of MSPA's are discussed like good bandwidth, low weight, small size, generation of all kinds of polarization and low cost of fabrication. Some mathematical detail here may be provided regarding the inverse relationship between the linear size of an antenna and its frequency of operation based on the fact that the far field antenna pattern is the spatial Fourier transform of its current distribution.

Chapter 2 presents a thorough literature survey on modifications of MSPA's to meet the needs of larger bandwidth for 5G applications. This survey includes techniques like changing the nature of the feeds, use of fractal patch shapes having self-similar structures. Some stress is put on the fact that normal MSPA's are small in size and hence operate at narrow bandwidth around the high frequency region while fractal MSPA's have all dimensions built in owing to their self similar structure (ie, at each length scale, the fractal has the same shape), and hence cover a broader bandwidth around the high frequency region. I think that some mathematical explanation that connects bandwidth to size and shape

of the MSPA's and MSFA's (Microstrip fractal antenna) can be provided in this chapter. Some examples of fractal shapes can be included in this chapter especially those based on Mandelbrot's original work using non-linear difference equations in the complex plane.

Chapter 3 presents the design of a novel triangular shaped MSFA for the enhancement of gain-bandwidth product. Again, if all dimensions are present within a fractal structure (larger dimensions are generated by causing the fractal to be more kinky so that within a small region, the antenna boundary has a very large length), then a theoretical explanation of how the gain-bandwidth product will be enhanced may be provided here using standard retarded potential theory:

$$\mathbf{A}(\omega, \mathbf{r}) = \frac{\mu}{4\pi r} \exp(-j\omega r/c) \int_{\mathbf{r}' \in \text{antennabody}} \mathbf{J}(\omega, \mathbf{r}') \exp(j\omega \hat{r} \cdot \mathbf{r}'/c) d^3 \mathbf{r}'$$

Chapter 4 starts by noting that in Chapter 3, C-band and X-band applications had been covered using MSFA's. However, this design did not cover the entire C-band, had more mismatch losses, lower radiation efficiency, small gain-bandwidth product and low peak gain. Hence the author proposes in the current chapter, a new kind of MSFA which improves upon these performance parameters. The new design is based on placing a rectangular radiation patch on the substrate and a partial ground plane on its reverse side. Complete C-band coverage is demonstrated via experiments and improvement in the other parameters mentioned above is demonstrated. However the peak gain and the gain-bandwidth product are not improved that much in this design and hence the author in the next chapter designs MSFA's that improve in addition, these parameters too. I feel that in this chapter, the author must clearly define the terms "fractional gain bandwidth" and "impedance bandwidth" clearly and illustrate mathematical methods for calculating these for a surface antenna. The computation of input impedance as a function of frequency for an antenna is based on giving a feed current and using Pocklington's integral equation based on vanishing boundary conditions on the tangential electric field on the antenna surface, calculate the surface current distribution on the antenna the the electric field generated by this distribution between the feed gap.

Chapter 5 presents the design of a low cost MSPA to enhance peak gain and gain-bandwidth product for ultra wide band applications. The radiation patch in this design comprises two rectangles and one triangle with two embedded rectangular slots to enhance peak gain and minimize losses. The design is further extended to two linear arrays. The chapter is well written but I would like to see some mathematical analysis about how one computes the electromagnetic fields within the cavity of an MSPA when the shape of the patch is given (for example a fractal shape) and the cavity depth is also given. For example, a finite element method (FEM) may be adopted here for solving Helmholtz' equations with probe sources and boundary conditions for the cavity fields:

$$(\nabla^2 + \omega^2 \epsilon \mu) [\mathbf{E}(\omega, \mathbf{r}), \mathbf{H}(\omega, \mathbf{r})] =$$

$$[-j\omega\mu\mathbf{J}(\omega, \mathbf{r}), \nabla \times \mathbf{J}(\omega, \mathbf{r})],$$

$$\mathbf{E}_{tan}|_{cavity boundary} = \mathbf{0}, \mathbf{H}_{nor}|_{cavity boundary} = \mathbf{0}$$

The finite elements must be 3-D like say cubes or tetrahedra which fill up the cavity.

Chapter 6 presents a rectangular microstrip patch antenna having broadband properties and yielding maximum depth of return loss having applications in amateur radio, terrestrial broadband, radio astronomy etc. Some mathematical expressions based on transmission line models for the return loss may be included in this chapter, like when an electromagnetic wave is incident on a boundary within the slot cavity, then what percentage of the incident power flux is absorbed and what percentage is reflected may be provided here. The analysis for computing return loss may be based on the technique used in standard electromagnetic field theory for computing the reflection and transmission coefficients when a plane wave is incident upon a boundary separating two dielectric media.

In Chapter 7, narrowband antenna design is improved to wideband antenna design, with mismatch loss much smaller and improved depth of return loss. I suggest that the author relate mismatch theory to matching theory for transmission lines which removes all the reflected wave components, thereby causing all the power to be transmitted to the load.

In chapter 8, the author designs a PMFA (Printed monopole fractal antenna) having high gain-bandwidth product in the given band. It has enormous impedance bandwidth. I suggest that the author as before, look at the computation of the input impedance to antenna of a given shape by solving for the electric field in between the probe terminals for a given current estimation. For this, he may as stated earlier, use the Pocklington integral equation (Balanis' book) for computing the antenna surface current density by equating the tangential components of the electric field on the antenna to zero with the electric field obtained by applying the retarded Greens' function to the surface current density. The Pocklington integral equation can be solved by discretizing the antenna surface into pixels and using the method of moments.

In summary, I feel that the candidate has performed a very good piece of research and well deserves the award of the PhD degree but after he answers some of the queries raised in the above report during the time of the viva voce examination. I also feel that if the above mentioned mathematical remarks are incorporated into the thesis, then it will become worthy of publication as a book on "Fractal microstrip antennas-theory and applications in the modern wireless age" in a reputed publishing company like Springer or Cambridge University Press.

6.59 Training a DNN with stochastic inputs with analysis of the robustness against input process and weight matrix fluctuations

$x(t) \in \mathbb{R}^d$ is the input process. $y(t) \in \mathbb{R}^{N \times d}$ is the desired output process. It may for example be obtained from the input via nonlinear distortions with additive noise:

$$y(t) = \psi(x(t)) + v(t)$$

The DNN output, assuming that it has just one layer is

$$z(t) = f(Wx(t) + b)$$

where $W \in \mathbb{R}^{M \times d}$ is the weight matrix, $b \in \mathbb{R}^M$ is the bias vector and $f : \mathbb{R}^M \rightarrow \mathbb{R}^N$ is a nonlinear map. If it has two layers then the output will be of the form

$$z(t) = f_2(W_2f_1(W_1x(t) + b_1) + b_2)$$

More generally, if it has L layers, then

$$z(t) = f_L(W_Lf_{L-1}(W_{L-1}f_{L-2}(\dots f_2(W_2f_1(W_1x(t) + b_1) + b_2)\dots) + b_L)$$

We write this as

$$z(t) = f(W, b, x(t))$$

where

$$W = (W_1, \dots, W_L), b = (b_1, \dots, b_L)$$

Suppose the bias b is fixed so we can write

$$z(t) = f(W, x(t))$$

We choose W so that

$$\mathbb{E}[(y(t) - z(t))^2]$$

is a minimum. Assume that this minimum is attained when $W = W_0$. Now suppose W_0 gets perturbed slightly to $W_0 + \delta W$ and simultaneously the input process $x(t)$ gets perturbed slightly to $x(t) + \delta x(t)$. We wish then to design the weight matrix perturbation δW so that the mean square error $\mathbb{E}(y(t) - f(W_0 + \delta W, x(t) + \delta x(t)))^2$ is still a minimum. Using Taylor series, we expand

$$\begin{aligned} f(W_0 + \delta W, x + \delta x) &= \exp((\delta W, \nabla_W) + (\delta x, \nabla_x))f(W_0, x) \\ &\approx f(W_0, x_0) + [(\delta W, \nabla_W) + (\delta x, \nabla_x)]f(W_0, x_0) \\ &\quad + (1/2)[(\delta W, \nabla_W) + (\delta x, \nabla_x)]^2f(W_0, x_0) \end{aligned}$$

In this approximation, we have retained only upto quadratic order terms in $\delta W, \delta x$. Then,

$$\begin{aligned} E(\delta W) &= \mathbb{E}(y(t) - f(W_0 + \delta W, x(t) + \delta x(t)))^2 \\ &\approx \mathbb{E}[(\epsilon(t) - ((\delta W, \nabla_W) + (\delta x(t), \nabla_x))f(w_0, x(t))) \\ &\quad - (1/2)[(\delta W, \nabla_W) + (\delta x, \nabla_x)]^2f(W_0, x(t))]^2 \end{aligned}$$

6.60 Quantum Boltzmann equation

An approximate derivation. The mixed N particle state $\rho(t) \in \mathcal{S}(\mathcal{H}^{\otimes N})$ is assumed to have all same marginals of any given order. Thus, if $k \geq 1$ and $i_1 < i_2 < \dots < i_k \leq N$ and (j_1, \dots, j_{N-k}) is the complement of $\{i_1, \dots, i_k\}$, then we are assuming that

$$\rho_{1\dots k}(t) = Tr_{k+1,\dots,N}(\rho(t)) = Tr_{j_1,\dots,j_{N-k}}(\rho(t))$$

The Hamiltonian of the system is

$$H = \sum_{a=1}^N H_a + \sum_{1 \leq a < b \leq N} V_{ab}$$

where H_a acts in \mathcal{H}_a while V_{ab} acts in $\mathcal{H}_a \otimes \mathcal{H}_b$. Here, $\mathcal{H}_a, a = 1, 2, \dots, N$ are identical copies and so are $V_{ab}, 1 \leq a < b \leq N$. The Schrodinger-Von-Neumann evolution equation for the density $\rho(t)$ is

$$i\rho'(t) = [H, \rho(t)]$$

This gives on taking partial trace over $2, 3, \dots, N$,

$$i\rho'_1(t) = [H_1, \rho_1(t)] + (N-1)Tr_2[V_{12}, \rho_{12}(t)]$$

Again taking the trace over $3, 4, \dots, N$ gives

$$i\rho'_{12}(t) = [H_1 + H_2 + V_{12}, \rho_{12}(t)] + (N-2)Tr_3[V_{13} + V_{23}, \rho_{123}(t)]$$

We write

$$\rho_{12}(t) = \rho_1(t) \otimes \rho_1(t) + g_{12}(t),$$

$$\rho_{123} = \rho_{12} \otimes \rho_1 + g_{123} = \rho_1 \otimes \rho_1 \otimes \rho_1 + g_{12} \otimes \rho_1 + g_{123}$$

Then,

$$i\rho'_{12}(t) = i(\rho'_1 \otimes \rho_1 + \rho_1 \otimes \rho'_1 + g_{12'}) =$$

$$[H_1 + H_2 + V_{12}, \rho_1 \otimes \rho_1 + g_{12}] + (N-2)Tr_3[V_{13} + V_{23}, \rho_{12} \otimes \rho_1 + g_{123}]$$

Using the equation of motion for ρ_1 , this gives the following exact equation followed by the approximate equation based on treating g_{123} as small and of the second order of smallness as compared to V_{ab} ,

$$\begin{aligned} ig'_{12} &= [H_1 + H_2 + V_{12}, g_{12}] + (N-2)Tr_3[V_{13} + V_{23}, \rho_{12} \otimes \rho_1 + g_{123}] \\ &\approx [H_1 + H_2 + V_{12}, g_{12}] + (N-2)Tr_3[V_{13} + V_{23}, \rho_1 \otimes \rho_1 \otimes \rho_1 + g_{12} \otimes \rho_1] \\ &= [H_1 + H_2 + V_{12}, g_{12}] + (N-2)Tr_3[V_{13} + V_{23}, g_{12} \otimes \rho_1] \end{aligned}$$

If we no regard g_{12} as the same order of smallness as V_{ab} , then this last equation further approximates to

$$ig'_{12} = [H_1 + H_2 + V_{12}, g_{12}]$$

which gives

$$g_{12}(t) = \exp(-itad(H_1 + H_2 + V_{12}))(g_{12}(0)) = T_{12}(t)(g_{12}(0))$$

say. Then, the equation for ρ_1 becomes with this approximation

$$\begin{aligned} i\rho'_1 &= [H_1, \rho_1] + (N - 1)Tr_2[V_{12}, \rho_1 \otimes \rho_1 + g_{12}] \\ &= [H_1, \rho_1] + (N - 1)Tr_2[V_{12}, \rho_1 \otimes \rho_1] + (N - 1)Tr_2[V_{12}, T_{12}(t)(g_{12}(0))] \end{aligned}$$

This may be termed as the quantum Boltzmann equation.

6.61 List of Ph.D scholars supervised by Harish Parthasarathy with a brief summary of their theses

[1] S.N.Sharma: Applications of nonlinear filtering theory to certain problems in classical mechanics.

(NSIT-DU, 2004)

Summary: The stochastic two body gravitational problem in the presence of interplanetary dust modeled using Brownian motion has the form

$$\begin{aligned} r''(t) - r(t)\theta'^2(t) &= -GM/r^2 + f_r(t), \\ r(t)\theta''(t) + 2r'(t)\theta'(t) &= f_\theta(t) \end{aligned}$$

where f_r, f_θ are white noise processes, ie, formal differentials of Brownian motion. These equations are cast in the from of four coupled nonlinear Ito stochastic differential equations and mean and covariance propagation equations for this system are obtained by expanding the variables $r(t), \theta(t), r'(t), \theta'(t)$ about their mean values and applying Ito's formula for Brownian motion to obtain the covariance propagation equations. Further, by taking noisy measurements on the position of the body, the EKF is applied to obtain real time estimates of the trajectory. Sharma then applies the Kushner nonlinear filtering theory to the nonlinear Van-der-Pol and Duffing oscillator taking cubic corrections into account, ie, he goes a step further than the EKF which is based on expanding upto quadratic terms. More accurate filtering results are obtained by considering the joint evolution of the first three conditional moments.

[2] Vipin Behari Vats: Parameter estimation algorithms in nonlinear systems using nonlinear LMS algorithm with a study of the behaviour of the Lyapunov exponents of autonomous nonlinear systems for small initial perturbations around a fixed point.

(NSIT-DU, 2007)

[3] Tarun K.Rawat: Applications of stochastic nonlinear filtering theory to trajectory manoeuvring of spacecrafts and convergence analysis of least mean

phase algorithms using stochastic differential equations driven by Brownian motion.

(NSIT-DU, 2009)

[4] Sudipta Majumdar: Modeling and parameter estimation in nonlinear transistor circuits using Volterra approximations combined with wavelet based compression for data storage for the purpose of estimation.

(NSIT-DU, 2010)

[5] Arathi Vaish: Finite element method for determining the modes in waveguides having various kinds of cross section and with inhomogeneous and anisotropic media filling the guide taking into account background gravitational perturbations in the form of a curved space-time metric.

(NSIT-DU, 2011)

[6] Akash Rathee: Study of higher harmonic generation in nonlinear transistor circuits using Fourier series and perturbation theory.

(NSIT-DU, 2011)

[7] Rajeev Srivastava: Image modeling, smoothing and enhancement using partial differential equations with emphasis on diffusion equations with intensity dependent diffusion matrix coefficient.

(NSIT-DU, 2009)

[8] Rajveer S. Yaduvanshi: Magneto-hydrodynamic antenna construction analysis using Navier-Stokes and Boltzmann kinetic transport equation.

(NSIT-DU, 2010)

[9] Lalit Kumar: Studies in transmission line and waveguide analysis taking hysteresis and capacitive nonlinearities and quantum mechanical effects of transmission line and waveguide fields on atoms and quantum harmonic oscillators.

(NIT, 2018)

[10] Kumar Gautam: Quantum gate design by perturbing real quantum systems with electromagnetic fields.

(NSIT-DU, 2017)

[11] Rohit Singla: Studies in robot trajectory tracking and dynamic parameter estimation in the presence of noise and in master-slave teleoperation based tracking using adaptive control algorithms.

(Yet to submit, has published three technical papers in the impact factor-4 Springer journal "Nonlinear Dynamics")

[12] Navneet Sharma: Quantum parameter estimation using search algorithms with applications to quantum communication and quantum gate design.

Given an unperturbed Hamiltonian H_0 and a time dependent perturbation operator $V(t) = \sum_{k=1}^p \theta_k V_k(t)$, the problem is to estimate the parameter vector $\theta = (\theta_k)$ based on measuring observables $X_a, a = 1, 2, \dots, n$ on the system state at different times taking into account the collapse postulate of quantum mechanics. For example, let $U(t|\theta)$ denote the unitary evolution:

$$iU'(t, s|\theta) = (H_0 + \sum_{k=1}^p \theta_k V_k(t))U(t, s|\theta), t \geq s,$$

$$U(s, s|\theta) = I$$

Let $\{M_a : a = 1, 2, \dots, r\}$ be a POVM. Then, if $\rho(0)$ is the initial state of the system, the probability of measuring the outcomes a_1, a_2, \dots, a_s at times $t_1 < t_2 < \dots < t_s$ is given by

$$P(a_1, \dots, a_s, t_1, \dots, t_s | \theta) =$$

$$\text{Tr}(E_{a_s} U(t_s, t_{s-1}) E_{a_{s-1}} \dots U(t_2, t_1) E_{a_1} U(t_1, 0) \rho(0) U(t_1, 0)^* E_{a_1}^* U(t_2, t_1)^* \dots U(t_s, t_{s-1})^* E_{a_s}^*)$$

where

$$E_a = \sqrt{M_a}$$

We then estimate θ by applying the maximum likelihood method to this probability.

(Yet to defend, all examiner reports recommend award of degree)

[13] Pravin Malik: Antenna design using numerical solution of integral equations arising from the boundary conditions on the antenna surface.

[14] Manisha Khulbe: Studies in electromagnetic wave propagation in inhomogeneous, anisotropic and field dependent(nonlinear) media with applications to estimating the medium parameters from discrete measurements of the electromagnetic field at different space-time points. Applications to antenna design for wave propagation in nonlinear media are also considered. For this, perturbative expansions of the medium permittivity and permeability as Taylor series in the electric and magnetic fields as well as expansion of the field independent coefficients of the permittivity and permeability in terms of basis functions is performed. These expansions are substituted into the Maxwell equations to obtain a sequence of linear equations for each perturbative order. Boundary conditions of the electromagnetic fields on the antenna surface are applied to derive integral equations for the induced surface current density.

6.62 The problem of determining the surface current density induced on an antenna surface placed in a nonlinear inhomogeneous and anisotropic medium taking gravitational effects into account

Statement of the problem: Let $\epsilon_{\alpha\beta}^{\mu\nu}(\omega, r, \mathbf{F})$ denote the field dependent permittivity-permeability tensor. Here,

$$F_{\mu\nu} = A_{\nu,\mu} - A_{\mu,\nu}$$

is the covariant em field tensor with A_μ as the covariant em four potential. the dependence of $\epsilon_{\alpha\beta}^{\mu\nu}$ on the em field tensor $\mathbf{F} = ((F_{\mu\nu}))$ shows that the medium

is nonlinear and its dependence on the space-time coordinates x and the fact that this tensor is generally non-diagonal shows that the medium is nonlinear, inhomogeneous and anisotropic. The Maxwell field equations in such a medium are

$$(\epsilon_{\alpha\beta}^{\mu\nu}(\omega, r, \mathbf{F}) F^{\alpha\beta} \sqrt{-g}),_{\nu} = J^{\mu}(\omega, x)$$

yields the nonlinear wave equation in such a medium taking space-time curvature into account. We assume the Lorentz gauge conditions

$$(A^{\mu} \sqrt{-g}),_{\mu} = 0$$

Here,

$$A^{\mu} = g^{\mu\nu} A_{\nu}, F^{\mu\nu} = g^{\mu\alpha} g^{\nu\beta} F_{\alpha\beta}$$

We are here assuming that the gravitational field metric tensor is time independent and hence we are operating completely in the temporal frequency domain. Thus where-ever ∂_0 occurs in the above system, we replace it by the multiplication operator $j\omega$.

$$A_{\mu} = A_{\mu}(\omega, r),$$

$$F_{0r} = j\omega A_r - A_{0,r},$$

$$g_{\mu\nu} = g_{\mu\nu}(r)$$

where $r = (x, y, z)$. The above Maxwell equations can be expressed as

$$j\omega \epsilon_{\rho\sigma}^{\mu 0}(\omega, r, \mathbf{F}) F^{\rho\sigma} \sqrt{-g} + (\epsilon_{\rho\sigma}^{\mu m} F^{\rho\sigma} \sqrt{-g}),_m = 0$$

To simplify further, we must make use of the gauge condition:

$$(A_{\nu} \sqrt{-g} g^{\mu\nu}),_{\mu} = 0$$

gives

$$j\omega g^{0\nu} \sqrt{-g} A_{\nu} + (A_{\nu} g^{m\nu} \sqrt{-g}),_m = 0$$

If we use the synchronous reference system, then $g_{0m} = 0$ and $g_{00} = 1$. Then the metric can be expressed as

$$d\tau^2 = dt^2 - \gamma_{rs}(r) dx^r dx^s$$

This equation implies that $g^{0m} = 0$ and $g^{00} = 1$. So we get for the gauge condition,

$$j\omega \sqrt{\gamma} A_0 - (A_s \gamma^{ms} \gamma),_m = 0$$

where $((\gamma^{rs})) = ((\gamma_{rs}))^{-1}$.

Reference: L.D.Landau and E.M.Lifshitz, "The classical theory of fields", Butterworth and Heinemann.

Acknowledgements: I am grateful to Prof.Malay Ranjan Tripathi for suggesting me to write this article for a workshop conducted by him.